COMING, GOING, GONE? POPULATION CONNECTIVITY AND EXTINCTION RISK IN RESTRICTED RANGE CORAL REEF FISHES ON ISOLATED ISLANDS

PhD thesis submitted by

Martin H. van der Meer (BSc Hons)

On

10th May 2013

For the degree of Doctor of Philosophy

School of Marine and Tropical Biology

James Cook University

Townsville, Queensland, Australia

STATEMENT ON THE CONTRIBUTION OF OTHERS

The chapters in this thesis are also manuscripts that have been published, submitted or are in preparation for submission. Several researchers have made contributions to these manuscripts and it is necessary to recognize their contributions.

Chapter 3 is now published as: van der Meer MH, Hobbs J-PA, Jones GP, van Herwerden L (2012) Historic hybridisation between two Australian anemonefish species *Amphiprion* and present−day patterns of connectivity. *Ecology and Evolution,* 2, 1592−1604. M.H.vdM conceived the ideas, helped with funding, collected specimens, analysed the data, led the writing; J-P.A. H helped conceiving ideas, collecting specimens and writing; G.P. Jones helped with funding and writing; L. vH helped with conceiving ideas, funding and writing.

Chapter 4 is now published as: van der Meer MH, Hobbs J-PA, Jones GP, van HerwerdenL (2012) Genetic connectivity among and self-replenishment within island populations of a restricted range subtropical reef fish. *PLoS ONE,* 7(11), e49660. doi:10.1371/journal.pone.0049660. M.H.vdM conceived the ideas, helped with funding, collected specimens, analysed the data, led the writing; J-P.A. H helped conceiving ideas, collecting specimens and writing; G.P. Jones helped with funding and writing; L. vH helped with conceiving ideas, funding and writing.

Chapter 5 is now published as: van der Meer MH, Horne JB, Gardner MG, Hobbs J-PA, Pratchett M, van Herwerden L (2012) Limited demographic gene flow and high self-

replenishment drives peripheral isolation in an endemic coral reef fish. *Ecology and Evolution*, DOI: 10.1002/ece3.584.. M.H.vdM conceived the ideas, helped with funding, collected specimens, analysed the data, led the writing; J.B. H helped with writing; M.G. G helped with writing; J-P.A. H helped conceiving ideas, collecting specimens and writing; M. P helped with funding and writing; L. vH helped with conceiving ideas and writing.

Chapter 6 is a manuscript currently in preperation for submission and is co-authored by M. Berumen, J-P.A. Hobbs and L. van Herwerden. M.H.vdM conceived the ideas, helped with funding, collected specimens, analysed the data, led the writing; M. B helped with funding and writing; J-P.A. H helped conceiving ideas, collecting specimens and writing; L. vH helped with conceiving ideas and writing.

Appendix 1 is now published as: van der Meer, M.H., Gardner, M.G., Hobbs, J-P.A, Jones, G. & van Herwerden, L. (2012) Identification of seventeen microsatellite markers for conservation genetic studies of the endemic anemonefish *Amphiprion mccullochi*. *Conservation Genetic Resources*, 4, 247−250. M.H.vdM conceived the ideas, helped with funding, collected specimens, analysed the data, led the writing; M.G. G helped with funding and writing; J-P.A. H helped conceiving ideas, collecting specimens and writing; L. vH helped with funding, conceiving ideas and writing.

Appendix 2 is now published as: van der Meer MH, Gardner MG, Hobbs J-PA, Pratchett MS, van Herwerden L (2012) Identification of twenty one microsatellite loci for conservation genetic studies of the endemic butterflyfish *Chaetodon tricinctus*. *Conservation Genetic Resources*, 4,

243−246. M.H.vdM conceived the ideas, helped with funding, collected specimens, analysed the data, led the writing; M.G. G helped with funding and writing; J-P.A. H helped conceiving ideas, collecting specimens and writing; M.S. P helped with funding and writing; L. vH helped with funding, conceiving ideas and writing.

Appendix 3 is now published as: van der Meer MH, Gardner MG, Berumen M, Hobbs J-PA, van Herwerden L (2012) Identification of seventeen microsatellite loci for conservation genetic studies of the endemic wrasse *Coris bulbifrons*. *Conservation Genetic Resources*, 4, 243−246. M.H.vdM conceived the ideas, helped with funding, collected specimens, analysed the data, led the writing; M.G. G helped with funding and writing; M. B helped with funding and writing; J-P.A. H helped conceiving ideas, collecting specimens and writing; L. vH helped with writing.

I was supported by an International Postgraduate Award Scholarship including full tuition waiver from James Cook University. Funding for the study came from Internal Research Allocation Award to M. vdM (JCU); Graduate Research Award to M. vdM (JCU); Sessional Award to M. vdM (JCU); Great Barrier Reef Marine Parks Authority Science for Management Award to M. vdM; Australian National Network in Marine Science Award to M. vdM; Griffith University/James Cook University Collaborative Grants Scheme Award to M. vdM, J-P.A H, G.P. J, J. H and L. vH; ARC Centre of Excellence for Coral Reef Studies to G. P, M. P; Envirofund Australia (Natural Heritage Trust) to, Lord Howe Island Board, Lord Howe Island Marine Park and, Australian Department of the Environment and Water Resources.

ACKNOWLEDGEMENTS

I thank all the people that have assisted me over the years, not just with this PhD but with guidance in my life, and I am particularly grateful to the people mentioned below. For those that I have forgotten to mention below, my apologies, but I know that I have thanked you along the way.

Firstly, I want to say a very big thanks to my mother (Hazel van der Meer) who without her continued support, patience and guidance, I would not be where I am today. As a single parent she raised me by herself and gave me every opportunity that arose to explore who I was. Her unconditional love has meant that I have travelled far and wide, been on many adventures (often by myself in the middle of nowhere and reliant on my backpack and the medical equipment therein) and had the freedom to turn into the man I am today. Without her financial support for the first year of my MSc (before being awarded a scholarship and upgrading to a PhD), this would never have happened. Thanks mom. In the hard times, my family has always been there for me, so thank you to Aunty Carol and Uncle Robert, Aunty Brenda and Uncle Fanie, Robyn, Wendy and of course Granny for all the walks in national parks and reserves! Gramp's, I know you would be proud if you were still here, thanks for all the fishing lessons and all days pulled at the beach fishing, swimming and exploring rock pools. Lastly, a special thanks to my partner Ebonney, for the many hours spent pouring over each manuscript correcting all the fine details and all the encouragement along the way.

I am grateful to all the people I have met that encouraged me from a little kid to go outdoors, explore, enjoy wildlife, take photos and spend time, for it is only when you immerse yourself in nature will you unlock its secrets.

For the owners at the scuba diving centres, reptile parks, dairy farms, and wildlife adventure tours I have worked at, your input into my life has been invaluable. Also, all the people I have helped with their own research, ranging from small terrestrial mammals to large marine reptiles, your ideas, approaches and techniques have given me the skills I needed to get ahead and succeed.

I thank two people at JCU in particular, my primary supervisor Lynne van Herwerden and my mentor (and subsequent supervisor) JP Hobbs. Lynne has been instrumental in my progress as a molecular geneticist; she has encouraged me to go beyond my best but most importantly has always had her door open, for both research and personal meetings. It is this quality; to be able to talk openly that makes her one of the best teachers I have met. Her guidance without frustration has been exceptional; especially since I had no genetics background before undertaking this PhD. JP has been tireless in his pursuit for me to excel in everything from fieldwork, presentations, grants and scientific writing. His attention to detail and teaching in the field has far exceeded my expectations. He has challenged and guided my thoughts and opened up the way I view nature. I extend my thanks to my two secondary supervisors, GP Jones and M Pratchett who, without which, I would not have had the funding to get to these remote locations. Their help in obtaining permits, funding and bringing their knowledge and experience into editing chapters has been superb. There are also members of the MEEL laboratory I would like to thank for assistance

either in the lab or with analytical software, B Danastas, H Veilleux, R Huerlimann, P Cowman, J Hodge, J Horne, D Jones, E Todd, C Smith-Keune, M Kayan. My thanks also extend to JCU for granting me a scholarship and tuition waiver, which allowed me to complete this PhD.

Lastly, I thank these people for their help and assistance in the field, S Gudge and I Kerr from Lord Howe Island Marine Park and, D Biggs (Charter Marine), J Edward (Bounty Divers), D Creek, M Smith, J Marges, K Christian, and J and P Davidson (Reserves and Forestry) for their assistance at Norfolk Island.

ABSTRACT

In the terrestrial environment the highest rates of extinction have been in endemic species on isolated islands. In recent years, threats to coral reefs on isolated islands have increased causing both local and global extinctions. Effective management that maintains restricted range species must be based on identified important source and/or sink populations. This requires knowledge of abundance, distribution, habitat specialization, patterns and levels of population connectivity and levels of genetic diversity within and among inhabited locations for all species of interest. Here I evaluate the genetic risk of extinction in six species of coral reef fishes with different distribution range sizes (widespread, restricted, endemic), found in Australian coastal waters or on remote offshore islands. Population structure, connectivity and levels of genetic diversity were obtained and compared over spatial- and time- scales using mtDNA and species specific msatDNA (generated explicitly for this study). The six species examined were the Great Barrier Reef anemonefish (*Amphiprion akindynos*), McCulloch's anemonefish (*A. mccullochi*), Three-striped butterflyfish *(Chaetodon tricinctus*), large Doubleheader wrasse (*Coris bulbifrons*), Black cod (*Epinephelus daemelii*) and Galapagos shark (*Carcharhinus galapagensis*).

Chapter 1 defines endemic species; discusses what the biggest threats faced by endemic species are; defines the IUCN criteria for assessing extinction risk; establishes how genetic traits - especially population connectivity and genetic diversity - over different time scales (using both mt- and msatDNA) are vital to adequately evaluate realised dispersal.

Chapter 2 outlines general Material and Methods, provides species descriptions and general laboratory techniques and analyses used in each chapter; whilst species specific Material and Methods details are retained in individual thesis chapters.

Chapter 3 discovers historical hybridization between two sister species of anemonefish, the widespread Great Barrier Reef anemonefish (*Amphiprion akindynos*) and the endemic McCulloch's anemonefish (*A. mccullochi*), informed by a diverse range of mtDNA and msatDNA data analyses.

Chapter 4 reveals patterns and levels of population connectivity and genetic diversity in the endemic McCulloch's anemonefish (*A. mccullochi*), informed by a diverse range of mtDNA and msat DNA data analyses.

Chapter 5 reveals patterns and levels of population connectivity and genetic diversity in the endemic three-striped butterflyfish (*Chaetodon tricinctus*) using similar approaches to those used in the previous chapters.

Chapter 6 reveals patterns and levels of population connectivity and genetic diversity in the endemic doubleheader wrasse (*Coris bulbifrons*) using similar approaches to those used in the previous chapters.

Chapter 7 does three things. Firstly, it combines data from previous chapters with data published on black cod (*Epinephelus daemelii*) and Galapagos shark (*Carcharhinus galapagensis*) populations structure to determine whether patterns and levels of population connectivity and levels of genetic diversity of the diverse assemblage of fish species from five different families (Pomacentridae, Chaetodontidae, Labridae, Epinephelinae, Carcharhinidae) (Chapters 3 – 6, A1 – 3), within the restricted distribution range of the Lord Howe Island and Norfolk Island rises are the same or not. Secondly, it assesses the local or global genetic extinction risk of each examined species in an IUCN framework by ranking relevant traits in an extinction risk matrix to generate a relative genetic extinction risk for each species. Finally, this

chapter establishes if a single or multiple spatial management strategies are required to protect the unique biodiversity at the isolated South-west Pacific Ocean islands they inhabit (Chapter 7).

Appendix A1 describes a suite of new msatDNA markers that I developed for McCulloch's anemonefish (*A. mccullochi*) population genetic studies.

Appendix A2 describes a suite of new msatDNA markers that I developed for three-striped butterflyfish (*Chaetodon tricinctus*) population genetic studies.

Appendix A3 describes a suite of new msatDNA markers that I developed for doubleheader wrasse (*Coris bulbifrons*) population genetic studies.

Both anemonefishes and the wrasse were more susceptible to extinction than the other three species, due to their relatively low levels of historic and demographic gene flow. All species had high genetic diversity, suggesting high levels of adaptive capacity and thus some resilience to environmental change. However, the high levels of self-replenishment (and consequently low levels of recent migration) are a cause for concern on one hand, as it suggests that most locations are largely self-sustaining. Although isolated reefs/islands are less prone to major disturbances than coastal reefs, when they do occur, their remoteness may increase the risk of extinction faced by endemic species, as a viable population may not exist following a disturbance (especially if populations have high self-replenishment). On the other hand, high levels of self-replenishment are crucial for the persistence of populations on demographic time scales. The similar levels of population connectivity and high self-replenishment shown across this taxonomically wide range of species that vary in ecological and life history traits, suggests that environment shapes genetic traits in species at this study system. This is encouraging as a single management plan for the design of effective marine reserves in the region may be feasible. This management strategy involving a network of MPAs that protects part of each location in the

geographic range of endemics is likely to be effective at conserving the unique biodiversity of endemism hotspots. However, computer models incorporating ecology, life history, population connectivity, oceanography and different environmental change scenarios are needed to further our understanding of extinction risk at these remote islands, now and in the future.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: **General Introduction**

"To do science is to search for repeated patterns, not simply to accumulate facts" (MacArthur 1972).

Endemic species, by definition, are species unique to a defined geographic location (Gaston 1994). Small range endemics are thus species restricted to small regions, as a result of limited dispersal or ecological barriers. In the terrestrial environment, the highest rates of extinction over the last 400 years have been in endemic species on isolated islands (Whittaker & Fernández-Palacios 2007). This is because these endemics have a small geographic range combined with ecological, biological and genetic traits that increase their risk of extinction (Gaston 1994, Frankham 1997, 1998). Furthermore, human activities, such as the introduction of species, habitat destruction and over-exploitation (Reid & Miller 1989), combined with the introduction of diseases (Warner 1968, Flesness 1989) have all been major causes of extinctions on islands. In the marine environment, numerous population declines and local extinctions have been recorded in recent times, but relatively few global extinctions have been observed (Dulvy et al. 2003). In the tropical marine environment, isolated islands are endemism hotspots, that is, a high proportion of the community is comprised of endemic species (Jones et al. 2002). If patterns of extinction risk follow those of the terrestrial environment, then isolated islands may represent locations where there are a high proportion of vulnerable marine species. As for terrestrial animals,

marine island endemics may exhibit an elevated extinction risk through a range of ecological and genetic factors.

The diversity, frequency and severity of threats to coral reefs have increased greatly in recent years (Hoegh-Guldberg 1999, Hughes et al. 2003, Bellwood et al. 2004), affecting both small and large range species. A key emerging threat to coral reefs is global warming where increases in ocean temperature cause "mass bleaching events" which result in widespread mortality to corals (McClanahan 2004, Glynn 2012) and anemones (Scott & Hill 2011, Hattori 2002). With future impacts, particularly coral bleaching, expected to increase in severity and frequency (Harvell et al. 1999, Hoegh-Guldberg 1999, Sheppard 2003), habitat loss is likely to cause local extinctions of specialist fishes that rely on highly vulnerable coral habitats. Indeed, extensive or selective coral loss has caused population declines or local extinctions of some coral reef fishes, particularly species reliant on live coral (Kokita & Nakazano 2001, Graham et al. 2006, Pratchett et al. 2008), while the loss of anemones during bleaching events has resulted in local extinctions of anemonefish (Hattori 2002).

Due to their isolation, coral reefs on remote islands are potentially isolated from threats or anthropogenic disturbances and, despite limited connectivity, can recover from major disturbances (Gilmour et al. 2013). However, these remote reefs may be even more vulnerable when disturbances do occur, due to the lack of connectivity. The ability for coral reef fishes to recover from local extinction will

be dependent on the regeneration of coral reefs and the arrival of larvae from distant source populations. Models predict that gene flow of a few recruits per generation between locations over historical time scales will not sustain populations (Cowen et al. 2000, Cowen et al. 2002). Widespread species inhabiting isolated locations may have limited resilience, because recolonisation may be slow due to greater distances to nearest source populations (Ayre & Hughes 2004). In contrast, populations of restricted range species may simply fail to recover if no viable source population remains following a disturbance. This comes in light of recent studies showing how populations are maintained by high levels of self-recruitment (i.e. the supply of new individuals into a population) in a range of coral reef fishes (bluehead wrasse in the U.S. Virgin Islands, Swearer et al. 1999; anemonefish and butterflyfish in Papua New Guinea, Jones et al. 2005, Almany et al. 2007, Planes et al. 2009; doubleheader wrasse at Lord Howe Island, Patterson & Swearer 2007; stripey snapper and coral trout at Keppel Islands, Harrison et al. 2012). Consequently, there is an urgent need to understand gene flow among, and genetic diversity at, locations inhabited by endemic reef fishes for ongoing monitoring and conservation, and to determine their recolonisation ability and resilience.

The International Union for Conservation of Nature (IUCN) assesses the conservation status of all species using the IUCN Red List Categories and Criteria (IUCN 2012, http://www.iucnredlist.org). For the purpose of regional conservation assessments, species extinction risks are published as Red Lists

within a specific geographically defined area. This would usually involve delineating species boundaries, identifying subpopulations, and threats specific to that geographic area. These assessments use a range of ecological (i.e. abundance, range size, specialization), life history (i.e. reproduction, growth rate) and genetic traits (i.e. population structure, gene flow) where avaliable.

The IUCN Red List defines subpopulations as "geographically, or otherwise, distinct groups in the population between which there is little demographic or genetic exchange (typically one successful migrant individual or gamete per year or less)". To apply this definition ecologists usually turn to genetic or mark recapture techniques. Increasingly, ecologists are encouraged to use genetic approaches to estimate connectivity (gene flow via migrant exchange and/or self-recruitment) and population size; given massive advances in both genetic techniques and post processing software (Selkoe & Toonen 2006). However, to date two alternative perspectives on dispersal and connectivity have been predominantly applied: either historical or demographic.

Evolutionary biologists use mtDNA (mitochondrial) and/or nDNA (nuclear DNA) to determine genetic exchange between populations that may have been isolated over thousands of generations (i.e. over historical time scales). They detect relatively rare chance events of genetic exchange that introduce foreign alleles that influence modes of speciation, and spread adaptive change (Hellberg 2009). Ecologists use msatDNA (microsatellites) to determine genetic exchange over a

single or few generations (i.e. demographic time scales) and are predominantly interested in two issues (i) estimating self replenishment (i.e. indirect approach) or self-recruitment (i.e. direct approach) to determine if populations are demographically open, closed or intermediate (e.g. Swearer et al. 1999, Jones et al. 2005, Almany et al. 2007, Planes et al. 2009, Berumen et al. 2012) – often used to assess the effectiveness of conservation measures such as the placement of marine protected areas (MPAs); (ii) estimating demographic independence or interdependence of populations, which are considered independent if gene flow (*m*) is under 10 % and interdependent if gene flow is over 10 % (Waples & Gaggiotti 2006). Thus, these two perspectives – evolutionary and ecological, represent dispersal over a range of time scales (i) when compared relative to each other and (ii) through the use of specific analyses designed to answer questions using the marker, and hence time scale, of choice. The best approach should combine both evolutionary and ecological perspectives (i.e historical and demographic time scales; Hellberg 2009) since this will provide a more realistic overview of dispersal (Leis et al. 2011). Henceforth, I discuss genetic exchange in the context of the relative difference in time scales between these two perspectives (i.e. where evolutionary genetic exchange of mtDNA represents historical time scales relative to ecological genetic exchange of msatDNA which represents demographic time scales).

In addition to population connectivity, genetic diversity is an important indicator of vulnerability to extinction. In terrestrial environments, endemics and isolated

populations of widespread species are expected to have low genetic diversity due to limited gene flow to other populations (Diamond 1984, Frankham 2002) and their small population size (Gaston 1994, Gaston et al. 1997), which allows for a loss of genetic diversity through genetic drift (Frankham 1996). These factors increase risk of extinction (Frankham 1996, 2002) because low genetic diversity reduces a species potential to adapt to changing environmental conditions. The IUCN understands the importance of genetic diversity to extinction risk and considers conservation of genetic diversity a priority (McNeely et al. 1990, Frankham 2002). If endemic marine species or isolated populations of widespread species also have low genetic diversity, then this would increase their risk of extinction.

The remote Australian offshore coral reefs in the south-west Pacific ocean, comprising Middleton Reef (MR), Elizabeth Reef (ER), Lord Howe Island (LHI) and Norfolk Island (NI; Fig 1.1) are a regional hotspot for endemic coral reef fishes (Marine Parks Authority 2010), ranking 4th in the Indo-Pacific for endemism (7.2% endemism), behind Hawaii (25%), Easter Island (22.2%) and the Marquesas (11.6) (Randall 1998; 2001; 2007). This region provides an ideal study system as endemic reef fish occur only at these four discrete locations that are separated by deep ocean, resulting in connectivity that is likely to be totally restricted to pelagic larval dispersal over known distances (45 - 600 km). The connectivity among subpopulations at these locations is still poorly understood spatially and at historical and demographic time scales, despite the importance of

this information for effective management (e.g. Leis et al 2011) and conservation of endemic species. MPAs have been established at three locations (MR, ER, LHI) to conserve reef fishes, but no such protection exists at NI. It is not known if the placement of these MPAs is effective and without the guidance of empirical data on connectivity, managers are often forced to implement strategies based on "best guesses" (McCook et al. 2009).

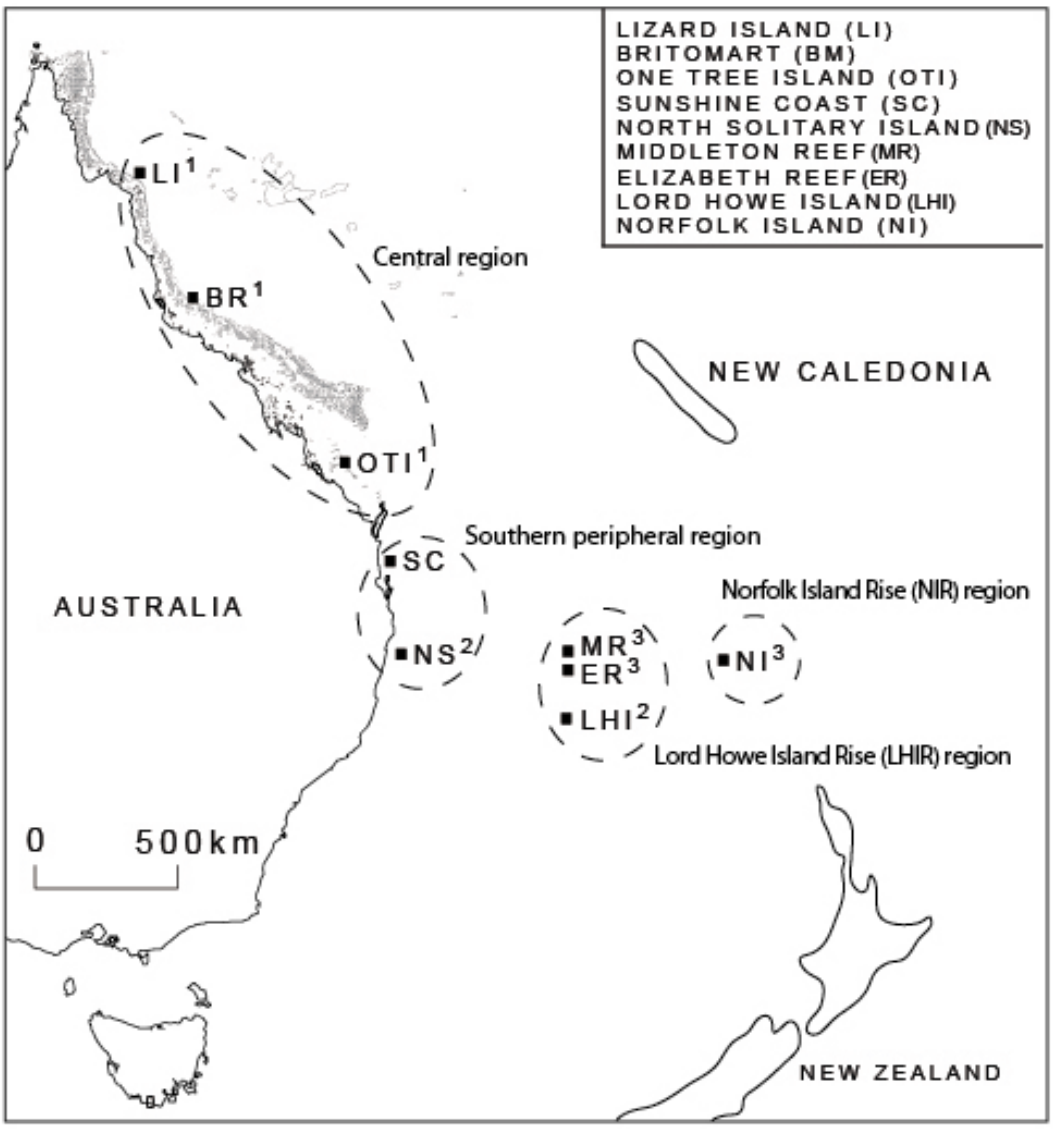

Figure 1.1 Sampling locations, management agencies and identified regions. Sample locations: Lizard Island (LI), Britomart (BM), One Tree Island (OTI), Sunshine Coast (SC), North Solitary Island (NS), Middleton Reef (MR), Elizabeth Reef (ER), Lord Howe Island (LHI), Lord Howe Island Lagoon (LHIL) and Norfolk Island (NI). Management agencies: Great Barrier Reef Marine Parks Authority (GBRMPA)[1], New South Wales Department of Primary Industries (NSWDPI)[2] and the Department of Sustainability, Environment, Water,

Population and Communities (SEWPaC)[3]. Identified regions: central, southern peripheral, Lord Howe Island Rise (LHIR) and Norfolk Island Rise (NIR).

Previous marine genetic research in this region examined population connectivity in two restricted range species, the IUCN listed Near Threatened black cod (Epinephelus daemelii; http://www.iucnredlist.org/details/61337/0, Appleyard & Ward 2007, van Herwerden et al. 2009) and the IUCN listed Near Threatened Galapagos shark (Carcharhinus galapagensis; http://www.iucnredlist.org/details/41736/0, van Herwerden et al. 2008). The black cod, distributed from the NSW coastline (Australia) east to MR, ER, LHI, NI and the Kermadec Islands (New Zealand), is listed as 'vulnerable' by the New South Wales (NSW) Fisheries Management Act 1994, due to population declines potentially caused by overfishing. Likewise, the Galapagos shark, with a worldwide distribution that is restricted to remote offshore islands and present within Australian waters at MR, ER, LHI and NI only, is 'near threatened' (Bennet et al. 2003) due to population declines, also potentially caused by overfishing. These two species probably have highest risk of extinction in this study system as both are more intensively targeted by fisheries, than any other species, and sharks particularly reproduce slowly and in low numbers compared to fishes. Furthermore, local extinctions have already been reported for the Galapagos shark (Luiz & Edwards 2011). However, neither study sampled these species across their range (leaving unsampled 'ghost' populations). Moreover, both studies suffered from data limitations with mtDNA and if used, inadequate

msatDNA loci, to establish connectivity. Unsampled 'ghost' populations can affect key demographic estimates (e.g. population size, genetic diversity, migration rate; Beerli 2004), whilst limited loci reduces the statistical power to resolve connectivity among locations (see Selkoe & Toonen 2006).

Likewise data on black cod and Galapagos shark ecological and life history traits are lacking. Thus, these studies provide insufficient information to managers. Furthermore, body size and life history of black cod and Galapagos sharks are not typical of the majority of reef fishes in this region. Therefore, connectivity studies of other endemic reef fishes in this region are required to determine if taxonomically distinct species show similar levels of population connectivity and genetic diversity, and to determine if a single management strategy can inform management, rather than having to rely on "best guesses".Future research should focus on collecting black cod and Galapagos shark samples from throughout their ranges to determine rangewide genetic traits; while for both these species and those in this thesis, ecological and life history traits need to be determined in order for a more holistic understanding of extinction risk in this study system.

To address this knowledge gap, albeit with limitations noted above, I sampled three endemic reef fishes: McCulloch's anemonefish (*Amphiprion mccullochi*), three-striped butterflyfish (*Chaetodon tricinctus*), and doubleheader wrasse (*Coris bulbifrons*) from their entire range (MR, ER, LHI, NI; Fig 1.1) to determine the patterns and levels of connectivity over spatial- and time- scales, and levels of

genetic diversity (van der Meer et al. 2012a, b, c; van der Meer et al. 2013a, b, Chapters 3 to 6). Three suites of new msatDNA loci were designed (van der Meer et al. 2012 d, e, f; Chapters A1 to A3) to be used in population genetic studies. In addition to these endemic species, the Great Barrier Reef (GBR) anemonefish (*Amphiprion akindynos*) was also examined from its Australian (GBR) distribution. This species is restricted to areas adjacent to the LHI region and historically hybridised with its sister species *A. mccullochi,* an endemic to the LHI region (van der Meer et al. 2012a), as was established in the course of the present study. Furthermore, preliminary phylogenetic and population genetic analyses undertaken during this study indicate that *A. akindynos* shows peripheral speciation at the southern edge of its range (data not shown) – a process documented in other species (corals, Budd & Pandolfi 2010; coral reef fish, Bowen et al. 2013; zooplankton, Dawson & Hamner 2005).

In this thesis I combine prior findings (black cod, Galapagos shark) with findings from four additional coral reef fishes examined here (GBR anemonefish, McCulloch's anemonefish, three-striped butterflyfish, doubleheader wrasse). Combining genetic data of species with differing life histories, ecologies and dispersal abilities enables us to determine if a single, or multiple spatial management strategies are required for conserving endemic marine species in this region. This is important for conservation of biodiversity, particularly as three government/state agencies manage the waters inhabited by these species (Fig 1.1): the Great Barrier Reef Marine Park Authority (GBRMPA – GBR), the New South

Wales Department of Primary Industries (NSWDPI – North Solitary and state waters of Lord Howe Islands) and the Department of Sustainability, Environment, Water, Population and Communities (SEWPaC – Middleton and Elizabeth Reefs, Norfolk Island, Commonwealth waters of LHI).

Thus my aims were threefold, to: (i) determine whether patterns and levels of population connectivity (historical connectivity, demographic connectivity, self-replenishment and recent migration) and levels of genetic diversity within this restricted distribution range are the same or not, for the diverse assemblage of species from five different families: Pomacentridae (Chapters 3 – 4, A1), Chaetodontidae (Chapter 5, A2), Labridae (Chapter 6, A3), Epinephelinae and Charcharhinidae (Chapter 7), (ii) assess the local or global genetic extinction risk of each species in an IUCN framework (Chapter 7) and, (iii) establish if a single or multiple spatial management strategies are required to protect these species at the isolated south west Pacific Ocean islands they inhabit (Chapter 7).

CHAPTER 2: **General Material and Methods**

2.1. Study system (Elizabeth/Middleton Reefs and Lord Howe/Norfolk Islands)

Middleton Reef (MR) lies 600 km east of the Australian coast (29°S 159°E) in the northern Tasman Sea, with Elizabeth Reef (ER) 45 km due south of MR, Lord Howe Island (LHI; 31°S 159°E) 160 km south of ER and, Norfolk Island a further 600 km to the east of MR/ER (29°S 168°E; Fig 1.1). These remote coral reefs/islands rise independently from deep water (>2000m) as seamounts and are separated by deep oceanic waters. All locations (especially NI) are far removed from the nearest major coral formations of the southern Great Barrier Reef (GBR), which can be found 900 km to the north (Choat et al. 2006).

The Elizabeth and Middleton Reefs Marine National Nature Reserve covers an area of 188 000 hectares, and includes both the ER (~5,100 ha) and MR (~3,700 ha) which are the southern-most open-ocean platform reefs in the world (Oxley et al. 2004). Lord Howe Island sits on the western margin of the basaltic Lord Howe Rise (along with MR and ER), is crescent in shape rising 800 m above sea level (asl), 11 km long and 4 km wide and has a long, sheltered coral fringed lagoon on the western coast (Edgars et al. 2010). Norfolk island lies on the Norfolk Island Rise, is roughly circular in shape rising 319 m asl, 8 km long and 5km wide and, has no true coral reefs but rather a limestone reef (1 km long) capped with corals at Emily Bay (Francis & Randall 1993).

The boundary between the tropical Coral Sea and the temperate Tasman Sea lies close to these high latitude reef systems, known as the Tasman Front, which oscillates in a north-south direction leading to seasonally warmer or cooler sea temperatures (Suthers et al. 2011). The prevailing currents are from the southward flowing East Australian Current (EAC) and the tropical convergence (Middleton et al. 2006), which provide a mechanism for colonisation by larvae from the extensive area of coral reefs to the north. Thus, this area provides habitats for both tropical and subtropical/warm temperate species and represents transitions from biologically complex tropical habitats, dominated by scleractinian corals, to temperate habitats with a greater biomass of macroscopic algae (Crossland 1988, Johannes et al. 1983). As a result, this area harbors a number of endemic species including the ones in this thesis: the McCulloch's anemonefish (*Amphiprion mccullochi*), the three-striped butterfly fish *(Chaetodon tricinctus)* and the doubleheader wrasse (*Coris bulbifrons*).

For simplicity, I refer to the four main geographic regions in this study as central region (GBR), southern peripheral region (North Solitary, NS), Lord Howe Island Rise region (LHIR – MR, ER, LHI; remnants of 6.7 Million year old (Ma) volcanoes) and Norfolk Island Rise region (NIR; remnants of a 2.3 to 3.05 Ma volcano) (Fig 1.1). Details of study species, family name, geographic range, sample locations and, the genetic marker used for mtDNA analyses and the number of genetic markers used in msatDNA analyses are provided in Table 2.1.

Table 2.1 Details of study species, family name, geographic range, sample locations and, the genetic marker used for mtDNA analyses and the number of genetic markers used in msatDNA analyses. Lizard Island (LI), Britomart Reef (BM), One Tree Island (OTI), Sunshine Coast (SC), North Solitary Island (NS), Middleton Reef, (MR), Elizabeth Reef (ER), Lord Howe Island (LHI), Lord Howe Island Lagoon (LHIL), Norfolk Island (NI).

Species	**Family**	**Geographic range**	**Sample locations**	**Genetic markers**	
				mtDNA	msatDNA
Amphiprion akindynos	Pomacentridae	Great Barrier Reef (GBR) south to the North Solitary Islands (NS) and extending out to New Caledonia, but not including MR, ER, LHI, NI.	LI, BM, OTI, SC, NS	D Loop	17
Amphiprion mccullochi	Pomacentridae	MR, ER, LHI	MR, ER, LHI, LHIL	D Loop	17
Chaetodon tricinctus	Chaetodontidae	MR, ER, LHI, NI	MR, ER, LHI, NI	Cyt b	20
Coris bulbifrons	Labridae	MR, ER, LHI, NI	MR, ER, LHI, NI	D Loop	17
Epinephelus daemelii	Epinephelinae	NSW coastline (Australia) east to MR, ER, LHI, NI and the Kermadec Islands (New Zealand)	MR, ER, LHI	D Loop	6
Carcharhinus galapagensis	Carcharhinidae	Remote offshore locations around the world but in Australia it is only found at MR, ER, LHI and NI	MR, ER, LHI	D Loop	3 (none were informative)

2.2. Study species

2.2.1. McCulloch's anemonefish (*Amphiprion mccullochi*)

As a subfamily, Amphiprioninae comprise of 28 species with a distribution ranging longitudinally between the Philippines and the GBR, and latitudinally between Sumatra and Melanesia (Santini & Polacco 2006). They are among the most specialised groups of reef fishes having an obligate relationship with their host anemones (Allen 1972). Within this group, the endemic McCulloch's anemonefish (*Amphiprion mccullochi*) is important as it is potentially at risk of extinction because (i) its geographic range is among the smallest for coral reef fishes (MR, ER, LHI), (ii) it's an extreme habitat specialist due to its obligate relationship with only one host species of anemone (Fautin & Allen 1992) and; (iii) throughout its range it has very low abundance (Choat et al. 2006, Hobbs & Feary 2007, Hobbs et al. 2009), except for an extremely small area of habitat at Lord Howe Island Lagoon (LHIL), which supports 92% of the world's *A. mccullochi* population (Hobbs et al. 2007).

2.2.2. Three-striped butterflyfish (*Chaetodon tricinctus*)

Butterflyfishes, are eye-catching and iconic inhabitants of coral reef environments, with over 130 species in the family being represented in all coral reef regions (Allen et al. 1998). Within this family, the three-striped butterflyfish (*C. tricinctus*) is one of the 41butterflyfish species that feed directly on scleratinian corals (Cole et al. 2008, Rotjan & Lewis 2008). *C. tricinctus* is important as it is potentially at risk of extinction because of (i) its small geographic range (MR, ER, LHI, NI) and (ii) it is a habitat specialist that feeds exclusively on live corals (Kuiter 2002) and is mostly found in close association with corals of the genus *Acropora* (Hobbs et al. 2009). Furthermore, the abundance of *C. tricinctus* is positively linked to the abundance of *Acropora*

spp., indicating that a loss of this coral could cause decreases in abundance and potential local extinction of *C. tricinctus* (Hobbs et al. 2009).

2.2.3. Doubleheader wrasse (*Coris bulbifrons*)

The wrasse family is large and diverse consisting of about 400 species in 60 genera, many of which are brightly colored (Nelson 1976) and readily recognized by their thick lips, the inside of which is sometimes curiously folded (Choat and Bellwood 1998). Within this family, the doubleheader wrasse (*Coris bulbifrons*) is important as it is vulnerable due to overfishing and overexploitation by the IUCN (Choat & Pollard 2010) because (i) of its small geographic range (MR, ER, LHI, NI), (ii) it is targeted for food in recreational line fisheries, especially at LHI and ER, (iii) its locally abundant in sheltered habitats of MR, ER and LHI (Choat et al. 2006a, Hobbs & Feary 2007, Hobbs et al. 2009) but is rare at NI (*pers. Obs. MvdM and J-PH*) and (iv) declining number of mature individuals.

2.3. Genetic analyses

I applied a range of frequency and Bayesian based molecular tools to establish mtDNA and msatDNA levels of phylogenetic and population genetic structure. This resulted in reliable estimates of gene flow in this study system and used together, these tools provided a well-rounded view of dispersal (Leis et al. 2011). I used a large number of polymorphic microsatellite loci and sampled all known locations of all three species, to compensate for the small sample sizes used in this study (see Selkoe & Toonen 2006). Furthermore, I recognize that my estimates for "self-replenishment" inferred indirectly from genetic markers are merely a proxy for self-recruitment, which is typically assessed using more direct methods (e.g. natural or artificial

otolith tags), such as those used by Swearer et al. (1999), Jones et al. (2005) and Almany et al. (2007). Nevertheless, direct destructive sampling approaches are not feasible for this study, without negatively impacting populations, due to the large sample sizes typically required for such parentage-based studies (see Berumen et al. 2012, Harrison et al. 2013). Therefore, I believe that my indirect estimates of self-replenishment represent the best possible substitute for realised self-recruitment obtainable for these species.

NOTE: For this thesis, only two changes were made to already published papers (Chapters 3, 4, 6, A1 - 3): (i) I present general genetic methods applicable to all chapters hereunder, whilst retaining critical analyses information in each chapter. Chapters 5 (in review) and 7 (ready for submission) are presented word for word. (ii) I have removed references from all chapters and combined them into a single bibliography at the end of the thesis.

2.3.1. mtDNA analyses

DNA extraction. Total genomic DNA was extracted from approximately 1mm^3 of tissue using standard chelex-proteinase K digestion extraction procedures (see Walsh et al. 1991; used in Chapter 3) and a Qiagen Gentra Puregene extraction protocol (Qiagen, Australia; used in all subsequent Chapters). For the latter extraction protocol, sample tissue 6-8 mm^2 in size, was blotted on KimWipes (Kimberley-Clark, USA) to remove any Ethanol residues, placed into a 2 ml Eppendorf tube (Eppendorf, Australia) containing a solution of 200µl of Qiagen Gentra Puregene Cell Lysis Solution (Qiagen, Australia) and 1.5µl (20mg/ml) of Proteinase K (Qiagen, Australia). Tissue was subsequently macerated using spring scissors. Samples were incubated in a rotating oven pre-heated to 55^{o}C for 12 hours. Following this, samples were removed and 2-

5µl RnASE (Qiagen, Australia) was added to the samples and returned to oven for a further 30 minutes to 1 hour. Samples were cooled on ice for 10 minutes upon removal from oven. 66µl of Qiagen Gentra Puregene Protein Precipitation Solution (Qiagen, Australia) was added to each sample and vortexed at high speed for 15 seconds. Samples were centrifuged at 13 000 rpm for 6 minutes. The resulting supernatant was transferred by gently pouring it into a clean 2mL Eppendorf tube containing 200µl of 100% Isopropanol (the tube containing the protein pellet was then discarded). The sample was mixed by gently inverting 50 times and then centrifuged at 13 000 rpm for 2 minutes. The Isopropanol supernatant was carefully poured off, ensuring that the DNA pellet remained in the tube. The tube was gently blotted on clean Kim Wipe's to remove excess Isopropanol. The pellet was washed by adding 200µl of 70% Ethanol. The tube was again inverted gently 10 times and then centrifuged at 13 000 rpm for 2 minutes. The Ethanol supernatant was discarded and the tube containing the DNA pellet was gently blotted on clean Kim Wipe's to remove excess Ethanol. The tube was covered with clean Kim Wipe's and left open to evaporate remaining Ethanol for 2-4 hours. DNA was rehydrated into a primary stock with 50-100µl of a 1 x Tris-Edta Buffer (pH 8.0, filter sterilised) overnight. Gel electrophoresis was used to quantify DNA with a DNA Size Standard - Lambda standards (Bio-Rad, USA) and determine molecular weight with Easy Ladder 1 (Bioline, UK). Nano-Drop (Thermo Scientific, USA) was used to determine DNA quality. Primary stocks were frozen at -20 ^{o}C.

PCR reactions. Each 20 µl mtDNA PCR amplification reaction contained 20ng DNA, 0.2mM DNTP, 1 unit of Bioline Biotaq Red DNA Polymerase, 0.5µM of each primer and variable $MgCl_2$. Primers were tested and optimised using a Bio-Rad C1000 Thermal Cycler.

Amplifications followed the same basic cycling protocol: 40 s at 94°C, 40 s at primer specific annealing temperatures and 40 s at 72°C. The cycling profile was flanked by an initial 3 min denaturing step (94°C) and a 5 min terminal extension phase (72°C). I genotyped each species using microsatellite markers that I developed (see Appendexes 1, 2, 3). Six multiplex reactions were performed and each 10 µl PCR amplification reaction contained 10 ng DNA, 5 µl (2x) Type-it Multiplex PCR Master Mix, 2 µl multiplex primer mix (at a concentration of 2 µM). Amplifications followed the same basic cycling protocol: 30 s at 95°C, 90 s at 58°C and 40 s at 72°C. The cycling profile was flanked by an initial 5 min denaturing step (94°C) and a 30 min terminal extension phase (72°C). PCR products were purified by ethanol ammonium acetate precipitation for genotyping (Genetic Analysis Facility, James Cook University, Townsville).

Phylogenetic tree. Forward sequences were automatically aligned using the plugin CLUSTAL W in Geneious Pro 4.7 (Drummond et al. 2009), conservatively trimmed to minimise the amount of missing data, and manually edited, inserting gaps where required and checking for ambiguities. Differences between individual sequences were determined for the following characters: A, G, T, C and IUB symbols (Nomenclature Committee 1985). Sequence data were obtained from GenBank for species to act as outgroups.

The data from all chapters followed the same phylogenetic analyses: (1) Bayesian inference (MB) in MrBayes 3.1 (Huelsenbeck et al. 2001, 2002) with 10 million generations of Monte Carlo Markov Chains (MCMC); (2) Ten independent Maximum Likelihood (ML) analyses, followed by an 100 bootstrap replicate analysis using GARLI 0.951 (Zwickl 2006), from which a 50% majority rule ML consensus tree was constructed in PAUP* 4.10b (Swofford 2001); (3)

Maximum Parsimony (MP) was performed in MEGA 4.0 (Tamura et al. 2007) with 1000 bootstrap replicates, from which a 50% majority rule consensus tree was constructed and; (4) Bayesian inference in BEAST V1.6.1 (Drummond & Rambaut 2007) was tested using a strict clock model (estimated clock rate, uniform prior distribution, HKY+G site model for 5 million MCMC chains with sampling at every 5000 trees) and a Relaxed clock model (uncorrelated lognormal clock model, with the same parameters noted previously). Tracer V1.5, identified no significant difference between clock models based on the Bayes Factor (BF) evaluation of the models, (Kass & Raftery 1995, Suchard et al. 2001), BF = 0. I therefore used the Speciation (Yule process) to construct a Bayesian skyline plot using a Strict clock model (as above, but with 50 million MCMC chains sampled every 5000^{th} tree. Maximum clade credibility trees (MCCT) were constructed in TreeAnnotator V1.6.1 after discarding the initial burn-in of 10%. The MCCT was viewed separately in FigTree V1.3.1 (available at http://tree.bio.ed.ac.uk/software/figtree/). The best outgroup rooted ML tree from GARLI was selected to reconstruct the evolutionary history with bootstrap values for each clade from all four analyses, if present.

Minimum Spanning Tree (MST). An MST was generated based on output obtained from ARLEQUIN 3.5 (Excoffier et al. 2005) to explicitly identify shared haplotypes between the four locations (MR, ER, LHI, NI).

Quantifying the level of mtDNA gene flow. mtDNA migration rates and effective population sizes were estimated between or within each location using MIGRATE-n 2.4.3 (http://popgen.sc.fsu.edu/Migrate-n.Html; Beerli & Felsenstein 2001, Beerli 2004). I tested a combination of various migration priors (F_{st} and own: isolation-by-distance), custom-migration

models (Stepping-stone, Island-n and variable Theta only) and a geographic matrix - all with a constant mutation rate.

2.3.2. msatDNA analyses

Spatial population partitioning. Three molecular analytical tools were used to establish spatial population partitioning in msatDNA: (i) Discriminant Analysis of Principal Components (DAPC; Jombart et al. 2010) uses allelic states to discriminate between the four locations, yielding scatterplots of discriminant functions based on the spatial distributions of microsatellite genotypes. DAPC also provided posterior probabilities of population assignments for each individual; (ii) a likelihood-based assignment method was used in GeneClass2 (Paetkau et al. 1995, Paetkau et al. 2004, Piry et al. 2004) to determine significant inter-location gene flow and (iii) STRUCTURE V2.3 (Pritchard et al. 2000, Hubisz et al. 2009) placed individuals into clusters that minimize Hardy-Weinberg Equilibrium (HWE) and can be used to identify demographic gene flow between the four locations. To determine the "best value" for K, I followed the method suggested by Pritchard et al. (2000), which involved comparing mean log-likelihoods penalised by one-half of their variances (see Hubisz et al. 2009). The final run consisted of an Admixture model with 2 M iterations and a 100 k iteration burn-in.

