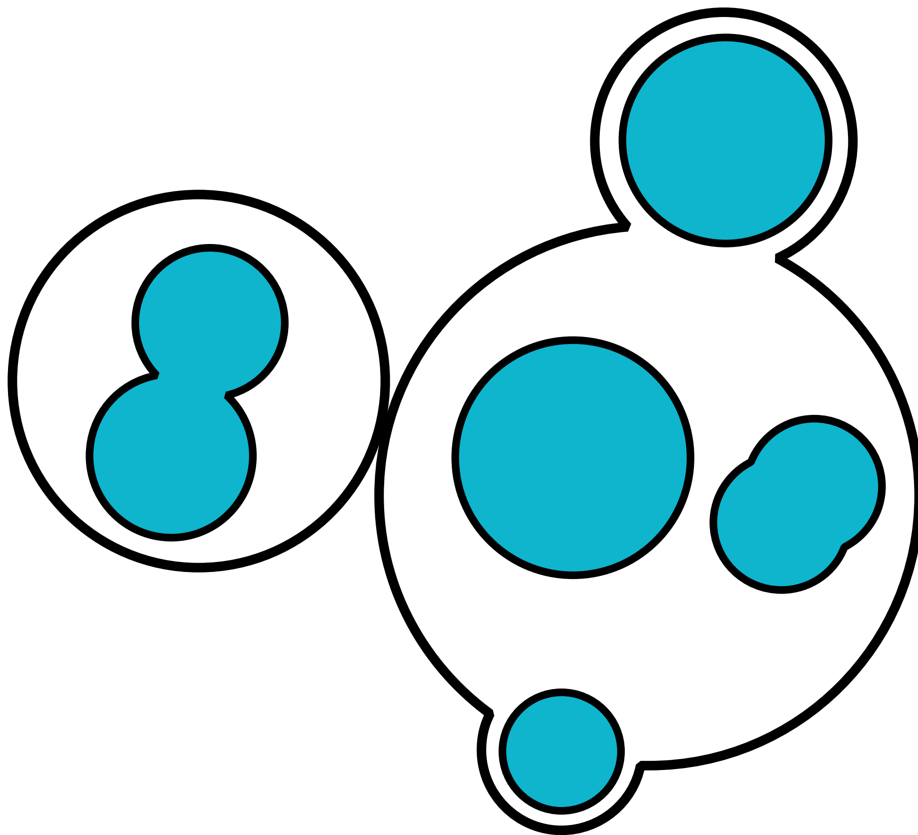
The Emergence of Biological Diversity Through Organisational Scale



Thomas Keggin

DISS. ETH NO. 29300

DISS. ETH NO. 29300

The emergence of biological diversity through organisational scale

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

Presented by

Thomas Keggin

Master of Science (Research)
Durham University

Born on 04.03.1992
Citizen of the British Islands, Isle of Man
Citizen of the United States of America

Submitted on the recommendation of

Prof. Dr. Loïc Pellissier
Prof. Dr. Stéphanie Manel
Prof. Dr. Nadir Alvarez

2023

This thesis is dedicated to my brother

Charlie Keggin

who believed in me and showed me how to believe in myself.

Contents

Acknowledgements 1

Summary 1

Résumé 2

Introduction 5

- Context 5
- Biological units 5
- Biological diversity 6
- Ecology and evolution 7
- Unifying ecological and evolutionary theory 7
- How to explore this new conceptual framework 9
- The study system 9
- This thesis – experimentation and observation 10
- References 11

Chapter 1: Diversity across organisational scale emerges through dispersal ability and speciation dynamics in tropical fish. 17

- Authors 17
- Author affiliations 17
- Abstract 17
- Introduction 17
- Methods 20
- Results 24
- Discussion 27
- Acknowledgements 35
- References 35
- Supplementary 41

Chapter 2: Patterns of genetic diversity differ between the Western Indian Ocean and Caribbean 49

- Authors 49
- Author affiliations 49
- Abstract 49
- Introduction 50
- Methods 52
- Results 57
- Discussion 60
- Ethics 62

Acknowledgements..62
References..63
Supplementary..68
Chapter 3: The emergence of diversity through organisational scale is associated with regional differences and biological traits..79
Authors..79
Author Affiliations..79
Abstract..79
Introduction..79
Methods..84
Results..87
Discussion..91
Conclusion..94
Acknowledgements..95
References..95
Supplementary..100
Discussion..102
Overview..102
Unified processes generate emergent diversity patterns. 102
How universal processes drive the emergence of diversity through organisational scale..103
Future studies and conservation implications..106
Methodological contribution..107
Challenges and reflections..108
Conclusion..109
References..109
Curriculum Vitae..113
Employment History..113
Education..114
Conferences..114
Publications..115

Acknowledgements

Firstly, I would like to thank Loïc Pellissier for tolerating me and my lack of wine knowledge for over 4 years. It’s been scary to inherit such a cool genetic data set and be entrusted with expanding it through sampling, lab work, and analysis. It can be difficult to appreciate the huge amount of effort underlying the sampling scheme and I have found it at (many) times overwhelming. Being trusted with it in an atmosphere where I felt “free to fail” has been liberating, personally meaningful, and must have taken a leap of faith in me which I very much appreciate. This is extended to Stéphanie Manel who gracefully stepped in as a second supervisor and has provided consistent and insightful feedback on all my messy work, as well as to Camille Albouy for your humour and supervisory support.

Technically, there is no way I could have completed the DNA extractions, library preparations, sequencing, and bioinformatics in the time frame that it was done without the endless patience and freely given support of Silvia Kobel, Niklaus Zemp, and Aria Minder of the Genetic Diversity Centre, ETH Zürich. The course of my PhD and my half-baked, sporadic requests for help have been chaotic, but always received suspiciously warmly by Team GDC. Whilst I don’t understand how someone can have so much patience, I am very appreciative.

I cannot claim most of the conceptualisation in thesis as purely my own as it has grown through many engaging discussions with many people. But in particular, Maurine Vilcot, with whom we have had sporadic and frenzied periods of unnaturally productive work – both in the lab for chapter 2 and in the conceptualisation, planning, and execution of the chapter 3 analysis. Much of my thinking on the topic has been shaped through our chats.

Without the friendship, support, and always constructive criticism of Conor Waldock and Alex Skeels I think I would have quit the course years ago and this document would not exist. I don’t know how to put this one into words, other than I consider myself unreasonably fortunate in meeting two such excellent people and look forward to your continued friendship.

I want to thank the “Grubs”, Eilísh Richards, Conor Waldock, and Hannah Matthews for our covid lockdown support group. Living alone during the pandemic could have been much worse.

I want to thank, in general, everyone who I have overlapped with in the Landscape Ecology / Ecosystems and Landscape Evolution group. In particular Charles Novaes de Santana for the friendly enthusiasm when I started and Victor Boussange and Flora Desmet for letting me awkwardly live with them on arriving in Zürich. I want to thank Lydian Boschman, your bluntness and direction I find inspiring, as well as Giulia Donati for your grace in passing on the data set and for encouraging me to get back into horse riding. Oskar Hagen, I would like to thank for your general aura of friendly chaos and dubious-yet-effective finishing-your-PhD advice. Ian McFadden, Joan Casanelles Abella, Yohann Chauvier, Conor and Alex again; the sloppy mix of personal and professional support has been a rock for me.

I want to thank Daniel Earley for making the poor decision of becoming flatmates 6 months before my deadline and Sarah Schmid for making a similarly poor choice. Having a stressful time with company is better than having a stressful time alone.

I want to thank my mum and dad, Sarah and James Keggin for praising enthusiasm over achievement and I would like to thank my sister Georgie for always believing in me regardless of how much I over share and annoy her. I want to thank Tristan Jaensch, Liam Jackson, and Dave “Brew King” Santiago

for the food, boxing, and beer, respectively – you guys have kept me straight on this bent journey. Similarly, to Sam Hibdige and Emily Guest for their hilarious and invigorating PhD support song, with further thanks to Jamie Ranaldi and Leah Kelly for the dungeons and dragons which provided much needed escapism. The same goes for Alex Jørgensen, Xavier Rossiter, Daniel Smith, and Patrick Evans who were overly available to play games over the years.

Finally, I would like to especially thank Georgie Keggin, Sam Hibdige, and Sarah Schmid for sacrificing too much of your spare time proofreading large swathes of this thesis and helping translate the summary. I am indebted to all three of you.

Summary

Our understanding of biological diversity is often segregated into fields of study, each focussing on a particular set of patterns, dynamics, or scale. This makes the overwhelming complexity of life much more manageable, but at the risk of losing a holistic overview of the single living system we are studying. Two closely entwined fields, ecology and evolution, have common roots and attempt to explain similar patterns and processes. Despite this, over the last century or so they have remained relatively disconnected; the main separation being that evolution often has a focus on patterns and processes at the organisational level of individuals and populations, and ecology with a focus on species and communities. However, these organisational levels comprise a much larger hierarchical scale of organisation, from nucelotides to ecosystems, with each level higher being an aggregate of the biological units from the levels below: a population is an aggregation of individuals, a species is an aggregation of populations, a community is an aggregation of species, and so on. To fully understand how patterns of diversity form across these aggregated levels of organisation, we need to embrace the fact that they comprise a single, unified system. What happens at one level will have cascading consequences at all other levels. Throughout this thesis I argue, along with others, that the division between ecology and evolution, and the associated study of organisational levels in isolation, needs revision. I put forward that the formation of diversity across organisational scale is explainable through universal processes and that we have the tools available to support this unified conceptual framework. The thesis is organised into an introduction, three chapters, and a discussion.

In the introduction, diversity is defined as variation between biological units at different levels of organisation, and how these fit into a nested, hierarchical scale. I then outline how a part of this scale has historically been tackled by the fields of ecology and evolution and how I believe they can be unified into a single conceptual framework. This is followed by an overview of how this could be done using a parallel experimental and observational approach using tropical reef fishes as a suitable study system.

In chapter 1, we apply a mechanistic model to this conceptual framework by simulating the formation of tropical reef fish diversity over the last 200 mya using population level processes. From only population-level processes, patterns of both population and species diversity emerge which we can then investigate holistically. To better understand the dynamics of diversification at both the population and species levels of organisation, we apply an existing species-genetic diversity correlation approach. We also develop and apply a “continuity” metric to measure the relative ratio between levels of diversity and use it to investigate the role of model parameters, corresponding to biological traits, on the emergence of diversity from the population level to the species level. We find correlations between population and species diversity patterns, as well as significant impacts of biological traits. From these results, we propose a diversity partitioning mechanism from the population level to the species level through speciation.

In chapter 2, we expand an existing population genetics data set from the Western Indian Ocean to the Caribbean. The sampling scheme consists of 42 species across both regions encompassing a wide range of biological traits and phylogenetic history, whilst retaining closely related, comparable species between regions. We analyse this data set to understand the roles of biological traits and seascape in influencing patterns of genetic diversity in this study system. We find that the effect of seascape is dominant over that of traits, with the larger, more connected habitat patches in the Caribbean corresponding to greater diversity, increased gene-flow, and reduced inbreeding. We conclude that the impacts of biological traits are dependent on their environmental context and that this has implications for species responses to climate change.

In chapter 3, we validate the patterns and expectations of process developed in chapter 1 by comparing the population genetics data set of chapter 2 to species level data from the literature. Following the same but expanded conceptual and analytical framework as chapter 1, we find correlations between genetic and species levels of diversity that correspond to the existing literature. We expand the analysis by applying the continuity metric developed in chapter 1 and find that the emergence of diversity across organisational levels is influenced by both biological traits and the environment. We find support for the diversity partitioning mechanism through speciation proposed in chapter 1.

In the discussion, the results of the three chapters are put into context, with an emphasis on the comparison between the experimental approach of chapter 1 and the observational approach of chapter 3. The contribution of these chapters is put into the context of a conceptual eco-evolutionary framework based on unified processes across organisational scale. The exact action of the processes identified and validated in chapters 1 and 3 are examined in more detail with speculation as to likely dynamics and areas of remaining uncertainty. I propose avenues of future work I believe would be fruitful to follow and discuss the conservation implications of this work. This is followed by an overview of the methodological contribution of the continuity metric and how it has improved, and should facilitate, work in eco-evolutionary study.

Overall, this thesis presents and applies a conceptual framework of eco-evolutionary diversification that bridges levels of organisational scale through universal processes. By implementing a twinned empirical approach of experimentation and observation, a foundation is laid for further reconciliation between the merging fields of ecology and evolution to build a clearer, more holistic understanding of biological diversity.

Résumé

Notre compréhension de la diversité biologique est souvent divisée en domaines d'étude, chacun se concentrant sur un ensemble particulier de modèles, de dynamiques ou d'échelles. Cela rend l'extraordinaire complexité de la vie beaucoup plus gérable, mais au risque de perdre une vue d'ensemble du système vivant unique que nous étudions. Deux domaines étroitement liés, l'écologie et l'évolution, ont des racines communes et tentent d'expliquer des schémas et des processus similaires. Malgré cela, elles sont restées relativement déconnectées au cours du siècle dernier, la principale différence étant que l'évolution se concentre souvent sur les schémas et les processus au niveau organisationnel des individus et des populations, tandis que l'écologie se concentre sur les espèces et les communautés. Toutefois, ces niveaux d'organisation constituent une échelle hiérarchique beaucoup plus vaste, allant des nucléotides aux écosystèmes, chaque niveau supérieur étant un agrégat des unités biologiques des niveaux inférieurs : une population est un agrégat d'individus, une espèce est un agrégat de populations, une communauté est un agrégat d'espèces, et ainsi de suite. Pour bien comprendre comment les schémas de diversité se forment à travers ces niveaux d'organisation agrégés, nous devons admettre qu'ils constituent un système unique et unifié. Ce qui se passe à un niveau engendrera une cascade de conséquences à tous les autres niveaux. Tout au long de cette thèse, je défends l'idée – également soutenue par d'autres scientifiques – que la division entre l'écologie et l'évolution ainsi que l'étude associée des niveaux d'organisation isolés, doivent être révisées. Je suggère que la formation de la diversité à travers l'échelle organisationnelle peut être expliquée par des processus universels et que nous disposons des outils nécessaires pour soutenir ce cadre conceptuel unifié. La thèse est constituée d'une introduction, trois chapitres et une discussion.

Dans l'introduction, la diversité est définie comme la variation entre les unités biologiques à différents niveaux d'organisation, et la manière dont ces unités s'inscrivent au sein d'une hiérarchie constituée de différents niveaux imbriquées les uns dans les autres. Je décris ensuite comment une partie de cette hiérarchie a été historiquement abordée par les domaines de l'écologie et de l'évolution et comment je pense que ces deux domaines peuvent être unifiés dans un cadre conceptuel unique. Cette présentation est suivie d'une vue d'ensemble de la manière dont cela pourrait être réalisé en utilisant une approche expérimentale et observationnelle parallèle, en utilisant les poissons de récifs tropicaux comme système d'étude.

Dans le chapitre 1, nous appliquons un modèle mécaniste à ce cadre conceptuel en simulant la formation de la diversité des poissons des récifs tropicaux au cours des 200 derniers millénaires en se basant sur des processus intervenant au niveau de la population. À partir des seuls processus au niveau de la population, des modèles de diversité des populations et des espèces émergent, que nous pouvons ensuite étudier de manière holistique. Pour mieux comprendre la dynamique de la diversification au niveau des populations et des espèces, nous appliquons une approche existante de corrélation entre la diversité génétique et la diversité des espèces. Nous développons et appliquons également une métrique de "continuité" pour mesurer le rapport relatif entre les niveaux de diversité et nous l'utilisons pour étudier le rôle des paramètres du modèle, correspondant aux traits biologiques, sur l'émergence de la diversité du niveau de la population au niveau de l'espèce. Nous trouvons des corrélations entre les modèles de diversité des populations et des espèces, ainsi que des impacts significatifs des traits biologiques. À partir de ces résultats, nous proposons un mécanisme de répartition de la diversité du niveau de la population au niveau de l'espèce par le biais de la spéciation.

Dans le chapitre 2, nous enrichissons un jeu de données de génétique des populations existant de l'océan Indien occidental aux Caraïbes. Le plan d'échantillonnage comprend 42 espèces dans les deux régions, englobant une large gamme de traits biologiques et d'histoire phylogénétique, tout en conservant des espèces étroitement liées et comparables entre les régions. Nous analysons cet ensemble de données pour comprendre le rôle des caractéristiques biologiques et du paysage marin dans l'influence des schémas de diversité génétique dans ce système d'étude. Nous constatons que l'effet du paysage marin est plus important que celui des traits. En effet, les parcelles d'habitat plus grandes et plus connectées dans les Caraïbes correspondent à une plus grande diversité, à un flux génétique accru et à une consanguinité réduite. Nous concluons que les traits biologiques dépendent de leur contexte environnemental et que cela a des implications pour les réponses des espèces au changement climatique.

Dans le chapitre 3, nous validons les modèles et les attentes du processus développés dans le chapitre 1 en comparant l'ensemble des données de génétique des populations du chapitre 2 aux données sur les espèces tirées de la littérature. En suivant le même cadre conceptuel et analytique élargi que celui du chapitre 1, nous trouvons des corrélations entre les niveaux de diversité génétique et de diversité des espèces qui correspondent à la littérature existante. Nous élargissons l'analyse en appliquant la métrique de continuité développée au chapitre 1 et nous constatons que l'émergence de la diversité à travers les niveaux organisationnels est influencée à la fois par les caractéristiques biologiques et par l'environnement. Nous étayons ainsi le mécanisme de répartition de la diversité par spéciation proposé au chapitre 1.

Dans la discussion, les résultats des trois chapitres sont mis en contexte, en mettant l'accent sur la comparaison entre l'approche expérimentale du chapitre 1 et l'approche observationnelle du chapitre 3. La contribution de ces chapitres est replacée dans le contexte d'un cadre conceptuel éco-évolutif basé sur des processus unifiés à l'échelle de l'organisation. L'action exacte des processus

identifiés et validés dans les chapitres 1 et 3 est examinée plus en détail, avec des spéculations sur les dynamiques probables et les zones d'incertitude qui subsistent. Je propose des pistes de travail futures que je considère comme potentiellement fructueuse et je discute des implications de ce travail en matière de conservation. Ceci est suivi d'une vue d'ensemble de la contribution méthodologique de la métrique de continuité et de la façon dont elle a amélioré, et devrait faciliter, les différents travaux de recherche éco-évolutive.

Dans l'ensemble, cette thèse présente et applique un cadre conceptuel de la diversification éco-évolutive qui relie les niveaux d'échelle organisationnelle par le biais de processus universels. En mettant en œuvre une approche empirique jumelée d'expérimentation et d'observation, une base est posée afin de réconcilier davantage les domaines de l'écologie et de l'évolution permettant ainsi de construire une compréhension plus claire et plus holistique de la diversité biologique.

Introduction

One of the key steps for future research will be to develop a general theoretical framework for eco-evolutionary dynamics—and then to quantify these dynamics in natural populations.

Pelletier et al. (2009)

Context

Like all species, we are an intrinsic part of our environment and as objective as we might believe ourselves to be, the environment has shaped how we think and relate to our surroundings. The discussion of humanity's relationship with nature over the millennia is outside the scope of this thesis, but it is important to be aware of our own philosophical context and constraints when trying to understand what is going on around us. Ultimately, whatever we perceive is only a model of reality, built on our flawed sensory perceptions and the conceptual foundation we adopt (Descartes et al., 1954; Lupyan, 2017). Whichever philosophical foundation we choose, we should remain aware of our axioms and assumptions, and remind ourselves that anything we learn is likely an incomplete representation of the truth. We are only human, and as famously put by JBS Haldane in 1927, "…the universe is not only queerer than we suppose, but queerer than we can suppose." (Haldane, 1927). With these considerations in mind, this thesis is written in the context of western scientific philosophy, with the focus on understanding the formation of biodiversity mostly through the lens of the evolutionary synthesis developed over the last 100 years, and its associated literature (Lewens, 2019). Through this thesis I would like to express my own perspectives of the diversification of life, try to highlight how it might differ to established assumptions in my own cultural context, and hopefully contribute to improving our unavoidably flawed models of the natural world. I would like to start by working through what I believe to be the logic and flaws of our current study of diversity, then outline how I have, with help, attempted to address some of these issues.

Biological units

The term "diversity" is often thrown around on the assumption that the reader knows what the writer is referring to depending on context and usage. This is problematic and it is worth defining our fundamental concepts at the beginning. Biodiversity, here, is the variation between and within biological units (Gaston & Spicer, 2013). These units are essentially groupings of biological material and their associated information that aggregate through levels in an organisational scale. At the basal level, we have molecules carrying heritable information – mainly nucleotides (Charlesworth et al., 2017). These nucleotides are aggregated into genes, then chromosomes, then genomes – one or more of which are contained within an individual organism. These individuals can form their own spatial aggregations, called populations, which comprise a species – commonly defined as an aggregation of individuals that can sustainably reproduce with one another (Mayr, 1963). Through speciation, reproductively isolated individuals can (mostly) no longer back-cross (Dobzhansky, 1974; Hibdige et al., 2021; Mayr, 1963), but retain their shared evolutionary history. We can attempt to aggregate species together based on their similarities (Rabosky et al., 2018) through taxonomic classification. However, creating discrete levels out of the continuous scale of inter-species differentiation is ultimately flawed (Laurin & Arntzen, 2010). Alternatively, we can continue to aggregate diversity spatially based on the co-occurrence of individuals. This inevitably becomes complex, as an aggregation of individuals belonging to different species within a single geographic area, which we call a community, doesn't necessarily encompass all the individuals of the entire

species. A community is dependent on the boundaries of the area, a problem similar in defining populations and much easier to resolve in island systems (MacArthur & Wilson, 2016; Waples & Gaggiotti, 2006). Either way, we can see how diversity quickly becomes confused depending on how we choose to aggregate our observations.

Importantly, not all biological units are as discrete as others. The mitochondrial, chloroplast, and nuclear genomes in Eukaryotes are distinct structural entities, as are single nucleotides. Genes, which were originally thought of as a unit of abstract inheritance, are functional groupings of nucleotides that can be difficult to define (Portin & Wilkins, 2017). Individuals are much more obviously discrete, and species identity can be confirmed through testing the reproductive success of different individuals but can become messy depending on reproductive mode and the extent of horizontal gene transfer (Padial & De la Riva, 2021). Populations and communities are defined subjectively (Waples & Gaggiotti, 2006), and are really more arbitrary breakpoints between spatially distributed individuals. After some consideration, we see that biological units are on a continuum of "discreteness" which we must navigate. Through this framework of understanding we must keep in mind two things: that not all biological units are equally coherent; and that apart from the nucleotide, each biological unit is comprised of an aggregation of smaller biological units. This is the nested scale of biological organisation.

Biological diversity

Now that we have defined biological units and the nested structure in which we have placed them, defining what we mean by diversity becomes simpler. Following the framework of Tucker et al. (2017), we can think of and measure diversity according to three questions that correspond to three facets of diversity:

How much/many? **Richness.**
How different? **Divergence.**
How regular? **Regularity.**

These can be applied to any of our biological units that we've aggregated through organisational scale. How many genes? How many chromosomes? How different are populations (Meirmans & Hedrick, 2011; Wright, 1949)? How regular are communities? Each question corresponds to a conceptual facet of diversity that is applicable to all units through scale. What remains is to understand how to measure them. This returns us to the nested scale of organisation: to compare two biological units, we must measure and contrast their constituent parts. For example, we want to compare the diversity of two spatially distinct communities. For richness, we can ask: how many species? For divergence we can ask: how different are species? And for regularity we can ask, how regular are species? To compare communities, we acknowledge that they are an aggregate of species and define their diversity through species quantification. Now we compare two species: how many individuals? How different are individuals? How regularly are they distributed? We drop a level of aggregation in both the target unit and units that we quantify: to understand communities, we measure species; to understand species, we measure individuals; to understand individuals, we measure genomes; all the way down to the humble nucleotide. This conceptual framework provides us with both biological units to measure, and an idea of how to measure them. The next step is to define what we would like to learn which, motivations aside, can be reduced down to a question of, "how have observed patterns of diversity (richness, divergence, and regularity) come about?". The question of how is a question of process, and process is where things become complex.

Ecology and evolution

Understanding the processes that explain diversity has been tackled in the last century by two closely linked concurrent disciplines: ecology, and evolution. This divide is despite the inherent interconnectivity of biological units through organisational scale and it is surprising that the study of these different units has been so disparate (Bailey et al., 2009; Coulson et al., 2006; Pelletier et al., 2009), especially considering the frequency of applications of theory from one field to the other (Hubbell, 2001; Kimura & Weiss, 1964; Laurin & Arntzen, 2010). This divide is historical. Evolution as we know it today was kickstarted with Darwinism (Darwin, 1859), then neo-Darwinism or the Modern Synthesis (Hancock et al., 2021), now moving into the still developing Evolutionary Synthesis or Extended Evolutionary Synthesis. Despite this very recent insight in the long history of human record, attempts to explain how exactly the diversity of life has come about, as a consequence of the environment, goes back arguably as far as 400 BCE (Zirkle, 1941). Ecology is often claimed to be a younger field dating back only mid-way through the 20th century (McIntosh, 1986). But the term "ecology" was coined in 1866 by a purported admirer of Darwin, Ernst Haeckel, to encompass the relationship between organisms and their environment (Haeckel, 1866). However, the concept of ecology has been argued to go back as far as ancient Greece under similar reasoning as the concept of evolution (Egerton, 2001). Regardless, many modern ecological concepts were blurred with burgeoning Darwinian theory (Egerton, 2013). The difference between them often stated that ecology seeks to understand the relationships between organisms, their environment, and one another; whilst evolution seeks to understand the processes that drive change in organisms over time (Lewens, 2019). But consider that changes in a species' traits impacts the dynamics of an ecosystem (Yoshida et al., 2003) and that competition over a dynamic landscape results in changes in species' traits (Brockhurst et al., 2014). It is apparent that these dynamics occur in the same arena, and the separation of ecological and evolutionary study seems arbitrary. Yet instead of formally acknowledging this unity, evolutionary and ecological dynamics are still commonly subdivided and treated as separate entities that instead interact with one another (Leibold et al., 2022; Segar et al., 2020). This eco-evolutionary divide is despite the overlapping objectives and theory since their modern inception in the second half of the 1800s (Darwin, 1859; Haeckel, 1866), increasing acknowledgement that the two disciplines should be unified (Lamy et al., 2017; Pelletier et al., 2009; Schmidt et al., 2022), and concerted efforts to link them (Ware et al., 2019). For me, I believe the theoretical separation of the two disciplines is an interesting 20th century trend that has corresponded to an incredible surge in knowledge, but has impeded the shared objective of holistically understanding biological diversity.

Unifying ecological and evolutionary theory

Efforts to merge ecology and evolution centres around the idea that evolutionary processes and patterns often concern genomes, individuals, and populations; and ecological processes and patterns concern communities and ecosystems (Vellend, 2005; Ware et al., 2019). Evolutionary theory has mostly developed through the advent of population genetics and the understanding that four main processes drive changes in organisms over time: genetic drift, gene-flow, mutation, and selection (Hamilton, 2021). Diversification across populations is of interest, and so patterns one aggregative level lower in individuals and their genotypes are the units of measurement (Hamilton, 2021). Ecological theory has been concerned with the diversification of communities and ecosystems, and similarly patterns one aggregative level lower across species and their traits are the units of measurement (Jørgensen & Fath, 2014). Efforts to link ecology and evolution therefore often focus on trying to incorporate patterns and process at one level (population and individual) with

those at the other level (community and species; Bailey et al., 2009; Des Roches et al., 2018; Pelletier et al., 2009; Ware et al., 2019; Whitham et al., 2006).

This has been framed comparatively, for example by Vellend and Geber (2005) who identify the parallel processes between community ecology and population genetics to construct a framework of reciprocal feedback between species diversity (community) and genetic diversity (individual and population). Identifying parallel processes reminds us of the close link between ecology and evolution and reduces the apparent complexity of community ecology. According to this framework (Antonovics, 2003; Vellend & Geber, 2005), the four main population genetic processes have parallels at the community level: gene-flow between populations is analogous to dispersal between communities – both processes homogenise biological units; stochastic genetic drift in allele frequencies in isolated populations corresponds to random changes in species composition between communities; selection acts upon both alleles and species; and mutation generates new variants in a population, much like speciation generates new species in communities. This framework established a set of simple expectations: that if processes at both the species and genetic levels are comparable, then they should produce the same patterns (Antonovics, 1976, 2003; Chave, 2004; Huston & Huston, 1994; Vellend, 2005, 2010; Vellend et al., 2014). This can be directly tested by fitting correlations between the two levels of organisation, termed Species-Genetic Diversity Correlations (SGDCs). Interestingly, these SGDCs are highly variable (Kahilainen et al., 2014; Lawrence & Fraser, 2020; Manel et al., 2020; Taberlet et al., 2012). Either the conceptual framework is correct and we have yet to fully identify feedbacks between organisational levels, or the conceptual framework needs revision. Through SGDCs, environmental variables such as habitat fragmentation (Reisch & Hartig, 2021) and habitat heterogeneity (Lamy et al., 2013; Schmidt et al., 2022) have been proposed to create a discordance between patterns of diversity at each organisational level. But if we reflect on the nested scale of biological organisation, we should remind ourselves that: 1. not all biological units are equally robust; and 2., apart from the basal nucleotide, all biological units are aggregations of smaller biological units. If we deconstruct these “parallel processes” (Lamy et al., 2017; Vellend & Geber, 2005) we could reframe this eco-evolutionary separation. Gene-flow and dispersal between communities: population gene-flow is the reproductive result of the movement of individuals or their propagules between populations, species dispersal is the movement and persistence of individuals or propagules between communities. Underlying both gene-flow and dispersal is the movement of individuals and their constituent alleles across a land- or sea-scape. There is a single process at play – individual dispersal. Similarly, drift is the stochastic change in frequency of alleles within a population (Wright, 1943, 1949) or species within a community (Vellend & Geber, 2005). Underlying both is a single process across both levels – random persistence of individuals and their constituent alleles. In population genetics, selection is the non-random survival of alleles in a population (Darwin, 1859; Hamilton, 2021); in community ecology, selection is the non-random survival of species in a community (Jørgensen & Fath, 2014). But reduced down, there is only one process: the non-random fitness of individuals and their constituent alleles. An exception to this reduction is mutation and speciation: mutation and subsequent generation of new genetic variants from existing biological material is a result of molecular replicative machinery and reproductive mode (Ellegren & Galtier, 2016). Similarly, speciation is the generation of diversity from existing material as a result of reproductive isolation between individuals (Mayr, 1963). However, I would still like to posit that whilst similar in action and intrinsically linked – speciation occurs through sufficient genetic differentiation through mutation, drift, and selection to disrupt genomic compatibility (Wolf et al., 2010) – mutation and speciation should not be described as parallel. Mutation within the constituent species of communities will have emergent effects on community dynamics (Ware et al., 2019), and speciation may have consequences for the individuals and alleles

contained within both the parent and offspring species. Overall, conceptually, the framework becomes simpler: there are unified processes that have consequences across multiple levels of organisation. Diversity patterns that we measure in ecology and those that we measure in evolution should be the end result of unified eco-evolutionary processes across organisational scale.

Table 1: Unified processes across organisational scale.

Population genetic	**Community ecology**	**Unified**
Gene flow	Dispersal	*Individual dispersal*
Drift	Drift	*Random survival of individuals*
Selection	Selection	*Non-random survival of individuals*
Mutation		*Mutation*
	Speciation	*Speciation*

How to explore this new conceptual framework

The idea of unified processes producing emergent patterns across the various levels of organisational scale should at this point seem plausible. However, to my knowledge, it has not yet been explicitly explored through either experimental or observational study. Although, I believe some conceptual frameworks come quite close by incorporating population genetic models interacting across multiple species, then linking them directly to emergent species-environment dynamics (Ware et al., 2019). To combat this lack of proof, a two-pronged empirical approach is required to satisfy the scientific method – observation and experimentation. First, conducting experiments should allow for us to establish a new model framework for how diversity emerges through levels of organisation, as well as to establish some expectations of subsequent diversity patterns across organisational scale. We can then follow this with observations in the natural world to verify our results. But before this is possible, we must first choose an appropriate biological study system.

The study system

For the study system, we focus on tropical reef-associated fishes as they make an easy fit for testing our conceptual framework. They comprise a relatively recent and very diverse radiation (Rabosky et al., 2018), are mostly restricted to islands of shallow water habitat, are distributed throughout the tropics globally (Parravicini et al., 2021), have relatively strict environmental limits (Waldock et al., 2019), and have been of conservation concern for decades due to their perceived natural beauty and vulnerability to climate change (Hatcher et al., 1989). We also have a relatively clear idea of their evolutionary development through a fossil record associated with reef formation dating back to the late Cretaceous (Near et al., 2013; Sallan & Friedman, 2012; Bellwood et al., 2015). Evidence suggests that diversification of modern lineages and assemblages started after the K/Pg extinction event mostly in historical hotspot of the Tethys Sea between the Eurasian and African continental plates (Cowman & Bellwood, 2011; Cowman & Bellwood, 2013; Cowman et al., 2009). The Eocene-Oligocene boundary coincides with high extinction rates as the Tethys began to close and temperature regimes changed (Cowman & Bellwood, 2011). This coincided with the development of the Indo-Australasian Archipelago (IAA) hotspot which saw rapid diversification and remains today the most species rich tropical reef fish area (Parravicini et al., 2021; Renema et al., 2008). From this, we know their phylogenetic history (Rabosky et al., 2018), an estimation of their historical distributions (Bellwood et al., 2015), and their current global distributions, limits, and abundances (Edgar et al., 2020; Parravicini et al., 2021; Waldock et al., 2019). The fact that reefs consist of a

global set of discrete island habitats reduces the complexity of investigating multiple levels of biological organisation – especially if we remind ourselves that both populations and communities are sometimes arbitrary and non-discrete units (Spalding et al., 2007; Waples & Gaggiotti, 2006; Warren et al., 2015). This is not the first attempt to leverage tropical reef fishes against eco-evolutionary processes across organisational scale. A positive relationship between species richness and mitochondrial diversity across both freshwater and marine fishes (Manel et al., 2020), and between mitochondrial diversity and species richness in tropical Pacific fishes (Messmer et al., 2012). A positive relationship between measures of inter-population divergence derived from putatively neutral single nucleotide polymorphisms and species richness was also found in the Western Indian Ocean (Vilcot et al., 2023) – indicating that diversity patterns across organisational scale are likely further dependent on the spatial partitioning considered.

This thesis – experimentation and observation

I believe we are now equipped with a conceptual framework of unified eco-evolutionary processes acting across biological organisational scale, and a suitable study system to test it on – what remains is the how. This thesis includes three chapters which include: *in silico* experimentation as a proof of concept and establishment of expectations, generation of a macro-genetics dataset for tropical coral-reef associated fishes, and *in vivo* observations of diversity patterns across both a population macro-genetics dataset and corresponding species-based community data. These three chapters are outlined below:

Chapter 1

The *in silico* experimentation was carried out using a bottom-up mechanistic model that simulates explicit population to species level objects through the population-level manipulation of abundance, trait evolution, speciation, and dispersal – supporting dynamic environmental landscapes (O. Hagen et al., 2021). Whilst a young model, it, and its precursors, have been used to accurately simulate process and pattern at both the population (Leugger et al., 2022) and species levels (Boschman et al., 2021; Oskar Hagen et al., 2021). Further, the bottom-up approach is a good fit for exploring unified process to generate emergent patterns across organisational levels. We are limited to the population and species levels for biological units, but this is sufficient to provide a proof of concept and results enough to set expectations of emergent diversity patterns through at least two levels in organisational scale. This chapter simulates the emergence of tropical reef fish diversity over a tectonically and thermally dynamic seascape covering the last 200 million years. We find that this simulation does a surprisingly accurate job of replicating global species richness patterns for this system. More pertinently, we find correlations between the population and species levels in different facets of diversity and develop the “continuity” metric to measure the ratio of relative diversity between organisational levels. We find differences in the emergence of diversity between facets across organisational levels as well as associations between the continuity metric and simulation parameters emulating heritable biological traits. These associations provide patterns and dynamics we could expect to observe in natural systems, and align with the existing literature (Manel et al., 2020; Schmidt et al., 2022; Vilcot et al., 2023). In particular, we highlight the importance of a potential speciation diversity partitioning mechanism in influencing the emergence of diversity from populations to species.

Chapter 2

Whilst many data already exist for tropical reef systems at the species level (Albouy et al., 2019; Boettiger et al., 2012; Edgar et al., 2020; Manel et al., 2020; Parravicini et al., 2021; Rabosky et al., 2018), we still lack a unified genetic sampling scheme across geographic space, biological traits, and

reef fish phylogeny. In this chapter, we expand an existing single nucleotide polymorphism (SNP) data set in the Western Indian Ocean (Donati et al., 2021) by sampling, sequencing, and genotyping comparable species in the Caribbean. Species are sampled across the teleost phylogeny and, where possible, have representative close relatives in both the Western Indian Ocean and Caribbean. This allows us to compare the effect of both seascape and biological traits on patterns of population genetic diversity. We find that the effect of seascape is dominant, with higher observed heterozygosity and lower differentiation in the Caribbean compared to the Western Indian Ocean. This corresponded to increased levels of inbreeding in the Western Indian Ocean. We infer that these patterns are likely due to the smaller and more isolated habitat patches found in the Western Indian Ocean. The results support classical neutral population genetics theory (Hamilton, 2021; Kimura & Weiss, 1964; Nei & Takahata, 1993; Wright, 1943, 1949) and provide a suitable foundation for our *in vivo* observation of diversity patterns across organisational scale in chapter 3.

Chapter 3

We take the conceptual and analytical framework developed *in silico* in chapter 1 and apply it to both the SNP dataset produced in chapter 2 and various sources of community-level biodiversity data including species distributions (Parravicini et al., 2021), species phylogeny (Rabosky et al., 2018), and species traits (Boettiger et al., 2012; Luiz et al., 2013). We measure SGDCs across multiple facets of diversity which provide limited information on the emergence of diversity patterns through organisational levels. We then apply the continuity metric developed in chapter 1 to investigate associations between the ratios of species to genetic diversity and lineage traits and the effect of seascape. We find that the differing seascapes of the Caribbean and Western Indian Ocean play a role in how diversity emerges from the population/individual level of diversity to the species level. We also find that dispersal ability, abundance, and speciation rate are all significantly related to patterns of discordance in diversity between the species and genetic levels of organisation. The patterns observed here align with expectations derived from the experimental approach in chapter 1, providing support for our conceptual and analytical framework of universal process across scale.

References

Albouy, C., Archambault, P., Appeltans, W., Araujo, M. B., Beauchesne, D., Cazelles, K., Cirtwill, A. R., Fortin, M. J., Galiana, N., Leroux, S. J., Pellissier, L., Poisot, T., Stouffer, D. B., Wood, S. A., & Gravel, D. (2019). The marine fish food web is globally connected [Article]. *Nature Ecology & Evolution*, *3*(8), 1153-+. https://doi.org/10.1038/s41559-019-0950-y

Antonovics, J. (1976). The input from population genetics: "The new ecological genetics". *Systematic Botany*, *1*(3), 233-245. https://doi.org/10.2307/2418718

Antonovics, J. (2003). Toward community genomics? *Ecology*, *84*(3), 598-601. https://doi.org/10.1890/0012-9658(2003)084[0598:Tcg]2.0.Co;2

Bailey, J. K., Hendry, A. P., Kinnison, M. T., Post, D. M., Palkovacs, E. P., Pelletier, F., Harmon, L. J., & Schweitzer, J. A. (2009). From genes to ecosystems: an emerging synthesis of eco-evolutionary dynamics. *New Phytologist*, *184*(4), 746-749. https://doi.org/https://doi.org/10.1111/j.1469-8137.2009.03081.x

Bellwood, D. R., Goatley, C. H., Cowman, P. F., & Bellwood, O. (2015). The evolution of fishes on coral reefs: fossils, phylogenies and functions. *Ecology of fishes on coral reefs*, 55-63.

Boettiger, C., Lang, D. T., & Wainwright, P. (2012). rfishbase: exploring, manipulating and visualizing FishBase data from R. *Journal of Fish Biology*, *81*(6), 2030-2039.

Boschman, L. M., Cassemiro, F. A., Carraro, L., de Vries, J., Altermatt, F., Hagen, O., Hoorn, C., & Pellissier, L. (2021). South American freshwater fish diversity shaped by Andean uplift since the Late Cretaceous. *BioRxiv*, 2021.2005. 2014.444133.

Brockhurst, M. A., Chapman, T., King, K. C., Mank, J. E., Paterson, S., & Hurst, G. D. D. (2014). Running with the Red Queen: the role of biotic conflicts in evolution. *Proceedings of the Royal Society B-Biological Sciences*, *281*(1797), Article 20141382. https://doi.org/10.1098/rspb.2014.1382

Charlesworth, D., Barton, N. H., & Charlesworth, B. (2017). The sources of adaptive variation. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1855), 20162864. https://doi.org/doi:10.1098/rspb.2016.2864

Chave, J. (2004). Neutral theory and community ecology [Review]. *Ecology Letters*, *7*(3), 241-253. https://doi.org/10.1111/j.1461-0248.2003.00566.x

Coulson, T., Benton, T., Lundberg, P., Dall, S., & Kendall, B. (2006). Putting evolutionary biology back in the ecological theatre: a demographic framework mapping genes to communities. *Evolutionary Ecology Research*, *8*(7), 1155-1171.

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 [Article]. *Journal of Evolutionary Biology*, *24*(12), 2543-2562. https://doi.org/10.1111/j.1420-9101.2011.02391.x

Cowman, P. F., & Bellwood, D. R. (2013). The historical biogeography of coral reef fishes: global patterns of origination and dispersal. *Journal of Biogeography*, *40*(2), 209-224.

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*(3), 621-631.

Darwin, C. (1859). On the origin of species. *published on*, *24*, 1.

Des Roches, S., Post, D. M., Turley, N. E., Bailey, J. K., Hendry, A. P., Kinnison, M. T., Schweitzer, J. A., & Palkovacs, E. P. (2018). The ecological importance of intraspecific variation. *Nature Ecology & Evolution*, *2*(1), 57-64. https://doi.org/10.1038/s41559-017-0402-5

Descartes, R., Anscombe, G. E. M., & Geach, P. T. (1954). Philosophical writings.

Dobzhansky, T. (1974). Chance and creativity in evolution. *Studies in the philosophy of biology: Reduction and related problems*, 307-338.

Donati, G. F. A., Zemp, N., Manel, S., Poirier, M., Claverie, T., Ferraton, F., Gaboriau, T., Govinden, R., Hagen, O., Ibrahim, S., Mouillot, D., Leblond, J., Julius, P., Velez, L., Zareer, I., Ziyad, A., Leprieur, F., Albouy, C., & Pellissier, L. (2021). Species ecology explains the spatial components of genetic diversity in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1959), 20211574. https://doi.org/doi:10.1098/rspb.2021.1574

Edgar, G. J., Cooper, A., Baker, S. C., Barker, W., Barrett, N. S., Becerro, M. A., Bates, A. E., Brock, D., Ceccarelli, D. M., Clausius, E., Davey, M., Davis, T. R., Day, P. B., Green, A., Griffiths, S. R., Hicks, J., Hinojosa, I. A., Jones, B. K., Kininmonth, S., . . . Stuart-Smith, R. D. (2020). Establishing the ecological basis for conservation of shallow marine life using Reef Life Survey. *Biological Conservation*, *252*, 108855. https://doi.org/https://doi.org/10.1016/j.biocon.2020.108855

Egerton, F. N. (2001). A History of the Ecological Sciences: Early Greek Origins. *Bulletin of the Ecological Society of America*, *82*(1), 93-97. http://www.jstor.org/stable/20168519

Egerton, F. N. (2013). History of Ecological Sciences, Part 47: Ernst Haeckel's Ecology. *The Bulletin of the Ecological Society of America*, *94*(3), 222-244. https://doi.org/https://doi.org/10.1890/0012-9623-94.3.222

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422-433. https://doi.org/10.1038/nrg.2016.58

Gaston, K. J., & Spicer, J. I. (2013). *Biodiversity: an introduction*. John Wiley & Sons.

Haeckel, E. (1866). *Generelle Morphologie der Organismen* (Vol. 2). Georg Reimer.

Hagen, O., Fluck, B., Fopp, F., Cabral, J. S., Hartig, F., Pontarp, M., Rangel, T. F., & Pellissier, L. (2021). gen3sis: A general engine for eco-evolutionary simulations of the processes that shape Earth's biodiversity [Article]. *PLOS Biology*, *19*(7), 31, Article e3001340. https://doi.org/10.1371/journal.pbio.3001340

Hagen, O., Skeels, A., Onstein, R. E., Jetz, W., & Pellissier, L. (2021). Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. *Proceedings of the National Academy of Sciences*, *118*(40), e2026347118. https://doi.org/doi:10.1073/pnas.2026347118

Haldane, J. B. S. (1927). *Possible worlds*. Routledge.

Hamilton, M. B. (2021). *Population genetics*. John Wiley & Sons.

Hancock, Z. B., Lehmberg, E. S., & Bradburd, G. S. (2021). Neo-darwinism still haunts evolutionary theory: A modern perspective on Charlesworth, Lande, and Slatkin (1982). *Evolution*, *75*(6), 1244-1255. https://doi.org/10.1111/evo.14268

Hatcher, B., Johannes, R., & Robinson, A. (1989). Review of the research relevant to the conservation of shallow tropical marine ecosystems. *Oceanography and Marine Biology*, *27*, 337-414.

Hibdige, S. G. S., Raimondeau, P., Christin, P. A., & Dunning, L. T. (2021). Widespread lateral gene transfer among grasses. *New Phytologist*, *230*(6), 2474-2486. https://doi.org/10.1111/nph.17328

Hubbell, S. (2001). The Unified Neutral Theory of Biodiversity and Biogeography. *Princeton University Press, Princeton, New Jersey, USA. Hughes, RG (1986) Theories and models of species abundance. American Naturalist*, *128*, 879-899.

