2016

Gene flow and genetic structure of the seagrass Thalassia hemprichii in the Indo-Australian Archipelago

Udhi Eko Hernawan
Edith Cowan University

Edith Cowan University

Copyright Warning

Gene flow and genetic structure of the seagrass *Thalassia hemprichii* in the Indo-Australian Archipelago

Udhi Eko Hernawan

BSc (Life Sciences)

MSc (Marine Biology and Ecology)

This thesis is presented in fulfilment of the requirements for the degree of Doctor of Philosophy

School of Science

Edith Cowan University

October 2016

ABSTRACT

How genetic variation is distributed across space (genetic structure) and what factors influence the spatial genetic structuring is one of the primary questions in population genetics. The interaction between species biology (e.g. life-history traits) and physical processes operating in the seascape over time, including palaeo-historical events (e.g. sea level fluctuations) and contemporary processes (e.g. ocean currents), have been predicted to influence the extent of gene flow and the spatial genetic structuring in marine organisms. However, the relative contribution of each factor in governing the genetic pattern remains unclear. This study examined the pattern of genetic structure and the factors influencing this using multiple approaches across different temporal and spatial scales in the Indo-Australian Archipelago (IAA), the world's hotspot of marine biodiversity.

By comparing population genetic data of co-distributed marine species (e.g. fishes, molluscs, etc.), this study shows that for marine organisms, the interaction between species biological traits and the physical/environmental processes (habitat variability, water current, etc.) are the greatest drivers of genetic structure in the IAA. Since the physical/environmental processes fluctuate over time, spanning from hours to millennia, the temporal scale (palaeo-historical vs contemporary) at which physical/environmental processes generate genetic structure were examined using seascape genetic analysis. To minimise the effect of different biological traits, the seascape genetic analysis focused only on one species, *Thalassia hemprichii*, one of the dominant seagrass species in the IAA.

The analysis revealed that both palaeo-historical processes (vicariance due to Pleistocene sea level fluctuations) and more contemporary processes (ocean currents) strongly influence the pattern of genetic structure at a regional scale (>300 km). At this spatial scale, the influence of contemporary ocean currents is much smaller than that of historical vicariance. This finding contrasts with previous studies highlighting a strong effect of ocean currents in seagrass connectivity. Only when the effect of historical vicariance was minimised by spatially down-scaling the study from a

regional (>300 km) to local (<75 km) scale, contemporary processes, including ocean currents and habitat heterogeneity, were shown to strongly influence the pattern of genetic structure.

This study also revealed that significant genetic structure can occur at both regional and local scales. At the regional scale, the genetic clusters span distances of at least 500 km, suggesting that genetic connectivity of *T. hemprichii* populations occurs over very large geographic scales. At the local scale, significant spatial genetic structure was detected, negating the prediction of a single panmictic population. The strong genetic structuring occurring at both large and small spatial scales suggests that predicting seagrass connectivity solely based on geographic distance is inaccurate, and the relevant distance between populations in the marine system is not purely geographic, but rather determined by other factors operating on the seascape setting such as water currents and habitat heterogeneity. Thus, seascape setting is very important in seagrass gene flow and structure.

Based on the pattern of gene flow, genetic structure, and genetic diversity, this research provides recommendations for seagrass conservation management in the IAA, including spatial design of conservation reserves and restoration including transplantation.

DECLARATION PAGE

I certify that this thesis does not, to the best of my knowledge and belief:

i. incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution of higher education;

ii. contain any material previously published or written by another person except where due reference is made in the text of this thesis;

iii. contain any defamatory material; or

iv. contain data that have not been collected in a manner consistent with ethics approval.

Udhi Eko Hernawan

October 2016

ACKNOWLEDGMENT

The thesis would never have been accomplished without the help and support of many people. Foremost, I am very thankful to Kathryn McMahon for all the guidance, help and support. She has inspired me to tackle difficulties and kept me focused on the goal during this PhD journey. I also thank Gary Kendrick for his help, suggestions and enthusiasm, which helped strengthen my confidence in completing this PhD. My gratitude goes to Kor-jent van Dijk, who has dedicated his time and expertise to guide me in completing all genetic lab work and analyses. I also thank Paul Lavery, his insightful comments and suggestions helped me to improve my thesis.

I acknowledge the significant contribution from Ming Feng (CSIRO, hydrodynamic simulation), Oliver Berry (CSIRO, genetic analysis), Edward Biffin (The University of Adelaide, single nucleotide polymorphism/SNP screening), and Christopher Kavazos (seascape genetics). Others also contributed in many ways for field assistance, collecting samples, lab work or comments and suggestions: L. Ruiz-Montoya, J.P Hobbs, A.Z. Perez, F. Webster, A. Isaac, S. Isaac, T. Stumpagee, M. Waycott (the State Herbarium of South Australia and the University of Adelaide), Flavia Tarquinio, Maryam Abdolahpour, Rachele Bernasconi, Anna Lafratta and all other CMER members, Abigail Moore (STPL Palu), Bahtiar & K. Togaishamu (Unhalu Kendari), Wahyu Adi & Nur Annis (UBB Bangka), Try Al Tanto (LPSDKP Padang), Raismin Kotta, Yulius Suni, Ucu Y. Arbi, Susi Rahmawati, Supono, Alan, Alvi Sitepu, Ahmad Ainarwowan, Aliyadi, and Kadir Yamko. I also thank the Bardi Jawi community, rangers and the traditional owners for access and field assistance in the Kimberley.

My family has always been very supportive of my choices and ambitions. A special thank goes to my wife, Miyati, for her support, patience and care during the many months spent away from home and the kids.

I would like to thank to the South Australian Regional Facility for Molecular Evolution and Ecology (SARFMEE) for enabling genetic lab work; and to the Research Centre for Oceanography (P2O-LIPI), UPT LIPI Tual, UPT LIPI Biak, UPT LIPI Mataram, P2LD LIPI Ambon, UPT LIPI Pulau Pari, UPT LIPI Bitung, UPT LIPI Bitung, STPL Palu, LPSDKP Padang, Universitas Bangka Belitung and Universitas Haluoleo, Indonesia for providing supports and assistance during field work in Indonesia.

Finally, I am very grateful to the financial support of ECU-IPRS; the G100379 project of the Department of Education and Training, Collaborative Research Network Program (Funding Agreement CRN2011:5, Edith Cowan University and University of Western Australia); the Western Australian Marine Science Institution (WAMSI Kimberley Research Program: Project 1.1.3 Ecological Connectivity (KM) and 2.2.4 Primary Producers (GAK)); and the Indonesia Endowment Fund for Education (LPDP-Indonesia).

TABLE OF CONTENTS

ABSTRACT iii
DECLARATION PAGE v
ACKNOWLEDGMENT vi
TABLE OF CONTENTS viii
List of Figures xi
List of Tables xiii
List of Publications xiv
List of Conference Presentations xv
Chapter 1 1
GENERAL INTRODUCTION 1
1.1 Population connectivity 1
1.1.1 Connectivity, gene flow, and genetic structure 1
1.1.2. Connectivity and seascape 3
1.2 The Indo-Australian Archipelago 4
1.2.1 Hotspot of biodiversity 5
1.2.2 Geological history 5
1.2.3 Connectivity and genetic structure in the IAA 6
1.3 Seagrass importance and conservation 8
1.3.1 Significance of seagrass 8
1.3.2 Significance of genetics and connectivity for seagrass conservation 8
1.4 *Thalassia hemprichii* 9
1.5. Summary and aim of the study 11
1.5.1. The overarching aim 12
Chapter 2 15
PREDICTORS OF MARINE GENETIC STRUCTURE 15
2.1 Introduction 16
2.2 Materials and Methods 19

2.2.1 Literature survey 19
2.2.2 Data extraction 19
2.2.3 Statistical analysis 21
2.3 Results 23
2.3.1 Literature survey 23
2.3.2 Predictors of genetic structure 23
2.3.3 Influence of the important predictors 29
2.4 Discussion 30
2.4.1 The influence of dispersal-related traits 30
2.4.2 Influence of genetic markers on F_{ST} 31
2.4.3 The importance of ecoregion 32
2.4.4 Implications for marine conservation 33
Chapter 3 36
HISTORICAL AND CONTEMPORARY GENETIC STRUCTURE OF THE SEAGRASS *THALASSIA HEMPRICHII* 36
3.1 Introduction 37
3.2 Materials and Methods 40
3.2.1 *Thalassia hemprichii* 40
3.2.2 Study sites and sampling design 40
3.2.3 SNP screening 41
3.2.4 Microsatellite amplification 42
3.2.5 Phylogeography using SNP 42
3.2.6 Genetic structure using microsatellite 43
3.2.7 Oceanographic connectivity 44
3.2.8 Relative roles of palaeo-historical processes and contemporary oceanographic connectivity in gene flow 45
3.3 Results 46
3.3.1 Phylogeography using SNP 46
3.3.2 Genetic structure using microsatellites 47
3.3.3 The role of historical processes and oceanography 51
3.4. Discussion 52
3.4.1 Phylogeography 52
3.4.2 Genetic structure at more recent divergence time 54

3.4.3 Genetic diversity 56
Chapter 4 59
CONTEMPORARY GENETIC STRUCTURE OF THE SEAGRASS *THALASSIA HEMPRICHII* AT A LOCAL SCALE 59
4.1 Introduction 60
4.2. Materials and Methods 63
4.2.1 Study site and sampling 63
4.2.2 Microsatellite amplification and genotyping 64
4.2.3 Genetic analysis 65
4.2.4 Seascape genetic analysis 66
4.3 Results 69
4.3.1 Microsatellite data properties 69
4.3.2 Genetic differentiation and connectivity 71
4.3.3 The drivers of genetic differentiation 73
4.4 Discussion 74
4.4.1 Genetic differentiation and connectivity 74
4.4.2 What is driving genetic differentiation? 75
Chapter 5 79
GENERAL DISCUSSION 79
5.1 Factors influencing genetic structure 80
5. 2 Seagrass connectivity in the IAA: Scaling from regional to local 83
5.3 Management implications 85
5.3.1 Spatial design of conservation units 85
5.3.2 Seagrass restoration 88
5.4 Directions for future research 90
REFERENCES 93
APPENDIX 116

List of Figures

Figure 1.1 Connectivity and genetic differentiation 2

Figure 1.2 Map showing the Indo-Australian Archipelago with major ocean currents across the region (black arrows) 6

Figure 1.3 General morphology and meadows of *Thalassia hemprichii*..... 10

Figure 1.4 The multiple approaches used to understand genetic structure and the factors driving it for marine organisms and, in greater detail, for *Thalassia hemprichii* 14

Figure 2.1 Marine ecoregions in the Indo-Australian Archipelago based on Spalding *et al*. (2007) 18

Figure 2.2 Relative evidence weight of predictors generated for genetic cluster and F_{ST} using full dataset, species with PLD>0, sessile and sedentary species, and mobile species 28

Figure 3.1 Map showing the seagrass *Thalassia hemprichii* sampling sites throughout the Indo-Australian Archipelago. 39

Figure 3.2 Bootstrap consensus tree constructed from SNP dataset using MEGA7 (Kumar *et al.* 2016) and estimates of migration calculated from MIGRATE-N (Beerli & Felsenstein 2001)..... 47

Figure 3.3 Genetic diversity of the seagrass *Thalassia hemprichii* in the Indo-Australian Archipelago calculated from microsatellite data 48

Figure 3.4 Bayesian assignment of individuals calculated in STRUCTURE and migration estimates calculated in MIGRATE 51

Figure 4.1 Map of the sampling sites and the pattern of gene flow based 63

Figure 4.2 Pairwise of F_{ST} between populations of *Thalassia hemprichii* in the Kimberley 71

Figure 4.3 Cluster of populations resulted from STRUCTURE analysis (Pritchard *et al.* 2000) 73

Figure 5.1 Seagrass management units suggested by the connectivity patterns 87

Figure 5.2 Proposed protected-seagrass sites in the Kimberley, Western Australia based on the genetic patterns 88

List of Tables

Table 2.1 Criteria and variables extracted from peer-reviewed studies 21

Table 2.2 Best models generated for F_{ST} using full and restricted dataset. Only models within the lowest two AIC units are shown in table 25

Table 2.3 Best models generated for genetic cluster using full and restricted dataset. Only models within the lowest two AIC units are shown in table 26

Table 2.4 Multiple comparisons of means on the influence of the important categorical variables 29

Table 3.1 Allelic richness and genetic diversity of *Thalassia hemprichii* in the Indo-Australian Archipelago obtained from 16 microsatellite loci standardized at 24 MLGs 49

Table 4.1 Genetic diversity of *Thalassia hemprichii* in the Kimberley obtained from 16 microsatellite loci, standardized at 14 MLGs 70

Table 4.2. Variation partitioning on genetic differentiation (F_{ST}) into components accounted for the explanatory variables 74

List of Publications

At the point of submitting this thesis, Chapter 2, 3 and 4 have been submitted for the following publications:

1. **Hernawan UE**, McMahon K, Kendrick GA, van Dijk KJ, Lavery PS. Predictors of marine genetic structure in the Indo-Australian Archipelago. In review: *Journal of Biogeography* (revisions submitted March 2016)

2. **Hernawan UE**, McMahon K, van Dijk KJ, Kendrick GA, Feng M, Biffin E, Lavery PS**.** Historical processes and contemporary ocean currents drive genetic structure in the seagrass *Thalassia hemprichii* in the Indo-Australian Archipelago. In review: *Molecular Ecology* (submitted May 2016)

3. **Hernawan UE**, McMahon K, van Dijk KJ, Kendrick GA, Feng M, Berry O, Kavazos C. Extreme ocean currents and habitat characteristics drive genetic divergence in a tropical seagrass. In review: *Molecular Ecology* (submitted March 2016)

In addition, this thesis have contributed to the following publications:

1. Book chapter:
   McMahon K, Sinclair E, Sherman C, van Dijk KJ, **Hernawan U**, Verduin J, Waycott M. Genetic connectivity in tropical and temperate Australian seagrass species. *In:* Seagrass of Australia (Chapter 8). In review (March 2016).

2. Reports
   McMahon, K, Lavery, P, Statton, J, McCallum, R and **Hernawan, U** (2015). Current state of knowledge regarding the effects of dredging-related 'pressure' on seagrasses Final Report of Theme 5 Project 1a of the Dredging Science Node of the Western Australian Marine Science Institution, WAMSI, Perth, Western Australia. 42 pp.

   McMahon K, **Hernawan U**, van Dijk K-J, Waycott M, Biffin E, Evans R and Lavery P (2015). Genetic Variability of Seagrass in NW Australia. Final Report for Project 5.2 of the WAMSI Dredging Science Node of the Western Australian Marine Science Institution, WAMSI, Perth, Western Australia. 35 pp.

   McMahon K, **Hernawan U**, van Dijk K-J, Waycott M, and Dawkins K (2016). Ecological connectivity of Seagrass in the Kimberley, Australia. Final Report for Project 1.1.3 of the WAMSI Kimberley Science Node of the Western Australian Marine Science Institution, WAMSI, Perth, Western Australia. 40 pp.

List of Conference Presentations

Hernawan UE, McMahon K, Kendrick GA, van Dijk KJ, Lavery PS. Predictors of marine genetic structure in the Indo-Australian Archipelago. ASLO Aquatic Science Meeting, 22-27 February 2015, Granada Spain. Oral Presentation.

Hernawan UE, McMahon K, Kendrick GA, van Dijk KJ, Lavery PS, Berry O, Travers M, Underwood, J. Genetic connectivity of the seagrass *Thalassia hemprichii* in the Kimberley and Pilbara, Western Australia. AMSA (Australian Marine Science Association) Conference: Estuaries to Oceans. 5-9 July 2015, Geelong, Australia. Oral Presentation.

Hernawan UE, McMahon K, Kendrick GA, van Dijk KJ, Lavery PS. Predictors of genetic structure in marine organisms in the Indo-Australian Archipelago: Generalisable patterns and a seagrass-case study. ECU Research Week- CMER Symposium: Coastal Connectivity Implications for Marine Management. 14-18 September 2015, Joondalup, Australia. Oral Presentation.

Hernawan UE, McMahon K, Kendrick GA, van Dijk KJ, Lavery PS. So near, yet so far: Genetic connectivity of the seagrass *Thalassia hemprichii* in tropical Australia. Coastal and Estuarine Research Federation (CERF) 2015: Grand Challenges in Coastal and Estuarine Science. 8-12 November 2015, Portland Oregon, USA. Oral Presentation.

McMahon K, **Hernawan U**, van Dijk K-J, and Waycott M. Molecular ecology of seagrasses: tools for conservation and management. International Seminar on Natural Resources Biotechnology (Universitas Atma Jaya Yogyakarta). 8-9 September 2015, Yogyakarta, Indonesia. Oral Presentation.

Chapter 1

GENERAL INTRODUCTION

1.1 Population connectivity

1.1.1 Connectivity, gene flow, and genetic structure

Most populations are spatially separated from one another by unsuitable habitat, but are interconnected to a lesser or greater extent by movement of individuals. Under the island model (Wright 1951), the immigrants settle and reproduce in the recipient habitat, facilitating gene flow from the natal populations to the recipient populations. The degree of gene flow among populations that affects evolutionary processes within a population is described as genetic connectivity (Lowe & Allendorf 2010). The level of gene flow, together with mutation, selection and genetic drift, can influence the spatial distribution of genetic variation within and among populations (genetic structure). High gene flow homogenizes genetic variation by counteracting the effect of mutation, selection and genetic drift, while low gene flow leads to genetic differentiation (Wright 1951; Slatkin 1987; Charlesworth *et al.* 2003).

From an evolutionary perspective, movement and reproduction of the immigrant in the recipient population has direct consequences on gene flow and genetic structure. A population is comprised of n subpopulations ($n \geq 2$) and mating is random and without any restriction within each subpopulation (Figure 1.1). In terms of connectivity, two contrasting scenarios can be identified. When all individuals have the potential to move freely to any place within the entire population, and thus are likely to mate with other members, the subpopulations are in panmixia (Figure 1.1A). As mating is random, there would be no genetic differentiation among the subpopulations.

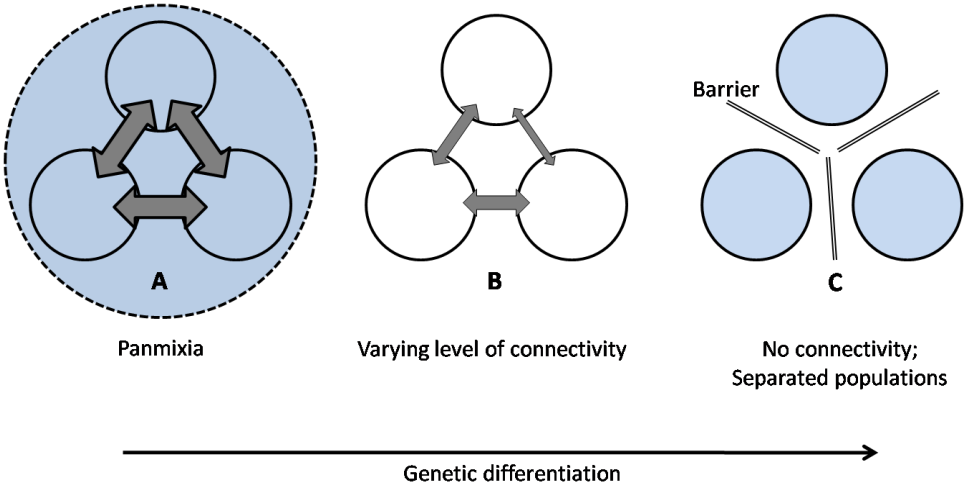

Figure 1.1 Connectivity and genetic differentiation. A small circle represents a subpopulation. A double arrow represents linkage among subpopulations (gene flow through migration). A group of subpopulations can be in panmixia (A) or in complete separation/independence (C). Figure (B) represents a varying degree of connectivity among *n* number of subpopulations. The pattern (level and direction) of gene flow depends on dispersal ability, life history, distance between subpopulations, and the nature of the intervening environment (e.g. water current, etc.).

The opposite scenario is that subpopulations are completely isolated (Figure 1.1C). In this situation, microevolutionary forces that operate within population (i.e. mutation and genetic drift) predominantly drive genetic differentiation (Figure 1.1C). It is only in the context of very long timescales, when at some time there was gene flow, that they might be considered a single population (Waples & Gaggiotti 2006). The intermediate situation lies between the contrasting scenarios (Figure 1B). This depiction likely represents the most common situation under natural circumstances where subpopulations are connected to a lesser or greater extent by gene flow, leading to a specific pattern of genetic structure.

The most common measure of genetic differentiation is F_{ST} as defined by Wright (1943), with values ranging from 0 (no differentiation/panmixia) to 1 (complete genetic isolation). There has been considerable debate about the accuracy of F_{ST} (e.g. Jost 2008; Whitlock 2011) however it is still recommended as a reliable measure of genetic differentiation (Meirmans & Hedrick 2011).

1.1.2. Connectivity and seascape

Connectivity and the resulting spatial genetic structuring in the marine environment is governed by the interactions of species biology (e.g. life-history traits, dispersal ability), physical processes (e.g. environmental variability, water currents) and biological processes (e.g. competition) operating in the seascape over time (e.g. López-Duarte *et al.* 2012). For a long time, the understanding of connectivity in the marine system has focused on early life stage dispersal (Shanks *et al.* 2003; Shanks 2009; Selkoe & Toonen 2011). Under this paradigm, movement of larvae or propagules is predicted to be the major driver of dispersal and gene flow among populations, particularly for sessile species, like corals and seagrass. However, the complexities of processes involved in realized dispersal challenged this paradigm. In some cases, the role of larval dispersal is often tempered by environmental factors, such as habitat specificity (e.g. Ayre *et al.* 2009).

Furthermore, physical processes in the seascape such as the direction and flow of water currents combined with island and shoreline configuration influence the extent of connectivity (Riginos & Liggins 2013). Even in a uniform, advective environment (e.g. a linear coastline with uniform water flow), gene flow may not be homogeneous and populations can be genetically structured due to environmental variability that drives local selection (Pringle & Wares 2007). For example, in the intertidal mussel *Perna perna*, different temperature regimes are responsible for driving strong genetic divergence despite high connectivity by water currents, suggesting environmentally-mediated selection (Zardi *et al.* 2011).

In addition, since the seascape attributes affecting gene flow fluctuate over time, spanning from hours to hundreds of thousands of years, the observed genetic pattern will reflect historical and/or contemporary connectivity, and one task of population geneticists is to tease apart these components (Riginos & Liggins 2013). For instance, strong genetic divergence in the estuarine seaweed *Fucus ceranoides*, despite a lack of the present-day gene flow barriers (e.g. habitat discontinuities or prominent ecological/oceanographic barriers), reflect historical vicariance rather than

the present-day connectivity (Neiva *et al.* 2012). In another case, genetic clustering in the intertidal mussel *Perna perna* best represents the present-day connectivity since the genetic pattern resulted from the specific responses of the genetic clusters to different temperature regimes in the present time (Zardi *et al.* 2011).

Due to the temporal component of gene flow, disentangling historical *vs* contemporary influences is necessary to better understand the complexities of processes influencing gene flow and genetic structure. For long-lived species like seagrass (Arnaud-Haond *et al.* 2012), the temporal component of genetic variation is important to consider when predicting gene flow against a changing seascape, because it takes multiple generations (can be up to 50 generation times) for more than one allelic change to accumulate in population (Epperson 2005). To address the time-space dependence of genetic variation, a number of approaches are available, such as employing genetic markers with different mutation rates, incorporating historical/contemporary data into the genetic analysis, and conducting the analysis (with the same or different markers) over different spatial scales from within a local population to among populations across different regions (Anderson *et al.* 2010; Epps & Keyghobadi 2015).

This thesis examined the pattern of genetic structure and the factors influencing this using multiple approaches across different temporal and spatial scales in the Indo-Australian Archipelago (IAA), the world's hotspot of marine biodiversity.

1.2 The Indo-Australian Archipelago

The Indo-Australian Archipelago (the IAA), also known as the Malay Archipelago, is a mosaic of island arcs in Southeast Asia, stretching between 95°E and 140°E across the equator (Figure 1.2). There are more than 17,000 islands, stretching over a distance of about 5,000 km, linking the Asian mainland with the Greater Australian Continent. The IAA connects the Pacific and the Indian Ocean through the Indonesian Throughflow (ITF). The ITF originates from the North Equatorial Current (NEC) of the Pacific Ocean, passes through the Indonesian Archipelago and

joins the South Equatorial Current (SEC) of the Indian Ocean. It is the only low-latitude link, transporting water from the Pacific Ocean into the Indian Ocean (Molcard *et al.* 2001; Sprintall 2009; Du & Qu 2010).

1.2.1 Hotspot of biodiversity

Situated between two of the major biogeographic provinces, the Pacific and the Indian Ocean, the IAA is home to megadiversity of both terrestrial and marine life (Myers *et al.* 2000) and is a hotspot of marine biodiversity (Renema *et al.* 2008). Although it occupies a very small portion of the planet's surface, more than 50% of the world's reef-building coral species can be found in the IAA, particularly in the heart of this zone, the Coral Triangle. The IAA is the epicentre of biodiversity of not only corals, but many other marine organisms, including fishes, echinoderms, molluscs, crustaceans (Hoeksema 2007) and seagrasses (Short *et al.* 2007). Despite its remarkable biodiversity, many habitats and species in this region are at risk from anthropogenic pressures (McLeod *et al.* 2010; Hoegh-Guldberg 2010), making it one of the global priorities for marine biodiversity conservation (Selig *et al.* 2014).

1.2.2 Geological history

A major geological event commonly believed to have great impacts on the biodiversity in the IAA is the Pleistocene climatic oscillations (Figure 2.1). During the Pleistocene epoch (starting from about 2.5 Mya), the Earth's global climate fluctuated and was dominated by glaciation periods where the sea level fell approximately 118-135 m lower than today's level (Clark & Mix 2002). The dramatic sea level decline during the glaciation had a great impact on the coastline configuration and marine habitat in the IAA (Figure 1.1), significantly changing the pattern of marine connectivity in the region. In the west, the Sunda Shelf was exposed, creating a single broad sub-continent that bridged Java, Borneo and Sumatra with the mainland of Southeast Asia. In the east, the Sahul Shelf was also exposed, forming a single continuous landmass extending from New Guinea to Tasmania. Sulawesi, however, remained separated from Borneo by a narrow but deep ocean trench (Voris 2000). As the Sunda and Sahul Shelves emerged, the water

flow from the Pacific to the Indian Ocean had significantly reduced (Kuhnt *et al.* 2004).