Quantifying the level of msatDNA gene flow. Demographic migration rates and effective population sizes were tested and estimated between each four locations using MIGRATE-n 2.4.3 as above.

Inferred levels of self-replenishment and recent migration. I did not sample new recruits for any species in order to determine self-recruitment as in Jones et al. (2005). However, I used BAYESASS v3 (Wilson & Rannala 2003), a program specifically designed for population genetic studies that estimates recent migration rates (past 2 - 3 generations) between populations (or locations). Conversely, this program also has the ability to estimate any individuals not migrating (i.e. self-replenishing). To determine if the dataset was appropriate for use in BAYESASS, I used the results from both DAPC and STRUCTURE. Used together, these programs are likely to be better than F_{st} values (Faubet et al. 2007) at determining the appropriateness of a dataset for BAYESASS because they extract more information from the genetic data than frequency based fixation indices. However, estimates of migration rates are accurate when model assumptions are not violated (i.e. migration rates are low, loci are in linkage equilibrium) and genetic differentiation is not too low ($F_{st} \geq 0.05$). If model assumptions are violated, then accurate estimates are obtained only when migration rates are very low (m = 0.01) and genetic differentiation is high ($F_{st} \geq 0.10$; Faubet et al. 2007). Lastly, when migration rates fall below 10%, populations can probably be considered demographically independent (Waples & Gaggiotti 2006).

2.3.3. Population genetic diversities

Molecular diversity indices for mtDNA (haplotype diversity, h; nucleotide diversity, π) and for msatDNA (genetic diversity, *gd*) were estimated in ARLEQUIN 3.5 (Excoffier et al. 2005). Haplotype and nucleotide diversities of the data were interpreted as either low with specified cut-off values of h and π (%) < 0.5 or high if values of h and π (%) were > 0.5 (Grant & Bowen 1998).

CHAPTER 3: **Historic hybridisation and introgression between two iconic Australian anemonefishes and demographic patterns of population connectivity**

(Published as: van der Meer MH, Hobbs J-PA, Jones GP, van Herwerden L (2012) Historic hybridisation between two Australian anemonefish species *Amphiprion* and contemporary patterns of connectivity. *Ecology and Evolution,* 2, 1592−1604)

3.1. ABSTRACT

Endemic species on islands are considered at risk of extinction for several reasons, including limited dispersal abilities, small population sizes and low genetic diversity. I used mitochondrial DNA (D-Loop) and seventeen microsatellite loci to explore the historical relationship between an endemic anemonefish, *Amphiprion mccullochi* (restricted to isolated locations in subtropical eastern Australia) and its more widespread sister species, *A. akindynos*. An mtDNA phylogram showed reciprocal monophyly was lacking for the two species, with two supported groups, each containing representatives of both species, but no shared haplotypes and up to twelve species- but not location-specific management units (MU). Population genetic analyses suggested historical connectivity among samples of each species (mtDNA), whilst ecological connectivity was only evident among populations of the endemic, *Amphiprion mccullochi*. This suggests higher dispersal between endemic anemonefish populations at both historical and ecological timeframes, despite separation by hundreds of kilometres. The complex mtDNA structure results from historical hybridisation and introgression in the historical past of these species, validated by msat analyses (NEWHYBRIDS, STRUCTURE and DAPC). Both species had high genetic diversities

(mtDNA $h > 0.90$, $\pi = 4.0\%$; msat genetic diversity $gd > 0.670$). Whilst high genetic diversity and connectivity reduce extinction risk, identifying and protecting populations implicated in generating reticulate structure amongst these species should be a conservation priority.

3.2. INTRODUCTION

Remote islands often contain a disproportionate number of endemic species (Ceballos & Brown 1995, Randall 1998, Gillespie et al. 2008) and genetically distinct populations of species with broad geographic ranges (Avise 1992, Slatkin 1993). The uniqueness of island communities makes them a high conservation priority (Gillespie & Roderick 2002), even more so given the extremely high rates of local and global extinctions of species inhabiting islands (Steadman 1995). Island endemics face an elevated risk of extinction because they often have vulnerable traits, such as, limited dispersal abilities, small population sizes and low genetic diversity due to genetic drift and inbreeding (Frankham 1997, 1998). A sound knowledge of the historical history and adaptive capacity of endemic species or isolated populations is required in order to understand what increases their risk of extinction, thereby enabling the development of appropriate conservation measures (Faith 1992, Moritz 2002).

Throughout the world's tropical oceans, coral reef organisms are distributed on islands and reefs that represent varying degrees of isolation. For coral reef fishes, it is the most isolated locations that contain the highest levels of endemism (Lessios et al. 2001, Jones et al. 2002, Allen 2008). Isolated islands also support genetically differentiated populations of some widespread reef fish species (Muss et al. 2001, Planes & Fauvelot 2002, Winters et al. 2010). Islands communities are also characterised by a high proportion of vagrants, which can result in a high incidence of hybridisation due to a scarcity of conspecific mates (Hobbs et al. 2009b, Hobbs et al. in press). Reef fishes on isolated islands also appear to more vulnerable to extinction as evidence by recent extinctions (Roberts & Hawkins 1999, Dulvy et al. 2003). These extinctions highlight the

need to examine the genetic characteristics of island fish faunas to help determine why this group may be vulnerable. Processes that influence genetic resilience and promote species persistence times may be identified by examining spatial and temporal patterns of gene flow.

Coral reef anemonefishes (genus *Amphiprion*) represent a useful model system for understanding historical histories and population genetic structures of island fish faunas. Although the genus has a broad Indo-Pacific distribution, more than 25% of species are endemic to isolated islands or have peripheral populations at these remote locations (Fautin & Allen 1997). Although these island endemics are often closely related to more broadly distributed sister species, the historic colonization and speciation processes, and current levels of population differentiation are not well known. In a detailed phylogenetic study of 23 of the 28 anemonefish, Santini and Polacco (2006) suggested the group (Family: Pomacentridae, subfamily Amphiprioninae) originated some 5 to 13 million years ago in the Indo-Pacific, with many of the endemics being of recent origin. A more detailed investigation of their historic relationships with putative sister species, and current levels of gene flow among locations, is necessary to understand how endemic anemonefishes originate and persist. This information will be useful for predicting how these species will persist in the future and aid management strategies aimed at conserving these iconic coral reef fishes.

Using phylogenetic and population genetic analyses I reconstructed the historical history of two Australian anemonefishes: a small-range species, *Amphiprion mccullochi*

(endemic to Middleton Reef, Elizabeth Reef and Lord Howe Island) and its more widespread sister species, *A. akindynos,* found on the Great Barrier Reef, New Caledonia and the subtropical east coast of Australia. The endemic McCulloch's anemonefish is of particular conservation concern because it has three characteristics known to elevate the risk of extinction - a very small geographic range (Coleman 1980, Fautin & Allen 1992, Hobbs et al. 2009), extreme habitat specialisation (one species of host anemone - *Entacmaea quadricolor;* Fautin and Allen 1992) and relatively small local populations (Choat et al. 2006, Hobbs & Feary 2007). *Entacmaea quadricolor* is distributed from Micronesia and Melanesia to East Africa and the Red Sea and to Japan and Australia (Fautin & Allen 1992), but no further south than the Solitary islands (NSW) on the east coast of Australia. *Amphiprion mccullochi* is thought to have arisen by divergence from a more widespread most recent common ancestor (mrca) shared with *A. akindynos*, its sister species, on the eastern Australian coast (Santini & Polacco 2006). However, the mode of speciation and current levels of genetic connectivity among the three populations of *A. mccullochi* are unknown. *Amphiprion akindynos* may be less of a conservation concern because it is more widely distributed, inhabits 6 species of anemones (Fautin & Allen 1997), and can be locally abundant (e.g. Richardson 1999).

This study addressed the following specific questions: (1) What is the historical relationship between the sister species, *A. mccullochi* and *A akindynos,* based on mtDNA? (2) How many management units (MU, *sensu* Moritz 1994) can be identified for *A. mccullochi* and *A. akindynos*? (3) What is the demographic relationship of connectivity between and within species, based on msat DNA? 4) What are the genetic diversities of these sister species and do they suggest resilience or susceptibility to

extinction and environmental change? Identifying MUs, including the direction of connectivity between these units, is essential information to ascertain best practice management and maximise biodiversity conservation of these important southernmost coral reefs. I discuss the implications of the underlying genetic structure to extinction risk and the conservation of these remote populations.

3.3. MATERIAL AND METHODS

3.3.1. Study system

McCulloch's anemonefish (*Amphiprion mccullochi*) has the smallest geographic range of any of the 28 species of anemonefish and is endemic to 3 isolated, oceanic locations more than 600 km off the east Australian coastline (Elizabeth Reef, ER; Middleton Reef, MR; and Lord Howe Island, LHI). It is only found living in close association with its host anemone, *Entacmaea quadricolor,* and occurs at depths between 2 and 45 m (Fautin & Allen 1992). Its colouration of black body with a whitish snout, caudal peduncle and caudal fin makes it easily recognizable. Its sister species, *A. akindynos* (Santini & Polacco 2006) is more widespread, ranging from the Great Barrier Reef (GBR) south to the Solitary Islands and extending out to New Caledonia (but not including Elizabeth and Middleton Reefs or Lord Howe Island). It lives amongst its host anemones *Entacmaea quadricolor, Heteractis aurora, H. crispa, H. magnifica, Stichodactyla haddoni* and *S. mertensii* (Fautin & Allen 1992), at depths between 1 m and 25 m (Allen 1991). It has two white bars on its body and a colour transition from a dark brown/orange body to a whitish caudal fin (Fautin & Allen 1992).

3.3.2. Sampling locations and procedures

Finclips from 60 *A. mccullochi* individuals were collected from two out of the three known populations at MR (n = 30; Choat et al. 2006) and from LHI (n = 30; Hobbs et al. 2009a), and preserved in 70% alcohol. These two populations (MR and LHI) represent either end of the entire geographic range of this species, spanning 200 km. Bay et al. (2006) sampled *A. akindynos* at two GBR locations - a central, Lizard Island (LI) population (n = 20) and a southern peripheral, One Tree Island (OTI) population (n = 24), spanning 1200 km. The geographic ranges of the two sister species are separated by at least 600 km of deep open ocean habitat.

3.3.3. Genetic techniques

The non-coding regions of the mtDNA were selectively amplified by Polymerase Chain Reaction (PCR) using the following primers for the damselfish *Acanthachromis polyacanthus*: dLoop F (5′-CATATATGTRTTATCAACATTA-3′) and CR-E H16498R (5′-CCTGAAGTAGGAACCAGATG-3′) (Bay *et al.* 2006).

3.3.4. mtDNA analysis

Sequence data were obtained from GenBank for the following five species which acted as outgroups: *A. clarkii* (DQ343928.1), *A. clarkii* (orange; DQ343929.1), *A. chrysopterus* (DQ343927.1), *A. latezonatus* (DQ343933.1) and *A. leucokranos* (DQ343934.1). All *A. akindynos* sequences (DQ250449.1 to DQ250492.1) from Bay *et al.* (2006) were analysed together with *A. mccullochi* sequences based on the close genetic relationship between these two species (Santini & Polacco 2006). One GBR *A. akindynos* sequence from an unspecified location was also included (DQ343924.1).

jModeltest (Posada 2008) identified an HKY + I model under Akaike Information Criterion with gamma = 0.271. Fifty-three of the 317 nucleotides sequenced for *A. mccullochi* were parsimony informative. The transition (ts) : transversion (tv) substitution ratio was approximately 6 : 1. There were a total of 65 variable sites excluding 5 single base indels. The nucleotide composition was AT-biased with 71.11% AT : 28.89% GC (*A. mccullochi* Dloop), 71.60% AT : 28.40% GC (*A. akindynos* Dloop) and 71.36% AT : 28.64% GC (combined *A. mccullochi* and *A. akindynos* Dloop), which is consistent with fish mitochondrial DNA (McMillan & Palumbi 1997).

Bayesian skyline plots in BEAST V1.6.1 were constructed to evaluate the presence of demographic stability and/or expansions from coalescent analyses using Strict clock model (parameters as above). I also evaluated population stasis using Fu's Fs parameter for population stasis (Fu 1997) and Tajima's test for selective neutrality of mtDNA, D (Tajima 1983) both of which accepted the hypothesis of a static population under an assumption of selective neutrality for *A. mccullochi* (D = 1.578, p = 0.96 and Fs = 2.80, p = 0.85) however, for *A. akindynos* selective neutrality of the mtDNA was accepted (D = 0.103, p = 0.61), whilst population stasis was rejected (Fs = -24.42, $p < 0.0001$), suggesting a spatial expansion by *A. akindynos*. Following 90000 resamplings of the data, the F-statistics (fixation indices, Φ_{st}, Φ_{ct}, Φ_{sc}) were determined using an AMOVA (Excoffier et al. 1992) to detect population genetic partitioning between either regions (LI and OTI vs MR and LHI) or genetically distinct lineages regardless of location or species (ESUs were compared to each other; Table 1). Haplotype (h) and nucleotide diversities (□%) of the data were interpreted based on Grant and Bowen (1998). Population pairwise F_{st} comparisons (measured in Arlequin ver 3.5) initially identified no differences between

locations for either species, which informed further inter- and intra-specific analyses between clades (Table 2). Isolation by distance between reefs was tested using a Mantel Test in IBD ver1.4 with 10000 permutations (Bohonak 2002).

3.3.5. Microsatellite analyses

GENEPOP 4.0 (Rousset 2008) was used to perform exact tests of departures from Hardy-Weinberg equilibrium (HWE) for each locus per sampled location (i.e. 17 loci x 4 sampled locations) and to test for linkage disequilibrium (LD) between the 17 loci within each of the two study species (i.e. 17 x 17 (-17) = 272 tests for each species), using the Markov chain algorithm. If departure from HWE was observed, the program MICRO-CHECKER 2.2.3 (van Oosterhout *et al.* 2004) was used to detect the presence of null alleles, large allele dropout and other scoring errors. I conducted 20 batches with 5000 iterations per batch. A False Discovery Rate (FDR, Benjamini & Hochberg 1995) correction was applied to all HWE and LD results, using the program QVALUE (Storey 2002). Significant single-locus departures from HWE were detected in eight of sixty-eight tests at the population level before FDR correction and five afterwards (Am1, Am11, Am14, Am17, Am19). Null alleles were identified in LI (Am11, Am17, Am19), OTI (Am1, Am7, Am21), MR (Am11, Am14, Am17) and LHI (Am11, Am17, Am19) as indicated in MICROCHECKER. Loci that were not in HWE and had null alleles (i.e. Am1, Am11, Am14, Am17, Am19) were not used in subsequent analysis (ARLEQUIN, STRUCTURE, NEWHYBRIDS and MIGRATE-n). Of 544 locus x locus exact tests of linkage disequilibrium, (272 per species), only thirteen were significant before FDR and none after FDR correction (Benjamini & Hochberg 1995), indicating that loci are independently assorted.

Following 90 000 resamplings of the data, F-statistics (fixation indices, Φ_{st}, Φ_{ct}, Φ_{sc}) were determined using an AMOVA (Excoffier et al. 1992) to detect population genetic partitioning between regions (LI and OTI vs MR and LHI; Table 3.1). Microsatellite genetic diversity (*gd*) of differentiated populations was determined in ARLEQUIN 3.5 (Excoffier *et al.* 2005), using 90 000 permutations (Table 3.2). Isolation by distance was tested using a Mantel Test in IBD ver1.4 with 10000 permutations (Bohonak 2002), as noted for mtDNA analyses.

Given contrasting results found with mtDNA and msatDNA, I specifically tested for interbreeding using two programs: 1) NEWHYBRIDS (Anderson & Thompson 2002), which implements a Bayesian method aimed at detecting the presence of hybrids from a sample of individuals of mixed origin. I used a Markov chain Monte Carlo procedure with a 150 000 burn-in and 150 000 steps and, 2) MIGRATE-n 2.4.3 (http://popgen.sc.fsu.edu/Migrate-n.Html; see Beerli & Felsenstein 2001, Beerli 2004) to estimate long term gene flow between MUs and short term gene flow (nuclear DNA) between locations. I set the migration rate parameter for mtDNA (Theta and M to a maximum of 0.1 and 5 000, respectively) and msatDNA (both Theta and M to a maximum of 100). I conducted 10 replicates of a Bayesian analysis with one long chain sampling every 100 trees of 100 000 sampled and a 20 000 iteration burn-in for mtDNA and a Bayesian analysis with one long chain sampling every 100 trees of 5 000 sampled and a 1 000 thousand iteration burn-in for msatDNA . All parameters converged and fell within the 90% CI yielding values for θ and m for each locus per population. Finally, a discriminant analysis of principal components (DAPC) was also used to determine

demographic gene flow between populations as per Jombart et al. (2010). This required the use of the program R 2.12 (http://www.Rproject.org). I retained 75 principal components (PCs) comprising 95% of the total genetic information as predictors for DA.

3.3.6. Population genetic diversity

As per Chapter 2.3.3.

3.4. RESULTS

3.4.1. Historical relationship

The best outgroup rooted ML phylogram and the MST identified two strongly supported groups and this was supported by all additional phylogenetic analyses (Fig 3.1a). It appeared that either one group emerged from the other or that there was a splitting to form two sister groups, depending on the analysis. The two groups did not represent the two species *per se*, and lacked reciprocal monophyly, because individuals of both species were detected in both groups. Despite the lack of species-specific partitions, there were no shared haplotypes between species. This suggests a complex historical history between these recognised species or incomplete lineage sorting. Sixteen synapomorphic substitutions characterised the split between these two phylogenetic groups (1 and 2). Group 1, which emerges from Group 2, contains 56% of all sampled *Amphiprion mccullochi*, whereas the apparently older Group 2 contains 70% of all sampled *A. akindynos* represented.

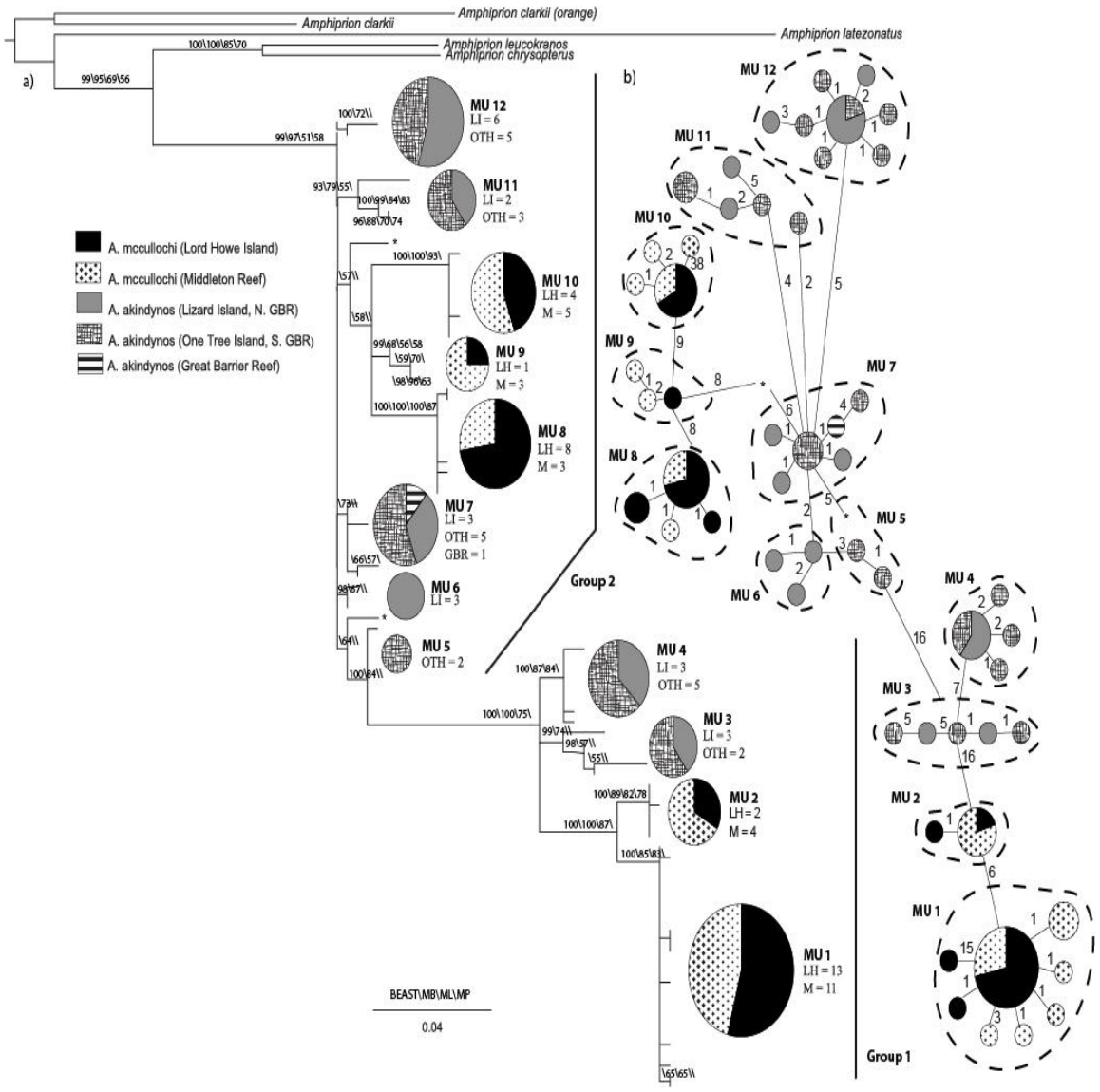

Figure 3.1 a) An outgroup rooted phylogram of D Loop sequences from 50 *Amphiprion mccullochi* individuals from Middleton Reef and Lord Howe Island and 44 *Amphiprion akindynos* individuals from Lizard and One Tree Islands on the Great Barrier Reef. *A. clarkii* was specified as the outgroup species. This represents the best ML tree from 10 individual runs in GARLI. Asterisks identify individuals of *A. akindynos* that fall outside a clade. Numbers on branches indicate support for each clade, BEAST, MB, ML and MP. b) Haplotype minimum spanning tree (MST) with number of substitutions between haplotypes indicated on connectors. Different shading represents each of the four locations as shown on the key to the figure

3.4.2. Management Units (MUs)

The two groups contained a total of seven and possibly as many as twelve Management Units (MUs), of which three to four were in Group 1 (MU1 to 4) and four to eight (MU5 to 12) were in Group 2. Haplotype sharing was observed between locations within species: MR – LHI in all five *A. mccullochi* MUs; LI – OTI in five of the seven *A. akindynos* MUs. When samples were structured by MU irrespective of sampling location (MU 1 to 12) during population genetic analyses, pairwise F_{st} values indicating significant genetic differentiation were obtained between all but two MU pairs (pairwise F_{st} = 0.455 to 0.983, $p < 0.00001$ to 0.04800). Exceptions were MU5 - 6 (pairwise F_{st} = 0.634, $p = 0.10$) and MU5 - 9 (F_{st} =0.860, $p = 0.10$). Lack of significance among these MU-pairs, despite very large pairwise F_{st} values may result from smaller sample sizes in these three clades than in the remaining nine clades (see Figure 3.1a). AMOVA results confirmed genetic partitioning between MUs, Φ_{st} =0.894 ($p < 0.001$), however, this only explained 10.56% of the variation, whilst 49.56% and 39.88% of the variation occurred among Groups (1 vs 2) and among populations within Groups respectively Φ_{ct} =0.496, $p < 0.001$; Φ_{sc} = 0.791, $p < 0.001$; Table 3.1).

I further examined levels of gene flow between locations and/or species using pairwise F_{st} values for a number of inter-species and mixed-species comparisons, consistent with the phylogenetic signals described above. There was no mtDNA differentiation between sample locations (LHI vs MR or OTI vs LI) within either species (*A. mccullochi* or *A. akindynos*, F_{st} = -0.005 and -0,025, $p = 0.414$ and 0.865, respectively). Given this and

disregarding species identity, to compare fish from offshore to fish from continental shelf locations (OTI or LI vs LHI or MR), all combinations of comparisons showed significant pairwise mtDNA F_{st} values (F_{st} = 0.293 to 0.349, $p < 0.00001$). The more statistically rigorous AMOVA of samples structured by region (LHI and MR vs OTI and LI) confirmed mtDNA genetic partitioning with more than two thirds (68.31%) of the genetic variation within locations, Φ_{st} =0.317 ($p < 0.0001$), and less than one third of the variation detected among regions (offshore vs GBR continental locations), which was not significant, Φ_{ct} = 0.327 (p = 0.341, Table 3.1). Further, none of the variation occurred among locations within regions (i.e. LHI vs MR; OTI vs LI), Φ_{sc} = -0.014 (p = 0.596). There was no isolation by distance (IBD) based on mtDNA from all locations sampled, using a Mantel's test of pairwise geographic (km) and genetic (F_{st}) distance between locations (z = 2079.08, R^2 = 0.31, p = 0.337).

Both Fu's Fs parameter for population stasis (Fu 1997) and Tajima's test for selective neutrality of mtDNA, D (Tajima 1983) accepted the hypothesis of a static population under an assumption of selective neutrality for *A. mccullochi* (D = 1.578, p = 0.96 and Fs = 2.80, p = 0.85). In contrast, population stasis was rejected for *A. akindynos* (D = -24.42, $p < 0.0001$; Fs = -11.757, p = 0.02), suggesting that this species has undergone spatial expansion. This was confirmed by both Bayesian and mismatch distribution analyses (data not shown). When mtDNA data of specimens from all locations, regardless of species partition, were considered collectively (MR vs LHI vs OTI vs LI), total mtDNA suggested selective neutrality (D = 0.103 to 0.636, p = 0.61 to 0.80) and appeared to be

either evolutionarily stable or to represent admixture of previously differentiated lineages.

Table 3.1: AMOVA analysis for the genetic data from *Amphiprion mccullochi* and *A. akindynos* structured into a) mtDNA - partitioned by geographic region (GBR vs offshore locations) and b) mtDNA – partitioned by ESU's and c) msat DNA – partitioned by geographic region (GBR vs offshore locations).

Source of variation	d.f	SSD	Variance component	Percentage of variation	F-statistics fixation indices (p-value)
a) Region					
Among groups	1	170.34	3.53	32.63	F_{ct} = 0.326 (0.341 ± 0.014)
Among populations within groups	2	10.10	-0.10	-0.93	F_{sc} = -0.014 (0.596 ± 0.011)
Within populations	90	665.51	7.39	68.31	F_{st} = 0.317 **(<0.001 ± 0.000)**
b) Clades					
Among groups	1	352.29	6.38	49.56	F_{ct} = 0.496 **(0.007 ± 0.003)**
Among populations within groups	10	367.64	5.13	39.88	F_{sc} = 0.791 **(<0.001 ± 0.000)**
Within populations	80	108.74	1.36	10.56	F_{st} = 0.894 **(<0.001 ± 0.000)**
c) Microsatellite					
Among groups	1	30.54	0.25	5.01	F_{ct} = 0.05 (0.337 ± 0.002)
Among populations within groups	2	13.684	0.04	0.85	F_{sc} = 0.009 (0.006 ± 0.000)
Within populations	192	917.64	4.78	94.14	Fst = 0.05 **<0.001 ± 0.000)**

3.4.3. Demographic connectivity

Msat pairwise F_{st} values were largely consistent with mtDNA results: There was no genetic differentiation between *A. mccullochi* populations, LHI-MR (F_{st} = 0.005, p = 0.081), but significant differentiation existed between *A. akindynos* populations, LI and OTI (F_{st} = 0.016, p = 0.036). Importantly, the more statistically rigorous AMOVA of samples structured by region (LHI and MR vs OTI and LI) identified msat genetic partitioning within locations, accounting for 94.14% of the genetic variation; Φ_{st} =0.06 (p < 0.001). Only 5.01% of the variation occurred among regions (offshore vs GBR continental locations), but this was not significant, Φ_{ct} = 0.05 (p = 0.337), and almost none (0.85%) of the variation, albeit significant, occurred among locations within regions (i.e. LHI vs MR, OTI vs LI), Φ_{sc} = 0.009 (p = 0.006, Table 3.1). In contrast to mtDNA, there was strong evidence of isolation by distance (IBD) based on msat DNA using a Mantel's test of log pairwise geographic (km) and log genetic (F_{st}) distance between all sampled locations, regardless of species (z = -23.11, R^2 = 0.89, p = 0.048).

NEWHYBRIDS and STRUCTURE analyses indicated intra-specific but not inter-specific gene flow between species and the likelihood of the marginal posterior probability distribution was greatest when K = 2 (Fig 3.2a, b). DAPC was consistent with this and *A. mccullochi* populations (MR and LHI) appeared to have overlapping genotypic profiles, as per mtDNA, while *A. akindynos* populations (OTI and LI on the GBR) were genetically distinct from each other (Fig 3.2c), unlike mtDNA results. Using

the four sampled populations as *a priori* population criteria, DAPC assigned 96.94% of all individuals to the population where they were sampled (assignment per population, MR = 80%, LHI = 80%, LI = 100% and OTI = 100%). MIGRATE-n indicated high levels of historical gene flow relative to demographic gene flow. Levels of historical gene flow between lineages (i.e. Group 1 vs Group 2; 4Nm values ranged from 28 to 62) were an order of magnitude lower compared to within lineage geneflow: Group 1 (MU 1.1 to 1.4; 4Nm values ranged from 980 to 2 326) and Group 2 (MU 2.1 to 2.8; 4Nm values ranged from 241 to 1 247). Demographic gene flow was a few orders of magnitude less than historical gene flow, with 4Nm values ranging from 1 to 6.

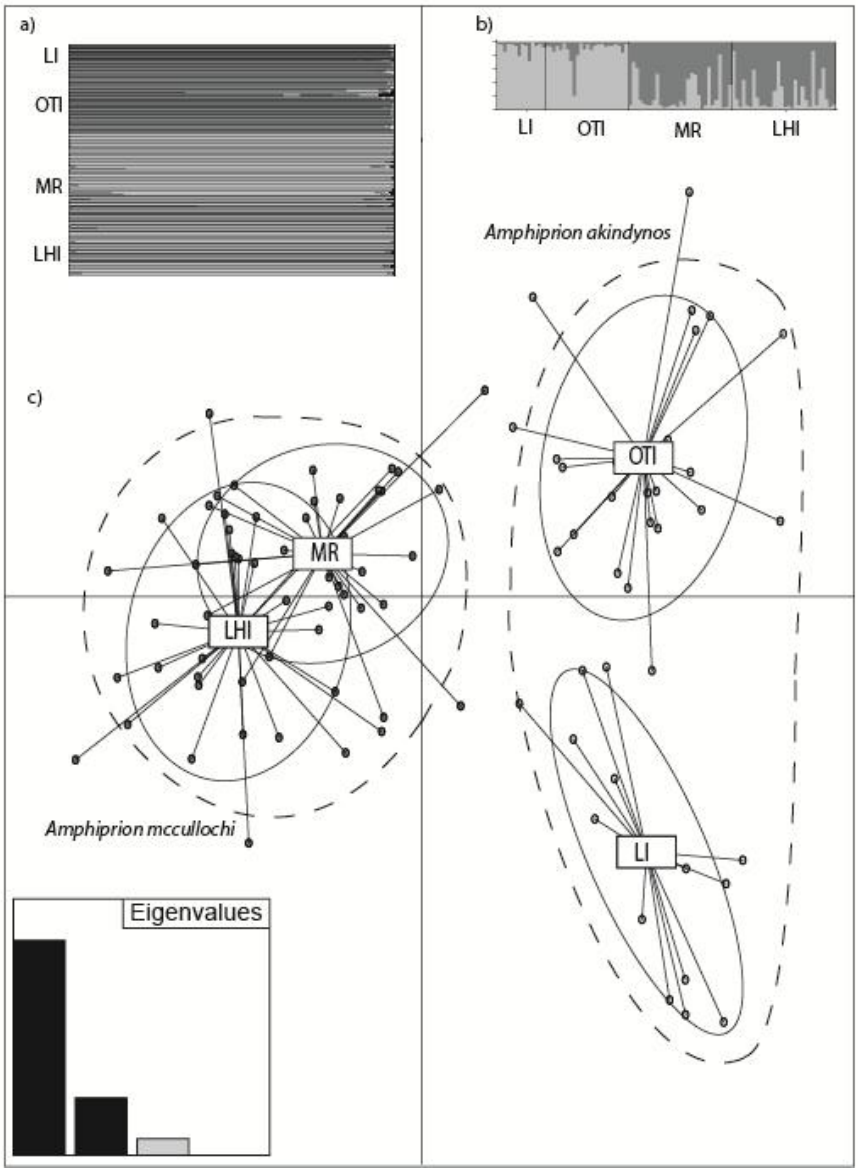

Figure 3.2 Separation of *Amphiprion mccullochi* and *A. akindynos* based on various analyses of msat loci: a) NEWHYBRID analysis showing pure (grey) and F1 hybrid (black) status, b) STRUCTURE cluster analysis and c) Scatterplots of the discriminant analysis of principal components (DAPC) of msat data for two *A. mccullochi* and two *A. akindynos* populations using geographic sample site as priors for genetic clusters. Populations are named and individual genotypes appear as dots surrounded by 95% inertia ellipses. Eigenvalues show the amount of genetic information contained in each successive principal component with X and Y axes constituting the first two principal components, respectively.

3.4.4. Genetic diversity

Genetic diversity was high for both species, evident from high haplotype (*h*), nucleotide (%π) and genotypic (*gd*) diversities for each sampled location (*h* = 0.846 to 0.982, %π = 3.97 to 5.11, *gd* = 0.670 to 0.718, for each species). Total haplotype, nucleotide and genotypic diversity were also high when species were combined (*h* = 0.905 to 0.976, %π = 4.03 to 5.67 and *gd* = 0.690 to 0.706; Table 3.2). When samples were grouped according to mtDNA groups, irrespective of species or location, haplotype diversity was still high (*h* = 0.600 to 1.000), but nucleotide diversity was at least one third (%π = 0.11 to 1.76) that of the afore-mentioned values for each species (Table 3.2).

Table 3.2: Genetic diversity estimates for *Amphiprion mccullochi* and *A. akindynos*. Sample size (n), number of haplotypes (n_h), haplotypes diversity ± SE (*h*), nucleotide diversity ± SE (%π) and genetic diversity (*gd*). Genetic diversity estimates for one *A. akindynos* sample of unknown location on the GBR was omitted.

Site	**n**	**n_h**	**h**	**%π**	***gd***
A. mccullochi					
Middleton Reef	22	14	0.952 ± 0.026	4.90 ± 2.53	0.691 ±0.360
Lord Howe Island	26	11	0.846 ± 0.054	5.11 ± 2.62	0.670 ± 0.351
Total	48	21	0.905 ± 0.027	5.00 ± 2.50	0.690 ± 0.370
Clade 1.1	22	9	0.701 ± 0.103	1.76 ± 0.98	
Clade 1.2	6	3	0.600 ± 0.215	0.11 ± 0.14	
Clade 2.4	11	4	0.600 ± 0.154	0.22 ± 0.20	
Clade 2.5	3	3	1.000 ± 0.272	0.63 ± 0.60	
Clade 2.6	8	4	0.643 ± 0.184	0.24 ± 0.22	
A. akindynos					
Lizard Island	20	14	0.952 ± 0.026	4.90 ± 2.53	0.718 ± 0.371
One Tree Island	24	11	0.846 ± 0.054	5.11 ± 2.62	0.693 ± 0.360
Total	44	34	0.976 ± 0.013	4.03 ± 2.06	0.706 ± 0.484
Clade 1.3	5	5	1.000 ± 0.127	1.46 ± 1.00	
Clade 1.4	8	4	0.643 ± 0.184	0.40 ± 0.32	
Clade 2.1	2	2	1.000 ± 0.500	0.32 ± 0.45	
Clade 2.2	3	3	1.000 ± 0.272	0.63 ± 0.60	
Clade 2.3	8	6	0.893 ± 0.111	0.76 ± 0.52	
Clade 2.7	5	4	0.900 ± 0.161	1.20 ± 0.85	
Clade 2.8	11	7	0.818 ± 0.119	0.56 ± 0.40	
A. mccullochi* and *A. akindynos					
Total	93	55	0.969 ± 0.009	5.67 ± 2.81	0.691 ± 0.370

3.5. DISCUSION

Phylogenetic analysis of mtDNA revealed a complex historical history where *A. mccullochi* and the more widespread *A. akindynos* were not partitioned into monophyletic lineages as anticipated, but were mixed into two paraphyletic lineages that did not correspond to either species or geographic location *per se*. These analyses also revealed that the species did not share any haplotypes with each other and identified up to 12 species-specific MUs distributed among the four geographic locations sampled for the two species. Population genetic analyses based on mtDNA confirmed the absence of genetic partitioning by location for both species examined. Importantly, despite this complex historical history, msat DNA analyses revealed no demographic gene flow existed between species, and geneflow between populations was limited. Populations connected by such low levels of gene flow are effectively demographically independent and significant to conservation as these MUs are fundamental to effective demographic management and are the logical unit for demographic study and population monitoring (Moritz 1994). Finally, the high genetic diversity of both species may result in greater adaptive capacity to cope with future environmental change, insofar as greater genetic diversity provides more "raw material" for selection to act on.

3.5.1. Historical relationship

There are three possible interpretations of the combined genetic results. Firstly, the two species represent different colour variants of a single species rather than different species *per se*. However, this contradicts existing taxonomic classification (Santini & Polacco 2006) and requires further ecological and experimental data to explore this. The second possibility is that incomplete lineage sorting may be responsible for the apparent lack of reciprocal mtDNA

monophyly. However, as mtDNA lineages suffer incomplete lineage sorting for a much shorter period of time (25%) than do nuclear DNA lineages, our msat DNA results do not support this scenario, because msat DNA did partition the species. Which leads to the third possible interpretation - that there has been historical bi-directional hybridisation.

I argue that the high level of mtDNA genetic partitioning observed - whilst not identifying species-specific lineages or location-specific MUs - is most consistent with an historical history of at least two reticulate events between *A. akindynos* and *A. mccullochi*. Reticulate events have been documented for numerous other reef fish species (e.g. McMillan et al. 1999, van Herwerden & Doherty 2006, van Herwerden et al. 2006, Yaakub et al. 2006, Marie et al. 2007), including secondary contact between differentiated lineages of *A. akindynos* on range edges as documented by Bay and Caley (2011). Hybridisation and reticulate evolution is also common in other coral reef organisms (e.g. corals: Willis et al. 2006). Several lines of evidence support a scenario of historical hybridisation: (i) two paraphyletic mtDNA lineages exist, each comprising of both *A. mccullochi* and *A. akindynos*, where each lineage may represent a mtDNA lineage of one of the species prior to historic hybridization; (ii) there are no shared mtDNA haplotypes between present day populations of the two species; demographic geneflow is very limited between the species, as measured by msat DNA and Migrate-n; (iii) there is population genetic partitioning between the two species, evident from both mtDNA and msatDNA, even if treated as populations rather than species and; (iv) msat genotypes of almost 97% of individuals were assigned back to the population they were collected from and there was strong and total partitioning between species following DAPC analysis of msat DNA. Together, this suggests that

geneflow between species occurred in the historical past but is either no longer occurring, or is happening at a level not detectable in this study.

Historical gene flow between species via hybridisation is increasingly being documented among coral reef fishes (e.g. Yaakub et al. 2006, Hobbs et al. 2009b, Hobbs et al. in press) and has been reported in a number of anemonefish species, including *A. akindynos* (Sea Read 2009, Fautin & Allen 1997). Timm et al. (2008) reports hybridization between *Amphiprion* species, as a possible explanation for sequence sharing, especially since, within the genus *Amphiprion*, several species have similar colouration and overlapping variation at otherwise diagnostic morphological characters. Colouration is very pertinent in the present case, as juvenile *A. akindynos* and *A. mccullochi* are almost morphologically indistinguishable (Richardson 1998). Settlement to an anemone occurs during the juvenile phase and given that both species use the same host anemone (*Entacmaea quadricolor*; Fautin & Allen 1992) and have similar colouration, this could lead to the formation of heterospecific social groups, and possible interbreeding. Furthermore, although the two species occur in allopatry, two vagrant individuals of *A. akindynos* have been recorded at LHI (Crean et al. 2009). The arrival of vagrants into the distributional range of an allopatric sister species is thought to promote interbreeding between species due to the low availability of conspecific partners for the vagrants (Hobbs et al. 2009b). The East Australian Current flows through the ranges of the two study species and the strength and direction is influenced by climate (Philander et al. 1990, Middleton et al. 2009). Historical changes in this current may have facilitated the arrival of vagrants, which may have resulted in contact and hybridisation between the two species. Taken together, all these factors suggest hybridisation is a likely scenario in the historical history of these two species.

3.5.2. Management Units (MUs)

Two groups (evolutionary units) and at least seven MUs were identified in this study, with at least two *A. akindynos* MUs and five *A. mccullochi* MUs. While each MU represents one of the study species exclusively, most are not partitioned by geography. This complex underlying phylogenetic structure may have occurred because anemonefish have both self-recruitment at demographic scales (Jones et al. 2005) and inter-population connectivity at longer historical timescales (Timm 2008).

3.5.3. Demographic connectivity

Microsatellite loci detected very limited geneflow between the sister species *A. akindynos* and *A. mccullochi* which strengthens our suggestion of historical hybridisation as current day populations are not mixing. Likewise, Lizard Island and One Tree Island *A. akindynos* populations were genetically distinct which is consistent with Bay and Caley (2011), indicating that the distance between these two populations may be too great for geneflow to occur. In contrast, there was a lack of genetic partitioning for *A. mccullochi* populations. The difference between species in the level of gene flow between populations probably represent different geographic distances between the sample populations (McCulloch populations were 160km apart and the akindynos sample populations were 1200km apart). This gene flow between populations is promising for *A. mccullochi* since it facilitates recolonisation if one population was to go locally extinct. However, only two locations in each species were sampled and further samples of *A. akindynos* from the Great Barrier Reef, southern reefs (e.g. Solitary Islands), and New

Caledonia, as well as samples of *A. mccullochi* from Elizabeth Reef will be needed to fully quantify gene flow between all populations of these species.

3.5.4. Genetic diversity

In both species mtDNA data showed high genetic variability, *h* almost double and □ close to an order of magnitude greater than the cut-off defined by Grant and Bowen (1998; both *h* and □ > 0.5). Similarly, high levels of genetic diversity for msatDNA were also found in both species. This suggests that populations of both species are either large and stable with long historical histories, or that there has been secondary contact between differentiated lineages (Grant & Bowen 1998). Given that *A. mccullochi* does not have a large population (Choat et al. 2006, Hobbs & Feary 2007), secondary contact between differentiated lineages is more likely and is consistent with reticulate evolution between these sister species. Given that mtDNA diversity tracks with nuclear genetic diversity in many marine species (reviewed by Johannesson & Andre 2006), such high genetic diversity in *A. mccullochi* and *A. akindynos* is encouraging as it suggests that both species may have a greater adaptive capacity to deal with environmental change than if they had low genetic diversity. However, a cautious approach is still warranted given that quantitative trait loci under selection, can have no genetic diversity in peripheral populations despite high neutral genetic diversity (Kellermann et al. 2009).

Conserving genetic diversity is considered a priority by the IUCN (McNeely et al. 1990), and even more so for *A. mccullochi* given its vulnerability to extinction due to other traits (low abundance, small geographic range, and ecological specialisation). The high genetic diversity of *A. mccullochi* appears to have arisen through complex reticulate evolution involving historical

(but not demographic) hybridisation. High genetic diversity and population connectivity may reduce extinction risk in *A. mccullochi*, but identifying and protecting populations implicated in generating the complex reticulate structure amongst this species should be a conservation priority. Conserving *A. mccullochi* would best be achieved by protecting each of its MUs and by minimising threats to population size, such as habitat loss (e.g. anemone bleaching) and collection for the aquarium trade.

CHAPTER 4: **Genetic connectivity among and self-replenishment within island populations of a restricted range subtropical reef fish**

(Published as: van der Meer MH, Hobbs J-PA, Jones GP, van Herwerden L (2012) Genetic connectivity among and self-replenishment within island populations of a restricted range subtropical reef fish. *PLoS ONE*, 7(11), e49660. doi:10.1371/journal.pone.0049660.)

4.1. ABSTRACT

Marine protected areas (MPAs) are increasingly being advocated and implemented to protect biodiversity on coral reefs. Networks of appropriately sized and spaced reserves can capture a high proportion of species diversity, with gene flow among reserves presumed to promote long term resilience of populations to spatially variable threats. However, numerically rare small range species distributed among isolated locations appear to be at particular risk of extinction and the likely benefits of MPA networks are uncertain. Here I use mitochondrial and microsatellite data to infer historical and demographic gene flow among isolated locations as well as levels of self-replenishment within locations of the endemic anemonefish *Amphiprion mccullochi*, restricted to three MPA offshore reefs in subtropical East Australia. I infer high levels of gene flow and genetic diversity among locations over historical time, but limited demographic gene flow amongst locations and high levels of self-replenishment (68 to 84%) within locations over demographic time. While long distance dispersal explained the species' integrity in the past, high levels of self-replenishment suggest locations are predominantly maintained by local replenishment. Should local extinction occur, demographic rescue effects through large scale connectivity are unlikely. For isolated islands with large numbers of endemic

species, and high local replenishment, there is a high premium on local species-specific management actions.

4.2. INTRODUCTION

It is widely accepted that life within the world's oceans, especially within highly diverse coral reefs, is under an increasing threat in the 21st century (Steffen et al. 2007). New management strategies are being developed in a bid to protect marine life from a range of anthropogenic impacts (Botsford et al. 1997, Pikitch et al. 2004). One of the most popular approaches has been the establishment of no-take Marine Protected Areas (MPAs) whose efficacy in conserving biodiversity continues to be debated. While appropriately designed MPA networks can encompass a high proportion of species (Hooker & Gerber 2004) and genetic diversity Miller & Ayre 2008), the degree to which reserves contribute to the long term persistence of locations and maintain natural evolutionary processes is uncertain. A major factor that dictates how well MPAs work, is the extent of larval connectivity among locations (Sale et al. 2005), including links between protected and unprotected areas and among different nodes in MPA networks (Planes et al. 2009, Harrison et al. 2012).