Huston, M. A., & Huston, M. A. (1994). *Biological diversity: the coexistence of species*. Cambridge University Press.

Jørgensen, S. E., & Fath, B. (2014). *Encyclopedia of ecology*. Newnes.

Kahilainen, A., Puurtinen, M., & Kotiaho, J. S. (2014). Conservation implications of species-genetic diversity correlations [Review]. *Global Ecology and Conservation*, *2*, 315-323. https://doi.org/10.1016/j.gecco.2014.10.013

Kimura, M., & Weiss, G. H. (1964). The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance. *Genetics*, *49*(4), 561-576. https://doi.org/10.1093/genetics/49.4.561

Lamy, T., Jarne, P., Laroche, F., Pointier, J. P., Huth, G., Segard, A., & David, P. (2013). Variation in habitat connectivity generates positive correlations between species and genetic diversity in a metacommunity [Article]. *Molecular Ecology*, *22*(17), 4445-4456. https://doi.org/10.1111/mec.12399

Lamy, T., Laroche, F., David, P., Massol, F., & Jarne, P. (2017). The contribution of species-genetic diversity correlations to the understanding of community assembly rules. *Oikos*, *126*(6), 759-771. https://doi.org/10.1111/oik.03997

Laurin, M., & Arntzen, J. W. (2010). The subjective nature of Linnaean categories and its impact in evolutionary biology and biodiversity studies. *Contributions to Zoology*, *79*(4), 131-146. https://doi.org/https://doi.org/10.1163/18759866-07904001

Lawrence, E. R., & Fraser, D. J. (2020). Latitudinal biodiversity gradients at three levels: Linking species richness, population richness and genetic diversity [Review]. *Global Ecology and Biogeography*, *29*(5), 770-788. https://doi.org/10.1111/geb.13075

Leibold, M. A., Govaert, L., Loeuille, N., De Meester, L., & Urban, M. C. (2022). Evolution and Community Assembly Across Spatial Scales. *Annual Review of Ecology Evolution and Systematics*, *53*, 299-326. https://doi.org/10.1146/annurev-ecolsys-102220-024934

Leugger, F., Broquet, T., Karger, D. N., Rioux, D., Buzan, E., Corlatti, L., Crestanello, B., Curt-Grand-Gaudin, N., Hauffe, H. C., & Rolečková, B. (2022). Dispersal and habitat dynamics shape the genetic structure of the Northern chamois in the Alps. *Journal of Biogeography*, *49*(10), 1848-1861.

Lewens, T. (2019). The Extended Evolutionary Synthesis: what is the debate about, and what might success for the extenders look like? *Biological Journal of the Linnean Society*, *127*(4), 707-721. https://doi.org/10.1093/biolinnean/blz064

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*, *110*(41), 16498-16502. https://doi.org/doi:10.1073/pnas.1304074110
Lupyan, G. (2017). How Reliable Is Perception? *Philosophical Topics*, *45*(1), 81-106. https://www.jstor.org/stable/26529426
MacArthur, R. H., & Wilson, E. O. (2016). Princeton University Press. https://doi.org/doi:10.1515/9781400881376
Manel, S., Guerin, P. E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., & Pellissier, L. (2020). Global determinants of freshwater and marine fish genetic diversity. *Nature Communications*, *11*(1). https://doi.org/ARTN 69210.1038/s41467-020-14409-7 Mayr, E. (1963). Animal Species and Evolution (Belknap, Cambridge, MA).

McIntosh, R. P. (1986). *The background of ecology: concept and theory*. Cambridge University Press.
Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources*, *11*(1), 5-18. https://doi.org/https://doi.org/10.1111/j.1755-0998.2010.02927.x
Messmer, V., Jones, G. P., Munday, P. L., & Planes, S. (2012). Concordance between genetic and species diversity in coral reef fishes across the Pacific Ocean biodiversity gradient. *Evolution*, *66*(12), 3902-3917. https://doi.org/10.1111/j.1558-5646.2012.01725.x
Near, T. J., Dornburg, A., Eytan, R. I., Keck, B. P., Smith, W. L., Kuhn, K. L., Moore, J. A., Price, S. A., Burbrink, F. T., & Friedman, M. (2013). Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. *Proceedings of the National Academy of Sciences*, *110*(31), 12738-12743.
Nei, M., & Takahata, N. (1993). Effective population size, genetic diversity, and coalescence time in subdivided populations. *Journal of Molecular Evolution*, *37*(3), 240-244. https://doi.org/10.1007/BF00175500
Padial, J. M., & De la Riva, I. (2021). A paradigm shift in our view of species drives current trends in biological classification. *Biological Reviews*, *96*(2), 731-751. https://doi.org/https://doi.org/10.1111/brv.12676
Parravicini, V., Bender, M. G., Villéger, S., Leprieur, F., Pellissier, L., Donati, F. G. A., Floeter, S. R., Rezende, E. L., Mouillot, D., & Kulbicki, M. (2021). Coral reef fishes reveal strong divergence in the prevalence of traits along the global diversity gradient. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1961), 20211712. https://doi.org/doi:10.1098/rspb.2021.1712
Pelletier, F., Garant, D., & Hendry, A. P. (2009). Eco-evolutionary dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1523), 1483-1489. https://doi.org/doi:10.1098/rstb.2009.0027
Portin, P., & Wilkins, A. (2017). The evolving definition of the term “gene”. *Genetics*, *205*(4), 1353-1364.
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*(7714), 392-395. https://doi.org/10.1038/s41586-018-0273-1
Reisch, C., & Hartig, F. (2021). Species and genetic diversity patterns show different responses to land use intensity in central European grasslands. *Diversity and Distributions*, *27*(3), 392-401. https://doi.org/10.1111/ddi.13199
Renema, W., Bellwood, D. R., Braga, J. C., Bromfield, K., Hall, R., Johnson, K. G., Lunt, P., Meyer, C. P., McMonagle, L. B., Morley, R. J., O'Dea, A., Todd, J. A., Wesselingh, F. P., Wilson, M. E. J., & Pandolfi, J. M. (2008). Hopping hotspots: Global shifts in marine Biodiversity. *Science*, *321*(5889), 654-657. https://doi.org/10.1126/science.1155674
Sallan, L. C., & Friedman, M. (2012). Heads or tails: staged diversification in vertebrate evolutionary radiations. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1735), 2025-2032.

Schmidt, C., Munshi-South, J., Dray, S., & Garroway, C. J. (2022). Determinants of genetic diversity and species richness of North American amphibians. *Journal of Biogeography*, *49*(11), 2005-2015. https://doi.org/10.1111/jbi.14480

Segar, S. T., Fayle, T. M., Srivastava, D. S., Lewinsohn, T. M., Lewis, O. T., Novotny, V., Kitching, R. L., & Maunsell, S. C. (2020). The Role of Evolution in Shaping Ecological Networks. *Trends in Ecology & Evolution*, *35*(5), 454-466. https://doi.org/10.1016/j.tree.2020.01.004

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*(7), 573-583. https://doi.org/10.1641/b570707

Taberlet, P., Zimmermann, N. E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Coldea, G., Mirek, Z., Moilanen, A., Ahlmer, W., Marsan, P. A., Bona, E., Bovio, M., Choler, P., Cieslak, E., Colli, L., Cristea, V., Dalmas, J. P., . . . IntraBioDiv, C. (2012). Genetic diversity in widespread species is not congruent with species richness in alpine plant communities [Article]. *Ecology Letters*, *15*(12), 1439-1448. https://doi.org/10.1111/ele.12004

Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., Grenyer, R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Purschke, O., Redding, D. W., Rosauer, D. F., Winter, M., & Mazel, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews*, *92*(2), 698-715. https://doi.org/https://doi.org/10.1111/brv.12252

Vellend, M. (2005). Species diversity and genetic diversity: Parallel processes and correlated patterns. *American Naturalist*, *166*(2), 199-215. https://doi.org/10.1086/431318

Vellend, M. (2010). Conceptual synthesis in community ecology [Review]. *Quarterly Review of Biology*, *85*(2), 183-206. https://doi.org/10.1086/652373

Vellend, M., & Geber, M. A. (2005). Connections between species diversity and genetic diversity. *Ecology Letters*, *8*(7), 767-781. https://doi.org/10.1111/j.1461-0248.2005.00775.x

Vellend, M., Lajoie, G., Bourret, A., Murria, C., Kembel, S. W., & Garant, D. (2014). Drawing ecological inferences from coincident patterns of population- and community-level biodiversity. *Molecular Ecology*, *23*(12), 2890-2901. https://doi.org/10.1111/mec.12756

Vilcot, M., Albouy, C., Donati, G. F. A., Claverie, T., Julius, P., Manel, S., Pellissier, L., & Leprieur, F. (2023). Spatial genetic differentiation correlates with species assemblage turnover across tropical reef fish lineages. *Global Ecology and Biogeography*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/geb.13637

Waldock, C., Stuart-Smith, R. D., Edgar, G. J., Bird, T. J., & Bates, A. E. (2019). The shape of abundance distributions across temperature gradients in reef fishes [Article]. *Ecology Letters*, *22*(4), 685-696. https://doi.org/10.1111/ele.13222

Waples, R. S., & Gaggiotti, O. (2006). INVITED REVIEW: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, *15*(6), 1419-1439. https://doi.org/https://doi.org/10.1111/j.1365-294X.2006.02890.x

Ware, I. M., Fitzpatrick, C. R., Senthilnathan, A., Bayliss, S. L. J., Beals, K. K., Mueller, L. O., Summers, J. L., Wooliver, R. C., Van Nuland, M. E., Kinnison, M. T., Palkovacs, E. P., Schweitzer, J. A., & Bailey, J. K. (2019). Feedbacks link ecosystem ecology and evolution across spatial and temporal scales: Empirical evidence and future directions. *Functional Ecology*, *33*(1), 31-42. https://doi.org/https://doi.org/10.1111/1365-2435.13267

Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilee, R., Condamine, F. L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernandez-Palacios, J. M., Hengl, T., Norder, S. J., Rijsdijk, K. F., Sanmartin, I., Strasberg, D., Triantis, K. A., . . . Thebaud, C. (2015). Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters*, *18*(2), 200-217. https://doi.org/10.1111/ele.12398

Whitham, T. G., Bailey, J. K., Schweitzer, J. A., Shuster, S. M., Bangert, R. K., Leroy, C. J., Lonsdorf, E. V., Allan, G. J., DiFazio, S. P., Potts, B. M., Fischer, D. G., Gehring, C. A., Lindroth, R. L., Marks, J. C., Hart, S. C., Wimp, G. M., & Wooley, S. C. (2006). A framework for community and ecosystem genetics: from genes to ecosystems [Review]. *Nature Reviews Genetics*, *7*(7), 510-523. https://doi.org/10.1038/nrg1877

Wolf, J. B. W., Lindell, J., & Backström, N. (2010). Speciation genetics: current status and evolving approaches. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1547), 1717-1733. https://doi.org/doi:10.1098/rstb.2010.0023

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

Wright, S. (1949). The genetical structure of populations. *Annals of eugenics*, *15*(1), 323-354.

Yoshida, T., Jones, L. E., Ellner, S. P., Fussmann, G. F., & Hairston, N. G. (2003). Rapid evolution drives ecological dynamics in a predator–prey system. *Nature*, *424*(6946), 303-306. https://doi.org/10.1038/nature01767

Zirkle, C. (1941). Natural Selection before the "Origin of Species". *Proceedings of the American Philosophical Society*, *84*(1), 71-123. http://www.jstor.org/stable/984852

Chapter 1: Diversity across organisational scale emerges through dispersal ability and speciation dynamics in tropical fish.

Authors
Thomas Keggin[1,2], Conor Waldock[4,5], Alex Skeels[1,2], Oskar Hagen[6], Camille Albouy[1,2], Stéphanie Manel[3], Loïc Pellissier[1,2]

Author affiliations
1. Ecosystems and Landscape Evolution, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland
2. Unit of Land Change Science, Swiss Federal Research Institute WSL
3. CEFE, Univ. Montpellier, CNRS, EPHE- PSL University, Montpellier, France
4. Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland
5. Department of Fish Ecology and Evolution, Center for Ecology, Evolution and Biogeochemistry, Eawag - Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland
6. Evolution and Adaptation, German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Germany.

All work was conducted at 1 and 2.

Abstract
Biodiversity exists at different levels of organisation: e.g., individual, population, species, and community. These levels all exist within the same system, with diversity patterns emerging across organisational scale through overlapping processes. Despite this inherent interconnectivity, observational studies reveal that diversity patterns across levels are not always consistent. However, the underlying mechanisms for variable continuity in diversity across levels remain elusive. To investigate these mechanisms, we apply a spatially explicit simulation model to reconstruct the global diversification of tropical reef fishes at both the population and species levels. We find significant relationships between the population and species levels of diversity which vary dependent on both the measure of diversity and the spatial scale considered. In turn, these population-species relationships are driven by modelled biological trait parameters, especially the divergence threshold at which populations speciate. To explain these multi-level diversity patterns, we propose a simple, yet novel, population-to-species diversity partitioning mechanism which occurs through speciation. We predict that in real-world systems this mechanism is driven by the molecular dynamics that determine genetic incompatibility, and therefore reproductive isolation between individuals. We highlight the importance of considering diversity across all levels of organisation in furthering our understanding of the emergence of biodiversity across organisational scale.

Introduction
Biological diversity is measured as variation within and between different levels of organisation; from nucleotides, genes, individuals, populations, species, through to whole meta-communities (Mayr, 1963). The processes shaping diversity at these different organisational levels are often studied in isolation, however, it has been predicted that similar processes should operate across

levels to jointly shape these patterns. For example, processes such as drift and migration which cause populations to fragment and diverge from one another (Ellegren & Galtier, 2016; Hamilton, 2021; Wright, 1931; Wright, 1943) – increasing population level diversity – are also expected to be the same as those that eventually cause isolated populations to become new species (Seehausen et al., 2014). As such, measures of diversity are expected to be positively correlated between the population and species level (Vellend & Geber, 2005). Empirically, however, there is inconsistency in both the direction and the strength of the diversity relationships between the population and species levels of organisation, with some systems showing strong positive correlations (Blum et al., 2012; Manel et al., 2020; Messmer et al., 2012), others weak and negative correlations (Schmidt et al., 2022), and some remaining ambiguous (Taberlet et al., 2012). Furthermore, diversity can be measured in various ways, capturing different aspects of this variation (Tucker et al., 2017), and so we might also expect variation across diversity metrics and observational datasets. This mismatch between expectation and empirical observation underlies as-of-yet unidentified mechanisms driving the emergence of diversity across organisational scale.

Predictions for a positive relationship between diversity across the population and species levels comes from a minimal evolutionary framework which proposes four analogous processes which play out at both the species- and population-levels to shape diversity: drift, gene-flow, selection, and mutation/speciation (Antonovics, 1976; Hamilton, 2021; Vellend, 2010; Vellend & Geber, 2005). At the population level (i.e., considering variation between geographically isolated groups of individuals), drift is the random variance in allelic frequencies in populations over time, gene-flow is the exchange of alleles between populations as a consequence of migration, selection is the non-random retention of alleles as a consequence of environmental pressure, and mutation is the generation of new alleles (Hamilton, 2021). At the species level (i.e., considering variation between species within communities), drift is the stochastic fluctuation in species abundances and composition in communities over time (Hubbell, 2001; Nee, 2005), dispersal results in the exchange of species between communities, selection determines survivability of species in communities, and speciation results in the generation of new species (Antonovics, 1976; Chave, 2004; Vellend, 2010). Given that these processes are analogous at both levels of organisation, it is expected that patterns in diversity across these levels will be positively correlated. There is some support for this, for example, a positive covariation in microsatellite allelic- and species-richness across sites in a freshwater system was found (Blum et al., 2012). Yet whilst these processes are analogous, variation in the observed species- and population-level diversity relationships suggests that this framework is insufficient to completely explain observable biodiversity patterns.

Unexpected variation in observed species- and population-level diversity patterns may occur if biotic and abiotic factors operate with different strengths at different levels of biological organisation. In contrast to the predicted positive relationship between population- and species-level diversity, many studies have also found no, or negative, relationships (Kahilainen et al., 2014). For example, alpine plant communities across the Alps and Carpathians showed incongruence between genetic markers and species diversity indices (Taberlet et al., 2012). Whilst population richness, population-level genetic diversity, and species richness do not covary across a range of vertebrate species across in the North America (Lawrence, 2020). This decoupling of diversity between organisational levels has been explained by the environment and/or species biological traits, including dispersal ability, environmental niche width (Decocq et al., 2021), geographic range size (Lawrence, 2020), biotic interactions (Schmidt et al., 2022), and landscape features and dynamics (Reisch & Hartig, 2021; Reisch & Schmid, 2019). These factors have been proposed to modulate population- and species-diversity independently. For example, density dependent environmental feedbacks have been

suggested to lead to a mismatch (Schmidt et al., 2022). These mismatches likely occur because highly productive regions support many species, but due to the subsequent increase in inter-species competition effective population sizes are reduced – reducing genetic diversity (Connell et al., 1971; Janzen, 1970; Nei & Takahata, 1993). A similar effect is predicted by habitat patchiness, with high patchiness decreasing effective population size, but increasing species diversity (Jaquiery et al., 2011). These examples suggest that both the spatial and temporal scales in which processes occur differ across organisational scale leading to incongruent diversity patterns at different levels. However, we still lack mechanistic understanding of how different ecological or evolutionary processes act through populations and species to shape diversity pattern formation across organisational levels.

Commonly, relationships between population and species diversity are explored with correlative methods – which are limited only to describing patterns found in nature, but mechanistic models are a powerful approach to allow ecological and evolutionary processes to be explicitly simulated and emergent patterns to be validated against observational data to draw inferences about biological systems (Gotelli et al., 2009; Hagen, 2022; Leprieur et al., 2016; Pilowsky et al., 2022). For example, population- and species-level patterns have been explored mechanistically at local patch scales (Vellend, 2005), which found a neutral positive correlative expectation between organisational levels made variable by introducing selection. More recently, mechanistic models have included both deep-time evolutionary process and shallow-time ecological processes alongside broad-scale environmental information, integrating eco-evolutionary dynamics more completely with landscape dynamics (Descombes et al., 2018; Gaboriau et al., 2019; Leprieur et al., 2016; Pellissier et al., 2014). This approach offers the opportunity to explore various processes including drift, dispersal, mutation, and speciation across a dynamic landscape within a unified modelling framework.

In this study, we investigate the processes shaping patterns of diversity at and between the population and species levels at a global scale using a spatially explicit model of diversification on tropical reef-associated fishes. Reef fishes provide a suitable model system to study this continuity in diversity across organisational scale as they are highly diverse and have a wealth of spatial, phylogenetic, and trait information available. At the species level, tropical reef-associated fishes have spatially structured diversity patterns, with a centre of diversity in the Indo-Australian Archipelago that roughly follows a longitudinal negative gradient away from this major hotspot (Cowman & Bellwood, 2011; Hillebrand, 2004; Kinlock et al., 2018). Similarly, genetic diversity studies find spatial diversity patterns relate to seascape structure, barriers to dispersal, historical effects, and dispersal abilities (Eble et al., 2015). These population- and species-level diversity patterns have been investigated in this system showing mixed relationships. A positive relationship was observed between genetic diversity and species richness in tropical Pacific fishes (Messmer et al., 2012), and between species richness and nucleotide diversity across both freshwater and marine fishes (Manel et al., 2020). A positive relationship between the population and species levels was also found in the Western Indian Ocean, but only in pairwise comparisons between sites (β-diversity) and not at the local or global scales (α- or γ-diversity) (Vilcot et al., 2023) – indicating that diversity patterns across organisational scale are likely further dependent on the spatial partitioning considered.

To simulate the diversification of tropical reef associated fishes across organisational levels, we used biological traits and palaeogeological information over the last 200 million years corresponding to the early establishment of the Euteleost radiation. We implemented this in the spatially explicit eco-evolutionary simulation model, gen3sis (O. Hagen et al., 2021). We consider different measures of diversity (richness, phylogenetic diversity, and mean pairwise distance) and spatial partitioning: γ,

the global diversity generated within the system, and β, the diversity differences between geographically distinct regions. Through this, we aim to address the following questions:

1. What is the relationship between the population and species levels of diversity?
2. Is this relationship consistent across spatial partitioning (β- and γ-diversity), and amongst different measures of diversity (richness, phylogenetic diversity, and mean pairwise distance)?
3. Which biological processes amongst dispersal, evolutionary rate, competition, and speciation traits drive variation in population-species diversity relationships?
4. How do population-species diversity relationships relate to clade properties such as range size and endemism?

Methods

To model the diversification of tropical reef fishes, we used the mechanistic simulation model, gen3sis (O. Hagen et al., 2021). Gen3sis is configured with species objects with information down to the population level and runs over a spatially explicit landscape – which can be customised with both paleoenvironmental reconstructions and biological configurations.

Paleo-environmental reconstructions

As input, the gen3sis model requires both a physical landscape with which modelled species interact, and a distance matrix to determine the cost of dispersal across the landscape (O. Hagen et al., 2021). The landscape consists of marine bathymetry and sea surface temperature at a 1x1° resolution at 166.7 ka time steps back to 200 Ma (Scotese, 2021; Scotese et al., 2021). The extent of the input data is global, but habitable cells are restricted to those above a mean temperature of 17°C and shallower than 2000m. These cut offs were chosen based on modelled thermal ranges of extant coral reef fishes (Waldock et al., 2019) and then visually matched with current coral reef distributions (UNEP-WCMC, 2010). The distance matrices allow free movement in all marine cells, and no movement across terrestrial cells.

Bathymetry was derived from an elevation model based on a mixture of plate tectonic modelling and geological evidence, described in detail by Scotese (2021). To match the model time steps here, these existing time steps were temporally interpolated using a linear function. Cells above sea level were removed. Temperature data are derived from a model based on oxygen isotope information, lithologic indicators, and the bio-geological record described in Scotese et al. (2021). As published, these data describe average tropical temperature change from the present (delta temperature) in 5 Ma time intervals into the past. These values are then modified geographically based on reconstructed climatic bands (paleo-Köppen belts). To generate one degree resolution sea surface temperature estimates, the boundaries of the climatic belts were first smoothed using the focal function in the R raster package (Hijmans et al., 2015) using a focal window of 81 cells. Boundary values for the north and south poles where the focal window exceeded the limits of the global extent were set to -20 °C, matching the temperature values of the polar climate bands. From these smoothed 5 Ma intervals, smoothed spatial climate distributions were generated for each 166.7 ka time step using linear interpolation. Further, delta temperature values were calculated for each time step by linearly interpolating the 1 Ma interval values provided by Scotese et al. (2021) and applied to the new geographically smoothed time steps. Finally, corrections are made to account for climatic fluctuations associated with recent glacial maxima (Annan & Hargreaves, 2013). Cost distance values between habitable cells in the reconstructed landscapes were calculated using the *transition* function in the gdistance package in R (van Etten, 2017). The shortest path between each pair of cells was calculated and the distance between all pairs stored in a distance matrix. Paths were

calculated using an 8-direction adjacency scheme whereby cells are deemed adjacent if they are in contact vertically, horizontally, or diagonally. Each cell is also given a conductance value representing ease of travel across that cell. All marine cells were given a value of 1 (passable), whilst terrestrial cells were given a value of 0 (impassable).

Biological configuration

The species object within our gen3sis model contains the values for each species' traits, abundance, and cell-to-cell differentiation across all inhabited cells. The species traits include a thermal optimum, a competitive niche value, and a niche width determining the competitive range of a population; these are summarised in Table 2. Each simulation was seeded with a single species occupying all habitable cells in the first time-step with the trait values described above and run with the following functions at each time step. The speciation threshold parameter represents allopatric speciation and is simulated through the use of divergence between geographically distinct adjacent cell clusters within a species. Geographic cells that experience no dispersal between them in a time step will increase their pairwise divergence by 1. Cells that experience dispersal will decrease their divergence by 1. If all the divergence values between two cells exceed the speciation threshold, then a new species will form.

Each time step, for every pair of inhabited to habitable cells, a potential dispersal event is calculated. The dispersal distance parameter is drawn from a Weibull distribution; if the dispersal distance exceeds the geographical distance between cells, the dispersal attempt is successful. On a successful dispersal attempt, if the target cell is already occupied, then the pairwise divergence value between those two cells is reduced, simulating gene-flow. If the target cell is unoccupied by that species, a colonisation event occurs. In the case of colonisation, the starting abundance is reduced to the initial abundance parameter value, allowing for incumbency effects.

Every time step, the competitive niche and thermal optimum of each species is subject to change. Firstly, the traits are modified by the addition of a random value drawn from a Gaussian distribution of mean 0 and a standard deviation that varies between simulations, but is common between traits. Once the traits of each species in each cell have been modified, traits of geographically adjacent clusters of cells within species are homogenised by assigning the mean trait values. The ecology function determines the abundance values (0–1) of each species within each cell. This is done through a simulation of temperature tolerance and competition. At the start of each time step, the abundance value is at the maximum of 1. It is then reduced based on the difference between the environmental temperature and the thermal optimum of the population. The reduction is proportional to the magnitude of the probability density of a Gaussian distribution function with a mean equal to the environmental temperature value and a standard deviation of 2 °C. Once the abundances of the species within a cell have been adjusted by temperature, interspecies competition is carried out. Each species has a competitive niche value between 0 and 1, representing an abstract competitive space. They also have a competitive width value which determines the amount of that competitive space on either side of the niche value in which that species competes, e.g., if one species has a niche value of 0.3 and another with 0.4, and the competitive width is 0.2, then those two species will experience competition with one another. Species with overlapping niches will compete proportional to both their respective abundances and the size of the overlap. I.e., a species with a high abundance will exert a greater competitive pressure than a species with a low abundance. Abundances are then also further reduced by the proportion of their competitive space that exceeds the 0-1 bounds. Finally, species whose abundances have been reduced to a value less than 0.1 are reduced to 0, causing local extinction in that cell.

Through modifying these parameters, we explored the impact of biological traits on the relationship between the species and population levels of organisation. This was done through varying the parameters summarised in Table 2 using tropical reef-fish values taken from the literature. We generated 3000 unique parameter combinations using the quasi-random Sobol sequence number generation approach (Prowse et al., 2016). Each set of parameters feeds into one simulation. We removed simulations with fewer than 20 extant species as the patterns generated with too few species lack discriminatory power. We compared the remaining simulations to real-world observed patterns of species richness aggregated to a 1-degree resolution (Albouy et al., 2019). The richness was therefore summed across all simulations which was then normalised, along with observed richness, between 0 and 1 to be comparable.

Calculation of clade properties

Conceptually, we considered each simulation as representing a clade of fish with differing biological traits for which clade characteristics can be defined. These characteristics were calculated from the species object trait values and are summarised in Table 3. Our analyses comprise metrics at only the species level, only the population level, and at both levels.

At the species level we calculated the species richness per cell, the total extant and extinct species across all time steps, species range size, temporal species turnover,

$$Temporal\ species\ turnover = \frac{Extant\ species}{Extant\ species\ +\ Extinct\ species} \tag{1}$$

and weighted endemism (Crisp et al., 2001) per cell. Throughout the simulation, gen3sis calculates a species phylogeny based on pairwise species divergence times. From this species phylogeny we calculated Faith's phylogenetic diversity estimated as the total branch length within a phylogeny (Faith, 1992); and mean pairwise distance between species as the mean distance between pairs of objects within a phylogeny (Webb, 2000). We calculated the diversification rate from the simulated phylogeny as the inverse of the evolutionary distinctiveness following the fair proportions framework (Isaac et al., 2007; Jetz et al., 2012; Redding & Mooers, 2006). As measures of functional trait diversity, we calculated the mean, maximum, minimum, range, evenness (Mouillot et al., 2005), and diversity (Leps et al., 2006) of the thermal and competitive niche traits using in-house functions in R (R Core Team, 2022).

At the population level, we calculated the total number of geographic cell clusters per simulation (Supplementary Figure 4) across all species as well as the phylogenetic diversity (PD) and mean pairwise distance (MPD). To calculate PD and MPD at the population level, the divergence values between inhabited cells within each species was taken and aggregated into geographic clusters. The mean divergence value between each cluster is then calculated and decomposed into a cluster-to-cluster divergence matrix. A phylogeny object from this cluster divergence matrix was calculated using a hierarchical clustering approach implemented by *hclust* in the R stats package (R Core Team, 2022). From this cluster phylogeny, phylogenetic diversity is calculated using the *pd* function in the phylomeasures R package (Tsirogiannis & Sandel, 2016). The mean value from each simulation was then taken to make values comparable to the species level phylogeny. Similarly, mean pairwise distance was calculated as the mean pairwise distance between these geographic clusters of cells.

Table 2: Summary of simulation parameters

Parameter	***Description***	***Parameter space***
Initial abundance	When a new cell is colonised, it is seeded with an initial abundance (whereafter the abundance returns to 1 with each time step).	**0.11 – 1**. From the minimum value before extinction to full abundance on colonisation.
Thermal optimum	The thermal optimum of the root species at the start of the simulation was varied across the entire temperature range present in all habitable cells across the entire simulation.	**17 - 31.4 °C**. Values from (Waldock et al., 2019).
Dispersal distance	The distance a species can disperse from cell-to-cell at each time step. This determines inter-population connectivity and colonisation events. These values are taken from a Weibull distribution approximating the probability distribution of dispersal events.	The scale of the Weibull dispersal kernel was varied from **100 to 5000 km** based on long term movement observations reported by (Green et al., 2015) for non-pelagic coral reef fishes. The shape was set to 2.5.
Speciation threshold	The divergence threshold at which two populations will speciate.	**12 – 600 timesteps**, equivalent to between **20 ka and 1 ma**. The divergence required for two populations to allopatrically speciate is complex (Seehausen et al., 2014). Here, we simply explore as wide a range of values as possible.
Mutation rate	The standard deviation of the normal distribution around the thermal and competitive nice traits from which new trait values are picked at each time step.	**0.01 to 0.15**. These values were based on estimation based on preliminary pilot simulations.
Competitive niche width	The amount of competitive space around the competitive niche trait value within which other species will compete.	**0.02 to 0.50**. The competitive niche width was varied from 0.02 to 0.50 based on preliminary simulations.

We focus on three different measures of diversity: richness, phylogenetic diversity (PD), and mean pairwise distance (MPD). Despite these metrics being conceptually related and occasionally correlated (Tucker & Cadotte, 2013), they capture different aspects of biological diversity (Tucker et al., 2017). The relationship between the species and population levels of these diversity metrics, or the continuity across levels, was calculated. This was done by first normalising the constituent metrics across simulations (species richness/PD/MPD, cluster richness/PD/MPD) to between 0 and 1, making metrics relative measures across organisational levels. The species level metrics were then divided by their corresponding cluster level metrics, e.g., species richness / cluster richness. These values were then log-transformed, giving positive values where species diversity was relatively higher than cluster diversity and negative values where it was lower. Formalised, this metric of continuity across levels was calculated as,

$$Continuity = \log\left(\frac{species\ diversity}{population\ diversity}\right) \tag{2}$$

This total diversity across simulations we defined as γ-diversity. To allow a β-diversity metric in our analyses, we divided the habitable cells in the model into bioregions, defined as realms by Spalding et al. (2007); Central Indo-Pacific, Eastern Indo-Pacific, Tropical Atlantic, Tropical Eastern Pacific, and Western Indo-Pacific (Supplementary Figure 3). Once subset into these bioregions, all diversity metrics described above were also calculated for each bioregion. β-diversity values are then the mean Euclidean distances between the continuity values amongst all pairs of bioregions.

Exploration of continuity patterns

We compared the relationship between the species and population levels of diversity in our simulations across the three facets of diversity: richness, phylogenetic diversity, and mean pairwise distance. For each facet comparison, a simple linear model was fit using the *lm* function in the R stats package (R Core Team, 2022). The models' normal distribution assumption was satisfied using a log transformation for all diversity measures, except for species MPD. These continuity relationships were then investigated in light of biological parameter values: initial abundance, thermal optimum, dispersal distance, speciation threshold, mutation rate, and competitive niche width. For the continuity metrics of γ- and β-diversity, we fitted multiple linear regression models using the biological parameter values as predictors. These model variables were then reduced using a forward and backward stepwise model selection based on AIC scores using the *step* function in the R stats package (R Core Team, 2022). Finally, we correlated the continuity metrics to the calculated clade properties: species range, thermal and niche trait evenness, weighted endemism, species turnover, and diversification rate. This was done with the Hmisc package in R (Harrell Jr, 2019) using Spearman's Rank Correlation Coefficient to capture non-linear relationships between variables. P-values were Bonferroni corrected for multiple testing. This was visualised using a scaled PCA implemented in the R stats package (R Core Team, 2022).

Results

We varied model parameter values across model simulation runs, with each model simulation conceptually considered to be one clade of fish with the parameters. These parameters define the clade's biological traits and properties, and our simulations reproduced variation in diversity across these. From 3000 simulations, 316 were retained that contained 20 or more extant species (median = 51). There was a wide range of diversity values at both the species and population levels; in richness (species, 20-1837; population, 1-101), Faith's phylogenetic diversity (species, 3412-661201, population, 4-1145), and mean pairwise distance (species, 761-2260, population, 2-232). This

Table 3: Summary of metrics

Level	*Metric*	*Description*
Species	Surviving species	The total number of extant species within a simulation.
	Species phylogenetic diversity	The total branch length in the phylogeny object, calculated using the the phylomeasures R package (Faith, 1992; Tsirogiannis & Sandel, 2016).
	Species mean pairwise distance	The mean pairwise distance between extant species in the phylogeny object, calculated using the Phylomeasures package (Tsirogiannis & Sandel, 2016; Webb, 2000).
	Total species	The total number of extinct and extant species within a simulation.
	Species range	The mean number of occupied cells for all extant species.
	Species turnover	The number of extant species over the sum of extant and extinct species.
	Species richness	The mean simulation species richness per cell.
	Diversification rate	Calculated from the simulation phylogeny as the reciprocal of the evolutionary distinctiveness (Jetz et al., 2012). Evolutionary distinctiveness was calculated using the evol_distinct() function in the phyloregion R package (Daru et al., 2020) following the fair proportions framework described by (Isaac et al., 2007).
	Weighted endemism	Weighted endemism for each cell was calculated as the number of species occupying that cell divided by the total ranges of those occupying species (Crisp et al., 2001). From this, the mean was taken.
Population	Total clusters	The total number of extant clusters of adjacent inhabited cells within all species in the simulation.
	Cluster phylogenetic diversity	Faith's phylogenetic diversity (Faith, 1992) calculated from population the population phylogeny.
	Cluster mean pairwise distance	The mean pairwise distance between populations in the population phylogeny.
Both	Continuity	The log-value of species diversity divided by the population diversity.
	Thermal traits	The mean, maximum, minimum, and range, evenness (Leps et al., 2006), and diversity (Mouillot et al., 2005).
	Competitive niche	The mean, maximum, minimum, and range, evenness (Leps et al., 2006), and diversity (Mouillot et al., 2005).

variation was also true of diversity values across geographic regions and in clade properties such as species turnover and diversification rate both globally and regionally (Supplementary Table 1).

Continuity across facets of diversity

In all three diversity metrics we found a negative relationship between γ-diversity at the population and species levels with effect sizes being greatest in mean pairwise distance (MPD), then phylogenetic diversity (PD), and finally richness which was not significant (richness, $\beta = -0.10$, $t = -1.8$, $p = 0.07$; PD, $\beta = -0.11$, $t = -2.4$, $p < 0.05$; MPD, $\beta = -0.12$, $t = -10.9$, $p < 0.01$; Figure 2). In the majority of simulations, MPD values were relatively higher at the species level than at the population level, whilst richness and PD had a similar distribution of relative values at both the population and species levels (Supplementary Figure 2). For measures of β-diversity, we found a positive relationship between the species and population levels (richness, $\beta = 0.39$, $t = 5.9$, $p < 0.01$; PD, $\beta = 0.31$, $t = 7.0$, $p < 0.01$; MPD, $\beta = 0.13$, $t = 2.7$, $p < 0.01$; Figure 2). An increase in the difference between regions at the population level was associated with an increase in the difference between regions at the species level, with the strongest relationship occurring with the richness metric, then PD, followed by MPD.

The impact of biological parameters on continuity

Continuity metrics of all three aspects of γ-diversity were significantly associated with biological parameters: richness (Adj. $R^2 = 0.42$, $F = 59.0$, $p < 0.001$), PD (Adj. $R^2 = 0.68$, $F = 137.1$, $p < 0.001$), and MPD (Adj. $R^2 = 0.79$, $F = 292.5$, $p < 0.001$). For each parameter, a positive coefficient indicates that increasing a parameter increases the amount of species diversity relative to population diversity. Conversely, a negative coefficient indicates that increasing a parameter value increases the amount of population diversity relative to species diversity. The speciation threshold parameter had a consistently strong negative relationship across all three diversity continuity metrics (richness, $\beta = -1.0$, $t = -12.7$, $p < 0.001$; PD, $\beta = -1.4$, $t = -24.4$, $p < 0.001$; MPD, $\beta = -1.3$, $t = -28.4$, $p < 0.001$; Figure 3). The parameters dispersal range, speciation threshold, and competitive niche size had a negative relationship with richness continuity (Supplementary Table 2; Figure 3), whilst initial colonisation abundance had positive relationships (richness, $\beta = 0.3$, $t = 5.8$, $p < 0.001$; PD, $\beta = 0.2$, $t = 4.6$, $p < 0.001$). The parameters dispersal range, speciation threshold, competitive niche size, and thermal optimum had a negative relationship with PD continuity (Supplementary Table 2; Figure 3), whilst the initial colonisation abundance had positive relationships. Speciation threshold, dispersal range, and starting thermal optimum (Supplementary Table 2; Figure 3) were negatively related to MPD continuity. Trait mutation rate was found to not be significantly associated with each of the three measures of continuity in diversity and was removed from all the models in the stepwise variable selection.

Association of continuity with clade properties

The relationships between the clade properties and each continuity metric were evaluated with pairwise Spearman's rank correlations and visualised with a principal components analysis for each facet of diversity. For richness, increasing thermal evenness ($r(314) = -0.53$, $p < 0.001$) and competitive evenness ($r(314) = -0.55$, $p < 0.001$), and species turnover ($r(314) = -0.32$, $p < 0.001$) were correlated with increasing population diversity relative to species diversity. The converse was true for thermal diversity ($r(314) = 0.22$, $p < 0.01$) and competitive diversity ($r(314) = 0.47$, $p < 0.001$) which was associated with an increase in species diversity relative to population diversity. These patterns were the same for both phylogenetic diversity and mean pairwise distance, except for the association with thermal niche diversity which was non-significant (Supplementary Table 3). There

were no significant relationships between continuity across levels and species range, weighted endemism, and diversification rate (Supplementary Table 3).

In the PCAs for all three diversity metrics, the first component accounted for between 38 – 40% of the variance, whilst the second component accounted for between 23 – 25% of the variance. For richness, the first component was contributed to mostly by competitive evenness and both thermal and competitive diversity (Supplementary Table 4), whilst the second component was mostly driven by species range, thermal evenness, species turnover and weighted endemism (Supplementary Table 4). For phylogenetic diversity, trait evenness and diversity contributed most to the first component (Supplementary Table 4), whilst the first component for mean pairwise distance was mostly contributed to by competitive evenness (Supplementary Table 4).

Discussion

We used a spatially explicit simulation model to investigate the drivers of continuity across the species and population levels of diversity. We show that the strength and direction of this relationship is variable and dependent on the metrics considered. These results help our understanding of the widely different and sometimes contradictory patterns found in empirical data which are based on various metrics and spatial scales (Manel et al., 2020; Schmidt et al., 2022). In particular, we found a negative relationship between population and species diversity in γ-diversity metrics (total diversity at the population and species levels). This was most heavily influenced by the speciation threshold, i.e., the divergence threshold that triggers speciation, which determines the frequency of diversity partitioning from the population level to the species level. Conversely, we found that the population-species diversity relationship was positively correlated for β-diversity metrics, demonstrating how geographic structuring plays out similarly across organisational scale. Finally, we describe the association between organisational continuity in diversity clade traits which connects trait-based functional diversity measures (Mason et al., 2005; Mouillot et al., 2005) to the emergence of contrasting diversity patterns across scale (Manel et al., 2020; Pfeiffer et al., 2018; Reisch & Hartig, 2021; Reisch & Schmid, 2019; Schmidt et al., 2022).

We uncover how population and species diversity are not necessarily positively related (Kahilainen et al., 2014) and can even show negative relationships (Schmidt et al., 2022), despite expectations of both levels of organisation being driven by the same fundamental processes (Antonovics, 1976; Vellend, 2005). In the simulated data, when considering the total global diversity (i.e., γ-diversity), we found negative relationships across three diversity measures: richness, phylogenetic diversity (PD), and mean pairwise distance (MPD), the latter two of which were significant. This negative relationship was mostly strongly explainable by the speciation threshold parameter, which controls the time it takes for diverging populations to become reproductively isolated and speciate. We infer that this negative relationship between species- and population-level diversity is the consequence of a partitioning effect of the total diversity across the two levels of organisation (Figure 1). In the simulation model, population-level diversity arises as populations migrate to new areas and eventually become isolated through environmental change. Eventually, isolated populations become new species at a rate modulated by the speciation threshold. Speciation does not remove diversity from the system, rather the diversity which was formerly between populations becomes diversity between species. As such, diversity has been directly transferred from the population level to the species level, decreasing the diversity at one level whilst increasing it at the other. This partitioning of diversity from populations to species could explain the negative correlation between population- and species-levels of diversity. This is supported by the strong negative relationship (the higher the speciation threshold, the more population diversity there is relative to species diversity) we find in

our simulations between the speciation threshold and continuity in all three diversity metrics. Here, the time required for speciation to occur controls the rate at which diversity is partitioned between levels, with a shorter speciation threshold leading to a faster rate of partitioning. This model parameter is a proxy for several real-world interacting genomic processes which shape the accumulation of incompatibilities and eventual speciation of populations (Feder et al., 2012; Ravinet et al., 2017; Seehausen et al., 2014). The amount of differentiation required for speciation to occur is a suite of compounding incompatibilities that accrue over time (Seehausen et al., 2014). The rate in absolute time at which these phenotypic incompatibilities accrue is determined by various traits such as generation time (Potts, 1984), background mutation rate (Nei et al., 1983), genomic architecture (Ellegren, 2014), and the complexity of life history traits (Bromham, 2011; Gavrilets et al., 2000; Martin et al., 2019; Palumbi, 1994) which are all inherited biological characteristics that vary across lineages (Seehausen et al., 2014; Singhal et al., 2018). To summarise, the most important determinant of continuity in diversity across biological organisation in our model was the speciation threshold – the amount of divergence between populations required for speciation to occur. This threshold corresponds to the genomic mechanisms that determine how quickly reproductive isolation can become established under similar levels of isolation. The correspondence between the model and observed process highlights a fundamental mechanism at play likely shaping the emergence of diversity across levels of biological organisation.