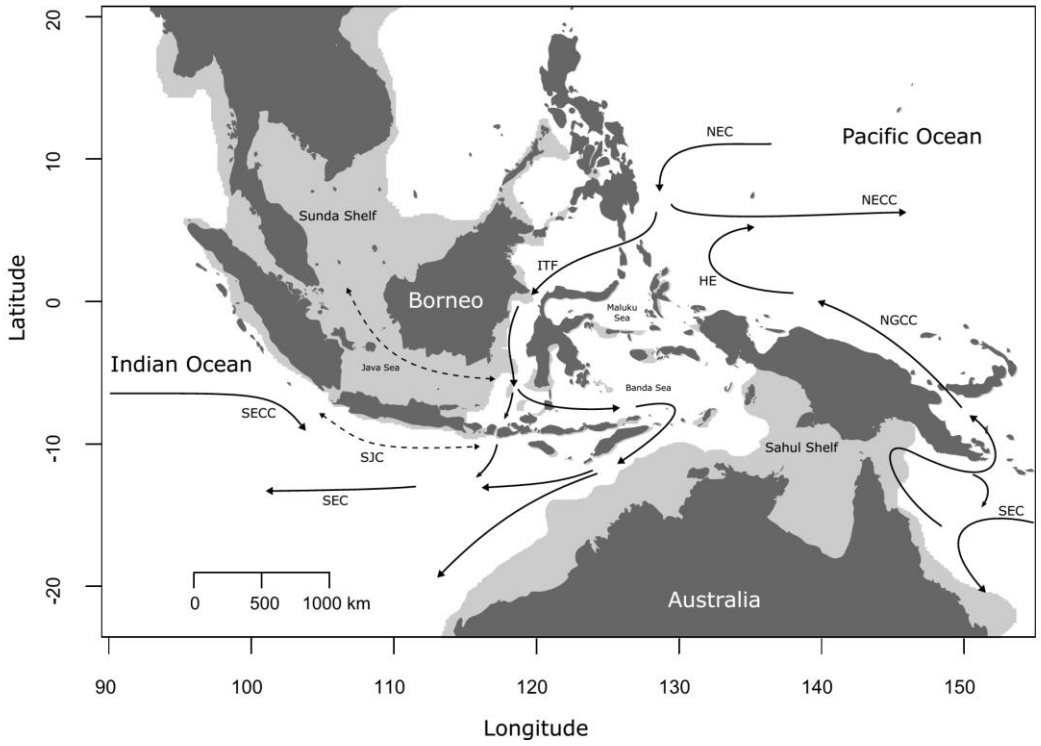

Figure 1.2 Map showing the Indo-Australian Archipelago with major ocean currents across the region (black arrows). Dark grey areas indicate present-day island configuration and light grey areas indicate landmass configuration during the Pleistocene low sea-level stands. Seasonally reversing currents are shown by dashed arrows. SEC: South Equatorial Current; NGCC: New Guinea Coastal Current, NECC: North Equatorial Counter Current; NEC: North Equatorial Current; SJC: South Java Current; SECC: South Equatorial Counter Current.

1.2.3 Connectivity and genetic structure in the IAA

Our understanding of connectivity and genetic structure in the IAA is based almost solely on genetic studies of fishes and invertebrates (Carpenter *et al.* 2011) which show a variety of connectivity patterns. The patterns of gene flow are not always concordant among taxa, partially due to different species biology and sampling methodologies. On one hand, some species exhibit high levels of connectivity across the region, for example the gastropod *Nerita plicata* (Crandall *et al.* 2008a) and three

species of reef fishes from the genus *Naso* (Horne *et al.* 2008). On the other hand, limited connectivity and strong genetic structure is demonstrated for some other species, such as stomatopods (Barber *et al.* 2006), giant clams (de Boer *et al.* 2008), seastars (Crandall *et al.* 2008b) and the mottled spinefoot fishes (Ravago-Gotanco & Juinio-Meñez 2010).

The variety of connectivity patterns found in the IAA probably represents the complexity of historical geology combined with contemporary oceanographic currents interacting with a range of ecological and life-history traits, particularly species dispersal-related traits. To date, studies on marine connectivity in the IAA have focused only on animals (Carpenter *et al*, 2011). As most marine animals have at least one mobile phase in their life cycle, the observed connectivity patterns are partially determined by the mobility characteristics and the duration of the larvae and adults. Analyses of corals, crustaceans, molluscs, bryozoans, tunicates and algae found that the more time propagules are in a planktonic state the further they tend to be dispersed. Animals with larval duration less than 100 h tend to disperse up to 1 km, while those with longer larval duration (more than 300 h) can disperse more than 20 km (Shanks *et al*, 2003). Seagrass, on the other hand, passively disperse via (i) propagules (seeds or fruits), (ii) vegetative growth, and (iii) vegetative fragments (unrooted shoots), ranging from a few metres to several hundreds of kilometres (Kendrick *et al.* 2012), thus different genetic patterns can be expected.

In addition, connectivity studies in the IAA are mostly limited to inferences from slowly-mutating markers (e.g. mitochondrial genes) (Carpenter *et al*, 2011). The use of markers with different mutation rates (e.g, single nucleotide polymorphisms/SNP vs microsatellites) may also reveal new patterns, because they detect genetic changes that occur at different divergence time. Mutation frequency in SNPs are generally lower than microsatellites, thus SNPs can provide better insights into evolutionary history, while microsatellites provide better resolution to detect changes in allele frequency that occur over more recent divergence time (Haasl & Payseur 2011). This study uses a tropical seagrass, *Thalassia hemprichii,* as a focal species to infer connectivity and genetic structure of marine plants in the IAA.

1.3 Seagrass importance and conservation

1.3.1 Significance of seagrass

Seagrasses occupy only a small portion of the world's ocean surface (less than 0.2%), yet they provide essential ecosystem goods and services (Duarte 2002; Cullen-Unsworth & Unsworth 2013). Some of the ecosystem services they provide are: primary production in food web dynamics (e.g. Duffy 2006, Vonk *et al*. 2008), nutrient recycling (McGlathery *et al.* 2007), and habitat provision for numerous associated species (e.g. Gartner *et al*. 2013, Hutchinson *et al*. 2014, Ávila *et al*. 2015). As many species found in seagrass meadows are commercially-valuable, seagrasses significantly support commercial and recreational fisheries (de la Torre-Castro *et al.* 2014; Tuya *et al.* 2014; Jackson *et al.* 2015). Furthermore, due to their high primary production rates combined with trapping of sediment and particles and sediment stabilisation, seagrass meadows are significant carbon sinks (Fourqurean *et al.* 2012; Duarte *et al.* 2013). However, despite their critical roles in coastal ecosystems, seagrass beds are declining globally at an accelerating rate, 110 km^2y^{-1} since 1980, mostly due to anthropogenic factors, particularly urban/industrial runoff, urban/port infrastructure development, agricultural runoff, and dredging (Waycott *et al.* 2009; Grech *et al.* 2012).

1.3.2 Significance of genetics and connectivity for seagrass conservation

In a rapidly changing environment, maintaining ecosystem resilience has become a major conservation goal. Ecosystem resilience is defined as the capacity of a system to absorb disturbance and to return to its original state without losing functions and services (Côté & Darling 2010). Ecosystem resilience can be maintained by protecting biodiversity, including genetic diversity, as this maximizes functional redundancy and response diversity in the ecosystem. Functional redundancy and response diversity are critical to compensate species/habitat loss, to buffer against any environmental changes and to facilitate successful reorganisation of ecological systems following disturbances (Mori *et al.* 2012). In seagrass, it has been shown that genetic diversity is critical for maintaining seagrass resilience and ecosystem functioning following disturbances (Hughes & Stachowicz 2004; Reusch *et al.* 2005;

Ehlers *et al.* 2008; Jahnke *et al.* 2015). In addition, restoration success can be enhanced by carefully selecting genetically appropriate provenance (Reynolds *et al.* 2013) and enriching genetic diversity (Reynolds *et al.* 2012). Genetic diversity determines plant fitness, stability and the functioning of the ecosystem at a community level. On the contrary, loss of genetic diversity may reduce the adaptive response to disturbances and environmental changes (Williams 2001; Reed & Frankham 2003).

Connectivity between and within seagrass populations is also an important component of resilience as it enhances their capacity for recovery and self-organization after disturbances. The exchange or dispersal of propagules, including adults, among populations may increase or maintain genetic diversity and also reduce the risk of local extinction because it allows continuous recruitment from other sites (Bernhardt & Leslie 2013). In the case of the source-sink model initially coined by Pulliam (1988), immigration from "source" populations can be critical for the persistence of "sink" populations by compensating for low recruitment in the "sink" populations (Lowe & Allendorf 2010). Additionally, loss of "source" populations may lead to subsequent loss in "sink" populations. Thus, identification of population connectivity is essential for site selection and for determining the size of the panmictic unit or a management unit (Procaccini *et al.* 2007). Given the practical difficulties of measuring dispersal directly at large spatial scales, genetic analysis offers tools, despite its limitations, for estimating population connectivity (Lowe & Allendorf 2010; Kendrick *et al.* 2016).

1.4 *Thalassia hemprichii*

The tropical seagrass *Thalassia hemprichii* (Ehrenberg) Ascherson, 1871 (Figure 1.3) is widely-distributed across the Indo-west Pacific and the Indian Ocean (Short *et al.* 2007). Fossil evidence suggests that this species evolved from a common ancestor of the genus *Thalassia* in the Tethyan Sea, with the divergence of *T. hemprichii* estimated to have occurred during the Miocene about 15 Mya (van Tussenbroek *et*

al. 2006). This coincides with the period of considerable geological changes and island formation in the IAA (Hall 2009).

This species is one of the dominant, meadow-forming seagrass species in the IAA. As a clonal organism, *T. hemprichii* is capable of expanding meadows by vegetative growth (van Tussenbroek *et al.* 2006). In most locations throughout its distribution range, sexual reproduction of *T. hemprichii* is seasonal and the reproductive season varies in different geographical areas (van Tussenbroek *et al.* 2006). The fruits are positively buoyant and *in situ* measures showed that floating fruits can disperse 23 – 74 km within 2-7 days (Lacap *et al.* 2002). Vegetative fragment can be afloat for months and still viable (Wu *et al.* 2016). The seeds have no dormancy period and may begin to germinate before they are released from the mature fruits (Kuo *et al.* 1991; Rollon *et al.* 2003). It is likely that the floating fruits would enable dispersal and high levels of connectivity among subpopulations, as has been observed in the sister species, *Thalassia testudinum,* in the Caribbean (van Dijk *et al.* 2009).

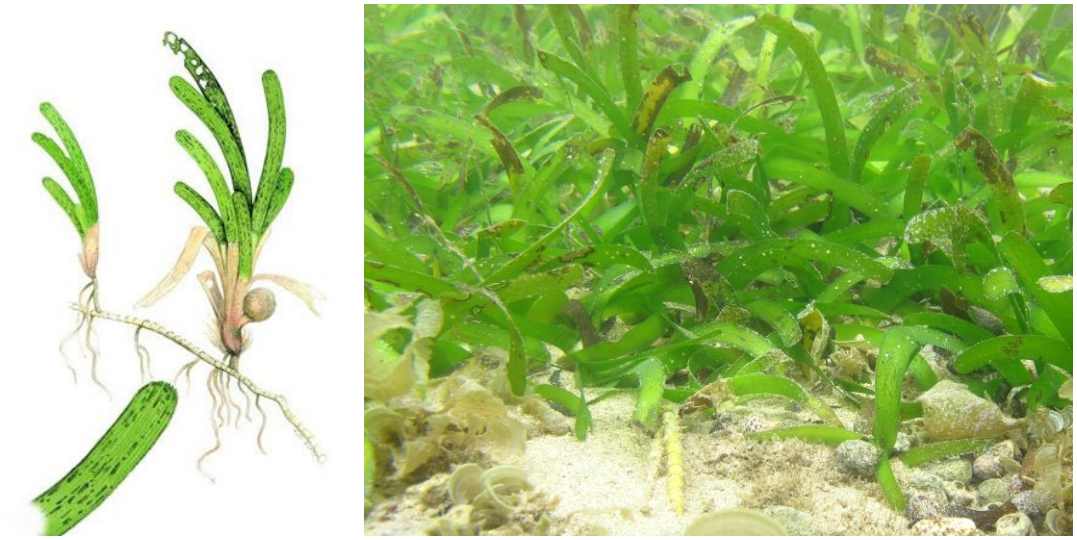

Figure 1.3 General morphology of *Thalassia hemprichii* (figure obtained from http://www.seagrasswatch.org) and the seagrass meadows in Bungus, Padang, Indonesia.

At present, connectivity and population genetic studies on seagrass have been limited mostly to temperate species (Procaccini *et al.* 2007; Kendrick *et al.* 2012). Patterns of gene flow and structure vary among species, and significant genetic structure can occur across a range of spatial scales; within a meadow, between sites in a location and among locations. Studies have reported panmictic populations occurring over large geographical distances (hundreds of kilometres) in widely distributed seagrasses such as *Z. noltii* (Coyer *et al.* 2004; Jahnke *et al.* 2016), *Z. marina* (Olsen *et al.* 2004), and *P. oceanica* (Arnaud-Haond *et al.* 2007; Serra *et al.* 2010). Genetic surveys of these species across their geographical range finds that they are generally connected up to 100 km (Kendrick *et al.* 2012). However, highly differentiated populations at small spatial scales (tens of kilometres) have been also reported, for example *Z. marina* (Muñiz-Salazar *et al.* 2006; Tanaka *et al.* 2011; Olsen *et al.* 2013), *Z. muelleri* (Sherman *et al.* 2016).

In the IAA, information on seagrass connectivity and population genetics is limited to two studies, *Enhalus acoroides* (Nakajima *et al.* 2014) and *C. serrulata* (Arriesgado *et al.* 2015). These studies reported significant genetic connectivity over very large spatial scales (thousands of km) and suggested that ocean currents might be responsible for the observed patterns. Since dispersal potential in seagrass vary with species (Kendrick *et al.* 2012; McMahon *et al.* 2014), further studies on other seagrass species can provide a more complete understanding of seagrass connectivity in the IAA. This study provides baseline information about connectivity and genetic structure of *T. hemprichii* populations. This information is important for conservation management of the seagrass in the future and understanding the evolutionary processes driving the biodiversity in the IAA.

1.5. Summary and aim of the study

The varying patterns of gene flow and genetic structure observed in the IAA probably represents the complexity of the palaeo-historical and contemporary environmental processes interacting with the varying ecological and life-history traits, particularly species dispersal-related traits. How genetic variation is distributed

across space and what factors predominantly influence spatial distribution of genetic variation is one of the primary questions related to marine conservation in the IAA (Barber 2009; Barber *et al.* 2011; Carpenter *et al.* 2011). Many phylogeographic studies, which mostly focused on animals using slowly-mutating markers (e.g. mitochondrial genes) have addressed this question but their conclusions vary depending on the taxon and methodology (Carpenter *et al.* 2011). Furthermore, these studies often highlighted the potential role of historical processes (e.g. Pleistocene sea-level fluctuations) and contemporary physical oceanography in the pattern of gene flow and genetic structure in marine populations. However, how each component contributes to the genetic pattern remains unclear, particularly in seagrass populations in the IAA. Teasing apart the relative contribution of each component to the genetic pattern is crucial to improve our understanding on the complexities of the evolutionary processes in the IAA.

To disentangle the relative contribution of historical and contemporary influences in genetic variation, studies should be designed to capture genetic variations at different spatial and temporal scales (Epps & Keyghobadi 2015). The inference of historical processes can be obtained by contrasting genetic data from, (i) different markers and (ii) a set of populations that experiences different historical processes. Such inference often requires genetic sampling over large spatial scales or across biogeographic regions because at this spatial scale populations are more likely to experience different demographic processes such as local extinction and colonization. In contrast, inference of contemporary processes requires: (i) rapidly mutating markers, to detect variation at a more recent divergence time; and (ii) genetic sampling at smaller spatial scales (local populations) because at this spatial scale, the populations are more likely to experience similar historical processes.

1.5.1. The overarching aim

The general aim of this dissertation is to investigate how genetic variation is distributed across space and how external factors (i.e. palaeo-historical events and contemporary processes) influence the pattern of genetic variation in the IAA. There are three specific objectives:

1. to identify the main factors that influence the spatial distribution of genetic variation across multiple marine taxa in the IAA,

2. to investigate how historical geological processes (e.g. Pleistocene glaciations) and contemporary physical oceanography drive the patterns of genetic structure in the IAA, and

3. to investigate how contemporary physical factors such as seascape features (e.g. water current and environmental variability) influence genetic patterns at a local spatial scales (less than 100 km).

For the first objective, population genetic data of co-distributed marine species was collated to estimate factors most important in driving genetic structure and differentiation across the IAA. The hypothesis is that the interaction of species biology (dispersal-related traits) and and physical processes (habitat heterogeneity, oceanographic-geologic features) determine the spatial distribution of genetic variation across a broad range of marine taxa in the IAA. This part is presented in Chapter 2 in the thesis (Figure 1.4).

For the second objective, the temporal scale (palaeo-historical *vs* contemporary processes) at which physical/environmental factors influence gene flow and genetic structure was investigated at a large spatial scale (>300 km, Chapter 3) using a seascape genetic approach. To minimise the effect of different biological traits, the seascape genetic analysis focused only on one species, *Thalassia hemprichii*, one of the dominant seagrass species in the IAA. The analysis used two genetic markers that have different divergence-time resolution: SNPs (relatively slow mutation rate) and microsatellites (relatively fast mutation rate) (Haasl & Payseur 2011).

For the third objective, the study focused on gene flow and genetic structure of the seagrass at more recent divergence time (using microsatellite) on a local spatial scale (<75 km, Chapter 4). The role of contemporary physical factors such as water current and habitat variability in driving the genetic patterns at this spatial scale was also examined using a seascape genetic approach.

Finally, the significance and implications of the results from all objectives are summarised and discussed in Chapter 5.

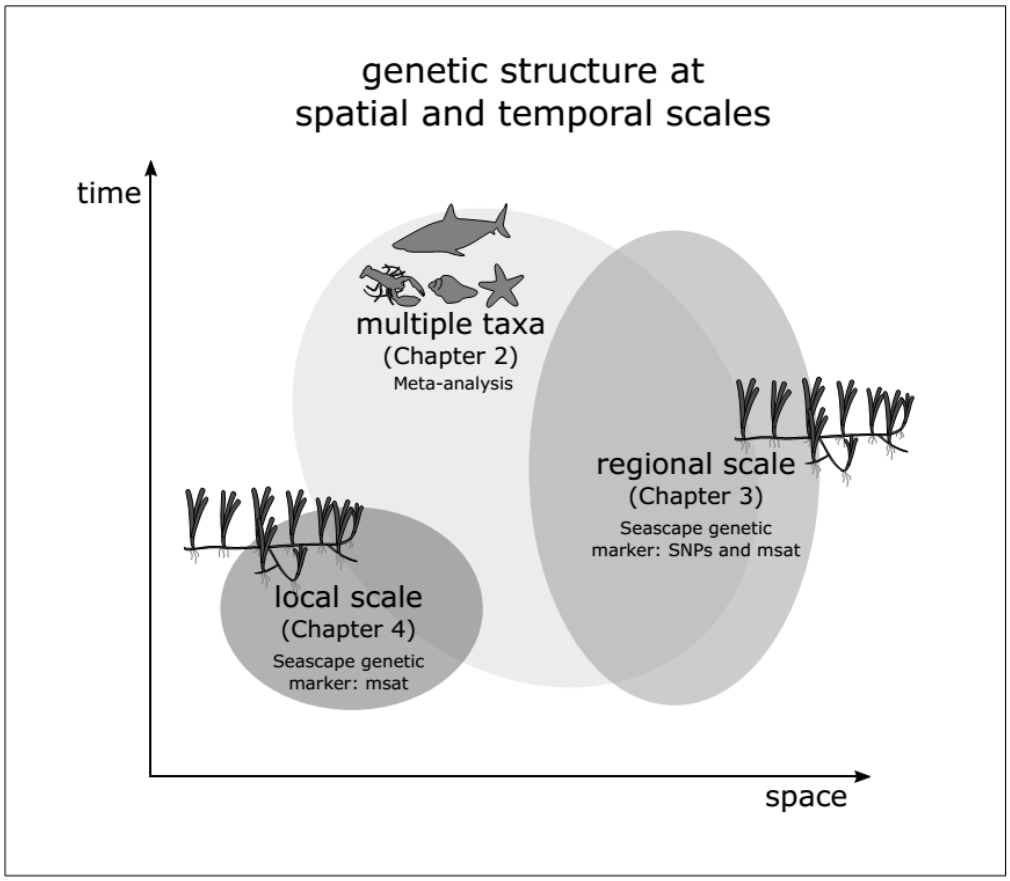

Figure 1.4. The multiple approaches used to understand genetic structure and the factors driving it for marine organisms and, in greater detail, for the seagrass *T. hemprichii*, across different spatial and temporal scales in the IAA, with reference to the relevant chapters in this thesis.

Chapter 2

PREDICTORS OF MARINE GENETIC STRUCTURE

Abstract

Among the factors affecting genetic structure of marine organisms across their ranges are habitat heterogeneity, oceanographic/geological features and dispersal-related traits of the species. The objective of this study is to assess the importance of these factors in determining genetic structure of marine species in the Indo-Australian Archipelago (IAA). The study collated data on 65 marine species from six taxa groups (fishes, molluscs, crustaceans, echinoderms, corals and marine plants) and used generalized linear models (GLMs) to estimate the best predictors of genetic structure. Genetic structure was characterized by F_{ST} and the number of genetic clusters. Predictors tested were: the type of genetic markers; habitat heterogeneity and oceanographic-geological features, represented by the number of marine ecoregions; species dispersal-related traits (e.g. pelagic larval duration-PLD) and geographic distance separating populations. The analysis indicated that the type of genetic markers significantly influenced F_{ST}, but not genetic clusters. The use of mtDNA resulted in higher F_{ST} than nuclear markers (microsatellite and allozyme). The most important predictors for F_{ST} were reproductive strategy, adult life habits, and PLD; while for genetic clusters the number of marine ecoregions and PLD were most important. The results generally indicates that dispersal biology and regional physical/environmental processes influence population genetic structure in a similar manner across a broad range of marine species in the IAA. The significance of marine ecoregions in predicting genetic clusters suggests that a marine ecoregion may not only be unique from an ecological perspective, but also from an evolutionary perspective. The finding emphasizes the importance of incorporating genetic cohesiveness into the designation of marine ecoregions.

2.1 Introduction

Most species are composed of spatially separated populations that are connected by dispersal. Successful dispersal, that is when migrants settle and interbreed with members of a recipient population, results in exchange of genetic material (gene flow). The level of gene flow, together with mutation, selection and genetic drift, can influence the spatial distribution of genetic variation within and among populations (genetic structure). High gene flow homogenizes genetic variation by counteracting the effect of mutation, selection and genetic drift, while low gene flow leads to genetic differentiation. Barriers to gene flow among populations generate genetic clustering, such that they no longer behave as a single, randomly mating population (panmictic population) (Slatkin 1987; Charlesworth *et al.* 2003; Hedrick 2012).

Spatial genetic structuring is a consequence of the interaction between intrinsic (e.g., life-history traits) and extrinsic factors (e.g., habitat heterogeneity and dispersal barriers) over time (Lowe *et al.* 2004). Among the intrinsic factors, the duration of early life stages (pelagic larval duration–PLD) has been highlighted as a key factor in determining genetic structure. A longer PLD increases the species' dispersal potential as larvae or dispersules are transported by currents for longer (Shanks *et al.* 2003; Shanks 2009). As dispersal should theoretically facilitate gene flow (Wright 1931; Slatkin 1987), PLD should be inversely correlated with genetic structure (Palumbi 1992; Doherty *et al.* 1995; Siegel *et al.* 2003). However, a number of recent analyses found weak or no correlation between PLD and genetic structure, hence other dispersal-related traits (e.g. reproductive strategy and adult mobility) could also be important for influencing spatial genetic structure (Bradbury *et al.* 2008; Galarza *et al.* 2009; Weersing & Toonen 2009; Kelly & Palumbi 2010; Riginos *et al.* 2011; Selkoe *et al.* 2014).

Extrinsic factors influencing genetic structure include geological history, past and/or contemporary oceanography, and habitat heterogeneity. Historical geologic processes could generate biogeographic barriers that restrict gene flow in many marine species. For example, the exposed continental shelves during periods of low sea level in the

Pleistocene caused vicariance in marine species, e.g. giant clams (Kochzius & Nuryanto 2008) and groupers (Gaither *et al.* 2011). Barriers to gene flow could also emerge from contemporary oceanographic processes, e.g. ocean eddy preventing gene flow in giant clams (de Boer *et al.* 2008), although water currents could also be the dispersal vector, facilitating gene flow among populations (e.g. asymmetric pattern of gene flow in the brown seaweed *Sargassum fusiforme* driven by the Kuroshio current, Hu *et al.* 2013). Habitat heterogeneity could act as a driver of local selection and adaptation, thus contributing to patterns of genetic structure (Riginos & Liggins 2013; Wang & Bradburd 2014).

The Indo-Australian Archipelago–IAA (Figure 2.1) comprises more than 20,000 islands situated in the Central Indo-Pacific and is one of the most geologically dynamic and complex regions on earth (Lohman *et al.* 2011). Although it occupies only about 4% of the planet's land surface (Lohman *et al.* 2011) the IAA is the epicentre of biodiversity; not only of corals, but also fishes, echinoderms, molluscs, crustaceans and seagrasses (Hoeksema 2007; Short *et al.* 2007). Despite its importance, many habitats and species in this region are threatened with extinction under current and predicted future anthropogenic pressures (McLeod *et al.* 2010; Hoegh-Guldberg 2010). A meta-analysis by Selig *et al* (2014) highlighted this region as one of the global priorities for marine biodiversity conservation, yet research efforts that support conservation management are lacking (Fisher *et al.* 2011).

