Phylogeographic history and evolutionary diversification of pygmy perches (Teleostei: Percichthyidae)

BY

Sean J. Buckley
B.Sc. (Animal Behaviour) [Flinders University] 2015
B.Sc. (Hons) [Flinders University] 2016

Molecular Ecology Lab at Flinders University (MELFU)
Thesis submitted to Flinders University for the degree
of Doctor of Philosophy
College of Science and Engineering, Flinders University
Adelaide, South Australia
Primary supervisor: Prof. Luciano Beheregaray
Secondary supervisor: Assoc. Prof. Luciana Möller
Adjunct supervisor: Dr Peter Unmack
04/12/2020

Table of contents

Contents

List of figures

List of tables

List of appendices

Summary

Analysing the environmental and evolutionary histories of species provides an empirical framework for testing biogeographic hypotheses and understanding factors underlying past, current and future demographic responses. Expanding these analyses to a comparative framework also allows for assessing the role of species-specific traits on biogeographic patterns, identifying whether species responses might be shared or not. The temperate zone of Australia is divided into two contrasting regions in the southeast and southwest corner of the mainland, with substantial spatial and temporal environmental variability in the southeast but with pronounced topographic and climatic stability since the Pliocene in the southwest. Several biological groups, such as pygmy perches, are common to both biogeographic regions, and present suitable systems to evaluate and compare the interaction of environmental and evolutionary histories. Additionally, all pygmy perch species are of conservation concern, and understanding the contribution of evolutionary history to their adaptive potential is important for guiding management strategies.

This work aims to investigate the phylogeographic histories of several pygmy perch lineages across the two temperate bioregions and to identify environmental factors underlying patterns of divergence and persistence. To achieve those aims, I first generated genome-wide (based on ddRAD[1]) data to assess phylogenetic, demographic and population genetic patterns in pygmy perches from southeastern and southwestern Australia. These results were then integrated with environmental

[1] ddRAD is a reduced representation sequencing method that stands for “double digest restriction-site associated DNA”.

modelling for the two temperate regions to provide a more comprehensive assessment of their evolutionary histories. Complex coalescent modelling using 7,958 ddRAD loci containing 7,780 putatively unlinked SNPs[2] revealed the spatial complexity of Plio-Pleistocene environmental changes on the evolutionary patterns of the southern pygmy perch (*Nannoperca australis*), a lineage from southeast Australia. The findings indicate that aridification played a stronger role in isolating inland populations compared to sea level changes in coastal populations.

The *Nannoperca australis* patterns were compared with those inferred for the Yarra pygmy perch (*Nannoperca obscura*) across their co-distributed range. This analysis used 21,051 ddRAD loci and 17,389 SNPs for *N. australis* and 19,428 ddRAD loci and 15,715 SNPs for *N. obscura*. It revealed concordant demographic histories across a shared habitat refugium, but significant species-wide discordance in distribution patterns and connectivity. This disparity likely reflects differences in long-term standing genetic variation between species and may be a contributing factor to the variable success of recent conservation efforts for these threatened lineages.

Remarkably contrasting results were obtained for lineages endemic to the southwest Australia biodiversity hotspot in Western Australia. Here, populations belonging to the western pygmy perch species complex (*Nannoperca vittata* and *Nannoperca pygmaea*) were analysed using 19,426 ddRAD loci and 18,177 SNPs. The analyses revealed a number of anciently isolated lineages, including cryptic species, and reduced environmental fluctuations in this biodiversity hotspot, which allowed the

[2] SNP is a genomic marker that stands for “single nucleotide polymorphism”

accumulation and maintenance of pygmy perch diversity. A novel functional enrichment approach using fixed SNP differences was developed, which demonstrated potential mechanisms for reproductive isolation between cryptic species based on intracellular differences associated with chromosome arrangement.

Overall, these results illustrate the contrasting interplay between species evolutionary histories and environmental dynamics, with climatic stability allowing divergent lineages to persist but climatic variability creating more complex and fluctuating patterns of connectivity and persistence. Species responses to shared environmental change were not fully concordant and highlighted the role of intrinsic factors, such as genetic diversity, in allowing species to adapt to changing habitats. Altogether, environmental and species trait variation intersect to create an intricate network defining past species responses, adding to the complex nature of predicting adaptive resilience to environmental and climatic change. These findings provide important implications for the management of freshwater biota more widely, as well as specific to the pygmy perches, urging for protection of habitat for the newly identified lineages and informing *ex situ* and *in situ* conservation efforts.

Declaration

I, Sean J. Buckley, certify that this thesis:

1. does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and
2. to the best of my knowledge and belief, does not contain any material previously published or written by another person except where due reference is made in the text.

Acknowledgments

I would like to thank my supervisors, Professor Luciano Beheregaray, Associate Professor Luciana Möller and Dr. Peter Unmack for their support throughout this candidature. Your guidance has helped steer this research and my professional development in many ways, and I am grateful for your leadership.

I would also like to my colleagues within MELFU (Chris, Cat, Yuma, Katie, Kimberley, Andrea (AB), Andrea (ABC), Isabella, and Elly) and Flinders University more broadly (Bec and René). Your support – both academic and 'moral' – has made me feel at home in the lab and I am glad to be part of such a strong and supportive network. Additionally, I would like to thank Mark Adams and Michael Hammer for their assistance in sample collection, without whom this project would not be able to proceed.

I am grateful for the financial support provided by the Australian Research Council via a Linkage project (LP100200409) and a Future Fellowship project (FT130101068) to Luciano Beheregaray, as well as funding provided by the Australian Government Research Training Program Scholarship.

Lastly, but certainly not least, I would like to thank my beloved fiancé Benji, without whom this thesis would not be possible. Your unwavering support and patience, even during the toughest times, helped me keep pushing through, and I consider this thesis as much a reflection of your work as it is mine.

Chapter 1: General Introduction

Climate Change and Species Persistence

Anthropogenic climate change is placing unprecedented selective pressures on species globally, leading to unprecedented extinction rates (Le Roux *et al.* 2019). Whether species will be able to track suitable climates as they shift (Bridle & Vines 2007), or persist in new environments through genetic adaptation (Bradshaw & Holzapfel 2006; Healy *et al.* 2018) or phenotypic plasticity (Murren *et al.* 2015; Sandoval-Castillo *et al.* 2020) will largely depend on their intrinsic characteristics (Aitken *et al.* 2008; Waldvogel *et al.* 2020), including genetic (Fordham *et al.* 2014), demographic (Pearson *et al.* 2014; Williams *et al.* 2008) and ecological traits (Somero 2010; Travis *et al.* 2013). For example, species with low dispersal capacity are unlikely to be able to sufficiently track their ecological niche under rapid climate change (Hoffmann & Sgro 2011; Travis *et al.* 2013), and species with low genetic diversity may not possess the adaptive potential to respond to local environmental changes (Williams *et al.* 2008). Understanding the interaction of intrinsic species traits and extrinsic environmental conditions is key to determining how, which, or if species will be able to adapt to ongoing climate change (Harrisson *et al.* 2014).

Phylogeography – From Genetics to Genomics

The relationship between the historical environment and the genealogy of species is often used as a proxy for species responses to future climatic changes (Dawson *et al.* 2011; Fenderson *et al.* 2020; MacDonald *et al.* 2008). This approach falls within the domain of phylogeography, which often seeks to combine population genetic, phylogenetic and environmental analyses to provide a multifaceted overview of species' evolutionary histories (Cutter 2013; Edwards *et al.* 2016; Rissler 2016). Particularly, phylogeographic studies often focus on determining the relative role and

impact of past Earth history events on the evolution and diversification of species (Beheregaray 2008; Riddle 2016; Waters *et al.* 2020). The usage of historical environmental changes as a proxy for modern climate change is supported by analogous prehistoric climatic patterns, with ongoing contemporary climate change now analogous to patterns last observed in the Pliocene (Burke *et al.* 2018).

Phylogeographic investigations of diversification span the population-speciation continuum (Coates *et al.* 2018), ranging from the delineation of intraspecific genetic units (Moritz 1994) to delimiting species (Beheregaray & Caccone 2007; Carstens *et al.* 2013; De Queiroz 2007). This framework has helped to identify 'cryptic species', which lack the typical morphological, ecological or behavioural divergence that may ordinarily define species but entail highly genetically divergent lineages that may be reproductively isolated (Coates *et al.* 2018). Cryptic species represent an important component of biodiversity and provide insights into the mechanisms and patterns of speciation (Fišer *et al.* 2018; Seehausen *et al.* 2014; Struck *et al.* 2018). Thus, phylogeography provides an avenue to evaluate a breadth of factors underlying both micro- and macroevolutionary patterns (Avise *et al.* 2016; Papadopoulou & Knowles 2017).

Extending phylogeographic analyses from taxon-specific studies to a comparative framework provides an approach to determine whether species responses are shared (Potter *et al.* 2018). This enables addressing underlying mechanisms driving species responses to environmental change, such as particular ecological traits (Avise *et al.* 2016; Riddle 2016). Typically, comparative phylogeography has

operated based on a dichotomy of shared (concordant) or individual (discordant) evolutionary histories across species (Papadopoulou & Knowles 2016). Concordance is typically associated with either the ubiquity of an historical environmental change (Avise *et al.* 2016; Barrow *et al.* 2018; Zamudio *et al.* 2016) or the role of shared ecological traits driving species responses (Paz *et al.* 2015; Rincon-Sandoval *et al.* 2019). Contrastingly, discordance has been attributed to variation in species traits (Bell *et al.* 2017; Massatti & Knowles 2014; Zamudio *et al.* 2016) or an artefact of demographic stochasticity within lineages (Papadopoulou & Knowles 2016). Thus, a comparative phylogeographic framework allows for the investigation of the role of specific traits in their response to shared environmental change, an important factor in determining future adaptive potential (Healy *et al.* 2018; Pearson *et al.* 2014; Williams *et al.* 2008).

Phylogeography has traditionally relied upon organellar markers (mitochondrial or chloroplast DNA) to infer demographic and phylogenetic patterns (Brumfield *et al.* 2003; Fahey *et al.* 2014; Riddle *et al.* 2008), owing to their relatively rapid mutation rates, lack of recombination, and broad orthology across taxa (Brito & Edwards 2009; Wang 2010). However, genome-wide heterogeneity in evolutionary patterns (Cutter 2013; Degnan & Rosenberg 2009), due to processes such as incomplete lineage sorting (Maddison 1997; Nakhleh 2013) or horizontal transfer (Edwards *et al.* 2016), may create significant discordance across markers and between genetic lineages and the underlying population history (Bravo *et al.* 2019). This gene tree discordance might not be detected when using a single locus or a few loci (Brito & Edwards 2009), potentially biasing phylogeographic studies based on these markers.

Recent advancements in analytical approaches and data collection may overcome these limitations, and have significantly improved the ability of phylogeographic studies to explore the relationship between the historical environment and species evolution (Beheregaray 2008; Carstens *et al.* 2013; McCormack *et al.* 2013). The development of next-generation sequencing technologies over the last decade have made the collection of thousands of genome-wide markers feasible, even for non-model species (Carstens *et al.* 2012; Rissler 2016). For example, reduced representation approaches such restriction-associated DNA sequencing (RADseq) provide a cost-effective method for obtaining multi-locus data suitable for addressing complex phylogeographic questions (McCormack *et al.* 2013; Nunziata & Weisrock 2018). In tandem, recent advancements in coalescent modelling, informed by detailed geological and ecological history, have improved the ability to provide more nuanced inferences of phylogeographic patterns (Cutter 2013; Hickerson *et al.* 2010). For example, coalescent modelling approaches have been extended to incorporate gene flow (Hey 2010), large numbers of populations (Excoffier *et al.* 2013) or detect selection within lineages (Hickerson *et al.* 2010). Similarly, the development of the multispecies coalescent as a statistical approach to species delimitation has improved the robustness of using genetic data to delimit species (Degnan & Rosenberg 2009; Edwards *et al.* 2016). Together, these developments have reinvigorated the field of phylogeography, providing an avenue to test more complex questions about the interaction of environmental variation and species evolution.

Temperate Australia

Temperate regions across the globe demonstrate a history of variable environmental and climatic changes, driving complex patterns of evolutionary history (Smith *et al.* 2017; Sommer & Zachos 2009). Similarly, anthropogenic climate change is expected to significantly impact many temperate regions across the planet through increasing temperatures and more variable precipitation patterns (Bonada *et al.* 2007). Phylogeographic studies in temperate regions are thus important to understand both the role of environmental variation in the evolutionary histories of species, as well as inform future projections in these bioregions. Within mainland Australia, the temperate zone forms two anciently isolated and contrasting bioregions at similar latitudes within the southwest and southeast corner (Figure 1.1). These disparate temperate regions are separated by a vast, arid expanse known as the Nullarbor Plain, which originated ~15 million years ago following major marine transgressions and a tectonic uplift event (Benbow 1990). The lack of connected drainages across the plain is a major barrier to the connectivity of water-dependent species (Byrne *et al.* 2011; Cook *et al.* 2015; Crisp & Cook 2007; Rix *et al.* 2015), and the Nullarbor Plain's formation represents a key vicariance event in Australia's environmental history. Despite their broadly similar climatic profiles, the two temperate regions demonstrate contrasting environmental histories, making them ideal for testing various hypotheses about the role of environmental history on the evolution, diversification, and persistence of temperate species.

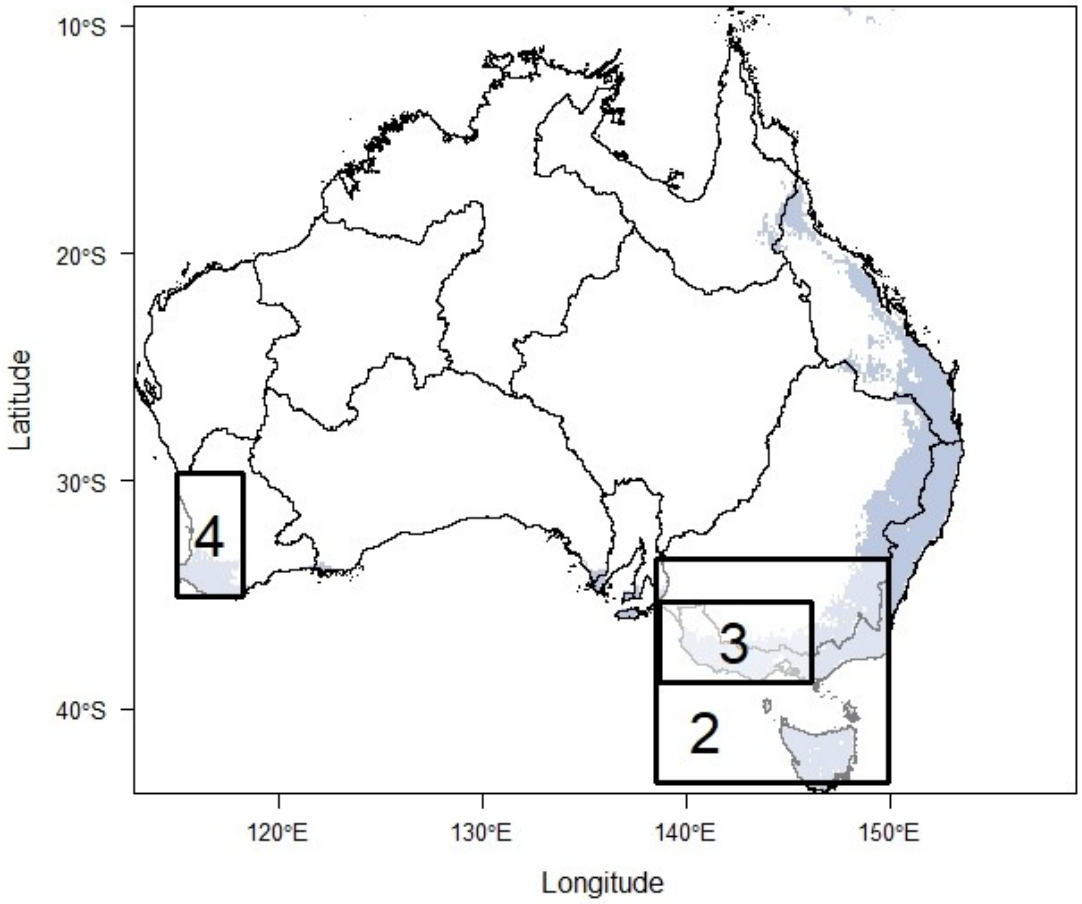


Figure 1.1: The Australian temperate ecosystems, and regions studied in this thesis. Solid black lines indicate major drainage basin divides. Temperate regions are defined by the shaded area, derived from a recent global re-evaluation of Köppen-Geiger climatic profiles at 1km resolution (Beck et al. 2018). Boxes indicate study regions for each chapter (numbered by chapter) of this thesis.

The Variable Southeast

The southeast temperate zone spans from the centre of Australia's southern coast towards the northeast, eventually developing into more tropical habitat as it shifts northward. The region is spatially variable, with heterogeneous topography (Figure 1.2A), multiple major drainage basins (Dickinson *et al.* 2002) and a strong hydroclimatic gradient (Figure 1.2B; Brauer *et al.* 2016). Much of the region is dominated by sclerophyllous woodlands and open vegetation, progressing to shrubland and desert further inland (Byrne *et al.* 2008; Chapple *et al.* 2005). The

region also shows strong temporal variability, and has significantly changed since the Pliocene including intense aridification in inland regions (Byrne *et al.* 2011), sea level changes with associated cyclic island connections (Porter-Smith *et al.* 2012), and tectonic shifts (Unmack 2001). Together, these events have driven large-scale alterations to the ecosystem, including the formation and demise of a 90,000 km^2 lake (McLaren *et al.* 2011) and cyclic emergence and submergence of similarly-sized land bridge (Blom & Alsop 1988; Porter-Smith *et al.* 2012). Additionally, warmer and wetter conditions in the Miocene sustained historical subtropical rainforest along the continental coastline (Byrne *et al.* 2011; Byrne *et al.* 2008), highlighting the transformative effect of climatic changes on the ecosystem. These significant alterations have made the region a prime study area for investigating the impact of past environmental changes on the phylogeography of terrestrial (e.g. Ansari *et al.* 2019; Cooper *et al.* 2000; Dubey & Shine 2010; Joseph *et al.* 2008; Kawakami *et al.* 2009; Kreger *et al.* 2019; Neal *et al.* 2019; Norgate *et al.* 2009), freshwater (e.g. Adams *et al.* 2014; Coleman *et al.* 2010; Hammer *et al.* 2019; Hodges *et al.* 2015; Murphy & Austin 2004; Schultz *et al.* 2008; Unmack *et al.* 2012; Unmack *et al.* 2013; Waters *et al.* 2019) and marine (e.g. Colgan 2016; Overeem *et al.* 2008; Shaddick *et al.* 2011; Waters 2008; York *et al.* 2008) organisms. Thus, the temperate southeast of Australia remains a suitable location to investigate the relative role of various past environmental changes on phylogeographic patterns.

The Stable Southwest

The characteristics in the southeast contrast with the temperate southwest of Australia, which has experienced significantly less environmental variation over similar time scales. The region features a relatively simplistic landscape with limited

topographic variation (Figure 1.2C; Funnekotter *et al.* 2019; Wheeler & Byrne 2006), no major river drainage divides, stable geology (Hopper & Gioia 2004) and little climatic variation since the Pliocene (Spooner *et al.* 2011). A simple bioclimatic gradient defines the region (Figure 1.2D), with increasing aridification and temperatures further inland, with the temperate region primarily constricted to areas with >300mm annual rainfall (Hopper & Gioia 2004). Despite the lack of topographic or environmental variation that might typically support high biodiversity through narrow-range endemics (Cowling & Lombard 2002; Hickerson *et al.* 2010), the region is one of only two global biodiversity hotspots declared for Australia, commonly referred to as the Southwest Western Australia (SWWA) hotspot (Hopper & Gioia 2004; Myers *et al.* 2000). This status is based on the region's high floristic diversity and endemism, with >8,000 species of plants of which approximately half are endemic (Gioia & Hopper 2017), and extensive habitat loss, with ~70% of native land vegetation cleared primarily for agricultural purposes (Habel *et al.* 2019; Monks *et al.* 2019). These aspects make the SWWA a region of high conservation concern and targeted by studies that investigated the underlying evolutionary mechanisms and the future of biodiversity within the hotspot (reviewed in Byrne *et al.* 2011; Crisp & Cook 2007; Hopper & Gioia 2004; Rix *et al.* 2015).

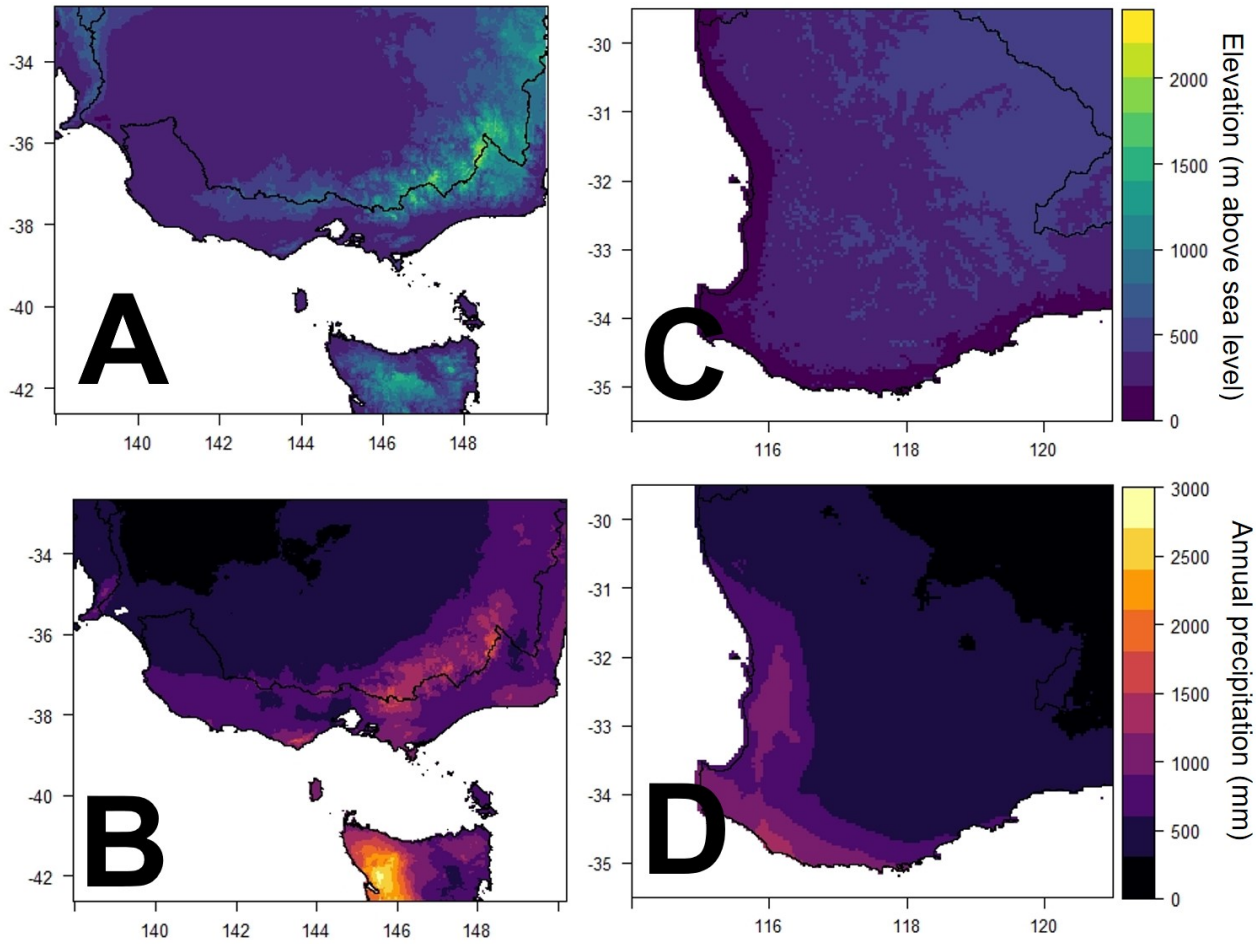


Figure 1.2: Comparison of spatial environmental heterogeneity between southeast (**A – B**) and southwest (**C – D**) Australia. **A, C:** Elevation, derived from the Etopo1 Global Relief Model (Amante & Eakins 2009). **B, D:** Annual precipitation, derived from the WorldClim database (Hijmans *et al.* 2005). Colours and break points are shared across both bioregions, with higher maximum values in southeast Australia.

Freshwater Species as Phylogeographic Models

Investigating the effect of environmental changes on evolutionary history and diversification requires a suitable model taxon. To this end, freshwater-dependent species are important indicators of historical environmental changes (Beheregaray *et al.* 2002; Burridge *et al.* 2007; Murphy & Austin 2004) given their reliance on limited habitat and often reduced capacity for dispersal (Davis *et al.* 2018; Inoue *et al.* 2014;

Thomaz *et al.* 2017). Even minor alterations to hydrology or topology can have strong impacts on the evolution of freshwater species (Inoue *et al.* 2014; Thomaz *et al.* 2017; Wallis *et al.* 2017): for example, riverine rearrangements can facilitate dispersal and shifting distributional patterns of aquatic species (Murphy & Austin 2004; Waters *et al.* 2020). Additionally, freshwater species show a high propensity for cryptic speciation (Avise 2000; Gouws *et al.* 2006; Murphy & Austin 2004), providing the ability to investigate the interaction of Earth history, biotic traits, and diversification more thoroughly (Burridge & Waters 2020). These same characteristics also make freshwater species highly susceptible to extinction under contemporary climate change, causing them to be some of the most globally threatened taxa (Collen *et al.* 2014; Darwall *et al.* 2011). Understanding how past environmental changes have impacted the evolution of freshwater species thus provides a template for predicting how ongoing and future climate change might similarly impact their persistence.

Pygmy Perches

The pygmy perches (Percichthyidae), consisting of the genus *Nannoperca* (six named species) and *Nannatherina balstoni*, embody these traits. These small obligate freshwater fishes (max size ≈ 10cm) are habitat specialists (Wedderburn *et al.* 2012), preferring ephemeral and slow-moving streams or floodplains over major river systems (Hammer *et al.* 2013) and showing limited dispersal capacity (Saddlier *et al.* 2013). Pygmy perches are relatively short-lived (maximum age of 3 – 6 years) and reach sexual maturity after a single year (Attard *et al.* 2016a; Humphries 1995). Their ancient origins and poor dispersal ability suggest that their long-term persistence has been influenced by local adaptation to environmental instability, a

pattern supported by hydroclimatic adaptation in southern pygmy perch (Brauer *et al.* 2016; Brauer *et al.* 2017; Morrongiello *et al.* 2010; Morrongiello *et al.* 2012).

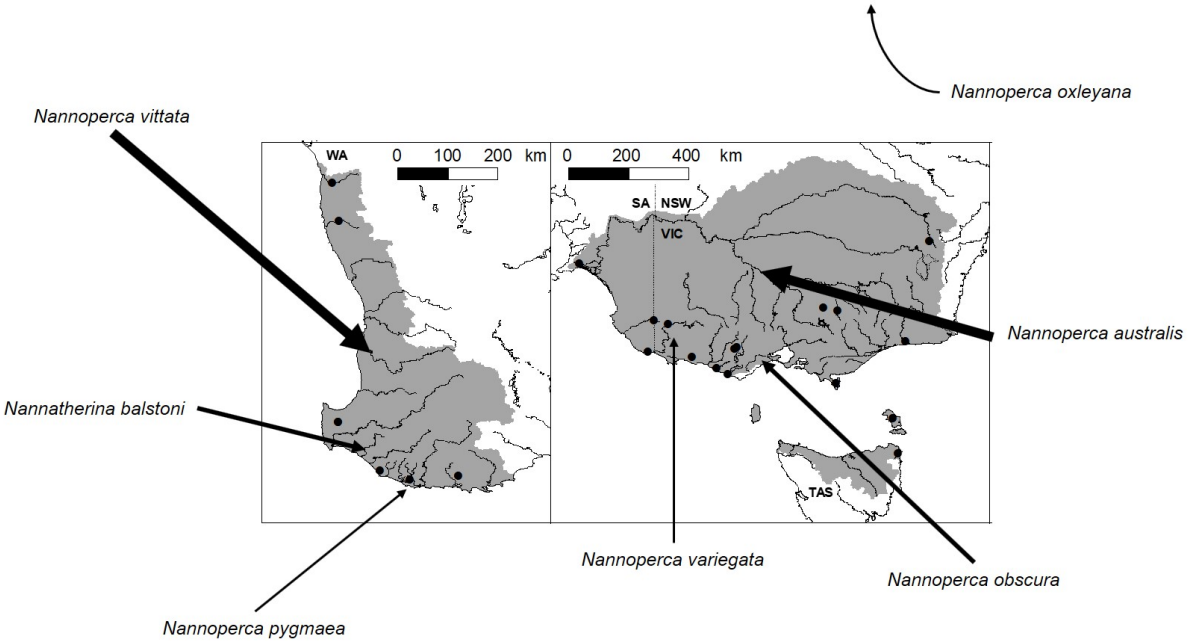


Figure 1.3: Distribution of pygmy perch species in southeast and southwest Australia, adapted from Buckley *et al.* (2018). Shaded regions denote the combined pygmy perch distribution for the region, excluding *N. oxleyana* which occupies a small distribution further north. The weight of the arrows between each species and the map indicates the relative extent of their distribution, with thicker arrows indicating more widespread species. Images of each species have been removed due to copyright restrictions. Distribution map reproduced with permission for open access articles by Royal Society Publishing.

As a group, pygmy perch are endemic to both temperate bioregions (Figure 1.3), with multiple divergent lineages within each region originating as early as the Miocene – Oligocene boundary (Figure 1.4). Despite the contemporary isolation of southwest and southeast temperate Australia, pygmy perches demonstrate a complex history involving multiple migrations across continental Australia (Buckley *et al.* 2018). Although species-specific phylogeographic patterns have been studied for some species, these have either used previous generation markers (Hammer *et al.*

2010; Unmack *et al.* 2013) or have not been representative of the full species distribution (Cole *et al.* 2016). These studies have identified several divergent intraspecific lineages within their respective study species, including evolutionarily significant units (ESUs), reflecting historical structure within species.

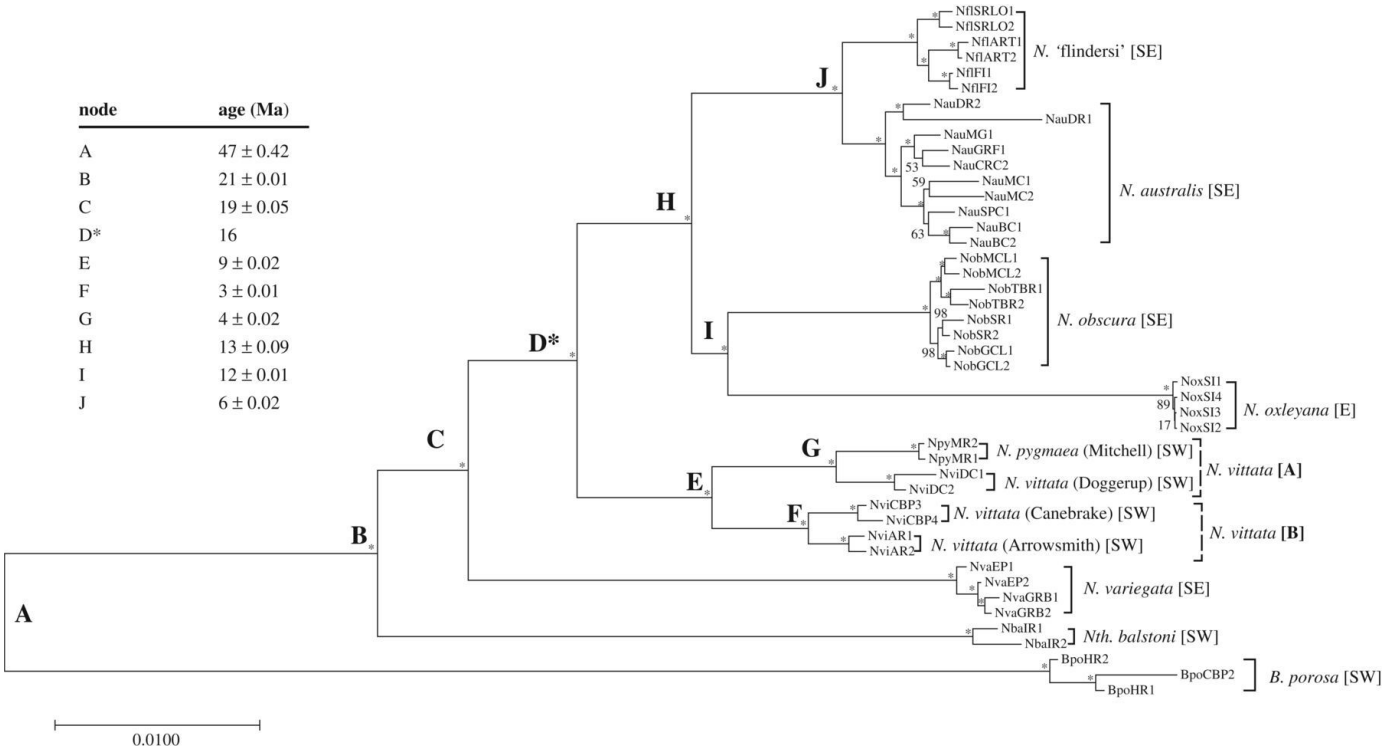


Figure 1.4: Phylogenomic tree of all described pygmy perch species based on 13,991 ddRAD loci, adapted from Buckley *et al.* 2018. Nodes are labelled with bootstrap support (asterisks represent 100% bootstrap support) and divergence times (indicated by letters and insert table) based on a biogeographic calibration point (D*****). Clades are labelled by species (or 'superclades' for N. vittata [A] and [B]) and their relative distribution within Australia. Reproduced with permission for open access articles by Royal Society Publishing.

The taxonomy of the pygmy perch lineage is not fully resolved: in 2013, a morphologically and genetically divergent species (*Nannoperca* pygmaea) was described for southwest Australia (Morgan *et al.* 2013). Additionally, several cryptic species have recently been identified within the group, with the latest suggestion of

at least ten species of pygmy perch based on the delineation of *N.* 'flindersi' from *N. australis,* and a complex of three unnamed cryptic species within *N. vittata* (Buckley *et al.* 2018; Unmack *et al.* 2011). However, these studies used geographically sparse data within lineages, may not have detected additional cryptic species and could not accurately infer geographic boundaries between putative species.

All species of pygmy perch are of conservation concern and listed in state and national legislation as well as within the International Union for Conservation of Nature's (IUCN) Red List (Table 1.1). Common threats across species include predation by invasive fish species (Beatty & Morgan 2013; Saddlier *et al.* 2013), a major drought event at the start of the millennium (Hammer *et al.* 2013; Wedderburn *et al.* 2012) and habitat fragmentation from anthropogenic modification of hydrological systems (Allen *et al.* 2020; Brauer *et al.* 2016; Cole *et al.* 2016). Several pygmy perch species show patterns indicative of a significantly more widespread and connected distribution prior to European settlement in Australia ~200 years ago (Cole *et al.* 2016; Hammer *et al.* 2010).

In response to these growing threats, several ongoing conservation programs have been enacted to protect pygmy perches. Many of these have focused on southeastern species such as *Nannoperca australis* and *N. obscura*, which have been the subject of long-term monitoring (Hammer *et al.* 2013; Wedderburn *et al.* 2012) and extensive genetic-based captive breeding programs (Attard *et al.* 2016a; Attard *et al.* 2016b; Beheregaray *et al.* in press). Reintroductions following multiple generations of captive breeding have had variable success, with some indication of

recovery in *N. australis* but not in *N. obscura*, which has recently gone locally extinct within part of its range (Beheregaray *et al.* in press). While less attention has been given to southwestern species, ongoing monitoring and conservation efforts for *N. pygmaea* aim to protect this narrowly distributed species (Allen *et al.* 2020). Phylogeographic studies of pygmy perch species can offer critical information for their conservation management by providing a historical reference regarding population connectivity, climatic responses, and accurate delineations of species.

Table 1.1: International, national and state conservation status of all described pygmy perch species. All statuses are as listed in the relevant Act or documentation as of 13/08/2020. National status is based on species classification under the Environmental Protection and Biodiversity Conservation Act, 1999. States: SA = South Australia; VIC = Victoria; NSW = New South Wales; QLD = Queensland; WA = Western Australia. Status: UL = Unlisted; NT = Near Threatened; VU = Vulnerable; TR = Threatened; EN = Endangered; CR = Critically Endangered. Relevant state legislation is noted by symbol and footnote.

Genus	Species	Common name	State	State status	National status	IUCN Red List
Nannatherina	*balstoni*	Balston's pygmy perch	WA*	VU	VU	EN
Nannoperca	*australis*	Southern pygmy perch	SA†	EN	UL	NT
			VIC‡	VU		
			NSW§	EN		
	obscura	Yarra pygmy perch	SA†	CR	VU	EN
			VIC‡	TR		
	oxleyana	Oxleyanan pygmy perch	NSW§	EN	EN	EN
			QLD#	VU		
	vittata	Western pygmy perch	WA*	UL	UL	VU
	pygmaea	Little pygmy perch	WA*	EN	EN	EN
	variegata	Variegated pygmy perch	SA†	CR	VU	EN
			VIC‡	TR		

*Biodiversity Conservation Act, 2016
†Action Plan for South Australian Freshwater Fishes, 2009; Fisheries Management Act, 2007
‡Flora and Fauna Guarantee Act, 1988
§Fisheries Management Act, 1994
#Nature Conservation Act, 1992; Fisheries Act, 1994

Significance and Justification

Predicting how species may respond to anthropogenic climate change is complicated by spatial heterogeneity in resultant environmental change and its interaction with species-specific traits. For threatened and dispersal-limited taxa, such as freshwater fishes, this poses a significant challenge for determining risk in the near future and appropriate conservation management. Using phylogeographic analyses to infer historical species responses to past environmental changes provides an empirical basis for determining potential responses to contemporary climate change. The development of genomic technologies and complex coalescent methods has positioned the field of phylogeography to investigate nuance in these interactions.

This thesis aims to evaluate the role of biogeographic history on species evolution in a comparative framework that spans ecologically similar species and disparate biogeographic regions. For Australian freshwater fishes, it entails some of the first phylogeographic research using genomics, and arguably the first to use genomics within a comparative phylogeography framework. Here, phylogeographic histories of pygmy perch species in southeast and southwest Australia are explored in detail using a complementary suite of genomic, modelling and environmental approaches. Inferred patterns are contrasted between disparate bioregions and co-distributed species to highlight the complex nature of evolutionary responses to past environmental change. The findings of this thesis represent an original contribution to biogeography and evolutionary theory, and inform on direct practical suggestions for the ongoing conservation of threatened freshwater fishes and potentially other similar taxa.

This thesis ties into several wider programs of biogeographic and conservation research within the Molecular Ecology Lab at Flinders University. Specifically, this thesis has connections to an Australian Research Council Linkage project (LP100200409: “Restoration genetics of five endangered fish species from the Murray-Darling Basin”) and a Future Fellowship project (FT130101068: “Evolution, adaptation and resilience of Australian freshwater fishes”). Although these projects have reached their funding completion, this thesis extends their legacy and utilises some of the data previously collected as part of these projects.

Within this thesis, genomic sequences for 90 individuals were generated across all the study species using tissue samples acquired from the South Australian Museum. These samples were collected over the last decade by collaborators as part of other related projects, and there is an established database which connects all sequences generated herein with specimens lodged in the South Australian Museum. Accession numbers for these samples are provided in the relevant section of each data chapter.

A portion of the work within this thesis builds upon my prior Honours research on the phylogenomic and biogeographic patterns of all pygmy perches (Buckley *et al.* 2018). Some aspects of that research – e.g. additional ancestral area estimation, phylogenetic analyses, drafting of the manuscript, submission and revisions – were done within the duration of this PhD candidature. This paper has been provided in Appendix 1 for further reference.

Thesis Outline

This thesis contains a general introduction, three data chapters and a concluding chapter. This first chapter provides a brief introduction to the relevant topics of the thesis, including general information about the study regions and species as well as the broader significance of the work. Each of the three data chapters are intended as stand-alone manuscripts for publication and contain detailed background relevant to the chapter within their own introduction sections. The final concluding chapter synthesises the results across the data chapters to broadly summarise the work and outline the ongoing and future research arising from the thesis.

Chapter 2: *The roles of aridification and sea level changes in the diversification and persistence of freshwater fish lineages*[3].

A combination of phylogenetic and coalescent approaches with environmental modelling are used to determine the relative role of various Earth history events on the diversification of *N. australis* and *N.* 'flindersi' across their full distributions and spanning back to the Pliocene. Strong hierarchical phylogenetic structure was observed, demonstrating divergences spanning species, intraspecific lineages, and historical population structure. For inland populations, these divergences were largely driven by ancient aridification events, whilst phylogeographic patterns for coastal population were mostly influenced by sea level changes during the Pleistocene.

[3] Chapter publication status: submitted to *Molecular Ecology* (MEC-20-0685), with major changes required after initial round of revision. Revised manuscript to be sent back to *Molecular Ecology* in late September.

Chapter 3: *Variation in intraspecific demographic history drives localised concordance but species-wide discordance in responses to Plio-Pleistocene climatic change*[4].

A suite of population genetics, historical demographic modelling and environmental approaches are used to compare the phylogeographic histories of *N. australis* and *N. obscura* across their co-distributed range. Despite showing concordant demographic histories within a shared long-term climatic refugium, the two species show highly discordant changes in distribution, with *N. australis* demonstrating a large range expansion during glacial maxima. The lack of expansion in *N. obscura* is likely attributed to a lack of standing genetic variation, which might underlie the variable success in their conservation management.

Chapter 4: *Long-term climatic stability drives accumulation of divergent lineages in a temperate biodiversity hotspot*[5].

Phylogenetic analyses, species delimitation, species distribution modelling and a novel enrichment method of fixed differences are used to investigate the phylogeographic history of *N. vittata* and *N. pygmaea* within the SWWA biodiversity hotspot. Lineages showed ancient divergences dating to the Miocene/Pliocene, including the formation of two cryptic species within *N. vittata* which show evidence for reproductive isolation through enrichment of genes putatively linked to

[4] Chapter publication status: this work will be submitted to *Proceedings of the Royal Society B: Biological Sciences*.
[5] Chapter publication status: this manuscript will be prepared for submission to a journal later this year, possibly *Global Ecology and Biogeography.*

reinforcement. Long-term climatic stability is reflected in stable species distributions over time, allowing divergent lineages to accumulate within a biodiversity hotspot.

Chapter 5: *Phylogeographic history and evolutionary diversification of pygmy perches* (Teleostei: Percichthyidae)

This concluding chapter briefly summarises the major findings of the previous data chapters to provide an overview of biogeographic history across temperate Australia. Some comparisons between the two disparate bioregions, as well as across co-distributed taxa, are made to infer broader inferences of the results. The implications of this work are summarised within the concluding remarks, both in terms of understanding evolutionary responses to environmental change as well as the inferences for the conservation management of pygmy perches. This chapter is only intended to succinctly summarise the work of the thesis, and does not contain a lengthy discussion of the literature.

Chapter 2: The roles of aridification and sea level changes in the diversification of freshwater fish lineages

Sean J Buckley[1], Chris Brauer[1], Peter Unmack[2], Michael Hammer[3], Luciano B. Beheregaray[1*]

[1]Molecular Ecology Laboratory, College of Science and Engineering, Flinders University, Adelaide, SA 5001, Australia

[2]Institute for Applied Ecology, University of Canberra, ACT 2601, Australia

[3]Natural Sciences, Museum and Art Gallery of the Northern Territory, Darwin, NT 0801, Australia

This chapter has been submitted to Molecular Ecology and is reproduced with permission. I am the primary author, with Dr. Chris Brauer, Dr. Peter Unmack, Dr. Michael Hammer and Prof. Luciano Beheregaray as co-authors. I was responsible for approximately half the data collection, all analyses and drafting the manuscript. Chris Brauer provided some sequences from previous work. All co-authors provided comments on the manuscript. Luciano Beheregaray additionally supervised the project.

Abstract

While the influence of Pleistocene climatic changes on divergence and speciation has been well-documented across the globe, complex spatial interactions between hydrology and eustatics over longer timeframes may also determine species evolutionary trajectories. Across mainland Australia, glacial cycles were not associated with changes in ice cover and instead largely resulted in fluctuations from moist to arid conditions across the landscape. Here, we investigate the role of hydrological and coastal topographic changes brought about by Plio-Pleistocene climatic changes on the biogeographic history of a small Australian freshwater fish, the southern pygmy perch *Nannoperca australis*. Using 7,958 ddRAD-seq loci and 45,104 filtered SNPs, we combined phylogenetic, coalescent and species distribution analyses to investigate the various roles of aridification, sea level and tectonics and associated biogeographic changes across southeast Australia. A hierarchy of divergence events were apparent across the clade, ranging from the initial formation and maintenance of a cryptic species (*N.* 'flindersi') in the Miocene to population fragmentation during the Holocene. Environmental factors driving divergences varied spatially, with sea-level changes since the Pliocene impacting coastal populations and the reduction or disappearance of large waterbodies throughout the Pleistocene driving divergence in inland populations. Isolated climatic refugia and fragmentation due to lack of connected waterways maintained the identity and divergence of inter- and intraspecific lineages. Our historical conclusions suggest that the impacts of environmental change across species distributions are spatially heterogeneous, a findings with implications for large-scale predictions of species responses to climate change. Conservation approaches should consider this spatial variation when

developing management plans, especially regarding how aridification and eustatic changes may act heterogeneously across species distributions.

Introduction

Dramatic changes in climate, hydrology and topography have long been recognised to have lasting impacts on the diversity, distribution and divergence of species and populations (Pelletier *et al.* 2015). Understanding the relationship between the historical environment and the genealogy of species remains critical for interpreting how contemporary climate change may impact on species currently and in the near future. Most notably, increasing aridification and rising sea-levels predicted by climate change projections call into question the adaptive capacity and resilience of organisms, especially those with poor dispersal potential and narrow ranges (Davis *et al.* 2013; Falkenmark 2013; Grummer *et al.* 2019). However, applying broad-scale inferences about environmental changes to understand biodiversity resilience in the future is further complicated by spatial variation in environmental factors that might impact on how within-species responses vary across their ranges (Razgour *et al.* 2019). For example, historical eustatic changes likely had a larger influence on coastal or marine ecosystems while aridification played a stronger role further inland (Beheregaray *et al.* 2002; Pinceel *et al.* 2013). Thus, understanding the relative role of different environmental changes between regions is important in more accurately predicting species' responses.

Complex impacts of climatic change are particularly exacerbated in freshwater ecosystems, as increasing temperature and aridity alters the stability and structure of hydrological systems (Blöschl *et al.* 2019; Middelkoop *et al.* 2001; Nijssen *et al.* 2001; Pinceel *et al.* 2013). With limited dispersal capability and reliance on available freshwater for survival, aquatic-dependent species often demonstrate strong evolutionary associations with hydrological changes. Even minor alterations to

hydrologic structure can have profound impacts on the evolution of a diverse array of freshwater taxa (Inoue *et al.* 2014; Thomaz *et al.* 2017; Wallis *et al.* 2017). For example, tectonic activity can reshape waterways, leading to river capture across new areas and shifting distributional patterns of water-dependent species (Murphy & Austin 2004; Waters *et al.* 2001). Thus, freshwater biodiversity functions as an important indicator of the impact of historical environmental changes.

Inferences of phylogeographic responses to past environmental change relies upon a combination of genetic, spatial and modelling approaches. Determining the relative role of past climatic events is difficult when resolution is low due to few genetic markers or limited model capability (Carstens *et al.* 2012; Cutter 2013; Nakhleh 2013). To this end, the collection of thousands of genetic markers enables better capture of the diverse array of demographic processes influenced by Earth history (Carstens *et al.* 2012; Edwards *et al.* 2016). In tandem, recent advancements in coalescent modelling, informed by detailed information of geological and ecological history, have improved the ability to provide more nuanced inferences (Cutter 2013; Excoffier *et al.* 2013). This combination of greater data and sophisticated modelling techniques provides the analytical framework to address questions about the spatial variance of species responses to climate change.

A suitable biogeographic setting to test hypotheses of spatial and temporal variation of Earth history on evolution is one including both inland and coastal regions. In this regard, the temperate southeast of Australia is well-suited given it has been influenced by aridification across the continent, as well as by shifts in landmass attributed to eustatic changes (Chapple *et al.* 2011; Faulks *et al.* 2010; McLaren &

Wallace 2010). This region is characterised by complex geography and geology, affected by a history of uplift, subsidence and volcanism (Unmack 2001). The region is subdivided by the Great Dividing Range, which runs parallel to the coastline from northern Australia to the southern coast. This range acts as a barrier that separates the inland Murray-Darling Basin (MDB) from coastal areas and is a key biogeographic feature of the region (Figure 2.1b; Chapple *et al.* 2011; Unmack 2001).

Within the MDB, major hydrological changes in the past were mostly associated with aridification and tectonics, such as the formation and decline of the paleo megalake Bungunnia during the Plio-Pleistocene (from 3 Mya until ~700 Kya), which spanned 90,000 km^2 across the lower section of the MDB at its largest size (McLaren *et al.* 2011; McLaren *et al.* 2012). Lake Bungunnia has been suggested to have acted as a barrier for some terrestrial species (Ansari *et al.* 2019; Cooper *et al.* 2000; Joseph *et al.* 2008; Kawakami *et al.* 2009; Neal *et al.* 2019), and inversely may have acted as a conduit for aquatic species (Waters *et al.* 2019). Contrastingly, environmental changes across the coastal habitats south of the MDB were more associated with eustatic changes. Particularly, the formation and submergence of the Bassian Isthmus, which connected the island of Tasmania to the mainland during glacial maxima (Blom & Alsop 1988; Porter-Smith *et al.* 2012), is a well-documented driver of biogeographic patterns for a variety of terrestrial, marine and freshwater taxa (Schultz *et al.* 2008; Waters 2008). This combination of relevant Earth history factors across southeast Australia, and their potentially interactive nature, provides a scenario to investigate the relative role of different past environmental changes on phylogeographic patterns.

An ideal system for studying biogeographic changes in southeast Australia is the southern pygmy perch, *Nannoperca australis* (Percichthyidae). This small-bodied (<80mm), habitat-specialist fish prefers slow flowing and vegetated ephemeral streams (Hammer *et al.* 2013; Wedderburn *et al.* 2012). It is distributed throughout the temperate southeast Australia region, occupying the MDB, coastal Victoria and northern Tasmanian rivers. Previous phylogenetic work indicated that southern pygmy perch from eastern Victoria, Flinders Island and north-eastern Tasmania belong to a genetically distinct cryptic species referred to as *Nannoperca* 'flindersi' (Buckley *et al.* 2018; Unmack *et al.* 2013). Estimates of divergence time using a biogeographic calibration point suggest this split occurred ~6 Mya (Figure 1.4), but the biogeographic forces driving this speciation remain unknown. Being an ancient lineage, *N. australis* has likely responded to a variety of environmental changes across inland and coastal habitats since the Miocene. For instance, landscape genomics (Brauer *et al.* 2016) and ecological transcriptomic studies (Brauer *et al.* 2017) have shown that the strong contemporary east-west hydroclimatic gradient of the MDB has led to the evolution of several adaptively divergent populations in *N. australis*. Furthermore, *N. australis* is threatened, particularly within the MDB, due to extreme pressure from anthropogenic changes to water flow, introduced predators and contemporary climate change (Balcombe *et al.* 2011; Brauer *et al.* 2016). The low dispersal capability, small population sizes and high genetic structure of the species makes their survivability of great concern (Brauer *et al.* 2016; Brauer *et al.* 2017; Cole *et al.* 2016). In fact, the species has been the target of multiple genetic-based captive-breeding and reintroduction programs in the MDB (Attard *et al.* 2016a; Beheregaray *et al.* in press). Here we used genome-wide data to assess the relative

roles of hydrological and coastal topographic changes as drivers of evolutionary diversification and lineage persistence. We hypothesised that demographic changes and lineage diversification linked to aridification would be older (Miocene – Pliocene) and stronger for populations from inland basins, whereas changes linked to eustatic variation would be comparatively younger (Pleistocene) and common for populations from coastal or island habitats. We tested the impact of these factors using a hierarchical framework that incorporates complex, hypothesis-driven coalescent modelling, model-free demographic analyses and spatial (species distribution) modelling.

Materials & Methods

Sample Collection and Genomic Library Preparation

A total of 109 samples across 21 known populations of *N. australis* and three populations of *N.* 'flindersi' (*n* = 4 – 5 individuals per population) were used (Table 2.1). This sample spans the full geographic range of the species and includes at least one population from each management unit identified in previous population genetic and genomic studies (Figure 2.1a; Cole *et al.* 2016; Unmack *et al.* 2013). The sister species *N. obscura* (Figure 1.4) was included as the outgroup for phylogenetic analyses (*n* = 5). Specimens were collected using a combination of electrofishing, dip-, fyke- or seine-netting. Specimens (either caudal fin or entire specimen) were stored dry at −80°C at the South Australian Museum, or in 99% ethanol at Flinders University.

Table 2.1: Locality data for samples used in this study. Abbreviations described in the table were those used for further analyses, while n refers to the number of individuals sequenced per population. *Nannoperca obscura* samples were only included as an outgroup in the phylogenetic analysis.

Species	Population	Abbreviation	Field code	*n*
N. australis	Angas R., Strathalbyn	NauANG	F-FISH84	5
	Lake Alexandrina	NauALE	SPPBrA*	4
	Middle Ck, Warrenmang, Avoca	NauAVO	F-FISH75: PU99-33SPP	5
	Jew Harp Ck, Sidonia	NauJHA	F-FISH78: PU00-01SPP	3
	Tributary to Seven Creeks	NauSEV	PU13-65SPP	4
	Merton Ck, Goulburn Rvr.	NauMER	F-FISHY6: PU09-01SPP	5
	Broken R., Lima South	NauBRO	F-FISHY6: PU09-02SPP	5
	King R., Cheshunt, Ovens Rvr.	NauKIN	F-FISHY6: PU09-06SPP	4
	Spring Ck, Mitta Mitta	NauSPR	F-FISHY6: PU09-13SPP	5
	Gap Ck, Kergunyah, Kiewa	NauGAP	F-FISHY6: PU09-12SPP F-FISH77: PU99-81SPP	5
	Murray R. lagoon, Albury	NauALB	F-FISH53: IW94-47	4
	Coppabella Ck, Coppabella	NauCOP	F-FISH75: PU99-82SPP	5
	Blakney Ck, Lachlan Rvr.	NauLRT	F-FISH98: LPP-*	5
	Glenelg R., Glenisla	NauGRG	F-FISH78: PU0014-SPP	5
	Merri R., Grassmere	NauMRG	F-FISH78: PU00-22SPP	5
	Curdies R., Curdie	NauCRC	F-FISH78: PU00-24SPP	4
	Gellibrand R. floodplain	NauGRF	F-FISH97: PU02-92SPP	5
	Barongarook Ck, Colac	NauBAR	SPP08-13	4
	Mundy Gully	NauMG	F-FISHY8: PU08-11SPP	4
	Gnarkeet Ck, Hamilton	NauGCH	F-FISHY2: PU00-27SPP	4
	Wilsons Promontory	NauWP	F-FISH97: PU02-70SPP	5
***N.* 'flindersi'**	Snowy R. lagoon, Orbost	NflSRLO	F-FISH77: PU99-85SPP	5
	Flinders Island	NflFI	F-FISH84: FI-*	4
	Anson R. tributary	NflANS	F-FISH82: HT-2*	5
N. obscura	Lake Alexandrina	*Outgroup*	YPBR*	5
Total		**24**		**109 (*114*)**

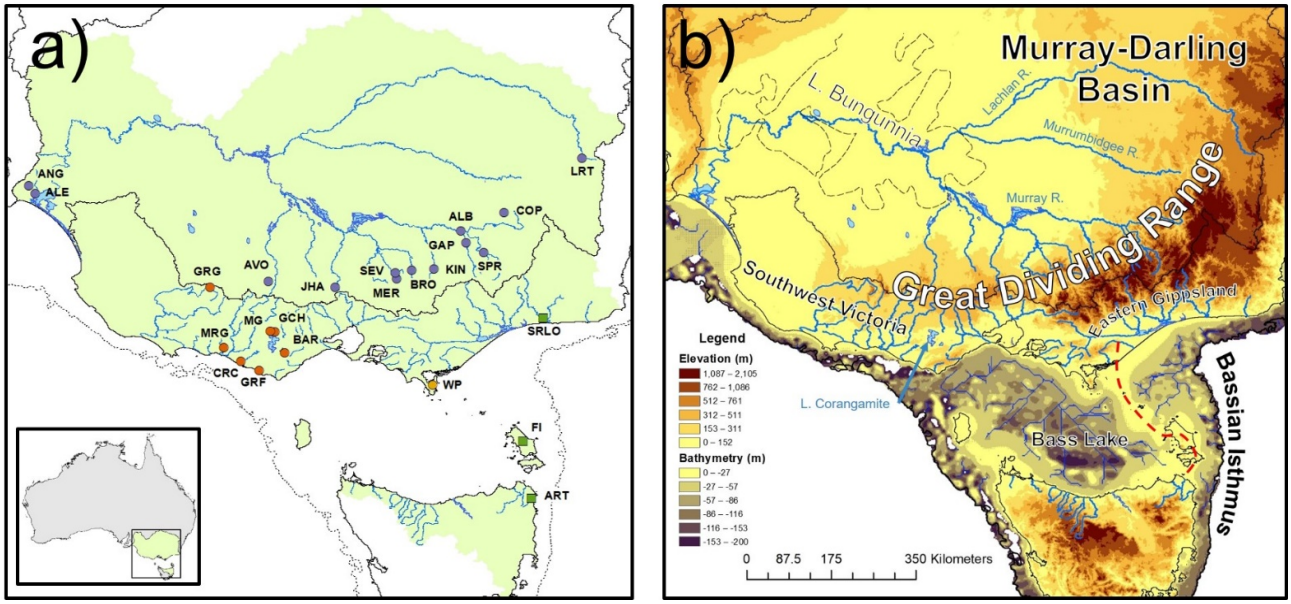


Figure 2.1: a) Distribution and sampling map for southern pygmy perch. Inset depicts extent of distribution within Australia. The shaded area denotes the putative distribution of the species, spanning multiple major basins (black lines). Locality abbreviations are detailed in Table 2.1. Colours denote major clades explored within coalescent models (refer to Results) whilst shapes denote 'species' (circles = *N. australis*; squares = *N.* 'flindersi'). The extent of the continental shelf (-121m), which was exposed during glacial periods, is indicated by the dotted line. **b)** Topographic map (including bathymetry) of southeast Australia, highlighting topographic heterogeneity and major biogeographic regions across the area. Solid black lines indicate major basin boundaries whilst the red dashed line indicates the drainage divide across the Bassian Isthmus. The maximum extent of Lake Bungunnia (at 1.2 Mya) is also indicated with a narrow dashed line.

DNA was extracted from muscle tissue or fin clips using a modified salting-out method (Sunnucks & Hales 1996) or a Qiagen DNeasy kit (Qiagen Inc., Valencia, CA, USA). Genomic DNA quality was assessed using a spectrophotometer (NanoDrop, Thermo Scientific), 2% agarose gels, and a fluorometer (Qubit, Life Technologies). All ddRAD genomic libraries were prepared in-house following (Peterson *et al.* 2012), with modifications as described in Brauer *et al.* 2016. Of the 109 samples, 73 were previously paired-end sequenced as part of a landscape

genomics study (Brauer *et al.* 2016) using an Illumina HiSeq 2000 at Genome Quebec (Montreal, Canada). The additional 36 samples were single-end sequenced on an Illumina HiSeq 2500 at the South Australia Health and Medical Research Institute.

Bioinformatics

The resultant reads (forward reads only for paired-end samples) were filtered and demultiplexed using the 'process_radtags' module of Stacks 1.29 (Catchen *et al.* 2013), allowing ≤ 2 mismatches in barcodes. Barcodes were removed and reads trimmed to 80 bp to remove low- quality bases from the ends. Cut reads were then aligned using PyRAD 3.0.6 (Eaton 2014), and further filtered by removing reads that had > 5 bp with a Phred score of < 20. Loci were retained at a minimum sequencing depth of 5 and occurring in at least ~90% of samples (103). The final alignment contained 7,958 ddRAD loci and 45,104 SNPs. Although tree-based analyses predominantly used a concatenation of this alignment, coalescent-based analyses that formed the basis of more thorough investigations of demography and divergence were based upon unlinked SNPs (see Brauer *et al.* 2016 for details).

Phylogenetic Analysis

To determine evolutionary relationships as a basis for phylogeographic modelling, a maximum likelihood (ML) phylogeny was estimated using RAxML 8.2.11 (Stamatakis 2014) and the 7,958 concatenated ddRAD loci dataset. This was done using rapid hill-climbing and 1,000 resampling estimated log-likelihood (Pante *et al.* 2015) bootstraps under a GTR+Γ substitution model. Additionally, we estimated a ML tree in RAxML by partitioning the alignment by ddRAD locus, as well as inferring gene trees separately per ddRAD locus, using the same parameters. Concordance

between individual gene trees from the latter approach and the concatenated tree were summarized using site concordance factors (Minh *et al.* 2020a) in IQ-TREE2 (Minh *et al.* 2020b). These approaches were done to better account for the potential role of genome-wide rate heterogeneity and incomplete lineage sorting (Liu *et al.* 2015). The resultant phylogenetic trees were visualised using MEGA 7 (Kumar *et al.* 2016) and rooted with *N. obscura* as the outgroup.

To determine if dendritic river hierarchy alone could explain phylogenetic patterns across the MDB lineage, and to clarify coalescent models (see Results), linear correlations between genetic and riverine distance were estimated using StreamTree (Kalinowski *et al.* 2008). StreamTree models genetic divergence across a dendritic river system and assigns a cost to each riverine segment, comparing this modelled distance with the empirical data. While StreamTree is often used with pairwise F_{ST} values (e.g. Brauer *et al.* 2018) to assess contemporary spatial patterns, we used uncorrected genetic distances (*p*-distance) as this more likely contains signal of historic patterns of divergence (Nei 2001). Pairwise *p*-distances between individuals were estimated using PAUP* 4 (Swofford 2002) and averaged per population for all 13 MDB populations.

Divergence Time Estimation

We estimated divergence times using r8s 1.81 (Sanderson 2003). Given the lack of suitable fossils for pygmy perches, we calibrated the node between *N. australis* and *N.* 'flindersi' at 5.9 – 6.1 Mya based on a previous estimate that includes all pygmy perch species (Figure 1.4). We applied a broader range around the calibrated node to accommodate potential variation in dating not captured by the methods in

(Buckley *et al.* 2018). Divergence times for each node were estimated using a penalized-likelihood model under a truncated Newton algorithm (Nash 2000), which uses a parametric branch substitution rate model with a nonparametric roughness penalty (Sanderson 2003). Cross-validation was used to determine the best value of the smoothing parameter for the roughness penalty between $\log_{10}$ 0 and $\log_{10}$ 100. The optimum smoothing parameter of $\log_{10}$41 (1.613), with a chi-square error of -12836.285, was used to estimate divergence times between populations and higher order clades across the lineage.

Ancestral Range Estimation

We used a phylogenetic tree-based method to estimate ancestral ranges across the ML tree with the R package BioGeoBEARS (Matzke 2013). The ML tree was collapsed down to individual populations using the R package ape (Paradis *et al.* 2004). Given the paraphyletic nature of the Albury population (NauALB) within the phylogenetic tree (Figure 2.3), this population was pruned from the tree. The tree was then converted to ultrametric format using the divergence time estimates from r8s.

Tips were assigned to one of six main geographic regions based on current hydrogeology (McLaren *et al.* 2011) and biogeographic regions of interest. These were the MDB, coastal (western) Victoria, Wilson's Promontory, Snowy River (i.e. eastern coastal Victoria), Flinders Island and northern Tasmania. Individual *N.* 'flindersi' populations were assigned to unique geographic states given their current isolation and to allow for the explicit testing of vicariance vs. dispersal scenarios across the Bassian Isthmus. Ranges spanning multiple states were filtered to only

those composed of neighbouring ranges (total number of possible ranges = 21). Furthermore, given the historical marine inundation of the MDB which would have precluded the presence of southern pygmy perch, time-stratification was used to exclude 'MDB' as a geographic state prior to 5 Mya or 3 Mya. Although freshwater drainages may have existed further inland, the limited range and connectivity of these habitats would be unlikely to sustain southern pygmy perch populations over long-term. These times reflect a conservative estimate of marine inundation (the most recent time at which marine sediments have been accurately identified within paleolake Bungunnia; McLaren *et al.* 2011) and a more relaxed estimate that is possibly the most recent time marine water could have been present within the basin. Ancestral ranges were estimated under all six available models (DEC, DIVA-LIKE and BAYAREA-LIKE, with and without founder-event speciation, +J). All models were run under both time- stratification scenarios and compared using the Akaike Information Criterion (AIC) within each set.

Biogeographic Hypothesis Testing Using Coalescent Modelling

We expanded upon our tree-based approaches using a suite of population-level SNP-based analyses to more thoroughly investigate historical patterns of demography and divergence. These analyses provided a more robust analytical framework for testing specific hypotheses of the impact of Earth history on the evolution of southern pygmy perches. Specific hypotheses based on biogeographic events were tested using a coalescent framework within FastSimCoal 2.6 (Excoffier *et al.* 2013). These hypotheses expand on the broad interpretations based on phylogenetic analyses. These models were focused around particular divergence

events across the lineage, with hypotheses built around the separation of major clades within the phylogeny (Figure 2.2).

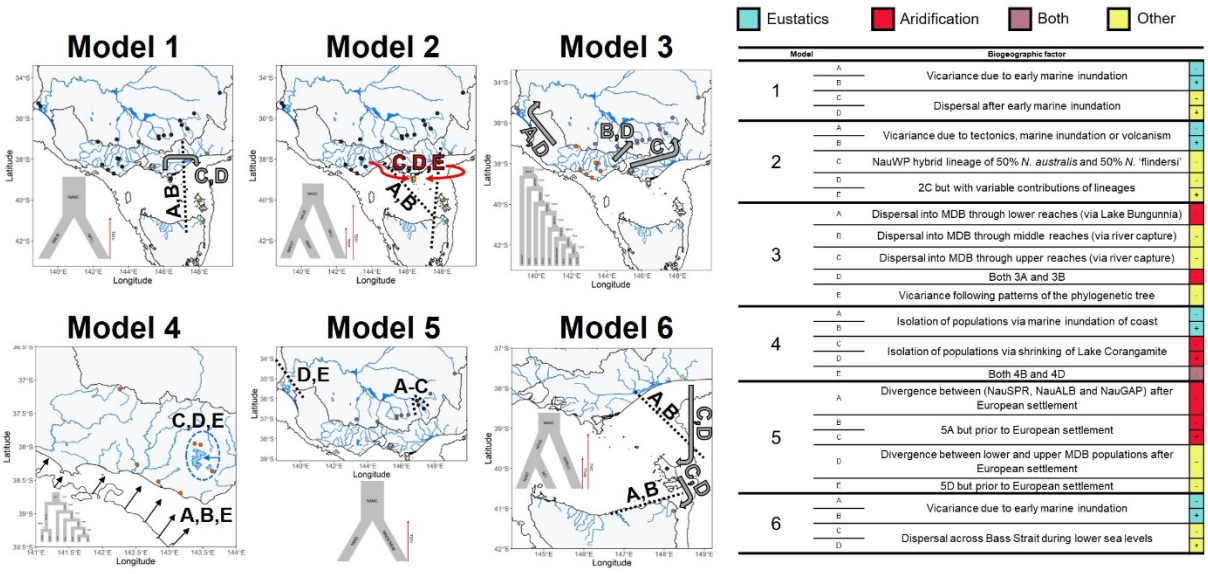


Figure 2.2: Representation of coalescent hypotheses. Left: Maps depicting the biogeographic scale and patterns of each set of models, with dashed lines indicating vicariance-based models (and their associated breaks), grey arrows indicating dispersal-based models, and red arrows indicating hybridisation-based models (with associated models labelled in letters). Insets depict a representative model (Model 'A') for each model set, which are described in greater detail in the Appendices. Right: table listing the specific hypotheses per model, and whether these relate to eustatic changes (blue), aridification (red), both eustatic changes and aridification (purple) or other biogeographic mechanisms such as tectonics (yellow). Models with post-divergence gene flow are indicated a "+" in the right column.

For each major divergence event, we estimated coalescent models considering various competing hypotheses that may have driven the divergence (Figure 2.2). These models primarily varied in the order of population coalescence (akin to topology), the timing of divergence events and other demographic parameters such as the presence of bottlenecks and bidirectional gene flow, based on the expected

impact of biogeographic patterns previously identified within the literature (Appendix 2) and information from both prior and current phylogenetic and coalescent analyses (Buckley *et al.* 2018). For models around species divergence, we tested whether vicariance or dispersal (through the absence or presence of a bottleneck) was associated with the initial divergence of *N.* 'flindersi' and whether post-isolation gene flow had occurred or not (Appendix 3). Although bottlenecks may occur under either vicariance or dispersal scenarios, the *absence* of a strong bottleneck is unlikely to occur under dispersal. For the divergence of the NauWP population, we tested various hypotheses of biogeographic causes of lineage isolation by varying the timing of divergence in association with known environmental changes (Appendix 4; Model 2A – B). We also tested whether this lineage originated as a hybrid of *N. australis* and *N.* 'flindersi' or as an isolated *N. australis* population (Appendix 4; Models 2C – E). For the separation of the MDB and coastal Victoria lineages, we tested colonization pathways from a coastal ancestor by altering the topology of the MDB populations in the model following the expected migratory pathways (Appendix 5; Models 3A – C), including a scenario of more than one colonization (Model 3D) and one following the inferred topology of the phylogenetic tree (Model 3E). Within the coastal lineage, we tested whether populations were isolated as a result of marine inundation during Pleistocene sea level rise, the shrinking of Lake Corangamite, or a combination of these factors (Appendix 6). Within the MDB lineage, we tested whether pre-European settlement (>200 years ago) isolation was evident between either putatively 'upper' and 'lower' populations, or between the most divergent set of populations (NauGAP, NauALB and NauSPR) and the rest of the MDB populations (Appendix 7). Within *N.* 'flindersi', we tested whether isolation of populations was associated with early Pleistocene marine inundation of the

Bassian Isthmus or as a result of isolated dispersal events, including testing for post-isolation gene flow (Appendix 8). Specific biogeographic hypotheses for each divergence event, and their predictive impacts on the evolution and demography of southern pygmy perch, are described in the Appendices.

Estimating Effective Population Size Changes

As a model-free and exploratory approach to clarify demographic history, changes in effective population size (*Ne*) over time were estimated using the site frequency spectrum (SFS) and coalescent modelling in a stairway plot (Liu & Fu 2015). Given the strong population genomic structure reported for southern pygmy perch (e.g. F_{ST} = 0 – 0.798; Brauer *et al.* 2016) including for populations sampled in this study, and to account for the biasing effect of population structure on the SFS (Stadler *et al.* 2009; Xue & Hickerson 2015), loci were re-aligned for each population independently. As missing data can bias the distribution of the SFS (Shafer *et al.* 2017), only loci present in all samples for each population were retained. Unlinked biallelic SNPs from each independent alignment were then used to generate the single-population SFS. Stairway plots were estimated assuming a mutation rate of 10^{-8} – 10^{-9} per site per generation based on genome-wide estimates of SNP mutation rates (Brumfield *et al.* 2003), which have been applied to other fishes (Stobie *et al.* 2018), and a generation time of one year based on the annual breeding of southern pygmy perches and the observation that most individuals do not live beyond one to two years in the wild (i.e. a single breeding season; Attard *et al.* 2016a; Humphries 1995).

Species and Lineage Distribution Modelling

The distribution of the species was modelled using MaxEnt 3.4 (Phillips *et al.* 2017) and 19 BioClim variables from WorldClim v1.4 (Hijmans *et al.* 2005), summarising precipitation and temperature – two groups of climatic variables known to impact on local adaptation and distribution of southern pygmy perches (Brauer *et al.* 2016). To account for non-climatic environmental aspects that may limit the distribution of the species (Paz *et al.* 2015), elevation (extracted from the Etopo1 combined bathymetry and topography dataset; (Amante & Eakins 2009) and topographic wetness index (extracted from the ENVIREM database; Title & Bemmels 2018) were also included. Species occurrence data was collected from the Atlas of Living Australia (http://www.ala.org.au), with low geographic resolution (>30km) occurrences and geographic outliers removed (6,106 occurrences) based on expert knowledge (P. Unmack, pers. comm). However, this dataset did not include the Murrumbidgee River and mid-Murray River where *N. australis* previously occurred but has recently been extirpated due to post-European settlement habitat modification across the MDB (Cole *et al.* 2016). Duplicates from the same coordinate point were removed to minimise the biasing effect of uneven sampling effort (Elith *et al.* 2011), reducing the dataset to 2,528 unique occurrences. Additional tests of spatial autocorrelation were performed for the 19 BioClim variables using a Pearson's pairwise correlation test in SDMToolbox (Appendix 9; Brown *et al.* 2017). Highly correlated ($|r| > 0.8$) variables were removed to avoid overfitting of the model (Dormann *et al.* 2013), reducing the environmental data down to 9 bioclimatic variables and the two topographic variables. A subset of 25% of occurrence sites were used to train the model.

Climatic data from the Last Glacial Maximum (LGM; 22 Kya) were extrapolated from the WorldClim 1.4 database (Hijmans *et al.* 2005) to project the historic distribution of

the species. To evaluate environmental conditions more reflective of the divergence between the two species, the species distribution model (SDM) was also projected back to the mid-Pliocene (~3.2 Mya) using a subset of 6 of the previous 9 BioClim variables (excluding variables bio2, bio3 and bio6) from the PaleoClim database (Brown *et al.* 2018). The fit of each SDM was determined using the area under the receiver operating curve (AUC).

A lineage-specific distribution model (LDM) method described in Rosauer *et al.* 2015 was used to determine the relative distributions of each lineage over time; this was done with two 'intraspecific' lineages of *N. australis* and *N.* 'flindersi'. A total of 72 site localities (*n* = 61 for *N. australis*; *n* = 11 for *N.* 'flindersi') were used based on genetic assignment to a 'species' within this study, as well as based on mitochondrial DNA results (Unmack *et al.* 2013). We estimated the LDMs for both species across all three time periods (current, LGM, and Pliocene). Although the location of intraspecific lineages is unlikely to remain constant in time, this method allows the inference of probable relative distributions of each 'species' under past climatic conditions.

Results

Bioinformatics

A total of 340,950,849 reads resulted in a dataset of 7,958 ddRAD loci with 45,104 variable sites (SNPs), 30,485 parsimony-informative sites and an average of 2.34% (±3.31% SD) missing data per individual. For coalescent analyses, SNPs were reduced to a single SNP per ddRAD locus, resulting in 7,780 biallelic SNPs in the joint SFS.