The continuity between the population and species levels of diversity depended on the measure of biodiversity used (i.e., species, richness, phylogenetic, and MPD). As such, ignoring the multifaceted nature of diversity may overlook how common evolutionary mechanisms drive variation amongst biological levels of organisation (Tucker et al., 2017). As a metric, MPD is skewed heavily towards the species level, with simulated clades typically having more divergence at the species-level relative to divergence at the population-level (Supplementary Figure 2). The cause of this species-level skew in MPD, rather than PD, is likely driven by fundamental differences between populations and species in each measure of diversity and the aspect of diversity each metric is measuring. Phylogenetic diversity is a sum of the total branch length in a phylogeny and is heavily influenced by the number of objects present in the system (i.e., richness in populations or species; (Tucker et al., 2017)), whilst the mean pairwise distance controls for this effect by averaging the number of objects and representing only the distances between them. This difference plays out in the partitioning of diversity between levels through speciation, and homogenisation of populations through gene flow. Specifically, regarding PD, migration between populations shares alleles (Wright, 1943) and migration between communities shares species (MacArthur & Wilson, 2016), homogenising the number of units present (richness) at both levels. For MPD on the other hand, some processes that decrease diversity at the population-level do not have a similar effect at the species-level. Migration between populations homogenises them through gene-flow which slows divergence and therefore decreases MPD values – whilst at the species level, migration between communities does not decrease species-species divergence (except for instances of introgression and horizontal gene transfer which are not explored in our model; Hibdige et al., 2021; Payseur & Rieseberg, 2016; Figure 1). Additionally reflected in MPD, highly divergent populations eventually become new species - removing them from the population level as they are partitioned into species-level objects through speciation. Whilst there is some evidence that evolutionarily distinct clades may be at higher risk of extinction (Dinnage et al., 2020), which may selectively remove highly divergent branches from the species-level phylogeny, whether this is widespread is unclear. This lack of removal of high divergence values between species allows species-level MPD to increase uninhibited. The result is

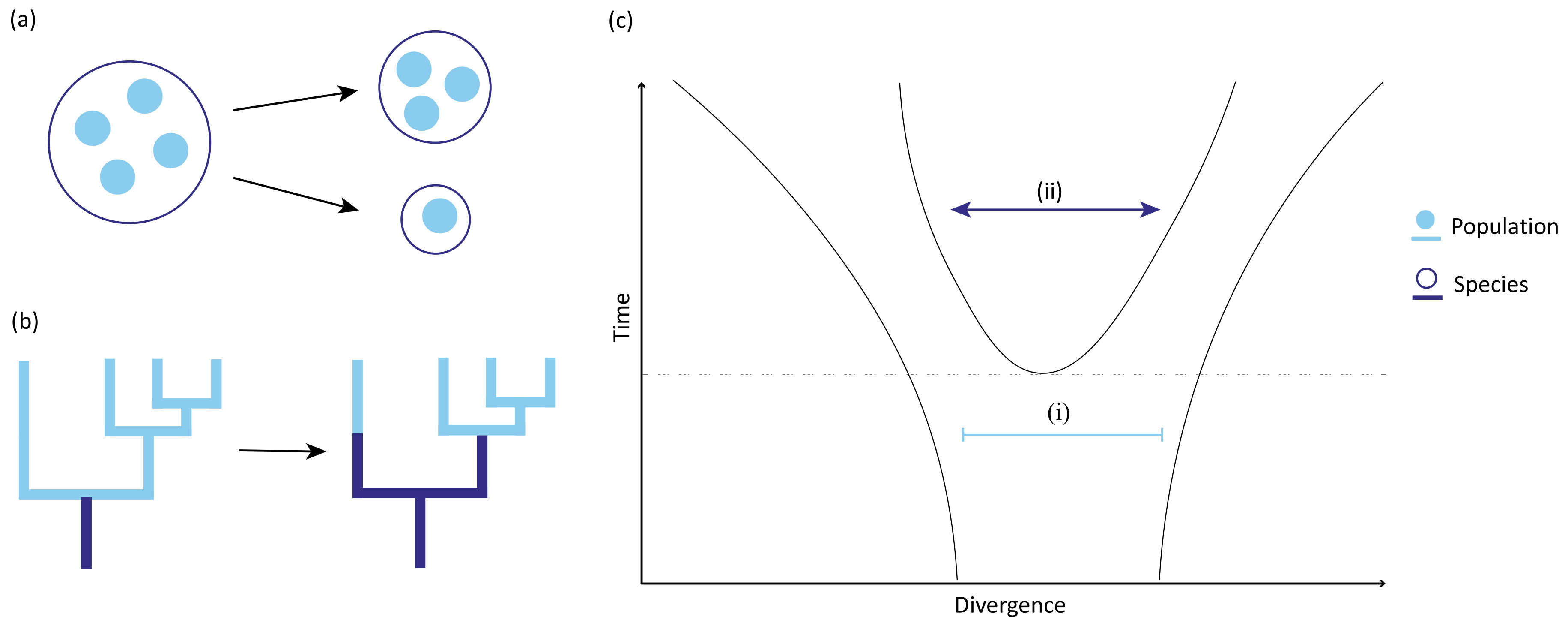


Figure 1: Conceptual diagram of the partitioning of diversity through speciation and the accumulation of divergence across levels of organisation. The species and population levels of organisation are represented by dark and light blue, respectively. (a) For richness, the total number of populations stays the same when speciation occurs, but another species is added to the system – 1 species and 4 populations becomes 2 species and 4 populations. This creates an uneven increase in the number of objects at each level. (b) Phylogenetic trees constructed from population objects are nested within species phylogenies. For phylogenetic diversity and mean pairwise distance, the phylogenetic tree topography, and therefore total diversity, is conserved throughout the speciation process. However, some of the population level diversity is partitioned from the population level to the species level. Since speciation does not add or remove total diversity from the system, but rather transfers it directly from one to another, this dynamic drives a negative relationship between the species and population levels of diversity. (c) For divergence, (i) at the population level the upper limit is determined by the speciation threshold and divergence is slowed by gene-flow. (ii) at the species level there is no upper limit and few brakes inhibiting divergence between species over time. The dotted horizontal line represents a speciation event.

two-fold: at the population level, divergence is both capped by the speciation threshold and slowed by gene-flow; whilst at the species level, divergence has few brakes and is limited only by increasing extinction probability over time, such that divergence values are limited to the sets,

$$\nabla\text{species} = \{\rho \leq \nabla\text{species} < \infty\} \quad (3)$$

$$\nabla\text{population} = \{0 \leq \nabla\text{population} < \rho\} \quad (4)$$

where ∇ denotes divergence and ρ the speciation threshold (Figure 1). These differing processes across measures of diversity highlight an important consideration in our study of continuity of diversity across organisational scale – we must be careful when comparing different organisational levels to ensure what we are measuring is actually comparable, and be mindful that different metrics behave and interact in interestingly different ways across organisational scale.

Considering the geographic segregation of diversity (i.e., β-diversity) can highlight distinct patterns at both the population (Donati et al., 2021) and species levels (Whittaker, 1960) that differ to patterns in total (γ) diversity. In our simulations the β-diversity metrics do not follow the same pattern as the γ-diversity metrics, with β-diversity values at the population and species levels showing a positive relationship (Figure 2). The simulated positive relationships reflect those found in a tropical reef fish system showed corresponding patterns in genetic and species β-diversity in the Western Indian Ocean (Vilcot, in press). Similarly, a stronger relationship between genetic and species rarity was found in alpine plants (Taberlet et al., 2012), which account more for the geographic distribution of species and alleles than total diversity metrics. The cause is likely due to β-diversity being a measure of segregation of diversity across sites (Tuomisto, 2010; Whittaker, 1960) and is therefore scaled for the absolute diversity in the system. Since β-diversity is a relative measure of diversity segregation, the partitioning effect of speciation is reduced. This allows parallel processes across levels such as drift to generate diversity patterns unimpeded. Through this, our simulation results support the important role of diversity measurement in explaining seemingly contradictory relationships between the population and species levels of organisation in empirical studies of these dynamics.

Biological traits modulate the eco-evolutionary processes that should in turn influence diversification across organisational scale (Lawrence, 2020; Lawrence & Fraser, 2020; Schmidt et al., 2022; Stewart et al., 2016). Dispersal range impacted continuity across all three facets of diversity (richness, phylogenetic diversity, and mean pairwise distance) with higher values driving more diversity at the population level relative to the species level. This is expected in an allopatric speciation model as higher dispersal increases range connectivity in a finite geographic space providing fewer opportunities for inter-population divergence to occur (O. Hagen et al., 2021; Oskar Hagen et al., 2021). Further, we explored how diversification across scale related to emergent clade properties. For example, high temporal species turnover is correlated with increasing population level to species level diversity. This pattern relates to the idea that, unlike population diversity, species diversity is theoretically uncapped (Harmon & Harrison, 2015; Equations 3,4) – apart from the age of the of the simulation (or perhaps even real systems), there are no hard limits to the maximum divergence between species, nor the complexity of their relationships (Rabosky, 2020). In finite, bounded real-world systems however, this may not be the case as limiters to species richness are well documented (Fine, 2015; Leprieur et al., 2016; Mihaljevic et al., 2017; Rabosky et al., 2018; Rabosky & Hurlbert, 2015). Our interpretation of the patterns found here is that extinction dynamics likely impact populations and species differently. The difference being the absolute values of diversity at each scale – relatively, diversity takes longer to accrue to the maximum at the species level than the population level (Equations 3,4). It follows that relative diversity at the population level is much

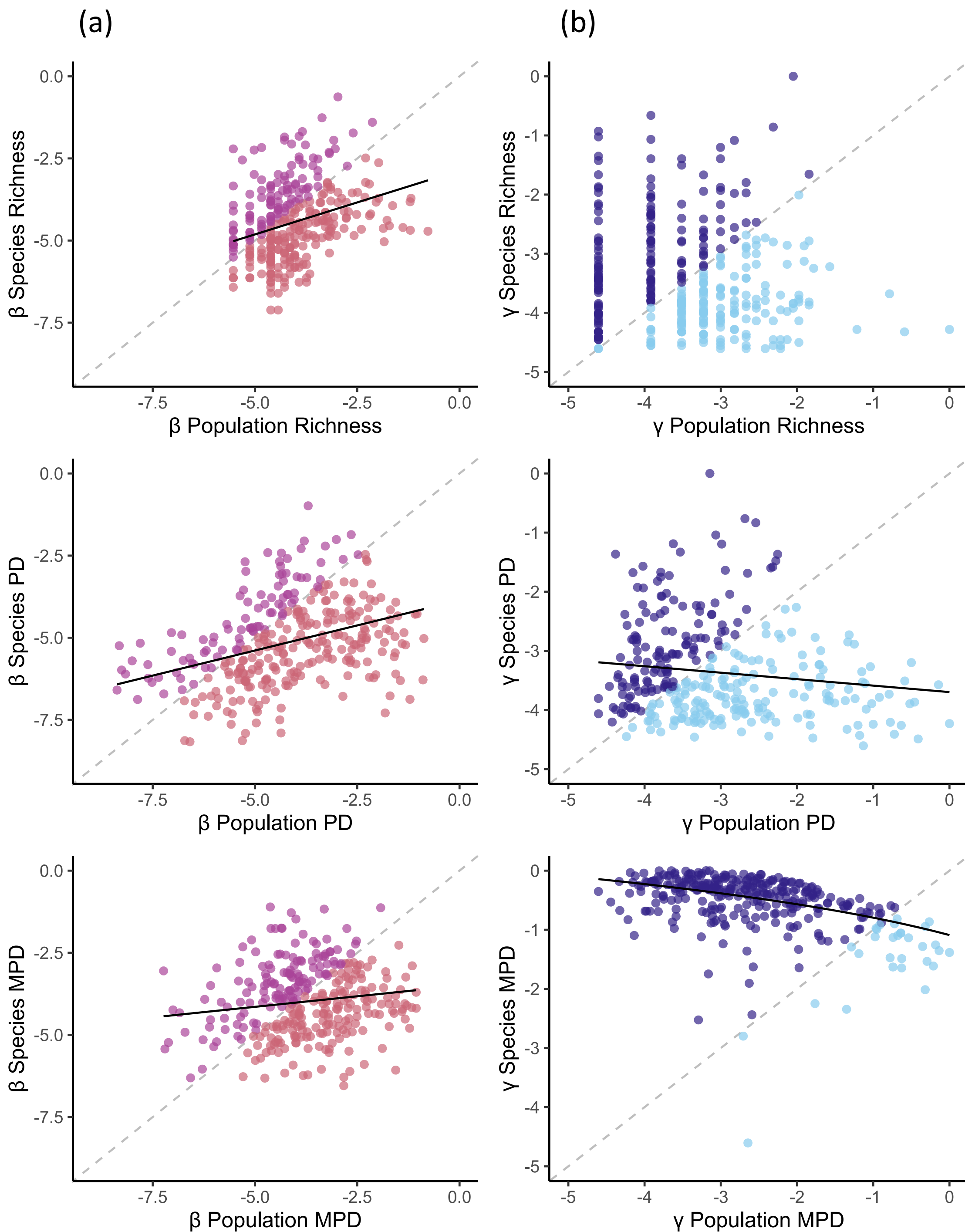


Figure 2: Simulated β- and γ-diversity relationships between the population and species levels of organisation across three measures of diversity; richness, phylogenetic diversity (PD), and mean pairwise distance (MPD). The grey dashed line represents the expected positive relationship between the two levels, whilst the black solid lines represent the simulated relationship found through a significant ($p < 0.05$) simple linear regression. (a) All β-diversity relationships are positive, and (b) all γ-diversity relationships are negative. Dark colours represent higher relative species diversity and lighter colours represent higher relative population diversity. All diversity measures have been log-transformed for both the regressions and visualisation.

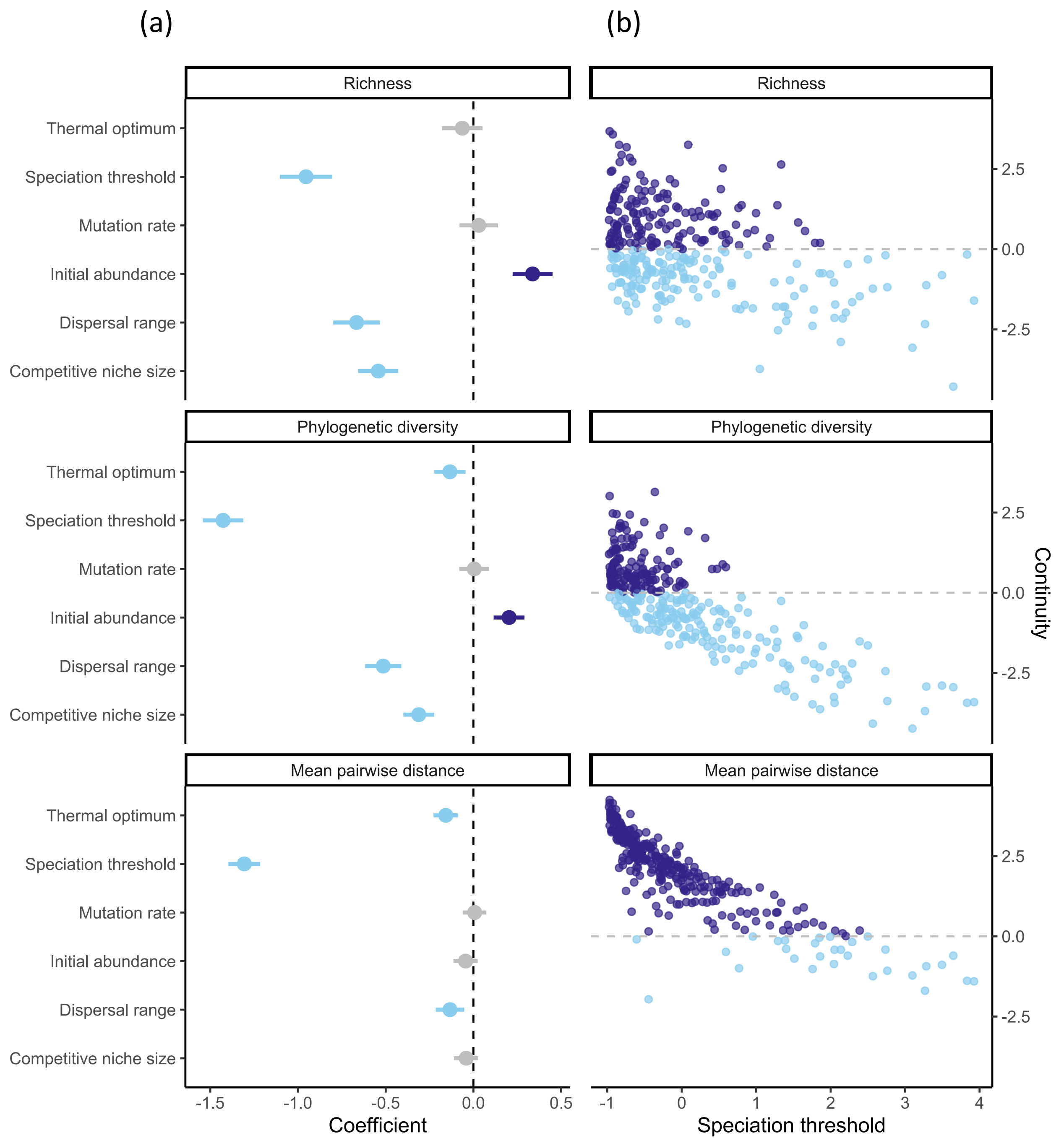


Figure 3: (a) Plots of multiple linear regression predictor coefficients showing the direction and magnitude of impact on population-species continuity metrics across each facet of diversity. Negative values (light blue) indicate that increasing the parameter drives the relative diversity towards the population level, whilst positive values (dark blue) drives diversity to the species level. Horizontal bars represent the standard error. Greyed parameters are less significant ($p > 0.05$). (b) Scatterplot of continuity metrics against the most significant parameter, the speciation threshold. Positive values (dark blue) indicate relatively more species diversity and negative values (light blue) indicate relatively more population diversity.

more robust to extinction than at the species level. These complex dynamics are difficult to validate empirically, but we hope conceptualising them here is a first step in understanding how they develop across organisational scale.

Limitations

We attempt to investigate mechanisms driving diversification across scale through a modelling approach for which there are clear limitations. The greatest being spatial scale to which we are limited to γ- and β-diversity comparisons across organisation (Whittaker, 1960) whilst in the knowledge that continuity in α-, β-, and γ-diversity behave differently (Vilcot, in press). In turn, we should also acknowledge that whilst our model is rooted in the real-world system of tropical reef-associated fishes, the goal is to meaningfully implement process, not recreate patterns perfectly. Despite this, the mechanistic modelling approach applied here shows that even with a relatively simple representation of biological processes, observed patterns can broadly be reproduced (Supplementary Figure 1). These include the Indo-Australasian Archipelago major hotspot, and Indian Ocean and Caribbean minor hotspots, as well as the latitudinal gradient of low equatorial richness followed by tropical increase and eventual temperate decrease (Albouy et al., 2019; Hillebrand, 2004). Key differences between simulated and observed patterns are likely a result of the model resolution and exclusion of key oceanographic dynamics. The low-resolution results in the Red Sea being isolated from the Indian Ocean and the Indo-Malayan archipelago fusing into an impermeable barrier. We decided to leave these inaccuracies that emerge in the final time step of our model in to remain more consistent with the accuracy of timesteps into the past. The simulation also did not account for the Eastern Pacific Barrier (Romero-Torres et al., 2018), the Benguela Current in the Eastern Atlantic (Floeter et al., 2008) which inhibit shallow water coral reef formation and dispersal (Parravicini et al., 2013), and the obstructive fresh-water outflow from major river basins (Floeter et al., 2008). Further, it is likely the latitudinal gradient remained under-developed due to the hard temperature and depth limits used to compile the landscape inputs. This prevents potential back and forth colonisation of tropical reef-fish clades to colder and/or deeper waters (Bongaerts et al., 2010). Given these considerations, we have confidence that the parameters and seascapes we did implement performed well in emulating process, and that these are viable for inferring the fundamental processes that shape diversity across organisational scale that we aim to explore.

Conclusions

We reveal the speciation threshold as an important potential driver of the formation of counter-intuitive continuity in diversity patterns across organisational levels, as explored here in reef-associated fish. In turn, this speciation threshold parameter is a proxy for a vast world of mechanisms below the population and species levels of organisation – at the scale of the individual and gene – indicating that to fully understand these patterns we must consider mechanisms across the full breadth of organisational scale and that our focus on species-population continuity in diversity patterns is only a start. We also highlight that metric choice can drive differences in the continuity between organisational levels of biodiversity which likely contribute to the variable continuity patterns seen in empirical studies so far. Finally, we uncover covariation between continuity in diversity across organisational scale and common ecological descriptors which we hope helps provide context for these dynamics in the larger field of eco-evolutionary study.

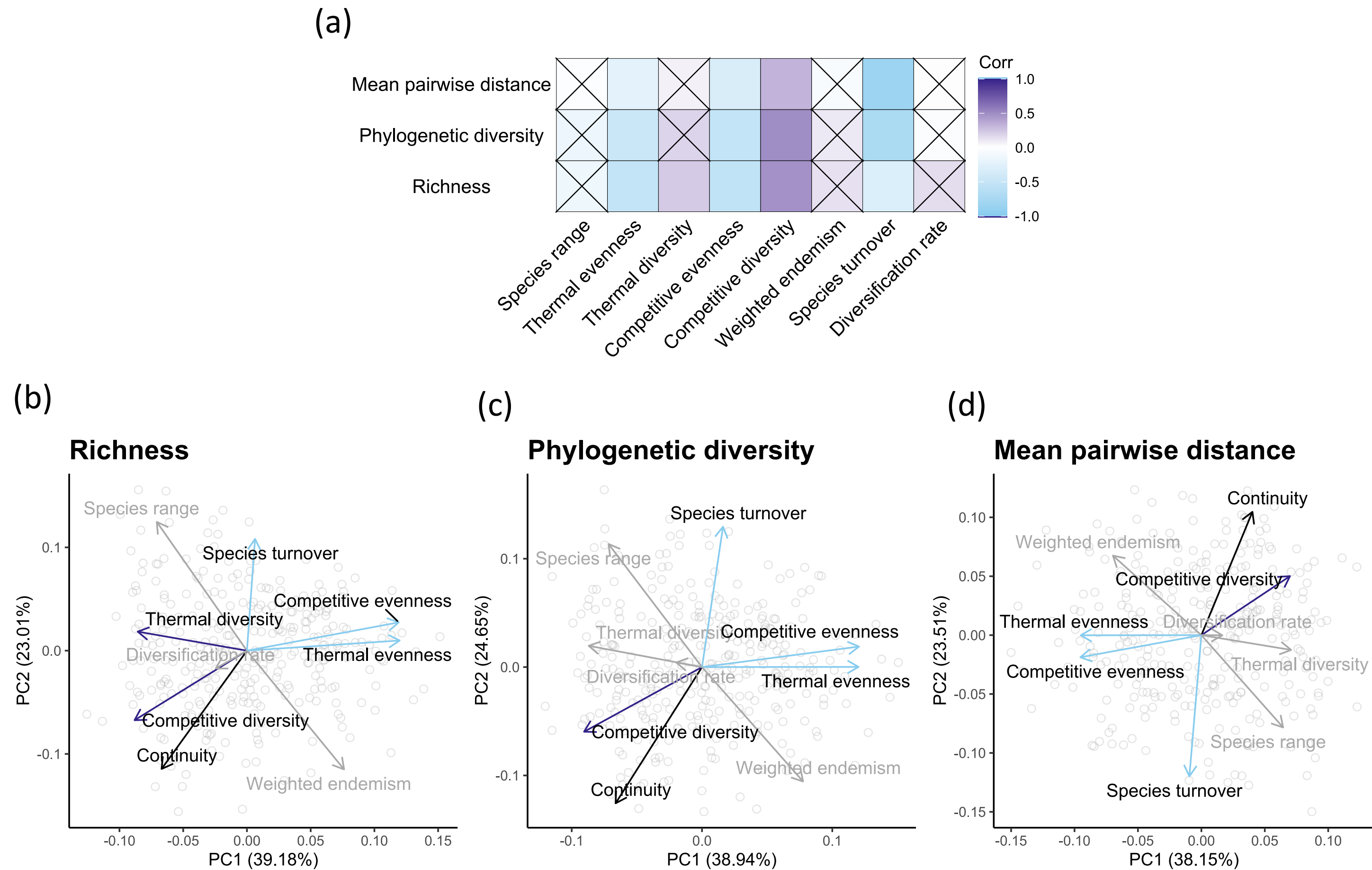


Figure 4: a) plot of the correlations between diversity continuity metrics and clade properties. Light blue indicates that increasing the clade property value is associated with an increased relative population level diversity compared to species level diversity, and vice versa for dark blue. Crosses indicate non-significant values. b-d) PCA plots of each continuity metric and clade properties. Dark blue arrows indicate a significant correlation between the clade property and a relative increase in species level to population level diversity. Light blue indicates a significant correlation in the opposite direction. Grey clade properties had no significant relationship with the continuity metric.

Acknowledgements

We thank Charles N.D. Santana, Victor Boussange, Benjamin Flück, and Lydian Boschman for excellent support and discussions through conceptualisation, technical troubleshooting, and for help navigating paleo-reconstructions.

Funding was supported by the ETH research grant financing the salary of Thomas Keggin (ETH Research Grant ETH-34 18-1).

References

Albouy, C., Archambault, P., Appeltans, W., Araujo, M. B., Beauchesne, D., Cazelles, K., Cirtwill, A. R., Fortin, M. J., Galiana, N., Leroux, S. J., Pellissier, L., Poisot, T., Stouffer, D. B., Wood, S. A., & Gravel, D. (2019). The marine fish food web is globally connected [Article]. *Nature Ecology & Evolution*, *3*(8), 1153-+. https://doi.org/10.1038/s41559-019-0950-y

Annan, J. D., & Hargreaves, J. C. (2013). A new global reconstruction of temperature changes at the Last Glacial Maximum. *Clim. Past*, *9*(1), 367-376. https://doi.org/10.5194/cp-9-367-2013

Antonovics, J. (1976). The input from population genetics: "The new ecological genetics". *Systematic Botany*, *1*(3), 233-245. https://doi.org/10.2307/2418718

Blum, M. J., Bagley, M. J., Walters, D. M., Jackson, S. A., Daniel, F. B., Chaloud, D. J., & Cade, B. S. (2012). Genetic diversity and species diversity of stream fishes covary across a land-use gradient [Article]. *Oecologia*, *168*(1), 83-95. https://doi.org/10.1007/s00442-011-2078-x

Bongaerts, P., Ridgway, T., Sampayo, E. M., & Hoegh-Guldberg, O. (2010). Assessing the 'deep reef refugia' hypothesis: focus on Caribbean reefs. *Coral Reefs*, *29*(2), 309-327. https://doi.org/10.1007/s00338-009-0581-x

Bromham, L. (2011). The genome as a life-history character: why rate of molecular evolution varies between mammal species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1577), 2503-2513. https://doi.org/doi:10.1098/rstb.2011.0014

Chave, J. (2004). Neutral theory and community ecology [Review]. *Ecology Letters*, *7*(3), 241-253. https://doi.org/10.1111/j.1461-0248.2003.00566.x

Connell, J., Den Boer, P., & Gradwell, G. (1971). Dynamics of populations. *Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands. chapter On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees*, 298-313.

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 [Article]. *Journal of Evolutionary Biology*, *24*(12), 2543-2562. https://doi.org/10.1111/j.1420-9101.2011.02391.x

Crisp, M. D., Laffan, S., Linder, H. P., & Monro, A. (2001). Endemism in the Australian flora. *Journal of Biogeography*, *28*(2), 183-198. https://doi.org/https://doi.org/10.1046/j.1365-2699.2001.00524.x

Daru, B. H., Karunarathne, P., & Schliep, K. (2020). phyloregion: R package for biogeographical regionalization and macroecology. *Methods in Ecology and Evolution*, *11*(11), 1483-1491.

Decocq, G., Guiller, A., Kichey, T., Van de Pitte, K., Gallet-Moron, E., Honnay, O., & Closset-Kopp, D. (2021). Drivers of species and genetic diversity within forest metacommunities across agricultural landscapes of different permeability. *Landscape Ecology*. https://doi.org/10.1007/s10980-021-01296-6

Descombes, P., Gaboriau, T., Albouy, C., Heine, C., Leprieur, F., & Pellissier, L. (2018). Linking species diversification to palaeo-environmental changes: A process-based modelling approach [Article]. *Global Ecology and Biogeography*, *27*(2), 233-244. https://doi.org/10.1111/geb.12683

Dinnage, R., Skeels, A., & Cardillo, M. (2020). Spatiophylogenetic modelling of extinction risk reveals evolutionary distinctiveness and brief flowering period as threats in a hotspot plant genus.

Proceedings of the Royal Society B: Biological Sciences, *287*(1926), 20192817. https://doi.org/doi:10.1098/rspb.2019.2817
Donati, G. F. A., Zemp, N., Manel, S., Poirier, M., Claverie, T., Ferraton, F., Gaboriau, T., Govinden, R., Hagen, O., Ibrahim, S., Mouillot, D., Leblond, J., Julius, P., Velez, L., Zareer, I., Ziyad, A., Leprieur, F., Albouy, C., & Pellissier, L. (2021). Species ecology explains the spatial components of genetic diversity in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1959), 20211574. https://doi.org/doi:10.1098/rspb.2021.1574
Eble, J. A., Bowen, B. W., & Bernardi, G. (2015). Phylogeography of coral reef fishes. *Ecology of fishes on coral reefs*, 64-75.
Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, *29*(1), 51-63. https://doi.org/10.1016/j.tree.2013.09.008
Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422-433. https://doi.org/10.1038/nrg.2016.58
Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*(1), 1-10. https://doi.org/https://doi.org/10.1016/0006-3207(92)91201-3
Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*(7), 342-350. https://doi.org/10.1016/j.tig.2012.03.009
Fine, P. V. A. (2015). Ecological and Evolutionary Drivers of Geographic Variation in Species Diversity. In D. J. Futuyma (Ed.), *Annual Review of Ecology, Evolution, and Systematics, Vol 46* (Vol. 46, pp. 369-392). https://doi.org/10.1146/annurev-ecolsys-112414-054102
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*(1), 22-47. https://doi.org/https://doi.org/10.1111/j.1365-2699.2007.01790.x
Gaboriau, T., Albouy, C., Descombes, P., Mouillot, D., Pellissier, L., & Leprieur, F. (2019). Ecological constraints coupled with deep-time habitat dynamics predict the latitudinal diversity gradient in reef fishes [Article]. *Proceedings of the Royal Society B-Biological Sciences*, *286*(1911), 10, Article 20191506. https://doi.org/10.1098/rspb.2019.1506
Gavrilets, S., Li, H., & Vose, M. D. (2000). Patterns of parapatric speciation [Article]. *Evolution*, *54*(4), 1126-1134. <Go to ISI>://WOS:000089317900005
Gotelli, N. J., Anderson, M. J., Arita, H. T., Chao, A., Colwell, R. K., Connolly, S. R., Currie, D. J., Dunn, R. R., Graves, G. R., Green, J. L., Grytnes, J. A., Jiang, Y. H., Jetz, W., Lyons, S. K., McCain, C. M., Magurran, A. E., Rahbek, C., Rangel, T., Soberon, J., . . . Willig, M. R. (2009). Patterns and causes of species richness: a general simulation model for macroecology [Review]. *Ecology Letters*, *12*(9), 873-886. https://doi.org/10.1111/j.1461-0248.2009.01353.x
Green, A. L., Maypa, A. P., Almany, G. R., Rhodes, K. L., Weeks, R., Abesamis, R. A., Gleason, M. G., Mumby, P. J., & White, A. T. (2015). Larval dispersal and movement patterns of coral reef fishes, and implications for marine reserve network design [Review]. *Biological Reviews*, *90*(4), 1215-1247. https://doi.org/10.1111/brv.12155
Hagen, O. (2022). Coupling eco-evolutionary mechanisms with deep-time environmental dynamics to understand biodiversity patterns. *Ecography*, *n/a*(n/a), e06132. https://doi.org/https://doi.org/10.1111/ecog.06132
Hagen, O., Fluck, B., Fopp, F., Cabral, J. S., Hartig, F., Pontarp, M., Rangel, T. F., & Pellissier, L. (2021). gen3sis: A general engine for eco-evolutionary simulations of the processes that shape Earth's biodiversity [Article]. *PLOS Biology*, *19*(7), 31, Article e3001340. https://doi.org/10.1371/journal.pbio.3001340
Hagen, O., Skeels, A., Onstein, R. E., Jetz, W., & Pellissier, L. (2021). Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. *Proceedings of the National Academy of Sciences*, *118*(40), e2026347118. https://doi.org/doi:10.1073/pnas.2026347118
Hamilton, M. B. (2021). *Population genetics*. John Wiley & Sons.

Harmon, L. J., & Harrison, S. (2015). Species Diversity Is Dynamic and Unbounded at Local and Continental Scales. *The American Naturalist*, *185*(5), 584-593. https://doi.org/10.1086/680859
Harrell Jr, F. E. (2019). Hmisc: Harrell Miscellaneous. *CRAN2018*, *2022*, 235-236.
Hibdige, S. G. S., Raimondeau, P., Christin, P. A., & Dunning, L. T. (2021). Widespread lateral gene transfer among grasses. *New Phytologist*, *230*(6), 2474-2486. https://doi.org/10.1111/nph.17328
Hijmans, R. J., Van Etten, J., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J. A., Lamigueiro, O. P., Bevan, A., Racine, E. B., & Shortridge, A. (2015). Package 'raster'. *R package*, *734*.
Hillebrand, H. (2004). Strength, slope and variability of marine latitudinal gradients. *Marine Ecology Progress Series*, *273*, 251-268. http://www.jstor.org/stable/24867490
Hubbell, S. (2001). The Unified Neutral Theory of Biodiversity and Biogeography. *Princeton University Press, Princeton, New Jersey, USA. Hughes, RG (1986) Theories and models of species abundance. American Naturalist*, *128*, 879-899.
Isaac, N. J. B., Turvey, S. T., Collen, B., Waterman, C., & Baillie, J. E. M. (2007). Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny. *Plos One*, *2*(3), e296. https://doi.org/10.1371/journal.pone.0000296
Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, *104*(940), 501-528.
Jaquiery, J., Broquet, T., Hirzel, A. H., Yearsley, J., & Perrin, N. (2011). Inferring landscape effects on dispersal from genetic distances: how far can we go? *Molecular Ecology*, *20*(4), 692-705. https://doi.org/10.1111/j.1365-294X.2010.04966.x
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*(7424), 444-448. https://doi.org/10.1038/nature11631
Kahilainen, A., Puurtinen, M., & Kotiaho, J. S. (2014). Conservation implications of species-genetic diversity correlations [Review]. *Global Ecology and Conservation*, *2*, 315-323. https://doi.org/10.1016/j.gecco.2014.10.013
Kinlock, N. L., Prowant, L., Herstoff, E. M., Foley, C. M., Akin-Fajiye, M., Bender, N., Umarani, M., Ryu, H. Y., Sen, B., & Gurevitch, J. (2018). Explaining global variation in the latitudinal diversity gradient: Meta-analysis confirms known patterns and uncovers new ones [Article]. *Global Ecology and Biogeography*, *27*(1), 125-141. https://doi.org/10.1111/geb.12665
Lawrence, E. R. (2020). *Synthesizing vertebrate population richness and genetic diversity across the American continents* Concordia University].
Lawrence, E. R., & Fraser, D. J. (2020). Latitudinal biodiversity gradients at three levels: Linking species richness, population richness and genetic diversity [Review]. *Global Ecology and Biogeography*, *29*(5), 770-788. https://doi.org/10.1111/geb.13075
Leprieur, F., Colosio, S., Descombes, P., Parravicini, V., Kulbicki, M., Cowman, P. F., Bellwood, D. R., Mouillot, D., & Pellissier, L. (2016). Historical and contemporary determinants of global phylogenetic structure in tropical reef fish faunas [Article]. *Ecography*, *39*(9), 825-835. https://doi.org/10.1111/ecog.01638
Leps, J., de Bello, F., Lavorel, S., & Berman, S. (2006). Quantifying and interpreting functional diversity of natural communities: practical considerations matter [Article]. *Preslia*, *78*(4), 481-501. <Go to ISI>://WOS:000242535900008
MacArthur, R. H., & Wilson, E. O. (2016). Princeton University Press. https://doi.org/doi:10.1515/9781400881376
Manel, S., Guerin, P. E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., & Pellissier, L. (2020). Global determinants of freshwater and marine fish genetic diversity. *Nature Communications*, *11*(1). https://doi.org/ARTN 69210.1038/s41467-020-14409-7

Martin, S. H., Davey, J. W., Salazar, C., & Jiggins, C. D. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLOS Biology*, *17*(2), Article e2006288. https://doi.org/10.1371/journal.pbio.2006288

Mason, N. W. H., Mouillot, D., Lee, W. G., & Wilson, J. B. (2005). Functional richness, functional evenness and functional divergence: the primary components of functional diversity [Article]. *Oikos*, *111*(1), 112-118. https://doi.org/10.1111/j.0030-1299.2005.13886.x

Mayr, E. (1963). Animal Species and Evolution (Belknap, Cambridge, MA).

Messmer, V., Jones, G. P., Munday, P. L., & Planes, S. (2012). Concordance between genetic and species diversity in coral reef fishes across the Pacific Ocean biodiversity gradient. *Evolution*, *66*(12), 3902-3917. https://doi.org/10.1111/j.1558-5646.2012.01725.x

Mihaljevic, M., Korpanty, C., Renema, W., Welsh, K., & Pandolfi, J. M. (2017). Identifying patterns and drivers of coral diversity in the Central Indo-Pacific marine biodiversity hotspot. *Paleobiology*, *43*(3), 343-364. https://doi.org/10.1017/pab.2017.1

Mouillot, D., Mason, W. H. N., Dumay, O., & Wilson, J. B. (2005). Functional regularity: a neglected aspect of functional diversity [Article]. *Oecologia*, *142*(3), 353-359. https://doi.org/10.1007/s00442-004-1744-7

Nee, S. (2005). The Neutral Theory of Biodiversity: Do the Numbers Add up? *Functional Ecology*, *19*(1), 173-176. http://www.jstor.org/stable/3599286

Nei, M., Maruyama, T., & Wu, C.-I. (1983). Models of evolution of reproductive isolation. *Genetics*, *103*(3), 557-579. https://doi.org/10.1093/genetics/103.3.557

Nei, M., & Takahata, N. (1993). Effective population size, genetic diversity, and coalescence time in subdivided populations. *Journal of Molecular Evolution*, *37*(3), 240-244. https://doi.org/10.1007/BF00175500

Palumbi, S. R. (1994). Genetic Divergence, Reproductive Isolation, and Marine Speciation. *Annual Review of Ecology and Systematics*, *25*, 547-572. http://www.jstor.org/stable/2097324

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*(12), 1254-1262. https://doi.org/https://doi.org/10.1111/j.1600-0587.2013.00291.x

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Molecular Ecology*, *25*(11), 2337-2360. https://doi.org/10.1111/mec.13557

Pellissier, L., Leprieur, 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*(6187), 1016. https://doi.org/10.1126/science.1249853

Pfeiffer, V. W., Ford, B. M., Housset, J., McCombs, A., Blanco-Pastor, J. L., Gouin, N., Manel, S., & Bertin, A. (2018). Partitioning genetic and species diversity refines our understanding of species-genetic diversity relationships. *Ecology and Evolution*, *8*(24), 12351-12364. https://doi.org/10.1002/ece3.4530

Pilowsky, J. A., Colwell, R. K., Rahbek, C., & Fordham, D. A. (2022). Process-explicit models reveal the structure and dynamics of biodiversity patterns. *Science Advances*, *8*(31), eabj2271. https://doi.org/doi:10.1126/sciadv.abj2271

Potts, D. C. (1984). Generation times and the quaternary evolution of reef-building corals. *Paleobiology*, *10*(1), 48-58. https://doi.org/10.1017/s0094837300008010

Prowse, T. A. A., Bradshaw, C. J. A., Delean, S., Cassey, P., Lacy, R. C., Wells, K., Aiello-Lammens, M. E., Akçakaya, H. R., & Brook, B. W. (2016). An efficient protocol for the global sensitivity analysis of stochastic ecological models. *Ecosphere*, *7*(3), e01238. https://doi.org/https://doi.org/10.1002/ecs2.1238

R Core Team. (2022). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. https://www.R-project.org/

Rabosky, D. L. (2020). Speciation rate and the diversity of fishes in freshwaters and the oceans. *Journal of Biogeography*, *47*(6), 1207-1217. https://doi.org/10.1111/jbi.13839

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*(7714), 392-395. https://doi.org/10.1038/s41586-018-0273-1
Rabosky, D. L., & Hurlbert, A. H. (2015). Species Richness at Continental Scales Is Dominated by Ecological Limits. *The American Naturalist*, *185*(5), 572-583. https://doi.org/10.1086/680850
Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlovic, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, *30*(8), 1450-1477. https://doi.org/10.1111/jeb.13047
Redding, D. W., & Mooers, A. Ø. (2006). Incorporating Evolutionary Measures into Conservation Prioritization. *Conservation Biology*, *20*(6), 1670-1678. https://doi.org/https://doi.org/10.1111/j.1523-1739.2006.00555.x
Reisch, C., & Hartig, F. (2021). Species and genetic diversity patterns show different responses to land use intensity in central European grasslands. *Diversity and Distributions*, *27*(3), 392-401. https://doi.org/10.1111/ddi.13199
Reisch, C., & Schmid, C. (2019). Species and genetic diversity are not congruent in fragmented dry grasslands. *Ecology and Evolution*, *9*(1), 664-671. https://doi.org/10.1002/ece3.4791
Romero-Torres, M., Treml, E. A., Acosta, A., & Paz-García, D. A. (2018). The Eastern Tropical Pacific coral population connectivity and the role of the Eastern Pacific Barrier. *Scientific Reports*, *8*(1), 9354. https://doi.org/10.1038/s41598-018-27644-2
Schmidt, C., Dray, S., & Garroway, C. J. (2022). Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity. *Evolution*, *76*(1), 86-100. https://doi.org/https://doi.org/10.1111/evo.14407
Scotese, C. R. (2021). An Atlas of Phanerozoic Paleogeographic Maps: The Seas Come In and the Seas Go Out. *Annual Review of Earth and Planetary Sciences*, *49*(1), 679-728. https://doi.org/10.1146/annurev-earth-081320-064052
Scotese, C. R., Song, H., Mills, B. J. W., & van der Meer, D. G. (2021). Phanerozoic paleotemperatures: The earth's changing climate during the last 540 million years. *Earth-Science Reviews*, *215*, 103503. https://doi.org/https://doi.org/10.1016/j.earscirev.2021.103503
Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., Peichel, C. L., Saetre, G. P., Bank, C., Brannstrom, A., Brelsford, A., Clarkson, C. S., Eroukhmanoff, F., Feder, J. L., Fischer, M. C., Foote, A. D., Franchini, P., Jiggins, C. D., Jones, F. C., . . . Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*(3), 176-192. https://doi.org/10.1038/nrg3644
Singhal, S., Huang, H., Grundler, M. R., Marchán-Rivadeneira, M. R., Holmes, I., Title, P. O., Donnellan, S. C., & Rabosky, D. L. (2018). Does Population Structure Predict the Rate of Speciation? A Comparative Test across Australia's Most Diverse Vertebrate Radiation. *The American Naturalist*, *192*(4), 432-447. https://doi.org/10.1086/699515
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*(7), 573-583. https://doi.org/10.1641/b570707
Stewart, L., Alsos, I. G., Bay, C., Breen, A. L., Brochmann, C., Boulanger-Lapointe, N., Broennimann, O., Bultmann, H., Bocher, P. K., Damgaard, C., Daniels, F. J. A., Ehrich, D., Eidesen, P. B., Guisan, A., Jonsdottir, I. S., Lenoir, J., le Roux, P. C., Levesque, E., Luoto, M., . . . Pellissier, L. (2016). The regional species richness and genetic diversity of Arctic vegetation reflect both past glaciations and current climate [Article]. *Global Ecology and Biogeography*, *25*(4), 430-442. https://doi.org/10.1111/geb.12424
Taberlet, P., Zimmermann, N. E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Coldea, G., Mirek, Z., Moilanen, A., Ahlmer, W., Marsan, P. A., Bona, E., Bovio, M., Choler, P.,

Cieslak, E., Colli, L., Cristea, V., Dalmas, J. P., . . . IntraBioDiv, C. (2012). Genetic diversity in widespread species is not congruent with species richness in alpine plant communities [Article]. *Ecology Letters*, *15*(12), 1439-1448. https://doi.org/10.1111/ele.12004
Tsirogiannis, C., & Sandel, B. (2016). PhyloMeasures: a package for computing phylogenetic biodiversity measures and their statistical moments. *Ecography*, *39*(7), 709-714. https://doi.org/https://doi.org/10.1111/ecog.01814
Tucker, C. M., & Cadotte, M. W. (2013). Unifying measures of biodiversity: understanding when richness and phylogenetic diversity should be congruent. *Diversity and Distributions*, *19*(7), 845-854. https://doi.org/10.1111/ddi.12087
Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., Grenyer, R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Purschke, O., Redding, D. W., Rosauer, D. F., Winter, M., & Mazel, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews*, *92*(2), 698-715. https://doi.org/https://doi.org/10.1111/brv.12252
Tuomisto, H. (2010). A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, *33*(1), 2-22. https://doi.org/10.1111/j.1600-0587.2009.05880.x
UNEP-WCMC, W. C., WRI, TNC. (2010). *Global distribution of warm-water coral reefs, compiled from multiple sources including the Millennium Coral Reef Mapping Project* Version 4.0). https://doi.org/https://doi.org/10.34892/t2wk-5t34
van Etten, J. (2017). R Package gdistance: Distances and Routes on Geographical Grids. *Journal of Statistical Software*, *76*(13), 1 - 21. https://doi.org/10.18637/jss.v076.i13
Vellend, M. (2005). Species diversity and genetic diversity: Parallel processes and correlated patterns. *American Naturalist*, *166*(2), 199-215. https://doi.org/10.1086/431318
Vellend, M. (2010). Conceptual synthesis in community ecology [Review]. *Quarterly Review of Biology*, *85*(2), 183-206. https://doi.org/10.1086/652373
Vellend, M., & Geber, M. A. (2005). Connections between species diversity and genetic diversity. *Ecology Letters*, *8*(7), 767-781. https://doi.org/10.1111/j.1461-0248.2005.00775.x
Vilcot, M. (in press). Species-genetic diversity correlations in tropical reef fishes.
Vilcot, M., Albouy, C., Donati, G. F. A., Claverie, T., Julius, P., Manel, S., Pellissier, L., & Leprieur, F. (2023). Spatial genetic differentiation correlates with species assemblage turnover across tropical reef fish lineages. *Global Ecology and Biogeography*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/geb.13637
Waldock, C., Stuart-Smith, R. D., Edgar, G. J., Bird, T. J., & Bates, A. E. (2019). The shape of abundance distributions across temperature gradients in reef fishes [Article]. *Ecology Letters*, *22*(4), 685-696. https://doi.org/10.1111/ele.13222
Webb, C. O. (2000). Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees. *Am Nat*, *156*(2), 145-155. https://doi.org/10.1086/303378
Whittaker, R. H. (1960). Vegetation of the Siskiyou mountains, Oregon and California. *Ecological Monographs*, *30*(3), 279-338.
Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, *16*(2), 97-159. https://doi.org/10.1093/genetics/16.2.97
Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

Supplementary

Supplementary Table 1: Summary of median diversity and clade properties.

Region	*Species richness*	*Species PD*	*Species MPD*	*Population richness*	*Population PD*	*Population MPD*
Eastern Indo-Pacific	9.5	7835	1760.1618	1	1.8889	1.7632
Tropical Eastern Pacific	17	8016	1733.2557	1	0.3246	0.3246
Tropical Atlantic	27	10643.5	1726.419	2	10.5853	7.6351
Western Indo-Pacific	27	10358	1742.3737	2	14.7713	8.3729
Central Indo-Pacific	25	9834	1732.9517	3	23.6243	11.0183
Global	**51**	**15413.5**	**1761.3478**	**3**	**36.1665**	**16.3198**

Region	*Species range*	*Thermal evenness*	*Thermal diversity*	*Competitive evenness*	*Competitive diversity*	*Mean species richness*	*Weighted endemism*	*Species turnover*	*Diversification rate*
Eastern Indo-Pacific	5.093	0.9321	0.0017	0.8425	0.044	4.0909	0.003	0.0183	0.0026
Tropical Eastern Pacific	19.3167	1.0001	0.0016	0.9262	0.0498	5.0859	0.0112	0.0543	0.0023
Tropical Atlantic	42.6061	0.9865	0.0016	0.9143	0.051	4.6273	0.0125	0.0874	0.0028
Western Indo-Pacific	48.3364	1.0514	0.0017	0.9815	0.0475	3.6938	0.0046	0.0924	0.0029
Central Indo-Pacific	148.7826	1.0028	0.0014	0.929	0.0451	4.4347	0.003	0.0773	0.0027
Global	**113.2523**	**1.0166**	**0.0016**	**0.938**	**0.0468**	**4.2047**	**0.006**	**0.1714**	**0.0039**

Supplementary Table 2: Summary of multiple linear regression models predicting population-species level continuity using clade properties as predictor variables.