What factors affect spatial distribution of genetic variation is one of the primary questions related to marine conservation in the IAA (Barber 2009; Barber *et al.* 2011; Carpenter *et al.* 2011). Many phylogeographic studies have addressed this question but conclusions vary depending on the focal taxon and methodology. A recent work by Carpenter *et al* (2011) has attempted to reveal commonalities in patterns in the IAA across a broad range of marine taxa. This work used a qualitative approach examining the phylogeographic signal of each species but did not take dispersal-related traits into account. Here, a quantitative approach (generalized linear modelling–GLM) was employed to address the question of what factors are most important for influencing genetic structure of marine organisms in the IAA. The hypothesis that habitat heterogeneity, oceanographic-geologic features and dispersal-

related traits would significantly determine genetic structure in a range of marine species was tested by comparing data of co-distributed species across the IAA, within the Central Indo Pacific biogeographic realm (Spalding *et al.* 2007). This study identifies a more generalized picture of genetic structure across a broad range of marine taxa in the IAA using GLM.

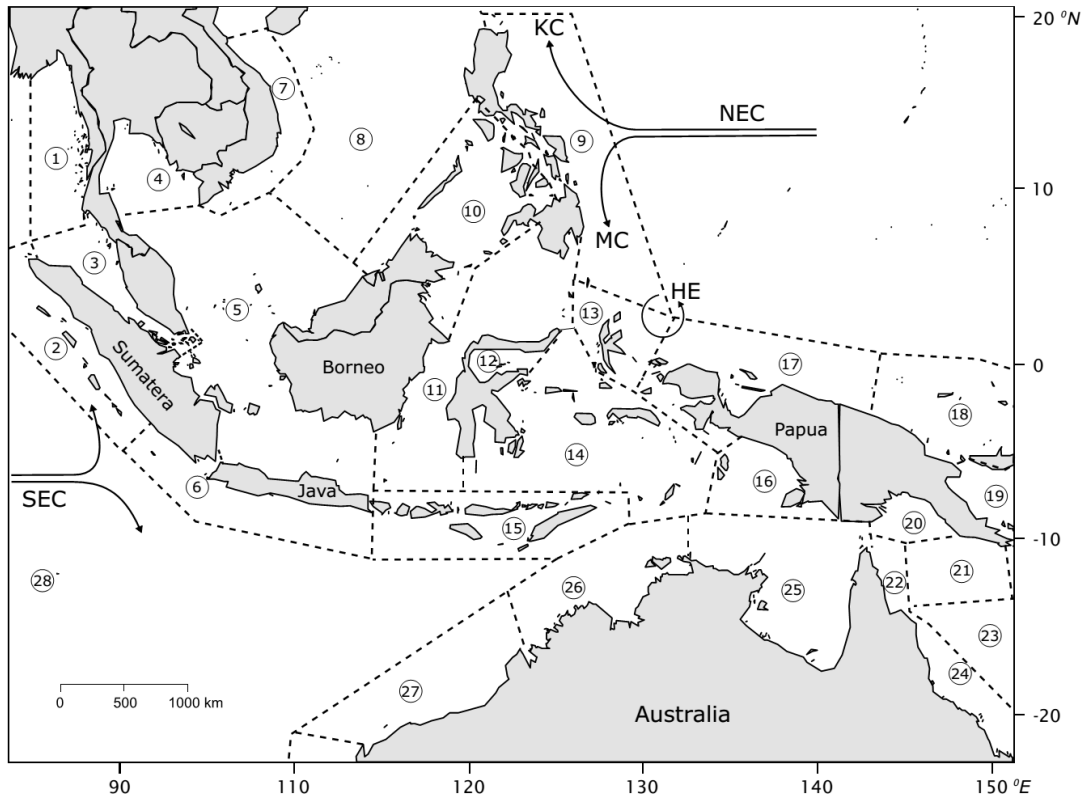

Figure 2.1 Marine ecoregions in the Indo-Australian Archipelago based on Spalding *et al*. (2007). Dashed lines correspond to boundaries of each ecoregion. (1)-Andaman Sea Coral Coast, (2)-Western Sumatra, (3)-Malacca Strait, (4)-Gulf of Thailand, (5)-Sunda Shelf/Java Sea, (6)-Southern Java, (7)-Southern Vietnam, (8)-South China Sea Oceanic Islands, (9)-Eastern Philippines, (10) Palawan/North Borneo, (11)-Sulawesi Sea/Makassar Strait, (12)-Northeast Sulawesi, (13)-Halmahera, (14)-Banda Sea, (15)-Lesser Sunda, (16)-Arafura Sea, (17)-Papua, (18)-Bismarck Sea, (19)-Solomon Sea, (20)-Gulf of Papua, (21)-Southeast Papua New Guinea, (22)-Torres Strait Northern GBR, (23)-Coral Sea, (24)-Central and Southern GBR, (25)-Arnhem Coast-Gulf of Carpentaria, (26)-Bonaparte Coast, (27)-Exmouth to Broome, (28)-Cocos-Keeling/Christmas Island. NEC=North Equatorial Current, KC=Kuroshio Current, MC=Mindanao Current, HE=Halmahera Eddy, SEC=South Equatorial Counter current.

2.2 Materials and Methods

2.2.1 Literature survey

Peer-reviewed publications reporting population genetic structure of marine species in the IAA were searched using the Web of Science and Google Scholar databases (November 2014). The search terms included: “gene flow”; “genetic structure”; “phylogeography”; and “population genetics”. As this study was spatially limited to the IAA, the search results were refined using these following terms: ”Indo-Australian Archipelago”; “East Indies”; “Coral Triangle”; “Indo-Malay”; “Indonesia”; “Malaysia”; “Philippines”; or “Australia”. This yielded 94 publications. Publications that contained all of the following: 1) marine species; 2) sampling locations within the IAA; and 3) data from which spatially explicit genetic structure could be determined were selected for further analysis (Appendix, Table A2.1).

2.2.2 Data extraction

Based on the hypothesis, four sets of variables were extracted: measures of genetic structure, habitat heterogeneity and oceanographic-geologic features, geographic distance and species dispersal-related traits (Table 2.1). Two measures of genetic structure were used: (i) global F_{ST} and (ii) the number of genetic clusters (*cluster*) derived from examining the *K*-value of STRUCTURE analyses, the number of significant clusters in a principal coordinate analysis, or the number of distinct clades in a phylogenetic tree and/or haplotype network provided in each study. While F_{ST} is a common measure of genetic structure (Meirmans & Hedrick 2011), genetic clustering is commonly used to identify managemet units relevant in conservation management and cannot be fully addressed using only F_{ST}. While many reviews or meta-analyses only used F_{ST} to measure genetic structure (e.g., Bradbury *et al*., 2008; Galarza *et al*., 2009; Weersing & Toonen, 2009; Kelly & Palumbi, 2010), the use of genetic clustering provides an alternative measure of genetic structure.

Habitat heterogeneity and oceanographic-geological features were represented by the number of marine ecoregions covered by each study (*ecoregion*) as defined from the

Marine Ecoregions of the World system (MEOW, Spalding *et al*. 2007) (Figure 2.1). A marine ecoregion is characterized by a distinct suite of oceanographic or topographic features, such as isolation, upwelling, nutrient inputs, freshwater influx, temperature regimes, exposure, sediments, currents, and bathymetric or coastal complexity (Spalding *et al.* 2007). The variable geographic distance (*distance*) was calculated by measuring the pairwise minimum distance by sea (without crossing any landmass) among all sampling sites in each study in Google Earth v7.1.2.2041. Then, the largest pairwise minimum distance was included in the analysis as the maximum geographic distance to represent the spatial scale of the study.

The dispersal-related traits examined were obtained from peer-reviewed publications, IUCN Redlist (iucnredlist.org), FishBase (fishbase.org) and LarvalBase (larvalbase.org); and included pelagic larval duration (*PLD*), adult life habit (with respect to adult mobility), reproductive strategy (with respect to how sperm and eggs are released) and egg type (related to how fertilized eggs are dispersed).

The variable *PLD* was defined as the maximum pelagic larval duration in hours for each species. For marine plants (seagrasses and mangroves) the *PLD* was determined based on the maximum viability of the seed before settlement. Adult life habit (*adult*) represents the species motility in the adult phase, which has the potential to influence dispersal and genetic structure. This variable was classified into sessile (e.g. corals), sedentary (restricted movement, e.g. sea urchin), motile (freely moving/swimming e.g. fishes) and migratory (e.g. the skipjack tuna *Katsuwonus pelamis*) (e.g., Maguire *et al*., 2006; de Juan *et al*., 2009). The reproductive strategy (*rep. strategy*) pertaining to the mode that sperm and eggs are released was classified into broadcaster and brooder. The brooders potentially exhibit greater genetic structure than the broadcast-spawning species due to the lack of a planktonic dispersive stage (Foggo *et al.* 2007; Bradbury *et al.* 2008). The variable egg type (*egg*) related to the mode that fertilized eggs are dispersed, either in the pelagic or benthic zone or as direct development (e.g. some sharks). Pelagic eggs are predicted to have a greater dispersal potential (Bradbury *et al.* 2008; Riginos *et al.* 2011). Species that mouth-/pouch brood (e.g. seahorses) or guards their eggs (e.g. *Amphiprion ocellaris*) were classified as benthic eggs (Table 2.1). For marine plants, seed buoyancy (buoyant=pelagic or

sinking=benthic) was used to define the category for the variable *egg*. Additionally, the genetic markers in the study were also recorded as "sequences" for mtDNA, "allozyme" for allozyme, and "other" for microsatellite and other markers.

Table 2.1. Criteria and variables extracted from peer-reviewed studies.

Criteria	Variable
Genetic structure	• F_{ST} • Number of genetic clusters (*cluster*)
Habitat heterogeneity and oceanographic-geologic features	• Number of marine ecoregions (*ecoregion*)
Geographic distance	• Maximum distance among sampling sites (*distance*)
Dispersal-related traits	• Pelagic larval duration (*PLD*)
	• Adult life habit: sessile, sedentary, motile, migratory (*adult*)
	• Reproductive strategy: broadcaster, brooder, mixed (*rep. strategy*)
	• Egg type: pelagic, benthic, direct developers (*egg*)
Genetic marker	• Type of genetic markers used: mtDNA, allozyme, msat/other (*marker*)

2.2.3 Statistical analysis

Generalized linear models were used to investigate which variables best predicted genetic structure. This approach accommodates non-linear distributions and different data types (e.g. continuous and categorical). The basic formula of GLMs for each response variable (F_{ST} and *cluster*) included the predictor variables *ecoregion*, *distance*, *PLD*, *adult*, *rep. strategy*, and *egg*. The variable *marker* was treated as a fixed factor in the model due to differences in attributes and sensitivity of the genetic markers to detect genetic variation (Parker *et al.* 1998; Schlötterer 2004).

Four different sets of models were run, the first on the full dataset to examine general patterns across all species (1) and three sets using subsets of the data. In set 2, all

records with no pelagic life-history phase were removed (n=15), to minimise the potential bias from the absence of larval duration in the analysis. In subset 3 only sessile and sedentary species (n=48) were used and subset 4 only mobile species (n=36) were included. This approach allowed us to examine if the drivers of genetic structure were consistent across the adult life habits. Mobile species have a greater potential to move over their life-cycle compared to sedentary species, hence different predictors may be important for genetic structure.

A key assumption for GLMs is independence among continuous predictor variables (Fox & Weisberg 2011). This was tested for *ecoregion*, *distance*, and *PLD* using Hoeffding's D test in function *hoeffd* of **Hmisc** 3.15-0 package (Harrell Jr 2015), confirming independence (max. Hoeffding's D value of pairwise comparison= 0.2). Multicollinearity was also not detected from the variable inflation factor-VIF (*PLD*= 1.05; *ecoregion*= 1.58; *distance*= 1.53) calculated using **car** 2.0-25 package (Fox *et al.* 2015).

The R package **glmulti** 1.0.7 was used to calculate the GLMs (Calcagno & Mazancourt 2010). It generates all possible model formulas and fits them with a GLM. This approach does not require 'a priori' selection of candidate models, which is needed in other packages (e.g., **MuMIn**). In the case of missing data, **glmulti** excluded the corresponding variable from the calculation. For F_{ST}, log linearized F_{ST} ((F_{ST}+0.001)/(1-(F_{ST}+0.001)) was used to improve the approximation of linearity and the GLMs was run using the Gaussian distribution family with an identity link function. For the response variable *cluster*, the GLM was performed using the Poisson distribution family with a log link function since *cluster* is count data. Model selection was based on Akaike's Information Criterion (AIC). Models within the lowest two AIC units are considered best at explaining the response variable (Burnham & Anderson 2002). To examine the contribution of each predictor in determining genetic structure, relative evidence weight of the predictor was calculated as the sum of the relative evidence weights of all models in which the predictor appears (Calcagno & Mazancourt 2010).

McFadden's pseudo-R^2 was calculated to represent the explained variation in GLMs since the models had non-linear distributions. Values of McFadden's pseudo-R^2

between 0.2 - 0.4 are considered to be indicative of very good model fits (Louviere *et al.* 2000). The influence of the important predictors resulting from **glmulti** (if any) was examined using the best models. If the predictor was continuous data, the influence of the predictor was derived directly from the coefficient value in the best model. If the predictor was categorical, multiple comparison of means in the package **multcomp** was used to test the effect of different categories (Hothorn *et al.* 2008). All statistical analysis was done in the statistical computing environment, R version 3.2.2 (R Development Core Team, 2015) and RSTUDIO version 0.98.1103.

2.3 Results

2.3.1 Literature survey

Literature survey resulted in 55 filtered-publications. There are some publications, each of which reported more than one species. There are also publications reporting similar species, but they used different markers and covered different regions. In total, there are 65 marine species (six taxonomic groups: fishes, molluscs, crustaceans, echinoderms, corals and marine plants). The full dataset comprised 84 records, in which fishes contribute almost 50% of the records. After filtering the full dataset, the subset data contained 69 records of species with pelagic larval state (PLD>0), 48 records of sessile and sedentary species and 36 records of mobile (including migratory) species (Figure A2.1, Appendix).

2.3.2 Predictors of genetic structure

Using F_{ST} as the response variable, the GLMs indicated that *adult*, *rep. strategy* and *PLD* were consistently the most important predictors of genetic structure. These variables were observed in all best models from the full dataset (all species), subset 2 (PLD>0) and subset 3 (sessile and sedentary; see Table 2.2). The importance of these predictors was also indicated by the fact that the R^2 values changed little when other predictors were excluded. Although the variable *ecoregion* and *distance* did not consistently appear in the best models in all datasets, inclusion of these increased the

explained variation, with the highest pseudo-R^2 *(= 0.42*) for the full dataset and all explanatory variables except *egg*.

In contrast to the F_{ST} response variable, the variables *ecoregion* and *PLD* were the most important predictors for the response variable *cluster*; in the full dataset (pseudo-R^2*=0.22-0.23*), subset 2 (pseudo-R^2*=0.25*) and subset 3 (pseudo-R^2=0.07-0.20; see Table 2.3). The R^2 values also changed little when other predictors were excluded, indicating the importance of *PLD* and *ecoregion*. In subset 4 (mobile species), **glmulti** returned similar patterns for both response variable F_{ST} and *cluster*, where all best models showed weak model fits (pseudo-R^2<0.20). In addition, the model *Fst ~ marker* was among the best supported models, within the lowest two AIC units.

Table 2.2. Best models generated for F_{ST} using full and restricted dataset. Only models within the lowest two AIC units are shown in table

Model	ΔAIC	AIC Weight	*Pseudo-R^2*
Full dataset (n=84)			
Fst ~ (marker) + adult + rep. strategy + PLD + ecoregion	0	0.286	0.42
Fst ~ (marker) + adult + rep. strategy + PLD + distance	0.592	0.213	0.42
Fst ~ (marker) + adult + rep. strategy + PLD	0.663	0.206	0.40
Fst ~ (marker) + adult + rep. strategy + PLD + ecoregion + distance	1.557	0.131	0.42
Species with a pelagic larval stage / PLD > 0 (n=69)			
Fst ~ (marker) + adult + rep. strategy + PLD + distance	0	0.309	0.35
Fst ~ (marker) + adult + rep. strategy + PLD + ecoregion	1.754	0.128	0.33
Fst ~ (marker) + adult + rep. strategy + PLD + egg + distance	1.770	0.127	0.35
Fst ~ (marker) + adult + rep. strategy + PLD + ecoregion + distance	1.892	0.120	0.35
Fst ~ (marker) + adult + rep. strategy + PLD	1.989	0.114	0.31
Sessile and sedentary (n=48) excluding predictor adult life habit			
Fst ~ (marker) + rep. strategy + PLD + egg + ecoregion	0	0.219	0.40
Fst ~ (marker) + rep. strategy + PLD	0.897	0.139	0.29
Fst ~ (marker) + rep. strategy + PLD + egg + distance	0.946	0.136	0.22
Fst ~ (marker) + rep. strategy + PLD + ecoregion	1.168	0.122	0.32
Fst ~ (marker) + rep. strategy + PLD + egg	1.384	0.110	0.35
Fst ~ (marker) + rep. strategy + PLD + distance	1.529	0.102	0.32
Fst ~ (marker) + rep. strategy + PLD + egg + ecoregion + distance	1.904	0.085	0.40
Mobile species (n=36), excluding predictor adult life habit			
Fst ~ (marker) + rep. strategy	0	0.134	0.16
Fst ~ (marker) + egg	1.091	0.078	0.19
Fst ~ (marker) + rep. strategy + ecoregion	1.105	0.077	0.19
Fst ~ (marker) + rep. strategy + distance	1.217	0.073	0.18
Fst ~ (marker)	1.744	0.056	0.06
Fst ~ (marker) + rep. strategy + PLD	1.861	0.053	0.17

Table 2.3. Best models generated for genetic cluster using full and restricted dataset. Only models within the lowest two AIC units are shown in table.

Model	ΔAIC	AIC Weight	*Pseudo-R^2*
Full dataset (n=84)			
cluster ~ (marker) + PLD + ecoregion	0	0.285	0.22
cluster ~ (marker) + PLD + ecoregion + rep. strategy	1.693	0.122	0.23
cluster ~ (marker) + PLD + ecoregion + distance	1.986	0.105	0.22
Species with a pelagic larval stage / PLD > 0 (n=69)			
cluster ~ (marker) + PLD + ecoregion	0	0.258	0.25
cluster ~ (marker) + PLD + ecoregion + rep. strategy	1.996	0.096	0.25
cluster ~ (marker) + PLD + ecoregion + distance	1.996	0.095	0.25
cluster ~ (marker) + PLD + ecoregion + egg	1.999	0.258	0.25
Sessile and sedentary (n=48) excluding predictor adult life habit			
cluster ~ (marker) + PLD	0	0.201	0.07
cluster ~ (marker) + PLD + distance	0.391	0.165	0.17
cluster ~ (marker) + PLD + ecoregion	0.531	0.154	0.16
cluster ~ (marker) + PLD + rep. strategy	1.858	0.079	0.07
cluster ~ (marker) + PLD + rep. strategy + ecoregion	1.939	0.076	0.20
Mobile species (n=36), excluding predictor adult life habit			
cluster ~ (marker)	0	0.177	0.002
cluster ~ (marker) + ecoregion	0.615	0.130	0.15
cluster ~ (marker) + PLD	1.765	0.073	0.03
cluster ~ (marker) + rep. strategy	1.785	0.072	0.03
cluster ~ (marker) + distance	1.982	0.066	0.004

The importance of predictors was indicated by the model-averaged importance of terms, which provides relative evidence weights (x-axis) for each predictor, ranging from 0 to 1 (Figure 2.2). A predictor with a relative evidence weight higher than 0.8 is considered to be highly significant (Calcagno & Mazancourt 2010). For the response variable F_{ST} in the full dataset, as well as subset 2 (PLD>0) the variables *adult*, *rep. strategy*, and *PLD* were consistently very close to 1, these predictors explaining genetic structure best (Figure 2.2, right-hand panel). As the variable *adult* was excluded in subset 3, *rep. strategy* and *PLD* consistently showed high relative evidence weight. For the response variable *cluster*, *PLD* and *ecoregion* had relative

evidence weights close to 0.8 or more in the full dataset and subset 2 (Figure 2.2, left-hand panel). Furthermore, genetic structure of subset 3 was predominantly affected by *PLD*, other predictors contributing less.

For subset 4 the relative evidence weights of all predictors fell far below the threshold 0.8 for both the response variables. None of the predictors show any significance in explaining the genetic structure of mobile species (Figure 2.2).

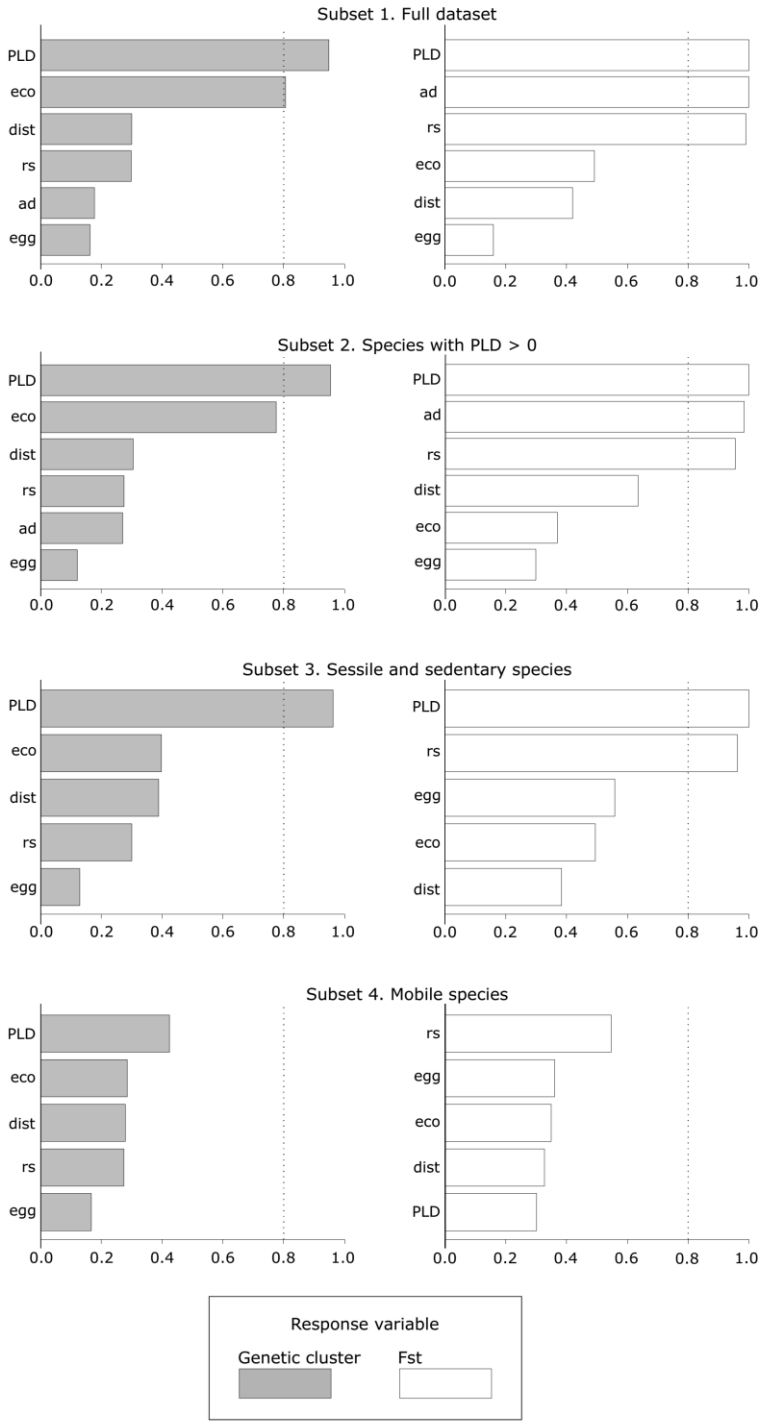

Figure 2.2 Relative evidence weight of predictors generated for genetic cluster (left-hand panel) and F_{ST} (right-hand panel) using the full dataset, species with PLD>0, sessile and sedentary species, and mobile species. The x-axis indicates relative evidence weight. A vertical dashed line at evidence weight 0.8 is the threshold to determine the importance of terms. Abbreviations, eco= *ecoregion*, dist= *distance*, rs= *rep. strategy*, ad= *adult* (adult life habit), egg= *egg* type.

2.3.3 Influence of the important predictors

The influence of each important predictor (*marker*, *adult*, *rep. strategy*, *PLD* and *ecoregion*) was derived from the model coefficient in the best GLMs with the lowest AIC. The type of genetic *marker* significantly influenced the predictions for F_{ST}, but not for genetic clustering. Multiple comparison of means showed that the influence of mtDNA sequences in determining F_{ST} was significantly higher than that of allozyme (p-value=0.032) and microsatellite (p-value=0.003), indicating that F_{ST} estimated from mtDNA is higher than that estimated from nuclear markers (Table 2.4), while no significant differences were found between allozyme and microsatellite.

Table 2.4. Multiple comparisons of means on the influence of the important categorical variables.

	Differences between category	**p-value**
Genetic markers		
msat &other > allozyme	0.2608	0.93711
mtDNA > allozyme	1.7854	0.03191
mtDNA > msat&other	1.5246	0.00352
Adult life habits		
migratory < free swimming	-2.5207	0.00265
Sedentary < free swimming	-0.5008	0.41243
Sessile > free swimming	0.8056	0.41243
Sedentary > migratory	2.0199	0.04027
Sessile > migratory	3.3264	< 0.001
Sessile > sedentary	1.3065	0.11375
Reproductive strategy		
Brooder > broadcaster	1.6380	0.00047

The adult life habit significantly influenced F_{ST}, with smaller F_{ST} values in migratory species compared with all other life habits (Table 2.4). Among the sessile, sedentary, and free-swimming species, no significant differences in F_{ST} prediction were found.

Reproductive strategy significantly influenced F_{ST} ($p<0.001$), brooders having higher F_{ST} than broadcasters (Table 2.4).