Historically, the pelagic larval stage of most marine species was thought to result in broad scale larval dispersal aided by ocean currents (Roberts 1997). This holds true over historical time scales where the occasional long distance dispersal of pelagic larvae acting as agents of gene flow, have connected distant locations (Graham et al. 2008), maintained high levels of genetic diversity (Klanten et al. 2007, Horne et al. 2008) and thereby helped reduce a species risk of extinction (Frankham 1998). However, a growing number of studies focusing on demographic time scales show high levels of self-recruitment (Planes et al. 2009, Harrison et al. 2012, Jones et al. 2009). Although none of these studies show 100% self-recruitment, and the scales of demographic connectivity are only just beginning to be assessed (Berumen et al. 2012), this

finding suggests that the appropriate scale and distance between MPAs may indeed be smaller than previously assumed (Harrison et al. 2012, James et al. 2002, Gerlach et al. 2007). Thus, connectivity operates over two time scales: historical and demographic. Most traditional population genetic studies infer historical connectivity (Lowe & Allendorf 2010) (gene flow) using mtDNA to capture the longer term signals of dispersal Postma & van Noordwijk 2005). In recent years, a range of new statistical software (e.g. STRUCTURE, Pritchard et al. 2000; DAPC, Jombart et al. 2010; Migrate-n, Beerli & Felsenstein 2001; BAYESASS, Wilson & Rannala 2003) has become available and is increasingly being applied to population genetic studies (Hellberg 2007) to infer demographic connectivity using msatDNA to capture the shorter term signals of dispersal. Sometimes there is a 'lack of congruence' between connectivity operating over different time scales (historical and demographic). For example, coral trout (*Plectropomus maculatus)* and stripey snapper (*Lutjanus carponotatus*) lack spatial genetic structure along the GBR (spanning more than 1000 km) on historical time scales using mtDNA (Evans et al. 2010), while comprehensive parentage analyses of very large sample sizes at much smaller spatial scales using msatDNA, identified high levels of both local- and self-recruitment for both species (Harrison et al. 2012). The lack of spatial genetic structure in mtDNA sequence data is not uncommon (Eble et al. 2011, Leis et al. 2011) since only a few recruits per generation are sufficient to maintain spatial genetic homogeneity on historical time scales (Shulman 1998, Planes 2002). Together these tools and analyses (mtDNA and msatDNA) are useful because they provide a more holistic picture of connectivity (gene flow) and retention over a range of spatial and temporal scales (Miller & Ayre 2008, Willis et al. 2003, Anderson & Meikle 2010, Burbrink 2010, Harris & Taylor 2010, Horne et al. 2011, Scascitelli et al. 2010).

The evolution of island faunas is interesting because they are clearly punctuated with historical periods of colonization and gene flow, evidenced by the wide distribution of the same species across isolated locations (Ciofi et al. 2006), yet they have presumed low levels of demographic gene flow. Isolated islands are a conservation priority due to their high level of endemism and high rates of extinction (Whittaker & Fernandez-Pealacios 2007). Species endemic to isolated islands have an increased risk of extinction because they often exhibit a number of vulnerable biological (e.g. flightlessness) (Carlquist 1974, Ciofi et al. 2006, Whittaker & Fernandez-Pealacios 2007) ecological (e.g. small populations, habitat specialists) (Munday 2004) and genetic traits (e.g. low gene flow and genetic diversity) (Cowen et al. 2002). Management plans identify endemic species as a conservation priority; however, effective protection of vulnerable species requires estimates of gene flow (historical and demographic) between isolated locations and estimates of genetic diversity (Drechsler et al. 2003). Likewise, effective management strategies need to conserve both species and genetic diversity in order to maximise ecosystem and population resilience (Johannesson & Andre 2006). Conservation of genetic diversity is an IUCN priority (McNeely et al. 1990) as it provides the raw material for the maintenance of species over historical time scales and provides a basis for responses to rapid environmental change and natural selection (McNeely et al. 1990, Johannesson & Andre 2006, Bell & Okamura 2005) where a reduced genetic diversity has been correlated with decreased fitness (Hoelzel et al 2002).

In this study I examine historical and demographic levels of gene flow in the McCulloch's anemonefish (*Amphiprion mccullochi*), an endemic to three isolated locations in the South-West Pacific Ocean, 600 km off Australia's east coast (Fig 4.1a − e). Due to this species being found

at only three locations, I was able to sample all known locations leaving no 'ghost' populations un-sampled, giving us a high level of confidence and statistical power in our estimates of gene flow. This species is important as it is potentially at risk of extinction because (i) its geographic range is among the smallest for coral reef fishes (ii) it's an extreme habitat specialist due to its obligate relationship with only one host species of anemone (Fautin and Allen 1992) and; (iii) throughout its range it has very low abundance (Chaot et al. 2006, Hobbs et al. 2007, Hobbs et al. 2009), except for an extremely small area of habitat at Lord Howe Island Lagoon (LHIL), which supports 92% of the world's *A. mccullochi* population (Hobbs et al. 2009). Lord Howe Island is a World Heritage Area because it accommodates significant ongoing biological and ecological processes in the development and evolution of coastal, terrestrial, freshwater and marine ecosystems (Heath 2008). The island is an endemism hotspot and contains significant habitats for *in-situ* conservation of biological diversity, including threatened species of exceptional conservation value (Heath 2008). The efficacy of reserves to reduce extinction risk will depend on historical and demographic levels of gene low among these isolated locations. I was particularly interested in whether the high abundance of McCulloch's anemonefish in the lagoon at Lord Howe Island (LHIL) will act to export migrants and help replenish other low abundance locations at greater risk of local extinction.

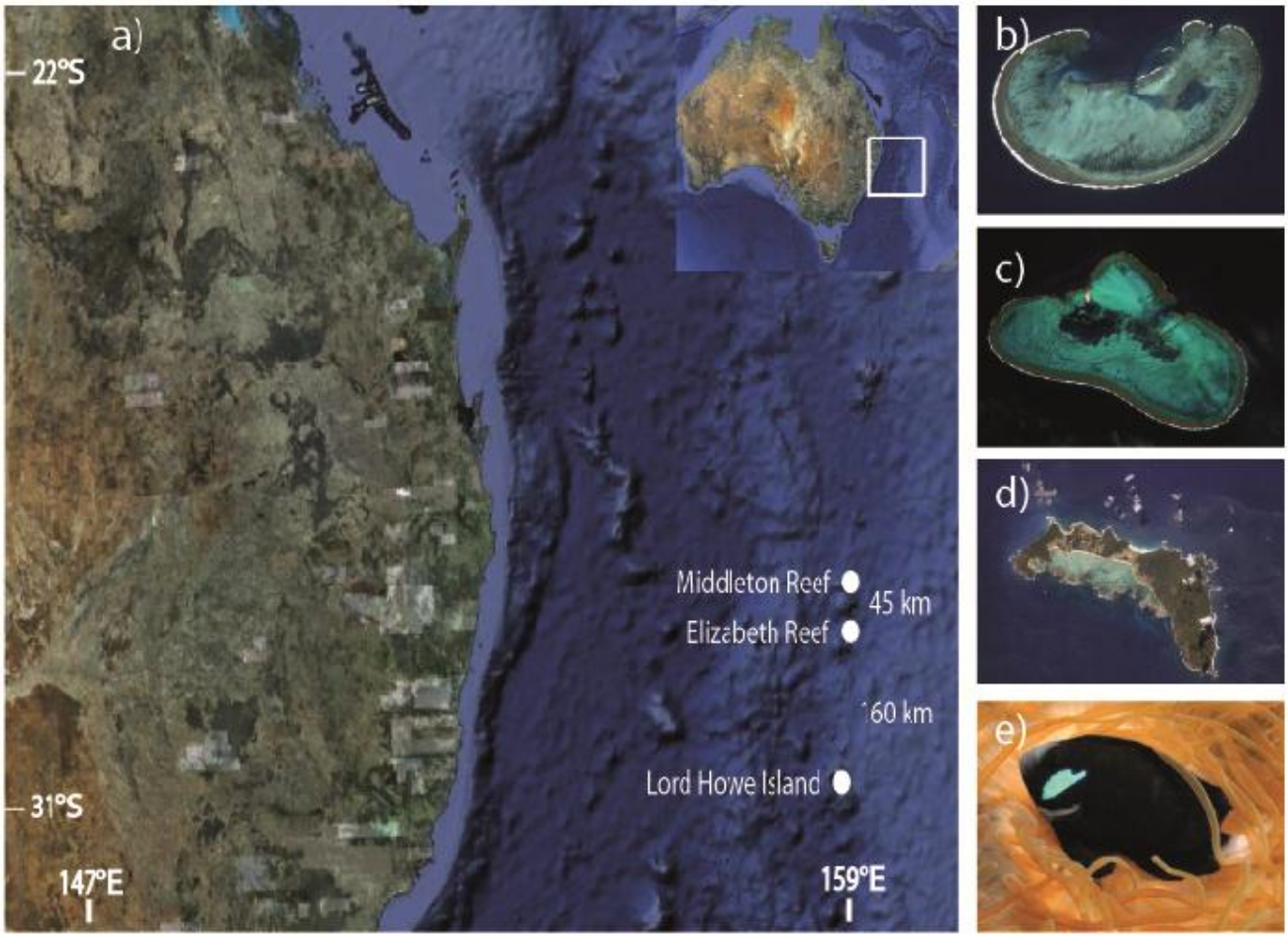

Figure 4.1 Location maps and focal species. (A) Goole Earth image of eastern Australia showing Middleton Reef (MR), Elizabeth Reef (ER) and Lord Howe Island (LHI) in the Southwest Pacific Ocean, to the southeast of the Great Barrier Reef. Aerial photographs of MR (B); ER (C) and LHI (D), indicating both the outside (LHI) and Lagoon (LHIL) sample sites. (E) *Amphiprion mccullochi* in its host anemone *Entacmaea quadricolor* (Photo courtesy of Justin Gilligan).

The aims of this study are fourfold: (i) to determine the patterns and levels of gene flow between locations over historical time scales; (ii) to determine the patterns and levels of gene flow between locations over demographic time scales; (iii) to infer levels of self-replenishment (as a proxy for self-recruitment) and recent migration (iv) to measure population genetic diversities at all locations as an indicator of potential resilience of populations to environmental change and extinction.

4.3. MATERIAL AND METHODS

118 *A. mccullochi* fin clips were taken from four locations, MR (n = 30) (Choat et al. 2006) ER (n = 25) (Choat et al. 2006) outside the lagoon at LHI (LHI, n = 33) and within the LHI Lagoon (LHIL, n = 30) (Hobbs et al. 2009).

4.3.1. Gene flow between locations - historical time scales

The mtDNA phylogenetic analysis. I assigned well supported distinct phylogenetic lineages as management units (MU) (Moritz 1994). A MU is a population that lacks reciprocal monophyly for mtDNA haplotypes, yet has divergent haplotype frequencies [55], as found here.

Quantifying the level of historical gene flow. Due to the previously identified secondary contact between *A. mccullochi* and *A. akindynos* (van der Meer et al. 2012a, Chapter 3) and since MU were not differentiated geographically, both the Stepping-stone and Island-n migration models were not appropriate as priors for the dataset; rather Migrate-n input files had to be modified and customised. I split the mtDNA data in three ways (i) two groups representing the two admixed lineages: Group 1 (MU 1 - 2) and Group 2 (MU 3 - 5) to estimate historical migration between lineages; migration was then compared within Groups (ii) between MU 1 and 2 in Group 1 and; (iii) between MU 3, 4 and 5 in Group 2. I set the datatype to an F84 mutation model and the migration rate parameters for mtDNA (θ and M to a maximum of 0.1 and 1000, respectively) to conduct Bayesian analysis using one long chain that sampled every 100th of 100 k sampled trees and applied a 20 k iteration burn-in. All parameters converged and fell within the 90% CI yielding values for θ and M (mutation-scaled migration rate) per location.

4.3.2. Gene flow between locations - demographic time scales

Quantifying the level of demographic gene flow. I set the datatype to Microsatellite (a simple electrophoretic ladder model with stepwise mutation) and the migration rate parameters for msatDNA (θ and M were both set to a maximum of 100) to conduct Bayesian analysis using one long chain that sampled every 100th of 100 k sampled trees and applied a 20 k iteration burn-in. All parameters converged and fell within the 90% CI yielding values for θ and M (mutation-scaled migration rate) for each locus per location.

4.3.3. Inferred levels of self-replenishment and recent migration

I used BAYESASS v3 to estimate both self-replenishment (as a proxy for self-recruitment) and recent migration between locations; with a MCMC chain, consisting of a total of 11 M steps, a 2 M step burn in and a sampling interval of 100 k, with prior values for migration rate, allele frequency and inbreeding coefficient of 0.95, 0.95 and 0.95, respectively. These priors were selected because they gave acceptance rates of between 20 and 40% (Faubet et al. 2007). Ten separate runs assessed convergence of the MCMC to evaluate consistency of the results obtained from these inferences.

4.3.4. Population genetic diversities

As per Chapter 2.3.3.

4.4. RESULTS

Summary statistics. Three hundred and twenty-two base pairs of mtDNA D-loop were resolved for 105 *Amphiprion mccullochi* individuals. There were a total of forty-six polymorphic sites, of which forty were parsimony informative (six singletons). Allelic diversity was lowest at LHI-L and highest at LHI, whilst F_{IS} did not differ significantly across the three regions surveyed (F_{IS} = 0.07, p = 0.97; Table 4.S1). Elizabeth Reef had the most private alleles, 13 across 17 loci, while the remaining three populations had 12 private alleles each across all loci (Table 4.S1). Of the 17 msatDNA loci: (i) significant single-locus departures from HWE were detected in nine of sixty-eight tests at the population level before FDR correction and two afterwards (LHIL: Am1; ER: Am11); similarly, seven single-locus HWE departures were detected at the regional level before FDR and six afterwards (Table 4.S1); (ii) null alleles were identified in ER (Am6, Am7, Am11, Am19), MR (Am11, Am17), LHI (Am4, Am7) and LHI-L (Am11, Am19) and (iii) of 544 locus x locus exact tests for linkage disequilibrium (136 per population), only 17 were significant before FDR and one after FDR correction (Am6) (Benjamini and Hochberg 1995). Loci that were not in HWE and had null alleles (i.e. Am1, Am11, Am14, Am17, Am19) were not used in subsequent analyses (ARLEQUIN, STRUCTURE, and MIGRATE-n). Detailed summary statistics, mtDNA and msatDNA AMOVA between regions, msatDNA AMOVA by loci, pairwise population comparisons and genetic diversity indices are presented in Supporting Information (4.S1, 4.S2, 4.S3, 4.S4 and 4.S5 respectively).

4.4.1. Gene flow between locations - historical time scales

Synopsis: *A. mccullochi* mtDNA suggested the existence of two historical lineages (Groups) consisting of a total of five MU with each location being represented in all MU (Fig 4.2 a, b). High levels of historical gene flow were found between spatially intermixed MU but this was reduced between Groups 1 and 2, which themselves were also spatially intermixed. This suggests that historical gene flow exists between all locations occupied by *A. mccullochi*. The relative percentage of each geographic location within different MU suggests geographic structure and should guide future population monitoring and demographic studies to better inform management.

The mtDNA phylogenetic analysis (Fig 4.2a) showed two major groups and five distinct management units (MU): MU 1 (n = 41), MU 2 (n = 15), MU 3 (n = 12), MU 4 (n = 18) and MU 5 (n = 12) with a total of 30 haplotypes (Fig 4.2b). All locations were relatively evenly represented within the two groups: Group 1 (MR =23, ER = 21, LHI = 32, LHIL = 25) and Group 2 (MR = 21, ER = 21, LHI = 29, LHIL = 29). However, some locations had markedly different proportional representation within some MU (in bold) compared to others: MU 1 was relatively evenly represented by all locations (MR = 21, ER = 26, LHI = 24, LHIL = 29), but the remaining four MU (2, 3, 4 and 5) differed in representation of individuals from specific locations - MU 2 was LHI dominated (MR = 27, ER = 7, **LHI = 52**, LHIL = 13); MU 3 was under-represented by ER individuals (MR = 31, **ER = 7**, LHI =31 , LHIL = 31); MU 4 was ER dominated (MR = 17, **ER = 50**, LHI = 25, LHIL = 8) and MU 5 was LHIL dominated (MR = 18, ER = 12, LHI = 29, **LHIL = 41**). This indicates that three of the MU (2, 4 and 5) are overrepresented by three specific locations − LHI, ER and LHIL, respectively. In contrast, MR individuals were relatively evenly distributed across all five MU.

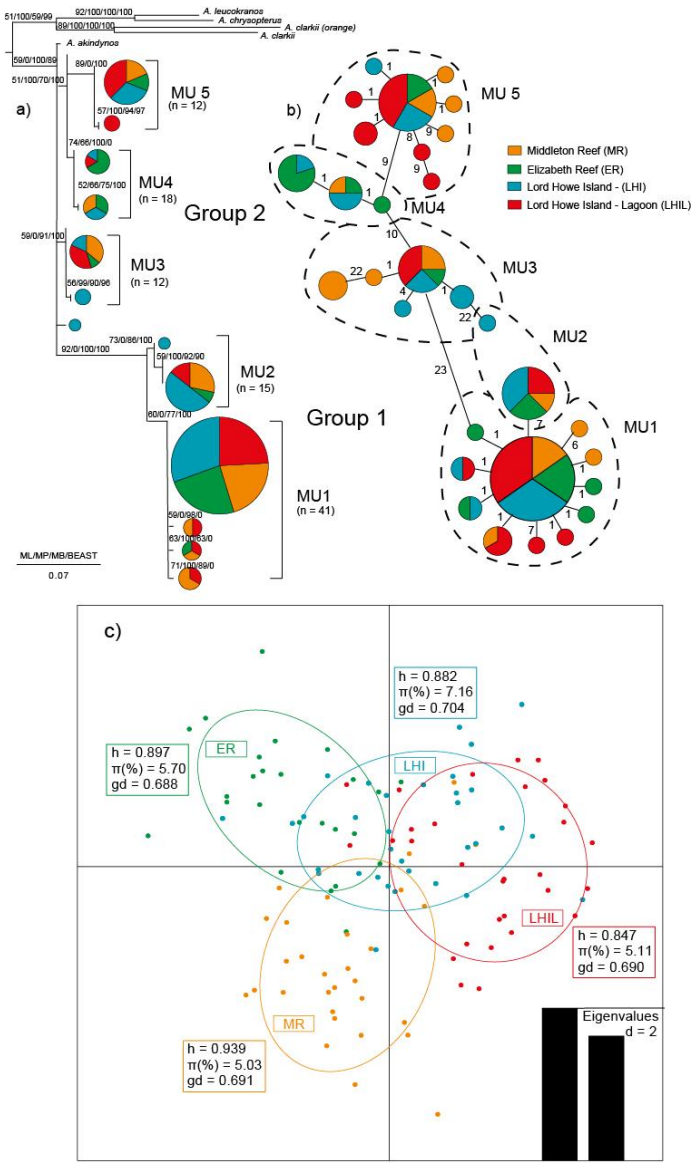

Figure 4.2 mtDNA and msatDNA gene tic analyses for *Amphiprion. mccullochi*. a) A phylogram of mtDNA (D-Loop) sequences from 118 *A. mccullochi* individuals from Elizabaeth Reef, Middleton Reef and Lord Howe Island. This represents the best ML tree from 10 individual analyses. Numbers on branches indicate support for each clade, based on phylogenetic analyses. b) Haplotype minimum spanning tree (MST) with the number of substitutions between haplotypes indicated on connectors. Different coloured fills represent each of the four populations from the three reefs as shown on the key to the figure. c) Scatterplots of the discriminant analysis of principal components of the microsatellite data for four *Amphiprion mccullochi* populations using geographic sample site as priors for genetic clusters. Individual genotypes appear as dots surrounded by 95% inertia ellipses. Eigenvalues show the amount of genetic information contained in each successive principal component with x and y axes constituting the first two principle components, respectively. Boxes indicate haplotype (*h*), nucleotide (π) and genetic diversity (gd) indices for *A. mccullochi.*

Population genetic analyses of mtDNA, based on an AMOVA, revealed two regional partitions (ER and MR vs LHI and LHIL) and all of the genetic variation (101.74%) was within locations, Φ_{st} = -0.017 (p = 0.8, Table 4.S2), however, this was not significant. Pairwise F_{st} comparisons subsequently revealed no mtDNA genetic differentiation between locations (MR, ER, LHI, LHIL; F_{st} = -0.0029 to -0.008, p = 0.513 to 0.973, Table S4) and is consistent with the phylogenetic results.

Quantifying the level of historical gene flow. Bayesian analysis, informed by the phylogenetic structure, was performed using MIGRATE-n, because analyses based on spatial structure failed. High levels of historical gene flow were indicated within - but less between groups: between Groups (i.e. Group 1 - Group 2) M ranged from 19 to 42 (Fig 4.3a). These values were 2- to 6- fold lower than historical gene flow within groups: Group 1 (MU 1 - 2) M ranged from 72 to 146 (Figure 3a) and Group 2 (MU 3, 4, 5) M ranged from 180 to 246 (Fig 4.3a).

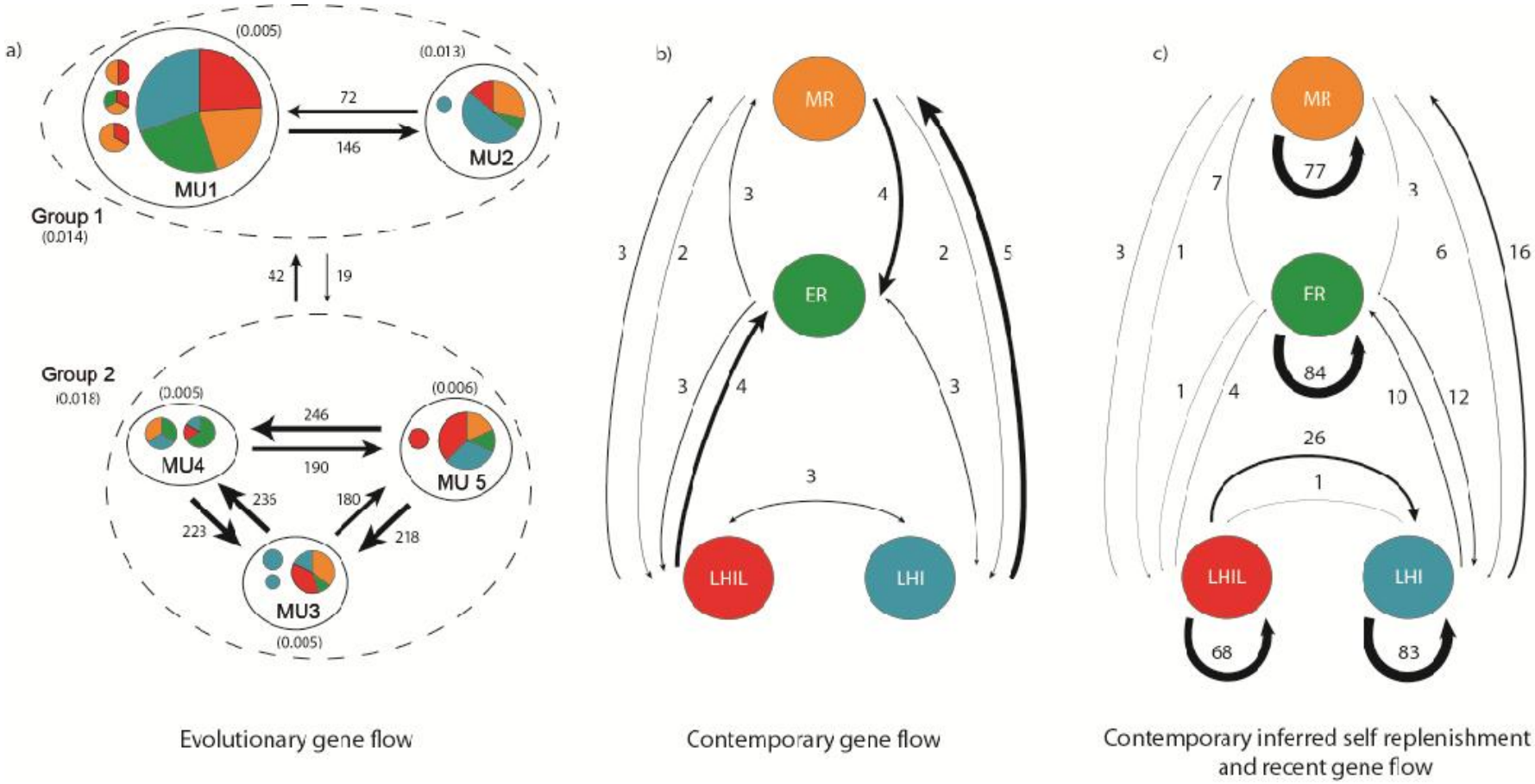

Figure 4.3 Migration rates among *Amphiprion mccullochi* locations. The thickness of the line is directly proportional to the number of migrants (M) and the colour of lines indicate predominant direction of gene flow. Population size (θ, within parentheses) is also shown for each location. a) Migrate-n historical gene flow (mtDNA), b) Migrate-n demographic gene flow (msatDNA) and c) BAYESASS analysis of self-replenishment (msatDNA) and recent migration shown as a percentage.

4.4.2. Gene flow between locations - demographic time scales

Synopsis: msatDNA allele frequencies, genotypic distributions in space, genotypic assignments and genotypic posterior probability distributions suggested significant spatial partitions between *A. mccullochi* from the four locations in the latter three of the four analyses. Low levels of demographic gene flow were detected between the four locations, consistent with the patterns of demographic gene flow and with the high levels of inferred self-replenishment evident at all four locations (next section). This is in stark contrast to the patterns and levels of historical gene flow.

Population genetic analyses of msatDNA. The statistically rigorous AMOVA found significant structure in the locus by locus msatDNA (Φ_{st} = -0.49 to 0.056, $p < 0.05$, Table 4.S3) and in the global AMOVA as a weighted average over all microsatellite loci (Φ_{st} = 0.007, p =0.015, Table 4.S2), with 99.34% of the genetic variation existing within locations. Raw msatDNA pairwise F_{st} comparisons also identified significant genetic partitioning between all locations (F_{st} = -0.004 to 0.026, p = 0.01 to 0.03, Table 4.S4), but ENA corrected pairwise F_{st} values showed significant differentiation only between two of the four locations, ER and LHI (F_{st} = 0.014, $p< 0.05$, Table 4.S4). Discriminant analysis of principal components (DAPC) partitioned *A. mccullochi* into four spatially structured populations (Fig 4.2c). Using the four locations as *a priori* population criteria, DAPC assigned 76 to 80% of all individuals to the location from which they were sampled (assignment per population: 76% each for ER and LHI; 80% each for MR and LHI, Fig 4.4). The 95% genotypic inertia ellipses (GIE) for ER and LHIL did not overlap, whilst the 95% GIE for MR overlapped with all 95% GIEs from the remaining

three locations. This is consistent with some ENA corrected pairwise F_{st} values and importantly, with the composition of MU 2, 4 and 5. Geographical structure in msatDNA data was also confirmed by GeneClass2 analyses, where only 5 individuals grouped with a location from which they were not sampled (MR = 1, ER = 1, LHIL = 3). Similarly, four geographically partitioned populations were identified by STRUCTURE analyses, as the likelihood of the marginal posterior probability distribution was highest when $K = 4$.

Quantifying the level of demographic gene flow. Demographic gene flow between locations was a few orders of magnitude lower than historical gene flow between locations using Migrate-n, with M values ranging from 2 to 5 (Fig 4.3b). This suggests that populations at each location are unlikely to be sustained from distant locations in the short term.

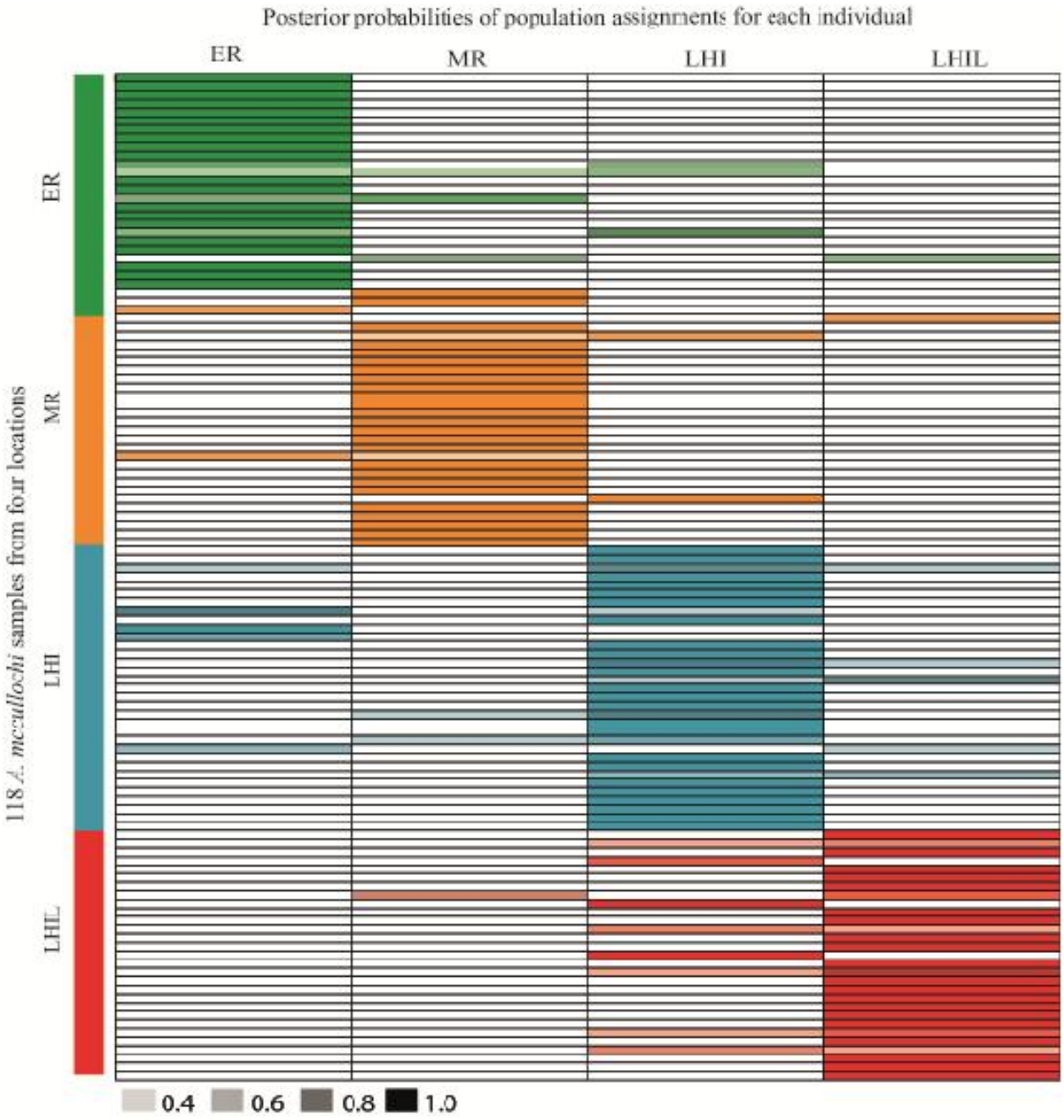

Figure 4.4 Posterior probability of assignment of each individual genotype to four *Amphiprion mccullochi* populations as indicated by DAPC. The names of the possible assignment populations are given on the x-axis. 118 genotypes are listed on the y-axis, along with the population from which they were sampled. Coloured bars corresponds to a 0.2 to 0.8 probability of assignment to a given population.

4.4.3. Inferred levels of self-replenishment and recent migration

Demographic independence is suggested for all location pairs except: LHIL to LHI (*m* = 26%), LHI to/from ER (*m* = 10 and 12%, respectively) and MR to LHI (*m* = 16%; Fig 4.3c). Conversely, high levels of self-replenishment (68 to 84%) were inferred at all four locations (Fig 4.3c). This indicates that each location is predominantly sustained by self-replenishment in the short term, rather than replenishment from distant locations.

4.4.4. Population genetic diversities

Amphiprion mccullochi from all four locations had high haplotype diversity (*h*), nucleotide diversity (%π) and genotypic diversity (*gd*): *h* = 0.846 to 0.939,%π = 5.03 to 7.16, *gd* = 0.690 to 0.736 (Fig 4.2c). Total haplotype, nucleotide and genotypic diversities were also high, *h* = 0.897,%π = 5.70 and *gd* = 0.688 (Table 4.S2) for this species. This is high genetic diversity and is unexpected for a low abundance endemic species, but is consistent with increased genetic diversity expected within locations when there is historical connectivity between them (i.e. within location − high genetic diversity; between locations − low genetic diversity).

4.5. DISCUSSION

Isolated islands are global hotspots of endemicity for a range of coral reef organisms (Roberts et al. 2002, Allen 2008) and determining the level and direction of gene flow (Palstra et al. 2007) between locations is a fundamental step in establishing MPA networks that effectively conserve unique marine biodiversity. In this study, *A. mccullochi* was found to have: (i) sufficient gene flow between locations resulting in a lack of geographic partitioning over historical time scales; (ii) genetically differentiated populations at all four sampled locations, due to low levels of demographic gene flow between locations, despite the historical homogenisation; (iii) demographic dependence between LHI and LHIL, LHI and ER and MR and LHI ,yet high levels of inferred self-replenishment at all four locations and; (iv) high genetic diversity at all locations, despite high levels of inferred self replenishment. This is consistent with inter-location gene flow at historical time scales.

4.5.1. Gene flow between locations - historical time scales

The identification of discrete phylogenetic lineages or management units (MU) is critical for developing effective management strategies (Schwartz et al. 2007). MU represent populations which rely on self regulation rather than immigration from external sources. Two distinct lineages with a total of 5 MU were suggested for *A. mccullochi* mtDNA. Despite this, the relative percentage of each location within MU suggests geographic structure. The occurrence of two lineages within a species has also been found for coral reef fishes on the Great Barrier Reef (GBR). Both *Plectropomus maculatus* and *Lutjanus carponotatus* show a lack of geographic partitioning along the GBR, yet display two distinct lineages, suggesting admixtures of differentiated lineages rather than stable populations (Evans et al. 2010). A lack of geographical

structure has also been found in endemic Hawaiian species *Chaetodon multicinctus*, *Chaetodon miliaris*, *Chaetodon fremblii* (Craig et al. 2010) and *Halichoeres ornatissimus* (Ludt et al. 2012) and in numerous other widespread coral reef fish species including *S. frenatus* (Dudgeon et al. 2000), *C. sordidus* (Bay et al. 2004), *Lethrinus miniatus* (van Herwerden et al. 2000), *Pseudochromis fuscus* (Messmer et al. 2005) and *Plectropomus leopardus* (van Herwerden et al. 2006).

A. mccullochi showed high historical gene flow between MU within lineages and to a far lesser extent, between lineages. Higher gene flow from Group 2 into Group 1 is clear, suggesting introgression of mtDNA (shown to be a result of historical hybridisation between *A. mccullochi* and its widespread sister species *A. akindynos*) (van der Meer et al. 2012a; Chapter 3). In a similar way, the levels of historical gene flow between three sympatric species pairs of three-spined stickleback (*Gasterosteus aculeatu*) have revealed natural hybridisation and break down of a species pair into a hybrid swarm (Gow et al. 2006). In addition, historical gene flow between locations has also been found in Red Sea reef fishes *Larabicus quadrilineatus, Chromis viridis* and *Pseudanthias squamipinnis* (Froukh 2007. Froukh & Kochzius 2007). Consequently, the lack of geographical structuring and observed spatial genetic homogeneity identified in this study of the endemic *A. mccullochi*, is likely due to high levels of historical gene flow, which is sufficient for all locations to be connected on historical time scales, thereby maintaining genetic homogeneity.

4.5.2. Gene flow between locations - demographic time scales

A. mccullochi showed strong demographic genetic differentiation between locations, consistent with other coral reef fish such as the Hawaiian endemic surgeonfish *Ctenochaetus strigosus* (Eble et al. 2009). Strong discrepancies between historical and demographic levels of gene flow in *A. mccullochi* are a direct result of different spatial and temporal time scales. Discrepancies in gene flow, between time scales, has also been shown for *Lutjanus synagris* (Gold et al. 2011), *Plectropomus maculatus* and *Lutjanus carponotatus* (Harrison et al. 2012, Evans et al. 2010).

As previously highlighted, only a few individuals are needed over historical time scales to ensure homogeneity across a species entire geographical range (Shulman 1998, Planes 2002). However, models predict that this level of gene flow is not sufficient to sustain local populations and as a consequence, local populations must sustain themselves via self-recruitment or self-replenishment (Cowen 2000, 2002). Thus, although historical gene flow is important, it is the dispersal rate of individuals that is of immediate interest to sustaining populations (Palsboll et al. 2007). *A. mccullochi* showed very low levels of gene flow at demographic timescales which is consistent with model prediction. The low levels of demographic gene flow in this system most likely result from the short pelagic larval duration of *A. mccullochi* and the geographical isolation between locations enhanced by predominant east to west oceanographic currents limiting north-south gene flow between locations (Gaither et al. 2010).

4.5.3. Inferred levels of self-replenishment and recent migration

Demographic independence results from gene flow between two locations falling below 10% (Waples & Gaggiotti 2006). Thus, the high abundance of the McCulloch's anemonefish residing within the LHI lagoon will not directly sustain other locations in the short term , except outside the lagoon at LHI. Rather LHIL will help replenish LHI, which in turn will replenish ER, whilst

both ER and MR will replenish LHI. This complex network of gene flow highlights the need to protect each location under one management strategy. Interestingly, the levels of inferred self-replenishment found in this study (≥68%) are remarkably similar to the estimated levels of self-recruitment in other congeneric anemonefish studies in Papua New Guinea (PNG) (Jones et al. 2005, Planes et al. 2009) These levels are also similar to those found in other reef fishes inhabiting islands including butterflyfish in PNG (Almany et al. 2007) and wrasse in the Caribbean (Swearer et al. 1999) whose estimates of self-recruitment ranged from 30 to 60%. Possibly, the higher self-replenishment in *A. mccullochi*, compared to the above studies, results from the complete sampling of all locations leaving no 'ghost ' populations un-sampled. However, further investigation using direct methods (e.g. by using natural or artificial otolith tags of newly recruited juveniles (Jones et al. 2005, Planes et al. 2009) is necessary to validate the inferred levels of self-replenishment in *A. mccullochi*. This approach may not be appropriate for endemic species with low abundance. Given the rarity of *A. mccullochi* at MR and ER, parentage studies involving otolith tagging and the sacrificing of a high proportion of individuals may lead to local extinction at these sites.

4.5.4. Population genetic diversities

A. mccullochi showed high genetic diversities despite its low abundance and high levels of inferred self replenishment. Similarly high genetic diversities have also been found in other coral reef fish including *Plectropomus maculatus, Lutjanus carponotatus* (Evans et al. 210), *Lethrinus miniatus* (van Herwerden et al. 2003) and damselfish on the Great Barrier Reef (Bay et al. 2004). In *A. mccullochi* this higher than expected genetic diversity is most likely driven by bi-

directional hybridisation with its sister species *A. akindynos* (van der Meer et al. 2012a, Chapter 3), a process which has also been documented in *Plectropomus leopardus* (van Herwerden et al. 2006). While high genetic diversities may provide some level of population resilience to environmental change, high levels of inferred self-replenishment make populations more vulnerable to extirpation due to low levels of replenishment from elsewhere via demographic gene flow. Additionally, a cautious approach is required to prevent population losses, even those with high genetic diversity (Chapman et al. 2011), as quantitative trait loci under selection at the peripheral edge of a species distribution range might have no genetic diversity remaining, despite neutral markers having relatively high genetic diversity in the same population (Kellerman et al. 2009). Therefore, low levels of demographic gene flow, coupled with high levels of self-replenishment have implications for the management, persistence and effective conservation of this endemic coral reef fish species – even if genetic diversity is high.

Threats and concerns Conserving endemic species such as *A. mccullochi* presents a unique challenge to management. Although remote islands are largely unaffected by the pressures experienced by coastal reefs, a variety of anthropogenic threats still exist. These include sewage leaks and anemone bleaching due to increased temperatures (Jones et al. 2004). The occurrence of these events at locations such as LHI lagoon (Coade et al. 2010) is a serious cause for concern (Hattori 2002) since 75% of *A. mccullochi* surveyed in 2009 resided in designated high-protection 'sanctuary zones' within the lagoon (Hobbs et al. 2009). It follows then that protecting critical habitat (i.e. *Entacmaea quadricolor* anemones) and keeping the natural genetically distinct sub-populations (MU) of endemic fish intact, should be a priority of management plans. In addition, isolated locations that are predominantly dependent on self-recruitment are unlikely

to be sustained by long distance transport over hundreds of kilometres (Cowen 2000, 2002) and therefore unlikely to recover fast (Graham et al. 2007, Smith et al. 2008). Lastly, small, isolated populations are subject to genetic deterioration and, if habitat fragmentation increases in the future (due to habitat loss from climate change), gene flow may be further restricted, leading to inbreeding and an increase in extinction risk with as much as 29% reduced persistence times (Liao & Reed 2008).

Climate change offers an additional suite of threats and concerns**.** LHI, like other isolated islands, is facing an escalation of threats (e.g. increasing intensity and frequency of cyclones, rising sea surface temperatures, ocean acidification) (Heath 2008), with negative effects on biodiversity expected within the region. In the case of the McCulloch's anemonefish and Hobbs et al. (2009) noted in their surveys of LHI coral reefs that some of the host anemones were bleached (typically a response to elevated sea temperatures) (Hill & Scott 2012). As sea temperatures continue to increase due to global warming, the intensity and frequency of bleaching events is likely to increase, directly threatening the persistence of this obligate habitat specialist and potentially other coral reef fish. High genetic diversity is unlikely to overcome the loss of habitat in the time frames expected, particularly if the quantitative trait associated with specialised host use already has limited or no genetic diversity. With the expected increase in strength of the EAC bringing warmer waters to subtropical regions (Suthers et al. 2011) these isolated island populations may at further risk of extinction if they can not tolerate elevated temperatures or extend their current geographic ranges.

Conclusion The present study highlights the importance of estimating both historical and demographic levels of gene flow (connectivity) due to the different spatial and temporal scales at

which these processes operate. While populations are primarily being maintained by self replenishment, exchange among islands over historical time is critical to understanding patterns of genetic diversity and differentiation. Locations with high levels of self-replenishment (e.g. MR, ER, LHI) each require protection as they receive few dispersing larvae from each other. Locations with lower levels of self-replenishment (e.g. LHIL) are just as important to protect as they provide a dual benefit because they are a source for their own and other populations, aiding in rescue effects of depleted/extinct populations and enhancing genetic diversity. Thus both predominantly self-replenishing and predominantly dispersing locations should ideally be protected, from activities such as aquarium collecting, to maximise biodiversity conservation in low abundance endemics living on isolated reefs and islands. Although this study focused on a single coral reef species at four locations in the South-West Pacific Ocean, the region harbours 16 other species of endemic marine fishes, as well as numerous other endemic marine species that have similar geographic distributions as our study species. Thus patterns of gene flow and self-replenishment in *A. mccullochi* may be representative of other endemic species.

Supplementary Online Information

Table 4.S1. Summary statistics for 17 microsatellite loci Am1-24. Sample sizes (N), observed number of alleles (Na), observed number of private alleles (Pa), observed heterozygosity (Ho), expected heterozygosity (He), the average inbreeding coefficient (F_{IS}), probability of departure from HWE (pHWE) and significance of departure after FDR correction FDR (pFDR) for each locus at each population (significance of departure in bold, $p < 0.05$).

Population	Am1	Am4	Am5	Am6	Am7	Am9	Am10	Am11	Am12	Am14	Am15	Am17	Am18	Am19	Am21	Am22	Am24
All (118)																	
N	115	115	113	114	111	114	115	115	110	109	91	95	103	83	85	93	110
Na	5	11	16	5	12	16	10	13	11	11	17	16	13	16	10	16	5
Pa	115	115	113	114	111	114	115	115	110	109	91	95	103	83	85	93	110
Ho	0.626	0.713	0.832	0.702	0.649	0.974	0.696	0.574	0.855	0.771	0.868	0.684	0.845	0.723	0.824	0.871	0.591
He	0.547	0.789	0.850	0.622	0.793	0.904	0.683	0.791	0.813	0.803	0.848	0.824	0.789	0.847	0.812	0.857	0.567
Fis	-0.157	0.098	0.020	-0.123	0.190	-0.073	-0.018	0.274	-0.054	0.036	-0.022	0.158	-0.073	0.144	-0.001	-0.015	-0.040
pHWE	**0.013**	0.367	0.998	0.646	0.130	0.789	0.999	**0.000**	0.647	**0.000**	1.000	**0.000**	**0.000**	**0.000**	0.433	**0.000**	0.400
pFDR	0.069	0.671	0.817	0.817	0.361	0.817	0.817	**0.000**	0.817	**0.000**	0.817	**0.000**	**0.000**	**0.000**	0.734	**0.000**	0.695
ER (25)																	
N	25	25	25	25	22	25	25	25	22	25	20	15	15	20	21	13	21
Na	4	10	10	4	6	13	5	9	8	9	13	10	7	11	8	13	4
Pa	0	0	2	0	0	0	0	1	0	0	2	2	2	2	0	2	0
Ho	0.800	0.640	0.840	0.680	0.500	0.880	0.720	0.560	0.909	0.880	0.850	0.800	0.867	0.600	0.905	0.846	0.524
He	0.554	0.774	0.868	0.566	0.729	0.912	0.756	0.805	0.789	0.806	0.845	0.784	0.767	0.824	0.804	0.891	0.490
Fis	-0.426	0.192	0.053	-0.181	0.335	0.055	0.068	0.323	-0.129	-0.071	0.020	0.015	-0.096	0.295	-0.101	0.090	-0.045
pHWE	0.263	**0.025**	0.825	0.744	**0.042**	0.698	0.095	**0.000**	0.777	0.988	0.996	0.572	**0.009**	**0.013**	0.116	0.543	0.979
pFDR	0.546	0.109	0.817	0.817	0.162	0.817	0.314	**0.000**	0.817	0.817	0.817	0.817	0.057	0.069	0.361	0.817	0.817

Table 4.S1 *(Continued)*

Population	Am1	Am4	Am5	Am6	Am7	Am9	Am10	Am11	Am12	Am14	Am15	Am17	Am18	Am19	Am21	Am22	Am24
MR (30)																	
N	33	31	29	33	32	33	33	33	32	28	23	28	31	23	24	28	32
Na	4	9	12	4	12	13	8	8	10	10	11	11	9	10	9	11	4
Pa	0	0	2	0	2	0	1	1	0	1	1	3	1	0	0	0	0
Ho	0.576	0.548	0.828	0.758	0.594	1.000	0.758	0.545	0.875	0.857	0.870	0.786	0.839	0.826	0.875	0.821	0.469
He	0.528	0.757	0.826	0.670	0.792	0.905	0.690	0.779	0.833	0.849	0.830	0.824	0.779	0.847	0.824	0.853	0.572
Fis	-0.076	0.291	0.015	-0.116	0.265	-0.090	-0.083	0.314	-0.035	0.009	-0.026	0.065	-0.061	0.047	-0.041	0.055	0.196
pHWE	0.876	**0.018**	0.670	0.460	0.273	**0.018**	0.998	0.154	0.910	0.698	0.998	0.270	0.937	0.784	0.627	0.704	0.595
pFDR	0.817	0.083	0.817	0.743	0.546	0.083	0.817	0.396	0.817	0.817	0.817	0.546	0.817	0.817	0.817	0.817	0.817
LHI (33)																	
N	33	31	29	33	32	33	33	33	32	28	23	28	31	23	24	28	32
Na	4	9	12	4	12	13	8	8	10	10	11	11	9	10	9	11	4
Pa	0	0	2	0	2	0	1	1	0	1	1	3	1	0	0	0	0
Ho	0.576	0.548	0.828	0.758	0.594	1.000	0.758	0.545	0.875	0.857	0.870	0.786	0.839	0.826	0.875	0.821	0.469
He	0.528	0.757	0.826	0.670	0.792	0.905	0.690	0.779	0.833	0.849	0.830	0.824	0.779	0.847	0.824	0.853	0.572
Fis	-0.076	0.291	0.015	-0.116	0.265	-0.090	-0.083	0.314	-0.035	0.009	-0.026	0.065	-0.061	0.047	-0.041	0.055	0.196
pHWE	0.876	0.018	0.670	0.460	0.273	**0.018**	0.998	0.154	0.910	0.698	0.998	0.270	0.937	0.784	0.627	0.704	0.595
pFDR	0.817	0.083	0.817	0.743	0.546	0.083	0.817	0.396	0.817	0.817	0.817	0.546	0.817	0.817	0.817	0.817	0.817

Table 4.S1 *(Continued)*

Population	Am1	Am4	Am5	Am6	Am7	Am9	Am10	Am11	Am12	Am14	Am15	Am17	Am18	Am19	Am21	Am22	Am24
LHIL (30)																	
N	27	30	30	27	28	27	30	29	28	28	27	27	30	27	27	27	28
Na	3	11	9	4	10	13	5	7	8	8	11	9	9	13	9	10	4
Pa	0	1	0	0	0	1	1	1	0	0	2	1	0	4	0	1	0
Ho	0.556	0.900	0.800	0.667	0.786	1.000	0.600	0.621	0.821	0.750	0.889	0.593	0.867	0.704	0.778	0.889	0.571
He	0.543	0.819	0.844	0.620	0.851	0.907	0.635	0.782	0.833	0.763	0.875	0.823	0.812	0.853	0.800	0.843	0.554
Fis	-0.005	-0.082	0.069	-0.056	0.095	-0.084	0.072	0.223	0.032	0.035	0.003	0.297	-0.050	0.193	0.046	-0.036	-0.013
pHWE	**0.007**	0.901	0.910	0.724	0.972	0.203	0.981	0.127	0.830	0.898	0.647	0.219	**0.034**	0.346	0.446	0.502	0.192
pFDR	**0.049**	0.817	0.817	0.817	0.817	0.470	0.817	0.361	0.817	0.817	0.817	0.491	0.139	0.650	0.738	0.793	0.460

Table 4.S2: AMOVA analysis for a) mtDNA sequences from *Amphiprion mccullochi* structured into geographic regions and b) global AMOVA weighted across all seventeen microsatellite loci.
Significant *p*-values are in bold.