	Richness Continuity			PD Continuity			MPD Continuity		
Predictors	*Estimates*	*Statistic*	*p*	*Estimates*	*Statistic*	*p*	*Estimates*	*Statistic*	*p*
(Intercept)	8.46	12.42	**<0.001**	9.48	17.16	**<0.001**	8.60	20.10	**<0.001**
Dispersal Range	-0.64	-10.03	**<0.001**	-0.51	-9.86	**<0.001**	-0.12	-3.07	**0.002**
Speciation threshold	-0.94	-12.68	**<0.001**	-1.43	-24.41	**<0.001**	-1.30	-28.41	**<0.001**
Competitive niche size	-0.54	-9.36	**<0.001**	-0.31	-7.01	**<0.001**			
Initial abundance	0.33	5.78	**<0.001**	0.20	4.55	**<0.001**	-0.06	-1.63	0.105
Thermal optimum				-0.13	-2.97	**0.003**	-0.16	-4.41	**<0.001**
Observations	316			316			316		
R^2 / R^2 adjusted	0.431 / 0.424			0.689 / 0.684			0.790 / 0.787		

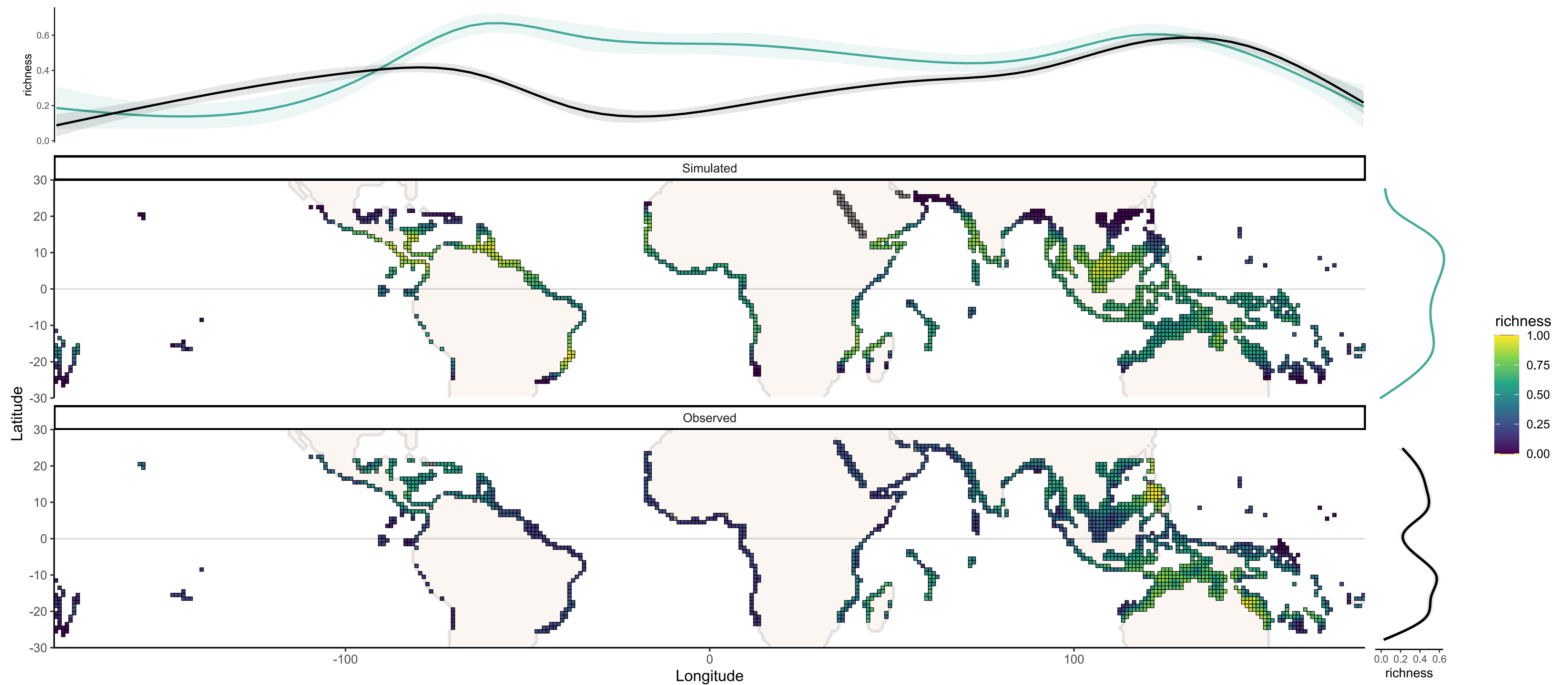
Supplementary Table 3: Correlation values between diversity continuity metrics and clade properties. p-values have been Bonferroni corrected and significant values are highlighted in bold.

r values			
Trait	*Richness*	*Mean pairwise distance*	*Phylogenetic diversity*
Species range	-0.16	-0.03	-0.17
Thermal evenness	**-0.53**	**-0.24**	**-0.47**
Thermal diversity	**0.22**	0.06	0.18
Competitive evenness	**-0.55**	**-0.34**	**-0.53**
Competitive diversity	**0.47**	**0.32**	**0.49**
Weighted endemism	0.13	-0.07	0.09
Species turnover	**-0.32**	**-0.85**	**-0.73**
Diversification rate	0.14	0.01	-0.05

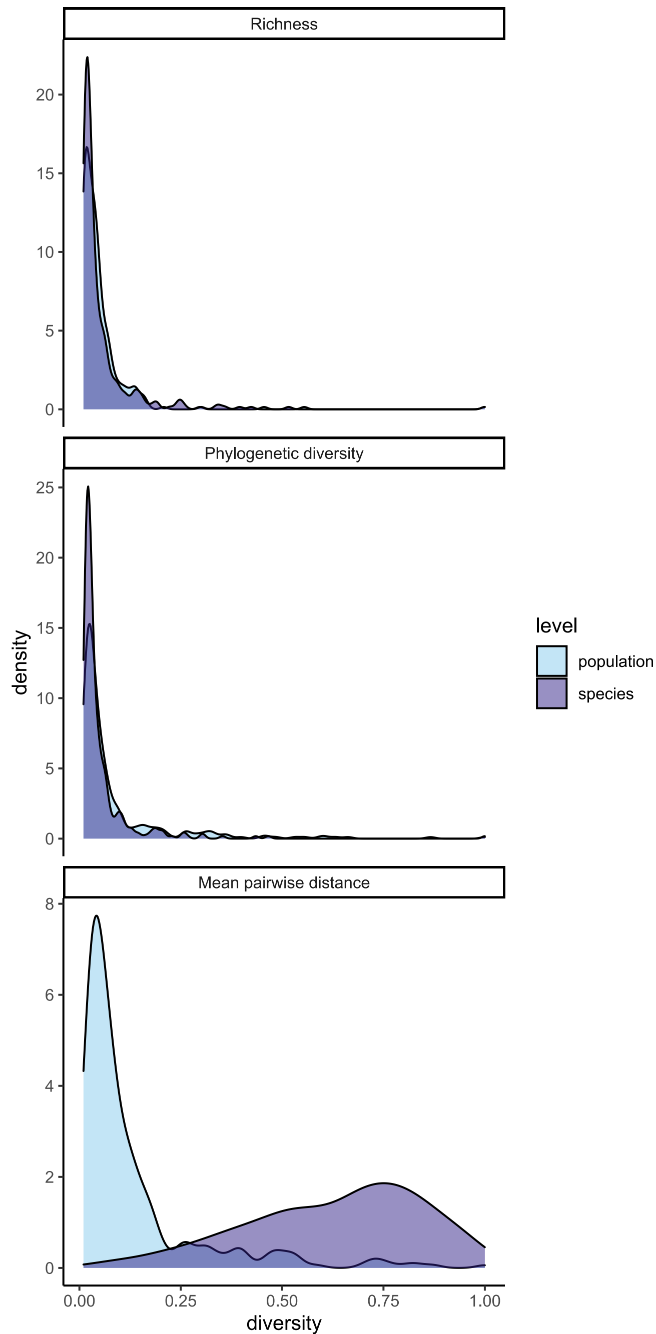
p values			
Trait	*Richness*	*Mean pairwise distance*	*Phylogenetic diversity*
Species range	0.28	33.30	0.17
Thermal evenness	**0.00**	**0.00**	**0.00**
Thermal diversity	**0.01**	17.89	0.07
Competitive evenness	**0.00**	**0.00**	**0.00**
Competitive diversity	**0.00**	**0.00**	**0.00**
Weighted endemism	1.08	10.62	6.88
Species turnover	**0.00**	**0.00**	**0.00**
Diversification rate	0.84	47.30	23.12

Supplementary Table 4: Table of PCA contributions to each variable corresponding to visualisation in Figure 4.

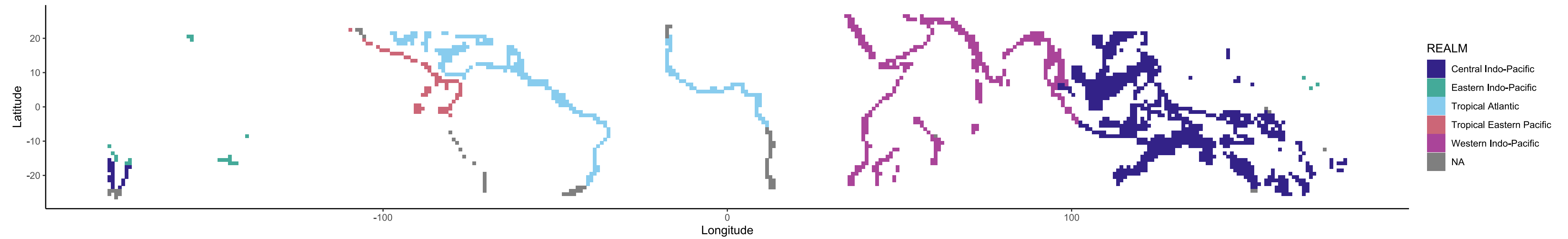
	Richness		***PD***		***MPD***	
Variable	*Component 1*	*Component 2*	*Component 1*	*Component 2*	*Component 1*	*Component 2*
Continuity	7.52	22.01	7.52	22.01	7.52	22.01
Species range	8.39	26.08	8.39	26.08	8.39	26.08
Thermal evenness	24.09	0.16	24.09	0.16	24.09	0.16
Thermal diversity	12.33	0.57	12.33	0.57	12.33	0.57
Competitive evenness	23.81	1.26	23.81	1.26	23.81	1.26
Competitive diversity	13.04	7.6	13.04	7.6	13.04	7.6
Weighted endemism	9.74	22.17	9.74	22.17	9.74	22.17
Species turnover	0.07	19.65	0.07	19.65	0.07	19.65
Diversification rate	1.01	0.5	1.01	0.5	1.01	0.5



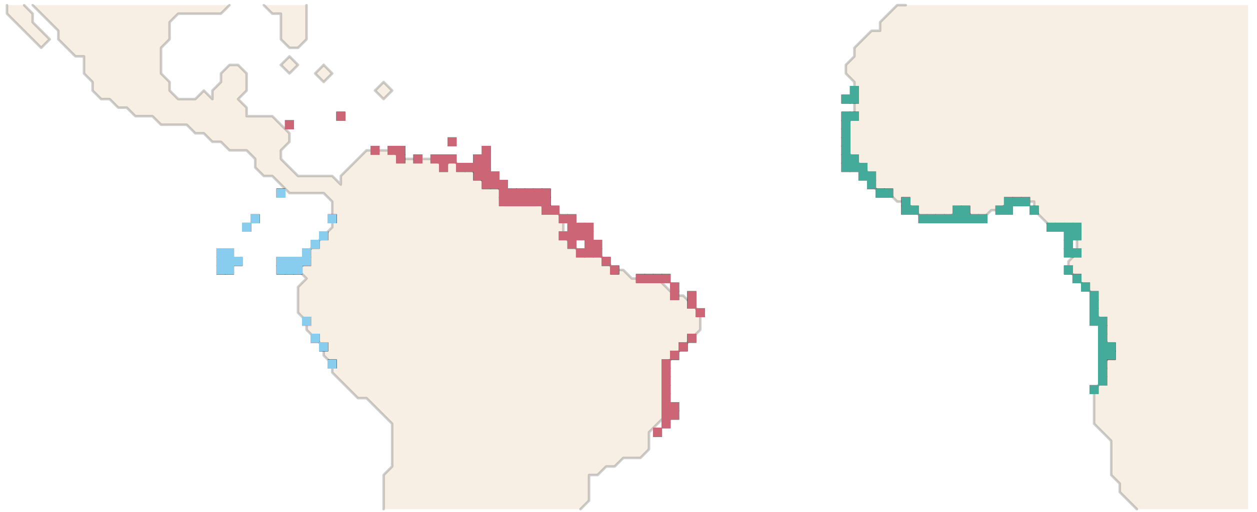
Supplementary Figure 1: Comparison of simulated and observed tropical fish species richness. Simulated richness is the summation of all species present in all retained simulations. Observed richness is taken from Albouy et al. (2019). Both datasets are normalised between 0 and 1 for comparability. Marginal plots are the mean richness values across latitude and longitude; green and black are simulated and observed, respectively. Grid cells are at 1° resolution.



Supplementary Figure 2: distribution of normalised diversity metrics at the species and population levels of organisation across retained simulations. For the richness metrics (richness and phylogenetic diversity), the distributions are similar across levels. For the divergence metric (mean pairwise distance), the diversity distribution is heavily skewed to the right at the species level and skewed to the left at the population level.



Supplementary Figure 3: To calculate β-diversity, all habitable cells in each simulation were assigned to one of the 5 tropical realms described by (Spalding et al. 2007). Cells that were not assigned to a tropical realm were given no designation (NA values) and were discarded from the β-diversity analyses.



Supplementary Figure 4: An example of population assignment in a simulation where each occupied cell has been clustered based on their dispersal distance and distance to one another. Each colour represents a geographic cluster which is then treated as the population object in the analyses. Each colour represents a geographic population.

Chapter 2: Patterns of genetic diversity differ between the Western Indian Ocean and Caribbean

Authors

Thomas Keggin[1,2], Maurine Vilcot[3], Giulia Donati[6], Conor Waldock[5], Niklaus Zemp[10], Giomar H. Borrero-Pérez[7], Maria Mutis Martinezguerra[7], Virginie Marques[1,2], Laure Velez[4], Juan David González Corredor[7], Andrés Acosta-Chaparro[7], Mark Vermeij[6], Régis Hocdé[4], Fabien Leprieur[3,4], David Mouillot[3], Andrea Polanco[8], Camille Albouy[1,2], Stéphanie Manel[3], Loïc Pellissier[1,2]

Author affiliations

1. Ecosystems and Landscape Evolution, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland
2. Unit of Land Change Science, Swiss Federal Research Institute WSL
3. CEFE, Univ. Montpellier, CNRS, EPHE- PSL University, Montpellier, France
4. MARBEC, CNRS, Ifremer, IRD, University of Montpellier, Montpellier, France
5. Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland
6. Department of Fish Ecology and Evolution, Center for Ecology, Evolution and Biogeochemistry, Eawag - Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland
7. Programa de Biodiversidad y Ecosistemas Marinos, Instituto de Investigaciones Marinas y Costeras-INVEMAR, Museo de Historia Natural Marina de Colombia (MHNMC), Santa Marta, Colombia
8. Fundación Biodiversa Colombia, Bogotá, D.C., Colombia
9. CARMABI Foundation, Curaçao
10. Genetic Diversity Centre (GDC), ETH Zürich, CH8092 Zürich, Switzerland

Abstract

Patterns of genetic diversity are a complex product of core population genetic processes such as gene flow, genetic drift, mutation, and selection. The action of these processes is tightly linked to a species' biological traits and its environmental context over time. Whilst many of these dynamics are theoretically well defined, the relative importance and interactions between these dynamics in natural populations remains unclear. In this study, we take an existing population genetic data from the Western Indian Ocean and expand the sampling and genotyping scheme to the Caribbean. We sample an additional 640 individuals within 15 species in the Caribbean across a range of biological traits in species distributed widely across the Teleost radiation. We select species that have phylogenetically close relatives in the Western Indian Ocean to allow meaningful comparison between oceans. This expanded data set allows us to investigate the relative associations of seascape and biological traits with local (α) and regional (γ) genetic diversity, as well as genetic differentiation (β). We find no significant relationship between all three measures of genetic diversity and the dispersal proxies of maximum body length and pelagic larval duration, as well as species abundance. We do, however, find higher values of observed heterozygosity in the Caribbean compared to the Western Indian Ocean. Conversely, we find higher levels of inbreeding and genetic differentiation in the Western Indian Ocean. We attribute these differences to the greater amount of habitat area and connectivity in the Caribbean and evaluate the relative importance of biological

traits and habitat configuration. We conclude that whilst traits play a role in determining patterns of genetic diversity, they are dependent on environmental context.

Introduction

Neutral intraspecific genetic diversity (Holderegger et al., 2006) varies widely across organisms, which has been related to either their biological traits (Donati et al., 2021; Romiguier et al., 2014), the properties of the habitat that they occupy (Manel & Holderegger, 2013; Selkoe et al., 2016), or both. The properties of a land-, or sea-scape has been widely accepted as a primary determinant of intraspecific genetic diversity since the founding of population genetic study with models such as isolation by distance (Wright, 1943). In recent decades, spatial data have been combined with population genetic (or genomic) markers in wild species to explicitly disentangle the various effects of the environment on genetic diversity and structure – referred to as "landscape genetics" (Manel et al., 2003), or "seascape genetics" for marine systems (Selkoe et al., 2008). General relationships have emerged from these studies, including the effect of isolation by distance on genetic distinctiveness, dispersal barriers (Rocha et al., 2007), the structure and size of the habitat landscape on the total genetic diversity (Lamy et al., 2012), and abiotic selective factors such as temperature and salinity (Geburzi et al., 2022; Lehnert et al., 2019). Particular to ocean systems, these effects are complicated by current systems which directionally facilitate or block dispersal events (Thompson et al., 2018), and depth which provides an often overlooked and complex third spatial dimension for species distributions (Gaither et al., 2018; Gaither et al., 2016). These explanatory variables do not, however, explain the inter-species variation in genetic diversity and structure found within a geographic location.

Genetic diversity (both heterozygosity and differentiation between individuals) can vary between marine species within the same seascape (Donati et al., 2021) – variation that cannot solely be explained by the environment. It follows that if the environment remains constant, the variation in observed genetic patterns must be either stochastic or driven by intrinsic differences between species (Leigh et al., 2021). The processes influencing genetic diversity and population structure such as patterns of genetic diversity are associated with the effective size of the population and its fluctuations over time – a dynamic linked with reproductive output (Ellegren & Galtier, 2016). Other reproductive traits such as generation time (Potts, 1984), levels of out-crossing, and prevalence of selfing also serve to enhance or reduce the genetic diversity of a system over time (Wright et al., 2013). Reproductive mode can even play a role, with diversity enhanced through introgression (Anderson & Hubricht, 1938; He et al., 2019) or through horizontal gene transfer events (Hibdige et al., 2021). These influences on diversity are all modulated through the mutation rate which determines the basal generation of new variation, and selection which in turn acts to fix or remove it (Hamilton, 2021). Diversity patterns are contained within the genome, whose architecture dictates linkage and recombination, and varies between lineages (Singhal et al., 2018). The variations generated by these biological processes is not necessarily evenly distributed across individuals, and diversity can increase or decrease depending on dispersal between isolated populations and subsequent gene flow – species with greater connectivity will have a larger effective meta-population size, increasing diversity (Frankham, 2015). Dispersal is also the key biological process that determines the genetic population structure of a species: species with greater dispersal ability will experience greater gene flow, and therefore reduced population structure (Donati et al., 2021). Population structure can be further complicated through the mode of dispersal: philopatry can result in vagile species experiencing greater isolation between populations (Ashe et al., 2015; Leis, 2020), and passive dispersal of propagules can result in uneven directionality of gene-flow depending on air or ocean currents (Hare et al., 2005). Understanding how the seascape in

conjunction with species traits shapes genetic diversity is central to our understanding of how intra-species diversity will persist under global change.

Species cannot be considered outside the context of their environment and the influence of a seascape and of species traits on genetic diversity are inherently linked, and dependent of the definition of diversity used (Donati et al., 2021; Gaston & Spicer, 2013). An important consideration is spatial scale, which is often measured across three different facets, α-, β-, and γ-genetic diversity. Genetic α-diversity can be defined as the diversity within a single group of co-occurring individuals, and γ-diversity as the diversity across all individuals within all groups. Genetic β-diversity can be represented as the differentiation between each group, or the residual diversity between the local (α) and total (γ) diversity values. This is an expression of an 80-year old hierarchical population structure framework (Wright, 1949), but with modern measures of α- and γ-diversity estimated from observed or expected heterozygosity, allelic richness, and nucleotide diversity, and β-diversity still being represented by Wright's F_{ST} and its derivatives (Meirmans & Hedrick, 2011; Wright, 1949). This β-diversity, or differentiation between groups, is linked to a species' dispersal ability and mode of transport (Donati et al., 2021). However, these traits would be irrelevant if the species was restricted to a small, homogeneous patch of habitat – there would be no populations to disperse between (Wright, 1943). If two species, one with a high dispersal capacity, and another with a low dispersal capacity were distributed across two patches of habitat in close proximity, then the amount of gene flow may be the same. However, if the distance between patches was increased the high-capacity disperser should be better able to remain connected across the seascape and maintain gene flow. The effect of dispersal capacity on genetic diversity and genetic population structure depends on the environmental context (White et al., 2011). This contextual effect of dispersal should also have ramifications on the mixing of genetic variants in the meta-population (α- and γ-diversity), the more connected the populations, the greater the gene flow and the larger the pool of effective interconnected individuals. The number of effective individuals, or the effective population size, is tightly linked with genetic diversity which is driven by life history traits such as reproductive mode (Ellegren & Galtier, 2016). The maximum capacity of the effective population size is a product of the carrying capacity, productivity, and selective pressure of the habitat over time (Hohenlohe et al., 2021) – genetic α- and γ-diversity is therefore influenced by both biotic and abiotic factors which are intrinsically linked. These interactions could be even more direct, with the environment driving changes in the biological processes that determine genetic diversity, for example, as postulated in the evolutionary rate hypothesis: higher environmental temperatures increase mutation rate and decrease generation time (Evans & Gaston, 2005). Intraspecies diversity is a product of both their biological traits and environmental conditions – what remains unclear is whether these processes have the same contribution in shaping species genetic diversity. To disentangle the effects of traits and environment, we need comparisons across both aspects, as well as controls for evolutionary history. In natural systems these comparisons are challenging, but a biogeographic comparison controlling for phylogenetic lineages might offer an approximate approach to such a comparison.

Here, we investigate the effect of seascapes and biological traits in shaping the intra-specific genetic diversity of 42 species of tropical fishes. Tropical fishes are an attractive natural study system to investigate the relative impacts of seascape and traits since they are globally distributed throughout the tropics within a well-defined habitat (Cowman & Bellwood, 2011; Renema et al., 2008). The current distribution consists of relatively discontinuous bioregions along longitude which contain distinct species assemblages (Spalding et al., 2007). Whilst distinct, these assemblages contain shared lineages from the recent global radiation (Rabosky et al., 2018) resulting in a set of phylogenetically related species assemblages covering a similar trait space across highly variable

seascape configurations (Kulbicki et al., 2014; Parravicini et al., 2021). These seascapes consist of many small islands of coral reef habitat amidst vast traversable areas of pelagic ocean, making it a suitable fit to traditional island-based models of population genetics and spatial ecology (Warren et al., 2015). By sampling species assemblages across bioregions, we have a natural experiment consisting of comparable communities with shared phylogenetic heritage and traits, distributed across meaningfully separated seascapes with variable habitat configurations. This comparison of seascapes, whilst retaining some control over trait variation in shared lineages, allows us to tease apart the roles of biological traits and the environment in forming genetic diversity and genetic population structure. We formulate these aims into the following questions:

1. What is the effect of seascape on genetic diversity (α- and γ-diversity) and population genetic structure (β-diversity)?
2. What is the effect of biological traits on genetic diversity (α- and γ-diversity) and population genetic structure (β-diversity)?
3. What are the relative contributions of these two factors to patterns of genetic diversity (α- and γ-diversity) and population genetic structure (β-diversity)?

Methods

Field sampling

Between 2016 and 2020, a total of 1753 individuals were sampled across the Western Indian Ocean (1096 individuals, 27 species) and the Caribbean Sea (640 individuals, 15 species) (Figure 1 A). These 42 reef fish species were selected from species in each ocean to ensure maximum representation in terms of body size and morphology and to be distributed across the Teleost phylogeny (Figure 1D). We also aimed to sample comparable species between oceans, that is, species of the same families if possible and of the same orders if not. For both oceans, four sites were sampled: the Caribbean Sea, including Providencia Island (Colombia), Santa Marta (Colombia), Curaçao, and Martinique (France); and the Western Indian Ocean including Mafia Island (Tanzania), Mayotte (France), Seychelles, and the Maldives (Figure 1A). We collected the small individual fish by scuba diving and using hand barrier nets whereas the large, and more difficult to catch were supplemented by samples taken from local fishermen. A combination of muscle tissue and fin clippings were used across samples. These tissue samples were stored in either 90% ethanol or RNAlater.

DNA extraction, library preparation and sequencing

The DNA extraction of tissue samples was done using sbeadex™ tissue purification kit by LGC following the manufacturer's protocol and the KingFisher Flex by Thermo Fisher Scientific for large extraction batches. We evaluated the DNA quantity by using a dsDNA HS Qubit assay on a Spark M10 plate reader. A subset of extracted samples was qualitatively evaluated using an agarose gel electrophoresis with GelRed as the UV dye.

Reduced representation libraries were generated for both ocean basins using a double digest restriction-associated DNA sequencing (ddRADseq) protocol adapted from Peterson et al. (2012). Libraries from the Western Indian Ocean were constructed as described in Donati et al. (2021), whilst those from the Caribbean were constructed as follows. For all Caribbean samples, the protocol started with 50 ng of DNA that has been reduced using the EcoRI (G/AATTC) and Taq1-v2 (T/CGA) restriction enzymes. Samples were 150 bp pair-end sequenced using the Illumina Novaseq 6000 on a single S1 flowcell. We multiplexed all samples using 48 custom barcodes and Illumina Dual Index Sequencing (5 x i5 indices, and 4 x i7 indices) allowing up to 960 unique index and barcode combinations (Supplementary table 4). Ultimately, samples were split across 19 uniquely indexed

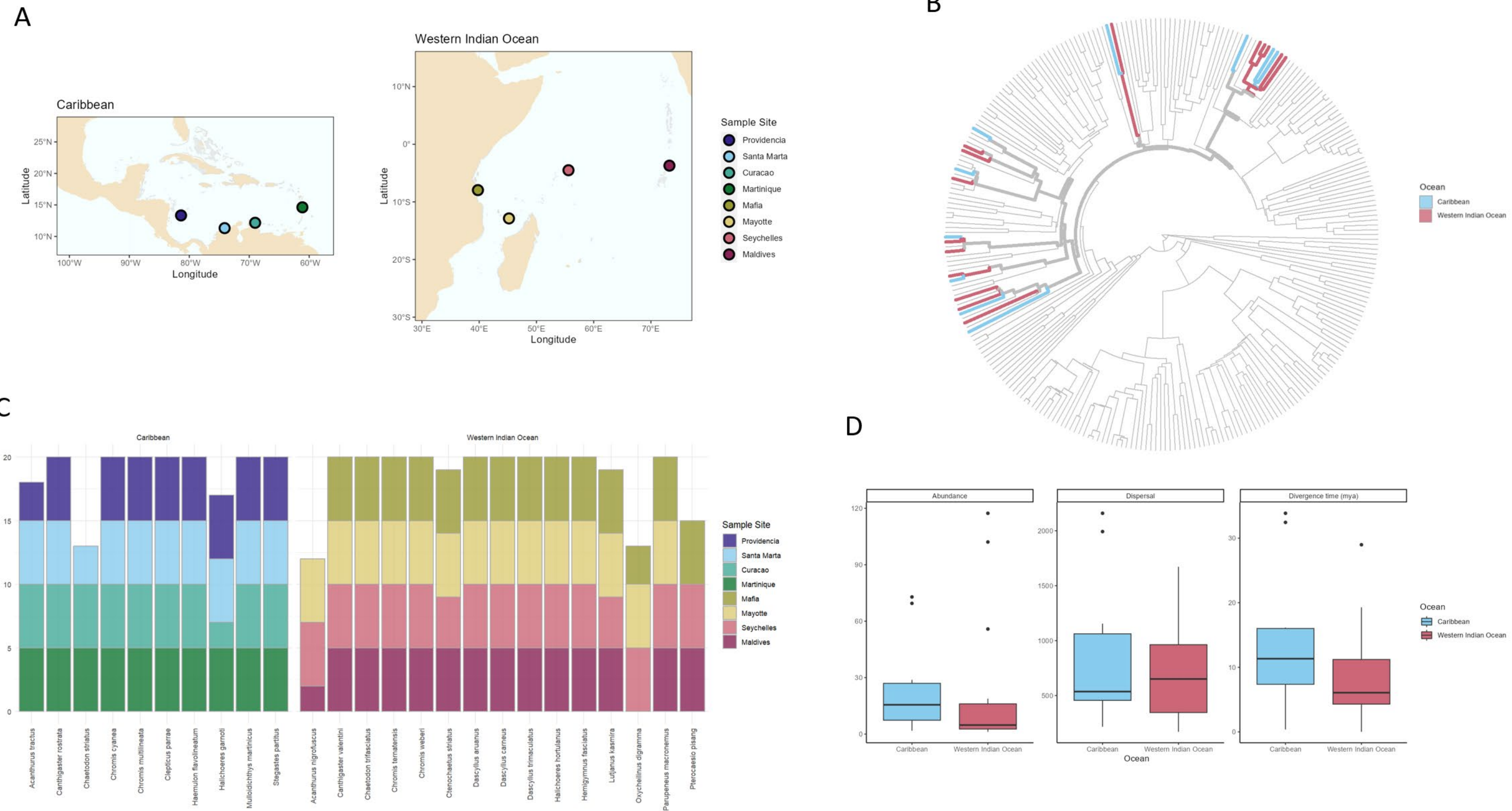


Figure 1: Summary of the sampling scheme. A) Plot of the sampling locations across the two bioregions. Light grey polygons represent reef coverage buffered to 7 km, as used in the seascape metric calculations. B) Phylogenetic tree plot of a randomly reduced subsample of the Teleost fish radiation, including the sampled fishes. Dark red indicate those from the Western Indian Ocean, light blue are those from the Caribbean. C) Plot of the number of samples collected, sequenced, and retained in this study for each sample site in both bioregions. D) Seascape metrics for the two bioregions. Values for each metric have been scaled from 0 to 1 by the largest value for each metric, making the larger value 1, and the lower value a proportion of the larger value.

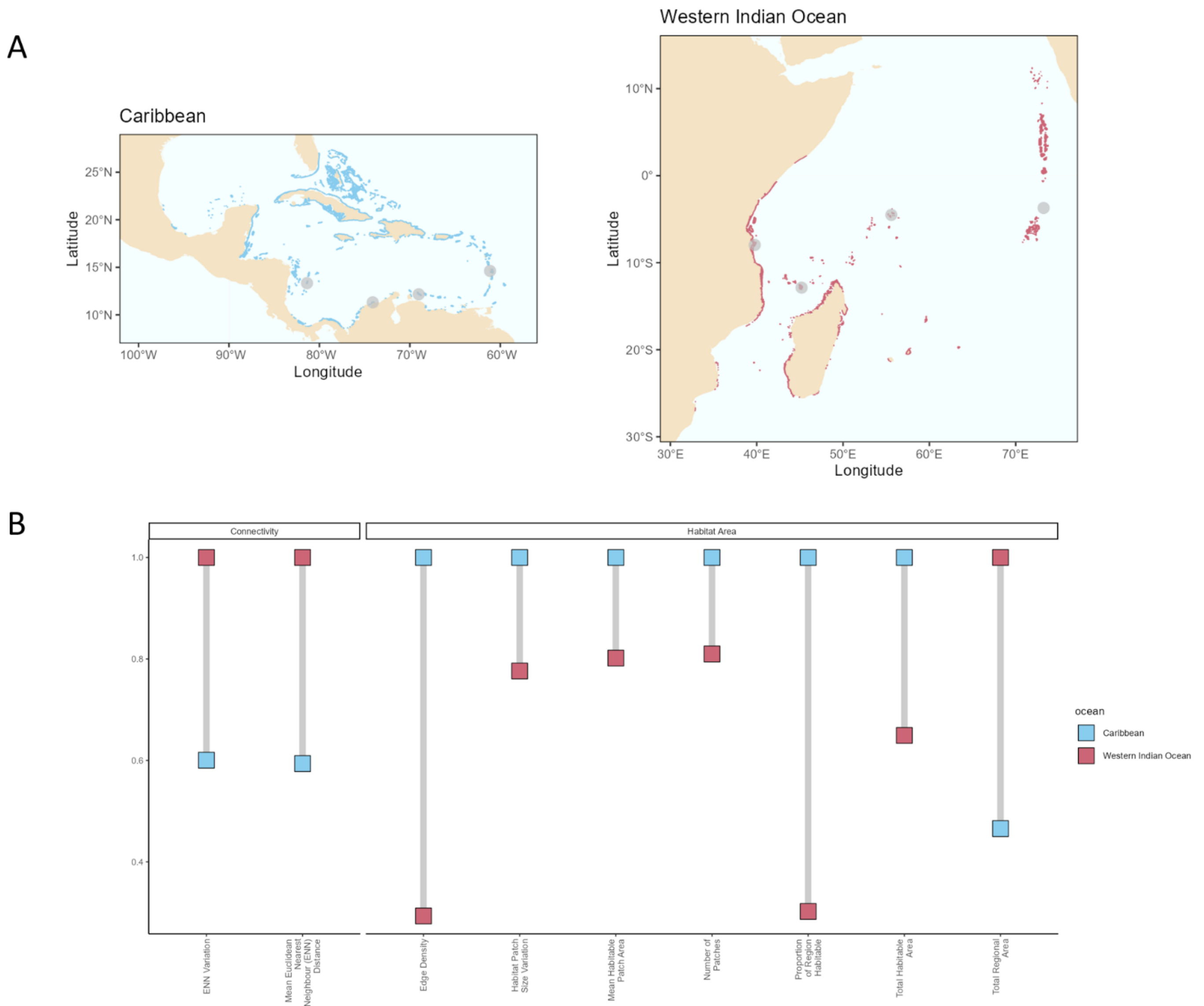


Figure 2: Map and summary of the Caribbean and Western Indian Ocean seascapes. A) Maps of the two seascapes with coloured polygons representing the distribution of shallow water coral reef habitat (UNEP WCMC, 2010). B) Proportional comparison between the difference seascape metrics calculated for each region. Values are all scaled to 1 for the higher seascape value and the lower as a proportion of the higher. The metrics are divided into those related to connectivity between habitable patches, and area of habitable patches. Light blue and dark red squares represent the Caribbean and Western Indian Ocean seascapes, respectively.

pools, with each pool containing only one species to control for variation in the number of restriction sites between species. The exception is a pool containing a split between *Chromis cyanea* and *Chromis multilineata* which we expected to have sufficiently similar genomes. Between pools we included negative controls and individual replicates to identify possible contamination and measure consistency in sequencing amongst samples. We aimed for a total fragment length of 550 bp to give an insert size of 370 bp, performing the size selection using AMPure XP beads (Greenwald et al., 2019). Because of the low amount of starting DNA, we increased the number of PCR cycles, but not beyond the limit recommended by (Peterson et al., 2012) despite the apparent robustness of RADseq genotyping to PCR bias (Euclide et al., 2020). Regardless, to mitigate against PCR duplicates we performed 4 replicates of the PCR amplification/adapter ligation step which were immediately recombined to a single pool. We evaluated each pool using the Agilent 4150 Tapestation for fragment size distribution, Qubit fluorometer for DNA quantity, and a subset run on qPCR to check correct binding of Illumina adapters. Using the Qubit quantification values and the relative number of individuals in each pool, we combined all the pools into a final library which was equimolar for each sample across pools. An aliquot of this final library was 150 bp pair-end sequenced on the Illumina iSeq 100 to return the total number of reads per pool. The disparity between pools in total number of reads was then used to reconcentrate the final library to better equalise the number of reads per pool (and therefore per species). After re-equalisation, the final library was sequenced on a single SP flow cell on the Illumina Novaseq 6000 platform at the Functional Genomics Centre Zürich. Notable differences between the library preparation between oceans was the use of an ambiguous barcode base (which were not subsequently implemented) and sequencing using the Illumina HiSeq 2500 platform for Western Indian Ocean samples.

Data set merging and cleaning

We made no alteration to the Western Indian Ocean pipeline until after genotyping, where we merge genotype calls from the two ocean basins. For the two parts of the data set, we follow the same protocol and genotyping as closely as possible. The following methods describe solely the Caribbean workflow until genotyping, after which they describe both the Caribbean and Western Indian Ocean. For details specific to the Western Indian Ocean, see Donati et al. (2021).

After genotype calling for each species separately, we cleaned and merged both ocean data sets. To ensure that signal from loci across the genome are represented equally by derived SNP markers, we limited the number of SNPs per locus to 1, choosing the SNP with the highest sequencing quality score per locus (Supplementary Figure 2). Once filtered, the VCF files were converted to the genind format using the “adegenet” R package (Jombart, 2008; Jombart & Ahmed, 2011). The genind files were processed in parallel from this point on. Given that different species have different genomic architectures, we expect different numbers of SNPs to be discovered. This disparity in the number of SNPs can bias the analyses, as more SNPs may capture more variation through greater sampling of the genome. To control for this potential bias, we randomly subsampled loci across species to the number of loci retained by the species with the minimum number of loci (n = 3715). Similarly, uneven numbers of sampled individuals between species and sites may result in uneven sampling of existing variation, so we subsampled each site to a maximum of 5 individuals per site per species. Given that random subsampling of loci across the genome creates a new source of variation in our pipeline we carried out the SNP subsampling step and all derived metric calculations iteratively 999 times. The mean metric values across these iterations were used in the final analyses after confirming an acceptably low level of variation between iterations (Supplementary Figure 4).

Population genetic metrics

Global population genetic metrics were calculated using the *global.stats* function in the hierfstat package in R (Goudet, 2005; R Core Team, 2022). We further calculated G_{ST}, G_{ST} max, G'_{ST}, G''_{ST}, H_{ST}, and Jost's F_{ST} manually from the basic statistics, as well as Hedrick's G''_{ST} using the mmod package (Winter, 2012). For visualisation of any genetic structure between individuals, we calculated a PCA for each species (function *dudi.pca*, ade4 R package; Dray et al. (2012)). From these metrics, we used observed and expected heterozygosity per subpopulation (respectively, H_O and H_S), total metapopulation expected heterozygosity (H_T), F_{IS}, and Hedrick's G''_{ST} – chosen as it works to standardise F_{ST} estimates across markers and species with varying effective population sizes, as is the case here, and when the number of sampled populations is low (Meirmans & Hedrick, 2011).

Seascape metrics

To explore the impact of seascape configuration on traits and genetic diversity, we computed several habitat metrics. The boundary of each sampling region/ocean was first derived from the province framework laid out by Spalding et al. (2007). For both the Western Indian Ocean and the Caribbean, the total habitat area was defined as the provinces containing the sample sites. Within these bounds, habitable areas were defined by the WCMC coral reef distribution polygons (UNEP-WCMC, 2010), buffered out by 7 km, projected to an Equal Area Cylindrical projection, and aggregated into a 1 km resolution raster. These habitable cells were then subdivided into patches using the *get_patches* function with 8 directions of connection in the landscapemetrics package in R (Hesselbarth et al., 2019). From these patches, in each region we calculated: Euclidian nearest neighbour, total number of patches, mean patch area, percentage of habitable area, variation in patch area, total regional area, edge density (amount of edge habitat vs total area), patch cohesion (level of patch aggregation), patch shape index (a comparison between patch shape and a regular square), and the edge to area ratio of patches. Distances are measured in km and areas are measured in km^2. For visualisation, we scaled these seascape metrics by their maximum and minimum values so that the higher seascape value is 1 and the lesser value is a proportion of the greater value.

Biological traits and species properties

Two traits were used as a proxy for dispersal capacity: maximum body length and pelagic larval duration (PLD). Maximum body length was extracted from fishbase (Boettiger et al., 2012), and PLD values were taken from both the GASPAR project (Parravicini et al., 2021) and Luiz et al. (2013). Species abundance values were extracted from the Reef Life Survey (Edgar et al., 2020; Edgar & Stuart-Smith, 2014), spanning the period 2009 to 2022. The relative total number of individuals recorded per dive survey was used as a proxy for abundance. The mean of these values was taken across the Tropical Atlantic and Western Indo-Pacific for each species retained in this study.

Determinants of α- and β-diveristy

We attempted to explain population genetic metric estimates of H_O, H_S, H_T, F_{IS}, and Hedrick's G''_{ST} with species' current abundance, most recent known divergence time, dispersal ability (body length x pelagic larval dispersal), and through the effect of seascape (Caribbean vs. Western Indian Ocean). To investigate the relationship between dispersal traits and the effect of region to population genetic diversity and structure, we implemented phylogenetic generalised least squares (PGLS). First, we fitted a multiple linear regression using the combined dispersal trait, divergence time, mean abundance and a binary dummy variable for each ocean as predictors, and H_O, H_S, H_T, and Hedrick's G''_{ST} as response variables in the R stats package (R Core Team, 2022). This linear regression assumes phylogenetic independence and is equivalent to lamba = 0 in the PGLS

framework. We then applied a stepwise variable selection algorithm to select variables for the best model fit. The remaining variables were then fitted to a PGLS model using the *phylolm* function in the phylolm package in R (Tung Ho & Ané, 2014) to take into account the phylogenetic autocorrelation of our sampled species. The phylogenetic tree used was a subset taken from the fish tree of life project (Rabosky et al., 2018), with the missing *Acanthurus tractus* being replaced by its close relative, *Acanthurus chirurgus* in the phylogeny.

Results

Seascape metrics

The two oceans had comparatively differing seascape characteristics. The Caribbean had a greater total habitable area (383,013 vs 248,661 km^2) as well as the proportion of that area that is habitable for the target species of this study (6 vs 2 % habitable). The Caribbean also had a greater number of habitat patches (163 vs 132), which were also on average larger (2,350 vs 1,884 km^2) and more variable in size (variation of 300 vs 233 km^2).

Table 1: Seascape metrics

Metric	**Caribbean**	**Western Indian Ocean**
Mean Euclidean Nearest Neighbour (ENN) Distance	23	39
ENN Variation	0.12	0.20
Mean Habitable Patch Area (km^2)	2,350	1,884
Total Habitable Area (km^2)	383,013	248,661
Number of Patches	163	132
Total Regional Area (km^2)	6,899,039	14,822,619
Percent of Region Habitable	6	2
Habitat Patch Size Variation	300	233
Edge Density	0.06	0.02

As well as having a greater patch size, the habitat patches were more irregular in shape reflected by their greater edge density (0.06 vs 0.02). Conversely, the Western Indian Ocean as a region covers a much greater pelagic area (6,899,039 vs 14,822,619 km^2) and has a much greater mean distance between patches (23 vs 39 km^2). These measures are summarised in Figure 1B and Table 1.

Sequencing results and α, β, and γ genetic diversity estimates

We sampled and genotyped an additional 15 species comprising 678 individuals at between 1326 to 26,705 loci per species. Of these 15 species, 5 species were removed: *Aulostomus maculatus*, *Rhinesomus triqueter*, *Sparisoma aurofrenatum*, and *Synodus intermedius* were removed from further analysis as they had no close phylogenetic comparison in the Western Indian Ocean; *Caranx ruber*, was removed due to insufficient sampling across sample sites. From the 27 species of the existing Western Indian Ocean data set, 13 species were removed: *Caranx melampygus*, *Myripristis violacea*, *Oxymonacanthus longirostris*, *Parapercis hexophtalma*, *Pseudanthias squamipinnis*, and *Zanclus cornutus* were removed as they lacked comparable species in the Caribbean; *Chromis atripectoralis* and *Gomphosus caeruleus* failed to sequence; *Naso brevirostris* had insufficient individuals per site; and *Lutjanus bengalensis*, *Monotaxis grandoculis*, and *Monotaxis heterodon* were discovered to contain many incorrect species designations. Species removals are summarised in Supplementary Table 3. Of the remaining Caribbean samples, the number of reads per sample per

species had a minimum of 1.19 million, a maximum of 10.43 million, and mean of 3.85 million. This corresponded roughly to a mean coverage per sample per species ranging from a minimum of 4.67, a maximum of 30.16, and a mean of 13.87 (Supplementary Table 2). In both the Caribbean and Western Indian Ocean, multiple SNPs were identified per locus, with a minimum species mean of 2, a maximum of 24, and a mean of 3 SNPs (Supplementary Figure 2). After retaining only one SNP per locus and iteratively subsetting down to the minimum number of SNPs per species, a total of 3715 SNP markers were retained (Supplementary Figure 2).