The Poisson model with a log link function showed that *ecoregion* positively influenced genetic clustering (coefficient=0.07). This positive effect suggests gene flow is limited across ecoregions. The *PLD* coefficient in the best models was negative but very small (-0.0000057). This is due to the large value discrepancy between *PLD* (ranging from 0-2400) and the number of genetic cluster (ranging from 1 to 11). As this is a Poisson model, genetic clustering is an exponential function of the explanatory variables (with the coefficient of the Intercept=0.47), therefore, *PLD* must be much smaller than *ecoregion*. For example, if we conducted a study across 10 marine ecoregions with PLD of 500 hours and mtDNA as the marker, we would expect to get at least 3 genetic clusters (y=exp(0.47+(10*0.07)+(500*-0.0000057)).

2.4 Discussion

2.4.1 The influence of dispersal-related traits

This study indicates that species dispersal biology (adult life habit, reproductive strategy, and PLD) influences population genetic structure and connectivity (as measured by F_{ST}) and this influence is similar for a broad range of marine species that inhabit the Indo-Australian Archipelago. The link between dispersal biology and population genetic structure lies in the success of dispersal, which is mediated by the movement of new genotypes and the distribution of these in space (Jordano 2010). In fact, dispersal biology essentially reflects the species' motion capacity (traits, or lack thereof, that enable the individuals to disperse) and navigation capacity (ability of individual to orient movement in space/time) (Nathan *et al.* 2008).

The analysis suggests that, all else being equal, species with higher motion capacity and/or navigation capacity (e.g. broadcast spawners or migratory species) will have higher genetic connectivity (smaller F_{ST}). These capacities increase the probability of moving across physical barriers, which restrict the dispersal path of species with limited motion or limited navigational capacities (e.g. brooder species and

sessile/sedentary species). As dispersal occurs predominantly during the early life stages in sessile and sedentary species (Cowen & Sponaugle 2009), reproductive strategy and pelagic larval duration were the most important predictors of F_{ST} for this subgroup.

Pelagic larval duration alone was not the best predictor of genetic structure, possibly explaining why previous studies found a weak relationship between PLD and genetic structure (e.g. Weersing & Toonen, 2009; Kelly & Palumbi, 2010). It is important to note that the analysis used only one PLD class, the maximum PLD, as Weersing & Toonen (2009) argued that the tails on the variation of larval duration were more informative than the mean PLD and genetic structure is influenced by multiple successful dispersal events over multiple generations (thus accounting for rare/extreme events). However, the PLD itself may not fully represent the 'true' scale of association between a species' dispersal and genetic structure, because it can vary greatly in space and time and is often estimated from a few individuals at one sampling site and generally under laboratory conditions (Wellington & Victor 1992; Macpherson & Raventos 2006; Weersing & Toonen 2009).

In contrast to sessile and sedentary species, no single factor was identified in predicting genetic structure in mobile species. As adults can disperse freely, the population connectivity is less likely constrained by dispersal barriers, larval dispersal and other dispersal-related traits, but more influenced by the behavioural ecology of the adults. For example, the spawning/reproductive behaviour and feeding migration have been shown to account for strong population connectivity in some species of salmonids, sharks, and herrings (Gaggiotti *et al.* 2009; Frisk *et al.* 2014).

2.4.2 Influence of genetic markers on F_{ST}

Across all studies, the study demonstrates that mtDNA tended to result in higher F_{ST} values than the other marker types. This corroborates previous studies showing differences between mtDNA and other marker types in measuring genetic structure (Weersing & Toonen 2009; Riginos *et al.* 2011) and supports the argument that direct comparisons of F_{ST} may be biased when the F_{ST} is estimated from different genetic marker types. The significant differences may result from: (i) the uniparental

inheritance of mtDNA leading to fixation faster than biparental inheritance of nuclear markers (thus higher F_{ST}), and (ii) differences in mutation rates, time to reach migration-drift equilibrium, and degree of polymorphisms among the marker types (reviewed in more details by Ballard & Whitlock, 2004; Weersing & Toonen, 2009). Highly polymorphic markers (such as microsatellite) would greatly reduce F_{ST} values (Meirmans & Hedrick 2011), because F_{ST} is inversely proportional to the total population heterozygosity. This may explain why F_{ST} derived from microsatellite markers is generally lower than mtDNA markers.

2.4.3 The importance of ecoregion

The importance of ecoregion was revealed in genetic clustering, but not in F_{ST}. This may be because global F_{ST} does not provide spatial information about genetic breaks, while genetic clustering contains information on where genetic breaks are and these might be congruent with ecoregion boundaries. Furthermore, the fact that this study focused only in the IAA, which is within the Central Indo Pacific biogeographic realm (Spalding *et al.* 2007), may also minimize the potential effect of major biogeographic transitions, which have been shown to affect F_{ST} (Riginos *et al.* 2011). Nevertheless, the significance of ecoregion revealed from genetic clustering indicates that different habitats and oceanographic-geological features among ecoregions act as barriers to gene flow in marine species.

While larval life history, habitat heterogeneity, and oceanographic-geological barriers on connectivity have been demonstrated to influence genetic structure for single taxa, such as corals (Baums *et al.* 2006), fishes (Galarza *et al.* 2009; Watson *et al.* 2010) and molluscs (Miller *et al.* 2013), the analysis confirms that they are important across a wide range of species from a number of taxonomic groups. Larval life history provides a means for dispersal, but the spatial scale and direction of dispersal is influenced by oceanographic or geologic barriers that may be contemporary or historical, like past changes in sea level and connectivity. For example, populations might be separated during Pleistocene glaciations, then re-joined as sea level rose, but genetic signatures of this historical separation can appear in genetic structure analysis. Even in the absence of dispersal barriers, individuals may reach new habitats, but local environmental selection may prevent them to

settle, recruit and reproduce thus preventing gene flow (Hunt & Scheibling 1997; Bierne *et al.* 2003; Marshall *et al.* 2010).

Due to the very complex configuration of oceanographic and geological features in the IAA (Lohman *et al.* 2011), the pattern of genetic structure cannot be explained simply as a function of geographic distance, which alone did not have a prominent or consistent effect on predicting genetic structure in the analysis (Figure 2.2). Two populations might be separated over geographic distances that are within the potential dispersal range of the species, but may still be genetically isolated from each other if gene flow barriers, such as water currents or local selection, are present. For example, the Halmahera Eddy (Figure 2.1- HE) is believed to be responsible for a pronounced genetic break in the redbelly yellowtail fusilier fish *Caesio cuning* between the east and the west of the Halmahera Island. This is despite the long PLD (37-47 days) of the fish and the highly mobile nature of the adult (Ackiss *et al.* 2013).

The observed increase in panmictic populations of marine species with the number of marine ecoregions sampled, indicates an association between genetic clustering and marine ecoregions. Thus, it can be proposed that, considering the PLD data and the absence of genetic data, marine ecoregions may be used to estimate population differentiation of marine species (McCreadie & Adler 2006; Muldoon & Goodman 2010). However, this does not substitute the need for population genetic studies, since the study has also shown that species traits (adult mobility, reproductive type, larval duration) may also influence the patterns of genetic structure.

2.4.4 Implications for marine conservation

The classification of Marine Ecoregions of the World (MEOW) does not explicitly consider a genetic dimension in its definition, despite the increasing awareness to consider genetic diversity in conservation planning (e.g. Sgrò *et al*., 2011; Rivers *et al*., 2014). MEOW focuses on the conservation of species, habitats, and ecological processes within a geographical space defined by natural characteristics thus is mainly based on ecological cohesiveness (Spalding *et al.* 2007). Despite the absence of a genetic dimension in its definition, the association between genetic structure and

marine ecoregions suggests that MEOW implicitly reflects genetic aspects in conservation planning for some marine species across the IAA. From the genetic perspective, this study validates the current implementation of ecoregion-based marine conservation. Genetic aspects should be of high priority in conservation because it contributes in population fitness (e.g. Reed & Frankham 2003, Vandewoestijne *et al*. 2008), influences species resistance to disease or disturbance (e.g. Altermatt & Ebert, 2008; Koh *et al*., 2012), and generally reflects the species' adaptive potential (e.g. Willi *et al*. 2006, Barrett & Schluter 2008).

Although this study shows the association between genetic structure and marine ecoregions, the absence of an explicit genetic dimension in MEOW could also mean that ecoregion-based marine conservation may still not representatively account for genetic variation essential for the long-term persistence of populations. Hence, as suggested by Carpenter *et al* (2011) based on a comparative examination of discordant and concordant phylogeographic breaks of marine taxa in the IAA, it can be proposed that population genetics be incorporated into the definition of a marine ecoregion as this would improve the applicability of marine ecoregions as a general framework for conservation management. Furthermore, the necessity of incorporating population connectivity and genetics in conservation management (Kool *et al.* 2013; Magris *et al.* 2014) strengthens the argument of including genetic cohesiveness into the designation of marine ecoregions.

The incorporation of genetic cohesiveness may lead to, (i) delineation or modification of current marine ecoregion boundaries, and (ii) establishment of new marine ecoregions. These changes should be considered if studies show genetic breaks that are not in concordance with the boundaries of marine ecoregions. For example, genetic breaks within the Eastern Philippines ecoregion (Figure 2.1, number 9) separating two clusters of populations were reported from several species and appears to be associated with the bifurcation of the North Equatorial Current into the Kuroshio Current and the Mindanao Current (Ravago-Gotanco *et al.* 2007; Ravago-Gotanco & Juinio-Meñez 2010; Nakajima *et al.* 2014).

Redefining marine ecoregions by incorporating genetic cohesiveness would require sufficient and well-represented genetic data. In the IAA, despite the increases in the

number of genetic studies in the last decade (Carpenter *et al.* 2011), there are still many gaps both spatially and within certain taxa (e.g. marine macrophytes). Over time these gaps will reduce and there may be more justification for assessing the incorporation of genetic cohesiveness into the marine ecoregion classification. To assist with this process, future genetic studies should sample sites that are representatively nested in each ecoregion within a marine province or a marine realm.

Chapter 3

HISTORICAL AND CONTEMPORARY GENETIC STRUCTURE OF THE SEAGRASS *THALASSIA HEMPRICHII*

Abstract

Understanding spatial patterns of gene flow and genetic structure is essential for the conservation of marine ecosystems. Contemporary ocean currents and historical isolation due to Pleistocene sea-level fluctuations have been predicted to influence the genetic structure in marine populations. In the Indo-Australian Archipelago (IAA), the world's hotspot of marine biodiversity, seagrasses are a vital component but information on their population genetics is very limited. Here, phylogeography and genetic structure of the seagrass *Thalassia hemprichii* in the IAA was examined based on single nucleotide polymorphisms (SNPs) and a panel of 16 microsatellite markers. Then, the relative importance of historical isolation and contemporary ocean currents were examined to determine their role in driving the patterns of genetic structure. Results from SNPs revealed three groups of populations: eastern Indonesia, western Indonesia (Sunda Shelf), and Indian Ocean; while the microsatellites showed five groups of populations (eastern Indonesia, Sunda Shelf, Lesser Sunda, Western Australia, and Indian Ocean). Asymmetrical gene flow among groups of populations with a trend of south-westward migration from eastern Indonesia was inferred from both SNPs and microsatellites. Genetic diversity was generally higher in eastern Indonesia (Biak and Tual) and decreased southwestward (Sunda Shelf, Kimberley, and Exmouth). The seagrass lineage divergence and migration pattern are consistent with the Centre of Origin and the Centre of Accumulation hypotheses. The analysis also demonstrated that the pattern of genetic

structure and connectivity is attributed partly to the Pleistocene sea level fluctuations modified to a smaller level by contemporary ocean currents. Based on genetic partitioning, migration pattern and genetic diversity, this study highlights conservation priorities for *T. hemprichii* populations in the IAA.

3.1 Introduction

Gene flow affects the distribution of genetic diversity across space, strongly influencing ecological and evolutionary processes such as species adaptive potential (Swindell & Bouzat 2006), population persistence (Palstra & Ruzzante 2008) and genetic integrity of species (Rieseberg & Burke 2001). Consequently spatial patterns of genetic structure should be an explicit consideration in conservation management (e.g. Crandall *et al.* 2000; Magris *et al.* 2014). However, understanding the observed patterns of gene flow and genetic structure in the marine environment is complicated as populations exist in complex seascapes, where dynamic features such as water currents and static attributes such as islands interact with species dispersal traits to influence genetic exchange over time (Riginos & Liggins 2013; Chapter 2 in this thesis). Furthermore, historical geological processes such as Pleistocene sea-level fluctuations can affect genetic exchange among populations in the past (e.g. Ravago-Gotanco & Juinio-Meñez 2010). Therefore, to fully understand processes driving genetic structure one needs to examine both the historical and contemporary drivers of gene flow.

The extraordinary marine biodiversity and endemism concentrated in the Indo-Australian Archipelago (IAA) has been the subject of phylogeographic studies for decades to attempt to understand the processes responsible for such biodiversity. Many of these studies have highlighted the potential role of Pleistocene sea-level fluctuations and oceanographic patterns in influencing the evolutionary processes that have led to the high biodiversity in the marine environment (e.g. Barber *et al.* 2006). Historically, the Pleistocene sea-level fluctuations had changed the configuration of coastlines and water bodies, oceanic currents, and availability of marine habitats in the region. The lowered sea levels exposed the Sunda and Sahul

Shelves leading to almost complete elimination of water flow from the Pacific to the Indian Oceans and the isolation of sea basins in the region (e.g. the Celebes and Maluku Seas), disconnecting populations that once freely exchanged migrants (Figure 3.1; Voris 2000; Lohman *et al.* 2011). It has been hypothesised that these historical isolation events caused not only genetic divergence between Indian and Pacific populations (e.g. Duda Jr & Palumbi 1999), but also regional genetic differentiation among populations across the IAA, for example in the boring giant clam (Kochzius & Nuryanto 2008), and the anemonefish (Timm & Kochzius 2008). In more recent times the oceanographic setting of the IAA (Figure 3.1) contributes to shaping and maintaining patterns of genetic structure, particularly for species with passive dispersal mechanisms, where water currents are the main dispersal vector. For example, a genetic break on the northern shores of Papua and across Halmahera Sea (which was still submerged during the Pleistocene low sea-level stands) was predicted to be due to the Halmahera Eddy acting as a barrier to larval dispersal westward into the Celebes, Maluku, and Banda Seas (Barber *et al.* 2006; de Boer *et al.* 2014).

Despite the essential role seagrasses play in ecosystem functioning and services (e.g. Duffy 2006; de la Torre-Castro *et al.* 2014), globally they are declining and urgent measures are required for effective conservation and management (e.g. Short *et al.* 2014). In the IAA, information on seagrass population genetics is limited to two studies on *Enhalus acoroides* (Nakajima *et al.* 2014) and *Cymodocea serrulata* (Arriesgado *et al.* 2015). Both species showed significant genetic structure over regional scales (>300 km) and ocean currents are believed to be responsible for the observed patterns. Here, this study focuses on the seagrass *Thalassia hemprichii* (Ehrenberg) Ascherson, 1871, a widely-distributed seagrass species in the Indo-West Pacific, and one of the dominant seagrass species in the IAA (Short *et al.* 2007).

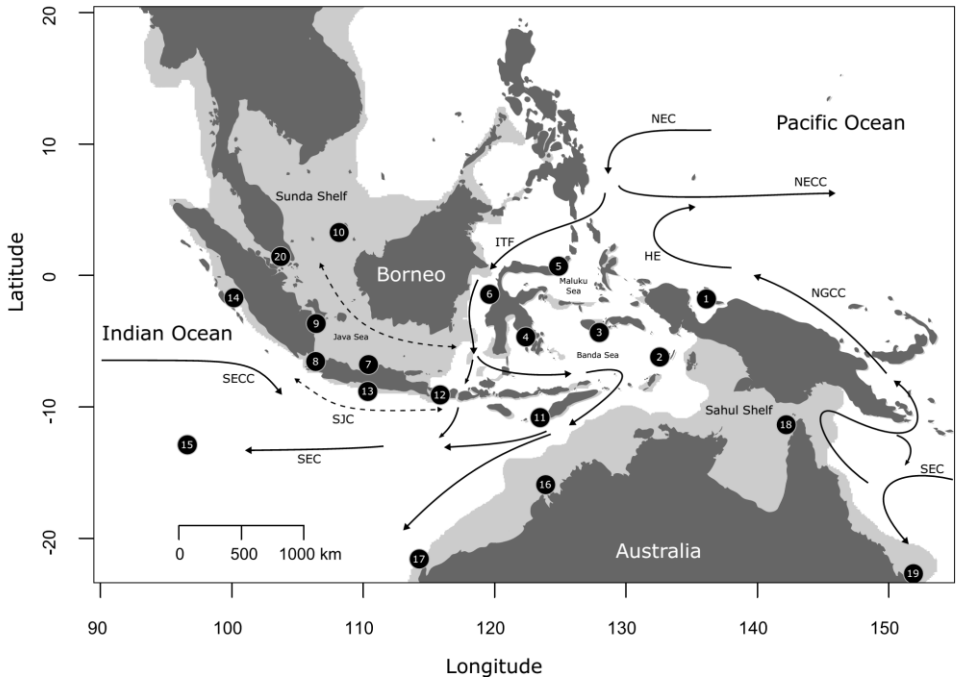

Figure 3.1. Map showing the seagrass *Thalassia hemprichii* sampling sites throughout the Indo-Australian Archipelago. Dark grey areas indicate present-day island configuration and light grey areas indicate landmass configuration during the Pleistocene low sea-level stands. Sampling site: 1. Biak; 2. Tual; 3. Ambon; 4. Kendari; 5. Bitung; 6. Palu; 7. Jepara; 8. Pari Is.; 9. Bangka; 10. Natuna, 11. Kupang; 12. Lombok; 13. Drini; 14. Padang; 15. Cocos Keeling, 16. Kimberley; 17. Exmouth; 18. Hammond Is.; 19. Magnetic Is.; 20. Semakau Is. Black arrowed lines indicate major ocean currents across the region and dashed arrows show seasonally reversing currents. HE: Halmahera Eddy; ITF: Indonesian Throughflow; SEC: South Equatorial Current; NGCC: New Guinea Coastal Current, NECC: North Equatorial Counter Current; NEC: North Equatorial Current; SJC: South Java Current; SECC: South Equatorial Counter Current.

To date, most studies examining the potential role of historical geological processes and contemporary physical oceanography in driving genetic structure in the IAA have been based on concordance between patterns of genetic structure and predicted patterns of population connectivity due to historical processes and oceanographic settings (Carpenter *et al.* 2011). The application of alternative approaches such as incorporating genetic data with oceanographic data and information on historical seascapes, combined with appropriate statistical analyses (e.g. redundancy analysis) can be very useful as it provides more information on how each predictor contributes to determining the observed genetic patterns (Meirmans 2015).

In the present study, the phylogeography of the seagrass *Thalassia hemprichii* in the IAA was examined using single nucleotide polymorphisms (SNPs) marker. SNPs marker was chosen because chloroplastidal-DNA (cpDNA) commonly used in angiosperms were rarely variable, especially in *Thalassia*. Beside the phylogeographic analysis, a more contemporary genetic structure and its relatedness to contemporary oceanography and seascape geological history was examined based on a panel of 16 microsatellite markers. Mutation frequency in SNPs are generally lower than microsatellites, thus SNPs can provide better insights into evolutionary history, while microsatellites provide better resolution to detect changes in allele frequency that occur over more recent divergence time (Haasl & Payseur 2011). Recent changes in allele frequency can also be detected by SNPs, however it would requires a random set of SNPs with many loci (>200 loci) compared to microsatellites (Liu *et al.* 2005; Haasl & Payseur 2011).

3.2 Materials and Methods

3.2.1 *Thalassia hemprichii*

The seagrass *T. hemprichii* has the potential for long distance dispersal due to its positively buoyant fruits that are able to float for up to 7 days, reaching a distance of 73 km (Lacap *et al.* 2002). Vegetative fragments of *T. hemprichii* could float for months and stil remains alive, potential to colonize new habitat (Wu *et al.* 2016). Although long distance dispersal is rare in *Thalassia* as most seeds are deposited within metres of the mother plants (e.g. van Dijk *et al.* 2009), the few dispersal events in which fruits are transported beyond the source population are significant in a population genetic context; one or a few successful migrants per generation are sufficient to maintain genetic connectivity between populations (Lowe & Allendorf 2010).

3.2.2 Study sites and sampling design

The spatial scale of this study extended from the western coast of Sumatra to the eastern coast of Australia (>1000 km, Figure 3.1). Populations in the middle of the

IAA approximately represent the centre of the seagrass distribution range, while Kimberley, Exmouth, Cocos Keeling, and Magnetic Island represent the southern limit of its distribution range (peripheral populations). During the Pleistocene low sea-level stands, marine populations are thought to have been fragmented into several groups of populations (Voris 2000; Carpenter *et al.* 2011; Collins 2011; Lohman *et al.* 2011; Figure A3.1, Appendix). The northern shores of Papua represent the Pacific Ocean populations. Several sea basins (e.g. Celebes Sea and Banda Sea) were likely to be isolated due to extended steep ridges and deep troughs, and shallow water habitats (Sunda Shelf, Sahul Shelf, and the north Western Australian continental shelf) were completely exposed.

Seventeen sites across this range were sampled for microsatellite analysis (Figure 3.1, excluding Site 18-20 due to logistical reasons) and sixteen sites for SNP analysis (Figure 3.1, excluding Site 3, 4, 6, and 11, due to logistical reasons). At each site, a total of 50 seagrass samples, separated by at least 2 m, were collected haphazardly over a 50 x 50 m area. Each sample consisted of two or three seagrass shoots connected by a rhizome (ramet). Epiphyte-free meristematic segments (2-4 cm) were excised from the shoots and immediately stored in plastic bags filled with silica gel for subsequent DNA extraction. DNA was extracted from 2-3 pieces (5-10 mm in size) of silica-dried samples using the AGRF extraction service (Australian Genomic Research Facility, www.agrf.org.au) employing the Nucleospin Plant II Kit (Machery-Nagel, Düren, Germany) with the PL2/PL3 buffer system.

3.2.3 SNP screening

Twelve samples were randomly selected from each of the sampling sites. SNP screening followed a developed method by Jardine *et al* (2015). This method combines complexity reduction using AFLP (amplified fragment length polymorphism) (Vos *et al.* 1995; van Orsouw *et al.* 2007) and then next generation sequencing technology. The SNP protocol was outlined in detail by Cross *et al.* (2016). In summary, it consisted of 5 main steps: (1) restriction/ligation, (2) pre-selective amplification, (3) selective amplification, (4) purification-size selection, and (5) SNP screening (Appendix A3.1). Next-generation eequencing was performed

on the Ion Torrent Proton (Life Technologies) at the Australian Cancer Research Foundation (ACRF), Cancer Genomics Facility in Adelaide.

Sequencing reads were analysed (de-multiplexing, trimming, and assembling) using CLC-Genomic Workbench (Qiagen, Venlo, The Netherlands) to generate a 'provisional reference genome' (Hird *et al.* 2011). Each individual's reads were mapped onto this reference and the consensus sequences were extracted. SNP loci were manually selected in Geneious R7 (Biomatters, Auckland, New Zealand) and identified 113 contigs with SNP loci. The selected contigs were trimmed, such that each contig contained a single SNP locus that was flanked with invariant sites of up to 15 bp length. Inclusion of invariant sites in the concatenation, rather than just SNP loci, may reduce acquisition bias and provide better accuracy (e.g. Leache *et al.* 2015). The contigs (including the invariant sites) were concatenated and the final SNP dataset consisted of 154 samples, each consisting of a 1435-bp nucleotide sequence. The final SNP dataset was used only for phylogeograhic analysis, because there were only 113 random SNP loci, which are not sufficient to detect allelic change at recent divergence times (Liu *et al.* 2005; Haasl & Payseur 2011).

3.2.4 Microsatellite amplification

Forty-eight samples were randomly selected from each of the sampling sites for microsatellite amplification on 16 previously developed polymorphic loci (Wainwright *et al.* 2013; van Dijk *et al.* 2014). Loci used included TH07, TH34, TH37, TH43, TH52, TH66, TH73, Thh1, Thh3, Thh5, Thh8, Thh15, Thh29, Thh34, Thh36, and Thh41. Fluorescently labelled primers were used for the amplification in three separate multiplex panels with the QIAGEN Type-it® microsatellite PCR Kit (10-μL reactions with ~1 ng of genomic DNA). Fragment analysis and capillary separation were run at GGF (Georgia Genomic Facility, USA) with GGF's size standard 500 ROX. Scoring of microsatellite alleles used the Microsatellite plugin v1.4 in GENEIOUS R7 v 7.1.7 (Biomatters Ltd).

3.2.5 Phylogeography using SNP

Since there are many models of nucleotide substitution for constructing phylogenetic trees, the best fitting model was determined using JMODELTEST 2.1.7 with three

substitution schemes including: all models with equal/unequal base frequencies (+F), with/without a proportion of invariable sites (+I), and with/without rate variation among sites (+G, 4 categories) (Darriba *et al.* 2012). Model selection was computed using the Akaike information criterion (AIC). The test showed that the best fitting model is HKY + G + I (Table A3.1, Appendix). The resulting model was implemented for constructing phylogenetic trees using a maximum likelihood approach in MEGA7.0.14 (Kumar *et al.* 2016). A bootstrap analysis of 100 replicates was run, each with the Nearest-Neighbour-Interchange branch swapping. The consensus tree was visualised using the online tool ITOL v3.0 (Letunic & Bork 2011; http://itol.embl.de/).

The degree of connectivity among phylogeographic clades was examined based on migration rate between multiple populations using MIGRATE-N 3.6.11 (Beerli & Felsenstein 2001). A Bayesian search strategy (Beerli 2006) was selected and the input parameters for this analysis were obtained from JMODELTEST (transition/tranversion ratio=7.2997; base frequency A=0.2650, C=0.2364, G=0.2331, T=0.2655; and gamma shape=1.7440). Samples were pooled based on the clades identified from the consensus tree generated in MEGA (Kumar *et al.* 2016).