Source of variation	Variance component	Percentage of variation	F-statistics fixation indices (p-value)
a) Region			
Amoung groups	-0.02	-0.17	F_{ct} = -0.002 (0.688)
Amoung populations within groups	-0.14	-1.57	F_{sc} = -0.016 (0.682)
Within populations	9.10	101.74	F_{st} = -0.017 (0.800)
b) Microsatellite			
Amoung groups	0.03	0.41	F_{ct} = 0.004 (0.037)
Amoung populations within groups	0.02	0.26	F_{sc} = 0.003 (0.235)
Within populations	6.70	99.34	F_{st} = 0.007 **(0.015)**

Table 4.S3: AMOVA fixation indices (Φ_{st}) for *Amphiprion mccullochi* across all populations surveyed

	Marker class and analysis		
	Raw msat	Msat corrected for null allele freq.	Standardised msat
Average	**0.007**	0.005	0.019
Am1	-0.004	**-0.001**	**-0.049**
Am4	-0.004	**-0.007**	**0.032**
Am5	0.008	0.007	0.076
Am6	0.007	0.009	**0.048**
Am7	0.001	**-0.005**	**0.021**
Am9	-0.003	**-0.002**	**0.053**
Am10	0.013	**0.013**	**0.043**
Am11	0.010	0.004	**0.025**
Am12	0.007	0.008	**-0.024**
Am14	0.011	**0.010**	**0.027**
Am15	0.006	0.006	**0.047**
Am17	0.010	**0.011**	**-0.016**
Am18	0.003	0.004	**-0.025**
Am19	0.002	**0.001**	**0.056**
Am21	-0.008	**-0.008**	**0.041**
Am22	0.004	0.004	**-0.031**
Am24	**0.031**	**0.031**	**-0.002**

Raw population differentiation from microsatellite allele frequencies for each individual loci and the average across all loci, population differentiation corrected for null allele frequencies using the ENA correction and standardized population differentiation for and across all loci (Φ_{st}). All values in bold are significant to the $p<0.05$ (i.e. 95% confidence interval).

Table 4.S4: Pairwise population Fst values for four populations of *Amphiprion mccullochi* using both d loop (mtDNA) and microsatellite (msat). Pairwise population structures (F_{st}) for four populations of *A. mccullochi*, using both d loop (mtDNA) and microsatellite (msat) loci showing raw and corrected Fst for null allele frequencies.

	D-loop				Raw msat				Corrected msat			
	ER	MR	LHI-N	LHI-L	ER	MR	LHI-N	LHI-L	ER	MR	LHI-N	LHI-L
ER		0.513	0.604	0.378		**0.010**	**0.010**	**0.010**		>0.05	>0.05	**<0.05**
MR	-0.013		0.973	0.490	0.023		**0.020**	**0.010**	0.005		>0.05	>0.05
LHI-N	-0.024	-0.029		0.550	0.016	0.007		**0.030**	0.005	0.002		>0.05
LHI-L	-0.008	-0.009	-0.018		0.026	0.023	0.004		0.014	0.005	0.000	

Fst values below diagonal. Significant p-vlaues are in bold (p > 0.05) above diagonal. Raw population differentiation from microsatellite allele frequencies and associated *p*-values. Corrected population differentiation for null allele frequencies at 95% CI (all *p*-values > 0.05) using the ENA correction. FDR correction of raw and corrected msat F_{st} *p*-values.

Table 4.S5: Sample sizes for D loop (total n = 105)

	n (D loop)	n_h	h	π(%)	n (msat)	gd	N_a	P_a	H_o	H_e	Fis
All	105	31	0.897	5.70	118	0.688	8.7	50	0.752	0.773	0.02
ER	21	12	0.914	5.13	25	0.736	8.5	13	0.753	0.763	0.04
MR	22	13	0.939	5.03	30	0.691	8.9	12	0.750	0.781	0.06
LHI-N	31	13	0.882	7.16	33	0.704	9.1	12	0.754	0.774	0.04
LHI-L	26	11	0.846	5.11	30	0.690	8.4	12	0.752	0.774	0.05

Number of haplotypes (n_h), haplotype diversity (h), nucleotide diversities (π) of D loop for all regions and populations of *Amphiprion mccullochi*. Sample sizes for msats (total n = 118), genetic diversity (gd) average number of alleles per locus (Na), observed number of private alleles (Pa), observed heterozygosity (H_o) and expected heterozygosity (H_e) and the inbreeding coefficient (Fis) averaged over seventeen microsatellite loci for four populations in three regions.

CHAPTER 5: **Limited demographic gene flow and high self-replenishment drives peripheral isolation in an endemic coral reef fish**

(Published as: van der Meer MH, Horne JB, Gardner MG, Hobbs J-PA, Pratchett M, van Herwerden L (2013) Limited contemporary gene flow and high self-replenishment drives peripheral isolation in an endemic coral reef fish. *Ecology and Evolution*, DOI: 10.1002/ece3.584.

5.1. ABSTRACT

Extensive ongoing degradation of coral reef habitats worldwide has lead to declines in abundance of coral reef fishes and local extinction of some species. Most vulnerable are ecological specialists and endemic species. Determining connectivity between locations is vital to understanding recovery and long-term persistence of species following local extinction. This study explored population connectivity in the ecologically-specialised endemic three-striped butterflyfish (*Chaetodon tricinctus*) using mt- and msat-DNA (microsatellites) to distinguish historical versus demographic gene flow, estimate self-replenishment and measure genetic diversity among locations at the remote Australian offshore coral reefs of Middleton Reef (MR), Elizabeth Reef (ER), Lord Howe Island (LHI) and Norfolk Island (NI). Mt- and msat-DNA suggested genetic differentiation of the most peripheral location (NI) from the remaining three locations (MR, ER, LHI). Despite high levels of mtDNA gene flow, there is limited msatDNA gene flow with evidence of high levels of self-replenishment ($\geq$ 76%) at all four locations. Taken together, this suggests prolonged population recovery times following population declines. The

NIR peripheral population (NI) is most vulnerable to local extinction due to its relative isolation, extreme levels of self-replenishment (95%) and low demographic abundance.

5.2. INTRODUCTION

Coral reef fishes have evolved in a close relationship with coral reef habitats to produce the most diverse vertebrate communities on earth (Bellwood 1996, Bellwood & Wainwright 2002, Wood 1999, Bellwood et al. 2010). However, coral reef habitats are coming under increasing pressure, facing a multitude of impacts including destructive and excessive fishing, sedimentation, pollution, disease, coral bleaching, ocean warming and acidification (Hoegh-Guldberg 1999, Hughes et al. 2003, Bellwood et al. 2004). These disturbances have combined to cause sustained and ongoing declines in the abundance of corals on reefs worldwide (e.g., Bellwood et al. 2004, Gardner et al. 2003) with approximately 20% of the world's coral reefs recently destroyed and a further 50% in decline (Wilkinson 2004), whilst coral cover on the Great Barrier Reef has halved in the last 27 years (De'ath et al. 2012). Given their strong reliance on live coral habitats, the abundance and diversity of reef fishes invariably declines with severe and/or prolonged declines in coral cover (Wilson et al. 2006, Jones et al. 2004, Graham et al. 2006, Pratchett et al. 2008). Extensive coral loss has resulted in the local extinction of some coral reef fishes, particularly those species that rely on live coral (Kokita & Nakazano 2001, Graham et al. 2006, Pratchett et al. 2008). Local extinction of coral dependent fishes are likely to increase if major disturbances that cause acute and extensive coral loss, such as coral bleaching, increase in incidence, as predicted (Hoegh-Guldberg 1999, Sheppard 2003).

In terrestrial habitats, endemic species (particularly on isolated islands) typically have higher rates of extinction and lower genetic diversity (Frankham 1997, Whittaker & Fernández-Palacios 2007). Coral reef fish communities on isolated islands tend to have a high proportion of endemics (Jones et al. 2002), and account for some of the most recent fish extinctions (Dulvy et

al. 2003). Endemic species may be particularly vulnerable to widespread disturbances with their inherent small geographical range and small population size (Gaston 1998). This risk of extinction is further increased if endemic species have specific dietary (Pratchett et al. 2006, Graham 2007) or specialist habitat (Munday 2004, Wilson et al. 2006, 2008) requirements. The ability for coral dependent fishes to recover from local extinction will be dependent on the regeneration of their coral resources and larval replenishment from distant locations as assessed by gene flow. Thus, there is an urgent need to understand gene flow between, and genetic diversity at, locations inhabited by endemic reef fishes for ongoing monitoring and conservation, and to determine their recolonisation ability and resilience.

To thoroughly understand gene flow it is important that both historical and demographic levels of gene flow are determined (i.e. at various time- and spatial-scales; Palstra et al. 2007) since some reef fish studies have shown discrepancies, up to an order of magnitude difference, in gene flow over these different time/spatial scales (i.e. high historical but limited demographic gene flow: Evans et al. 2010, Harrison et al. 2012, van der Meer et al. 2012b; Chapter 3, 4). Although determining levels of gene flow is important, of equal importance is conserving genetic diversity. Conserving genetic diversity is an *International Union for Conservation of Nature* (IUCN) priority (McNeely et al. 1990) for at least two reasons: (i) it provides the raw material for natural selection to act on over historical (Johanesson & Andre 2006) and demographic time scales (Bell & Okamura 2005); and (ii) low genetic diversity increases the risk of inbreeding depression (Reed & Frankham 2003).

Large data sets of highly polymorphic msatDNA loci (nuclear microsatellites) produced by next generation sequencing (e.g. Gardner et al. 2011) and advancements in statistical techniques (e.g. Pritchard et al. 2000, Beerli & Felsenstein 2001, Wilson & Rannala 2003, Excoffier et al. 2005, Jombart et al. 2010) have increased the sophistication of population genetic studies. However, to date, few such studies have been able to sample all existing locations across a species limited range. Unsampled "ghost" locations can affect key demographic estimates (i.e. population size, genetic diversity, migration rate; Beerli 2004). Here I investigate patterns of gene flow and measure population genetic diversities in an ecological specialist reef fish, the endemic three-striped butterflyfish (*Chaetodon tricinctus*), by complete sampling across its four geographically isolated locations (all found within Australian waters): Middleton Reef − MR, Elizabeth Reef − ER, Lord Howe Island − LHI, and Norfolk Island – NI (Fig 5.1).

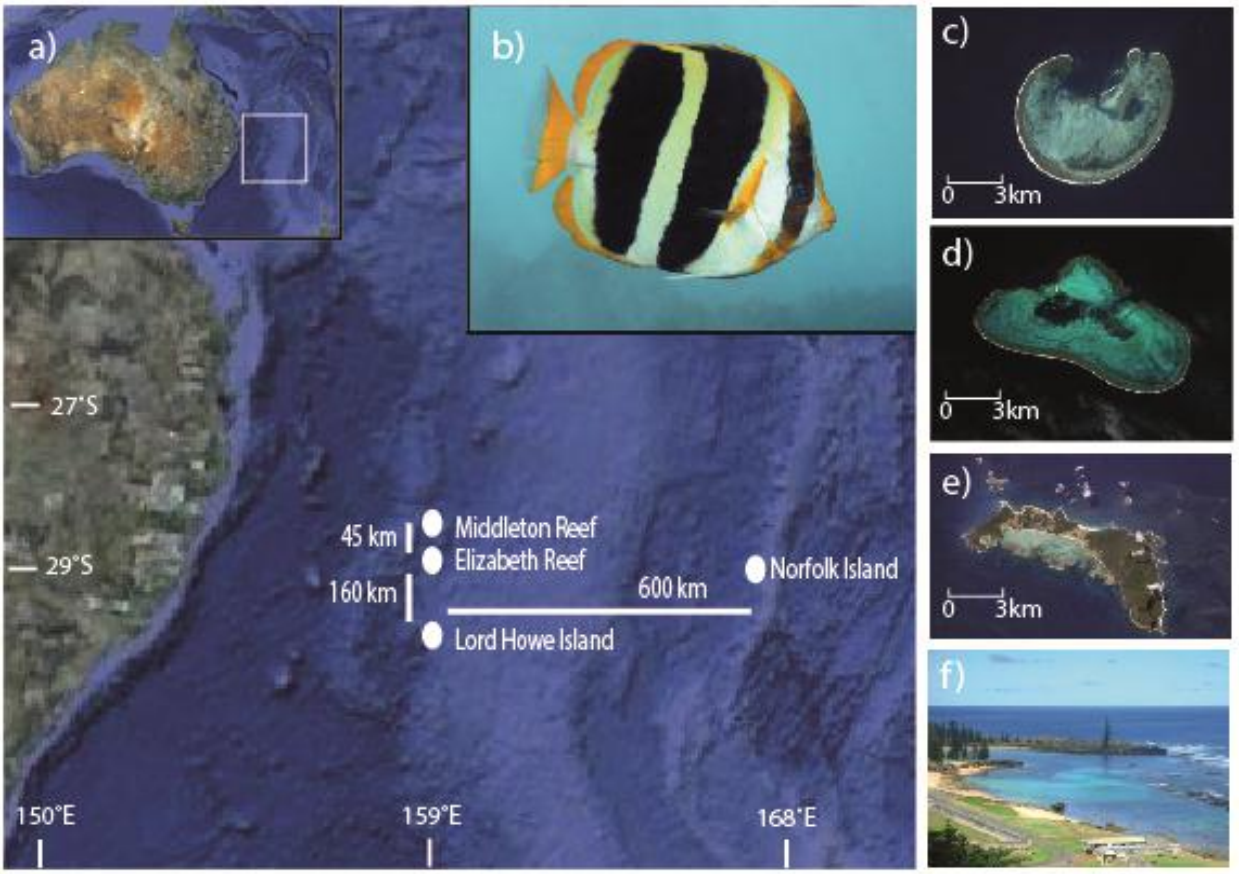

Figure 5.1 Location maps and focal species. (A) Google Earth image of eastern Australia showing Middleton Reef (MR), Elizabeth Reef (ER), Lord Howe Island (LHI) and Norfolk Island (NI) in the South West Pacific Ocean. (B) *Chaetodon tricinctus* swimming in the open (Photo courtesy of Justin Gilligan). Aerial photographs of MR (C); ER (D), LHI (E) and NI (F; the bay measures 1km in length).

The three-striped butterflyfish is an endemic to the LHI region (Randall 1976). This region is a hotspot for endemic coral reef fishes (Marine Parks Authority 2010) ranking fifth in the Indo-Pacific for percent endemism (7.2%, Randall 1998). Marine Protected Areas (MPAs) have been established to conserve reef fishes at three of these locations (MR, ER LHI), but no protection exists at Norfolk Island. This is an ideal study system since reef fishes occur on only four discrete islands/reefs that are separated by deep ocean water. Thus, connectivity of reef fish populations across the four locations is restricted to oceanic dispersal of pelagic larvae over known distances (e.g. 45 to 600 km).

Previous research on another endemic species in this system, the McCulloch's anemonefish (*Amphiprion mccullochi*), revealed limited demographic gene flow between ER, MR and LHI (van der Meer et al. 2012a, b; Chapters 3, 4). However, anemonefish have the shortest pelagic larval duration (PLD) of reef fishes (11 to 17 days: Victor 1986, Thresher et al. 1989, Wellington & Victor 1989, Victor & Wellington 2000) and their self-recruitment to natal areas has been well documented (Jones et al. 2005, Planes et al. 2009). While McCulloch's anemonefish provide a test of population connectivity in reef fishes at the lower limit of dispersal potential within the LHI region, determining the connectivity of reef fishes in general requires examining species from a common group with PLD's more typical of reef fish (20 to 50 days). Butterflyfishes (Chaetodontidae) are one of the 10 common families of fishes that are characteristic of coral reefs (Bellwood & Wainwright 2002). The PLD of *C. tricinctus* (mean = 35 days; authors unpublished data), is typical of butterflyfishes (26 to 53; e.g. Brothers et al. 1983, Brothers & Thresher 1985) and many other reef fishes. *C. tricinctus* is also one of the 41 butterflyfish species that feed directly on scleractinian corals (Cole et al. 2008, Rotjan & Lewis 2008),. Thus,

C. tricinctus provides a test of population connectivity in a common group of reef fishes, that is closely associated with coral reefs, and with a dispersal potential typical of most reef fishes.

C. tricinctus faces a higher risk of extinction as a consequence of its small geographic range, compared to its closest relatives *C. bennetti, C. plebeius* and *C. trifascialis* (Bellwood et al. 2010), which are distributed widely throughout the Indo-Pacfic (Allen et al. 1998). Moreover, *C. tricinctus* feeds exclusively on live corals (Kuiter 2002) and is mostly found in close association with corals of the genus *Acropora* (Hobbs et al. 2009). The abundance of *C. tricinctus* is positively linked to the abundance of *Acropora* spp., indicating that a loss of this coral could cause decreases in abundance and potential local extinction of *C. tricinctus* (Hobbs et al. 2009). The global abundance of *C. tricinctus* is likely to be much smaller than its widespread congenerics, and if it cannot alter its diet following coral loss, then these factors will compound upon its small geographic range and greatly increasing its vulnerability to local and possibly global extinction. Dramatic declines in abundance of several other butterflyfishes have occurred following extensive coral loss (Syms 1998, Pratchett et al. 2006), but some of the most vulnerable species have been spared from extinction due to their large geographic range (Lawton et al. 2011). Given that *C. tricinctus* exists at a few isolated locations and may be particularly vulnerable to local extinction, there is an obvious need to determine patterns of population connectivity and replenishment for this species.

The aims of this study were threefold: (i) to determine patterns and levels of gene flow between locations using mtDNA (mitochondrial DNA) and msatDNA; (ii) to estimate levels of self-replenishment (as a proxy for realised self-recruitment) and recent migration; and (iii) to measure

population genetic diversities at all locations as an indicator of potential resilience of populations to environmental change and extinction.

5.3. MATERIAL AND METHODS

Fin clip sample sizes ranged from 21 to 31: MR (n = 30), ER (n = 31), LHI (n = 26) and NI (n = 21).

5.3.1. Gene flow between locations - mtDNA

mtDNA phylogenetic analysis. mtDNA Cytochrome b (cyt b) sequence data were obtained from GenBank for the following three most closely related species which acted as outgroups: *C. trifascialis* (FJ167707.1), *C. plebius* (AF108602.1) and *C. bennetti* (FJ167686.1) based on the findings of Bellwood et al. (2009). jModeltest (Posada 2008) identified an TrN + G model under Akaike Information Criterion with gamma = 0.759.

Quantifying the level of mtDNA gene flow. A Log Maximum-Likelihood analysis (Ln ML) selected a migration prior (F_{st}), custom-migration model (migration model with variable Theta), constant mutation rate with an F84 mutation model, migration rate parameters (Theta and M to a maximum of 1 and 15,000, respectively) and a Bayesian analysis, using a heating search strategy of one long chain that sampled every 20th of 60 k sampled trees and applied a 20 k iteration burn-in. All parameters converged and fell within the 90% CI yielding values for θ and M (mutation-scaled migration rate) per location.

5.3.2. Gene flow between locations - msatDNA

Quantifying the level of msatDNA gene flow. I set datatype to Microsatellite (a simple electrophoretic ladder model), migration prior (F_{st}), custom-migration model

(migration model with variable Theta), constant mutation rate with a stepwise mutation, migration rate parameters (Theta and M to a maximum of 10 and 20, respectively) and a Bayesian analysis, using a heating search strategy of one long chain that sampled every 20th of 60 k sampled trees and applied a 20 k iteration burn-in. All parameters converged and fell within the 90% CI yielding values for θ and M (mutation-scaled migration rate) per location.

Inferred levels of self-replenishment and recent migration. I used BAYESASS v3 to estimate both self-replenishment (as a proxy for realised self-recruitment) and recent migration between locations; with a Markov chain Monte Carlo (MCMC) chain, consisting of a total of 11 M steps with a 2 M step burn in; prior values for migration rate, allele frequency and inbreeding coefficient were specified as 0.5, 0.6 and 0.6 respectively. These priors were selected because they gave acceptance rates within the 20 to 40% range showing convergence of the MCMC (Faubet et al. 2007). Ten independent runs separately assessed convergence of the MCMC (i.e. priors fell within the 20 to 40% range suggesting convergence) in order to evaluate the consistency of results obtained from these inferences.

5.3.3. Population genetic diversities

As per Chapter 2.3.3

5.4. RESULTS

Synopsis. Two hundred and eighty three base pairs of mtDNA (cyt b) were resolved for 97 *Chaetodon tricinctus* individuals; with a total of 15 polymorphic sites, of which three were parsimony informative. One small clade, Clade 1 (n = 4) contained exclusively individuals from the NIR peripheral location. The other, Clade 2 (n = 6), comprised of an equal frequency of all three LHIR locations (i.e. MR, ER, LHI; Fig 5.2a). A Minimum Spanning Tree (MST) identified 15 haplotypes in total, one of which was observed at high frequencies representing 82% (n = 80) of all individuals, and 12 of which were unique to single fish only in the sample examined here, nine of which were from the NIR peripheral location (Fig 5.2b). Mt- and msat-DNA Analysis of Molecular Variance (AMOVA) and pairwise F_{st} results indicate that there is population genetic differentiation between the LHIR locations and the NIR peripheral location, but there is no population genetic differentiation within the LHIR region (i.e. when MR, ER, LHI are grouped). Haplotype and genotype diversities were low (< 0.5) within the LHIR region, but high (> 0.5) at the NIR peripheral location (Fig 5.2c, Table 6.S1). Genotypic diversity (*gd*), in contrast, was high at three of the four locations, ER being the exception. Detailed genetic diversity, AMOVA, summary statistics, pairwise population comparisons and locus by locus AMOVA can be found in Supporting Information (Table 5.S1, 5.S2, 5.S3, 5.S4 and 5.S5 respectively).

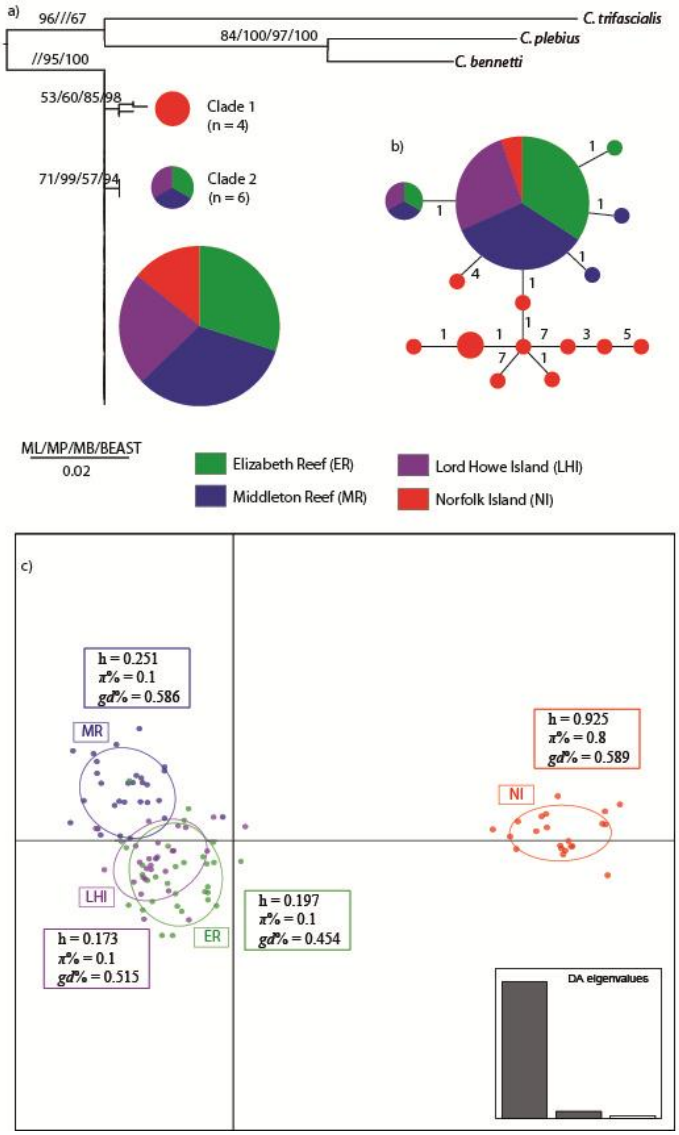

Figure 5.2 mt- and msat-DNA *C. tricinctus* analyses. a) A phylogram of mtDNA (cyt b) sequences from 97 *C. tricinctus* individuals from Middleton Reef, Elizabeth Reef, Lord Howe Island and Norfolk Island. This represents the best ML tree from 10 individual analyses. Numbers on branches indicate support for each clade, based on ML, MP, MB and BEAST analyses. b) Haplotype minimum spanning tree (MST) with number of substitutions between haplotypes indicated on connectors. Different fills represent each of the four locations as shown on the key to the figure and, c) Scatterplots of the discriminant analysis of principal components of the microsatellite data for four *C. tricinctus* locations using geographic sample site as priors for genetic clusters. Individual genotypes appear as dots surrounded by 95% inertia ellipses. Eigenvalues show the amount of genetic information contained in each successive principal component with X and Y axes constituting the first two principle components, respectively. Boxes indicate haplotype (*h*), nucleotide (*% π*) and genetic diversity (*gd*) indices for *C. tricinctus.*

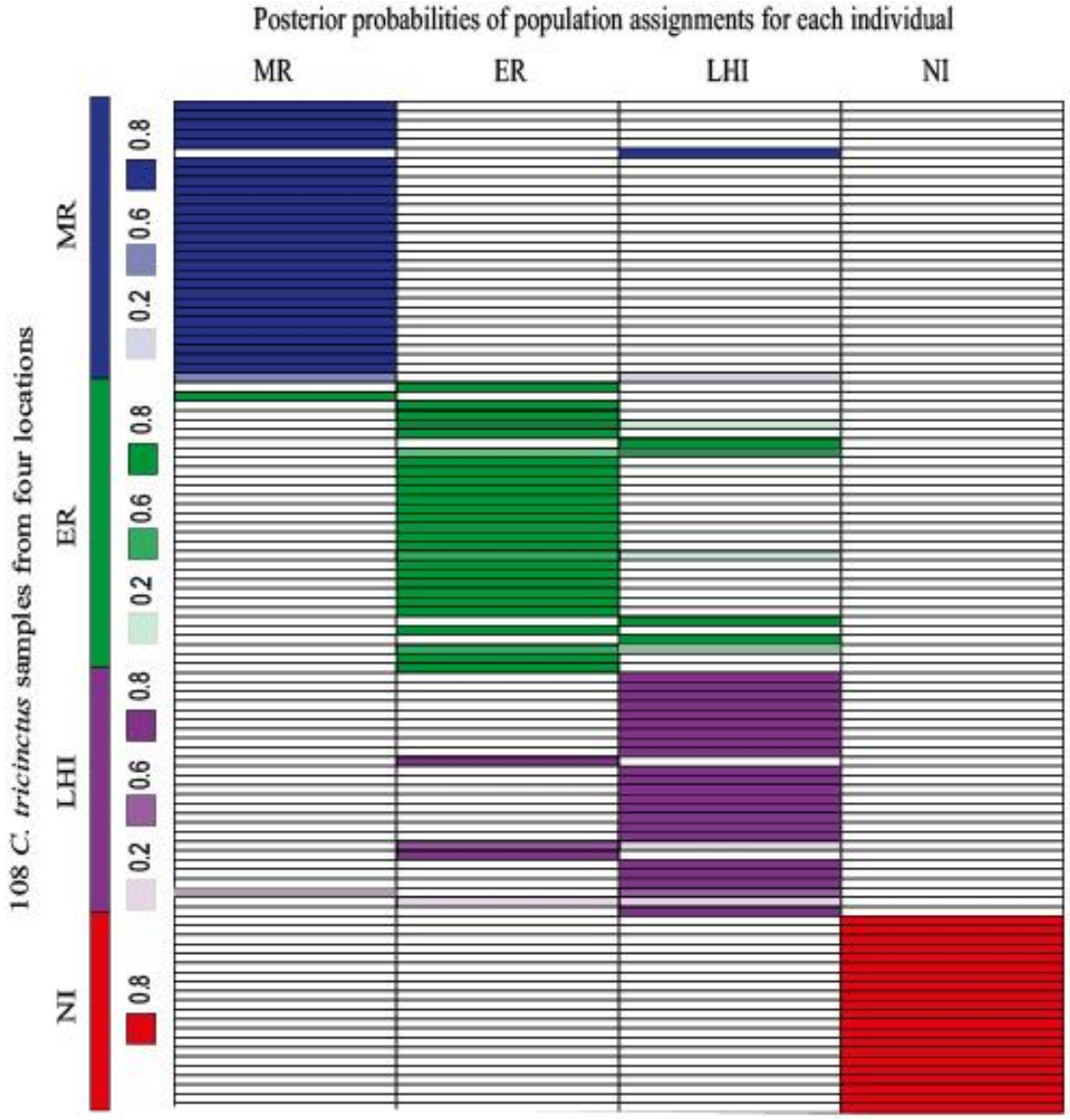

Figure 5.3 Posterior probability of assignment of each individual genotype to four *Chaetodon tricinctus* populations as indicated by DAPC. The names of the possible assignment populations are given on the x-axis. 108 genotypes are listed on the y-axis, along with the population from which they were sampled. Coloured bars correspond to a 0.2 to 0.8 probability of assignment to a given population.

Summary statistics. Heterozygote excess was evident from a negative inbreeding co-efficient (F_{IS}; Table 5.S1); although this was not significant between the four locations. The NIR peripheral location had the most private alleles, 25 across twenty loci, while the remaining three locations ranged from 6 to 15 private alleles across all loci (Table 5.S2). Of the 20 msatDNA loci: (i) significant single locus departures from HWE were detected in 11 of 80 tests at the location level before False Discovery Rate (FDR) correction and nine afterwards (ER: Ct4, Ct23, Ct24; MR: Ct16; NI: Ct3, Ct13, Ct17, Ct23, Ct24; Table 5.S2), similarly, six single locus HWE departures were detected before and after FDR when all locations were considered (Table 5.S2); (ii) null alleles were identified in MR (Ct18) and NI (Ct10, Ct16); and (iii) of the 212 locus by locus exact tests of linkage disequilibrium, 13 were significant before and 10 after FDR correction (Benjamini & Hochberg 1995). Loci that were not in HWE in more than one location (Ct23, Ct24) and had null alleles (Ct10, Ct16, Ct18) were not used in subsequent analysis (ARLEQUIN, STRUCTURE and MIGRATE-n) and, loci in linkage disequilibrium at all sites (Ct17, Ct21, Ct23) were not used in subsequent analysis (ARLEQUIN, STRUCTURE, BAYESASS). Thus 13 loci were used in the ARLEQUIN and STRUCTURE analyses, 15 loci were used in the MIGRATE -n analysis and 17 loci in the BAYESASS analysis.

5.4.1. Gene flow between locations - mtDNA

Patterns and levels of gene flow based on an mtDNA AMOVA indicated significant genetic variation (65.77%) within locations, Φ_{st} = 0.342 (p < 0.001, Table

5.S3). This was due to the NIR peripheral location mtDNA pairwise F_{st} differentiation from all three LHIR locations (pairwise F_{st} = 0.190 to 0.221, $p < 0.001$; Table 6.S4). Whilst there was no genetic differentiation amongst the three LHIR locations (pairwise F_{st} = -0.032 to -0.023, p = 0.865 to 0.991; Table 5.S4). A single regional partition was also suggested between the LHIR region and the NIR peripheral location, explaining 36.45% of the genetic variation, but this was not significant (Φ_{ct} = 0.365, p = 0.250; Table 5.S3).

Quantifying mtDNA gene flow using Migrate-n indicated high levels of mtDNA gene flow between all locations, with M values ranging from 7,035 to 10,385 (Fig 5.4a).

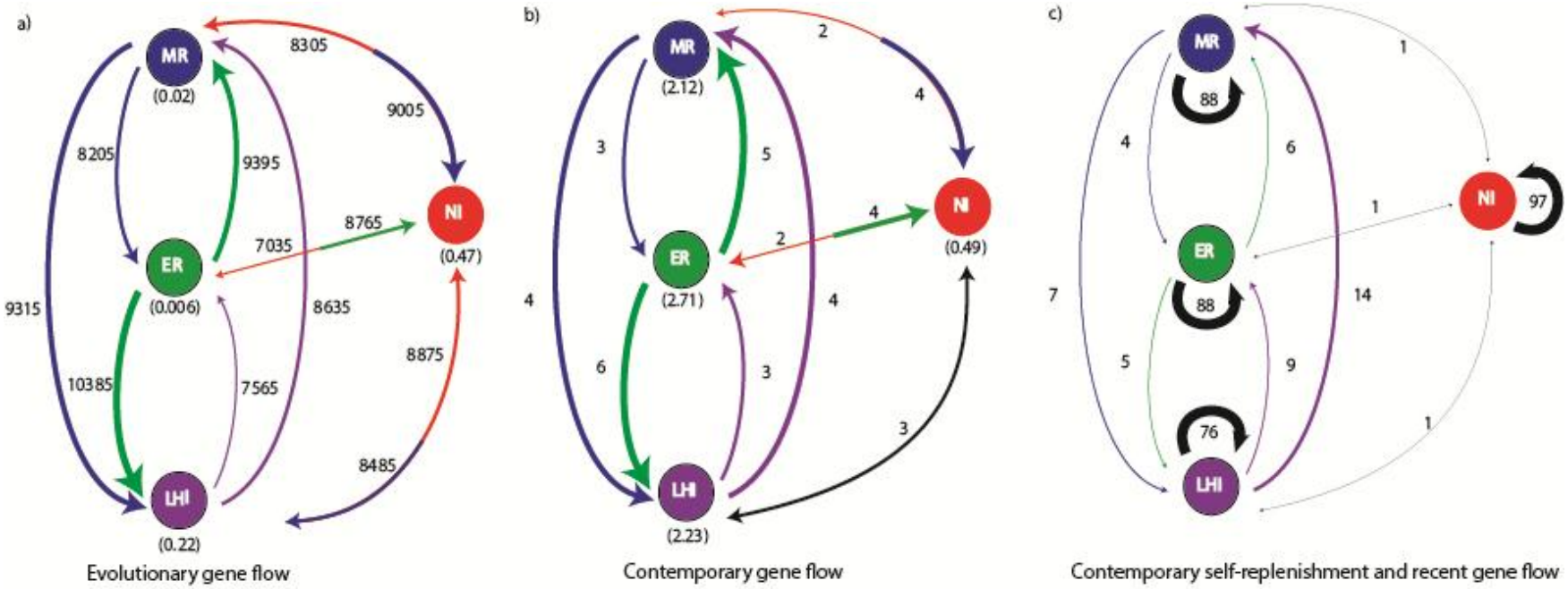

Figure 5.4 Migration rates among *Chaetodon tricinctus* locations. The thickness of the line is directionally proportional to the number of migrants (Nm) and the line colours indicate the predominant direction of gene flow. Population size (θ, within parentheses) is also shown for each location. a) Migrate-n historical gene flow (mtDNA), b) Migrate-n demographic gene flow (msatDNA) and (c) BAYESASS analysis of self-replenishment and recent migration rates (msatDNA) shown as a percentage.

5.4.2. Gene flow between locations - msatDNA

Patterns and levels of gene flow based on an msatDNA AMOVA indicated significant structure in 5 (of 20) locus by locus analyses corrected for null allele frequency (Φ_{st} = 0.001 to 0.368, $p < 0.05$; Table 5.S5), in 6 (of 20) locus by locus analyses corrected for standardised location differentiation (Φ_{st} = 0.004 to 0.852, $p < 0.05$; Table 5.S5) and in the global AMOVA as a weighted average over all microsatellite loci (Φ_{st} = 0.046, $p < 0.001$; Table 5.S3), with 95.39% of the genetic variation existing within locations. Raw msatDNA pairwise F_{st} comparisons showed very low to moderately significant genetic partitioning between the LHIR locations and the NIR peripheral location (F_{st} = 0.056 to 0.101, $p < 0.001$). In contrast, an Excluding Null Alleles (ENA) corrected msatDNA pairwise F_{st} value showed no significant genetic differentiation between any of the four locations as estimates of genetic differentiation between locations fell within 95% confidence intervals (F_{st} = 0.005 to 0.084, $p > 0.05$; Table 5.S4).

DAPC, GeneClass2 and STRUCTURE confirmed the presence of at least 3 distinct genetic populations corresponding to geographic location. Discriminant analysis of principal components (DAPC) partitioned *C. tricinctus* into the LHIR region and the NIR peripheral location (Fig 5.2c). Using the four locations as *a priori* population criteria, DAPC assigned 58 to 100% of all individuals to the location from which they were sampled (assignment per population, ER = 74%, MR = 90%, LHI = 58%, NI = 100%; Fig 5.3). Consistent with these assignments, with the allele frequencies and

genotypic assignments, the 95% Genotypic Inertia Ellipses (GIE) for ER and LHI overlap, whilst the 95% GIE for MR does not overlap with either ER or LHI and the 95% GIE for NI occupy a distant area of multivariate space, along the x-axis, from all three LHIR locations. Geographical structure in msatDNA data was confirmed by GeneClass2 analyses, where only 11 individuals were grouped with a location from which they were not sampled (MR = 1, ER = 7, LHIL = 3); thereby identifying 4 genetically differentiated populations. Similarly, four geographically partitioned populations were identified by STRUCTURE analyses, as the likelihood of the marginal posterior probability distribution was highest when $K = 4$.

Quantifying the level of msatDNA gene flow. Migrate-n indicated a few orders of magnitude lower levels of demographic gene flow between locations when compared to mtDNA gene flow, with M values ranging from 2 to 6 (Fig 6.4b).

Inferred levels of self replenishment and migrant exchange. Demographic independence is suggested for all location pairs except: LHI to MR (m = 14%) and possibly LHI to ER (m = 9%; Fig 5.4c). Conversely, high levels of self-replenishment (76 to 96%) were inferred at all four locations (Fig 5.4c). This further indicates that in the short term, each population is predominantly sustained by self-replenishment rather than replenishment from distant populations.

5.4.3. Population genetic diversities

Chaetodon tricinctus showed low haplotype diversity (h) and nucleotide diversity (% π) in all three LHIR locations (h = 0.173 to 0.251,% π = 0.1), whilst the NIR peripheral

location had four- to eight- fold higher *h* and % π (*h* = 0.925,% π = 0.8; Fig 5.2c), respectively. Three of the four locations (MR, LHI, NI) had high genetic diversity (*gd* = 0.515 to 0.589), ER being the exception (*gd* = 0.454; Fig 5.2c). Total haplotype, nucleotide and genotypic diversities were low (*h* = 0.384, %π = 0.2, *gd* = 0.490; Table 5.S1).

5.5. DISCUSION

Understanding both time- and spatial- scales of gene flow and the levels of genetic diversity, is vital to determine best practice management, maximise biodiversity conservation and evaluate the capacity of coral reef fishes to recover should they become locally extinct. In this study, *Chaetodon tricinctus* was found to have (i) sufficient mtDNA gene flow connecting all locations within the LHIR region, but low gene flow and consequent isolation of the NIR peripheral population from the LHIR locations; (ii) low msatDNA gene flow between all locations resulting in populations that are genetically differentiated; (iii) demographic dependence between LHI and MR (and possibly LHI and ER), yet high levels of inferred self-replenishment at all four locations; (iv) variable genetic diversities: low mtDNA genetic diversity at all three locations within the LHIR region, but not at the NIR peripheral location; and (v) high msatDNA genetic diversity at all four locations.

5.5.1. Gene flow between locations - mtDNA

Monophyly was suggested for *C. tricinctus* with the exception of 2 clades, one of which consisted exclusively of the NIR peripheral location. The lack of geographic population structure within the LHIR region may result if a small number of recruits per generation maintain spatial genetic homogeneity (Shulman 1998, Planes 2002). Similar genetic homogeneity has been found in studies on the endemic Hawaiian butterflyfishes *Chaetodon multicinctus, C. miliaris and C. fremblii (Craig et al. 2010)* and in numerous other coral reef fish species including the parrotfishes *S. frenatus* and *C. sordidus (*Dudgeon et al. 2000, Bay et al. 2004). In contrast, the genetic differentiation between the LHIR region and the NIR peripheral location likely results from limited gene flow due to geographic isolation (600 km of deep ocean separating each

region) and complicated ocean currents (which are seasonally stronger or weaker flowing west to east along the East Australian Current and seasonally migrating north or south along the Tasman Front; Suthers et al. 2011). Such strong genetic breaks at peripheral locations has been demonstrated in other reef fishes, including two widespread coral reef snappers *Lutjanus kasmira* and *L. fulvus* (Gaither et al. 2010) and two widespread parrotfishes *Scarus psittacus* (Winters et al. 2010) and *Chlorurus sordidus* (Bay et al. 2004).

Despite high mtDNA gene flow between all locations, conventional statistics (AMOVA and pairwise F_{st}) indicate that the three locations within the LHIR region and the NIR peripheral location are genetically differentiated. Although high mtDNA gene flow may provide some assistance to distant populations through recolonisation following local extinctions and increasing genetic diversity (Hanski 1999, Jones et al. 2009), benefits to the maintenance of distant populations may be minimal, especially if combined with little or no demographic gene flow on demographic time scales.

5.5.2. Gene flow between locations - msatDNA

C. tricinctus showed demographic genetic differentiation between all locations (with the possible exception of ER and LHI). The strong discrepancy between mtDNA and demographic levels of gene flow in *C. tricinctus* is increasingly being documented in other coral reef fishes such as snappers *Lutjanus carponotatus* (Evans et al. 2010, Harrison et al. 2012) and *Lutjanus synagris* (Gold et al. 2011), coral trout *Plectropomus maculatus* (Evans et al. 2010, Harrison et al. 2012) and in the endemic Lord Howe Island anemonefish *Amphiprion mccullochi* (van der Meer et al. 2012b; Chapter 4). This 'lack of congruence' between timescales may result from genetic homogeneity over historical time scales (Shulman 1998, Planes 2002) compared to substantial

amounts of self-recruitment over demographic time scales (Swearer et al. 1999, Jones et al. 2005, Almany et al. 2007, Planes et al. 2009).

The estimation of demographic gene flow is important for conservation because models predict that a few recruits per generation over historical time scales will not sustain populations (Cowen et al. 2000, Cowen et al. 2002). In light of this, Marine Protected Area's (MPAs) are designed to be large enough for locations to sustain themselves and yet spaced close enough so that larvae produced within an MPA can potentially be exported to unprotected areas (see Halpern & Warner 2003, Shanks et al. 2003, Jones et al. 2005, Harrison et al. 2012). In the case of *C. tricinctus*, it is unlikely that the current MPAs in the LHIR region will deliver any substantial recruitment to the NIR peripheral location due to the geographic isolation and complicated ocean currents around the Lord Howe Island and Norfolk Island Rise regions and, the high levels of larval retention possibly facilitated by natal homing (Botsford 2005, Hilborn et al. 2006). The high abundance of *C. tricinctus* at the LHIR locations (Choat et al. 2006, Hobbs et al. 2009) reduces the likelihood of local extinction, while higher levels of demographic gene flow, when compared to the NIR peripheral location, are likely to facilitate recovery following population declines (or local extinction). Given the extremely small population size of *C. tricinctus* at Norfolk Island (estimated to be less than 30 individuals), a slow recovery time is expected following population declines, due to intermittent pulse replenishment.