Variation in the calculation of population genetic metrics was small (Supplementary Figure 4) and subsequent analyses used the mean values across subsampling iterations (Supplementary Table 1). Corresponding to α- and γ-diversity facets, values of mean observed heterozygosity (H_O) across sites ranged from a low of 0.23 (*Lutjanus kasmira*, Lutjanidae) in the Western Indian Ocean to a high of 0.36 (*Chromis multilineata*, Pomacentridae) in the Caribbean. Mean expected heterozygosity (H_S) per site had a minimum value of 0.24 (*Lutjanus kasmira*, Lutjanidae) in the Western Indian Ocean and a maximum of 0.34 (*Oxycheilinus digramma*, Labridae) also in the Western Indian Ocean. Total expected heterozygosity (H_T) across sites was lowest at 0.24 (*Lutjanus kasmira*, Lutjanidae) in the Western Indian Ocean and highest at 0.35 (*Oxycheilinus digramma*, Labridae). Comparing H_O and H_S with the F_{IS} inbreeding index, we find a minimum value of -0.29 (*Chromis multilineata*, Pomacentridae) and a maximum value of 0.14 (*Acanthurus nigrofuscus*, Labridae). In the β-diversity facet, Hedrick's G''_{ST} had a minimum, effectively 0, value of -0.01 (*Mulloidichthys martinicus*), Mullidae) in the Caribbean, and a maximum value of 0.07 (*Hemigymnus fasciatus*, Labridae) in the Western Indian Ocean. We found the α- and γ-diversity facet metrics of H_S and H_T to be highly positively related in both oceans (PGLS: Caribbean, $\beta = 0.81$, $t = 11.29$, $p < 0.01$; Western Indian Ocean, $\beta = 0.99$, $t = 31.17$, $p < 0.01$), but no detectable relationships between Hedrick's G''_{ST} and any of the heterozygosity estimates in either ocean.

Determinants of population α, β, and γ genetic diversity

For α-diversity, only H_O and F_{IS} were significantly different between oceans, with H_O being greater in the Caribbean and F_{IS} greater in the Western Indian Ocean (H_s, $\beta = -0.01$, $t = -0.60$, $p = 0.56$; H_O, $\beta = -0.03$, $t = -2.61$, $p = 0.02$; F_{IS}, $\beta = 0.07$, $t = 2.30$, $p = 0.03$; Figure 2A). For β-diversity, Hedrick's G''_{ST} was greater in the Western Indian Ocean than in the Caribbean (Hedrick's G''_{ST}, $\beta = 0.01$, $t = 2.23$, $p = 0.04$; Figure 2A). Finally, there was no significant difference in the γ-diversity metric (H_T) between the two oceans (H_T, $\beta = 0.00$, $t = -0.55$, $p = 0.59$; Figure 2A). We further visualised the relationship between oceans in each metric by aggregating into orders (Figure 3B), which showed possibly more variable relationships between oceans. For H_O, all orders follow the trend of increased values in the Caribbean compared to the Western Indian Ocean, excluding Labriformes and Acanthuriformes. Similarly, F_{IS} values across orders showed the same increase from the Caribbean to the Western Indian Ocean, apart from Labriformes and Chaetodontiformes. For Hedrick's G''_{ST}, all orders reflected the global pattern. Unfortunately, the number of species within each order across ocean basins was insufficient to perform a statistical comparison per order (Figure 2B). We found no significant relationships between species traits and population genetic metrics representing the three facets of diversity in either ocean using the PGLS framework (Supplementary Figure 5).

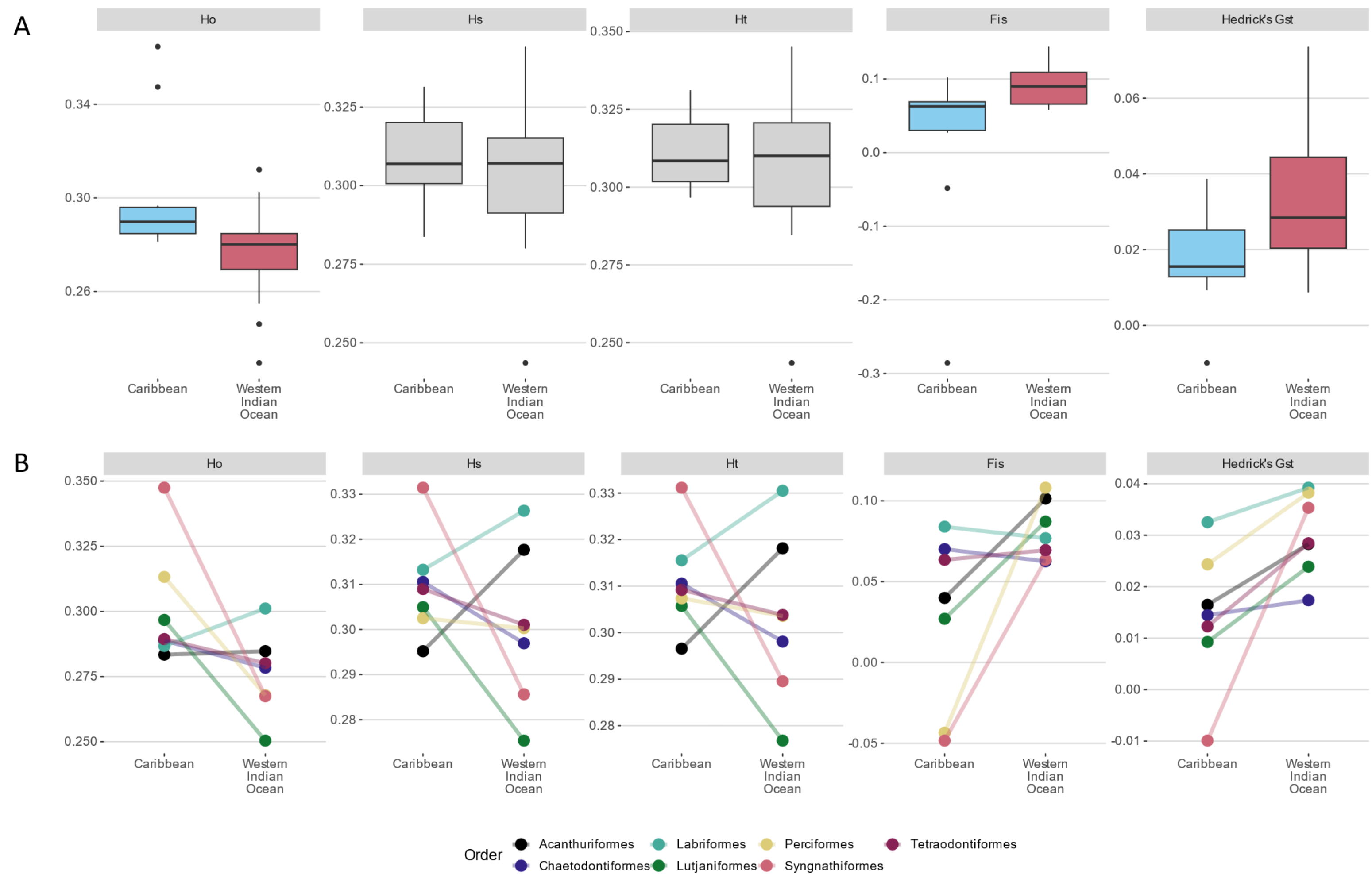


Figure 3: The effect of ocean bioregion on population genetic metrics of observed heterozygosity (H_O), expected mean heterozygosity across sample sites (H_S), expected total heterozygosity per ocean (H_T), the inbreeding fixation index F_{IS} and Hedrick's G_{ST}. A) global comparisons between oceans. B) Comparisons between oceans, subdivided by order. Each order comparison contains only one family, except Lutjaniformes, which contains two members of family Lutjanidae in the Western Indian Ocean, and one member of family Haemulidae in the Caribbean.

Discussion

We aimed to disentangle the relative influence of seascape and biological traits across three spatial facets of diversity, α, β, and γ. We found that that the influence of seascape was dominant across a range of phylogenetically comparable species, whilst biological traits (abundance and dispersal) were not. The influence of seascape on observed heterozygosity and the inbreeding coefficient, F_{IS}, aligned well with traditional expectations of neutral population genetics theory – larger, better connected habitats should support a larger effective population size leading to greater diversity and reduced inbreeding (Crooks et al., 2017; Hamilton, 2021). Similarly, genetic differentiation (Hedrick's G''_{ST}), was greater in the more fragmented seascape of the Western Indian Ocean providing greater barriers to gene-flow (Donati et al., 2021; Hamilton, 2021). The lack of relationship between abundance and dispersal with any of the diversity metrics is likely a result of the compounding of a truncated sampling scheme (Donati et al., 2021) and the relative importance of dispersal across varying seascapes.

The ocean effect was the sole significant predictor of observed heterozygosity (H_O), the sample site inbreeding estimator, F_{IS} (Wright, 1949), and Hedrick's G''_{ST} (Meirmans & Hedrick, 2011). In summary, H_O was overall greater in the Caribbean than in the Western Indian Ocean, whilst both F_{IS} and Hedrick's G''_{ST} were lower in the Caribbean than in the Western Indian Ocean. In the context of established neutral population genetics theory, these results fit with the expectations from the different habitat configurations of these two seascapes. The Caribbean seascape contains a much larger absolute area of reef habitat compared to the Western Indian Ocean as well as larger patch sizes (Figure 2B**,** Table 1). Given all else is equal, this should support a larger population census size, both in each patch and regionally. This greater effective population size is reflected in the elevated abundance values found in the Caribbean compared to the Western Indian Ocean (Edgar et al. 2020). Despite population genetic information containing signal across a much longer timeframe than contemporary visual surveys can capture (Bradburd and Ralph 2019), a greater census population size over time should support a greater effective population size, reducing the probability of inbreeding and increasing genetic diversity (Hamilton, 2021). At a smaller scale, this effect of more habitable area should have a parallel effect to that of having larger patch sizes. Larger patches should be able to support more diversity through larger effective population sizes which are less vulnerable to relatively large population fluctuations and the coupled increased likelihood of diversity loss or even local extinction (Crooks et al. 2017). These local effects of more habitat area per patch driving greater diversity are likely further bolstered by the greater numbers of these patches. Having more populations in a metapopulation reduces the severity of local extinction events (Warren et al., 2015), i.e., if a population is lost, the impact on the greater group of populations is lessened. In the Western Indian Ocean, the converse is true, with less habitat shared across fewer, smaller patches of habitat. With less habitat being able to support a smaller metapopulation consisting of fewer, smaller populations more vulnerable to local extinctions and greater absolute reduction in diversity through population fluctuations.

Hedrick's G''_{ST} showed an inverse pattern to observed heterozygosity and followed a decrease from the Western Indian Ocean to the Caribbean (Figure 2A). The mean greater distances between habitat patches in the Western Indian Ocean compared to the Caribbean (Figure 2**,** Table 1) makes dispersal more difficult between habitat patches, reducing gene-flow and therefore increasing population structure (Donati et al., 2021). This is likely compounded by the smaller patch sizes which in turn support fewer individuals and therefore fewer dispersal events (Poethke & Hovestadt, 2002). The increased population structure in the Western Indian Ocean compared to the Caribbean likely also compounds the difference in observed heterozygosity between the two regions. If populations

are more connected, migrants increase the effective population sizes of each population which should increase the level of genetic diversity across the metapopulation through a shift towards panmixia (Bradburd & Ralph, 2019). This emphasises that whilst we often think of populations as discrete units, it is not necessarily true (Waples & Gaggiotti, 2006). Depending on the level of connectivity between spatially distinct groups of individuals, they could be seen as many small populations, or one large population. The greater the connectivity between populations, the closer they are to being effectively a single, larger population generating greater genetic diversity. Interestingly, this is a reflection of the inverse spatial relationship between heterozygosity and population structure found in a single-species study system (Lamy et al., 2012). Further to the inverse pattern between oceans in genetic diversity and population structure, the Western Indian Ocean would be expected to experience a higher level of inbreeding compared to the Caribbean, as it has a smaller habitat area likely supporting a smaller and more fragmented effective population size. This is supported by lower estimates of F_{IS} in the Caribbean, indicating lower levels of inbreeding as compared to the Western Indian Ocean (Figure 2A). Whilst these dynamics are not new information in themselves, through a macro-genetics approach, they show that established population genetics theory applies across a range of both functionally and phylogenetically distinct fish families, and not just within species – these seascape-based dynamics have a level of universality across highly different species.

We found a lack of significant relationship between species' dispersal trait, last known divergence time, and recent abundance and our population genetic metrics: H_O, H_S, H_T, F_{IS}, and Hedrick's G''_{ST}. This is surprising given that based on a larger subset of the Western Indian Ocean, Donati et al. (2021) found a negative relationship between dispersal traits and both population structure and genetic diversity. The lack of these relationships makes it clear that the relative impact of the habitat configuration is, in this instance, much more prominent than the effect of functional traits and contemporary ecological characteristics. In the Western Indian Ocean, where smaller habitat patches are more isolated and dispersal between them more difficult, biological traits that determine this ability will be more influential. Whereas in the Caribbean where the sampled species have similar dispersal traits (Figure 3B), but where distances between habitat patches are much shorter, this trait will no longer play a dominant role in producing patterns of genetic diversity and population structure. This pattern has already been suggested in the Northwest Caribbean where low dispersal ability had minimal effect in increasing genetic structure across a connected seascape in *Stegastes partitus*, *Thalassoma bifasciatum*, *Haemulon flavolineatum*, *Hypoplectrus nigricans*, and *Chaetodon capistratus* – two of which are included in this study (Puebla et al., 2012). In other words, in the Caribbean context dispersal ability is no longer as relevant. Having an excess of dispersal capacity in a particular habitat configuration will have little influence on genetic patterns. By adding the Caribbean samples to the sampling scheme, we are adding noise to the dispersal-genetic metric relationships observed in the Western Indian Ocean and thereby obscuring them. Whilst a negative result, it highlights an important phenomenon – that biological processes are highly dependent on their environmental context.

Limitations

As with any genetic study, we are limited in our scope by our sampling and methodology. The Caribbean sampling and subsequent ddRADseq library generation was an extension of previous work which we aimed to stay consistent with Donati et al. (2021). However, the passage of time meant that the Illumina HiSeq 2500 platform was no longer a cost-effective option, and the Illumina Novaseq 6000 was chosen due to low sequencing costs and similar error profile to the HiSeq (Stoler & Nekrutenko, 2021). Since the ambiguous barcode bases were incorporated, but not implemented

in the Western Indian Ocean samples, this step was removed and should have a negligible influence on the final results. Analytically, incorporating ocean currents was outside the scope of this study, but given the co-occurrence of likely barriers to gene-flow and known current systems this would be a promising avenue for further work. Statistically, there could always be more extensive sampling; more species per family comparison would allow for within-family inter-ocean comparisons to better understand the generality of the global patterns we uncover in this study, and more species would also allow the statistical power to incorporate interaction effects between biological traits and habitat configurations. Finally, we are here limited to a comparison between only two bioregions, preventing us from identifying the exact characteristics of the seascapes that drive the described differences in genetic structure and diversity in each region. However, this is borderline fantasy given the considerable effort required to obtain the existing samples across multiple jurisdictions at such a broad geographic scale.

Conclusions

In conclusion, we found an increase in genetic diversity in the Caribbean compared to the Western Indian Ocean which reflected an opposite pattern in both population structure and an inbreeding index. By themselves these are likely a result of established population genetics processes playing out across a functionally and phylogenetically diverse set of tropical fish lineages – highlighting the universality of the influence of the environment on observed genetic patterns. Interestingly, expanding the sampling to the Caribbean from only the Western Indian Ocean (Donati et al., 2021), we lose the signal for the relationship between dispersal and genetic diversity. We attribute this to the greater connectivity of the Caribbean and the subsequent reduction in importance of dispersal traits in this habitat configuration – i.e., traits are highly dependent on environmental context.

Ethics

The collected data have no commercial value and cannot be used in a way that could be detrimental to local populations. They were generated in collaboration with the Genetic Diversity Centre (GDC), ETH Zurich.

Sampling was performed in accordance with local regulations and with local collaborators, research permits are the following: Maldives (OTHR) 30-D/INDIV/2016/538); Mayotte (06/UTM/2016); Seychelles (A0157); Tanzania (2017-242-NA-2017-87); Curaçao under sampling permit 2012 / 48584 provided to CARMABI (Piscaderabaai 00000, Willemstad, Curaçao; a registered CITES Institution under Code AN001) by the Curaçaon Government; Martinique, decision number 122 of the director of the sea of Martinique. Colombian samples were collected in collaboration with INVEMAR, through which all samples were collected and exported as stated in the following: “According to Paragraph 1, Article 2.2.2.2.8.1.2. Section 1 (Permits), Chapter 8 (Scientific Research), of Decree 1076 of 2015 "The Ministry of Environment and Sustainable Development, its attached entities, National Natural Parks of Colombia, the Regional Autonomous Corporations and/or of Sustainable Development and the Large Urban Centers shall not require the Permit to Collect specimens referred to in this decree (...)"; therefore, INVEMAR being an entity attached to the Ministry of Environment and Sustainable Development (see Article 1. 2.2.1., Title 2, of Decree 1076 of 2015), does not require permission to collect specimens of wild species. In this sense, the material to be exported comes from projects developed by INVEMAR, which supports the legal acquisition of the specimens to be exported.”

Acknowledgements

This project was supported by the Monaco Explorations Foundation, CORALINA, which provided support for entrance to the island of Providencia for the development of the project and maritime

support for the field sampling. We express special thanks to Nicasio and Casimiro for providing guidance at sea, as well as Oskar Franz and Alcides Martinez.

Sample extraction, sequencing library preparation and quantification, and bioinformatics steps was carried out in collaboration with the Genetic Diversity Centre, ETH Zürich. The sequencing for the Caribbean samples was carried out at the Functional Genomics Centre Zürich, ETH Zürich and University of Zürich. We would like to thank both facilities for their invaluable support, guidance, but most of all patience. In particular we would like to thank Silvia Kobel, Aria Minder, and Niklaus Zemp. We would also like to thank Alex Skeels for his freely given discussions, statistical help, and knowledge of phylogenetic methods.

Funding was supported by the ETH research grant financing the salary of Thomas Keggin (ETH Research Grant ETH-34 18-1) and the Reefish project funded by the Swiss National Science Foundation (grant number: 310030E-164294).

References

Anderson, E., & Hubricht, L. (1938). Hybridization in Tradescantia. III. The evidence for introgressive hybridization. *American Journal of Botany*, 396-402.

Ashe, J. L., Feldheim, K. A., Fields, A. T., Reyier, E. A., Brooks, E. J., O'Connell, M. T., Skomal, G., Gruber, S. H., & Chapman, D. D. (2015). Local population structure and context-dependent isolation by distance in a large coastal shark. *Marine Ecology Progress Series*, *520*, 203-216. https://doi.org/10.3354/meps11069

Boettiger, C., Lang, D. T., & Wainwright, P. (2012). rfishbase: exploring, manipulating and visualizing FishBase data from R. *Journal of Fish Biology*, *81*(6), 2030-2039.

Bradburd, G. S., & Ralph, P. L. (2019). Spatial Population Genetics: It's About Time. *Annual Review of Ecology, Evolution, and Systematics*, *50*(1), 427-449. https://doi.org/10.1146/annurev-ecolsys-110316-022659

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 [Article]. *Journal of Evolutionary Biology*, *24*(12), 2543-2562. https://doi.org/10.1111/j.1420-9101.2011.02391.x

Crooks, K. R., Burdett, C. L., Theobald, D. M., King, S. R. B., Di Marco, M., Rondinini, C., & Boitani, L. (2017). Quantification of habitat fragmentation reveals extinction risk in terrestrial mammals. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(29), 7635-7640. https://doi.org/10.1073/pnas.1705769114

Donati, G. F. A., Zemp, N., Manel, S., Poirier, M., Claverie, T., Ferraton, F., Gaboriau, T., Govinden, R., Hagen, O., Ibrahim, S., Mouillot, D., Leblond, J., Julius, P., Velez, L., Zareer, I., Ziyad, A., Leprieur, F., Albouy, C., & Pellissier, L. (2021). Species ecology explains the spatial components of genetic diversity in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1959), 20211574. https://doi.org/doi:10.1098/rspb.2021.1574

Dray, S., Pelissier, R., Couteron, P., Fortin, M. J., Legendre, P., Peres-Neto, P. R., Bellier, E., Bivand, R., Blanchet, F. G., De Caceres, M., Dufour, A. B., Heegaard, E., Jombart, T., Munoz, F., Oksanen, J., Thioulouse, J., & Wagner, H. H. (2012). Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs*, *82*(3), 257-275. https://doi.org/10.1890/11-1183.1

Edgar, G. J., Cooper, A., Baker, S. C., Barker, W., Barrett, N. S., Becerro, M. A., Bates, A. E., Brock, D., Ceccarelli, D. M., Clausius, E., Davey, M., Davis, T. R., Day, P. B., Green, A., Griffiths, S. R., Hicks, J., Hinojosa, I. A., Jones, B. K., Kininmonth, S., . . . Stuart-Smith, R. D. (2020). Establishing the ecological basis for conservation of shallow marine life using Reef Life Survey. *Biological Conservation*, *252*, 108855. https://doi.org/https://doi.org/10.1016/j.biocon.2020.108855

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. https://doi.org/10.1038/sdata.2014.7

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422-433. https://doi.org/10.1038/nrg.2016.58

Euclide, P. T., McKinney, G. J., Bootsma, M., Tarsa, C., Meek, M. H., & Larson, W. A. (2020). Attack of the PCR clones: Rates of clonality have little effect on RAD-seq genotype calls. *Molecular Ecology Resources*, *20*(1), 66-78. https://doi.org/https://doi.org/10.1111/1755-0998.13087

Evans, K., & Gaston, K. (2005). Can the evolutionary-rates hypothesis explain species-energy relationships? *Functional Ecology*, *19*(6), 899-915.

Frankham, R. (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*, *24*(11), 2610-2618. https://doi.org/https://doi.org/10.1111/mec.13139

Gaither, M. R., Gkafas, G. A., de Jong, M., Sarigol, F., Neat, F., Regnier, T., Moore, D., Gröcke, D. R., Hall, N., Liu, X., Kenny, J., Lucaci, A., Hughes, M., Haldenby, S., & Hoelzel, A. R. (2018). Genomics of habitat choice and adaptive evolution in a deep-sea fish. *Nature Ecology & Evolution*, *2*(4), 680-687. https://doi.org/10.1038/s41559-018-0482-x

Gaither, M. R., Violi, B., Gray, H. W. I., Neat, F., Drazen, J. C., Grubbs, R. D., Roa-Varón, A., Sutton, T., & Hoelzel, A. R. (2016). Depth as a driver of evolution in the deep sea: Insights from grenadiers (Gadiformes: Macrouridae) of the genus Coryphaenoides. *Molecular Phylogenetics and Evolution*, *104*, 73-82. https://doi.org/https://doi.org/10.1016/j.ympev.2016.07.027

Gaston, K. J., & Spicer, J. I. (2013). *Biodiversity: an introduction*. John Wiley & Sons.

Geburzi, J. C., Heuer, N., Homberger, L., Kabus, J., Moesges, Z., Ovenbeck, K., Brandis, D., & Ewers, C. (2022). An environmental gradient dominates ecological and genetic differentiation of marine invertebrates between the North and Baltic Sea. *Ecology and Evolution*, *12*(5), e8868. https://doi.org/https://doi.org/10.1002/ece3.8868

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*(1), 184-186.

Greenwald, W. W., Li, H., Benaglio, P., Jakubosky, D., Matsui, H., Schmitt, A., Selvaraj, S., D'Antonio, M., D'Antonio-Chronowska, A., Smith, E. N., & Frazer, K. A. (2019). Subtle changes in chromatin loop contact propensity are associated with differential gene regulation and expression. *Nature Communications*, *10*(1), 1054. https://doi.org/10.1038/s41467-019-08940-5

Hamilton, M. B. (2021). *Population genetics*. John Wiley & Sons.

Hare, M. P., Guenther, C., & Fagan, W. F. (2005). Nonrandom larval dispersal can steepen marine clines. *Evolution*, *59*(12), 2509-2517. https://doi.org/10.1554/05-150.1

He, F., Pasam, R., Shi, F., Kant, S., Keeble-Gagnere, G., Kay, P., Forrest, K., Fritz, A., Hucl, P., & Wiebe, K. (2019). Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nature Genetics*, *51*(5), 896-904.

Hesselbarth, M. H., Sciaini, M., With, K. A., Wiegand, K., & Nowosad, J. (2019). landscapemetrics: an open-source R tool to calculate landscape metrics. *Ecography*, *42*(10), 1648-1657.

Hibdige, S. G. S., Raimondeau, P., Christin, P. A., & Dunning, L. T. (2021). Widespread lateral gene transfer among grasses. *New Phytologist*, *230*(6), 2474-2486. https://doi.org/10.1111/nph.17328

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, *30*(1), 62-82. https://doi.org/https://doi.org/10.1111/mec.15720

Holderegger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, *21*(6), 797-807. https://doi.org/10.1007/s10980-005-5245-9

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403-1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070-3071. https://doi.org/10.1093/bioinformatics/btr521

Kulbicki, M., Parravicini, V., Bellwood, D. R., Arias-Gonzàlez, E., Chabanet, P., Floeter, S. R., Friedlander, A., McPherson, J., Myers, R. E., Vigliola, L., & Mouillot, D. (2014). Global Biogeography of Reef Fishes: A Hierarchical Quantitative Delineation of Regions. *Plos One*, *8*(12), e81847. https://doi.org/10.1371/journal.pone.0081847

Lamy, T., Pointier, J. P., Jarne, P., & David, P. (2012). Testing metapopulation dynamics using genetic, demographic and ecological data. *Molecular Ecology*, *21*(6), 1394-1410. https://doi.org/https://doi.org/10.1111/j.1365-294X.2012.05478.x

Lehnert, S. J., DiBacco, C., Van Wyngaarden, M., Jeffery, N. W., Ben Lowen, J., Sylvester, E. V. A., Wringe, B. F., Stanley, R. R. E., Hamilton, L. C., & Bradbury, I. R. (2019). Fine-scale temperature-associated genetic structure between inshore and offshore populations of sea scallop (Placopecten magellanicus). *Heredity*, *122*(1), 69-80. https://doi.org/10.1038/s41437-018-0087-9

Leigh, D. M., van Rees, C. B., Millette, K. L., Breed, M. F., Schmidt, C., Bertola, L. D., Hand, B. K., Hunter, M. E., Jensen, E. L., Kershaw, F., Liggins, L., Luikart, G., Manel, S., Mergeay, J., Miller, J. M., Segelbacher, G., Hoban, S., & Paz-Vinas, I. (2021). Opportunities and challenges of macrogenetic studies. *Nature Reviews Genetics*. https://doi.org/10.1038/s41576-021-00394-0

Leis, J. M. (2020). Measurement of swimming ability in larval marine fishes: comparison of critical speed with in situ speed. *Marine Ecology Progress Series*, *650*, 203-215. https://doi.org/10.3354/meps13233

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*, *110*(41), 16498-16502. https://doi.org/doi:10.1073/pnas.1304074110

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology & Evolution*, *28*(10), 614-621. https://doi.org/10.1016/j.tree.2013.05.012

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, *18*(4), 189-197. https://doi.org/10.1016/s0169-5347(03)00008-9

Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources*, *11*(1), 5-18. https://doi.org/https://doi.org/10.1111/j.1755-0998.2010.02927.x

Parravicini, V., Bender, M. G., Villéger, S., Leprieur, F., Pellissier, L., Donati, F. G. A., Floeter, S. R., Rezende, E. L., Mouillot, D., & Kulbicki, M. (2021). Coral reef fishes reveal strong divergence in the prevalence of traits along the global diversity gradient. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1961), 20211712. https://doi.org/doi:10.1098/rspb.2021.1712

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *Plos One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Poethke, H. J., & Hovestadt, T. (2002). Evolution of density–and patch–size–dependent dispersal rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*(1491), 637-645. https://doi.org/doi:10.1098/rspb.2001.1936

Potts, D. C. (1984). Generation times and the quaternary evolution of reef-building corals. *Paleobiology*, *10*(1), 48-58. https://doi.org/10.1017/s0094837300008010

Puebla, O., Bermingham, E., & McMillan, W. O. (2012). On the spatial scale of dispersal in coral reef fishes. *Molecular Ecology*, *21*(23), 5675-5688. https://doi.org/10.1111/j.1365-294X.2012.05734.x

R Core Team. (2022). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. https://www.R-project.org/

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*(7714), 392-395. https://doi.org/10.1038/s41586-018-0273-1

Renema, W., Bellwood, D. R., Braga, J. C., Bromfield, K., Hall, R., Johnson, K. G., Lunt, P., Meyer, C. P., McMonagle, L. B., Morley, R. J., O'Dea, A., Todd, J. A., Wesselingh, F. P., Wilson, M. E. J., & Pandolfi, J. M. (2008). Hopping hotspots: Global shifts in marine Biodiversity. *Science*, *321*(5889), 654-657. https://doi.org/10.1126/science.1155674

Rocha, L. A., Craig, M. T., & Bowen, B. W. (2007). Phylogeography and the conservation of coral reef fishes. *Coral Reefs*, *26*(3), 501-512. https://doi.org/10.1007/s00338-007-0261-7

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., Loire, E., Lourenco, J. M., Nabholz, B., Roux, C., Tsagkogeorga, G., Weber, A. A. T., Weinert, L. A., Belkhir, K., Bierne, N., . . . Galtier, N. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, *515*(7526), 261-263. https://doi.org/10.1038/nature13685

Selkoe, K. A., D'Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., von der Heyden, S., & Toonen, R. J. (2016). A decade of seascape genetics: contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, *554*, 1-19. https://doi.org/10.3354/meps11792

Selkoe, K. A., Henzler, C. M., & Gaines, S. D. (2008). Seascape genetics and the spatial ecology of marine populations. *Fish and Fisheries*, *9*(4), 363-377. https://doi.org/10.1111/j.1467-2979.2008.00300.x

Singhal, S., Huang, H., Grundler, M. R., Marchán-Rivadeneira, M. R., Holmes, I., Title, P. O., Donnellan, S. C., & Rabosky, D. L. (2018). Does Population Structure Predict the Rate of Speciation? A Comparative Test across Australia’s Most Diverse Vertebrate Radiation. *The American Naturalist*, *192*(4), 432-447. https://doi.org/10.1086/699515

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*(7), 573-583. https://doi.org/10.1641/b570707

Stoler, N., & Nekrutenko, A. (2021). Sequencing error profiles of Illumina sequencing instruments. *NAR Genomics and Bioinformatics*, *3*(1). https://doi.org/10.1093/nargab/lqab019

Thompson, D. M., Kleypas, J., Castruccio, F., Curchitser, E. N., Pinsky, M. L., Jonsson, B., & Watson, J. R. (2018). Variability in oceanographic barriers to coral larval dispersal: Do currents shape biodiversity? *Progress in Oceanography*, *165*, 110-122. https://doi.org/10.1016/j.pocean.2018.05.007

Tung Ho, L. s., & Ané, C. (2014). A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Systematic Biology*, *63*(3), 397-408.

UNEP-WCMC, W. C., WRI, TNC. (2010). *Global distribution of warm-water coral reefs, compiled from multiple sources including the Millennium Coral Reef Mapping Project* Version 4.0). https://doi.org/https://doi.org/10.34892/t2wk-5t34

Waples, R. S., & Gaggiotti, O. (2006). INVITED REVIEW: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, *15*(6), 1419-1439. https://doi.org/https://doi.org/10.1111/j.1365-294X.2006.02890.x

Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilee, R., Condamine, F. L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernandez-Palacios, J. M.,

Hengl, T., Norder, S. J., Rijsdijk, K. F., Sanmartin, I., Strasberg, D., Triantis, K. A., . . . Thebaud, C. (2015). Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters*, *18*(2), 200-217. https://doi.org/10.1111/ele.12398

White, T. A., Fotherby, H. A., & Hoelzel, A. R. (2011). Comparative assessment of population genetics and demographic history of two congeneric deep sea fish species living at different depths. *Marine Ecology Progress Series*, *434*, 155-164. https://doi.org/10.3354/meps09207

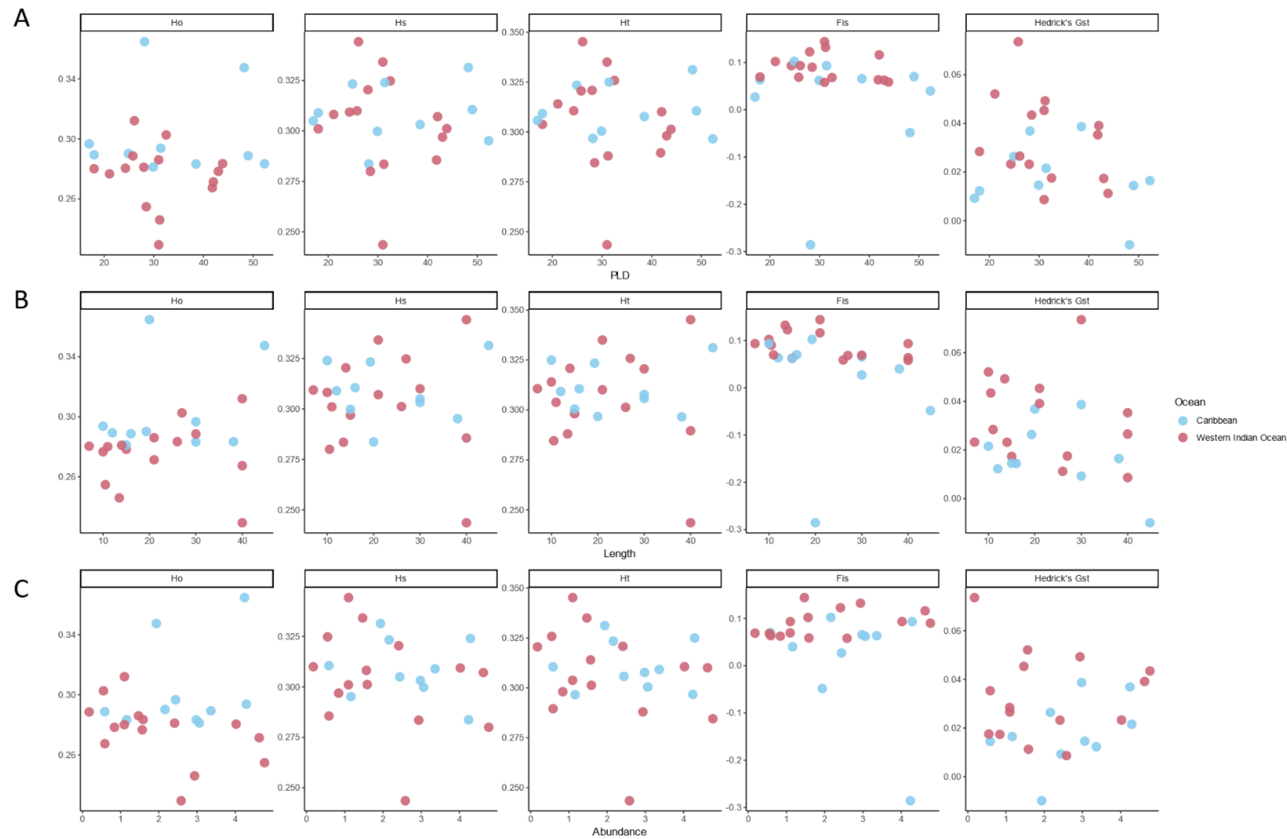
Winter, D. J. (2012). MMOD: an R library for the calculation of population differentiation statistics. *Molecular Ecology Resources*, *12*(6), 1158-1160.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

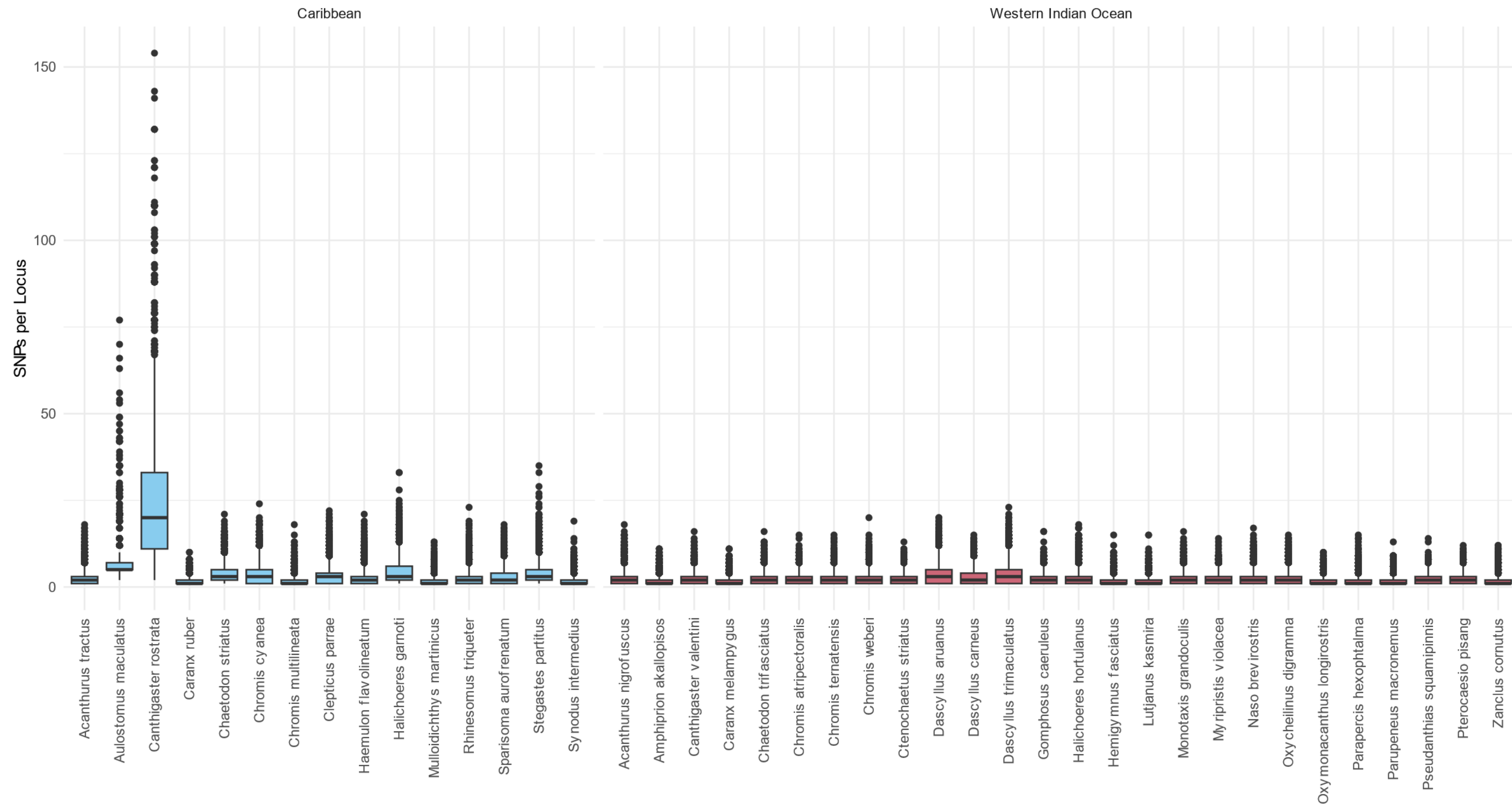
Wright, S. (1949). The genetical structure of populations. *Annals of eugenics*, *15*(1), 323-354.

Wright, S. I., Kalisz, S., & Slotte, T. (2013). Evolutionary consequences of self-fertilization in plants. *Proceedings of the Royal Society B-Biological Sciences*, *280*(1760), Article 20130133. https://doi.org/10.1098/rspb.2013.0133

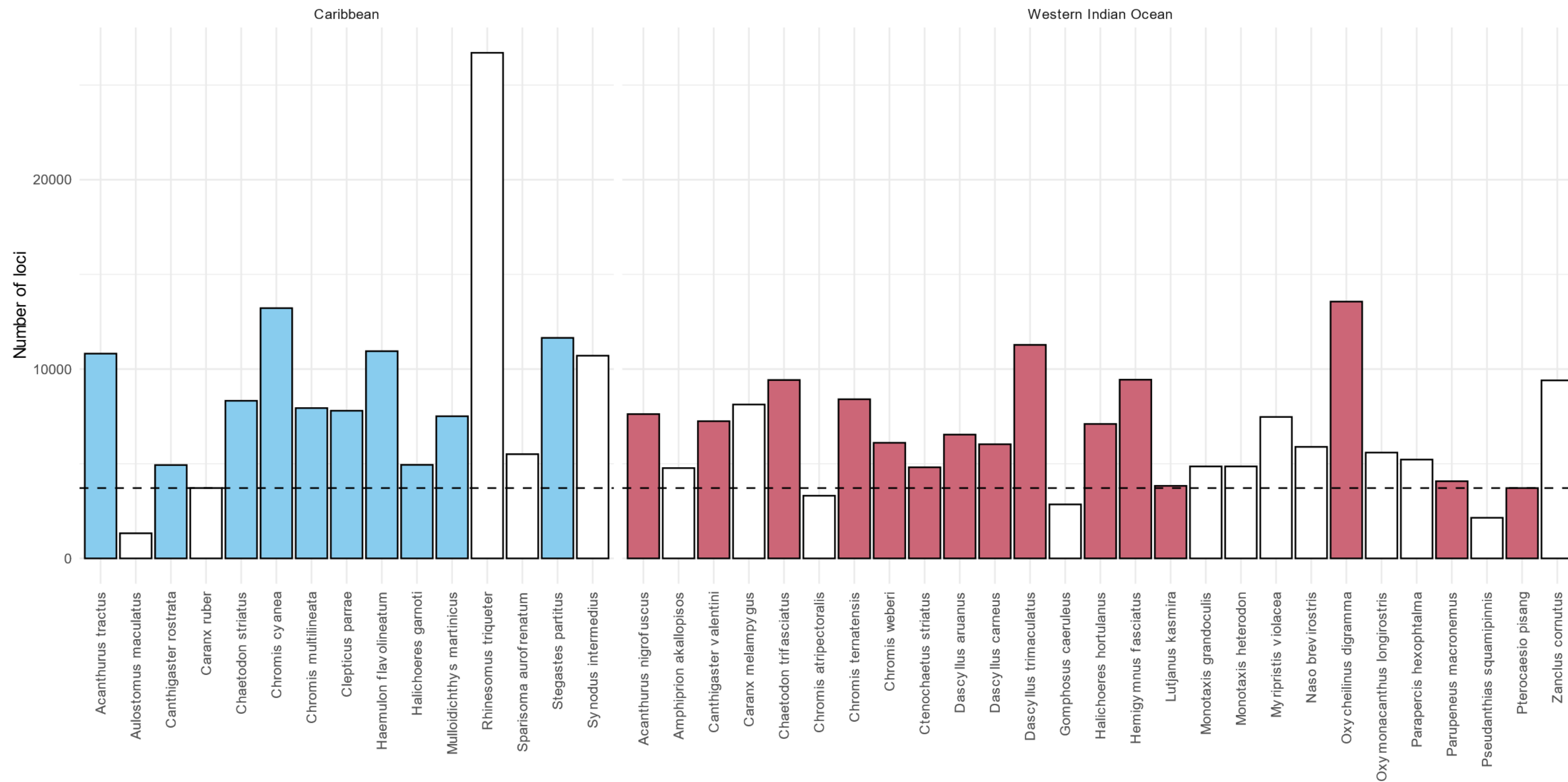
Supplementary



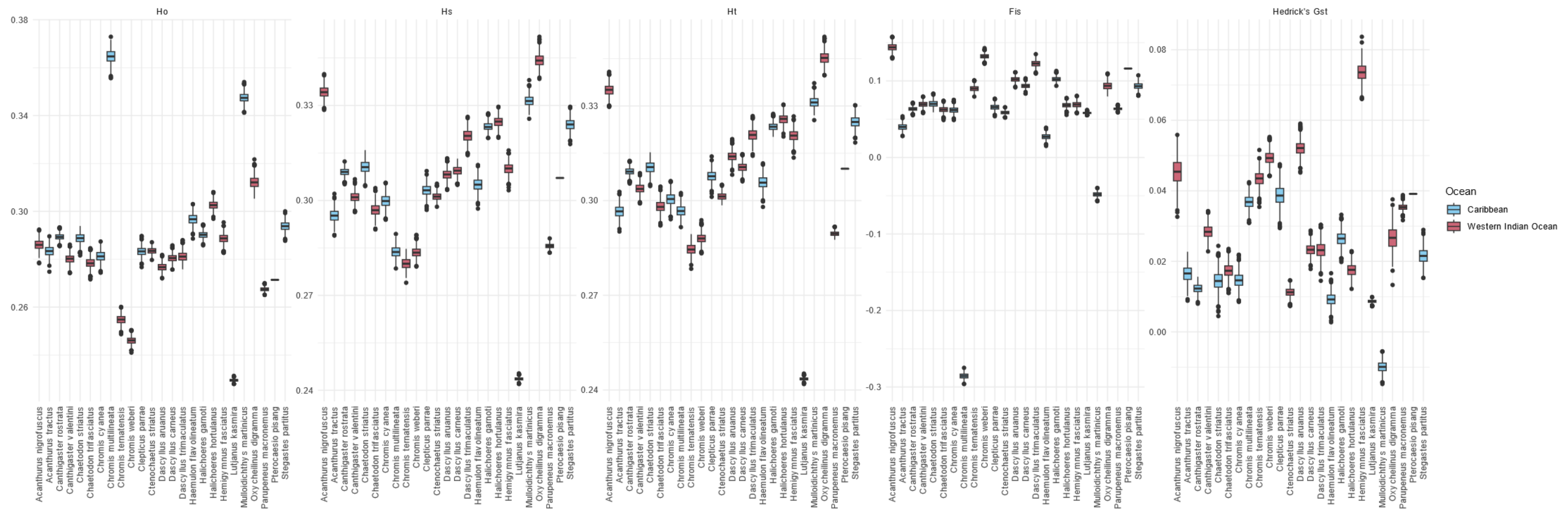
Supplementary Figure 1: Species characteristics plots. A) The pelagic larval duration in days against the population genetic metrics, of observed heterozygosity (H_O), expected mean heterozygosity across sample sites (H_S), expected total heterozygosity per ocean (H_T), the inbreeding fixation index F_{IS} and Hedrick's G_{ST}. B) The maximum body length against the same population genetic metrics as A. C) The same as A and B, but the trait is abundance. Light blue and dark red points represent species from the Caribbean and Western Indian Ocean, respectively.