3.2.6 Genetic structure using microsatellite

The presence of null alleles and scoring errors was investigated using ML-NULLFREQ with 100,000 randomizations (Kalinowski & Taper 2006). *Psex* values of MLG copies are mostly significant in each sampling sites, except Bangka, indicating that they were truly clones. For Bangka, most *psex* values were insignificant indicating that there may be more than one genet from the replicates of the same MLG (Appendix A3.2). However, as the allele composition was identical, we cannot differentiate between them and identify distinct genets from closely related MLGs. For further population genetic analysis, duplicates of multilocus genotypes (MLGs) were removed, resulting in a new dataset containing only unique MLGs using the package *poppr* in R (Kamvar *et al.* 2014). Estimate of linkage disequilibrium was based on the standardized index of association ($\bar{r}_d$, Agapow & Burt 2001) and calculated using the package *poppr* (Kamvar *et al.* 2014). Allelic richness (A_R) and private allele richness (P_A) standardised at 24 MLGs were calculated using HP-RARE

(Kalinowski 2005). Mean observed heterozygosity (H_O), mean unbiased expected heterozygosity (H_{NB}, Nei 1978) and inbreeding coefficient (F_{IS}) were calculated using GENETIX 4.05 (Belkhir *et al.* 2004). Genetic differentiation was examined based on F_{ST}, obtained from GENALEX v6.501 (Peakall & Smouse 2012). Population structure inferred from the number of distinct genetic clusters (*K*) among the populations was examined using a Bayesian assignment test in STRUCTURE v2.3.4 (Pritchard *et al.* 2000). Values of K from 1 to 17 were tested with 20 iterations for each K value and a burn-in of 10^5 and 10^6 replications. Determining the "true" *K* was based on Evanno *et al.* (2005). The Web server CLUMPAK was used to align multiple replicate analyses of the appropriate population clustering, and visualized the population structure (Kopelman *et al.* 2015). Based on the spatial pattern of population clustering from STRUCTURE, the degree of connectivity among clusters of populations was estimated in MIGRATE-N 3.6.11 (Beerli & Felsenstein 2001) with a Bayesian search strategy (Beerli 2006).

3.2.7 Oceanographic connectivity

Particle dispersal simulation was used to obtain a pairwise matrix of contemporary oceanographic connectivity among sampling sites. Daily surface velocity outputs from the Bluelink ReANalysis (BRAN) version 3p5 were used (Oke *et al.* 2013). BRAN is a data-assimilating hydrodynamic model product based on a 10-km horizontal resolution and 5-10 m vertical resolution (in the upper 300 m). BRAN assimilates along-track sea-level anomalies from satellite altimeters and sea level data from tidal gauges, satellite sea surface temperature, *in situ* temperature and salinity from Argo profiles and the Tropical Atmosphere Ocean (TAO) array (Oke *et al.* 2013). Surface current velocities from BRAN were used to drive a passive particle tracking model and a 4th-order Runga-Kutta sub-time-stepping scheme was used to update the particle locations (Feng *et al.* 2010). The model was demonstrated to reproduce key ocean current features in the IAA (He *et al.* 2015). A total of 1000 particles were released at each sampling site on the first day of the each month and were tracked over a one-year period. The random walk effect was considered by including a diffusivity of 3 m^2s^{-1} (Feng *et al.* 2010). A pairwise matrix of oceanographic connectivity was constructed from the source-sink relationship

defined as the number of days that particles released from the source site spend within a 30 x 30 km area around the destination site. The particle tracking model was run for 11 years, 1993-2003 to provide a range of different climate conditions and ocean circulation variations that could influence connectivity.

3.2.8 Relative roles of palaeo-historical processes and contemporary oceanographic connectivity in gene flow

To test the influence of the Pleistocene low sea-level stands on population differentiation in the SNP dataset, we performed Analysis of Molecular Variance (AMOVA) using ARLEQUIN 3.5 (Excoffier & Lischer 2010) with samples grouped according to the clades identified from the consensus tree generated in MEGA (Kumar *et al.* 2016). As we used SNPs to detect genetic changes over long evolutionary time scales, we did not examine the effect of contemporary oceanographic connectivity on the SNP dataset. For the microsatellite dataset, we also performed AMOVA with samples grouped according to the 5 clusters identified in STRUCTURE analysis.

Distance-based redundancy analysis (dbRDA) was used to determine the relative contribution of palaeo-historical processes (HP) and contemporary oceanographic connectivity (OC) in driving the patterns of a more recent genetic structure. The dbRDA performed a multivariate regression on the response variable, the pairwise genetic distance matrix (here, pairwise F_{ST} from microsatellite data) (Legendre & Anderson 1999). The variable HP is categorical and represents distinct biogeographic regions based on populations that were predicted to be isolated during the Pleistocene sea-level fluctuation (Voris 2000; Carpenter *et al.* 2011; Lohman *et al.* 2011; Figure A3.1, Appendix). For the variable OC, the 'raw' data of the oceanographic connectivity matrix was transformed into a weighted, directed network based on graph theory using the *igraph* package in R (Csardi & Nepusz 2015). Four network parameters were calculated: (i) *strength* -the total amount of connection (in and out) of a site (higher *strength* indicates higher degree of connectivity), (ii) *closeness* -the extent to which a site is connected to other sites, (iii) *betweenness* -the number of shortest connections between two sites that go through the site of interest, and (iv) *transitivity*, defined as the extent to which the adjacent

sites of a site are connected to each other (Table A3.2, Appendix). The calculation of *closeness* and *betweenness* treats the connection weights as 'cost', rather than a true 'connection strength', thus it represents the cost needed to connect nodes (higher *closeness* and *betweenness* indicates a higher degree of isolation) (Barrat *et al.* 2004; Csardi & Nepusz 2015). The values of these network parameters were scaled and normalized so that they have the same range and can be used as a predictor in the redundancy analysis. Finally, the dbRDA (F_{ST} ~ HP + OC) was performed with forward selection using the package *vegan* in R (Oksanen *et al.* 2015). Beside palaeo-historical processes and contemporary oceanographic connectivity in gene flow, spatial distance may also influence gene flow, and this was examined using Mantel test based on isolation by distance model (IBD).

3.3 Results

3.3.1 Phylogeography using SNP

A total of 154 samples from 16 sampling sites across the IAA were sequenced, with 113 SNP loci identified. The contigs were concatenated into 1435-bp length nucleotide sequences. The consensus tree from the maximum likelihood approach showed three distinct phylogenetic clades: (1) East clade, (2) Sunda clade and (3) Indian Ocean clade (Figure 3.2). The clustering pattern is concordant with the geographical positions of the sites. Sampling sites in the eastern part of Indonesia and Australia (Biak-1, Tual-2, Hammond Island-18 and Magnetic Island-19) form the East clade. Sampling sites in the western part of Indonesia grouped into the Sunda clade (Jepara-7, Pari-8, Bangka-9, Natuna-10, Lombok-12, and Drini-13), while the Indian Ocean clade consisted of Padang-14, Cocos Keeling-15 and two sites from the western coast of Australia (Kimberley-16 and Exmouth-17).

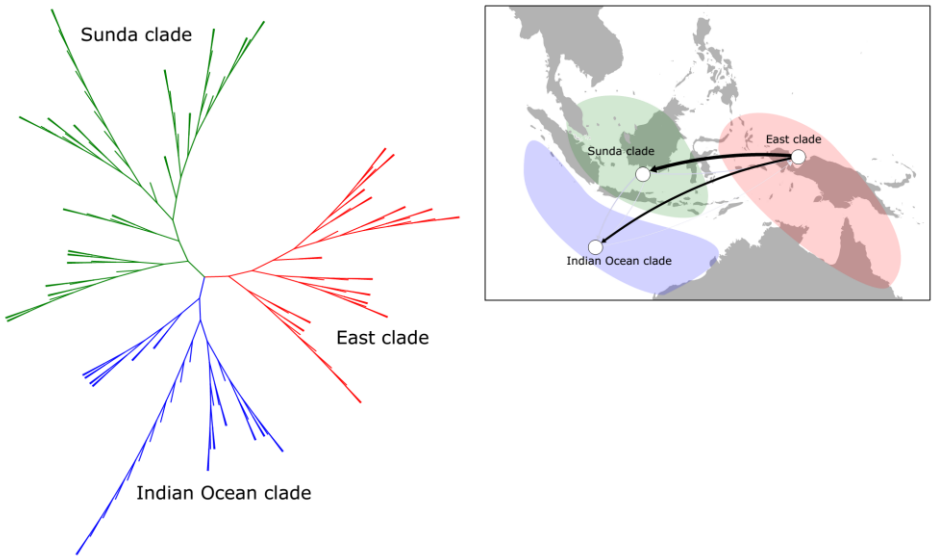

Figure 3.2. Bootstrap consensus tree constructed from SNP dataset using MEGA7 (Kumar *et al.* 2016) and estimates of migration calculated from MIGRATE-N (Beerli & Felsenstein 2001). The bootstrap consensus tree inferred from 100 replicates. Initial tree(s) for the heuristic search were obtained automatically by applying the Maximum Parsimony method. Black arrowed lines represent directional migration among groups of populations. Thickness of the lines are scaled to the mean numbers of migrants per generation.

Estimates of migration from MIGRATE analysis indicated asymmetrical gene flow among these populations with a trend of westward migration from the East clade (Figure 3.2). The greatest numbers of migrants were sourced from the East clade, moving to the Sunda clade (79.88) and to the Indian Ocean clade (50.83). In contrast, much smaller migration rates were found in the reverse direction (Table A3.3, Appendix).

3.3.2 Genetic structure using microsatellites

A total of 132 alleles were genotyped across 640 MLGs (17 sampling sites) based on 16 microsatellite loci. Genotyping and scoring errors were not detected in ML-NULLFREQ (estimate of genotyping error β <0.001). ML-NULLFREQ indicated the presence of null alleles in all loci, however the average frequency across populations was very low, ranging from 0.008 (Thh3) to 0.058 (TH66 and Thh15), thus all loci were retained for further analysis. The standardized index of association ($\bar{r}_d$) was

also very low (0.04, p-value=0.001), indicating a low probability of association between loci (Agapow & Burt 2001).

Genetic diversity was generally higher in the eastern part of Indonesia and decreasing in a southwestward direction (Figure 3.3). The highest total number of observed alleles, allelic richness, and unbiased expected heterozygosity was found in Biak-1, while the lowest was found in Bangka-9 and Exmouth-17. Private alleles were observed in all sites, except Bangka-9, Cocos Keeling-15 and Exmouth-17 (Table 3.1).

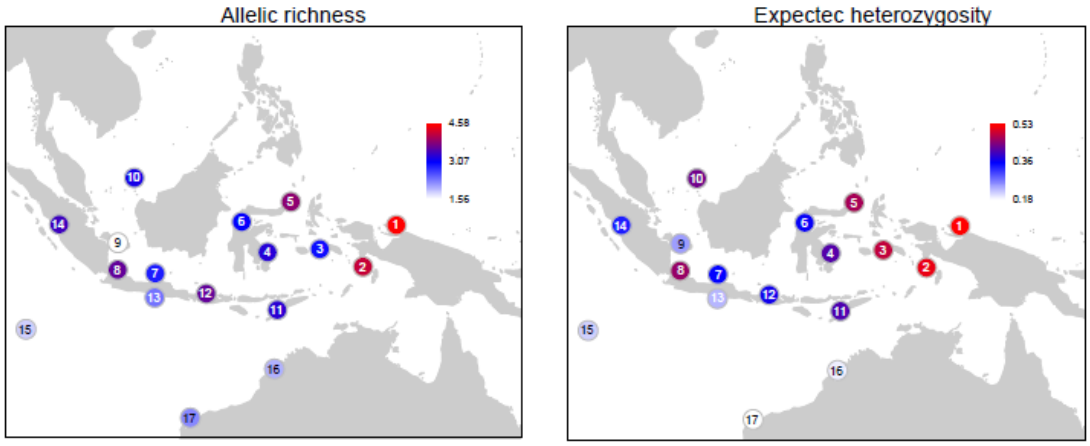

Figure 3.3. Genetic diversity of the seagrass *Thalassia hemprichii* in the Indo-Australian Archipelago calculated from microsatellite data. Allelic richness (left-hand panel); and unbiased expected heterozygosity (right-hand panel). Sampling site: 1. Biak; 2. Tual; 3. Ambon; 4. Kendari; 5. Bitung; 6. Palu; 7. Jepara; 8. Pari Is.; 9. Bangka; 10. Natuna, 11. Kupang; 12. Lombok; 13. Drini; 14. Padang; 15. Cocos Keeling; 16. Kimberley; and 17. Exmouth.

Table 3.1. Allelic richness and genetic diversity of *Thalassia hemprichii* in the Indo-Australian Archipelago obtained from 16 microsatellite loci with standardized of 24 MLGs. *N*: total number of samples examined; *G*: number of multilocus genotype; *R*: clonal richness (MLG-1/N-1); *nA*: observed alleles; A_R: allelic richness and P_A: private allele richness, standardized at 24 samples (average alleles per loci); H_O: observed heterozygosity; H_{NB}: non-biased heterozygosity-Nei, 1978. $^{++}$Site Bangka was not standardized to 24 MLGs for A_R and P_A, as there were only 5 MLGs found. Significance level, *=p<0.05, **=p<0.01

ID	Site	N	G	R	*nA*	A_R	P_A	H_{NB}	H_O	F_{IS}
Pacific Ocean										
1	Biak	48	48	1.00	85	4.58	0.22	0.534	0.462	0.136**
Banda Sea										
2	Tual	48	47	0.98	73	4.23	0.08	0.518	0.523	-0.010*
3	Ambon	48	24	0.49	48	3	0.01	0.489	0.552	-0.133**
4	Kendari	48	38	0.79	56	3.29	0.01	0.416	0.467	-0.126**
Celebes Sea										
5	Bitung	48	48	1.00	69	3.89	0.16	0.472	0.456	0.035*
6	Palu	48	39	0.81	53	3.08	0.03	0.363	0.389	-0.073**
Sunda Shelf										
7	Jepara	48	29	0.60	47	2.9	0.14	0.358	0.353	0.011**
8	Pari Is.	48	36	0.74	61	3.65	0.08	0.462	0.431	0.069**
9	Bangka^{++}	48	5	0.09	25	1.56	0	0.253	0.413	-0.772**
10	Natuna	48	37	0.77	53	3.2	0.07	0.437	0.419	0.042*
Lesser Sunda										
11	Kupang	48	43	0.89	56	3.32	0.13	0.415	0.430	-0.037*
12	Lombok	48	44	0.91	64	3.67	0.11	0.372	0.287	0.231**
Indian Ocean-onshore										
13	Drini	48	36	0.74	41	2.4	0.12	0.234	0.201	0.142**
14	Padang	48	40	0.83	60	3.45	0.17	0.339	0.258	0.241**
Indian Ocean-offshore										
15	Cocos Keeling	48	43	0.89	32	1.85	0	0.218	0.240	-0.102**
Western Australia										
16	Kimberley	48	44	0.91	36	2.02	0.02	0.196	0.215	-0.097*
17	Exmouth	48	39	0.81	38	2.26	0	0.184	0.188	-0.017

Strong overall genetic differentiation was observed among the seagrass populations (global F_{ST}=0.353, p=0.001). All pairwise differentiations between sampling sites were significant (p<0.01). Bayesian probability assignment in STRUCTURE revealed a spatial pattern of genetic structuring (Figure 3.4). Model evaluation with the delta *K* method (Evanno *et al.* 2005) supported two to four genetically divergent populations. At *K*=2, individuals sampled from Australia and the eastern part of Indonesia were assigned uniformly to one cluster (blue), while the other cluster (orange) consisted of most samples from the western part of Indonesia (except Padang) (Figure 3.4). At *K*=3, the blue cluster was then divided into two groups of populations (dark blue and blue). At *K*=4, as individuals from some populations showed mixed identity from two clusters, five genetically distinct regions was identified. The first region consisted of samples from the eastern part of Indonesia (1-6; green), the second from Sunda Shelf (7-10; dominant orange admixed with dark blue), the third from Lesser Sunda (11, 12, 14; dominant dark blue admixed with orange), the fourth from the south coast of Java and Cocos Keeling (13 and 15; orange), and the fifth from Australia (16-17; blue).

Based on these five genetic regions, MIGRATE analysis using microsatellite data indicated asymmetrical gene flow among groups of populations with a trend of westward, southward and southwestward migration, with eastern Indonesia as the main source population (Figure 3.4). This migration pattern is similar to that based on the SNP data. The greatest numbers of migrants were from the eastern part of Indonesia, moving to the other groups of populations (to Sunda Shelf=36.71, to Lesser Sunda=32.37, to Indian Ocean=26.66, and to Western Australia=25.60). Like the SNPs, smaller migration rates were found in the reverse direction (Table A3.4, Appendix).

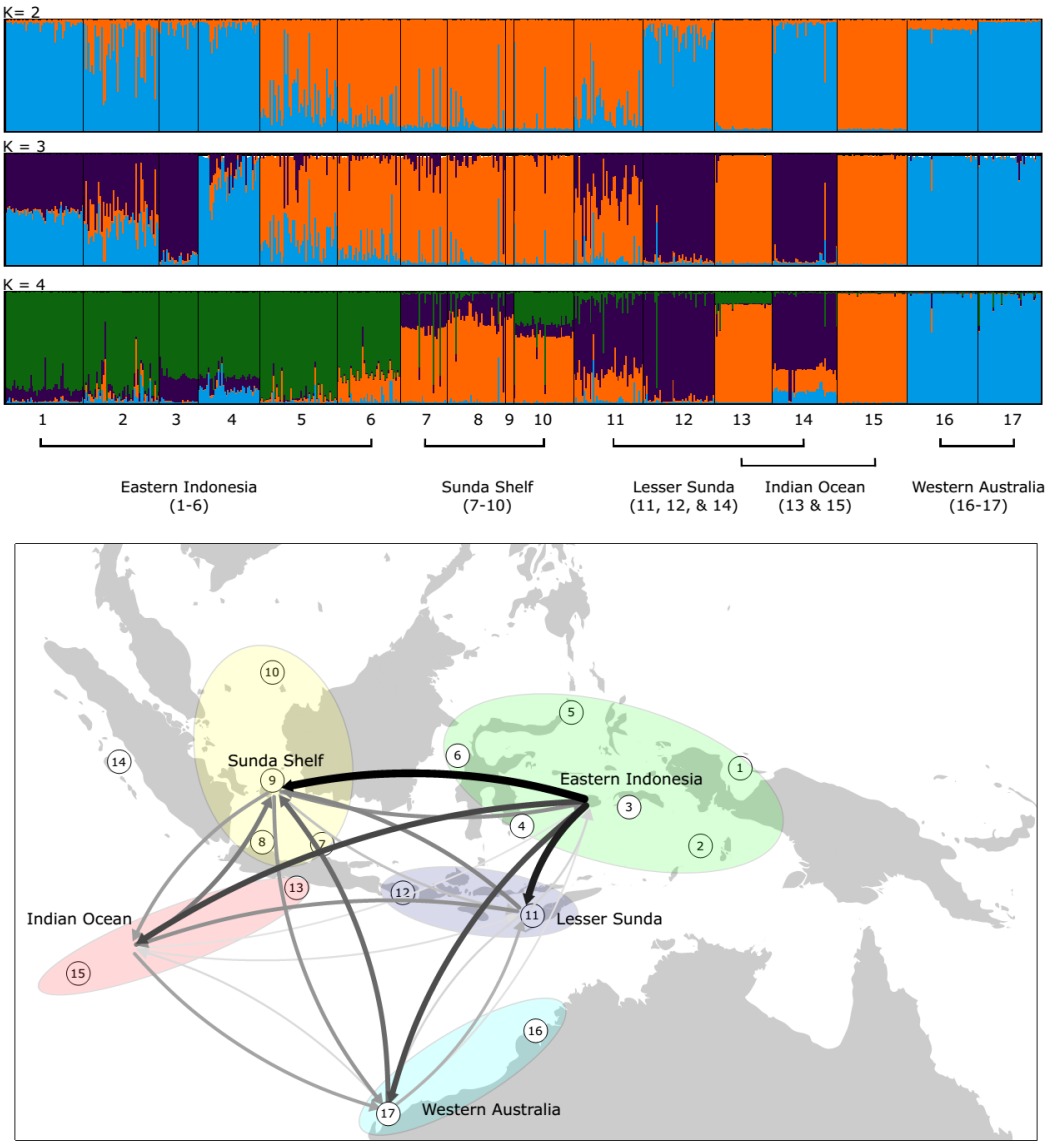

Figure 3.4. Bayesian assignment of individuals calculated in STRUCTURE, at *K*=2 to *K*=4 (Pritchard *et al.* 2000; top panel) and migration estimates based on microsatellite data calculated in MIGRATE (Beerli and Felsenstein, 2001; bottom panel). Samples were pooled based on five clusters of populations resulted from the Bayesian assignment at *K*=4. Sampling site: 1. Biak; 2. Tual; 3. Ambon; 4. Kendari; 5. Bitung; 6. Palu; 7. Jepara; 8. Pari Is.; 9. Bangka; 10. Natuna, 11. Kupang; 12. Lombok; 13. Drini; 14. Padang; 15. Cocos Keeling, 16. Kimberley; and 17. Exmouth

3.3.3 The role of historical processes and oceanography

AMOVA on the SNP dataset showed that the variation was mostly observed among sites within group (55.97%), while variation among groups of sampling sites was 31.25%. On the microsatellite dataset, the variation was mostly within sites

(65.58%), while variation among groups was 9.04 % (Table A3.5. Appendix). For more recent genetic divergence, the db-RDA analysis showed that palaeo-historical processes explained higher variation (20.3%, p<0.001) than contemporary oceanographic connectivity (8.4%), but the oceanographic effect was also significant (p=0.008, Table A3.6, Appendix). In addition, Mantel test on pairwise F_{ST} and geographic distance (km) was not significant (Mantel r=0.166, p=0.129) indicating that spatial distance may not have significant role in gene flow.

3.4. Discussion

3.4.1 Phylogeography

Phylogeographic analysis revealed regional patterns of genetic isolation with 3 distinct regions: eastern Indonesia (East clade), western Indonesia (Sunda Shelf clade), and the Indian Ocean. The broad pattern of genetic separation, with one in the Pacific (eastern Indonesia), the other in the Indian Ocean, and the transition of the Pacific and Indian Oceans (western Indonesia), broadly support similar patterns observed in populations of some marine invertebrates (Williams & Benzie 1998; Barber *et al.* 2006; de Boer *et al.* 2014) and several fish species (Borsa 2003; Timm & Kochzius 2008; Drew & Barber 2009; Ackiss *et al.* 2013). The concordance of phylogeographic patterns among different taxa at the broad scale suggests that common physical/environmental forces might have acted on marine population connectivity in the region. The significant amount of variation explained in the AMOVA analysis supports the hypothesis that vicariance among the Pacific and Indian Ocean basins during the Pleistocene low sea level contributed to genetic divergence between populations in the Pacific and Indian Oceans.

During periods of low sea level in the Pleistocene, the emergence of the Sunda and Sahul Shelves seems to have acted as a major barrier to connectivity between populations of *T. hemprichii* in the Pacific and Indian Oceans (Voris 2000; Lohman *et al.* 2011), and the finding suggests that this historical event might have fragmented the seagrass metapopulation in the IAA. When the sea level rose, ocean currents allowed seagrass to disperse and colonize available habitats. The asymmetrical

pattern of migration (Figure 3.2) is congruent with the major paleo- and contemporary ocean currents flowing from the Pacific Ocean to the Indian Ocean. This study also demonstrates that eastern Indonesia was, and still is, an important source of migrants that colonize marine habitats on the Sunda Shelf, once a dry and exposed land mass during the periods of low sea levels in the Pleistocene (Figure 3.2).

However, since Pleistocene vicariance explained a relatively small amount of the variation in AMOVA analysis (31.25%), it may not be the only factor driving genetic divergence. Microevolutionary forces that operate within populations such as genetic drift, selection, and mutation contribute to shape intraspecific genetic pattern. Demographic situations, such population size and bottleneck, could also lead to genetic divergence (Slatkin 1987; Charlesworth *et al.* 2003).

Although Indonesia has generally been recognised as a region of strong genetic breaks between the Pacific and Indian Oceans, the detailed genetic patterns can vary among species, often showing complex spatial structure with more than two distinct groups across Indonesia and the location of genetic breaks varying among species (Carpenter *et al.* 2011). In this study, a genetic break was observed in the Banda and Maluku Seas, separating the Eastern clade with the Sunda Shelf clade (Figure 3.2). This pattern matches those observed in the red bellied fusilier *Caesio cuning* (Ackiss *et al.* 2013) and the stomatopod *Haptosquilla glyptocerus* (Barber *et al.* 2006). In contrast, no genetic break was observed for *T. hemprichii* around the Bird's Head region of Papua, which has been identified as a genetic break for the the giant clam *Tridacna maxima* (Nuryanto & Kochzius 2009). These complex and varying regional substructures are likely attributed to historical distribution of species, their rates of colonisation and the species' biology interacting with the physical environment (Bradbury *et al.* 2008; Galarza *et al.* 2009; Selkoe *et al.* 2014).

The genetic break in the seagrass population between eastern and western Indonesia at Banda and Maluku Seas, a biogeographic region known as Wallacea, provides more evidence that Wallacea is not only a dispersal barrier for terrestrial species, but also for marine species. Although water bodies still existed in the Banda and Maluku ocean basins during the Pleistocene low sea-level stands, connectivity between

eastern and western Indonesia and across the ocean basins might have been restricted due to the exposed land mass (Voris 2000). The Wallace region, which is marked with several biogeographic delineations (i.e. Wallace's Line, Huxley's Line, Lydekker's Line and Weber's Line), was earlier recognised as a separation zone for terrestrial species between Southeast Asia and New Guinea-Australia (Mayr 1944; Raes & van Welzen 2009). Later, studies have shown that some marine species, for example stomatopods (Barber *et al.* 2000, 2006) and seahorses (Lourie & Vincent 2004), also followed the separation pattern (genetic divergences on the Wallace region).