Phylogenetic Analysis

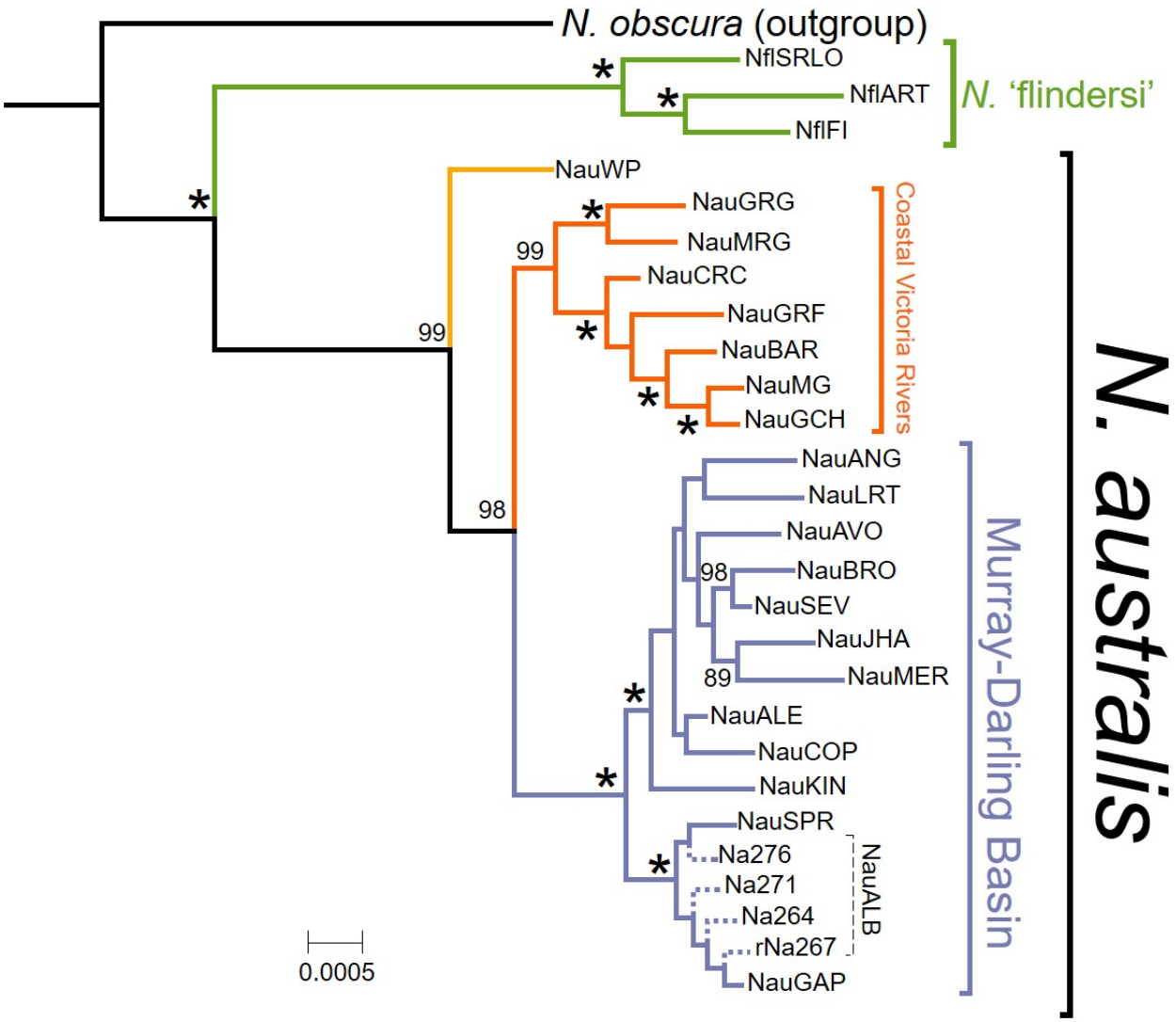


Figure 2.3: ML phylogeny of *N. australis* and *N.* 'flindersi' using 7,958 concatenated ddRAD loci containing 45,104 SNPs. As all samples within a population formed monophyletic clades (excluding NauALB, shown in dashed lines), the phylogeny was collapsed to individual populations. Branch colours denote major clades explored within coalescent models (refer to Results). Nodes with 100% bootstrap support are indicated by asterisks. The tree was rooted using *N. obscura* as the outgroup. The full phylogenetic tree with all 119 samples is shown in Appendix 10.

Partitioning the alignment per RAD locus had negligible impact on the topology and branch lengths across the phylogeny, but improved bootstrap support for several of the more conflicted nodes in the tree (Appendix 11). Site concordance factors

demonstrated more variable support across gene trees but broadly supported the same major delineations. For simplicity, we focus here on the phylogeny of the concatenated alignment. The ML phylogeny (Figure 2.3) separated *N.* 'flindersi' from the rest of *N. australis*, corroborating previous phylogenetic results (Buckley *et al.* 2018; Unmack *et al.* 2013). Within *N. australis*, three major lineages were delineated; one of the Wilson's Promontory population (NauWP), one of coastal Victoria populations and another of populations within the MDB. The coastal Victorian lineage showed relatively stronger phylogenetic structure than the MDB lineage, with its easternmost populations diverging more recently compared to westernmost coastal Victorian populations. The MDB clade, however, generally featured shorter branches and lower bootstrap support. Despite being geographically apart, lower MDB populations (Lake Alexandrina and Angas) shared a MRCA with upper Murray populations (Lachlan River and Coppabella Creek, respectively). The earliest branching clade of the MDB lineage contained populations from the upper Murray River (Spring Creek, Gap Creek, and Albury). Within this group, the Albury population was paraphyletic with the other two; this is expected based on previously described levels of admixture across the populations (Brauer *et al.* 2016). The breadth of the phylogenetic tree was well supported, with the majority of population-level and above divergences with bootstrap support of >80%.

The StreamTree results did not suggest that contemporary riverine hierarchy alone could explain patterns of historical phylogeographic divergence across the MDB (Figure 2.4; $R^2 = 0.464$). Assessing the fit of the StreamTree model by comparing the observed and expected genetic distance for each population individually demonstrated that this was likely driven by several outlier populations (NauANG,

NauALE, NauALB and NauCOP) with notably higher modelled genetic distance (Appendix 12). Removal of these four populations from the StreamTree model led to much higher correlation (R^2 = 0.982) with similar genetic distance penalties for river segments common to both sets of populations (Appendix 13).

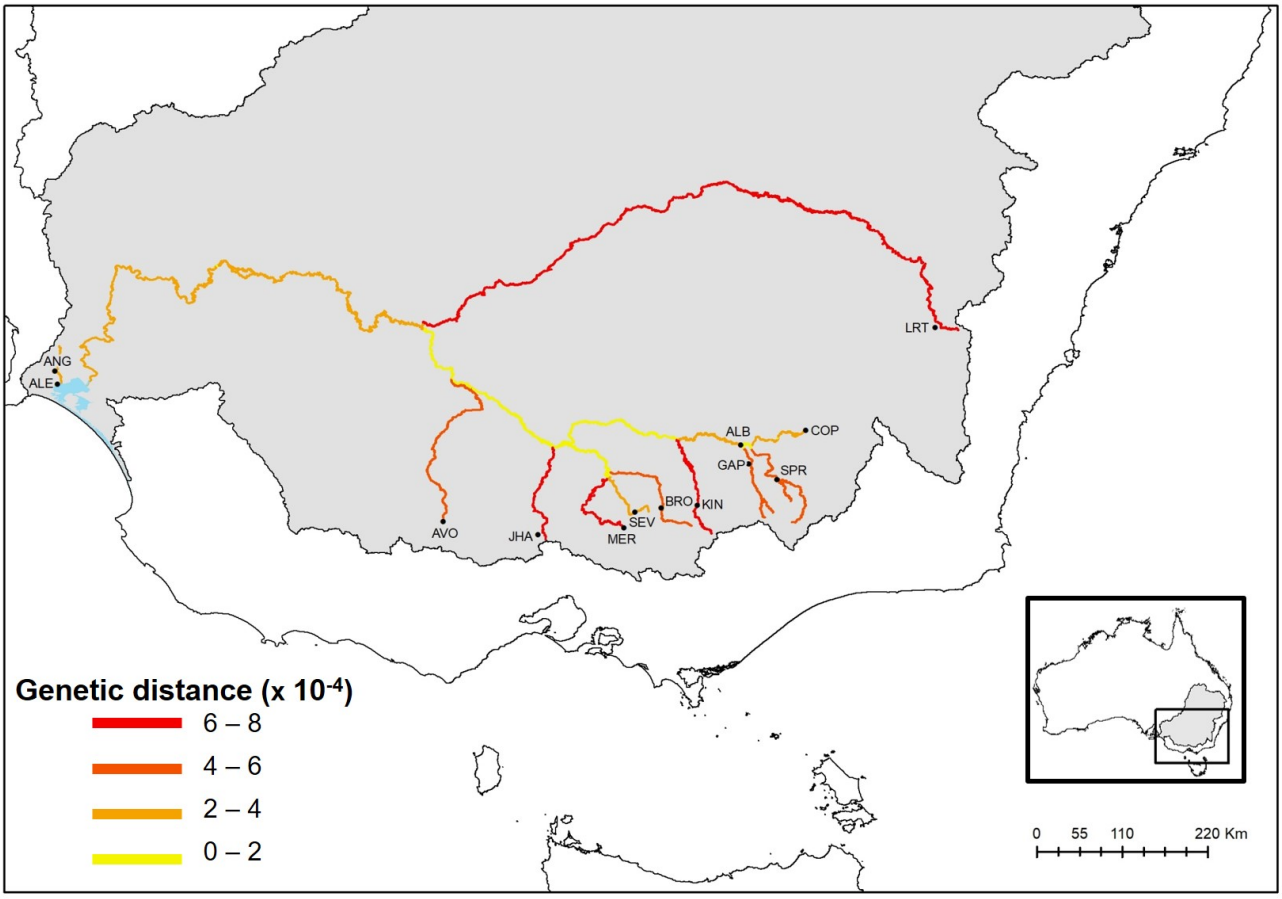


Figure 2.4: Dendritic riverine network of the MDB, with streams colour-coded according to the StreamTree model that determines the contribution (as a penalty) of each segment in driving genetic divergence across the basin. Segments coloured in yellow confer little penalty (i.e. genetic divergence between populations at either end of the segment is low) whereas red segments confer higher genetic differentiation.

Ancestral Range Estimation

Comparison of ancestral range estimations from BioGeoBEARS identified the DIVA-LIKE model as the best supported under both time-stratification scenarios (AIC = 24.16 and 59.1 for models excluding presence in the MDB until 5 Mya and 2 Mya, respectively). This model represents a likelihood approximation of the model implemented in DIVA (Ronquist, 1997) which broadly considers the relative role of dispersal and vicariance (but not sympatric mechanisms) in driving biogeographic patterns (Matzke 2013). Although patterns were similar across both time-stratification criteria (Appendix 14), we choose to focus on the more conservative (5 Mya constraint) results given the lack of precision in determining the end of marine inundation into the MDB (McLaren *et al.* 2011). This DIVA-LIKE model demonstrated strong patterns of vicariance, with weak contributions of dispersal (d = 1.46 x 10^{-2}) and extinction (e = 1 x $10^{-12)}$ and all major geographic changes associated with vicariance events (Figure 2.5). Including a parameter for founder event (+J) estimated very weak contributions of founding events and contributed to negligible change in log likelihood across either time-stratification scenario (Appendix 14). Ancestral states for major nodes were well resolved across the phylogeny.

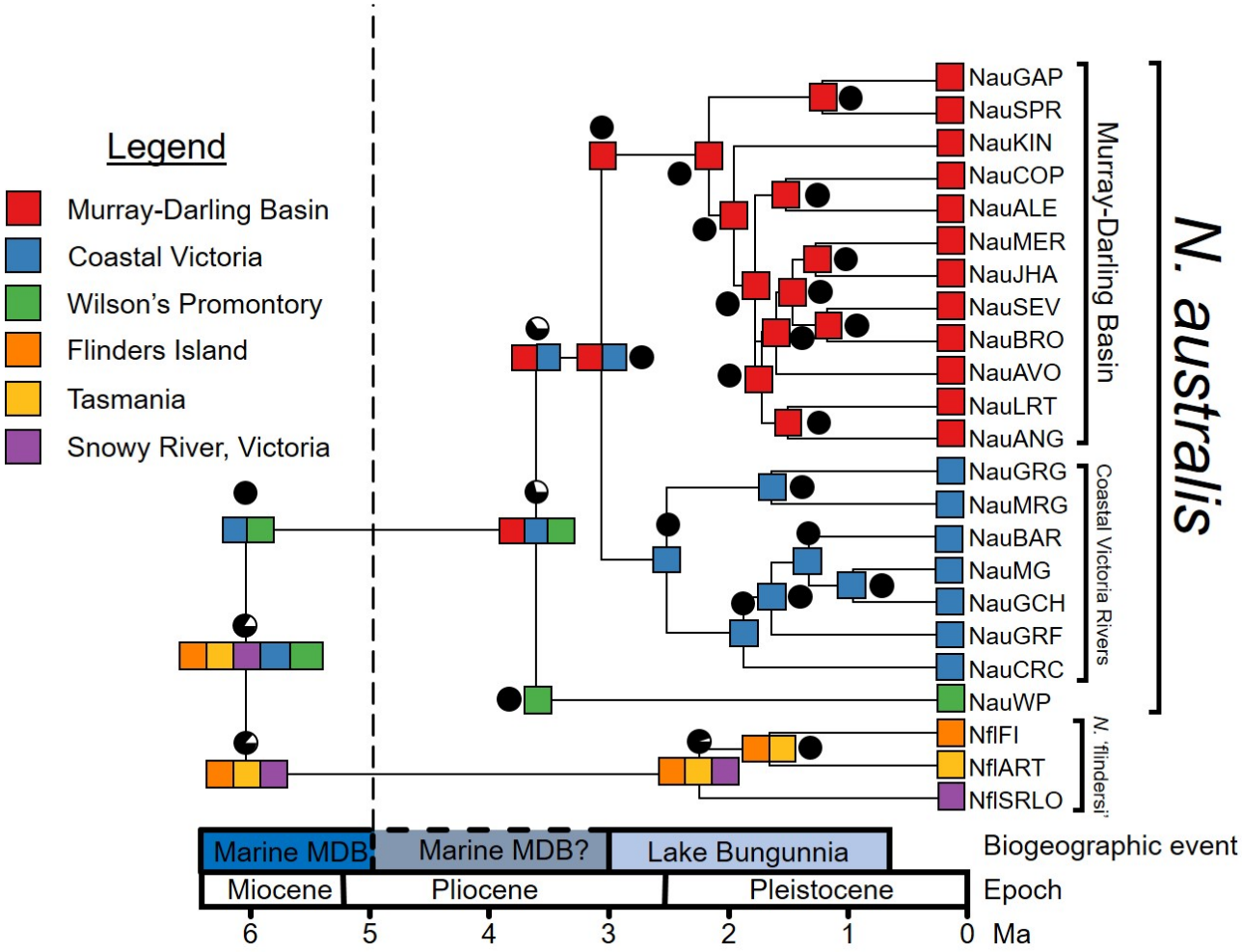


Figure 2.5: Most likely ancestral ranges under the best supported model (DIVA-LIKE), with presence in the MDB excluded until 5 Mya (indicated by the dashed line). A biogeographic timeline of major alterations to the MDB is included for reference. Colours denote one of six contemporary areas, or ranges combining more than one area, as described by the legend. Pie charts demonstrate the probability of the most likely range for each node in black.

Biogeographic Hypothesis Testing

Comparison of biogeographic hypotheses under coalescent models clearly supported one model over others for each focal divergence event (Figure 2.6): these results are detailed in the Appendix 15. In general, most models including post-isolation gene flow were better supported than those without, and models based on

vicariance due to hydrological changes were better supported than those invoking tectonic or dispersal mechanisms.

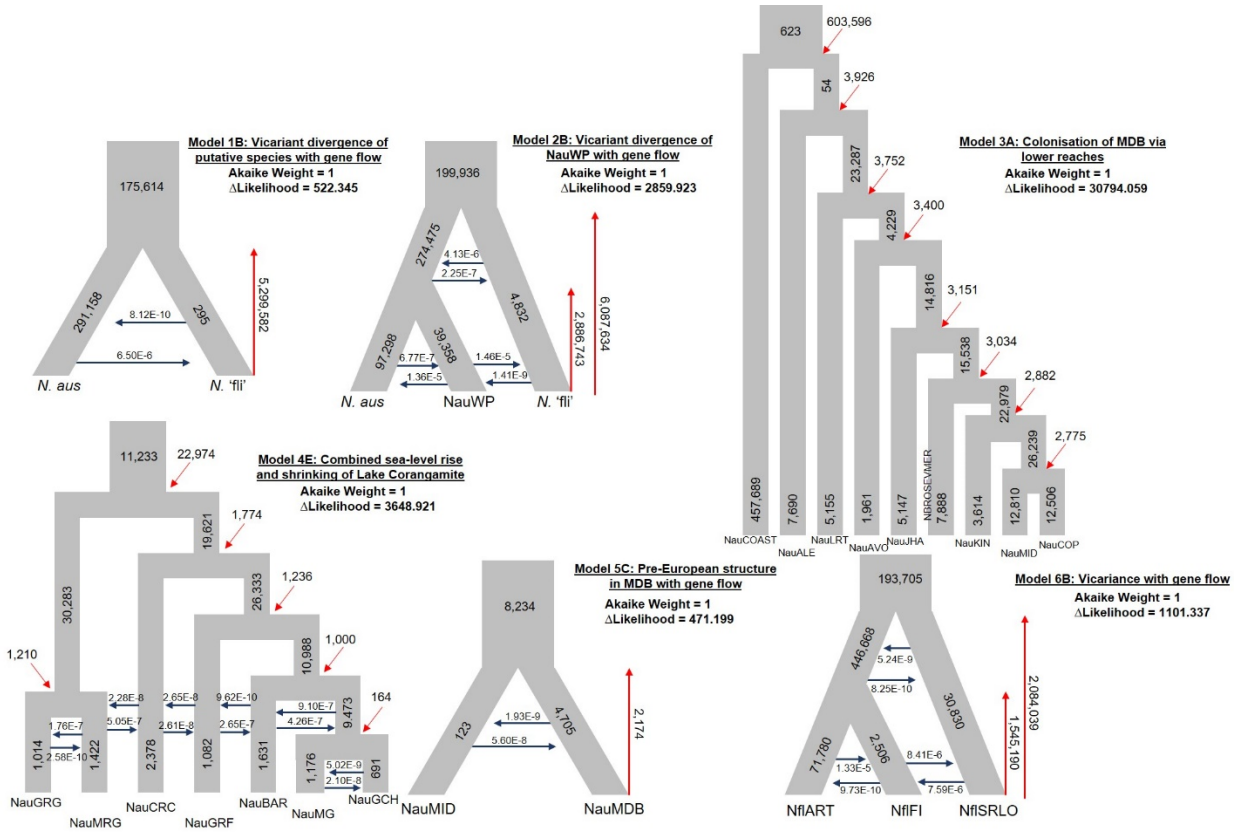


Figure 2.6: Representative diagrams of the best supported coalescent models under each model set. The full set of tested models and the biogeographic hypotheses underpinning them are described within the Appendices (Appendix 2 – 8). Red arrows denote divergence time parameters whilst blue arrows denote migration rate parameters. Population sizes are reported as the number of diploid individuals (N/2). Gene flow parameters are reported in terms of proportion of alleles moving in the direction of the arrow forward in time. ΔLikelihoods = difference between estimated (simulated) model likelihoods and observed (empirical) likelihoods.

Divergence of species. — Coalescent models including gene flow between the two species after divergence was better supported than models without, suggesting that secondary contact occurred at some point after their initial divergence (Figure 2.6;

Model 1B). Models including bottlenecks did not have substantially higher support than those without, and inferred bottlenecks were small in magnitude.

Divergence of Wilson's Promontory population. — Coalescent models of the Wilson's Promontory population and the two species separately also suggested likely gene flow across adjacent lineages, with divergence of the Wilson's Promontory population occurring ~3 Mya (Figure 2.6; Model 2B). However, this gene flow was not indicative of a hybridisation event and models simulating the Wilson's Promontory population as resulting from gene flow from either population coalesced nearly all alleles into the *N. australis* ancestral population. Thus, this population represents an anciently diverged population of *N. australis* that was isolated through vicariance.

Divergence of major N. australis lineages. — Testing migration pathways from the ancestral coastal population into the MDB suggested that colonisation most likely occurred through the lower sections of the MDB and upstream into the upper reaches (Figure 2.6; Model 3A). The timing of this dispersal event would pre-date ~600 Kya, the estimated time of divergence between the coastal and MDB populations within the best supported model. However, this model had only marginally better likelihood than one estimating the divergence time of the coast and MDB populations at ~1.2 Mya, suggesting that this estimate might not be overly precise.

Divergence within coastal Victoria lineage. — The coalescent model accounting for the effects of both sea-level changes leading to isolation of rivers and the shrinking

of Lake Corangamite isolating eastern lineages was better supported than models without these factors, or models only considering one (Figure 2.6; Model 4E).

Divergence within MDB lineage. — Coalescent modelling of the MDB lineage suggested that some phylogeographic structure pre-dated European settlement, with the divergence of the most basal lineage (containing the Spring Creek, Gap Creek and Albury populations) estimated to have occurred ~2 Kya, albeit with low levels of gene flow since divergence (Figure 2.6; Model 5C). Models partitioning putatively upper and lower populations into single demes did not converge, likely reflecting their paraphyletic nature.

Divergence within N. 'flindersi'. — Coalescent models including dispersal between adjacent populations gave much greater likelihood estimations than models without gene flow (Figure 2.6; Model 6B). Including bottlenecks indicative of a dispersal event did not improve likelihoods, supporting a range expansion and vicariance scenario. The central population of Flinders Island had a much smaller population size than either of the other two populations. Divergence time estimates for between populations suggest a relatively ancient split, with the Snowy River population separating from the other two lineages ~1.5 Mya and the secondary split between Flinders Island and Tasmania at ~1.3 Mya.

Historical Demography Reconstruction

One-dimensional site-frequency spectra estimated from single-population SNPs used a mean of 3,527 (1,045 – 7,389) SNPs (Appendix 16). Stairway plots indicated considerable declines in many populations of southern pygmy perch over the last 1

Myr (Figure 2.7). For many of these populations, gradual and concordant bottlenecks were apparent within clades. Non-declining populations were typically relatively stable, with none demonstrating an increase in *Ne* over this time. Almost all populations across both putative species demonstrated population growth deep in the past (~ 1 Mya), although this may reflect fewer ancient coalescent events within the data. Within *N.* 'flindersi', the mainland population of NflSRLO demonstrated a strong decline in *Ne* starting at ~10 Kya which contrasted with the more stable demographic histories of the island populations. This decline resulted in much lower estimates of recent *Ne* than the island counterparts, with the Tasmanian population NflART showing the highest consistent *Ne* of all *N.* 'flindersi' populations.

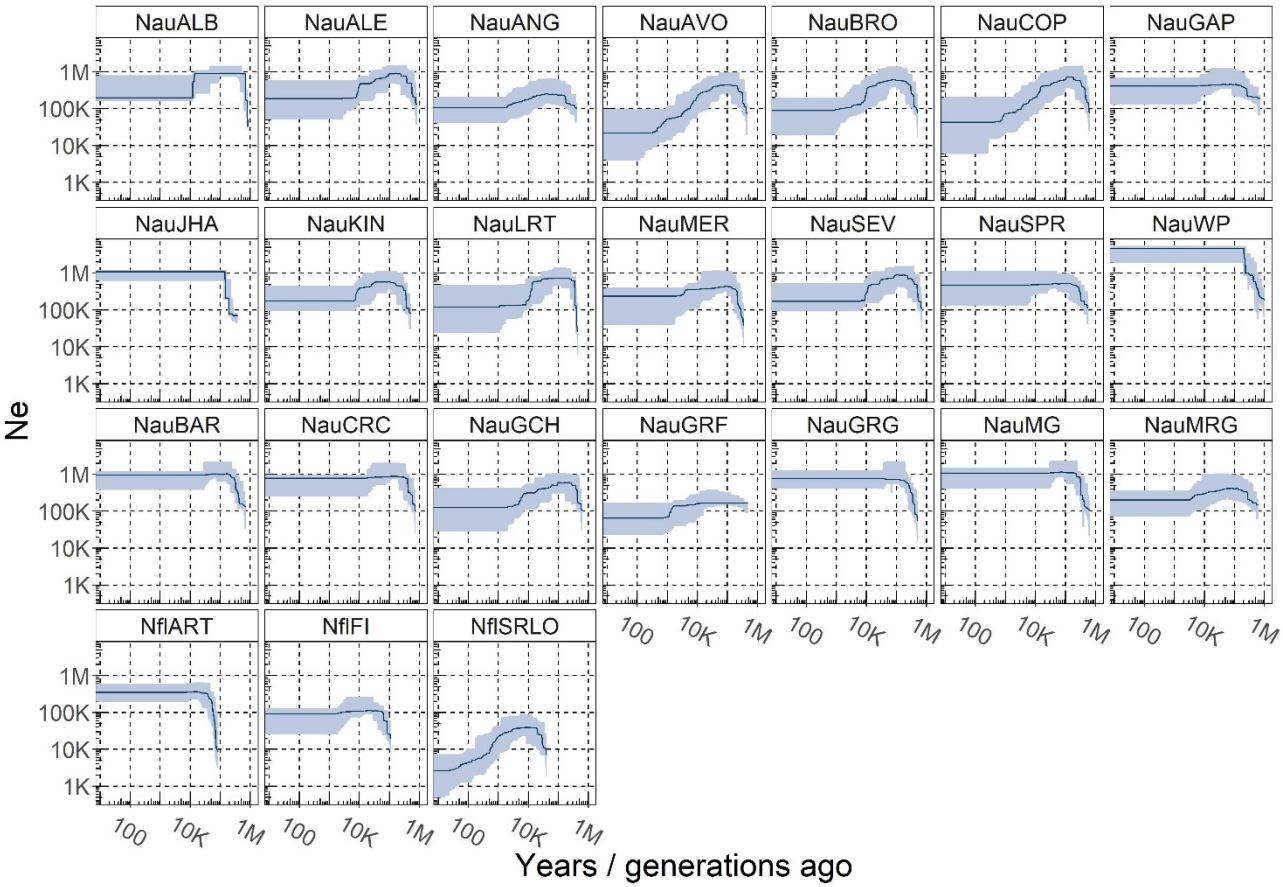


Figure 2.7: Stairway plot reconstructions of demographic history for individual populations. Both axes are reported in $\log_{10}$ scale. Dark blue lines indicate medians with 95% confidence intervals shaded.

Top two rows = MDB and Wilson's Promontory (NauWP) populations; third row = coastal Victoria populations; bottom row = *N.* 'flindersi' populations.

Demographic histories were variable across populations of *N. australis*. Within the coastal lineage, most populations demonstrated relatively stable demographic histories, with three showing weak declines in *Ne*. This contrasted with populations across the MDB, where notable population declines at ~100 Kya were observed in the majority of populations. Nevertheless, a few populations also demonstrated stable or weakly declining *Ne* over time, with NauJHA an outlier showing dramatic growth at ~ 100 Kya followed by stable *Ne.* However, this is likely driven by the lower sample size (*n* = 3) within this population compared to others across the MDB, resulting in few inferred coalescent events across the tree and a biased SFS. The highly divergent Wilson's Promontory population showed a sharp population increase at ~100 Kya followed by relatively stable and high *Ne* over time. Although all population stairway plots inferred no changes in *Ne* <5 Kya, this likely reflects a lack of very recent coalescent events within each population owing to small sample size.

Species and Lineage Distribution Modelling

The SDMs for southern pygmy perch based on the nine uncorrelated BioClim and 2 topographic variables effectively predicted the contemporary distribution for the species, showing highest habitat suitability along the Victorian coast, southern MDB and in northern Tasmania (including King and Flinders Island; Figure 2.8a). However, this SDM likely underestimates presence of *N. australis* within the connective center of the MDB, where downstream migration would have facilitated a

mosaic of intermediate populations prior to extensive flow abstraction and regulation over the last 200 years (Cole *et al.* 2016).

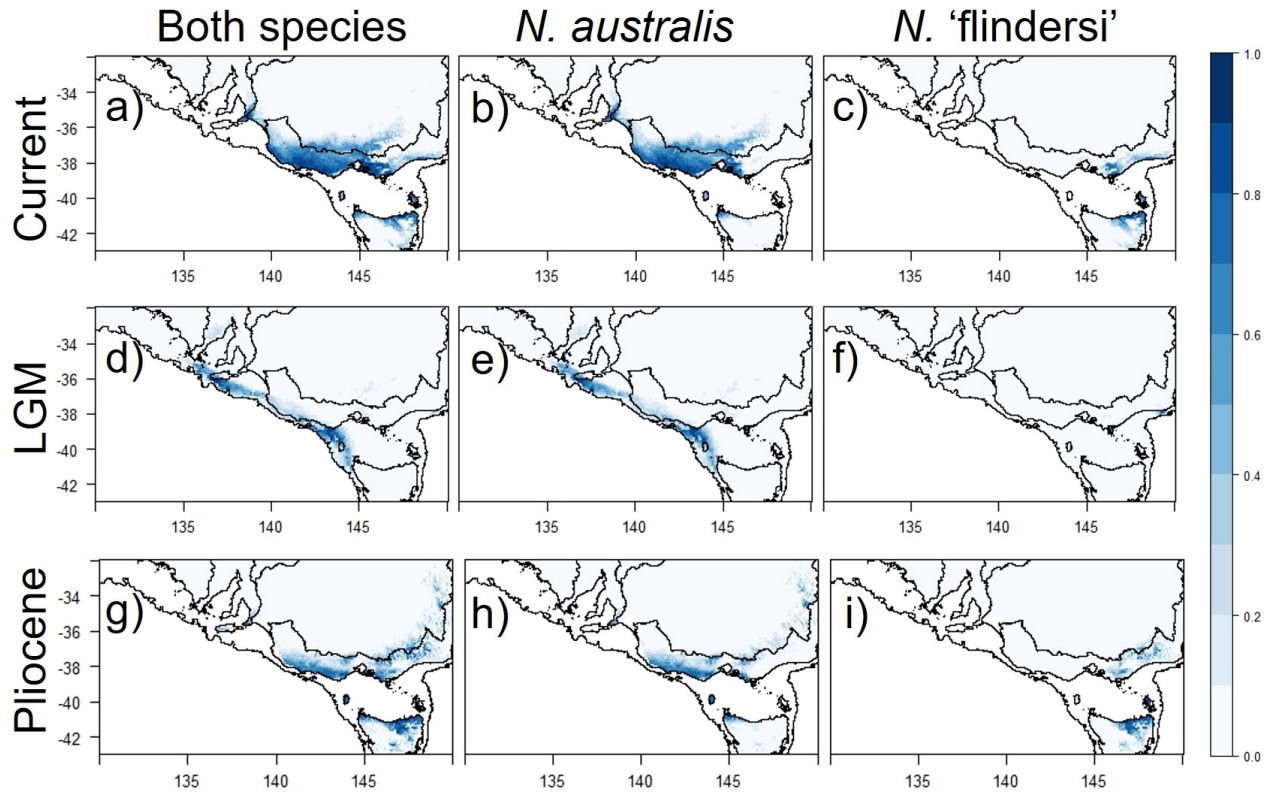


Figure 2.8: SDMs for all lineages and lineage-specific distribution models for each putative species based on 9 bioclimatic and 2 topographic variables. **a-c)** Distributions under contemporary climate conditions. **d-f)** Distributions under the LGM (22 Kya) climate conditions. **g-i)** Distributions under mid Pliocene (3.2 Mya) climate conditions.

Historic projections of the SDM for both species together highlighted two glacial refugia, one along the western coast of the distribution and another small and isolated refugium closer to the southeast corner of the mainland. Together, these results indicate an expansion in suitable habitat following the LGM (Figure 2.8d). The AUC of the model was estimated at 0.908, indicating a good model fit.

The LDMs under each time period demonstrated the disjunct spatial nature of the two species. Under contemporary conditions, the LDMs showed a geographic break near Wilson's Promontory, albeit with overlap at intermediate localities (Figure 2.8b-c). Additionally, the northern Tasmanian coastal habitat was delimited into two equally sized sections, with one per species. During the LGM, the disjunct LDMs of the two species in the east and west portion of the range supported their long-term isolation and provided evidence for an environmental barrier preventing contact between the species (Figure 2.8e-f). The Pliocene projection also showed refugial habitat along the Victorian coast and within the upper MDB (Figure 2.8g). Similarly to contemporary conditions, the LDMs showed a narrow division between species in northern Tasmania and central Victoria (Figure 2.8h-i). The subset data used for the Pliocene projection had marginally weaker support, with an AUC of 0.887.

Discussion

Establishing how aquatic-dependent lineages responded to past hydroclimatic changes contributes to our understanding about their contemporary ecological requirements and to predicting demographic responses under ongoing climate change. This study demonstrates the overarching impacts of varying hydrology due to Plio-Pleistocene climatic change (e.g. reduction of lake systems and rearrangement of river networks) on the evolution and diversification of a temperate freshwater-dependent fish clade. However, coalescent analyses and SDMs show that the evolutionary consequences of major shifts in sea level and hydroclimatic conditions varied substantially between coastal and inland environments. Aridification altered the demography of populations from inland river systems, whilst

eustatic changes and marine inundation were major evolutionary drivers of populations from coastal freshwater landscapes.

Aridification Drives Phylogeographic Structure of Inland Basins

Aridification of the Australian continent has dramatically altered the identity and stability of its ecosystems (Hawlitschek *et al.* 2012). For freshwater species, increasing aridification since the Pliocene may be responsible for a number of divergent clades, with water availability and river networks being critical for the long-term survival and evolution of freshwater lineages (Beheregaray *et al.* 2017; Faulks *et al.* 2010). Aridification and tectonics through the formation and demise of paleo-megalake Bungunnia was a major event impacting the evolutionary history of inland lineages of *N. australis*. Lake Bungunnia initially formed ~3 Mya when tectonics shifts along the Padthaway High resulted in major uplift across the region, damming the ancestral Murray River which approximately aligned with the current Glenelg River (McLaren *et al.* 2011; Waters *et al.* 2019). For many freshwater taxa across the southeast of the continent, isolation of lineages between the MDB and the southwest Victoria drainages has been associated with this tectonic shift in the Pliocene (Murphy & Austin 2004; Waters *et al.* 2019). Similar interpretations of tectonic changes influencing river capture have been proposed for movement across other sections of the Great Dividing Range into the MDB (Cook *et al.* 2006; Faulks *et al.* 2010; McGlashan 2001; Murphy & Austin 2004). The isolating effect of the tectonic changes between the two drainages is likely reflected in the strong differentiation between clades and the ancient nature of the phylogenetic-based divergence time of 3.03 Mya.

Contrasting to the phylogenetic-based divergence estimates, coalescent modelling instead points to more recent divergence (603 Kya), suggesting that the edges of Lake Bungunnia at its largest extent could have acted as suitable habitat for southern pygmy perch and potentially facilitated further dispersal into the inland basin. Given that the volume and extent of the lake fluctuated in accordance with glacial cycles and regional aridity (McLaren *et al.* 2011), connectivity was potentially only possible during its more expansive phases. With the eventual demise of the lake ~700 Kya (McLaren *et al.* 2011), this secondary bout of isolation disconnected the two basins fully, probably accounting for the results of coalescent models. Similar patterns of initial isolation by vicariance during the Pliocene, followed by Pleistocene secondary contact across the Great Dividing Range and into the MDB were observed within mountain galaxias (*Galaxias oliros* and *Galaxias olidus*), which share comparable ecological constraints to southern pygmy perch (Waters *et al.* 2019). Similar to Lake Bungunnia, the reduction of Lake Corangamite to one seventh of its original size over the course of the Holocene (White 2000) isolated several eastern coastal Victoria *N. australis* populations. The reduced size (~160 km^2) and hypersalinity (>50 g/L) of Lake Corangamite likely prevents connectivity between these populations under contemporary conditions (White 2000; Williams 1995).

Within the MDB clade there was weak evidence for historic phylogeographic structure, with coalescent models suggesting divergences dating as ~2 Kya. Correlating contemporary river structure and genetic distance *per se* did not predict genetic divergence between populations across the MDB. A combination of extensive flow and habitat modification since European settlement and naturally complex metapopulation dynamics (Brauer *et al.* 2016) are probably better proxies

for contemporary patterns observed within the MDB. This was reflected in the stairway plots, which showed a number of populations declining over the last 1 Myr but with variance in demographic histories across the basin. The most likely demographic scenarios include multiple waves of dispersal and colonisation, possibly in response to local extinction or during rare environmental events such as flooding, which altered patterns of genetic divergence.

Eustatic Changes Drive Phylogeography and Speciation Along Coastal Habitats

Sea level changes associated with interglacial periods and more arid climates played a critical role in the divergence of coastal lineages. Marine inundation across the East Gippsland region during the Mio-Pliocene (~6 Mya), prior to the climatic cycles of the Pleistocene, likely drove the initial divergence and speciation of *N.* 'flindersi'. Marine sediments and low elevation of the region indicates that this marine inundation was substantial (Gallagher *et al.* 2001; Holdgate *et al.* 2003), and was correlated with the onset of major aridification in the continent (Faulks *et al.* 2010; Garrick *et al.* 2004; McLaren & Wallace 2010). Ancient marine inundation of East Gippsland has been proposed to influence vicariant speciation in various terrestrial species (Chapple *et al.* 2005; Norgate *et al.* 2009). This low-lying region approximately forms the interface between the distribution of the two putative study species (Figure 2.1b) and the timing of this inundation corresponds well with the estimated molecular clock-based divergence time. This period of sea level rise is also associated with inundation of the lower parts of the MDB which ostensibly precluded the presence of pygmy perch (McLaren *et al.* 2011).

Previous hypotheses of the mechanisms driving the initial divergence of *N.* 'flindersi' and *N. australis* have suggested that the separation of drainages by tectonic shifts across the region (Dickinson *et al.* 2002; Gallagher *et al.* 2003) isolated populations following a dispersal event facilitated by river capture or flooding (Unmack *et al.* 2013). Regardless of the mechanism, divergence between Tasmanian and mainland lineages prior to the Pleistocene has been reported for birds (Lamb *et al.* 2019), lizards (Chapple *et al.* 2011; Dubey & Shine 2010; Kreger *et al.* 2019), butterflies (Norgate *et al.* 2009) and other freshwater fish (Coleman *et al.* 2010), suggesting that climatic oscillations during the LGM alone did not drive the speciation of *N.* 'flindersi'.

More recent sea-level changes also likely impacted within-basin phylogeographic patterns. Within *N.* 'flindersi', relatively ancient estimates of divergence times between populations (1.5 – 2 Mya) suggested that early glacial cycles of the Plio-Pleistocene resulted in strong differentiation. However, coalescent models suggested that gene flow across these disparate populations was possible during glacial maxima. At lowered sea levels, river systems occupied by *N.* 'flindersi' all drained eastward towards the continental shelf (Unmack *et al.* 2013), with shorter overland distances between river mouths than today (Figure 2.1b). Given the presence of a small glacial refugia in the far eastern extreme of the distribution, gene flow may have resulted from contraction into a singular locale followed by expansion back across the Bassian Isthmus during more favourable environmental conditions (Lambeck & Chappell 2001).

Within the coastal *N. australis* lineage, isolation of river catchments during aridification in the Pleistocene led to the strong structure observed within the phylogenetic tree. During glacial maxima, lowered sea levels substantially increased the extent of the mainland Australian coastline, particularly across the southeast corner (Williams *et al.* 2018). Although it does not appear that the current rivers of coastal Victoria coalesced before meeting the shoreline (Unmack *et al.* 2013), climatic modelling suggests that the low topographic relief and evaporation across this region would have allowed overland networks to form through small lakes and floodplains (Joyce *et al.* 2003; Williams *et al.* 2018). This environment is supported by the presence of historical cool temperate rainforest across the region (MacKenzie & Kershaw, 2000). This inferred glacial connectivity is further supported by limited divergence in mitochondrial haplotypes or allozymes across the region, in conjunction with multiple genetic exchanges across current drainage divides (Unmack *et al.* 2013). Sea-level rise and aridification during the Pleistocene inundated much of this habitat and reduced the extent of Lake Corangamite, subsequently isolating river systems from one another. Thus, along coastal Victoria aridification and eustatic changes demonstrated interactive effects on phylogeographic patterns. The more historic nature of divergences within the coastal lineage compared to the MDB lineage was corroborated by the well-resolved phylogenetic structure and relative stability of population sizes over time across the clade (as relatively recent divergences would be associated with detectable bottlenecks within the stairway plots).

Identity and Maintenance of Cryptic Species, N. 'flindersi'

Although the initial divergence between *N. australis* and *N.* 'flindersi' was associated with older biogeographic events during the Miocene, our results indicated weak post-divergence gene flow between the two species. Distribution modelling indicated a likely overlap in suitable habitat under contemporaneous conditions across the Victorian and Tasmanian coastlines, with little divergence in environmental ranges between the two species (Appendix 17). However, environmental changes during glacial maxima likely caused the two species to retract to isolated refugia. These factors together suggest a history of alternating periods of isolation and connectivity during glacial cycles, with isolated glacial refugia and weak interspecific interglacial gene flow limited to a narrow contact zone. Other studies of terrestrial species diversification across recurrently connected islands suggest patterns of gene flow in accordance with lower sea levels (Jordan & Snell 2008; Papadopoulou & Knowles 2017; Parent *et al.* 2008; Paulay & Meyer 2002). In this case, gene flow with *N. australis* does not appear to have impeded divergence and our results support the previous denotation of *N.* 'flindersi' as an independent species (Buckley *et al.* 2018; Unmack 2001; Unmack *et al.* 2013).

Implications for Conservation Management

Southern pygmy perch are currently listed as Near Threatened on the IUCN Red List and Vulnerable or Endangered within state government management lists (Hammer *et al.* 2013). Ongoing conservation management has sought to recover their numbers, particularly across the MDB (Attard *et al.* 2016a; Brauer *et al.* 2016; Cole *et al.* 2016). Across the species range, a number of clades have been identified and used as the basis for management practices (Cole *et al.* 2016; Unmack *et al.* 2013). Traditionally, conservation managers have adopted a 'local is best' paradigm,

maintaining independence of populations within captive breeding and translocation programs to prevent outbreeding depression (Frankham *et al.* 2011; Love Stowell *et al.* 2017). However, the complex nature of the southern pygmy perch populations across the MDB indicates a history likely dictated by metapopulation dynamics with natural patterns of local extinction, recolonization and sporadic gene flow (Cole *et al.* 2016). A growing body of literature suggests that the propensity and magnitude of outbreeding depression has been overestimated (Frankham *et al.* 2011). Given the low levels of genetic diversity and high imperilment of MDB populations (Brauer *et al.* 2016), as well as the recent pattern of within-basin divergence detected here, we argue that *in situ* and *ex situ* conservation efforts should use a basin-wide context when selecting populations as source for demographic and genetic rescue (e.g. captive breeding and translocations).

Implications Under Climate Change

Historic climatic fluctuations have often been used to predict future responses to anthropogenic climate change (Dawson *et al.* 2011; MacDonald *et al.* 2008). Many bioregions across the world are expected to increase in aridity with ongoing climate change (Christensen *et al.* 2007), impacting on the stability and structure of freshwater ecosystems globally (Middelkoop *et al.* 2001; Nijssen *et al.* 2001), including for the MDB (Cai & Cowan 2008; Pittock & Finlayson 2011). Our results demonstrate the strong isolating effect of aridification, particularly in the reduction or loss of major connecting waterbodies. For coastal and island regions, marine inundation similarly isolated local populations and resulted in large-scale reductions in available habitat.

Projections of sea level rise associated with melting glacial ice predict major inundation of coastal habitats globally (Rotzoll & Fletcher 2012). Although sea level rises of 1 – 2m under contemporary climate change are unlikely to impact many of the coastal *N. australis* populations, associated increases in the frequency and size of storm surges may amplify this threat (McInnes *et al.* 2011). Increased sea levels are likely to also have other secondary environmental impacts, such as salinisation of freshwater habitats, which may further impact upstream populations (Herbert *et al.* 2015). Indeed, terrestrial species extinctions have already been directly linked to inundation of island habitats (Waller *et al.* 2017). Together, these findings demonstrate the interactive effects of hydrological changes resulting from climate change on threatened freshwater ecosystems (Pittock *et al.* 2008). However, the precise impacts of climatic change on the availability and reliability of water resources are uncertain (Middelkoop *et al.* 2001), as it is also the case on the influence of hydrological change on the evolution and persistence of species.

Conclusions

Our study highlights how spatial variation in the role and extent of environmental changes may result in variable impacts on the demography, distribution and divergence of populations. Particularly, we show how aridification of inland waterbodies and sea level rise leading to marine inundation impacted different regions across the distribution of a freshwater fish, operating on different timescales and to different extents. These environmental changes caused strong divergence across the clade, resulting in a hierarchy of lineages spanning from a cryptic species to intraspecific clades. While further increases in temperature will directly impact on the long-term survival of many species broadly, additional impacts on hydrological

systems through aridification will have compounding effects on freshwater species. These findings indicate that ongoing impacts from anthropogenic climate change may be complex in nature and vary across biogeographic regions depending on the role and identity of environmental forces that operate locally. We suggest that future management scenarios should consider this spatial variation in prediction of responses to climate change, particularly in how aridification or eustatic changes may act heterogeneously across species distributions. For example, aridification of inland aquatic systems may further fragment populations and reduce connectivity, highlighting the need for environmental watering and rehabilitation of river reaches. Concordantly, marine inundation of coastal regions may significantly reduce available habitat for terrestrial and freshwater species, making translocations necessary in the immediate future. The combination of these factors should be considered within management frameworks, particularly for species that may be impacted by both simultaneously.

Chapter 3: Variation in intraspecific demographic history drives localised concordance but species-wide discordance in responses to Plio-Pleistocene climatic change

Sean J. Buckley[1], Chris Brauer[1], Peter J. Unmack[2], Michael P, Hammer[3], Luciano B. Beheregaray[1]

[1]Molecular Ecology Laboratory, College of Science and Engineering, Flinders University, Adelaide, SA 5001, Australia

[2]Centre for Applied Water Science, Institute for Applied Ecology, University of Canberra, ACT 2601, Australia

[3]Natural Sciences, Museum and Art Gallery of the Northern Territory, Darwin NT 0801, Australia

This chapter is under preparation for submission to the Proceedings of the Royal Society B: Biological Sciences. I am the primary author, with Dr. Chris Brauer, Dr. Peter Unmack, Dr. Michael Hammer and Prof. Luciano Beheregaray as co-authors. I was responsible for the majority of data collection, all analyses and drafting the manuscript. Chris Brauer provided some sequences from previous work. Luciano Beheregaray provided comments and edits on the manuscript and supervised the project.

Abstract

Understanding how species biology may facilitate resilience to contemporary climate change remains a critical factor in detecting and protecting species at risk of extinction. Many studies have focused on the role of particular ecological traits in driving species responses, but less so on demographic history and levels of standing genetic diversity. Here, we used environmental and genomic datasets to reconstruct the phylogeographic histories of two ecologically similar and largely co-distributed freshwater fishes (southern pygmy perch, *Nannoperca australis*, and Yarra pygmy perch, *Nannoperca obscura*) to assess the degree of concordance in their responses to Plio-Pleistocene climatic changes across southeast Australia. Although several co-occurring populations demonstrated concordant post-glacial bottlenecks, idiosyncratic population size changes were found at the range edges of the more spatially restricted species. However, *N. australis* demonstrated a significant historical expansion in range during the last glacial maximum whilst the *N. obscura* distribution remained constant throughout the Plio-Pleistocene. These discordant responses to climatic changes between species were associated with low levels of standing genetic variation in peripheral populations of *N. obscura* which might have hindered their adaptive potential. Our results highlight both the role of spatial scale in the degree of concordance in species responses to climate change, and the importance of standing genetic variation in facilitating range shifts under climate change. Even when ecological traits are similar between species, long-term genetic diversity and historical population demography may lead to discordant responses to ongoing and future climate change.

Introduction

Anthropogenic climate change is one of the greatest threats to global biodiversity. Understanding how or whether species may be able to adapt to current and future climatic changes is critical for conservation management of threatened taxa (Waldvogel *et al.* 2020). However, given that the majority of taxa across the globe are understudied, predicting the susceptibility and extent of species loss due to climate change remains a challenge. To this end, many studies have instead sought to determine ecological traits that may confer resilience or susceptibility to climate change across various taxa (Healy *et al.* 2018). Ecological and physiological traits such as thermal tolerance and dispersal capacity have been shown to be critical in driving adaptation to climatic changes (Somero 2010; Travis *et al.* 2013). Additionally, demographic and genetic traits such as population size, stability and standing genetic diversity are also important in facilitating adaptation to new environmental stressors (Pearson *et al.* 2014), and likely play a major role in species responses to climate change (Williams *et al.* 2008).

From a genetic perspective, adaptation to novel climatic conditions more often relies upon standing genetic variation (SGV) than *de novo* mutations (Hermisson & Pennings 2005; Lai *et al.* 2019; Morris *et al.* 2018). Thus, the maintenance of genetic diversity is critical to allow species to adapt to the changing climate (Fordham *et al.* 2014; Waldvogel *et al.* 2020). The degree to which SGV is maintained within species or a population varies significantly across taxa and is influenced by a complex of demographic, ecological and environmental factors. For example, populations occurring at the edge of a species range often have lower genetic diversity and connectivity than their more central counterparts (Aitken *et al.* 2008; Eckert *et al.* 2008). Conversely, less-connected edge populations exposed or responding to contrasting environments can accrue or maintain unique components of SGV (Antunes *et*

al. 2006; Eckert *et al.* 2008; Hampe & Petit 2005). Relatedly, spatial variation in local adaptation across populations may influence the ability of species to respond to climate change. In marginal populations, persistence is driven by the balance of the steepness of the selective environment and the effectiveness of selection relative to genetic drift (Polechova & Barton 2015). These components may contrast with populations occurring at the core of the distribution, where larger carrying capacities and SGV allow populations to persist closer to their selective optimum (Bridle *et al.* 2010). Thus, the interaction of neutral (demographic) and adaptive (ecological) traits, and their spatial variability, are critically important in understanding how species ranges may shift under climate change (Angert *et al.* 2008; Hoffmann & Sgro 2011).

Understanding factors underlying species responses to historical climatic fluctuations provides an empirical framework for determining how species may respond to current and future environmental changes (Fordham *et al.* 2014; Prates *et al.* 2016). Phylogeographic studies often seek to do this through a combination of population genetic, phylogenetic and environmental analyses to synthesise a multifaceted overview of species histories (Fenderson *et al.* 2020; Fordham *et al.* 2014). Recent analytical advancements in statistical approaches to inferring phylogeographic history, particularly through coalescent theory (Cutter 2013; Hickerson *et al.* 2010; Xue & Hickerson 2015), have substantially improved the power and resolution for detecting past changes in key demographic parameters (Beheregaray 2008; Garrick *et al.* 2019). Extending phylogeographic analyses from taxon-specific studies to assessments of how species assemblages have responded to past climatic changes provides an approach to estimating the ubiquity of species responses (Potter *et al.* 2018). Often, this framework takes a comparative approach to assess how particular ecological traits spread over disparate and divergent taxa may confer similar

responses to climate change (Papadopoulou & Knowles 2016; Riddle 2016; Sullivan *et al.* 2019). Similar species responses (concordance) across disparate taxa often indicate that shared ecological traits underlie the response (Paz *et al.* 2015; Rincon-Sandoval *et al.* 2019), or demonstrates the ubiquity in impact of the particular environmental change in question (Avise *et al.* 2016; Barrow *et al.* 2018). Contrastingly, idiosyncratic responses (discordance) are often attributed to either variation in species-specific traits (Massatti & Knowles 2014; Zamudio *et al.* 2016), which more strongly influence how each species has been impacted by the environmental change, or considered a nuisance aspect owing to stochasticity over time (Papadopoulou & Knowles 2016).

The simplified 'concordance vs discordance' paradigm has been criticised, particularly for its overemphasis on the assessment of concordance (Papadopoulou & Knowles 2016). Demographic characteristics such as population size, connectivity and genetic diversity are often spatially variable (Vucetich & Waite 2003), and the factors that underlie this variation might not be the same across species. By extension, intraspecific variation in demography may lead to spatial variation in the degree of concordance, even across ecologically similar species (Figure 3.1). For example, the interactive role of demographic aspects and adaptive potential may lead to intraspecific variation at local scales, even if species-wide patterns are concordant across taxa or vice versa (DeChaine & Martin 2005; Papadopoulou & Knowles 2016; Roe *et al.* 2011). These patterns may be reflected within shifts in species ranges over time, where intraspecific variation in demographic or ecological traits at range margins may drive interspecific discordance in species responses to environmental change.

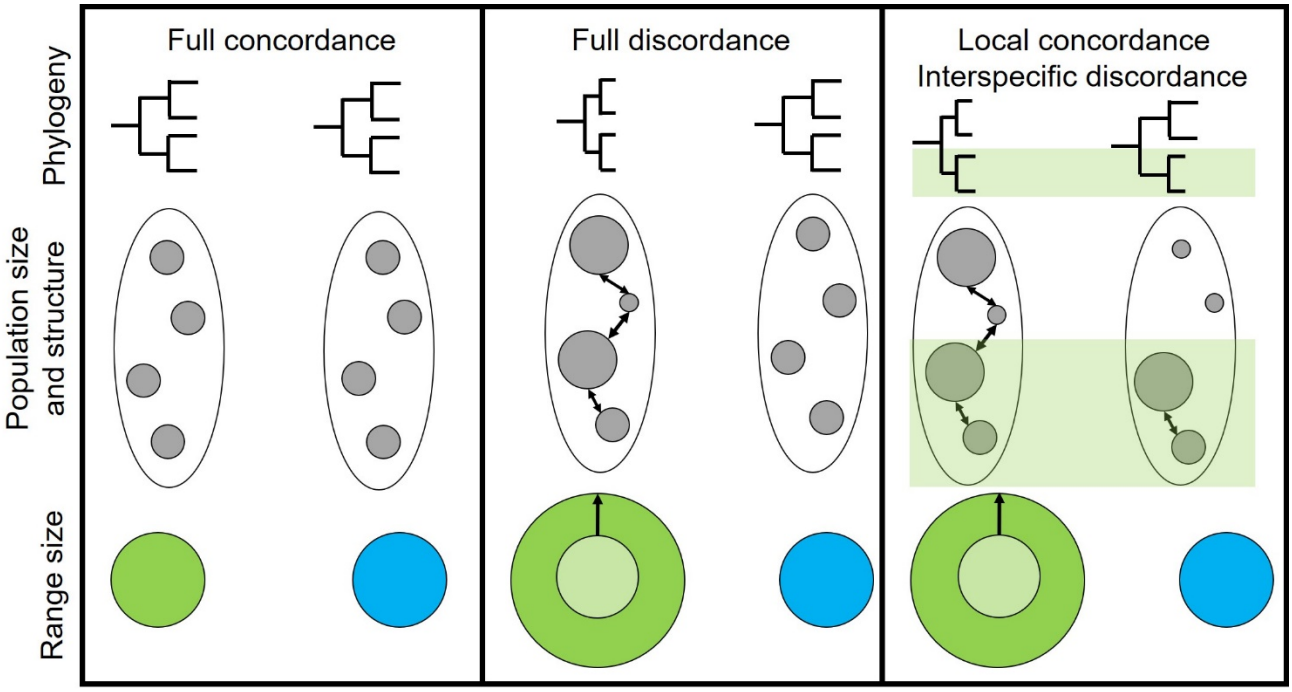


Figure 3.1: Diagrammatic representation of concordance and discordance paradigm across two co-distributed species. Top row: phylogenetic relationships and branch lengths of populations per species. Middle row: each circle indicates a single population, with the size of the circle denoting population size and arrows indicating connectivity amongst populations. Bottom row: species ranges, with arrows indicating an expansion in range. Left column: full concordance between species across all three measures. Middle column: full discordance across all three measures. Right column: localised concordance in the highlighted lineage/populations, but discordance in the other lineage, leading to discordance in species ranges.

Biogeographic regions that experienced large variation in the degree of aridity and sea level changes are particularly useful for studying and predicting species responses to climate change (Courchamp *et al.* 2014; Davis *et al.* 2013; Menon *et al.* 2010). In this regard, the southeast Australian temperate zone provides a model region to test how species have responded to major environmental changes owing to a combination of aridification and eustatic changes. Continental Australia has experienced significant environmental changes since the late Miocene, which heralded the onset of major aridification (Byrne *et al.* 2008). During the Pliocene, a brief humid period temporarily shifted the landscape back to wetter

conditions, and allowed the formation of large waterbodies such as the paleo-megalake Bungunnia (McLaren & Wallace 2010). However, the progression into the Pleistocene saw a return of aridification; while glacial periods in this region were not directly associated with the formation of glaciers, major changes in precipitation and temperature shifted ecosystem types to more arid conditions (Duckett *et al.* 2013). Concordantly, glacial maxima also drove eustatic changes, expanding much of the continental shelf as sea levels dropped (Schultz *et al.* 2008). The complex environmental history, and the role of aridification and eustatic changes on the evolution of temperate species, has been demonstrated by a number of phylogeographic studies (e.g. Ansari *et al.* 2019; Chapple *et al.* 2011; Neal *et al.* 2019).

Freshwater-dependent species are important indicators of historical environmental changes given their reliance on suitable habitat and often limited capacity for dispersal (Davis *et al.* 2018; Hughes *et al.* 2009). Within temperate southeast Australia, the often co-distributed diminutive percichthyids, southern (*Nannoperca australis*) and Yarra (*Nannoperca obscura*) pygmy perches provide an ideal comparative study system. Both species possess highly similar morphology, reproductive biology, salinity tolerance and habitat preferences, and also display similar patterns of metapopulation structure and dynamics (Attard *et al.* 2016a; Brauer *et al.* 2016; Brauer *et al.* 2013; Cole *et al.* 2016; Hammer *et al.* 2013; Wedderburn *et al.* 2012). While both species are relatively old (e.g. their lineages are thought to have coalesced around 13 million years ago; Buckley *et al.* 2018), there is some evidence of historical introgression (estimated at ~1 Mya) across them in a few localities (Hammer *et al.* 2010; Unmack *et al.* 2013). Within these populations, complete replacement of the mitogenome of *N. obscura* with the mitogenome of *N. australis* (Prosdocimi *et al.* 2012; Unmack *et al.* 2011) may reflect their close ecological relationship.

Both species have low dispersal capacity, and show little to no connectivity across catchments (Brauer *et al.* 2013; Cole *et al.* 2016; Wedderburn *et al.* 2012). Both species show strong population structure, with two evolutionary significant units (ESUs) separating MDB populations and coastal populations in *N. australis* (Unmack *et al.* 2013), and two clades each containing two ESUs in *N. obscura*: one of Eastern and Merri lineages, and another containing MDB and Central lineages (Hammer *et al.* 2010). Given their isolated populations, it is expected that their long-term persistence across environmental gradients depends on spatial variation of locally adaptive traits. This hypothesis is consistent with studies of *N. australis* that show that patterns of adaptation in traits related to reproductive fitness (Morrongiello *et al.* 2010; Morrongiello *et al.* 2012), in levels of adaptive genetic diversity (Brauer *et al.* 2016) and in variance of gene expression (Brauer *et al.* 2017) are strongly associated with patterns of hydroclimatic gradients at both local and regional scales.

Despite their close relationship and ecological similarities, the two species demonstrate marked differences in conservation status, genetic diversity and total distribution range. While both species are of conservation concern (*N. australis* as Near Threatened and *N. obscura* as Endangered) within the IUCN Red List (Beheregaray *et al.* in press) and in state conservation legislation (Table 1.1), some lineages of *N. obscura* are considered at higher risk due to their narrow range and extremely low genetic diversity (Attard *et al.* 2016a; Brauer *et al.* 2013). These factors are implicated in the local extirpation of *N. obscura* within the MDB in the last five years, following failed reintroductions across the area after a large-scale water shortage and drought impacted the region (Beheregaray *et al.* in press). The relatively low genetic diversity of *N. obscura* has not been associated with a particularly severe past bottleneck (Attard *et al.* 2016a), complicating determining factors underlying

this disparity. Additionally, it remains unclear whether the historical absence of *N. obscura* in some regions where *N. australis* is found is the result of historical local extinctions or a failure to initially colonise the habitat.

Here, we applied a comparative phylogeographic framework to explore the relative roles of ecological and demographic traits on evolutionary history. We used genomic (ddRAD-seq) datasets to estimate genetic diversity, phylogenetic relationships and demographic history of these two freshwater fishes, in conjunction with SDMs. Then, we statistically evaluated local-scale concordance across co-occurring populations to assess whether the two species shared demographic responses to environmental changes over the Plio-Pleistocene. We predicted that phylogenetic patterns, demographic histories and distribution changes would be concordant across the two species if ecological factors played a relatively strong role in determining species responses to past climatic changes, with current differences between species owing to more recent factors. Contrastingly, discordant evolutionary histories would indicate that genetic diversity and demography played a relatively larger role on disparate species responses and underpinned their contemporary differences in conservation status. Our framework also includes differentiation of local-scale (intraspecific, population-level) and wide-scale (species-level) responses to assess the role of intraspecific patterns in driving lineage responses.

Materials & Methods

Sample Collection and Library Preparation

The distribution of both species spans the southwest Victoria biogeographic province and the lower reaches of the MDB (Wedderburn *et al.* 2012). *Nannoperca australis* is more widely distributed and is also found across eastern Victorian drainages, northern Tasmania

and the upper reaches of the southern MDB (Unmack *et al.* 2013). The selection of individuals for this work was based on previous comprehensive population genetic studies (e.g. 578 individuals and 45 localities for *N. australis*; Cole *et al.* 2016); 541 individuals and 27 localities for *N. obscura*; Brauer *et al.* 2016), as well as on information from phylogeographic and landscape genomic studies of both species (Brauer *et al.* 2013; Unmack *et al.* 2013). The final sample contains all known genetically distinct populations and conservation management units of both species (as well as recently extirpated populations), including populations covering their full co-distributed range but excluding those where past hybridisation may confound patterns of species-specific demographic histories (Table 3.1; Unmack *et al.* 2013). This equals to seven populations of *N. obscura* and nine populations of *N. australis* occurring across all major drainages of the region (Figure 3.2). An additional 10 and 15 *N. obscura* and *N. australis* (respectively) from Lake Alexandrina within the lower MDB were also included for more targeted demographic reconstruction of these populations. For phylogenetic analyses, an additional 5 samples of a sister species (*Nannoperca vittata*) were included as outgroup (Figure 1.4). Specimens were collected using electrofishing, dip-, fyke- or seine-netting. Either the caudal fin or the entire specimen was stored at −80°C at the South Australian Museum, or in 99% ethanol at Flinders University.

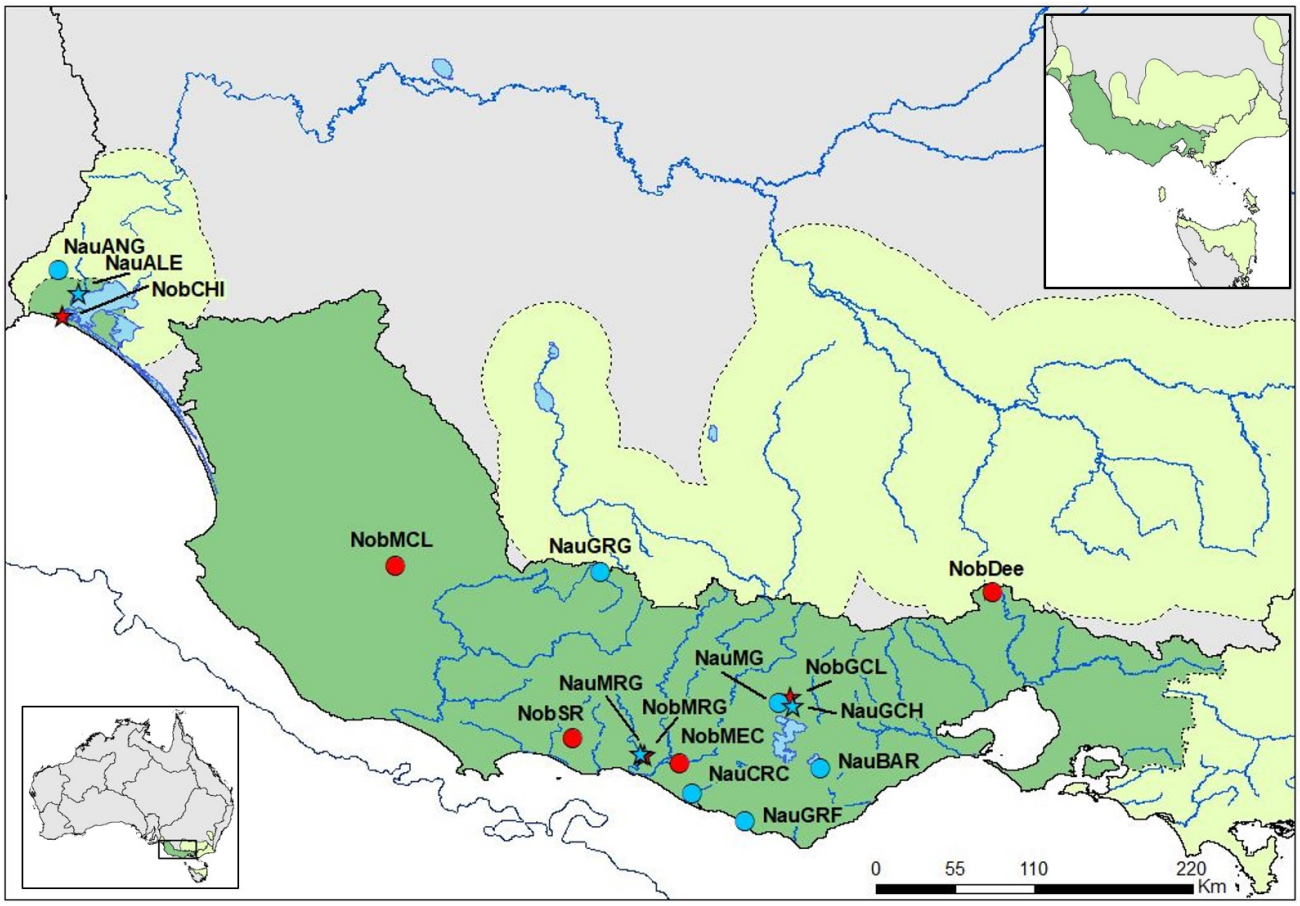


Figure 3.2: Contemporary distribution and sampling map for *N. australis* and *N. obscura*. *Nannoperca australis* sampling sites are indicated in blue, and *N. obscura* sites indicated in red. The distribution of *N. australis* is indicated with light green shading and dashed borders, with the distribution of *N. obscura* (and subsequently the region of co-occurrence) in darker green. The solid black line indicates the boundary of major drainage basins, and the dark blue line demonstrates the approximate shoreline during glacial maxima. Bottom left inset depicts study region and major drainage basins in Australia. Top right inset depicts the full extent of species distributions across southeast Australia.

DNA was extracted from muscle tissue or fin clips using a modified salting-out method (Cole *et al.* 2016) or a Qiagen DNeasy kit (Qiagen Inc., Valencia, CA, USA). Genomic DNA was checked for quality using a spectrophotometer (NanoDrop, Thermo Scientific), integrity using 2% agarose gels, and quantity using a fluorometer (Qubit, Life Technologies). The ddRAD genomic libraries were prepared in-house at the Molecular Ecology Lab of Flinders University following (Brauer *et al.* 2016), including modifications described therein. The

majority of the samples (56/98) were paired-end sequenced on an Illumina HiSeq 2000 at Genome Quebec (Montreal, Canada). The remaining 38 samples were single-end sequenced on an Illumina HiSeq 2500 at the South Australia Health and Medical Research Institute (SAHMRI).

Table 3.1: Locality data for samples used in this study. Abbreviations described in the table were those used for further analyses, while n refers to the number of individuals sequenced per population. *Nannoperca vittata* samples were only included as an outgroup in the phylogenetic analysis.

Species	Population	Abbreviation	Field code	Latitude	Longitude	N
N. australis	Angas, Middle Ck. Junction	NauANG	F-FISH84	-35.25	138.887	5
	Lake Alexandrina	NauALE	SPPBrA*	-35.395	139.008	4
	Glenelg Rvr., Glenisla	NauGRG	F-FISH78: PU0014-SPP	-37.1369	142.262	5
	Meriri Rvr., Grassmere	NauMRG	F-FISH78: PU00-22SPP	-38.2638	142.5151	5
	Curdies Rvr., Curdie	NauCRC	F-FISH78: PU00-24SPP	-38.5189	142.836	4
	Gellibrand Rvr. floodplain	NauGRF	F-FISH97: PU02-92SPP	-38.6919	143.1675	5
	Barongarook Ck., Colac	NauBAR	SPP08-13	-38.36	143.6404	4
	Mundy Gully	NauMG	F-FISHY8: PU08-11SPP	-37.9523	143.3794	4
	Gnarkeet Ck., Hamilton	NauGCH	F-FISHY2: PU00-27SPP	-37.9706	143.4674	4
N. obscura	Cleland, Hamilton Island	NobCHI	YPPBr*	-35.537	138.905	5
	Mt. Emu Ck.	NobMEC	PU02-112YPP	-38.325	142.759	5
	Mosquito Ck., Langkoop	NobMCL	PU00-16YPP	-37.0938	140.9841	4
	Shaw Rvr.	NobSR	PU02-113YPP	-38.1699	142.0902	2
	Merri Rvr., Grassmere	NobMRG	PU02-111YPP	-38.275	142.542	4
	Gnarkeet Ck., Lismore	NobGCL	PU00-27YPP	-37.9102	143.4469	5
	Deep Ck., Lancefield	NobDee	F-FISH97: PU02-106YPP	-37.259	144.713	2
N. vittata	Chesapeake Rvr.	*Outgroup*	F-FISHx2: V22, V64, V83, V85, V185	-34.8362	116.333	5
Total		**16**				**72 (77)**

Sequence Filtering, Alignment and SNP Calling

Sequences were demultiplexed using the 'process_radtags' module of Stacks 1.29 (Catchen *et al.* 2013), allowing up to 2 mismatches in barcodes. Barcodes were removed and sequences trimmed to 80 bp to remove low-quality bases from the end of the reads. Cut reads were aligned using PyRAD 3.0.6 (Eaton 2014), and further cleaned by removing

reads with >5 bp with a Phred score < 20. Loci were retained if they occurred in at least ~80% of samples (22 in *N. obscura*; 30 in *N. australis)* within the phylogenetic datasets.

Given the power of unlinked SNPs to reconstruct demographic histories (Brito & Edwards 2009), particularly within simulation-based frameworks where full genomic sequences are often intractable (Excoffier *et al.* 2013), we subsampled our ddRADseq data for some population-specific analyses (e.g. some measures of genetic diversity and coalescent-based demographic histories). Loci were re-aligned and SNPs called separately for each population (excluding those with $n < 3$ due to low sample size) using PyRAD. This was done as most SNPs called within each species were monomorphic within individual populations and would have significant effects on downstream analyses (Stadler *et al.* 2009). Only loci present in all individuals were kept to prevent missing data from biasing the SFS, used in all demographic analyses (Shafer *et al.* 2017).

Contemporary Genetic Diversity

Summaries of population-level genetic diversity parameters (allelic richness and gene diversity) were compared across populations within species, and between species, using the R package *hierfstat* (Goudet 2005). Given uneven sample sizes across populations, rarefaction was used ($n = 4$) to estimate mean values per locus per population. Due to the larger sample sizes available for Lake Alexandrina populations, genetic diversity parameters were also calculated using $n = 15$ rarefaction and loci aligned separately within each Lake Alexandrina population. For these populations, we also calculated effective population size (*Ne*) using NeEstimator (Do *et al.* 2014) and a lower allele frequency threshold of 0.02. Additionally, nucleotide diversity (π) within each population was estimated using dnaSP 6.1 (Rozas *et al.* 2017). Differences in population means of genetic

diversity parameters between the two species were statistically evaluated using t-tests (either a two-tailed t-test if data were distributed normally or Wilcox test if not) using R.

Phylogenetic and Historical Migration Analyses

The ML phylogenies of each species were estimated using RAxML 8.2.11 (Stamatakis 2014) with the concatenated ddRAD alignments to estimate evolutionary relationships. Phylogenies were estimated under the GTR-GAMMA model of evolution and 1,000 RELL bootstraps for each species. Additionally, we estimated gene trees for each RAD locus using IQ-TREE2 (Minh *et al.* 2020b) to account for heterogeneity across the genome and the potential impact of incomplete lineage sorting. Gene and site concordance factors (Minh *et al.* 2020a) were estimated by comparing individual gene trees to the concatenated RAxML tree.

As historical migration may impact on the shape and bifurcating nature of a phylogenetic tree (Edwards *et al.* 2016), and to infer historical population connectivity, we used the allele frequency-based covariance estimator TreeMix (Pickrell & Pritchard 2012). For this analysis, we iteratively increased the number of migrations from 0 – *n* for each species (nine in *N. australis;* seven in *N. obscura*) and evaluated the fit of each tree by comparing the standard error of the covariance matrix (W) and the overall likelihood of the tree. We further assessed the fit of the tree estimated under each number of migrations allowed by calculating the percentage of variance explained per model (Github = wlz0726/treemix). The best supported number of migrations was determined by the asymptote of the likelihood, where additional migrations did not substantially increase model likelihood.

Comparative Demographic Inference

Long-term demographic histories for each population of the two species were estimated using stairway plots and the SFS. One-dimensional SFS were calculated for each independent population alignment using an in-house script. Stairway plots were estimated assuming a mutation rate of 10^{-8} mutations per site per generation, and a generation time of one year for both species (Attard *et al.* 2016a; Humphries 1995). Although both species reproduce annually, most individuals do not live beyond one to two years in the wild (Humphries 1995). We then analysed co-distributed populations of *N. australis* and *N. obscura* under two coalescent frameworks to statistically evaluate the degree of concordant demographic history. The populations of Gnarkeet Creek (NauGCH and NobGCL), Merri River (NauMRG and NobMRG) and Lake Alexandrina (NauALE and NobCHI) were selected based on their contemporary co-occurrence and to represent the geographic range of the overlap in species distributions (Figure 3.2). We first investigated most likely demographic syndrome (e.g. growth, decline, stable population size) independently for each population, and then performed a hierarchical comparison across all six populations.