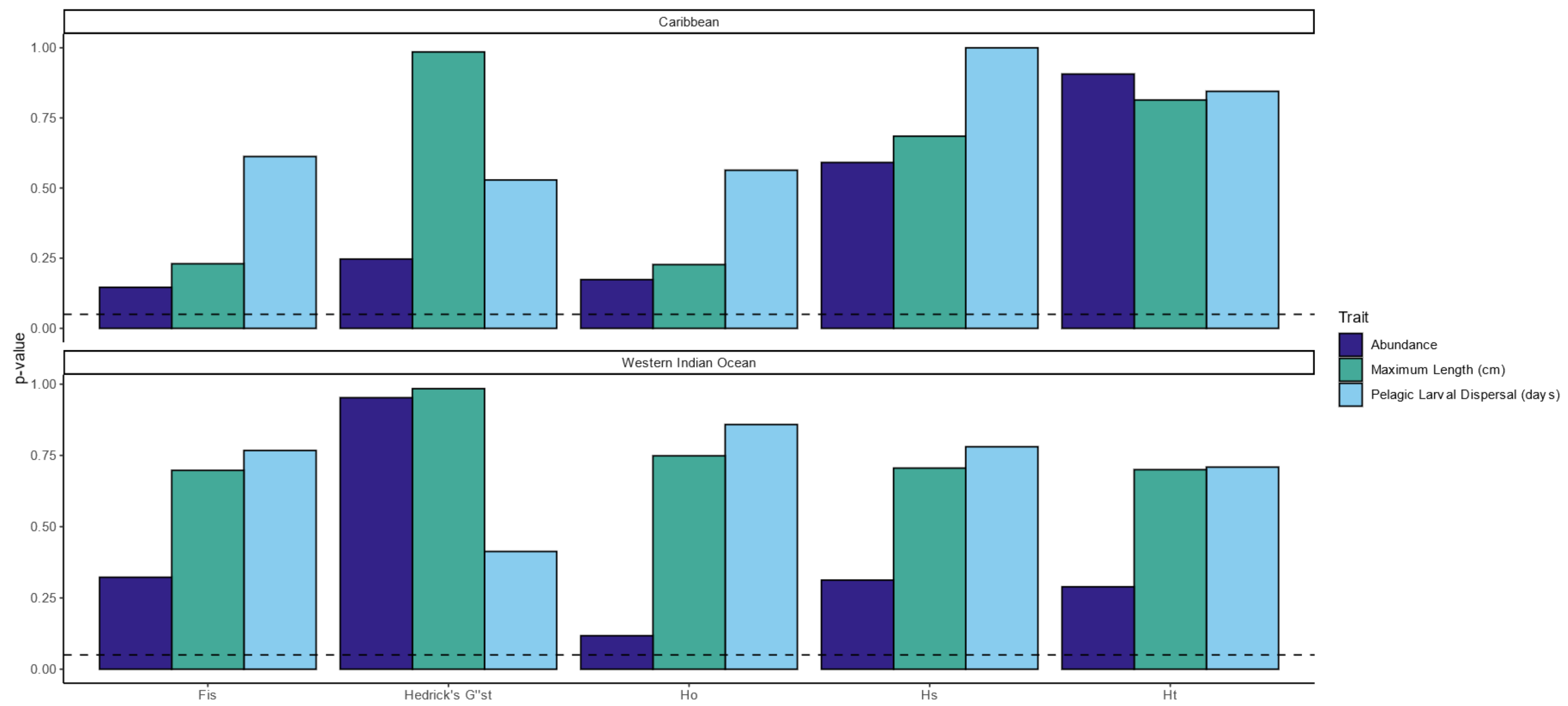
Supplementary Figure 2: Variation in the number of SNPs per locus before equalising to 1 SNP per locus. Keeping this variation would bias metric signal to loci that are overrepresented by SNPs.



Supplementary Figure 3: The number of loci per species before equalisation. Blank columns represent species that were dropped from the analysis. The horizontal dashed line corresponds to the minimum number of loci (*Pterocaesio pisang*) in the retained species to which all species were iteratively subsampled.



Supplementary Figure 4: Variation in metric estimates between iterations of the SNP subsampling step.



Supplementary Figure 5: p-values for the relationships between species traits and the population genetic metrics in the Caribbean and Western Indian Ocean. The horizontal dashed line corresponds to a p-value of 0.05.

Supplementary Table 1: Trait and population genetic metric data used for all parts of the analysis.

Ocean	Family	Species	Pelagic Larval Duration (days)	Maximum Length (cm)	Dispersal	Abundance	Ho	Hs	Ht	Fis	Hedrick's Gst
Caribbean	Acanthuridae	*Acanthurus tractus*	52	38	1993	3.1	0.283	0.295	0.297	0.040	0.016
Caribbean	Tetraodontidae	*Canthigaster rostrata*	18	12	216	41.6	0.289	0.309	0.309	0.064	0.012
Caribbean	Chaetodontidae	*Chaetodon striatus*	49	16	784	1.7	0.289	0.311	0.311	0.070	0.014
Caribbean	Pomacentridae	*Chromis cyanea*	30	15	449	22.7	0.281	0.300	0.300	0.062	0.015
Caribbean	Pomacentridae	*Chromis multilineata*	28	20	563	81.7	0.364	0.283	0.296	-0.285	0.037
Caribbean	Labridae	*Clepticus parrae*	39	30	1155	19.0	0.283	0.303	0.308	0.066	0.039
Caribbean	Haemulidae	*Haemulon flavolineatum*	17	30	510	12.5	0.297	0.305	0.306	0.027	0.009
Caribbean	Labridae	*Halichoeres garnoti*	25	19	481	9.4	0.290	0.323	0.323	0.103	0.027
Caribbean	Mullidae	*Mulloidichthys martinicus*	48	45	2159	6.1	0.348	0.332	0.331	-0.048	-0.010
Caribbean	Pomacentridae	*Stegastes partitus*	31	10	314	77.1	0.294	0.324	0.325	0.093	0.022
Western Indian Ocean	Acanthuridae	*Acanthurus nigrofuscus*	31	21	651	4.3	0.286	0.334	0.335	0.144	0.046
Western Indian Ocean	Tetraodontidae	*Canthigaster valentini*	18	11	198	2.9	0.280	0.301	0.304	0.070	0.028
Western Indian Ocean	Chaetodontidae	*Chaetodon trifasciatus*	43	15	645	2.4	0.278	0.297	0.298	0.063	0.017
Western Indian Ocean	Pomacentridae	*Chromis ternatensis*	29	11	299	114.5	0.255	0.280	0.284	0.090	0.043
Western Indian Ocean	Pomacentridae	*Chromis weberi*	31	14	421	17.0	0.246	0.283	0.288	0.132	0.049
Western Indian Ocean	Acanthuridae	*Ctenochaetus striatus*	44	26	1141	5.3	0.283	0.301	0.301	0.059	0.011
Western Indian Ocean	Pomacentridae	*Dascyllus aruanus*	21	10	211	4.9	0.277	0.308	0.314	0.102	0.052
Western Indian Ocean	Pomacentridae	*Dascyllus carneus*	24	7	170	55.0	0.281	0.310	0.311	0.093	0.023
Western Indian Ocean	Pomacentridae	*Dascyllus trimaculatus*	28	14	392	11.2	0.281	0.320	0.321	0.123	0.023
Western Indian Ocean	Labridae	*Halichoeres hortulanus*	33	27	878	1.8	0.303	0.325	0.326	0.068	0.018
Western Indian Ocean	Labridae	*Hemigymnus fasciatus*	26	30	774	1.2	0.288	0.310	0.321	0.072	0.073
Western Indian Ocean	Lutjanidae	*Lutjanus kasmira*	31	40	1240	13.3	0.230	0.244	0.244	0.058	0.009
Western Indian Ocean	Labridae	*Oxycheilinus digramma*	26	40	1044	3.0	0.312	0.344	0.345	0.093	0.027
Western Indian Ocean	Mullidae	*Parupeneus macronemus*	42	40	1672	1.8	0.268	0.286	0.290	0.063	0.035
Western Indian Ocean	Lutjanidae	*Pterocaesio pisang*	42	21	882	118.5	0.271	0.307	0.310	0.116	0.039

Supplementary Table 2: Overview of the number of fragments sequenced for each species and their respective sequencing coverage after mapping to *de novo* catalogue loci.

Species	**Maximum Reads (million)**	**Minimum Reads (million)**	**Mean Reads (million)**	**Maximum Coverage**	**Minimum Coverage**	**Mean Coverage**
Acanthurus tractus	4.85	0.86	2.56	15.43	2.35	8.06
Aulostomus maculatus	13.03	0.14	2.95	49.19	0.42	11.02
Canthigaster rostrata	7.82	0.32	3.77	53.41	1.14	24.09
Caranx ruber	18.09	4.20	10.43	50.69	11.55	29.41
Chaetodon striatus	3.00	0.36	1.19	13.40	1.70	5.53
Chromis cyanea	6.72	0.81	2.77	32.45	3.48	12.98
Chromis multilineata	6.93	1.42	3.65	24.84	4.96	12.63
Clepticus parrae	8.81	0.59	3.73	33.53	2.28	13.45
Haemulon flavolineatum	5.26	0.94	2.63	17.05	3.02	8.34
Halichoeres garnoti	14.83	3.26	8.51	53.85	11.44	30.16
Mulloidichthys martinicus	15.11	0.80	4.09	77.32	4.31	20.86
Rhinesomus triqueter	6.78	1.81	3.90	17.16	4.22	9.60
Sparisoma aurofrenatum	7.29	0.85	2.40	14.24	1.62	4.67
Stegastes partitus	5.38	0.87	2.39	15.38	2.48	7.15
Synodus intermedius	11.90	0.49	2.77	43.85	1.70	10.07

Supplementary Table 3: Overview of the decisions to remove species from the analyses. Mis-identified species were removed as the number of mis-identified individuals reduced the sample number per population to unusable numbers.

Ocean	Species	Reason
Caribbean	*Aulostomus maculatus*	No oceanic comparison
Caribbean	*Caranx ruber*	Insufficient sampling
Caribbean	*Rhinesomus triqueter*	No oceanic comparison
Caribbean	*Sparisoma aurofrenatum*	No oceanic comparison
Caribbean	*Synodus intermedius*	No oceanic comparison
Western Indian Ocean	*Amphiprion akallopisos*	Insufficient sampling
Western Indian Ocean	*Caranx melampygus*	No oceanic comparison
Western Indian Ocean	*Chromis atripectoralis*	Problematic sequencing
Western Indian Ocean	*Gomphosus caeruleus*	Problematic sequencing
Western Indian Ocean	*Lutjanus bengalensis*	Mis-identification
Western Indian Ocean	*Monotaxis grandoculis*	Mis-identification
Western Indian Ocean	*Monotaxis heterodon*	Mis-identification
Western Indian Ocean	*Myripristis violacea*	No oceanic comparison
Western Indian Ocean	*Naso brevirostris*	Insufficient sampling
Western Indian Ocean	*Oxymonacanthus longirostris*	No oceanic comparison
Western Indian Ocean	*Parapercis hexophtalma*	No oceanic comparison
Western Indian Ocean	*Pseudanthias squamipinnis*	No oceanic comparison
Western Indian Ocean	*Zanclus cornutus*	No oceanic comparison

Supplementary Table 4: Oligonucleotide sequences used for the ddRADseq P1 and P2 adapters and custom barcodes, as well as the dual-indexed Illumina indices.

Name	OligoSequence
GCATG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCATG
AACCA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACCA
CGATC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATC
TCGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGAT
TGCAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCAT
CAACC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACC
GGTTG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTTG
AAGGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGGA
AGCTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTA
ACACA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACACA
AATTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAATTA
ACGGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGGT
ACTGG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGG
ACTTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTTC
ATACG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATACG
ATGAG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATGAG
ATTAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATTAC
CATAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATAT
CGAAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGAAT
CGGCT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGCT
CGGTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGTA
CGTAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTAC
CGTCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTCG
CTGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGAT
CTGCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGCG
CTGTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGTC
CTTGG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGG
GACAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGACAC
GAGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGAT
GAGTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGTC
GCCGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCGT
GCTGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTGA
GGATA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGATA
GGCCA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCCA
GGCTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCTC
GTAGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTAGT
GTCCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCCG
GTCGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCGA
TACCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACCG
TACGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACGT
TAGTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTAGTA
TATAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTATAC
TCACG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCACG
TCAGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCAGT
TCCGG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCCGG
TCTGC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGC
TGGAA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGAA

TTACC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTACC
GCATG_EcoRI_P1.2	/5Phos/AATTCATGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AACCA_EcoRI_P1.2	/5Phos/AATTTGGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGATC_EcoRI_P1.2	/5Phos/AATTGATCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCGAT_EcoRI_P1.2	/5Phos/AATTATCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TGCAT_EcoRI_P1.2	/5Phos/AATTATGCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CAACC_EcoRI_P1.2	/5Phos/AATTGGTTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGTTG_EcoRI_P1.2	/5Phos/AATTCAACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AAGGA_EcoRI_P1.2	/5Phos/AATTTCCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AGCTA_EcoRI_P1.2	/5Phos/AATTTAGCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACACA_EcoRI_P1.2	/5Phos/AATTTGTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AATTA_EcoRI_P1.2	/5Phos/AATTTAATTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACGGT_EcoRI_P1.2	/5Phos/AATTACCGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACTGG_EcoRI_P1.2	/5Phos/AATTCCAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACTTC_EcoRI_P1.2	/5Phos/AATTGAAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATACG_EcoRI_P1.2	/5Phos/AATTCGTATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATGAG_EcoRI_P1.2	/5Phos/AATTCTCATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATTAC_EcoRI_P1.2	/5Phos/AATTGTAATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CATAT_EcoRI_P1.2	/5Phos/AATTATATGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGAAT_EcoRI_P1.2	/5Phos/AATTATTCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGCT_EcoRI_P1.2	/5Phos/AATTAGCCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGTA_EcoRI_P1.2	/5Phos/AATTTACCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGTAC_EcoRI_P1.2	/5Phos/AATTGTACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGTCG_EcoRI_P1.2	/5Phos/AATTCGACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGAT_EcoRI_P1.2	/5Phos/AATTATCAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGCG_EcoRI_P1.2	/5Phos/AATTCGCAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGTC_EcoRI_P1.2	/5Phos/AATTGACAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTTGG_EcoRI_P1.2	/5Phos/AATTCCAAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GACAC_EcoRI_P1.2	/5Phos/AATTGTGTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GAGAT_EcoRI_P1.2	/5Phos/AATTATCTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GAGTC_EcoRI_P1.2	/5Phos/AATTGACTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCCGT_EcoRI_P1.2	/5Phos/AATTACGGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCTGA_EcoRI_P1.2	/5Phos/AATTTCAGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGATA_EcoRI_P1.2	/5Phos/AATTTATCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGCCA_EcoRI_P1.2	/5Phos/AATTTGGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGCTC_EcoRI_P1.2	/5Phos/AATTGAGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTAGT_EcoRI_P1.2	/5Phos/AATTACTACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTCCG_EcoRI_P1.2	/5Phos/AATTCGGACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTCGA_EcoRI_P1.2	/5Phos/AATTTCGACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TACCG_EcoRI_P1.2	/5Phos/AATTCGGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TACGT_EcoRI_P1.2	/5Phos/AATTACGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TAGTA_EcoRI_P1.2	/5Phos/AATTTACTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TATAC_EcoRI_P1.2	/5Phos/AATTGTATAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCACG_EcoRI_P1.2	/5Phos/AATTCGTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCAGT_EcoRI_P1.2	/5Phos/AATTACTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCCGG_EcoRI_P1.2	/5Phos/AATTCCGGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCTGC_EcoRI_P1.2	/5Phos/AATTGCAGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TGGAA_EcoRI_P1.2	/5Phos/AATTTTCCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TTACC_EcoRI_P1.2	/5Phos/AATTGGTAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

library	i7_name	i7_index	i5_name	i5_index
Lib_13	P2_UDI0032_AATGCCTC	AATGCCTC	P1_UDI0029_CAGTGGAT	CAGTGGAT
Lib_14	P2_UDI0053_CAACAATG	CAACAATG	P1_UDI0029_CAGTGGAT	CAGTGGAT
Lib_15	P2_UDI0045_ATGGCATG	ATGGCATG	P1_UDI0029_CAGTGGAT	CAGTGGAT
Lib_16	P2_UDI0072_GTCTACAC	GTCTACAC	P1_UDI0029_CAGTGGAT	CAGTGGAT
Lib_17	P2_UDI0032_AATGCCTC	AATGCCTC	P1_UDI0036_GACGAGAG	GACGAGAG
Lib_18	P2_UDI0053_CAACAATG	CAACAATG	P1_UDI0036_GACGAGAG	GACGAGAG
Lib_19	P2_UDI0045_ATGGCATG	ATGGCATG	P1_UDI0036_GACGAGAG	GACGAGAG
Lib_20	P2_UDI0072_GTCTACAC	GTCTACAC	P1_UDI0036_GACGAGAG	GACGAGAG
Lib_21	P2_UDI0032_AATGCCTC	AATGCCTC	P1_UDI0047_AATTCTGC	AATTCTGC
Lib_22	P2_UDI0053_CAACAATG	CAACAATG	P1_UDI0047_AATTCTGC	AATTCTGC
Lib_23	P2_UDI0045_ATGGCATG	ATGGCATG	P1_UDI0047_AATTCTGC	AATTCTGC
Lib_24	P2_UDI0072_GTCTACAC	GTCTACAC	P1_UDI0047_AATTCTGC	AATTCTGC
Lib_25	P2_UDI0053_CAACAATG	CAACAATG	P1_UDI0066_GCTTGCGC	GCTTGCGC
Lib_26	P2_UDI0032_AATGCCTC	AATGCCTC	P1_UDI0066_GCTTGCGC	GCTTGCGC
Lib_27	P2_UDI0045_ATGGCATG	ATGGCATG	P1_UDI0066_GCTTGCGC	GCTTGCGC
Lib_28	P2_UDI0072_GTCTACAC	GTCTACAC	P1_UDI0066_GCTTGCGC	GCTTGCGC
Lib_29	P2_UDI0032_AATGCCTC	AATGCCTC	P1_UDI0096_GTGTAGAC	GTGTAGAC
Lib_30	P2_UDI0053_CAACAATG	CAACAATG	P1_UDI0096_GTGTAGAC	GTGTAGAC
Lib_31	P2_UDI0045_ATGGCATG	ATGGCATG	P1_UDI0096_GTGTAGAC	GTGTAGAC

Chapter 3: The emergence of diversity through organisational scale is associated with regional differences and biological traits.

Authors

Thomas Keggin[1,2], Maurine Vilcot[3], Conor Waldock[5], Stéphanie Manel[3], Loïc Pellissier[1,2]

Author Affiliations

1. Ecosystems and Landscape Evolution, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland
2. Unit of Land Change Science, Swiss Federal Research Institute WSL
3. CEFE, Univ. Montpellier, CNRS, EPHE- PSL University, Montpellier, France
4. Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland
5. Department of Fish Ecology and Evolution, Center for Ecology, Evolution and Biogeochemistry, Eawag - Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland

Abstract

Patterns of diversity are often studied separately across levels of organisational scale, such as the species and genetic levels. This separation is despite general acceptance that similar, or the same, processes determine patterns of diversity at both levels of organisation. Because of these overlaps it is also expected that patterns of diversity should be positively correlated across organisational levels – which in reality is not always the case. Reasons for non-positive, or commonly non-significant, correlations in diversity patterns at the species and genetic levels of diversity have been associated with environmental variables and biological traits. These inferences, however, remain limited by correlative methods and conceptual frameworks that are not unified in process across organisational levels. In this study, we compare patterns of tropical reef fish diversity in species richness, turnover, and phylogeny to corresponding measures of genetic diversity in a population genetics data set covering diverse biological traits and phylogenetic history. The sampling scheme covers both the Caribbean and Western Indian Ocean. We complement existing correlative measures with the continuity metric (the ratio of diversity between the species and genetic levels) and find support for a conceptual framework that allows for the emergence of diversity through universal processes across organisational scale. We find that both seascape configuration and biological traits influence how diversity emerges across organisational levels. We also find support for a diversity partitioning mechanism from the genetic level to the species level through speciation.

Introduction

Biodiversity is a multifaceted measure of the variability within and between biological units (Gaston & Spicer, 2013) that exists across the nested levels of biological organisational scale. These units range from small to large in a nested hierarchy of aggregated organisational levels. Levels within this scale range from nucleotides, then aggregate into genes, chromosomes, genomes, individuals, populations, species, communities, ecosystems, biomes and are ultimately grouped into the entire planet (Whitham et al., 2006). Biodiversity can be measured at each of these levels separately, and most studies focus specifically on one of these levels to ease the overwhelming complexity in studying life. Observing them separately does not disconnect them however, and efforts to reconcile

patterns and processes we have learnt at each level is required to improve our understanding of the organisation of biological systems. Amongst the multiple levels of organisation, the comparison of species diversity and genetic diversity has been the focus of multiple studies (Chave, 2004; Gaggiotti et al., 2018; Kahilainen et al., 2014; Lamy et al., 2013; Lamy et al., 2017; Overcast et al., 2019; Schmidt, Munshi-South, et al., 2022; Vellend, 2010; Vilcot et al., 2023; Whitham et al., 2006). Whilst species and genetic diversity have been extensively investigated, mostly through ecological and phylogenetic study at the species level, and population genetics at the individual level (Pelletier et al., 2009), the two levels are represented by theories that do not necessarily overlap (Antonovics, 2003; Schmidt, Munshi-South, et al., 2022; Vellend, 2005; Vellend & Geber, 2005). However, given that these two levels of organisation comprise the same biological system, we should be able to reconcile these bodies of understanding.

Existing theory connecting the species and genetic levels of diversity has mostly emerged from the study of species-genetic diversity correlations (SGDCs). Attempts to apply well-established population genetics principles to the field of ecology (Antonovics, 1976), and then later drawing parallels between the two (Antonovics, 2003; Vellend, 2010; Vellend et al., 2014). These parallels were drawn by comparing biological processes described at both the genetic and species levels and their resulting patterns, including dispersal, drift, selection, and mutation/speciation. Identification of similar processes across levels led to the null expectation that patterns should also be the same across levels, i.e., there should be a positive correlation between diversity at the two levels of organisation (Chave, 2004; Huston & Huston, 1994; Vellend & Geber, 2005). However, empirical studies have shown that the strength and direction of these SGDCs can be highly variable (Kahilainen et al., 2014; Lawrence & Fraser, 2020; Manel et al., 2020; Taberlet et al., 2012). To understand the discrepancies in patterns across organisational scale, various mechanisms have been proposed that may drive divergent patterns between the two levels of organisation. These include direct feedback effects, such as greater genetic diversity in a community enhancing species survivability and increasing species diversity; or increasing species diversity with static resource availability may reduce population sizes and reduce genetic diversity (Vellend & Geber, 2005). Moreover, extrinsic factors related to the physical landscape have been proposed to have opposite effects on species and genetic diversity, e.g., increasing variation in resources increases intra-species specialisation, reducing gene-flow, but also decreases inter-species competition, allowing more species to coexist (Schmidt, Dray, et al., 2022). Ultimately, these works are limited to environmental variables, not biological processes - integrating the joint effect of biological and landscape processes is likely required for understanding continuity across levels of organisation.

The direct comparisons of pattern and process between organisational levels lays an important foundation, and frames the puzzle of differing species and genetic diversity patterns as being driven by parallel or interacting processes (Figure 2A; Vellend & Geber, 2005; Vellend et al., 2014). This idea of understanding continuity in diversity patterns across scale through parallel processes has been established in a paradigm of disentangling ecological vs evolutionary processes and feedbacks (Bailey et al., 2009; Pelletier et al., 2009; Schoener, 2011) and our thinking remains mostly (Leibold et al., 2022; Segar et al., 2020) but not completely (Lamy et al., 2017; Schmidt, Dray, et al., 2022) entrenched within it. Alternatively, we can accept that patterns can be measured separately at each level, but assume that processes are universal (Figure 2B). Gene flow and migration are both a consequence of the movement of individuals (Hamilton, 2021; MacArthur & Wilson, 2016; Nei et al., 1983), the process is dispersal of individuals. Drift is the stochastic differentiation of isolated populations and communities over time through uncoupled loss and gain of individuals, resulting in loss and fixation of alleles and species (Hamilton, 2021; Vellend & Geber, 2005). Selection of alleles

and species is the result of non-random reproductive success of individuals, the process is ultimately selection on individual phenotypes (Hamilton, 2021). If we accept the case of universal process, the question changes from “which processes are analogous across levels and what are the feedbacks?” to “through which individual processes do diversity patterns emerge through organisational scale?”. This is not a new concept, but instead embraces perspectives voiced over 75 years ago (Hancock et al., 2021). If we readopt this paradigm, the complexity of the puzzle is reduced from various processes at each organisational level and their interactions, to just one universal set.

Measuring the correspondence between species and population metrics is aided by the development of continuity metrics (Keggin et al., Chapter 1). SGDCs are useful in revealing the relationship between levels of organisational diversity, but are limited in their explanatory power. The pattern is often a single relationship representing the comparison of multiple species per geographic site (Manel et al., 2020; Schmidt, Munshi-South, et al., 2022; Vilcot et al., 2023), which can be compared to environmental variables corresponding to each of those sites. These relationships between species and genetic diversity are valuable, but cannot be used if variation between these levels does not follow a correlated relationship. SGDCs are also difficult to leverage against lineage-specific variables likely important in driving diversity through both levels, such as dispersal ability and competitive traits (Keggin et al., Chapter 1), requiring sampling across many geographic sites to attain enough statistical power (Vilcot et al., 2023). However, if we decompose the ratio of diversity across levels into a single metric per species or lineage, we can: as seen *in silico* (Keggin et al. Chapter 1). In this case the discordance in the relationship between the population and species levels was linked to universal processes: ease of speciation, dispersal, abundance, and competition. Notable was the proposal of a partitioning effect between levels of diversity through speciation where the new species takes a portion of individuals, and their respective diversity, from the parent species – both the parent and offspring species have reduced genetic diversity, but species diversity increases. This mechanism is neither present at one level of organisation nor the other, nor is it a strict feedback, it is a mechanism dictating how diversity emerges through organisational scale (Keggin et al., Chapter 1). This analytical method allows us to incorporate more eco-evolutionary processes, but until now lacks a comparison to real world systems.

Our aim is to apply this framework, of unified process across organisational scale, to a natural system of tropical reef fishes in two geographically separated oceans. The use of universal process as a conceptual framework and the continuity metric as a measure have been demonstrated *in silico* (Keggin et al., Chapter 1). Here we apply this framework *in vivo* to a SNP-based macrogenetics dataset (Keggin et al., Chapter 2) of the study system that was simulated, tropical reef fishes; ideal for their suitability to island model frameworks (Warren et al., 2015) and rich diversity (Parravicini et al., 2021) – sampled across the Caribbean and Western Indian Ocean. Using SGDCs, tropical reef fishes have been found to have a positive relationship between species richness and non-neutral mitonchondrial control region (Messmer et al., 2012), as well as between non-neutral CoI mitonchondrial diversity and species richness (Manel et al., 2020). Similarly, Vilcot et al. (2023) found a positive relationship between putatively neutral RADseq SNP marker differentiation (a subset of the data applied here) and species turnover. These existing studies focus on patterns within a single comparison between the two levels of diversity. Here, we embrace the multidimensionality of diversity and measure continuity across 6 distinct facets of diversity

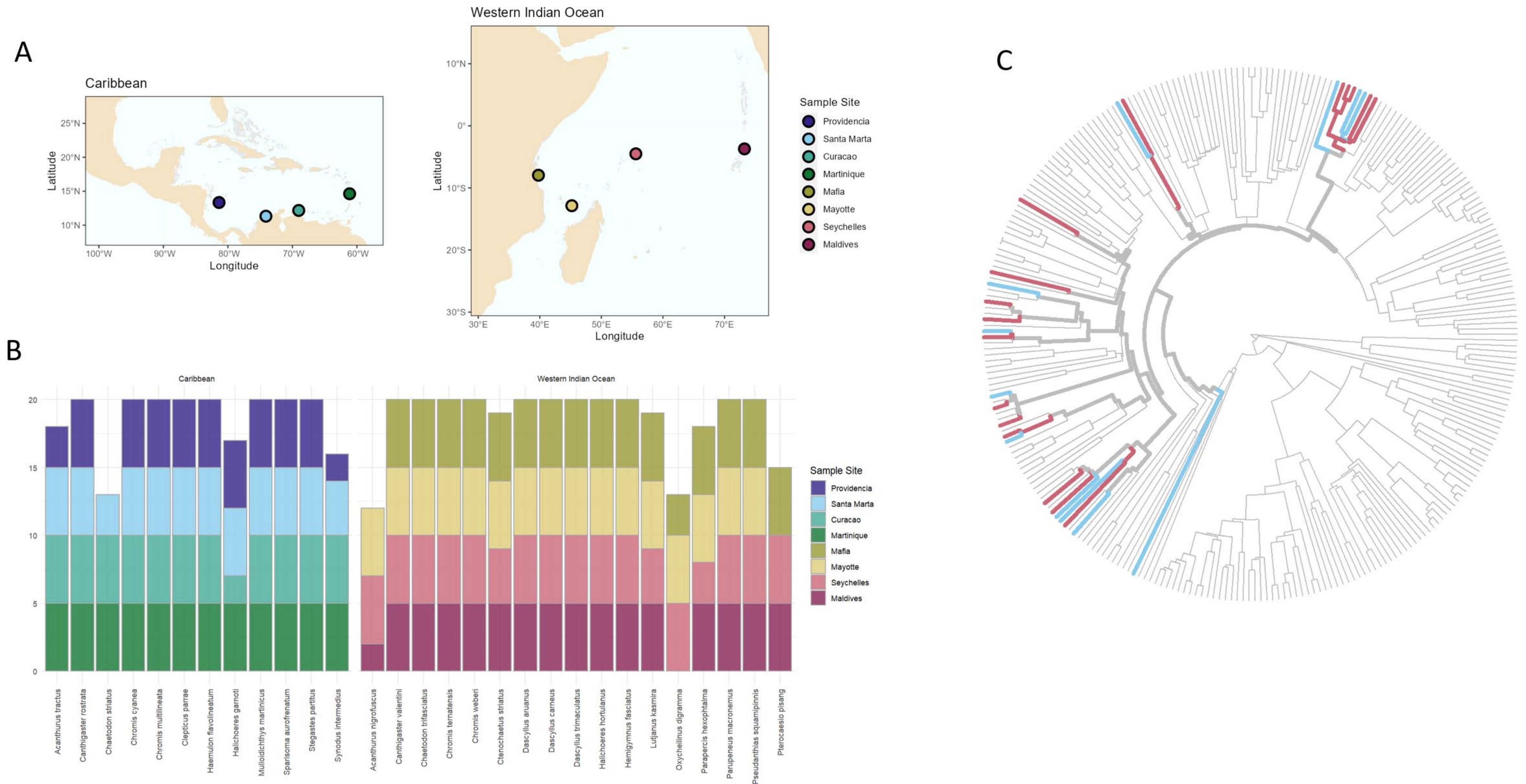


Figure 1: Genetic sampling scheme. A) maps of the 8 sites split across the two biogeographical realms. B) The number of individuals sampled and genotyped for each species at each site. C) A phylogeny demonstrating the phylogenetic sampling scheme. Light blue are Caribbean species, dark red are Western Indian Ocean species. Grey branches are other species randomly subsampled across the Teleost radiation to provide evolutionary context to the sampling.

encompassing both phylogenetic and richness derived components. We relate diversity across levels of organisation and their continuity to seascape configuration, dispersal, abundance, and speciation rate – variables identified likely to be important factors in influencing diversity emergence through organisational levels (Keggin et al., Chapter 1). To give structure to our application of this framework we specifically ask the following questions:

1. What is the relationship between species and genetic/individual/population diversity and are these consistent across diversity components? The SGDC approach.
2. Does the seascape impact the correspondence of diversity through organisational scale?
3. Using the continuity metric, can we relate biological processes to the correspondence of diversity through organisational scale?

For the second two questions, by measuring diversity values at the species and genetic levels separately and having continuity values available, we should be able to categorise our significant results into the following scenarios:

Scenario 1. The variable represents a process that drives the emergence of diversity patterns at only one level enough to create a sufficient discordance between levels to be detected in the continuity metric.

Scenario 2. The variable represents a process that drives the emergence of diversity patterns at both levels in opposite directions enough to create a sufficient discordance between levels to be detected in the continuity metric.

Scenario 3. The variable represents a process which is itself a mechanism that determines how diversity emerges through organisational scale. I.e., the influence of the variable on each level of diversity is negligible compared to the much greater influence it has on the emergence of diversity through the levels.

Scenario 4. The variable influences either diversity level, but has no impact on how diversity emerges through organisational scale.

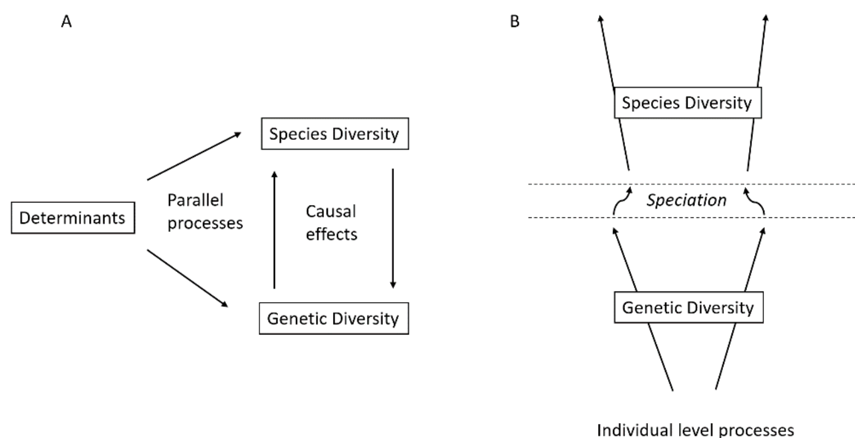


Figure 2: Conceptual frameworks of the action of eco-evolutionary processes. A) An representative framework from Vellend and Geber (2005) illustrating how processes running parallel at each level of diversity drive patterns of species and genetic diversity. This framework includes direct feedback effects between levels of diversity. B) The conceptual framework introduced here whereby individual-level processes drive patterns of both species and genetic diversity through an emergence through organisational scale. Speciation, in this instance, does not drive diversity directly at either level, but determines how it transitions from the genetic level to the species level of organisation.

Methods

Genetic data

All genetic data were taken from Keggin *et al.* (Chapter 2), including both genetic metrics and genotypes where a full description of the methodology used to sample the fishes, extract the DNA, prepare the libraries, carry out the sequencing, identify genotypes, and calculate population genetic metric estimates is provided. Briefly, between 2016 and 2020, a total of 1753 individuals were sampled across the Western Indian Ocean (1096 individuals, 27 species) and the Caribbean Sea (640 individuals, 15 species). These 42 reef fish species were selected from species in each ocean to ensure maximum representation in terms of body size and morphology, and to be distributed across the Teleost phylogeny (Figure 1C). We also aimed to sample comparable species between oceans: species of the same families if possible and of the same orders if not. For both oceans, four sites were sampled: the Caribbean Sea, including Providencia Island (Colombia), Santa Marta (Colombia), Curaçao, and Martinique (France); and the Western Indian Ocean, including Mafia Island (Tanzania), Mayotte (France), Seychelles, and the Maldives (Figure 1A). From these samples, various population genetic metrics were calculated, but here we retained mean expected heterozygosity across sample sites, total expected heterozygosity across sample sites, and global Hedrick's G''_{ST} per species (Meirmans & Hedrick, 2011). These metrics are all heterozygosity-based and we compare them to richness-based metrics at the species level.

Phylogenetic metrics were calculated at the genetic level through calculating genetic distances between individuals based on their SNP data, then by constructing phylogenetic trees using these based distance matrices. To calculate the genetic distance matrices, SNP data were first converted to Adegenet's genind format then processed using the *dist.gene* function in the ape package in R (Paradis & Schliep, 2019) using pairwise distances and removal of missing loci. From these individual-based distance matrices, intra-species phylogenetic trees were constructed through a neighbour-joining tree estimation approach using the *nj* function in the ape package in R (Paradis & Schliep, 2019; Saitou & Nei, 1987). From these intra-species phylogenetic trees, three phylogenetic metrics were calculated: phylogenetic diversity *sensu* Faith using the *pd* function in the picante R package (Faith, 1992; Kembel et al., 2010); mean pairwise distance using the *mpd* function in the picante package; and variation in pairwise distances using the base R *var* function on the calculated mean pairwise distances (R Core Team, 2022). These within-species phylogenies were plotted and visually checked for correspondence to expected population genetic structure.

Species data

As with the genetic data, two classes of species diversity metric were calculated per family: richness based-metrics (comparable to the heterozygosity-based metrics at the genetic level) and phylogenetic metrics. For all metrics, species were filtered based on their inclusion in families of interest, and their presence in the areas sampled at the genetic level. From a global marine species presence/absence database (Parravicini et al., 2013), we extracted a PA matrix for each family, as grouped by the NCBI taxonomy (Schoch et al., 2020). The resultant family matrices were then filtered to those families represented in the genetic sampling scheme. These global family matrices were further subdivided geographically into the eight sample sites to correspond to the genetic sampling scheme. Since the genetic sampling locations are point coordinates, we expanded these point locations by a 111 km buffer radius to capture local species richness per family per sample site. From these by-family and by-site presence/absence matrices, we derived species lists per family per site from which we calculated both the ecological and phylogenetic metrics. Three richness metrics were calculated per family: mean local species richness per site (alpha richness); total species

richness per family per region (the Caribbean and Western Indian Ocean); and community turnover of species between sites per region, as defined by the turnover component of Jaccard's dissimilarity index calculated using the *beta.multi* function of the betapart package in R (Baselga & Orme, 2012). To calculate the phylogenetic metrics, we took the backbone phylogeny from the Fish Tree of Life Project (Rabosky et al., 2018) and subset the species using the species lists derived for the ecological metrics using the *keep.tip* function in the R ape package (Paradis & Schliep, 2019). From these subset phylogenies, we calculated three phylogenetic metrics identical to the genetic-level phylogenies: phylogenetic diversity, mean pairwise distance, and variation in pairwise distance. We substituted *Acanthurus tractus*, which was not present in the fish tree of life phylogeny, with the congeneric species, *Acanthurus chirurgus*.

Continuity metrics

Following the framework proposed by Keggin *et al.* (Chapter 1), we decomposed the ratio of diversity between the species and genetic levels into six continuity metrics – one for each facet of diversity: alpha richness, beta richness, gamma richness, phylogenetic diversity, mean pairwise distance, and variation in mean pairwise distance (Table 1). This decomposition was done per family using pairs of comparable metrics across organisational levels according to the following:

$$Continiuty = log\left(\frac{species\ diversity}{genetic\ diversity}\right)$$

Where species and genetic diversity refer to the following pairs of diversity metrics: per site expected heterozygosity and alpha species richness; global expected heterozygosity and gamma species richness; Hedrick's G''_{ST} and Jaccard's Turnover component; genetic and species phylogenetic diversity; genetic and species mean pairwise distance; and genetic and species variation in pairwise distance (Table 1). Before this decomposition was carried out, all metrics were scaled between 0.01 and 1. Where data are aggregated to the family level, all metrics are decomposed into the mean for that family before continuity was calculated.

Table 1: The six facets of diversity and their constituent metrics.

Facet	**Genetic Metric**	**Species Metric**
Alpha richness	H_S	Mean per site species richness
Beta richness	Hedrick's G''_{ST}	Jaccard's Turnover
Gamma richness	H_T	Total species richness
Phylogenetic Diversity	PD	PD
Mean Pairwise Distance	MPD	MPD
Variation in Pairwise Distance	VPD	VPD

Biological traits

We compared continuity metrics to a set of functional traits and phylogeny-derived evolutionary characteristics inherent to each species. We extracted the following traits: reproductive mode, spawning behaviour, and maximum body length from fishbase (Boettiger et al., 2012); pelagic larval duration from Luiz et al. (2013); and time from last known divergence event, diversification rate, and variable time speciation rate from the fish tree of life project (Rabosky et al., 2018). Given that maximum body length and pelagic larval duration can both be considered proxies for dispersal ability, and are positively correlated (linear regression; $\beta = 0.41$, $t = 2.32$, $p = 0.03$), we combined the

Table 2: Analysis sets. A) The analysis sets which aggregate by family. B) The analysis sets which do not aggregate.

A

Family	Ocean	Trait
Acanthuridae	✓	✓
Chaetodontidae	✓	✓
Haemulidae	✓	✓
Labridae	✓	✓
Lutjanidae	✓	✓
Mullidae	✓	✓
Pinguipedidae		✓
Pomacentridae	✓	✓
Scaridae		✓
Serranidae		✓
Synodontidae		✓
Tetraodontidae	✓	✓

B

Family	Species	All	Lineage
Acanthuridae	*Acanthurus nigrofuscus*	✓	✓
Acanthuridae	*Acanthurus tractus*	✓	✓
Acanthuridae	*Ctenochaetus striatus*	✓	✓
Chaetodontidae	*Chaetodon striatus*	✓	✓
Chaetodontidae	*Chaetodon trifasciatus*	✓	✓
Haemulidae	*Haemulon flavolineatum*	✓	✓
Labridae	*Clepticus parrae*	✓	✓
Labridae	*Halichoeres garnoti*	✓	✓
Labridae	*Halichoeres hortulanus*	✓	✓
Labridae	*Hemigymnus fasciatus*	✓	✓
Labridae	*Oxycheilinus digramma*	✓	✓
Lutjanidae	*Lutjanus kasmira*	✓	✓
Lutjanidae	*Pterocaesio pisang*	✓	✓
Mullidae	*Mulloidichthys martinicus*	✓	✓
Mullidae	*Parupeneus macronemus*	✓	✓
Pinguipedidae	*Parapercis hexophtalma*	✓	
Pomacentridae	*Chromis cyanea*	✓	✓
Pomacentridae	*Chromis multilineata*	✓	✓
Pomacentridae	*Chromis ternatensis*	✓	✓
Pomacentridae	*Chromis weberi*	✓	✓
Pomacentridae	*Dascyllus aruanus*	✓	✓
Pomacentridae	*Dascyllus carneus*	✓	✓
Pomacentridae	*Dascyllus trimaculatus*	✓	✓
Pomacentridae	*Stegastes partitus*	✓	✓
Scaridae	*Sparisoma aurofrenatum*	✓	
Serranidae	*Pseudanthias squamipinnis*	✓	
Synodontidae	*Synodus intermedius*	✓	
Tetraodontidae	*Canthigaster rostrata*	✓	✓
Tetraodontidae	*Canthigaster valentini*	✓	✓

two variables into a dispersal trait computed as the product of the two. For family level comparisons, the mean family trait value per ocean was used, except for the divergence time, for which the oldest divergence time for that family was used. Additionally, we extracted abundance estimates from the Reef Fish Survey (Edgar et al., 2020). This was done by extracting the relative total number of individuals sighted in visual dives for the target species within our bioregions of interest (Spalding et al., 2007). We grouped values by site and by year, took the mean values, then for each species used those means as our final global species values.

Statistical analyses

Four subsets of species and families were used in the analysis of continuity: the all set, the lineage set, the ocean set, and the traits set (Table 2). The lineage set contains species for which there are more than one sample species per family to allow for inter-family comparisons in continuity metric values. The ocean set contains a subset of species for which there are within-family comparisons between oceans, allowing a seascape comparison in continuity metric values. The traits set contains additional families that do not have comparable species across oceans, but provide more data for the inference of the impact of species traits on continuity. Both the ocean and traits sets were aggregated by family. Finally, the “all” was identical to the traits set, but was not aggregated by family.

We first followed the SGDC approach. For each of the six facets of diversity (Table 1) the relationship between each pair of diversity metrics across levels of organisation was fit with and phylogenetic generalised least squares (PGLS) model using the *phylolm* function in the phylolm package in R (Tung Ho & Ané, 2014). For this we used the "all" data set, aggregating all species values to the family level, then also for each ocean separately (Table 2). Since the data were aggregated to the family level, the phylogeny was also reduced to a single species representing each family, controlling for evolutionary relationships between families. The phylogeny used was taken from Rabosky et al. (2018) and implemented using Pagel's lambda (Pagel, 1999). We compared levels of continuity between families visually, as there were insufficient species within each family to apply a phylogenetic ANOVA. Underlying these comparisons is the "lineage" set, which retains only species for which their family contains more than one sampled member (Table 2). The influence of seascape was modelled using a simple linear regression in R (R Core Team, 2022), with the two oceans represented by a binary variable. Models were fit for each of the six facets of diversity at both the genetic and species levels of diversity, as well as for continuity. The ocean comparison was done using the "ocean" set, containing only species for which there are within-family comparisons between oceans, and aggregated by family (Table 2). We fit a PGLS model to three biological traits (dispersal, speciation rate, and abundance) against each facet of diversity at the genetic and species levels, as well as for continuity. Again, we fit the PGLS using the species-level phylogeny from Rabosky et al. (2018) reduced to a single representative species per family, and implemented using Pagel's lambda in the phylolm package in R (Pagel, 1999; R Core Team, 2022; Tung Ho & Ané, 2014). Both the dispersal trait and the abundance were log transformed to meet the normal distribution assumption. This was done with using the "trait" set which included all species aggregated by family (Table 2).

Results

Species-genetic diversity correlations

We assessed possible relationships (SGDCs) between the species and genetic levels of diversity across the 6 facets (Table 2). Of all the species-genetic diversity correlations, only the comparison for the richness-based beta diversity metrics (Jaccard's Turnover and Hedrick's G''_{ST}) was significant (PGLS, $\beta = 0.23$, $t = 3.47$, $p = 0.01$; Figure 3), finding a positive association between the two levels – families with higher levels of species turnover between sites were comprised of species with high levels of differentiation between sites. However, when subdivided into the two oceans, only the association in the Western Indian Ocean was significant (PGLS, $\beta = 0.34$, $t = 4.21$, $p < 0.01$).