3.4.2 Genetic structure at more recent divergence time

In more contemporary times the patterns of regional divergence are overlain with patterns of contemporary isolation, combined with LLD events, resulting in shared alleles and unique alleles across the region. Samples from Lesser Sunda and Sunda Shelf were separated into different clusters (Figure 3.4). The Sunda Shelf-Lesser Sunda divergence might be a signature of post-Last Glacial Maximum (LGM) colonization in the Sunda Shelf. During the LGM, the Sunda Shelf was dry land (Voris 2000) and when the sea level rose, individuals from the Lesser Sunda might have colonised the available habitat in the Sunda Shelf (Java Sea). This hypothesis is supported by the east-to-west migration pattern, showing more migrants moving from the Lesser Sunda (and eastern Indonesia) to the Sunda Shelf (Figure 3.4; Table A3.4, Appendix) than in reverse direction. Interestingly, this asymmetry is not present in the contemporary ocean currents. Seasonal surface currents flow westward in the Java Sea during the southeast monsoon (June-September) and eastward during the northwest monsoon (November-February) (Wyrtki 1961; Figure 3.1), while *T. hemprichii* was reported to flower and fruit throughout the year, especially in the east side of Sunda Shelf (Verheij & Erftemeijer 1993). There may be influence of surface wind drag (windage) that contributes to the migration pattern as windage can significantly determine seagrass dispersal direction and distance (Ruiz-Montoya & Lowe 2012; Ruiz-Montoya *et al.* 2015). However, monsoon wind also seasonally reverses its directions, blowing westward during the southeast monsoon and eastward during the northwest monsoon over the Java Sea (Wyrtki 1961).

Besides the Sunda Shelf-Lesser Sunda divergence, the scenario of post-LGM colonization might also explain the divergence among the Western Australian sites (Kimberley-16 and Exmouth-17), Cocos Keeling-15 and Padang-14 (Figure 3.4; top panel). As this divergence was observed in microsatellite but absent in SNP, it is likely to have occured in more recent times (post-LGM). During the LGM, the wide continental shelf of north Western Australia was emerged land (Collins 2011), and the migration pattern (Figure 3.4; bottom panel) indicated that when the sea level rose, seagrass migrants from either the north or the west might have colonised available inshore habitat in the region.

In the Indian Ocean, one might predict that samples from Padang-14 would form a cluster with other samples from Drini-13 and Cocos Keeling-15 (the Indian Ocean cluster), however the Bayesian assignment analysis showed Padang-14 is more related to Lesser Sunda populations than to Drini-13 and Cocos Keeling-15. Ocean currents may partly explain this counterintuitive pattern. The surface current off Java, called the South Java Current (SJC) (Figure 3.1), may facilitate connectivity among marine benthic populations in the south-western coast of Sumatera (e.g. Padang), southern coast of Java (e.g. Drini), and southern coasts of islands in Lesser Sunda (Quadfasel & Cresswell 1992), while the South Equatorial Current (SEC) on the Indian Ocean side might result in Drini being genetically more connected to Cocos Keeling than to other populations. Further studies with more sampling sites along the southern coast of Java and western coast of Sumatera are needed to clarify this.

Although the microsatellite analysis provides evidence of more recent microevolutionary processes (post-LGM) generating genetic divergence in the Indo-Australian Archipelago, the genetic imprint of more ancient processes (Pleistocene sea-level fluctuations) is still strong as indicated by the dbRDA analysis. The smaller imprint of contemporary processes in the microsatellite dataset might be due to the time lag between the occurrence of a driving force (e.g. water current) and the point at which the genetic response becomes detectable (close to or reaching equilibrium), such that recent changes in population structure, size or connectivity may not be immediately reflected in genetic metrics (Epps & Keyghobadi 2015). The time lag occurs because it requires multiple generations for allelic changes to reach

equilibrium (Bolliger *et al.* 2014). Seagrass can have extremely long life span, potentially millennia (Arnaud-Haond *et al.* 2012), increasing the likelihood of long time lags (Epps & Keyghobadi 2015).

3.4.3 Genetic diversity

The highest genetic diversity based on microsatellite data was observed in the heart of the Coral Triangle (eastern Indonesia) reaffirming the status of this region as the hotspot of tropical marine biodiversity. The Coral Triangle was first recognised for its extreme marine biodiversity at a species level (Hoeksema 2007), and the finding supports this from the perspective of intraspecific genetic diversity. Maximum genetic diversity in the Coral Triangle has been previously documented in the pearl oyster *Pinctada maxima* (Lind *et al.* 2007) and other marine species (reviewed in Gaither & Rocha 2013). Two hypotheses may explain the high genetic diversity of seagrass populations in the region, the Centre of Origin hypothesis or the Centre of Accumulation hypothesis (Bowen *et al.* 2013). Under the Centre of Origin hypothesis, the ancestral populations with highest genetic diversity should be in the centre of the hotspot, and subsequent dispersal radiates from the centre towards the peripheral populations. The asymmetrical migration pattern (Figure 3.4; bottom panel), showing most migrants were sourced from eastern Indonesia to the other regions, supports this hypothesis. However, since the seagrass populations on the eastern and northern side of the Coral Triangle (e.g., Philippines and Micronesia) have not been sampled, it is not possible to conclude that the ancestral populations of the seagrass would be in the hotspot. Alternatively, the Centre of Accumulation suggests that the ancestral populations are not in the core region of the Coral Triangle, but in the peripheral regions and extend their ranges into the Coral Triangle, presumably by ocean currents. Considering the migration pattern (Figure 3.4), the ancestral populations might reside somewhere in the Western Pacific, and the Coral Triangle islands would then act as stepping stones for migration towards the Indian Ocean. To examine these hypotheses more samples from the Western Pacific islands need to be examined.

The seagrass populations in the Kimberley-16, Exmouth-17 and Cocos Keeling-15 showed lower levels of genetic diversity compared to those from the eastern part of

Indonesia (e.g. Tual-3). The lower genetic diversity is likely due to these populations being at the southern periphery of the species' distribution range. Based on the central-marginal hypothesis-CMH, peripheral populations are predicted to have lower genetic diversity compared to those in the middle of the distribution range (Eckert *et al.* 2008). The CMH suggests that effective population size and gene flow should be depleted at range margins. In addition, peripheral populations are generally assumed to be subjected to greater selection pressures from less favourable biotic and abiotic environments, thus only a small number of genotypes survive (Hoffmann & Blows 1994). Lower genetic diversity in peripheral populations has been observed in other tropical seagrass species, e.g. *Enhalus acoroides* (Nakajima *et al.* 2014) and *Cymodocea serrulata* (Arriesgado *et al.* 2015) and temperate species e.g. *Zostera marine* (Diekmann & Serrão 2012). To determine whether the CMH truly explain the pattern of genetic diversity in *T. hemprichii,* future population genetic surveys should cover the seagrass entire distribution range, include population at other extremes of the distribution.

In conclusion, this study reveals that the pattern of genetic structure and connectivity of the seagrass *T. hemprichii* in the IAA is attributed partly to the Pleistocene sea level fluctuations modified to a smaller level by contemporary ocean currents. While the isolating/facilitating effect of contemporary processes such as water currents in seagrass connectivity is evident in experimental studies (e.g. Ruiz-Montoya *et al.* 2015), the analysis showed that the signature of this contemporary process is less strongly reflected in the genetic pattern. This contrasting finding may be due to the time lag between a perturbation occurring and the genetic response become detectable (Epps & Keyghobadi 2015), thus highlighting the importance of investigating the role of time lags in seascape genetics, particularly in seagrass populations. This study also revealed that the megadiverse IAA is not only a hotspot of marine species diversity but also intraspecific genetic diversity. The seagrass lineage divergence and migration pattern is consistent with both the Centre of Origin and the Centre of Accumulation hypotheses.

Genetic patterns revealed in this study provide information useful for management of seagrass in the IAA. Five genetically distinct regions (eastern Indonesia, Sunda

Shelf, Lesser Sunda, Western Australia, and Indian Ocean) were identified and should be managed as independent conservation units. As the genetic variability in Eastern Indonesia was the greatest and also the source of migrants for other populations, this area should be considered of high importance for conservation. Marine habitats of the Sunda Shelf populations, relative to the rest of the IAA have the lowest genetic diversity and are among the most over-exploited and degraded ecosystems in the world with high levels of anthropogenic disturbance (Vo *et al.* 2013). Therefore these are the most vulnerable and conservation programmes should also prioritize the management of these populations in the Sunda Shelf.

Chapter 4

CONTEMPORARY GENETIC STRUCTURE OF THE SEAGRASS *THALASSIA HEMPRICHII* AT A LOCAL SCALE

Abstract

Understanding patterns of gene flow and processes driving genetic differentiation is highly informative for a broad range of conservation practices. In marine organisms, genetic differentiation among populations is influenced by a range of spatial, oceanographic and environmental factors that are attributed to the seascape. The relative influences of these factors may vary in different locations and can be measured using seascape genetic approaches. In this study, a seascape genetic approach was applied to populations of the seagrass *Thalassia hemprichii* at a fine spatial scale (less than 100 km) in the Kimberley coast, Western Australia, a highly complex seascape with strong, multidirectional currents greatly influenced by extreme tidal ranges (up to 11m, the world's largest tropical tides). The seascape approach incorporated genetic data from a panel of 16 microsatellite markers, over-water distance, oceanographic data derived from predicted passive dispersal on a 2 km-resolution hydrodynamic model, and habitat characteristics from each meadow sampled. Significant spatial genetic structuring and asymmetric gene flow were detected, in which meadows 12-14 km apart were less connected than ones 30-50 km apart. This pattern was significantly explained by oceanographic connectivity and differences in habitat characteristics, suggesting a combined scenario of dispersal limitation and facilitation by ocean current with local adaptation. This finding adds to the growing evidence for the key role of seascape attributes in driving spatial patterns of gene flow. Despite the seagrass potential for long distance dispersal, there was

significant genetic structuring over small spatial scales highlighting the importance of implementing local-scale conservation and management measures.

4.1 Introduction

The extent of gene flow among populations influences how available genetic diversity is distributed in space, also referred to as genetic structure. When gene flow is absent, populations will differentiate over time due to drift, inbreeding and local selection/adaptation, which can lead to allopatric speciation (Hey 2006; Nosil 2008). A minimum level of gene flow maintains genetic cohesiveness of species and reduces the deleterious effect of inbreeding and genetic drift. Furthermore, gene flow may spread species' adaptive potential throughout populations as species evolve to adapt to environmental change (Frankham 2005; Charlesworth & Willis 2009; Lowe & Allendorf 2010). Considering the role of gene flow for population persistence, understanding genetic structure and differentiation can be highly informative for a broad range of conservation practices, for example identification of units to conserve, spatial boundaries of the units, and source/sink populations (Frankham *et al.* 2010; Beger *et al.* 2014; Magris *et al.* 2014).

As indicated in Chapter 2 and 3, patterns of genetic differentiation between populations in marine organisms are driven by the interaction of species biology (e.g. the capacity of the organism to disperse and recruit) and factors attributed in the seascape over time, including physical drivers like currents and ocean barriers, geographic distance between populations, and availability of environments to colonise. Traditionally, isolation by distance (IBD) has been the most simple model used to explain genetic differentiation, whereby gene flow decreases with distance from the original population (Wright 1943). However, in the marine environment, genetic differentiation often cannot be explained simply as a function of geographic distance (Riginos & Liggins 2013). As dispersal of most marine organisms can be either facilitated or limited by oceanographic features (such as currents) regardless of the geographic distances, patterns of genetic differentiation often follow a function of resistance to gene flow, known as 'isolation by resistance' (IBR; Thomas *et al.*

2015), or ‘isolation by oceanographic distance’ (IBOD; Alberto *et al.* 2011). Alternatively, local habitat characteristics and ecological processes (e.g. competition) can also limit gene flow by preventing successful settlement and recruitment (Ranta *et al.* 2009; Marshall *et al.* 2010; Orsini *et al.* 2013). Despite high levels of migration, a population may still be genetically isolated, if settlement is restricted or post-recruitment mortality is high. In this case, genetic differentiation follows a model of ‘isolation by environment’/IBE, whereby gene flow is affected by differential environmental conditions and ecological processes between populations (Wang & Bradburd 2014).

Those models are not mutually exclusive because in natural situations gene flow is often governed by a combination of scenarios described above (e.g. Sjöqvist *et al.* 2015). Furthermore, the relative roles of each component vary in different systems and can be measured using seascape genetics (e.g. Selkoe *et al.* 2010; Giles *et al.* 2015). A seascape genetics approach incorporates species biological traits, oceanography, habitat characteristics and other types of data into population genetic analyses to understand processes determining gene flow in the marine environment (Riginos & Liggins 2013).

The Kimberley coast on the Australian North West Shelf is rich in biodiversity, yet the least scientifically studied in Australia (Wilson 2013). It is located within one of the least human-impacted regions in the world (Halpern *et al.* 2008). The coast is a highly complex seascape with thousands of islands subjected to an extreme tidal range, up to 11m, the world’s largest tropical tides (Wilson 2013). The local currents are heavily influenced by tide, and override the broader scale, outer continental shelf currents (Condie & Andrewartha 2008). Currents around the islands are multidirectional and can exceed 1 ms^{-1}, producing spectacular ocean conditions including whirlpools and extreme standing waves (Cresswell & Badcock 2000; Wilson 2013; Lowe *et al.* 2015).

The growing interest in industrial use of the Kimberley marine environment and the lack of even a basic baseline understanding of important processes, such as population connectivity, urge scientific investigations on dispersal and gene flow. In this region, the influence of ocean currents on gene flow has been examined at a

large scale (100s – 1000s km) between oceanic coral reefs on the shelf margin (Condie & Andrewartha 2008; Underwood 2009; Underwood *et al.* 2012) where southward gene flow has been detected, aligned with the dominant southward moving Holloway Current (Wilson 2013). To date, no studies have been published on dispersal and gene flow at fine spatial scales (less than 100 km) along the coastal margin of the shelf that are primarily influenced by tidal currents. One could expect that strong, multidirectional tidal flow would facilitate connectivity, thus homogenizing the spatial distribution of genetic variation. However, the seascape complexity could also result in isolation that might lead to genetic differentiation.

The spatial pattern of gene flow, which can vary widely, have been well documented in some seagrass populations (e.g. Alberto *et al.* 2006; van Dijk *et al.* 2009; Olsen *et al.* 2013), however how external factors drive the genetic pattern in a seagrass population is still unclear. Here, this study examined (i) the pattern of genetic differentiation and structure among populations of the tropical seagrass *Thalassia hemprichii* (Ehrenberg) Ascherson, 1871 and (ii) how spatial, oceanographic and environmental factors influence the pattern. The study incorporated genetic data from a panel of 16 microsatellite markers, a biophysical dispersal simulation, and environmental data into a seascape genetic approach using partial RDA analysis (Meirmans 2015). The seagrass *T. hemprichii* is widely-distributed in the Indo-west Pacific (Short *et al.* 2007). It is dominant in the intertidal areas of the Kimberley (Wilson 2013; Lowe *et al.* 2015) and an important food source for megagrazers, like turtles and dugongs (André *et al.* 2005). The species is a useful model for characterising gene flow in the nearshore islands of the Kimberley, because surface currents are the main vector for its dispersal (Kendrick *et al.* 2012; McMahon *et al.* 2014) and it has a long-distance dispersal (LDD) potential (positively buoyant fruits traveling up to 73 km within 2-7 days of flotation time) (Lacap *et al.* 2002). At a fine spatial scale (< 75km), a single panmictic population is expected because: 1) the spatial scale of the study is well within the dispersal potential of *T. hemprichii,* 2) strong, multidirectional tidal flow would facilitate connectivity, 3) a previous study on its sister species *T. testudinum* reported a single panmictic population covering a much larger spatial scale, i.e. 350 km (van Dijk et al 2009).

4.2. Materials and Methods

4.2.1 Study site and sampling

Thirteen sites were sampled in August and October 2014 around the Buccaneer Archipelago (four sites), the Sunday Islands group including mainland sites (seven sites), and intermediate sites between the first two groups (two sites) (Table 4.1; Figure 4.1). All sampling sites were in the high intertidal shallow lagoonal environments where *T. hemprichii* predominantly occurs. Pairwise over-water distances between sampling sites, defined as the shortest path between two locations without crossing boundaries of any landmass (at mean sea level), ranged from 2 km to 73 km. The distance calculation was performed based on high-resolution bathymetric data from the US National Ocean and Atmospheric Administration (NOAA) using the package *marmap* in R (Pante & Simon-Bouhet 2013).

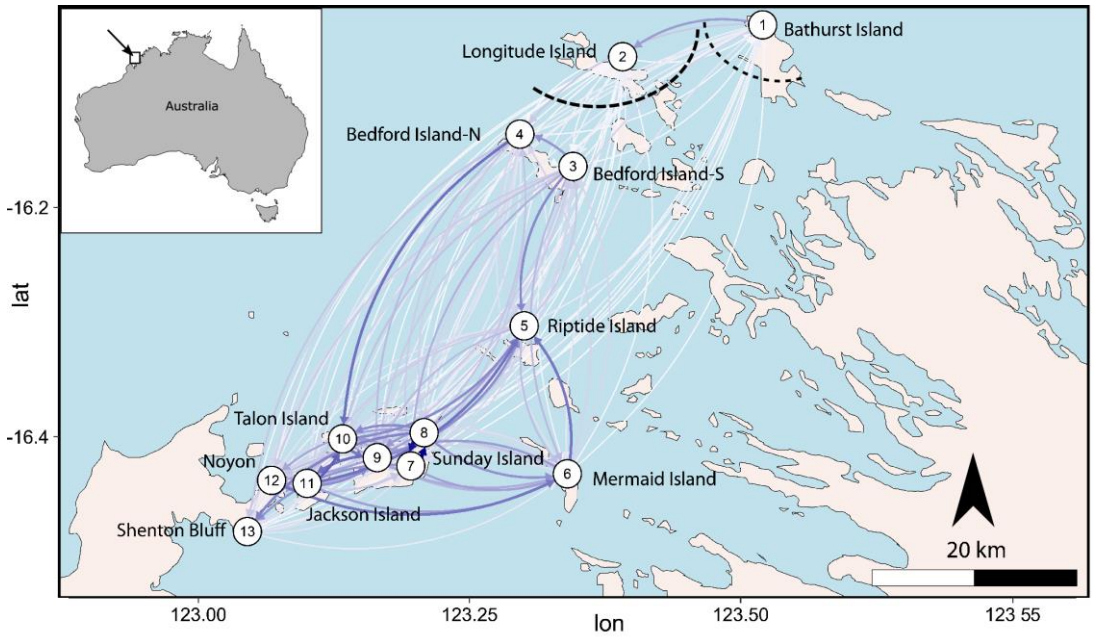

Figure 4.1. Map of the sampling sites and the pattern of gene flow based on $\widehat{Nm}$ (relative number of migrants per generation- Alcala et al. 2014). Sampling sites (populations) were represented by numbers within circles (referred to Table 4.1). Levels of $\widehat{Nm}$ among sampling sites were represented by curved lines. The thicker the lines, the higher level of gene flow between populations. Barriers of gene flow based on Monmonier algorithm were indicated by dashed line.

At each sampling site, five variables were recorded to define environmental conditions used in the seascape genetic analysis: (i) water depth (m) relative to mean sea level, determined from Google Earth version 7.1.2.2041; (ii) dominant sediment type through visual assessment using Wentworth's categories (boulder, cobble, pebble, granule, sand, silt, and clay) (Wentworth 1922); (iii) geomorphic habitat type (reef terrace, reef lagoon); (iv) number of other seagrass species, and (v) the presence/absence of corals (Table A4.1, Appendix). For geomorphic habitat, reef terrace is defined as a step-like reef flat, while reef lagoon is a shallow body of water separated from the main water body by fringing reefs.

A total of 650 samples (meristematic region of seagrass leaves) were collected from all sampling sites. At each site, a total of 50 samples, separated by at least 2m, were collected in two ways, either randomly in a circular area with a diameter of ~50m and 5-10 25m-transects starting from the centre of the circle, or if this was not possible, due to the patchy nature of the meadow, then samples were collected haphazardly across the patches in a similar area. Clear meristematic sections of the leaf were cut into strips and then immediately stored in silica to rapidly dry and preserve the DNA. DNA was extracted from 2-3 pieces (5-10 mm in size) of silica dried samples using AGRF extraction service (Australian Genomic Research Facility, www.agrf.org.au). Extractions were done using the Nucleospin Plant II Kit (Machery-Nagel, Düren, Germany) with the PL2/PL3 buffer system.

4.2.2 Microsatellite amplification and genotyping

Microsatellite amplification was conducted on 16 polymorphic microsatellite loci previously developed: Thh5, Thh34, Thh15, TH66, TH37, TH43, TH43, Thh8, TH34, Thh41 TH52, TH07, Thh29, Thh1, Thh36, and Thh3 (Wainwright *et al.* 2013; van Dijk *et al.* 2014). These were amplified using fluorescently labelled primers in three separate multiplex panels with the QIAGEN Type-it® microsatellite PCR Kit (10-µL reactions with ~1 ng of genomic DNA). PCR conditions were set as: an initial 10 min denaturation at 95 ^{0}C, followed by 38 cycles of 30 s at 94 ^{0}C, 45 s at an annealing temperature of 55 ^{0}C, 45 s at 72 ^{0}C with a final extension of 72 ^{0}C for 30 min. Fragment analysis and capillary separation were run at GGF (Georgia Genomic Facility, USA) with GGF's size standard 500 ROX. Microsatellite alleles were

manually scored and checked for errors using the Microsatellite plugin v1.4 in GENEIOUS R7 v 7.1.7 (Biomatters Ltd).

4.2.3 Genetic analysis

Microsatellite data properties

Genotyping errors and the presence of null alleles were tested using a maximum likelihood approach implemented in ML-NULLFREQ with 100,000 randomizations (Kalinowski & Taper 2006). This has been shown to be the overall best performing method for null allele detection (Dąbrowski *et al.* 2015). *Psex* values of MLG copies are mostly insignificant in each sampling sites indicating that there may be more than one genet from the replicates of the same MLG (Appendix A4.1). However, as the allele composition was identical, we cannot differentiate between them and identify distinct genets from closely related MLGs. A new dataset containing only unique multilocus genotypes (MLGs) was generated using the package *poppr* in R (Kamvar *et al.* 2014) for further analysis.

Linkage disequilibrium across multiple loci was examined based on the standardized index of association (r_D) accounting for different sample sizes using the package *poppr* (Kamvar *et al.* 2014). Departure from Hardy-Weinberg Equilibrium (HWE) was based on the inbreeding coefficient (F_{IS}) calculated in GENETIX 4.05 (Belkhir *et al.* 2004). Genetic diversity at each site was expressed as four parameters: (1) allelic richness (A_R), (2) private allele richness (P_A), (3) observed heterozygosity (H_O), and (4) unbiased expected heterozygosity (H_{NB}, Nei 1987). Allelic richness (A_R) and private allele richness (P_A) were calculated using HP-RARE (Kalinowski 2005) from a standardized set of 14 samples. One site (Shenton Bluff) had a sample size lower than 14 and was excluded from the rarefaction. Mean observed heterozygosity (H_O) and mean unbiased expected heterozygosity (H_{NB}, Nei 1978) were calculated using GENETIX 4.05 (Belkhir *et al.* 2004).

Genetic differentiation and structure

F_{ST} was used as a measure of genetic differentiation. Since the mutation rate can affect differentiation, the use of F_{ST} with highly polymorphic markers, such as

microsatellites, can lead to bias in estimating genetic differentiation. The mutational effects on genetic differentiation were examined using the correlation coefficient between G_{ST} and H_S across loci (r_{GH}) in the program CODIDI (Wang 2015). The r_{GH} was positive and not significant, inferring F_{ST} was not underestimated (0.2878, *p*-value=0.279, Figure A4.1, Appendix) (Wang 2015) and that the genetic differentiation measure was not affected by mutation rate. The population-pairwise F_{ST} was obtained from GENALEX v6.501 (Peakall & Smouse 2012).

Population structure was examined using a Bayesian assignment test in STRUCTURE v2.3.4 (Pritchard *et al.* 2000). This allows us to identify the number of panmictic clusters (K) among the populations. The number of panmictic clusters (K) tested was set from K=1 to K=13, with 20 iterations for each K value, a burn-in of 10^5 and 10^6 replications. Determining the "true" K was based on Evanno *et al.* (2005) and the Web server CLUMPAK was used to align multiple replicate analysis of the appropriate K (Kopelman *et al.* 2015).

Migration

The relative number of migrants per generation ($\widehat{Nm}$, Alcala et al. 2014) was estimated using the function 'divMigrate' in the package *diveRsity* in R (Keenan *et al.* 2013). As there is no evidence that locus mutation rates were significantly affecting genetic differentiation (positive and insignificant r_{GH} as described above), the $\widehat{Nm}$ values were calculated across all loci. The migration pattern was visualized using the package *qgraph* (Epskamp *et al.* 2012). Barriers to gene flow were analysed based on the Monmonier algorithm using F_{ST} with Delaunay Triangulation for the neighbour network using the package *adegenet* in R (Jombart 2008).

4.2.4 Seascape genetic analysis

Oceanographic connectivity

A biophysical dispersal modelling based on Regional Ocean Modelling System (ROMS - M. Feng, unpublished project report) with 2 km resolution was performed to construct a site-pairwise matrix of oceanographic connectivity. The model was nested within the Ocean Forecasting for Australia Model 3 (OFAM3) simulation

(Zhang *et al.* 2016) and forced by 3-hourly meteorological measures derived from Kobayashi *et al* (2015). The model simulation occurred over the 2009-2012 time period. Hourly sea surface current velocities (0-5 m) were extracted from the model output and used for particle tracking modelling. A total of 100 particles were seeded in seagrass sampling sites during austral spring-summer (September-January), at 3-day intervals. This particle release period was chosen to represent the fruiting season of the seagrass based on field observations (A. Z. Perez, *personal communication*). A 4^{th}-order Runga-Kutta sub-time-stepping scheme was used to update the particle locations every hour (Feng *et al.* 2010). Using the random walk effect of 1 m^2s^{-1}, particles were tracked for 7 days based on the potential dispersal duration of the seagrass fruits (Lacap *et al.* 2002). The grid size for tracking the particles from each sampling site was set to 500m x 500m. Connectivity among sampling sites was estimated as the average number of particles released from site *i* that were tracked to be in site *j*, this ranged from 0 to 7.49 per release period, based on 48 simulation replicates in each year of the 4-year time period. The oceanographic connectivity matrix was visualized using the package *qgraph* (Epskamp *et al.* 2012) (Figure A4.2, Appendix).