Given the reduced power of stairway plots to detect recent changes in *Ne*, we used FastSimCoal2 (Excoffier *et al.* 2013) to simulate model-based demographic histories over the last 30 Kyr. Simulations of *Ne* were conducted for each population under five different single population demographic scenarios: stable *Ne* over time, exponential growth, exponential decline, and fluctuating growth/decline (Figure 3.3). Parameters were estimated using 40 optimisation cycles with 500,000 simulations per demographic scenario and the fit of the models estimated using AIC and Akaike weights. Confidence intervals for the parameters specified in the best supported demographic model per population were estimated by simulating 100 SFS and re-simulating point estimates using 500,000 iterations per SFS.

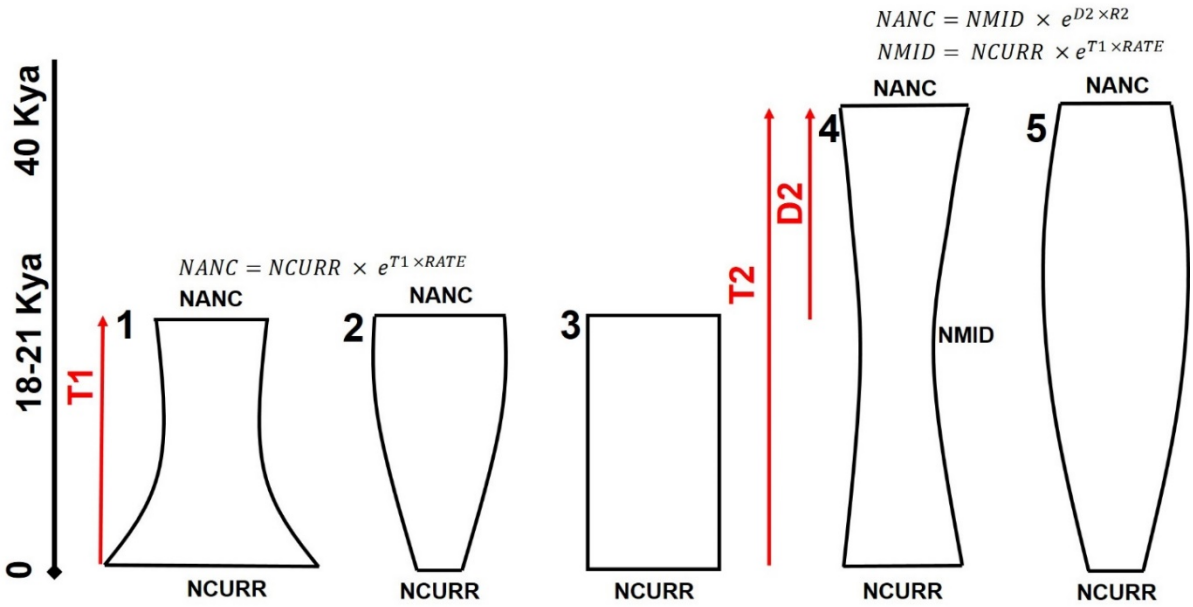


Figure 3.3: Diagrammatic representations of demographic syndromes tested per population selected for codemographic analysis, using FastSimCoal2. Model numbers correspond to the same numbers in Results and Appendices. The width of each diagram corresponds to *Ne* over time, changing from current *Ne* (at the bottom of the figure) back in time (moving upwards). *RATE* and *R2* parameters indicate rates used to calculate exponential growth or decay of *Ne* over time, with formula for *Ne* at given points indicated above each model. Red parameters indicate timing of the end of growth or decay rates, depending on the model. Priors for all parameters are defined in Appendix 18.

Additionally, we ran co-demographic model-based simulations using the aggregate SFS (aSFS) and the *R* package Multi-DICE (Xue & Hickerson 2017) to determine if demographic histories were congruent across co-distributed populations. Multi-DICE implements a hierarchical approximate Bayesian computation (hABC) approach to simulate demographic histories based on the aSFS, including hyperparameters to determine congruence of demographic parameters whilst allowing species-specific parameters to vary. The aSFS is an extension of other variants of the SFS which combines multiple SFS from separate populations, sorting them within each frequency bin in descending order across all taxa (and thus not preserving the original order of each SFS). Since the aSFS requires equal

sample sizes across the combined taxa (Xue & Hickerson 2015), individual SFS for all populations were down-projected to the smallest population sample size (n = 8 haploid samples) using the easySFS Python pipeline (github: isaacovercast/easySFS.git) and DaDi (Gutenkunst *et al.* 2010). A single model of exponential growth followed by exponential decline was applied to all populations using broad uniform priors (Figure 3.4), based on results from individual coalescent simulations (see Results). Broad priors for contemporary *Ne* (10K – 100K) and rate of population size changes (ε1 = 5 – 20 and ε2 = 0.1 – 1 for exponential decline and growth, respectively) were set to capture the array of results from individual FastSimCoal2 models. The timing of the change in population size patterns (from exponential growth to decline) was set (τ) using a broad prior of 10 – 100 Kya to account for difficulties in detecting the timing of events. Population contractions were considered asynchronous if they occurred more than 1,000 years apart (β = 1,000). We first tested the proportion of co-contracting taxa (ξ = 0.167 – 1), and then fixed this hyperparameter to better explore the remaining parameters. A “leave-one-out” approach using 50 pseudo-observed datasets was used to generate a confusion matrix, with the most likely proportion of co-contracting taxa determined using the top 1,500 simulations and Bayes Factors. Parameters were estimated using 1.5 million coalescent simulations and posterior distributions estimated using the top 100 simulations and the *abc* R package (Csilléry *et al.* 2012). We further tested whether demographic syndromes were broadly consistent across populations by comparing the posterior distributions for *Ne* and bottleneck strength (ε).

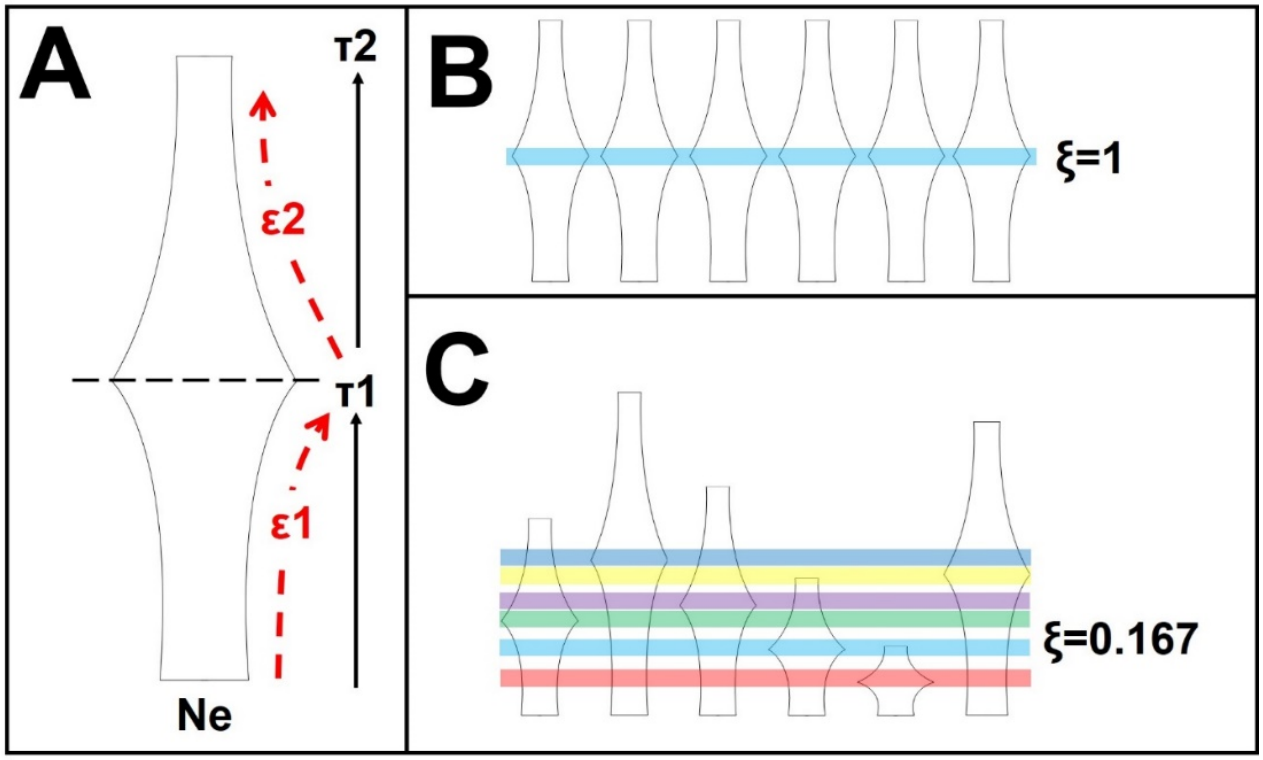


Figure 3.4: Diagrammatic representations of codemographic models within Multi-DICE. **A**) Example of parameters within a single taxon of the demographic model, containing two population size changes over time. A single broad prior distribution is set for parameters across all taxa, but posteriors may vary freely across taxa within the model. Arrows for τ parameters indicate direction backwards in time (i.e. τ1 is more recent in the past compared to τ2). **B)** An example of a fully synchronous model, when the hyperparameter ξ (the proportion of taxa belonging to a single synchronous event) is one. All populations share the timing of the population size change event, within the range of buffer β. **C**) An example of a fully asynchronous model, when ξ = 1/6 taxa per event. All populations have different timings of events, with each event >β generations from one another. The hyperparameter allows intermediate variations of ξ, which are not demonstrated here. We used a set of models to first determine the most likely value of ξ, before fixing this hyperparameter to better explore the remaining parameters. All other parameters and priors were kept constant between the two steps.

Contemporary and Paleoclimatic Environmental Modelling

SDMs were estimated using an ensemble modelling approach within the *R* package biomod2 (Thuiller *et al.* 2009). We estimated SDMs for both species across eleven time

slices ranging from contemporary conditions to the Pliocene using the PaleoClim database (Brown *et al.* 2018). This database contains 19 bioclimatic variables obtained from WorldClim (Hijmans *et al.* 2005) and the HadCM3 circulation model (Gordon *et al.* 2000) to reconstruct historical climatic conditions. A total of 6,021 *N. australis* and 852 *N. obscura* occurrences were collated from a combination of sampled sites within this and past studies (Brauer *et al.* 2013; Buckley *et al.* 2018; Unmack *et al.* 2013), as well as from the Atlas of Living Australia (http://www.ala.org.au). We removed geographic outliers based on known species distributions (P. Unmack, pers. comm.) and occurrences with low geographic resolution (>30km) and filtered this dataset to a single occurrence per environmental raster cell reduce to impact of spatial autocorrelation (Elith *et al.* 2011), resulting in final datasets of 1,021 and 163 observations for *N. australis* and *N. obscura,* respectively.

Highly correlated environmental variables ($r < |0.8|$) were removed from the dataset based on a Pearson's correlation test in SDMToolbox (Brown *et al.* 2017), resulting in eight bioclimatic layers (Appendix 19). These were annual mean temperature (Bio1), mean diurnal range (Bio2), isothermality (Bio3), temperature seasonality (Bio6), mean temperature of the wettest quarter (Bio8), mean temperature of the driest quarter (Bio9), annual precipitation (Bio12) and precipitation seasonality (Bio15). For the three oldest (Late Pleistocene-Pliocene) time periods, Bio2, Bio3 and Bio6 were unavailable and thus not included within the projections. SDMs were estimated using the MaxEnt model (Phillips *et al.* 2017), random forest (RF) and a generalised linear model (GLM), and an ensemble model generated per time period using the weighted mean of all models. All models were evaluated using both the area under the receiver operating curve (ROC) and the true skill statistic (TSS). This combination of models and evaluation methods was chosen to avoid implicit biases associated with any single method alone (Barbet-Massin *et al.* 2012). We

quantitatively assessed the relative stability of species distributions over time by estimating the mean and standard deviation of suitability across all time periods within each species. Differences in distributional ranges between species across time were estimated by converting ensemble and individual model SDMs to binary presence-absence maps using a minimum presence threshold based on the minimum suitability of the top 90% of putative occurrences under current conditions (Appendix 20) and calculating the number of putative presence cells.

Results

Bioinformatics

The combined sequencing runs returned a total of 86.14M reads for *N. obscura* (average of 2.27M per sample) and a total of 120.14M reads for *N. australis* (average of 2.18M per sample). We obtained 21,051 ddRAD loci containing 53,334 filtered SNPs for *N. obscura* and 19,428 ddRAD loci containing 69,264 filtered SNPs for *N. australis*, with low missing data in both alignments (Appendix 21)**.** Genetic diversity differed remarkably between the two species, with allelic richness, gene diversity, nucleotide diversity and number of SNPs per population alignment being significantly higher ($p \leq 0.01$) in *N. australis* (Table 3.2).

Table 3.2: Population genetic summaries of *N. australis* and *N. obscura*. Populations with n ≤ 3 were excluded due to low sample size. Where applicable, values are reported as means of means across all loci ± standard deviation under rarefaction (n = 4 samples) using the species-wide alignment SNPs. For Lake Alexandrina populations (NauALE and NobCHI), gene diversities were also calculated for individual population alignments and rarefaction of n = 15 samples to take advantage of larger sample sizes (reported second). Totals are reported as species-level averages across the alignment(s), excluding SNPs which refers to the total number of SNPs in the species-wide alignment.

Alignment			**Species-wide**		**Individual populations**		
Species	**Population**	**N**	**AR**	**Hs**	**Ne**	**Π**	**SNPs**
N. australis	NauANG	5	1.138 (± 0.339)	0.029 (± 0.106)		1.59×10^{-4}	2,446
	NauALE	5 20	1.354 (± 0.480)	0.062 (± 0.165) 0.179 (± 0.142)	71.3 [35.9 – 520.9]	1.23×10^{-3}	1,198
	NauGRG	5	1.123 (± 0.313)	0.035 (± 0.108)		8.84×10^{-5}	5,282
	NauMRG	5	1.116 (± 0.312)	0.037 (± 0.119)		1.62×10^{-4}	3,835
	NauGRF	5	1.092 (± 0.285)	0.027 (± 0.106)		7.84×10^{-5}	2,969
	NauBAR	4	1.240 (± 0.428)	0.061 (± 0.145)		2.02×10^{-4}	5,084
	NauMG	4	1.216 (± 0.413)	0.056 (± 0.138)		1.57×10^{-4}	5,357
	NauGCH	4	1.181 (± 0.387)	0.047 (± 0.136)		1.88×10^{-4}	4,272
	Total	56	1.241	0.0439		2.83×10^{-4}	17,389
N. obscura	NobCHI	5 15	1.151 (± 0.347)	0.038 (± 0.117) 0.136 (± 0.108)	14.9 [4.6 – 652.2]	1.69×10^{-4}	721
	NobMEC	5	1.030 (± 0.166	0.011 (± 0.069)		2.11×10^{-5}	1,002
	NobMCL	4	1.100 (± 0.290)	0.026 (± 0.094)		4.18×10^{-5}	1,633
	NobMRG	4	1.058 (± 0.234)	0.020 (± 0.093)		3.23×10^{-5}	1,454
	NobGCL	5	1.070 (± 0.250)	0.019 (± 0.088)		2.49×10^{-5}	1,350
	Total	33	1.168	0.0237		5.78×10^{-5}	15,715
T-test			T = 2.38 ($p < 0.04$)	T = 2.93 ($p = 0.01$)		W = 56 ($p = 0.01$)	T = 4.63 ($p < 0.01$)

N = total number of samples per population. AR = rarefied allelic richness. Hs = rarefied gene diversity. Ne = effective population size, with 95% confidence intervals estimated by jack-knifing in square brackets. π = nucleotide diversity. Populations means across species were compared using either an unpaired two-samples t-test (T) or Wilcoxon rank test (W).

Phylogenetic Analysis

Phylogenetic analysis of both datasets returned a highly supported phylogenetic tree for each species. Site concordance factors broadly supported these patterns, although gene concordance factors were low across both trees (Appendix 22) – this is not unexpected when gene trees are estimated from short and relatively uninformative individual loci (Minh *et al.* 2020a). For southern pygmy perch, the topology of this phylogenetic tree mirrored the geographic range of the samples, with a clear division between the MDB lineage and the

coastal lineage within the tree (Figure 3.5A). Within the coastal clade, populations diverged in a longitudinal manner, with eastern-most populations as the most recently diverged. In contrast, the phylogenetic tree for *N. obscura* did not demonstrate the same precise patterns, with populations not diverging in an exactly longitudinal manner. However, this was driven by a single outlier population and may instead be an artefact of short ancestral branches within the tree.

Migrations inferred by TreeMix indicated a greater number of events within *N. australis* (four) than *N. obscura* (one event; Figure 3.5B; Appendix 23). Within *N. australis*, migrations were inferred both across populations of the coastal lineage as well as into the MDB. The strongest migrations were between eastern coastal populations, and from the ancestor of the westernmost coastal population into the ancestor of the MDB lineages. For *N. obscura*, the single migration inferred suggested historical gene flow from the easternmost population to a more central population. Trees and migration edges for both species were well supported by covariance matrices, with low pairwise residuals (Appendix 24) and standard errors of < 1 for any given population for both species (Appendix 25).

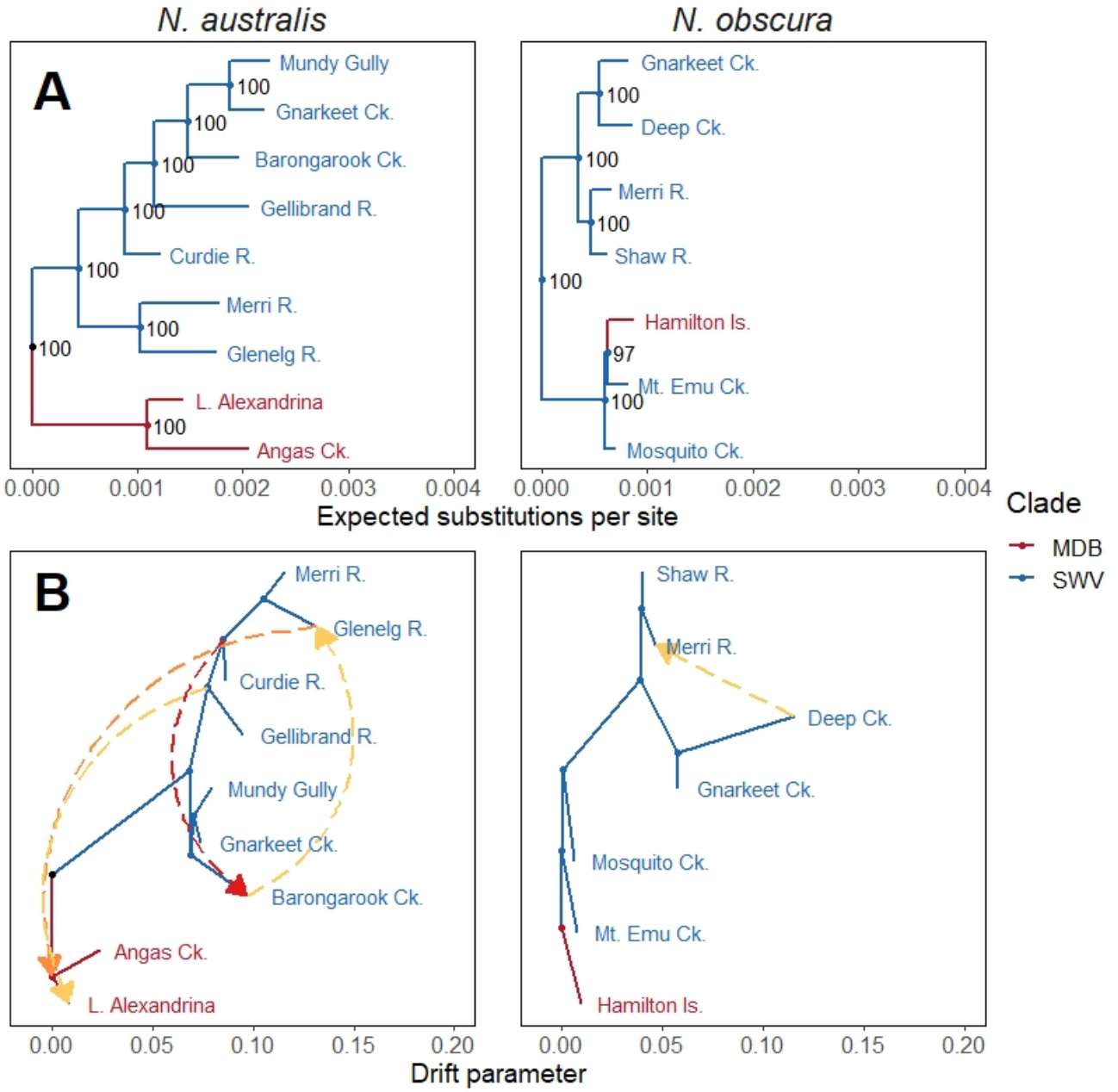


Figure 3.5: Phylogenetic histories and migration patterns in *N. australis* and *N. obscura*. **A:** ML phylogenetic trees estimated with RAxML based on ddRAD loci. All populations were reciprocally monophyletic, so clades were collapsed down to the population level for simplicity (full trees are shown in Appendices 26 and 27). Both trees were rooted using *N. vittata* as the outgroup, which was dropped for visualisation. Node values show bootstrap support from 1,000 RELL bootstraps. Branch colours indicate the drainage basin of origin for each population or clade. **B:** Best supported ancestral migration patterns inferred using TreeMix based on SNP datasets. All displayed migrations were statistically significant ($p < 0.05$). Arrows denote the direction of inferred migrations, with the colour indicating their relative weights.

Comparative Demography

Stairway plots demonstrated broadly similar demographic histories across the two species, with most populations relatively stable or declining slightly over the last 1 Mya (Figure 3.6A). Populations within both species demonstrated variable demographic histories, although populations of *N. australis* appeared generally more stable over time. Both Lake Alexandrina populations (NobCHI and NauALE) showed significant historical increases in *Ne* >200 Kya, and long-term stable population sizes following this expansion.

Most populations chosen for comparative analysis demonstrated fluctuating demographic histories (Figure 3.6B), with a period of pre-LGM growth followed by a post-LGM decline (Figure 3.7). However, the eastern *N. obscura* population (NobGCL) contrasted this pattern, with a model of low but constant population size more supported than other demographic histories (Model 3). Strong post-glacial declines were present in Lake Alexandrina populations of both species, with weaker declines in the more eastern population pairs.

A confusion matrix suggested that the co-demographic model was more likely to infer fully synchronous ($\xi = 1$) or fully asynchronous ($\xi = 0.167$) co-contractions over intermediate proportions of taxa (Appendix 28). Despite this, Bayes Factors supported a fully synchronous model over more asynchronous models, and so ξ was fixed to 1 to better explore other parameters (Figure 3.6C). Contemporary population sizes were inferred to be relatively small across all populations with relatively weak post-glacial bottleneck strength (Figure 3.6D; Appendix 30). These bottlenecks were similar in magnitude across populations, as indicated by low values of the dispersion index (Figure 3.6E). However, Multi-DICE did not recover the same timing of the bottleneck, possibly due to relatively low resolution within the aSFS (Figure 3.6F). Overall, these results support a widespread concordant bottleneck across the six co-distributed populations.

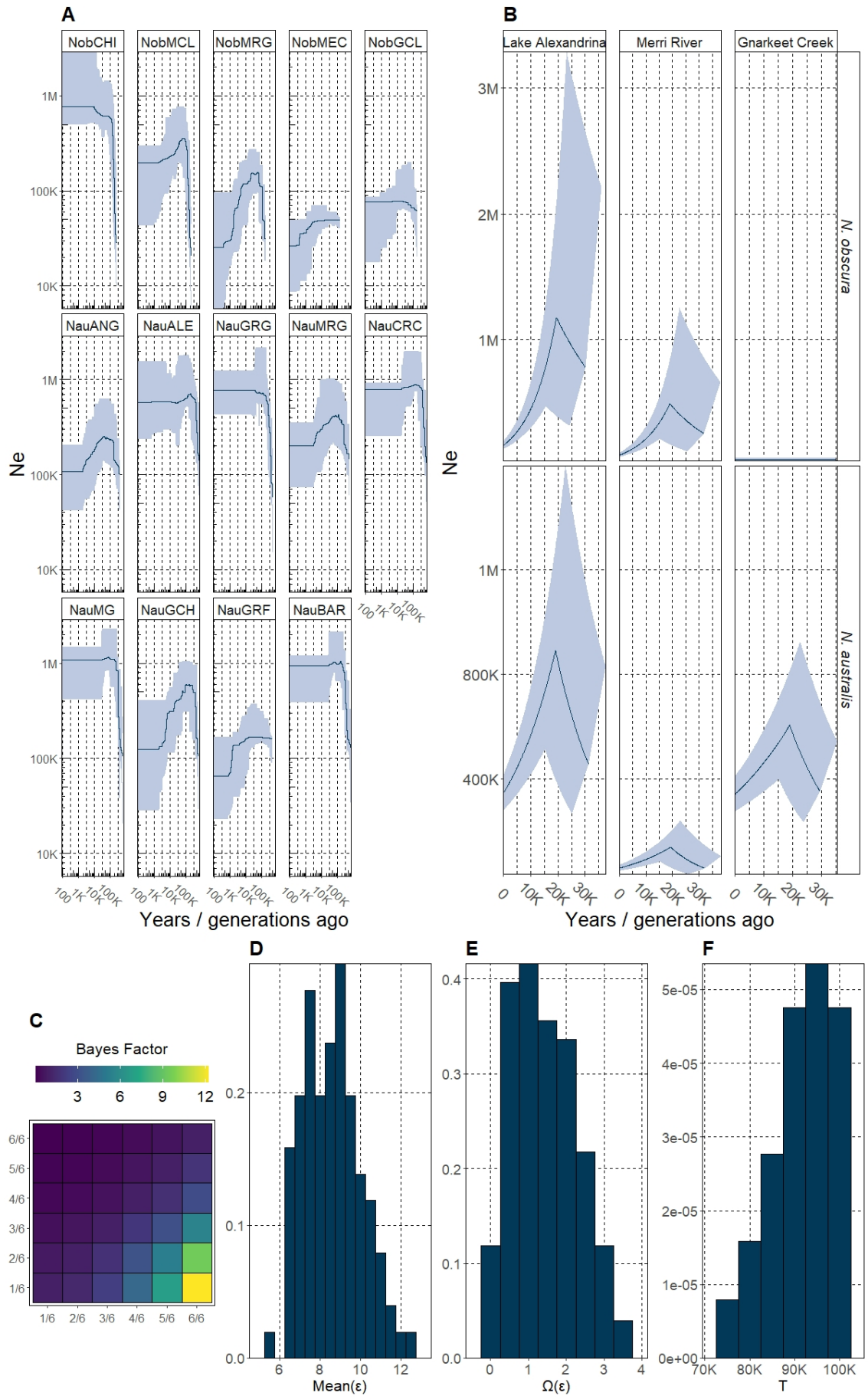

A
NobCHI
NobMCL
NobMRG
NobMEC
NobGCL
NauANG
NauALE
NauGRG
NauMRG
NauCRC
NauMG
NauGCH
NauGRF
NauBAR
1M
100K
10K
Ne
100
1K
10K
100K
Years / generations ago
B
Lake Alexandrina
Merri River
Gnarkeet Creek
3M
2M
1M
800K
400K
N. obscura
N. australis
0
10K
20K
30K
Ne
Years / generations ago
C
Bayes Factor
3
6
9
12
6/6
5/6
4/6
3/6
2/6
1/6
D
Mean(ε)
E
Ω(ε)
F
T
0e+00
1e-05
2e-05
3e-05
4e-05
5e-05
70K
80K
90K
100K

Figure 3.6: Demographic histories of *N. australis* and *N. obscura* populations. **A:** Stairway plot reconstructions of demographic history. Top row = *N. obscura* populations; bottom two rows = *N. australis* populations. Individual populations are arranged longitudinally (from westernmost to easternmost) within each species. Both axes are presented in $\log_{10}$ scale. **B:** Most likely individual demographic histories for co-occurring populations over the Pleistocene, simulated using FastSimCoal2. Dark blue lines indicate mean *Ne* over time, calculated based on the means of current *Ne*, rates of change and timing of switching rates. Shaded areas indicate 95% confidence intervals, similarly based on the 97.5% and 2.5% probability estimates for the combination of parameters. **C:** Bayes Factor matrix of the proportion of populations showing synchronised bottlenecks (ξ) within a co-demographic model using Multi-DICE. Each cell indicates the Bayes Factor comparing the model in the column with the model in the row, with brighter colours indicating greater support for the column. **D:** Posterior distribution of mean bottleneck strength (ε) across all six populations. **E:** Posterior distribution of dispersion index of bottleneck strength (Var(ε)/Mean(ε)) across all six populations. **F:** Posterior distribution of the timing of the bottleneck event, in generations (equal to years).

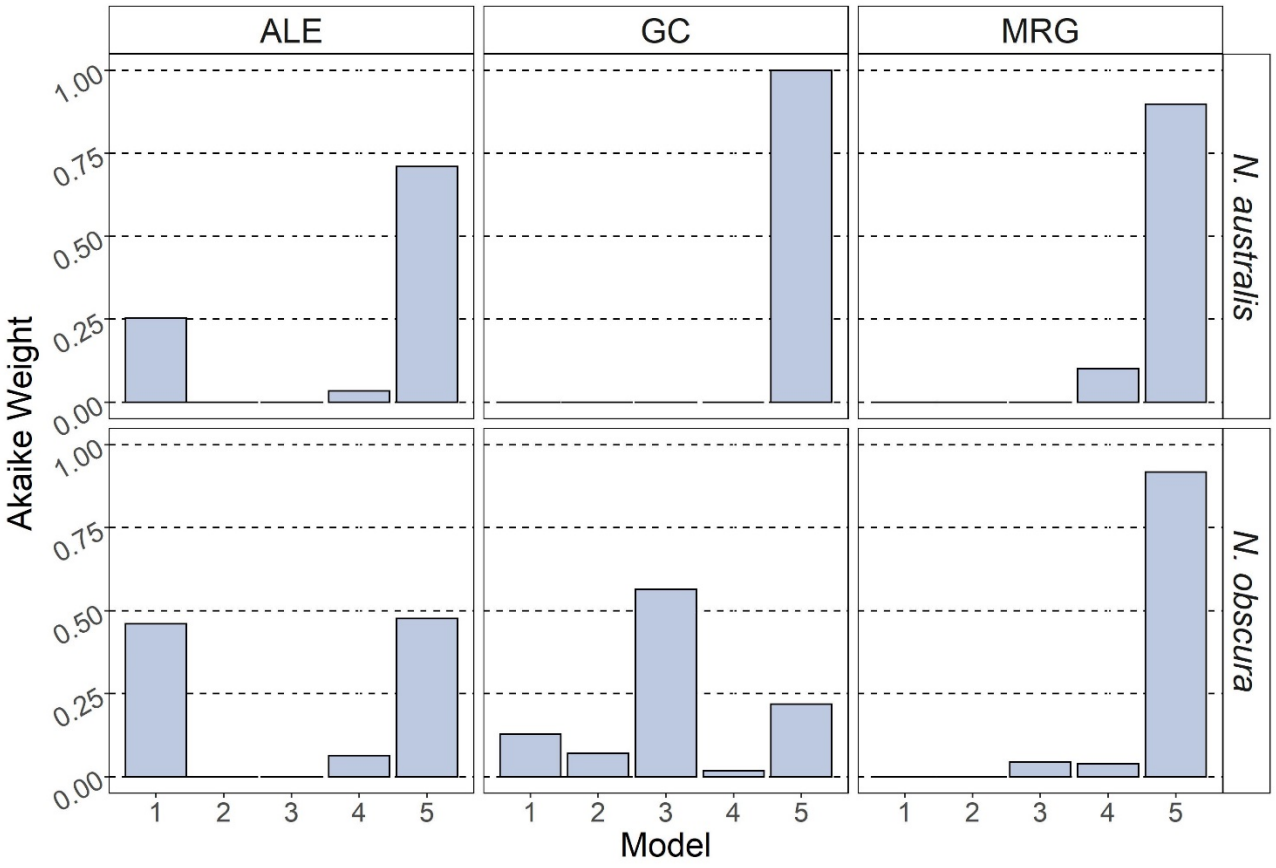


Figure 3.7: Akaike weights for demographic syndrome models using FastSimCoal2. Individual demographic models relate to the descriptions and numbering in Figure 3.3.

Species Distribution Modelling

Comparing the SDMs of the two species indicated much greater maximum distribution and variation in distributional range in *N. australis* than in *N. obscura*. *Nannoperca obscura* demonstrated long-term isolation to a relatively small region of southwest Victoria, whilst *N. australis* demonstrated a significant range expansion event throughout the early Pleistocene with a more recent contraction in the Holocene (Figure 3.8). Despite these differences, both species maintained a shared climatic refugium in southwest Victoria, highlighted by a region of high mean suitability in both species (Figure 3.9B).

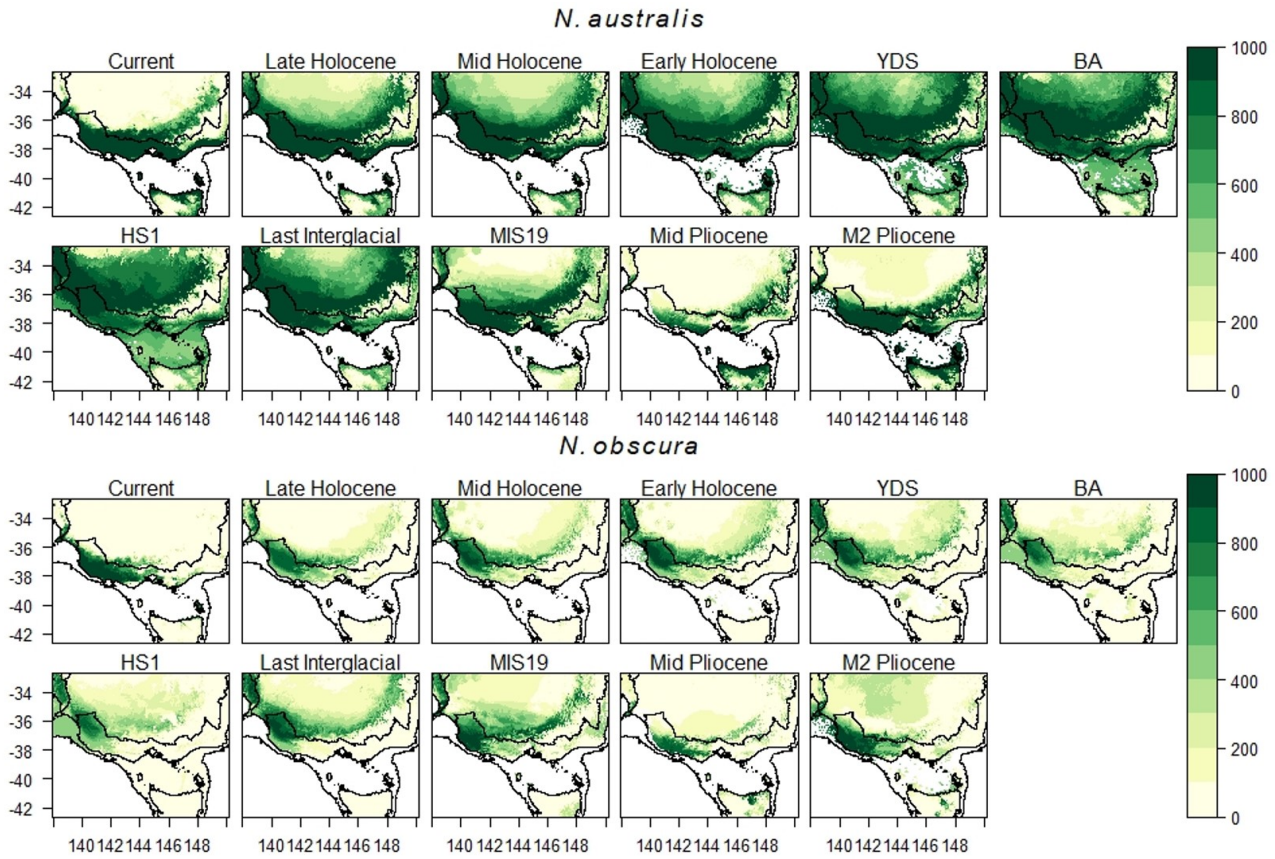


Figure 3.8: Ensemble SDMs projected from current conditions to the mid-Pliocene for both *N. australis* and *N. obscura*. SDMs are based on 1,021 and 163 occurrences, respectively, and 10 environmental variables (9 bioclimatic variables + elevation).

Additionally, the broader distribution of *N. australis* over time was further exemplified by the total presence area estimations. *Nannoperca australis* demonstrated significantly larger distributions throughout the Pleistocene compared to the relatively stable range of *N. obscura*, with the former spanning a range approximately twice as large as the latter during the mid-Pleistocene (Figure 3.9A). These patterns were similarly reflected within the standard deviations across timeslices per species, with *N. australis* showing much higher variation over a larger area (Figure 3.9C). While there was significant variation in estimated area across the different methods, ensemble models approximately captured the mean of

all models. Comparisons across the different methods indicated that RandomForest was more conservative in estimating area (Figure 3.9A).

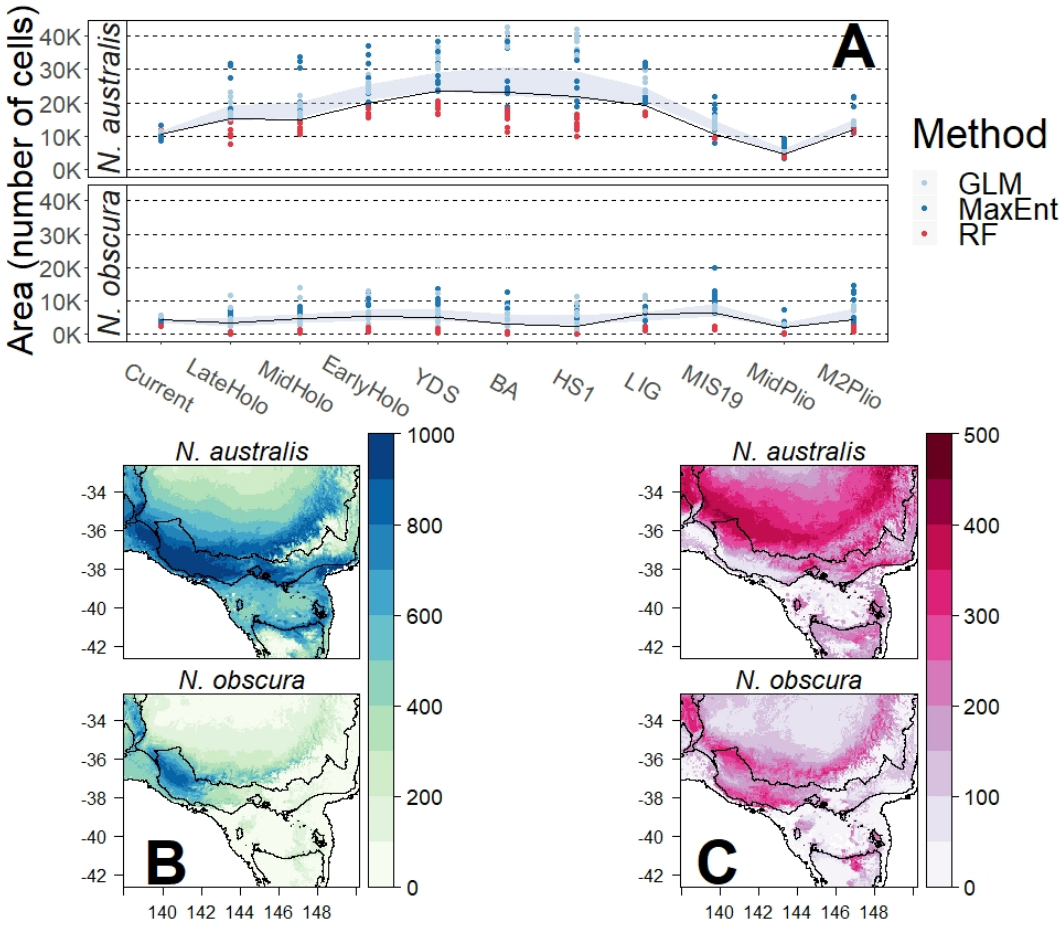


Figure 3.9: Comparisons of summaries of distributional changes over eleven time periods spanning the Plio-Pleistocene. Top row = *N. australis;* bottom row = *N. obscura*. **A:** Distribution extent per species. Area estimates were based on 90% minimum training presence threshold calculated separately for each individual model and the ensemble (*n* = 28 models total) under current conditions to convert each model to a binary presence-absence raster. From this, the total number of inferred presence cells were counted. Individual models are indicated by points, with SDM method indicated by colour. The 95% confidence interval of all 27 individual models is shown by the pale blue ribbon. The ensemble model is represented by a solid black line. **B:** Average (mean) cell suitability across all time periods. **C:** Variation (standard deviation) in cell suitability across all time periods. Overall, these results demonstrate the consistently more widespread distribution of *N. australis* compared to *N. obscura*, with large variation owing to a range expansion throughout the early Pleistocene.

Discussion

Our results demonstrate how spatial variation in demographic history may drive species-wide discordant responses to past climatic changes, even when local-scale impacts are concordant and species' ecological traits are similar. Specifically, we show that in a region that acted as a shared climatic refugium for two co-distributed and ecologically similar freshwater fishes, populations experienced concordant post-glacial bottlenecks. Contrastingly, species distributions varied dramatically over history, with a significant expansion in the range of *N. australis* during glacial maxima but a stable distribution in *N. obscura*. Towards the edges of this refugium demographic histories decreased in concordance, suggesting that habitat fringe effects prevented one species (*N. obscura*) from expanding its distribution under more favourable climatic conditions. Together, our findings demonstrate the importance of intraspecific, population-level characteristics such as standing genetic diversity in driving species-wide responses to climate change.

Phylogeographic History of Southeast Australia

The temperate zone of southeast Australia has undergone significant environmental change since the Pliocene, owing to a combination of continent-wide aridification (Hawlitschek *et al.* 2012), eustatic changes (Chapple *et al.* 2011) and major hydrological rearrangements (Waters *et al.* 2019). These various aspects likely had significant impacts on the persistence and connectivity of freshwater lineages across the region (Chapter 2). This was supported by the high level of phylogenetic structure within *N. australis,* and inferred migration pathways correspond well to those previously suggested through ancient hydrological conduits (Waters *et al.* 2019).

Although phylogenetic patterns in *N. obscura* did not directly match the longitudinal gradient of populations (namely, NobMEC fit within a clade of more western samples), previous phylogenetic analyses using allozymes and mitochondrial DNA showed a similar pattern (Hammer *et al.* 2010). This disjunction was previously attributed to potential historical connections from Mount Emu Creek into more western populations (Brauer *et al.* 2013), although short branch lengths and low genetic diversity across the species may also indicate incomplete lineage sorting as a factor (Degnan & Rosenberg 2009; Liu *et al.* 2015). Indeed, the difference in lineage sorting between the two species may be intrinsically linked to their differences in overall genetic diversity. Greater genetic diversity in *N. australis* provides more variation to differentiate lineages, leading to longer branches and easier lineage sorting compared to the short internal branches of *N. obscura*. For both species, we denote two major clades: one of MDB populations and another of coastal populations in *N. australis*, as previously suggested (Unmack *et al.* 2013), and two clades each containing two previously identified ESUs in *N. obscura* (Hammer *et al.* 2010).

Within the SDMs, a region of southwest Victoria was highlighted as a climatic refugium for both species throughout the Plio-Pleistocene. This region was consistently identified as suitable habitat for both species across all time slices, highlighted by the high mean and low variation in suitability. Previous palaeontological work has suggested that much of southeast Australia was likely wetter and colder during glacial maxima, resembling rainforest habitat (Williams *et al.* 2018). High water availability, lower temperature and dense overgrowth all provide suitable habitat for pygmy perch species (Woodward & Malone 2002), supporting the identity of this stable refugium. Co-occurring populations within this shared refugium demonstrated highly congruent demographic histories at both more ancient (>1 Myr) and more recent (since the LGM) temporal scales. This concordance

is expected when ecological traits, habitat preferences and environmental stability are shared across the species in question (Zamudio *et al.* 2016). Although individual populations within each species demonstrated spatially variable demographic histories, comparisons across the two species showed similar patterns of *Ne* over time for most directly co-occurring populations.

Local Concordance but Widespread Discordance

While many studies inferring concordant responses to glacial climatic shifts report congruent scenarios of glacial bottlenecks (Provan & Bennett 2008), both pygmy perch species demonstrated temporally synchronous expansions during the LGM with post-glacial contractions across central populations. This is not unexpected given the higher degree of population connectivity and suitable habitat inferred under glacial conditions (Williams *et al.* 2018), which would have facilitated the growth and spread of populations across southwest Victoria. This is further supported by the importance of wetland and floodplain habitats for pygmy perch recruitment (Woodward & Malone 2002), a habitat type which would have been more abundant under these conditions. However, the degree of observed concordance reduced for pairwise populations that occurred closer to the edge of this shared refugium (Gnarkeet Creek), suggesting the species had discordant demographic responses at the fringe of the range. Similarly, phylogenetic patterns at the species-wide level varied between the two species, with clearer geographic sorting and historical migration across *N. australis* lineages compared to *N. obscura*.

Spatial variation in demographic history, and its concordance across taxa, may result from a number of different mechanisms (Aitken *et al.* 2008; Papadopoulou & Knowles 2016). Particularly for narrowly distributed species, edge-of-range effects acting on populations

occurring close to the ecological tolerance threshold of the species may result in highly divergent patterns of demographic history and genetic diversity compared to more central populations (Aitken *et al.* 2008; Bridle & Vines 2007; Eckert *et al.* 2008; Razgour *et al.* 2013). Similarly, populations occupying the edge of the distribution of a species often have reduced fitness and gene flow due to lower densities and asymmetrical gene flow (Aitken *et al.* 2008). By extension, the ecological range of species may be a strong factor driving discordance when particular locations are at the periphery of the distribution of one species, but not another.

Given the broad similarity in ecological traits between the two species and their co-occurring nature (Woodward & Malone 2002), it is unlikely that discordance in species-wide responses to past climatic changes is solely a result of different ecologies. However, it is important to note that there is some variation in microhabitat preference between the two species, with *N. obscura* limited to larger, lowland channels and floodplains whereas *N. australis* is also found in streams and dense swamps (Woodward & Malone 2002). These patterns suggest greater habitat specialisation in *N. obscura*, which might drive lower SGV (or result from it) and impede range expansions. Thus, we cannot completely rule out some role of ecology and its interactions with genetic diversity in driving discordant species responses. The lower genetic diversity in *N. obscura* could not be directly attributed to notable and widespread genetic bottlenecks, suggesting instead that the species suffered from a consistent pattern of being genetically depauperate. The reduced adaptive capacity and long-term SGV of *N. obscura* relative to *N. australis* is supported by an order of magnitude lower *Ne* over the last 5,000 generations in the former in a shared environment (Attard *et al.* 2016a). Combined, these factors suggest that long-term SGV may be a key

factor driving the temporally and spatially widespread discordance in response to Pleistocene climate changes.

A number of studies have suggested that adaptive responses, particularly in scenarios of range expansion (Prentis *et al.* 2008; Szűcs *et al.* 2017), are more frequently driven by soft sweeps of SGV than *de novo* novel mutations (Hermisson & Pennings 2005; Lai *et al.* 2019; Morris *et al.* 2018). While many studies focus on rapid adaptation from SGV in terms of invasive species colonising new habitats (Chen *et al.* 2018; Prentis *et al.* 2008), similar dynamics can be expected to play a role in range expansions of native taxa (Williams *et al.* 2019). In regards to range shifts across the Pleistocene, higher SGV may have predisposed *N. australis* to capitalise on the broadly wetter environment of glacial periods and subsequently expand. Similarly, the higher degree of historical connectivity present across now-isolated river drainages (Unmack *et al.* 2013) likely facilitated interpopulation gene flow, which may have further bolstered SGV and adaptive potential (Aitken *et al.* 2008; Bridle & Vines 2007; De Carvalho *et al.* 2010). The presence of interpopulation gene flow in *N. australis* may have also facilitated range expansion if locally adaptive alleles were transferred into edge populations (Krehenwinkel & Tautz 2013; Prentis *et al.* 2008). Contrastingly, a lack of long-term SGV within *N. obscura* may have prevented them from expanding under these conditions, leading to the species-wide discordance in responses to past climatic change. The spatial variation in concordance, with discordance occurring at the edge of the *N. obscura* pre-glacial refugium, supports this conclusion.

Implications of Demographic History for Conservation Management

Discordant species-wide responses to past climatic change may play an important role in contemporary genetic diversity and, by extension, current conservation efforts. Observed

genetic diversity in current populations can be thought of as the combination of contemporary and historical processes (Barrow *et al.* 2018), a factor which can strongly impact the resilience of species and populations long-term (Eizaguirre & Baltazar-Soares 2014; Frankham 2009). More specifically, historical bottlenecks can simultaneously drive contemporary inbreeding depression and low genetic diversity (Bouzat 2010; Dobrynin *et al.* 2015; Frankham 2009). Additionally, the parallels between historical range expansion scenarios and current reintroductions to conserve species demonstrates how historical processes may inform current practices (Szűcs *et al.* 2017). For example, the lack of adaptive capacity to drive range expansion in *N. obscura* may also contribute to the limited success of reintroduction programs for the species and their local extirpation from isolated edge populations within the MDB (Beheregaray *et al.* in press). Inferring aspects which might have historically prevented range expansion in endangered species can help to determine the success of future reintroductions, particularly in terms of adaptation from SGV (Szűcs *et al.* 2017).

Predicting Species Responses to Contemporary Climate Change

Understanding how, and which species, may be able to adapt under contemporary climate change remains a critical aspect of evolutionary biology (Healy *et al.* 2018; Krehenwinkel & Tautz 2013). Determining historical responses to past climatic changes, and the factors that drove those responses, may allow for prediction of traits important for particular aspects of climatic response (Healy *et al.* 2018). Typically, this framework has focused on understanding how ecological traits such thermal tolerance (Deutsch *et al.* 2008; Somero 2010), seasonality (Bradshaw & Holzapfel 2006) or dispersal ability (Travis *et al.* 2013) may underpin individual species responses to climatic change (Williams *et al.* 2008). However, demographic parameters such as population sizes, distribution area and population

structure are also critical components for species susceptibility to contemporary climate change (Pearson *et al.* 2014). Here, we demonstrate that intraspecific SGV may also be a critical component of whether species can respond to climatic changes, particularly in fringe populations at the edge of the distribution. This corroborates studies indicating that adaptive potential is largely driven by SGV prior to the origination of major selective pressure (Lai *et al.* 2019) and suggests that broad ecology alone may not be a strong predictor of species' ability to respond. Thus, understanding how the demographic history of individual population may predispose, or hinder, adaptive potential in species is an important component in future conservation management of threatened species. For species with low SGV, pro-active measures such as assisted gene flow and maintenance of effective population size may assist in their long-term conservation (Pavlova *et al.* 2017).

Conclusions

Long-term SGV drove discordance in the response of closely related and ecologically similar freshwater fishes to historical climate change, by facilitating range expansion and spread of one species but not the other. However, in the centre of a shared habitat refugium, demographic histories were concordant, suggesting that spatial variation in the degree of concordance is linked to the interaction of SGV and distribution edge effects. Together, this demonstrates the importance of the maintenance SGV for adaptive potential in response to climatic changes and the role of non-ecological traits in driving patterns of concordance or discordance.

Chapter 4: Long-term climatic stability drives accumulation and maintenance of divergent lineages in a temperate biodiversity hotspot

Sean J Buckley[1], Chris Brauer[1], Peter Unmack[2], Michael Hammer[3], Jonathon Sandoval-Castillo[1], Luciano B. Beheregaray[1*]

[1]Molecular Ecology Laboratory, College of Science and Engineering, Flinders University, Adelaide, SA 5001, Australia

[2]Institute for Applied Ecology, University of Canberra, ACT 2601, Australia

[3]Natural Sciences, Museum and Art Gallery of the Northern Territory, Darwin, NT 0801, Australia

This chapter is under preparation for submission to an as-yet decided journal, with Dr. Chris Brauer, Dr. Peter Unmack, Dr. Michael Hammer, Dr. Jonathan Sandoval-Castillo and Prof. Luciano Beheregaray as co-authors. It is currently in draft status and has not yet been reviewed by co-authors or formatted for a specific journal. I was responsible for the data collection, all analyses and drafting the manuscript. Jonathon Sandoval-Castillo assembled the southern pygmy perch genome and wrote the scripts for annotation and enrichment. Luciano Beheregaray provided comments on the manuscript and supervised the project.

Abstract

Anthropogenic climate change is likely to drive regional climate disruption and instability across the globe. This issue is likely to be exacerbated within biodiversity hotspots, both due to the greater potential for species loss but also to the possibility that endemic lineages might not have the evolutionary potential to respond to rapid climate change. Here, we assessed the role of climatic stability on the accumulation and persistence of lineages in a freshwater fish group (genus *Nannoperca*) endemic to the SWWA biodiversity hotspot. Using a combination of genomic and environmental approaches, we investigated population divergence, phylogenetic relationships, species delimitation and ecological niche modelling in SWWA pygmy perches. We identified divergent lineages, including at least two cryptic species, that showed long-term patterns of isolation and persistence. We also developed a novel approach that estimates functional enrichment of diagnostic loci between putative species. This approach pointed to reproductive isolation between putative species potentially linked to genes associated with chromosomal segregation, cytokinetic functions and metabolic processes. Together, these findings suggest a revision of taxonomic and conservation status within *N. vittata* is required to adequately maintain unique evolutionary lineages into the future. Our results demonstrated strong divergence within the group and highly stable modelled distributions, highlighting the role of climatic stability in allowing the persistence of isolated lineages in the SWWA. This biodiversity hotspot is under compounding threat from recent and ongoing climate change, and habitat modification, which may further threaten previously undetected cryptic diversity across the region.

Introduction

Global biodiversity is increasingly threatened by anthropogenic climate change, with contemporary rates of extinction substantially higher than much of geological history (Le Roux *et al.* 2019). As new climatic regimes exceed the ecological tolerances of species,

organisms must shift their ranges to track suitable climate (Bridle & Vines 2007), adapt to rapidly changing selective environments (Healy *et al.* 2018), or risk extinction (Pearson *et al.* 2014). However, predicting species responses to climate change is challenging due to the complex nature of adaptive resilience (Waldvogel *et al.* 2020) and to spatial variation in climate change impacts (Pimm 2009).

The impacts of anthropogenic climate change are expected to be particularly exacerbated within regions that experienced long-term climatic stability due to a number of convergent factors (Harrison & Noss 2017). Firstly, species that have not been exposed to major climatic changes over their evolutionary histories are less likely to have evolved traits allowing adaptive responses to contemporary climate (Jansson 2003; Sandoval-Castillo *et al.* 2020). For example, species occupying regions with low temporal climatic variation show lower dispersal ability (Sandel *et al.* 2011), greater ecological specialisation (Dynesius & Jansson 2000) and lower thermal tolerance (Addo-Bediako *et al.* 2000) than species from more historically variable regions. This reflects the inverse outcome of the 'climate variability hypothesis', which states that populations and species that evolved in more climatically variable environments will possess greater adaptive potential to changing environments (Gutiérrez-Pesquera *et al.* 2016; Sandoval-Castillo *et al.* 2020). Under the strain of contemporary climate change, species in historically stable environments might thus not be pre-disposed to adapt to shifting climatic regimes. Secondly, variation in long-term climatic stability is often suggested to be a primary mechanism driving the spatial heterogeneity of biodiversity (Harrison & Noss 2017; Sandel *et al.* 2011). This pattern is particularly apparent in the high diversity of equatorial regions [the "latitudinal gradient": (Mittelbach *et al.* 2007; Rabosky *et al.* 2018)] and in hotspots of biodiversity (Habel *et al.* 2019) since both show relatively stable climatic histories. Over geological timescales,

regional climatic stability is predicted to promote the persistence and speciation of endemic lineages, driving higher cumulative diversity (Fjeldsaå *et al.* 1997; Sundaram *et al.* 2019). However, several alternative (but not exclusive) biogeographic mechanisms, relating to increased rates of speciation or decreased rates of extinction (Cardillo & Pratt 2013; Moritz *et al.* 2000), may also drive regional diversification (Cook *et al.* 2015; Rabosky *et al.* 2018). Lower extinction rates could be facilitated by biogeographic characteristics such as ancient environments or spatial area (Jansson 2003; Mittelbach *et al.* 2007), whereas increased speciation rates could result from allopatric speciation of narrow range endemic lineages (Barraclough *et al.* 2000; Goldberg *et al.* 2005). Thus, determining the relative role of climatic stability on biodiversity is important to determine extinction risk under climate change.

The combination of conservation concerns within highly biodiverse regions is epitomised within global 'biodiversity hotspots'. These regions are delineated by high species diversity and endemism (>1,500 endemic species), and the degree of habitat loss (>70% of primary vegetation; (Myers *et al.* 2000). A total of 36 global biodiversity hotspots have been proposed that contain more than half of all endemic plant and terrestrial vertebrate species diversity, but comprise only 2.5% of Earth's land surface (Habel *et al.* 2019). Together, these aspects make biodiversity hotspots regions of high conservation concern.

The disproportionate risk of extinction in biodiversity hotspots is also exacerbated by the amount of potentially undiscovered biodiversity globally (Joppa *et al.* 2011). This hidden diversity includes not only undescribed, morphologically-divergent lineages, but also cryptic species (Scheffers *et al.* 2012). Cryptic species represent unique lineages which are genetically divergent without associated phenotypic or ecological differentiation

(Beheregaray & Caccone 2007; Struck *et al.* 2018). Although the biogeographic mechanisms driving divergence and subsequent speciation are complex in nature (Edwards *et al.* 2005; Funk *et al.* 2006), the formation of reproductively-isolated cryptic species may be associated with both neutral and selective factors (Ravinet *et al.* 2018; Rundle *et al.* 2000). For example, population divergences in allopatry may lead to reinforcement in secondary contact if selection acts against maladaptive hybrids, resulting in reproductive isolation (Hoskin *et al.* 2005; Nosil & Crespi 2006; Rundle & Schluter 1998). The long-term maintenance of anciently isolated lineages may be important for the persistence of cryptic species and the accumulation of biodiversity (Bickford *et al.* 2007). Accurately delimiting species remains a critical component of conservation management at the taxon-specific (through revision of species classifications and associated legislation) and the regional (through identifying hotspots of cryptic diversity) scales.

Traditional approaches to species delimitation have relied on either few DNA markers, and primarily on mitochondrial DNA (Hickerson *et al.* 2010; Hurst & Jiggins 2005), or on simplistic measures of genetic differentiation (Lefébure *et al.* 2006). However, the pervasive nature of evolutionary heterogeneity across the genome (Cutter 2013; Edwards *et al.* 2016), owing to processes such as incomplete lineage sorting (Degnan & Rosenberg 2009), heterogeneous introgression (Wallis *et al.* 2017) and variable temporal resolution (Wang 2010) may impact the effectiveness of few genetic markers in delimiting species. The development of next-generation sequencing and new analytical approaches have greatly improved the ability to explore the relationship between the historical environment, cryptic species and diversification (Beheregaray 2008; Carstens *et al.* 2013; McCormack *et al.* 2013). Genome-wide sequences can be used to better capture the underlying demographic and evolutionary processes driving divergence and subsequent speciation (Dupuis *et al.*

2012; Fontaneto *et al.* 2015; Fujita *et al.* 2012). Timely, the improvement of complex modelling approaches such as the multi-species coalescent (Cutter 2013; Yang & Rannala 2014) allows for a more robust and empirical framework to delineating species (Coates *et al.* 2018), and inferring intrinsic aspects of reproductive isolation (Nosil *et al.* 2017). These approaches predominantly capitalise on genome-wide data to infer neutral signals of divergence and speciation, which might be later corroborated with adaptive (Stanton *et al.* 2019) or ecological (Edwards & Knowles 2014; Rissler & Apodaca 2007) information. However, genomic sequences maintain signals of selection (Via 2009), which could be theoretically explored within the same species delimitation framework (Seehausen *et al.* 2014). For example, functional enrichment of putative loci under selection provides an avenue for investigating traits underlying adaptive divergence between lineages (Brauer *et al.* 2016) and could be extended to species comparisons (Funk *et al.* 2016; Nosil *et al.* 2008; Wang *et al.* 2018). These combined developments allow for thorough investigation of the process of speciation and its relationship to the historical environment, an important component in understanding biodiversity hotspots.

The first, of only two, biodiversity hotspots to be declared for Australia was the temperate southwest, commonly referred to as the Southwest Western Australia hotspot or the Southwest Australia Floristic Region (Hopper & Gioia 2004; Myers *et al.* 2000). This region features high floristic diversity and endemism, with >8,000 species of plants recorded and >4,000 endemic species (Gioia & Hopper 2017). While species diversity appears lower for animal groups (~500 vertebrate species), relatively high rates of endemism (~23%) suggest that this pattern extends to non-plant taxa as well (Rix *et al.* 2015). The region's classification as a biodiversity hotspot is also driven by extensive habitat loss and deforestation, with ~70% of native land vegetation cleared primarily for agricultural

purposes (Habel *et al.* 2019; Monks *et al.* 2019). Despite the breadth of biodiversity, the SWWA features a simplistic landscape with limited topographic variation (Funnekotter *et al.* 2019; Wheeler & Byrne 2006), no major river drainage divides, stable geology (Hopper & Gioia 2004) and little climatic variation since the Pliocene (Spooner *et al.* 2011). Within the region, biogeographic subdivisions have been delineated primarily based on rainfall, including the High Rainfall Province (HRP; >600mm annual rainfall) and the Transitional Rainfall Province (TRP; 300 – 600mm annual rainfall; Hopper & Gioia 2004; Rix *et al.* 2015).

Several biogeographic mechanisms have been proposed to explain the high biodiversity of the SWWA, originating from the Mesozoic through to the Pleistocene (detailed in Rix *et al.* 2015). These include clades of ancient Gondwanan lineages that diverged throughout the Mesozoic until the late Eocene ~100 Mya (Gouws & Stewart 2007; Hopper *et al.* 2009); vicariantly-isolated mesic lineages that diverged from eastern Australian lineages with the formation of the Nullarbor Plain during the Miocene (14 – 16 Mya; Buckley *et al.* 2018; Crisp & Cook 2007; Rix & Harvey 2012); as well as *in situ* diversification (Gouws & Stewart 2007; Hopper & Gioia 2004). Despite the apparent lack of topographic or environmental barriers (Cowling & Lombard 2002; Hickerson *et al.* 2010), SWWA studies have demonstrated both interspecific (speciation) and intraspecific (phylogeographic structure) diversification primarily associated with late Miocene – early Pliocene aridification and contraction of mesic refugia (Byrne *et al.* 2011; Rix *et al.* 2015; Rix & Harvey 2012). Intraspecific phylogeographic structure has been described in several plants (Dalmaris *et al.* 2015; Millar *et al.* 2017; Nistelberger *et al.* 2014), narrow range terrestrial invertebrates (Cooper *et al.* 2011; Dadour & Johnson 1983; Harms 2018), and freshwater fauna (Edwards *et al.* 2008; Gouws & Stewart 2007; Gouws *et al.* 2006; Klunzinger *et al.* 2020; Unmack *et al.* 2012),

suggesting that traits such as limited dispersal and ecological specialisation may play a role in driving phylogeographic patterns. Spatial patterns of speciation in some lineages reflect intraspecific phylogeographic structure in others, implying similar underlying biogeographic mechanisms interacting differently with species traits in driving the patterns and degree of divergence (Rix *et al.* 2015). Understanding how these factors may drive diversification within the region is key to predicting how biodiversity may respond, or not, to anthropogenic climate change. If climatic stability played a major role in the persistence of lineages in the SWWA, then contemporary climate change may disproportionately threaten the region as individual isolated lineages struggle to adapt to novel and changing environments. However, if past environmental changes such as aridification drove divergence across the region, then further fragmentation might be expected under contemporary climate change as aridity and temperature intensifies within the region (Chapter 2). Predicting how species will respond to ongoing climate change requires an understanding of the environmental history of the region and how this interacts with their ecological traits.

Freshwater taxa are ideal models for investigating the role of climatic history on diversification given their limited dispersal capacity (Davis *et al.* 2018; Inoue *et al.* 2014; Thomaz *et al.* 2017) and reliance on constrained habitats, making them effective indicators of historical environmental change (Burridge *et al.* 2007; Murphy & Austin 2004). Freshwater species show a high propensity for cryptic diversity (Avise 2000; Gouws *et al.* 2006; Murphy & Austin 2004), enabling thorough investigations of the interaction of Earth history and diversification (e.g. Chapter 2; Beheregaray *et al.* 2002; Schultz *et al.* 2008; Waters *et al.* 2020). Within SWWA, the western (*Nannoperca vittata*) and little (*Nannoperca pygmaea*) pygmy perches demonstrate these traits, given their ancient origins (Buckley *et al.* 2018), limited dispersal capacity (Beatty *et al.* 2011), and role as ecological specialists

(Allen *et al.* 2020). These two small freshwater fishes are found in sympatry in some parts of their ranges. *Nannoperca vittata* is significantly more widespread and occurs throughout the HRP, whereas *N. pygmaea* occurs in only a handful of southern rivers of the HRP (Allen *et al.* 2020). Although no studies have assessed adaptation in response to climatic changes in either species, landscape genomic and ecological transcriptomic studies on a related pygmy perch species (*Nannoperca australis*) have demonstrated adaptive divergence associated to hydroclimatic variation (Brauer *et al.* 2016; Brauer *et al.* 2017). While both species show some similar ecological characteristics such as body plan and reproductive strategy, the two show marked differences in total body size, morphology, growth rate, salinity tolerance and reproductive timing (Allen *et al.* 2020). These factors, in conjunction with strong genetic differentiation in allozymes (Morgan *et al.* 2013), mitochondrial genes (Unmack *et al.* 2011) and genome-wide markers (Buckley *et al.* 2018), and with the lack of observed hybridisation with *N. vittata* (Allen *et al.* 2020), corroborates the identity of the recently described *N. pygmaea* as a distinct species. This is reflected by their strong evolutionary distinctiveness, with studies suggesting that the divergence between the two species occurred during the Pliocene ~4 million years ago (Figure 1.4; Unmack *et al.* 2011). In addition, cryptic species have been suggested within *N. vittata* based on ancient divergence times and coalescent modelling using mitochondrial DNA and two nuclear genes (Unmack *et al.* 2011), and genome-wide data (Buckley *et al.* 2018). Within this genomic study, *N. vittata* was proposed to consist of up to three distinct cryptic species, with a minimum of two based on the paraphyletic nature of the lineage with *N. pygmaea* (Figure 1.4). However, these studies were either lacking in genetic resolution (Unmack *et al.* 2011) or geographic sampling (Buckley *et al.* 2018), and as such a more thorough investigation of phylogeographic patterns across the region is required to refine hypotheses of species delineations.

Nannoperca pygmaea is currently listed as Endangered within the IUCN Red List (Allen *et al.* 2020; Beatty & Morgan 2019), as well as state (WA Biodiversity Conservation Act) and national (Environment Protection and Biodiversity Conservation Act) legislation, whilst *N. vittata* remains unprotected within Australian legislation (Table 1.1). However, the identification of cryptic species within *N. vittata* may significantly reduce population size and range of each species and therefore should drive a re-evaluation of this listing. Additionally, freshwater fishes within the SWWA are currently impacted by secondary salinization of rivers (Beatty *et al.* 2011) and land clearing leading to habitat alteration (Andrich & Imberger 2013). These threats are exacerbated by reduced rainfall and increased temperatures reported over the last few decades (Smith & Power 2014), which are predicted to worsen under anthropogenic climate change (Christensen *et al.* 2007; Pittock *et al.* 2008). Thus, freshwater taxa within the SWWA appear to be ubiquitously threatened and currently unlisted species may be more susceptible to extinction than their current conservation status suggests.

Here, we investigate the role of long-term climatic stability in the SWWA biodiversity hotspot on the diversification and accumulation of lineages in a group of freshwater fishes. We used high-throughput genomic (ddRAD-seq) data to characterise divergence across the lineage based on genetic differentiation, phylogenetic patterns and species delimitation. We expanded upon these methods by developing a novel approach that uses functional enrichment of fixed differences to investigate the potential role of selection on divergence and maintenance of species. Additionally, we used environmental data and modelling analyses to reconstruct species and lineage distributions since the Pliocene. We hypothesised that long-term climatic stability across the biodiversity hotspot would lead to

both the maintenance of divergent and isolated genetic lineages, including cryptic species, and long-term stable species distributions. Alternatively, if localized biogeographic events resulted in the diversification of pygmy perches in the SWWA, we expect that divergence events would coincide with ancient environmental changes and fluctuating lineage and species distributions over time. These divergence events may be shared across multiple independent lineages, including other co-distributed taxa, if vicariant forces drive diversification. To the best of our knowledge, this study represents one of the first regional investigations of phylogeography and species delimitation within the SWWA using genome-wide data (see also Binks *et al.* 2019), and the first such study for an animal taxon.

Methods

Sample Collection and Library Preparation

Sampling sites were selected to capture the full distribution of the lineages spanning the HRP, as well as a potentially relict population found in the northernmost extreme of the distribution in the TRP (Figure 4.1; Buckley *et al.* 2018; Unmack *et al.* 2011). This phylogeographic-driven sampling design effectively covers all the described taxonomic diversity of *Nannoperca* in the SWWA, including all previously suggested cryptic species (Buckley *et al.* 2018). A total of 25 *N. vittata* from six sites and eight *N. pygmaea* from two sites were sampled (Table 4.1). An additional four Yarra pygmy perch (*N. obscura*) were included as outgroup for phylogenetic analyses (Buckley *et al.* 2018). Specimens were collected using electrofishing, dip-, fyke- or seine-netting. Either the caudal fin or the entire specimen was stored at −80°C at the South Australian Museum, or in 99% ethanol at the Molecular Ecology Lab at Flinders University.

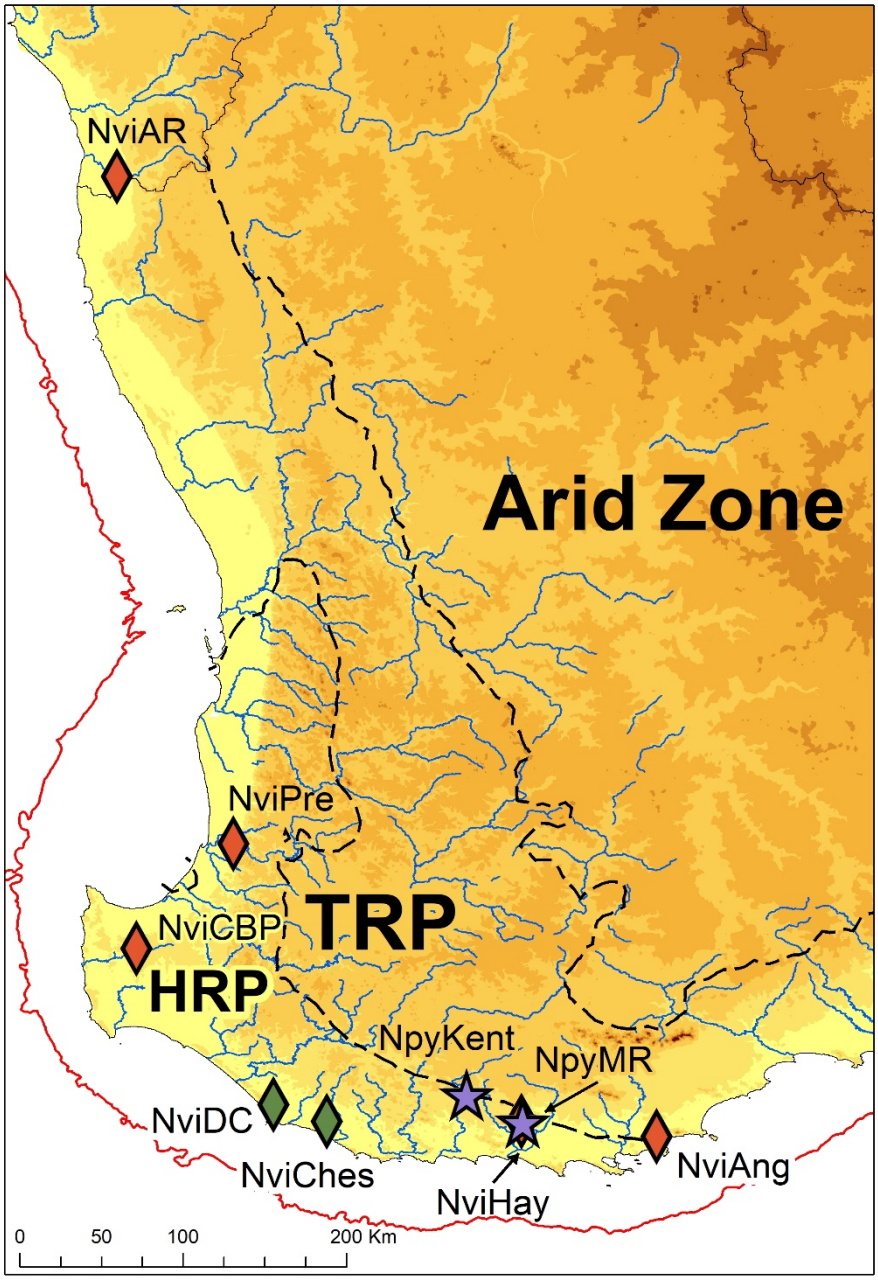


Figure 4.1: Map of sampling sites used within this study, and including major rivers and water bodies across the SWWA. Localities for *N. vittata* are indicated by diamonds and localities for *N. pygmaea* are indicated with stars. This includes one site of *N. vittata* (NviHay) and one site of *N. pygmaea* (NpyMR) which occur within the same river, and hence overlap on the map. Locality colours indicate putative species (see Results). Colouration of the map indicates elevation above sea-level. Solid black lines indicate major drainage basin boundaries and the dashed black lines indicate biogeographic province boundaries. Solid red line indicates the maximum landmass extent exposed during the LGM. TRP = Transitional Rainfall Province. HRP = High Rainfall Province.

DNA was extracted from muscle tissue or fin clips using a combination of a modified salting-out method (Sunnucks & Hales 1996) and a Qiagen DNeasy kit (Qiagen Inc., Valencia, CA, USA). Genomic DNA was checked for quality using a spectrophotometer (NanoDrop, Thermo Scientific), integrity using 2% agarose gels, and quantity using a fluorometer (Qubit, Life Technologies). The ddRAD genomic libraries were prepared in house following (Brauer *et al.* 2016). Of the 33 samples, eight were paired-end sequenced on an Illumina HiSeq 2000 at Genome Quebec (Montreal, Canada) as part of a previous phylogenomic study (Buckley *et al.* 2018). The remaining 25 samples were single-end sequenced on a single lane of Illumina HiSeq 2500 at the South Australia Health and Medical Research Institute in Adelaide.