Ocean effect on continuity

To assess the influence of seascape on diversity and resultant continuity values, differences in diversity values between oceans was compared across all 6 diversity facets at the genetic and species levels, as well as for continuity. These inter-ocean differences showed higher values in either ocean depending on both the facet and level of organisation (Figure 4). At the genetic level, variation in phylogenetic diversity (VPD) and Hedrick's G_{ST} showed significant differences between oceans, with phylogenetic variation being larger in the Caribbean (linear regression, $\beta = -1.4$, $t = -2.30$, $p = 0.04$), and the opposite pattern found for genetic differentiation based on Hedrick's G''_{ST} (linear regression, $\beta = 0.32$, $t = 2.72$, $p = 0.02$). At the species level, phylogenetic diversity (PD), alpha species richness and gamma species richness were all significant, with greater phylogenetic diversity and more species found in the Western Indian Ocean than the Caribbean (PD, $\beta = 1.56$, $t = 2.78$, $p = 0.02$; alpha richness, $\beta = 0.38$, $t = 3.24$, $p < 0.01$; gamma richness, $\beta = 0.36$, $t = 2.94$, $p = 0.01$). Decomposing the relationship between the genetic and species level metrics into the continuity

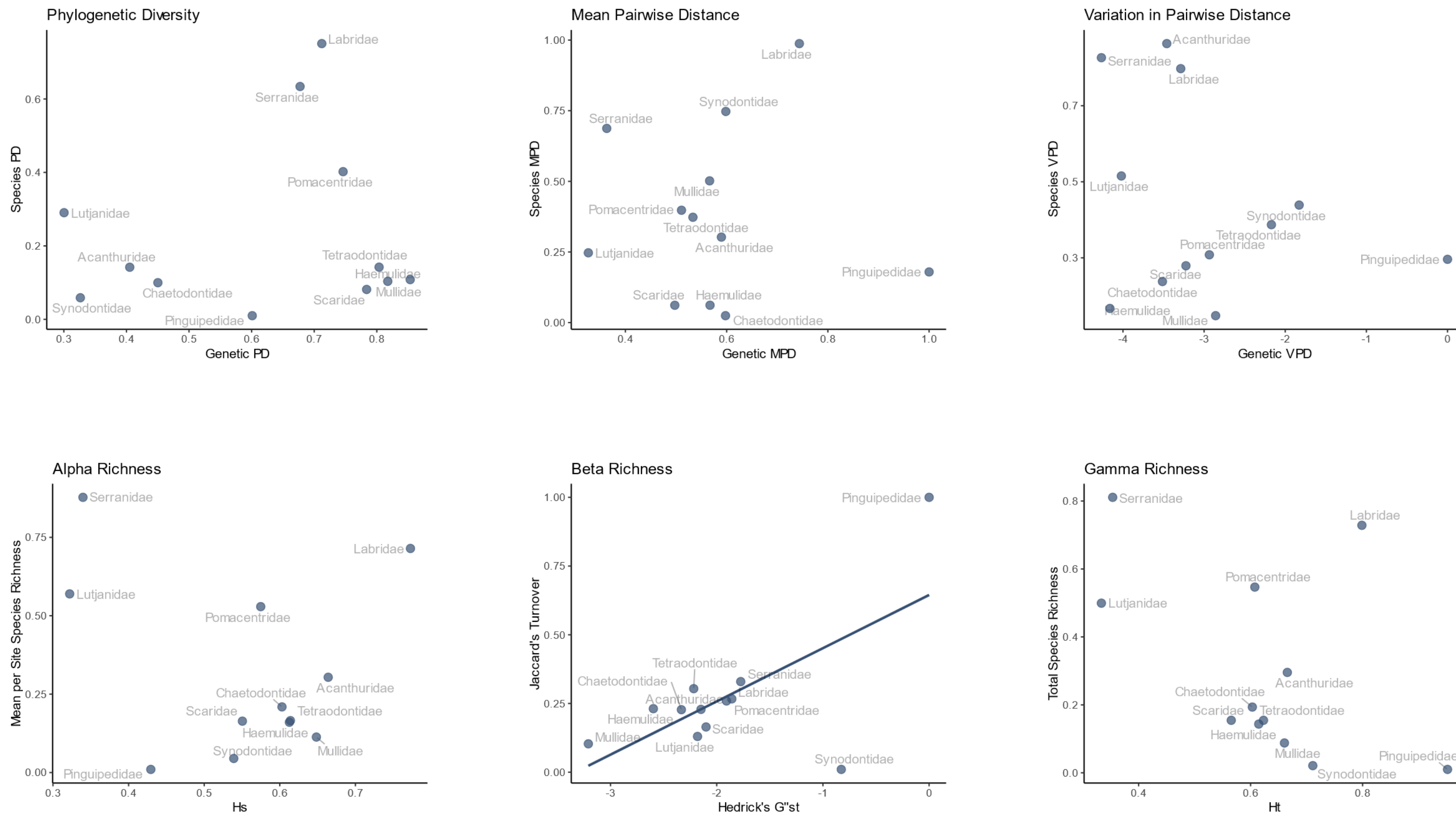


Figure 3: Species-genetic diversity correlations for each of the 6 diversity facets, solid trend lines indicate a significant relationship between the two variables (β = 0.23, t = 3.47, p = 0.01). A) Phylogenetic diversity (PD), species PD against genetic PD. B) Mean pairwise distance (MPD), species MPD vs genetic MPD. C) Variation in pairwise distance (VPD), species VPD vs genetic VPD. D) Alpha richness, alpha species richness, or mean species richness per site, vs mean expected heterozygosity per site. E) Beta richness, Jaccard's Turnover component vs Hedrick's G_{ST}. F) Gamma richness, total species richness per bioregion vs mean overall expected heterozygosity.

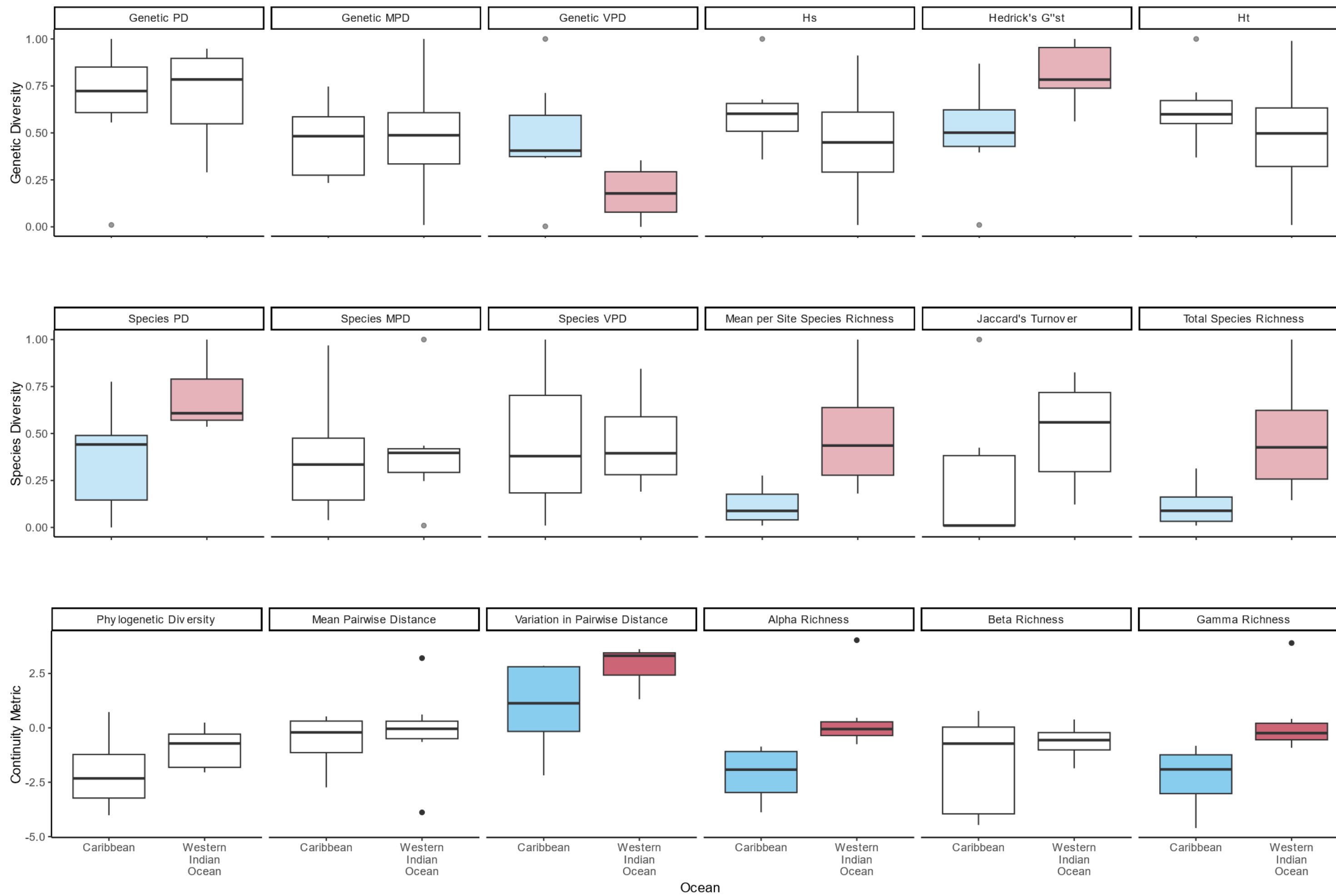


Figure 4: The effect of ocean on genetic diversity metrics, species diversity metrics, and continuity metrics for each facet. Coloured comparisons indicate a significant difference between groups. For the continuity metric, lower values indicate more population diversity proportional to species diversity within the sampling scheme.

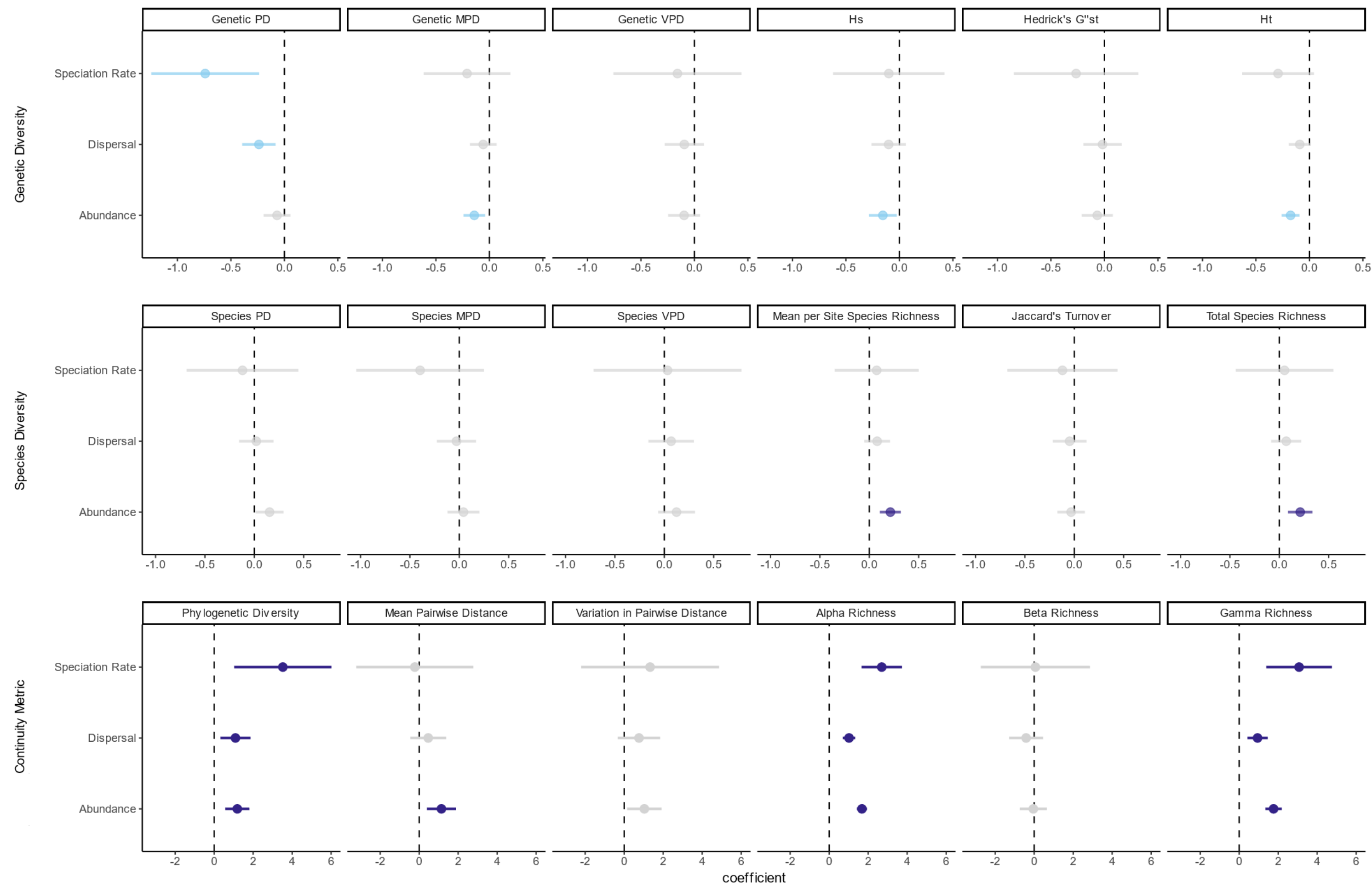


Figure 5: Coefficient plot derived from the phylogenetic generalised least squares models fitting genetic diversity, species diversity, and the continuity metric to the three biological traits: abundance, dispersal, and speciation rate. The top row represents the association of traits with genetic diversity, the middle row the association with species diversity, and the bottom row the association with the continuity metric. Light blue point and lines indicate significant negative relationships ($p < 0.05$), whilst dark blue point and lines indicate significant positive relationships ($p < 0.05$).

metric, we found significant between-ocean differences in VPD, alpha richness, and gamma richness (VPD, $\beta = 1.84$, $t = 2.22$, $p < 0.05$; alpha richness, $\beta = 2.55$, $t = 3.30$, $p < 0.01$; gamma richness, $\beta = 2.56$, $t = 3.14$, $p < 0.01$) – whereby diversity in these three metrics all contained proportionally more diversity at the species level than the genetic level in the Western Indian Ocean than the Caribbean. Both MPD and beta richness have mean continuity values close to 0 in both oceans – indicating that there was a relative equal split in the amount of diversity at the species and genetic levels across the sampled families.

Effect of traits on continuity

Finally, we tested the influence of lineage traits and their associated processes to the formation of diversity across the 6 facets at both the species and genetic levels, and on continuity. All three biological traits – dispersal, abundance, and speciation rate – were associated with different measures of diversity and continuity between them, as summarised in Figure 5. At the genetic level, both an increase in the speciation rate and dispersal was associated with a decrease in genetic PD (speciation rate; $\beta = -0.74$, $t = -2.88$, $p = 0.02$; dispersal, $\beta = -0.24$, $t = -2.99$, $p = 0.02$). Similarly, higher abundance values were associated with lower genetic MPD, alpha expected heterozygosity, and gamma expected heterozygosity (abundance: MPD, $\beta = -0.14$, $t = -2.73$, $p = 0.03$; H_s, $\beta = -0.15$, $t = -2.33$, $p < 0.05$; H_T, $\beta = -0.18$, $t = -4.11$, $p < 0.01$). At the species level, higher abundance values were associated with an increase in alpha and gamma species richness, whilst all other associations were not significant (abundance: alpha richness, $\beta = 0.21$, $t = 3.94$, $p < 0.01$; gamma richness, $\beta = 0.21$, $t = 3.36$, $p < 0.01$). After decomposing the relationship between the species and genetic levels of diversity into the continuity metric, we found multiple associations with the biological trait metrics. Increasing abundance values was associated with an increase in species diversity compared to genetic diversity for PD, MPD, alpha richness, and gamma richness (abundance: PD, $\beta = 1.19$, $t = 3.73$, $p < 0.01$; MPD, $\beta = 1.14$, $t = 2.99$, $p = 0.02$; alpha richness, $\beta = 1.38$, $t = 12.72$, $p < 0.01$; gamma richness, $\beta = 1.76$, $t = 8.22$, $p < 0.01$). For increasing dispersal values, there was a significant shift towards the species level compared the genetic level for PD, alpha richness, and gamma richness (dispersal: PD, $\beta = 1.09$, $t = 2.77$, $p = 0.02$; alpha richness, $\beta = 1.01$, $t = 6.22$, $p < 0.01$; gamma richness, $\beta = 0.94$, $t = 3.54$, $p < 0.01$). This same pattern was present for increasing lambda speciation rate values, with a corresponding shift towards the species level from the genetic level in PD, alpha richness, and gamma richness (speciation rate: PD, $\beta = 3.52$, $t = 2.77$, $p = 0.02$; alpha richness, $\beta = 2.70$, $t = 5.10$, $p < 0.01$; gamma richness, $\beta = 3.07$, $t = 8.22$, $p > 0.01$).

Discussion

In this study, we apply a novel conceptual and analytical framework to explore how universal eco-evolutionary processes result in the emergence of various diversity patterns across the genetic and species levels of organisational diversity. We show that universal biological traits drive the emergence of diversity at both the species and genetic organisational levels of diversity, and also how the speciation rate is related to this emergence of diversity through organisational scale – despite not being directly associated with diversity patterns at either level. We also find that seascape, or habitat configuration, is important, echoing patterns and hypotheses derived from species-genetic diversity correlative methods (Reisch & Hartig, 2021; Schmidt, Dray, et al., 2022) – patterns which we replicate and complement (Vilcot et al., 2023). Combined, our results demonstrate how this expanded framework complements the SGDC approach and advances our conceptual understanding of how diversity patterns form across levels of biological organisation.

The effect of biological traits

All three biological variables included here were significant in influencing patterns in continuity metrics: dispersal, speciation rate, and abundance (Figure 5). Notably, all three were significantly associated with patterns of continuity in phylogenetic diversity (PD), alpha richness, and gamma richness. Only the association of abundance with continuity in mean pairwise distance was significant in the remaining facets (Figure 5). Interestingly, these three facets, PD, and alpha and gamma richness can all be considered "richness" metrics across the many metrics devised for studying biodiversity (Tucker et al., 2017) as they are all quantifications of the amount of biological units within a system. Species richness represents the number of species and expected heterozygosity is a proxy for allelic richness – the number of genetic variants within a genome. PD here is the summed branch length of a phylogeny, or the number of units of diverged existence across all species or individual lineages (Faith, 1992). Our results here suggest that these three facets of diversity representing varying types of biological richness may be more prone to the influence of biological traits in creating discordance in patterns of diversity across levels of biological organisation. How exactly this discordance arises is more complex. Abundance is negatively associated with alpha and gamma richness at the genetic level – which is converse to expectations from neutral theory, but appears to be a peculiarity of this study system (Donati et al., 2021; Vilcot et al., 2023) – but positively associated as expected at the species level (Bock et al., 2007; Hakkila et al., 2021; Figure 4). This aligns with the expectations we would have in our scenario 2, whereby opposite associations at each level correspond to an imbalance in continuity. I.e., reducing genetic diversity and increasing species diversity results in more diversity at the species level compared to that at the genetic with greater abundance values. Interestingly, both speciation rate and, with a lesser effect size, dispersal were associated with continuity in the alpha and gamma richness facets, but not with either the genetic or species levels independently (Table 1**,** Figure 5). We interpret this as expectations for our scenario 3 – that dispersal and speciation rate have a negligible influence on the formation of heterozygosity and species richness compared to the role they play in generating discordant patterns between the two. Both large dispersal capacities and fast speciation rates were associated with more species diversity relative to genetic diversity, a pattern that may point towards the same mechanism. As postulated by Keggin *et al.* (Chapter 1), speciation should act as a partitioning mechanism of diversity from the genetic level to the species level as a portion of the parent species, and its constituent individuals with their diversity, is removed. Both the parent and daughter species suffer reduced genetic diversity, but the species pool is increased. A high speciation rate should reflect this process, and is reflected in these data. We might expect the opposite of this effect in the dispersal ability whereby lesser dispersal ability increases probability of allopatric speciation (Heinz et al., 2009; Pellissier, 2015). However, the interaction between the movement of individuals over a sea- or land-scape over evolutionary time and the diversification of lineages is complex and sensitive to both spatial and temporal scale (Lenoir et al., 2012). Perhaps over longer periods of time and a dynamic evolutionary seascape longer dispersing species are exposed to greater environmental heterogeneity (Battisti et al., 2019; Monaco et al., 2020; Sunday et al., 2015), maintain larger ranges more likely to suffer hard barriers to gene-flow (Alzate & Onstein, 2022; Bowman et al., 2002; Lester et al., 2007), and through greater effective population sizes – derived from a more connected meta-population – produce more lasting offspring species. If these were true, over a longer temporal scale greater dispersal could enable more speciation and subsequent partitioning of diversity from the genetic level to the species level. The result is a mix of likely interacting trait effects that would decouple patterns of species and genetic diversity.

Ocean differences

Measures of continuity across facets of diversity also varied between oceanic realms. Our sampling scheme consisted of two biogeographic realms, the Caribbean, and the Western Indian Ocean, representing differing seascape characteristics and species assemblages, but sharing recently diverged species (Rabosky et al., 2018; Renema et al., 2008; Spalding et al., 2007; Keggin *et al.,* Chapter 2). For variation in pairwise distance, alpha richness, and gamma richness, we find significant differences between oceanic realms, with all three representing relatively more species diversity in the Western Indian Ocean as compared to the Caribbean, and *vice versa* for genetic diversity. Again, we can categorise these patterns into the four scenarios discussed above: one, sufficient emergence of pattern at one organisational level; two, two levels in opposite directions; three, indicative of a mechanism directly influencing patterns of diversity emergence; or four, influence one or both levels of organisation without affecting diversity emergence. Through the significant difference between oceanic realms of only one level of organisational diversity for each facet of diversity, all three facets fit the expectations for scenario one: that the pattern of continuity is representative of the emergence of diversity at a single level of organisation sufficient to produce a difference in diversity patterns between levels. Although genetic differences in both alpha and gamma expected heterozygosity visually hold the opposite pattern to species diversity – a scenario 2 expectation. For alpha and gamma diversity, this pattern could be expected considering the habitat configuration of each realm. The Caribbean contains a larger, better connected habitable area than the Western Indian Ocean. For genetic diversity, the expectation for larger, more connected habitat is greater genetic diversity through greater effective population sizes over time (Bradburd & Ralph, 2019; Donati et al., 2021; Poethke & Hovestadt, 2002). Which contrasts the species level where the opposite is expected to be true as greater connectivity between populations should increase dispersal, resultant gene-flow and therefore reduce the likelihood of allopatric speciation (Mayr, 1963). This dynamic is similar to the resource-based driven relationship proposed by Schmidt, Munshi-South, et al. (2022), and carries the same signal as the diversity partitioning through speciation hypothesis discussed above, but this time in a geographical context (Keggin *et al.*, Chapter 1). This expected increased gene-flow in the Caribbean is reflected in our estimate of genetic differentiation (Hedrick's G''_{ST}), which is significantly greater in the Western Indian Ocean (Figure 4). Despite this large difference in at the genetic level, there is almost no difference in the continuity values for encompassing beta richness facet. In this case, whilst not significant, the pattern between oceans in the beta richness facet is similar at the species level (Figure 4). This, again, is expected through the positive correlation we find using the SGDC approach (Figure 3), what we expect for beta comparisons in different systems (Lamy et al., 2017), simulations (Keggin *et al.*, Chapter 1), and the continuity values approaching 0 for this facet – reflecting the null expectation of a positive correlation between the two levels of diversity (Vellend, 2005; Vellend & Geber, 2005).

Combining SGDCs with Continuity

This suite of compounding effects on continuity through both the seascape and biological traits may explain the noisiness and lack of correlations that we usually see in SGDCs. This is contrary to our beta richness measure which appears undisturbed by these variables (Figure 4,5) and positively correlated across levels (Figure 3). This contradiction suggests that the continuity metric characterises relationships between levels of diversity that we are unable to capture using an SGDC approach, and *vice versa* – when evaluating the uses of the two approaches, we find that they are complementary. With SGDCs, confident inference is only possible when statistically significant relationships are uncovered, but in this study we only find a significant relationship between genetic species diversity in the beta richness facet of our analysis (Figure 3). However, by applying the

continuity metric, we find significant signal for various patterns in all 5 of the other facets associated with both the environment and biological traits (Figure 5) – traits which are difficult to approach through SGDCs. Each SGDC value is a general relationship between many lineages at a site, or a single lineage across many sites (Figure 3; Schmidt, Munshi-South, et al., 2022; Vilcot et al., 2023). These general patterns allow us to infer useful information across a geographical landscape, i.e., local environmental variables vs local SGDC (Lamy et al., 2013; Lamy et al., 2017; Lamy et al., 2012; Laroche et al., 2015; Lawrence, 2020; Lawrence & Fraser, 2020; Manel et al., 2020; Schmidt, Dray, et al., 2022; Schmidt, Munshi-South, et al., 2022; Vilcot et al., 2023), but require extensive sampling schemes to provide the required statistical power to find meaningful relationships across sites within a lineage (Vilcot et al., 2023). This extension of our analytical toolset opens up an exciting new avenue for understanding these dynamics.

Limitations

Whilst a useful new tool, the continuity metric used here has limitations. The main consideration which must be brought forward in future work is the relative nature of continuity values. They are relative to the values at each level of organisation in the sampling scheme and the absolute continuity values are not comparable between datasets. This however does not apply to the direction and strength of relationships between continuity and predictor variables being explored. The second caution is in the choice of metric comparisons. In this study, we selected six facets of diversity to compare genetic and species metrics we believe to be meaningfully comparable, but careful consideration of marker types, genome (e.g. mitochondrial vs nuclear; Manel et al., 2020; Schmidt, Munshi-South, et al., 2022) and nature of the measurement. Continued discussion of appropriate metrics for comparison is required. In terms of our dataset, despite the extensive data collection carried out, the results are limited in our sampling scheme. A binary inter-ocean comparison prevents us from associating specific environmental variables with either SGDCs or the continuity metric values. We also lack sufficient species per family to highlight the inter-lineage differences in continuity metric values, which does not allow us to highlight the likely significant differences that SGDCs cannot detect.

Conclusion

This study takes a novel conceptual and analytical framework developed *in silico* and leverages it against an extensive macrogenetics database (Keggin *et al.*, Chapter 2; Donati et al., 2021), species occurrence (Parravicini et al., 2021), abundance (Edgar et al., 2020), traits (Boettiger et al., 2012; Luiz et al., 2013), and phylogenetic information (Rabosky et al., 2018). Through this application, we find patterns in-line with dynamics derived from existing SGDC-based literature (Schmidt, Munshi-South, et al., 2022; Vilcot et al., 2023) and expand our knowledge to match expectations developed experimentally *in silico* (Keggin *et al.*, Chapter 1) – in particular, patterns supporting the diversity partitioning hypothesis through speciation. Instead of treating diversification processes as parallel between levels of organisation (Schoener, 2011; Vellend & Geber, 2005; Vellend et al., 2014), our framework allows us to explore the emergence of diversity patterns through organisational scale from unified processes with both environmental and lineage-based variables. We believe that this framework and its application contributes meaningfully to the need to "…develop a general theoretical framework for eco-evolutionary dynamics—and then to quantify these dynamics in natural populations" (Pelletier et al., 2009), representing a significant step towards synthesising eco-evolutionary study.

Acknowledgements

We would like to thank Alex Skeels for his freely given phylogenetic knowledge, and critical and friendly discussion, as well as Camille Albouy for helping source the various data sets included in the analyses.

Funding was supported by the ETH research grant financing the salary of Thomas Keggin (ETH Research Grant ETH-34 18-1).

References

Alzate, A., & Onstein, R. E. (2022). Understanding the relationship between dispersal and range size. *Ecology Letters*, *25*(10), 2303-2323. https://doi.org/10.1111/ele.14089

Antonovics, J. (1976). The input from population genetics: "The new ecological genetics". *Systematic Botany*, *1*(3), 233-245. https://doi.org/10.2307/2418718

Antonovics, J. (2003). Toward community genomics? *Ecology*, *84*(3), 598-601. https://doi.org/10.1890/0012-9658(2003)084[0598:Tcg]2.0.Co;2

Bailey, J. K., Hendry, A. P., Kinnison, M. T., Post, D. M., Palkovacs, E. P., Pelletier, F., Harmon, L. J., & Schweitzer, J. A. (2009). From genes to ecosystems: an emerging synthesis of eco-evolutionary dynamics. *New Phytologist*, *184*(4), 746-749. https://doi.org/https://doi.org/10.1111/j.1469-8137.2009.03081.x

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

Battisti, C., Fanelli, G., Amori, G., Luiselli, L., & Capizzi, D. (2019). Do disturbance-sensitive and habitat-specialized species have a smaller range size? Evidence for a set of common mammals at regional scale. *Ethology Ecology & Evolution*, *31*(5), 479-490. https://doi.org/10.1080/03949370.2019.1626487

Bock, C. E., Jones, Z. F., & Bock, J. H. (2007). Relationships between species richness, evenness, and abundance in a southwestern Savanna. *Ecology*, *88*(5), 1322-1327. https://doi.org/10.1890/06-0654

Boettiger, C., Lang, D. T., & Wainwright, P. (2012). rfishbase: exploring, manipulating and visualizing FishBase data from R. *Journal of Fish Biology*, *81*(6), 2030-2039.

Bowman, J., Jaeger, J. A. G., & Fahrig, L. (2002). Dispersal distance of mammals is proportional to home range size. *Ecology*, *83*(7), 2049-2055. https://doi.org/10.2307/3071786

Bradburd, G. S., & Ralph, P. L. (2019). Spatial Population Genetics: It's About Time. *Annual Review of Ecology, Evolution, and Systematics*, *50*(1), 427-449. https://doi.org/10.1146/annurev-ecolsys-110316-022659

Chave, J. (2004). Neutral theory and community ecology [Review]. *Ecology Letters*, *7*(3), 241-253. https://doi.org/10.1111/j.1461-0248.2003.00566.x

Donati, G. F. A., Zemp, N., Manel, S., Poirier, M., Claverie, T., Ferraton, F., Gaboriau, T., Govinden, R., Hagen, O., Ibrahim, S., Mouillot, D., Leblond, J., Julius, P., Velez, L., Zareer, I., Ziyad, A., Leprieur, F., Albouy, C., & Pellissier, L. (2021). Species ecology explains the spatial components of genetic diversity in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1959), 20211574. https://doi.org/doi:10.1098/rspb.2021.1574

Edgar, G. J., Cooper, A., Baker, S. C., Barker, W., Barrett, N. S., Becerro, M. A., Bates, A. E., Brock, D., Ceccarelli, D. M., Clausius, E., Davey, M., Davis, T. R., Day, P. B., Green, A., Griffiths, S. R., Hicks, J., Hinojosa, I. A., Jones, B. K., Kininmonth, S., . . . Stuart-Smith, R. D. (2020). Establishing the ecological basis for conservation of shallow marine life using Reef Life Survey. *Biological Conservation*, *252*, 108855. https://doi.org/https://doi.org/10.1016/j.biocon.2020.108855

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*(1), 1-10. https://doi.org/https://doi.org/10.1016/0006-3207(92)91201-3

Gaggiotti, O. E., Chao, A., Peres-Neto, P., Chiu, C.-H., Edwards, C., Fortin, M.-J., Jost, L., Richards, C. M., & Selkoe, K. A. (2018). Diversity from genes to ecosystems: A unifying framework to study variation across biological metrics and scales. *Evolutionary Applications*, *11*(7), 1176-1193. https://doi.org/10.1111/eva.12593

Gaston, K. J., & Spicer, J. I. (2013). *Biodiversity: an introduction*. John Wiley & Sons.

Hakkila, M., Johansson, A., Sandgren, T., Uusitalo, A., Monkkonen, M., Puttonen, P., & Savilaakso, S. (2021). Are small protected habitat patches within boreal production forests effective in conserving species richness, abundance and community composition? A systematic review. *Environmental Evidence*, *10*(1), Article 2. https://doi.org/10.1186/s13750-020-00216-6

Hamilton, M. B. (2021). *Population genetics*. John Wiley & Sons.

Hancock, Z. B., Lehmberg, E. S., & Bradburd, G. S. (2021). Neo-darwinism still haunts evolutionary theory: A modern perspective on Charlesworth, Lande, and Slatkin (1982). *Evolution*, *75*(6), 1244-1255. https://doi.org/10.1111/evo.14268

Heinz, S. K., Mazzucco, R., & Dieckmann, U. (2009). Speciation and the evolution of dispersal along environmental gradients. *Evolutionary Ecology*, *23*(1), 53-70. https://doi.org/10.1007/s10682-008-9251-7

Huston, M. A., & Huston, M. A. (1994). *Biological diversity: the coexistence of species*. Cambridge University Press.

Kahilainen, A., Puurtinen, M., & Kotiaho, J. S. (2014). Conservation implications of species-genetic diversity correlations [Review]. *Global Ecology and Conservation*, *2*, 315-323. https://doi.org/10.1016/j.gecco.2014.10.013

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P., & Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, *26*(11), 1463-1464.

Lamy, T., Jarne, P., Laroche, F., Pointier, J. P., Huth, G., Segard, A., & David, P. (2013). Variation in habitat connectivity generates positive correlations between species and genetic diversity in a metacommunity [Article]. *Molecular Ecology*, *22*(17), 4445-4456. https://doi.org/10.1111/mec.12399

Lamy, T., Laroche, F., David, P., Massol, F., & Jarne, P. (2017). The contribution of species-genetic diversity correlations to the understanding of community assembly rules. *Oikos*, *126*(6), 759-771. https://doi.org/10.1111/oik.03997

Lamy, T., Pointier, J. P., Jarne, P., & David, P. (2012). Testing metapopulation dynamics using genetic, demographic and ecological data. *Molecular Ecology*, *21*(6), 1394-1410. https://doi.org/https://doi.org/10.1111/j.1365-294X.2012.05478.x

Laroche, F., Jarne, P., Lamy, T., David, P., & Massol, F. (2015). A Neutral Theory for Interpreting Correlations between Species and Genetic Diversity in Communities. *American Naturalist*, *185*(1), 59-69. https://doi.org/10.1086/678990

Lawrence, E. R. (2020). *Synthesizing vertebrate population richness and genetic diversity across the American continents* Concordia University].

Lawrence, E. R., & Fraser, D. J. (2020). Latitudinal biodiversity gradients at three levels: Linking species richness, population richness and genetic diversity [Review]. *Global Ecology and Biogeography*, *29*(5), 770-788. https://doi.org/10.1111/geb.13075

Leibold, M. A., Govaert, L., Loeuille, N., De Meester, L., & Urban, M. C. (2022). Evolution and Community Assembly Across Spatial Scales. *Annual Review of Ecology Evolution and Systematics*, *53*, 299-326. https://doi.org/10.1146/annurev-ecolsys-102220-024934

Lenoir, J., Virtanen, R., Oksanen, J., Oksanen, L., Luoto, M., Grytnes, J.-A., & Svenning, J.-C. (2012). Dispersal ability links to cross-scale species diversity patterns across the Eurasian Arctic tundra. *Global Ecology and Biogeography*, *21*(8), 851-860. https://doi.org/https://doi.org/10.1111/j.1466-8238.2011.00733.x

Lester, S. E., Ruttenberg, B. I., Gaines, S. D., & Kinlan, B. P. (2007). The relationship between dispersal ability and geographic range size. *Ecology Letters*, *10*(8), 745-758. https://doi.org/10.1111/j.1461-0248.2007.01070.x

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*, *110*(41), 16498-16502. https://doi.org/doi:10.1073/pnas.1304074110

MacArthur, R. H., & Wilson, E. O. (2016). Princeton University Press. https://doi.org/doi:10.1515/9781400881376

Manel, S., Guerin, P. E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., & Pellissier, L. (2020). Global determinants of freshwater and marine fish genetic diversity. *Nature Communications*, *11*(1). https://doi.org/ARTN 69210.1038/s41467-020-14409-7

Mayr, E. (1963). Animal Species and Evolution (Belknap, Cambridge, MA).

Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources*, *11*(1), 5-18. https://doi.org/https://doi.org/10.1111/j.1755-0998.2010.02927.x

Messmer, V., Jones, G. P., Munday, P. L., & Planes, S. (2012). Concordance between genetic and species diversity in coral reef fishes across the Pacific Ocean biodiversity gradient. *Evolution*, *66*(12), 3902-3917. https://doi.org/10.1111/j.1558-5646.2012.01725.x

Monaco, C. J., Bradshaw, C. J. A., Booth, D. J., Gillanders, B. M., Schoeman, D. S., & Nagelkerken, I. (2020). Dietary generalism accelerates arrival and persistence of coral-reef fishes in their novel ranges under climate change. *Global Change Biology*, *26*(10), 5564-5573. https://doi.org/10.1111/gcb.15221

Nei, M., Maruyama, T., & Wu, C.-I. (1983). Models of evolution of reproductive isolation. *Genetics*, *103*(3), 557-579. https://doi.org/10.1093/genetics/103.3.557

Overcast, I., Emerson, B. C., & Hickerson, M. J. (2019). An integrated model of population genetics and community ecology. *Journal of Biogeography*, *46*(4), 816-829. https://doi.org/10.1111/jbi.13541

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, *401*(6756), 877-884.

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526-528.

Parravicini, V., Bender, M. G., Villéger, S., Leprieur, F., Pellissier, L., Donati, F. G. A., Floeter, S. R., Rezende, E. L., Mouillot, D., & Kulbicki, M. (2021). Coral reef fishes reveal strong divergence in the prevalence of traits along the global diversity gradient. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1961), 20211712. https://doi.org/doi:10.1098/rspb.2021.1712

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*(12), 1254-1262. https://doi.org/https://doi.org/10.1111/j.1600-0587.2013.00291.x

Pelletier, F., Garant, D., & Hendry, A. P. (2009). Eco-evolutionary dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1523), 1483-1489. https://doi.org/doi:10.1098/rstb.2009.0027

Pellissier, L. (2015). Stability and the competition-dispersal trade-off as drivers of speciation and biodiversity gradients. *Frontiers in Ecology and Evolution*, *3*, Article 52. https://doi.org/10.3389/fevo.2015.00052

Poethke, H. J., & Hovestadt, T. (2002). Evolution of density–and patch–size–dependent dispersal rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*(1491), 637-645. https://doi.org/doi:10.1098/rspb.2001.1936

R Core Team. (2022). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. https://www.R-project.org/

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*(7714), 392-395. https://doi.org/10.1038/s41586-018-0273-1

Reisch, C., & Hartig, F. (2021). Species and genetic diversity patterns show different responses to land use intensity in central European grasslands. *Diversity and Distributions*, *27*(3), 392-401. https://doi.org/10.1111/ddi.13199

Renema, W., Bellwood, D. R., Braga, J. C., Bromfield, K., Hall, R., Johnson, K. G., Lunt, P., Meyer, C. P., McMonagle, L. B., Morley, R. J., O'Dea, A., Todd, J. A., Wesselingh, F. P., Wilson, M. E. J., & Pandolfi, J. M. (2008). Hopping hotspots: Global shifts in marine Biodiversity. *Science*, *321*(5889), 654-657. https://doi.org/10.1126/science.1155674

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*(4), 406-425.

Schmidt, C., Dray, S., & Garroway, C. J. (2022). Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity. *Evolution*, *76*(1), 86-100. https://doi.org/https://doi.org/10.1111/evo.14407

Schmidt, C., Munshi-South, J., Dray, S., & Garroway, C. J. (2022). Determinants of genetic diversity and species richness of North American amphibians. *Journal of Biogeography*, *49*(11), 2005-2015. https://doi.org/10.1111/jbi.14480

Schoch, C. L., Ciufo, S., Domrachev, M., Hotton, C. L., Kannan, S., Khovanskaya, R., Leipe, D., Mcveigh, R., O'Neill, K., Robbertse, B., Sharma, S., Soussov, V., Sullivan, J. P., Sun, L., Turner, S., & Karsch-Mizrachi, I. (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database*, *2020*. https://doi.org/10.1093/database/baaa062

Schoener, T. W. (2011). The Newest Synthesis: Understanding the Interplay of Evolutionary and Ecological Dynamics. *Science*, *331*(6016), 426-429. https://doi.org/10.1126/science.1193954

Segar, S. T., Fayle, T. M., Srivastava, D. S., Lewinsohn, T. M., Lewis, O. T., Novotny, V., Kitching, R. L., & Maunsell, S. C. (2020). The Role of Evolution in Shaping Ecological Networks. *Trends in Ecology & Evolution*, *35*(5), 454-466. https://doi.org/10.1016/j.tree.2020.01.004

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*(7), 573-583. https://doi.org/10.1641/b570707

Sunday, J. M., Pecl, G. T., Frusher, S., Hobday, A. J., Hill, N., Holbrook, N. J., Edgar, G. J., Stuart-Smith, R., Barrett, N., Wernberg, T., Watson, R. A., Smale, D. A., Fulton, E. A., Slawinski, D., Feng, M., Radford, B. T., Thompson, P. A., & Bates, A. E. (2015). Species traits and climate velocity explain geographic range shifts in an ocean-warming hotspot. *Ecology Letters*, *18*(9), 944-953. https://doi.org/10.1111/ele.12474

Taberlet, P., Zimmermann, N. E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Coldea, G., Mirek, Z., Moilanen, A., Ahlmer, W., Marsan, P. A., Bona, E., Bovio, M., Choler, P., Cieslak, E., Colli, L., Cristea, V., Dalmas, J. P., . . . IntraBioDiv, C. (2012). Genetic diversity in widespread species is not congruent with species richness in alpine plant communities [Article]. *Ecology Letters*, *15*(12), 1439-1448. https://doi.org/10.1111/ele.12004

Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., Grenyer, R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Purschke, O., Redding, D. W., Rosauer, D. F., Winter, M., & Mazel, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews*, *92*(2), 698-715. https://doi.org/https://doi.org/10.1111/brv.12252

Tung Ho, L. s., & Ané, C. (2014). A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Systematic Biology*, *63*(3), 397-408.