Disentangling the drivers of genetic differentiation

Variation partitioning based on partial redundancy analysis (partial-RDA) was used to determine the relative contribution of spatial components (geographic distance, GD), oceanographic connectivity (OC) and environmental factors (habitat characteristics, EN) in driving genetic differentiation (GS). As this analysis required both the response and explanatory variables to be single or multicolumn numeric matrices, the 'raw' data of GS, GD, OC, and EN were transformed into new data frames suitable for the analysis. Firstly, a new matrix for the response variable (GS) was constructed by retaining all positive axes derived from a principal coordinate analysis (PCoA) on the linearized F_{ST} (Rousset 1997). Secondly, spatial eigenvectors of a principal coordinate neighbourhood matrix (PCNM) on the pairwise geographic distances were retained to construct the GD matrix. As the first four (out of eight PCNM variables) did not display collinearity, those eigenvectors were used to construct the GD matrix.

For the OC data frame, the pairwise matrix of oceanographic connectivity were transformed into a weighted, directed network based on graph theory using the *igraph* package in R (Csardi & Nepusz 2015). Here, four network parameters were calculated: (i) *strength* -defined as the total amount of the weighted connection coming into and out from a sampling site (higher *strength* indicates higher degree of connectivity), (ii) *closeness* -defined as the degree to which a site is connected to other sites in a network, (iii) *betweenness* -the number of shortest connections between two sites that go through the site of interest, and (iv) *transitivity*, defined as the extent to which the adjacent sites of a site are connected to each other. For calculating *closeness* and *betweenness*, the package treats the connection weights as 'cost' instead of 'connection strength', thus it represents the cost needed to connect nodes (higher *closeness* and *betweenness* imply a higher degree of isolation) (Barrat *et al.* 2004; Csardi & Nepusz 2015). The network parameters indicated that Bathurst Island and Longitude Island were oceanographically isolated from the other sites (Table A4.2, Appendix). The network parameters were used for the seascape genetic analysis. A principal component analysis (PCA) was run on the centred and scaled values of the network parameters. The OC data frame was constructed based on the on the first 3 PCA axes representing 90% of the variance of the data.

For the EN data frame, the categorical variables (sediment type, habitat type, and the presence of corals) was transformed into dummy variables, and combined with the numerical data (water depth and number of other seagrass species). Then, a correspondence analysis (CA, unconstrained ordination) was run on the transformed environmental data. The ordination plot showed that all sites, with the exception of Bathurst Island and Longitude Island, clustered together, indicating that these two sites were different to the remainder (Figure A4.3, Appendix). The variable most responsible in driving the environmental differentiation was sediment type. The EN data frame was constructed based on the first 3 CA axes from the correspondence analysis representing 96% of the variance of the data.

Finally, the basic formula for the partial RDAs was 'GS ~ GD + OC + EN'. The analysis decomposed the variation in the response variable GS into components accounted for by the explanatory variable GD, OC and EN. The adjusted-R^2 was

calculated to determine the amount of variation attributed to each explanatory variable controlling the effect of the other variables (the conditional effect) and without controlling the effect of the other ones (the marginal effect), and the shared fraction of variation by any combination of explanatory variables (Peres-Neto *et al.* 2006). This approach is more robust to decompose spatially structured genetic variation than a Mantel test and its derived forms (Legendre & Fortin 2010; Guillot & Rousset 2013; Meirmans 2015). The analysis was performed using the package *vegan* in R to perform the variation partitioning analysis (Oksanen *et al.* 2015).

4.3 Results

4.3.1 Microsatellite data properties

Over the entire study area, a total of 605 samples were amplified, and clonal correction on each site resulted in 336 MLGs. From these MLGs, a total of 65 alleles were observed across 16 microsatellite loci. ML-NULLFREQ did not detect scoring errors in all loci (estimate of genotyping error $\beta < 0.001$). A significant heterozygote deficit was detected in some sampling sites (positive F_{IS}, Table 4.1). Furthermore, heterozygote deficits were detected in six loci (Thh34, Thh15, TH73, TH43, Thh1, and Thh3). ML-NULLFREQ indicated the presence of null alleles in those loci, although the average frequency was relatively low (Thh34=0.145, Thh15=0.097, TH73=0.115, TH43=0.133, Thh1=0.116, and Thh3=0.120). After these loci were removed, similar populations still showed heterozygote deficits, thus the heterozygote deficit is likely attributed to biological factors, such as inbreeding and the Wahlund effect (reduction in observed heterozygosity due to subpopulation structures), rather than technical issues like the presence of null alleles (Dharmarajan *et al.* 2013), thus all loci were retained for further analysis. The test for linkage disequilibrium across multiple loci showed a small standardised index of association ($\bar{r}_d = 0.0217$, $p=0.001$), indicating a low chance of association between loci (Agapow & Burt 2001).

Genotypic diversity (*R*) varied greatly from 0.09 to 0.94 (Table 4.1). The total number of observed alleles (*nA*) ranged from 21 (Shenton Bluff) to 36 (Mermaid Island), while allelic richness (A_R) ranged from 1.47 (Bedford Island – North) to 1.84 (Talon Island and Mermaid Island). The highest expected heterozygosity (H_{NB}) was found at Longitude Island (0.216), with the lowest at Noyon (0.107). Most sites in the Buccaneer Archipelago exhibited significant excess of heterozygotes (negative value of F_{IS}), except Bedford Island-South. In the Sunday Island group and mainland, significant excess of heterozygotes was only detected at Talon Island (Table 4.1).

Table 4.1. Genetic diversity of *Thalassia hemprichii* in the Kimberley obtained from 16 microsatellite loci, standardized at 14 MLGs. *N*: total number of individuals examined; *G*: number of multilocus genotype; *R*: clonal richness (MLG-1/N-1); *nA*: observed alleles; A_R: allelic richness and P_A: private allele richness, standardized at 14 samples (average alleles per loci); H_O: observed heterozygosity; H_{NB}: non-biased heterozygosity-Nei, 1978. $^{++}$Shenton Bluff population was not standardized to 14 MLGs for A_R and P_A, as there were only 5 MLGs found. Significance level, *=p<0.05, **=p<0.01

Region	Site	Site ID	*N*	*G*	*R*	*nA*	A_R	P_A	H_O	H_{NB}	F_{IS}
Buccaneer Archipelago	Bathurst Is.	1	30	14	0.48	24	1.50	0	0.232	0.168	-0.408**
	Longitude Is.	2	48	23	0.49	30	1.83	0.06	0.291	0.216	-0.357**
	Bedford Is. –South	3	48	37	0.77	28	1.65	0.01	0.120	0.139	0.138**
	Bedford Is. –North	4	48	23	0.47	24	1.47	0	0.133	0.133	-0.004
Intermediate sites	Riptide Is.	5	48	43	0.94	31	1.82	0.03	0.199	0.211	0.059*
	Mermaid Is.	6	48	44	0.91	36	1.84	0.09	0.215	0.196	-0.097**
Sunday Island group, (including mainland sites)	Sunday Is. –South	7	47	20	0.43	27	1.58	0.05	0.119	0.132	0.105**
	Sunday Is. –North	8	48	27	0.57	27	1.56	0.12	0.130	0.131	0.009
	Halls Pool	9	48	32	0.66	27	1.61	0.07	0.104	0.171	0.399**
	Talon Is.	10	48	18	0.36	31	1.84	0.12	0.208	0.180	-0.162**
	Jackson Is.	11	48	33	0.68	31	1.73	0.08	0.135	0.141	0.047
	Noyon	12	48	17	0.36	26	1.59	0	0.092	0.107	0.143*
	Shenton Bluff^{++}	13	48	5	0.09	21	1.31	0.06	0.175	0.133	-0.366

4.3.2 Genetic differentiation and connectivity

Overall, significant genetic differentiation was detected among the sampling sites (global F_{ST} 0.201, p=0.001). Pairwise F_{ST} between sampling sites varied by more than an order of magnitude (0.022 between two Sunday Island populations; to 0.495 between Longitude Island and Bedford Island-North) (Figure 4.2). All pairwise F_{ST} were significantly greater than zero (p<0.01), except between the two Sunday Island populations (p=0.071). The highest genetic differentiation was found in pairs of populations that were separated by only 12 km (Longitude Island and Bedford Island-South, F_{ST}=0.431) and 14 km (Longitude Island and Bedford Island-North, F_{ST}=0.495, whereas more distant sampling sites such as (Bedford Island-South and Noyon) had F_{ST} of 0.07.

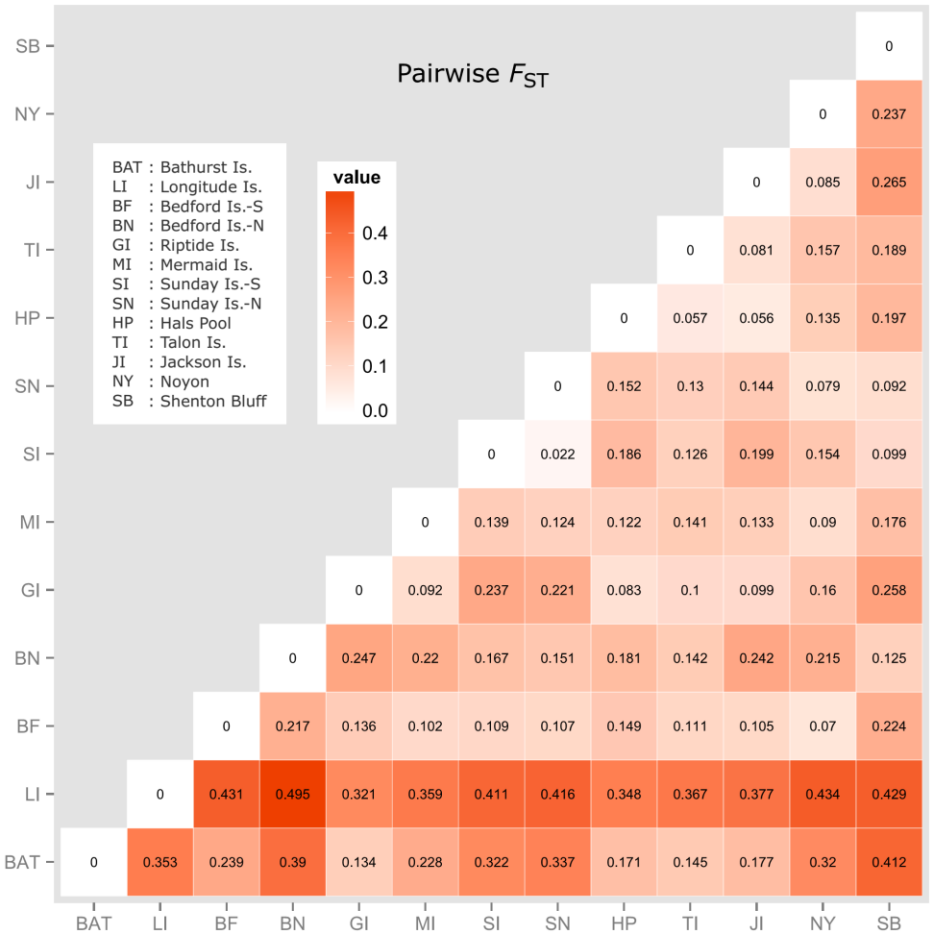

Figure 4.2. Pairwise of F_{ST} between populations of *Thalassia hemprichii* in the Kimberley. All pairwise F_{ST} were significant (p<0.01), except between the two Sunday Island populations (SI *vs* SN; F_{ST}=0.022, p=0.071). Global F_{ST} = 0.201, p=0.001.

The relative number of migrants ($\widehat{Nm}$) among the seagrass populations ranged from 0.014 to 1 with meadows 12-14 km apart (Longitude Island and the two Bedford Island populations) less connected than ones 30-50 km apart such as Bedford Island-North and Talon Island (Figure 4.1). Gene flow was asymmetrical, predominantly in a south-westward direction, from the Buccaneer Archipelago to the Sunday Island group. The highest level of gene flow ($\widehat{Nm}$ = 1) was observed from Sunday Island-South to Sunday Island-North. Low levels of gene flow were detected from Bathurst Island and Longitude Island to all other sites, suggesting that the two populations were relatively isolated from the other populations. This was supported from the genetic barrier analysis based on the Monmonier algorithm (Figure 4.1).

Bayesian probability assignment conducted in STRUCTURE revealed a spatial pattern of genetic differentiation (Figure 4.3). Model evaluation with the deltaK method (Evanno *et al.* 2005) indicated two to four populations were best supported, of which *K*=3 had the highest support. At *K*=2, individuals sampled from Bathurst Island and Longitude Island were relatively uniformly assigned with high probability to one cluster. Individuals from the remaining sampling sites were either assigned strongly to the other cluster or exhibited high admixture between the two clusters. At *K*=3, individuals sampled from Bathurst and Longitude Islands formed a distinct and uniform cluster. Individuals from the remaining sites were either strongly assigned to one cluster (Sunday Island and Noyon), or were highly admixed between the two remaining clusters. At *K*=4, individuals from Bathurst Island became distinct from those collected at Longitude Island, but the clustering pattern of the remaining individuals did not change significantly (Figure 4.3).

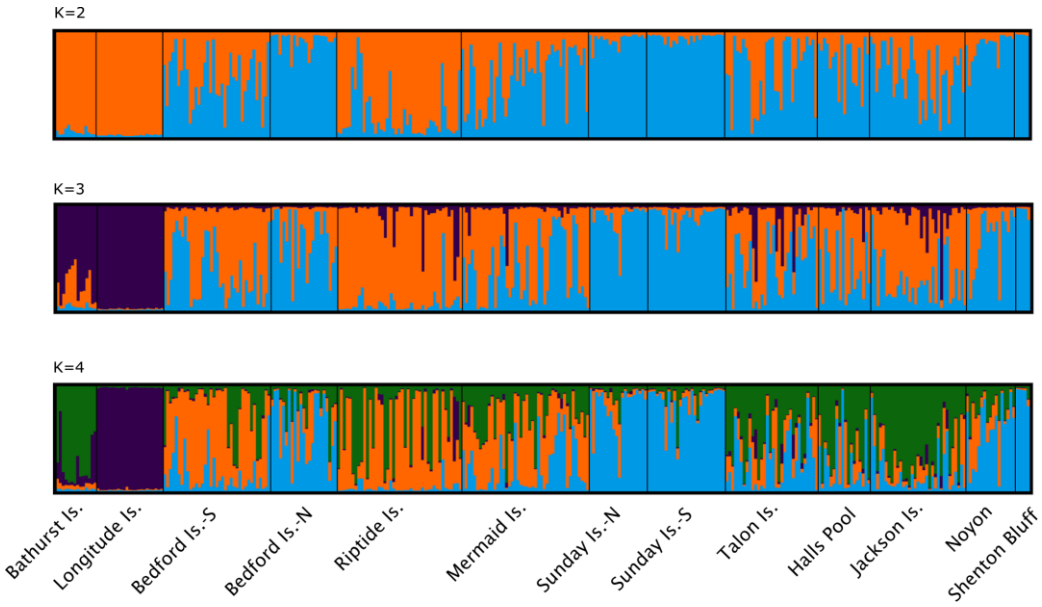

Figure 4.3. Cluster of populations resulted from STRUCTURE analysis (Pritchard *et al.* 2000). Each individual is represented by a thin vertical line, which is partitioned into K segments that represent its estimated population group membership fractions. Each colour represents a distinct population. Black lines separate individuals from geographical site locations.

4.3.3 The drivers of genetic differentiation

The variance partitioning analysis revealed that oceanographic connectivity was the most significant driver of genetic differentiation, followed by environmental factors (Table 4.2). The marginal effect of oceanographic connectivity and environmental factors was significant accounting for 62.5% and 54.5% of the variation in genetic differentiation among the seagrass populations, respectively (Table 4.2). In contrast, geographic distance accounted for a smaller proportion of the variation (10%) and the effect was not significant (p=0.292). About a third of total variation (28.2%) was not explained by any of the variables. When each individual effect was conditionally estimated by controlling the other explanatory variables, the effects became non-significant (p>0.05), therefore variable oceanographic connectivity and environment do not explain the genetic differentiation independently but in combination.

Table 4.2. Variation partitioning on genetic differentiation (F_{ST}) into components accounted for the explanatory variables: GD (geographic distance), OC (oceanographic connectivity) and EN (environmental factors). Fraction of variation is expressed as a percentage from R^2_{adj} values. df_{mod}: degrees of freedom of model; df_{res}: degrees of freedom of residuals

	R^2_{adj} (%)	df_{mod}	df_{res}	F	*p-value*
Marginal					
EN	54.51	3	9	5.793	0.025
OC	62.46	3	9	7.655	0.002
GD	10.02	4	8	1.334	0.292
Residual	28.19				
Conditional					
EN | (OC + GD)	17.49	3	2	2.034	0.278
OC | (EN + GD)	23.00	3	2	2.359	0.297
GD | (OC + GD)	13.11	4	2	1.697	0.338

The shared fraction among the three explanatory variables explained 36.98% of the total variation, while the shared fraction between oceanographic connectivity and environmental factors explained 21.29% of the total variation. The shared fractions could not be tested for significance as they had zero degree of freedom. There were two negative values associated with the shared fractions, i.e. between geographic distance and environmental factors (-21.26) and between the geographic distance and oceanographic connectivity (-18.81) indicating either: (i) strong, direct and opposing effects of the explanatory variables on the response variable, or (ii) correlations between the explanatory variables (see more details in Legendre & Legendre 1998).

4.4 Discussion

4.4.1 Genetic differentiation and connectivity

The spatial setting of this study, a maximum distance between sites of 73 km, is within the potential dispersal range of this species (Lacap *et al.* 2002), but despite

this, there was strong genetic structuring of the seagrass *T. hemprichii*. However, patterns of differentiation were idiosyncratic with respect to geographic proximity, with some of the closely situated sampling sites being the most differentiated and assigned to a different population cluster. This is in contrast to previous studies carried out on its sister species *T. testudinum,* although both species have similar dispersal mechanisms (floating propagules and no seedbank). Over a similar spatial scale (max. pairwise geographic distance <100 km), high genetic connectivity and a single panmictic population was observed in *T. testudinum* in Florida Bay (Bricker *et al.* 2011). Even at a much larger spatial scale (>1000 km), *T. testudinum* also showed high levels gene flow and low genetic differentiation in the Atlantic coast of Mexico (van Dijk *et al.* 2009). Also, this finding did not match with other seagrass species such as *Posidonia australis* in Western Australia which is less differentiated and more connected at similar spatial scales (Sinclair *et al.* 2014; Ruiz-Montoya *et al.* 2015). Differences in hydrodynamic regimes and meadow condition (for example, island and patchy meadows in the Kimberley *vs* more extensive meadows in Florida Bay and the Atlantic coast of Mexico) might be responsible for this contrast. However, significant genetic differentiation at small spatial scales is not uncommon in seagrass populations, for example in the eelgrass *Zostera marina* (Muñiz-Salazar *et al.* 2006; Tanaka *et al.* 2011; Olsen *et al.* 2013), *Z. japonica* (Hodoki *et al.* 2013), *Z. muelleri* (Sherman *et al.* 2016), and *Cymodocea nodosa* (Alberto *et al.* 2006). In addition, small spatial scale population differentiation and structure is often found in other marine organisms, particularly sessile and sedentary inshore species, for example seaweeds (van der Strate *et al.* 2003; Neiva *et al.* 2012), fishes (Hoffman *et al.* 2005; Ciannelli *et al.* 2010), molluscs (Johnson & Black 2006; Temby *et al.* 2007; Hoffman *et al.* 2012), and corals (Whitaker 2004; Bongaerts *et al.* 2010).

4.4.2 What is driving genetic differentiation?

The main findings of this study emphasize the potential key role of seascape attributes (oceanographic settings and environmental factors) in governing patterns of gene flow. Population differentiation is the consequence of barriers of gene flow imposed by the interactions between species biological traits and their environment, supported in this study by a combined oceanographic connectivity and environment

relationship with gene flow. The finding that the marginal effect of oceanographic connectivity and environmental factors were significant, but the conditional ones were not (Table 4.2), indicates that genetic differentiation in a complex seascape subjected to extreme tidal currents is driven by a combination of water currents constraining or facilitating dispersal combined with differential selection from the environment. In the present study, the oceanographic isolation of Bathurst Island and Longitude Island is in concordance with the genetic isolation of these two sites. Furthermore, the south-westward direction of gene flow (from the Buccaneer Archipelago to the Sunday Island group) was also in concordance with the tide-driven current generally flowing southward into King Sound when the tide is rising, and the direction reverses toward the open ocean during the falling tide. These concordances may imply that oceanographic settings can have profound consequences on gene flow in marine populations, as has been documented in other marine studies, including seagrass (e.g., Johansson *et al.* 2015; Ruiz-Montoya *et al.* 2015). Populations may be geographically close to each other, but could still be genetically isolated if the oceanographic setting prevents dispersal; or geographically distant populations can be highly connected if currents facilitate dispersal among them.

Even if dispersal barriers are absent, pre- and/or post-settlement selection in the recipient populations could still prevent gene flow. The significant effect of environmental characteristics observed in this study indicates that selection against migrants from non-matching natal environments might occur. The seagrass propagules may reach new locations by floating with the water current for up to seven days, but, following arrival in the new location, the local environments might filter these migrants either through inhibiting the migrants' settlement and survival or their ability to sexually reproduce (Nosil *et al.* 2005; Wang & Bradburd 2014). This scenario has been observed in the diatom *Skeletonema marinoi* (Sjöqvist *et al.* 2015), the anemonefish *Amphiprion bicinctus* (Nanninga *et al.* 2014) and the intertidal mussel *Perna perna* (Zardi *et al.* 2011). However, this hypothesis needs to be tested using experimental approaches. This study shows that a significant proportion of genetic differentiation was associated with differences in environmental characteristics, especially sediment type. Sediment conditions and grain size can

influence seagrass growth, survival, and species composition, possibly by controlling nutrient availability and providing a physical matrix for the roots and rhizomes to anchor (Short 1987; Terrados *et al.* 1999; van Katwijk & Wijgergangs 2004; Tanaka & Kayanne 2007). To determine whether local adaptation/selection truly drives genetic differentiation, one could examine population-specific fitness across different sediment types in reciprocal transplants, multiple common garden or provenance trial experiments (Wang & Bradburd 2014). In addition, genetic studies employing a larger panel of markers at functional loci would provide better statistical resolution to examine the hypothesis of local adaptation/selection (Tiffin & Ross-Ibarra 2014).

Beside the local selection scenario, it is also possible that a reduced effective gene flow among populations observed was partially due to colonization history (founder events). This study found significant heterozygote deficits in some sampling sites which is a microevolutionary consequence of founder events as the founder population generally represent a small proportion of the genetic variation from a larger source population (Mayr 1963). Under this scenario, the first few founders to colonize available habitats bring small, but sufficient genetic variation. These founders monopolize the habitat and prevent settlement of new migrants. Consequently, the gene frequency is resistant to decay of genetic exchange, leading to genetic differentiation (de Meester *et al.* 2002; Orsini *et al.* 2013).

This study revealed that geographic distance poorly explained the pattern of gene flow. At a small spatial scale, within the dispersal range of the organism and in a system with multidirectional currents, a stepping-stone model of dispersal, in which gene flow is limited only by geographic distance, is not likely to occur (e.g. the coral *Acropora spicifera*, Houtman Abrolhos Islands (Thomas *et al.* 2015), the whelk *Kelletia kelletii*, Santa Barbara Channel (White *et al.* 2010). In contrast, the significant effect of geographic distance in limiting gene flow is often found at a larger spatial scale in marine systems, for example, in a series of local populations current along the coast (Thiel & Haye 2006; Couceiro *et al.* 2007).

From an analytical standpoint, this study highlights the integration of physical and environmental data into population genetic studies to fully understand processes governing gene flow. Oceanographic connectivity data generated from biophysical

simulation is potentially very useful to predict population connectivity in the absence of genetic data. The seascape approach in this study can be applied elsewhere where genetic studies are lacking. For terrestrial settings, a similar approach using a biophysical simulation based on wind-mediated dispersal of pollen and seed has also improved the understanding of processes influencing gene flow (Schueler & Schlunzen 2006; Kuparinen *et al.* 2007).

In conclusion, this chapter presents evidence of significant genetic differentiation among populations of the seagrass *T. hemprichii* over a relatively small spatial scale in an extreme tidal environment, and this was not explained by geographic distance. A seascape genetic approach showed that oceanographic connectivity, in combination with environmental factors, explained the patterns in genetic differentiation, and the effects of these components cannot be separated. These findings add to the growing evidence for the significant contribution of oceanography and environmental factors in governing the pattern of genetic differentiation in marine populations and that survival of marine species is a complex interaction between connectivity among populations of a species and environment. As the environment rapidly changes under anthropogenic-driven climate change, populations are at risk, even those with high levels of connectivity (Bernhardt & Leslie 2013). Clearly, understanding population connectivity should be a priority for marine conservation and management, and as shown here, the expectations of high genetic connectivity in a species with floating propagules and a capacity for LDD were not met. Instead high levels of genetic differentiation were observed among populations 10s of kilometres apart and populations are possibly locally adapted, suggesting a local conservation management program to be more appropriate.

Chapter 5

GENERAL DISCUSSION

Understanding how genetic variation is distributed across space and how external factors (e.g. palaeo-historical events and contemporary processes) influence the spatial pattern has been a central focus in population genetics, underpinning research in evolutionary biology, conservation biology and related disciplines. Understanding genetic structure and the potential driving forces promoting population divergence and adaptation not only provides insights into population connectivity, but also aids prediction of how populations will respond to a changing environment, which are both important for conservation management.

The patterns in genetic structure and the factors influencing this structure using multiple approaches across different temporal and spatial scales have been examined in this thesis (Figure 1.3). Chapter 2 collated population genetic data (number of genetic clusters and global F_{ST}) of co-distributed marine species to estimate the most important factors driving population structure and differentiation across the Indo-Australian Archipelago, focusing on a regional scale (>300 km). Chapter 3 investigated further the temporal scale (palaeo-historical vs contemporary processes) at which physical and environmental factors influence genetic structure across the IAA in the seagrass *Thalassia hemprichii*. A seascape genetic approach was implemented with two genetic markers that have different resolution of the divergence-time: SNPs (relatively slower mutation rate) and microsatellites (relatively faster mutation rate) (Haasl & Payseur 2011). Then, Chapter 4 focused on how contemporary physical processes such as the movement of water currents and habitat variability influence genetic patterns, at local spatial scales (less than 75 km). In this chapter, the effect of historical processes was minimised by spatially down-scaling the study from regional (>300 km) to local (<75 km).