Less than 10% gene flow between populations suggests demographic independence (Waples & Gaggiotti 2006) and high levels of self-recruitment, which is vital for populations to persist (Hastings & Botsford 2006). However, trying to classify populations as ‘open’ or ‘closed’ may not be appropriate (Largier 2003, Mora & Sale 2002). Rather, populations that have 80% of the successful recruits generated internally, will take substantially longer to recover following

local extinction than ones with only 20% self-recruitment (Miller & Shanks 2004) and should be considered *largely closed* or *largely open*, respectively. All locations appear demographically independent and may be considered largely closed. However, both MR and to a lesser extent ER, receives some gene flow from LHI, suggesting that the population at LHI is important for management and continued protection because it exports individuals to MR and ER. Levels of inferred self-replenishment found in *C. tricinctus* (≥ 76%) are highly similar to the estimated levels of self-recruitment in other congeneric butterflyfish in Papua New Guinea (PNG, Almany et al. 2007) and other island populations of coral reef fishes (Swearer et al. 1999, Jones et al. 2005, Planes et al. 2009). The consistency of results between the indirect methods of the present study and the direct methods of former studies suggest that self-replenishment can be used to approximate self-recruitment in coral reef fish populations, given a sufficient number of unlinked loci, high detectable levels of self-replenishment and no unsampled ghost locations. Moreover, estimates of self-replenishment in *C. tricinctus* tended to be slightly higher than estimates of self-recrutiment in the above studies. It is unlikely that this difference is due entirely to methodological considerations, given that indirect genetic methods are thought to overestimate gene flow (Hellberg et al. 2002). Rather, high levels of self-recruitment in C. tricinctus might be inherent of its small geographic range and thus high self-recruitment is needed to sustain isolated populations in this system (Hobbs et al. 2011). Alternatively, studies sampling new recruits are also estimating self-recruitment during the post-settlement mortality period. Consequently, the genetic makeup of the recruit cohort that survives through to adulthood is changed creating a disparity between estimates of self-replenishment and self-recruitment

5.5.3. Population genetic diversities

C. tricinctus showed high mtDNA genetic diversity at the NIR peripheral location and low diversity at the LHIR locations. While msatDNA genetic diversity was high at three of the four locations (MR, LHI, NI), ER being the exception. Similar genetic diversities have been found in other coral reef fish using cytochrome b including the endemic Hawaiian butterflyfishes *Chaetodon fremblii, C. miliaris* and *C. multicinctus* (Craig et al. 2010) and in the more widespread butterflyfishes *C. lunulatus, C. trifascialis* and *C. trifasciatus* (Lawton et al. 2011, Montanari et al. 2011). Species with high genetic diversity may have some resilience to extinction as a decrease in genetic diversity is generally associated with decreased fitness (Hoelzel et al. 2002). Thus the high overall genetic diversity at the NIR peripheral location, resulting from pulse recruitment periodically bringing new genetic material into the population resident here and the occurrence of rare haplotypes (see below), is encouraging, since it may buffer a small, demographically isolated population against some impacts. However the reverse patterns occur at the LHIR locations, where the risk of extinction associated with low genetic diversity is counteracted by high population abundances. Of interest are the rare haplotypes seen only at the NIR peripheral location that may represent either historical polymorphisms (a relic or refugium population) or mutation accumulation. Given the high abundance of *Acropora* at this location (authors unpublished data) but extremely low abundance of *C.tricinctus* individuals, it is likely that self-recruitment is limiting population numbers. If unique genetic diversity is a feature of NI populations of endemics within the region, then protecting these populations and the habitats they rely on is vitally important.

Conclusion Given the low demographic gene flow between LHIR and NIR peripheral locations (and high self-replenishment) in *C. tricinctus,* the MPAs at the LHIR region are of limited

benefit to the unprotected NIR peripheral location (NI). Therefore, the NIR peripheral location requires some protective management strategies to conserve its genetically unique population of *C. tricinctus*. However, within the LHIR region, LHI is an important source of gene flow to both MR and ER and as such, warrants continued MPA protection and monitoring. Similar patterns of gene flow between locations has also been found for the endemic McCulloch's anemonefish, *Amphiprion mccullochi* (van der Meer et al. 2012a, b; Chapter 3, 4) and may be indicative of generalised patterns of gene flow of all endemics in the region. Given the importance of the LHI region as an endemism hotspot, determining patterns of gene flow across a number of endemic species with varying biological and ecological characteristics will be crucial for developing effective conservation management strategies.

Supplementary Online Material

Table 5.S1: Sample sizes for mtDNA (cytochrome b, total $n = 97$)

	n (Cyt b)	n_h	h	π(%)	n (msat)	gd	Ave N_a	P_a	H_o	H_e	F_{is}
All	97	15	0.384	0.2	108	0.490	6.15		0.662	0.634	-0.052
ER	29	3	0.197	0.1	31	0.454	6.35	11	0.699	0.656	-0.041
MR	30	4	0.251	0.1	30	0.568	6.30	15	0.652	0.635	-0.008
LHI	22	2	0.173	0.1	26	0.515	5.85	6	0.663	0.631	-0.028
NI	16	11	0.925	0.8	21	0.589	6.10	25	0.670	0.646	-0.036

cyt b, cytochrome b.

Number of samples (n), number of haplotypes (n_h), haplotype diversity (h), nucleotide diversities (π) of cyt b for *Chaetodon tricinctus* from all locations. Sample sizes for msats (total n = 108), genetic diversity (gd), average number of alleles per locus (Na), observed number of private alleles (Pa), observed heterozygosity (H_o), expected heterozygosity (H_e) and the inbreeding coefficient (F_{IS}) averaged over twenty-one microsatellite loci for four locations.

Table 5.S2. Summary statistics for twenty microsatellite loci (Ct2-24) from *Chaetodon tricinctus*. Sample sizes (N), observed number of alleles (Na), observed number of private alleles (Pa), observed heterozygosity (H_o), expected heterozygosity (H_e), the average inbreeding coefficient (F_{IS}), probability of departure from HWE (pHWE, Peakall and Smouse 2006) and significance of departure after FDR correction FDR (pFDR, Storey 2002) for each locus at all and each location (significance of departure in bold, $p < 0.05$).

Population	Ct2	Ct3	Ct4	Ct5	Ct7	Ct8	Ct9	Ct10	Ct11	Ct12	Ct13	Ct14	Ct16	Ct17	Ct18	Ct20	Ct21	Ct22	Ct23	Ct24
All (108)																				
N	71	60	77	62	70	72	79	73	70	77	62	78	75	68	77	66	72	69	76	80
Na	4	6	8	6	11	8	6	6	11	5	11	7	9	8	5	10	16	10	6	5
Pa	1	1	2	1	4	4	1	3	2	1	2	1	4	3	1	2	3	2	2	2
Ho	0.423	0.400	0.675	0.629	0.986	0.694	0.709	0.521	0.857	0.558	0.823	0.705	0.600	0.721	0.403	0.758	0.833	0.826	0.868	0.175
He	0.465	0.400	0.650	0.688	0.777	0.631	0.640	0.524	0.758	0.586	0.824	0.745	0.574	0.722	0.420	0.812	0.882	0.857	0.658	0.217
F_{IS}	0.092	0.000	-0.039	0.085	-0.269	-0.100	-0.107	0.008	-0.130	0.047	0.002	0.053	-0.045	0.003	0.042	0.067	0.055	0.036	-0.321	0.193
pHWE	0.092	0.765	**0.000**	0.251	0.040	0.926	0.770	0.877	0.995	0.943	**0.000**	0.646	0.150	0.715	0.663	0.565	**0.000**	**0.003**	**0.000**	**0.000**
pFDR	0.229	0.951	**0.000**	0.479	0.120	0.990	0.951	0.990	0.995	0.990	**0.000**	0.951	0.315	0.951	0.951	0.951	**0.000**	**0.011**	**0.000**	**0.000**
ER (31)																				
N	19	19	24	20	20	20	28	24	16	26	21	26	24	18	25	14	22	15	24	29
Na	3	5	7	6	9	4	5	5	7	5	11	6	5	5	4	7	14	9	5	4
Pa	0	0	1	1	3	0	0	1	0	1	2	0	2	0	0	1	1	1	1	2
Ho	0.474	0.421	0.708	0.750	1.000	0.650	0.714	0.583	0.813	0.577	0.905	0.692	0.708	0.778	0.600	0.786	0.773	0.867	0.917	0.103
He	0.514	0.357	0.713	0.681	0.764	0.625	0.679	0.588	0.752	0.577	0.864	0.702	0.577	0.702	0.473	0.770	0.870	0.849	0.661	0.162
F_{IS}	0.078	-0.178	0.006	-0.101	-0.309	-0.040	-0.052	0.007	-0.081	0.000	-0.047	0.014	-0.227	-0.108	-0.269	-0.020	0.112	-0.021	-0.388	0.363
pHWE	0.655	0.999	**0.001**	0.668	0.893	0.389	0.897	0.465	0.807	0.643	0.096	0.582	0.776	0.567	0.597	0.559	0.137	0.155	**0.000**	**0.000**
pFDR	0.647	0.737	**0.005**	0.647	0.695	0.647	0.695	0.647	0.694	0.647	0.307	0.647	0.694	0.647	0.647	0.647	0.343	0.343	**0.000**	**0.000**

Table 5.S2 *(Continued)*

Population	Ct2	Ct3	Ct4	Ct5	Ct7	Ct8	Ct9	Ct10	Ct11	Ct12	Ct13	Ct14	Ct16	Ct17	Ct18	Ct20	Ct21	Ct22	Ct23	Ct24
MR (30)																				
N	30	24	30	24	28	30	27	29	30	30	24	28	30	28	28	29	28	30	28	27
Na	3	5	7	5	7	7	6	4	9	4	8	7	7	7	4	9	13	8	4	3
Pa	1	1	1	0	1	3	1	1	0	0	0	1	2	2	1	0	1	0	0	0
Ho	0.300	0.417	0.633	0.583	0.964	0.733	0.741	0.586	0.833	0.533	0.667	0.750	0.567	0.643	0.214	0.793	0.857	0.833	0.929	0.185
He	0.383	0.444	0.618	0.694	0.654	0.654	0.628	0.499	0.767	0.563	0.778	0.736	0.553	0.719	0.339	0.822	0.889	0.846	0.675	0.230
F_{IS}	0.216	0.063	-0.024	0.159	-0.474	-0.121	-0.179	-0.175	-0.086	0.052	0.143	-0.019	-0.025	0.106	0.367	0.035	0.036	0.015	-0.375	0.194
pHWE	0.544	0.821	**0.047**	0.446	0.280	0.950	0.993	0.532	0.983	0.307	0.945	0.254	**0.000**	0.728	0.004	0.860	0.006	0.438	0.002	0.621
pFDR	0.879	0.993	0.197	0.879	0.806	0.993	0.993	0.879	0.993	0.806	0.993	0.806	**0.000**	0.993	0.028	0.993	0.032	0.879	0.021	0.932
LHI (26)																				
N	22	17	23	18	22	22	24	20	24	21	17	24	21	22	24	23	22	24	24	24
Na	3	3	5	5	7	5	5	4	10	4	8	6	4	6	4	9	11	9	5	3
Pa	0	0	0	0	0	1	0	0	2	0	0	0	0	1	0	0	1	1	1	0
Ho	0.545	0.353	0.696	0.556	1.000	0.682	0.667	0.350	0.917	0.571	0.941	0.667	0.524	0.773	0.417	0.696	0.864	0.792	0.750	0.250
He	0.501	0.372	0.602	0.603	0.791	0.569	0.595	0.471	0.708	0.602	0.785	0.739	0.541	0.723	0.433	0.784	0.856	0.845	0.613	0.260
F_{IS}	-0.089	0.051	-0.155	0.079	-0.264	-0.198	-0.121	0.257	-0.294	0.051	-0.198	0.098	0.031	-0.069	0.038	0.112	-0.008	0.063	-0.224	0.037
pHWE	0.836	0.935	0.624	**0.039**	0.247	0.846	0.643	0.701	1.000	0.157	0.192	0.288	0.947	0.436	0.334	0.495	0.681	0.184	0.126	0.441
pFDR	0.987	0.994	0.920	0.410	0.741	0.987	0.920	0.920	1.000	0.672	0.672	0.756	0.994	0.842	0.779	0.866	0.920	0.672	0.672	0.842

Table 5.S2 *(Continued)*

Population	Ct2	Ct3	Ct4	Ct5	Ct7	Ct8	Ct9	Ct10	Ct11	Ct12	Ct13	Ct14	Ct16	Ct17	Ct18	Ct20	Ct21	Ct22	Ct23	Ct24
NI (21)																				
N	17	21	20	21	21	20	21	17	20	17	21	20	17	21	20	20	21	20	16	21
Na	4	5	9	6	1	5	7	3	8	4	9	5	6	11	6	6	9	8	6	4
Pa	1	1	3	1	1	1	1	0	2	0	0	0	6	2	0	0	0	2	1	0
Ho	0.529	0.333	0.950	0.524	0.000	0.650	0.667	0.294	0.600	0.588	0.810	0.800	0.529	0.952	1.000	0.650	0.714	0.800	1.000	1.000
He	0.507	0.465	0.738	0.592	0.000	0.625	0.604	0.524	0.756	0.637	0.831	0.735	0.727	0.800	0.646	0.760	0.813	0.843	0.693	0.625
F_{IS}	-0.044	0.283	-0.288	0.115	0.000	-0.040	-0.103	0.439	0.207	0.076	0.026	-0.088	0.271	-0.190	-0.547	0.145	0.121	0.050	-0.442	-0.601
pHWE	0.981	**0.000**	0.997	0.660	na	0.853	0.981	0.062	0.014	0.565	**0.000**	0.143	**0.018**	0.199	0.172	0.109	0.202	0.204	**0.001**	**0.002**
pFDR	0.997	**0.000**	0.997	0.836	na	0.997	0.997	0.168	0.053	0.767	**0.000**	0.298	0.057	**0.298**	0.298	0.259	0.298	0.298	**0.006**	**0.01**

Table 5.S3: AMOVA analysis for a) mtDNA (Cyt b) sequences from *Chaetodon tricinctus* structured into the westernmost location (MR, ER, LHI) vs the NIR peripheral location (NI) and b) global AMOVA weighted across all twenty microsatellite loci.

Source of variation	Variance component	Percentage of variation	F-statistics fixation Indices (p-value)
a) Location			
Among groups	0.147	36.45	F_{ct} = 0.365 (p = 0.250)
Among populations within locations	-0.009	-2.22	F_{sc} = -0.035 (p = 0.983)
Within populations	0.266	65.77	F_{st} = 0.342 (p < 0.001)
b) Microsatellite			
Among groups	0.24	3.53	F_{ct} = 0.035 (p = 0.047)
Among populations within locations	0.07	1.07	F_{sc} = 0.011 (p = 0.004)
Within populations	6.46	95.39	F_{st} = 0.046 (p < 0.001)

Table 5.S4. Pairwise population structures (F_{st}) generated for mtDNA (cytochrome b, n = 97) and for twenty microsatellite loci (n = 108) from four *Chaetodon tricinctus* locations showing raw and corrected F'_{st} for null allele frequencies.

	cyt b				Raw msat				Corrected msat			
	ER	MR	LHI	NI	ER	MR	LHI	NI	ER	MR	LHI	NI
ER		0.865	0.991	**0.000**		0.991	0.413	**0.000**		p > 0.05	p > 0.05	p > 0.05
MR	-0.023		0.991	**0.009**	-0.040		0.991	**0.000**	0.011		p > 0.05	p > 0.05
LHI	-0.032	-0.031		**0.009**	-0.030	0.001		**0.000**	0.005	0.006		p > 0.05
NI	0.216	0.221	0.190		0.056	0.101	0.088		0.070	0.084	0.072	

Raw locations differentiation from microsatellite allele frequencies and associated *p*-values using Arlequin ver 3.5 (Excoffier et al. 2005). Corrected location differentiation for null allele frequencies at 95% CI (all *p*-values > 0.05) using the ENA correction of Chapuis and Estoup (2007). FDR correction of raw F'_{st} *p*-values in QVALUE (Storey 2002). The use of both p-values and Confidence Intervals (CI) is responsible for the discrepancy between the FST values. While the p-value is used to test if the sample is significant or not from zero (i.e. is the FST different from zero), the CI is used to construct confidence around the estimate [i.e.is there significant genetic differentiation (FST) between locations].Thus using the raw msatDNA data I find that the p-value is significant from zero. However, when I correct for null alleles (ENA), our confidence estimate suggests that there is no genetic differentiation (FST) between the locations.

Table 5.S5: AMOVA fixation indices (F'_{st}) for *Chaetodon tricinctus* across all locations surveyed

	Marker class and analysis		
	Raw msat	Msat corrected for null allele freq.	Standardised msat (F'_{st})
Average	**0.046**	0.040	0.113
Ct2	0.001	-0.001	**0.004**
Ct3	-0.006	-0.004	-0.017
Ct4	0.000	**0.004**	**0.006**
t5	0.019	0.025	0.068
Ct7	**0.397**	**0.368**	**0.852**
Ct8	-0.001	-0.002	-0.005
Ct9	-0.006	-0.005	-0.014
Ct10	-0.005	-0.006	-0.017
Ct11	0.003	**0.005**	**0.013**
Ct12	**0.052**	0.036	0.091
Ct13	0.016	0.011	0.073
Ct14	0.011	0.018	0.070
Ct16	**0.035**	0.040	0.091
Ct17	0.024	0.019	0.077
Ct18	**0.077**	0.059	0.130
Ct20	0.001	-0.005	-0.016
Ct21	0.001	-0.002	-0.017
Ct22	0.006	**0.001**	**0.007**
Ct23	**0.043**	0.039	0.117
Ct24	0.184	**0.127**	**0.222**

Raw locations differentiation from microsatellite allele frequencies for each individual locus and as the average across all loci, locations differentiation corrected for null allele frequencies using the ENA correction of Chapuis & Estoup (2007) and standardized locations differentiation for and across all loci (F'_{st}). All values in bold are significant to the $p<0.05$ (i.e. 95% confidence interval).

CHAPTER 6: **Population connectivity and the effectiveness of a marine reserve network in protecting a vulnerable, exploited and endemic coral reef fish**

(In preparation for submission, word for word: van der Meer MH, Berumen ML, Hobbs J-PA, van Herwerden L. Population connectivity and the effectiveness of a marine reserve network in protecting a vulnerable, exploited and endemic coral reef fish.

6.1. ABSTRACT

Multiple anthropogenic impacts, including overfishing, have negatively impacted coral reefs worldwide. Marine Protected Areas (MPAs) have been widely advocated to mitigate these impacts by conserving biodiversity and preventing overfishing. MPA effectiveness depends on population connectivity patterns between protected and non-protected areas. Remote islands are endemism hotspots for coral reef fishes and also provide rare examples of coral reefs with limited fishing pressures. This study explored population connectivity across a network of protected and non-protected areas for an endemic wrasse, *Coris bulbifrons*, IUCN listed as "vulnerable" due to its small, declining geographic range and declining abundance. We used mt- and msat-DNA (microsatellites) to distinguish between historic and demographic gene flow, estimate self-replenishment and measure genetic diversity among all locations in the species range – the remote Australian offshore locations of Middleton Reef (MR), Elizabeth Reef (ER), Lord Howe Island (LHI), and Norfolk Island (NI). MtDNA detected no genetic differentiation among locations, whilst msatDNA suggested genetic differentiation between the most peripheral

(NI) and all remaining locations (MR, ER, LHI). Despite high gene flow between locations (mtDNA = 259–1144), msatDNA based gene flow was limited (*Nm* = 3–9), with high self-replenishment (68–93%) at all locations. Limited demographic gene flow among locations is seen in other endemic fishes, indicating that a network of MPAs that protect part of each location would provide adequate protection for endemic species. Existing MPAs at MR and LHI appear adequate, but greater protection is required at ER and NI. For fishes vulnerable to overfishing, such as *C. bulbifrons*, population monitoring in areas open to fishing is required to determine if fishing restrictions need adjustment to prevent overfishing.

6.2. INTRODUCTION

Coral reefs worldwide have been impacted by disease, hurricanes, human overpopulation, eutrophication and global climate change (Hughes et al. 2003; Bellwood et al. 2004; Lesser 2007). However, fishing has had the most direct and wide-ranging influence on coral reefs and other marine ecosystems (Jennings & Kasier 1998, Jackson et al. 2001, Wilson et al. 2010). Over the past 500 years, fishing has systematically removed large marine species from the world's oceans (NRC 1995, Jackson et al. 2001, Pitcher 2001, Block 2011) to an extent where the populations of some important species (such as apex predators, Block 2011) are functionally absent with numbers too low for them to be able to perform their former ecological role (Dayton 1998). With anthropogenic pressures increasing (Steffen et al. 2007), natural resource managers have established Marine Protected Areas (MPAs) to conserve biodiversity and protect fisheries stocks from overfishing. Fisheries managers also use restrictions of catch (gear restrictions or imposition of quotas) and effort (closed seasons, temporary area closures, limited numbers of fishers or vessels) to prevent overfishing (Pitcher & Hart 1982, Worm et al. 2009). Although these approaches have been successful in many cases, they have not always prevented overexploitation of stocks or even collapse of some fisheries (Roughgarden & Smith 1996, Mullon et al. 2005). When designing MPAs, managers rarely have empirical evidence of levels of connectivity or gene flow among locations, forcing implementations based on "best guesses" (McCook et al. 2009).

Remote islands often provide rare examples of coral reefs with limited fishing pressure and provide a unique opportunity to assess how coral reef ecosystems function without major human impacts. For example, the remote and lightly fished North-West Hawaiian Islands support

significantly more fish biomass than the heavily fished main Hawaiian Islands (Friedlander & DeMartini 2002). Remote islands are also hotspots for coral reef fish endemicity (Jones et al. 2002, Roberts et al. 2002), with a high proportion of their communities comprised of endemic species. Whilst terrestrial endemics on remote islands are well known for their vulnerability to overexploitation (Whittaker 1998), much less is known about the vulnerability of their marine counterparts. In the Indo-Pacific, the locations with the greatest proportion of endemic coral reef fishes include: Hawaii (25% endemism), Easter Island (22.2%), the Marquesas (11.6%), Lord Howe and Norfolk Islands (7.2%) and Rapa (5.5%) (Randall 1998; 2001; 2007). The Lord Howe Island region in the south-west Pacific Ocean consists of four oceanic features: Middleton Reef (MR), Elizabeth Reef (ER), Lord Howe Island (LHI) and Norfolk Island (NI) (Fig 1). These remote locations harbour tropical habitats dominated by scleractinian corals that transition into temperate habitats dominated by macroalgae (Crossland 1988, Johannes et al. 1983). The remoteness and transition between habitats makes these islands and reefs endemism hotspots for both coral and algae reef fishes (Marine Parks Authority 2010).

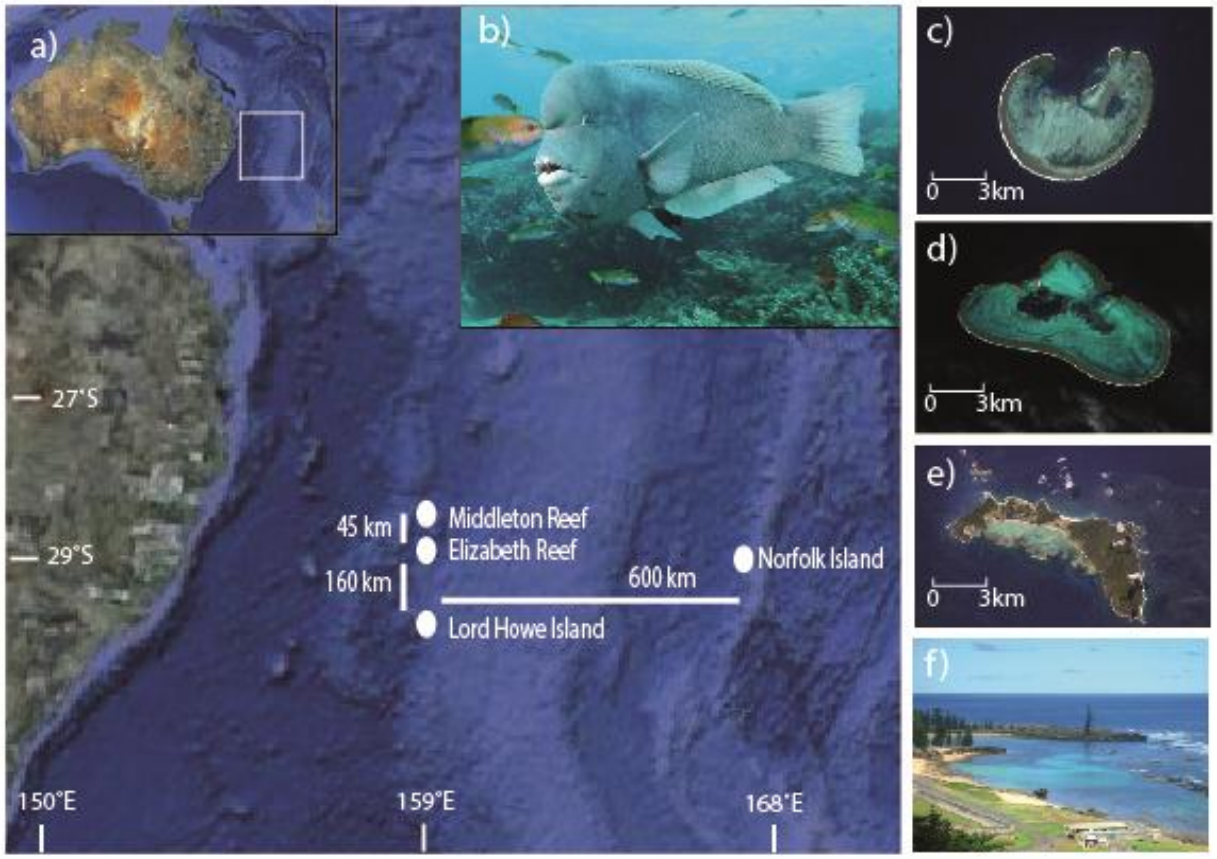

Figure 6.1 Location maps and focal species. (A) Google Earth image of eastern Australia showing Middleton Reef (MR), Elizabeth Reef (ER), Lord Howe Island (LHI), and Norfolk Island (NI) in the South West Pacific Ocean. (B) *Coris bulbifrons* swimming in the open (Photo courtesy of Justin Gilligan). Aerial photographs of MR (C); ER (D), LHI (E) and NI (F; the bay measures 1km in length).

The doubleheader wrasse (*Coris bulbifrons*, Randall & Kuiter 1982) is an iconic reef fish endemic to the Lord Howe region (MR, ER, LHI, NI; Francis 1993). This large wrasse (maximum total length = 65 cm, Choat et al. 2006a) is targeted for food in recreational fisheries, especially at LHI and ER. It is locally abundant in sheltered habitats at MR, ER and LHI (Choat et al. 2006a; Hobbs & Feary 2007; Hobbs et al. 2009), but rare at NI (*authors pers. obs. MvdM, J-PAH*). Furthermore, there are clear colour differences between regions with the NI population showing distinct patterns and bands around the eye, not found elsewhere (*authors pers. obs. MvdM, J-PAH*). The current network of MPAs provides a diversity of protection to *C. bulbifrons*: MR is fully protected with no fishing allowed, ER allows recreational line and spearfishing but no commercial or charter fishing. In state waters at LHI, no take zones represent 27% of the area, and outside of these line fishing is allowed with a limit of one *C. bulbifrons* per person per day but no size limit. No-take zones cover 31% of commonwealth waters at LHI and outside of these zones spear and line fishing are allowed with a bag limit of 20 fish per person per day. At NI, there are no MPAs and no restrictions on fishing for *C. bulbifrons*.

C. bulbifrons is listed as vulnerable by the IUCN largely due to (i) the small area of occupancy (< 2000 km^2), (ii) the severely fragmented distribution (occurs at only 4 isolated locations) (iii) a declining area of occupancy and (iv) the declining number of mature individuals (Choat & Pollard 2010). Determining the temporal and spatial scales of connectivity (gene flow) between the four locations (incorporating the whole range of this species distribution) is essential to establish how populations are maintained and replenished, and, should local extinction occur, if there are any possible rescue effects. It is clear that this vulnerable species requires effective

conservation management, especially in light of evidence demonstrating rapid population declines in large wrasses following minimal fishing (Choat et al. 2006b).

Complete sampling throughout the whole range of a species distribution is rare in the coral reef fish literature despite its importance to: accurately estimate gene flow (historic and demographic); establish which locations export and import migrants; and identify those subpopulations that have the highest genetic diversity (as an indicator of potential resilience to environmental change and extinction). This information can then be used by management agencies to target conservation efforts by determining which location(s) need the greatest level of protection. Previous population genetic research on two other endemic species found that within this system (MR, ER, LHI, NI), the anemonefish *Amphiprion mccullochi* and the butterflyfish *Chaetodon tricinctus*, showed similar patterns and levels of gene flow, self-replenishment and genetic diversity, despite different life history traits (van der Meer et al 2012a; b; 2013). Combining these previous studies with the present one, provides an important framework to test population connectivity across the same study system using species that differ substantially in life history traits, mating systems, dispersal abilities and ecologies. Determining whether patterns of population connectivity within this restricted distribution range are the same for these different fishes is important to establish if a single spatial management strategy would protect a wide range of species.

The primary aims of this study were threefold, to estimate for *C. bulbiforns*: (i) patterns and levels of gene flow among locations/subpopulations using mtDNA and msatDNA; (ii) levels of self-replenishment (as a proxy for self-recruitment) based on msat DNA assignment and

exclusion analyses; (iii) population genetic diversities at all locations/subpopulations as a measure of genetic resilience to environmental change. We then place these results into a general framework that compares gene flow patterns among different species from this endemism hotspot. This information will help determine if a single spatial management stratergy is feasible for conserving endemic reef fishes within the Lord Howe region.

6.3. MATERIAL AND METHODS

We applied a range of frequency and Bayesian based molecular tools to establish phylogenetic (mtDNA) and population genetic (mt- and msatDNA) structure as well as levels of self-replenishment (msatDNA). This achieved a reliable estimate of gene flow in this study system and together these tools provided a holistic view of dispersal and retention over time (Leis et al. 2011). However, due to the large number of analyses, we present only methods specific to this study below, whilst standard Materials and Methods (i.e. genetic and laboratory techniques) follow those reported by van der Meer et al. (2012a; b; c; 2013). Since three of the four locations are World (LHI) or National Heritage (MR, ER) listed, and given the difference in abundance of the species per location, only 10 to 37 fin clips were taken from each location to estimate gene flow: MR (n = 20), ER (n = 10), LHI (n = 37), and NI (n = 16).

Ethics statement

The main aim of this study was to determine gene flow between, self-replenishment within and genetic diversity at all locations in the distibution range of the endemic doubleheader wrasse *(Coris bulbifrons)*. Fin clip samples were obtained from adult fishes (>100mm) either by spearfishing or by anaesthetising fish with clove oil, fin clipped *in situ* and released alive (Permit Numbers: LHIMP08/R01, 003-RRRWN-110211-02, P11/0035-1.0; Animal ethics: A1605).

Study system

MR, ER ,and LHI are referred to as the "western region" for *C. bulbifrons* because they occur on the same geographic feature, the Lord Howe Island Rise; they are relatively close to each other (Fig. 1) and they support higher abundances (Choat et al. 2006a; Hobbs et al. 2009). In contrast,

NI is referred to as the "peripheral location" for this species because it is the only location situated on a separate geographic feature, the Norfolk Island Rise, which is isolated by more than 600 km from the western region (Fig 1) and has a lower relative abundance of the study species (*authors pers. obs.*). *C. bulbifrons* inhabits shallow (< 40 m) reef habitats and adults will not traverse deep oceanic waters to disperse between the four locations. However, *C. bulbifrons* has a pelagic larval phase (mean duration 36 days; *authors unpubl data*) that is capable of dispersing between locations and this may be aided by oceanographic currents. Preliminary analyses indicate that *C. bulbifrons* and *C. agyula* are sister species, and that there is evidence of possibly mtDNA hybridization between these species (*unpubl data*).

Gene flow between locations - mtDNA

mtDNA phylogenetic analysis The non-coding (control region (CR) containing the D loop segment) was sequenced to determine population genetic structure in *C. bulbifrons* as per van der Meer et al. (2012a; b). Fin clips from three *C. gaimard* individuals were collected from Christmas Island to use as an outgroup (as indicated by Barber & Bellwood 2005). jModeltest (Posada 2008) identified a GTR+G model under AIC with gamma = 0.271. The three most commonly used phylogenetic analyses: Maximum Likelihood (ML), Maximum Parsimony (MP), and Bayesian Inference (MrBayes, MB; BEAST) were performed on aligned mtDNA sequence data. A Minimum Spanning Tree (MST) was generated based on output obtained from ARLEQUIN 3.5 (Excoffier et al. 2005) to explicitly identify shared haplotypes between *C. bulbifrons* from the four locations (MR, ER, LHI, NI).

Quantifying the level of mtDNA gene flow. To obtain reliable estimates of mtDNA gene flow, samples from ER were pooled with MR, given that pairwise $F_{st,}$ DAPC, STRUCTURE and GeneClass analyses could not genetically differentiate the two populations (also see msatDNA section hereunder). Thus, *C. bulbifrons* mtDNA migration rates were estimated among, and effective population sizes were estimated within, each of the three areas (MR-ER, LHI, NI) using MIGRATE-n 2.4.3 (http://popgen.sc.fsu.edu/Migrate-n.Html; Beerli & Felsenstein 2001; Beerli 2004). We tested a combination of various migration priors (F_{st} and OWN: isolation-by-distance) and custom-migration models (Stepping-stone, Island-n, and variable θ only) - all with a constant mutation rate. A Log Maximum-Likelihood analysis (Ln ML) comparing all possible combinations was selected: migration prior (F_{st}), custom-migration model (migration model with variable θ), constant mutation rate with a F84 mutation model, migration rate parameters (θ and M set to a maximum of 0.8 and 8000, respectively). Finally, a Bayesian analysis was performed using a heating search strategy of one long chain that sampled every 30th of 60 k sampled trees and applied a 20 k iteration burn-in. All parameters estimated by the Bayesian analysis converged and fell within the 90% CI, yielding values for θ and M (mutation-scaled migration rate) per location.

Gene flow between locations - msatDNA

Patterns of gene flow (msatDNA). We used a large number of polymorphic microsatellite loci (n = 17) and sampled all known locations, to compensate for small sample sizes used in this study (see Selkoe & Toonen 2006). Furthermore, we recognize that our estimates for "self-replenishment" inferred indirectly from genetic markers are merely a proxy for self-recruitment,

which is typically assessed using more direct methods (e.g. by using natural or artificial otolith tags), such as those used by Swearer et al. (1999), Jones et al. (2005), and Almany et al. (2007). Nevertheless, direct approaches may not be feasible for our study species, without negatively impacting populations, due to large sample sizes typically required for such parentage-based studies (see Berumen et al. 2012, Harrison et al. 2013). Therefore, we believe that our indirect estimates of self-replenishment represent the best possible substitute for self-recruitment obtainable for this species without substantially impacting its populations.

Three molecular analytical tools were used to establish spatial population partitioning based on msatDNA: (i) discriminant analysis of principal components (DAPC; Jombart et al. 2010) uses allelic states to discriminate between the four locations, yielding scatterplots of discriminant functions based on the spatial distributions of microsatellite genotypes. DAPC also provided posterior probabilities of population assignments for each individual; (ii) a likelihood-based assignment method was used in GeneClass2 (Paetkau et al. 1995; Paetkau et al. 2004; Piry et al. 2004) to determine significant inter-location gene flow and (iii) STRUCTURE V2.3 (Pritchard et al. 2000; Hubisz et al. 2009) places individuals into clusters that minimise Hardy-Weinberg Equilibrium (HWE) and can be used to identify contemporary gene flow among the four locations. To determine the "best value" for K (the number of subpopulations in the distribution range), we followed the method suggested by Pritchard et al. (2000), which involved comparing mean log-likelihoods penalised by one-half of their variances (see Hubisz et al. 2009). The final run consisted of an Admixture model with 2 M iterations and a 100 k iteration burn-in.

Quantifying the level of msatDNA gene flow. As above, samples from ER were pooled with MR to estimate *C. bulbifrons* migration rates among, and effective population sizes (within)

each of the three populations (MR/ER, LHI, NI) based on msatDNA, using MIGRATE-n 2.4.3 as above. We set datatype to Microsatellite (a simple electrophoretic ladder model), migration prior (F_{st}), custom-migration model (migration model with variable θ), constant mutation rate with a stepwise mutation, migration rate parameters (θ and M to a maximum of 60 and 40, respectively), and a Bayesian analysis, using a search strategy of one long chain that sampled every 20th of 50 k sampled trees and applied a 20 k iteration burn-in. All parameters converged and fell within the 90% CI yielding values for θ and M (mutation-scaled migration rate) per population.

Inferred levels of self-replenishment and recent migration. This study did not sample new wrasse recruits in order to determine self-recruitment as in Jones et al. (2005). However, we used BAYESASS v3 (Wilson & Rannala 2003), a program specifically designed for population genetic studies that estimates recent migration rates (past 2 - 3 generations) between populations. Conversely, this program also has the ability to estimate any individuals not migrating (i.e. self-replenishing). We used BAYESASS v3 to estimate both self-replenishment and recent migration between locations with an MCMC chain, consisting of a total of 11 M steps with a 2 M step burn in. Prior values for migration rate, allele frequency, and inbreeding coefficient were specified as 0.9, 0.9, and 0.9 respectively. These priors were selected because they gave acceptance rates of between 20 and 40% suggesting congruence (Faubet et al. 2007). Ten separate runs assessed convergence of the MCMC in order to evaluate the consistency of results obtained from these inferences.

Population genetic diversities

Molecular diversity indices for mtDNA (haplotype diversity, h; nucleotide diversity, π) and for msatDNA (genetic diversity, *gd*) were estimated in ARLEQUIN 3.5 (Excoffier et al. 2005). Haplotype and nucleotide diversities of the data were interpreted as either low with specified cut-off values of *h* and π (%) < 0.5 or high if values of *h* and π (%) were > 0.5 (Grant & Bowen 1998).

6.4. RESULTS

Synopsis. Three hundred and forty one base pairs of mtDNA (D Loop) were resolved for 81 *C. bulbifrons* individuals with a total of 56 polymorphic sites, of which 37 were parsimony informative. A total of 13 spatially intermixed clades were identified (bootstrap values >50%), with Clade 13 comprising 38% of the entire sample (n = 31/81) suggesting that mtDNA gene flow exists between all locations occupied by *C. bulbifrons* (Fig 6.1a). A Minimum Spanning Tree (MST) identified 56 haplotypes in total, many of which were unique (n = 44) (Fig 6.1b). mt- and msat-DNA AMOVA and (corrected msat) pairwise F_{st} results indicate that there is no population genetic differentiation between locations within the western region (MR, ER, LHI) or between the western region locations and the peripheral location (NI; 6.S1, 6.S2, 6.S3, 6.S4). However, DAPC, STRUCTURE, and GeneClass msatDNA analyses partitioned *C. bulbifrons* into two differentiated regions - a western region (MR/ER, LHI) and a peripheral location (NI; Fig 6.2c, 6.3). High levels of self-replenishment (68 to 93%) were inferred at all four locations (Fig 4c). Haplotype (h), nucleotide (π), and genotypic diversities (*gd*) were high (> 0.5) at all locations (Fig 6.2c, Table 6.S4). Detailed AMOVA, locus-by-locus AMOVA, pairwise population comparisons, and genetic diversity can be found in Supporting Information (6.S1, 6.S2, 6.S3, and 6.S4 respectively).

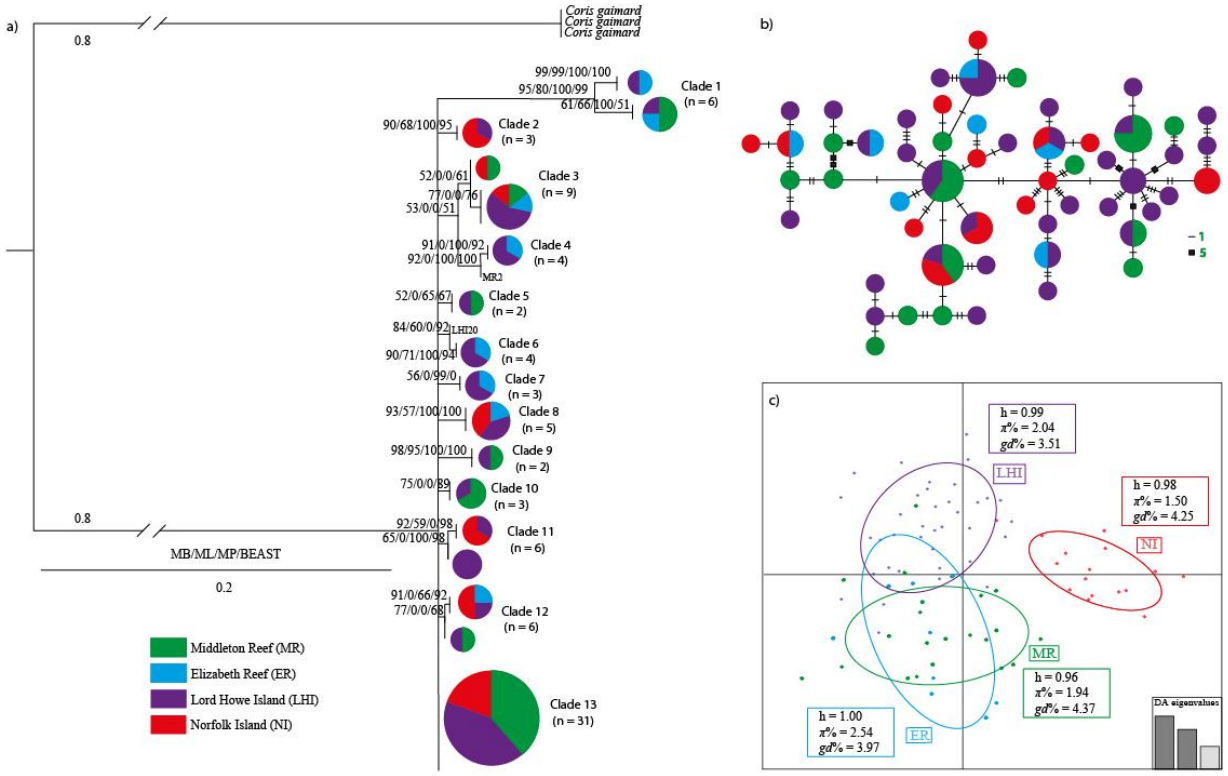

Figure 6.2 *C. bulbifrons* mt- and msat-DNA analyses. a) A phylogram of mtDNA (D loop) sequences from 81 *C. bulbifrons* individuals from Middleton Reef, Elizabeth Reef, Lord Howe Island, and Norfolk Island. This represents the best ML tree from 10 individual analyses. Numbers on branches indicate support for each clade, based on ML, MP, MB, and BEAST analyses. b) Haplotype minimum spanning tree (MST) with number of substitutions between haplotypes indicated on connectors. Different fills represent each of the four locations as shown on the key to the figure. c) Scatterplots of the discriminant analysis of principal components of the microsatellite data for the four locations where *C. bulbifrons* occurs globally, using geographic sample site as priors for genetic clusters. Individual genotypes appear as dots surrounded by 95% inertia ellipses. Eigenvalues show the amount of genetic information contained in each successive principal component with the x- and y-axes constituting the first two principle components, respectively. Boxes indicate haplotype (*h*), nucleotide (*% π*), and genetic diversity (*gd*) indices for *C. bulbifrons*.

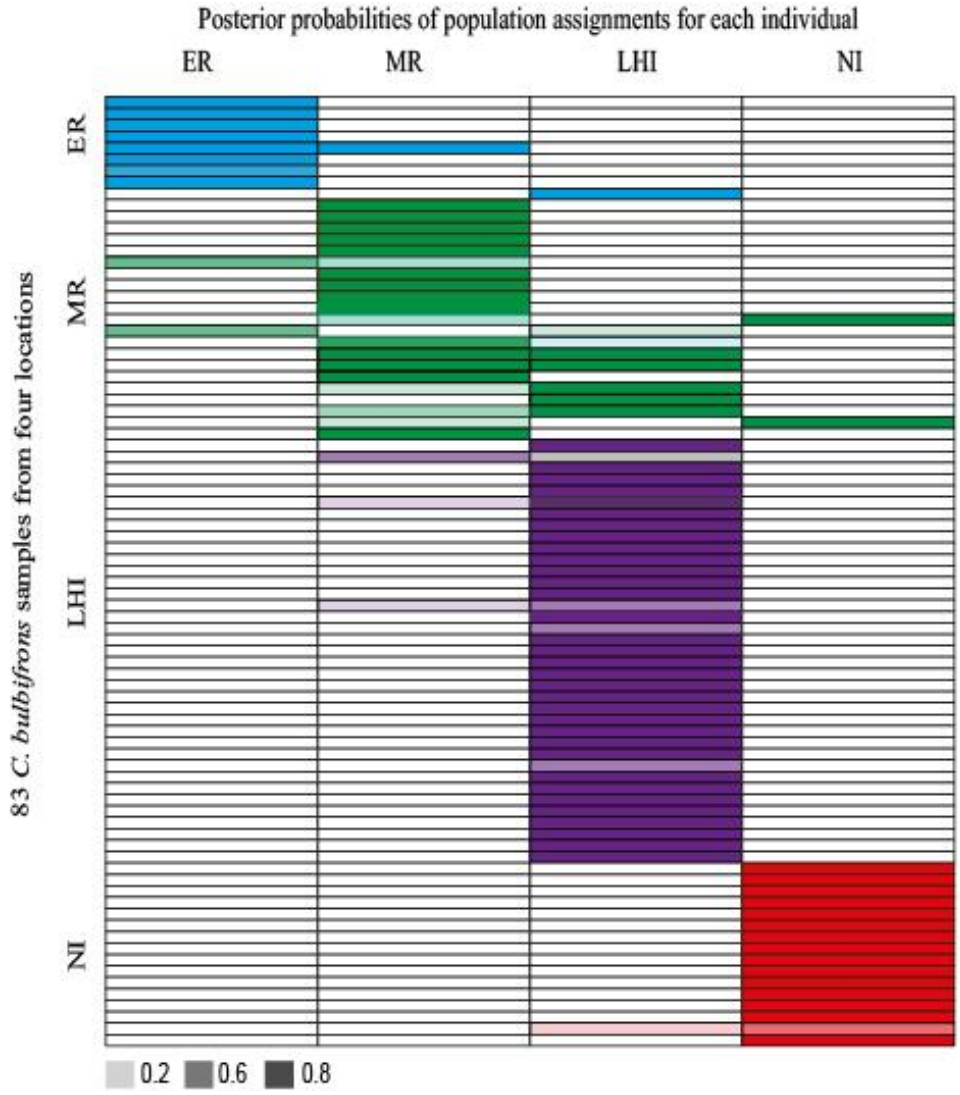

Figure 6.3 Posterior probability of assignment of each individual genotype to four *Coris bulbifrons* populations as indicated by DAPC. The names of the possible assignment populations are given on the x-axis. 83 genotypes are listed on the y-axis, along with the population from which they were sampled. Coloured bars in different shades correspond to a 0.2 to 0.8 probability of assignment to a given population.