Vellend, M. (2005). Species diversity and genetic diversity: Parallel processes and correlated patterns. *American Naturalist*, *166*(2), 199-215. https://doi.org/10.1086/431318

Vellend, M. (2010). Conceptual synthesis in community ecology [Review]. *Quarterly Review of Biology*, *85*(2), 183-206. https://doi.org/10.1086/652373

Vellend, M., & Geber, M. A. (2005). Connections between species diversity and genetic diversity. *Ecology Letters*, *8*(7), 767-781. https://doi.org/10.1111/j.1461-0248.2005.00775.x

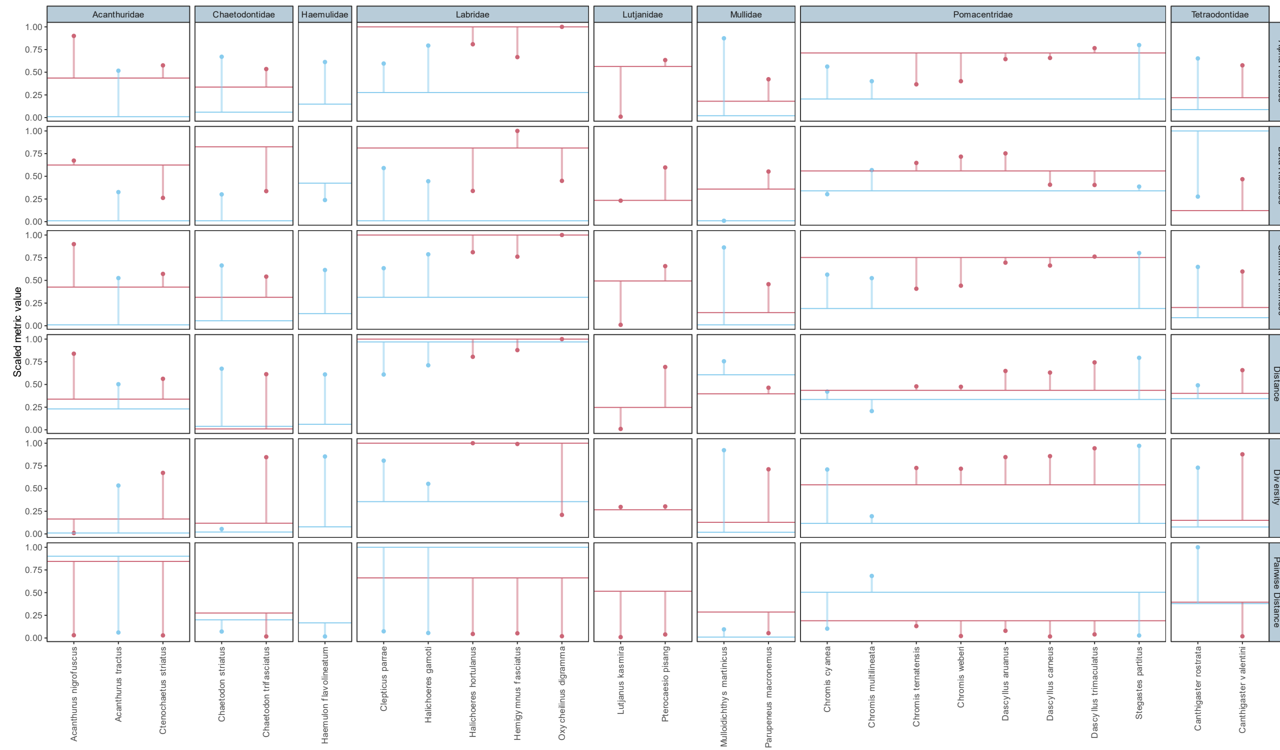
Vellend, M., Lajoie, G., Bourret, A., Murria, C., Kembel, S. W., & Garant, D. (2014). Drawing ecological inferences from coincident patterns of population- and community-level biodiversity. *Molecular Ecology*, *23*(12), 2890-2901. https://doi.org/10.1111/mec.12756

Vilcot, M., Albouy, C., Donati, G. F. A., Claverie, T., Julius, P., Manel, S., Pellissier, L., & Leprieur, F. (2023). Spatial genetic differentiation correlates with species assemblage turnover across tropical reef fish lineages. *Global Ecology and Biogeography*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/geb.13637

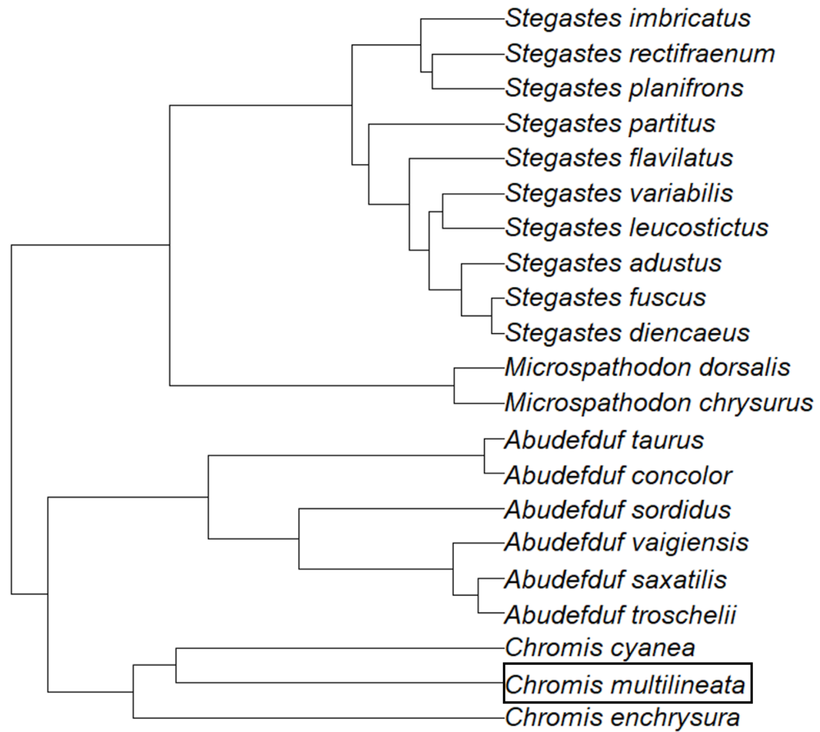
Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilee, R., Condamine, F. L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernandez-Palacios, J. M., Hengl, T., Norder, S. J., Rijsdijk, K. F., Sanmartin, I., Strasberg, D., Triantis, K. A., . . . Thebaud, C. (2015). Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters*, *18*(2), 200-217. https://doi.org/10.1111/ele.12398

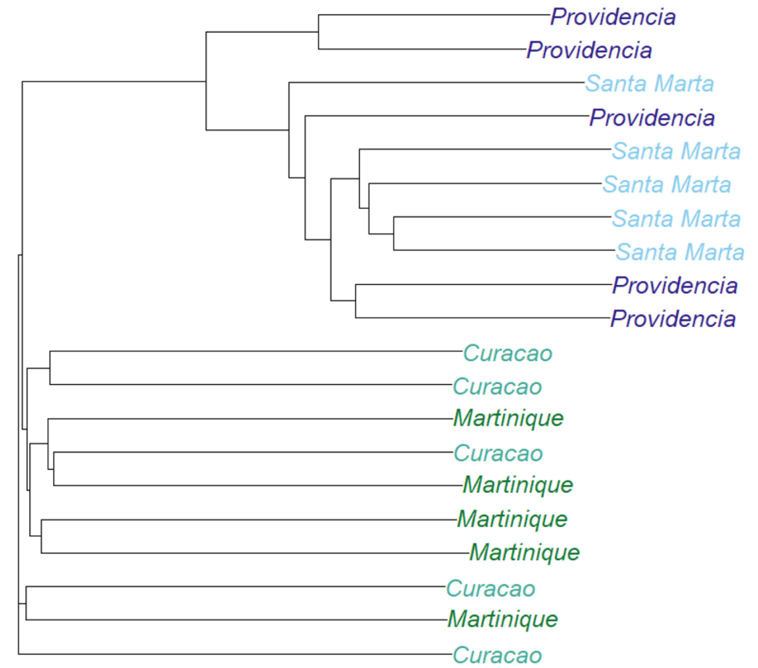
Whitham, T. G., Bailey, J. K., Schweitzer, J. A., Shuster, S. M., Bangert, R. K., Leroy, C. J., Lonsdorf, E. V., Allan, G. J., DiFazio, S. P., Potts, B. M., Fischer, D. G., Gehring, C. A., Lindroth, R. L., Marks, J. C., Hart, S. C., Wimp, G. M., & Wooley, S. C. (2006). A framework for community and ecosystem genetics: from genes to ecosystems [Review]. *Nature Reviews Genetics*, *7*(7), 510-523. https://doi.org/10.1038/nrg1877

Supplementary



Supplementary Figure 1: Constituent metrics of the continuity metric, with species grouped by family along the x-axis, and diversity metrics grouped by diversity facet in the y-axis. This plot is a visual aid in understanding the relationship between the species and genetic diversity metrics and the continuity metric. Light blue represents the Caribbean, whilst dark red represents the Western Indian Ocean. Horizontal lines represent the family diversity value for each bioregion. Each dot represents the genetic diversity metric for each species. If we were to calculate the continuity metric for each species, it would be proportional to the length of the vertical lines, or the difference between the species diversity value and genetic diversity value. Values at both levels have been plotted on the same y-axis after scaling.





Supplementary Figure 2: Phylogenies at both the species and genetic levels of diversity. A) The phylogeny for Pomacentridae family members present in the Caribbean, derived from Rabosky et al. (2018). B) The within-species phylogeny for Pomacentridae member, Chromis multilineata, derived from per individual SNP genotypes.

Discussion

Overview

The aim of this thesis has been to develop and evaluate a conceptual framework of diversification through organisational scale. By doing so, I hoped to contribute towards removing the arbitrary boundaries between the fields of ecology and evolution and make the argument that there are universal processes that drive the emergence not at, but through, the differing levels within the scale of biological organisation – at least from the genetic to the community levels. The approach has been empirically two-pronged, using *in silico* experimentation, and *in vivo* observation of a tropical reef fish study system. In chapter 1, we applied *in silico* experimentation to clarify the conceptual framework, formulate a new "continuity" metric, and establish expectations for pattern and process within this new framework. Through this modelling exercise we expanded on existing correlative methods and generated hypotheses for the roles of abundance, dispersal, environmental tolerances, and competition. We also posited that speciation acts as a diversity partitioning mechanism between the population and species levels of diversity. To validate our simulation results, and derived inferences, we compared simulated patterns and processes to *in vivo* observations. To make the comparison between the genetic and species levels of organisation, we needed comparable data sets at both levels in our tropical reef fish study system. At the species level, we compiled pre-existing community and phylogenetic data from the literature. At the genetic level, we expanded a ddRADseq-derived (Peterson et al., 2012) single nucleotide polymorphism (SNP) data set from the Western Indian Ocean (Donati et al., 2021), tuning the sampling scheme to allow for comparison between oceanic regions as well as biological traits and phylogenetic relationships. This exercise was insightful in itself and highlighted the importance of the environmental context when investigating the role of biological traits on patterns of genetic diversity. We now had a dataset covering both the species and genetic levels that we could use to validate against the results of the *in silico* experimentation in chapter 1. Mirroring and expanding on the analytical approach of chapter 1, we were able to compare the expectations for pattern and process established *in silico* with *in vivo* observations. Overall, we found a high level of correspondence between both the two approaches and with expectations derived from the literature. The joint results support the validity of a conceptual framework with universal processes driving diversity patterns across organisational scale, and a re-evaluation of speciation as an eco-evolutionary mechanism.

Unified processes generate emergent diversity patterns.

Eco-evolutionary processes are often divided across levels of organisational scale, with evolutionary mechanisms acting within species at the population/individual levels and ecological mechanisms measured at the community level between species and the environment (Hamilton, 2021; Jørgensen & Fath, 2014). To reconcile these processes, the approach has been to couple these processes at each level (Pelletier et al., 2009; Ware et al., 2019). Alternatively, we can view these processes as occurring across all levels of organisation simultaneously, with their subsequent patterns emerging through each level. In chapter 1 we replicate observed patterns of species richness, including relationships between the population and species levels (Manel et al., 2020; Schmidt, Munshi-South, et al., 2022; Vilcot et al., 2023). Through the model, we generated these patterns by configuring processes at the population level using observed species traits and allowing patterns of species diversity to emerge through speciation (Chapter 1, Supplementary Figure 1). Mismatches between simulated and observed species richness patterns are present and are likely due to a lack of resolution and input inaccuracies such as freshwater outflow and cold water upwellings (Floeter et

al., 2008). This particular model has been previously applied to study how lower-level population dynamics produce emergent diversity patterns at the species level of biological organisation (Boschman et al., 2021; Hagen, 2022; Oskar Hagen et al., 2021). Despite this, the conceptual significance of a single model framework that produces patterns concordant with both the population level (Leugger et al., 2022) and the species level (Boschman et al., 2021) has gone understated. In the model framework implemented in gen3sis (O. Hagen et al., 2021), ecological and evolutionary processes are not run in parallel nor coupled at each level of organisation, but are explicitly simulated through the same dynamics at the population level (O. Hagen et al., 2021). Dispersal between communities and gene-flow between geographically distinct populations are both simulated through the same dispersal function – i.e., the abstract movement of individuals from one grid cell to another (O. Hagen et al., 2021). Trait evolution of species is the emergent property of the mutation and selection of trait values, per-population, aggregated to the species level (O. Hagen et al., 2021). The novelty of chapter 1 was to highlight the existing, under-emphasised unity of process across organisational scale (Ware et al., 2019). This was achieved by directly comparing the emergent patterns of diversity at the population and species levels, and finding explanatory variables that drive discordant patterns. Although this is not a new goal (Antonovics, 1976; Lamy et al., 2013; Lamy et al., 2017; Laroche et al., 2015; Schmidt, Dray, et al., 2022; Schmidt, Munshi-South, et al., 2022; Vellend, 2005, 2010; Vellend & Geber, 2005; Vellend et al., 2014; Vilcot et al., 2023), by simplifying the eco-evolutionary framework down to a single set of interacting processes, it has become much easier to identify factors associated with the variable patterns of diversity we observe between levels of organisation (Decocq et al., 2021; Lawrence, 2020; Reisch & Hartig, 2021; Reisch & Schmid, 2019; Schmidt, Dray, et al., 2022).

In the introduction of this thesis, these processes were presented whereby dispersal/gene-flow, drift, and selection are unified (Introduction, Table 1), but mutation and speciation are not. Speciation is a complex process that spans a world of literature of its own (Seehausen et al., 2014), and is often considered to be the result of reproductive isolation between populations (Mayr, 1963). It involves the transition from one organisation level to another, with two divergent populations becoming two sibling species, unlike mutation which does not necessarily involve a shift of diversity between organisation levels (but consider chromosomal rearrangements and genome duplication). For me, speciation as a macro eco-evolutionary mechanism for transitioning diversity is the most impactful insight derived from the modelling exercise in chapter 1. Conceptually, it reinforces the idea that diversity patterns are not formed in parallel across levels of organisation, but emerge up from lower to higher levels (e.g., the dispersal of individuals) via transitionary processes such as speciation. Modelling, however, does not necessarily correspond to reality, and we need to validate our models against reality – we must therefore place the explanatory variables identified *in silico* into the context of the *in vivo* observed patterns uncovered in both chapter 3 and the literature.

How universal processes drive the emergence of diversity through organisational scale.

Chapter 1 and Chapter 3 were planned in conjunction, with the modelled tropical reef fish study system being the same as the observed. Through their comparison, the biological traits found to be significantly associated with the relationship between the genetic and species levels of organisation in the observed dataset corresponded with the significant model parameters in chapter 1 (Chapter 1, Figure 3; Chapter 3, Figure 5). In both the modelled and observed datasets, these associations were uncovered through the comparison of traits and seascape to the continuity metrics calculated for different facets of diversity (Chapter 3, Table 1) – the continuity metric being the scaled relative ratio of diversity at the genetic/population and species levels of organisation (Chapter 1, equation

2). We also found correspondence between the significant positive correlation between the genetic/population level and the species level in the β-diversity facet (Chapter 1, Figure 2A; Chapter 3, Figure 3).

In the observed dataset, all three biological variables (abundance, dispersal, and speciation rate) included in the analysis were significantly associated with the continuity metric across different facets of diversity (Chapter 3, Figure 5). These roughly correspond to initial abundance, dispersal, and speciation threshold parameters in the simulation model (Chapter 1, Figure 3). The significance of these variables in both *in silico* experiments and *in vivo* observations gives confidence in the result (Pelletier et al., 2009), but their interpretation remains complex. Directional relationships between a variable and the continuity metric for a facet of diversity can be interpreted as an association of increasing variable values with a relative shift of diversity towards the genetic or species level. Continuity values of 0 would indicate that genetic and species metric values are relatively the same, i.e., they are positive correlated. For example, in the observed dataset, we find that an increase in the speciation rate (Rabosky et al., 2018) corresponds to more species diversity relative to the amount of genetic diversity in the system – at least for phylogenetic diversity, alpha richness, and gamma richness (Chapter 3, Figure 5). But how do we gain understanding of the mechanism generating this pattern? For each significant biological variable, we remind ourselves of four possible scenarios:

Scenario 1. The variable represents a process that drives the emergence of diversity patterns at only one level enough to create a sufficient discordance between levels to be detected in the continuity metric.

Scenario 2. The variable represents a process that drives the emergence of diversity patterns at both levels in opposite directions enough to create a sufficient discordance between levels to be detected in the continuity metric.

Scenario 3. The variable represents a process which is itself a mechanism that determines how diversity emerges through organisational scale. I.e., the influence of the variable on each level of diversity is negligible compared to the much greater influence it has on the emergence of diversity through the levels.

Scenario 4. The variable influences either diversity level, but has no impact on how diversity emerges through organisational scale.

By looking also at the relationship between each biological variable and the constituent diversity metrics (the genetic metric and the species metric), we can gain a better idea of which of the three scenarios is most likely. In the case of alpha and gamma richness (expected heterozygosity vs. species richness), increasing abundance is associated with a decrease in expected heterozygosity at the genetic level, and with increasing species richness at the species level. In this case we find that the effect of the variable is opposite for each level of diversity, likely driving the large shift towards the species level as reflected in the continuity metric (Chapter 3, Figure 5). In this instance we would expect scenario 2 to be at play, despite the counter-intuitive relationship between expected heterozygosity and abundance which may be a peculiarity of the study system (Donati et al., 2021; Hamilton, 2021; Chapter 2). Despite this unexpected relationship between abundance and heterozygosity, we recover the same directionality in the continuity metric in the modelled results (Chapter 1, Figure 3). This pattern of continuity in the alpha and gamma richness facets (Chapter 3, Table 1) and abundance can be contrasted to the relationship between abundance and observed mean pairwise distance (MPD) across organisational levels, where increasing abundance is associated with a decrease in genetic MPD, but no discernible association at the species level

(Chapter 3; Figure 5). In this instance we could suspect scenario 1, although this pattern was not produced in the simulated dataset (Chapter 1, Figure 3). Finally, we have evidence for scenario 3 (association with continuity between organisational levels, but not with the underlying diversity metrics) in the relationship between speciation rate and increasing species diversity relative to genetic diversity in alpha and gamma richness (Chapter 3; Figure 5). Speciation rate is not a significant variable in determining either expected heterozygosity or species richness, yet it is significantly associated with the ratio between the two (Chapter 3; Figure 5). This same pattern was uncovered in in the abstracted "speciation threshold" (Chapter 1, Figure 3). In the modelled data, the speciation threshold was interpreted as an abstract proxy for a lineage's "ease of speciation" or disposition to reproductive isolation, as is heritably variable across lineages (Feder et al., 2012; Ravinet et al., 2017; Seehausen et al., 2014). The speciation threshold model parameter can be seen as the inverse of the observed tip-specific speciation rate, which can also be interpreted as a measure of a species' "ease of speciation" (Rabosky et al., 2018). In the simulations, it was speculated that speciation works as a partitioning mechanism of diversity between the genetic and species levels of diversity (see Chapter 1 discussion), whereby the individuals that become reproductively isolated into a new species effectively take their diversity with them: a new species is formed, but both the old and the new species contain reduced genetic diversity through their isolation – an effective population bottleneck for both (Hamilton, 2021). The patterns we find in both the modelled and observed data support the speciation diversity partitioning hypothesis.

Interestingly, the effect of dispersal capacity was the opposite in the modelled and observed datasets, with increasing dispersal corresponding to more population diversity relative to species diversity in the modelled data and more species diversity relative to population diversity in the observed data (Chapter 1, Figure 3; Chapter 3, Figure 5). In both instances they are significant, so it seems likely that dispersal plays a role in influencing the emergence of diversity through organisational scale, but the exact mechanism is cryptic. In the modelled data, we hypothesised that higher dispersal should result in more gene-flow between populations (Donati et al., 2021), thereby reducing the chance of allopatric speciation through population isolation and genetic drift (Hamilton, 2021). To explain the reverse effect uncovered in the observed data, we proposed that species with greater dispersal capacities should be exposed to greater environmental heterogeneity (Battisti et al., 2019; Monaco et al., 2020; Sunday et al., 2015) and maintain larger range sizes that are more likely to become disconnected by the development of hard barriers to dispersal such as tectonic movements (Alzate & Onstein, 2022; Bowman et al., 2002; Lester et al., 2007). This mismatch is interesting, and might be related to the timescale at which the consequent patterns of dispersal emerge – which has been an historical argument for the separation of ecological and evolutionary processes (Pelletier et al., 2009). I think this highlights a major limitation in the modelling approach implemented in chapter 1 – that all the processes act at the same timescale in a stepwise fashion (O. Hagen et al., 2021). It was a discussion point during the development of the model and is mitigated against through the flexibility of each functional module (speciation, dispersal, evolution, and ecology). Dispersal, however, is an abstract connection between grid cells that results in either colonisation of new cells, or homogenisation of inhabited cells. The homogenising effect of dispersal through gene-flow, which can be rapid even with few migrants (Åkesson et al., 2016; Vilà et al., 2003), could occur at a much faster rate than differentiation through genetic drift which is dependent on effective population size (Jorde & Ryman, 2007), whilst they occur at the same rate in the chapter 1 model configuration. This oversight in the model set up could have resulted in the mismatch. This is likely compounded by a similar comparison between rapid homogenisation through gene-flow and the difficulty for a species to become established in a new community – although I believe this to be less likely as incumbency effects are incorporated in the chapter 1

model set up through the initial abundance parameter and competition function. Through the twin experimental modelling and observational approaches, we found mostly overlapping results which gives us confidence in how we interpret the emergence of diversity through organisational scale. The mismatches, however, point to a greater complexity than what we have been able to address here, such as temporal scale, leaving plenty of future avenues to investigate.

Future studies and conservation implications

The original plan for this thesis intended to expand further down the scale of biological organisation to the genome level. In our modelling approach, the basal unit is a geographic population (O. Hagen et al., 2021), for our observed data, we settled on a RADseq approach (Baird et al., 2008; Peterson et al., 2012) which allowed a trade-off between the effort and cost of comparably sampling and genotyping many phylogenetically distant species (Chapter 2, Figure 1) across many sample sites against genomic resolution. This use of many non-model organisms made, for us, whole genome sequencing unfeasible. The most interesting next step, in my opinion, would be to take the conceptual framework and accompanying dynamics, explored from the individual level through to the community level, and expand it further down the organisational scale to within individuals. Individuals are not homogenous units, comprised of multiple genomes which themselves are aggregations of many chromosomes, genes and ultimately nucleotides. Each genomic landscape is shaped by a world of molecular dynamics, which have cascading repercussions up the scale of biological organisation (Bailey et al., 2009; Ellegren & Galtier, 2016; Seehausen, 2004; Seehausen et al., 2014; Singhal et al., 2018; Ware et al., 2019). Known variation across the genome and between genomes across chromosomes and chromosome sets[1] is influenced by effective population size, differential selection across the genome, variants introduced by dispersal, and of course the many faces of mutation (Ellegren & Galtier, 2016). We could ask questions such as, "do patterns of diversity between chromosomes correspond to patterns of diversity at the individual, population and species levels of organisation?". I believe the conceptual framework would likely change as we investigate further complexity, but the comparative method of correlating between levels, and use of continuity metrics should still be applicable. Through further expansion, we would gain a more holistic understanding of biodiversity formation across the full spectrum of life.

Another, less ambitious, direction would be to split diversity measures further, into neutral and non-neutral categories (Hubbell, 2001; Kimura & Weiss, 1964) to see if we can tease apart specific selective dynamics that may affect how diversity emerges through scale. This would be interesting, particularly within genomes when dealing with non-discrete linkage and selective genomic landscapes (Ellegren & Galtier, 2016). In our current conceptual framework of unified eco-evolutionary process across organisational scale, we imply that most processes occur between species (e.g., dispersal is the movement of individuals or their propagules). But perhaps if we start digging into the more basal aggregates of biological units such as genes, we will find other processes are more influential at these levels – similar to postulates introduced by Dawkins (1976) in the Selfish Gene, which remain under-explored in the literature (Gardner & Welch, 2011).

In a different direction, this framework has conservation implications. Conservation is often focussed at the species level, attempting to prioritise hotspots of richness, phylogenetic diversity, functional diversity, and endemism (Bottrill et al., 2008). However, there is a growing field of conservation genetics aiming to also highlight the importance, and enable the conservation, of genomic diversity

[1] As far as I am aware, there is no single word for a set of chromosomes. I.e., a haploid genome has one set, a diploid two, a hexaploid six. I would like to propose we call a set of chromosomes a "ploid". Polyploid, or "many ploids". In the scale of biological organisation, a ploid would be one level.

and resilience to future change (Hohenlohe et al., 2021). In our framework we holistically investigate both levels of organisational diversity as well as the processes that are important in driving their emergence. If we better understand how this plays out spatially – which is an important factor (Chapter 3, Figure 4,5; Lamy et al., 2013; Schmidt, Dray, et al., 2022; Schmidt, Munshi-South, et al., 2022; Vilcot et al., 2023) – we will be able to apply more holistic management schemes. Further, it should also be possible to leverage the modelling approach of chapter 1 to future predictions of joint population and species diversity under climate change. The time frame at which the gen3sis model is run is highly dependent on the configuration of the different processes and the interpretation of what the model represents (O. Hagen et al., 2021). In chapter 1, we configured the model to run over deep geological time (200 ma to present) at 166.7 ky time steps covering dynamics at large spatial and temporal scales. This is not required, and instead we could ask the question, "how would we expect population and species level diversity, and the continuity between the two, to respond to climate change?". We could produce a set of static contemporary seascapes with dynamic sea surface temperature regimes run at yearly intervals over the next 100-200 years in accordance with different climate projections and seed the model with known species distributions, traits, and phylogenetic relationships. Using mechanistic models to predict species richness responses would be innovative in itself, but further innovation would be to simultaneously attempt to predict population structure as in Leugger et al. (2022), incorporating continuity across levels of organisation – at least assessing the expected spatial correspondence between the two would have pertinent implications for conservation genetics (Hohenlohe et al., 2021).

Methodological contribution

An unexpected outcome of the work throughout this thesis has been the development of the continuity metric in chapter 1, which stemmed from the limited explanatory power of correlating measures of diversity across levels of organisation (Schmidt, Dray, et al., 2022; Vilcot et al., 2023). Prior to this thesis, studies have been limited to correlative approaches whereby multiple comparisons between the species and genetic levels of diversity are taken across species, geographic locations, or both (Antonovics, 1976; Lamy et al., 2013; Lamy et al., 2017; Laroche et al., 2015; Schmidt, Dray, et al., 2022; Schmidt, Munshi-South, et al., 2022; Vellend, 2005, 2010; Vellend & Geber, 2005; Vellend et al., 2014; Vilcot et al., 2023). In statistical terms, the response variable is not each single species-genetic comparison, it is the overall relationship of species per site, or it is the overall relationship of a single species across multiple sites – the response is the characteristic of a fitted model. Nevertheless, this correlative method provides valuable information as to whether a direct association exists between levels at all, and if so, the direction and strength of that relationship under different environmental conditions (Lamy et al., 2013) or biological traits (Vilcot et al., 2023). However, there must be a significant relationship to allow inference. In chapter 3, we uncovered an overall positive correlation between the species and genetic levels of organisation in the beta richness facet of diversity (Chapter 3, Figure 3). But after partitioning the dataset into the Caribbean and Western Indian Ocean, we found that the association was strong only in the Western Indian Ocean and non-significant in the Caribbean. We could no longer view the trend as global and can infer only that something particular to the Western Indian Ocean allows a positive correlation to emerge. From the Caribbean data, all we know is that the there is no association. This drawback is even more limiting in understanding biological traits. At a single geographic site, we can only retrieve the correlation derived from multiple lineages (Schmidt, Munshi-South, et al., 2022). To compare how each lineage, and their associated traits, differ, we need multiple geographic sampling locations to see how each lineage responds differently to changing environmental conditions – fitting a correlation across sites per lineage. This approach requires a significant sampling effort to achieve

sufficient statistical power, especially when correlations are already weak (Vilcot et al., 2023). On the other hand, the continuity metric decomposes the species-genetic comparison per lineage into a ratio between the two and can be associated with the respective lineage's traits. This means that the investigator no longer needs multiple sampling sites and consistently significant relationships between levels of organisation to compare to either biological traits or environmental variables. Importantly, there does not have to be a direct relationship between levels of organisation to understand how diversity emerges through them. It allows us to think more flexibly about diversity across levels of organisation. The questions we can ask go from, "do patterns of diversity at different organisational levels correlate, and why?" to "what is associated with unbalanced ratios of diversity across levels of organisation and what are the causal mechanisms?". It is important to make clear that the continuity metric is not a replacement for correlative approaches, but a complementary method. In fact, the strength is that the null expectation of a test of continuity vs explanatory variables is that there is a 1:1 positive correlation. If both organisational levels of diversity were to relatively increase at the same rate across lineages or geographic sites, the continuity metric would be 0 for each data point – this is seen in the beta richness facet in the observed dataset where we find a positive correlation between the species and genetic levels (Chapter 3, Figure 3) and continuity metric values tightly centred around 0 (Chapter 3, Figure 4, Figure 5). But, from these zero continuity values alone it would not be possible to convincingly infer the positive relationship. Another caveat of the continuity metric is that values are relative to the sampling scheme. This relativity means that absolute continuity values are meaningless when compared between studies. It also means that outliers will have a strong effect on the directionality of the metric, as seen for the variation in pairwise distance (VPD) facet in the observed dataset (Chapter 3, Supplementary Figure 1). In the case of VPD, *Canthigaster rostrata* and *Chromis multilineata* have very high genetic-level VPD values that depress values across the other species. These outliers minimise any variation that exists between the rest of the sampled species as well as subsequent continuity values. Despite these caveats, the application of the continuity metric allows us to explore the influence of both biological traits and environmental variables directly. We no longer have to rely on the abstraction of a significant model fit to investigate eco-evolutionary dynamics across levels of biological organisation – this advance should significantly facilitate future studies.

Challenges and reflections

The development of this thesis has been a learning process and the chapters were compiled mostly in serial in the order they are arranged here. Many of the concepts presented by the end of chapter 3 and in the discussion are the product of a continuous learning process, meaning that the conceptual and analytical framework has become more comprehensive as time has gone on. I believe that the incorporation of more facets of diversity and the deconstruction and comparison of continuity to its constituent metrics in chapter 3 reflects this. In a perfect world, the comparison between chapter 1 and 3 could be much more complete, but future revisions of these works should rectify this. I should also admit that attempting to review, criticise, and conceptually reorganise over 100 years of two biological traditions is an overambitious task, and is far from over. I believe the conceptual framework of unified process through organisational scale, the insights into the partitioning of diversity through speciation in particular, and the methodological contribution of the continuity metric are significant improvements in the field. But, with such a vast body of literature to wrangle, I am sure there are excellent published works out there that I have, to the detriment of my arguments, overlooked. For this negligence I apologise, and I look forward to corrections, improvements, and further development of our understanding.

Conclusion

In conclusion, this thesis set out to challenge the current division between ecological and evolutionary study by demonstrating that diversity across levels of organisational scale can be understood as a product of universal processes. To achieve this, a revised conceptual framework of diversification was introduced and developed whereby singular processes cascade up through the various aggregations of organisational levels, being modified along the way by transitionary processes such as speciation. Whilst this framework is far from the last word in the merging of eco-evolutionary study, it is a step in the correct direction. Through the development of the continuity metric, future work on understanding the dynamic emergence of diversity from the genetic to species level of organisation is facilitated. The continuity metric should also serve to expand this line of investigation further down the organisational scale into the depths of within-individual patterns of diversity. Pushing this diversification framework across more levels of biological organisation will likely necessitate its revision, but should allow a much more holistic understanding of the formation of biological diversity. Only by accepting the notion that the life is "…queerer than we can suppose" can we stomach the flexibility of thought required to understand it as a single, unified system; from the smallest nucleotide to the mightiest ecosystem, in all its wondrously confusing diversity.

References

Åkesson, M., Liberg, O., Sand, H., Wabakken, P., Bensch, S., & Flagstad, Ø. (2016). Genetic rescue in a severely inbred wolf population. *Molecular Ecology*, *25*(19), 4745-4756.

Alzate, A., & Onstein, R. E. (2022). Understanding the relationship between dispersal and range size. *Ecology Letters*, *25*(10), 2303-2323. https://doi.org/10.1111/ele.14089

Antonovics, J. (1976). The input from population genetics: "The new ecological genetics". *Systematic Botany*, *1*(3), 233-245. https://doi.org/10.2307/2418718

Bailey, J. K., Hendry, A. P., Kinnison, M. T., Post, D. M., Palkovacs, E. P., Pelletier, F., Harmon, L. J., & Schweitzer, J. A. (2009). From genes to ecosystems: an emerging synthesis of eco-evolutionary dynamics. *New Phytologist*, *184*(4), 746-749. https://doi.org/https://doi.org/10.1111/j.1469-8137.2009.03081.x

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *Plos One*, *3*(10), Article e3376. https://doi.org/10.1371/journal.pone.0003376

Battisti, C., Fanelli, G., Amori, G., Luiselli, L., & Capizzi, D. (2019). Do disturbance-sensitive and habitat-specialized species have a smaller range size? Evidence for a set of common mammals at regional scale. *Ethology Ecology & Evolution*, *31*(5), 479-490. https://doi.org/10.1080/03949370.2019.1626487

Boschman, L. M., Cassemiro, F. A., Carraro, L., de Vries, J., Altermatt, F., Hagen, O., Hoorn, C., & Pellissier, L. (2021). South American freshwater fish diversity shaped by Andean uplift since the Late Cretaceous. *BioRxiv*, 2021.2005. 2014.444133.

Bottrill, M. C., Joseph, L. N., Carwardine, J., Bode, M., Cook, C., Game, E. T., Grantham, H., Kark, S., Linke, S., McDonald-Madden, E., Pressey, R. L., Walker, S., Wilson, K. A., & Possingham, H. P. (2008). Is conservation triage just smart decision making? *Trends in Ecology & Evolution*, *23*(12), 649-654. https://doi.org/https://doi.org/10.1016/j.tree.2008.07.007

Bowman, J., Jaeger, J. A. G., & Fahrig, L. (2002). Dispersal distance of mammals is proportional to home range size. *Ecology*, *83*(7), 2049-2055. https://doi.org/10.2307/3071786

Dawkins, R. (1976). *The selfish gene*. Oxford university press.

Decocq, G., Guiller, A., Kichey, T., Van de Pitte, K., Gallet-Moron, E., Honnay, O., & Closset-Kopp, D. (2021). Drivers of species and genetic diversity within forest metacommunities across agricultural landscapes of different permeability. *Landscape Ecology*. https://doi.org/10.1007/s10980-021-01296-6

Donati, G. F. A., Zemp, N., Manel, S., Poirier, M., Claverie, T., Ferraton, F., Gaboriau, T., Govinden, R., Hagen, O., Ibrahim, S., Mouillot, D., Leblond, J., Julius, P., Velez, L., Zareer, I., Ziyad, A., Leprieur, F., Albouy, C., & Pellissier, L. (2021). Species ecology explains the spatial components of genetic diversity in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1959), 20211574. https://doi.org/doi:10.1098/rspb.2021.1574

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422-433. https://doi.org/10.1038/nrg.2016.58

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*(7), 342-350. https://doi.org/10.1016/j.tig.2012.03.009

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*(1), 22-47. https://doi.org/https://doi.org/10.1111/j.1365-2699.2007.01790.x

Gardner, A., & Welch, J. J. (2011). A formal theory of the selfish gene. *Journal of Evolutionary Biology*, *24*(8), 1801-1813. https://doi.org/https://doi.org/10.1111/j.1420-9101.2011.02310.x

Hagen, O. (2022). Coupling eco-evolutionary mechanisms with deep-time environmental dynamics to understand biodiversity patterns. *Ecography*, *n/a*(n/a), e06132. https://doi.org/https://doi.org/10.1111/ecog.06132

Hagen, O., Fluck, B., Fopp, F., Cabral, J. S., Hartig, F., Pontarp, M., Rangel, T. F., & Pellissier, L. (2021). gen3sis: A general engine for eco-evolutionary simulations of the processes that shape Earth's biodiversity [Article]. *PLOS Biology*, *19*(7), 31, Article e3001340. https://doi.org/10.1371/journal.pbio.3001340

Hagen, O., Skeels, A., Onstein, R. E., Jetz, W., & Pellissier, L. (2021). Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. *Proceedings of the National Academy of Sciences*, *118*(40), e2026347118. https://doi.org/doi:10.1073/pnas.2026347118

Hamilton, M. B. (2021). *Population genetics*. John Wiley & Sons.

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, *30*(1), 62-82. https://doi.org/https://doi.org/10.1111/mec.15720

Hubbell, S. (2001). The Unified Neutral Theory of Biodiversity and Biogeography. *Princeton University Press, Princeton, New Jersey, USA. Hughes, RG (1986) Theories and models of species abundance. American Naturalist*, *128*, 879-899.

Jorde, P. E., & Ryman, N. (2007). Unbiased Estimator for Genetic Drift and Effective Population Size. *Genetics*, *177*(2), 927-935. https://doi.org/10.1534/genetics.107.075481

Jørgensen, S. E., & Fath, B. (2014). *Encyclopedia of ecology*. Newnes.

Kimura, M., & Weiss, G. H. (1964). The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance. *Genetics*, *49*(4), 561-576. https://doi.org/10.1093/genetics/49.4.561

Lamy, T., Jarne, P., Laroche, F., Pointier, J. P., Huth, G., Segard, A., & David, P. (2013). Variation in habitat connectivity generates positive correlations between species and genetic diversity in a metacommunity [Article]. *Molecular Ecology*, *22*(17), 4445-4456. https://doi.org/10.1111/mec.12399

Lamy, T., Laroche, F., David, P., Massol, F., & Jarne, P. (2017). The contribution of species-genetic diversity correlations to the understanding of community assembly rules. *Oikos*, *126*(6), 759-771. https://doi.org/10.1111/oik.03997

Laroche, F., Jarne, P., Lamy, T., David, P., & Massol, F. (2015). A Neutral Theory for Interpreting Correlations between Species and Genetic Diversity in Communities. *American Naturalist*, *185*(1), 59-69. https://doi.org/10.1086/678990

Lawrence, E. R. (2020). *Synthesizing vertebrate population richness and genetic diversity across the American continents* Concordia University].

Lester, S. E., Ruttenberg, B. I., Gaines, S. D., & Kinlan, B. P. (2007). The relationship between dispersal ability and geographic range size. *Ecology Letters*, *10*(8), 745-758. https://doi.org/10.1111/j.1461-0248.2007.01070.x

Leugger, F., Broquet, T., Karger, D. N., Rioux, D., Buzan, E., Corlatti, L., Crestanello, B., Curt-Grand-Gaudin, N., Hauffe, H. C., & Rolečková, B. (2022). Dispersal and habitat dynamics shape the genetic structure of the Northern chamois in the Alps. *Journal of Biogeography*, *49*(10), 1848-1861.

Manel, S., Guerin, P. E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., & Pellissier, L. (2020). Global determinants of freshwater and marine fish genetic diversity. *Nature Communications*, *11*(1). https://doi.org/ARTN 692

10.1038/s41467-020-14409-7

Mayr, E. (1963). Animal Species and Evolution (Belknap, Cambridge, MA).

Monaco, C. J., Bradshaw, C. J. A., Booth, D. J., Gillanders, B. M., Schoeman, D. S., & Nagelkerken, I. (2020). Dietary generalism accelerates arrival and persistence of coral-reef fishes in their novel ranges under climate change. *Global Change Biology*, *26*(10), 5564-5573. https://doi.org/10.1111/gcb.15221

Pelletier, F., Garant, D., & Hendry, A. P. (2009). Eco-evolutionary dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1523), 1483-1489. https://doi.org/doi:10.1098/rstb.2009.0027

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *Plos One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

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*(7714), 392-395. https://doi.org/10.1038/s41586-018-0273-1

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlovic, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, *30*(8), 1450-1477. https://doi.org/10.1111/jeb.13047

Reisch, C., & Hartig, F. (2021). Species and genetic diversity patterns show different responses to land use intensity in central European grasslands. *Diversity and Distributions*, *27*(3), 392-401. https://doi.org/10.1111/ddi.13199

Reisch, C., & Schmid, C. (2019). Species and genetic diversity are not congruent in fragmented dry grasslands. *Ecology and Evolution*, *9*(1), 664-671. https://doi.org/10.1002/ece3.4791

Schmidt, C., Dray, S., & Garroway, C. J. (2022). Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity. *Evolution*, *76*(1), 86-100. https://doi.org/https://doi.org/10.1111/evo.14407

Schmidt, C., Munshi-South, J., Dray, S., & Garroway, C. J. (2022). Determinants of genetic diversity and species richness of North American amphibians. *Journal of Biogeography*, *49*(11), 2005-2015. https://doi.org/10.1111/jbi.14480

Seehausen, O. (2004). Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, *19*(4), 198-207. https://doi.org/10.1016/j.tree.2004.01.003

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., Peichel, C. L., Saetre, G. P., Bank, C., Brannstrom, A., Brelsford, A., Clarkson, C. S., Eroukhmanoff, F., Feder, J. L., Fischer, M. C., Foote, A. D., Franchini, P., Jiggins, C. D., Jones, F. C., . . . Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*(3), 176-192. https://doi.org/10.1038/nrg3644

Singhal, S., Huang, H., Grundler, M. R., Marchán-Rivadeneira, M. R., Holmes, I., Title, P. O., Donnellan, S. C., & Rabosky, D. L. (2018). Does Population Structure Predict the Rate of

Speciation? A Comparative Test across Australia's Most Diverse Vertebrate Radiation. *The American Naturalist*, *192*(4), 432-447. https://doi.org/10.1086/699515

Sunday, J. M., Pecl, G. T., Frusher, S., Hobday, A. J., Hill, N., Holbrook, N. J., Edgar, G. J., Stuart-Smith, R., Barrett, N., Wernberg, T., Watson, R. A., Smale, D. A., Fulton, E. A., Slawinski, D., Feng, M., Radford, B. T., Thompson, P. A., & Bates, A. E. (2015). Species traits and climate velocity explain geographic range shifts in an ocean-warming hotspot. *Ecology Letters*, *18*(9), 944-953. https://doi.org/10.1111/ele.12474

Vellend, M. (2005). Species diversity and genetic diversity: Parallel processes and correlated patterns. *American Naturalist*, *166*(2), 199-215. https://doi.org/10.1086/431318

Vellend, M. (2010). Conceptual synthesis in community ecology [Review]. *Quarterly Review of Biology*, *85*(2), 183-206. https://doi.org/10.1086/652373

Vellend, M., & Geber, M. A. (2005). Connections between species diversity and genetic diversity. *Ecology Letters*, *8*(7), 767-781. https://doi.org/10.1111/j.1461-0248.2005.00775.x

Vellend, M., Lajoie, G., Bourret, A., Murria, C., Kembel, S. W., & Garant, D. (2014). Drawing ecological inferences from coincident patterns of population- and community-level biodiversity. *Molecular Ecology*, *23*(12), 2890-2901. https://doi.org/10.1111/mec.12756

Vilà, C., Sundqvist, A. K., Flagstad, Ø., Seddon, J., rnerfeldt, S. B., Kojola, I., Casulli, A., Sand, H., Wabakken, P., & Ellegren, H. (2003). Rescue of a severely bottlenecked wolf (<i>Canis lupus</i>) population by a single immigrant. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1510), 91-97. https://doi.org/doi:10.1098/rspb.2002.2184

Vilcot, M., Albouy, C., Donati, G. F. A., Claverie, T., Julius, P., Manel, S., Pellissier, L., & Leprieur, F. (2023). Spatial genetic differentiation correlates with species assemblage turnover across tropical reef fish lineages. *Global Ecology and Biogeography*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/geb.13637

Ware, I. M., Fitzpatrick, C. R., Senthilnathan, A., Bayliss, S. L. J., Beals, K. K., Mueller, L. O., Summers, J. L., Wooliver, R. C., Van Nuland, M. E., Kinnison, M. T., Palkovacs, E. P., Schweitzer, J. A., & Bailey, J. K. (2019). Feedbacks link ecosystem ecology and evolution across spatial and temporal scales: Empirical evidence and future directions. *Functional Ecology*, *33*(1), 31-42. https://doi.org/https://doi.org/10.1111/1365-2435.13267

Curriculum Vitae

thomaskeggin@hotmail.com

Employment History

PhD Candidate in Eco-Evolutionary Biology

ETH Zürich, Switzerland
03.2019 – 04.2023

The PhD project focussed on the formation of tropical reef fish biodiversity – from genes to ecosystems. Eco-evolutionary processes that generate biodiversity were simulated over deep geological time in R, then validated against real-world genetic, species, phylogenetic, and biological trait data. The result has been a re-evaluation of the conceptual frameworks underlying our understanding of how biodiversity arises.

The project involved planning and carrying out SCUBA-based fieldwork, a large-scale next generation sequencing project, and extensive model-construction, data management, and analysis in R.

Side projects have included the development of a genetic database for fish, teaching responsibilities, and participation in various data collection field trips.

Research Assistant

Durham University, United Kingdom.
11.2018 – 02.2019

Developing a genetic test (ddRADseq SNP panel) to identify anglerfish (*Lophius piscatorius*) to Northeast Atlantic populations of origin for fisheries management.

- Design of sampling scheme.
- Planning of laboratory and bioinformatics pipeline.
- Extraction of DNA and partial genomic library preparation.

Quality Assurance Officer

Astrea Bioseparations, Isle of Man.
02-06.2018

QA at a manufacturing site producing protein purification filtration systems.

- Management, review, and upkeep of standard operating procedures.
- Batch manufacturing record processing and release.
- Review of employee performance statistics.
- Employee training database development and implementation.
- Assigned to an internal project to set up the Laboratory Information Management System.

Fisheries Intern

Isle of Man Government, Isle of Man.
10-11.2017

- Horse mussel (*Modiolus modiolus*) reef mapping using video analysis.
- Electrofishing as part of the assisted breeding program for salmon (*Salmo salar*).
- Scallop (*Pecten maximus*) processing (age, weight, size, reproductive stage).

Agricultural Technician

Limagrain, United Kingdom.
07-10.2017

Seasonal harvesting labour for crop variety selection – harvesting wheat varieties and quantifying their yield.

Junior Health, Safety, and Environment Professional

Schlumberger, Kuwait.
10.2014 – 01.2015

Desert-based land seismic crew in the Burgan Oilfield.

- Inventory management and distribution of safety critical equipment.
- Accurate upkeep of crew chemical safety records.
- Monitoring and quality control of un-exploded ordnance release.
- Shared responsibility for safe working conditions for approximately 400 field-based personnel.
- Resolving workforce procedural conflicts in a highly multicultural and multilingual setting.
- Enforcement of environmental responsibilities.

Research Assistant

Durham University, United Kingdom
08-09.2013

Assisting in river-based data collection.

- Electrofishing for the mark and recapture of freshwater fish.
- Fish tagging using subcutaneous fluorescent gel injections and RFID chips.

Education

Master of Science (Research)

Durham University, United Kingdom
09.2015 – 12.2017

Thesis: Population genomics of two deep sea sharks: *Centroselachus crepidater* and *Deania calcea*.

- ddRADseq derived SNP genotyping.
- DNA extraction, library preparation, bioinformatics (*de novo* stacks pipeline), analyses.

Bachelor of Science in Biology

Durham University, United Kingdom
09.2011 – 05.2014

Conferences

INTECOL. 2022

Geneva, Switzerland

Presentation: Diversity patterns across biological organisation are determined by an interplay of biological traits and landscape dynamics through time.

Publications

A. Skeels, L.M. Boschman, I. R. McFadden, E.M. Joyce, O. Hagen, O. Jiménez Robles, W. Bach, V. Boussange, T. Keggin, W. Jetz, L. Pellissier. *Accepted*. **Paleoenvironments shaped the exchange of terrestrial vertebrates across Wallace's Line**. Science.

Keggin, T., C. Waldock, A. Skeels, O. Hagen, C. Albouy, S. Manel, L. Pellissier. *In revision*. **Diversity across organisational scale emerges through dispersal ability and speciation dynamics in tropical fish**. BMC Biology.

Polanco F, A., C. Waldock, T. Keggin, V. Marques, R. Rozanski, A. Valentini, T. Dejean, S. Manel, M. Vermeij, and C. Albouy. 2022. **Ecological indices from environmental DNA to contrast coastal reefs under different anthropogenic pressures.** Ecology and Evolution. 12:e9212.

Stauffer, S., M. Jucker, T. Keggin, V. Marques, M. Andrello, S. Bessudo, M. C. Cheutin, G. H. Borrero-Pérez, E. Richards, and T. Dejean. 2021. **How many replicates to accurately estimate fish biodiversity using environmental DNA on coral reefs?** Ecology and Evolution 11:14630-14643.

Keggin, T. 2017. **Population genomics of two deep sea sharks: *Centroselachus crepidater* and *Deania calcea***. Durham University.