Table 4.1: Locality data for *N. vittata* and *N. pygmaea* samples. Abbreviations described in the table refer to those used in further analyses and the sampling map, with *n* denoting the number of individuals sequenced per population. *Nannoperca obscura* samples were only included as an outgroup for phylogenetic analyses.

Species	Population	Abbreviation	Field code	Latitude	Longitude	N
N. vittata	Arrowsmith Rvr.	NviAR	F-FISHY6: DM146,147	-29.626506	115.174854	4
	Preston Rvr.	NviPre	F-E292: MA08-72NV	-33.3027	115.8176	4
	Angove Rvr.	NviAng	F-FISHY4: EV-1:5	-34.917	118.15	3
	Hay Rvr.	NviHay	F-FISHY6: PU09-37NV	-34.834	117.406	4
	Canebrake Pool, Margaret Rvr.	NviCBP	F-FISHY6: PU09-58NV	-33.880372	115.282701	2
	Doggerup Ck.	NviDC	F-FISHY6: PU09-49NV	-34.741106	116.038006	4
	Chesapeake Brook	NviChes	F-FISHx2: V**	-34.833	116.333	4
N. pygmaea	Mitchell Rvr.	NpyMR	F-FISHY6: ESP0900*	-34.836211	117.412321	4
	Kent	NpyKent	F-FISHx2B: SB14 – SB23	-34.6857	117.10272	4
N. obscura	Merri Rvr., Grassmere	NobMRG	PU02-111YPP	-38.275	142.542	4
Total			**10**			**37**

Filtering and Alignment

Sequences were demultiplexed using the 'process_radtags' module of Stacks 1.29 (Catchen *et al.* 2013), allowing up to 2 mismatches in the 6 bp barcodes. Barcodes were removed and sequences trimmed to 80 bp to remove low-quality bases from the end of the reads. Cut reads (forward only for paired-end) were aligned using PyRAD 3.0.6 (Eaton 2014), and further cleaned by removing reads with >5 bp with a Phred score < 20. Loci were retained if they occurred in at least ~80% of samples (30) within the dataset, resulting

in a final alignment of 19,426 ddRAD loci containing 81,251 SNPs. For SNP-based analyses (i.e. principal coordinates analysis (PCoA), species delimitation, historical migration and fixed differences), a single SNP per ddRAD locus was subsampled (18,177 SNPs) to reduce the impact of linkage disequilibrium (details in Brauer *et al.* 2016).

Population Divergence and Population Clustering

We first assessed population divergence by estimating pairwise uncorrected genetic distances (*p*-distances) between individuals using PAUP* 4 (Swofford 2002) based on the full sequences of all characterised ddRAD loci (i.e. 19,426 loci, see Results). Mean *p*-distances were calculated between all pairwise populations (i.e. sampled sites). Population clustering was also assessed using a PCoA of potentially unlinked biallelic SNPs (i.e. 18,177 SNPs, see Results) in dartR (Gruber *et al.* 2018), with individuals assigned to populations based on sampling locality. We also used dartR to determine the number of fixed differences between pairwise population comparisons, as well as hierarchically across the ML phylogenetic tree (see Results) by iteratively collapsing lineages. The SNPs were considered fixed differences at a threshold of 0.05 (i.e. >95% frequency in one population and <5% in the other) to account for sequencing errors or 'near fixation' (Gruber *et al.* 2018; Johnson & Slatkin 2007).

Phylogenetic Analysis

A ML phylogeny of the lineage was estimated using RAxML 8.2.11 (Stamatakis 2014) and the concatenated sequence alignment to determine evolutionary relationships and to guide species delimitation analyses. The ML phylogeny was estimated under the GTR+Γ model of evolution and 1,000 resampling estimated log-likelihood bootstraps (Kishino *et al.* 1990). A ML phylogeny with the alignment partitioned by ddRAD locus was also estimated, as well as estimating individual gene trees per locus, using IQ-TREE2 (Minh *et al.* 2020b). This

was done to account for the potential impact of genome-wide heterogeneity in evolutionary rates (Liu *et al.* 2015). Concordance between gene trees and the partitioned phylogeny was estimated using site and gene concordance factors (Minh *et al.* 2020a), with a summary species tree estimated using ASTRAL-III (Zhang *et al.* 2018) and assuming each population as an individual 'species' (Rabiee *et al.* 2019).

Species Delimitation and Divergence Time Estimates

The species tree and delimitation were estimated using SNAPP 1.5.0 (Bryant *et al.* 2012) within BEAST 2.6.1 (Bouckaert *et al.* 2019). We iteratively tested nine different scenarios of species composition, ranging from two to nine species (i.e. each population as a separate species) based on phylogenetic patterns (see Figure 4.2D). Given that SNAPP can resolve species identities with only a few thousand SNPs (Leaché *et al.* 2014), and to reduce computational time, we subsampled the alignment down to the two individuals with the lowest missing data per population and 5,000 randomly selected SNPs. We used broad priors with gamma distributions for speciation rate (λ = G(3 ,2.5)) and population sizes (θ = G(2.85, 955.27)) based on sequence divergence for all scenarios to cover possible parameter ranges. Mutation rate priors were left at their default settings (u + v = 1). Two separate chains were run per model to assess convergence of parameter estimates. Models were run for at least 10 million generations and/or until ESS > 200 was reliably achieved. Model traces were visualised using Tracer v1.5 (Rambaut & Drummond 2009) with the first two million generations discarded as burn-in. Species composition likelihoods were estimated using an AIC through Markov chain Monte Carlo analysis (AICM; Raftery *et al.* 2006). Although other methods such as reverse-jump MCMC via path-sampling or stepping-stone analysis are more widely used (Grummer *et al.* 2013), given the large size

of the dataset and number of models tested we opted to use AICM based on its reasonable performance and reduced computational demand (Baele *et al.* 2012).

Divergence times across the clade were inferred by using an extension of SNAPP described in Stange *et al.* 2018. We calibrated the oldest divergence in the tree using a normal distribution with mean of 9.27 Ma and standard deviation of 0.51 (95% CI = mean ± 1 Ma) based on previously obtained divergence estimates for all pygmy perches (Figure 1.4). We applied broader ranges around calibration nodes than previously suggested to accommodate potential variation in calibration age not captured by the methods in that study. The topology of the tree was fixed using the ML tree and the model run for one million generations. Confidence intervals of divergence times were inferred using TreeAnnotator v2.6 (Drummond & Rambaut 2019) and 95% posterior probabilities.

Historical Admixture and Introgression

We tested whether lineages found across the clade, especially for species inferred from the delimitation analyses, were historically genetically isolated or not using two different allele frequency-based approaches. First, we determined historical connectivity across populations based on the 18,177 SNP dataset using TreeMix (Pickrell & Pritchard 2012). The number of migrations within the model were iteratively increased from none to nine, with the fit of each model estimated by comparing the resultant covariance matrices and the overall tree likelihood. Additionally, we calculated the percentage of variation explained per migration model (Github = wlz0726/treemix), with the best supported number of migrations determined by the asymptote of likelihood. We also determined whether substantial introgression occurred between putative species (see Results) suggested by species delimitation approaches using an ABBA-BABA test (Martin *et al.* 2014) in D-Suite (Malinsky 2019). Introgression was determined assuming a pattern of divergence following the

phylogenetic tree (i.e. *N. vittata* [A] and *N. pygmaea* as sister lineages), and summarised based on Patterson's D (Patterson *et al.* 2012) and the fraction of introgressed alleles (f_4-ratio).

Functional Divergence of Cryptic Species

Given the lack of relevant physiological or ecological datasets to independently assess species identities inferred from species delineation methods, we developed a novel approach that uses functional enrichment of fixed differences. Under genetic drift alone, it is expected that fixed differences would be randomly distributed throughout the genome (Johansson *et al.* 2010). However, if selection played a role in the accumulation of fixed differences, whether as part of the speciation process or as a consequence of it, then fixed differences should be overrepresented in genes underlying important divergent traits. This was done by carrying out annotation of the ddRAD loci that contained SNPs fixed for alternate alleles between each of the three identified putative species and all other populations combined (i.e. one allele fixed within all populations of the putative species and the alternate allele fixed in all populations of *both* other putative species). While fixed differences may not represent true ecological characteristics, the SNPs used here were diagnostic for putative species and may underlie physiological traits associated with speciation and divergence. We identified and aligned 1,466, 968 and 3,847 ddRAD loci containing differentially fixed SNPs between *N. pygmaea*, Nvi[A], and Nvi[B] and all other populations (Appendix 31), respectively, to a newly assembled southern pygmy perch (*Nannoperca australis*) genome (N50 size = 10,039 bp and total scaffold length = 573,250,588 bp; Beheregaray *et al.*, unpublished data). Additionally, the full set of 19,600 ddRAD loci were also aligned to the *N. australis* genome as a reference for enrichment analyses. The 80 bp sequence of each ddRAD locus was mapped to the genome and 300

bp from both flanking regions were extracted. The resultant 680 bp were annotated using the Osteichthyes protein database in UniProt (UniProt Consortium) and blastx v 2.2.28, with a minimum alignment of 100 amino acids (300 bp), similarity > 70% and an e-value threshold of 10^{-6}. Where multiple alignments were possible, the alignment with the best e-value was selected in order to retain a single alignment per ddRAD locus. The annotation and gene ontology (GO) terms were extracted from the UniProt database using a custom script (Appendix 32). Enrichment of functions within *N. pygmaea, N. vittata* [A], and *N. vittata* [B] was assessed using a GO term analysis in the topGO R package (Alexa & Rahnenfuhrer 2020). Significant enrichment of GO terms ($p < 0.01$) was determined using the full ddRAD seq dataset as the background gene set and a minimum node size of five in Fisher's exact tests (e.g. Alexa & Rahnenfuhrer 2010; Sandoval-Castillo *et al.* 2018).

Species and Lineage Distribution modelling

SDMs were estimated using an ensemble modelling approach within biomod2 (Thuiller *et al.* 2009). SDMs were estimated across eleven time periods ranging from contemporary conditions to the Pliocene using the PaleoClim database (Brown *et al.* 2018). These time periods included the Late Holocene (4.2 – 0.3 Kya), Mid Holocene (8.326 – 4.2 Kya), Early Holocene (11.7 – 8.326 Kya), Younger Dryas Stadial (12.9 – 11.7 Kya), Bølling-Allerød (14.7 – 12.9 Kya), Heinrich Stadial 1 (17.0 – 14.7 Kya), Last Interglacial (~130 Kya), MIS19 (~787 Kya), mid Pliocene warm period (3.205 Mya), and M2 (~3.3 Mya) phases. Occurrence records for all lineages were obtained from a combination of sampled sites within this and past studies (Buckley *et al.* 2018; Unmack *et al.* 2011), as well as from the Atlas of Living Australia (http://www.ala.org.au). We further filtered the occurrence data to reduce the impact of spatial autocorrelation (Elith *et al.* 2011) by removing occurrences outside the known distribution of species (P. Unmack, pers. comm.; Allen *et al.* 2020) and

sampling a single occurrence per environmental raster cell. This resulted in a final dataset of 114 observations used within the SDMs.

Highly correlated climatic variables ($r < |0.8|$) were pruned from the dataset based on a Pearson's correlation test within SDMToolbox (Brown *et al.* 2017), resulting in eight climatic layers used across the models (Appendix 33). The final input variables were annual mean temperature (bio1), isothermality (bio3), mean temperature of the wettest quarter (bio8), mean temperature of the driest quarter (bio9), mean temperature of coldest quarter (bio11), annual precipitation (bio12), precipitation of the driest month (bio14), precipitation seasonality (bio15) and sea-level corrected elevation**.** For the three oldest time periods (Late Pleistocene-Pliocene), bio3 was unavailable and thus not included within the projections. We generated three replicates of 1,000 pseudoabsences randomly from the background >30km from occurrences to reduce the likelihood of generating false absences within habitable areas. Each dataset was replicated three times, with 80% of sites independently and randomly subset to train the model (*n* = nine datasets).

The SDMs were estimated using five separate models: classification tree analysis (CTA), a generalised boosting model (GBM), multiple adaptive regression splines (MARS), maximum entropy (MaxEnt) and random forest (RF) (*n* = 45 models total). These approaches cover a range of different statistical approaches (Buisson *et al.* 2010; Hao *et al.* 2019), and model averaging in an ensemble framework is expected to reduce bias associated with any single method (Chapter 3; Barbet-Massin *et al.* 2012; Marmion *et al.* 2009). Each model was evaluated using both the ROC and the TSS. Individual methods were checked for consistency by generating ensemble models per method under contemporary conditions and visually comparing results. Ensemble SDMs amalgamating the results of all SDMs per

time period were generated, excluding models with TSS < 0.7. The mean and standard deviation of suitability were calculated across ensemble models to assess long-term habitat suitability trends. Changes in distribution extent were quantified by converting SDMs per time period to binary presence-absence maps using two separate approaches: using the minimum occurrence value of the top 90% of training presences as the presence/absence threshold (Appendix 20), and using a maximised TSS approach (the default within biomod2). The area of each binary SDM was estimated by the number of presence cells in the raster.

Ensemble SDMs were then subdivided into individual lineage-specific distributions following the cost-distance approach described in Rosauer *et al.* 2015. Various LDMs were determined by assigning the seven *N. vittata* localities to three separate sets of LDMs (one with two lineages, one with three lineages, and one with four lineages) following the results from species delimitation and phylogenetics (see Results). Since this approach requires geographic (or environmental) distances between sites to accurately assign cells to lineages, *Nannoperca pygmaea* was not included as a lineage given its narrow current range and sympatry with *N. vittata*.

Results

Bioinformatics

The combined sequencing runs returned a total of 71.44M reads, with an average of 2.16M reads per sample. After quality control and alignment of sequences, a dataset of 19,426 ddRAD loci was obtained across all putative species. This alignment contained 18,177 unlinked biallelic SNPs and an average of 4.93 (±1.76)% missing data per sample (Appendix 34).

Population Divergence

Pairwise genetic distances and fixed differences inferred highly divergent structure across the clade. Pairwise population averages of genetic differences ranged from 0.03% to 1.55% whist the number of pairwise fixed differences ranged from 8 to 5,947 SNPs (Figure 4.2A). Progressively collapsing the phylogeny to compare fixed differences across the hierarchy of the tree resulted in high numbers of fixed differences (up to 3,261 SNPs for the deepest node; Appendix 35). A PCoA demonstrated three major and highly divergent clusters, with the two primary axes explaining a combined total of 77.6% of the variation (Figure 4.2B). Each of the three clusters represented the populations of *N. pygmaea, N. vittata* [A] and *N. vittata* [B]. Similar clusters could be observed within the pairwise *p*-distances (>0.5% between clusters) and by the number of fixed differences (>2000 between clusters).

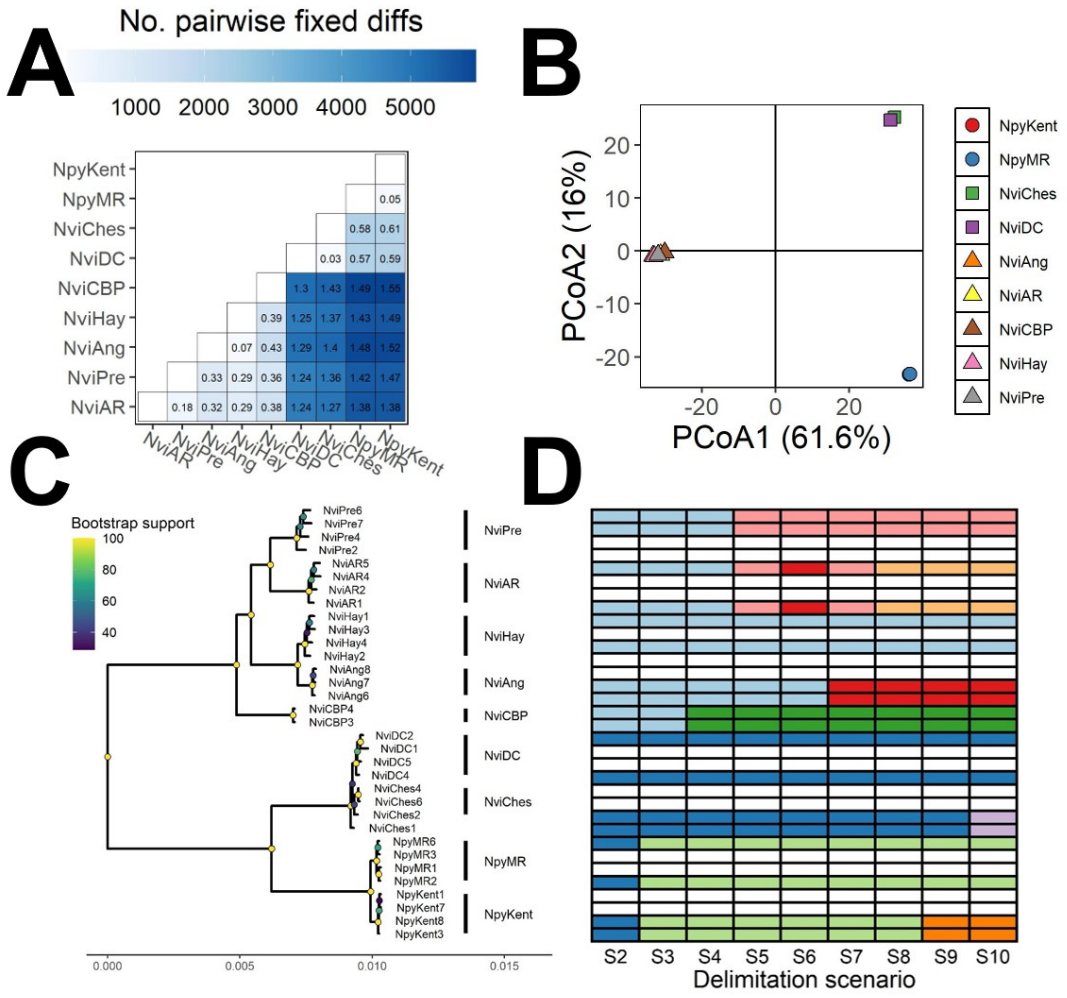


Figure 4.2: Population divergence and clustering. **A:** Heatmap of pairwise population summaries of divergence. Numbers within cells indicate pairwise uncorrelated genetic distances (*p*-distance; %) based on 19,600 ddRAD loci, whilst colours indicate the number of fixed differences within the 18,177 unlinked SNPs, accounting for false positives based on sample sizes. **B:** PCoA of 18,177 unlinked SNPs. Each point represents the centroid per population with colours indicating assignment to a cluster (i.e. a putative species: circles = *N. pygmaea*, squares = *N. vittata* [A] and triangles = *N. vittata* [B]). **C:** ML tree estimated using 19,426 concatenated ddRAD loci and RAxML. Node colours indicate bootstrap support estimated using 1,000 RELL bootstraps. **D:** Allocation of samples to species delimitation scenarios in SNAPP. Each column represents a single model, with each row corresponding to the aligned sample in the phylogenetic tree. Cell colours indicate allocation to species, with cells of the same colour indicating one species. Blank rows indicate samples that were not used in species delimitation.

Phylogenetic Analysis

Partitioning the alignment by ddRAD loci did not affect the topology or branch lengths of the tree, and site concordance factors supported this tree (Appendix 36A – 36C). The summary species tree likewise showed strong support for the separation of lineages, with posterior probabilities of one for almost all nodes (Appendix 36D). Thus, interpretations of the phylogenetic patterns here are based on the concatenation tree. The highly divergent nature of lineages across the group was supported by the ML phylogeny, with all population-level and above nodes supported by 100% bootstrap values (Figure 4.2C).The phylogeny demonstrated three particularly divergent lineages: one containing *N. pygmaea* populations, a sister lineage of two *N. vittata* populations and another paraphyletic *N. vittata* lineage containing the remaining five populations. Despite their co-occurrence, the sympatric populations of *N. vittata* and *N. pygmaea* (NviHay and NpyMR) were not closely associated within the phylogenetic tree. Instead, the sister *N. vittata* lineage (denoted as *N. vittata* [A]) to *N. pygmaea* contained populations from the central southern portion of the distribution.

Species Delimitation and Divergence Time

Species delimitation models increased in likelihood with an increasing number of species, with the highest likelihood and lowest AICM for a delimitation scenario considering all populations as separate species (Figure 4.2D; S10). However, there was a noticeable plateau in AICM values after S5, with greater numbers of species conferring relatively small increases in model likelihood beyond this point (Appendix 37). Divergence time estimates suggested that most lineages diverged from one another between 1 and 5 Mya, with most populations diverging more recently, within the last 300 Kya (Figure 4.3). For all nodes, there was minimal overlap in 95% confidence intervals.

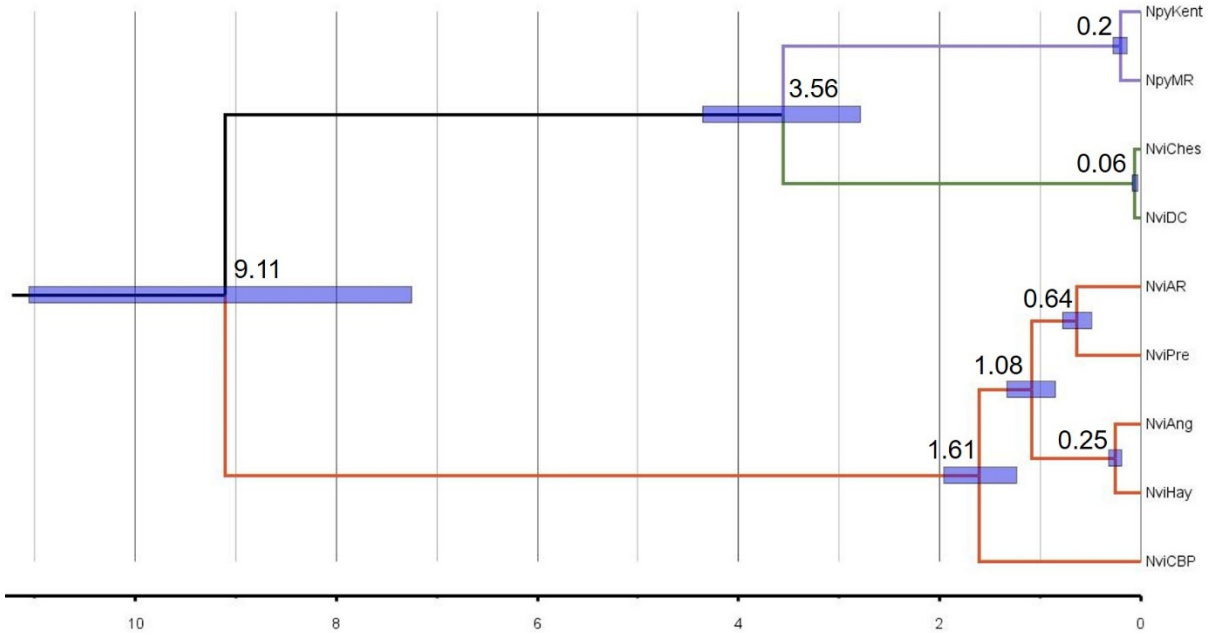


Figure 4.3: Chronogram of divergence time estimates across the clade using SNAPP. Error bars show 95% posterior probabilities out of 1M simulations. The phylogeny was calibrated at the most basal divergence in the tree based on a previous estimation of divergence time using a biogeographic calibration point (Figure 1.4).

Gene Flow and Introgression

No signal of historical population connectivity was inferred across the clade, with TreeMix best supporting a tree that contained no migrations over models with any (Figure 4.4A). This non-migratory model explained 99.99% of the variation in allele frequencies (Figure 4.4B – C). Similarly, D-Suite results showed negligible evidence of introgression across putative species, with a Patterson's *D* of 0.20 and f_4-ratio of 0.028.

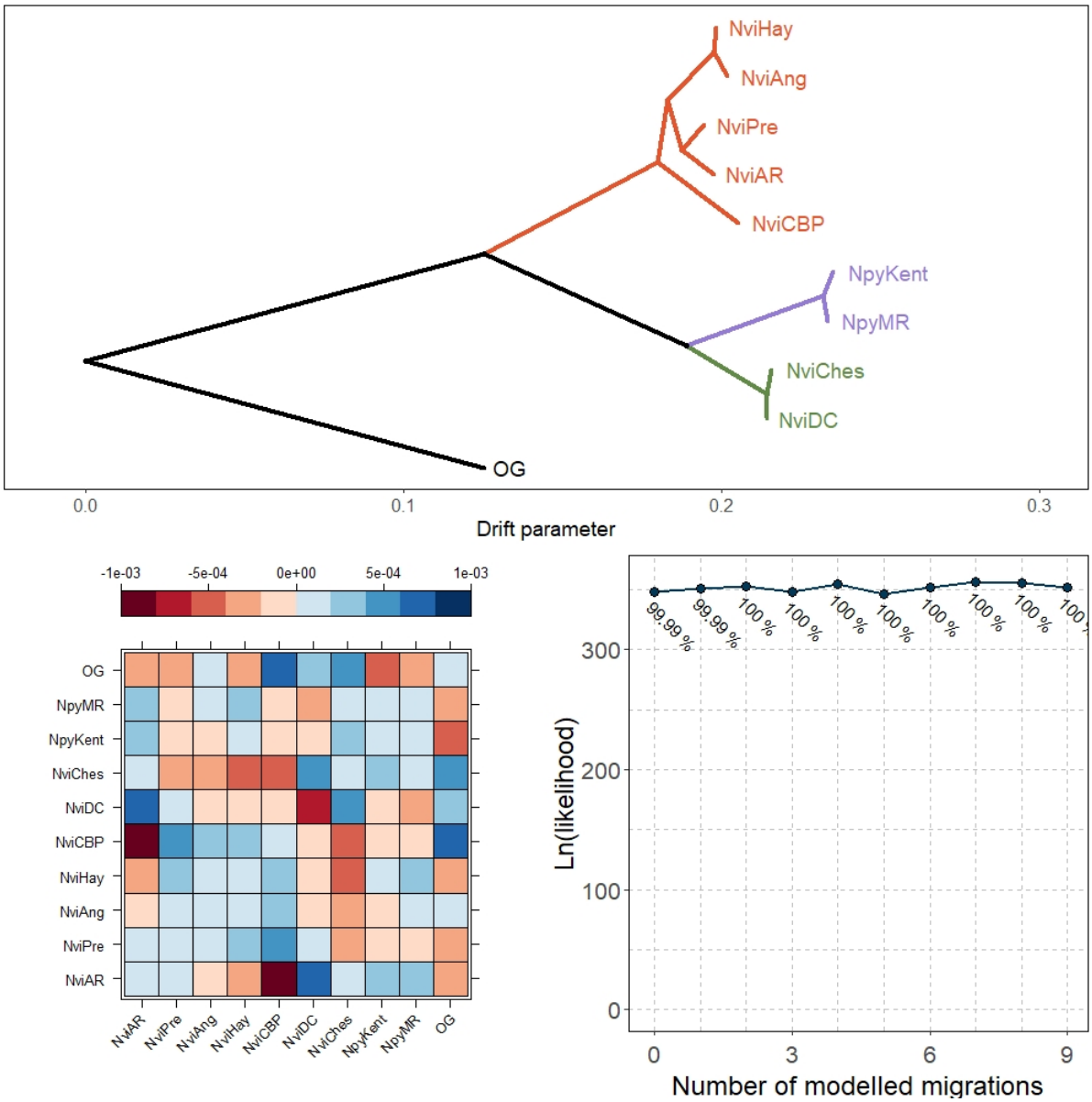


Figure 4.4: Historical migration results using TreeMix. Top figure shows the best supported model, with no modelled migrations and lineages coloured by putative species. The standard error of the covariance matrix (bottom left) shows that variance in allele frequencies are well captured in the tree. Considering additional migration did not improve the likelihood of the model (bottom right), with the zero-migration model explaining 99.99% of the variation.

Enrichment of Fixed Differences

All four sets of ddRAD loci were aligned to the southern pygmy perch genome with 76.14 – 82.89% success, resulting in 16,246 background aligned loci. For each putative species, 1,172 *N. pygmaea*, 739 *N. vittata* [A] and 2,929 *N. vittata* [B] loci were aligned to the

genome. For the background dataset, 7,956 (48.97%) of aligned loci were successfully annotated, with slightly lower annotation rate for putative species datasets (*N. pygmaea* = 463 loci, 39.51%; *N. vittata* [A] = 252, 34.1%; *N. vittata* [B] = 1,039 (35.47%; Appendix 38). For *N. pygmaea*, *N. vittata* [A] and *N. vittata* [B], GO enrichment analyses showed significant enrichment ($p < 0.01$) for two, one and three molecular functions and ten, zero and nine biological processes, respectively (Figure 4.5; Appendix 39). There was no enrichment for cellular components for any putative species. Collectively, these GO terms were underpinned by 96 unique enriched genes (Appendix 40). Within *N. pygmaea*, GO terms were largely associated with DNA helicase activity, development and organisation of cytoskeleton, kidney development and metabolic processes. Within *N. vittata* [B], almost half (5 of 12; 41.67%) of the significantly enriched GO terms were related to heart or circulatory functions, although these were underpinned by only 5 of 28 (17.86%) enriched genes. Other enriched GO terms were related to sister chromatid segregation, and amide and vitamin binding. Only the Protein Binding GO term was enriched within *N. vittata* [A], although this comprised 42 different genes relating to a variety of functions including transport proteins, transcription factors and cytoskeleton elements.

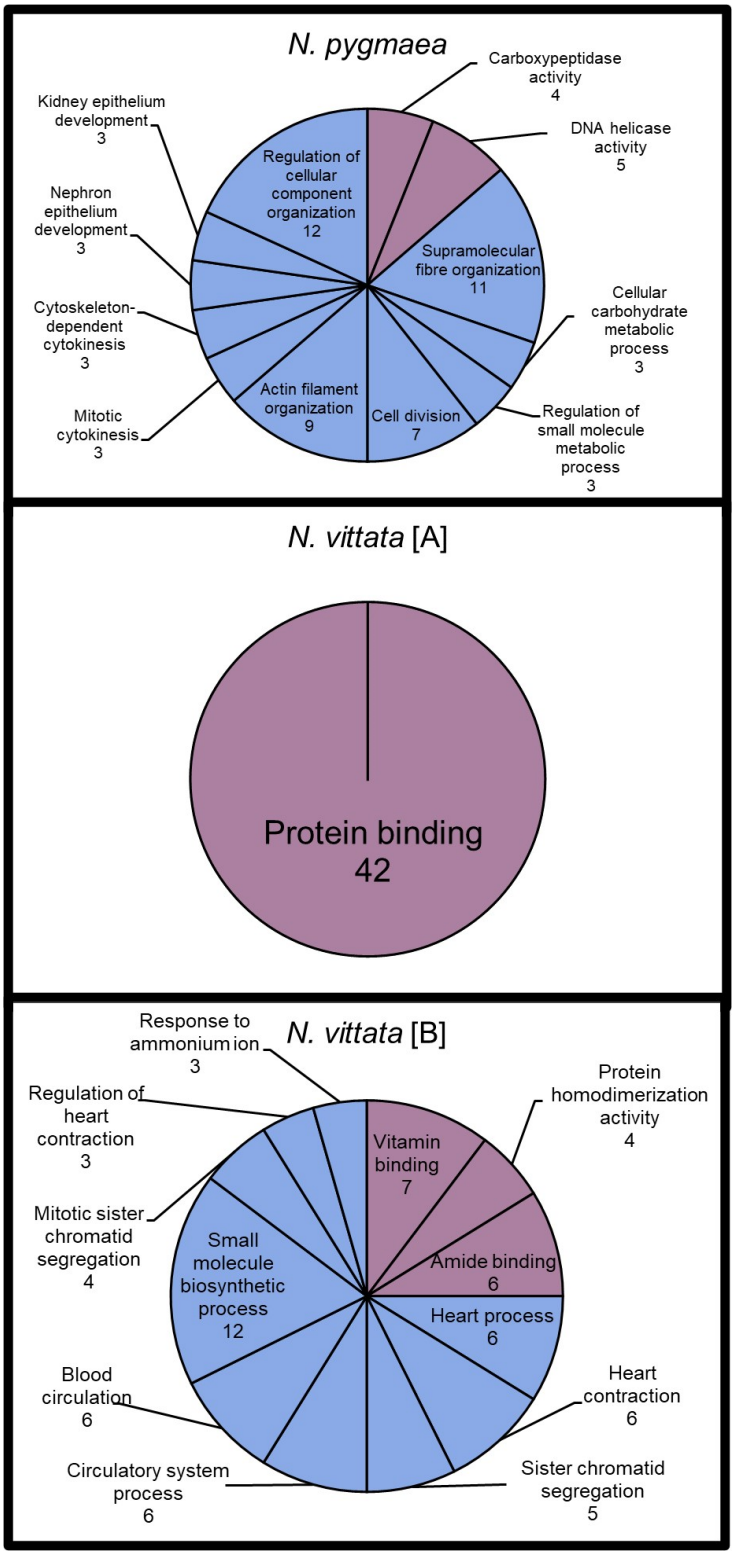


Figure 4.5: Pie charts of significantly enriched GO terms for fixed differences between species. The size of each slice indicates the number of underlying genes (also indicated by the labelled number) and colour denotes the ontology category (Molecular Functions in purple, Biological Processes in blue).

Species Distribution Modelling

All models demonstrated high fit to the data, accurately capturing the full distribution of the species complex (Appendix 41A). Ensemble models built separately for each method demonstrated highly similar areas of suitability, suggesting that variation among methods was minimal (Appendix 41C). Annual precipitation (bio12) was the strongest driving variable across all models, with approximately double the variable importance of all others (Appendix 41B). The ensemble projections of the SDMs over time highlighted the environmental stability of SWWA since the Pliocene, with little fluctuation in distribution extent (Figure 4.6A). Across all time periods, the most significant portion of suitable habitat tended towards the coastal edges of the region, with some expansion across the continental shelf during lower sea levels but not inland. This long-term distribution stability was similarly reflected by high mean suitability (Figure 4.6B) and low standard deviation (Figure 4.6C) over time. Relatedly, there was a clear correlation between distribution area and sea level contours (Appendix 42).

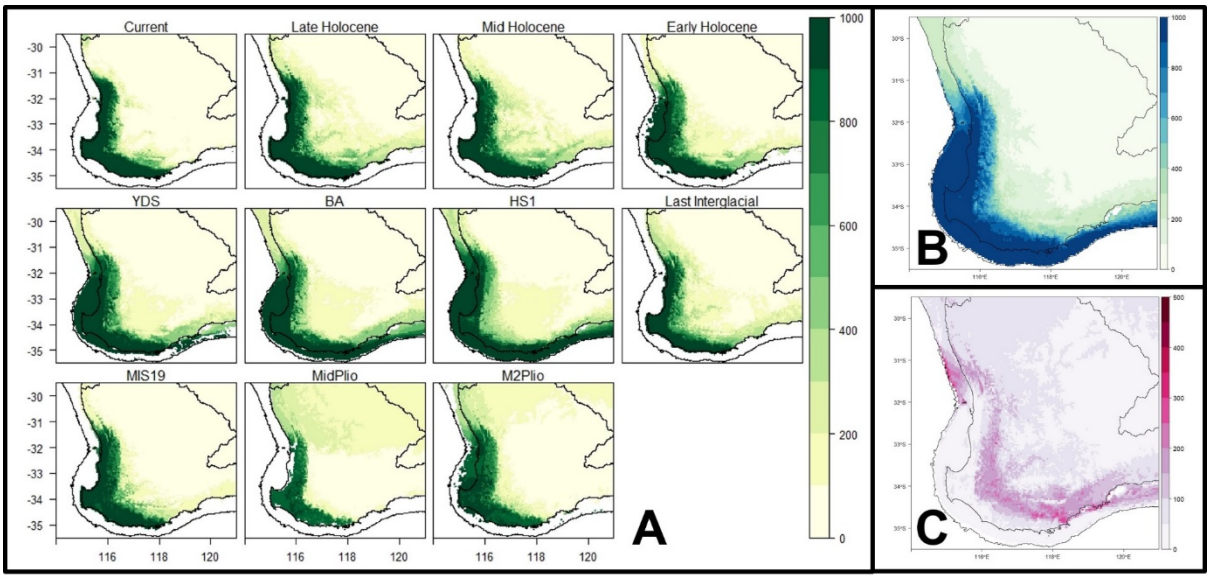


Figure 4.6: SDMs across eleven time periods, ranging from current conditions back to the Pliocene. **A:** Ensemble SDMs and projections summarising 45 individual models. **B:** Mean cell suitability across all eleven time periods. **C:** Standard deviation of cell suitability across all eleven time periods.

The LDMs under the more conservative species delimitation scenarios (S3 & S4) inferred a fragmented distribution for the *N. vittata* [B] lineage (Appendix 43). Only under when *N. vittata* was considered four separate lineages (S5) did each LDM become a singular contiguous range. Regardless of delimitation, lineages formed tightly clustered distributions with minimal overlap. Similar to the SDMs, LDMs of cryptic species appeared stable over time, excluding during the mid-Pliocene (3.2 Mya) where there was more significant overlap and a northward shift in southern lineages (Appendix 44).

Discussion

We identified several divergent lineages across a species complex of freshwater fish endemic to a temperate biodiversity hotspot, including at least two cryptic species. We also showed that these lineages probably persisted in isolation since the late Miocene (9 – 3.5 Mya) over a climatically stable region. Current and projected climatic conditions disproportionately threaten this region because its high lineage diversity appears to have evolved in the absence of major climatic and geological changes.

Diversification and Taxonomy of Western Pygmy Perches

Our results highlight a hierarchy of divergence across Western Australia's *Nannoperca* clade, ranging from divergent populations and cryptic species to morphologically differentiated lineages. No historical migration nor introgression was detected across the clade, indicating that the inferred evolutionary distinctiveness is a product of long-term isolation. Summaries of population divergence largely identified three major clusters, consisting of *N. pygmaea* and two separate clusters of *N. vittata* lineages. This was supported with phylogenetic analysis, which highlighted the paraphyletic nature of *N. vittata*. Additionally, co-occurring populations of *N. vittata* and *N. pygmaea* (populations NviHay

and NpyMR) were not sister lineages within the tree, suggesting that these may instead represent secondary contact following allopatric divergence rather than sympatric speciation.

Species delimitation results implied the presence of a higher number of putative species within the clade, with greatest support for a model considering each population a separate species, and an approximate plateau in likelihood at five species. Multispecies coalescent approaches such as SNAPP have been demonstrated to over-split lineages (Chambers & Hillis 2020; Jackson *et al.* 2017; Sukumaran & Knowles 2017), particularly when populations are highly structured and divergent (Derkarabetian *et al.* 2019). Thus, we instead suggest that three species are present in the clade, denoted *N. pygmaea*, *N. vittata* [A] and *N. vittata* [B]: this is the minimum number of realistic species within the clade given the paraphyly of *N. vittata* and reduces the risk of over-splitting (Coates *et al.* 2018). Neither of these *N. vittata* species could reliably be assigned to the *N. vittata* name as the holotype only denotes “interior of Western Australia” as its origin.

A previous phylogenomic study of all pygmy perch species, including species delimitation using a multi-species coalescent approach, suggested that *N. vittata* was a complex of three cryptic species (Buckley *et al.* 2018). However, this was based on few populations within SWWA, all of which were included in this study. Two of these species correspond to the most divergent lineages within *N. vittata* [B] (NviCBP and NviAR), suggesting that the higher number of species in the previous study was attributed to a lack of intermediate sampling within the lineage.

The delineation of two divergent and cryptic species within *N. vittata* suggests that a taxonomic revision of the lineage is required. This is critical given that species are the most common operational unit of conservation management (Stanton *et al.* 2019), and particularly, are the focus of protective legislation (Coates *et al.* 2018). For example, the identification of the narrowly distributed *N. vittata* [A] lineage implies that it may be of greater conservation concern than the currently unlisted status of *N. vittata* would denote (Bickford *et al.* 2007; Buckley *et al.* 2018; Funk *et al.* 2012). Similarly, *N. pygmaea* is listed as one of the most imperilled freshwater fishes in the country, owing to its small distribution and high extinction risk (Lintermans *et al.* 2020). Given that *N. vittata* [A] and *N. pygmaea* occupy narrow distributions under threat from salinisation and aridification (Allen *et al.* 2020), *ex situ* conservation efforts such as translocation and the formation of insurance populations may be required in the immediate future to protect them from rapid habitat loss. These programs will need to maximise the maintenance of genetic diversity within these depauperate lineages and may rely on genetic rescue (detailed below) across populations of putative species defined here.

Mechanisms Maintaining Reproductive Isolation Across Putative Species

In this study, we developed a novel approach to investigate the role of selection on diagnostic SNPs between putative species by leveraging information from a newly assembled pygmy perch genome. While this approach provides an additional line of evidence for species delimitation, and complements other assessments of species divergence, it has several limitations. As fixed differences can be spuriously inferred with low sample sizes (Gruber *et al.* 2018), larger sample sizes may reveal that some SNPs are not fixed within populations. Additionally, inferring the ecological function of annotated genes remains a challenge, especially in non-model taxa (Pavey *et al.* 2012). On a more

theoretical note, this approach cannot distinguish between SNPs under selection prior to speciation and those selected post-speciation events (as part of reinforcement, for example; Aeschbacher *et al.* 2017; Langerhans & Riesch 2013; Matute & Ortiz-Barrientos 2014). Additionally, it is implicitly assumed that fixed difference SNPs relate to functional changes in the proteins of the genes they occur in, which may not necessarily be the case.

Some extensions of the approach may help to address some of these caveats. For example, determining whether fixed differences convey structural changes in proteins (Cingolani *et al.* 2012) may provide greater confidence in their importance. Additionally, relating significantly enriched gene ontologies to pathways may make connections between different functions and their biological importance clearer (Jassal *et al.* 2020). As with other studies on the genomic architecture of adaptation, expanding reduced representation datasets (such as the ddRAD data used here) to whole genome datasets would provide better coverage of fixed differences that may underlie species divergence (Rice *et al.* 2011), and allow testing of other components of reproductive isolation such as 'genomic islands of speciation' (Malinsky *et al.* 2015; Turner *et al.* 2005) and chromosomal rearrangements (Butlin 2005; Rieseberg 2001).

Annotation of loci containing fixed differences between putative species suggested that functional differences in important physiological traits may have reinforced their initial divergence. For example, enrichment of genes related to metabolism and kidney development in *N. pygmaea* might link to its morphological differentiation (Clarke & Johnston 1999; Gonzalez & Mcdonald 1992; Passow *et al.* 2015), particularly in terms of smaller body size, slower growth rate and lower salinity tolerance (Allen *et al.* 2020). Additionally, reproductive isolation between putative species might be reinforced by

differences in genes related to chromosome arrangement and segregation during mitosis. Gene functions associated with cytoskeletal elements, including chromosomal cytoskeleton proteins, were enriched within *N. pygmaea* and may also play a functional role in mitosis. Similarly, while *N. vittata* [A] only showed enrichment for protein-binding functions, a number of these were associated with transport (actin-based) proteins that influence intracellular arrangements (Lodish *et al.* 2000; Pantaloni *et al.* 2001). Hybrid inviability in crosses of *Drosophila melanogaster* and *Drosophila simulans* has been associated with failure of chromatids to correctly segregate during anaphase of mitosis, resulting from incompatible alleles in loss-of-function genes (Blum *et al.* 2017; Ferree & Prasad 2012; Matute *et al.* 2014; Thomae *et al.* 2013). Thus, differences in proteins responsible for chromatid segregation may cause hybrids of putative species to be reproductively isolated through similar postzygotic mechanisms, further reinforcing the observed lack of hybridisation between sympatric *N. vittata* and *N. pygmaea* populations (Allen *et al.* 2020; Morgan *et al.* 2013). This lack of hybridisation is also likely reinforced by their staggered breeding seasons (Allen *et al.* 2020), although no enriched functions appeared related. Additionally, various circulatory functions were significantly enriched within *N. vittata* [B]: although not directly associated with speciation, these functions have been linked to migratory ability (Eliason *et al.* 2011), food deprivation and hypoxia (Gamperl & Farrell 2004) and cold adaptation (Tiitu 2001) in fishes. These traits may be correlated with the broader distribution of the species relative to *N. vittata* [A], although associated phenotypic differences could not be explored here.

Furthermore, the strong genetic differentiation between putative species may further contribute to reproductive isolation through Bateson-Dobzhansky-Muller Incompatibilities (BDMIs), which result from the epistatic effect and evolutionary mismatch of alleles in hybrid

lineages (Lafon-Placette & Kohler 2015; Lima 2014). Given that BDMIs are driven by fixed differences between species interacting within hybrid individuals (Van Belleghem *et al.* 2018), the divergent nature of putative species here may also drive reduced hybrid viability (Singhal & Moritz 2013).

Maintenance of Highly Divergent Lineages Through Climatic Stability

The highly divergent and ancient nature of lineages within the clade was likely facilitated by the long-term climatic stability of the region, with divergence events spanning the Miocene and Pliocene. Although major environmental changes associated with aridification during the Pliocene impacted the biota of the SWWA (Hopper & Gioia 2004), temperature and rainfall remained relatively constant throughout the glacial cycles of the Pleistocene, especially within the HRP (Rix *et al.* 2015). The stable species distribution reflected these relatively constant climatic conditions since the Pliocene (Spooner *et al.* 2011), with a coastward expansion across exposed landmass when sea levels were lower during glacial maxima. Similar patterns of a coastward expansion, but limited variation in inland regions, have been shown for several coastal plant species in the SWWA (Nevill *et al.* 2014; Nistelberger *et al.* 2014), suggesting that distribution extent was consistently limited by aridity and temperature in inland habitats (Brouwers *et al.* 2012). These spatial patterns reflect the intensity of aridification further inland compared to the relatively benign HRP (Rix *et al.* 2015), and likely placed significant limitations on inland range expansions of *N. vittata* due to intense selective pressure. This hypothesis is consistent with a meta-analysis of estimated selection gradients that showed that regional variation in precipitation is a key factor driving patterns of selection around the globe (Siepielski *et al.* 2017).

While causation of divergences within the lineage was not inferred here, the historical demographic isolation of lineages, the low genetic diversity and low dispersal capacity for pygmy perches (Brauer *et al.* 2016; Buckley *et al.* 2018; Unmack *et al.* 2011) together suggest that their persistence must owe, in part, to perpetual stable habitat. Overall, these results add to a growing body of evidence that broad-scale and long-term climatic stability in the SWWA has driven the accumulation of regional biodiversity (e.g. Binks *et al.* 2019; Cardillo & Pratt 2013; Cooper *et al.* 2011; Funnekotter *et al.* 2019; Galaiduk *et al.* 2017; Gouws & Stewart 2007; Harms 2018; Laity *et al.* 2015; Millar *et al.* 2017; Nevill *et al.* 2014; Nistelberger *et al.* 2014; Rix *et al.* 2015; Sundaram *et al.* 2019). These endemic species encompass an array of lineages ranging from Gondwanan relicts originating as early as the Cretaceous (Giribet *et al.* 2012; Murienne *et al.* 2014; Schultz *et al.* 2009) to late Miocene – early Pliocene diversification events (Hopper *et al.* 2009; Unmack *et al.* 2012), reflecting both ancient ('paleo-endemic') and recent ('neo-endemic') lineages (Rix *et al.* 2015; Sundaram *et al.* 2019). These concerted patterns thus demonstrate how biodiversity in SWWA has persisted and accumulated over millions of years due to the region's climatic stability.

Phylogeographic Structure Within the SWWA

Intraspecific phylogeographic structure was also detected in the most widespread putative species (*N. vittata* [B]), with three highly divergent lineages possibly representing evolutionarily-significant units (ESU; Moritz 1994). This is based on their relatively ancient divergence (>1 Myr) and long-term demographic and spatial isolation. Within this delineation, the Margaret River population (NviCBP) represented a unique ESU: this is corroborated by observations of genetic divergence across other aquatic taxa in the region, with a unique ESU of West Australian freshwater mussel (*Westralunio carteri;* Klunzinger *et*

al. 2020) and a morphologically distinct species of freshwater crayfish (the hairy marron, *Cherax tenuimanus;* Vercoe *et al.* 2009) endemic to Margaret River. The remaining two ESUs captured disjunct pairs of populations at opposing extremes of the species distribution (in the north and the southeast).

Traditionally, the role of ESUs as independent conservation units has led to their separate and isolated management (Funk *et al.* 2012; Weeks *et al.* 2016). This approach often assumes that admixture across ESUs will disrupt locally adaptive genetic variation through outbreeding depression (Weeks *et al.* 2016). However, a growing number of studies have advocated for the role of inter-population genetic mixing – referred to as 'genetic rescue' (Ralls *et al.* 2018) or 'assisted gene flow' (Pavlova *et al.* 2017) – given the elevated extinction risk in small and isolated populations (Frankham 2015). A number of these studies argue that the risk of outbreeding depression, especially relative to the risk of inbreeding depression and population declines, has been somewhat exaggerated (Chan *et al.* 2019; Frankham 2015; Pavlova *et al.* 2017). As such, proposals of genetic rescue have extended to between subspecies or ESUs where necessary (Love Stowell *et al.* 2017; Weeks *et al.* 2016; for examples see Beheregaray *et al.* in press, Harrisson *et al.* 2016 and Johnson *et al.* 2010). Although these examples demonstrate weaker genetic divergence than between ESUs of *N. vittata* [B], it is important to note that predicting the relative probability of inbreeding and outbreeding depression during genetic rescue remains challenging (Robinson *et al.* 2020), and may interfere with effective conservation management (Pavlova *et al.* 2017; Weeks *et al.* 2016). Thus, although these ESUs currently represent independent conservation units, there may be a need to re-assess the potential for genetic rescue as the fates of individual ESUs become clearer.

Implications for Conservation of Biodiversity Hotspots

Freshwater biodiversity is considered one of the most threatened groups globally (Collen *et al.* 2014; Lintermans *et al.* 2020). Within the SWWA, freshwater species are currently threatened by several convergent issues, including ongoing aridification since the mid-last century (Smith & Power 2014), habitat clearing primarily for agricultural development (Andrich & Imberger 2013), secondary salinization of rivers (Allen *et al.* 2020) and invasive species (Beatty & Morgan 2013). These threats will likely be exacerbated by anthropogenic climate change and their own interactive effects (Dinnage *et al.* 2020): projections of climate change alterations alone predict 40 - 50% loss of habitat and 10 – 44% loss of endemic plants (294 – 1,293 species) in the SWWA (Habel *et al.* 2019). However, these projections worsen to 56 – 65% habitat loss when considering the effect of increased agronomic pressure from population growth, demonstrating how climate change and land clearing pressures interact. Additionally, spatial heterogeneity in the effects of climate change may affect some lineages more than others (Hansen *et al.* 2019; Ashcroft *et al.* 2009), such as due to a northern contraction of suitable mesic habitat (Klausmeyer & Shaw 2009). Relatedly, a reduction in fitness associated with heterogeneous climate change has already been observed in some SWWA plant species (Brouwers *et al.* 2012; Dalmaris *et al.* 2015; Monks *et al.* 2019). Conservation programs must therefore consider both the spatial heterogeneity of these impacts, and how they may interact in increasing extinction risk, especially for freshwater fauna.

Disruption of climatic regimes across the globe, in conjunction with broadly increased temperatures (Jeremias *et al.* 2018), aridity (Davis *et al.* 2013; Falkenmark 2013) and sea

levels (Rotzoll & Fletcher 2012) is already impacting species persistence globally. These effects will be exacerbated within regions of low historical climate variability due to both low species resilience and high species diversity, with compounding effects for coastal temperate ecosystems (Chapter 2). Species that evolved in benign climates are not expected to have the same adaptive resilience as those in more variable environments (Gutiérrez-Pesquera *et al.* 2016), a hypothesis supported by reduced dispersal capacity (Sandel *et al.* 2011), adaptive plasticity (Sandoval-Castillo *et al.* 2020) and thermal tolerance (Dynesius & Jansson 2000). Additionally, the high biodiversity of historically stable regions may mean they become hotspots of extinction as swathes of taxa struggle to respond to the selective challenges of anthropogenic climate change (Brooks *et al.* 2002; Waldvogel *et al.* 2020). Given that biodiversity hotspots may also have an excess of undiscovered and cryptic biodiversity (Joppa *et al.* 2011), current estimates of extinction rates may be underestimated (Le Roux *et al.* 2019; Scheffers *et al.* 2012). Together, these components highlight why biodiversity hotspots are regions of high conservation concern and should be priorities in conservation management to maximise preservation of at-risk species. Two of the three putative species delineated here have small ranges, a pattern shared by other narrow-range endemics in biodiversity hotspots (Goldberg *et al.* 2005). Maintenance of critical habitats such as wetlands and ephemeral streams through environmental watering may help protect these species (Allen *et al.* 2020), while translocations and captive breeding programs targeting those most at risk of extinction (e.g. Attard *et al.* 2016a; Thiele *et al.* 2020) may prevent further loss of biodiversity within this hotspot.

Conclusions

Long-term climatic stability has driven the accumulation of divergent and isolated lineages of freshwater fishes in SWWA, providing a mechanism for its status as a biodiversity hotspot. This included the divergence of likely reproductively-isolated cryptic species and a morphologically divergent species, which show evidence of selective reinforcement through differences in intracellular and physiological traits. Given that regional climates are being disrupted globally by climate change, and that species that evolved in historically stable regions may lack the adaptive mechanisms to respond, biodiversity hotspots like SWWA may be exceptionally threatened. These findings only further solidify the conservation need for these regions, with preservation of threatened habitat such as rivers and refugial pools crucial for the persistence of a suite of endemic species. For those species most at risk, adaptive conservation management strategies such as the formation of insurance populations and captive breeding programs may become necessary to protect lineages with lower adaptive resilience to rapidly changing environments.

Chapter 5: Future Directions and Conclusions

General Conclusions

Environmental change associated with both past and future climatic changes are thought to be spatially variable (Ackerly *et al.* 2010; Ashcroft *et al.* 2009) and expected to have complex influences on the evolution and persistence of species globally (Aitken *et al.* 2008; Hoffmann & Sgro 2011). The potential effects of changing environments might be exacerbated for freshwater lineages, where weak dispersal ability, habitat loss by aridification and inundation by sea level rise coincide to increase extinction risk (Blöschl *et al.* 2019; Middelkoop *et al.* 2001; Nijssen *et al.* 2001; Pinceel *et al.* 2013). This thesis aimed to determine how historical environmental changes have impacted the evolution, distribution, and diversity of a group of freshwater fishes by comparing phylogeographic histories across two disparate temperate regions. The thesis entails the most spatially extensive and analytically thorough investigations of phylogeographic histories in the study lineages to date.

On one hand, spatial heterogeneity in major environmental changes since the Miocene had substantial impacts on intraspecific variation in evolutionary history within southeast Australia. On the other hand, in southwest Australia environmental stability is thought to have facilitated the accumulation of divergent lineages and diversification with a hotspot. Although biogeographic changes across southeast Australia were extensive, ecologically-similar species demonstrated discordant evolutionary responses despite local-scale concordance in demographic history. Together, these findings highlight how environmental and demographic variability complexify evolutionary responses and may influence the ability to predict species responses to climate change. Here, I briefly summarise again the major findings of each data chapter and highlight their contributions to our understanding of regional biogeographic history, species responses to climate change and conservation

management. I expand upon the common themes of the thesis to briefly discuss how the various factors covered in the different data chapters – such as environmental history, genetic variation, and ecological traits – interact to contribute to species responses to environmental change. Lastly, I outline future developments which might expand upon the work within this thesis.

Chapter 2: The roles of aridification and sea level changes in the diversification and persistence of freshwater fish lineages.

Determining the resilience of species to contemporary climate change is complicated by spatial heterogeneity in environmental factors and how these might impact on intraspecific responses across their ranges (Razgour *et al.* 2019). Relatedly, understanding the relative role of different environmental changes across regions is critical to more accurately predict species' responses (Zamudio *et al.* 2016). Chapter two of this thesis evaluated the relative role of environmental changes, especially aridification and sea level change, on the evolution and diversification of *Nannoperca australis* across both inland and coastal basins. Genome-wide data was used to investigate evolutionary history using a combination of phylogenetic and coalescent methods, in conjunction with SDMs. A hierarchy of divergence was evident across the clade, spanning species divergences in the Miocene to Holocene fragmentation. For inland populations, these divergences were primarily associated with ancient aridification events leading to the contraction or demise of connective waterbodies, whereas coastal populations were more affected by changes in habitat extent and structure with fluctuating sea levels. This included the divergence of a cryptic species (*Nannoperca* 'flindersi'), which was likely reinforced by isolated refugia during glacial periods despite potential connectivity and interglacial gene flow. For some regions, these two effects were interactive, where the reduction of waterbodies and coastal inundation both contributed to

historical population structure. This highlights the spatially heterogeneous nature of environmental change across species distributions and how applying large-scale predictions of species responses to climate change may not accommodate all locally relevant factors (Pittock *et al.* 2008; Razgour *et al.* 2019). Future management scenarios should consider this spatial variation in prediction of responses to climate change, particularly in how aridification or eustatic changes may act heterogeneously across species distributions.

Chapter 3: Variation in intraspecific demographic history drives localised concordance but species-wide discordance in responses to Plio-Pleistocene climatic change.

Understanding how species traits may facilitate resilience to contemporary climate change remains a critical factor in detecting and protecting species at risk of extinction (Waldvogel *et al.* 2020). Adaptive potential in response to climatic changes likely relies on both ecological (Somero 2010; Travis *et al.* 2013) and demographic (including genetic) traits (Pearson *et al.* 2014; Williams *et al.* 2008), although intraspecific and spatial variation in demographic traits may significantly impact how even ecologically similar species respond (Bridle *et al.* 2010; Hoffmann & Sgro 2011; Vucetich & Waite 2003). Chapter three examined concordance in species responses to Plio-Pleistocene climatic changes in *Nannoperca obscura* and *N. australis*. Genomic data was used to investigate evolutionary histories of the two species, primarily focusing on demographic histories, across their shared range. This included a comparative test to determine whether demographic changes in co-distributed populations were statistically concordant, and environmental modelling to examine species-wide distribution changes. Co-distributed populations largely showed concordant demographic changes, with a signal of glacial expansion and post-glacial

bottleneck across most populations. However, an expansion in range was only observed in *N. australis*, and the degree of concordance was reduced near the edge of their long-term shared refugium. Together, these results point to a history of low SGV in *N. obscura,* which might not have pre-disposed them to expand during more favourable conditions unlike *N. australis*. Indeed, this difference in genetic variation may underscore vastly different successes in reintroduction programs for the two species (Beheregaray *et al.* in press). This disparity highlights the importance of SGV in species responses to climatic change (Szűcs *et al.* 2017; Williams *et al.* 2019), even when ecological traits alone should not suggest such discordance. For genetically depauperate species such as *N. obscura*, increasing genetic diversity through adaptive introgression (genetic rescue) might provide a strong avenue to improve their adaptive potential, especially under climate change (Love Stowell *et al.* 2017; Pavlova *et al.* 2017; Robinson *et al.* 2020).

Chapter 4: Long-term climatic stability drives accumulation of divergent lineages in a temperate biodiversity hotspot.

Long-term climatic stability is often inferred as a primary mechanism for the accumulation of biodiversity within hotspots and in tropical regions (Harrison & Noss 2017; Mittelbach *et al.* 2007), including the SWWA hotspot (Hopper & Gioia 2004). However, alternative hypotheses have been proposed, and the relative role of stability on the process of diversification is debatable (Cook *et al.* 2015). In chapter four, the role of historical climatic stability on the accumulation of divergent lineages – including cryptic species – in *Nannoperca vittata* and *N. pygmaea* was investigated within this biodiversity hotspot. Stable climatic profiles since the Pliocene drove a stable species distribution over time and allowed for the persistence of divergent and anciently isolated lineages. The formation of cryptic species within this group was supported by fixation in genes potentially related to

reinforcement, promoting reproductive isolation through intracellular mechanisms associated with chromosome segregation. These findings support the identity of SWWA as a biodiversity hotspot driven by millions of years of relatively stable climate (Binks *et al.* 2019; Cardillo & Pratt 2013; Cook *et al.* 2015), and highlight the compounding risk of climate change due to high species diversity and low adaptive resilience (Brooks *et al.* 2002; Gutiérrez-Pesquera *et al.* 2016; Habel *et al.* 2019). Within the pygmy perches, the denotation of cryptic species and historically isolated lineages requires re-evaluation of taxonomy and conservation management and suggests that extinction risk might be underestimated even for more widespread 'species'.

Biogeographic Comparison

The two temperate zones investigated here demonstrated contrasting temporal environmental variation, which likely drove the disparate evolutionary histories in ecologically similar species. In the southwest, long-term climatic stability spanning from the Pliocene maintained anciently isolated and highly divergent lineages, whilst environmental variability in the southeast lead to variable patterns of divergence including fluctuating population connectivity and spatial variation in the timing and extent of divergence. Additionally, there appeared to be significant differences in the degree of historical interpopulation connectivity between the two regions, with southwest species demonstrating less connectivity. While this variation may seem contradictory to expectations, it is important to note that environmental changes may also facilitate reconnection and gene flow between previously isolated lineages through secondary contact (Cros *et al.* 2020; Garrick *et al.* 2019; Provan & Bennett 2008). For poorly dispersing freshwater organisms, range expansions and gene flow are often facilitated by tectonic shifts driving river capture (Waters *et al.* 2020) or the presence of large water bodies (Miura *et al.* 2019), both of which

have been historically lacking in southwest Australia (Edwards *et al.* 2008; Hopper & Gioia 2004). These rearrangements occurred throughout southeast Australia from the Miocene onwards (McLaren *et al.* 2012; Waters *et al.* 2019) and were important drivers of connectivity for pygmy perches, underlying the disparate patterns between the temperate zones. Similarly contrasting phylogeographic patterns between southeast and southwest Australian lineages of widespread clades have been noted in other taxa, through variable divergence times (Cook *et al.* 2015; Rix & Harvey 2012; Schultz *et al.* 2009) or degree of phylogeographic structure (Byrne 2008; Unmack *et al.* 2012). The link between historical environmental variation and variable patterns of divergence has been widely demonstrated (Barrow *et al.* 2018; Bell *et al.* 2017; Waters *et al.* 2020), with the timing and nature of divergence dependant on the interaction of Earth history and species traits (Bell *et al.* 2017; Papadopoulou & Knowles 2016; Paz *et al.* 2015; Zamudio *et al.* 2016). These findings add to an extensive body of literature that demonstrates contrasting phylogeographic patterns between biogeographic regions with variable or static histories (e.g. Bell *et al.* 2017; Carnaval *et al.* 2014; Moussalli *et al.* 2009; Smith *et al.* 2017).

Despite their contrasting environments, some aspects of biogeographic history were shared between the two temperate regions. Chiefly, the presence of relatively stable climatic refugia – even within the more historically variable southeast – were key to the maintenance and persistence of lineages. Within southeast Australia, isolated refugia enabled the persistence of *Nannoperca* 'flindersi' as a divergent cryptic species despite post-divergence gene flow. Similarly, the shared climatic refugium in coastal Victoria used by both *N. obscura* and *N. australis* appeared to be critical for their persistence, particularly for the former which showed likely reduced adaptive capacity owing to low SGV. The role of climatic refugia in the long-term persistence of species is a well-documented phenomenon

(Gavin *et al.* 2014; Hewitt 2000; Provan & Bennett 2008; Shafer *et al.* 2010; Sommer & Zachos 2009) and provides a historical reference for determining future species persistence. For example, future projections of climate change, in conjunction with known history of past refugia, might help indicate where – if any – climatic refugia will exist to buffer the selective stress of contemporary climate change (Gavin *et al.* 2014; Keppel *et al.* 2012). However, historical refugia are often associated with protecting species from extreme cold climates, contrasting with a contemporarily warming world (Keppel *et al.* 2012). Additionally, the ubiquity of climatic refugia has been debated, with species-specific traits playing a key role in determining the location, extent, and presence of past refugia (Keppel *et al.* 2012; Stewart *et al.* 2010). These aspects place severe challenges on identifying future climatic refugia and highlight the requirement for proactive conservation measures to preserve at-risk species.

Discordance Across Species

Although contrasting environmental histories in the temperate zones likely drove contrasting phylogeographic patterns, species responses within a bioregion were not fully concordant. The discordance in *N. australis* and *N. obscura* responses to past climatic change were likely linked to differences in SGV, an important factor underlying adaptive potential to environmental change (Fordham *et al.* 2014; Waldvogel *et al.* 2020; Willi & Hoffmann 2009). Even when environmental barriers are shared between species, interactions between species-specific traits and environmental variation may drive interspecific discordance (Bell *et al.* 2017; Papadopoulou & Knowles 2016; Paz *et al.* 2015). Thus, caution should be exercised when applying assumptions of phylogeographic patterns to other co-distributed taxa, even when they appear to be ecologically similar and are hypothesised to be concordant.

Idiosyncratic species responses to environmental changes across the globe appear to be linked to two clear sources of variation: environmental (both temporally and spatially) and individual species traits, including ecological or physiological aspects as well as demographic characteristics such as genetic diversity. The various chapters of this thesis have contributed to the understanding of how phylogeographic patterns may be influenced by both environmental and demographic variability, especially in freshwater species with poor dispersal capacity.

Future Directions and Concluding Remarks

The results from this thesis highlight the complex interaction of extrinsic and intrinsic factors in adaptation to climatic changes and provide some of the most thorough investigations of phylogeography using genomics in Australian freshwater fishes. For *N. obscura* and *N. vittata*, the relevant chapters within this thesis represent their first genomic phylogeographic studies, and the most comprehensive such study for *N. australis*. This work thus paves the way for further investigations of how environmental variation and history have influenced evolutionary patterns and adaptive potential in the study species. Additionally, this thesis contributes to the ongoing conservation efforts for these species, by providing baseline historical patterns upon which actions such as genetic rescue, translocation or the delineation of conservation units may be based.

Despite the applications of this thesis, it by no means provides an exhaustive overview of all relevant factors underlying evolutionary history in pygmy perches. As such, expanding the work presented here may provide greater insight into the role of evolutionary and environmental history in the persistence of threatened lineages, especially for pygmy

perches. The comparative analyses for *N. australis* and *N. obscura* could be further extended to investigate differences in adaptive potential. For example, candidate loci associated with adaptation to hydroclimatic variation based on genotype-environment associations (Brauer *et al.* 2016) and differential expression (Brauer *et al.* 2017) have been identified in *N. australis.* These same loci could be analysed within *N. obscura* to test if variation in these ecologically important adaptive genes is low in the species. Another similar approach may involve assessing the potential for adaptive introgression across ESUs within *N. obscura*, a process which is currently underway. This could include a spatial assessment of genomic vulnerability (Bay *et al.* 2018), whereby the least vulnerable populations can be considered as a source for genetic rescue (Razgour *et al.* 2019; Robinson *et al.* 2020).

A genetic rescue program for the MDB lineage of *N. obscura* – now extinct in the wild – is in its early stages. Remnant populations of the MDB lineage, obtained from three separate dams, have been kept within the MELFU for approximately ten years. In 2019, Urrbrae and Investigator Colleges (both within South Australia) established their own *in situ* captive breeding programs using a subset of individuals from the MELFU, resulting in a handful of the first putative crosses between MDB and coastal *N. obscura* lineages (Brauer & Beheregaray, unpublished data). The low number of hybrids within these two college programs was attributed to the collection of fish late in the breeding season, and it is expected that further generations in the following year will improve the number of hybrids. The detected hybrids show increased genetic diversity (heterozygosity) compared to the parental lineages, suggesting that genetic rescue is possible.

Complementary to these breeding programs, a separate experiment was established within the MELFU in 2018. This involved the creation of several separate breeding groups, including tanks of pure coastal fish, pure MDB fish, and reciprocal combinations of the two (i.e. MDB males with coastal females, and vice versa). Like the college arrangements, these tanks were set up late into the *N. obscura* breeding season and did not produce offspring in 2018/2019. However, the following season was highly successful, with hundreds of offspring generated in late 2019/early 2020. Genomic sequencing and parentage analysis will be done in 2021 to determine the level of hybridisation between the two lineages and assess the potential for genetic rescue in the MDB lineage. The findings of this thesis contribute to this ongoing program by providing a historical framework for future crosses and demonstrating how historically low genetic diversity likely played a strong role in their reduced distribution and conservation success.

Although the *N. australis* draft genome provided a reference for the alignment and annotation of ddRAD loci, improvements to the quality of the genome are already underway at the MELFU using extended PacBio reads. These are expected to reduce the number of scaffolds comprising the genome, leading to improve alignment and annotation. Additionally, this improved genome will provide the framework for building a linkage map, which could then be used to determine genomic islands of speciation (Malinsky *et al.* 2015; Turner *et al.* 2005), for example.

As mentioned within Chapter 4, the functional enrichment of fixed differences approach itself could be extended to delve deeper into the functional role of these diagnostic traits. Corroborating these findings with physiological, ecological, or experimental evidence may provide stronger inferences of reproductive isolation and species delineation through

integrative taxonomy (Dayrat 2005). The applicability of this enrichment approach to infer the role of selection in cryptic species can be validated using a suite of different case studies, ranging from weakly differentiated cryptic species to morphologically-defined or putatively reproductively isolated species. Given that this method is based on fixed differences, it is likely conservative regarding detecting species boundaries, and these further studies may help to determine how sensitive the approach is to species divergence. Several applicable datasets already existed within the MELFU (such as golden perch, Pavlova *et al.* 2017; rainbowfishes, Sandoval-Castillo *et al.* 2020; and river blackfishes, Unmack *et al.* 2017), and additional datasets may be sourced from the literature.

The combined results from the three data chapters of this thesis contribute to our understanding of how spatiotemporal environmental variation and species traits interact to influence species responses to environmental change. Complexity in species responses owing to both regional environmental heterogeneity and intrinsic species characteristics may complicate predictions of species vulnerability to anthropogenic climate change. Historically stable bioregions are especially at risk under future climate change by the compounding impact of high species diversity and reduced adaptive capacity, although historically variable regions are not unthreatened. Particularly, aridification and eustatic changes will likely place compounding impacts on freshwater lineages through population isolation and habitat loss, both of which are expected to worsen under anthropogenic climate change. Thus, considering heterogeneity in environmental shifts associated with broader climatic change is critically important to identify lineages at higher risk of extinction and inform appropriate conservation management measures.

Appendices

Appendix 1: Full manuscript of Buckley *et al.* (2018) “Phylogenomic history of enigmatic pygmy perches: implications for biogeography, taxonomy and conservation.” *Royal Society Open Science*. 5:172125. Reproduced with permission for open access articles by Royal Society.

ROYAL SOCIETY OPEN SCIENCE

rsos.royalsocietypublishing.org

Research

Cite this article: Buckley SJ, Domingos FMCB, Attard CRM, Brauer CJ, Sandoval-Castillo J, Lodge R, Unmack PJ, Beheregaray LB. 2018 Phylogenomic history of enigmatic pygmy perches: implications for biogeography, taxonomy and conservation. *R. Soc. open sci.* **5**: 172125. http://dx.doi.org/10.1098/rsos.172125

Received: 8 December 2017
Accepted: 30 April 2018

Subject Category:
Biology (whole organism)

Subject Areas:
evolution/biogeography/genomics

Keywords:
cryptic species, ddRAD-seq, freshwater fish, historical biogeography, phylogeography, *Nannoperca*

Author for correspondence:
Luciano B. Beheregaray
e-mail: luciano.beheregaray@flinders.edu.au

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4108328.

THE ROYAL SOCIETY PUBLISHING

Phylogenomic history of enigmatic pygmy perches: implications for biogeography, taxonomy and conservation

Sean J. Buckley[1], Fabricius M. C. B. Domingos[1,2], Catherine R. M. Attard[1], Chris J. Brauer[1], Jonathan Sandoval-Castillo[1], Ryan Lodge[1], Peter J. Unmack[3] and Luciano B. Beheregaray[1]

[1]Molecular Ecology Laboratory, College of Science and Engineering, Flinders University, GPO Box 2100, Adelaide, South Australia 5001, Australia
[2]Instituto de Ciências Biológicas e da Saúde, Universidade Federal de Mato Grosso, Pontal do Araguaia, MT 78698-000, Brazil
[3]Institute for Applied Ecology, University of Canberra, Canberra, Australian Capital Territory 2601, Australia

CRMA, 0000-0003-1157-570X; LBB, 0000-0003-0944-3003

Pygmy perches (Percichthyidae) are a group of poorly dispersing freshwater fishes that have a puzzling biogeographic disjunction across southern Australia. Current understanding of pygmy perch phylogenetic relationships suggests past east–west migrations across a vast expanse of now arid habitat in central southern Australia, a region lacking contemporary rivers. Pygmy perches also represent a threatened group with confusing taxonomy and potentially cryptic species diversity. Here, we present the first study of the evolutionary history of pygmy perches based on genome-wide information. Data from 13 991 ddRAD loci and a concatenated sequence of 1 075 734 bp were generated for all currently described and potentially cryptic species. Phylogenetic relationships, biogeographic history and cryptic diversification were inferred using a framework that combines phylogenomics, species delimitation and estimation of divergence times. The genome-wide phylogeny clarified the biogeographic history of pygmy perches, demonstrating multiple east–west events of divergence within the group across the Australian continent. These results also resolved discordance between nuclear and mitochondrial data from a previous study. In addition, we propose three cryptic

species within a southwestern species complex. The finding of potentially new species demonstrates that pygmy perches may be even more susceptible to ecological and demographic threats than previously thought. Our results have substantial implications for improving conservation legislation of pygmy perch lineages, especially in southwestern Western Australia.

1. Introduction

The biogeographic histories of species contain information about their past distribution and evolutionary trajectories that can clarify key aspects about community assemblages, biotic exchanges and environmental determinants of biodiversity [1–4]. In fact, coherent patterns of distribution and evolutionary history are often identified in analyses of multiple codistributed taxa, which points to commonalities in biogeographic history [5–7]. On the other hand, biogeographic distributions that cannot be related to ecology or explained by dispersal or vicariance offer enigmatic puzzles that require additional investigation (e.g. [8,9]). Biogeographic knowledge can be further used in applied research, such as assessments of taxonomy and improving conservation legislation [10]. Species delimitation, for example, is a type of phylogenetic analysis that uses computational modelling to determine the number of putative species in a group by using coalescent methods [11,12]. This can lead to revisions of taxonomic uncertainties, a critical aspect of conservation legislation. For example, the identification of independently evolving but cryptic lineages within a single threatened taxon indicates that conservation strategies should manage these lineages separately and avoid outbreeding depression and hybridization among lineages [13].

With the expansion of next-generation sequencing (NGS) technologies, the ability to sequence a representative subset of the genome or even full genomes has become a feasible process in phylogenetics [14–17]. It allows analyses of thousands or more DNA markers, providing high power for circumventing locus-specific biases and inferring phylogenies. Furthermore, no *a priori* genomic resources are required [18–20]. One popular method for obtaining genomic data for non-model species is restriction site-associated DNA sequencing (RAD-seq), which uses restriction enzymes to break down the genome to easily sequenced fragments [21] that can be reconstructed into thousands of DNA loci for estimating robust phylogenetic trees [22,23]. The applicability of RAD-seq in phylogenetics has been recently assessed, with overall support for the method, especially for studies involving recently diverged taxa and complex phylogenies [20,24–26].