Summary statistics. Heterozygote deficiency was evident from a positive inbreeding co-efficient (F_{IS}) at MR, ER, and LHI; while heterozygote excess was evident at NI from a negative inbreeding co-efficient (Table 6.S4). Lord Howe Island had the most private alleles, nine across 17 loci, while the remaining three locations ranged from two to six private alleles across all loci (Table 6.S4). This most likely results from a sampling bias at this location. Of the 17 msatDNA loci: (i) significant single locus departures from HWE were detected in 9 of 68 tests at the location level before FDR correction and only one afterwards (NI: Cb2); (ii) null alleles were identified only at 1 locus, Cb5 in 1 location, ER and (iii) of the 272 locus-by-locus exact tests of linkage disequilibrium, 10 were significant before and two after FDR correction (Cb13, Cb30; Benjamini & Hochberg 1995). Loci that had null alleles (Cb5) were not used in subsequent analysis (ARLEQUIN, STRUCTURE and MIGRATE-n) and loci in linkage disequilibrium at all sites (Cb13, Cb30) were not used in subsequent analysis (ARLEQUIN, STRUCTURE, BAYESASS). Thus 14 loci were used in the ARLEQUIN and STRUCTURE analyses, and 16 loci were used in MIGRATE-n and BAYESASS analyses.

Gene flow between locations - mtDNA

Patterns and levels of gene flow based on an mtDNA AMOVA structured into regions: western region (MR, ER and LHI) vs peripheral location (NI), indicated no significant genetic variation (0.19%) between regions, $\Phi_{ct} = 0.002$ ($p = 0.752$), or amongst locations within regions, $\Pi_{sc} = -0.004$ ($p = 0.618$; Table 6.S1). All genetic variation occurs within locations (100.23%), $\Phi_{st} = -0.002$ ($p = 0.553$; Table 6.S1), although this is not significant. Likewise, mtDNA pairwise F_{st} showed no genetic differentiation amongst any of the paired locations (pairwise $F_{st} = -0.018$ to 0.012, $p = 0.234$ to 0.694; Table 6.S3).

Quantifying mtDNA gene flow using Migrate-n indicated high levels of historic gene flow between all locations (MR/ER, LHI, NI), with M values ranging from 291 to 1144 (Fig 6.4a).

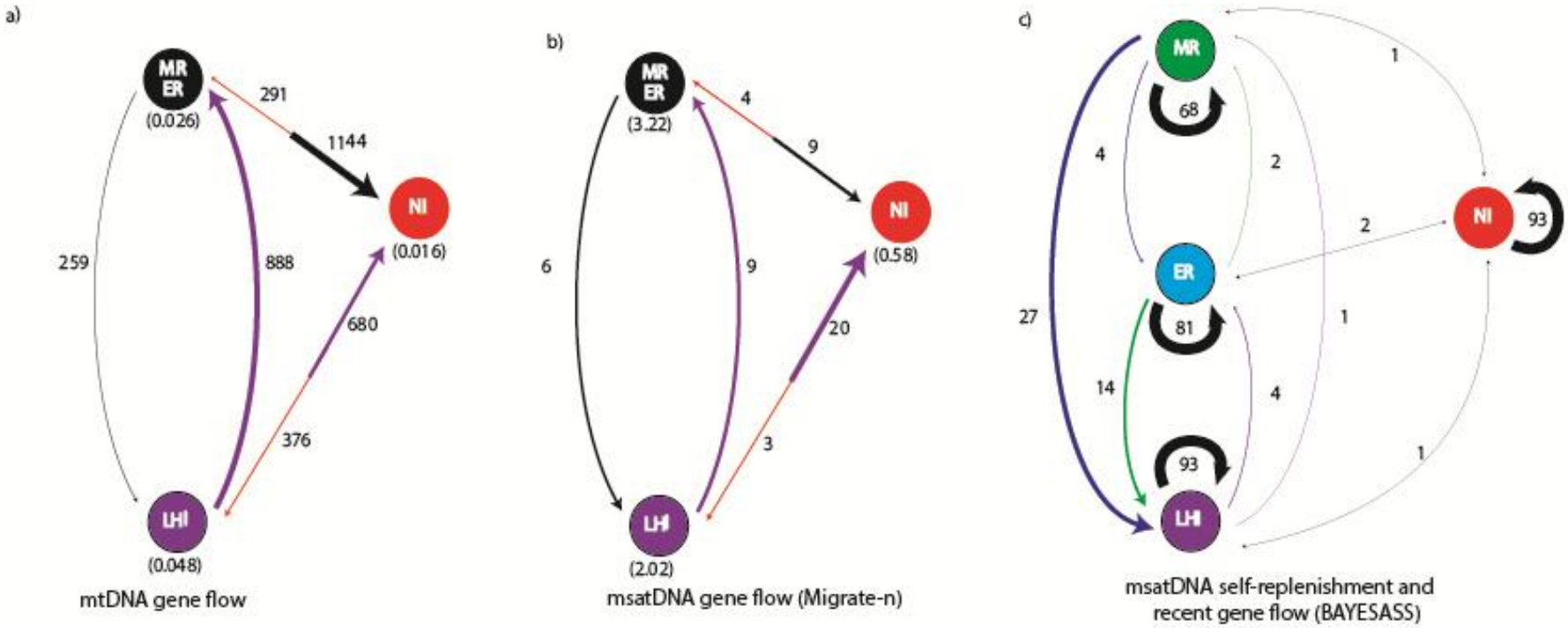

Figure 6.4 Migration rates among *Coris bulbifrons* locations. The thickness of the arrowed line is directionally proportional to the number of migrants (M) and the line colours indicate the predominant direction of gene flow. Population size (θ, within parentheses) is also shown for each location. a) Migrate-n evolutionary gene flow (mtDNA), b) Migrate-n contemporary gene flow (msatDNA), and (c) BAYESASS analysis of self-replenishment and recent migration rates (msatDNA) shown as a percentage.

Gene flow between locations - msatDNA

Patterns and levels of gene flow based on an msatDNA AMOVA, indicated significant structure in 7 (of 17) locus-by-locus analyses corrected for null allele frequency (Φ_{st} = 0.001 to 0.190, $p < 0.05$; Table 6.S2), corrected for standardised population differentiation (Φ_{st} = 0.006 to 0.268, $p < 0.05$; Table 6.S2), and in the global AMOVA as a weighted average over all microsatellite loci (Φ_{st} = 0.025, $p < 0.001$; Table 6.S1), with 97.53% of the genetic variation existing within populations. Raw msatDNA pairwise F_{st} comparisons showed low non-significant genetic partitioning between populations (F_{st} = 0.007 to 0.027, $p > 0.144$); with the exception of LHI and ER (F_{st} = 0.044, $p = 0.003$) as well as LHI and NI (F_{st} = 0.043, $p < 0.001$; Table 6.S3). In contrast, the Excluding Null Alleles (ENA) corrected msatDNA pairwise F_{st} values showed no significant genetic differentiation between populations (F_{st} = 0.011 to 0.099, $p > 0.05$; Table 6.S3).

DAPC, GeneClass2 and STRUCURE analyses all supported at least three geographically distinct populations. DAPC partitioned *C. bulbifrons* into the western region (MR, ER, LHI) and the peripheral location (NI, Fig 6.2c). Using the four locations as *a priori* population criteria, DAPC assigned 70 to 94% of all individuals to the populations from which they were sampled (assignment per population, ER = 70%, MR = 85%, LHI = 92%, NI = 94%; Fig 6.3). Consistent with these assignments, with the allele frequencies and genotypic assignments, the 95% genotypic inertia ellipses (GIE) for MR and ER overlap, whilst the 95% GIE for ER overlaps with LHI, and the 95% GIE for NI occupies a distant area of multivariate space, along the x-axis, from all three western populations. Geographical structure in msatDNA data was confirmed by GeneClass2 analyses, where only 16 individuals were grouped with populations from which they were not sampled (MR = 3, ER = 4, LHIL = 6, NI = 3). Similarly, three geographically

partitioned populations were identified by STRUCTURE analyses (MR/ER, LHI and NI), as the likelihood of the marginal posterior probability distribution was highest when $K = 3$ (i.e. three distinct populations), data not shown.

Quantifying the level of msatDNA gene flow. Migrate-n indicated a few orders of magnitude lower levels of msatDNA gene flow between populations (MR/ER, LHI, NI) when compared to mtDNA gene flow, with *Nm* values ranging from 3 to 20 (Fig 6.4b). Directions of gene flow were consistent between the different markers with most of the gene flow towards NI and towards MR-ER from LHI within the western region.

Inferred levels of self replenishment and migrant exchange. Despite weak genetic differentiation (F_{st}) between populations, both DAPC and STRUCTURE partitioned the data into 3 distinct populations (MR/ER, LHI and NI). Used together, these programs are likely to be better than F_{st} values (Faubet et al. 2007) at determining the appropriateness of a dataset for BAYESASS because they extract more information from the genetic data than frequency based fixation indices. However, estimates of migration rates are accurate when model assumptions are not violated (i.e. migration rates are low, loci are in linkage equilibrium) and genetic differentiation is not too low ($F_{st} \geq 0.05$). If model assumptions are violated, then accurate estimates are obtained only when migration rates are very low ($m = 0.01$) and genetic differentiation is high ($F_{st} \geq 0.10$; Faubet et al. 2007). Lastly, when migration rates fall below 10%, populations can probably be considered demographically independent (Waples & Gaggiotti 2006). Demographic independence is suggested for all population pairs except: ER to LHI (m = 14%) and MR to LHI (m = 27%; Fig 6.4c). Conversely, high levels of self-replenishment (68 to 93%) were inferred at all four populations (Fig 6.4c). This further indicates that in the short term,

each population is predominantly sustained by self-replenishment rather than replenishment from distant populations.

Population genetic diversities

Coris bulbifrons showed high haplotype (*h*), nucleotide (% π), and genetic diversities at all locations (*h* = 0.98 to 1.00, % π = 1.94 to 2.54, *gd* = 3.51 to 4.37; Fig 6.2c). Total haplotype, nucleotide, and genotypic diversities were also high (*h* = 0.99, %π = 1.95, *gd* = 0.384; Table 6.S1).

6.5. DISCUSSION

Studying endemic species that occur in discrete and remote locations offers unique opportunities to examine gene flow throughout a species entire geographic range. Island systems such as that examined in this study provide useful empirical data on gene flow (historic and demographic) among isolated locations, separated by known distances and deep oceanic waters. In this study, *Coris bulbifrons* was found to have (i) sufficient mtDNA gene flow among locations to indicate a lack of geographic partitioning between locations; (ii) low levels of msatDNA gene flow among locations (except between MR/ER) resulting in genetically differentiated populations; (iii) demographic interdependence between LHI and MR/ER, yet high levels of inferred self-replenishment at all four locations; and (iv) high genetic diversities (mtDNA and msatDNA) at all four locations.

Gene flow among locations - mtDNA

A complete lack of spatial genetic structure was suggested for *C. bulbifrons*. This is likely the result of a small number of recruits per generation maintaining spatial genetic homogeneity (Shulman 1998; Planes 2002). Interestingly, the peripheral location (NI) shows no genetic break from the western region, in contrast with peripheral locations of other widespread reef fishes at larger spatial scales; e.g. parrotfish at Hawaii (Bay et al. 2004), damselfishes and wrasses at Fiji (Drew et al. 2008), snappers at the Marquesas (Gaither et al 2010) and parrotfish at Cocos Keeling, Hawaii, Marquesas and Seychelles (Winters et al. 2010). Thus mtDNA gene flow has been sufficient to connect all locations over historic time scales or insufficient time has passed since the species expanded its range, for neutral markers to accumulate genetic differences due to genetic drift.

The high level of inferred mtDNA gene flow evident among locations is encouraging since gene flow to distant locations, especially to the peripheral location (NI) which is not protected by any MPA, may assist with recovery of depleted populations, recolonisation following local extinction, and increasing genetic diversity (Hanski 1999; Jones et al. 2009). Despite the clear benefit of high levels of historical mtDNA gene flow among locations, if this is combined with limited gene flow on demographic time scales, then the benefit may be tempered by minimal, if any replenishment between locations in the short term.

Gene flow among locations - msatDNA

C. bulbifrons showed msatDNA genetic differentiation between MR/ER, LHI, and NI on demographic timescales. The contrast between high mtDNA and low msatDNA levels of gene flow evident in *C. bulbifron*s has been documented in other endemic coral reef fishes within the region: the endemic Mcculloch's anemonefish, *Amphiprion mccullochi* (van der Meer et al. 2012a; b) and the endemic three-striped butterflyfish, *Chaetodon tricinctus* (van der Meer et al. 2013), indicating that this pattern is consistently found in endemic species examined from this region to date. This apparent discrepancy between mtDNA and msatDNA likely results because few recruits per generation may maintain mtDNA genetic homogeneity over historical timescales (Shulman 1998; Planes 2002), whereas populations at isolated locations require substantial amounts of self-recruitment on demographic timescales to maintain viable populations. Most studies attempting empirical measurement indeed find similar levels of self-recruitment, for example Caribbean wrasses (Swearer et al. 1999); Papua New Guinea anemonefish (Jones et al. 2005; Planes et al. 2009; Berumen et al. 2012); and Papua New Guinea butterflyfish (Almany et al. 2007; Berumen et al. 2012). Lastly, the high levels of self-recruitment and low levels of

contemporary gene flow, together with restricted gene flow at longer time scales at the peripheral location, may underlie observed phenotypic differences (stripes and patterns around the eye) seen in NI individuals, indicating that NI is at the very least, a genetically distinct and unique subpopulation (*sensu* Drew et al. 2008). This may further indicate that *C. bulbifrons* at NI is at an early stage of peripheral speciation (*sensu* Bowen et al. 2013). Despite mtDNA homogeneity, phenotypic differences such as colour may evolve faster than mtDNA markers and be the first traits to emerge in a diverging species (e.g. McCafferty et al. 2002; Rocha, 2004; but see Messmer et al. 2005; Schultz et al. 2007; DiBattista et al. 2012).

Models predict that gene flow of a few recruits per generation between locations over historical timescales will not sustain populations (Cowen et al. 2000; Cowen et al. 2002). Thus, the major benefits of MPAs include protection of biodiversity and fisheries stocks within MPAs and the net larval export of larvae to unprotected areas to help sustain exploited populations (Halpern & Warner 2003; Shanks et al. 2003; Jones et al. 2005; Harrison et al. 2012). Currently, the MPAs at MR, ER, and LHI encompass suitable habitat for *C. bulbifrons* (and other coral reef fish). However, NI remains unprotected and hence open to overfishing and exploitation. With extremely low levels of msatDNA gene flow between the western region and the peripheral location, it is unlikely that larvae produced within the MPAs of the western region will be of any real demographic benefit to NI. The very low abundance of *C. bulbifrons* at NI and the lack of migrants from other locations increase the risk of local extinction at this peripheral location. If local extinction occurs, recolonisation and recovery will be slow. Local extinction at NI will result in the loss of a unique colour variant and unique genotypes from this species' eastern range edge. In contrast, in the western region the greater abundance of *C. bulbifrons* (Choat et al. 2006a; Hobbs et al. 2009) and higher levels of recent msatDNA gene flow among populations

(MR, ER, LHI) is encouraging, since both serve to decrease the risk of local extinction and facilitate recovery should populations decline or go locally extinct. However, with high levels of self-replenishment at all populations (see below), and fishing pressure at ER and in some areas at LHI, populations may still decline with minimal fishing effort (Choat et al. 2006b). The abundance of *C. bulbifrons* should be monitored in areas open to fishing to determine if fishing restrictions need to be adjusted to prevent overfishing.

All populations appear demographically independent, except LHI, as gene flow among populations was less than 10% (*sensu* Waples & Gaggiotti 2006). With such high levels of self-replenishment (> 68%), all populations should be considered largely closed and recovery following population decline or local extinction is still likely to take some time (Miller & Shanks 2004). The input of gene flow from populations at both MR and ER to LHI highlights the importance of continued population monitoring and protection at these two locations (MR, ER), especially at ER where spearfishing is allowed. If LHI faces population decline from anthropogenic factors such as overfishing, then populations at either or both MR and ER will aid in rescue effects, thereby reducing recovery times. The high levels of self-replenishment estimated in this study for *C. bulbifrons* using molecular techniques are supported by direct estimates of self-recruitment based on otolith chemistry at LHI for this species (Patterson & Swearer 2007) and also for other coral reef fish populations at other islands, e.g. bluehead wrasse in the U.S. Virgin Islands (Swearer et al. 1999), anemonefish and butterflyfish in Papua New Guinea (Jones et al. 2005; Almany et al. 2007; Planes et al. 2009).

Population genetic diversities

Maintaining genetic diversity is an IUCN priority (McNeely et al. 1990), for at least two reasons: (i) it provides the raw material for natural selection to act on over historical (Johanesson & Andre 2006) and demographic (Bell & Okamura 2005) timescales, and (ii) low genetic diversity increases the risk of inbreeding depression (Reed & Frankham 2003). *C. bulbifrons* showed high genetic variability (mtDNA, h and π) compared to other marine fishes (Grant and Bowen, 1998) and high levels of genetic diversity for msatDNA. Given that mtDNA diversity tracks nuclear genetic diversity in many marine species (reviewed by Johannesson & Andre 2006), including *Coris bulbifrons*, this is encouraging as species with high genetic diversity may have some resilience compared to species with decreased genetic diversity, which tend to have decreased fitness (Hoelzel et al. 2002). Additionally, species with high genetic diversity may have a greater adaptive capacity to deal with the impacts of environmental change, than those with low genetic diversity (Avise 2000).

Population connectivity in endemic fishes within the LHI region

The LHI region is a hotspot for endemic coral reef fishes (Marine Parks Authority 2010) with the 4th highest percent endemism (7.2%) in the Indo-Pacific (Randall 1998; 2001; 2007). Currently, an MPA network consisting of three isolated locations (MR, ER, LHI) aims to protect this unique diversity. Previous population genetic research on an anemonefish (*A. mccullochi*) and butterflyfish (*C. tricinctus*)(van der Meer et al. 2012a; b; 2013), both endemics to the region, found similar patterns and levels of gene flow over spatial- and time- scales as reported here for *C. bulbifrons*. The peripheral location (NI) was also identified as genetically distinct from the western region (LHI region) for the endemic anemonefish and butterflyfish species, indicating

limited demographic gene flow. While the current network of MPAs within the western region may provide adequate protection for endemic reef fishes, the lack of protection at NI is concerning. Furthermore, many of the endemics at NI have low abundance (author's pers obs). Given the elevated risk of extinction of endemics at NI, and the genetic uniqueness of these NI populations, establishing protective measures (MPAs and fishing regulations) should be a management priority at this location. We also recommend that in the western region, some areas of ER be closed to all types of fishing to provide some protection from overfishing at this location.

Genetic studies across three taxonomically distinct groups having different ecological (anemone dwelling planktivore, corallivore and herbivore, respectively) and life history traits (mean Pelagic Larval Duration: *A. mccullochi* = 12, *C. tricinctus* = 35, *C. bulbifrons* = 36), show similar patterns of population connectivity and genetic diversity. This indicates that a single management strategy within this region may be appropriate for the design of MPAs to protect endemic reef fishes in the LHI region. If other remote islands with high levels of reef fish endemism (e.g. Easter Island, the Marquesas, Rapa) also show similar patterns of population connectivity and replenishment among endemics, then this advocates for a single spatial management strategy at remote island groups. This management strategy involving a network of MPAs that protects part of each location in the geographic range of endemics is likely to be effective at conserving the unique biodiversity of endemism hotspots.

Conclusion Despite high mtDNA gene flow over historical timescales in *C. bulbifrons*, low levels of msatDNA gene flow over demographic timescales demonstrate that the distribution of this endemic species is comprised of at least three demographically isolated populations. While the high abundance at the three western locations and high genetic diversities at all locations may help to ameliorate impacts of some disturbances, the low contemporary gene flow between populations should be of concern for management and conservation of this species. This is particularly concerning for NI, which has a very small population of *C. bulbifrons*, receives limited contemporary gene flow and is currently unprotected by an MPA or fishing restrictions of any description. Although isolation buffers remote locations from many anthropogenic impacts, it can also increase vulnerability because of limited connectivity and replenishment between populations, as shown here. Given that remote reefs are also hotspots of endemicity for coral reef fishes (Jones et al. 2002), if widespread impacts occur, endemic species may be particularly vulnerable due to their small and fragmented geographic distribution, so urgent management action is therefore recommended to protect this and other endemic species from local extinction. This is of particular importance during the Anthropocene when coral reef species are experiencing increased stresses (Steffen et al. 2007), even if they are in remote places, due to acidification and rising temperatures and how these factors impact coral reef habitats.

Supplementary Online Material

Table 6.S1. AMOVA analysis for a) mtDNA (D Loop) sequences from *Coris bulbifrons* structured into LHIR region (MR, ER, LHI) vs NIR peripheral location (NI) and b) global AMOVA weighted across all 17 microsatellite loci.

Source of variation	Variance component	Percentage of variation	F-statistics fixation indices (p-value)
a) mtDNA			
Among locations	0.006	0.19	F_{ct} = 0.002 (0.752)
Among populations within locations	-0.014	-0.42	F_{sc} = -0.004 (0.618)
Within populations	3.23	100.23	F_{st} = -0.002 (0.553)
b) Microsatellite			
Among locations	-0.011	-0.49	F_{ct} = -0.005 (0.664)
Among populations within locations	0.069	2.96	F_{sc} = 0.029 (0.06)
Within populations	2.26	97.53	F_{st} = 0.025 (0.001)

Table 6.S2. AMOVA fixation indices (F_{st}) for *Coris bulbifrons* across all populations surveyed

	Marker class and analysis		
	Raw msat	Msat corrected for null allele freq.	Standardised msat (F_{st})
Average	**0.043**	0.035	0.061
Cb1	0.002	**0.001**	**0.006**
Cb2	**0.155**	**0.155**	**0.233**
Cb4	-0.002	0.022	-0.003
Cb5	**0.047**	**0.062**	**0.115**
Cb6	–	0.040	**0.146**
Cb7	–	**0.010**	0.041
Cb8	–	-0.011	-0.018
Cb10	0.031	0.035	0.085
Cb11	-0.008	-0.007	-0.022
Cb12	0.026	**0.006**	**0.024**
Cb13	0.005	0.015	0.000
Cb15	0.033	0.037	0.047
Cb18	–	-0.019	-0.040
Cb23	0.010	**0.010**	0.039
Cb26	**0.070**	0.070	0.091
Cb30	**0.187**	**0.190**	**0.268**
Cb36	**0.054**	0.049	**0.116**

Raw population differentiation from microsatellite allele frequencies for each individual locus and the average across all loci, population differentiation corrected for null allele frequencies using the ENA correction of Chapuis & Estoup (2007) and standardized population differentiation for and across all loci (F_{st}). All values in bold are significant to the $p<0.05$ (i.e. 95% confidence interval).

Table 6.S3. Pairwise population structures ($\prod_{st}$ and F_{st}) generated for mtDNA (D Loop, n = 82) and for 17 microsatellite loci (n = 83), respectively from four *Coris bulbifrons* sample locations showing raw and corrected F'_{st} for null allele frequencies.

	D Loop				Raw msat				Corrected msat			
	MR	ER	LHI	NI	MR	ER	LHI	NI	MR	ER	LHI	NI
MR		0.414	0.234	0.694		0.144	0.151	0.180		>0.05	>0.05	>0.05
ER	-0.001		0.333	0.613	0.027		**0.003**	0.149	0.041		>0.05	>0.05
LHI	0.012	-0.001		0.442	0.010	0.044		**0.000**	0.011	0.045		>0.05
NI	-0.018	-0.015	-0.013		0.007	0.021	0.043		0.046	0.099	0.064	

Raw population differentiation from microsatellite allele frequencies below the diagonal and associated *p*-values above the diagonal, using Arlequin ver 3.5 (Excoffier et al. 2005). Corrected population differentiation for null allele frequencies at 95% CI (all *p*-values > 0.05) using the ENA correction of Chapuis and Estoup (2007). FDR correction of raw F'_{st} *p*-values in QVALUE (Storey 2002).

Table 6.S4. Sample sizes for mtDNA (D Loop, *n* = 81 and msat DNA, n = 83)

	n (D Loop)	n_h	*h*	*π(%)*	*n* (msat)	*gd*	N_a	P_a	H_o	H_e	Fis
All	81	56	0.99	1.95	83	3.84	3.44		0.392	0.408	0.038
MR	20	14	0.96	1.94	20	4.37	4.35	6	0.501	0.503	0.005
ER	8	8	1.00	2.54	10	3.97	3.05	2	0.356	0.408	0.127
LHI	37	33	0.99	2.04	37	3.51	4.41	9	0.383	0.436	0.121
NI	16	13	0.98	1.50	16	4.25	1.94	2	0.284	0.295	-0.157

D Loop, non-coding region D Loop
Number of haplotypes (n_h), haplotype diversity (*h*), nucleotide diversities (π) of D Loop for all regions and populations of *Coris bulbifrons*. Sample sizes for msats (total n = 83), genetic diversity (*gd*), average number of alleles per locus (Na), observed number of private alleles (Pa), observed heterozygosity (H_o) and expected heterozygosity (H_e) and the inbreeding coefficient (F_{IS}) averaged over 17 microsatellite loci for four populations.

CHAPTER 7: **General Discussion**

Population connectivity (fundamentally, the exchange of individuals among locations) over historical time scales plays an important role in genetic structure and the evolution of species. However, high levels of self-replenishment, especially at isolated reefs, are crucial for population persistence on demographic time scales (Robertson 2001, Hastings & Botsford 2006). Without population persistence on demographic time scales, evolutionary processes acting over historical time scales are not possible. Thus, the very high levels of self-replenishment (and low levels of recent migration) seen in all study species for which sufficient data was available (GBR anemonefish, Mcculloch's anemonefish, three-striped butterflyfish, doubleheader wrasse), is a double-edged sword: It is a strength - allowing populations to persist through time and it is a weakness, because if/when significant disturbances do occur, populations may be lost. Regardless, endemic species in this study system clearly persist through time despite high levels of habitat or food specialization, low abundance at some locations, and high self-replenishment. This capacity of endemic species to persist against the odds may be assisted by low levels of gene flow over time, which generates the observed high genetic diversity and increased persistence times that would be expected otherwise. Mixing that results from low levels of gene flow over time, either amongst populations within a species or amongst close relatives (hybridization) has also been documented for various widespread species of reef fish, e.g. numerous Indo-Pacific surgeonfishes (Horne et al. 2008, Horne & van

Herwerden 2013). Lastly, hybridization in up to three of the four species examined here (GBR and Mccullochs anemonefish; possibly the doubleheader wrasse, where initial phylogentic analyses suggested hybridization with a widespread sister species, *Coris agyula* – data not shown), has injected new genetic variants into otherwise isolated populations, thereby increasing genetic diversity and the capacity to adapt to new environments (reviewed by Abbott et al. 2013). Notwithstanding genetic extinction risk, the process of extinction is complex and involves many extrinsic (e.g. overexploitation or habitat clearing) and intrinsic (e.g. demographic stochasticity or inbreeding depression; O'Grady et al. 2006, Szulkin & Sheldon 2006) forces. However, genetic information, along with ecological and life history trait information, will allow managers now and in the future, to carefully assess and comprehensively weigh species extinction risks (e.g. Dunham et al. 1999).

Are the population genetic characteristics of connectivity and diversity similar across species?

I compared four genetic traits across a range of taxonomically divergent coral reef fishes, with different ecological and life history traits, and found clear support for similar patterns and levels of population connectivity and genetic diversity across species. This suggests that the environment has been a major force shaping the genetic traits seen in endemic species at the Lord Howe- and Norfolk Rise study system, which may influence other coral reef species in this study system similarly.

(i) Historical connectivity (mtDNA) among locations is sufficient to maintain genetic homogeneity (i.e. there is no geographic basis to genetic structure) in the majority of species examined here (*A. akindynos, A. mccullochi, C. bulbifrons, E. daemellii*). However, in at least two species, regional structure was evident: *C. tricinctus* at the NIR location (Norfolk Island) were genetically differentiated from the LHIR region (MR, ER, LHI), whilst *C. galapagensis* at MR/ER were genetically differentiated from LHI within the LHIR region. Furthermore, historical hybridisation is evident between the two anemonefish species, the GBR and McCulloch's anemonefish and possibly between *Coris bulbifrons* and its' sister *C. aygula*.

The observed mtDNA genetic homogeneity in several species examined, may result from a small number of migrants per generation (Slatkin 1993, Shulman 1998, Planes 2002) as has been described in other species; e.g. the endemic Hawaiian butterflyfishes *Chaetodon multicinctus, C. miliaris and C. fremblii* (Craig et al. 2010) and in numerous other coral reef fish species including the parrotfishes *S. frenatus* and *C. sordidus (*Dudgeon et al. 2000; Bay et al. 2004). In contrast, genetic differentiation of peripheral regions likely results from limited gene flow due to geographic isolation (> 100's km separate suitable habitat or coral reefs) and complicated ocean currents (which vary seasonally in strength and migrate north or south along the Tasman Front seasonally; Suthers et al. 2011). Genetic differentiation of peripheral populations has been found in other widespread reef fishes at larger spatial scales; e.g. parrotfish at Hawaii (Bay

et al. 2004); Cocos Keeling, Marquesas and Seychelles (Winters et al. 2010); damselfishes and wrasses at Fiji (Drew et al. 2008); snappers at the Marquesas (Gaither et al. 2010). Lastly, historic hybridization evident between *A. akindynos* and *A. mccullochi* is thought to be facilitated by the sporadic arrival of vagrants (*A. akindynos*) into the distribution range of allopatric sister species (*A. mccullochi*) promoting interbreeding between species due to the rarity of conspecific partners for the vagrants (e.g. Hobbs et al. 2009b). Historical gene flow between species via hybridisation is increasingly being documented among coral reef fishes (e.g. McMillan et al. 1999, van Herwerden & Doherty 2006, van Herwerden et al. 2006, Yaakub et al. 2006, Marie et al. 2007, Hobbs et al. 2009b) and has been reported in a number of anemonefish species, including *A. akindynos* (Sea Read 2009, Fautin & Allen 1997, Timm et al. 2008, Bay & Caley 2011). Hybridisation and reticulate evolution is also common in other coral reef organisms, particularly corals (Willis et al. 2006).

(ii) Demographic connectivity (msatDNA). Demographic connectivity among locations is an order of magnitude lower than historical connectivity (across the same spatial scale) and geographical genetic structure is therefore evident in all four species examined in this study. The remaining two of six species considered herein (*E. daemellii* or *C. galapagensis*) cannot yet be evaluated at this level due to insufficient microsatellite data presently available.

Both *A. akindynos* and *C. bulbifrons*, show genetic differentiation of peripheral regions, which matches colour differentiation and could signal early

peripheral speciation of *A. akindynos* at the southern (SC, NS) edge from its central (GBR) region and, of *C. bulbifrons*, at the eastern (NIR) region from the western (LHIR) region. Peripheral speciation (or budding) has been evident for numerous other endemic reef fishes, many in Hawaii (see review Rocha & Bowen 2008, Bowen et al. 2013) and at the Lord Howe rise (Hodge et al 2012 and references therein). Specific examples include Indo-Pacific damselfish (McCafferty et al. 2002) and western Atlantic wrasse (Rocha 2004), but see Schultz et al. (2007), DiBattista et al. (2012) and Messmer et al. (2005) for contrasting cases where strong geographic components in the distribution of colour do not correspond to genetic partitioning for flame angelfish, pygmy angelfish and dottybacks, respectively.

Models predict that gene flow of a few recruits per generation between locations over historical time scales will not sustain populations (Cowen et al. 2000; Cowen et al. 2002), which must rely on substantial amounts of self-recruitment over demographic time (Swearer et al. 1999; Jones et al. 2005; Almany et al. 2007; Planes et al. 2009). As such, this strong 'lack of congruence' between historical and demographic connectivity is a result of time scale differences in levels of connectivity, as shown for other coral reef fishes; e.g. snappers *Lutjanus carponotatus* (Evans et al. 2010; Harrison et al. 2012) and *Lutjanus synagris* (Gold et al. 2011); coral trout *Plectropomus maculatus* (Evans et al. 2010; Harrison et al. 2012) and in all three endemic species to this region (McCulloch's anemonefish, three-striped butterflyfish, doubleheader wrasse).

(iii) Self-replenishment and recent migration (msatDNA). Self-replenishment at each location is high (68 to 94%) in all four species for which sufficient msatDNA data was available. Additionally, recent migration (msatDNA) suggests a complex pattern of gene flow within the LHIR region (MR, ER, LHI, LHIL). In contrast, the NIR region (NI) receives virtually no gene flow and is considered insular and effectively a closed system on demographic (and possibly historical) time scales.

High levels of self-recruitment are vital for populations to persist (Hastings & Botsford 2006). This is especially true for species that live on isolated islands separated by large distances (e.g. Robertson 2001). Thus, the high levels of self-replenishment (an indirect estimate of self-recruitment as measured by direct approaches, such as otolith tagging or parentage-offspring analyses) in the four comprehensively examined species (excluding black cod and Galapagos sharks), is not surprising. Similar levels of self-recruitment have been shown in other coral reef fishes; e.g. Caribbean wrasses (Swearer et al. 1999); Papua New Guinea anemonefish (Jones et al. 2005, Planes et al. 2009, Berumen et al. 2012); Papua New Guinea butterflyfish (Almany et al. 2007, Berumen et al. 2012). High levels of self-replenishment, coupled with the complex patterns of recent migration within the LHIR region may result from a range of factors, including: (i) propagule trapping eddies (Condie et al. 2011, Mullaney et al. 2011); (ii) spatial and temporal scales of sampling (James et al. 2002); (iii) tidal nature and weak currents around MR/ER (Middleton 2006), which limit water movement and aid self-replenishment; (iv) larval behaviour (e.g. Atema et al. 2002, Dixson

et al. 2008, Horne et al 2011) and (v) local selection pressures such as differences in water temperature (Rocha et al. 2002, 2005, Santos et al. 2006), where locally derived recruits mostly persist, but recruits from different thermal environments rarely persist.

(iv) Genetic diversity (mtDNA – haplotype and nucleotide diversity; msatDNA – genotypic diversity) is high in all species except black cod (very low nucleotide diversity) and Galapagos sharks (low MR-ER region specific mtDNA diversity). I explore possible reasons and give examples for these patterns and levels of genetic diversity below.

Genetic variation is the raw material for historical change upon which natural selection can act (Frankel & Soule 1981), allowing populations to evolve in response to environmental change. However, there are two types of genetic diversity, neutral and adaptive, with important distinctions between them. While neutral genetic diversity can be used to investigate gene flow, migration or dispersal, it is selectively neutral and thus gives no information on the adaptive or evolutionary potential of a population or a species. Adaptive potential is usually assessed in quantitative genetic experiments (Holderegger et al. 2006) where selective traits are examined. In this thesis I used putative neutral genetic diversity (assessed by mtDNA and msatDNA) as an indicator of high total genetic diversity, where mtDNA diversity tracks with nuclear genetic diversity in many marine species (reviewed by Johannesson & Andre 2006), as was shown for the study species here. However, this may not capture the true genetic diversity at

quantitative trait loci that may be particularly important in maintaining a population. For example, specialist species dependent on particular environments or habitats (e.g. specific anemone hosts) may have specific genetic traits that limit their distribution (upper limit: Chevin et al. 2010, Mitchell & Hoffman 2010; lower limit: Kellerman et al. 2009). Thus, populations at peripheral locations or specialist species (and their hosts) may simply lack genetic variation in key traits (e.g. tolerance to low or high temperatures), thus limiting their ability to adapt to conditions beyond their preferred habitat and constraining their evolutionary responses to change. Therefore, a cautious approach is required to prevent population losses, even when high genetic diversity is evident at neutral loci (Chapman et al. 2011), as quantitative trait loci under selection at the peripheral edge of a species distribution range might have no genetic diversity remaining, despite neutral markers displaying relatively high genetic diversity in the same population (Kellerman et al. 2009).

The IUCN Red Data List, genetic traits, the IUCN framework and combined threats to study species

I used four genetic traits (where available) to score a species genetic extinction risk, (i) **historical connectivity** (mtDNA – phylogeny, Minimum Spanning Tree, population genetic indices, migration rates and directions); (ii) **demographic connectivity** (msatDNA – population genetic structure, migration rates and directions); (iii) **self-replenishment and recent migration** (msatDNA). Using these three genetic traits provided a comprehensive understanding of gene flow in

space and time, which when used together, provided an overview of dispersal (Leis et al. 2011); (iv) **genetic diversity** (mtDNA – haplotype and nucleotide diversity; msatDNA – genotypic diversity) which, when combined with connectivity traits (i to iii) above, provides insights into species resilience/extinction risk. In addition to the four genetic traits used here, the analyses used in most of these studies applied the latest analytical software to narrow the gap between indirect and direct estimates of connectivity, as encouraged by Hellberg (2009).

Each trait (historical connectivity, demographic connectivity, self-replenishment and recent migration, genetic diversity) was ranked based on either the level of gene flow or genetic diversity, which resulted in a relative genetic extinction risk matrix. Low levels of gene flow among locations (historic and demographic connectivity), together with high self-replenishment (and consequently low levels of recent migration) and low genetic diversity would all serve to increase extinction risk. In contrast, high levels of gene flow, together with low levels of self-replenishment and high genetic diversity would decrease extinction risk. Thus I ranked the traits as follows; Low = 1, Medium = 2 or High = 3. Genetic trait scores were then summed to produce an overall relative genetic extinction risk score for each species. Overall scores were subsequently ranked (on a minimum to maximum scale of 4 (when all genetic traits scored 1) to 12 (when all genetic traits scored 3) and were identified as more susceptible, S^{+}; less susceptible, S^{-}; not ranked, NR due to data not available, NA or Data Deficient,

DD. Genetic trait(s) that made a species more susceptible (i.e. the weakest link in the chain) were also identified.

The four genetic traits were used to produce a genetic extinction risk matrix that identified *A. akindynos, A. mccullochi* (anemonefishes) and *C. bulbifrons* as more susceptible to extinction than *C. tricinctus*, due largely to the lower levels of historical and demographic connectivity in the former three species (Table 7.1). In contrast, *C. tricinctus* and tentatively, *E. daemelii,* are less susceptible to extinction than the afore-mentioned three species, largely due to higher historical and demographic connectivities of less susceptible species. Notwithstanding the relevance of genetic traits, this represents an incomplete extinction risk matrix, as insufficient sampling of locations throughout the distribution range and insufficient msat data are available for *E. daemelii* and *C. galapagensis* to assess their extinction risk. Thus, for the latter two species further genetic research that builds on existing work by Appleyard & Ward (2007) and van Herwerden et al. (2009) is required. Additional careful consideration of the extinction risk should include differences in life history parameters (specifically generation time, age at maturity, lifespan and fecundity) and differences in fishing pressure (specifically weighting species that are heavily fished compared to species that are less vulnerable to fishing pressure).

Table 7.1 Genetic extinction matrix for six species of coral reef fishes. Genetic traits are ranked as Low = 1, Medium = 2 or High = 3. All genetic trait scores were summed for each species to give an overall extinction risk score for that species. This overall score was subsequently ranked as either S^+ = more susceptible or S^- = less susceptible relative to each other, not ranked, NR due to data not available, NA or Data Deficient, DD.

Common name	Species name	Genetic traits				Total score	Extinction risk	Threats
		Historical connectivity	*Demographic connectivity*	*Self-replenishment and recent migration*	*Genetic diversity*			
Great Barrier Reef anemonefish[1,2]	*Amphiprion akindynos*	2	1	1	3	7	**S^+**	Bleaching events, coastal development
McCullochs anemonefish[3]	*Amphiprion mccullochi*	2	1	1	3	7	**S^+**	Bleaching events
Three striped butterflyfish[4]	*Chaetodon tricinctus*	3	1.5	1	3	8.5	S^-	Bleaching events
Doubleheader wrasse[5]	*Coris bulbifrons*	2	1	1	3	7	**S^+**	Overfishing
Black cod[6, 7]	*Epinephelus daemelii*	2.5 (DD)	2 (DD)	NA	2 (DD)	NR	NR	Overfishing
Galapagos shark[8]	*Carcharhinus galapagensis*	1.5 (DD)	NA	NA	2 (DD)	NR	NR	Overfishing

Genetic data for these species was obtained from a range of studies, as indicated: [1]van der Meer et al. 2012a, [2]van der Meer et al. 2013a, [3]van der Meer et al. 2012b, [4]van der Meer et al. 2012c, [5]van der Meer et al. 2013b, [6]Appleyard & Ward 2007, [7]van Herwerden et al. 2009, [8]van Herwerden et al. 2008.

I discuss below the IUCN Red Data List and the categories and criteria used to assess extinction risk of species, how genetic traits are included into these assessments and finally, I summarise the various threats affecting the study species extinction risk.

(i) The IUCN Red Data List.

The IUCN system is a probabilistic assessment that aims to classify species into vulnerability categories (in the absence of conservation interventions; (i) not evaluated, (ii) least concern, (iii) near threatened, (iv) vulnerable, (v) endangered, (vi) critically endangered, (vii) extinct, (viii) extinct in the wild and (ix) data deficient; http://www.iucnredlist.org/technical-documents/categories-and-criteria) based on certain Criteria (A to E), given that a large proportion of species in the high threat category will go extinct over short periods, usually 10 years or three generations, whichever is longer (Mace et al. 2008). Where possible all available data on ecological, life history and genetic traits are used in assessing a species risk of extinction. Only by incorporating genetic information, along with complementary ecological and life history traits, will managers now, and in the future, be able to comprehensively assess species extinction risks (e.g. Dunham et al. 1999).

(ii) Genetic traits and the IUCN framework.

The four genetic traits used in this study (historical connectivity, demographic connectivity, self-recruitment and recent migration, genetic diversity), considered inclusively, are important to understand patterns and levels of connectivity over both time- and spatial- scales, and the levels of genetic diversity within each population.. They

provide empirical data on population connectivity and genetic diversity, which can guide mangers of marine reserves. Moreover, these genetic traits, when used together, assist reviewers when considering species for inclusion in IUCN Red Lists, since the exchange of individuals among subpopulations over longer time scales influences their likelihood of extinction (Hastings and Botsford 2006), while demographic isolation of geographically isolated populations is particularly important for extinction risk assessments and low genetic diversity reduces a species potential to adapt to changing environmental conditions. Never has this been more pertinent than now, in light of increasingly fragmented habitats and changing global climate (see Jones et al. 2007).

(iii) Combined threats affecting extinction risk of study species

Coral reef habitats are under increasing pressure, concurrently facing a raft of impacts including destructive and excessive fishing; sedimentation - often associated with changes in land use practices and coastal development; pollution in the form of excessive nutrients, plastics and toxic chemicals, most of which enter the marine environment via runoff; disease; coral bleaching; ocean warming and acidification (Hoegh-Guldberg 1999, Hughes et al. 2003, Bellwood et al. 2004). These concurrent disturbances cause sustained and ongoing declines in the abundance of corals on reefs worldwide (e.g. Bellwood et al. 2004, Gardner et al. 2003) with approximately 20% of the world's coral reefs recently destroyed and a further 50% in decline (Wilkinson 2004). Even coral cover on the iconic, massive Great Barrier Reef has halved in the last 27 years (De'ath et al. 2012), a truly concerning phenomenon given the level of protection, management and

relatively small human population along the north Queensland coast compared to most other coral reef environments in the Indo Pacific.

For *A. akindynos* with its Australian distributional stronghold on the GBR, this is concerning as it faces multiple anthropogenic impacts given its close proximity to the coastline. In the case of endemic - *A. mccullochi, C. tricinctus, C. bulbifrons;* restricted range - *E. daemelli* and widespread - *C. galapagensis*, species also examined here, the remote islands that they inhabit are largely unaffected by pressures experienced by coastal reefs. However, a variety of anthropogenic threats may still exist, including sewage leaks into LHIL (Coade et al. 2010); anemone bleaching (Scott & Hill 2011) and coral bleaching due to warming (Jones et al. 2004), affecting all coral reef fish species, especially corallivores such as *C. tricinctus* (Lawton et al, 2011); fishing pressure, which predominantly impacts large apex predator populations including sharks (e.g. Myers et al. 2007) and large cods, such as *E. daemelii* (as shown for the large endemic Hawaiian cod, *Epinephelus quernus),* even if only fished recreationally (Friedlander & DeMartini 2002). The combined impact of these events within this region is a serious cause for concern since studies have shown local extinction following habitat destruction; e.g. in the pink skunk anemonefish (*A. perideraion*) at Sesko Island, Okinawa, Japan (Hattori 2002).

Marine life surveys of MR and ER found that both *A. mccullochi* and *C. tricinctus* were restricted by their habitat requirements to just a few sites at both reefs (Choat et al. 2006, Hobbs and Feary 2007, Hobbs et al. 2009) and 75% of *A. mccullochi* surveyed in 2009 resided in designated high-protection 'sanctuary zones' within the LHI lagoon (Hobbs et al. 2009). Therefore, these species have a high risk of global extinction because they suffer a double (three-striped butterflyfish, *C. tricinctus*) or triple (McCullochs

anemonefish, *A. mccullochi*) jeopardy due to being restricted range endemics, which have low abundances and which are habitat specialists (at a time when bleaching threatens the health of their habitat). In the case of the doubleheader wrasse (*C. bulbifrons*), these populations could quickly decline with minimal fishing effort, especially since specific sampling by Choat et al. (2006) demonstrated that the greatest abundance occurred in shallow sheltered waters, which are most accessible to human activity such as line fishing. The doubleheader wrasse, like most large wrasses, appears to be a fast growing species; however, this may not be an adequate buffer, especially in light of evidence demonstrating rapid population decline with minimal fishing in another large wrasse, *Cheilinus undulatus* (e.g. Choat et al. 2006b). Like *C. bulbifrons*, the locally rare *E. daemelli* (Choat et al. 2006) and *C. galapagensis* face the most direct threat - fishing. Compared to areas with complete protection, cod and shark numbers have declined substantially in partially protected areas open to recreational fishing in both the GBR and Hawaii (Friedlander and DeMartini 2002, Robbins et al 2006). Additionally, local extinctions were reported for Galapagos sharks following the commencement of fishing at the St Pauls Rocks Archipelago in the equatorial Atlantic (Luiz & Edwards 2011). Thus, the abundance and size of apex predators such as black cod and Galapagos sharks at MR and ER are likely to be depleted even if they are not target species (van Herwerden et al. 2008, 2009). Finally, the patchy geographical distribution and unknown level of demographic connectivity (in the form of self-replenishment) between isolated locations where both of these apex predators occur, greatly increases their risk of local extinction due to overfishing, particularly if they are predominantly self-replenishing, as was shown in this study for the anemonefishes, the endemic butterflyfish and the endemic wrasse.

Is there a single spatial management strategy for conserving unique biodiversity found at the LHI and NI rises?