Phylogenomics can be applied to complex and conflicting biogeographic assessments in a more robust manner than traditional phylogenetic analyses. One long-standing biogeographic puzzle to be addressed using phylogenomics relates to the currently disjunct distribution of populations and closely related lineages of several aquatic-dependent taxa across southern Australia [27,28]. These taxa, which include crustaceans, fishes and frogs, occur in a distinct temperate zone in the west (i.e. southwest Western Australia) and the east (i.e. the Murray–Darling Basin and southeast coast) of southern Australia, which are isolated by a vast expanse of arid habitat and the limestone-rich Nullarbor Plain in central southern Australia. The palaeoclimate of the area is relatively well-understood: the aridity of central southern Australia developed starting in the Oligocene [29,30] followed by the formation of the Nullarbor Plain 14–16 Ma in the mid-Miocene [27,31]. By the mid-Miocene aridity was firmly established based on the lack of connected drainages across the Nullarbor Plain [27]. Given this, as one would parsimoniously expect, aquatic-dependent lineages on either side of the plain typically show reciprocal monophyly based on genetic data [28,29,32]. However, there are exceptions to this pattern, such as that observed in pygmy perches (Teleostei, Percichthyidae) [27].

Pygmy perches are composed of species endemic to the southeast (e.g. Murray–Darling Basin and coastal drainages) and the southwest (figure 1): six recognized species within *Nannoperca*, namely *N. australis*, *N. obscura*, *N. oxleyana*, *N. pygmaea*, *N. variegata*, and *N. vittata* [27,33,34], and the sole member of *Nannatherina*, *Nth. balstoni*. A previous phylogeny of pygmy perches supported multiple divergences or east–west migrations prior to the formation of the Nullarbor divide [27]. That study used limited genetic data (the cytochrome *b* gene, three nuclear loci, and allozymes), which could limit their ability to determine accurate inferences of species relationships. It is further compounded by the discordance of phylogenetic trees across the molecular markers used, which showed contrasting phylogenetic relationships both within and across marker types. For instance, there were unresolved phylogenetic relationships between *N. australis*, *N. obscura*, *N. oxleyana* and *N. variegata*, possibly resulting from the complete mitochondrial introgression of *N. obscura* by *N. australis* [27,35].

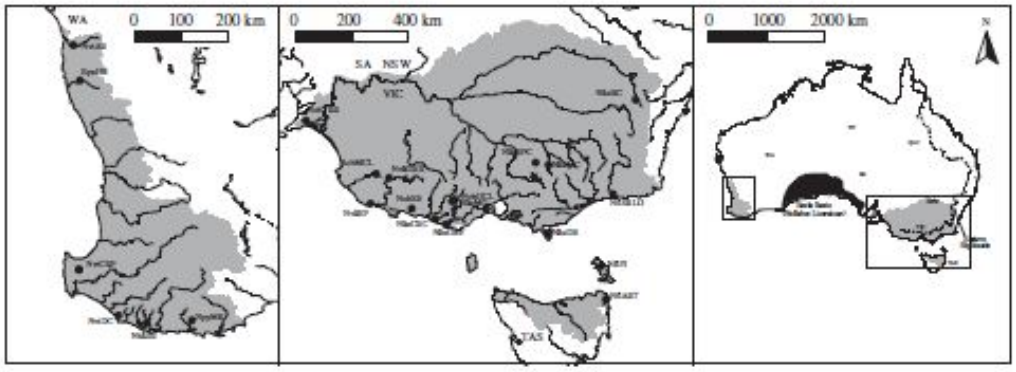


Figure 1. Contemporary distributions of pygmy perch species and populations used in this study. Population abbreviations are denoted within table 1. Population locations are not shown for *N. oxleyana*, which occupies a small region of lower Queensland (distribution indicated in the right figure). The Nullarbor Plain barrier is indicated by the black section of the right figure.

Issues with the pygmy perch phylogeny also extend to taxonomy, with some studies suggesting there are cryptic species in *Nannoperca*. Within *N. australis*, populations from the eastern portion of its range have been suggested as a separate species informally referred to as *N.* 'flindersi' that is awaiting formal taxonomic review [27,36]. Within *N. vittata*, Unmack *et al.* [27] suggested that a second cryptic species was present. Additionally, *N. pygmaea* was only recently formally described as a distinct species from its putative sister species *N. vittata* based on the morphological and allozyme assessment by Morgan *et al.* [34]. Most pygmy perch species are currently threatened or endangered, thus confirming their taxonomy and identifying cryptic species is imperative for well-informed conservation. Five species (*N. obscura*, *N. oxleyana*, *N. pygmaea*, *N. variegata* and *Nth. balstoni*) are listed as threatened in national legislation [33], all species excluding *N. vittata* are listed in their respective state legislation, and several species have also been listed from vulnerable to endangered in the International Union for Conservation of Nature (IUCN) Red List (IUCN, 2015). Declines in populations of these species have been linked largely to habitat degradation as a result of human-altered hydrological conditions (such as damming and agricultural irrigation) and competition or predation by introduced fish species [27,33,37–40]. These pressures are expected to be exacerbated by a drying climate due to contemporary climate change, with large impacts on ecologically specialized species or those with limited distributions, such as pygmy perches [33,38,39,41].

Here, we capitalize on NGS technology to address uncertainties in the evolutionary history of pygmy perches, to assess the taxonomic identity of recently suggested or described species, and to investigate the possibility of further cryptic diversity within pygmy perch lineages. We present phylogenetic trees based on genome-wide data and use a species delimitation framework and divergence estimates encompassing samples from all described species of pygmy perches. Our findings have implications for clarifying the puzzling biogeography of southern Australia and for the taxonomy and conservation management of a threatened group of endemic freshwater fishes.

2. Material and methods

2.1. Sample collection and ddRAD library preparation

Specimens were selected to represent the breadth of currently described pygmy perch species, totalling 45 individual samples across seven formally recognized species and all known potential cryptic forms (table 1). Our samples include populations spanning each species range and representatives of all known lineages, including known evolutionarily significant units (ESUs) within each taxon [27,36,39,40,42] (figure 1). The most closely related extant sister lineage to pygmy perches, the nightfish *Bostockia porosa* was included as the outgroup [27]. Specimens were collected using electrofishing, dip-, fyke- or seine-netting. Either the caudal fin or the entire specimen was stored at −80°C at the South Australian Museum, or in 99% ethanol at Flinders University.

DNA was extracted from muscle tissue or fin clips using a modified salting-out method [43] or a Qiagen DNeasy kit (Qiagen Inc., Valencia, CA, USA). Genomic DNA was checked for quality using a spectrophotometer (NanoDrop, Thermo Scientific), integrity using 2% agarose gels, and quantity

Table 1. Locality data for all species and individuals examined. Population abbreviations described in the table were those used for further analyses, while *n* refers to the number of individuals sequenced per locality.

species	location (*n*)	population abbreviation	field code
Bostockia porosa	Hill River (2)	BpoHR	F-FISHY6:DM184+
	Canebrake Pool (1)	BpoCBP	F-FISHx2:P21+
Nannatherina balstoni	Inlet River (2)	NballR	F-FISHx2:B6+
Nannoperca variegata	Glenelg River (2)	NvaGRB	PU00-15VPP
	Ewen Ponds (2)	NvaEP	F-FISH83:HS63+
Nannoperca vittata sp. 1	Arrowsmith River (2)	NviAR	F-FISHY6:DM150+
Nannoperca vittata sp. 2	Doggerup Creek (2)	NviDC	PU09-49NV
Nannoperca vittata sp. 3	Canebreak Pool (2)	NviCBP	PU09-58NV
Nannoperca pygmaea	Mitchell River (2)	NpyMR	PU09-37
Nannoperca oxleyana	Stradbroke Island (4)	NoxSI	F-FISH93:BF1+, CF1+
Nannoperca obscura	Gnarkeet Creek (2)	NobGCL	PU00-27YPP
	Shaw River (2)	NobSR	PU02-113YPP
	Tookyaerta [breeders] (2)	NobTBR	YPBr+
	Mosquito Creek (2)	NobMCL	PU00-16YPP
Nannoperca australis	Darby River (2)	NauDR	PU02-70SPP
	Gellibrand River (1)	NauGRF	PU02-92SPP
	Mundy Gully (1)	NauMG	SPP08-11
	Curdies River (1)	NauCRC	PU00-24SPP
	Swanpool Creek (1)	NauSPC	PU09-03SPP
	Blakney Creek (2)	NauBC	F-FISH98:LPP-3+
	Meadows Creek (2)	NauMC	rCB1301+
Nannoperca 'flindersi'	Snowy River (2)	NflSRLO	PU99-85SPP
	Anson River (2)	NflART	F-FISH82:HT-20+
	Flinders Island (2)	NflFI	F-FISH84:FI-3+

using a fluorometer (Qubit, Life Technologies). The ddRAD genomic libraries were prepared for the 45 samples in house following Peterson *et al*. [44], with modifications as described in Brauer *et al*. [39]. Genomic libraries were paired-end sequenced on an Illumina HiSeq 2000 at Genome Quebec (Montreal, Canada).

2.2. Sequence filtering and alignment

The resultant reads were filtered and cleaned to create a set of aligned concatenated sequences for statistical analyses. The raw sequences were demultiplexed using the 'process_radtags' module of STACKS 1.29 [45], allowing up to 2 mismatches in the 6 bp barcodes. Barcodes were removed and sequences trimmed to 80 bp to remove low-quality bases from the end of the reads. These cut reads were then aligned using the software package PyRAD 3.0.6 [46], and further cleaned by removing reads that had more than 5 bp with a PHRED score of less than 20. As PyRAD is a de novo assembly pipeline, it is particularly effective for large-scale, divergent sequences which may cause some sequences to 'drop out' due to indels or mutations [46].

To account for the effect of proportions of missing data in the alignment on downstream analyses [20,47], two alignment criteria were used based on minimum coverage: a strict dataset with a minimum coverage of 40 individuals per locus (approx. 89%), and a relaxed dataset with a minimum coverage of 32 individuals per locus (approx. 70%).

2.3. Sequence divergence analysis

The genetic distance between concatenated ddRAD sequences from different individuals was calculated using an uncorrected (*P*) distance matrix generated in PAUP* 4 [48], with mean values across lineages (species and, in the case of *N. vittata* and *N. pygmaea*, geographical groups and pairs of lineages). This was done for each filtered dataset (strict and relaxed) to compare the effects of the filtration process on downstream analyses.

2.4. Phylogenetic analysis

In order to determine evolutionary relationships among pygmy perch samples, maximum-likelihood (ML) phylogenies were estimated using RAxML for both the strict (4381 loci) and relaxed (13 991 loci) dataset [49]. RAxML remains one of the best options available to estimate a ML phylogeny in genome-wide datasets [49]. This was done using rapid hill-climbing and 1000 resampling estimated log-likelihood [50] bootstraps with a GTRGAMMA model, using the online service CIPRES and the supercomputer XSEDE [51]. Invariable sites (+I) were not included in the model as they are intrinsically linked to other factors such as rate categories and are unlikely to be true biologically (as all sites within a sequence are likely to have some, if negligible, mutation rate) [52,53]. This model composition is the recommended choice by the developers of RAxML [49].

To test for consistency of results, Bayesian estimation of the pygmy perch phylogeny was also conducted using PhyloBayes 4.1 [54], using a CAT substitution model and a discrete gamma distribution of site rate heterogeneity model. A Markov chain Monte Carlo chain was run for a total of 1480 cycles consisting of 91 584 tree generations until the log likelihood demonstrated a stable equilibrium. Bayesian analysis was limited to the strict dataset due to computational limitations of the software. The resultant phylogenetic trees of both methods were visualized using MEGA 7 [55], using *B. porosa* as the outgroup.

2.5. Species delimitation

Species were delimited using the software package Bayesian Phylogenetics and Phylogeography (BPP 3.2) [56]. Delimitation was limited to the smaller, strict dataset due to computational constraints. BPP uses a coalescent modelling method to estimate species limits in a Bayesian framework, and estimates multiple species tree hypotheses simultaneously with the delimitation to create a more robust and statistically sound analysis by providing posterior probabilities of both.

The unguided species delimitation method (analysis A11 as described within the manual) was used to simultaneously estimate both the species tree and the delimitation of the designated species, avoiding biases associated with using a fixed guide phylogeny [56,57]. As BPP can only coalesce but not split input species, all three populations of *N. vittata* were input as separate species on account of their divergent and paraphyletic nature in the maximum-likelihood phylogeny (see Results below), giving a total of 11 input species. Biologically reasonable priors for the gamma distributions of population sizes of lineages (θs) and the divergence time of the root of the phylogeny (τs) were adjusted until convergence of species delimitations were found over two runs for both incorporated algorithms ($n = 4$) to confidently provide consistent results. Other divergence time parameters were estimated using a Dirichlet prior [58, eqn (2)].

A lower number of replications was initially used for computational efficiency [10 000 burn-in + (100 sample freq * 1000 *n*sample) = 110 000 generations]. Priors were progressively altered upwards from G(2, 0) for the τs prior and G(30, 1) for the θs prior until convergence of results was found. Convergence of the runs was found using a τs prior of G(10, 10 000), with Dirichlet prior for other divergence times, and a θs prior of G(2, 100), using cleandata = 0 to account for gaps and ambiguous characters in the sequence. Once convergence was found, a final species delimitation and species tree estimation analysis was done using a larger number of generations under each algorithm [2 replicates per algorithm; 10 000 burn-in + (5 sample freq * 98 000 *n*sample) = 500 000 generations].

2.6. Molecular dating

Divergence time estimates were obtained using a maximum clade credibility (MCC) pipeline. The MCC method allows for the summation of thousands of phylogenetic trees into a single, consensus tree by selecting the tree containing the most common clades, rather than building a tree from each most common clade that may never have been generated in the initial analysis [59]. As ambiguous or heterozygous sites can drastically impact divergence time estimates [60,61], haplotypes were generated

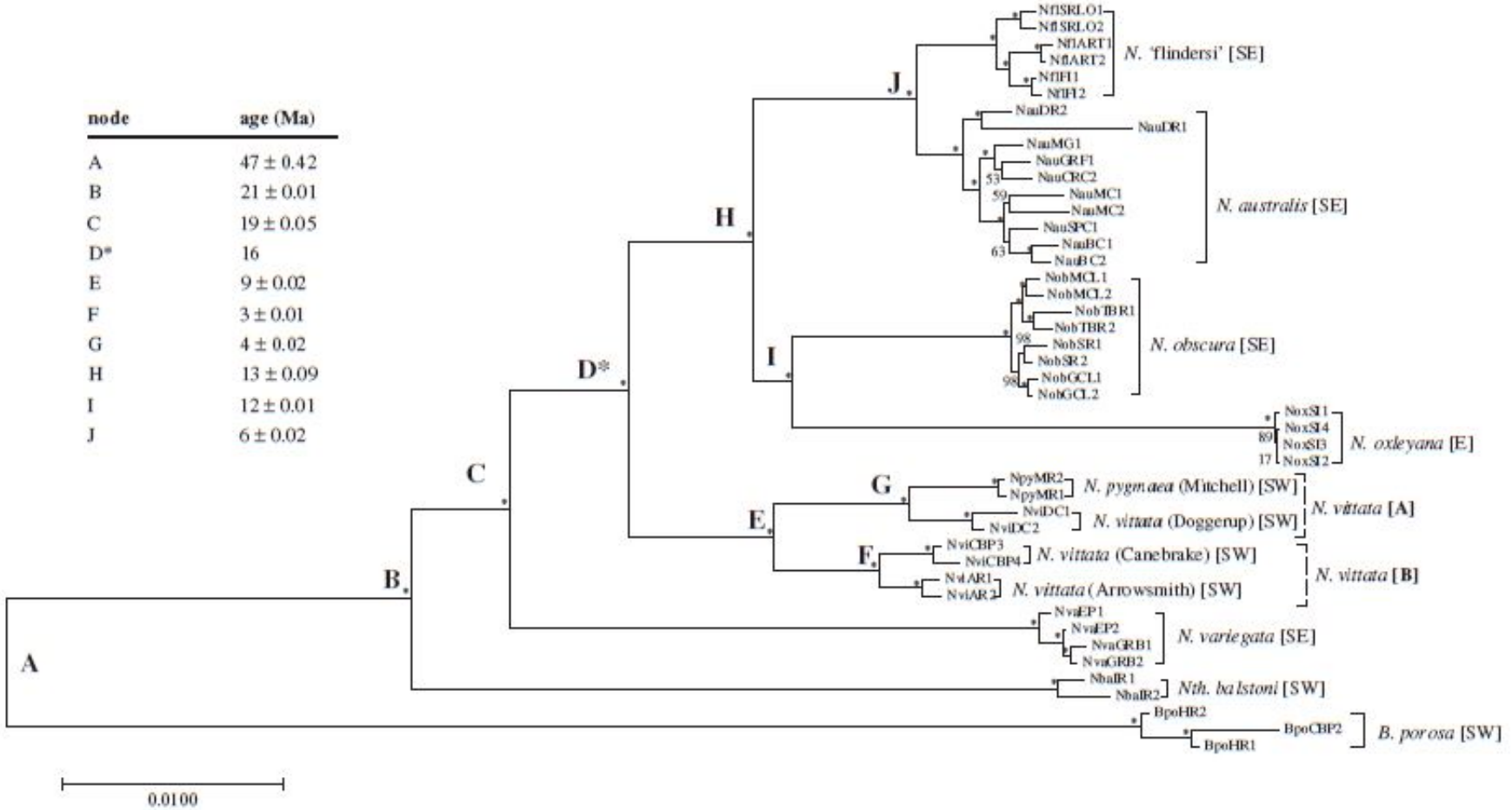


Figure 2. Maximum clade credibility (MCC) based on 3000 random repeated haplotypes (RRHS) of a 1 075 734 bp sequence (13 991 concatenated ddRAD loci) and divergence estimates from r8s. Bootstraps are calculated from 1000 bootstraps per RRHS summarized using the 'best tree' output of RAxML. All 3000 'best trees' were summarized using the consense function of ExaBayes and the resultant phylogeny plotted in MEGA 7.0. All major species and population divisions had 100% bootstrap support; all bootstraps with 100% support are represented by small asterisks. Codes for individuals and localities relate to the abbreviations in table 1. Node D denotes the node used to calibrate for divergence time estimation. Divergence estimates are reported for the MCC phylogenetic tree, with associated confidence intervals from 100 subsampled RRHS trees (± s.d.).

using repeated random haplotype sampling (RRHS), which randomly assigns a particular base to each heterozygous site within an unphased sequence [60]. A total of 3000 sets of RRHS sequences were created for divergence estimates using an edited RRHS Java command line script provided by [60], with each set of haplotype sequences analysed in a ML phylogeny using 1000 bootstraps in RAxML. The 'best tree' output for each RAxML run was unrooted using the R package *ape* [62] and summarized into a single MCC tree using the *consense* protocol of ExaBayes [63]. Due to issues with the formation of a polytomy within the root of the phylogeny which prevents divergence time estimation, a *B. porosa* outgroup sample (BpoHR1) was removed from the tree using the *prune* command of *ape*. This had no effect on branch lengths or topology within the rest of the phylogeny.

The software r8s 1.81 was then used to estimate divergence times in the MCC phylogeny [64]. R8s estimates absolute rates of divergence times across branches of a given phylogenetic tree based on branch lengths, by using an estimation of relative rate across branches and relating this to a set age calibration of at least one given node [64]. A single calibration point was placed at the split between the eastern clade of *N. australis–N. obscura–N. oxleyana* and the western clade of *N. vittata–N. pygmaea* at 14–16 Ma. This date represents the formation of the current day Nullarbor Plain which is associated with the cessation of potential connectivity between the eastern and western freshwater fauna [27]. Divergence times for each node were estimated using a penalized-likelihood model under a truncated Newton algorithm [65], which uses a parametric branch substitution rate model with a nonparametric roughness penalty [64]. A 'smoothing parameter' determines the contribution of the roughness penalty aspect: a cross-validation procedure was used to determine the best value between $\log_{10}$ 0 and $\log_{10}$ 100. The optimum $\log_{10}$ smoothing parameter of 44.00, with a chi-square error of 139 049.74, was used to estimate divergence times for defined major species divisions (described in figure 2). Confidence intervals for the age estimations were calculated by sub-sampling 100 independent RRHS trees from the 3000 trees dataset. These trees were randomly selected from a pool of 2388 trees that showed matching topology to the MCC phylogenetic tree (determined using the all.equal.phylo function in *ape* [62]). Each tree was dated using the same methods as above and the distribution of node ages for all major species divisions calculated using the *profile* function of r8s. Using topologically identical trees with varying branch lengths to estimate confidence intervals is the suggested approach within the r8s manual [64]. Estimates for mutation rates across all branches, as well as the average across the phylogeny, were also calculated using r8s. Including fossil-calibrated reconstructions to estimate divergence time in this study was hampered due to the absence of known suitable fossils from pygmy perches or closely related taxa. Although evolutionarily divergent fossils exist for Centrarchidae—a freshwater family endemic to North America thought to have split from the Australian Percichthyidae over 61 million years ago [66]—applying poorly constrained fossil evidence with questionable placement in the phylogeny is not a recommended option [67,68].

2.7. Ancestral area reconstruction analysis

In order to statistically evaluate competing biogeographic hypotheses about multiple dispersal or vicariance events (see Discussion), we estimated ancestral areas using the R package *BioGeoBEARS* [69]. As *BioGeoBEARS* requires a time-calibrated ultrametric tree as an input file, the MCC phylogenetic tree was first collapsed down to species using the collapseTree command of the R package *phytools* [70] and calibrated using the MCC r8s divergence estimates and the chronos command of *ape*. Species were assigned to one of two biogeographic regions (East or West) with a max range of 2 (i.e. both regions) for ancestral lineages. Ancestral areas were estimated under all six available models (DEC, DIVA-LIKE and BAYAREA-LIKE, as well as their +J counterparts) and compared using the Akaike information criterion (AIC) to assess the fit of each model.

3. Results

3.1. Sequence filtering and alignment

The strict dataset (approx. 11% missing data) of 4381 ddRAD loci produced a concatenated sequence of 334 936 bp, and the relaxed dataset (approx. 30% missing data) of 13 991 ddRAD loci produced a concatenated sequence of 1 075 734 bp (table 2). Both datasets contained a large number of informative sites, with 35 856 parsimony-informative sites within the 41 067 single nucleotide polymorphisms (SNPs) of the strict dataset, and 123 252 parsimony-informative sites within the 142 476 SNPs of the relaxed dataset.

Table 2. Summary data for ddRAD loci and percentage of missing data (%) based on two separate filtering criteria. Strict = minimum of 40 individuals (approx. 89%) per locus; relaxed = minimum of 32 (approx. 70%) individuals per locus. PIS = parsimony-informative sites.

sequence subset	factor	strict dataset	relaxed dataset
total sequence	sequence length (bp)	334 936	1 075 734
	number of ddRAD loci	4381	13 991
	number of SNPs	41 067	142 476
	number of PIS	35 856	123 252
all individuals	mean no. of ddRAD loci	4156	12 043
	missing data (%)	5.09	13.92
ingroup only	mean no. of ddRAD loci	4278	12 459
	missing data (%)	2.34	10.95
outgroup only	mean no. of ddRAD loci	2470	6220
	missing data (%)	43.62	55.54

3.2. Sequence divergence analysis

For simplicity, we discuss here only the relaxed dataset sequence divergence values as both datasets demonstrated similar patterns across the group, although exact values for each pairwise comparison were slightly underestimated in the strict dataset (table 3). *Nannoperca vittata* was composed of two divergent groups (referred herein as 'superclades' [A] and [B]). Superclade A consisted of the Doggerup Creek lineage of *N. vittata* and *N. pygmaea*; superclade B consisted of the remaining two *N. vittata* lineages (Arrowsmith River and Canebrake Pool); these superclades were highly divergent to one another ($P = 1.26\%$). Marked divergence was also observed within *N. australis* (table 3). These two taxa had the greatest levels of genetic heterogeneity, with higher intraspecific mean uncorrected (*P*) genetic distance than all other intraspecific comparisons (0.2–0.44%). In contrast, *N. oxleyana*, *N. variegata* and each individual *N. vittata* lineage (DC, AR, CBP) demonstrated remarkably low within species genetic differentiation (0.01–0.04%).

Furthermore, individual clades of *N. vittata* within each superclade showed greater genetic distance to one another (0.36–0.55%) than that found within any other species. These levels of divergence approached the genetic distance between previously suggested species, such as between *N.* 'flindersi' and *N. australis* (0.68%). When comparing *N. vittata* [A] and *N. vittata* [B], the divergence seen between these (1.26%) exceeds the genetic differentiation between *N.* 'flindersi' and *N. australis*. Finally, the eastern species *N. oxleyana* appeared as the most divergent *Nannoperca* lineage, showing the greatest pairwise genetic distance to all other pygmy perches (2.50–4.59%). Similar levels of divergence were shown for *Nth. balstoni* to all other pygmy perches (2.93–4.59%).

3.3. Phylogenetic analysis

Visual inspection of the ML phylogenies demonstrated no effect of missing data or loci number, with both datasets showing identical topology for all major clades within pygmy perches and similar branch lengths (electronic supplementary material, figures S1 and S2). Similarly, Bayesian analysis of the strict dataset showed no variation in topology compared to the ML phylogeny of the same dataset and returned very high nodal support (0.95-1 posterior probability for all nodes of tree; electronic supplementary material, figure S1). The MCC tree demonstrated near identical topology to the previously generated trees, albeit with significantly higher bootstrap values and a greater difference in branch lengths (figure 2). This difference is expected due to the pseudo-phasing of data and larger number of bootstraps (3000 individual trees). Topological differences were limited to changes in the position of samples NauMC1, NvaEP2 and BpoHR1 within their respective lineages. Distinctive branch length variations were observed in Darby River *N. australis* samples (NauDR), within individual *N. vittata* locality groups (NviAR, NviDC and NviCBP), within *Nth. balstoni* and the outgroup *B. porosa*. Bootstrap support varied slightly across datasets, with the strict dataset giving higher support for some nodes (within *N. australis*, *N.* 'flindersi' and *N. variegata*), but lower support for others (within and between

Table 3. Uncorrected (*P*) mean percentage genetic distance matrix of pygmy perch lineages based on two genomic datasets. The top right section of the distance matrix represents the 'strict' (4381 loci) dataset while the bottom left section represents the 'relaxed' (13 991 loci) dataset. Comparisons between a lineage and itself (*italicized*, in boxes) represent within-lineage mean genetic distance (denoted as 'relaxed/strict'). Comparisons between a *N. vittata* superclade (A or B) and its containing lineages are not shown.

lineage		1	2	3	4	5	6	7	8	9	10	11	12	13
N. 'flindersi'	1	*0.17/0.14*	0.53	1.47	2.23	2.07	1.91	1.99	1.74	1.85	1.79	2.87	3.24	5.61
N. australis	2	0.68	*0.2/0.15*	1.34	2.07	1.92	1.76	1.84	1.58	1.69	1.64	2.61	3.09	5.85
N. obscura	3	1.89	1.76	*0.09/0.06*	2.07	2.04	1.88	1.96	1.75	1.82	1.78	2.85	3.25	5.63
N. oxleyana	4	2.73	2.59	2.5	*0/0*	2.77	2.59	2.68	2.48	2.55	2.51	3.57	3.96	6.26
N. pygmaea	5	2.67	2.53	2.63	3.42	*0.01/0.01*	0.44		1.03	1.02		2.76	3.17	5.63
N. vittata (DC)	6	2.5	2.36	2.46	3.25	0.55	*0.03/0.02*		0.91	0.9		2.42	2.98	5.45
N. vittata [A]	7	2.58	2.44	2.55	3.33			*0.44/0.44*			0.97	2.59	3.08	5.54
N. vittata (CBP)	8	2.3	2.16	2.32	3.1	1.33	1.2		*0.01/0.01*	0.28		2.51	2.93	5.41
N. vittata (AR)	9	2.43	2.29	2.39	3.17	1.31	1.19		0.36	*0.04/0.02*		2.55	2.95	5.45
N. vittata [B]	10	2.37	2.23	2.35	3.14			1.26			*0.25/0.28*	2.53	2.94	5.43
N. variegata	11	3.46	3.2	3.44	4.17	3.31	2.88	3.09	3.07	3.1	3.08	*0.03/0.02*	3.38	5.71
Nth. balstoni	12	3.89	3.75	3.9	4.59	3.76	3.58	3.67	3.53	3.57	3.55	3.86	*0.04/0.03*	5.55
B. porosa	13	6.23	5.99	6.24	6.81	6.28	6.09	6.18	6.07	6.09	6.08	6.24	6.14	*0.2/0.17*

Table 4. Statistical evaluation of biogeographic models implemented in *BioGeoBEARS*. Extant species were assigned to an eastern or western geographical range with a maximum range of 2 (i.e. ancestors could occupy both areas) for all analyses, with the accuracy of each model assessed and compared using the Akaike information criterion (AIC).

model	log likelihood	dispersal (*d*)	extinction (*e*)	founder event (*J*)	AIC
DEC	−5.38	1.00×10^{-12}	1.00×10^{-12}	0	14.75
DEC+J	−4.36	3.50×10^{-12}	1.60×10^{-10}	0.1	14.71
DIVALIKE	−6.14	0.0083	1.00×10^{-12}	0	16.27
DIVALIKE+J	−3.96	1.00×10^{-12}	1.00×10^{-12}	0.097	13.93
BAYAREALIKE	−9.29	1.00×10^{-12}	0.029	0	22.58
BAYAREALIKE+J	−4.96	1.00×10^{-12}	1.00×10^{-12}	0.13	15.92

N. obscura and *N. oxleyana*). This most likely reflects a bias towards *N. australis*-specific loci in the strict dataset due to the higher number of samples for this taxon.

Four of the five species from eastern Australia (*N. australis*, *N.* 'flindersi', *N. obscura* and *N. oxleyana*) were reciprocally monophyletic. This clade was sister to a western clade comprised of *N. vittata* (including all putative cryptic forms within this species) and *N. pygmaea*. The combined eastern and western clade was sister to an eastern species, *N. variegata*, and these together were sister to a western species, *Nth. balstoni*. Additionally, the phylogeny included the distinct separation of *N.* 'flindersi' from its sister taxon *N. australis* and divergence of *N. pygmaea* and clades of *N. vittata*.

Finer scale population-level patterns could also be observed within several taxa, with *N.* 'flindersi' separating into one Victorian (Snowy River Lagoon, Orbost) and one Tasmanian (Flinders Island + Anson River) clade. Similarly, *N. australis* separated into three clades: one composed of Murray–Darling Basin individuals, one of fish from the southeast coast and one of a distinct population at Darby River. Shallower phylogeographic patterns were also observed within *N. obscura*, with geographical clusters of populations forming distinct clades.

3.4. Species delimitation

Coalescent modelling of species delimitation resulted in a total of 11 species within the phylogeny (posterior probability = 1 for all runs), with each clade within *N. vittata* being recognized as an independent species. All previously described or suggested species were similarly recognized as independent of one another, including *N.* 'flindersi' and *N. pygmaea*. The species tree output by BPP matched the topology of all ML and Bayesian trees and the MCC tree.

3.5. Molecular dating

Divergence time estimates from r8s revealed that pygmy perches are an ancient lineage, with the root of the group estimated at 20 (±0.01) Ma (figure 2). This time of origin is similar to other 'ancient' teleosts from the northern hemisphere [71,72]. Each inferred east–west divergence had a different estimated age within the pygmy perch radiation. Most lineages within the major eastern clade showed comparatively recent divergences, except for *N. oxleyana* which showed an older divergence time of 12 (±0.01) Ma from *N. obscura* (node I), reflected by the high genetic distance to all other pygmy perches. More closely related species had relatively young ages such as 3–4 Ma (nodes F and G) within the *vittata* clade and 6 (±0.02) Ma between *N. australis* and *N.* 'flindersi' (node J). All times of divergences were found to be within or close to the ranges proposed by Unmack *et al.* [27]. All estimations of node ages were highly consistent, as indicated by low standard deviation of all node estimations using 100 RRHS trees (electronic supplementary material, table S1; figure 2). This is expected as individual RRHS trees do not excessively vary in branch lengths and indicates that even pseudo-phasing of the data can produce consistent results.

Estimations of rate variation across sequences had a mean rate of 9.67×10^{-4} ($\pm 4.81 \times 10^{-7}$ standard deviation) substitutions per site per million years. Very small differences in rate variation were seen across lineages, as highlighted by the low standard deviation of total rate variation.

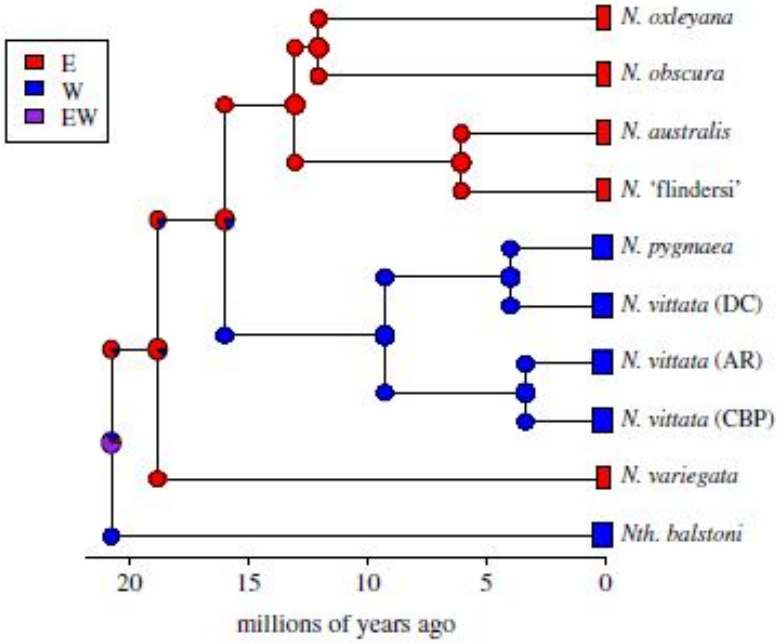


Figure 3. The most supported ancestral area reconstruction model (DIVALIKE+J), estimated within the R package *BioGeoBEARS*. Colours denote geographical range of lineages, with pie charts representing the relative probability of geographical range of ancestors. Lineages were collapsed down to the species level and the phylogeny time calibrated based on the MCC divergence estimates from r8s.

3.6. Ancestral area reconstruction

Evaluation of the six potential biogeographic models within *BioGeoBEARS* using AIC suggested that the DIVALIKE+J model was best representative of the data (table 4, figure 3). This model demonstrated low probability of dispersal or extinction events (d and $e = 1 \times 10^{-12}$) but showed an effect of a founder event ($J = 0.0972$). Furthermore, nearly all ancestral lineages demonstrated a likely eastern geographical range (excluding the ancestor of the two pygmy perch genera). These results suggest that a founder event (such as a rare, long-distance dispersal event) from the east was a likely driving factor in the colonization of the major western clade containing the *N. vittata* species complex (including *N. pygmaea*).

4. Discussion

We performed a genome-wide study using phylogenomic and species delimitation methods to clarify evolutionary and biogeographic history and to assess cryptic diversification in all known lineages of pygmy perches. We confirmed the biological relevance of previously inferred phylogenies and recently suggested species. Additionally, with this much larger dataset, we provided greater phylogenetic resolution and discovered extra cryptic species. Our findings reveal complex biogeographic patterns in southern Australia and point to the importance of in-depth taxonomic analyses for appropriate conservation management of threatened biodiversity.

4.1. Phylogenomics of pygmy perches

Our phylogenetic trees corroborate the combined tree presented by Unmack *et al.* [27] and went further by providing substantially improved phylogenetic resolution. There was a major improvement in bootstrap support for all nodes in the genomic phylogeny, highlighting the ability of ddRAD to avoid locus-specific biases such as mitochondrial introgression, which may have obscured some findings of the previous study. These phylogenetic trees enabled the resolution of conflicting nodes across different markers from Unmack *et al.* [27]: all phylogenies here indicate that *N. obscura* is sister to *N. oxleyana* and not to *N. australis*, and that *N. variegata* is the first branching lineage within *Nannoperca*. Furthermore, the paraphyletic nature of *N. vittata* further suggests that the taxonomy of the species is currently unresolved. Of particular note is the high level of divergence observed between individual, geographically isolated

lineages within *N. vittata* reflected in both the genetic distance and divergence estimates results (table 3, figure 2). The level of divergence between these *N. vittata* lineages is similar to that between *N. vittata* and *N. pygmaea*, possibly indicating cryptic species within *N. vittata* (see below). This study also enabled a much greater resolution of intraspecific population structure within *N. australis* and *N.* 'flindersi' than suggested by prior studies (e.g. [36]).

A species delimitation framework, in combination with other qualitative genetic analyses, supported the proposed *N. pygmaea*, which was previously only based on morphological assessments and limited allozyme data [27,73]. We also revealed two new cryptic lineages within *N. vittata*. It is recommended that a more thorough analysis of *N. vittata* should be conducted in accordance with an integrative taxonomic approach [74–76] before taxonomic changes are implemented and used in conservation management [77,78]. Assessing all known populations of *N. vittata*, with more comprehensive sampling than achieved here, may reveal the full geographical range and ecological nature of the delimited species. Additional research focused on morphological and ecological divergences is important for establishing species boundaries which would help resolve their taxonomy. Clearly, further taxonomic studies of pygmy perches are needed to better inform their conservation management, particularly for those lineages without formal description and legislative protection such as *N.* 'flindersi'. An understanding of species limits and local adaptation would provide a framework for targeted conservation actions in pygmy perches, such as genetic rescue [79].

While species delimitation provides a statistical framework for testing hypotheses of species identity, there are limitations in the outcomes. Simulations have suggested that BPP has a tendency to exaggerate the number of species within a phylogeny, providing high support for divergent lineages which are not biologically true species [80]. Despite this criticism, a solely genetic basis for species description has rarely (if ever) been used [81]. Taxonomic changes require a complement of various forms of analyses, including morphological, ecological and behavioural data [12,76] to infer species identity and reproductive isolation. Thus, while we do not present our delimitations as fully putative species, we conservatively suggest that those identified may possibly reflect truly cryptic species.

4.2. Biogeographic interpretation

Our study provides support for multiple east–west movement events across southern Australia involving both older (i.e. *Nth. balstoni* and *N. variegata*), as well as younger pygmy perch lineages. The ancestral area reconstruction analysis did not support a simple model of vicariant divergence of eastern and western pygmy perches as the result of the Nullarbor Plain. Instead, the most supported model suggested a founder event from an eastern ancestor into the west (the predecessor of the *N. vittata* species group). Given that the climatic and geological changes associated with the formation of the Nullarbor Plain were likely gradual, it is possible that a peripheral population of the eastern ancestor became isolated during the Miocene. Alterations to hydrology and geology, such as river capture, may have disconnected this peripheral population from the rest of the species range, causing it to divert westward and found the western pygmy perch group. This is particularly relevant for freshwater species which inhabit dendritic systems as geological and climatic changes can significantly alter hydrological connections, thereby having significant impacts on phylogeographic structure and interpretation [82]. Thus, we suggest that range expansion, followed by subsequent isolation due to a vicariant barrier, is likely the major mechanism driving this geographical separation. Nonetheless, we assume that the formation of the Nullarbor Plain was still pivotal to the complete separation between younger western (*N. vittata*) from younger eastern (*N. australis*, *N. obscura* and *N. oxleyana*) lineages. This scenario is similar to that proposed by Unmack *et al.* [27], who suggested that the lack of a singular east–west split of pygmy perches was indicative of multiple migrations across central southern Australia prior to the formation of the Nullarbor Plain (deemed the 'Multiple Invasion Hypothesis'; figure 4). While projections of historical climate suggest that the hydrology of southern Australia may have been suitable for the migration, little support for multiple migrations has been found in other aquatic taxa [2,29,32]. Of these, many are more dispersive than pygmy perches, suggesting they should likewise have been able to migrate across southern Australia multiple times. While this may seem counterintuitive, little is known about the historic ecology of aquatic biota to propose a mechanism or reason for this disparity. The historically widespread distributions of *N. vittata* and *N. australis* have been suggested to predispose them to being tolerant to a range of habitats [27]: thus, it is hypothetically possible that pygmy perches were able to tolerate intermediate habitats between the east and west. Additionally, historical metapopulation dynamics suggested for pygmy perches may have allowed them to respond to temporally unfavourable habitats [40].

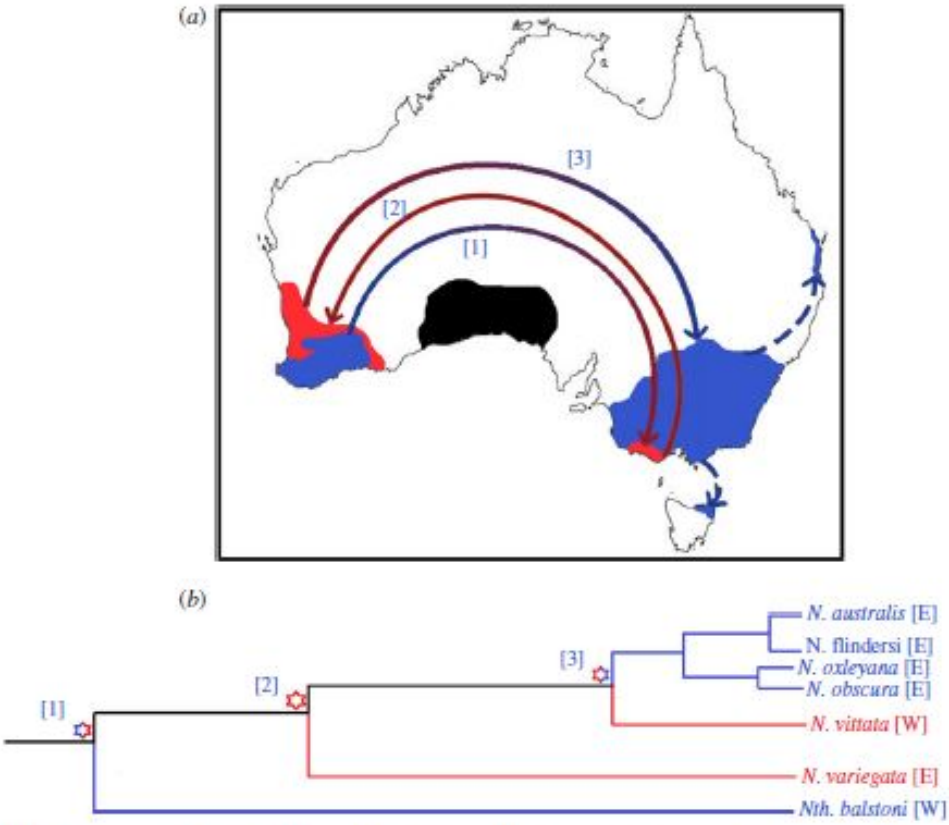


Figure 4. Hypothetical model of the Multiple Invasion Hypothesis and the biogeography of pygmy perches. (*a*) Graphical representation of Multiple Invasion (under the maximum number of theoretical migrations) as a possible mechanism for the biogeography of pygmy perches. The position of the Nullarbor Plain is demonstrated in black. Migrations of pygmy perches are indicated by the directionality and colour of the arrows, with contemporary distributions demonstrated by the filled regions. Dashed arrows represent secondary migrations (into *N. oxleyana* and *N.* 'flindersi') without crossing the continent. (*b*) Phylogenetic diagram of geographical divergences in pygmy perches under Multiple Invasion. Nodes representing migration events are denoted within the phylogeny by asterisks, with the numbers corresponding to a particular migration on the map. Divergence [3] represents the last possible migration event before the arrival of the Nullarbor Plain as a barrier to dispersal.

However, the Multiple Invasion Hypothesis for pygmy perches was later criticized, with Ladiges *et al.* [83] instead proposing that geographical paralogy may explain the lack of a singular east–west split. Geographical paralogy is when two (or more) independent but closely related lineages experience a biogeographic split at the same time (figure 5). In the case of pygmy perches, the geographical split of one lineage containing the ancestor of *Nth. balstoni* and *N. variegata*, and one lineage containing the ancestor of *N. australis*, *N. obscura*, *N. oxleyana* and *N. vittata*, would produce a similar phylogenetic pattern to that shown here. These independent divergence events could have been vicariantly caused by the formation of the Nullarbor Plain, or could reflect independent splitting events (both in mechanism and timing). There are unfortunately no known fossil records of pygmy perches across southern Australia to demonstrate the necessary ubiquity of pygmy perches prior to these splits [84].

The Multiple Invasion Hypothesis is instead more strongly supported than geographical paralogy by the current study. This is based on the lack of a sister relationship between *N. variegata* and *Nth. balstoni* required for geographical paralogy (electronic supplementary material, figures S1 and S2; figures 2 and 5). This lack of sister group relationships would likely only be incorrect if there are unaccountable artefacts from extinction of lineages or difficulties in detecting relationships in anciently diverged lineages. Additionally, the ancestral area reconstruction modelling did not suggest a simple vicariant separation of eastern and western pygmy perches from a widely distributed ancestor, but instead a founder event (e.g. long-distance dispersal) from the east to the west (figure 3). While there are significant limitations and assumptions with historical biogeographic analysis, alternative models that included many widespread ancestors were not as well supported by the data (table 4).

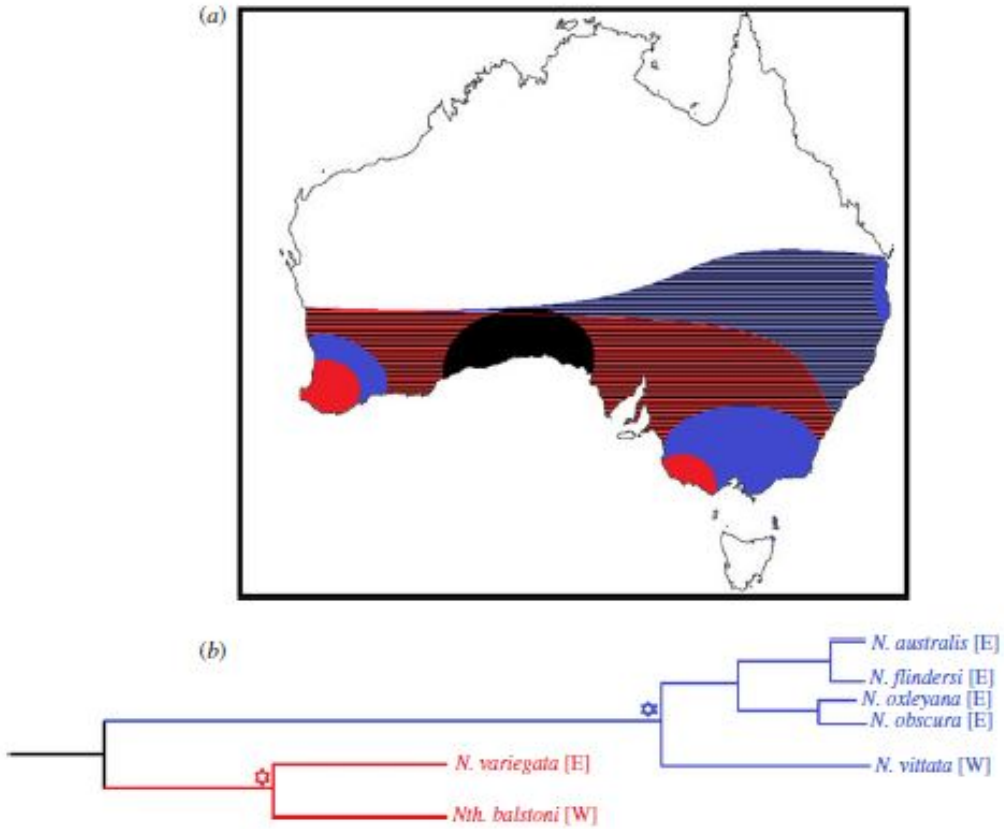


Figure 5. Hypothetical model of geographical paralogy and the biogeography of pygmy perches. (*a*) Graphical representation of geographical paralogy as a possible mechanism for the biogeography of pygmy perches. The position of the Nullarbor Plain is demonstrated in black. Speculative historic ranges of the two ancestral lineages required for geographical paralogy to occur are demonstrated by the striped regions, with contemporary distributions demonstrated by the filled regions. (*b*) Phylogenetic diagram of geographical divergences in pygmy perches under geographical paralogy. Nodes representing E/W divergences are denoted within the phylogeny by asterisks, with the second (blue) divergence representing the arrival of the Nullarbor Plain as a barrier to dispersal.

4.3. Biodiversity hotspot in southwestern Australia

The highly divergent nature of populations in endemic species (*N. vittata* spp. and *Nth. balstoni*) of southwestern Western Australia (figure 2) indicates that this region is a significant driver of evolution and speciation. This is consistent with other studies that showed high levels of endemism and species diversity within the region, leading to its internationally and nationally recognized status as a biodiversity hotspot [28,29]. Our findings reinforce the high conservation value of the region.

The high levels of *in situ* speciation in southwestern Australia have often been linked to historical climate changes [29,30]. The drivers of population diversification within pygmy perches probably relate to allopatric isolation as a result of fragmentation of wetter regions, with all species of western pygmy perches limited to areas of higher annual rainfall (greater than 600 mm) [29]. Allopatric speciation has been noted throughout southwest Western Australia for a range of taxonomic groups with limited dispersal capabilities, including many invertebrate groups (e.g. spiders, isopods and crayfishes) and some frog species [29,85,86].

It is likely that a combination of long-term isolation associated with formation of the Nullarbor Plain, as well as *in situ* speciation within southwest Western Australia, are major factors accounting for the diversification and evolution of pygmy perches. Further assessments focusing on biogeographic history are required to elucidate the complexity of the pygmy perch phylogeny. In this regard, species distribution modelling (SDM) appears as a robust approach to addressing competing biogeographic theories when combined with statistical phylogeographic methods [87,88]. For pygmy perches, SDM could be applied to southwestern Australia to determine what role environmental factors played in driving the divergence of *N. vittata* lineages.

4.4. Conservation concerns

The identification of cryptic species has profound implications for their conservation management, particularly for legislation. While *N. vittata* is currently unlisted within the IUCN Red List, our results suggest that *N. vittata* is potentially three independent species, each with a small number of narrow endemic populations that are at risk of losing genetic diversity and local extirpation. Thus, a thorough assessment of *N. vittata* throughout their currently recognized range and associated species descriptions are required to fully elucidate the taxonomic intricacies within this lineage and to incorporate these into conservation management practices.

Our study also identified intraspecific lineages that are not yet divergent enough to be considered as different species (as opposed to the divergent lineages within *N. vittata*). Some of these lineages correlated with previously identified ESUs [89], such as in *N. obscura* [42]. Additional intraspecific genetic structure in *N.* 'flindersi' and *N. obscura* suggests that conservation managers should recognize different conservation units. This includes three geographically isolated lineages of *N.* 'flindersi'. Given the isolation of these lineages by unpassable saltwater barriers (Bass Strait) and environmental differences between the regions [90,91], they may be adaptively divergent and are perhaps precursors of incipient species. A thorough assessment of population structure using genome wide data in each region is required for improving and defining appropriate units for management within species.

Both the taxonomic issues demonstrated by this study, as well as the low genetic diversity and dispersal capabilities of many pygmy perches [33,37,39,40,42], raise concerns for their conservation management under future climate change. Their low diversity potentially reduces their capacity to adapt to changing environments, and their low dispersal capabilities inhibit them from easily moving to more favourable environments. In addition, low genetic variation and human-induced adaptive divergence in habitat fragments threaten the species in the Murray–Darling Basin [39] (see also [40]). Genetic-based captive breeding and restoration efforts have already been required for *N. australis* and *N. obscura* in the lower Murray River [37,92], and local extirpations in various parts of the Murray–Darling Basin are already recorded [38,40]. Our novel phylogeographic findings are expected to inform future genetic-based breeding programmes in pygmy perches, particularly for *N. vittata* spp., as such programmes should take into account the evolutionary history and historical demography of threatened lineages to maximize their success [37].

We used genomics to identify and characterize biodiversity patterns across southern Australia and within the pygmy perches. Specifically, we aimed to improve our understanding of southern Australian biogeography and the taxonomy of pygmy perches, with associated implications for conservation. Using a powerful dataset, we have resolved previously uncertain phylogenetic relationships in pygmy perches, identified and confirmed additional cryptic species within *N. vittata* through a species delimitation framework, and estimated divergence times for the pygmy perch phylogeny using a ddRAD molecular clock. These findings have strengthened our understanding of biological concepts such as southern Australian biogeography, the need for further taxonomic research into pygmy perches and important conservation suggestions. The indication of high levels of cryptic speciation within southwest Western Australia further corroborates its high conservation importance and biodiversity hotspot identity, while the identification of independent species suggests modification for current conservation statuses and management.

Ethics. Collections were obtained under permits from various state fisheries agencies and research was under Flinders University Animal Welfare Committee approvals E313 and 396.

Data accessibility. The concatenated sequence input files under both alignment conditions are available in PHYLIP format in Dryad (http://dx.doi.org/10.5061/dryad.3d6g1) [93].

Authors' contributions. S.J.B. contributed to all sections of data analysis as well as drafting the manuscript. R.L., J.S.-C. and C.J.B. contributed to laboratory work. F.M.C.B.D., J.S.-C., C.J.B., C.R.M.A., P.J.U. and L.B.B. contributed to data analysis and interpretation. P.J.U. contributed to fieldwork and editing the manuscript. L.B.B. conceived and coordinated the study, and contributed to drafting the manuscript.

Competing interests. The authors declare no competing interests.

Funding. This project was funded by an Australian Research Council Future Fellowship (FT130101068 to L.B.B.).

Acknowledgements. We acknowledge the many researchers that provided fish samples or participated in field expeditions, especially Michael Hammer, David Morgan and Mark Adams. This work received logistic support from Flinders University, University of Canberra, South Australian Museum. We also thank Minami Sasaki for assisting with tissue sampling and DNA extractions.

References

1. Avise JC. 2003 The best and worst of times for evolutionary biology. *BioOne* **53**, 247–255. (doi:10.1641/0006-3568)
2. Morgan MJ, Roberts JD, Keogh JS. 2007 Molecular phylogenetic dating supports an ancient endemic speciation model in Australia's biodiversity hotspot. *Mol. Phylogenet. Evol.* **44**, 371–385. (doi:10.1016/j.ympev.2006.12.009)
3. Lorenzen ED, Heller R, Siegismund HR. 2012 Comparative phylogeography of African savannah ungulates. *Mol. Ecol.* **21**, 3656–3670. (doi:10.1111/j.1365-294X.2012.05650.x)
4. Dawson MN *et al.* 2013 An horizon scan of biogeography. *Front. Biogeogr.* **5**, fb_18854.
5. Beheregaray LB. 2008 Twenty years of phylogeography: the state of the field and the challenges for the southern hemisphere. *Mol. Ecol.* **17**, 3754–3774. (doi:10.1111/j.1365-294X.2008.03857.x)
6. Avise JC. 1992 Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos* **63**, 62–76. (doi:10.2307/3545516)
7. Arbogast BS, Kenagy GJ. 2001 Comparative phylogeography as an integrative approach to historical biogeography. *J. Biogeogr.* **28**, 819–825. (doi:10.1046/j.1365-2699.2001.00594.x)
8. Hedin M, McCormack M. 2017 Biogeographical evidence for common vicariance and rare dispersal in a southern Appalachian harvestman (Sabaconidae, *Sabacon cavicolens*). *J. Biogeogr.* **44**, 1665–1678. (doi:10.1111/jbi.12973)
9. Smith BT, Seeholzer GF, Harvey MG, Cuervo AM, Brumfield RT. 2017 A latitudinal phylogeographic diversity gradient in birds. *PLoS Biol.* **15**, e2001073. (doi:10.1371/journal.pbio.2001073)
10. Beheregaray LB, Caccone A. 2007 Cryptic biodiversity in a changing world. *J. Biol.* **6**, 9. (doi:10.1186/jbiol60)
11. Rannala B, Yang Z. 2017 Efficient Bayesian species tree inference under the multispecies coalescent. *Syst. Biol.* **66**, 823–842. (doi:10.1093/sysbio/syw119)
12. Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C. 2012 Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol. Evol.* **27**, 480–488. (doi:10.1016/j.tree.2012.04.012)
13. Huff DD, Miller LM, Chizinski CJ, Vondracek B. 2011 Mixed-source reintroductions lead to outbreeding depression in second-generation descendents of a native North American fish. *Mol. Ecol.* **20**, 4246–4258. (doi:10.1111/j.1365-294X.2011.05271.x)
14. Bininda-Emonds ORP. 2005 Supertree construction in the genomic age. *Methods Enzymol.* **395**, 745–757. (doi:10.1016/S0076-6879(05)95038-6)
15. Dupuis JR, Roe AD, Sperling FA. 2012 Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Mol. Ecol.* **21**, 4422–4436. (doi:10.1111/j.1365-294X.2012.05642.x)
16. Faircloth BC, Sorenson L, Santini F, Alfaro ME. 2013 A phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of ultraconserved elements (UCEs). *PLoS ONE* **8**, e65923. (doi:10.1371/journal.pone.0065923)
17. Buerki SB, Baker WJ. 2016 Collections-based research in the genomic era. *Biol. J. Linn. Soc.* **117**, 5–10. (doi:10.1111/bij.12721)
18. Leache AD, Banbury BL, Nieto-Montes de Oca A, Stamatakis A. 2015 Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Syst. Biol.* **64**, 1032. (doi:10.5061/dryad.t9r3g)
19. Recknagel H, Jacobs A, Herzyk P, Elmer KR. 2015 Double-digest RAD sequencing using Ion proton semiconductor platform (ddRADseq-Ion) with nonmodel organisms. *Mol. Ecol. Resour.* **15**, 1316–1329. (doi:10.1111/1755-0998.12406)
20. DaCosta JM, Sorenson MD. 2016 DdRAD-seq phylogenetics based on nucleotide, indel, and presence-absence polymorphisms: analyses of two avian genera with contrasting histories. *Mol. Phylogenet. Evol.* **94**, 122–135. (doi:10.1016/j.ympev.2015.07.026)
21. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008 Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **3**, e3376. (doi:10.1371/journal.pone.0003376.g001)
22. Cruaud A *et al.* 2014 Empirical assessment of RAD sequencing for interspecific phylogeny. *Mol. Biol. Evol.* **31**, 1272–1274. (doi:10.1093/molbev/msu063)
23. Rubin BER, Ree RH, Moreau CS. 2012 Inferring phylogenies from RAD sequence data. *PLoS ONE* **7**, e33394. (doi:10.1371/journal.pone.0033394)
24. Gonen S, Bishop SC, Houston RD. 2015 Exploring the utility of cross-laboratory RAD-sequencing datasets for phylogenetic analysis. *BMC Res. Notes* **8**, 299. (doi:10.1186/s13104-015-1261-2)
25. Takahashi T, Moreno E. 2015 A RAD-based phylogenetics for *Orestias* fishes from Lake Titicaca. *Mol. Phylogenet. Evol.* **93**, 307–317. (doi:10.1016/j.ympev.2015.08.012)
26. Henning F, Meyer A. 2014 The evolutionary genomics of cichlid fishes: explosive speciation and adaptation in the postgenomic era. *Annu. Rev. Genomics Hum. Genet.* **15**, 417–441. (doi:10.1146/annurev-genom-090413-025412)
27. Unmack PJ, Hammer MP, Adams M, Dowling TE. 2011 A phylogenetic analysis of pygmy perches (Teleostei: Percichthyidae) with an assessment of the major historical influences on aquatic biogeography in southern Australia. *Syst. Biol.* **60**, 797–812. (doi:10.1093/sysbio/syr042)
28. Crisp MD, Cook LG. 2007 A congruent molecular signature of vicariance across multiple plant lineages. *Mol. Phylogenet. Evol.* **43**, 1106–1117. (doi:10.1016/j.ympev.2007.02.030)
29. Rix MG, Edwards DL, Byrne M, Harvey MS, Joseph L, Roberts JD. 2015 Biogeography and speciation of terrestrial fauna in the south-western Australian biodiversity hotspot. *Biol. Rev.* **90**, 762–793. (doi:10.1111/brv.12132)
30. Byrne M *et al.* 2011 Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. *J. Biogeogr.* **38**, 1635–1656. (doi:10.1111/j.1365-2699.2011.02535.x)
31. Benbow MC. 1990 Tertiary coastal dunes of the Eucla Basin, Australia. *Geomorphology* **3**, 9–29. (doi:10.1016/0169-555X(90)90029-P)
32. Munasinghe DHN, Burridge CP, Austin CM. 2004 Molecular phylogeny and zoogeography of the freshwater crayfish genus *Cherax* Erichson (Decapoda: Parastacidae) in Australia. *Biol. J. Linn. Soc.* **81**, 553–563. (doi:10.1111/j.1095-8312.2003.00299.x)
33. Saddler S, Koehn JD, Hammer MP. 2013 Let's not forget the small fishes: conservation of two threatened species of pygmy perch in south-eastern Australia. *Mar. Freshw. Res.* **64**, 874. (doi:10.1071/mf12260)
34. Morgan DL, Beatty SJ, Adams M. 2013 *Nannoperca pygmaea*, a new species of pygmy perch (Teleostei: Percichthyidae) from Western Australia. *Zootaxa* **3637**, 401. (doi:10.11646/zootaxa.3637.4.1)
35. Prosdocimi F, de Carvalho DC, de Almeida RN, Beheregaray LB. 2012 The complete mitochondrial genome of two recently derived species of the fish genus *Nannoperca* (Perciformes, Percichthyidae). *Mol. Biol. Rep.* **39**, 2767–2772. (doi:10.1007/s11033-011-1034-5)
36. Unmack PJ, Hammer MP, Adams M, Johnson JB, Dowling TE. 2013 The role of continental shelf width in determining freshwater phylogeographic patterns in south-eastern Australian pygmy perches (Teleostei: Percichthyidae). *Mol. Ecol.* **22**, 1683–1699. (doi:10.1111/mec.12204)
37. Attard C, Möller LM, Sasaki M, Hammer MP, Bice CM, Brauer CJ, Carvalho DC, Harris JO, Beheregaray LB. 2016 A novel holistic framework for genetic-based captive-breeding and reintroduction programs. *Conserv. Biol.* **30**, 1060–1069. (doi:10.1111/cobi.12699)
38. Hammer MP, Bice CM, Hall A, Frears A, Watt A, Whiterod NS, Beheregaray LB, Harris JO, Zampatti BP. 2013 Freshwater fish conservation in the face of critical water shortages in the southern Murray-Darling basin, Australia. *Mar. Freshw. Res.* **64**, 807. (doi:10.1071/mf12258)
39. Brauer CJ, Hammer MP, Beheregaray LB. 2016 Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. *Mol. Ecol.* **25**, 5093–5113. (doi:10.1111/mec.13830)
40. Cole TL, Hammer MP, Unmack PJ, Teske PR, Brauer CJ, Adams M, Beheregaray LB. 2016 Range-wide fragmentation in a threatened fish associated with post-European settlement modification in the Murray–Darling basin, Australia. *Conserv. Genet.* **17**, 1377–1391. (doi:10.1007/s10592-016-0868-8)
41. Wedderburn SD, Hammer MP, Bice CM. 2012 Shifts in small-bodied fish assemblages resulting from drought-induced water level recession in terminating lakes of the Murray-Darling Basin, Australia. *Hydrobiologia* **691**, 35–46. (doi:10.1007/s10750-011-0993-9)
42. Brauer CJ, Unmack PJ, Hammer MP, Adams M, Beheregaray LB. 2013 Catchment-scale conservation units identified for the threatened Yarra pygmy perch (*Nannoperca obscura*) in highly modified river systems. *PLoS ONE* **8**, e82953. (doi:10.1371/journal.pone.0082953.g001)
43. Sunnucks P, Hales DF. 1996 Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus Sitobion (Hemiptera: Aphididae). *Mol. Biol. Evol.* **13**, 510–524. (doi:10.1093/oxfordjournals.molbev.a025612)
44. Peterson B, Weber J, Kay E, Fisher H, Hoekstra H. 2012 Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model

and non-model species. *PLoS ONE* **7**, e37135. (doi:10.1371/journal.pone.0037135)

45. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013 Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140. (doi:10.1111/mec.12354)
46. Eaton DA. 2014 PyRAD: assembly of *de novo* RADseq loci for phylogenetic analyses. *Bioinformatics* **30**, 1844–1849. (doi:10.1093/bioinformatics/btu121)
47. Whelan NV. 2011 Species tree inference in the age of genomics. *Trends Evol. Biol.* **3**, 5. (doi:10.4081/eb.2011.e5)
48. Swofford DL. 2002 *PAUP*. Phylogenetic analysis using parsimony (*and other methods)*. Sunderland, MA: Sinauer Associates.
49. Stamatakis A. 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. (doi:10.1093/bioinformatics/btu033)
50. Pante E *et al.* 2015 Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Mol. Ecol.* **24**, 525–544. (doi:10.1111/mec.13048)
51. Miller MA, Pfeiffer W, Schwartz T. 2010 Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proc. Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA, 14 November 2010*, pp. 1–8.
52. Jia F, Lo N, Ho SY. 2014 The impact of modelling rate heterogeneity among sites on phylogenetic estimates of intraspecific evolutionary rates and timescales. *PLoS ONE* **9**, e95722. (doi:10.1371/journal.pone.0095722)
53. Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A. 2014 Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol. Biol.* **14**, 82. (doi:10.1186/1472-2148-14-82)
54. Lartillot N, Lepage T, Blanquart S. 2009 PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* **25**, 2286–2288. (doi:10.1093/bioinformatics/btp368)
55. Kumar S, Stecher G, Tamura K. 2016 MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874. (doi:10.1093/molbev/msw054)
56. Yang Z. 2015 The BPP program for species tree estimation and species delimitation. *Curr. Zool.* **61**, 854–865. (doi:10.1093/czoolo/61.5.854)
57. Leache AD, Fujita MK. 2010 Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proc. R. Soc. B* **277**, 3071–3077. (doi:10.1098/rspb.2010.0662)
58. Yang Z, Rannala B. 2010 Bayesian species delimitation using multilocus sequence data. *Proc. Natl Acad. Sci. USA* **107**, 9264–9269. (doi:10.1073/pnas.0913022107)
59. Drummond AJ, Rambaut A. 2017 *Summarizing trees*. BEASTdoc [Internet]. (July). See http://beast.community/summarizing_trees.
60. Lischer HE, Excoffier L, Heckel G. 2014 Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of *Microtus* voles. *Mol. Biol. Evol.* **31**, 817–831. (doi:10.1093/molbev/mst271)
61. Fontaneto D, Flot JF, Tang CQ. 2015 Guidelines for DNA taxonomy, with a focus on the meiofauna. *Mar. Biodivers.* **45**, 433–451. (doi:10.1007/s12526-015-0319-7)
62. Paradis E, Claude J, Strimmer K. 2004 APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290. (doi:10.1093/bioinformatics/btg412)
63. Aberer AJ, Kobert K, Stamatakis A. 2014 ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. *Mol. Biol. Evol.* **31**, 2553–2556. (doi:10.1093/molbev/msu236)
64. Sanderson M. 2003 R8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**, 301–302. (doi:10.1093/bioinformatics/19.2.301)
65. Chen W, Lavoué S, Beheregaray LB, Mayden RL. 2014 Historical biogeography of a new antitropical clade of temperate freshwater fishes. *J. Biogeogr.* **41**, 1806–1818. (doi:10.1111/jbi.12333)
66. Parham JF *et al.* 2012 Best practices for justifying fossil calibrations. *Syst. Biol.* **61**, 346–359. (doi:10.1093/sysbio/syr107)
67. Bromham L, Duchêne S, Hua X, Ritchie AM, Duchêne DA, Ho SYW. 2018 Bayesian molecular dating: opening up the black box. *Biol. Rev.* **93**, 1165–1191. (doi:10.1111/brv.12390)
68. Bhatnagar I, Kim SK. 2010 Immense essence of excellence: marine microbial bioactive compounds. *Mar. Drugs* **8**, 2673–2701. (doi:10.3390/md8102673)
69. Matzke NJ. 2013 Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. PhD thesis, University of California, Berkeley.
70. Revell LJ. 2012 Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223. (doi:10.1111/j.2041-210X.2011.00169.x)
71. Hollingsworth Jr PR, Near TJ. 2009 Temporal patterns of diversification and microendemism in eastern highland endemic barcheek darters (Percidae: Etheostomatinae). *Evolution* **63**, 228–243. (doi:10.1111/j.1558-5646.2008.00531.x)
72. Santini F, Nguyen MTT, Sorenson L, Waltzek TB, Lynch Alfaro JW, Eastman JM, Alfaro ME. 2013 Do habitat shifts drive diversification in teleost fishes? An example from the pufferfishes (Tetraodontidae). *J. Evol. Biol.* **26**, 1003–1018. (doi:10.1111/jeb.12112)
73. Morgan DL, Beatty S. 2003 *Fish fauna of Margaret River Western Australia*. Perth, Australia: Margaret River Regional Environment Centre, Murdoch University.
74. Pires AC, Marinoni L. 2010 DNA barcoding and traditional taxonomy unified through integrative taxonomy: a view that challenges the debate questioning both methodologies. *Biota Neotrop.* **10**, 339–346. (doi:10.1590/S1676-06032010000200035)
75. De Queiroz K. 2007 Species concepts and species delimitation. *Syst. Biol.* **56**, 879–886. (doi:10.1080/10635150701701083)
76. Dayrat B. 2005 Towards integrative taxonomy. *Biol. J. Linn. Soc.* **85**, 407–415. (doi:10.1111/j.1095-8312.2005.00503.x)
77. Peterson AT, Navarro-Siguenza AG. 1999 Alternate species concepts as bases for determining priority conservation areas. *Conserv. Biol.* **13**, 427–431. (doi:10.1046/j.1523-1739.1999.013002427.x)
78. Mace GM. 2004 The role of taxonomy in species conservation. *Phil. Trans. R. Soc. Lond. B* **359**, 711–719. (doi:10.1098/rstb.2003.1454)
79. Pavlova A *et al.* 2017 Severe consequences of habitat fragmentation on genetic diversity of an endangered Australian freshwater fish: a call for assisted gene flow. *Evol. Appl.* **10**, 531–550. (doi:10.1111/eva.12484)
80. Sukumaran J, Knowles LL. 2017 Multispecies coalescent delimits structure, not species. *Proc. Natl Acad. Sci. USA* **114**, 1607–1612. (doi:10.1073/pnas.1607921114)
81. Pie MR, Bornschein MR, Ribeiro LF, Faircloth BC, McCormack JE. 2017 Phylogenomic species delimitation in microendemic frogs of the Brazilian Atlantic Forest. *BioRxiv*. (doi:10.1101/143735)
82. Waters JM, Craw D, Youngson JH, Wallis GP. 2001 Genes meet geology: fish phylogeographic pattern reflects ancient, rather than modern, drainage connections. *Evolution* **55**, 1844–1851. (doi:10.1111/j.0014-3820.2001.tb00833.x)
83. Ladiges PY, Bayly MJ, Nelson G. 2012 Searching for ancestral areas and artifactual centers of origin in biogeography: with comment on east–west patterns across southern Australia. *Syst. Biol.* **61**, 703–708. (doi:10.1093/sysbio/sys005)
84. Unmack PJ. 2001 Biogeography of Australian freshwater fishes. *J. Biogeogr.* **28**, 1053–1089. (doi:10.1046/j.1365-2699.2001.00615.x)
85. Guows G, Stewart B, Daniels S. 2006 Phylogeographic structure of a freshwater crayfish (Decapoda: Parastacidae: *Cherax preissii*) in south-western Australia. *Mar. Freshw. Res.* **57**, 837–848. (doi:10.1071/MF05248)
86. Edwards DL, Roberts JD, Keogh JS. 2008 Climatic fluctuations shape the phylogeography of a mesic direct-developing frog from the south-western Australian biodiversity hotspot. *J. Biogeogr.* **35**, 1803–1815. (doi:10.1111/j.1365-2699.2008.01927.x)
87. Ruegg KC, Hijmans RJ, Moritz C. 2006 Climate change and the origin of migratory pathways in the Swainson's thrush, *Catharus ustulatus*. *J. Biogeogr.* **33**, 1172–1182. (doi:10.1111/j.1365-2699.2006.01517.x)
88. Richards CL, Carstens BC, Knowles LL. 2007 Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *J. Biogeogr.* **34**, 1833–1845. (doi:10.1111/j.1365-2699.2007.01814.x)
89. Moritz C. 1994 Defining 'evolutionarily significant units' for conservation. *Trends Ecol. Evol.* **9**, 373–375. (doi:10.1016/0169-5347(94)90057-4)
90. Stein JL, Hutchinson MF, Stein JA. 2014 A new stream and nested catchment framework for Australia. *Hydrol. Earth Syst. Sci.* **18**, 1917–1933. (doi:10.5194/hess-18-1917-2014)
91. Kennard MJ, Pusey BJ, Olden JD, Mackay SJ, Stein JL, Marsh N. 2010 Classification of natural flow regimes in Australia to support environmental flow management. *Freshw. Biol.* **55**, 171–193. (doi:10.1111/j.1365-2427.2009.02307.x)
92. Attard CRM, Brauer CJ, Van Zoelen JD, Sasaki M, Hammer MP, Morrison L, Harris JO, Möller LM, Beheregaray LB. 2016 Multi-generational evaluation of genetic diversity and parentage in captive southern pygmy perch (*Nannoperca australis*). *Conserv. Genet.* **17**, 1469–1473. (doi:10.1007/s10592-016-0873-y)
93. Buckley S, Mata F, Attard C, Brauer C, Sandoval-Castillo J, Lodge R, Unmack P, Beheregaray L. 2018 Data from: Phylogenomic history of enigmatic pygmy perches: implications for biogeography, taxonomy and conservation. Dryad Digital Repository. (doi:10.5061/dryad.3d6gt)

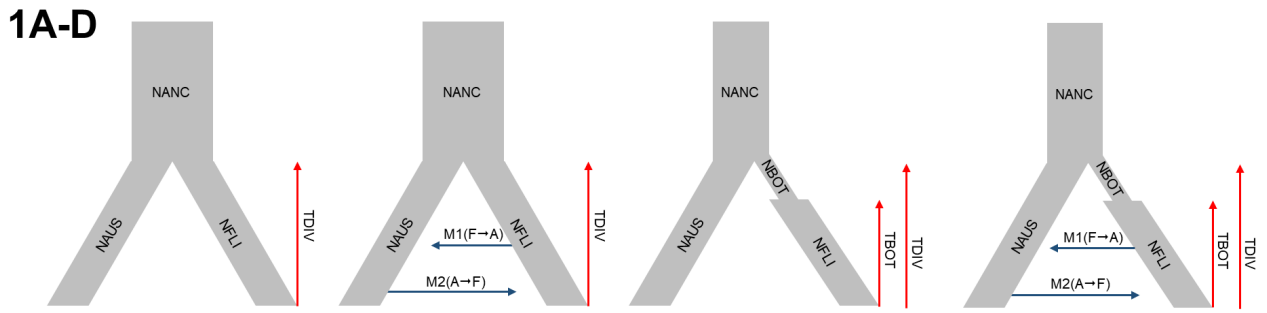
Appendix 2: Biogeographic hypothesis testing framework for coalescent modelling in FastSimCoal2.

Model		Biogeographic mechanism	Reference(s)	Fixed/constrained parameters	No. parameters
1	A	Range expansion and vicariant separation due to early cycles of marine inundation in E Gippsland	Chapple *et al.* 2005 Norgate *et al.* 2009 Coleman *et al.* 2010	N/A	4
	B	Range expansion and vicariant separation due to early cycles of marine inundation in E Gippsland with post-isolation migration	Chapple *et al.* 2005 Norgate *et al.* 2009 Coleman *et al.* 2010	M_1-M_2 = 10^{-10} – 10^{-5}	6
	C	Dispersal into Eastern Victoria after withdrawal of Eastern Gippsland inundation	N/A	Bottleneck	6
	D	Post-inundation dispersal with post-isolation migration between species	N/A	Bottleneck M_1-M_2 = 10^{-10} – 10^{-5}	8
2	A	Vicariant separation of NauWP due to tectonics, marine inundation or volcanism	Dickinson *et al.* 2002 Gallagher *et al.* 2001 Holdgate *et al.* 2003 Lesti *et al.* 2008	T_{WP} = 4.5 – 2, 6 – 2, or <5.3 Mya, respectively.	7
	B	Vicariant separation of NauWP with post-isolation migration	N/A	M_1-M_6 = 10^{-10} – 10^{-5}	13
	C	NauWP is 50% hybrid lineage of *N. aus* and *N.* 'fli'	N/A	$Coal_{NauWP→Fli}$ = 0.5; $Coal_{NauWP→Aus}$ = 0.5	7
	D	NauWP is hybrid lineage of *N. aus* and *N.* 'fli' with variable contributions of gene flow between species	N/A	$Coal_{NauWP→Fli}$ = 1^{-10} - 0.5 $Coal_{NauWP→Aus}$ = 1^{-10} - 0.5	9
	E	NauWP is hybrid lineage of *N. aus* and *N.* 'fli' with variable contributions of gene flow between species with additional post-isolation migration	N/A	$Coal_{NauWP→Fli}$ = 1^{-10} - 0.5 $Coal_{NauWP→Aus}$ = 1^{-10} - 0.5 M_1-M_6 = 10^{-10} – 10^{-5}	15
3	A	Dispersal through the lower reaches of the MDB (via Lake Bungunnia)	McLaren *et al.* 2012	T_{COAST} = 700 Kya, 1.2 Mya, 3 Mya.	25

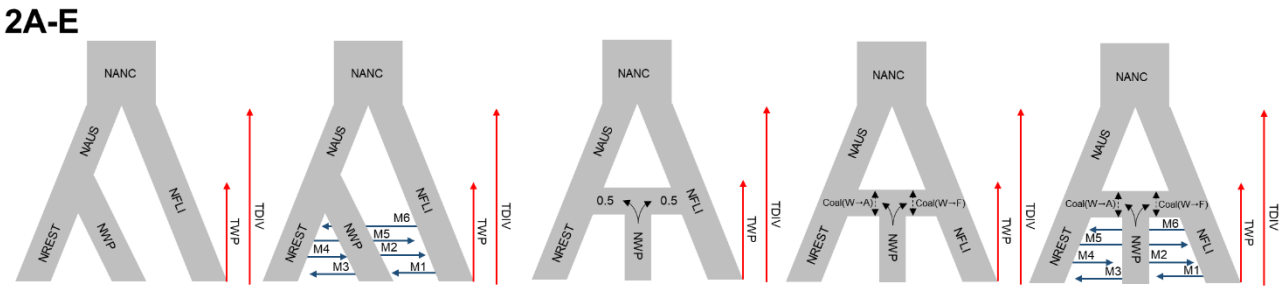
	B	Dispersal through the middle reaches of the MDB (across the GDR)	Unmack *et al.* 2001		25
	C	Dispersal through upper reaches of the MDB (across the GDR)	Unmack *et al.* 2001		25
	D	Two separate migration events into the MDB (via Lower and Mid)	N/A		24
	E	Vicariance following the patterns of the phylogenetic tree (i.e. MID most basal)	N/A		25
4	A	Sea-level rise sequentially disconnecting river systems along coast	Williams *et al.* 2018	T_{BASE} < 25,000 years ago Ts = 100 – 5 ,000 years apart	19
	B	Sea-level rise sequentially disconnecting river systems along coast, removing migration	Williams *et al.* 2018	T_{BASE} < 25 Kya Ts = 100 – 5,000 years apart $M_1 – M_{12} = 10^{-10} – 10^{-5}$	29
	C	Reduction of Lake Corangamite isolating eastern populations along the coast	White, 2000	T_{BAR} = 1,000 years ago $M_1 – M_{12} = 10^{-10} – 10^{-5}$	18
	D	Reduction of Lake Corangamite isolating eastern populations along the coast, removing migration	White, 2000	T_{BAR} = 1,000 years ago	26
	E	Both sea-level rise and reduction in Lake Corangamite contributing to disconnection of river systems along coast, removing migration	White, 2000	T_{BAR} = 1,000 years ago T_{BASE} < 25,000 years ago Ts = 100 – 5 ,000 years apart	31
5	A	Divergence between MID and the rest of the MDB populations prior to European settlement	Cole *et al.* 2016	T_{DIV} > 200 years ago	4
	B	Divergence between MID and the rest of the MDB populations after European settlement	Brauer *et al.* 2016	T_{DIV} < 200 years ago	4
	C	Divergence between MID and the rest of the MDB populations prior to European	Cole *et al.* 2016	T_{DIV} > 200 years ago	6

		settlement with post-isolation gene flow			
	D	Divergence between Upper and the Lower MDB populations prior to European settlement	Cole *et al.* 2016	T_{DIV} > 200 years ago	4
	E	Divergence between Upper and the Lower MDB populations after European settlement	Brauer *et al.* 2016	T_{DIV} < 200 years ago	4
6	A	Divergence of populations of *N.* 'flindersi' across the Bass Strait due to vicariance from early cycles of marine inundation	Coleman *et al.* 2010	T_{TAS} = 1.5 – 2 Mya T_{VIC} = 2 – 2.5 Mya	6
	B	Divergence of populations of *N.* 'flindersi' across the Bass Strait due to variance from early cycles of marine inundation with gene flow post-isolation	Coleman *et al.* 2010	$M_1 – M_6 = 10^{-10} – 10^{-5}$ T_{TAS} = 1.5 – 2 Mya T_{VIC} = 2 – 2.5 Mya	10
	C	Divergence of populations of *N.* 'flindersi' across the Bass Strait due to isolated dispersal events with associated bottleneck	N/A	T_{TAS} = 1.5 – 2 Mya T_{VIC} = 2 – 2.5 Mya $R_{SRLO} = -10^{-4} – 0$ $R_{ART} = -10^{-4} – 0$ $R_{FI} = -10^{-4} – 0$	7
	D	Divergence of populations of *N.* 'flindersi' across the Bass Strait due to isolated dispersal events with associated bottleneck and gene flow post-isolation	N/A	$M_1 – M_6 = 10^{-10} – 10^{-5}$ T_{TAS} = 1.5 – 2 Mya T_{VIC} = 2 – 2.5 Mya $R_{SRLO} = -10^{-4} – 0$ $R_{ART} = -10^{-4} – 0$ $R_{FI} = -10^{-4} – 0$	11

Appendix 3: Coalescent models testing alternate hypotheses for the divergence of *N. australis* and *N.* 'flindersi'. Black text indicates population size parameters, red arrows indicate timing parameters, and blue arrows indicate migration parameters. NAUS and NFLI consist of all *N. australis* and *N.* 'flindersi' populations combined, respectively.

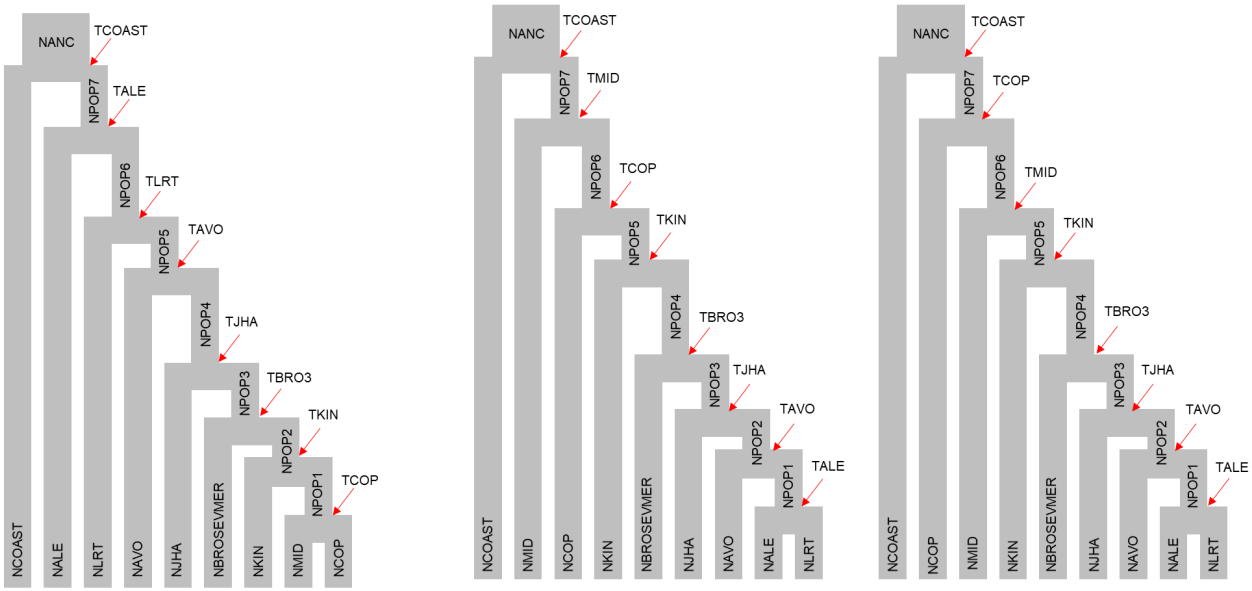


Appendix 4: Coalescent models testing alternate hypotheses for the divergence of the Wilson's Promontory population of *N. australis*. Black text indicates population size parameters, red arrows indicate timing parameters, and blue arrows indicate migration parameters. NAUS consists of all *N. australis* populations excluding NauWP, and NFLI consists of all *N.* 'flindersi' populations combined.

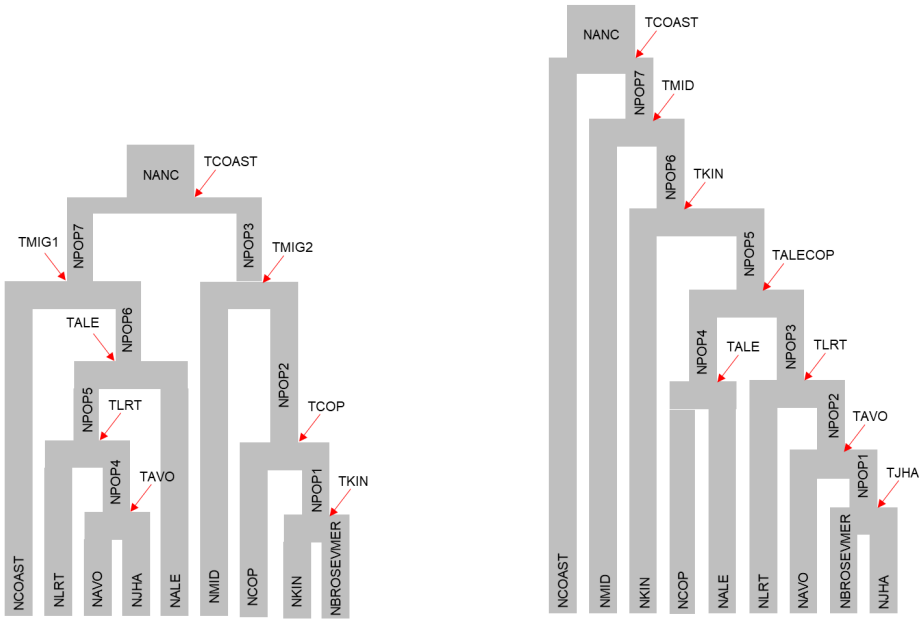


Appendix 5: Coalescent models testing alternate hypotheses for the divergence of the MDB and coastal Victoria clades. Black text indicates population size parameters, red arrows indicate timing parameters, and blue arrows indicate migration parameters. NCOAST consists of all coastal Victoria populations combined.

3A-C

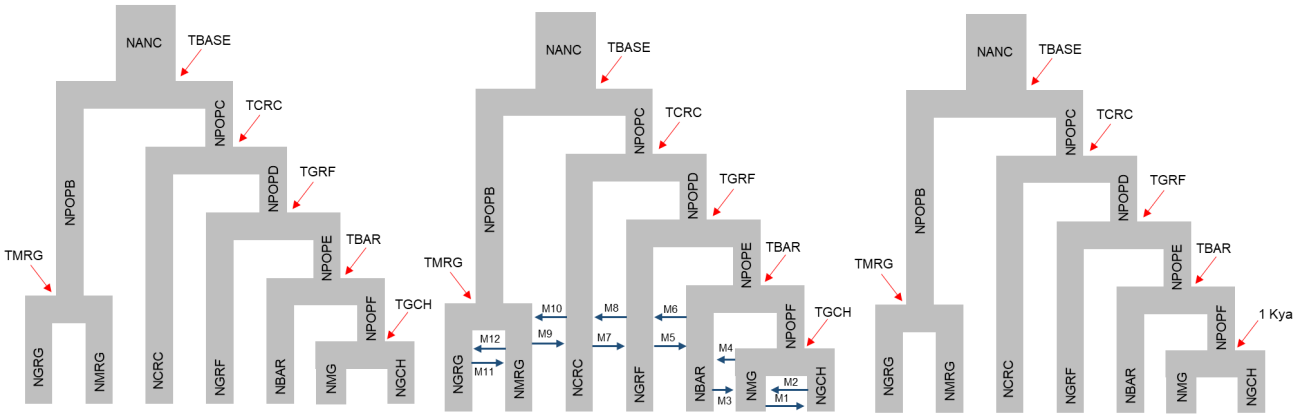


3D-E

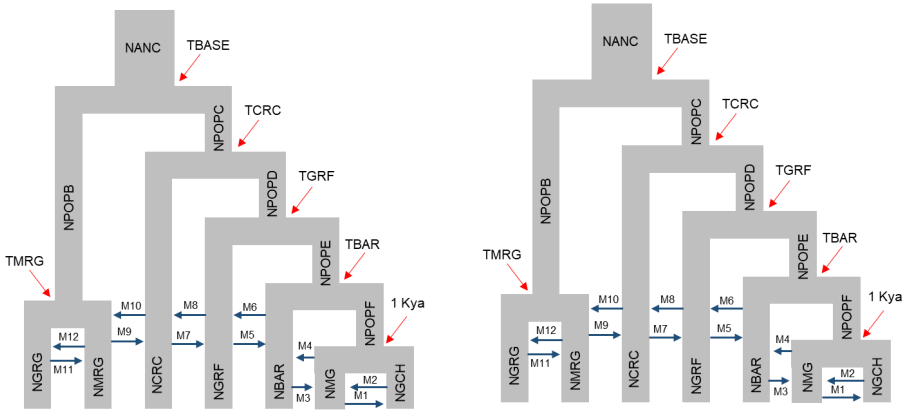


Appendix 6: Coalescent models testing alternate hypotheses for the divergence of populations within the coastal Victoria clade. Black text indicates population size parameters, red arrows indicate timing parameters, and blue arrows indicate migration parameters.

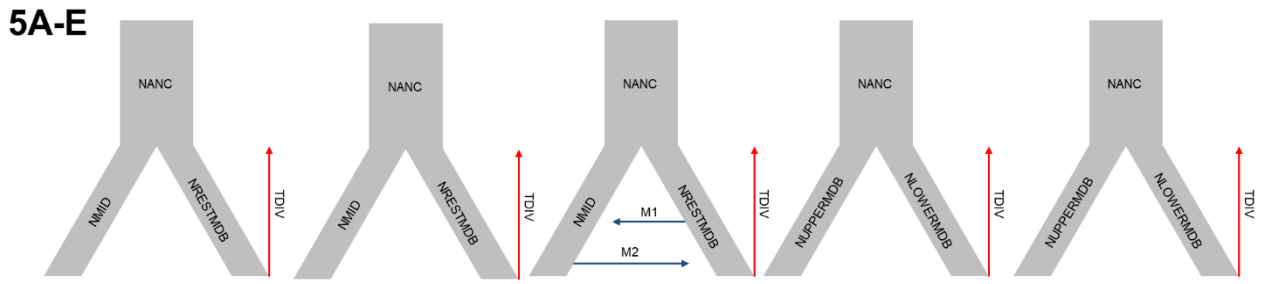
4A-C



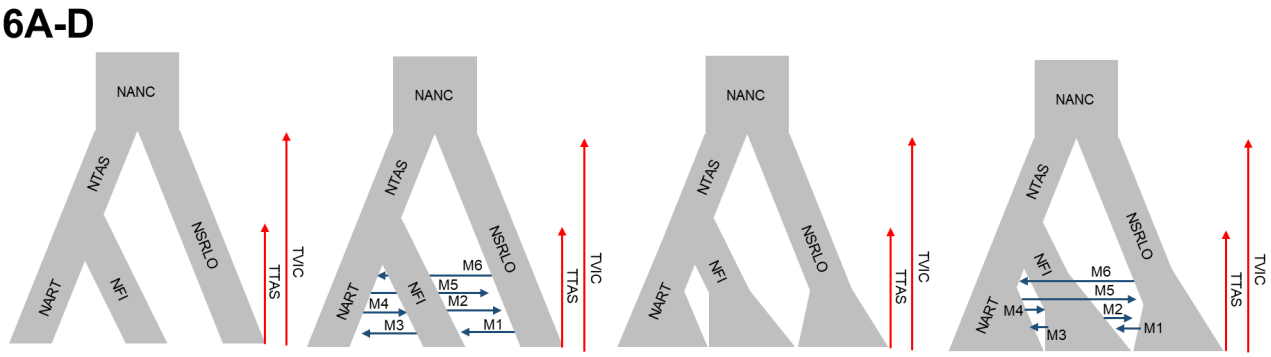
4D-E



Appendix 7: Coalescent models testing alternate hypotheses for divergence within the MDB clade. Black text indicates population size parameters, red arrows indicate timing parameters, and blue arrows indicate migration parameters. NMID consists of the NauGAP, NauALB and NauSPR populations combined, with NRESTMDB consisting of all other MDB populations. NLOWERMDB consists of the NauALE and NauANG populations combined, with NUPPERMDB consisting of all other MDB populations.



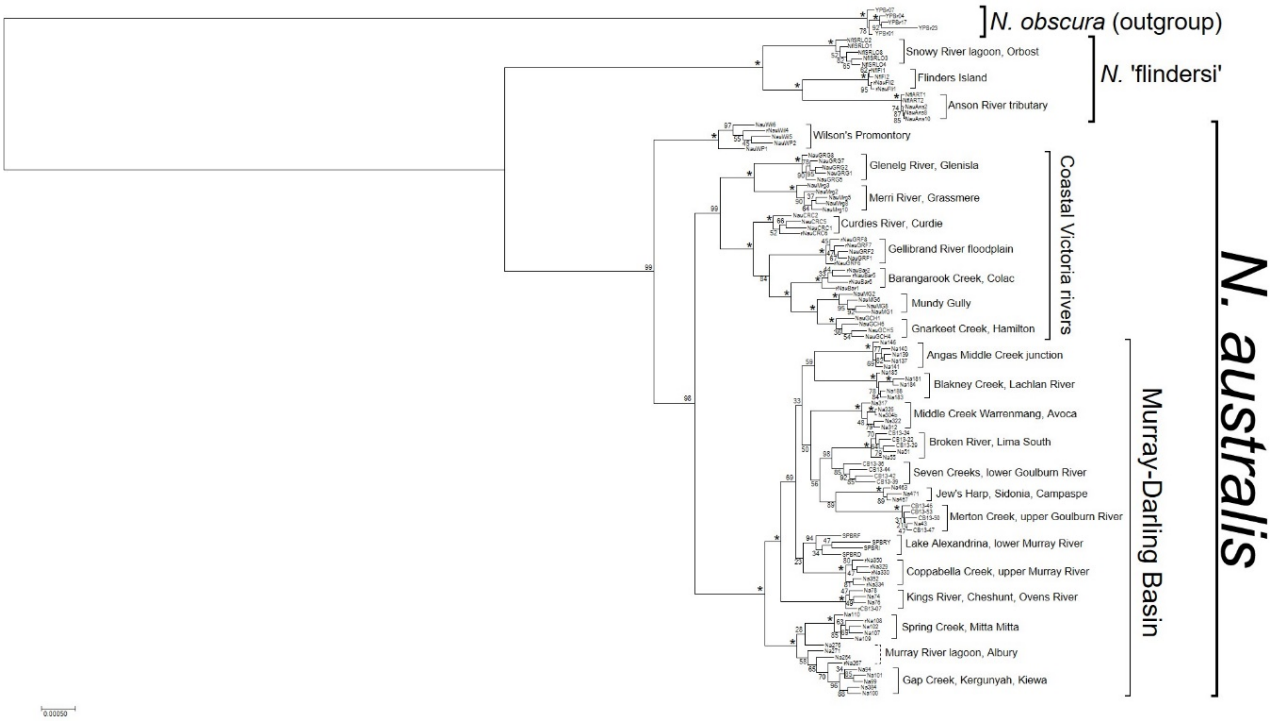
Appendix 8: Coalescent models testing alternate hypotheses for the divergence of *N. 'flindersi'* populations. Black text indicates population size parameters, red arrows indicate timing parameters, and blue arrows indicate migration parameters. Enlarged populations indicate exponential growth post-divergence.



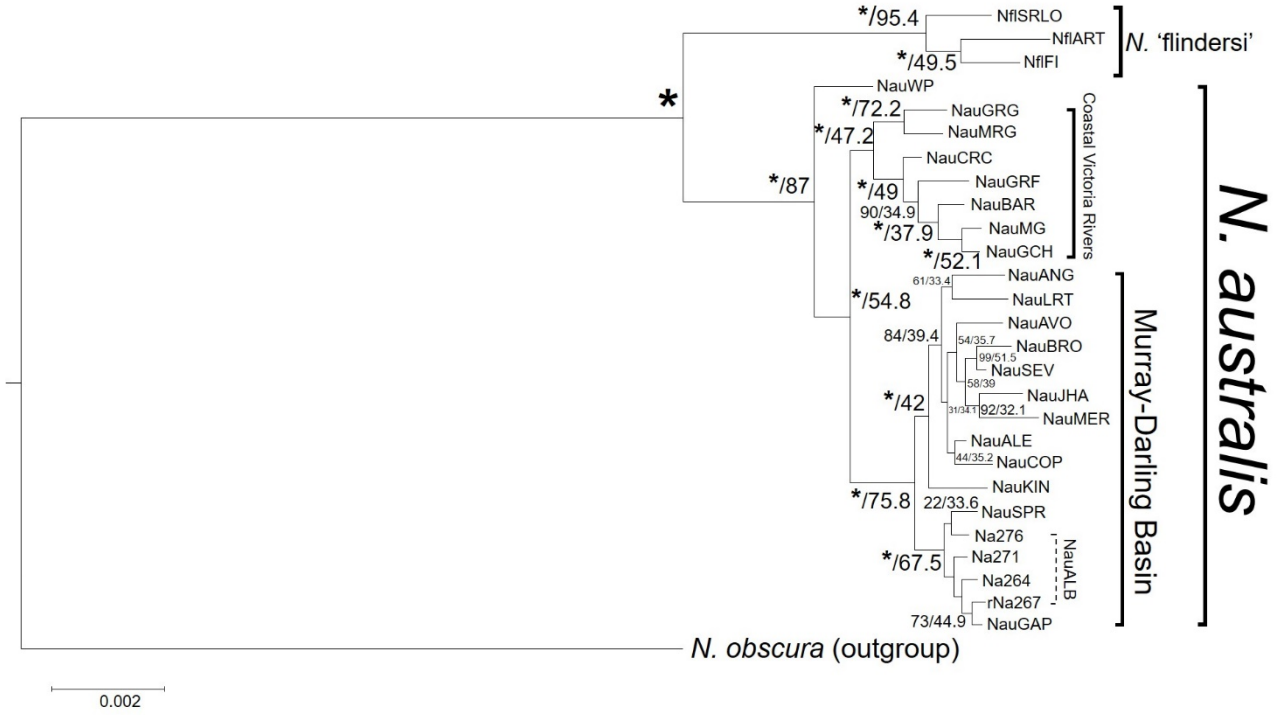
Appendix 9: Pearson's pairwise correlation for all 19 bioclimatic variables obtained from WorldClim v1.4. Selected uncorrelated variables (|R| < 0.8) are highlighted in bold, with correlation values between these variables highlighted in green.

Variable	**Bio1**	**Bio2**	**Bio3**	Bio4	Bio5	**Bio6**	Bio7	**Bio8**	**Bio9**	Bio10	Bio11	Bio12	Bio13	**Bio14**	**Bio15**	Bio16	Bio17	Bio18	**Bio19**
Bio1																			
Bio2	0.63																		
Bio3	-0.06	-0.16																	
Bio4	0.56	0.86	-0.63																
Bio5	0.90	0.88	-0.25	0.83															
Bio6	0.63	-0.12	0.40	-0.26	0.88														
Bio7	0.59	0.95	-0.45	0.97	0.88	-0.23													
Bio8	0.68	0.47	-0.27	0.55	0.62	0.21	0.52												
Bio9	0.31	0.18	0.22	0.02	0.29	0.41	0.09	-0.35											
Bio10	0.96	0.77	-0.27	0.78	0.98	0.39	0.79	0.70	0.25										
Bio11	0.89	0.29	0.27	0.13	0.63	0.90	0.18	0.50	0.38	0.73									
Bio12	-0.83	-0.68	-0.09	-0.50	-0.80	-0.46	-0.58	-0.50	-0.34	-0.81	-0.72								
Bio13	-0.81	-0.72	-0.02	-0.57	-0.82	-0.36	-0.65	-0.52	-0.30	-0.81	-0.66	0.98							
Bio14	-0.79	-0.50	-0.26	-0.28	-0.69	-0.64	-0.38	-0.35	-0.44	-0.70	-0.79	0.91	0.83						
Bio15	-0.81	-0.40	0.51	-0.56	-0.29	0.47	-0.52	-0.34	0.22	-0.26	0.21	0.10	0.27	-0.25					
Bio16	-0.81	-0.73	-0.01	-0.58	-0.82	-0.35	-0.65	-0.54	-0.28	-0.82	-0.66	0.98	1.00	0.83	0.27				
Bio17	-0.78	-0.52	-0.26	-0.70	-0.70	-0.62	-0.39	-0.35	-0.42	-0.70	-0.78	0.94	0.86	0.99	-0.21	0.86			
Bio18	-0.72	-0.46	-0.29	-0.22	-0.63	-0.64	-0.32	-0.21	-0.54	-0.63	-0.75	0.88	0.80	0.97	-0.24	0.80	0.97		
Bio19	-0.80	-0.73	0.04	-0.61	-0.82	-0.30	-0.68	-0.61	-0.20	-0.82	-0.62	0.96	0.99	0.78	0.32	0.99	0.81	0.73	

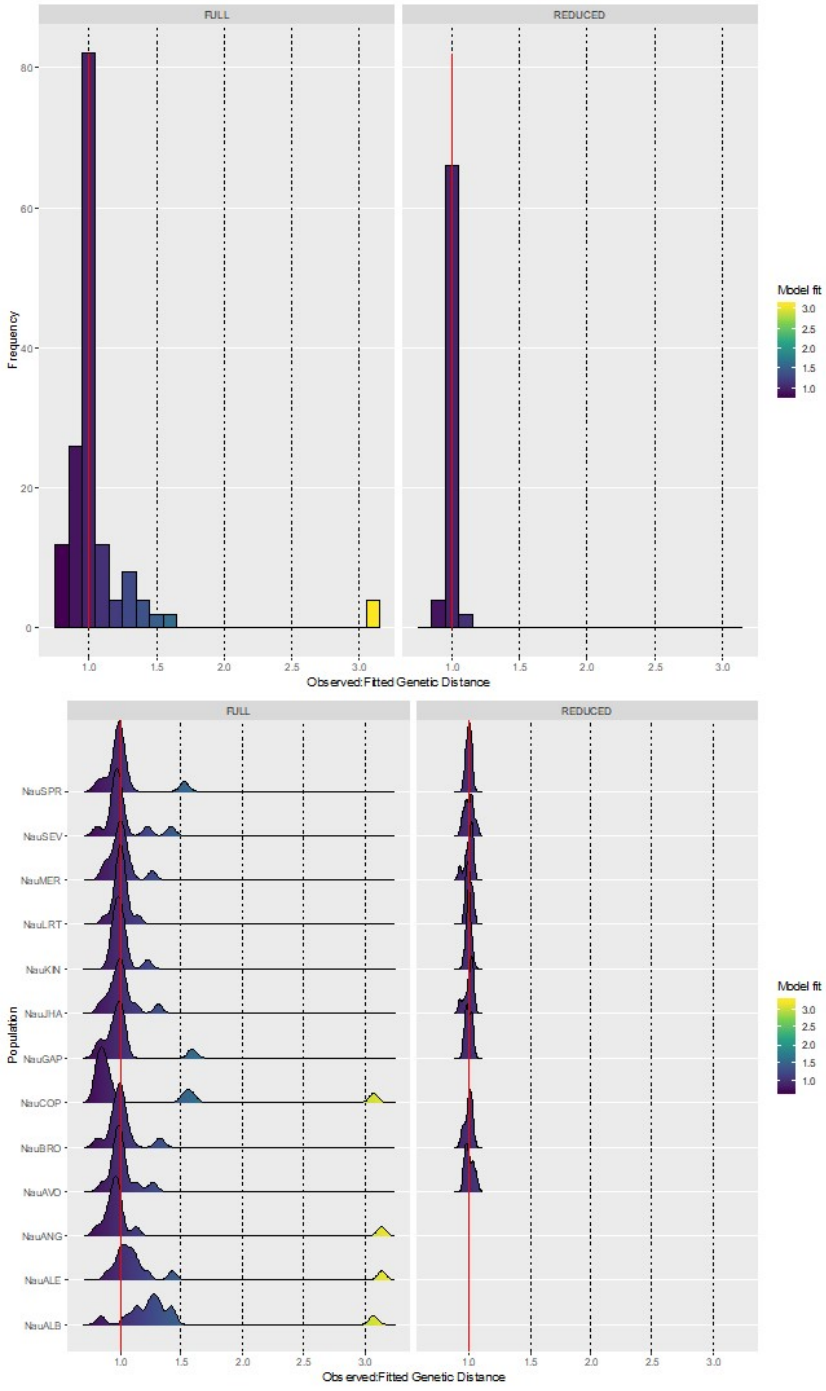
Appendix 10: The ML phylogeny of all 114 *N. australis* and *N.* 'flindersi' samples based on 7,780 concatenated RADseq loci. Node labels indicate bootstrap support from 1,000 RELL bootstraps, with asterisks denoting nodes with 100% support. Species, clades and populations are denoted by square brackets.



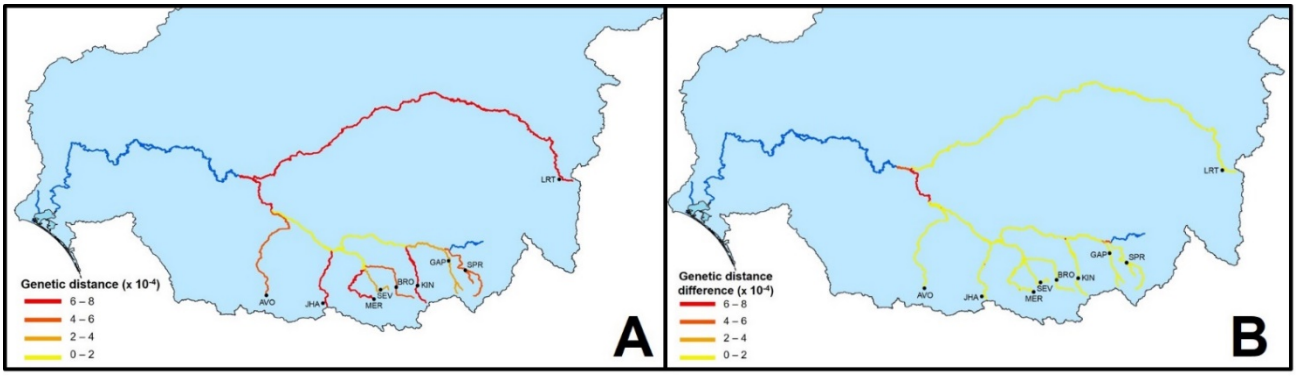
Appendix 11: The ML phylogeny of *N. australis* and *N.* 'flindersi' collapsed down to populations based on 7,780 partitioned ddRADseq loci. Node labels indicate bootstrap support from 1,000 RELL bootstraps (reported first), with asterisks denoting nodes with 100% support, as well as site concordance factors between individual ddRAD loci gene trees and the tree (reported second). Concordance factors were calculated using 100 random quartets per branch.



Appendix 12: Evaluation of *StreamTree* model fit across the entire set of population comparisons (top) and for comparisons involving each specific population (bottom). Density plots represent the probability distribution of mismatch (ratio) between observed (true) genetic distance and distance modelled by *StreamTree*, with the red line indicating perfect fit of the model (StreamTree value = observed genetic distance). The full dataset (left) is compared with the model with divergent outliers removed (right).



Appendix 13: Visual representation of the reduced *StreamTree* model (excluding 4 outlier populations) of genetic divergence across the MDB. **A:** Dendritic riverine network of the MDB, with streams colour-coded according to the reduced *StreamTree* model that determines the contribution (as a penalty) of each segment in driving genetic divergence (measured as mean uncorrected genetic distance per population (p) x 10^2) across the basin. Segments coloured in yellow confer little penalty (i.e. genetic divergence between populations at either end of the segment is low) whereas red segments confer higher genetic differentiation. **B:** Visual comparison of *StreamTree* models using the full (n = 13 populations) and reduced (n = 9) populations for stream segments that were considered under both models. Yellow segments demonstrate streams with similar associated penalties across both models whereas red segments showed more variable penalties between the two models.



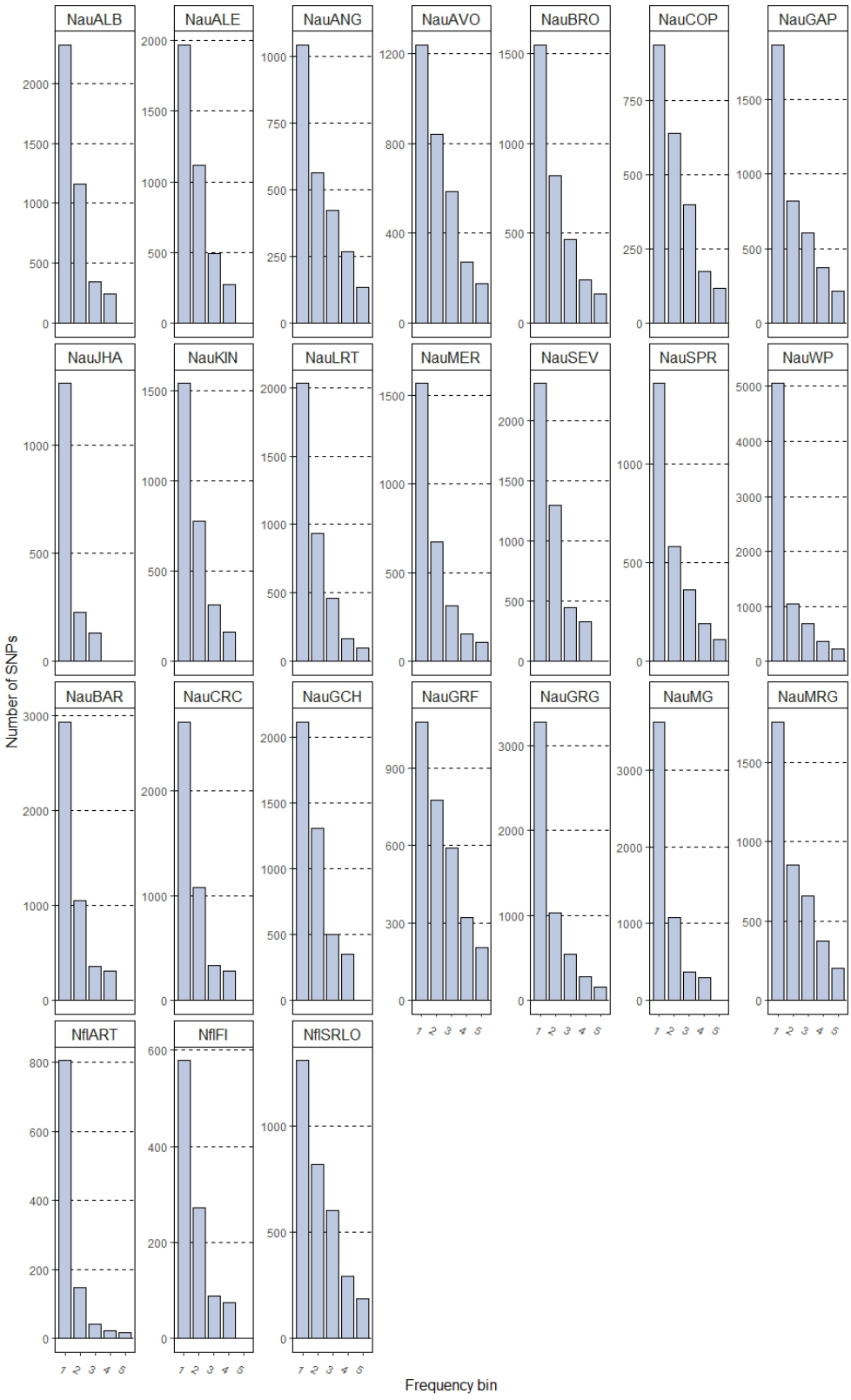
Appendix 14: Comparison of ancestral range estimation model likelihoods across six biogeographic models and two time stratification scenarios, and including a constraint on maximum range size. The most supported models are highlighted in green.

Constraint	Model	Log likelihood	Dispersal (*d*)	Extincton (*e*)	Founder event (*J*)	AIC	Relative AIC
MDB excluded prior to 5 Mya	DEC	-17.6	0.030	1.0×10^{-12}	0	39.19	0.0004
	DEC+J	-16.61	1.0×10^{-12}	1.0×10^{-12}	0.013	39.23	0.0004
	DIVALIKE	-10.08	0.015	1.0×10^{-12}	0	24.16	0.73
	DIVALIKE+J	-10.08	0.015	1.0×10^{-12}	1.0×10^{-5}	26.16	0.27
	BAYAREALIKE	-30.29	0.036	0.21	0	64.57	1.2×10^{-9}
	BAYAREALIKE+J	-20.08	1.0×10^{-12}	1.0×10^{-12}	0.029	46.16	1.2×10^{-5}
MDB excluded prior to 5 Mya + max range = 2 areas	DEC	-23.07	0.059	1.0×10^{-12}	0	50.14	0.0028
	DEC+J	-19.1	1.0×10^{-12}	1.0×10^{-12}	0.026	44.21	0.054
	DIVALIKE	-17.58	0.059	1.0×10^{-12}	0	39.15	0.67*
	DIVALIKE+J	-17.55	0.039	1.0×10^{-12}	0.008	41.1	0.25
	BAYAREALIKE	-32.1	0.093	0.13	0	68.2	3.30×10^{-7}
	BAYAREALIKE+J	-20.08	1.0×10^{-12}	1.0×10^{-12}	0.029	46.16	0.02
MDB excluded prior to 2 Mya	DEC	-30.35	0.050	0.062	0	64.69	0.042
	DEC+J	-30.35	0.050	0.062	1.0×10^{-5}	66.69	0.016
	DIVALIKE	-27.55	0.040	0.041	0	59.1	0.69*
	DIVALIKE+J	-27.55	0.040	0.041	1.0×10^{-5}	61.1	0.25
	BAYAREALIKE	-38.91	0.063	0.23	0	81.82	8.0×10^{-6}
	BAYAREALIKE+J	-36.64	0.037	0.080	0.023	79.28	2.9×10^{-5}

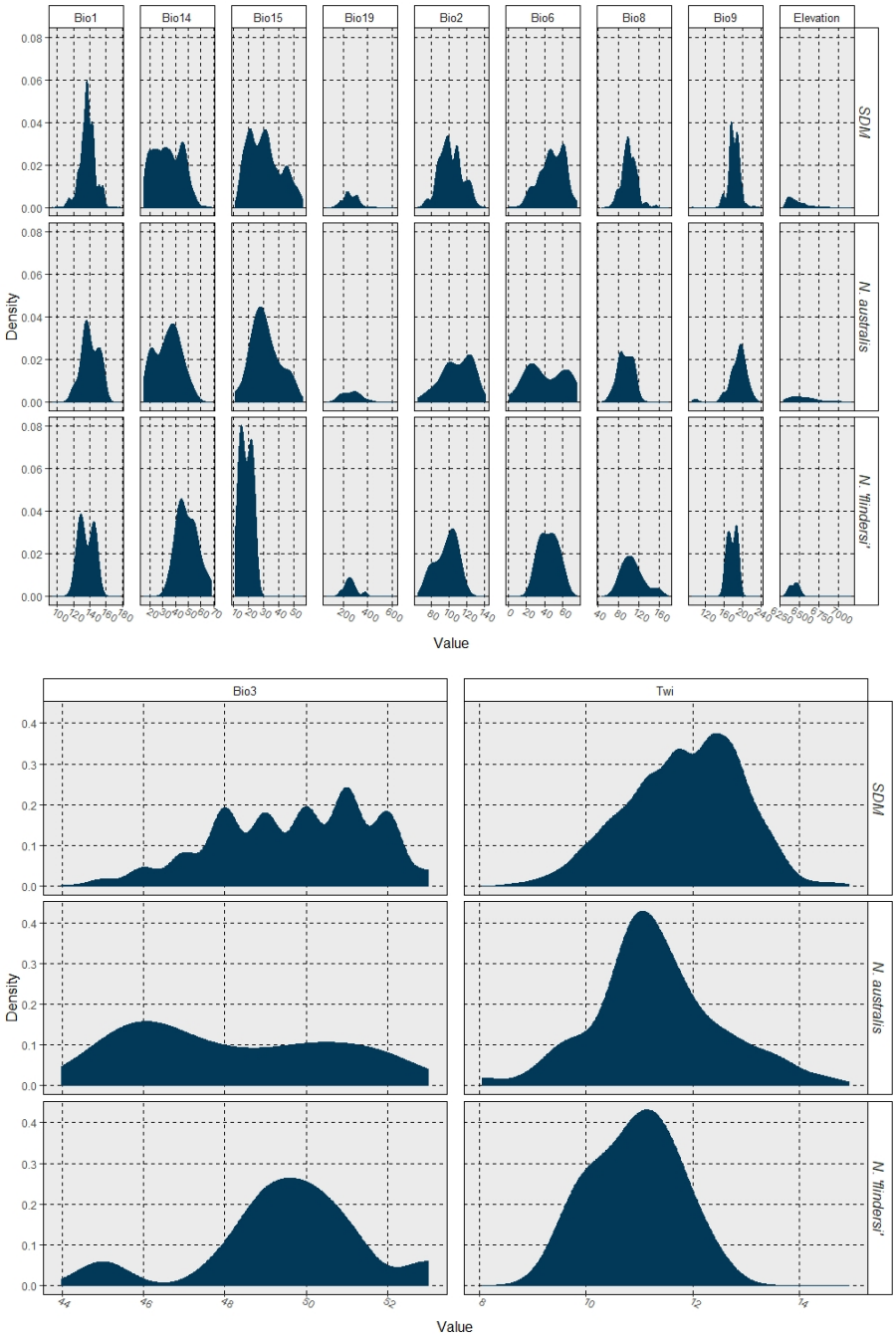
Appendix 15: Coalescent model likelihoods based on biogeographic hypotheses using FastSimCoal2. Individual model specifications and their underlying hypotheses are detailed in Table S2 and Figure S5. Model likelihoods (∆Likelihood) are reported as the difference between the maximum estimated likelihood and the observed likelihood of the SFS under each model. For each set of models, comparative likelihoods were calculated using AIC as [(2 x number of parameters) – (2 x log likelihood of the model)]. The difference between the AIC of each model and the lowest AIC of the model set (∆AIC) was also used to evaluate comparative likelihoods. Models were also compared using Akaike Weights to better characterise the fit of a model within a set. N/A denotes models which could not be estimated even under broad priors due to non-coalescence of demes. The best model of each set is highlighted in green.

Model	No. parameters	∆Likelihood	AIC	∆AIC	Akaike Weight
1A	4	939.081	38816.7869	1915.140203	0
1B	6	522.345	36901.6467	0	1
1C	6	6368.01	63821.92887	26920.28218	0
1D	8	719.992	37815.84477	914.1980717	3.0508×10^{-199}
2A	7	22822.537	228207.3438	91919.23483	0
2B	13	2859.923	136288.1089	0	1
2C	8	25908.698	244537.9509	108249.842	0
2D	10	22404.886	228406.3004	92118.19146	0
2E	16	3777.704	142636.9572	6348.848263	0
3A	25	30794.197	914430.059	0	1
3B	25	58640.469	1042666.881	128236.8216	0
3C	25	58671.987	1042812.026	128381.9674	0
3D	25	104816.136	1284161.485	369731.4258	0
3E	25	36531.741	940852.4256	26422.36657	0
4A	19	4427.185	234356.2201	3570.03817	0
4B	27	4456.823	234508.7081	3722.526204	0
4C	18	4298.586	233761.9998	2975.817889	0
4D	26	4218.219	233407.8961	2621.714177	0
4E	26	3648.921	230786.1819	0	1
5A	4	520.131	14400.71798	221.3401875	8.6415×10^{-49}
5B	4	531.54	14453.25837	273.8805742	3.36968×10^{-60}
5C	6	471.199	14179.3778	0	1
5D	4	N/A	N/A	N/A	N/A
5E	4	N/A	N/A	N/A	N/A
6A	7	3463.211	36803.27241	10847.18472	0
6B	13	1101.337	25938.44068	0	1
6C	10	3561.24	37260.71264	11304.62494	0
6D	16	1129.353	26073.45913	117.3714358	4.79862×10^{-30}

Appendix 16: One-dimensional site-frequency spectra for stairway plot analysis. SNPs were called independently for each population and hence are variable in number across the populations.



Appendix 17: Relative density plots of environmental variables within the SDM and each LDM individually. Variables Bio3 and Twi are plotted separately due to variance in maximum density. Environmental data across all variables were sampled from points using ArcMap based on 2,528 SDM occurrences, 61 *N. australis* LDM occurrences and 11 *N.* 'flindersi' LDM occurrences.



Appendix 18: Initial prior ranges for individual population demographic models estimated within FastSimCoal2. All priors were set using uniform distributions, with the same conditions across all six populations for the same model. The relationship of parameters to the model, and broad overviews of the models themselves, are demonstrated in Figure 3.2.

Parameter	Model				
	1	**2**	**3**	**4**	**5**
NCURR	10^3 - 10^4	10^3 - 10^4	10^3 - 10^4	10^3 - 10^4	10^3 - 10^4
RATE	-10^{-4} – 0	0 – 10^{-4}	NA	-10^{-4} – 0	0 – 10^{-4}
T1	18,000 – 21,000	18,000 – 21,000	NA	18,000 – 21,000	18,000 – 21,000
R2	NA	NA	NA	0 – 10^{-4}	-10^{-4} – 0
D2	NA	NA	NA	1,000 – 20,000	1,000 – 20,000
T2	NA	NA	NA	= T1 + D2	= T1 + D2

Appendix 19: Pearson's pairwise correlation for all 19 bioclimatic variables obtained from WorldClim v1.4. Selected uncorrelated variables (|R| < 0.8) are highlighted in bold, with correlation values between these variables highlighted in green.

Variable	**Bio1**	**Bio2**	**Bio3**	Bio4	Bio5	**Bio6**	Bio7	**Bio8**	**Bio9**	Bio10	Bio11	**Bio12**	Bio13	Bio14	**Bio15**	Bio16	Bio17	Bio18	Bio19
Bio1																			
Bio2	0.72																		
Bio3	0.08	0.06																	
Bio4	0.60	0.83	-0.47																
Bio5	0.93	0.89	-0.05	0.82															
Bio6	0.48	-0.16	0.41	-0.36	0.17														
Bio7	0.69	0.94	-0.24	0.96	0.89	-0.28													
Bio8	0.62	0.46	0.04	0.40	0.59	0.30	0.44												
Bio9	0.51	0.35	0.10	0.23	0.46	0.30	0.31	-0.19											
Bio10	0.95	0.83	-0.10	0.79	0.98	0.25	0.84	0.60	0.47										
Bio11	0.88	0.43	0.37	0.18	0.69	0.79	0.30	0.52	0.52	0.73									
Bio12	-0.80	-0.75	-0.15	-0.57	-0.82	-0.24	-0.69	-0.51	-0.42	-0.80	-0.66								
Bio13	-0.78	-0.77	-0.10	-0.61	-0.82	-0.18	-0.72	-0.51	-0.41	-0.80	-0.61	0.98							
Bio14	-0.76	-0.63	-0.31	-0.37	-0.71	-0.39	-0.51	-0.38	-0.50	-0.70	-0.73	0.91	0.84						
Bio15	-0.22	-0.35	0.36	-0.54	-0.35	0.27	-0.47	-0.36	0.09	-0.34	0.05	0.21	0.35	-0.14					
Bio16	-0.79	-0.77	-0.09	-0.63	-0.83	-0.18	-0.73	-0.54	-0.38	-0.81	-0.61	0.98	1.00	0.84	0.36				
Bio17	-0.77	-0.67	-0.26	-0.43	-0.74	-0.34	-0.57	-0.40	-0.49	-0.73	-0.71	0.94	0.88	0.99	-0.09	0.87			
Bio18	-0.71	-0.60	-0.28	-0.35	-0.67	-0.35	-0.49	-0.23	-0.62	-0.66	-0.69	0.87	0.81	0.95	-0.14	0.79	0.95		
Bio19	-0.75	-0.76	-0.07	-0.62	-0.81	-0.15	-0.72	-0.61	-0.27	-0.78	-0.57	0.96	0.98	0.79	0.41	0.99	0.83	0.71	

Appendix 20: Custom R script to calculate 90% minimum training presence thresholds for converting SDMs to binary presence-absence maps. A threshold is calculated for the ensemble model, and a separate threshold for each separate model (i.e. 27 thresholds in Chapter 3; 45 thresholds in Chapter 4).

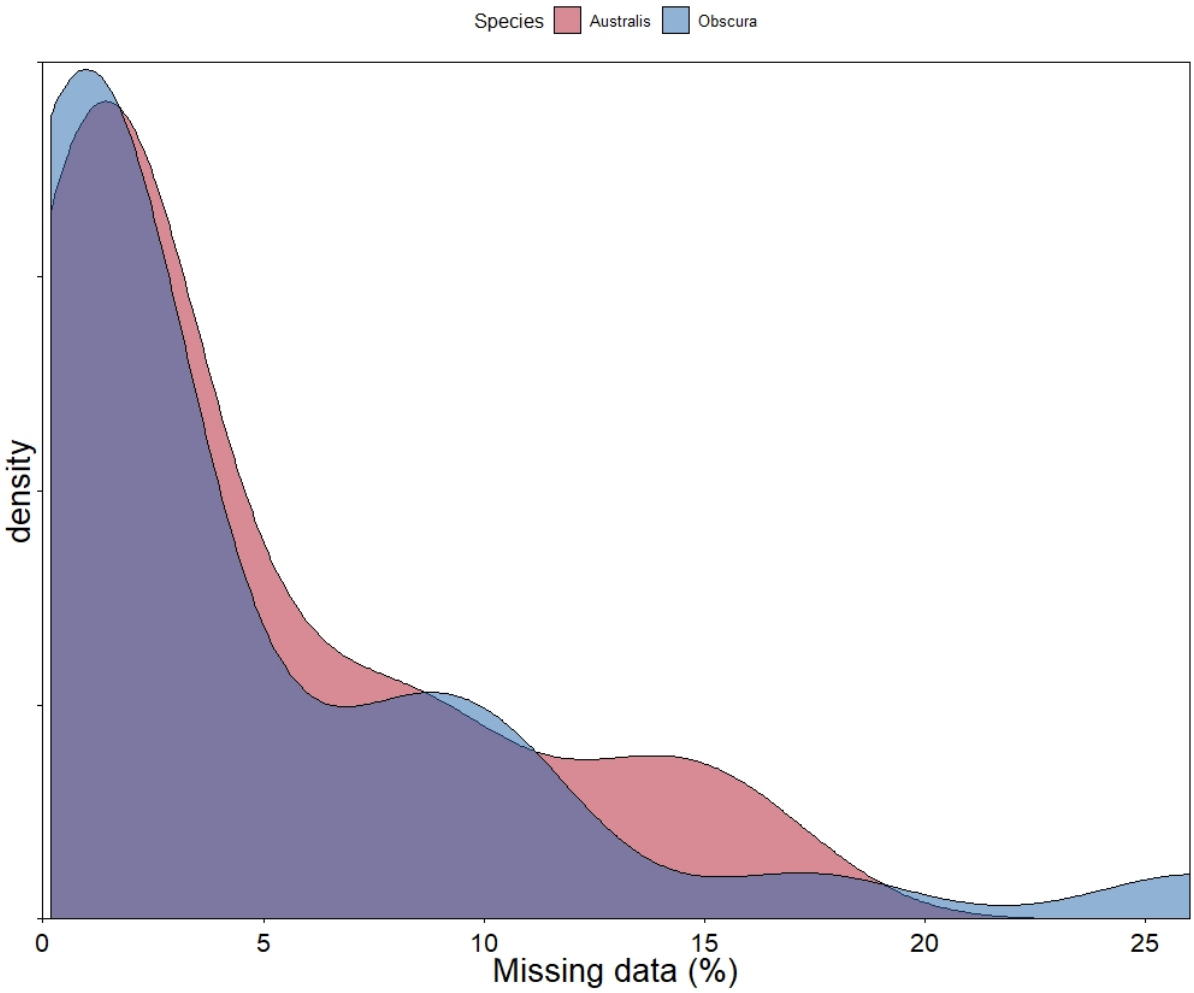
```
#!/usr/bin/env Rscript
library(biomod2)
library(raster)

#Used to loop over all time periods
species <- "Nannoperca.obscura"
projlist <- as.list(c("current","LateHolo","MidHolo","EarlyHolo","YDS","BA","HS1","LastInterglacial","MIS19","MidPlio","M2Plio"))
occs <- read.csv("SubsampledYarras.csv") #load occurrence data to base training

#Function for binary conversion of current SDM: adapted from https://babichmorrowc.github.io/post/2019-04-12-sdm-threshold/, accessed 10/10/2019
sdm_threshold <- function(sdm, occs, type = "mtp", binary = FALSE){
  occPredVals <- raster::extract(sdm, occs)
  if(type == "mtp"){
    thresh <- min(na.omit(occPredVals))
  } else if(type == "p10"){
    if(length(occPredVals) < 10){
      p10 <- floor(length(occPredVals) * 0.9)
    } else {
      p10 <- ceiling(length(occPredVals) * 0.9)
    }
    thresh <- rev(sort(occPredVals))[p10]
  }
  sdm_thresh <- sdm
  sdm_thresh[sdm_thresh < thresh] <- NA
  if(binary){
    sdm_thresh[sdm_thresh >= thresh] <- 1
  }
  return(sdm_thresh)
}

#script for binary conversion of projections as a function
```

```
#Function to loop the threshold calculation for all models
sdm_projthresh <- function(sdm, binary = FALSE){
  sdm_thresh <- sdm
  #thresh <- dfthresh(sdm[[i]])
  sdm_thresh[sdm_thresh < thresh] <- NA
  if(binary){
    sdm_thresh[sdm_thresh >= thresh] <- 1
  }
  return(sdm_thresh)
}

#actual scripts start here
#--------------------ENSEMBLE MODEL THRESHOLD-------------------------------
#1: define threshold for p10 based on current SDM and occs in ensemble model
species = "YPP"
outname <- paste(species, "Current", "MeanEM.asc", sep="")
r <- raster(outname)
occPredVals <- raster::extract(r, occs[,2:3]) #get suitability at all occurrence sites
if(length(occPredVals) < 10){
  p10 <- floor(length(occPredVals) * 0.9)
} else {
  p10 <- ceiling(length(occPredVals) * 0.9)
}
thresh <- rev(sort(occPredVals))[p10] #calculate suitability threshold above which at least 90% of training sites occur
print(thresh)

#2: loop over all ensemble projections using the previously determined threshold
for (proj in projlist) {
  outname <- paste(species, proj, "MeanEM.asc", sep="")
  r <- raster(outname)
  bin<-sdm_projthresh(r, binary = TRUE) #convert projections to binary using 10% training threshold
  counts <- freq(bin, digits = 0, value=1, useNA="no", merge=FALSE, progress='') #count number of presence cells
  print(counts)
  countname <- paste(species, proj, "EMcounts_2.csv", sep="")
  write.csv(counts, countname) #export results to a csv
}

#-------------------------INDIVIDUAL MODELS THRESHOLDS--------------------
```

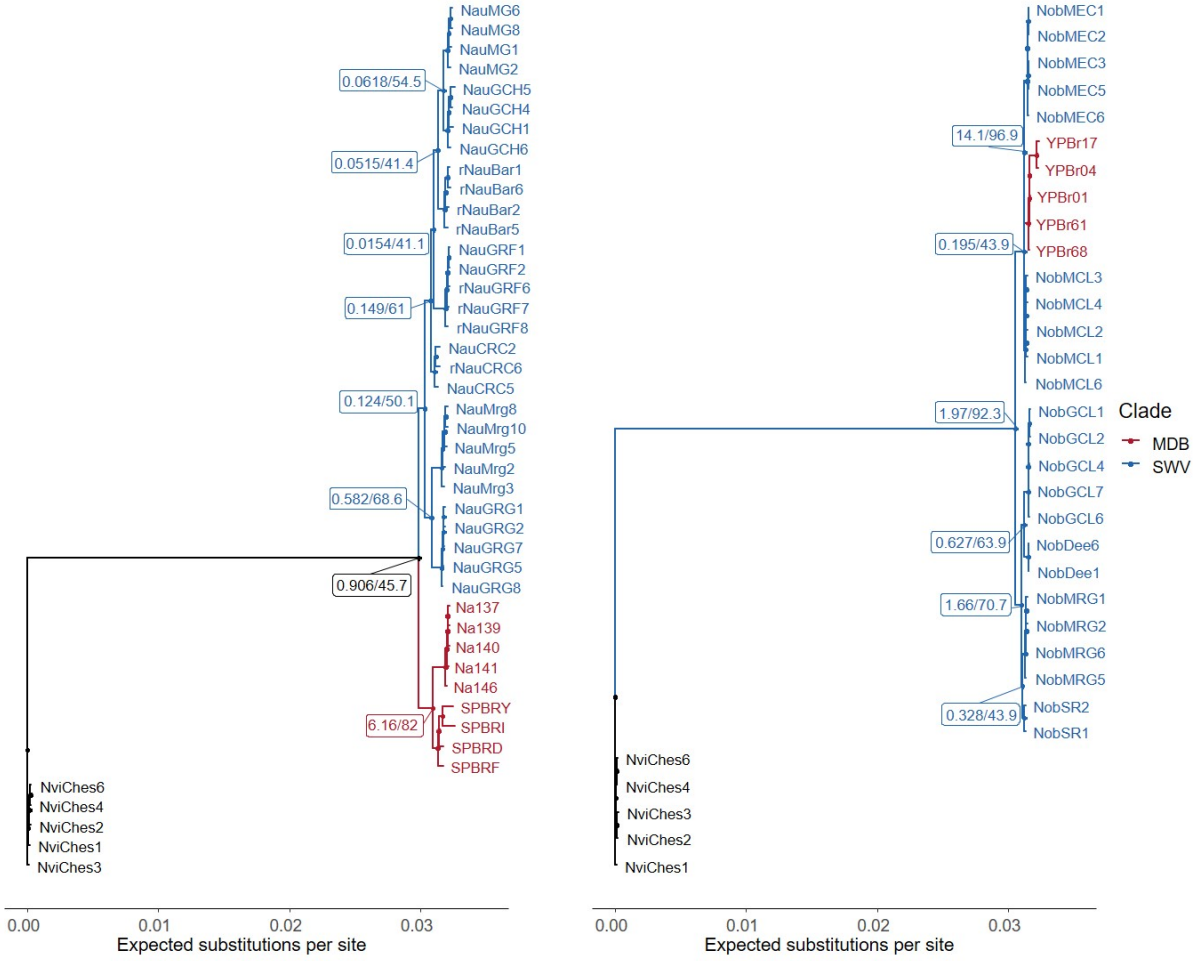
```
#Process to estimate binary counts per time period for each individual model
#1: Estimate binary thresholds for each model separately and concatenate to a data frame
dfthresh = NULL
outname <- paste(species, "/proj_current", "/",species, ".current.projection.out", sep="")
assign('out', get(load(outname)))
pred<-get_predictions(out)  #get predictions from file
for (i in 1:length(pred)) {
  sdm<-pred[[i]]
  occPredVals <- raster::extract(sdm, occs[,2:3]) #get suitability at all occurrence sites
  occPredValsTrim <- occPredVals[occPredVals != 0] #remove some values that were zeroes - possible error in some cells
  if(length(occPredValsTrim) < 10){
    p10 <- floor(length(occPredValsTrim) * 0.9)
  } else {
    p10 <- ceiling(length(occPredValsTrim) * 0.9)
  }
  thresh <- rev(sort(occPredValsTrim))[p10] #calculate suitability threshold above which at least 90% of training sites occur
  name <- names(sdm)
  dfthresh = rbind(dfthresh, data.frame(name, thresh, length(occPredValsTrim)))
}
rownames(dfthresh) <- dfthresh[,1]
dfthresh[,1] <- NULL #to set model names as row names and remove extra column

#2:Estimate binary counts across all models and timeslices, using the appropriate threshold per model
for (proj in projlist) {
  outname <- paste(species, "/proj_", proj, "/",species, ".", proj, ".projection.out", sep="")
  assign('out', get(load(outname)))
  pred<-get_predictions(out)  #get predictions from file
  for (i in 1:length(pred)) {
      sdm<-pred[[i]]
      rastname<-names(sdm)
      rastthresh <- dfthresh[rastname, thresh] #relate each model to a thresh value from the previous step
      bin<-sdm_projthresh(sdm, binary = TRUE)  #convert projections to binary using 10% training threshold
      counts <- freq(bin, digits = 0, value=1, useNA="no", merge=FALSE, progress='') #count number of presence cells
      print(counts)
      countres <- rbind(data.frame(rastname, counts))
      rownames(countres) <- countres[,1]
      countres[,1] <- NULL
      countname <- paste(species, proj, "indiv_counts.csv", sep="_")
```

```
    write.table(countres, countname, append = TRUE, sep=",", col.names = FALSE) #export results to a csv
  }
}
```

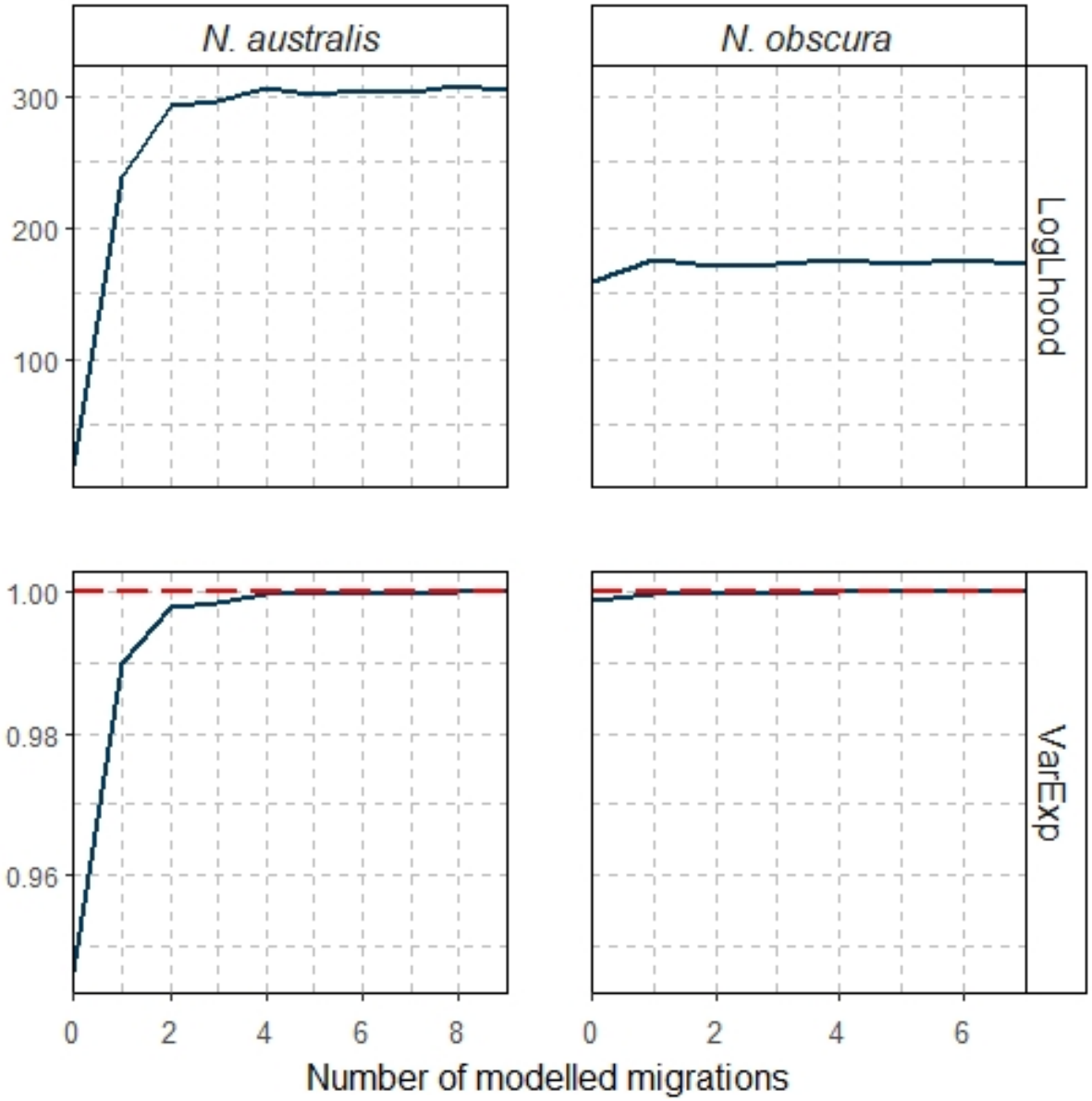
Appendix 21: Density plot of missing data per sample in the species-wide alignments. Colours denote species (*N. australis* in red, *N. obscura* in blue).



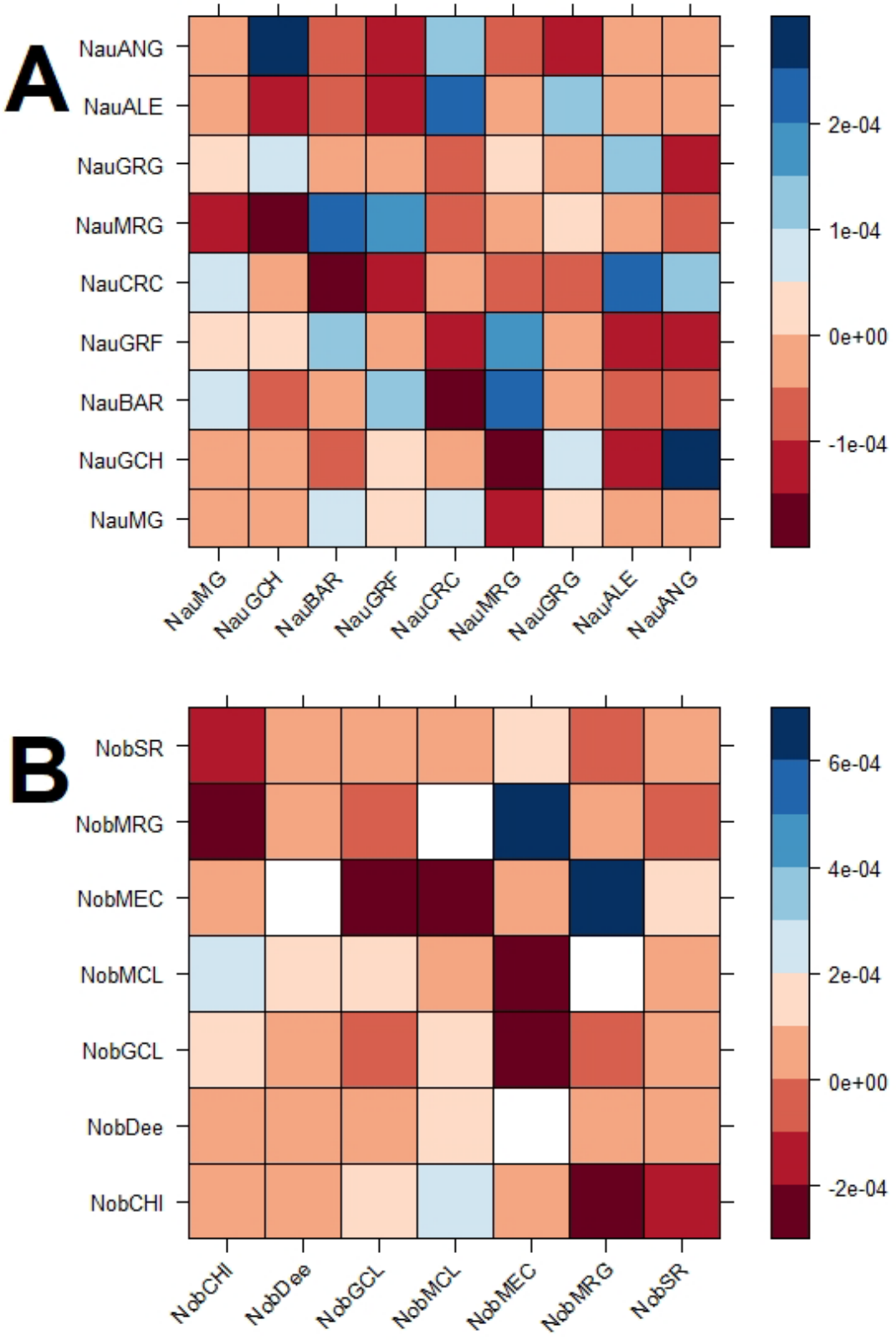
Appendix 22: Gene and site concordance factors between gene trees estimated with individual RAD loci and the concatenated phylogeny for *N. australis* (left) and *N. obscura* (right). Population-level divergences and above are labelled with concordance factors, with gene concordance factors reported first and site concordance factors reported second. Both factors are scaled out of 100.



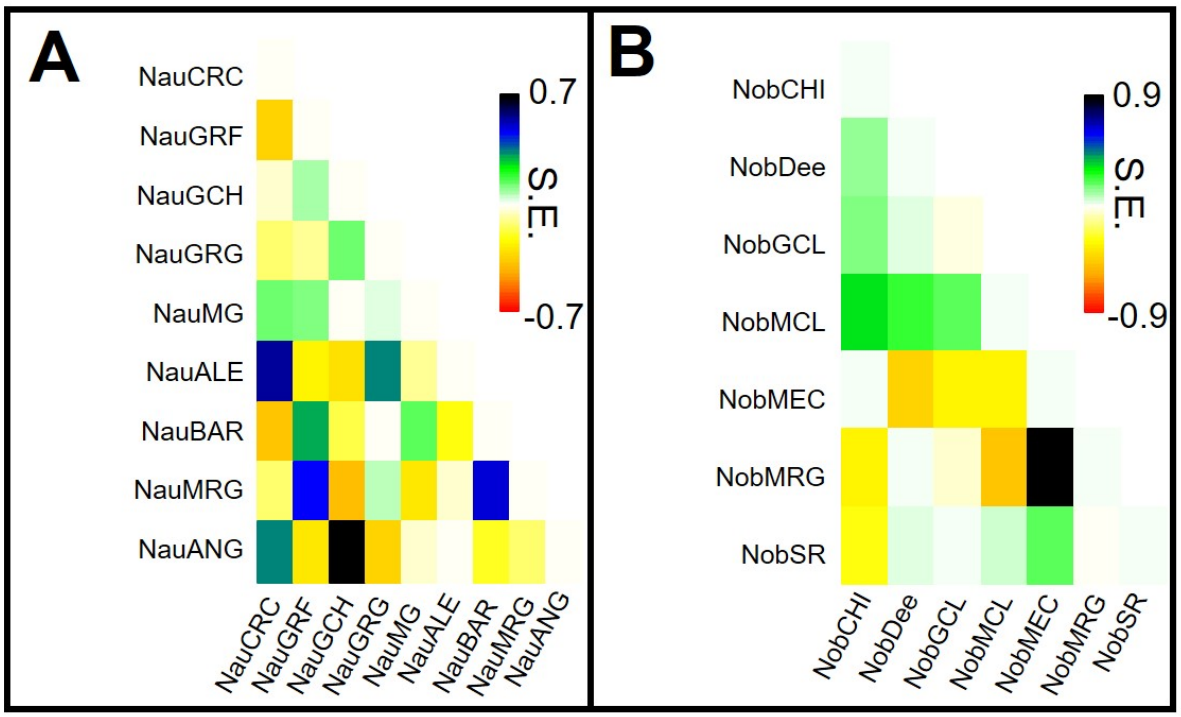
Appendix 23: Log likelihoods (top) and percentage of variation explained (bottom) of population mixtures and splits modelled with varying numbers of migration edges in TreeMix. Dashed red line indicates the asymptote of the variation explained.