With recent environmental and anthropogenic impacts on the rise, there is an urgent need to establish effective marine reserves that conserve biodiversity or maximise sustainable fishery yields (see Cowen et al. 2009). Population connectivity plays an important role in informing the conservation of marine populations as it provides a framework to test genetic exchange theory (Pringle & Wares 2006, Cowen et al. 2009). Given the wide range of species to be protected in the region, with varying ecological (e.g. habitat preferences, abundance, range size), life history (e.g. growth rate, maximum length, maximum age, age at first reproduction) and genetic (e.g. historical and demographic connectivity, self-recruitment/replenishment and recent migration, genetic diversity) traits, a single management solution (i.e. setting aside priority areas or zones, or complete protection of reefs with no-take zones enforced) is unlikely. Despite the variability in traits, I show that comparable patterns and levels of historical and demographic connectivity as well as self-replenishment levels emerge from the multiple species investigated within the MR - ER - LHI – NI regions. Likewise, I show comparable levels of genetic diversity across study species within this region. Together, this encourages effective protection of many endemic species found in the region by implementing a single or few management strategies. This management strategy, involving a network of MPAs that protect part of each location in the geographic range of endemics, is likely to be effective at conserving the unique biodiversity of this and possibly other endemism hotspots. However, computer models incorporating ecology, life history, population

connectivity, genetic diversity, oceanography and different environmental change scenarios are needed to further our understanding of extinction risk at these remote islands, now and in the future. Unfortunately, the five locations in this region are managed by three different agencies - GBRMPA manages the GBR; NSWDPI manages NS and LHI; SEWPaC manages MR, ER and NI - and historical population connectivity is largely not differentiated geographically, so there is an urgent need for managers to work together and share information to better conserve the unique biodiversity of the region as a whole. I found that the peripheral (southern peripheral and NIR) regions house the most unique diversity (genetic and colour variants), making them particularly valuable, as they contribute an important resource to be conserved and managed (see review by Johannesson & Andre 2006). Furthermore, peripheral habitats (such as isolated islands) may generate new species and subsequently export biodiversity, so they shouldn't be considered evolutionary graveyards (Bowen et al. 2013). Thus, it is vital that (i) a marine reserve be implemented at Norfolk Island, the NIR region, to protect these rare endemic species with unique colour and genetic features and (ii) all other locations in the region continue to be protected, with increased protection for fishing-sensitive species - doubleheader wrasse, black cod and Galapagos sharks at MR, ER and LHI.

Conclusion. The more we understand about population connectivity, close the gap between direct and indirect approaches, and integrate this knowledge with other traits (e.g. ecology, life history) into sophisticated models that incorporate oceanography, the more effective our management efforts will become. This will ultimately lead to better design of marine reserves that takes into account their relative positions within the region

being managed. However, changing climate may affect key parameters, e.g. larval growth rates (Sponaugle & Pinkard 2004, Martell et al. 2005), larval survival, dispersal and recruitment (Munday et al. 2009, Dixson et al. 2008), as well as disrupting present day currents, which may alter and disrupt the self-recruitment mechanism on these islands. Therefore, we need to use data such as is presented in the present study, in combination with information being generated from research into the impact of climate change on populations, to best manage and maintain biodiversity for the future.

BIBLIOGRAPHY

Adams CE, Hamilton DJ, Mccarthy I, Wilson AJ, Grant A, Alexander G, Waldron S, Snorasson SS, Ferguson MM, Skulason S (2006) Does breeding site fidelity drive phenotypic and genetic sub−structuring of a population of Arctic charr? *Evolutionary Ecology*, 20, 11–26.

Allen GR, Hoese DF, Paxton JR, Randall JE, Russell BC, Starck WA, Talbot FH, Whitley GP (1976) Annotated checklist of the fishes of Lord Howe Island. *Records of The Australian Museum*, 30, 365−454.

Allen GR (1972) The anemonefishes: their classification and biology, 2nd ed. TFH Publications, Neptune City, New Jersey.

Allen GR, Steene R, Allen M (1998) A guide to angelfishes and butterflyfishes. Odyssey Publishing−Tropical Reef Research, Perth.

Allen GR (1991) Damselfishes of the world. Mergus Publishers, Melle, Germany.

Allen GR (2008). Conservation hotspots of biodiversity and endemism for Indo−Pacific coral reef fishes. *Aquatic Conservation*, 18, 541−556.

Almany GR, Berumen ML, Thorrold SR, Planes S, Jones GP (2007) Local Replenishment of Coral Reef Fish Populations in a Marine Reserve. *Science*, 316,742−744.

Anderson EC, Thompson EA (2002) A model−based method for identifying species hybrids using multilocus genetic data. *Genetics*, 160, 1217−1229.

Anderson CS, Meikle DB (2010) Genetic estimates of immigration and emmigration rates relative to population density and forest patch area in *Peromyscus leucopus*. *Conservation Genetics*, 11, 1593−1605.

Atema J, Kingsford MJ, Gerlach G (2002) Larval reef fish could use odour for detection, retention and orientation to reefs. *Marine Ecology Progress Series*, 241, 151−160.

Auld TD, Hutton I (2004) Conservation issues for the vascular flora of Lord Howe Island. *Cunninghamia*, 8, 490−500.

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.

Ayre DJ, Hughes TP (2004) Climate change, genotypic diversity and gene flow in reef−building corals. *Ecological Letters*, 7, 273–278.

Barber PH, Bellwood DR (2005) Biodiversity hotspots: Evolutionary origins of biodiversity in wrasses (*Halichoeres*) in the Indo−Pacific and New World Tropics. *Molecular Phylogenetics and Evolution*, 35, 235−253.

Bay LK, Choat JH, van Herwerden L, Robertson DR (2004) High genetic diversities and complex genetic structure in an Indo−Pacific tropical reef fish (*Chlorurus sordidus*): evidence of an unstable evolutionary past? Marine Biology, 144, 757–767.

Bay LK, Crozier RH, Caley MJ (2006) The relationship between population genetic structure and pelagic larval duration in coral reef fishes on the Great Barrier Reef. Marine Biology, 149, 1247−1256.

Bay LK, Caley MJ (2011) Greater Genetic Diversity in Spatially Restricted Coral Reef Fishes Suggests Secondary Contact among Differentiated Lineages. *Diversity*, 3, 483−502.

Bell JJ, Okamura B (2005) Low genetic diversity in a marine nature reserve: re−evaluating diversity criteria in reserve design. *Proceedings of the Royal Society B*, 272, 1067−1074.

Bellwood DR (1996) The Eocene fishes of Monte Bolca, the earliest coral reef fish assemblage. *Coral Reefs*, 15, 11−19.

Bellwood DR, Wainwright PW (2002) The history and biogeography of fishes on coral reefs In, Sale PF (ed) Coral reef fishes, dynamics and diversity in complex ecosystems. Academic Press, San Diego.

Bellwood DR, Hughes TP, Folke C, Nystrom M (2004) Confronting the coral reef crisis. *Nature*, 429, 827−833.

Bellwood DR, Klanten S, Cowman PF, Pratchett MS, Konow N, van Herwerden L (2010) Evolutionary history of the butterflyfishes (f, Chaetodontidae) and the rise of coral feeding fishes. *Journal of Evolutionary Biology*, 23, 335−349.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences*, 98, 4563−4568.

Beerli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology*, 13, 827−836.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57, 289−300

Bennett MB, Gordon I, Kyne PM (SSG Australia & Oceania Regional Workshop, March 2003) Carcharhinus galapagensis. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2. <www.iucnredlist.org>. Downloaded on 24 March 2013.

Berumen ML, Almany GR, Planes S, Jones GP, Saenz−Agudelo P, Thorrold SR (2012). Persistence of self−recruitment and patterns of larval connectivity in a marine protected area network. *Ecology and Evolution*, 2, 444–452.

Bohonak AJ (2002) IBD (Isolation By Distance): A program for analyses of isolation by distance. *Journal of Heredity*, 93, 153−154

Boland FM, Church JA (1981) The East Australian Current 1978. *Deep Sea Research*, 28, 937−957.

Botsford LW, Castilla JC, Peterson CH (1997) The management of fisheries and marine ecosystems. *Science*, 277, 509−515.

Botsford LW (2005) Potential Contributions of Marine Reserves to Sustainable Fisheries, Recent Modeling Results. *Bulletin of Marine Science*, 76, 245−260.

Bowen BW, Muss A, Rocha LA, Grant WS (2006) Shallow mtDNA coalescence in Atlantic pygmy angelfishes (genus *Centropyge*) indicates a recent invasion from the Indian Ocean. *Journal of Heredity*, 97, 1–12.

Bowen BW, Rocha LA, Toonen RJ, Karl SA, ToBo Laboratory (2013) The origins of tropical marine biodiversity. *Trends in Ecology and Evolution*, 1666, (in press).

Brothers EB, Thresher RE (1985) Pelagic duration, dispersal and the distribution of Indo−Pacific coral reef fishes. In, Reaka ML (ed) The ecology of coral reefs. Symposia Series for Undersea Research. National Oceanic and Atmospheric Administration, Washington, DC.

Brothers EB, Williams DM, Sale PF (1983) Length of larval life in twelve families of fishes at 'One Tree Lagoon' Great Barrier Reef, Australia. *Marine Biology*, 76, 319–324.

Buckley TR, Attanayake D, Bradler S (2009) Extreme convergence in stick insect evolution, phylogenetic placement of the Lord Howe Island tree lobster. *Proceedings of the Royal Society B*, 276, 1055−1062.

Budd AF, Pandolfi JM (2010) Evolutionary novelty is concentrated at the edge of coral species distributions. *Science*, 328, 1558− 1560.

Burbrink FT (2010) Historical versus contemporary migration in fragmented populations. *Molecular Ecology*, 19, 5321−5323.

Buston PM, Jones GP, Planes S, Thorrold SR (2012) Probability of successful larval dispersal declines fivefold over 1 km in a coral reef fish. *Proceedings of the Royal Society B*, 279, 1883−1888.

Carlquist S (1974) Island biology. Columbia University Press, New York.

Ceballos G, Brown JH (1995) Global patters of mammalian diversity, endemism and endangerment. *Conservation Biology*, 9, 559−568

Chapman DD, Simpfendorfer CA, Wiley TR, Poulakis GR, Curtis C, *et al.* (2011) Genetic Diversity Despite Population Collapse in a Critically Endangered Marine Fish: The Smalltooth Sawfish Pristis pectinata. *Journal of Heredity*, 102, 643-652.

Chevin LM, Lande R, Mace GM (2010) Adaptation, plasticity, and extinction in a changing environment, towards a predictive theory. *PLoS Biology*, 8, e1000357. doi, 10.1371/journal.pbio.1000357.

Choat JH, Bellwood DR (1998) Wrasses and Parrotfishes. In: Paxton, J.R. and Eschmeyer, W.N. (Eds.) Encyclopaedia Of Fishes. Second Edition. Academic Press, San Diego.

Choat JH, van Herwerden L, Robbins WD, Hobbs J−PA, Ayling AM (2006) A report on the ecological surveys conducted at Middleton and Elizabeth Reefs, February 2006. Report to the Australian Government Department of Environment and Heritage, Canberra, Pp 65 (unpublished report).

Choat JH, Davies CR, Ackerman JL, Mapstone BM (2006b) Demography of a large teleost, *Chelinus undulates*, with a review of size distribution in labrid fishes. *Marine Ecology Progress Series*, 318, 237−246.

Choat JH, Pollard D (2010). *Coris bulbifrons*. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.1. <www.iucnredlist.org>. Downloaded on 01 October 2012.

Ciofi C, Wilson GA, Beheregaray LB, Marquez C, Gibbs JP, *et al.* (2006) Phylogeographic History and Gene Flow Among Giant Gala´pagos Tortoises on Southern Isabela Island. *Genetics*, 172, 1727−1744.

Coade G, Pritchard T, Neilson J, Gudge S, Kerr I (2010) Eutrophication risk in Lord Howe Island lagoon: nutrient and stable isotope monitoring results. NSW Marine Parks Authority, Pp 30 (unpublished report).

Cole AJ, Pratchett MS, Jones GP (2008) Diversity and functional importance of coral−feeding fishes on tropical coral reefs. *Fish and Fisheries*, 9, 286−307.

Coleman N (1980) Australia Sea Fishes South of 30°S. Lane Cove, NSW: Doubleday Australia Pty Ltd.

Condie SA, Mansbridge JV Cahill ML (2011) Contrasting local retention and cross−shore transports of the East Australian Current and the Leeuwin Current and their relative influences on the life histories of small pelagic fishes. *Deep−Sea Research II*, 58, 606−615.

Coyne JA, Orr HA (2004). Speciation. Sunderland, MA: Sinauer Associates, Inc.

Cowen RK, Lwiza KMM, Sponaugle Su, Paris CB, Olson DB (2000) Connectivity of marine populations, open or closed? *Science*, 287, 857−859.

Cowen RK, Paris CB, Olson DB, Fortuna JL (2002) The role of long distance dispersal versus local retention in replenishing marine populations. *Gulf and Caribbean Research*, 1−10.

Cowen RK, Sponaugle S (2009) Larval Dispersal and Marine Population Connectivity. *Annual Review of Marine Science*, 1, 443−466.

Craig MT, Hastings PA, Pondella DJ, Robertson DR, Casian JAR (2006) Phylogeography of the flag cabrilla (*Epinephelus labriformis*, Serranidae): implications for the biogeography of the Tropical Eastern Pacific and the early stages of speciation in a marine shore fish. *Journal of Biogeography*, 33, 969–979.

Craig MT, Eble JA, Bowen BW (2010) Origins, ages and population histories: comparative phylogeography of endemic Hawaiian butterflyfishes (genus *Chaetodon*). J Biogeogr 37: 2125–2136.

Crossland CJ (1988). Latitudinal comparisons of coral reef structure and function. *Proceedures of the 6th International Coral Reef Symposium*, 1, 221–226.

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, 62, 1908−1920.

Crean AJ, Swearer SE, Patterson HM (2009) Larval supply is a good predictor of recruitment in endemic but not non−endemic fish populations at a high latitude coral reef. *Coral Reefs*, 29, 137−143.

Dayton P (1998) Reversal of the burden of proof in fisheries management. *Science*, 279, 821–822.

Dawson MN, Hamner WM (2005) Rapid evolutionary radiation of marine zooplankton in peripheral environments. *Proceeding of the National Academy of Sciences*, 102, 9235–9240.

De'ath G, Fabricius KE, Sweatman H, Puotinen M (2012) The 27–year decline of coral cover on the Great Barrier Reef and its causes. *Proceedings of the National Academy of Science*, 109, 17995–17999.

DiBattista JD, Waldrop E, Bowen BW, Schultz JK, Gaither MR, *et al.* (2012) Twisted sister species of pygmy angelfishes: discordance between taxonomy, coloration, and phylogenetics. *Coral Reefs*, 31, 839–851.

DiBattista JD, Berumen ML, Gaither MR, Rocha LA, Eble JA, *et al.* (2013). After continents divide: comparative phylogeography of reef fishes from the Red Sea and Indian Ocean. *Journal of Biogeography*. doi: 10.1111/jbi.12068

Dixson D, Jones GP, Munday PL, Planes S, Pratchett MS, *et al.* (2008) Coral reef fish smell leaves to find island homes. *Proceedings of the Royal Society B*, 275, 2831−2839.

Doherty PJ (2002) Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem. Sale PF (ed). Academic, San Diego.

Drechsler M, Frank K, Hanski I, O'Hara RB, Wissel C (2003) Ranking metapopulation extinction risk: from patterns in data to conservation management decisions. *Ecological Applications*, 13, 990−998.

Drew J, Allen GR, Kaufman LES, Barber PH **(**2008**)** Endemism and regional color and genetic differences in five putatively cosmopolitan reef fishes. *Conservation Biology*, 224, 965−975.

Drummond AJ, Rambaut A (2007) Bayesian evolution analysis by sampling trees. *BMC Evolution Biology*, 7, doi,10.1186/1471−2148−7−214.

Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, *et al.* (2009) Geneious v4.7. Available from http://www.geneious.com.

Dudgeon CL, Gust N, Blair D (2000) No apparent genetic basis to demographic differences in scarid fishes across continental shelf of the Great Barrier Reef. *Marine Biology*, 137, 1059–1066.

Dulvy NK, Sadovy Y, Reynolds JD (2003) Extinction vulnerability in marine populations. *Fish and Fisheries*, 4, 25−64.

Dunham J, Peacock M, Tracy CR, Nielsen J, Vinyard G (1999) Assessing extinction risk: integrating genetic information. *Conservation Ecology*, 3 (2).

Eble JA, Toonen RJ, Bowen BW (2009) Endemism and dispersal: comparative phylogeography of three surgeonfishes across the Hawaiian Archipelago. *Marine Biology*, 156, 689–698.

Eble JA, Rocha LA, Craig MT, Bowen BW (2010) Not All Larvae Stay Close to Home: Insights into Marine Population Connectivity with a Focus on the Brown Surgeonfish (*Acanthurus nigrofuscus*). *Journal of Marine Biology,* 2010, 1–12.

Edgar GJ, Davey A, Kelly G, Mawbey R and Parsons K (2009) Biogeographical and ecological context for managing threats to coral and rocky reef communities in the Lord Howe Island Marine Park, south−western Pacific. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 20, 378−396.

Evans RD, van Herwerden L, Russ GR, Frisch AJ (2010) Strong genetic but not spatial subdivision of two reef fish species targeted by fishers on the Great Barrier Reef. *Fisheries Research*, 102: 16–25.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131, 479−491

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 30), an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47−50.

Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Conservation*, 61, 1−10

Faubet P, Waples R, Gaggiotti O (2007) Evaluating the performance of a multi−locus Bayesian method for the estimation of migration rates. *Molecular Ecology*, 16, 1149−1166.

Fautin DG, Allen GR (1992) Anemone Fishes and their host Sea Anemones: A guide for aquarists and divers. Western Australian Museum, Perth.

Fautin DG, Allen GR (1997) Anemone fishes and their host sea Anemones, A guide for aquarists and divers. 2nd Edition. Western Australian Museum, Perth.

Forges BR, Koslow JA, Poore GCB (2000) Diversity and endemism of the benthic seamount fauna in the southwest Pacific. *Nature*, 405, 944−947.

Francis MP, Randall JE (1993) Further additions to the fish faunas of Lord Howe and Norfolk Islands, Southwest Pacific Ocean. *Pacific Science*, 47, 118−135.

Frankel OH, Soule ME (1981) Conservation and evolution. Cambridge University Press, Cambridge.

Frankham R (1997) Do island populations have less genetic variation than mainland populations. *Journal of Heredity*, 78, 311−327

Frankham R (1998) Inbreeding and extinction: Island populations. *Conservation biology*, 12, 665−675

Fricke HW (1979) Mating system, resource defense and sex change in the anemonefish *Amphiprion akallopisos*. *Z Tierpsychol*, 50, 313−326.

Friedlander AM, DeMartini EE (2002). Contrasts in density, size, and biomass of reef fishes between the northwestern and the main Hawaiian islands: the effects of fishing down apex predators. *Marine Ecology Progress Series*, 230, 253–264.

Froukh T (2007). Genetic population structure, gene flow, and evolutionary history of selected ornamental fish in the Red Sea, Ph. D. thesis, University of Bremen, Bremen, Germany.

Froukh T, Kochzius M (2007) Genetic population structure of the endemic fourline wrasse (Larabicus quadrilineatus) suggests limited larval dispersal distances in the Red Sea. *Molecular Ecology*, 16, 1359−1367.

Fulgione D, Guglielmi S, Odierna G, Rippa D, Caliendo MF, *et al.* (2008) Morphological Differentiation and Genetic Structure in Island Lizard Populations. *Zoological Science*, 25, 465−474.

Fu Y−X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147, 915−925

Gaither MR, Toonen RJ, Robertson DR, Planes S, Bowen BW (2010) Genetic evaluation of marine biogeographical barriers: perspectives from two widespread Indo−Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). *Journal of Biogeography*, 37, 133–147.

Gardner TA, Cote IM, Gill JA, Grant A, Watkinson AR (2003) Long−Term Region−Wide Declines in Caribbean Corals. Science 301 (5635): 958−960

Gardner MG, Fitch AJ, Bertozzi T, Lowe AJ (2011) Rise of the machines – recommendations for ecologists when using next generation sequencing for microsatellite development. *Molecular Ecology Resources*, 11, 1093–1101.

Gaston KJ (1998) Rarity as double jeopardy. *Nature*, 394, 229–230.

Gerlach G, Atema J, Kingsford MJ, Black KP, Miller−Sims V (2007) Smelling home can prevent dispersal of reef fish larvae. Proceedings of the National Academy of Sciences, 1043, 858−863.

Gillespie RG, Claridge EM, Roderick GK (2008) Biodiversity dynamics in isolates island communities: interaction between natural and human−mediated processes. *Molecular Ecology*, 17, 45−57.

Gillespie RG, Claridge EM, Roderick GK (2008) Biodiversity dynamics in isolates island communities: interaction between natural and human−mediated processes. *Molecular Ecology*, 17, 45−57.

Glynn PW (2012) Global Warming and Widespread Coral Mortality: Evidence of First Coral Reef Extinctions. *Saving a Million Species*, 103−119.

Gold JR, Saillant E, Cummings NJ, Renshaw MA (2011) Genetic Divergence and Effective Size among Lane Snapper in U.S. Waters of the Western Atlantic Ocean. North American Journal of Fisheries Management, 31, 209–223.

Gow JL, Peichel C, Taylor EB (2006) Contrasting hybridization rates between sympatricthree−spined sticklebacks highlight the fragility of reproductive barriers between evolutionarily young species. *Molecular Ecology*, 15, 739–752.

Graham NAJ, Wilson SK, Jennings S, Polunin NVC, Bijoux JP, and Robinson J (2006) Dynamic fragility of oceanic coral reef ecosystems. *Proceedings of the National Academy of Science*, 103, 8425–8429

Graham NAJ, Wilson SK, Jennings S, Polunin NVC, Robinson J, *et al.* (2007) Lag effects in the impacts of mass coral bleaching on coral reef fish, fisheries, and ecosystems. *Conservation Biology*, 215, 1291−1300.

Graham EM, Baird AH,Connolly SR (2008) Survival dynamics of scleractinian coral larvae and implications for dispersal. *Coral Reefs*, 273, 529−539.

Grant WS, Bowen BM (1998) Shallow population histories in deep evolutionary lineages of marine fishes, insights from the sardines and anchovies and lessons for conservation. *Journal of Heredity*, 89, 415−426.

Halpern BS, Warner RR (2003) Review Paper Matching marine reserve design to reserve objectives. *Proceedings of the Royal Society B*, 270, 1871−1871.

Hanski IA (1999) Metapopulation Ecology. Oxford University Press, Oxford.

Harriott VJ (1995) Is the crown−of−thorns starfish a threat to the reefs of Lord Howe Island? *Aquatic Conservation: Marine and Freshwater Ecosystems***,** 5, 179−190.

Harriott VJ, Banks SA (1995) Recruitment of scleractinian corals in the Solitary Islands Marine Reserve, a high latitude coral−dominated community in Eastern Australia. *Marine Ecology Progress Series*, 123, 155−161.

Harriott VJ, Banks SA (2002) Latitudinal variation in coral communities in eastern Australia: a qualitative biophysical model of factors regulating coral reefs. *Coral Reefs*, 21, 83−94.

Harris LN, Taylor EB (2010) Pleistocene glaciations and contemporary genetic diversity in a Beringian fish, the broad whitefish, *Coregonus nasus* (Pallas): inferences from microsatellite DNA variation. *Journal of Evolutionary Biology*, 23, 72−86.

Harrison HB, Williamson DH, Evans RD, Almany GR, Thorrold SR, *et al. (*2012) Larval Export from Marine Reserves and the Recruitment Benefit for Fish and Fisheries. *Current Biology*, 22, 1023−1028.

Harrison HB, Saenz−Agudelo P, Planes S, Jones GP, Berumen ML (2013) Relative accuracy of three common methods of parentage analysis in natural populations. *Molecular Ecology*, 22, 1158−1170.

Harvell CD, Kim K, Burkholder JM, Colwell RR, Epstein PR, *et al.* (1999) Emerging marine diseases−climate links and anthropogenic factors. *Science*, 285, 1505−1510.

Hastings A, Botsford LW (2006) Persistence of spatial populations depends on returning home. *Proceedings of the National Academy of Sciences*, 103, 6067−6072.

Hattori A (2002) Small and large anemonefishes can coexist using the same patchy resources on a coral reef, before habitat destruction. *Journal of Animal Ecology*, 71, 824−831.

Hattori A (2005) High mobility of the protandrous anemonefish *Amphiprion frenatus*: nonrandom pair formation in limited shelter space. *Ichthyological Research*, 52, 57−63.

Heath L (2008) Garnaut Climate Change Review: Impacts of climate change on Australia's World Heritage properties and their values. ANU Institute for Environment, The Australian National University, Pp 25.

Hellberg ME, Burton RS, Neigel JE, Palumbi SR (2002) Genetic assessment of connectivity among marine populations. *Bulletin of Marine Science*, 70, 273−290.

Hellberg ME (2007) Footprints on water: the genetic wake of dispersal among reefs. *Coral Reefs*, 26, 463−473.

Hill R, Scott A (2012) The influence of irradiance on the severity of thermal bleaching in sea anemones that host anemonefish. *Coral Reefs*, 31, 273−284.

Hilborn R, Micheli F, De Leo GA (2006) Integrating marine protected areas with catch regulation. *Canadian Journal of Fisheries and Aquatic Sciences*, 63, 642−649.

Hobbs J−PA, Feary DA (2007) Monitoring the ecological status of Elizabeth and Middleton Reefs Report to Australian Government Department of The Environment and Water Resources, Canberra, Pp 37 (Unpublished report).

Hobbs J−PA, Neilson J, Gilligan JJ (2009) Distribution, abundance, habitat association and extinction risk of marine fishes endemic to the Lord Howe Island region. Report to Lord Howe Island Marine Park, Pp 37 (Unpublished report).

Hobbs JP, Frisch AJ, Allen GR, Van Herwerden L (2009b) Marine hybrid hotspot at Indo−Pacific biogeographic border. *Biology Letters*, 5, 258−261.

Hobbs JP, Jones GP, Munday PL (2010). Rarity and extinction risk in coral reef angelfishes on isolated islands, interrelationships among abundance, geographic range size and specialisation. *Coral Reefs*, 29, 1−11.

Hobbs J−PA, Jones GP, Munday PL (2011) Extinction risk in endemic marine fishes. *Conservation Biology*, 25, 1053–1055.

Hobbs J−PA, van Herwerden L, Pratchett MS, Allen GR (in press). Hybridisation among coral reef butterflyfishes. In: Pratchett MS, Berumen ML, BG Kapoor (eds) Biology of Butterflyfish. Science publishers Inc.

Hodge JR, Read CI, van Herwerden L, Bellwood DR (2012) The role of peripheral endemism in species diversification: Evidence from the coral reef fish genus Anampses (Family: Labridae). *Molecular Phylogenetics and Evolution*, 62, 653–663.

Hoegh−Guldberg O (1999) Coral bleaching, Climate Change and the future of the world's Coral Reefs. *Marine and Freshwater Research*, 50, 839−866.

Hoelzel AR, Fleischer RC, Campagna C, Le Boeuf BJ, Alvord G (2002) Impact of a population bottleneck on symmetry and genetic diversity in the northern elephant seal. *Journal of Evolutionary Biology*, 15, 567−575.

Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, 21, 797–807.

Hooker SK, Gerber LR (2004) Marine Reserves as a Tool for Ecosystem−Based Management: The Potential Importance of Megafauna. *BioScience*, 54, 27−39.

Horne JB, van Herwerden L, Choat JH, Robertson DR (2008) High population connectivity across the Indo−Pacific: Congruent lack of phylogeographic structure in three reef fish congeners. *Molecular Phylogenetics and Evolution*, 49, 629−638.

Horne JB, Momigliano P, Welch DJ, Newman SJ, van Herwerden L (2011) Limited ecological population connectivity suggests low demands on self−recruitment in a tropical inshore marine fish (*Eleutheronema tetradactylum*: Polynemidae). *Molecular Ecology*, 20, 2291−2306.

Horne JB, van Herwerden L (2013) Long-term panmixia in a cosmopolitan Indo-Pacific coral reef fish and a nebulous genetic boundary with its broadly sympatric sister species. *Journal of Evolutionary Biology*, doi: 10.1111/jeb.12092.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9, 1322−1332.

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, 294, 2310−2314.

Huelsenbeck JP, Larget B, Miller RE, Ronquist F (2002) Potential applications and pitfalls of Bayesian inference of phylogeny. *Systematic Biology*, 51, 673−688.

Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, *et al.* (2003) Climate Change, Human Impacts, and the Resilience of Coral Reefs. *Science*, 301, 929−933.

James MK, Armsworth PR, Manson LB, Bode L (2002) The structure of reef fish metapopulations: modelling larval dispersal and retention patterns. *Proceedings of the Royal Society B*, 269, 2079−2086.

Jackson JBC, Kirby MX, Berger WH, Bjorndal KA, Botsford LW, *et al.* (2001) Historical overfishing and the recent collapse of coastal ecosystems. Science 293:629–638.

Jennings S, Kaiser MJ (1998) The effects of fishing on marine ecosystems. *Advances in Marine Biology*, 34, 201–352.

Johannes RE, Weibe WJ, Crossland CJ, Rimmer DW, Smith SV (1983) Latitudinal limits of coral reef growth. *Marine Ecology Progress Series*, 11, 105−111.

Johannesson K, Andre C (2006) Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Molecular Ecology*, 15, 2013−2029.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components, a new method for the analysis of genetically structured populations. *BMC Ge*netics, 11, 94.

Jones JG, McDougall I (1973) Geological history of Norfolk and Philip Islands, southwest Pacific Ocean. *Journal of the Geological Society of Australia*, 20, 239–254.

Jones GP, Caley MJ, Munday PL (2002) Rarity in coral reef fish communities. In: Sale PF (ed) Coral reef fishes: dynamics and diversity in a complex ecosystem. Academic Press, San Diego.

Jones G, McCormick MI, Srivivasan M, Eagle JV (2004). Coral decline threatens fish biodiversity in marine reserves. *Proceedings of the National Academy of Science*, 101, 8251–8253.

Jones GP, Planes S, Thorrold SR (2005) Coral reef fish larvae settle close to home. *Current Biology*, 15, 1314−1318.

Jones AM, Gardner S, Sinclair W (2008) Losing 'Nemo': bleaching and collection appear to reduce inshore populations of anemonefishes. *Journal of Fish Biology*, 73, 753–761.

Jones GP, Almany GR, Russ GR, Sale PF, Steneck RS, *et al.* (2009) *Larval retention and connectivity among populations of corals and reef fishes: history, advances and challenges. Coral Reefs*, 28, 307−325.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology, 16, 1099−1106.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association*, 90, **773**−795.

Keck BP, Near TJ (2009) Geographic and temporal aspects of mitochondrial replacement in Nothonotus darters (Teleostei: Percidae: Etheostomatinae). *Evolution*, 64, 1410−1428.

Kellermann V, van Heerwaarden B, Sgrò CM, Hoffmann AA (2009). Fundamental Evolutionary Limits in Ecological Traits Drive Drosophila Species Distributions. *Science*, 325, 1244−1246.

Klanten OS, Choat JH, van Herwerden L (2007) Extreme genetic diversity and temporal rather than spatial partitioning in a widely distributed coral reef fish. *Marine Biology*, 150, 659−670.

Kokita T, Nakazono A (2001) Rapid response of an obligately corallivorous filefish Oxymonacanthus longirostris (Monacanthidae) to a mass coral bleaching event. *Coral Reefs*, 20, 155−158.

Kuiter RH (1996) Guide to sea fishes of Australia. New Holland, Sydney.

Kuiter, RH (2002) Butterflyfishes, bannerfishes and their relatives: A comprehensive guide to Chaetodontidae and Microcanthidae. TMC Publishing, Chorleywood.

Largier J (2003) Considerations in estimating larval dispersal distances from oceanographic data. *Ecological Applications*, 13, 71−89.

Lawton RJ, Messmer V, Pratchett MS, Bay LK (2011) High gene flow across large geographic scales reduces extinction risk for a highly specialised coral feeding butterflyfish. *Molecular Ecology*, 20, 3584−3598.

Leis JM, van Herwerden L, Patterson HM (2011) Estimating connectivity in marine fish populations: What works best? *Oceanography and Marine Biology: An Annual Review*, 49, 193−234.

Lessios HA, Kessing BD, Pearse JS (2001) Population structure and speciation in tropical seas, global phylogeography of the sea urchin Diadema. *Evolution*, 55, 955−975.

Liao W, Reed DH (2008) Inbreeding–environment interactions increase extinction risk. Animal Conservation, 12, 54−61.

Lowe A, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology*, 19, 3038−3051.

Ludt WB, Bernal MA, Bowen BW, Rocha LA (2012) Living in the Past: Phylogeography and Population Histories of Indo−Pacific Wrasses (Genus *Halichoeres*) in Shallow Lagoons versus Outer Reef Slopes. PLoS ONE 7: e38042. doi:10.1371/journal.pone.0038042.

Mace GM, Collar NJ, Gaston KJ, Hilton−Taylor C, Resit H, *et al.* (2008) Quantification of Extinction Risk: IUCN's System for Classifying Threatened Species. *Conservation Biology*, 22, 1424–1442.

Marie AD, van Herwerden L, Choat JH, Hobbs JP (2007) Hybridisation of reef fish at the Indo−Pacific biogeographic barrier, a case study. *Coral Reefs*, 26, 841−850.

Marine Parks Authority (2010) Natural values of Lord Howe Island. NSW Marine Parks Authority (unpublished report).

Margulies M, Egholm M, Altman WE, (2005) Genome sequencing in microfabricated high−density picolitre reactors. *Nature*, 437, 376−380.

Marshall PA, Baird AH (2000) Bleaching of corals on the Great Barrier Reef: differential susceptibilities among taxa. *Coral Reefs*, 19, 155–163.

Martell DJ, Keiffer JD, Trippel EA (2005) Effects of temperature during early life history on embryonic and larval development and growth in haddock. *Journal of Fish Biology*, 66, 1558–1575.

Mayr E (1940) Speciation phenomena in birds. *The American Naturalist*, 74, 249−278.

Mayr E (1942). Systematics and the origin of species. NY: Columbia University Press, New York.

McClanahan TR (2004) The relationship between bleaching and mortality of common corals. *Marine Biology*, 144, 1239−1245.

McCook LJ, Almany GR, Berumen ML, Day JC, Green AL, *et al.* (2009) Management under uncertainty: guide−lines for incorporating connectivity into the protection of coral reefs. *Coral Reefs*, 28, 353−366.

McDougall I, Embleton BJJ, Stone DB (1981) Origin and evolution of Lord Howe Island, Southwest Pacific. *Journal of the Geological Society of Australia*, 28, 155–176.

McKinney ML (1997) Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annual Review of Ecology, Evolution and Systematics*, 28, 495–516.

McMillan WO, Palumbi SR (1997) Rapid rate of control−region evolution in Pacific butterflyfishes (Chaetodontidae). *Journal of Molecular Evolution*, 45, 473−484.

McMillan WO, Weigt LA, Palumbi, SR (1999) Color pattern evolution, assortative mating, and genetic differentiation in brightly colored butterflyfishes (Chaetodontidae). *Evolution*, 53, 247−260.

McNeely JA, Miller KR, Reid WV, Mittermeier RA, Werner TB (1990) *Conserving the world's biological diversity*. World Conservation Union, World Resources Institute, Conservation International, World Wildlife Fund−US, and the World Bank, Washington, D.C.

Meglecz E, Costedoat C, Dubut V, Gilles A, Malausa T, *et al.* (2009) QDD: a user−friendly program to select microsatellite markers and design primers from large sequencing projects. Bioinformatics Advance Access, Oxford University Press.

Messmer V, van Herwerden L, Munday PL, Jones GP (2005) Phylogeography of colour polymorphism in the coral reef fish *Pseudochromis fuscus*, from Papua New Guinea and the Great Barrier Reef. *Coral Reefs*, 24, 392–402.

Meyer CP, Geller JB, Paulay G (2005) Fine scale endemism on coral reefs: archipelagic differentitaion in turbinid gastropods. *Evolution*, 59, 113−125.

Middleton J, Baird M, Cox D, Estrade P, Everett J, *et al.* (2009) Topographic effects of the East Australia Current outflow at Lord Howe Island, Middleton Reef and Taupo Seamount. RV Southern Surveyor voyage SS06/2009 Voyage Plan and Voyage Summary. CSIRO Division of Marine and Atmospheric Research, Hobart, Tasmania.

Miller KJ, Ayre DJ (2008) Protection of Genetic Diversity and Maintenance of Connectivity among Reef Corals within Marine Protected Areas. *Conservation Biology*, 22, 1245–1254.

Miller JA, Shanks AL (2004) Evidence for limited larval dispersal in black rockfish (*Sebastes melanops*), implications for population structure and marine−reserve design. *Canadian Journal of Fisheries and Aquatic Sciences*, 61, 1723−1735.

Mitchell KA, Hoffman AA (2010) Thermal ramping rate influences evolutionary potential and species differences for upper thermal limits in Drosophila. *Functional Ecology*, 24, 694−700.

Mora C, Sale PF (2002) Are populations of coral reef fishes open or closed? *Trends Ecology and Evolution*, 17, 422−428 .

Montanari SR, van Herwerden L, Pratchett MS, Hobbs J−PA (2011) Reef fish hybridization, lessons learnt from butterflyfishes (genus Chaetodon). *Ecology and Evolution*, 2, 310 −328.

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

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, 51, 238−254.

Mullaney TJ, Miskiewicz AG, Baird ME, Burns PTP, Suthers IM (2011) Entrainment of larval fish assemblages from the inner shelf into the East Australian Current and into the western Tasman Front. *Fisheries Oceanography*, 20, 434−447.

Mullon C, Fréon P, Cury P (2005) The dynamics of collapse in world fisheries. *Fish and Fisheries,* 6, 111–120.

Munday PL (2004) Habitat loss, resource specialisation, and extinction on coral reefs. *Global Change Biology*, 10, 1642−1647.

Munday PL, Leis JM, Lough JM, Paris CB, Kingsford MJ, *et al.* (2009a) Climate change and coral reef connectivity. *Coral Reefs*, 28, 379−395.

Munday PL, Donelson JM, Dixson DL, Geoff GK (2009b) Effects of ocean acidification on the early life history of a tropical marine fish. *Proceedings of the Royal Society B,* 276, 3275−3283.

Muss A, Robertson DR, Stepien CA, Wirtz, P, Bowen BW (2001) Phylogeography of *Ophioblennius*, the role of ocean currents and geography in reef fish evolution. *Evolution*, 55, 561−572.

Nelson JS (1976) Fishes of the world. New York: Wiley and Sons.

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, 3, 489−495.

Nettel A, Dodd RS, Afzal−Rafil Z, Tovilla−Hernandez C (2008) Genetic diversity enhanced by ancient introgression and secondary contact in East Pacific black mangroves. *Molecular Ecology*, 17, 2680−2690.

NRC (US National Research Council) (1995) Understanding marine biological diversity. National Academy Press, Washington, DC.

O'Hara TD (2007) Seamounts: centres of endemism or species richness for ophiuroids? *Global Ecology and Biogeography*, 16, 720−732.

Oxley WG, Ayling AM, Cheal AJ, Osborne K (2004) Marine Surveys undertaken in the Elizabeth and Middleton Reefs Marine National Nature Reserve, December 2003. Australian Institute of Marine Science. www.aims.gov.au/reef−monitoring

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real−time estimation of migration rate, a simulation−based exploration of accuracy and power. *Molecular Ecology*, 13, 55−65.

Palsboll PJ, Berube M, Allendorf FW (2007) Identification of management units using population genetic data. *Trends in Ecology and Evolution*, 22, 11−16.

Palstra FP, O'Connell MF, Ruzzante DE (2007) Population structure and gene flow reversals in Atlantic salmon Salmo salar over contemporary and long−term temporal scales: effects of population size and life history. *Molecular Ecology*, 16, 4504−4522.

Patterson HM, Swearer SE (2007) Long−distance dispersal and local retention of larvae as mechanisms of recruitment in an island population of a coral reef fish. *Australia Ecology,* 32, 122−130.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295.

Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology and Evolution*, 24, 386−393.

Philander SG, Holton JR, Dmowska R (1990) El Nino, La Nina, and the Southern Oscillation, Volume 46 (International Geophysics). Academic Press.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, *et al.* (2004) GENECLASS2, a software for genetic assignment and first generation migrants detection. *Journal of Heredity*, 95, 536−539.

Pitcher TJ, Hart PJB (1982) Fisheries ecology. Croom Helm, London.

Pitcher TJ (2001) Rebuilding ecosystems as a new goal for fisheries management: reconstructing the past to salvage the future. *Ecological Applications*, 11, 601–617.

Planes S, Fauvelot C (2002) Isolation by distance and vicariance drive genetic structure of a coral reef fish in the Pacific ocean. *Evolution*, 56, 378−39

Planes S, Jones GP, Thorrold SR (2009) Larval dispersal connects fish populations in a network of marine protected areas. *Proceedings of the National Acadamy of Science*, 106, 5693−5697.

Posada D (2008) jModelTest,Phylogenetic Model Averaging. *Molecular Biology and Evolution*, 25, 1253−1256.

Postma E, van Noordwijk AJ (2005) Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature*, 433, 65−68.

Pikitch EK, Santora C, Babcock EA, Bakun A, Bonfil R, *et al.* (2004) Ecosystem−based fishery management. *Science*, 305, 346−347.

Pratchett MS, Wilson SK, Baird AH (2006) Declines in the abundance of Chaetodon butterflyfishes (Chaetodontidae) following extensive coral depletion. *Journal of Fish Biology*, 69, 1269−1280.

Pratchett MS, Wilson SK, Graham NAJ, Munday PM, Jones GP, *et al.* (2007) Multi−scale temporal effects of climate−induced coral bleaching on motile reef organisms. In, van Oppen M and Lough JM (eds) Coral bleaching, patterns and processes, causes and consequences. New York, Springer.

Pratchett MS, Munday PL, Wilson SK, Graham NAJ, Cinner JE, *et al.* (2008) Effects of climate induced coral bleaching on coral reef fishes−ecological and economic consequences. ***Oceanography and Marine Biology, An Annual Review*,** 46, 251−296.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155, 945−959.

Randall JE (1976) The endemic shore fishes of the Hawaiian Islands, Lord Howe Island, and Easter Island. Trav Doc O.R.S.T.O.M. (Commerson Colloq., Re´union), 47:49–73.

Randall JE, Kuiter RH (1982) Three New Labrid Fishes of the Genus *Coris* from the Western Pacific. *Pacific Science,* 36, 159−173.

Randall JE (1998) Zoogeography of shore fishes of the Indo−Pacific region. *Zoological Studies*, 37, 227−268.

Randall JE (2007). Reef and Shore Fishes of the Hawaiian Islands. Honolulu, HI: University of Hawaii.

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. *Conservation Biology*, 17, 230–237.

Rhymer JM, Simberloff D (1996) Extinction by hybridisation and introgression. *Annual Review of Ecology, Evolution, and Systematics*, 27, 83−109.

Richardson DL (1998) Descriptions of the colour patterns of juvenile anemonefishes (Pomacentridae, Amphiprion) from New South Wales and the Lord Howe − Norfolk Island Region. *Proceedings of the Linnean Society of NSW*, 120, 81−86.

Richardson DL (1999) Correlates of environmental variables with patterns in the distribution and abundance of two anemonefishes (Pomacentridae, Amphiprion) on an eastern Australian sub−tropical reef system. *Environmental Biology of Fish*, 55, 255−263

Roberts CM (1997) Connectivity and Management of Caribbean Coral Reefs. *Science*, 278, 1454−1457.

Roberts CM, Hawkins JP (1999) Extinction risk in the sea. *Trends in Ecology and Evolution*, 14, 241−246.

Roberts CM, McClean CJ, Veron JEN, Hawkins JP, Allen GR, *et al.* (2002) Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science*, 295, 1280–1284.

Robertson DR (2001) Population maintenance among tropical reef fishes: Inferences from small−island endemics. *Proceedings of the National Academy of Sciences*, 98, 5667–5670.

Rocha LA, Bass AL, Robertson DR, Bowen BW (2002) Adult habitat preferences, larval dispersal, and the comparative phylogeography of three Atlantic surgeonfishes Teleostei, Acanthuridae. *Molecular Ecology*, 11, 243−252.

Rocha LA (2003) Patterns of distribution and processes of speciation in Brazilian reef fishes. *Journal of Biogeography*, 30, 1161−1171.

Rocha LA (2004) Mitochondrial DNA and color pattern variation in three western Atlantic Halichoeres (Labridae), with the revalidation of two species. *Copeia*, 2004, 770–782.

Rocha LA, Robertson DR, Roman J, Bowen BW (2005) Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society*, 272, 573−579.

Rocha LA, Bowen BW (2008) Speciation in coral−reef fishes. *Journal of Fish Biology*, 72, 1101−1121.

Rotjan RD, SM Lewis (2008) Impact of coral predators on tropical reefs. *Marine Ecology Progress Series*, 367, 73−91.

Rousset F (2008) GENEPOP ' 007: a complete re−implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, 8, 103−106.

Roughgarden J, Smith F (1996) Why fisheries collapse and what to do about it. *Proceedings of the National Academy of Science*, 93, 5078–5083.

Sale PF, Cowen RK, Danilowicz BS, Jones GP, Kritzer JP, et al. (2005) Critical science gaps impede use of no−take fishery reserves. *Trends in Ecology and Evolution*, 20, 74−80.

Santini S, Polacco G (2006) Finding nemo, molecular phylogeny and evolution of the unusual life style of anemonefish. *Gene*, 385, 19−27.

Santos S, Herbert T, Farias IP, Schneider H, Sampaio I (2006) Population genetic structuring of the king weakfish, *Macrodon ancylodon* Sciaenidae, in Atlantic coastal waters of South America, deep genetic divergence without morphological change. *Molecular Ecology*, 152, 4361−4373.

Scascitelli M, Whitney KD, Randell RA, King M, Buerkle CA, *et al.* (2010) Genome scan hybridizing sunflowers from Texas (Helianthus annuus and H debilis) revelas asymmetric patterns of introgression and small islands of genomic differentiation. *Molecular Ecology*, 19, 521−541.

Schiel DR, Kingsford MJ, Choat JH (1986) Depth, distribution and abundance of benthic organisms and fishes at the subtropical Kermadec Islands. *New Zealand Journal of Marine and Freshwater Research*, 20, 521−535.

Schultz JK, Pyle RL, DeMartini E, Bowen BW (2007) Genetic connectivity among color morphs and Pacific archipelagos for the flame angelfish, *Centropyge loriculus*. *Marine Biology*, 151, 167–175.

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution*, 22, 25–33.

Sea Read (2009) World's first − hybrid fish at Reef HQ Aquarium. Issue 26. www.gbrmpa.gov.au/__data/.../SeaRead_26_Final_QXPscreen.pdf

Selkoe, KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, 9, 615–629.

Shanks AL, Grantham BA, Carr MH (2003) Propagule dispersal distance and the size and spacing of marine reserves. *Ecological Applications*, 13, 159−169.

Sheppard CRC (2003) Predicted recurrences of mass coral mortality in the Indian Ocean. *Nature*, 425, 294−297.