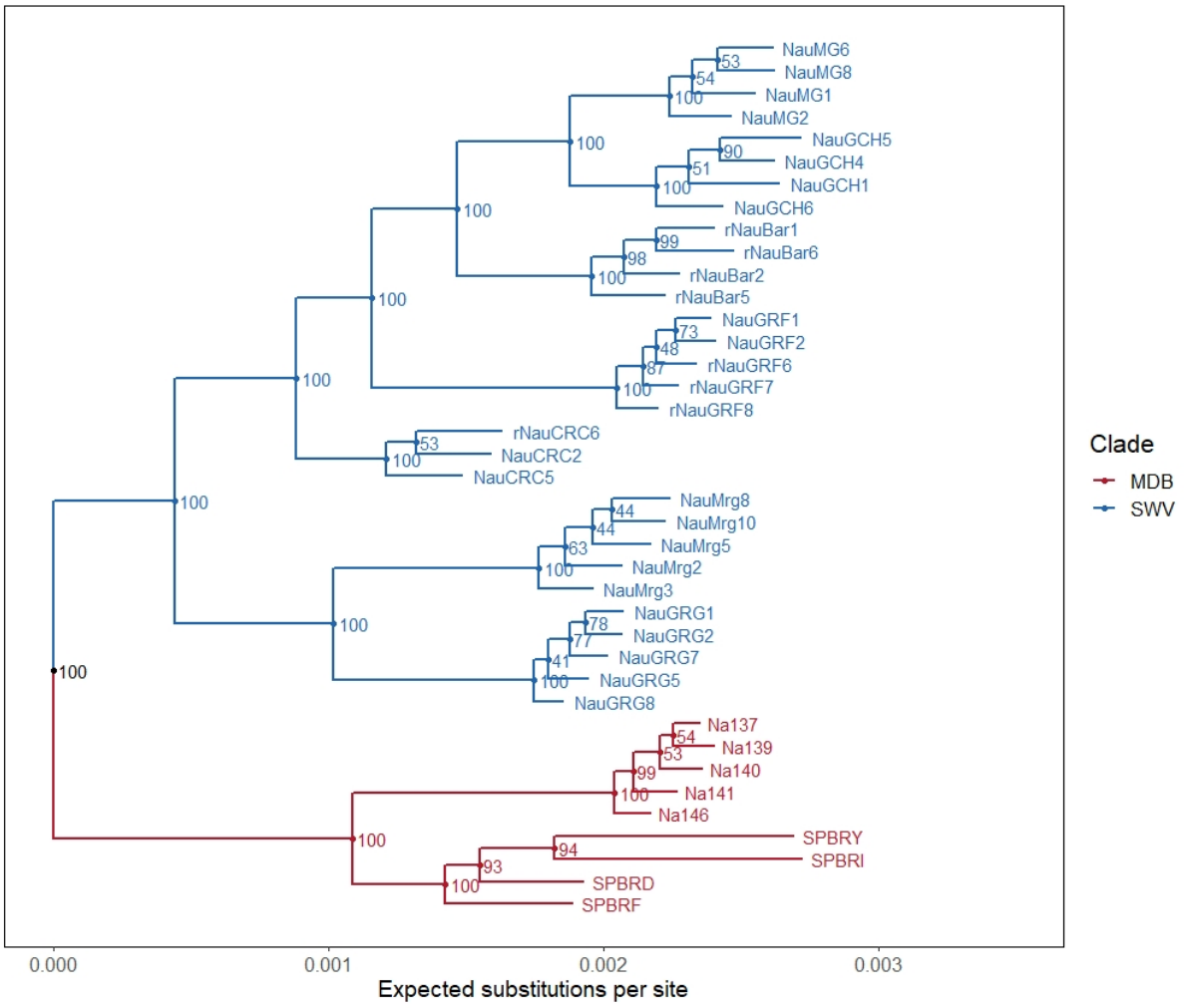
Appendix 24: Residual matrices for **A)** *N. australis* and **B)** *N. obscura* under the best supported TreeMix models.



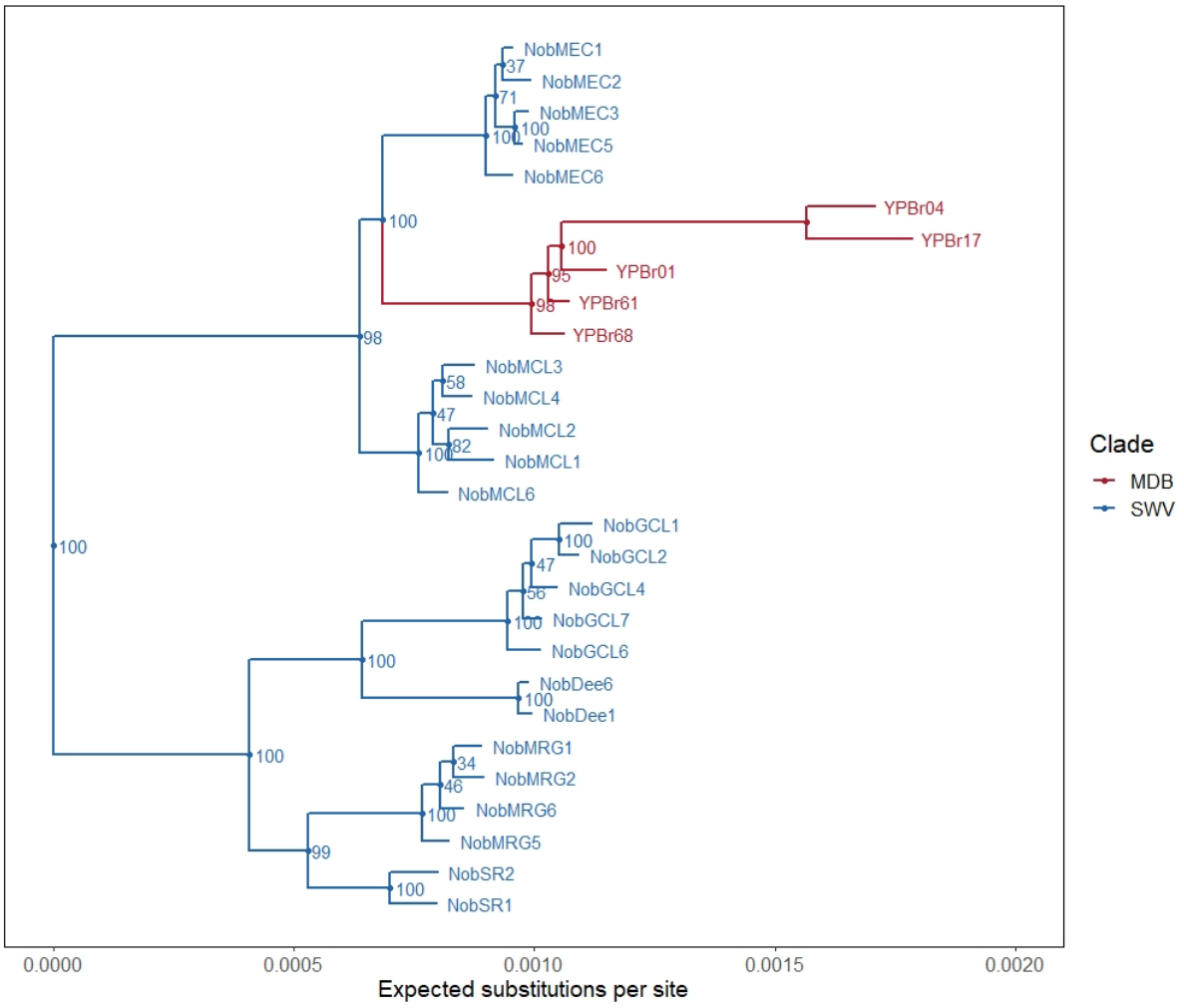
Appendix 25: Standard errors of the covariance matrices for **A)** *N. australis* and **B)** *N. obscura* under the best supported TreeMix models.



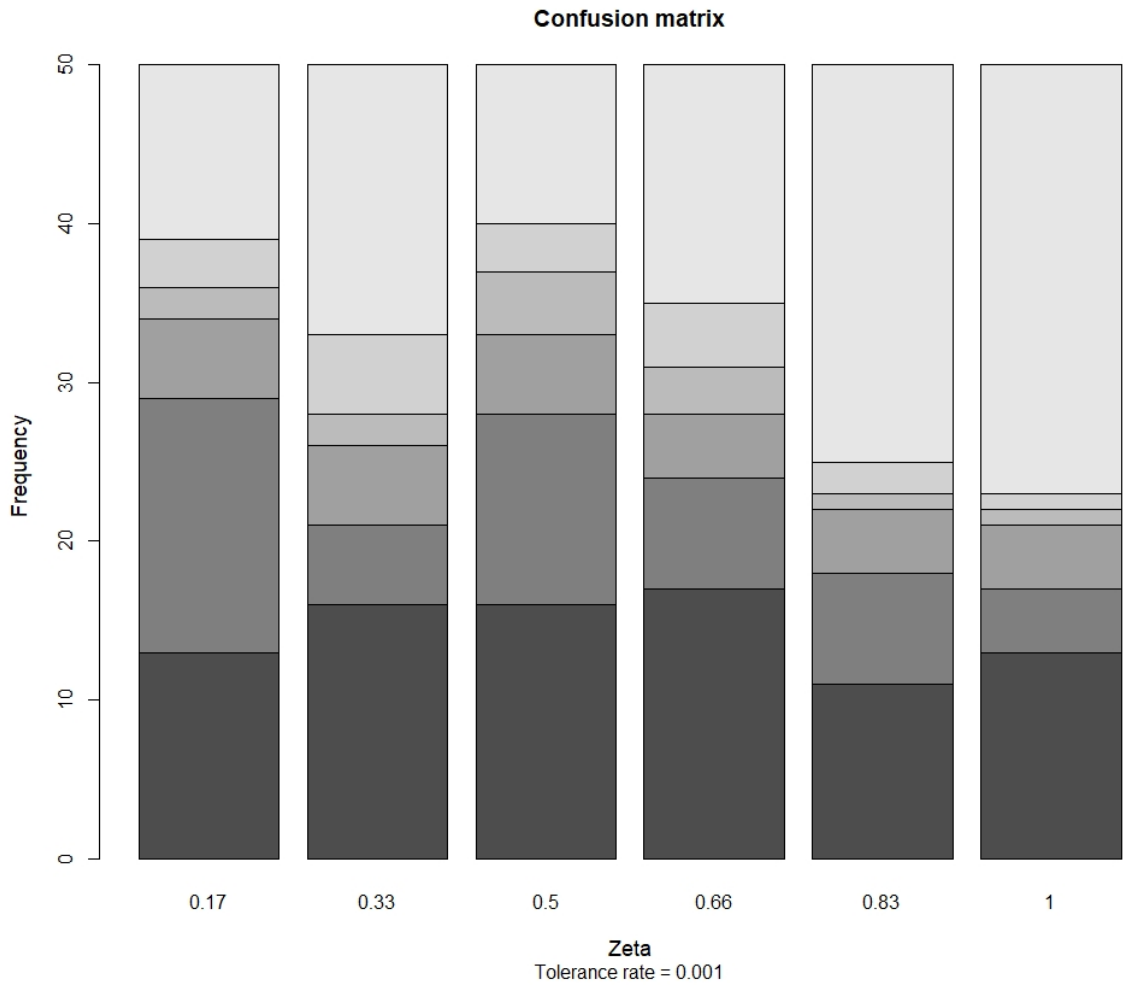
Appendix 26: Full ML phylogenetic tree for *N. australis* estimated using RAxML and based on 19,428 concatenated ddRAD loci. The tree was rooted using *N. vittata* as the outgroup. Node values show bootstrap support under 1,000 RELL bootstraps. Branch colours indicate the basin of origin for each clade (MDB = Murray-Darling Basin, SWV = southwest Victoria).



Appendix 27: Full ML phylogenetic tree for *N. obscura* estimated using RAxML and based on 12,705 concatenated ddRAD loci. The tree was rooted using *N. vittata* as the outgroup. Node values show bootstrap support under 1,000 RELL bootstraps. Branch colours indicate the basin of origin for each clade (MDB = Murray-Darling Basin, SWV = southwest Victoria).



Appendix 28: Confusion matrix of ξ hyperparameter within the co-demographic Multi-DICE model, estimated using 50 pseudo-observed datasets and the top 1,500 simulations (out of 1.5M simulations total). Colours range from dark grey (ξ = 0.17) to light grey (ξ = 1) for each possible value of ξ.



Appendix 29: Prediction errors of variables in co-demographic models within Multi-DICE. Tolerance = threshold for the proportion of the top number of simulations used in calculating errors, with the corresponding number of simulations. **τ** = timing of population size change (bottleneck). **ε1** = magnitude of exponential population size change (post-bottleneck). ***Ne*** = current effective population size. **Mean(ε1)** = average of ε1 across all six populations. **Ω(ε)** = dispersion index of ε1 (Var(ε1)/Mean(ε1)). Taxon-specific parameters (ε and *Ne*) are reported as the min – max range across all six populations.

		Parameter				
Tolerance	No. sims	τ	ε1	*Ne*	Mean(ε1)	Ω(ε1)
0.0067%	100	0.618	0.86 – 1.159	0.864 – 1.06	0.829	0.957
0.1%	1,500	0.622	0.901 – 1.214	0.851 – 1.043	0.798	0.969
0.5%	7,500	0.614	0.917 – 1.224	0.861 – 1.044	0.788	0.993
1%	15,000	0.611	0.925 – 1.237	0.863 – 1.047	0.789	1.000

Appendix 30: Posterior distributions of parameters from co-demographic models in Multi-DICE. Posterior probabilities were calculated using the top 0.0067% (*n* = 100) simulations out of 1.5M total simulations. **τ** = timing of population size change (bottleneck). **ε1** = magnitude of exponential population decline. ***Ne*** = current effective population size. **Mean(ε)** = average of ε across all six populations. **Ω(ε)** = dispersion index of ε (Var(ε)/Mean(ε)). Taxon-specific parameters (ε1 and Ne) are reported as the min – max range across all six populations.

	Parameter				
	τ	ε1	*Ne*	Mean(ε1)	Ω(ε1)
Min	73,751	5.009 – 5.061	10,026 – 10,542	5.673	0.019
2.5%	77,073.5	5.083 – 5.218	10,155 – 11,554	6.393	0.152
Mean	91,452.34	8.438 – 8.734	22,212 – 24,887	8.618	1.41
97.5%	99,851.5	16.012 – 18.843	58,988 – 72,268	11.390	2.991
Max	99,971	17.606 – 19.996	82,156 – 99,863	12.575	3.338

Appendix 31: Custom *R* script to detect fixed differences, adapted from the gl.fixed.diff() function of the dartR package.

```
#!/usr/bin/env Rscript

library(devtools)

library(dartR)

library(melfuR)

library(dplyr)

library(RADami)

WPPv <- read.structure(file = "WPP1_5vqualNviB.str", n.ind = 33, n.loc = 17145, onerowperind = FALSE, col.lab = 1, row.marknames = 1, col.pop = 2) #read Structure file

locnam <- read.table("WPP1_5vq.lmiss") #read loci names based on a VCF

locinames <- locnam[-1, 1]

WPPgl <- gi2gl(WPPv)

locNames(WPPgl) <- locinames

ftable <- gl.percent.freq(gl = WPPgl, v = 2) #calculate allele frequencies in the putative species + all other species

p1 <- ftable[ftable$popn == 1, ] #get allele frequencies of putative species

p2 <- ftable[ftable$popn == 2, ] #get allele frequency of other combined species

rownames(p2) <- p2$locus

rownames(p1) <- p1$locus

both = NULL

both <- cbind(p2$locus, p1$frequency, p2$frequency, p1$nmissing, p2$nmissing) #collate frequencies and missing data counts

both = as.data.frame(both)

colnames(both) = c("locus", "p1freq","p2freq","p1nmiss","p2nmiss")

#Select loci with fixed differences

wMD <- filter(both, (p1freq==0 & p2freq==100) | (p1freq==100 & p2freq==0))

write.csv(x = wMD, file = “species_gpf_loci.csv”)

##the column of loci numbers (“locus”) is then used to subsample the ddRAD loci from the “.loci” output of PyRAD using pyrad_filter.py (https://github.com/pimbongaerts/radseq/blob/master/pyrad_filter.py)

##this returns a subsampled “.loci” object with all samples for the given loci with fixed differences

##a consensus sequence per locus is then obtained using the following command in the RADami R package

Cons <- consensus.pyRAD(pyIn = “fixed_diff_loci.loci”, from = NA, to = NA, fastaNames = T, writeFile = “cons_loci”)

##this returns a FASTA file with the consensus sequence per locus containing fixed differences to be annotated
```

Appendix 32: Custom perl script to annotate ddRAD loci containing fixed differences and extract GO terms from the UniProt database.

```
#!/usr/bin/perl

while (@ARGV){

$_=shift @ARGV;

if ($_=~ /^-pdb$/){$pdb=shift @ARGV;}

elsif ($_=~ /^-ndb$/){$ndb=shift @ARGV;}

elsif ($_=~ /^-qry$/){$qry=shift @ARGV;}

elsif ($_=~ /^-nc$/){$nc=shift @ARGV;}

elsif ($_=~ /^-of$/){$out=shift @ARGV;}

}


if (not defined ($pdb && $qry && $out)){print "\nThis script annotate a fasta file to a protein local database (UniProt) or nucleotide database (Genbank) and then create a table with results and GO terms.\n\nUsage:\nAnnGO.pl\n\t-pdb <path to protein database>\n\t-qry <path to fasta file of loci to be annotated>\n\t-of <path folder to save all outputs>\n\t-nc <number cores to run in parallel>\n\nFor example:\n\nAGOUP.pl -pdb /pygmy/uniprot.fasta -qry /pygmy/PPcandiloci.fasta -of /pygmy/blastout/ -nc 10 \n\n"; exit;}

if (not defined ($nc)) {$nc= 10;}

unless (-d $out){`mkdir $out`;}

our $prefo = $qry;

$prefo =~ s/\.fasta//;

@prefc = split /\//, $prefo;

$pref = $prefc[-1];

our $qbase="$out/$pref";

##blast to protein database

`cat $qry | parallel -j $nc -k --block 3k --recstart '>' --pipe 'blastx -db $pdb -query - -evalue 1e-3 -outfmt 6' >> $qbase\.prot`;

`awk '{if (\$3>=70 && \$11<0.001)print }' $qbase\.prot | sort -k 1,1 -k 12,12rn | awk '!seen[\$1]++' > $qbase\_P.hit`;

`awk '{if (\$3>=70 && \$11<0.001)print }' $qbase\.prot | sort -k 1,1 -k 12,12rn | awk '!seen[\$1]++' | awk '{print \$1, \$2}' | awk '!seen[\$2]++' > $qbase\_PU.hit`;


##blast to nucleotide database

`cat $qry | parallel -j $nc -k --block 3k --recstart '>' --pipe 'blastn -db $ndb -query - -evalue 1e-3 -outfmt 6' >> $qbase\.gen`;

`awk '{if (\$3>=70 && \$11<0.001)print }' $qbase\.gen | sort -k 1,1 -k 12,12rn | awk '!seen[\$1]++' > $qbase\_G.hit`;

`awk '{if (\$3>=70 && \$11<0.001)print }' $qbase\.gen | sort -k 1,1 -k 12,12rn | awk '!seen[\$1]++' | awk '{print \$1, \$2}' | awk '!seen[\$2]++' > $qbase\_GU.hit`;

`awk 'FNR == NR { h[\$1]; next }; !(\$1 in h)' $qbase\_P.hit $qbase\_G.hit >> $qbase\_UG.hit`;




##Import the information of each gene
```

```
our $infgene="$out/genbank_info";

unless (-d $infgene){`mkdir $infgene`;}

`cat  $qbase\_PU.hit | while read i; do hit=\$(echo \$i | cut -d " " -f 2 | cut -d "|" -f 2); echo \"wget -q -O -
http://www.uniprot.org/uniprot/\${hit}.txt\ >> $infgene/\${hit}.info"; done > $out/gbcmd`;

`cat $out/gbcmd | parallel \"eval {}\"`;


##create annotation table

my $cmd="echo
Locus\'\\t\'GeneAccession\'\\t\'MapPosition\'\\t\'GeneFullName\'\\t\'GeneShortName\'\\t\'GeneAltNames\'\\t\'GeneSubName\'\\t\'GeneGo\'\\t\
'CellularComponent\'\\t\'MolecularFunction\'\\t\'BiologicalProcess >> $qbase\_annotation.csv";

system ($cmd);

our @hits=`cat  $qbase\_P.hit`;

foreach $hit (@hits){
        my $fname= "-";
        my $sname= "-";
        my $altname="-";
        my $subname="-";
        my $GOT="GO:";
        my $CC="CC:";
        my $MF="MF:";
        my $BP="BP:";
        my @hitinf=split /\t/, $hit;
        my @UP=split /\|/, $hitinf[1];
        my $UPa=$UP[1];
        our @geni=`cat $infgene/$UPa\.info`;
        foreach $line (@geni){
                chomp $line;
                if ($line =~ m/^DE   RecName: Full=/) {
                        $fname = $line;
                        $fname =~ s/DE   RecName: Full=//;
                        $fname =~ s/\{.*//;
                }
                elsif ($line =~ m/^GN   Name=/) {
                        $sname = $line;
                        $sname =~ s/GN   Name=//;
                        $sname =~ s/\{.*//;
                }
                elsif ($line =~ m/^DE   AltName: Full=/) {
                        $altname = $line;
```

```
                $altname =~ s/DE   AltName: Full=//;
                $altname =~ s/\{.*//;
        }
        elsif ($line =~ m/^DE   SubName: Full=/) {
                $subname = $line;
                $subname =~ s/DE   SubName: Full=//;
                $subname =~ s/\{.*//;
        }
        elsif ($line =~ m/^DR   GO;/) {
                my $GOTi = $line;
                $GOTi =~ s/DR   GO;//;
                my @GOTin= split /;/, $GOTi;
                #$GOTin[0]=~ s/GO\://;
                $GOTin[0]=~ s/\s//g;
                $GOT= $GOT.$GOTin[0].",";
                if ($GOTin[1]=~ m/C:/){
                $GOTin[1]=~ s/C\://;
                $CC=$CC.$GOTin[1].",";

                }
                if ($GOTin[1]=~ m/F:/){
                $GOTin[1]=~ s/F\://;
                $MF=$MF.$GOTin[1].",";
                }
                if ($GOTin[1]=~ m/P:/){
                $GOTin[1]=~ s/P\://;
                $BP=$GOTin[1];
                }
        }
    }
$GOT =~ s/, $//;
$MF =~ s/\'/\\'/g;
$CC =~ s/\'/\\'/g;
$BP =~ s/\'/\\'/g;
my $cmd2="echo $hitinf[0]\'\\t\'$UP[-1]\'\\t\'$hitinf[8]-
$hitinf[9]\'\\t\'$fname\'\\t\'$sname\'\\t\'$altname\'\\t\'$subname\'\\t\'$GOT\'\\t\'$CC\'\\t\'$MF\'\\t\'$BP >> $qbase\_annotation.csv";
$cmd2 =~ s/\(/\\(/g;
$cmd2 =~ s/\)/\\)/g;
```

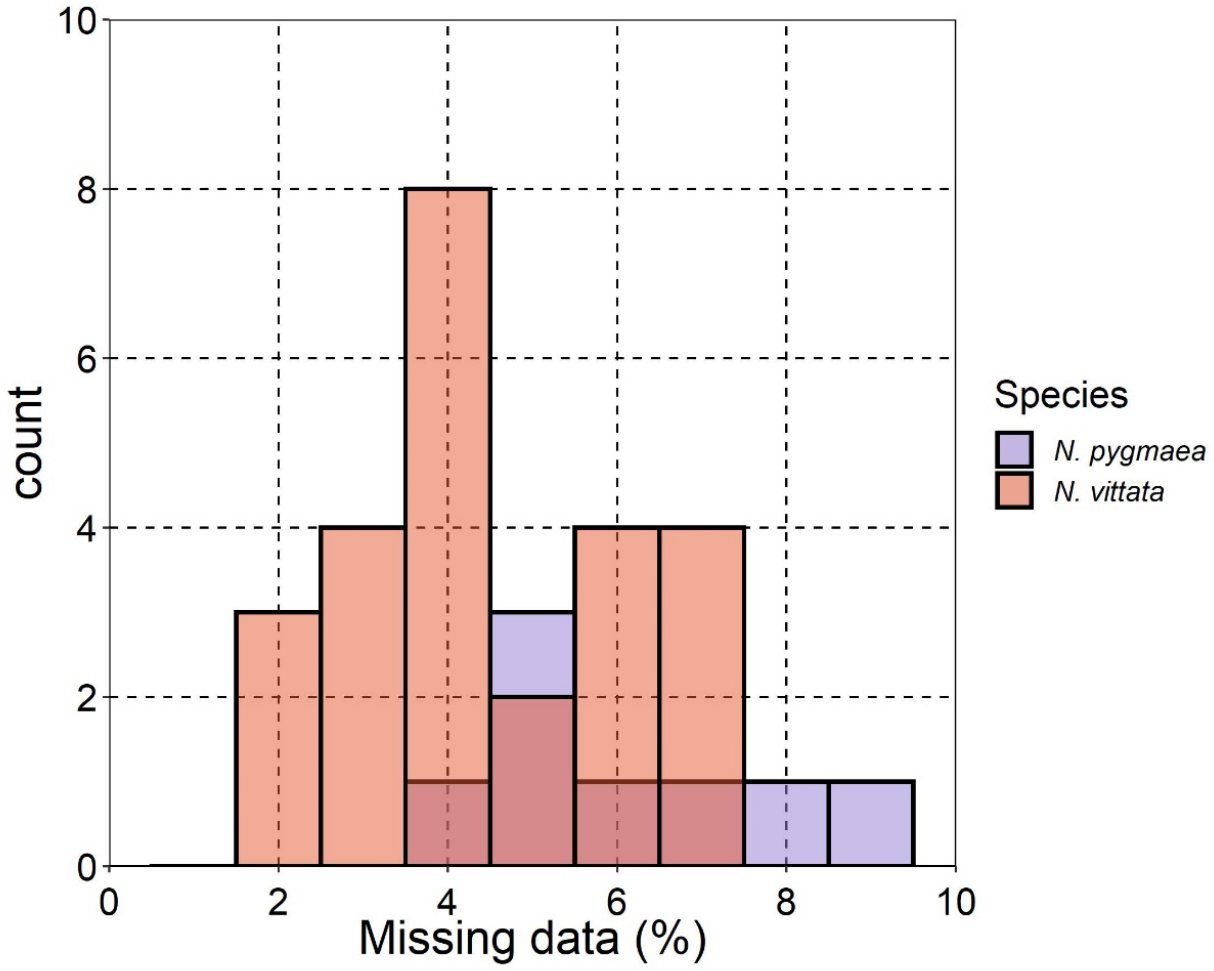
```
$cmd2 =~ s/\;//g;
my $cmd3="echo $UP[-1]\'\\t\'$GOT >> $qbase\_GO.tab";
print "$cmd2\n";
system ($cmd2);
system ($cmd3);
}
##add not annotated genes table
our @hits2=`cat  $qbase\_UG.hit`;
foreach $hit2 (@hits2){
        my $fname= "-";
        my $sname= "-";
        my $altname="-";
        my $subname="-";
        my $GOT="GO:";
        my $CC="CC:";
        my $MF="MF:";
        my $BP="BP:";
        my @hitinf=split /\t/, $hit;
        # my @UP=split /\|/, $hitinf[1];
        my $UPa=$hiting[1];
        our @geni=`cat $infgene2/$UPa\.info`;
        foreach $line (@geni){
                chomp $line;
                if ($line =~ m/^DEFINITION  /) {
                        $fname = $line;
                        $fname =~ s/DEFINITION  //;
                        $fname =~ s/\(.*//;
                }
                elsif ($line =~ m/\/gene=/) {
                        $sname = $line;
                        $sname =~ s/GN   Name=//;
                        $sname =~ s/\(.*//;
                }
                elsif ($line =~ m/^DE   AltName: Full=/) {
                        $altname = $line;
                        $altname =~ s/DE   AltName: Full=//;
                        $altname =~ s/\(.*//;
                }
```

```
                elsif ($line =~ m/^DE   SubName: Full=/) {
                        $subname = $line;
                        $subname =~ s/DE   SubName: Full=//;
                        $subname =~ s/\{.*//;
                }
                elsif ($line =~ m/^DR   GO;/) {
                        my $GOTi = $line;
                        $GOTi =~ s/DR   GO;//;
                        my @GOTin= split /;/, $GOTi;
                        $GOTin[0]=~ s/GO\://;
                        $GOT= $GOT.$GOTin[0].",";
                        if ($GOTin[1]=~ m/C:/){
                        $GOTin[1]=~ s/C\://;
                        $CC=$CC.$GOTin[1].",";

                        }
                        if ($GOTin[1]=~ m/F:/){
                        $GOTin[1]=~ s/F\://;
                        $MF=$MF.$GOTin[1].",";
                        }
                        if ($GOTin[1]=~ m/P:/){
                        $GOTin[1]=~ s/P\://;
                        $BP=$GOTin[1];
                        }
                }
        }
$MF =~ s/\'/\\'/g;
$CC =~ s/\'/\\'/g;
$BP =~ s/\'/\\'/g;
my $cmd3="echo $hitinf[0]\'\\t\'$UP[-1]\'\\t\'$hitinf[8]-
$hitinf[9]\'\\t\'$fname\'\\t\'$sname\'\\t\'$altname\'\\t\'$subname\'\\t\'$GOT\'\\t\'$CC\'\\t\'$MF\'\\t\'$BP >> $qbase\_annotation.csv";
$cmd3 =~ s/\(/\\(/g;
$cmd3 =~ s/\)/\\)/g;d
$cmd3 =~ s/\;//g;
print "$cmd3\n";
system ($cmd3);
}
```

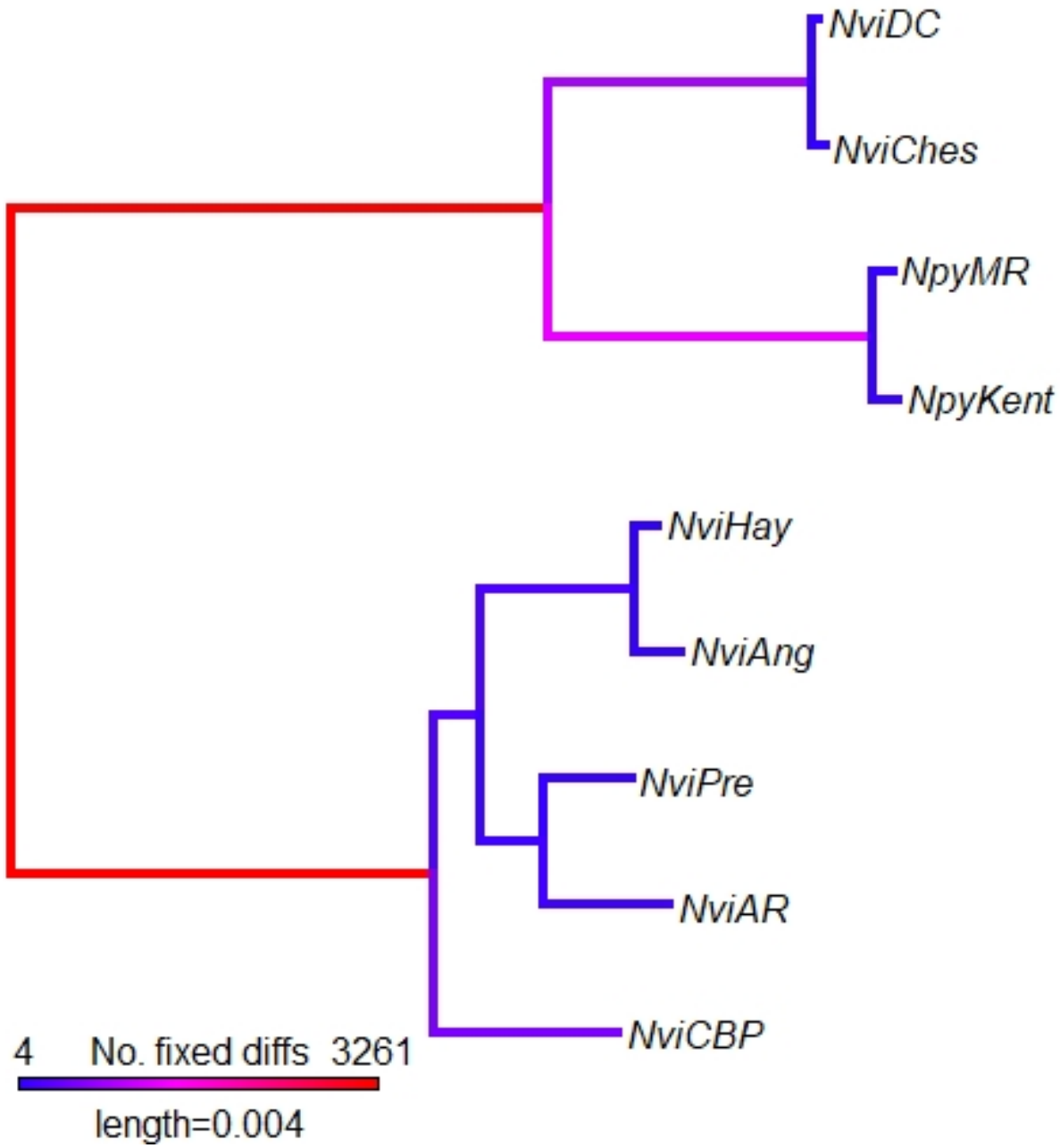
Appendix 33: Pearson's pairwise correlation for all 19 bioclimatic variables obtained from WorldClim v1.4. Selected uncorrelated variables (|R| < 0.8) are highlighted in bold, with correlation values between these variables highlighted in green.

Variable	**bio1**	bio2	**bio3**	bio4	bio5	bio6	bio7	**bio8**	**bio9**	bio10	**bio11**	**bio12**	bio13	**bio14**	**bio15**	bio16	bio17	bio18	bio19
bio1																			
bio2	0.56																		
bio3	-0.65	-0.56																	
bio4	0.74	0.84	-0.90																
bio5	0.91	0.81	-0.82	0.94															
bio6	0.12	-0.69	0.42	-0.55	-0.30														
bio7	0.68	0.93	-0.82	0.98	0.92	-0.64													
bio8	0.65	0.05	-0.12	0.15	0.40	0.53	0.10												
bio9	0.41	0.30	-0.18	0.24	0.36	0.08	0.26	0.17											
bio10	0.95	0.71	-0.81	0.91	0.98	-0.16	0.85	0.47	0.37										
bio11	0.59	-0.18	0.10	-0.10	0.22	0.83	-0.16	0.74	0.41	0.33									
bio12	-0.43	-0.81	0.55	-0.71	-0.67	0.65	-0.80	-0.06	-0.08	-0.57	0.28								
bio13	-0.30	-0.72	0.44	-0.60	-0.54	0.66	-0.70	-0.02	0.02	-0.43	0.35	0.98							
bio14	-0.61	-0.56	0.54	-0.61	-0.67	0.16	-0.60	-0.20	-0.51	-0.67	-0.24	0.39	0.21						
bio15	0.05	-0.35	0.18	-0.26	-0.15	0.55	-0.35	0.08	0.44	-0.06	0.51	0.69	0.80	-0.31					
bio16	-0.31	-0.73	0.45	-0.61	-0.55	0.66	-0.71	-0.03	0.04	-0.44	0.35	0.98	1.00	0.22	0.80				
bio17	-0.75	-0.60	0.67	-0.73	-0.80	0.15	-0.70	-0.30	-0.47	-0.81	-0.28	0.45	0.27	0.92	-0.24	0.28			
bio18	-0.51	-0.57	0.54	-0.58	-0.62	0.26	-0.60	-0.07	-0.63	-0.59	-0.13	0.44	0.28	0.84	-0.25	0.27	0.87		
bio19	-0.33	-0.71	0.46	-0.62	-0.56	0.64	-0.71	-0.05	0.04	-0.46	0.33	0.98	0.99	0.22	0.80	1.00	0.29	0.27	

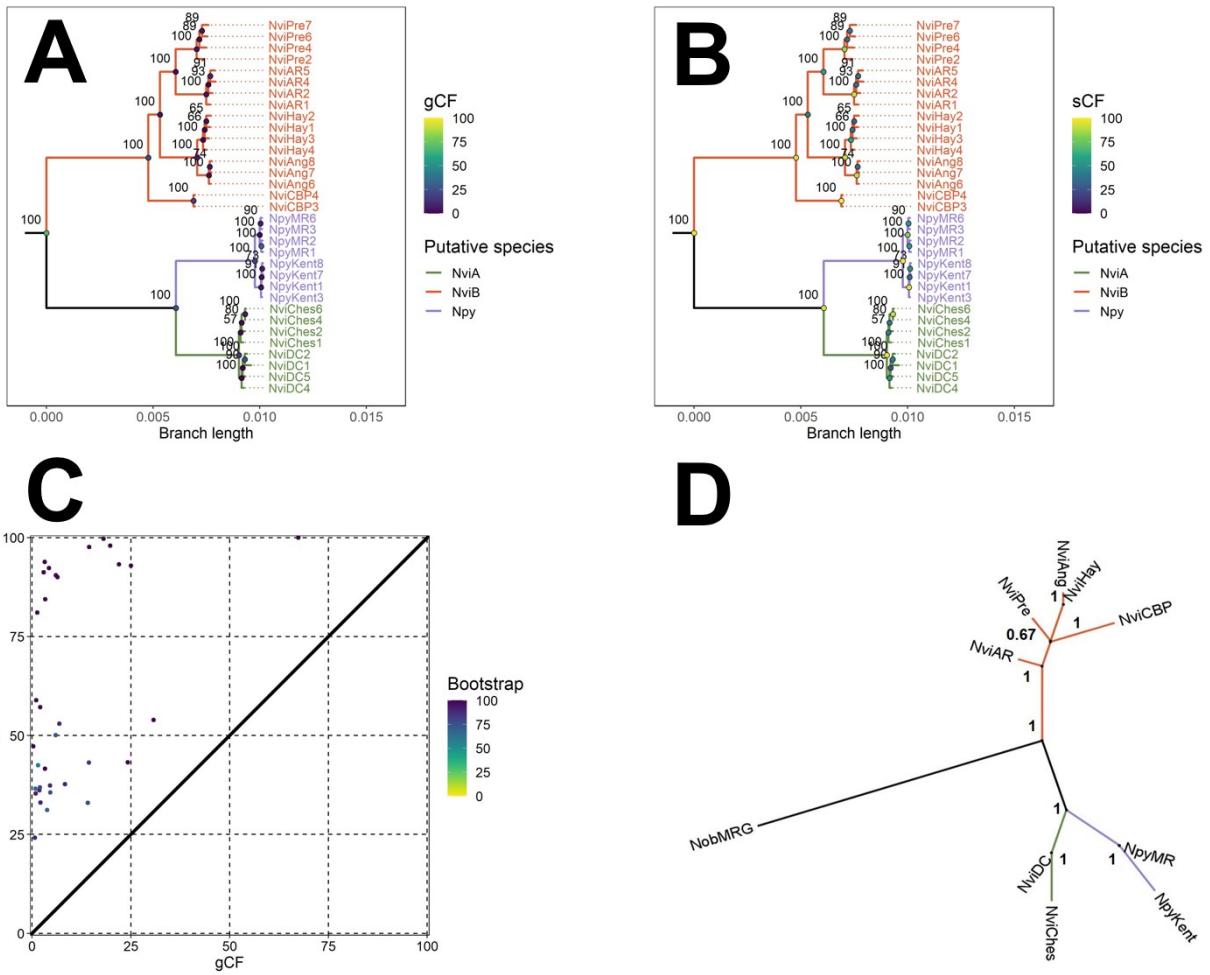
Appendix 34: Histogram of missing data per sample (%) for *N. vittata* and *N. pygmaea*.



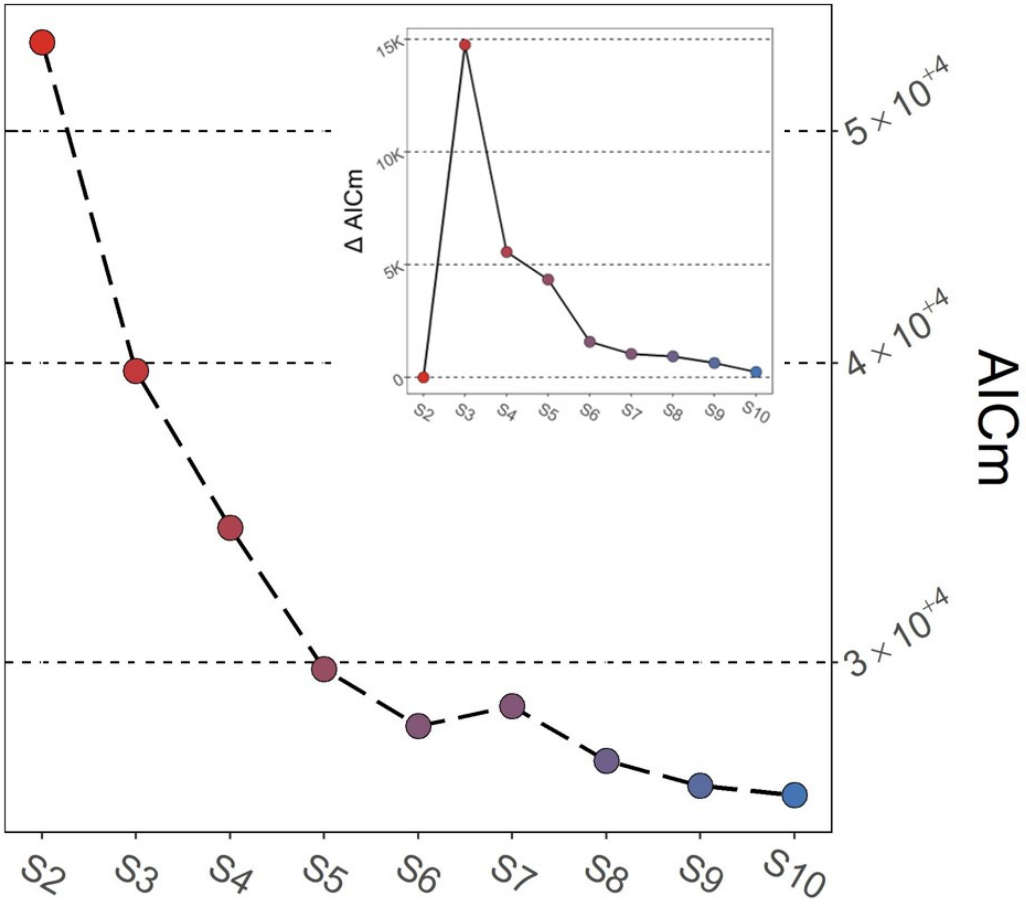
Appendix 35: The number of fixed difference SNPs hierarchically across the clade. Branch colours denote the number of fixed differences between the daughter lineage and all other lineages.



Appendix 36: Summary of results from IQ-TREE2 and ASTRAL phylogenetic analyses. The phylogenetic trees in A and B represent the ML tree estimated by IQ-TREE2 with the alignment partitioned per ddRAD locus, with node labels showing bootstrap support. Node point colours denote gene (gCF) and site (sCF) concordance factors in A and B, respectively. The relationship of partitioned bootstrap support, gCF and sCF is demonstrated in C. The species summary tree estimated by ASTRAL-III is shown in D, with nodes labelled according to local posterior probability that the branch is representative of the true species tree. For all phylogenetic plots, populations and branches are coloured according to putative species.



Appendix 37: AICM values for species delimitation scenarios using SNAPP. Colours denote the number of species in the given model, ranging from two (red) to nine (blue). Inset demonstrates the increase in AICM with increasing number of species included, with the greatest increase in AICM from two (S2) to three (S3) species.



Appendix 38: Breakdown of the number of ddRAD loci and enrichment GO terms containing fixed difference SNPs between putative species. “Number of loci” refers to ddRAD loci containing fixed differences per putative species or the total background set of all ddRAD loci. All loci were aligned to the assembled southern pygmy perch genome, and annotated using the UniProt database. MF = Molecular Function. BP = Biological Processes. CC = Cellular Components.

Species	Background	*N. pygmaea*	*N. vittata* [A]	*N. vittata* [B]
No. loci	19,600	1,466	968	3,847
No. loci aligned to genome	16,246	1,172	739	2,929
No. loci annotated	7,956	463	252	1,039
No. sig. enriched MF functions	-	2	1	3
No. sig. enriched BP functions	-	10	0	9
No. sig. enriched CC functions	-	0	0	0

Appendix 39: Significantly enriched GO terms for ddRAD loci containing fixed difference between each putative species and the other lineages. Significance was tested based on a background dataset of 19,600 ddRAD loci and a Fisher's exact test and a minimum node size of five.

Species	Ontology category	GO ID	Ontology term	Sig. genes	*p*
N. pygmaea	Molecular Function	GO:0004180	Carboxypeptidase activity	4	4.50×10^{-3}
		GO:0003678	DNA helicase activity	5	7.10×10^{-3}
	Biological Process	GO:0097435	Supramolecular fibre organization	11	2.10×10^{-3}
		GO:0044262	Cellular carbohydrate metabolic process	3	3.00×10^{-3}
		GO:0062012	Regulation of small molecule metabolic process	3	3.00×10^{-3}
		GO:0051301	Cell division	7	4.60×10^{-3}
		GO:0007015	Actin filament organization	9	5.10×10^{-3}
		GO:0000281	Mitotic cytokinesis	3	5.80×10^{-3}
		GO:0061640	Cytoskeleton-dependent cytokinesis	3	5.80×10^{-3}
		GO:0072009	Nephron epithelium development	3	5.80×10^{-3}
		GO:0072073	Kidney epithelium development	3	5.80×10^{-3}
		GO:0051128	Regulation of cellular component organization	12	5.80×10^{-3}
N. vittata [A]	Molecular Function	GO:0005515	Protein binding	42	8.80×10^{-3}
N. vittata [B]	Molecular Function	GO:0033218	Amide binding	6	1.00×10^{-3}
		GO:0019842	Vitamin binding	7	8.40×10^{-3}
		GO:0042803	Protein homodimerization activity	4	9.10×10^{-3}
	Biological Process	GO:0003015	Heart process	6	6.30×10^{-4}
		GO:0060047	Heart contraction	6	6.30×10^{-4}
		GO:0000819	Sister chromatid segregation	5	4.90×10^{-3}
		GO:0003013	Circulatory system process	6	5.36×10^{-3}
		GO:0008015	Blood circulation	6	5.36×10^{-3}
		GO:0044283	Small molecule biosynthetic process	12	8.76×10^{-3}

		GO:0000070	Mitotic sister chromatid segregation	4	9.26×10^{-3}
		GO:0008016	Regulation of heart contraction	3	9.34×10^{-3}
		GO:0060359	Response to ammonium ion	3	9.34×10^{-3}

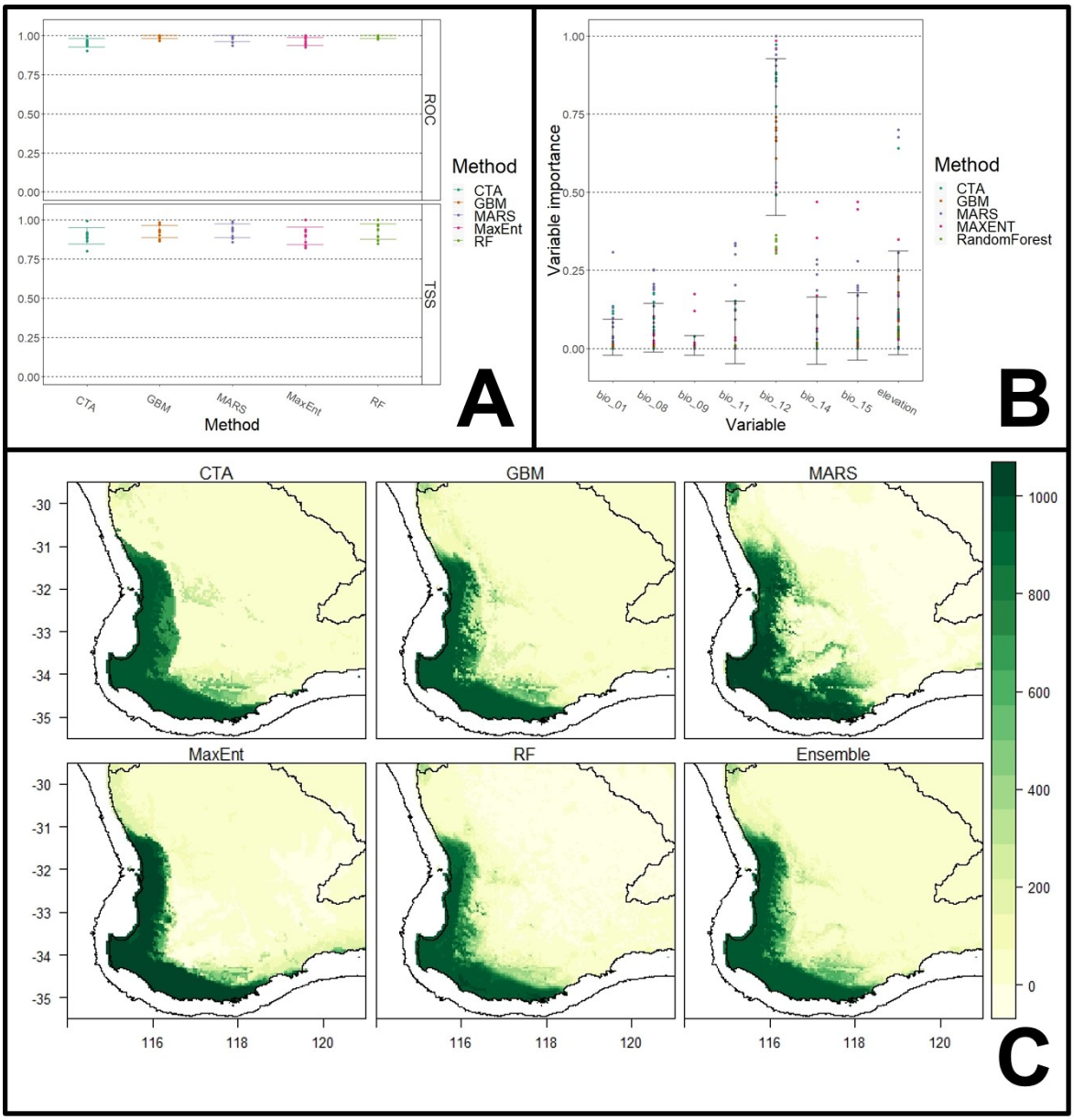
Appendix 40: List of unique genes within enriched GO terms, based on ddRAD loci containing fixed differences. Each gene is only listed once, with all associated GO terms (if more than one was significantly enriched) listed next to the gene.

Species	GO terms	Gene protein	Gene name
N. pygmaea	GO:0003678	A0A3B3DHB7_ORYME	Chromodomain-helicase-DNA-binding protein 8
	GO:0003678	A0A3P9GY40_ORYLA	DNA helicase
	GO:0003678	A0A3B3E1G7_ORYME	General transcription factor IIF subunit 2
	GO:0003678	A0A3Q1EHB4_9TELE	RuvB-like helicase
	GO:0004180	A0A1S3NAK0_SALSA	Angiotensin-converting enzyme
	GO:0004180	A0A3Q4HA80_NEOBR	Carboxypeptidase X, M14 family member 2
	GO:0004180	A0A3N0YUR9_ANAGA	Lysosomal Pro-X carboxypeptidase
	GO:0004180	A0A4W5MRD5_9TELE	Uncharacterized protein
	GO:0044262	A0A315V7B0_GAMAF	Phosphorylase b kinase regulatory subunit
	GO:0044262, GO:0062012	A0A3Q3M836_9TELE	Glucagon receptor b
	GO:0051128	A0A4W5LXA7_9TELE	Palmdelphin
	GO:0051128	H3BYY3_TETNG	RGM_C domain-containing protein
	GO:0051128	A0A4W5QEH8_9TELE	Zgc:56235
	GO:0051301	A0A3B5AK22_9TELE	Cell division control protein
	GO:0051301	A0A3B3YGP9_9TELE	Multifunctional fusion protein
	GO:0051301	A0A3P9I1I6_ORYLA	SKA2 domain-containing protein
	GO:0051301	A0A1S3REU7_SALSA	uncharacterized protein LOC106602330 isoform X2
	GO:0051301, GO:0000281, GO:0061460	E6ZHX6_DICLA	Leucine zipper putative tumor suppressor 2 homolog
	GO:0051301, GO:0061640, GO:0000281	A0A1A8BDA4_9TELE	Zinc finger, FYVE domain containing 26
	GO:0062012	A0A4W6CG56_LATCA	Electron transfer flavoprotein subunit alpha
	GO:0062012	A0A3B5A5C6_9TELE	Sodium channel protein
	GO:0072009, GO:0072073	A0A3Q3E067_HIPCM	Intraflagellar transport 74
	GO:0072009, GO:0072073	A0A3B1JSZ2_ASTMX	LIM homeobox transcription factor 1, beta b
	GO:0072009, GO:0072073, GO:0051128	A0A3Q3WIN1_MOLML	Crumbs cell polarity complex component 2b
	GO:0097435	A0A3Q3MSV2_9TELE	Actinin, alpha 2b
	GO:0097435	A0A3Q3AA25_KRYMA	Myosin XVIIIAa
	GO:0097435, GO:0007015	A0A3Q3K6N9_MONAL	Formin homology 2 domain containing 3a
	GO:0097435, GO:0007015	A0A3B3UHD3_9TELE	Kaptin (actin binding protein)
	GO:0097435, GO:0007015, GO:0051128	A0A4V6ARR2_COLLU	Calpain-2 catalytic subunit
	GO:0097435, GO:0007015, GO:0051128	A0A3Q2DJ26_CYPVA	Dishevelled associated activator of morphogenesis 1b

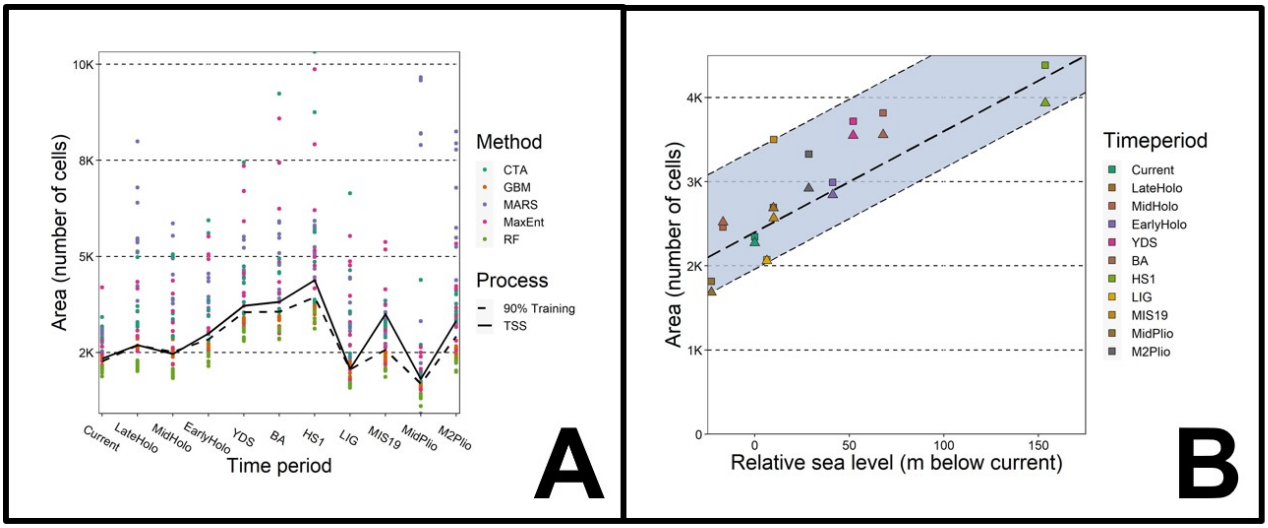
	GO:0097435, GO:00007015, GO:0051128	A0A3Q1H312_ANATE	Epidermal growth factor receptor pathway substrate 8a
	GO:0097435, GO:00007015, GO:0051128	A0A3P9HJN3_ORYLA	Leiomodin 3
	GO:0097435, GO:0007015, GO:0051128	Q52WX5_DANRE	Ras-like protein Rhobtb2a
	GO:0097435, GO:0007015, GO:0051128	A0A315VUH7_GAMAF	Spectrin beta chain
	GO:0097435, GO:0007015, GO:0051128	A0A0S7HHN3_9TELE	TWF2
N. vittata [A]	GO:0005515	A0A3Q1F0D0_9TELE	Actin binding Rho activating protein a
	GO:0005515	A0A3Q3NCX6_9TELE	ADP-ribosylation factor guanine nucleotide-exchange factor 1 (brefeldin A-inhibited)
	GO:0005515	A0A1A7ZNR0_NOTFU	Ankyrin repeat and BTB (POZ) domain containing 2
	GO:0005515	A0A3Q1K087_ANATE	Ankyrin repeat and BTB (POZ) domain containing 2a
	GO:0005515	A0A3Q1FNL0_9TELE	Anoctamin
	GO:0005515	A0A1A8DUH4_9TELE	Caldesmon 1 like
	GO:0005515	F1Q7X5_DANRE	Centromere protein F
	GO:0005515	A0A2D0Q5R2_ICTPU	CREB-regulated transcription coactivator 3 isoform X3
	GO:0005515	A0A0F8CNI0_LARCR	CRF domain-containing protein
	GO:0005515	A0A3Q2P1E8_FUNHE	Dedicator of cytokinesis 10
	GO:0005515	A0A0F8CVH6_LARCR	FERM, RhoGEF and pleckstrin domain-containing protein 2
	GO:0005515	A0A1A8I6T8_NOTKU	HtrA serine peptidase 1b
	GO:0005515	A0A3Q2CJL6_CYPVA	Kinesin family member 11
	GO:0005515	A0A4W6D0J8_LATCA	Kinesin motor domain-containing protein
	GO:0005515	A0A1A7XRW9_9TELE	Metastasis suppressor 1-like
	GO:0005515	A0A3B3HF18_ORYLA	Microtubule-associated protein
	GO:0005515	A0A3P9PB42_POERE	Myosin ID
	GO:0005515	A0A0P7YUT3_SCLFO	Myosin-6-like
	GO:0005515	A0A3N0YXD3_ANAGA	Nesprin-2
	GO:0005515	A0A3P9NVQ8_POERE	Phosphatase and actin regulator
	GO:0005515	A0A2U9CRD7_SCOMX	Putative transcription activator BRG1
	GO:0005515	O57396_ONCMY	RtSox23
	GO:0005515	A0A3Q2VM81_HAPBU	SEC7 domain-containing protein
	GO:0005515	A0A3B3D820_ORYME	Sema domain-containing protein
	GO:0005515	A0A3Q3WD23_MOLML	Spectrin beta chain
	GO:0005515	A0A0S7HMN7_9TELE	SPTB2
	GO:0005515	A0A1A8B561_NOTFU	T-cell lymphoma invasion and metastasis 2
	GO:0005515	A0A3Q1AH46_AMPOC	TNF receptor-associated factor
	GO:0005515	A0A3P8VZV9_CYNSE	Transportin 1
	GO:0005515	A0A3B3H3P4_ORYLA	t-SNARE coiled-coil homology domain-containing protein
	GO:0005515	A0A0S2MM99_SPAAU	Tyrosine-protein kinase receptor
	GO:0005515	A0A0F8BN41_LARCR	Uncharacterized protein
	GO:0005515	A0A3Q0RRI7_AMPCI	Wu:fi04e12
N. v	GO:0000819	A0A3B3YN51_9TELE	Structural maintenance of chromosomes 5

	GO:0000819, GO:0000070	A0A087Y092_POEFO	Charged multivesicular body protein 1A
	GO:0000819, GO:0000070	Q1ECZ7_DANRE	Hnrnpu protein
	GO:0000819, GO:0000070	A0A3B4EH51_PYGNA	MAU2 sister chromatid cohesion factor
	GO:0000819, GO:0000070	A0A3B4WID6_SERLL	Non-SMC condensin I complex, subunit G
	GO:0003015, GO:0060047, GO:0003013, GO:0008015	M4A8E9_XIPMA	Bridging integrator 1b
	GO:0003015, GO:0060047, GO:0003013, GO:0008015	A0A3Q1BZ38_AMPOC	Myocardial zonula adherens protein
	GO:0003015, GO:0060047, GO:0003013, GO:0008015, GO:0008016	A0A3Q2CS10_CYPVA	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma
	GO:0003015, GO:0060047, GO:0003013, GO:0008015, GO:0044283	A0A3Q2D6H5_CYPVA	Dihydrofolate reductase
	GO:0003015, GO:0060047, GO:0003013, GO:0008015, GO:0008016	A0A2D0SI98_ICTPU	RING finger protein 207 isoform X2
	GO:0019842	A0A0F8CG71_LARCR	Aromatic-L-amino-acid decarboxylase
	GO:0019842	A0A4W5MNY1_9TELE	Aspartate aminotransferase
	GO:0019842	A0A1A7XAM1_9TELE	Leprecan-like 2
	GO:0033218	A0A3P8UIF7_CYNSE	Amyloid beta (A4) precursor protein-binding, family A, member 2b
	GO:0033218	A0A3Q3WYM8_MOLML	Signal recognition particle subunit SRP68
	GO:0033218, GO:0019842, GO:0060359	A0A3B5MT62_9TELE	Coiled-coil domain containing 57
	GO:0033218, GO:0019842, GO:0044283	A0A1S3RIB8_SALSA	Pyruvate carboxylase
	GO:0033218, GO:0019842, GO:0044283	A0A3Q0T4B9_AMPCI	Pyruvate carboxylase b
	GO:0042803	I3KKA4_ORENI	Centromere protein F
	GO:0042803	A0A1A7X8I8_9TELE	Uncharacterized protein
	GO:0042803	A0A3B4G9U4_9CICH	V(D)J recombination-activating protein 1
	GO:0044283	A0A3Q3W7V5_MOLML	Fructose-bisphosphate aldolase
	GO:0044283	A0A3B5LMY6_9TELE	Hydroxysteroid (17-beta) dehydrogenase 7
	GO:0044283	A0A3B1JJD1_ASTMX	Inosine-5'-monophosphate dehydrogenase
	GO:0044283	I3K460_ORENI	Sepiapterin reductase b
	GO:0044283	A0A4W6DNN2_LATCA	Sterile alpha and TIR motif containing 1
	GO:0044283	A0A3Q3XPT5_MOLML	Uridine-cytidine kinase
	GO:0044283, GO:0060359	A0A2R8QPK9_DANRE	Acetyl-CoA carboxylase alpha

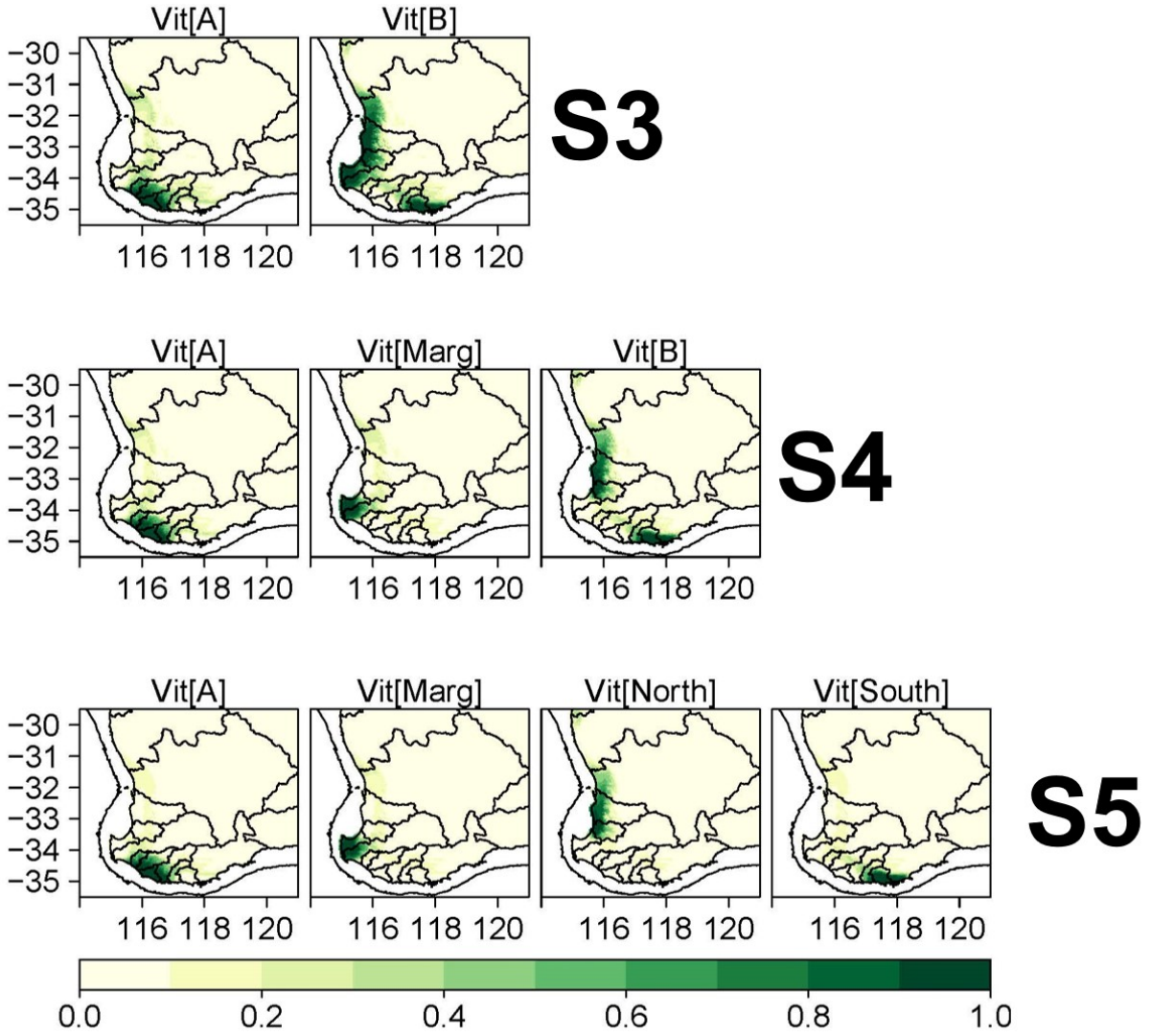
Appendix 41: Evaluations of SDM accuracy. **A**: Model fit per method (*n* = nine per method) using the ROC and the TSS. **B**: Estimates of variable importance across all models (*n* = 45 total) for all variables used in all projections (i.e. excluding bio3). Each point represents a single model, coloured by method, with error bars capturing the 95% confidence interval across all models. **C:** Maps of ensemble distribution models built from averaging contemporary distributions for each method separately (the first five), as well as the final ensemble (bottom right).



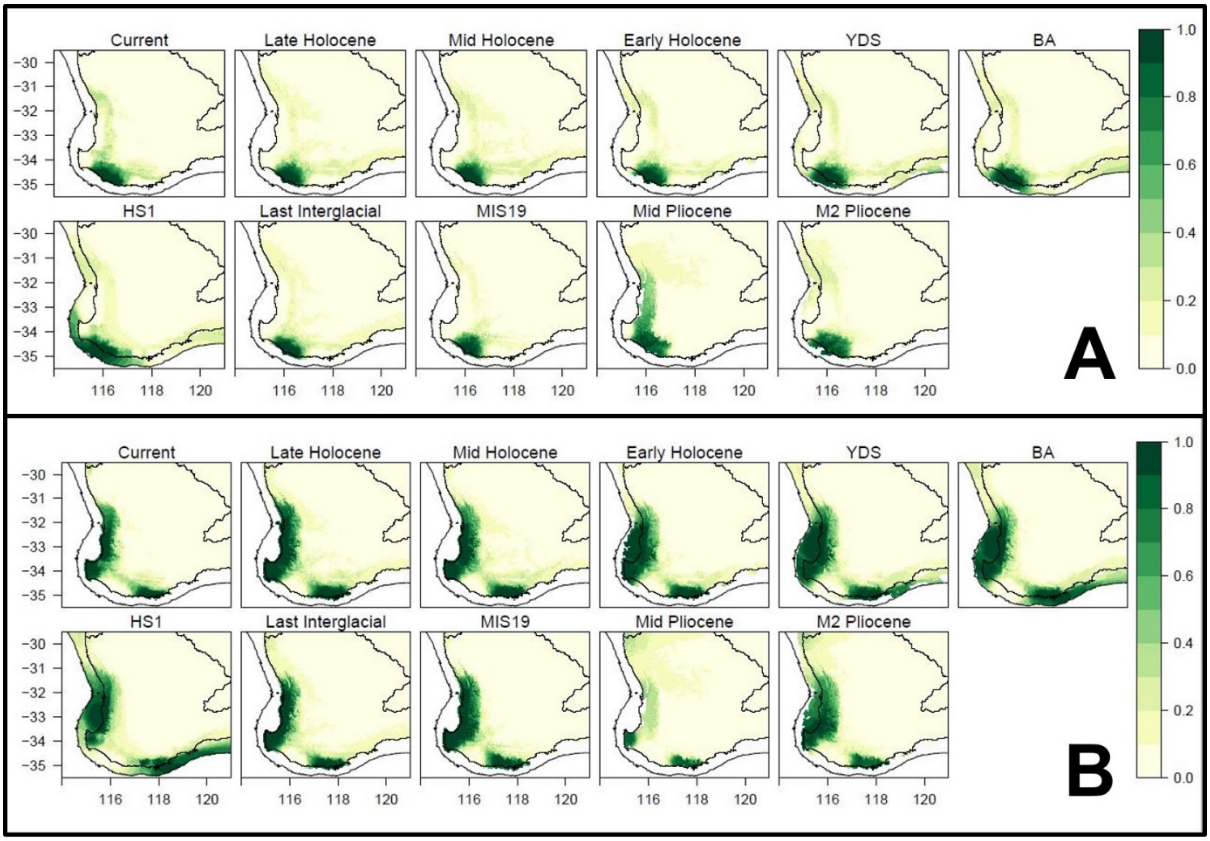
Appendix 42: Estimates of SDM area over time, based on converting individual and ensemble models to binary presence-absence format. **A:** Distribution area over time, with each point representing a single model with colour denoting the method used to generate the distribution. Solid line indicates the area estimated for the full ensemble model using the '90% minimum training presence' as the threshold, whilst the dashed line indicates the area estimated based on the true skill statistic method. All individual model areas were estimating using the 90% minimum training presence approach. **B:** Correlation of binary area of ensemble projections and the relative sea level, with triangles indicating area estimated using the 90% minimum training presence threshold and squares for area estimating using the TSS approach. Points are coloured by the time period of the projection. The dashed line represents the linear line of best fit, with the shaded area including all localities (i.e. min and max linear correlation).



Appendix 43: Contemporary LDMs for *N. vittata*, considering two to four separate lineages (akin to species delimitation model labelled beside each scenario). All LDMs were estimated using the contemporary ensemble SDM and a cost-distance approach.



Appendix 44: LDMs over time of putative cryptic species within *N. vittata*, estimated by subsetting the ensemble SDMs into separate lineages in a cost-distance approach. **A:** *N. vittata* [A] LDMs. **B:** *N. vittata* [B] LDMs.



References

Ackerly DD, Loarie SR, Cornwell WK, Weiss SB, Hamilton H, Branciforte R, Kraft NJB (2010) The geography of climate change: implications for conservation biogeography. *Diversity and Distributions* **16**, 476-487.

Adams M, Raadik TA, Burridge CP, Georges A (2014) Global biodiversity assessment and hyper-cryptic species complexes: more than one species of elephant in the room? *Systematic Biology* **63**, 518-533.

Addo-Bediako A, Chown SL, Gaston KJ (2000) Thermal tolerance, climatic variability and latitude. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**, 739-745.

Aeschbacher S, Selby JP, Willis JH, Coop G (2017) Population-genomic inference of the strength and timing of selection against gene flow. *Proceedings of the National Academy of Sciences* **114**, 7061-7066.

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* **1**, 95-111.

Alexa A, Rahnenfuhrer J (2010) topGO: Enrichment analysis for gene ontology version 2.

Alexa A, Rahnenfuhrer J (2020) topGO: Enrichment analysis for gene ontology 2.40.0.

Allen MG, Morgan DL, Close PG, Beatty SJ (2020) Too little but not too late? Biology of a recently discovered and imperilled freshwater fish in a drying temperate region and comparison with sympatric fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems* **30**, 1412-1423.

Amante C, Eakins B (2009) ETOPO1 1 Arc-Minute Global Relief Model: procedures, data sources and analysis (ed. NGDC-24 NTMN), National Geophysical Data Center, NOAA.

Andrich MA, Imberger J (2013) The effect of land clearing on rainfall and fresh water resources in Western Australia: a multi-functional sustainability analysis. *International Journal of Sustainable Development & World Ecology* **20**, 549-563.

Angert AL, Bradshaw HD, Jr., Schemske DW (2008) Using experimental evolution to investigate geographic range limits in monkeyflowers. *Evolution* **62**, 2660-2675.

Ansari MH, Cooper SJB, Schwarz MP, Ebrahimi M, Dolman G, Reinberger L, Saint KM, Donnellan SC, Bull CM, Gardner MG (2019) Plio-Pleistocene diversification and biogeographic barriers in southern Australia reflected in the phylogeography of a widespread and common lizard species. *Molecular Phylogenetics and Evolution* **133**, 107-119.

Antunes A, Faria R, Johnson WE, Guyomard R, Alexandrino P (2006) Life on the edge: The long-term persistence and contrasting spatial genetic structure of distinct brown trout life histories at their ecological limits. *Journal of Heredity* **97**, 193-205.

Ashcroft MB, Chisholm LA, French KO (2009) Climate change at the landscape scale: predicting fine-grained spatial heterogeneity in warming and potential refugia for vegetation. *Global Change Biology* **15**, 656-667.

Attard CR, Moller LM, Sasaki M, Hammer MP, Bice CM, Brauer CJ, Carvalho DC, Harris JO, Beheregaray LB (2016a) A novel holistic framework for genetic-based captive-breeding and reintroduction programs. *Conservation Biology* **30**, 1060-1069.

Attard CRM, Brauer CJ, Van Zoelen JD, Sasaki M, Hammer MP, Morrison L, Harris JO, Möller LM, Beheregaray LB (2016b) Multi-generational evaluation of genetic diversity and parentage in captive southern pygmy perch (*Nannoperca australis*). *Conservation Genetics* **17**, 1469-1473.

Avise J (2000) *Phylogeography: The History and Formation of Species* Harvard University Press, Cambridge, MA.

Avise JC, Bowen BW, Ayala FJ (2016) In the light of evolution X: Comparative phylogeography. *Proceedings of the National Academy of Sciences* **113**, 7957-7961.

Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution* **29**, 2157-2167.

Balcombe SR, Sheldon F, Capon SJ, Bond NR, Hadwen WL, Marsh N, Bernays SJ (2011) Climate-change threats to native fish in degraded rivers and floodplains of the Murray–Darling Basin, Australia. *Marine and Freshwater Research* **62**, 1099-1114.

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* **3**, 327-338.

Barraclough TG, Vogler AP, Associate Editor: Daniel S (2000) Detecting the geographical pattern of speciation from species-level phylogenies. *The American Naturalist* **155**, 419-434.

Barrow LN, Lemmon AR, Lemmon EM (2018) Targeted sampling and target capture: Assessing phylogeographic concordance with genome-wide data. *Systematic Biology* **67**, 979-996.