Shimizu M, Kosaka N, Shimada T, Nagahata T, Iwasaki H, *et al.* (2002) Universal fluorescent labeling (UFL) method for automated microsatellite analysis. *DNA Research*, 9, 173–178.

Shulman MJ (1998) What can population genetics tell us about dispersal and biogeographic history of coral−reef fishes. *Australian Journal of E*cology, 23, 216–225.

Slatkin M (1987) Gene Flow and the Geographic Structure of Natural Populations. *Science*, 236, 787–792.

Slatkin M (1993) Isolation by distance in equilibrium and non−equilibrium populations. *Evolution*, 47, 264−279

Smith LD, Gilmour JD, Heyward AJ (2008) Resilience of coral communities on an isolated system of reefs following catastrophic mass−bleaching. *Coral Reefs*, 27, 197−205.

Sokal RR, Rohlf FJ (1994) Biometry: the principles and practice of statistics in biological research, 3rd edn. W.H. Freeman, New York.

Spalding MD, Jarvis GE (2002) The impact of the 1998 coral mortality on reef fish communities in the Seychelles. *Marine Pollution Bulletin*, 44, 309−321.

Sponaugle S, Pinkard DP (2004) Impact of variable pelagic environments on natural larval growth and recruitment of a reef fish. *Journal of Fish Biology*, 64, 34–54.

Steadman DW (1995) Prehistoric extinctions of Pacific island birds− Biodiversity meets zooarchaeology. *Science*, 267, 1123−1131.

Steffen W, Crutzen PJ, McNeill JR (2007) The Anthropocene: Are Humans Now Overwhelming the Great Forces of Nature. *AMBIO: A Journal of the Human Environment*, 36, 614−621.

Storey JD (2002). A direct approach to false discovery rate. *Journal of the Royal Statistical Society Series B*, 64, 479−498.

Suchard MA, Weiss RE, Sinsheimer JS (2001) Bayesian selection of continuous time Markov chain evolutionary models. *Molecular Biology and Evolution*, **18**, 1001−1013.

Suthers IM, Young JM, Roughan MR, Everett JD, Brassington GB, *et al.* (2011) The strengthening East Australian Current, its eddies and biological effects−an introduction and overview. *Deep−Sea Research Part II*, 58, 538−546.

Swearer SE, Caselle JE, Lea DW, Warner RR (1999) Larval retention and recruitment in an island population of coral−reef fish. *Nature*, 402, 799–802.

Swofford DL (2001) PAUP*. Phylogenetic analysis using parsimony (and other methods), version 4.0b6. Sinauer Associates, Sunderland, Mass.

Syms C (1998) Disturbance and the structure of coral reef fish communities on the reef slope. *Journal of Experimental Marine Biology and Ecology*, 230, 151−167.

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics*, 105,437−460

Tamura K, Dudley J, Nei M (2007) MEGA4, Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution, 24, 1596−1599.

Thresher RE, Colin PL, Bell LJ (1989) Planktonic duration, distribution and population structure of western and central Pacific damselfishes (Pomacentridae). *Copeia*, 1989, 420–434.

Timm J, Figiel M, Kochzius M (2008) Contrasting patterns in species boundaries and evolution of anemonefishes (Amphiprioninae, Pomacentridae) in the centre of marine biodiversity. *Molecular Phylogenetics and Evolution*, 49, 268−276.

van der Meer MH, Hobbs J−PA, Jones GP, van Herwerden L (2012a) Historic hybridisation between two Australian anemonefish species *Amphiprion* and present−day patterns of connectivity. *Ecology and Evolution*, 2, 1592–1604.

van der Meer MH, Hobbs J−PA, Jones GP, van Herwerden L (2012b) Genetic connectivity among and self−replenishment within island populations of a restricted range subtropical reef fish. *PLoS ONE* 7(11), e49660.

van der Meer MH, Gardner MG, Hobbs J−PA, Jones GP, van Herwerden L (2012c) Identification of seventeen microsatellite markers for conservation genetic studies of the endemic anemonefish *Amphiprion mccullochi*. *Conservation Genetic Resources*, 4, 247−250.

van der Meer MH, Gardner MG, Hobbs J−PA, Pratchett MS, van Herwerden L (2012d) Identification of twenty one microsatellite loci for conservation genetic studies of the endemic butterflyfish *Chaetodon tricinctus*. *Conservation Genetic Resources*, 4, 243−246.

van der Meer MH, Gardner MG, Berumen M, Hobbs J−PA, van Herwerden L (2012e) Identification of seventeen microsatellite loci for conservation genetic studies of the endemic wrasse *Coris bulbifrons*. *Conservation Genetic Resources*, 5, 363−366.

van der Meer MH, Horne JB, Gardner MG, Hobbs J−PA, Pratchett M, van Herwerden L (2013a) Limited demographic gene flow and high self−replenishment drives peripheral isolation in an endemic coral reef fish. *Ecology and Evolution*, DOI: 10.1002/ece3.584.

van der Meer MH, Bay LK, Scott A, Hobbs J−PA, Jones GP, Hughes J, van Herwerden L (2013b) Species in the making: molecular and morphological evidence of peripatric speciation in two east Australian anemonefishes, *Journal of Biogeography* (in review).

van der Meer MH, Hobbs J−PA, Berumen M, van Herwerden L (2013c). Population connectivity and the effectiveness of a marine reserve network in protecting a vulnerable, exploited and endemic coral reef fish. *Conservation Genetics* (in review).

van Herwerden L, Benzie J, Peplow L, Davies C (2000) Microsatellite markers for coral trout *Plectropomus laevis* and red throat emperor *Lethrinus miniatus* and their utility in other species of reef fish. *Molecular Ecology*, 9, 1929–1931.

van Herwerden L, Benzie J, Davies C (2003) Microsatellite variation and population genetic structure of the red throat emperor on the Great Barrier Reef. *Journal of Fish Biology*, 62, 987–999.

van Herwerden L, Doherty PJ (2006) Contrasting genetic structures across two hybrid zones of a tropical reef fish, *Acanthochromis polyacanthus* (Bleeker 1855). *Journal of Evolutionary Biology*, 19, 239−252.

van Herwerden L, Choat JH, Dudgeon CL, Carlos G, Newman SJ, *et al.* (2006) Contrasting patterns of genetic structure in two species of the coral trout *Plectropomus* (Serranidae) from east and west Australia: introgressive hybridisation or ancestral polymorphisms. *Molecular Phylogenetics Evolution*, 41, 420–435.

van Herwerden L, Choat JH, Newman SJ, Leray M, Hillersøy G (2009) Complex patterns of population structure and recruitment of *Plectropomus leopardus* (Pisces, Epinephelidae) in the Indo−West Pacific, implications for fisheries management. *Marine Biology*, 156, 1595−1607.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO−CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4, 535–553.

Victor BC (1986) Duration of the planktonic larval stage of one hundred species of Pacific and Atlantic wrasses (family Labridae). *Marine Biology*, 90, 317–326.

Victor BC, Wellington GM (2000) Endemism and the pelagic larval duration of reef fishes in the eastern Pacific Ocean. *Marine Ecology Progress Series*, 205, 241–248.

Visser ME (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society B*, 275, 649−659.

Walsh PS, Metzger DA, Higuchi R (1991) Chelex−100 as a medium for simple extraction of DNA for PCR based typing from forensic material. *Biotechniques*, 10, 506−513.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, 15, 1419−1439.

Wellington GM, Victor BC (1989) Planktonic larval duration of one hundred species of Pacific and Atlantic damselfishes (Pomacentridae). *Marine Biology*, 101, 557–568.

West−Eberhard MJ (1983) Sexual selection, social competition, and speciation. *The Quaterly review of Biology*, 58, 155−183.

Whittaker RJ (1998) Islands biogeography: Ecology, evolution, and conservation. Oxford University Press, Oxford

Whittaker RJ, Fernández−Palacios JM (2007) Island Biogeography: Ecology, Evolution, and Conservation2nd edition. Oxford University Press.

Wilkinson C (2002) Status of Coral Reefs of the World, 2002. Townsville, Australia, Australian Institute of Marine Science.

Willis TJ, Millar RB, Babcock RC, Tolimieri N (2003) Burdens of evidence and the benefits of marine reserves: Putting Descartes before deshorse? *Environmental Conservation*, 30, 97−103.

Willis BL, Page CA, Dinsdale EA (2004) Coral disease on the Great Barrier Reef. In: Rosenberg E, Loya Y (eds) Coral disease and health. Springer, Berlin.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, 163, 1177−1191.

Wilson SK, Graham NAJ, Pratchett MS, Jones GP, Polunin NVC (2006) Multiple disturbances and the global degradation of coral reefs, are reef fishes at risk or resilient? *Global Change Biology*, 12, 2220−2234.

Wilson, SK, Fisher R, Pratchett M, Graham NAJ, Dulvy NK, Turner RA, Cakacaka A, Polunin NVC, Rushton SP (2008) Exploitation and habitat degradation as agents of change within coral reef fish communities. *Global Change Biology*, 14, 2796−2809.

Winters KL, van Herwerden L, Choat JH, Robertson DR (2010) Phylogeography of the Indo−Pacific parrotfish Scarus psittacus: isolation generates distinctive peripheral populations in two oceans. *Marine Biology*, 157, 1679−1691.

Yaakub SM, Bellwood DR, van Herwerden L, Walsh FM (2006) Hybridization in coral reef fishes, introgression and bi−directional gene exchange in Thalassoma (family Labridae). *Molecular Phylogenetics and Evolution*, 40, 84−100.

Zwickl DJ (2006**)** Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. thesis, The University of Texas at Austin, USA.

Wood R (1999) Reef Evolution. Oxford University Press, NewYork.

APPENDIX 1: **Identification of seventeen microsatellite markers for conservation genetic studies of the endemic anemonefish, *Amphiprion mccullochi***

(Published as: van der Meer, M.H., Gardner, M.G., Hobbs, J-P.A, Jones, G. & van Herwerden, L. (2012) Identification of seventeen microsatellite markers for conservation genetic studies of the endemic anemonefish *Amphiprion mccullochi*. *Conservation Genetic Resources*, 4, 247−250.)

A1.1 ABSTRACT

Endemic species at remote islands have a high risk of extinction because they often exhibit ecological, biological and genetic traits that make them particularly vulnerable to disturbances such as habitat loss and overharvesting. McCulloch's anemonefish (*Amphiprion mccullochi*) is endemic to a narrow oceanic region off Australia's east coast (containing Elizabeth Reef, Middleton Reef and Lord Howe Island) and is an extreme habitat specialist living on a single anemone host species, *Entacmaea quadricolor*. Using 454 shotgun sequencing, I developed primers for seventeen independent microsatellite loci to reveal the gene flow, population genetic structure and genetic diversity across these three reefs. Observed heterozygosities ranged from 0.556 to 1 and expected heterozygosities ranged from 0.543 to 0.907 in 30 individuals from Lord Howe Island. When cross tested with the close relative, *Amphiprion akindynos*, there was good amplification success and a high level of polymorphism. Therefore, these loci will be useful in *A. mccullochi*, *A. akindynos* and possibly other closely related anemonefish.

A1.2. INTRODUCTION

Remote offshore islands usually contain high numbers of endemic species (Ceballos and Brown 1995, Randall 1998, Gillespie et al. 2008), which makes them a high conservation priority (Gillespie & Roderick 2002). Conserving endemics is challenging because of their high extinction risk due to factors such as small population size, ecological specialisation, limited dispersal and low genetic diversity (McKinney 1997, Frankham 1998). Furthermore, endemic species are often highly sought after and exhibit traits that make them vulnerable to overharvesting.

In the marine environmnet, anemonefishes are an iconic group that contain mainly endemic species. This group, particularly the endemics, are highly sought after by commercial and recreational aquarium collecting trade. In addition to overharvesting, anemonefish populations are also at risk from global warming because increasing seawater termperatures cause their host anemone to bleach and die (Jones et al. 2008). Anemonefish cannot live without their host anemones and death of anemones during past bleaching events have resulted in the local extinction of anemonefish (Hattori 2002, 2005).

McCulloch's anemonefish (*Amphiprion mccullochi*) has arguably the smallest geographic range size of any anemonefish. It is endemic to Australian waters where it inhabits coral rich areas of lagoon and seaward reefs at Elizabeth Reef (ER), Middleton Reef (MR), Lord Howe Island (LHI; Coleman 1980). Moreover, it is a habitat specialist living on a single host anemone, *Entacmaea quadricolor* (Fautin & Allen 1992), which puts it a greater risk than other species with more host anemone choices. Previous studies found

that the McCulloch's anemonefish reaches its highest abundance at LHI and is relatively rare at ER and MR (Choat et al. 2006a, Hobbs & Feary 2007, Hobbs et al. 2009). Specifically, LHI supports 92% of the world's *A. mccullochi* population and most individuals are found in the shallow lagoon where some of the host anemones (*E. quadricolor* have been observed to bleach and die (Hobbs et al. 2009). Clearly, *A. mccullochi* has a high risk of extinction and determining the genetic characteristics of the 3 populations is a conservation priority.

To date, the gene flow, population genetic structure and genetic diversity of *A. mccullochi* have not been examined. This information will be vital to infer if the remnant high abundance populations at LHI could replenish or re-stock ER and MR after bleaching events and genetic diversity data will show how resilient these populations are to future environmental change. This paper describes the development of 17 polymorphic microsatellite markers for *A. mccullochi* collected from Lord Howe Island (n = 30) using 454 shotgun pyrosequencing on a Roche GS-FLX (Australain Genome Research Facility, AGRF, Brisbane, Australia). These markers will be useful to quantify population connectivity of *A. mccullochi*. Preliminary tests of cross-species amplification of these loci on *A. akindynos* are also described.

A1.3. MATERIAL AND METHODS

As per Chapter 2.3.1. (PCR reactions)

A1.4. RESULTS AND DISCUSSION

The resulting sequences (totalling 114,272 reads with an average sequence length of 350 and total GC content of 40.08%) were screened for pure di, tri, tetra, penta, and hexanucleotide microsatellite loci with six or more repeats using the default settings of QDD v1.3 (Meglecz et al. 2009) which identified 7224 (6.32% of sequences) microsatellite loci. PCR primers were successfully designed for 923 loci (12.78% of loci found). Of these, primers for the best of the best 24 loci were synthesised with a 5´ 17 base pair lambda tag (5-GGTGGCGACTCCTGGAG-3) to enable indirect fluorolabelling (Shimizu et al 2002) and minimize direct costs. Loci were initially tested for amplification success and specificity in eight individuals using Type-it microsatellite PCR kit (Qiagen). Individual amplifications were performed in 10 µL reactions, containing 1x Type-it Multiplex PCR Master Mix (Qiagen, Doncaster/Australia), 20-50 ng template, and 0.2 µM each primer (forward and reverse). Indirectly labelled reactions contained a tailed forward primer and a reporter primer (5' labelled with fluorescent dye modification HEX, TET or FAM) at a 1:4 ratio (total = 0.2 µM). All primers were tested and optimised using a Bio-Rad C1000 Thermal Cycler (see Table 1) with an initial denaturation of 94°C for 3 mins followed by 28 cycles of 94°C for 40 s, 58°C for 40 s and 72°C for 40s followed by 5 mins at 72°C in a Bio-Rad C1000 Thermal Cycler (Bio-Rad, Australia). PCR products were column purified using an Ammonium Acetate protocol. Genotypes were run on an Amersham Bioscinces Megabase Capillary Sequencer with a 400 bp size standard and scored using Fragment Profiler 1,2 (Amersham, Buckinghamshire, England).

Primer pairs for 17 loci reliably amplified products of the expected size, with no additional products and were polymorphic, representing one hexamer, two pentamers, eight tetramers, and four trimers. Directly-labelled forward primers (HEX, TET or FAM) were synthesised for the 17 polymorphic *A. mccullochi* loci with PCR multiplexing of three loci (Table 1). Loci were genotyped in directly labelled multiplex reactions in 30 *A. mccullochi* individuals from MR. DNA extraction protocol and PCR conditions as described above. Multiplex PCR combinations (Table 1) were designed and tested using PCR conditions described above.

Characteristics of the 17 loci are summarised in Table A1.1. GENALEX 6 (Peakall & Smouse 2006) was used to determine the number of alleles, observed and expected heterozygosities and, conformation to Hardy- Weinberg Equilibrium (HWE). CERVUS 3.0 (Kalinowski et al. 2007) was used to calculate polymorphic information content (PIC) for each locus, GENEPOP 4.0.10 (Rousset 2008) was used to test linkage disequilibrium and, MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) was used to detect the presence of null-alleles.

Table A1.1: Details for seventeen *Amphiprion mccullochi* microsatellite loci developed from 454 shotgun sequences. T_A, annealing temperature; *N*, sample size; N_A, number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; *p* HWE, Hardy-Weinberg equilibrium significance value at $p < 0.05$ after FDR correction; PIC, polymorphic information content. ^, locus may have null alleles.

Locus	Repeat motif	Primer sequence 5' to 3'	Size range	N	N_a	H_o	H_e	pHWE	PIC	Genbank
Amphiprion mccullochi										
Am1 [1A]	$(TG)_{12}$	[TET] ACAAAGCCTTCATGTGGGTC CGCAAGTGTTGCCTCATAGA	108-112	27	2	0.556	0.543	**0.007**	0.444	JN604050
Am4 [2A]	$(GT)_{13}$	[TET] AGCTGGTTGGGTGTTACCTG ATGGCTTCAGTCTGCTGGTT	120-156	30	11	0.900	0.819	0.901	0.802	JN604051
Am5 [2C]	$(AC)_{20}$	[FAM] CCACTAGAGGCTCCCTGTTG CGAGCATGAGCGAATGTATG	77-95	30	9	0.800	0.844	0.910	0.827	JN604052
Am6 [1C]	$(TGT)_{10}$	[HEX] AGCAGAGAGGAAAGAAGGGC CAAGTGCCTGGCAGAAGATT	262-274	27	4	0.667	0.620	0.724	0.542	JN604053

Am7 [3A]	$(ATG)_{19}$	[TET] TGTCGCTACGACAGACTGCT GCATGAGTGATTGGACCCTA	86-128	28	10	0.786	0.851	0.972	0.834	JN604054
Am9 [1B]	$(TTA)_{17}$	[FAM] TGCTGCACTCTGTCTATTTTGT GTGACTGAAGGCAAGGCAAT	151-202	27	13	1.000	0.907	0.203	0.899	JN604055
Am10 [4C]	$(ACAG)_{16}$	[HEX] GGAAGCAGCAATAAAGACGC AGAGACGCCTGATGGTGAGT	286-302	30	5	0.600	0.635	0.981	0.586	JN604056
Am11 [4B]	$(CTAT)_{12}$	[FAM] ATTCCCCGACGGAGAGTAGT TGTCGCTTTGTGACACCTTC	124-172^	29	7	0.621	0.782	0.127	0.751	JN604057
Am12 [3C]	$(TTCA)_{20}$	[HEX] ATGAGCAGCTTTGACGGAAT ACCTACATGGTTGGAGCCTG	182-210	28	8	0.821	0.833	0.830	0.811	JN604058
Am14 [2C]	$(GATG)_{13}$	[HEX] CAGCAGCCTCAAGTGACTGT GCAGCATTCTCACACACCAC	170-198	28	8	0.750	0.763	0.898	0.741	JN604059

Am15 [5A]	$(GTCT)_{12}$	[TET] ACTAGGCTCAGAGCAGGGTC CAAGTCAATCAAAGCAGCCA	100-160	27	11	0.889	0.875	0.647	0.863	JN604060
Am17 [6A]	$(AATA)_{14}$	[TET] GGCTGTCTGGGATGAGATGT TGTTCTGCAGATGGACTGTTTT	105-138	27	9	0.593	0.823	0.219	0.8	JN604061
Am18 [5B]	$(TGAA)_{19}$	[FAM] TGGTCCTAGCAGCTGTCTGT GGCTACATCTGCAACGACAA	87-119	28	8	0.929	0.790	0.809	0.76	JN604062
Am19 [5C]	$(TCCA)_{12}$	[HEX] CTGTAATGAATCCAAGGAGCTG TGGATAATGAAGAAATGGATGG	102-146^	27	13	0.704	0.853	0.346	0.837	JN604063
Am21 [5B]	$(TTCTA)_{11}$	[FAM] TCTCGTCTGGTGTTGACTGC CAATGGCTTTACTTTTCTCTGC	95-140	27	9	0.778	0.800	0.446	0.774	JN604064
Am22 [6C]	$(TGGGTC)_6$	[HEX] GCCGAATATGCCGTACAACT TATCTTCAGACCCACCTGGC	110-155	27	10	0.889	0.843	0.502	0.824	JN604065

Am24 [3B]	$(TCAGGA)_6$	[FAM] CTGCTGGATCAGGGTTAGGA ACCATGCCCAGGTACTGTCT	141-165	28	4	0.571	0.554	0.192	0.475	JN604066

[1A,B,C] *A. mccullochi* Multiplex corresponding to group 1 with three loci (A, B and C)

Fifteen of the 17 loci were in HWE (not Am11 or Am19) and few locus pairs (x out of y) with significant linkage disequilibrium were detected after FDR correction (Benjamini & Hochberg 1995): Am1 and Am5, Am1 and Am24, Am11 and Am14 and, Am12 and Am 14 (Table A1.1). In addition, null alleles were suggested for only two loci (Am11 and Am19) due to homozygote excess. The *A. mccullochi* markers displayed a high allelic richness (mean N_A = 8.35 ± 0.73, range 2-13) and high levels of expected heterozygosity (H_E = 0.773 ± 0.027, range 0.543-0.907). The probability of identity (PID) or the polymorphic information content (PIC) for the combined microsatellite loci was 0.724 indicating a high level of discrimination between individuals making them useful for studies of connectivity and population genetic structure in this taxon.

Loci were cross-tested on fourteen different individuals of *A. akindynos* from Lizard Island (Australia). All loci amplified successfully and were polymorphic with 4-11 alleles per locus (Table A1.2). The newly developed primers reported here will provide a useful tool to examine gene flow, population genetic structure and genetic diversity in the endemic *A. mccullochi* and potentially other related anemonefish species.

Table A1.2: Cross species amplification potential for the novel *Amphiprion mccullochi* microsatellite loci on the test species *A. akindynos*, number of successful amplifications, amplicon size range and number of alleles (N_a)

Locus	Amplification	Size range (bp)	N_a
Am4	1	118-166	10
Am5	1	71-113	11
Am6	6/7	262-274	4
Am7	4/5	86-134	10
Am9	13/14	151-223	15
Am10	13/14	286-298	4
Am11	1	116-148	6
Am12	13/14	166-222	11
Am14	13/14	162-198	7
Am15	1	100-144	11
Am17	13/14	100-144	10
Am18	13/14	108-141	7
Am19	13/14	99-139	11
Am21	1	94-142	11
Am22	1	90-160	11
Am24	13/14	115-180	4

APPENDIX 2: **Identification of twenty one microsatellite loci for conservation genetic studies of the endemic butterflyfish *Chaetodon tricinctus***

(Published as: van der Meer MH, Gardner MG, Hobbs J-PA, Pratchett MS, van Herwerden L (2012c) Identification of twenty one microsatellite loci for conservation genetic studies of the endemic butterflyfish *Chaetodon tricinctus*. *Conservation Genetic Resources*, 4, 243−246.)

A2.1. ABSTRACT

Coral reef habitats are increasingly under threat from global warming and this has influenced the associated reef fish communities. *Chaetodon tricinctus* is a butterflyfish endemic to the offshore reefs of Elizabeth Reef, Middleton Reef, Lord Howe Island and Norfolk Island off Australia's east coast. It is a highly specialised coral reef fish that is thought to rely on Acroporid coral species exclusively for food and shelter. I developed primers for twenty one microsatellite loci to reveal gene flow, population genetic structure and genetic diversity within and among these three reefs. Observed heterozygosities ranged from 0.185 to 0.964 and expected heterozygosities ranged from 0.230 to 0.889 in 30 individuals from Middleton Reef. When cross tested with *Chaetodon trifascialis*, a closely related species, there was poor amplification success and only a moderate level of polymorphism. Therefore, although these loci will be useful in *C. tricinctus*, it is unlikely that they can be used on other related butterflyfishes.

A2.2. INTRODUCTION

Coral reefs support an enormous diversity of species however, there is growing concern over the global decline in coral cover (Gardner et al. 2003, Hughes et al. 2003) and how this has negatively affected reef fish communities (Doherty et al. 2002, Spalding & Jarvis 2002). Reef fishes that specialise on one or a select few coral species for food or shelter appear to most at risk (Jones et al 2004).

The three-striped butterflyfish *Chaetodon tricinctus* appears to be one such habitat specialist. This species is endemic to Australian waters and inhabits coral rich areas of lagoon and seaward reefs at Elizabeth Reef (ER), Middleton Reef (MR), Lord Howe Island (LHI) and Norfolk Island (NI; Randall 1976). It is the most abundant butterflyfish in the region (Choat et al. 2006a, Hobbs & Feary 2007, Hobbs et al. 2009a). The feeding habits of this species are largely unknown, except that it feeds on polyps of hard corals (Kuiter 2002) and is closely associated with species of the scleractinian Acroporid genus (Hobbs et al. 2009). *Acropora* species are particularly susceptible to bleaching (Marshall & Baird 2000, Willis et al. 2004) and there is concern that local extinctions of *C. tricinctus* may occur (for evidence in other species see Graham et al. 2006) as the ability of this fish to switch diet to include other coral species or invertebrates is unclear (Hobbs et al. 2009).

Information on gene flow, population genetic structure and genetic diversity of *C. tricinctus* is lacking, however these data are vital for inferring if populations at the 4 different locations (ER, MR, LHI and NI) could replenish or re-stock each other following a major bleaching event. Additionally, genetic diversity data will help determine if these populations may or may not be resilient under future environmental conditions when oceans are expected to be warmer and

more acidic. This study describes the development of 21 polymorphic microsatellite markers and for *C. tricinctus* using 454 shotgun pyrosequencing on a Roche

A2.3. MATERIAL AND METHODS

As per Chapter 2.3.1. (PCR reactions).

A2.4. RESULTS AND DISCUSSION

The resulting sequences (totalling 113,794 reads with an average sequence length of 350 and total GC content of 42.53%) were screened for di, tri, tetra, penta, and hexanucleotide microsatellite loci with six or more repeats using the default settings of QDD1 (Meglecz et al. 2009) which identified 13,381 (11.76% of sequences) microsatellite loci. However, PCR primers were successfully designed for the best 1,664 (12.44%) of loci found. Of these, primers for the best of the best 24 loci were synthesised with a 5′ 17 base pair lambda tag (5-GGTGGCGACTCCTGGAG-3) to enable indirect fluorolabelling (Shimizu et al. 2002) and minimize costs. Loci were initially tested for amplification success and specificity in eight individuals using Type-it microsatellite PCR kit (Qiagen). Individual amplifications were performed in 10 µL reactions, containing 1x Type-it Multiplex PCR Master Mix (Qiagen, Doncaster/Australia), 20-50 ng template, and 0.2 µM each primer (forward and reverse). Indirectly labelled reactions contained a tailed forward primer and a reporter primer (5' labelled with fluorescent dye modification HEX, TET or FAM) at a 1:4 ratio (total = 0.2 µM). All primers were tested and optimised using a Bio-Rad C1000 Thermal Cycler (see Table 1) with an initial denaturation of 94°C for 3 mins followed by 28 cycles of 94°C for 40 s, 58°C for 40 s and 72°C for 40s followed by 5 mins at 72°C in a Bio-Rad C1000 Thermal Cycler (Bio-Rad, Australia). PCR products were column purified using an Ammonium Acetate protocol. Genotypes were run on an Amersham Bioscinces Megabase Capillary Sequencer with a 400 bp size standard and scored using Fragment Profiler 1,2 (Amersham, Buckinghamshire, England).

Primer pairs for 21 loci reliably amplified products of the expected size, with no additional products and were polymorphic, representing one hexamer, four pentamers and sixteen

tetramers. Directly-labelled forward primers (HEX, TET or FAM) were synthesised for the 21 polymorphic *C. tricinctus* loci to allow PCR multiplexing of at least three loci (Table 1). Loci were genotyped in directly labelled multiplex reactions in 30 *C. tricinctus* individuals from MR. DNA extraction protocol and PCR conditions as described above. Multiplex PCR combinations (Table 1) were designed and tested using PCR conditions described above.

Characteristics of the 21 loci are summarised in Table A2.1. GENALEX 6 (Peakall & Smouse 2006) was used to examine the number of alleles, observed and expected heterozygosities and conformation to Hardy-Weinberg Equilibrium (HWE). CERVUS 3.0 (Kalinowski et al. 2007) was used to calculate polymorphic information content (PIC) for each locus, GENEPOP 4.0.10 (Rousset 2008) and MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) were used to test linkage disequilibrium and the presence of null-alleles, respectively.

Table A2.1: Details for twenty one *Chaetodon tricinctus* microsatellite loci developed from 454 shotgun sequences. T_A, annealing temperature; N, sample size; N_A, number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; p HWE, Hardy-Weinberg equilibrium significance value at $p < 0.05$ after FDR correction; PIC, polymorphic information content. ^, locus may have null alleles.

Locus	Repeat motif	Primer sequence	Size range (bp)	N	N_a	H_o	H_e	pHWE	PIC	Genbank accession no.
Chaetodon tricinctus										
Ct2 [1C]	$(ACAG)_9$	[HEX]TGTCTTTGGATGTGAGCAGG ACCACAGGCCAGTTTTTCAC	237-245	30	3	0.300	0.383	0.544	0.321	JN604067
Ct3 [3C]	$(ATCT)_{11}$	[HEX]CCAAAGCTTGGCTTTTGTGT CTAAGTCTTGGCGAACAGCC	227-255	24	5	0.417	0.444	0.821	0.419	JN604068
Ct4 [1B]	$(CTAT)_{12}$	[FAM]AGCTGAGTTCAGGTCAGGGA GCATAGGACTCGGGAACAAG	130-158	30	7	0.633	0.618	0.047	0.569	JN604069
Ct5 [3B]	$(ATAG)_{11}$	[FAM]CTGACACAGCCAAACGATGT TACGGTCAGGTTCGGCTTAC	191-207	24	5	0.583	0.694	0.446	0.64	JN604070
Ct7 [7C]	$(TCTA)_{10}$	[HEX]CATTCATTCCTGCTGCCATA GGCCAAGTAACGTTTGCTGT	56-156	28	7	0.964	0.654	0.280	0.61	JN604071
Ct8 [1A]	$(TCCA)_{10}$	[TET]GCACTTATGTGCTCCATCCA GACCTCCAATCTGGCAACTC	102-140	30	7	0.733	0.654	0.950	0.591	JN604072
Ct9 [5C]	$(GATA)_{13}$	[HEX]TGGACCAGCACACTCAGAAG ACAGATGTGAATTCCCGTGG	203-223	27	6	0.741	0.628	0.993	0.587	JN604073
Ct10 [4C]	$(GATA)_{10}$	[HEX]TCGCAAGCTTGGATTATGAA GCCTCAGCAGTGCAAGTACA	191-203	29	4	0.586	0.499	0.532	0.453	JN604074
Ct11 [2A]	$(ATAG)_{10}$	[TET]TGACCAAAATGACAACCAGG TACTTTTATCGGCGGCAAAG	95-139	30	9	0.833	0.767	0.983	0.734	JN604075
Ct12 [4A]	$(TCAA)_{13}$	[TET]GTTTGTCAGCCGACTCACTG AAACAGTGTCTGGAGGCTGC	107-119	30	4	0.533	0.563	0.307	0.505	JN604076
Ct13 [3A]	$(GATA)_{13}$	[TET]ACAAGACAGCGAATGAGTGC AGGATCTCAGCCACGAAAGA	117-145	24	8	0.667	0.778	0.945	0.747	JN604077
Ct14 [6B]	$(TAGA)_9$	[FAM]GACATGACCACCAAGCAGC AGGGTGGGAGGCATAAAAGT	113-137	28	7	0.750	0.736	0.254	0.699	JN604078
Ct15 [5B]	$(ATAG)_{12}$	[FAM]TTTGTGCAACATGACAATGAA AATGATGGCTTTGCTCTGCT	120-164	27	9	0.926	0.848	0.509	0.83	JN604079

Ct16 [4B]	$(ATAG)_{13}$	[FAM]GTTAATGATGGCGTTGGAGC AACACTAAATTTCCCGGATCA	107-147	30	7	0.567	0.553	**0.000**	0.501	JN604080
Ct17 [7A]	$(TAGA)_{12}$	[TET]GCAAGATCTGGCAAGAAACC GGCTCACTACATTATGCACTGGT	108-136	28	7	0.643	0.719	0.728	0.682	JN604081
Ct18 [6A]	$(GTTT)_{9}$	[TET]TGTAAGGCTTGTAGTCAAGTCAGG GGAAAGCCTTTAGGGGACAG	89-101^	28	4	0.214	0.339	**0.004**	0.318	JN604082
Ct20 [2C]	$(AGGAG)_{9}$	[HEX]AGAGGCAGTCCTGGAAAATG TGTTTGGATGCTCAGGTCAG	150-300	29	9	0.793	0.822	0.860	0.799	JN604083
Ct21 [5B]	$(TTGTG)_{16}$	[FAM]GGCCCACAACTTCTCATTGT TACTCCACAGGGTTGAATGC	105-175	28	13	0.857	0.889	0.006	0.879	JN604084
Ct22 [2B]	$(TATCT)_{10}$	[FAM]CCTTCTGTGCCACTGGTTTT AGAACATGAGGACATCAGAGCA	140-180	30	8	0.833	0.846	0.438	0.828	JN604085
Ct23 [6C]	$(TTCAA)_{10}$	[HEX]CAAACCAACTTTTCTGTGAAGTG CCCCTCGTGAGACGAATAAA	135-150	28	4	0.929	0.675	**0.002**	0.618	JN604086
Ct24 [5A]	$(TAGATT)_{6}$	[TET]ATTTTAGGCTGTGGCCTGTG ACCAGCTGTCCTTGCAGAAG	90-102	27	3	0.185	0.230	0.621	0.211	JN604087

Nineteen of the 21 loci were in HWE (not Ct16 and Ct23) and few locus pairs (x out of y) with significant linkage disequilibrium were detected after FDR correction (Benjamini & Hochberg 1995): Ct2 and Ct21, Ct3 and Ct7, Ct5 and Ct18, Ct5 and Ct23, Ct7 and Ct9, Ct10 and Ct13, Ct14 and Ct23 and, Ct18 and Ct21 (Table 1). In addition, null alleles were suggested for only one loci (Ct18) due to homozygote excess. The *C. tricinctus* markers displayed high allelic richness (mean N_A = 6.5 ± 0.54, range 3-13) and high levels of expected heterozygosity (H_E = 0.635 ± 0.039, range 0.230-0.889). The polymorphic information content (PIC) for the combined microsatellite loci was 0.616 indicating a reasonable discrimination between individuals, making the loci useful for studies of connectivity and population genetic structure in this taxon.

All 21 loci were cross-tested on sixteen individuals of *Chaetodon trifascialis* (Heron Island, Australia). Fifteen loci amplified successfully and ten were polymorphic with 2–11 alleles per locus (Table A2.2). The newly developed primers reported here will provide a useful tool to examine the gene flow, population genetic structure and genetic diversity in *C. tricinctus* but may require optimization in other related butterflyfish species.

Table A2.2: Cross species amplification potential for the novel *Chaetodon tricinctus* microsatellite loci on the test species *C. trifascialis*, number of successful amplifications, amplicon size range and number of alleles (N_a)

Locus	**Amplification**	**Size range (bp)**	**N_a**
Ct2	7/8	221-265	10
Ct4	3/4	132-160	8
Ct5	7/8	179-279	11
Ct7	5/16	52	1
Ct8	15/16	132-160	8
Ct12	1	139-171	6
Ct13	1/16	162	1
Ct14	1/16	121	1
Ct15	1/8	180-189	2
Ct16	9/16	139-181	6
Ct18	15/16	67-71	2
Ct20	15/16	260-305	9
Ct21	1/16	95	1
Ct22	5/16	135-155	2
Ct24	15/16	78	1

APPENDIX 3: **Identification of seventeen microsatellite loci for conservation genetic studies of the endemic wrasse *Coris bulbifrons***

(Published as: van der Meer MH, Gardner MG, Berumen M, Hobbs J-PA, van Herwerden L (2012) Identification of seventeen microsatellite loci for conservation genetic studies of the endemic wrasse *Coris bulbifrons*. *Conservation Genetic Resources,* 4, 243−246.)

A3.1. ABSTRACT

Coral reefs around the world are in decline, in part due to various anthropogenic factors, including fishing pressure. *Coris bulbifrons* is a large wrasse endemic to only four oceanic locations off Australia's east coast: Middleton Reef, Elizabeth Reef, Lord Howe Island and Norfolk Island. The species is listed as vulnerable by the IUCN due to the potential threat of overfishing. Although these remote locations, some within Marine Protected Areas, experience limited fishing pressure, populations may quickly decline with minimal fishing effort as seen in the overfishing of other large wrasses. I developed primers for 17 microsatellite loci to examine gene flow, population genetic structure, and genetic diversity within and among these four locations. Observed heterozygosities ranged 0.126 – 0.752 in 37 individuals from Lord Howe Island indicating that these loci will be useful in *Coris bulbifrons* population genetic studies.

A3.2. INTRODUCTION

Coral reefs around the world are in decline from factors such as coastal development, pollution, and global climate change; however, fisheries are responsible for the most direct and wide-ranging influence on these and other marine ecosystems (Jennings & Kasier 1998, Jackson et al. 2001). Some remote locations and islands are rare examples of coral reefs with limited fishing pressures (Friedlander & DeMartini 2002). These same locations are also hotspots for coral reef endemicity (Jones et al. 2002, Roberts et al. 2002). Terrestrial endemics on remote islands are well known for their vulnerability to overexploitation (Whittaker 1998); however, much less is known of the vulnerability of their marine counterparts.

The doubleheader wrasse, *Coris bulbifrons*, is endemic to four locations in the South-West Pacific (approximately 600km off Australia's east coast). Three locations are situated along the Lord Howe Rise: Middleton Reef (MR), Elizabeth Reef (ER), Lord Howe Island (LHI). *Coris bulbifrons* is abundant in sheltered habitats at these three locations (Choat et al. 2006a, Hobbs & Feary 2007, Hobbs et al. 2009) and Marine Protected Areas (MPAs) have been established to promote population persistence. The fourth location (Norfolk Island - NI) is situated 600km to the east of the Lord Howe Rise and this NIR peripheral population of *C. bulbifrons* has the lowest densities of all four locations (unpublished data) and is not protected by any MPAs. Although *C. bulbifrons* is abundant at three of the four locations, these populations may quickly decline with minimal fishing effort as seen in the overfishing of other large wrasses (Choat et al. 2006b). For this reason *C. bulbifrons* is listed as vulnerable by the IUCN (Choat & Pollard 2010).

Information on gene flow, population genetic structure and genetic diversity of *C. bulbifrons* is lacking. These data are vital for inferring if populations at the four different locations (ER, MR, LHI, and NI) could replenish or re-stock each other following local extinction, and for determining population resilience under future environmental conditions.

This study describes the development of 17 polymorphic microsatellite markers for *C. bulbifrons* using 454 shotgun pyrosequencing on a Roche GS-FLX (Australain Genome Research Facility, AGRF, Brisbane, Australia).

A3.2. MATERIAL AND METHODS

As per Chapter 2.3.1 (PCR reactions)

A3.4. RESULTS AND DISCUSSION

The resulting sequences (totalling 103,719 reads) were screened for di, tri, tetra, penta, and hexanucleotide microsatellite loci with six or more repeats using the default settings of QDD1 (Meglecz et al. 2009). This process identified 8,878 microsatellite loci (within 8.56% of sequences obtained); PCR primers were successfully designed for 1,110 (1.07%) of loci found. Of these, directly-labelled forward primers (FAM, NED or VIC) were synthesised for 24 loci deemed the best candidates for performance and polymorphism. Loci were initially tested for amplification success and specificity in eight individuals using a Type-it microsatellite PCR kit (Qiagen, Doncaster, Australia). Individual amplifications were performed in 10 µL reactions, containing Type-it Multiplex PCR Master Mix, 20 – 50 ng DNA template, and 0.2 µM each primer (forward and reverse). All primers were tested and optimised using a C1000 Thermal Cycler, Bio-Rad, Australia (see Table 1) with an initial denaturation of 94˚C for 3 min followed by 28 cycles of 94˚C for 40 s, 58˚C for 40 s and 72˚C for 40 s followed by 30 min at 72˚C. PCR products were column purified using an Ammonium Acetate protocol. Genotypes were run on a 3730XL DNA Analyzer (Applied Biosystems, Saudi Arabia) at the King Abdullah University of Science and Technology (KAUST, Saudi Arabia) with a 550 bp size standard and scored using GeneMarker (SoftGenetics, USA). Primer pairs for 17 loci reliably amplified products of the expected size, with no additional products and were polymorphic, representing four dimer, two trimer, six tetramer, three pentamer and two hexamer simple sequence repeat (SSR) loci. Loci were pooled for PCR multiplex reactions consisting of at least three loci per multiplex (Table 1), using the same optimised conditions as above, in 37 *C. bulbifrons* individuals from LHI.

Characteristics of the 17 loci are summarised in Table A3.1. GENALEX 6 (Peakall & Smouse 2006) was used to examine the number of alleles, observed and expected heterozygosities and conformation to Hardy–Weinberg Equilibrium (HWE). CERVUS 3.0 (Kalinowski et al. 2007) was used to calculate polymorphic information content (PIC) for each locus. GENEPOP 4.0.10 (Rousset 2008) and MICROCHECKER 2.2.3 (Van Oosterhout et al. 2004) were used to test linkage disequilibrium and the presence of null-alleles, respectively.

Table A3.1: Details for seventeen *Coris bulbifrons* microsatellite loci developed from 454 shotgun sequences. T_A, annealing temperature; N, sample size; N_A, number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; p HWE, Hardy-Weinberg equilibrium significance value at $p < 0.05$ after FDR correction; PIC, polymorphic information content. ^, locus may have null alleles.

Locus	Repeat motif	Primer sequence	Size range (bp)	N	N_a	H_o	H_e	pHWE	PIC	Genbank accession no.
Coris bulbifrons										
Cb1 [1A]	$(AATCAG)_7$	[FAM]TCCAAAGAAGCTGGGGTTATT ACTTGGCAGATAAAGGCGAT	73-103	36	7	0.639	0.668	0.985	0.619	JX952210
Cb2 [2A]	$(AGGGTT)_5$	[FAM]GGGTTAAAATGGGACAAGGG GGTTTGGGTTAGGGTTAGGG	87-95	35	2	0.286	0.245	0.477	0.215	JX952211
Cb4 [1B]	$(GGAGA)_9$	[FAM]ATTTTGCAAAGTGTGGTCCC CTCTTCTGTTGCTCCCTGCT	200-210	34	3	0.118	0.164	**0.000**	0.157	JX952212
Cb5 [4A]	$(CTCAG)_8$	[FAM]ACAGCCCAGAGAAAACAGGA GTTGTGGAAAAACACCCACC	160-175	34	4	0.412	0.529	**0.000**	0.429	JX952213
Cb6 [1C]	$(AGGAG)_5$	[NED]GAGGAGAGGAGAGGAGGAGG GCTGATTGAACGGACCAGAT	78-118^	26	8	0.308	0.752	**0.000**	0.721	JX952214
Cb7 [4B]	$(CTAT)_{11}$	[NED]TAAGAGGTGTGCTGCGTTTG CTGACAGGGCAGCATTTGTA	150-182	29	5	0.448	0.487	**0.002**	0.431	JX952215
Cb8 [2B]	$(TGTC)_{10}$	[NED]TCATTTCCTTCCCCTGTCTG TGAGACTAAAGCAGCGAGCA	124-140	35	5	0.486	0.442	0.477	0.413	JX952216
C10 [3A]	$(AACA)_6$	[FAM]GAGCTGAAGGAAACGCAAAC AGCCTCTGGGAACATGAGAA	132-158	33	6	0.606	0.679	**0.000**	0.623	JX952217
Cb11 [2C]	$(ATGA)_6$	[VIC]GCCTCAGAGAAACAATTGGC CCTCCATCCTCTTTCATCCA	109-117	36	3	0.250	0.335	0.477	0.288	JX952218
Cb12 [1D]	$(AGGG)_6$	[VIC]GGATGGAGAAAAGAGGGGAG CAAGAGGTGTGAGCGACAAA	87-95	36	2	0.139	0.219	0.053	0.195	JX952219
Cb13 [4C]	$(AGG)_{13}$	[VIC]CTCAACGCATGAACTCCTGA TCCCCTGCTCGTCTAAGTTG	140-156^	34	5	0.382	0.505	**0.002**	0.470	JX952220
Cb15 [3B]	$(ATG)_{12}$	[NED]TGTAACAGCTTCATCAGGCG	125-134	32	4	0.406	0.436	0.160	0.409	JX952221

		GCTGACCCTCTCACACCATT								
Cb18 [1E]	$(AC)_{17}$	[NED]GGGCCATCAAAACACTCTGT	190-202	31	3	0.387	0.447	0.916	0.360	JX952222
		TGAGCAGAGTGGGGAGTTCT								
Cb23 [3C]	$(TG)_{10}$	[VIC]CCGTCACCCAAACTTTCACT	105-121	36	9	0.778	0.699	**0.000**	0.658	JX952223
		GGAACCTCCCTTCAAACACA								
Cb26 [5A]	$(ATTC)_7$	[FAM]CCTCTTCCTGCTTTTGGTGA	216-228	30	3	0.133	0.126	0.985	0.121	JX952224
		TTGGTATATGTGGCAGGCAA								
Cb30 [5B]	$(CA)_9$	[NED]TGTCTTGTCAGAGCCCACAG	184-188	34	3	0.118	0.138	**0.002**	0.132	JX952225
		ACACGCGTTGCTACAGACAC								
Cb36 [1F]	$(TC)_9$	[VIC]CATTCCTGAACCACAGCTCA	270-274	31	3	0.613	0.534	0.916	0.434	JX952226
		GTCTTCAAAGGTGACCGAGC								

Nine of the 17 loci were in HWE and no linkage disequilibrium between any locus pairs (zero out of 272) were detected after FDR correction (Benjamini & Hochberg 1995). In addition, null alleles were suggested for only two loci (Cb6 and Cb13) due to homozygote excess. The *C. bulbifrons* markers displayed medium-high allelic richness (mean N_a = 4.412 ± 2.063, range 2 – 9) and medium-high levels of expected heterozygosity (H_e = 0.436 ± 0.203, range 0.126 – 0.752). The polymorphic information content (PIC) for the combined microsatellite loci was 0.393 indicating moderate discrimination between individuals, making the loci useful for studies of connectivity and population genetic structure in this taxon. The newly developed primers reported here will provide a useful tool to examine the gene flow, population genetic structure and genetic diversity in *C. bulbifrons*.