Bay RA, Harrigan RJ, Underwood VL, Gibbs HL, Smith TB, Ruegg K (2018) Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* **359**, 83-86.

Beatty S, Morgan D (2013) Introduced freshwater fishes in a global endemic hotspot and implications of habitat and climatic change. *BioInvasions Records* **2**, 1-9.

Beatty S, Morgan D (2019) *Nannoperca pygmaea. The IUCN Red List of Threatened Species.*

Beatty SJ, Morgan DL, Rashnavadi M, Lymbery AJ (2011) Salinity tolerances of endemic freshwater fishes of south-western Australia: implications for conservation in a biodiversity hotspot. *Marine and Freshwater Research* **62**, 91-100.

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF (2018) Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data* **5**, 180214.

Beheregaray LB (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology* **17**, 3754-3774.

Beheregaray LB, Attard CR, Brauer CJ, Whiterod NS, Wedderburn SD, Hammer MP (in press) Conservation breeding and reintroduction of pygmy perches in the lower Murray-Darling Basin, Australia: two similar species, two contrasting outcomes. In: *IUCN Global Reintroduction Perspectives*.

Beheregaray LB, Caccone A (2007) Cryptic biodiversity in a changing world. *Journal of Biology* **6**, 9.

Beheregaray LB, Pfeiffer LV, Attard CRM, Sandoval-Castillo J, Domingos FMCB, Faulks LK, Gilligan DM, Unmack PJ (2017) Genome-wide data delimits multiple climate-determined species ranges in a widespread Australian fish, the golden perch (*Macquaria ambigua*). *Molecular Phylogenetics and Evolution* **111**, 65-75.

Beheregaray LB, Sunnucks P, Briscoe DA (2002) A rapid fish radiation associated with the last sea-level changes in southern Brazil: the silverside *Odontesthes perugiae* complex. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **269**, 65-73.

Bell RC, Parra JL, Badjedjea G, Barej MF, Blackburn DC, Burger M, Channing A, Dehling JM, Greenbaum E, Gvozdik V, Kielgast J, Kusamba C, Lotters S, McLaughlin PJ, Nagy ZT, Rodel MO, Portik DM, Stuart BL, VanDerWal J, Zassi-Boulou AG, Zamudio KR (2017) Idiosyncratic responses to climate-driven forest fragmentation and marine incursions in reed frogs from Central Africa and the Gulf of Guinea Islands. *Molecular Ecology* **26**, 5223-5244.

Benbow MC (1990) Tertiary coastal dunes of the Eucla Basin, Australia. *Geomorphology* **3**, 9-29.

Bickford D, Lohman DJ, Sodhi NS, Ng PK, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution* **22**, 148-155.

Binks RM, Gibson N, Ottewell KM, Macdonald B, Byrne M (2019) Predicting contemporary range-wide genomic variation using climatic, phylogeographic and morphological knowledge in an ancient, unglaciated landscape. *Journal of Biogeography* **46**, 503-514.

Blom WM, Alsop DB (1988) Carbonate mud sedimentation on a temperate shelf: Bass Basin, southeastern Australia. *Sedimentary Geology* **60**, 269-280.

Blöschl G, Hall J, Viglione A, Perdigão RAP, Parajka J, Merz B, Lun D, Arheimer B, Aronica GT, Bilibashi A, Boháč M, Bonacci O, Borga M, Čanjevac I, Castellarin A, Chirico GB, Claps P, Frolova N, Ganora D, Gorbachova L, Gül A, Hannaford J, Harrigan S, Kireeva M, Kiss A, Kjeldsen TR, Kohnová S, Koskela JJ, Ledvinka O, Macdonald N, Mavrova-Guirguinova M, Mediero L, Merz R, Molnar P, Montanari A, Murphy C, Osuch M, Ovcharuk V, Radevski I, Salinas JL, Sauquet E, Šraj M, Szolgay J, Volpi E, Wilson D, Zaimi K, Živković N (2019) Changing climate both increases and decreases European river floods. *Nature* **573**, 108 - 111.

Blum JA, Bonaccorsi S, Marzullo M, Palumbo V, Yamashita YM, Barbash DA, Gatti M (2017) The hybrid incompatibility genes LHR and HMR are required for sister chromatid detachment during anaphase but not for centromere function. *Genetics* **207**, 1457-1472.

Bonada N, DolÉDec S, Statzner B (2007) Taxonomic and biological trait differences of stream macroinvertebrate communities between mediterranean and temperate regions: implications for future climatic scenarios. *Global Change Biology* **13**, 1658-1671.

Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, du Plessis L, Popinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu C-H, Xie D, Zhang C, Stadler T, Drummond AJ (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology* **15**, e1006650.

Bouzat JL (2010) Conservation genetics of population bottlenecks: the role of chance, selection, and history. *Conservation Genetics* **11**, 463-478.

Bradshaw WE, Holzapfel CM (2006) Evolutionary response to rapid climate change. *Science* **312**, 1477-1478.

Brauer CJ, Hammer MP, Beheregaray LB (2016) Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. *Molecular Ecology* **25**, 5093-5113.

Brauer CJ, Unmack PJ, Beheregaray LB (2017) Comparative ecological transcriptomics and the contribution of gene expression to the evolutionary potential of a threatened fish. *Molecular Ecology* **26**, 6841-6856.

Brauer CJ, Unmack PJ, Hammer MP, Adams M, Beheregaray LB (2013) Catchment-scale conservation units identified for the threatened Yarra pygmy perch (*Nannoperca obscura*) in highly modified river systems. *PLoS One* **8**, e82953.

Brauer CJ, Unmack PJ, Smith S, Bernatchez L, Beheregaray LB (2018) On the roles of landscape heterogeneity and environmental variation in determining population genomic structure in a dendritic system. *Molecular Ecology* **27**, 3484-3497.

Bravo GA, Antonelli A, Bacon CD, Bartoszek K, Blom MPK, Huynh S, Jones G, Knowles LL, Lamichhaney S, Marcussen T, Morlon H, Nakhleh LK, Oxelman B, Pfeil B, Schliep A, Wahlberg N, Werneck FP, Wiedenhoeft J, Willows-Munro S, Edwards SV (2019) Embracing heterogeneity: coalescing the Tree of Life and the future of phylogenomics. *PeerJ* **7**, e6399.

Bridle JR, Polechova J, Kawata M, Butlin RK (2010) Why is adaptation prevented at ecological margins? New insights from individual-based simulations. *Ecology Letters* **13**, 485-494.

Bridle JR, Vines TH (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution* **22**, 140-147.

Brito PH, Edwards SV (2009) Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* **135**, 439-455.

Brooks TM, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, Rylands AB, Konstant WR, Flick P, Pilgrim J, Oldfield S, Magin G, Hilton-Taylor C (2002) Habitat loss and extinction in the hotspots of biodiversity. *Conservation Biology* **16**, 909-923.

Brouwers NC, Mercer J, Lyons T, Poot P, Veneklaas E, Hardy G (2012) Climate and landscape drivers of tree decline in a Mediterranean ecoregion. *Ecology and Evolution* **3**, 67-79.

Brown JL, Bennett JR, French CM (2017) SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ* **5**, e4095.

Brown JL, Hill DJ, Dolan AM, Carnaval AC, Haywood AM (2018) PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Scientific Data* **5**, 180254.

Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology & Evolution* **18**, 249-256.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution* **29**, 1917-1932.

Buckley SJ, Domingos FMCB, Attard C, Brauer CJ, Sandoval-Castillo J, Lodge R, Unmack P, Beheregaray LB (2018) Phylogenomic history of enigmatic pygmy perches: implications for biogeography, taxonomy and conservation. *Royal Society Open Science* **5**, e172125.

Buisson L, Thuiller W, Casajus N, Lek S, Grenouillet G (2010) Uncertainty in ensemble forecasting of species distribution. *Global Change Biology* **16**, 1145-1157.

Burke KD, Williams JW, Chandler MA, Haywood AM, Lunt DJ, Otto-Bliesner BL (2018) Pliocene and Eocene provide best analogs for near-future climates. *Proceedings of the National Academy of Sciences* **115**, 13288-13293.

Burridge CP, Craw D, Waters JM (2007) An empirical test of freshwater vicariance via river capture. *Molecular Ecology* **16**, 1883-1895.

Burridge CP, Waters JM (2020) Does migration promote or inhibit diversification? A case study involving the dominant radiation of temperate Southern Hemisphere freshwater fishes. *Evolution*.

Butlin RK (2005) Recombination and speciation. *Molecular Ecology* **14**, 2621-2635.

Byrne M (2008) Evidence for multiple refugia at different time scales during Pleistocene climatic oscillations in southern Australia inferred from phylogeography. *Quaternary Science Reviews* **27**, 2576-2585.

Byrne M, Steane DA, Joseph L, Yeates DK, Jordan GJ, Crayn D, Aplin K, Cantrill DJ, Cook LG, Crisp MD, Keogh JS, Melville J, Moritz C, Porch N, Sniderman JMK, Sunnucks P, Weston PH (2011) Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. *Journal of Biogeography* **38**, 1635-1656.

Byrne M, Yeates DK, Joseph L, Kearney M, Bowler J, Williams MA, Cooper S, Donnellan SC, Keogh JS, Leys R, Melville J, Murphy DJ, Porch N, Wyrwoll KH (2008) Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular Ecology* **17**, 4398-4417.

Cai W, Cowan T (2008) Evidence of impacts from rising temperature on inflows to the Murray-Darling Basin. *Geophysical Research Letters* **35**, L07701.

Cardillo M, Pratt R (2013) Evolution of a hotspot genus: geographic variation in speciation and extinction rates in *Banksia* (Proteaceae). *BMC Evolutionary Biology* **13**, 155.

Carnaval AC, Waltari E, Rodrigues MT, Rosauer D, VanDerWal J, Damasceno R, Prates I, Strangas M, Spanos Z, Rivera D, Pie MR, Firkowski CR, Bornschein MR, Ribeiro LF, Moritz C (2014) Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings of the Royal Society B: Biological Sciences* **281**, e20141461.

Carstens B, Lemmon AR, Lemmon EM (2012) The promises and pitfalls of next-generation sequencing data in phylogeography. *Systematic Biology* **61**, 713-715.

Carstens BC, Pelletier TA, Reid NM, Satler JD (2013) How to fail at species delimitation. *Molecular Ecology* **22**, 4369-4383.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology* **22**, 3124-3140.

Chambers EA, Hillis DM (2020) The multispecies coalescent over-splits species in the case of geographically widespread taxa. *Systematic Biology* **69**, 184-193.

Chan WY, Hoffmann AA, Oppen MJH (2019) Hybridization as a conservation management tool. *Conservation Letters* **12**, e12652.

Chapple DG, Chapple SNJ, Thompson MB (2011) Biogeographic barriers in south-eastern Australia drive phylogeographic divergence in the garden skink, *Lampropholis guichenoti*. *Journal of Biogeography* **38**, 1761-1775.

Chapple DG, Keogh JS, Hutchinson MN (2005) Substantial genetic substructuring in southeastern and alpine Australia revealed by molecular phylogeography of the *Egernia whitii* (Lacertilia: Scincidae) species group. *Molecular Ecology* **14**, 1279-1292.

Chen Y, Shenkar N, Ni P, Lin Y, Li S, Zhan A (2018) Rapid microevolution during recent range expansion to harsh environments. *BMC Evolutionary Biology* **18**, 187.

Christensen J, Hewitson B, Busuioc A (2007) *Regional climate projections. In Climate Change 2007: the physical science basis. Contribution of Working group 1 to the fourth assessment report of the Intergovernmental Panel on Climate Change. University Press, Cambridge*.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **6**, 80-92.

Clarke A, Johnston NM (1999) Scaling of metabolic rate with body mass and temperature in teleost fish. *Journal of Animal Ecology* **68**, 893-905.

Coates DJ, Byrne M, Moritz C (2018) Genetic diversity and conservation units: Dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution* **6**, 165.

Cole TL, Hammer MP, Unmack PJ, Teske PR, Brauer CJ, Adams M, Beheregaray LB (2016) Range-wide fragmentation in a threatened fish associated with post-European settlement modification in the Murray–Darling Basin, Australia. *Conservation Genetics* **17**, 1377-1391.

Coleman RA, Pettigrove V, Raadik TA, Hoffmann AA, Miller AD, Carew ME (2010) Microsatellite markers and mtDNA data indicate two distinct groups in dwarf galaxias, *Galaxiella pusilla* (Mack) (Pisces: Galaxiidae), a threatened freshwater fish from south-eastern Australia. *Conservation Genetics* **11**, 1911-1928.

Colgan DJ (2016) Marine and estuarine phylogeography of the coasts of south-eastern Australia. *Marine and Freshwater Research* **67**, 1597.

Collen B, Whitton F, Dyer EE, Baillie JE, Cumberlidge N, Darwall WR, Pollock C, Richman NI, Soulsby AM, Bohm M (2014) Global patterns of freshwater species diversity, threat and endemism. *Global Ecology and Biogeography* **23**, 40-51.

Cook BD, Baker AM, Page TJ, Grant SC, Fawcett JH, Hurwood DA, Hughes JM (2006) Biogeographic history of an Australian freshwater shrimp, *Paratya australiensis* (Atyidae): the role life history transition in phylogeographic diversification. *Molecular Ecology* **15**, 1083-1093.

Cook LG, Hardy NB, Crisp MD (2015) Three explanations for biodiversity hotspots: small range size, geographical overlap and time for species accumulation. An Australian case study. *New Phytologist* **207**, 390-400.

Cooper SJ, Harvey MS, Saint KM, Main BY (2011) Deep phylogeographic structuring of populations of the trapdoor spider *Moggridgea tingle* (Migidae) from southwestern Australia: evidence for long-term refugia within refugia. *Molecular Ecology* **20**, 3219-3236.

Cooper SJB, Adams M, Labrinidis A (2000) Phylogeography of the Australian dunnart *Sminthopsis crassicaudata* (Marsupialia : Dasyuridae). *Australian Journal of Zoology* **48**, 461-473.

Courchamp F, Hoffmann BD, Russell JC, Leclerc C, Bellard C (2014) Climate change, sea-level rise, and conservation: keeping island biodiversity afloat. *Trends in Ecology & Evolution* **29**, 127-130.

Cowling RM, Lombard AT (2002) Heterogeneity, speciation/extinction history and climate: explaining regional plant diversity patterns in the Cape Floristic Region. *Diversity and Distributions* **8**, 163-179.

Crisp MD, Cook LG (2007) A congruent molecular signature of vicariance across multiple plant lineages. *Molecular Phylogenetics and Evolution* **43**, 1106-1117.

Cros E, Chattopadhyay B, Garg KM, Ng NSR, Tomassi S, Benedick S, Edwards DP, Rheindt FE (2020) Quaternary land bridges have not been universal conduits of gene flow. *Molecular Ecology* **29**, 2692-2706.

Csilléry K, François O, Blum MGB (2012) abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution* **3**, 475-479.

Cutter AD (2013) Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Molecular Phylogenetics and Evolution* **69**, 1172-1185.

Dadour I, Johnson M (1983) Gentic differentiation, hybridization and reproductive isolation in *Mygalopsis marki* Bailey (Orthoptera: Tettigoniidae). *Australian Journal of Zoology* **31**, 353-360.

Dalmaris E, Ramalho CE, Poot P, Veneklaas EJ, Byrne M (2015) A climate change context for the decline of a foundation tree species in south-western Australia: insights from phylogeography and species distribution modelling. *Annals of Botany* **116**, 941-952.

Darwall WRT, Holland RA, Smith KG, Allen D, Brooks EGE, Katarya V, Pollock CM, Shi Y, Clausnitzer V, Cumberlidge N, Cuttelod A, Dijkstra K-DB, Diop MD, García N, Seddon MB, Skelton PH, Snoeks J, Tweddle D, Vié J-C (2011) Implications of bias in conservation research and investment for freshwater species. *Conservation Letters* **4**, 474-482.

Davis CD, Epps CW, Flitcroft RL, Banks MA (2018) Refining and defining riverscape genetics: How rivers influence population genetic structure. *Wiley Interdisciplinary Reviews: Water* **5**, e1269.

Davis J, Pavlova A, Thompson R, Sunnucks P (2013) Evolutionary refugia and ecological refuges: key concepts for conserving Australian arid zone freshwater biodiversity under climate change. *Global Change Biology* **19**, 1970-1984.

Dawson TP, Jackson ST, House JI, Prentice IC, Mace GM (2011) Beyond predictions: Biodiversity conservation in a changing climate. *Science* **332**, 53-58.

Dayrat B (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society* **85**, 407-417.

De Carvalho D, Ingvarsson PK, Joseph J, Suter L, Sedivy C, Macaya-Sanz D, Cottrell J, Heinze B, Schanzer I, Lexer C (2010) Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula* L.), a widespread forest tree. *Molecular Ecology* **19**, 1638-1650.

De Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology* **56**, 879-886.

DeChaine EG, Martin AP (2005) Historical biogeography of two alpine butterflies in the Rocky Mountains: broad-scale concordance and local-scale discordance. *Journal of Biogeography* **32**, 1943-1956.

Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* **24**, 332-340.

Derkarabetian S, Castillo S, Koo PK, Ovchinnikov S, Hedin M (2019) A demonstration of unsupervised machine learning in species delimitation. *Molecular Phylogenetics and Evolution* **139**, 106562.

Deutsch CA, Tewksbury JJ, Huey RB, Sheldon KS, Ghalambor CK, Haak DC, Martin PR (2008) Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences* **105**, 6668-6672.

Dickinson J, Wallace M, Holdgate G, Gallagher S, Thomas L (2002) Origin and timing of the Miocene-Pliocene unconformity in southeast Australia. *Journal of Sedimentary Research* **72**, 288-303.

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**, 20192817.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources* **14**, 209-214.

Dobrynin P, Liu S, Tamazian G, Xiong Z, Yurchenko AA, Krasheninnikova K, Kliver S, Schmidt-Küntzel A, Koepfli K-P, Johnson W, Kuderna LFK, García-Pérez R, Manuel Md, Godinez R, Komissarov A, Makunin A, Brukhin V, Qiu W, Zhou L, Li F, Yi J, Driscoll C, Antunes A, Oleksyk TK, Eizirik E, Perelman P, Roelke M, Wildt D, Diekhans M, Marques-Bonet T, Marker L, Bhak J, Wang J, Zhang G, O'Brien SJ (2015) Genomic legacy of the African cheetah, Acinonyx jubatus. *Genome Biology* **16**, 277.

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27-46.

Drummond AJ, Rambaut A (2019) TreeAnnotator 2.6.

Dubey S, Shine R (2010) Evolutionary diversification of the lizard genus *Bassiana* (Scincidae) across Southern Australia. *PLoS One* **5**, e12982.

Duckett PE, Stow AJ, Burridge C (2013) Higher genetic diversity is associated with stable water refugia for a gecko with a wide distribution in arid Australia. *Diversity and Distributions* **19**, 1072-1083.

Dupuis JR, Roe AD, Sperling FAH (2012) Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology* **21**, 4422-4436.

Dynesius M, Jansson R (2000) Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences* **97**, 9115-9120.

Eaton DA (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* **30**, 1844-1849.

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology* **17**, 1170-1188.

Edwards DL, Dale Roberts J, Scott Keogh J (2008) Climatic fluctuations shape the phylogeography of a mesic direct-developing frog from the south-western Australian biodiversity hotspot. *Journal of Biogeography* **35**, 1803-1815.

Edwards DL, Knowles LL (2014) Species detection and individual assignment in species delimitation: can integrative data increase efficacy? *Proceedings of the Royal Society B: Biological Sciences* **281**, 20132765.

Edwards SV, Kingan SB, Calkins JD, Balakrishnan CN, Jennings WB, Swanson WJ, Sorenson MD (2005) Speciation in birds: genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences* **102 Suppl 1**, 6550-6557.

Edwards SV, Potter S, Schmitt CJ, Bragg JG, Moritz C (2016) Reticulation, divergence, and the phylogeography-phylogenetics continuum. *Proceedings of the National Academy of Sciences* **113**, 8025-8032.

Eizaguirre C, Baltazar-Soares M (2014) Evolutionary conservation—evaluating the adaptive potential of species. *Evolutionary Applications* **7**, 963-967.

Eliason EJ, Clark TD, Hague MJ, Hanson LM, Gallagher ZS, Jeffries KM, Gale MK, Patterson DA, Hinch SG, Farrell AP (2011) Differences in thermal tolerance among sockeye salmon populations. *Science* **332**, 109-112.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17**, 43-57.

Excoffier L, Dupanloup I, Huerta-Sanchez E, Sousa VC, Foll M (2013) Robust demographic inference from genomic and SNP data. *PLOS Genetics* **9**, e1003905.

Fahey AL, Ricklefs RE, Dewoody JA (2014) DNA-based approaches for evaluating historical demography in terrestrial vertebrates. *Biological Journal of the Linnean Society* **112**, 367-386.

Falkenmark M (2013) Adapting to climate change: towards societal water security in dry-climate countries. *International Journal of Water Resources Development* **29**, 123-136.

Faulks LK, Gilligan DM, Beheregaray LB (2010) Clarifying an ambiguous evolutionary history: range-wide phylogeography of an Australian freshwater fish, the golden perch (*Macquaria ambigua*). *Journal of Biogeography* **37**, 1329-1340.

Fenderson LE, Kovach AI, Llamas B (2020) Spatiotemporal landscape genetics: Investigating ecology and evolution through space and time. *Molecular Ecology* **29**, 218-246.

Ferree PM, Prasad S (2012) How can satellite DNA divergence cause reproductive isolation? Let us count the chromosomal ways. *Genetics Research International* **2012**, 430136.

Fišer C, Robinson CT, Malard F (2018) Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology* **27**, 613-635.

Fjeldsaå J, Ehrlich D, Lambin E, Prins E (1997) Are biodiversity 'hotspots' correlated with current ecoclimatic stability? A pilot study using the NOAA-AVHRR remote sensing data. *Biodiversity & Conservation* **6**, 401-422.

Fontaneto D, Flot J-F, Tang CQ (2015) Guidelines for DNA taxonomy, with a focus on the meiofauna. *Marine Biodiversity* **45**, 433-451.

Fordham DA, Brook BW, Moritz C, Nogues-Bravo D (2014) Better forecasts of range dynamics using genetic data. *Trends in Ecology & Evolution* **29**, 436-443.

Frankham R (2009) *Introduction to conservation genetics*, 2nd ed. edn. Cambridge : Cambridge University Press, Cambridge.

Frankham R (2015) Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology* **24**, 2610-2618.

Frankham R, Ballou JD, Eldridge MDB, Lacy RC, Ralls K, Dudash MR, Fenster CB (2011) Predicting the probability of outbreeding depression. *Conservation Biology* **25**, 465-475.

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution* **27**, 480-488.

Funk DJ, Nosil P, Etges WJ (2006) Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proceedings of the National Academy of Sciences* **103**, 3209-3213.

Funk WC, Lovich RE, Hohenlohe PA, Hofman CA, Morrison SA, Sillett TS, Ghalambor CK, Maldonado JE, Rick TC, Day MD, Polato NR, Fitzpatrick SW, Coonan TJ, Crooks KR, Dillon A, Garcelon DK, King JL, Boser CL, Gould N, Andelt WF (2016) Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox *(Urocyon littoralis*). *Molecular Ecology* **25**, 2176-2194.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution* **27**, 489-496.

Funnekotter AV, Millar M, Krauss SL, Nevill PG (2019) Phylogeographic analyses of *Acacia karina* (Fabaceae) support long term persistence of populations both on and off banded iron formations. *Australian Journal of Botany* **67**, 194.

Galaiduk R, Halford AR, Radford BT, Moore CH, Harvey ES, Midgley G (2017) Regional-scale environmental drivers of highly endemic temperate fish communities located within a climate change hotspot. *Diversity and Distributions* **23**, 1256-1267.

Gallagher SJ, Greenwood DR, Taylor D, Smith AJ, Wallace MW, Holdgate GR (2003) The Pliocene climatic and environmental evolution of southeastern Australia: evidence from the marine and terrestrial realm. *Palaeogeography, Palaeoclimatology, Palaeoecology* **193**, 349-382.

Gallagher SJ, Smith AJ, Jonasson K, Wallace MW, Holdgate GR, Daniels J, Taylor D (2001) The Miocene palaeoenvironmental and palaeoceanographic evolution of the Gippsland Basin, Southeast Australia: a record of Southern Ocean change. *Palaeogeography, Palaeoclimatology, Palaeoecology* **172**, 53-80.

Gamperl AK, Farrell AP (2004) Cardiac plasticity in fishes: environmental influences and intraspecific differences. *Journal of Experimental Biology* **207**, 2539-2550.

Garrick RC, Banusiewicz JD, Burgess S, Hyseni C, Symula RE (2019) Extending phylogeography to account for lineage fusion. *Journal of Biogeography* **46**, 268-278.

Garrick RC, Sands CJ, Rowell DM, Tait NN, Greenslade P, Sunnucks P (2004) Phylogeography recapitulates topography: very fine-scale local endemism of a saproxylic 'giant' springtail at Tallaganda in the Great Dividing Range of south-east Australia. *Molecular Ecology* **13**, 3329-3344.

Gavin DG, Fitzpatrick MC, Gugger PF, Heath KD, Rodriguez-Sanchez F, Dobrowski SZ, Hampe A, Hu FS, Ashcroft MB, Bartlein PJ, Blois JL, Carstens BC, Davis EB, de Lafontaine G, Edwards ME, Fernandez M, Henne PD, Herring EM, Holden ZA, Kong WS, Liu J, Magri D, Matzke NJ, McGlone MS, Saltre F, Stigall AL, Tsai YH, Williams JW (2014) Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist* **204**, 37-54.

Gioia P, Hopper SD (2017) A new phytogeographic map for the Southwest Australian Floristic Region after an exceptional decade of collection and discovery. *Botanical Journal of the Linnean Society* **184**, 1-15.

Giribet G, Sharma PP, Benavides LR, Boyer SL, Clouse RM, De Bivort BL, Dimitrov D, Kawauchi GY, Murienne J, Schwendinger PJ (2012) Evolutionary and biogeographical history of an ancient and global group of arachnids (Arachnida: Opiliones: Cyphophthalmi) with a new taxonomic arrangement. *Biological Journal of the Linnean Society* **105**, 92-130.

Goldberg EE, Roy K, Lande R, Jablonski D (2005) Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *The American Naturalist* **165**, 623-633.

Gonzalez RJ, Mcdonald DG (1992) The relationship between oxygen consumption and ion loss in a freshwater fish. *Journal of Experimental Biology* **163**, 317-332.

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics* **16**, 147-168.

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

Gouws G, Stewart BA (2007) From genetic structure to wetland conservation: a freshwater isopod *Paramphisopus palustris* (Phreatoicidea: Amphisopidae) from the Swan Coastal Plain, Western Australia. *Hydrobiologia* **589**, 249-263.

Gouws G, Stewart BA, Daniels SR (2006) Phylogeographic structure of a freshwater crayfish (Decapoda: Parastacidae: *Cherax preissii*) in south-western Australia. *Marine and Freshwater Research* **57**, 837-848.

Gruber B, Unmack PJ, Berry OF, Georges A (2018) dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources* **18**, 691-699.

Grummer JA, Beheregaray LB, Bernatchez L, Hand BK, Luikart G, Narum SR, Taylor EB (2019) Aquatic landscape genomics and environmental effects on genetic variation. *Trends in Ecology & Evolution* **34**, 641-654.

Grummer JA, Bryson RW, Jr., Reeder TW (2013) Species delimitation using Bayes factors: Simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Systematic Biology* **63**, 119-133.

Gutenkunst R, Hernandez R, Williamson S, Bustamante C (2010) Diffusion Approximations for Demographic Inference: DaDi. *Nature Precedings*.

Gutiérrez-Pesquera LM, Tejedo M, Olalla-Tárraga MÁ, Duarte H, Nicieza A, Solé M (2016) Testing the climate variability hypothesis in thermal tolerance limits of tropical and temperate tadpoles. *Journal of Biogeography* **43**, 1166-1178.

Habel JC, Rasche L, Schneider UA, Engler JO, Schmid E, Rödder D, Meyer ST, Trapp N, Sos del Diego R, Eggermont H, Lens L, Stork NE (2019) Final countdown for biodiversity hotspots. *Conservation Letters* **12**, e12668.

Hammer MP, Adams M, Thacker CE, Johnson JB, Unmack PJ (2019) Comparison of genetic structure in co-occurring freshwater eleotrids (Actinopterygii: Philypnodon) reveals cryptic species, likely translocation and regional conservation hotspots. *Molecular Phylogenetics and Evolution* **139**, 106556.

Hammer MP, Bice CM, Hall A, Frears A, Watt A, Whiterod NS, Beheregaray LB, Harris JO, Zampatti BP (2013) Freshwater fish conservation in the face of critical water shortages in the southern Murray–Darling Basin, Australia. *Marine and Freshwater Research* **64**, 807.

Hammer MP, Unmack PJ, Adams M, Johnson JB, Walker KF (2010) Phylogeographic structure in the threatened Yarra pygmy perch *Nannoperca obscura* (Teleostei: Percichthyidae) has major implications for declining populations. *Conservation Genetics* **11**, 213-223.

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters* **8**, 461-467.

Hao T, Elith J, Guillera-Arroita G, Lahoz-Monfort JJ (2019) A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Diversity and Distributions* **25**, 839-852.

Hansen BB, Pederson ÅØ, Peeters B, Moullec ML, Albon SD, Herfindal I, Sæther BE, Grøtan V, Aanes R (2019) Spatial heterogeneity in climate change effects decouples the long-term dynamics of wild reindeer populations in the high Arctic. *Global Change Biology* **25,** 3656-3668.

Harms D (2018) The origins of diversity in ancient landscapes: Deep phylogeographic structuring in a pseudoscorpion (Pseudotyrannochthoniidae: *Pseudotyrannochthonius*) reflects Plio-Pleistocene climate fluctuations. *Zoologischer Anzeiger* **273**, 112-123.

Harrison S, Noss R (2017) Endemism hotspots are linked to stable climatic refugia. *Annals of Botany* **119**, 207-214.

Harrisson KA, Pavlova A, Gonçalves da Silva A, Rose R, Bull JK, Lancaster ML, Murray N, Quin B, Menkhorst P, Magrath MJL, Sunnucks P (2016) Scope for genetic rescue of an endangered subspecies though re-establishing natural gene flow with another subspecies. *Molecular Ecology* **25**, 1242-1258.

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P (2014) Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications* **7**, 1008-1025.

Hawlitschek O, Hendrich L, Espeland M, Toussaint EF, Genner MJ, Balke M (2012) Pleistocene climate change promoted rapid diversification of aquatic invertebrates in Southeast Australia. *BMC Evolutionary Biology* **12**, 142.

Healy TM, Brennan RS, Whitehead A, Schulte PM (2018) Tolerance traits related to climate change resilience are independent and polygenic. *Global Change Biology* **24**, 5348-5360.

Herbert ER, Boon P, Burgin AJ, Neubauer SC, Franklin RB, Ardón M, Hopfensperger KN, Lamers LPM, Gell P (2015) A global perspective on wetland salinization: ecological consequences of a growing threat to freshwater wetlands. *Ecosphere* **6**, 206.

Hermisson J, Pennings PS (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335-2352.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907-913.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution* **27**, 905-920.

Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* **54**, 291-301.

Hijmans R, E. Cameron S, Parra J, G. Jones P, Jarvis A (2005) Very high resolution interpolated climate surfaces of global land areas. *International Journal of Climatology* **25**, 1965-1978.

Hodges K, Donnellan S, Georges A (2015) Significant genetic structure despite high vagility revealed through mitochondrial phylogeography of an Australian freshwater turtle (*Chelodina longicollis*). *Marine and Freshwater Research* **66**, 1045.

Hoffmann AA, Sgro CM (2011) Climate change and evolutionary adaptation. *Nature* **470**, 479-485.

Holdgate GR, Wallace MW, Gallagher SJ, Smith AJ, Keene JB, Moore D, Shafik S (2003) Plio-Pleistocene tectonics and eustasy in the Gippsland Basin, southeast Australia: Evidence from magnetic imagery and marine geological data. *Australian Journal of Earth Sciences* **50**, 403-426.

Hopper SD, Gioia P (2004) The Southwest Australian Floristic Region: Evolution and conservation of a global hot spot of biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **35**, 623-650.

Hopper SD, Smith RJ, Fay MF, Manning JC, Chase MW (2009) Molecular phylogenetics of Haemodoraceae in the Greater Cape and Southwest Australian Floristic Regions. *Molecular Phylogenetics and Evolution* **51**, 19-30.

Hoskin CJ, Higgie M, McDonald KR, Moritz C (2005) Reinforcement drives rapid allopatric speciation. *Nature* **437**, 1353-1356.

Hughes JM, Schmidt DJ, Finn DS (2009) Genes in streams: Using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience* **59**, 573-583.

Humphries P (1995) Life history, food and habitat of southern pygmy perch, *Nannoperca australis*, in the Macquarie River, Tasmania. *Marine and Freshwater Research* **46**, 1159-1169.

Hurst GD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B: Biological Sciences* **272**, 1525-1534.

Inoue K, Monroe EM, Elderkin CL, Berg DJ (2014) Phylogeographic and population genetic analyses reveal Pleistocene isolation followed by high gene flow in a wide ranging, but endangered, freshwater mussel. *Heredity* **112**, 282-290.

Jackson ND, Carstens BC, Morales AE, O'Meara BC (2017) Species delimitation with gene flow. *Systematic Biology* **66**, 799-812.

Jansson R (2003) Global patterns in endemism explained by past climatic change. *Proceedings of the Royal Society B: Biological Sciences* **270**, 583-590.

Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, Loney F, May B, Milacic M, Rothfels K, Sevilla C, Shamovsky V, Shorser S, Varusai T, Weiser J, Wu G, Stein L, Hermjakob H, D'Eustachio P (2020) The reactome pathway knowledgebase. *Nucleic Acids Research* **48**, D498-d503.

Jeremias G, Barbosa J, Marques SM, Asselman J, Gonçalves FJM, Pereira JL (2018) Synthesizing the role of epigenetics in the response and adaptation of species to climate change in freshwater ecosystems. *Molecular Ecology* **27**, 2790-2806.

Johansson AM, Pettersson ME, Siegel PB, Carlborg O (2010) Genome-wide effects of long-term divergent selection. *PLOS Genetics* **6**, e1001188.

Johnson PLF, Slatkin M (2007) Accounting for bias from sequencing error in population genetic estimates. *Molecular Biology and Evolution* **25**, 199-206.

Johnson WE, Onorato DP, Roelke ME, Land ED, Cunningham M, Belden RC, McBride R, Jansen D, Lotz M, Shindle D, Howard J, Wildt DE, Penfold LM, Hostetler JA, Oli MK, O'Brien SJ (2010) Genetic restoration of the Florida panther. *Science* **329**, 1641-1645.

Joppa LN, Roberts DL, Myers N, Pimm SL (2011) Biodiversity hotspots house most undiscovered plant species. *Proceedings of the National Academy of Sciences* **108**, 13171-13176.

Jordan MA, Snell HL (2008) Historical fragmentation of islands and genetic drift in populations of Galápagos lava lizards (*Microlophus albemarlensis* complex). *Molecular Ecology* **17**, 1224-1237.

Joseph L, Dolman G, Donnellan S, Saint KM, Berg ML, Bennett AT (2008) Where and when does a ring start and end? Testing the ring-species hypothesis in a species complex of Australian parrots. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **275**, 2431-2440.

Joyce EB, Webb JA, Dahlhaus PG, Grimes KG, Hill SM, Kotsonis A, Martin J, Mitchell MM, Nielson JL, Orr ML, Peterson JA, Rosengren NJ, Rowan JN, Rowe RK, Sargeant I, Stone T, Smith BL, White S (2003) “Geomorphology, the evolution of Victorian landscapes.” In: *Geology of Victoria. Geological Society of Australia Special Publication 23* (ed. Birch WD), 533-561. Melbourne, VIC: Victorian Division, Geological Society of Australia.

Kalinowski ST, Meeuwig MH, Narum SR, Taper ML (2008) Stream trees: a statistical method for mapping genetic differences between populations of freshwater organisms to the sections of streams that connect them. *Canadian Journal of Fisheries and Aquatic Sciences* **65**, 2752-2760.

Kawakami T, Butlin RK, Adams M, Saint KM, Paull DJ, Cooper SJ (2009) Re-examination of a proposed case of stasipatric speciation: phylogeography of the Australian morabine grasshoppers (*Vandiemenella viatica* species group). *Molecular Ecology* **18**, 3429-3442.

Keppel G, Van Niel KP, Wardell-Johnson GW, Yates CJ, Byrne M, Mucina L, Schut AGT, Hopper SD, Franklin SE (2012) Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography* **21**, 393-404.

Kishino H, Miyata T, Hasegawa M (1990) Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *Journal of Molecular Evolution* **31**, 151-160.

Klausmeyer KR, Shaw MR (2009) Climate change, habitat loss, protected areas and the climate adaptation potential of species in mediterranean ecosystems worldwide. *PLoS One* **4**, e6392.

Klunzinger MW, Lopes-Lima M, Gomes-dos-Santos A, Froufe E, Lymbery AJ, Kirkendale L (2020) Phylogeographic study of the West Australian freshwater mussel, *Westralunio carteri*, uncovers evolutionarily significant units that raise new conservation concerns. *Hydrobiologia*.

Kreger K, Shaban B, Wapstra E, Burridge CP (2019) Phylogeographic parallelism: concordance of patterns in closely related species illuminates underlying mechanisms in the historically glaciated Tasmanian landscape. *bioRxiv*, 548446.

Krehenwinkel H, Tautz D (2013) Northern range expansion of European populations of the wasp spider Argiope bruennichi is associated with global warming-correlated genetic admixture and population-specific temperature adaptations. *Molecular Ecology* **22**, 2232-2248.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870-1874.

Lafon-Placette C, Kohler C (2015) Epigenetic mechanisms of postzygotic reproductive isolation in plants. *Current Opinion in Plant Biology* **23**, 39-44.

Lai Y-T, Yeung CKL, Omland KE, Pang E-L, Hao Y, Liao B-Y, Cao H-F, Zhang B-W, Yeh C-F, Hung C-M, Hung H-Y, Yang M-Y, Liang W, Hsu Y-C, Yao C-T, Dong L, Lin K, Li S-H (2019) Standing genetic variation as the predominant source for adaptation of a songbird. *Proceedings of the National Academy of Sciences* **116**, 2152-2157.

Laity T, Laffan SW, Gonzalez-Orozco CE, Faith DP, Rosauer DF, Byrne M, Miller JT, Crayn D, Costion C, Moritz CC, Newport K (2015) Phylodiversity to inform conservation policy: An Australian example. *Science of the Total Environment* **534**, 131-143.

Lamb AM, Goncalves da Silva A, Joseph L, Sunnucks P, Pavlova A (2019) Pleistocene-dated biogeographic barriers drove divergence within the Australo-Papuan region in a sex-specific manner: an example in a widespread Australian songbird. *Heredity* **123**, 608-621.

Lambeck K, Chappell J (2001) Sea level change through the last glacial cycle. *Science* **292**, 679-686.

Langerhans RB, Riesch R (2013) Speciation by selection: A framework for understanding ecology's role in speciation. *Current Zoology* **59**, 31-52.

Le Roux JJ, Hui C, Castillo ML, Iriondo JM, Keet JH, Khapugin AA, Medail F, Rejmanek M, Theron G, Yannelli FA, Hirsch H (2019) Recent anthropogenic plant extinctions differ in biodiversity hotspots and coldspots. *Current Biology* **29**, 2912-2918 e2912.

Leaché AD, Fujita MK, Minin VN, Bouckaert RR (2014) Species delimitation using genome-wide SNP data. *Systematic Biology* **63**, 534-542.

Lefébure T, Douady CJ, Gouy M, Gibert J (2006) Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution* **40**, 435-447.

Lima TG (2014) Higher levels of sex chromosome heteromorphism are associated with markedly stronger reproductive isolation. *Nature Communications* **5**, 4743.

Lintermans M, Geyle HM, Beatty S, Brown C, Ebner BC, Freeman R, Hammer MP, Humphreys WF, Kennard MJ, Kern P, Martin K, Morgan DL, Raadik TA, Unmack PJ, Wager R, Woinarski JCZ, Garnett ST (2020) Big trouble for little fish: identifying Australian freshwater fishes in imminent risk of extinction. *Pacific Conservation Biology*, PC19053 .

Liu L, Xi Z, Davis CC (2015) Coalescent methods are robust to the simultaneous effects of long branches and incomplete lineage sorting. *Molecular Biology and Evolution* **32**, 791-805.

Liu X, Fu YX (2015) Exploring population size changes using SNP frequency spectra. *Nature Genetics* **47**, 555-559.

Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell J (2000) Myosin: The actin motor protein. In: *Molecular Cell Biology. 4th edition*. WH Freeman.

Love Stowell SM, Pinzone CA, Martin AP (2017) Overcoming barriers to active interventions for genetic diversity. *Biodiversity and Conservation* **26**, 1753-1765.

MacDonald GM, Bennett KD, Jackson ST, Parducci L, Smith FA, Smol JP, Willis KJ (2008) Impacts of climate change on species, populations and communities: palaeobiogeographical insights and frontiers. *Progress in Physical Geography: Earth and Environment* **32**, 139-172.

Maddison WP (1997) Gene trees in species trees. *Systematic Biology* **46**, 523-536.

Malinsky M (2019) Dsuite - fast D-statistics and related admixture evidence from VCF files. *bioRxiv*, 634477.

Malinsky M, Challis RJ, Tyers AM, Schiffels S, Terai Y, Ngatunga BP, Miska EA, Durbin R, Genner MJ, Turner GF (2015) Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science* **350**, 1493-1498.

Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions* **15**, 59-69.

Martin SH, Davey JW, Jiggins CD (2014) Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Molecular Biology and Evolution* **32**, 244-257.

Massatti R, Knowles LL (2014) Microhabitat differences impact phylogeographic concordance of codistributed species: genomic evidence in montane sedges (*Carex* L.) from the Rocky Mountains. *Evolution* **68**, 2833-2846.

Matute DR, Gavin-Smyth J, Liu G (2014) Variable post-zygotic isolation in *Drosophila melanogaster*/*D. simulans* hybrids. *Journal of Evolutionary Biology* **27**, 1691-1705.

Matute Daniel R, Ortiz-Barrientos D (2014) Speciation: The strength of natural selection driving reinforcement. *Current Biology* **24**, R955-R957.

Matzke N (2013) *Probabilistic historical biogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing*, PhD thesis, University of California, 1-240.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution* **66**, 526-538.

McGlashan D (2001) Genetic evidence for historical continuity between populations of the Australian freshwater fish *Craterocephalus stercusmuscarum* (Atherinidae) east and west of the Great Dividing Range. *Journal of Fish Biology* **59**, 55-67.

McInnes KL, Macadam I, Hubbert G, O'Grady J (2011) An assessment of current and future vulnerability to coastal inundation due to sea-level extremes in Victoria, southeast Australia. *International Journal of Climatology* **33**, 33-47.

McKenzie GM, Kershaw AP (2000) The last glacial cycle from Wyelangta, the Otway region of Victoria, Australia. *Palaeogeography, Palaeoclimatology, Palaeoecology* **155**, 177-193.

McLaren S, Wallace MW (2010) Plio-Pleistocene climate change and the onset of aridity in southeastern Australia. *Global and Planetary Change* **71**, 55-72.

McLaren S, Wallace MW, Gallagher SJ, Miranda JA, Holdgate GR, Gow LJ, Snowball I, Sandgren P (2011) Palaeogeographic, climatic and tectonic change in southeastern Australia: the Late Neogene evolution of the Murray Basin. *Quaternary Science Reviews* **30**, 1086-1111.

McLaren S, Wallace MW, Reynolds T (2012) The Late Pleistocene evolution of palaeo megalake Bungunnia, southeastern Australia: A sedimentary record of fluctuating lake dynamics, climate change and the formation of the modern Murray River. *Palaeogeography, Palaeoclimatology, Palaeoecology* **317-318**, 114-127.

Menon S, Soberón J, Li X, Peterson AT (2010) Preliminary global assessment of terrestrial biodiversity consequences of sea-level rise mediated by climate change. *Biodiversity and Conservation* **19**, 1599-1609.

Middelkoop H, Daamen K, Gellens D, Grabs W, Kwadijk JCJ, Lang H, Parmet BWAH, Schädler B, Schulla J, Wilke K (2001) Impact of climate change on hydrological regimes and water resources management in the Rhine basin. *Climatic Change* **49**, 105-128.

Millar MA, Byrne M, Coates DJ, Roberts JD (2017) Comparative analysis indicates historical persistence and contrasting contemporary structure in sympatric woody perennials of semi-arid south-west Western Australia. *Biological Journal of the Linnean Society* **120**, 771-787.

Minh BQ, Hahn MW, Lanfear R (2020a) New methods to calculate concordance factors for phylogenomic datasets. *Molecular Biology and Evolution* **37**, 2727-2733.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020b) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**, 1530-1534.

Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB, Harrison SP, Hurlbert AH, Knowlton N, Lessios HA, McCain CM, McCune AR, McDade LA, McPeek MA, Near TJ, Price TD,

Ricklefs RE, Roy K, Sax DF, Schluter D, Sobel JM, Turelli M (2007) Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters* **10**, 315-331.

Miura O, Urabe M, Nishimura T, Nakai K, Chiba S (2019) Recent lake expansion triggered the adaptive radiation of freshwater snails in the ancient Lake Biwa. *Evolution Letters* **3**, 43-54.

Monks L, Barrett S, Beecham B, Byrne M, Chant A, Coates D, Cochrane JA, Crawford A, Dillon R, Yates C (2019) Recovery of threatened plant species and their habitats in the biodiversity hotspot of the Southwest Australian Floristic Region. *Plant Diversity* **41**, 59-74.

Morgan DL, Beatty SJ, Adams M (2013) *Nannoperca pygmaea*, a new species of pygmy perch (Teleostei: Percichthyidae) from Western Australia. *Zootaxa* **3637**, 401.

Moritz C (1994) Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology & Evolution* **9**, 373-375.

Moritz C, Patton JL, Schneider CJ, Smith TB (2000) Diversification of rainforest faunas: An integrated molecular approach. *Annual Review of Ecology and Systematics* **31**, 533-563.

Morris MRJ, Bowles E, Allen BE, Jamniczky HA, Rogers SM (2018) Contemporary ancestor? Adaptive divergence from standing genetic variation in Pacific marine threespine stickleback. *BMC Evolutionary Biology* **18**, 113.

Morrongiello JR, Bond NR, Crook DA, Wong BB (2010) Nuptial coloration varies with ambient light environment in a freshwater fish. *Journal of Evolutionary Biology* **23**, 2718-2725.

Morrongiello JR, Bond NR, Crook DA, Wong BBM (2012) Spatial variation in egg size and egg number reflects trade-offs and bet-hedging in a freshwater fish. *Journal of Animal Ecology* **81**, 806-817.

Moussalli A, Moritz C, Williams SE, Carnaval AC (2009) Variable responses of skinks to a common history of rainforest fluctuation: concordance between phylogeography and palaeo-distribution models. *Molecular Ecology* **18**, 483-499.

Murienne J, Daniels SR, Buckley TR, Mayer G, Giribet G (2014) A living fossil tale of Pangaean biogeography. *Proceedings of the Royal Society B: Biological Sciences* **281**, 20132648.

Murphy NP, Austin CM (2004) Phylogeography of the widespread Australian freshwater prawn, *Macrobrachium australiense* (Decapoda, Palaemonidae). *Journal of Biogeography* **31**, 1065-1072.

Murren CJ, Auld JR, Callahan H, Ghalambor CK, Handelsman CA, Heskel MA, Kingsolver JG, Maclean HJ, Masel J, Maughan H, Pfennig DW, Relyea RA, Seiter S, Snell-Rood E, Steiner UK, Schlichting CD

(2015) Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. *Heredity* **115**, 293-301.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* **403**, 853-858.

Nakhleh L (2013) Computational approaches to species phylogeny inference and gene tree reconciliation. *Trends in Ecology & Evolution* **28**, 719-728.

Nash SG (2000) A survey of truncated-Newton methods. *Journal of Computational and Applied Mathematics* **124**, 45-59.

Neal WC, James EA, Bayly MJ (2019) Phylogeography, classification and conservation of pink zieria (*Zieria veronicea*; Rutaceae): influence of changes in climate, geology and sea level in south-eastern Australia. *Plant Systematics and Evolution* **305**, 503-520.

Nei M (2001) Genetic Distance. In: *Encyclopedia of Genetics* (eds. Brenner S, Miller JH), pp. 828-832. Academic Press, New York.

Nevill PG, Bradbury D, Williams A, Tomlinson S, Krauss SL (2014) Genetic and palaeo-climatic evidence for widespread persistence of the coastal tree species *Eucalyptus gomphocephala* (Myrtaceae) during the Last Glacial Maximum. *Annals of Botany* **113**, 55-67.

Nijssen B, O'Donnell GM, Hamlet AF, Lettenmaier DP (2001) Hydrologic sensitivity of global rivers to climate change. *Climatic Change* **50**, 143-175.

Nistelberger H, Gibson N, Macdonald B, Tapper SL, Byrne M (2014) Phylogeographic evidence for two mesic refugia in a biodiversity hotspot. *Heredity (Edinb)* **113**, 454-463.

Norgate M, Chamings J, Pavlova A, Bull JK, Murray ND, Sunnucks P (2009) Mitochondrial DNA indicates late Pleistocene divergence of populations of *Heteronympha merope*, an emerging model in environmental change biology. *PLoS One* **4**, e7950.

Nosil P, Crespi BJ (2006) Ecological divergence promotes the evolution of cryptic reproductive isolation. *Proceedings of the Royal Society B: Biological Sciences* **273**, 991-997.

Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: “Isolation by adaptation” and multiple roles for divergent selection. *Evolution* **62**, 316-336, 321.

Nosil P, Feder JL, Flaxman SM, Gompert Z (2017) Tipping points in the dynamics of speciation. *Nature Ecology and Evolution* **1**, 1.

Nunziata SO, Weisrock DW (2018) Estimation of contemporary effective population size and population declines using RAD sequence data. *Heredity (Edinb)* **120**, 196-207.

Overeem RL, Peucker AJ, Austin CM, Dann P, Burridge CP (2008) Contrasting genetic structuring between colonies of the World's smallest penguin, *Eudyptula minor* (Aves: Spheniscidae). *Conservation Genetics* **9**, 893-905.

Pantaloni D, Clainche CL, Carlier M-F (2001) Mechanism of actin-based motility. *Science* **292**, 1502-1506.

Pante E, Puillandre N, Viricel A, Arnaud-Haond S, Aurelle D, Castelin M, Chenuil A, Destombe C, Forcioli D, Valero M, Viard F, Samadi S (2015) Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Molecular Ecology* **24**, 525-544.

Papadopoulou A, Knowles LL (2016) Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proceedings of the National Academy of Sciences* **113**, 8018-8024.

Papadopoulou A, Knowles LL (2017) Linking micro- and macroevolutionary perspectives to evaluate the role of Quaternary sea-level oscillations in island diversification. *Evolution* **71**, 2901-2917.

Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289-290.

Parent CE, Caccone A, Petren K (2008) Colonization and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **363**, 3347-3361.

Passow CN, Greenway R, Arias-Rodriguez L, Jeyasingh PD, Tobler M (2015) Reduction of energetic demands through modification of body size and routine metabolic rates in extremophile fish. *Physiological and Biochemical Zoology* **88**, 371-383.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D (2012) Ancient admixture in human history. *Genetics* **192**, 1065-1093.

Paulay G, Meyer C (2002) Diversification in the tropical pacific: Comparisons between marine and terrestrial systems and the importance of founder speciation. *Integrative and Comparative Biology* **42**, 922-934.

Pavey SA, Bernatchez L, Aubin-Horth N, Landry CR (2012) What is needed for next-generation ecological and evolutionary genomics? *Trends in Ecology & Evolution* **27**, 673-678.

Pavlova A, Beheregaray LB, Coleman R, Gilligan D, Harrisson KA, Ingram BA, Kearns J, Lamb AM, Lintermans M, Lyon J, Nguyen TTT, Sasaki M, Tonkin Z, Yen JDL, Sunnucks P (2017) Severe

consequences of habitat fragmentation on genetic diversity of an endangered Australian freshwater fish: A call for assisted gene flow. *Evolutionary Applications* **10**, 531-550.

Paz A, Ibanez R, Lips KR, Crawford AJ (2015) Testing the role of ecology and life history in structuring genetic variation across a landscape: a trait-based phylogeographic approach. *Molecular Ecology* **24**, 3723-3737.

Pearson RG, Stanton JC, Shoemaker KT, Aiello-Lammens ME, Ersts PJ, Horning N, Fordham DA, Raxworthy CJ, Ryu HY, McNees J, Akçakaya HR (2014) Life history and spatial traits predict extinction risk due to climate change. *Nature Climate Change* **4**, 217-221.

Pelletier TA, Crisafulli C, Wagner S, Zellmer AJ, Carstens BC (2015) Historical species distribution models predict species limits in western *Plethodon* salamanders. *Systematic Biology* **64**, 909-925.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**, e37135.

Phillips SJ, Anderson RP, Dudik M, Schapire RE, Blair ME (2017) Opening the black box: an open-source release of Maxent. *Ecography* **40**, 887.

Pickrell JK, Pritchard JK (2012) Inference of population splits and mixtures from genome-wide allele frequency data. *PLOS Genetics* **8**, e1002967.

Pimm SL (2009) Climate disruption and biodiversity. *Current Biology* **19**, R595-601.

Pinceel T, Brendonck L, Larmuseau MHD, Vanhove MPM, Timms BV, Vanschoenwinkel B (2013) Environmental change as a driver of diversification in temporary aquatic habitats: does the genetic structure of extant fairy shrimp populations reflect historic aridification? *Freshwater Biology* **58**, 1556-1572.

Pittock J, Finlayson CM (2011) Australia's Murray–Darling Basin: freshwater ecosystem conservation options in an era of climate change. *Marine and Freshwater Research* **62**, 232-243.

Pittock J, Hansen LJ, Abell R (2008) Running dry: Freshwater biodiversity, protected areas and climate change. *Biodiversity* **9**, 30-38.

Polechova J, Barton NH (2015) Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences* **112**, 6401-6406.

Porter-Smith R, Lyne VD, Kloser RJ, Lucieer VL (2012) Catchment-based classification of Australia's continental slope canyons. *Marine Geology* **303-306**, 183-192.

Potter S, Xue AT, Bragg JG, Rosauer DF, Roycroft EJ, Moritz C (2018) Pleistocene climatic changes drive diversification across a tropical savanna. *Molecular Ecology* **27**, 520-532.

Prates I, Xue AT, Brown JL, Alvarado-Serrano DF, Rodrigues MT, Hickerson MJ, Carnaval AC (2016) Inferring responses to climate dynamics from historical demography in neotropical forest lizards. *Proceedings of the National Academy of Sciences* **113**, 7978-7985.

Prentis PJ, Wilson JR, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. *Trends in Plant Science* **13**, 288-294.

Prosdocimi F, de Carvalho DC, de Almeida RN, Beheregaray LB (2012) The complete mitochondrial genome of two recently derived species of the fish genus *Nannoperca* (Perciformes, Percichthyidae). *Molecular Biology Reports* **39**, 2767-2772.

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution* **23**, 564-571.

Rabiee M, Sayyari E, Mirarab S (2019) Multi-allele species reconstruction using ASTRAL. *Molecular Phylogenetics and Evolution* **130**, 286-296.

Rabosky DL, Chang J, Title PO, Cowman PF, Sallan L, Friedman M, Kaschner K, Garilao C, Near TJ, Coll M, Alfaro ME (2018) An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392-395.

Raftery AE, Newton MA, Satagopan JM, Krivitsky PN (2006) Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. *Bayesian Statistics* **8**, 1-45.

Ralls K, Ballou JD, Dudash MR, Eldridge MDB, Fenster CB, Lacy RC, Sunnucks P, Frankham R (2018) Call for a paradigm shift in the genetic management of fragmented populations. *Conservation Letters* **11**, e12412.

Rambaut A, Drummond AJ (2009) Tracer v1.5

Ravinet M, Yoshida K, Shigenobu S, Toyoda A, Fujiyama A, Kitano J (2018) The genomic landscape at a late stage of stickleback speciation: High genomic divergence interspersed by small localized regions of introgression. *PLOS Genetics* **14**, e1007358.

Razgour O, Forester B, Taggart JB, Bekaert M, Juste J, Ibáñez C, Puechmaille SJ, Novella-Fernandez R, Alberdi A, Manel S (2019) Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences* **116**, 10418-10423.

Razgour O, Juste J, Ibanez C, Kiefer A, Rebelo H, Puechmaille SJ, Arlettaz R, Burke T, Dawson DA, Beaumont M, Jones G (2013) The shaping of genetic variation in edge-of-range populations under past and future climate change. *Ecology Letters* **16**, 1258-1266.

Rice AM, Rudh A, Ellegren H, Qvarnstrom A (2011) A guide to the genomics of ecological speciation in natural animal populations. *Ecology Letters* **14**, 9-18.

Riddle BR (2016) Comparative phylogeography clarifies the complexity and problems of continental distribution that drove A. R. Wallace to favor islands. *Proceedings of the National Academy of Sciences* **113**, 7970-7977.

Riddle BR, Dawson MN, Hadly EA, Hafner DJ, Hickerson MJ, Mantooth SJ, Yoder AD (2008) The role of molecular genetics in sculpting the future of integrative biogeography. *Progress in Physical Geography: Earth and Environment* **32**, 173-202.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution* **16**, 351-358.

Rincon-Sandoval M, Betancur RR, Maldonado-Ocampo JA (2019) Comparative phylogeography of trans-Andean freshwater fishes based on genome-wide nuclear and mitochondrial markers. *Molecular Ecology* **28**, 1096-1115.

Rissler LJ (2016) Union of phylogeography and landscape genetics. *Proceedings of the National Academy of Sciences* **113**, 8079-8086.

Rissler LJ, Apodaca JJ (2007) Adding more ecology into species delimitation: Ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). *Systematic Biology* **56**, 924-942.

Rix MG, Edwards DL, Byrne M, Harvey MS, Joseph L, Roberts JD (2015) Biogeography and speciation of terrestrial fauna in the south-western Australian biodiversity hotspot. *Biological Reviews* **90**, 762-793.

Rix MG, Harvey MS (2012) Phylogeny and historical biogeography of ancient assassin spiders (Araneae: Archaeidae) in the Australian mesic zone: evidence for Miocene speciation within Tertiary refugia. *Molecular Phylogenetics and Evolution* **62**, 375-396.

Robinson Z, Bell D, Dhendup T, Luikart G, Whiteley A, Kardos M (2020) Evaluating the outcomes of genetic rescue attempts. *Conservation Biology* **Accepted manuscript**.

Roe AD, Rice AV, Coltman DW, Cooke JE, Sperling FA (2011) Comparative phylogeography, genetic differentiation and contrasting reproductive modes in three fungal symbionts of a multipartite bark beetle symbiosis. *Molecular Ecology* **20**, 584-600.

Rosauer DF, Catullo RA, VanDerWal J, Moussalli A, Moritz C (2015) Lineage range estimation method reveals fine-scale endemism linked to Pleistocene stability in Australian rainforest herpetofauna. *PLoS One* **10**, e0126274.

Rotzoll K, Fletcher CH (2012) Assessment of groundwater inundation as a consequence of sea-level rise. *Nature Climate Change* **3**, 477.

Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sanchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34**, 3299-3302.

Rundle HD, Nagel L, Boughman JW, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**, 306-308.

Rundle HD, Schluter D (1998) Reinforcement of stickleback mate preferences: Sympatry breeds contempt. *Evolution* **52**, 200-208.

Saddlier S, Koehn JD, Hammer MP (2013) Let's not forget the small fishes – conservation of two threatened species of pygmy perch in south-eastern Australia. *Marine and Freshwater Research* **64**, 874.

Sandel B, Arge L, Dalsgaard B, Davies RG, Gaston KJ, Sutherland WJ, Svenning J-C (2011) The influence of Late Quaternary climate-change velocity on species endemism. *Science* **334**, 660-664.

Sanderson MJ (2003) r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**, 301-302.

Sandoval-Castillo J, Gates K, Brauer CJ, Smith S, Bernatchez L, Beheregaray LB (2020) Adaptation of plasticity to projected maximum temperatures and across climatically defined bioregions. *Proceedings of the National Academy of Sciences* **117**, 17112-17121.

Sandoval-Castillo J, Robinson NA, Hart AM, Strain LWS, Beheregaray LB (2018) Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliotis laevigata*), along a longitudinal environmental gradient. *Molecular Ecology* **27**, 1603-1620.

Scheffers BR, Joppa LN, Pimm SL, Laurance WF (2012) What we know and don't know about Earth's missing biodiversity. *Trends in Ecology & Evolution* **27**, 501-510.

Schultz MB, Ierodiaconou DA, Smith SA, Horwitz P, Richardson AM, Crandall KA, Austin CM (2008) Sea-level changes and palaeo-ranges: reconstruction of ancient shorelines and river drainages and the phylogeography of the Australian land crayfish *Engaeus sericatus* Clark (Decapoda: Parastacidae). *Molecular Ecology* **17**, 5291-5314.

Schultz MB, Smith SA, Horwitz P, Richardson AM, Crandall KA, Austin CM (2009) Evolution underground: a molecular phylogenetic investigation of Australian burrowing freshwater crayfish (Decapoda: Parastacidae) with particular focus on *Engaeus* Erichson. *Molecular Phylogenetics and Evolution* **50**, 580-598.

Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, Peichel CL, Saetre GP, Bank C, Brannstrom A, Brelsford A, Clarkson CS, Eroukhmanoff F, Feder JL, Fischer MC, Foote AD, Franchini P, Jiggins CD, Jones FC, Lindholm AK, Lucek K, Maan ME, Marques DA, Martin SH, Matthews B, Meier JI, Most M, Nachman MW, Nonaka E, Rennison DJ, Schwarzer J, Watson ET, Westram AM, Widmer A (2014) Genomics and the origin of species. *Nature Reviews Genetics* **15**, 176-192.

Shaddick K, Gilligan DM, Burridge CP, Jerry DR, Truong K, Beheregaray LB (2011) Historic divergence with contemporary connectivity in a catadromous fish, the estuary perch (*Macquaria colonorum*). *Canadian Journal of Fisheries and Aquatic Sciences* **68**, 304.

Shafer ABA, Cullingham CI, Côté SD, Coltman DW (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology* **19**, 4589-4621.

Shafer ABA, Peart CR, Tusso S, Maayan I, Brelsford A, Wheat CW, Wolf JBW, Gilbert M (2017) Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution* **8**, 907-917.

Siepielski AM, Morrissey MB, Buoro M, Carlson SM, Caruso CM, Clegg SM, Coulson T, DiBattista J, Gotanda KM, Francis CD, Hereford J, Kingsolver JG, Augustine KE, Kruuk LEB, Martin RA, Sheldon BC, Sletvold N, Svensson EI, Wade MJ, MacColl ADC (2017) Precipitation drives global variation in natural selection. *Science* **355**, 959-962.

Singhal S, Moritz C (2013) Reproductive isolation between phylogeographic lineages scales with divergence. *Proceedings of the Royal Society B: Biological Sciences* **280**, 20132246.

Smith BT, Seeholzer GF, Harvey MG, Cuervo AM, Brumfield RT (2017) A latitudinal phylogeographic diversity gradient in birds. *PLOS Biology* **15**, e2001073.

Smith I, Power S (2014) Past and future changes to inflows into Perth (Western Australia) dams. *Journal of Hydrology: Regional Studies* **2**, 84-96.

Somero GN (2010) The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers'. *Journal of Experimental Biology* **213**, 912-920.

Sommer RS, Zachos FE (2009) Fossil evidence and phylogeography of temperate species: 'glacial refugia' and post-glacial recolonization. *Journal of Biogeography* **36**, 2013-2020.

Spooner MI, De Deckker P, Barrows TT, Fifield LK (2011) The behaviour of the Leeuwin Current offshore NW Australia during the last five glacial–interglacial cycles. *Global and Planetary Change* **75**, 119-132.

Stadler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P (2009) The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. *Genetics* **182**, 205-216.

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313.

Stange M, Sanchez-Villagra MR, Salzburger W, Matschiner M (2018) Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (Ariidae) supports Miocene closure of the Panamanian Isthmus. *Systematic Biology* **67**, 681-699.

Stanton DWG, Frandsen P, Waples RK, Heller R, Russo I-RM, Orozco-terWengel PA, Pedersen C-ET, Siegismund HR, Bruford MW (2019) More grist for the mill? Species delimitation in the genomic era and its implications for conservation. *Conservation Genetics* **20**, 101-113.

Stewart JR, Lister AM, Barnes I, Dalen L (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences* **277**, 661-671.

Stobie CS, Oosthuizen CJ, Cunningham MJ, Bloomer P (2018) Exploring the phylogeography of a hexaploid freshwater fish by RAD sequencing. *Ecology and Evolution* **8**, 2326-2342.

Struck TH, Feder JL, Bendiksby M, Birkeland S, Cerca J, Gusarov VI, Kistenich S, Larsson KH, Liow LH, Nowak MD, Stedje B, Bachmann L, Dimitrov D (2018) Finding evolutionary processes hidden in cryptic species. *Trends in Ecology & Evolution* **33**, 153-163.

Sukumaran J, Knowles LL (2017) Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences* **114**, 1607-1612.

Sullivan J, Smith ML, Espíndola A, Ruffley M, Rankin A, Tank D, Carstens B (2019) Integrating life history traits into predictive phylogeography. *Molecular Ecology* **28**, 2062-2073.

Sundaram M, Donoghue MJ, Farjon A, Filer D, Mathews S, Jetz W, Leslie AB (2019) Accumulation over evolutionary time as a major cause of biodiversity hotspots in conifers. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20191887.

Sunnucks P, Hales DF (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Molecular Biology and Evolution* **13**, 510-524.

Swofford DL (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10* Sinauer Associates, Sunderland, Massachusetts.

Szűcs M, Vahsen ML, Melbourne BA, Hoover C, Weiss-Lehman C, Hufbauer RA (2017) Rapid adaptive evolution in novel environments acts as an architect of population range expansion. *Proceedings of the National Academy of Sciences* **114**, 13501-13506.

Thiele S, Adams M, Hammer M, Wedderburn S, Whiterod NS, Unmack PJ, Sasaki M, Beheregaray LB (2020) Range-wide population genetics study informs on conservation translocations and reintroductions for the endangered Murray hardyhead (*Craterocephalus fluviatilis*). *Aquatic Conservation: Marine and Freshwater Ecosystems* **30**, 1959-1974.

Thomae AW, Schade GO, Padeken J, Borath M, Vetter I, Kremmer E, Heun P, Imhof A (2013) A pair of centromeric proteins mediates reproductive isolation in *Drosophila* species. *Developmental Cell* **27**, 412-424.

Thomaz AT, Malabarba LR, Knowles LL (2017) Genomic signatures of paleodrainages in a freshwater fish along the southeastern coast of Brazil: genetic structure reflects past riverine properties. *Heredity* **119**, 287-294.

Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography* **32**, 369-373.

Tiitu V (2001) Cold adaptation suppresses the contractility of both atrial and ventricular muscle of the crucian carp heart. *Journal of Fish Biology* **59**, 141-156.

Title PO, Bemmels JB (2018) ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography* **41**, 291-307.

Travis JMJ, Delgado M, Bocedi G, Baguette M, Bartoń K, Bonte D, Boulangeat I, Hodgson JA, Kubisch A, Penteriani V, Saastamoinen M, Stevens VM, Bullock JM (2013) Dispersal and species' responses to climate change. *Oikos* **122**, 1532-1540.

Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLOS Biology* **3**, e285.

Unmack PJ (2001) Biogeography of Australian freshwater fishes. *Journal of Biogeography* **28**, 1053-1089.

Unmack PJ, Bagley JC, Adams M, Hammer MP, Johnson JB (2012) Molecular phylogeny and phylogeography of the Australian freshwater fish genus *Galaxiella*, with an emphasis on dwarf galaxias (*G. pusilla*). *PLoS One* **7**, e38433.

Unmack PJ, Hammer MP, Adams M, Dowling TE (2011) A phylogenetic analysis of pygmy perches (Teleostei: Percichthyidae) with an assessment of the major historical influences on aquatic biogeography in southern Australia. *Systematic Biology* **60**, 797-812.

Unmack PJ, Hammer MP, Adams M, Johnson JB, Dowling TE (2013) The role of continental shelf width in determining freshwater phylogeographic patterns in south-eastern Australian pygmy perches (Teleostei: Percichthyidae). *Molecular Ecology* **22**, 1683-1699.

Unmack PJ, Sandoval-Castillo J, Hammer MP, Adams M, Raadik TA, Beheregaray LB (2017) Genome-wide SNPs resolve a key conflict between sequence and allozyme data to confirm another threatened candidate species of river blackfishes (Teleostei: Percichthyidae: *Gadopsis*). *Molecular Phylogenetics and Evolution* **109**, 415-420.

Van Belleghem SM, Baquero M, Papa R, Salazar C, McMillan WO, Counterman BA, Jiggins CD, Martin SH (2018) Patterns of Z chromosome divergence among *Heliconius* species highlight the importance of historical demography. *Molecular Ecology* **27**, 3852-3872.

Vercoe P, Lawrence C, de Graaf M (2009) Rapid replacement of the critically endangered hairy marron by the introduced smooth marron (Decapoda, Parastacidae) in the Margaret River (Western Australia). *Crustaceana* **82**, 1469-1476.

Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences* **106 Suppl 1**, 9939-9946.

Vucetich J, Waite T (2003) Spatial patterns of demography and genetic processes across the species' range: Null hypotheses for landscape conservation genetics. *Conservation Genetics* **4**, 639-645.

Waldvogel AM, Feldmeyer B, Rolshausen G, Exposito-Alonso M, Rellstab C, Kofler R, Mock T, Schmid K, Schmitt I, Bataillon T, Savolainen O, Bergland A, Flatt T, Guillaume F, Pfenninger M (2020) Evolutionary genomics can improve prediction of species' responses to climate change. *Evolution Letters* **4**, 4-18.

Waller NL, Gynther IC, Freeman AB, Lavery TH, Leung LK-P (2017) The Bramble Cay melomys *Melomys rubicola* (Rodentia: Muridae): a first mammalian extinction caused by human-induced climate change? *Wildlife Research* **44**, 9-21.

Wallis GP, Cameron-Christie SR, Kennedy HL, Palmer G, Sanders TR, Winter DJ (2017) Interspecific hybridization causes long-term phylogenetic discordance between nuclear and mitochondrial genomes in freshwater fishes. *Molecular Ecology* **26**, 3116-3127.

Wang G-D, Zhang B-L, Zhou W-W, Li Y-X, Jin J-Q, Shao Y, Yang H-c, Liu Y-H, Yan F, Chen H-M, Jin L, Gao F, Zhang Y, Li H, Mao B, Murphy RW, Wake DB, Zhang Y-P, Che J (2018) Selection and environmental adaptation along a path to speciation in the Tibetan frog *Nanorana parkeri*. *Proceedings of the National Academy of Sciences* **115**, E5056-E5065.

Wang IJ (2010) Recognizing the temporal distinctions between landscape genetics and phylogeography. *Molecular Ecology* **19**, 2605-2608.

Waters JM (2008) Marine biogeographical disjunction in temperate Australia: historical landbridge, contemporary currents, or both? *Diversity and Distributions* **14**, 692-700.

Waters JM, Burridge CP, Craw D (2019) The lasting biological signature of Pliocene tectonics: Reviewing the re-routing of Australia's largest river drainage system. *Journal of Biogeography* **46**, 1494-1503.

Waters JM, Burridge CP, Craw D (2020) River capture and freshwater biological evolution: A review of galaxiid fish vicariance. *Diversity* **12**, 216.

Waters JM, Craw D, Youngson JH, Wallis GP (2001) Genes meet geology: Fish phylogeographic pattern reflects ancient, rather than modern, drainage connections. *Evolution* **55**, 1844-1851.

Wedderburn SD, Hammer MP, Bice CM (2012) Shifts in small-bodied fish assemblages resulting from drought-induced water level recession in terminating lakes of the Murray-Darling Basin, Australia. *Hydrobiologia* **691**, 35-46.

Weeks AR, Stoklosa J, Hoffmann AA (2016) Conservation of genetic uniqueness of populations may increase extinction likelihood of endangered species: the case of Australian mammals. *Frontiers in Zoology* **13**, 31.

Wheeler MA, Byrne M (2006) Congruence between phylogeographic patterns in cpDNA variation in Eucalyptus marginata (Myrtaceae) and geomorphology of the Darling Plateau, south-west of Western Australia. *Australian Journal of Botany* **54**, 17.

White ME (2000) *Running down: water in a changing land.* East Roseville, NSW: Kangaroo Press, East Roseville, NSW.

Willi Y, Hoffmann AA (2009) Demographic factors and genetic variation influence population persistence under environmental change. *Journal of Evolutionary Biology* **22**, 124-133.

Williams AN, Ulm S, Sapienza T, Lewis S, Turney CSM (2018) Sea-level change and demography during the last glacial termination and early Holocene across the Australian continent. *Quaternary Science Reviews* **182**, 144-154.

Williams JL, Hufbauer RA, Miller TEX (2019) How evolution modifies the variability of range expansion. *Trends in Ecology & Evolution* **34**, 903-913.

Williams SE, Shoo LP, Isaac JL, Hoffmann AA, Langham G (2008) Towards an integrated framework for assessing the vulnerability of species to climate change. *PLOS Biology* **6**, 2621-2626.

Williams WD (1995) Lake Corangamite, Australia, a permanent saline lake: Conservation and management issues. *Lakes & Reservoirs: Science, Policy and Management for Sustainable Use* **1**, 55-64.

Woodward GMA, Malone B (2002) Patterns of abundance and habitat use by *Nannoperca obscura* (Yarra pygmy perch) and *Nannoperca australis* (southern pygmy perch). *Proceedings of the Royal Society of Victoria* **114**, 61-72.

Xue AT, Hickerson MJ (2015) The aggregate site frequency spectrum for comparative population genomic inference. *Molecular Ecology* **24**, 6223-6240.

Xue AT, Hickerson MJ (2017) multi-dice: r package for comparative population genomic inference under hierarchical co-demographic models of independent single-population size changes. *Molecular Ecology Resources* **17**, e212-e224.

Yang Z, Rannala B (2014) Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution* **31**, 3125-3135.

York KL, Blacket MJ, Appleton BR (2008) The Bassian Isthmus and the major ocean currents of southeast Australia influence the phylogeography and population structure of a southern Australian intertidal barnacle *Catomerus polymerus* (Darwin). *Molecular Ecology* **17**, 1948-1961.

Zamudio KR, Bell RC, Mason NA (2016) Phenotypes in phylogeography: Species' traits, environmental variation, and vertebrate diversification. *Proceedings of the National Academy of Sciences* **113**, 8041-8048.

Zhang C, Rabiee M, Sayyari E, Mirarab S (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153.