5.1 Factors influencing genetic structure

This thesis illustrates that for marine organisms, the interaction between species biological traits and the environmental factors (habitat variability, water current, etc.) are the greatest drivers of the genetic structure in the IAA (Chapter 2). Since larval duration should correlate with species' dispersal potential (Shanks *et al.* 2003; Shanks 2009), and dispersal should, theoretically, facilitate gene flow, larval duration has been identified as most important for determining genetic structure (e.g. Faurby & Barber, 2012). However, the finding in Chapter 2 indicates that larval dispersal is not the only important factor influencing genetic structure, as suggested by some studies such as Riginos *et al.* (2011) and Bradbury *et al.* (2008). Other dispersal-related traits such as the ecology of the adult can have a strong influence as well. This is due to the fact that marine species have widely different dispersal capacities across their different life history stages (from eggs, to larvae and adults) and this can influence realised dispersal (e.g. Berry *et al.* 2012). Furthermore, external physical/environmental factors, including geologic/oceanographic features and habitat heterogeneity can have stronger influences on gene flow than larval dispersal (Hellberg 2009). Limited by a specific period of time, larvae may disperse up to a potential distance, however the spatial scale and dispersal direction can be regulated by, for example the speed and direction of water currents or physical barriers such as islands or deep canyons. Local environmental conditions can also pose selection forces preventing the dispersing larvae to settle or recruit, thus limiting gene flow (Hunt & Scheibling 1997; Bierne *et al.* 2003).

Consequently, the finding that the interaction between species biological traits and the environmental factors influences gene flow cautions against prediction of population connectivity based solely on larval dispersal. There is a need to understand and to predict population connectivity for conservation and management applications (Palumbi 2003; Almany *et al.* 2009) and measuring connectivity is challenging for most marine species. Ideally, connectivity is directly measured by tracking large numbers of individuals during their different life history stages and observing their successful movement, settlement and recruitment. However, direct

measurement is often not possible, for a range of practical reasons (e.g. the immense scale of the ocean in contrast with the minute size of the larvae). Instead, researchers use indirect approaches, one of which is using larval duration as a proxy for modelling or estimating population connectivity (e.g. Weersing & Toonen, 2009). However, since larval dispersal is taxon-specific and can also be influenced by physical/environmental factors such as habitat conditions (e.g. Bay *et al*., 2006) and dispersal barriers (e.g. Ayre *et al.* 2009), successful dispersal can be taxon, season or location specific (Sponaugle *et al.* 2002; Cowen & Sponaugle 2009). Based on the findings in Chapter 2, simply predicting spatial scales of population connectivity as a direct function of larval dispersal may not be sufficient since physical/environmental factors contribute to successful dispersal, which mediates genetic exchange. In fact, some species show limited connectivity despite their long larval duration, for example in the sea star *Patiria miniata* (Sunday *et al.* 2014) and the fish *Ophiodon elongatus* (Marko *et al.* 2007), while high connectivity across the sea basin scale was observed in species with short larval duration, like the diadromous fish *Kuhlia rupestris* (Feutry *et al.* 2013). Hence, combining information on larval dispersal with hydrodynamic modelling and genetic data can improve the estimation of population connectivity (e.g. Selkoe *et al.* 2010; Chapter 3 and 4 in this dissertation).

Following the finding that the interaction between species biological traits and physical/environmental factors influence genetic structure in marine populations (Chapter 2), the role of the temporal scale (palaeo-historical vs contemporary processes) at which physical/environmental factors influence genetic structure was further investigated in Chapter 3. This chapter demonstrates that palaeo-historical processes (vicariance due to Pleistocene sea level fluctuations) and a more contemporary process (the movement of ocean currents) strongly influence gene flow and genetic structure in this seagrass species. Using SNPs, the genetic imprint of historical vicariance was observed at a larger spatial scale (>300 km). At this spatial scale, populations are likely to have experienced different demographic processes such as local extinction and then colonization due to historical isolation. During the time of low sea-levels in the Pleistocene, local extinction would have occurred in some localities in western Indonesia such as the Sunda Shelf as, at that time, this region was dry land. But remnant seagrass populations likely persisted in

some water bodies in eastern Indonesia (e.g. Banda Sea) since this region was still submerged (Voris 2000).

As the sea level rose to the present-day level, the remnant populations in eastern Indonesia expanded and colonization occurred in the newly available marine habitat (Sunda Shelf). This is reflected in the gene flow direction from the east to the west and south, which aligns well with the present-day current flowing from the Pacific Ocean to the Indian Ocean and indicates ocean currents may have facilitated this post-LGM colonization. The role of ocean currents in facilitating gene flow has been highlighted in other seagrass species in the IAA, i.e. *Enhalus acoroides* (Nakajima *et al.* 2014) and *Cymodocea serrulata* (Arriesgado *et al.* 2015). Previous experimental studies (e.g. Ruiz-Montoya *et al.* 2015) have also shown a strong effect of water current in seagrass connectivity. Post-LGM colonization facilitated by water current was also suggested in other seagrass species from temperate region, e.g. *Posidonia oceanica* (Serra *et al.* 2010).

Despite these findings, the study also found that the influence of contemporary ocean currents in driving the genetic pattern, especially based on the microsatellite data, is not as strong as that of the palaeo-historical processes. The relatively weak genetic signature related to contemporary ocean currents might be due to the fact that changes in population demography and/or connectivity due to disturbances may not be immediately reflected in genetic metrics (Epps & Keyghobadi 2015). This is because it takes multiple generations for genetic variation to reach equilibrium after any change or perturbation, such that there is a lag between the time at which perturbation occurs and the time at which the genetic response become detectable (Bolliger *et al.* 2014; Epps & Keyghobadi 2015). The time lag 'problem' is a significant consideration in seagrasses because seagrass individuals can grow clonally and be long lived, from decades to potentially millennia (Arnaud-Haond *et al.* 2012). The time lag problem can make population genetic studies challenging to make temporal inferences from the genetic pattern, however it can be controlled, for example, by minimizing the effect of historical processes/changes in the seascape (Epps & Keyghobadi 2015).

When the effect of historical processes was minimized by spatially down-scaling the study from a regional (>300 km) to local (<75 km) scale, contemporary physical processes, including water currents and habitat heterogeneity were shown to strongly influence seagrass genetic structure (Chapter 4). Significant spatial genetic structuring with asymmetric gene flow was detected, in which meadows 12-14 km apart were less connected than meadows 30-50 km apart, indicating a phenomenon termed 'chaotic genetic patchiness' (Johnson & Black 1982) which has been observed in some marine populations, including seagrass (e.g. Arnaud-Haond *et al.* 2008). Using a seascape genetic approach, the 'chaotic' pattern was best explained by incorporating oceanographic and environmental data, suggesting a combined scenario of dispersal limitation and facilitation by ocean current with local selection/adaptation. Seagrass propagules may reach new habitats by floating with the ocean currents, but local environmental conditions in the new habitats may prevent the propagules to settle, recruit and reproduce, thus preventing gene flow. This finding adds to the growing understanding of the links between oceanography, ecology and genetic structure in the marine systems.

5. 2 Seagrass connectivity in the IAA: Scaling from regional to local

The main result in Chapter 2 provides a framework to predict regional population structure by estimating the number of genetically distinct populations. The number of genetically distinct populations (y) is an exponential function of pelagic larval duration (PLD) and the number of ecoregions covered by the study:

$$y = \exp^{\left(0.47+\left(ecoregion' \; 0.07\right)+\left(PLD' \; - 0.0000057\right)\right)}$$

Note that this function was based on slowly mutating markers, i.e. mitochondrial DNA. Here this study presents the model specifically for *T. hemprichii* populations and for the number of genetic clusters estimated in Chapter 3. These data were not included in the Chapter 2 data analysis, as it was not available at the time, and hence provides an independent test of the model generated from the synthesis. The seagrass fruits have the potential to float for 7 days (PLD=168 hours) (Lacap *et al.* 2002) and

the study covered 11 ecoregions, so by solving the equation we would expect 3 genetically distinct populations. Indeed, three (eastern Indonesia, western Indonesia-Sunda Shelf, and Indian Ocean) and five genetic clusters (eastern Indonesia, Sunda Shelf, Lesser Sunda, Western Australia, and Indian Ocean) were identified based on SNPs and microsatellites respectively (Chapter 3). Thus, the genetic data from SNPs in Chapter 3 supports the main results in Chapter 2. The higher number of genetic clusters found in microsatellite data is most likely due to its higher mutation frequency, while the above function was based on slower mutating markers.

Significant genetic structure can occur at both regional and local scales, as demonstrated in Chapter 3 and 4 (Figure 5.1). At the regional scale, the genetic clusters span distances of at least 500 km, suggesting that genetic connectivity of *T. hemprichii* populations occurs over very large geographic scales (Chapter 3). Genetic similarity over large spatial scales has been reported in other seagrasses; the tropical species *T. testudinum* (van Dijk *et al.* 2009), *Enhalus acoroides* (Nakajima *et al.* 2014) and a temperate seagrass, *Posidonia oceanica* (Arnaud-Haond *et al.* 2007). The large scale genetic similarity observed in several seagrass species suggests the probability of rare, but successful long-distance dispersal (LDD) events in seagrass populations. However, significant spatial genetic structuring in seagrass population is not a phenomenon occurring exclusively at large spatial scales, it can also occur at a smaller spatial scale as demonstrated in Chapter 4. Despite the fact that the spatial setting of the study in Chapter 4 is within the potential dispersal range of this species (Lacap *et al.* 2002), three genetic clusters were identified, negating the prediction of a single panmictic population. The strong genetic structuring occurring at both large and small spatial scales suggests that predicting seagrass connectivity solely based on geographic distance may be inaccurate, and the relevant distance between populations in the marine system may not be purely geographic, but rather determined by other factors operating on the seascape setting such as water currents and habitat heterogeneity (Hellberg 2009), as generally demonstrated in this thesis. Thus, seascape setting is very important for seagrass gene flow and genetic structure.

5.3 Management implications

Seagrasses have been declining in many localities in the IAA, such as the South China Sea (Vo *et al.* 2013), Singapore (Yaakub *et al.* 2014), and Indonesia (UNEP 2008). Unfortunately, management efforts specifically targeting seagrass lag behind those for other marine ecosystems (Orth *et al.* 2006). In the case of Indonesia, there is no legislation specifically addressing seagrass conservation. There are several pieces of legislation which refer to seagrass ecosystems (e.g., Act. No 31-2004 and No 27-2007), however they rarely provide protection for them (Nadiarty *et al.* 2012). This might, in part, be because seagrass science is very limited in the region (Ooi *et al.* 2011) and there is a general lack of public awareness about seagrass ecosystems (Orth *et al.* 2006). This thesis examined the population genetic aspects of *T. hemprichii* in the IAA, thus adding to the seagrass science in this region and does have implications in the context of conservation management. It provides insights for two main management issues: (1) spatial design of conservation units and (2) seagrass restoration and transplantation.

5.3.1 Spatial design of conservation units

Conservation units are specific regions with unique genetic and ecological attributes that enhance the potential for species survival, making them important to the overall conservation of a species. The units should be spatially designed to represent the network of gene flow among populations, capturing the adaptive potential within species (Crandall *et al.* 2000). Delineation of management units can be based upon the observed estimate of population genetic divergence. The threshold value of the genetic divergence at which populations should be assigned to different management units can vary, depending on the genetic methods, species, and conservation context (Palsboll *et al.* 2007).

For *T. hemprichii* in the IAA, the genetic patterns suggest five seagrass management units; eastern Indonesia, Sunda Shelf, Lesser Sunda, Western Australia and Indian Ocean (Figure 5.1). Populations in eastern Indonesia are assigned to one management unit since these populations harbour the highest genetic diversity and are the source of genetic diversity for other populations. Populations in the Sunda

Shelf are genetically distinct, resulting from the post-Last Glacial Maximum colonization and they are relatively degraded ecosystems compared to the other populations within the IAA (Halpern *et al.* 2008; Vo *et al.* 2013), thus warranting a separate management unit. Lesser Sunda is separated from the other regions and it may serve as a stepping stone for populations in eastern Indonesia to other regions. The Indian Ocean is assigned to a management unit as it is also genetically distinct. Western Australia deserves a separated management unit because it represents a peripheral population at the southern edge of the species distribution range, experiencing greater environmental pressure than the central populations. Peripheral populations tend to be ecologically and genetically divergent, and protection of these populations may be beneficial for novel evolutionary potential and future speciation events (Lesica & Allendorf 1995).

The management unit delineation described above, and based on the thesis findings, highlights the need for international cooperation in seagrass management since some management units extend across national borders. For example, seagrass populations in Cocos Keeling (Australia) are genetically more related to ones in the southern Java coast (Indonesia). As genetic diversity plays an essential role in species adaptive potential, the resilience of Cocos Keeling population may depend on both demographic and genetic exchange with populations in the southern coast of Java (Figure 5.1). Furthermore, the populations in eastern Indonesia were revealed to be the main source of genetic diversity for other populations in the west, including Western Australia. Demographic and genetic supply from eastern Indonesia could have an important role on the persistence of seagrass populations in Western Australia. This interdependence between populations across national borders emphasizes the importance of developing informed, multi-national management plans for seagrass.

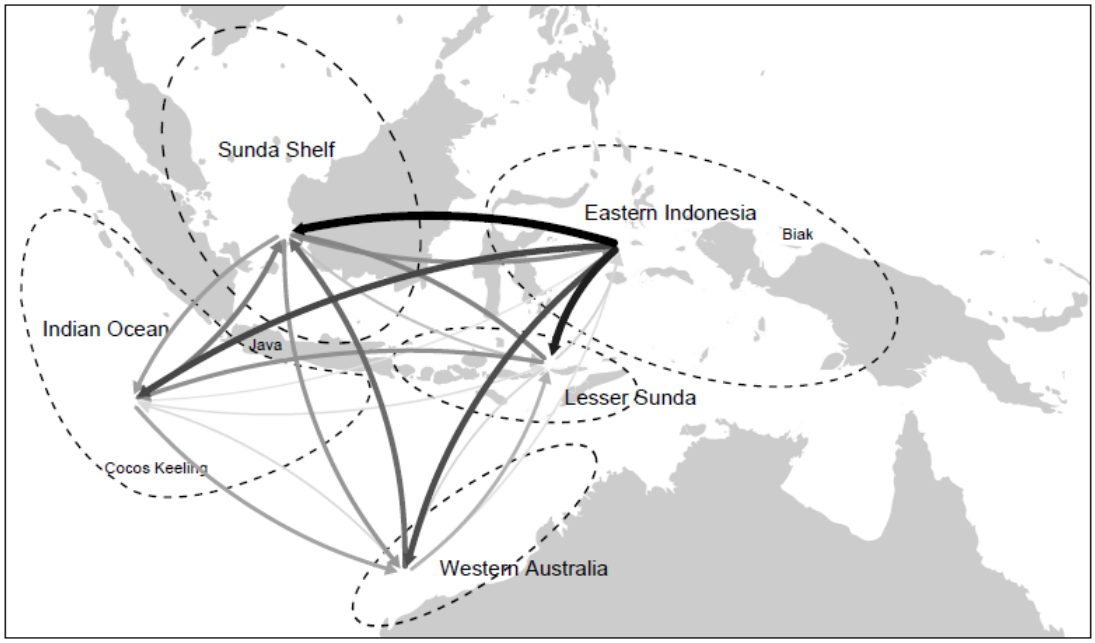

Figure 5.1 Seagrass management units (dashed circles) suggested by the connectivity patterns (arrows). Size indicates the level of connectivity.

As genetic divergence can occur even at a very small spatial scale (< 75 km), this study also emphasizes the importance of local-scale conservation and management measures to maintain population connectivity and resilience at a local level. Protected seagrass meadows can be established based on localised patterns of gene flow and genetic structure. For example, in the Kimberley (Western Australia), seagrass in the Buccaneer Archipelago is genetically separated from seagrass in the Sunday Island group and the gene flow occurs predominantly from the Buccaneer Archipelago to the Sunday Island group. Consequently, populations should be monitored and managed as multiple units that maintain connectivity between populations. At least four seagrass-protected sites are suggested and these are Bathurst Island, Longitude Island, Bedford Island and Talon Island (Figure 5.2). Bathurst Island and Longitude Island are genetically isolated, thus requiring separate protection measures. Within the Buccaneer Archipelago, seagrass meadows in the Bedford Islands seem to be the source of gene flow to the other meadows, highlighting them as a priority for protection. Within the Sunday Island group, Talon Island seems to be the stepping stone population, facilitating gene flow from

Buccaneer Archipelago to the rest of populations in the Sunday Island group. Therefore seagrass in Talon Island should be included in the list of protected sites.

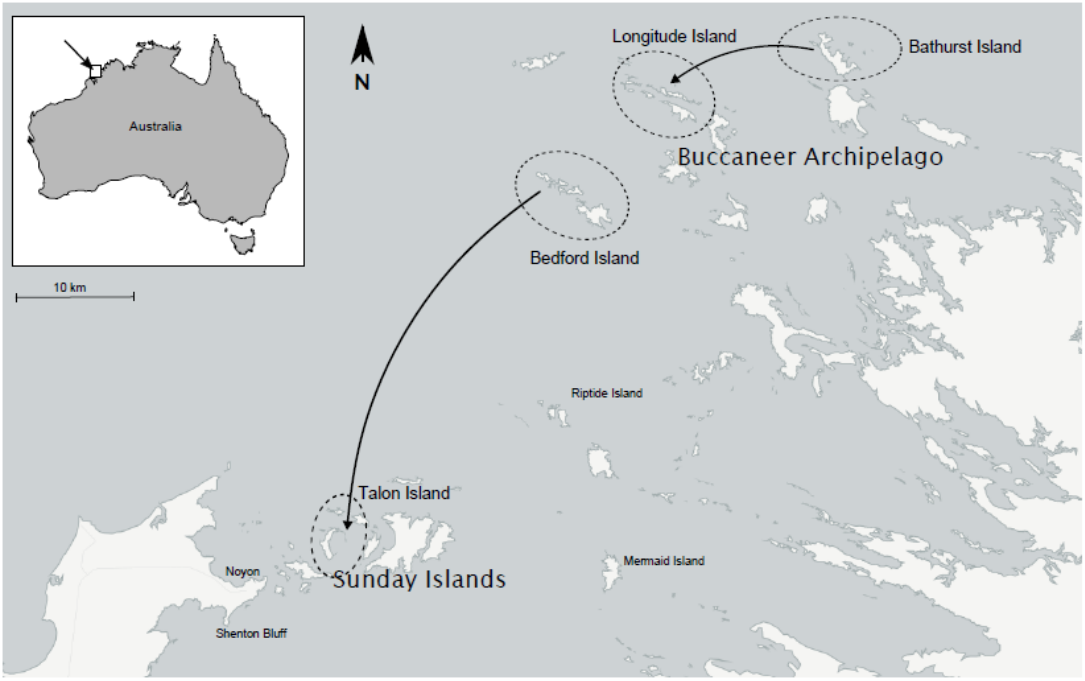

Figure 5.2 Proposed protected-seagrass sites (dashed circles) in the Kimberley, Western Australia based on the genetic patterns. Arrows indicates the general pattern of connectivity.

5.3.2 Seagrass restoration

Once a seagrass meadow is lost, it may not naturally recover, or the recovery may take place slowly (decades or centuries), particularly for slower-growing species like *Thalassia* (Erftemeijer & Robin Lewis 2006) that do not have a persistent seed bank. Seagrass conservation managers often conduct transplantation programs to restore populations or to speed up the rate of recovery (e.g. Paling *et al.* 2007). The success of the restoration depends on many important factors, including selection of appropriate donor meadows (Campbell 2002; van Katwijk *et al.* 2009). Generally, donor selection is based upon two attributes: ecological and genetic. The ecological attributes affect the ability of the plant from the candidate donor sites to survive and

expand at the transplantation site. This can be examined by observation of the environmental characteristic between the donor and transplantation sites. Transplanting from locations that are environmentally similar might enhance the transplantation success rate (van Katwijk *et al.* 2009). The genetic aspect, particularly the level of genetic diversity in the transplants, contribute to transplantation success in the long term (Williams 2001; Reynolds *et al.* 2012). The success mainly requires that the genetic contribution of the donor plants is sufficiently diverse to allow adaptation to environmental changes and avoid inbreeding in the transplantation site (van Katwijk *et al.* 2009). Since higher genetic diversity could enhance survivorship and reproductive success in seagrass (Hughes & Stachowicz 2004; Reusch & Hughes 2006; Ehlers *et al.* 2008), meadows with high genetic diversity, for example Biak, are ideally suitable as donor populations. However, it is equally important to consider the donor population's genetic assemblage, such that donor materials are genetically neither too dissimilar nor too similar to the transplant sites. If the donor materials are genetically dissimilar to the transplant meadows they may not successfully establish in the new site, or this could lead to, for example, outbreeding depression. On the other hand, if the donor materials are too genetically similar to the transplant meadows, they may successfully establish, but it would not increase the genetic diversity important for the meadows' resilience (van Katwijk *et al.* 2009; Evans *et al.* 2014).

Consequently, information on genetic diversity, gene flow and genetic structure, is necessary prior to transplantation attempts and this study provides this necessary information on *T. hemprichii*. For example, in the case of seagrass management in the Kimberley, if managers need to conduct transplantation in Noyon or Shenton Bluff, the best donor sites (based on the present genetic analysis) would be Talon Island as this population has relatively high genetic diversity and seems to be the stepping stone population facilitating gene flow among populations within the Sunday Island group (Chapter 4). Another example is in the case of seagrass transplantation in Bangka. Field observation for Chapter 3 suggests that seagrasses in Bangka represent the most degraded habitat among other regions within the IAA and transplantation may be urgently required to restore the population. The best donor

plants (based on the present study) would be seagrass from Pari Island since it is still in same genetic cluster, yet harbouring much higher genetic diversity (Chapter 3).

5.4 Directions for future research

While the number of research projects on genetic structure and connectivity has been increasing over the last decade in the IAA (Carpenter *et al.* 2011), seagrass connectivity is still poorly studied. This thesis is among the first to provide information on seagrass connectivity in the region. This work provides estimates of connectivity among *T. hemprichii* populations and factors that have shaped the pattern based on genetic measure across different spatial and temporal scales, including inter-basin level phylogeography and contemporary local-scale connectivity studies. The potential role of external factors in influencing gene flow in seagrasses, including water currents and habitat heterogeneity, is demonstrated in this study.

Important residual gaps in our knowledge are:

(i) The evolutionary history of *T. hemprichii*

While this study provides intraspecific genetic insight only within the IAA, the question of how intraspecific genetic divergences in *T. hemprichii* translates into macroevolutionary partitioning (leading to speciation) is still unresolved. Future research with sampling coverage across the entire distribution range of *T. hemprichii* can provide better insight into the evolutionary history of *T. hemprichii*. It can also resolve the competing hypotheses explaining the megadiversity in the IAA (Centre of Origin vs Centre of Accumulation);

(ii) The importance of contemporary demographic connectivity

Demographic connectivity, defined as the degree to which population growth and vital rates are affected by dispersal (Lowe & Allendorf 2010), is crucial in promoting stability in a metapopulation, and thus is of interest for

conservation management. This thesis provides estimates of seagrass connectivity in the evolutionary context (genetic connectivity) but it remains challenging to infer demographic connectivity from genetic data alone (Hedgecock *et al.* 2007; Lowe & Allendorf 2010). It has been suggested that information on local demographic rates (sexual and clonal reproduction, dispersal and recruitment) and how those rates vary over time combined with genetic data can improve the estimation of seagrass connectivity in both ecological and evolutionary contexts (Lowe & Allendorf 2010). The role of sexual and clonal reproduction, vegetative fragments, dispersal and recruitment in contributing to the connectivity of *T. hemprichii* is still unknown. Furthermore, information on dispersal potential and success of the seagrass propagules (vegetative fragments, pollen, seeds and fruit) is required to confirm the genetic pattern. In *T. hemprichii,* there is only one experimental study which investigated the dispersal potential of the propagules (Lacap *et al.* 2002) but how the dispersal events translate into successful recruitment remains a gap in our existing knowledge (McMahon *et al.* 2014; Kendrick *et al.* 2016; Wu *et al.* 2016).

(iii) Seagrass generation time

The generation time of *T. hemprichii* is unknown, but a previous study on *Posidonia oceanica* suggested that seagrass may live potentially for centuries (Arnaud-Haond *et al.* 2012). Since many genetic measurements rely on the number of generations, the species' generation time has implications for the ability to detect genetic change in the population. For example, a 1000-year old barrier to gene flow may not be detected in a species that has a millennial life-span. Information on how long the species can live can also help investigators to choose which genetic markers or measures would be best for studies that address the temporal scales of the genetic variation.

(iv) The role of local selection/adaption in genetic differentiation

This thesis has emphasized the potential role of habitat heterogeneity in driving genetic divergence through local selection. However, whether local

selection is driving this genetic divergence still needs further investigation. A neutrality test with the population markers identified can be performed to exclude potential outliers (if any); then genetic structure reassessment can be re-run to examine the potential role of local adapatation. Local adaptation can also be examined by multiple common garden or provenance trial experiments (Wang & Bradburd 2014) and by employing a larger panel of markers at functional loci (Tiffin & Ross-Ibarra 2014).

Further research could focus on filling these gaps in *T. hemprichii*, as well as in other species that have different dispersal traits for example *Halophila* and *Halodule*.

REFERENCES

Ackiss AS, Pardede S, Crandall ED *et al.* (2013) Pronounced genetic structure in a highly mobile coral reef fish, *Caesio cuning*, in the Coral Triangle. *Marine Ecology Progress Series*, **480**, 185–197.

Agapow P-M, Burt A (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, **1**, 101–102.

Alberto F, Arnaud-Haond S, Duarte C, Serrão E (2006) Genetic diversity of a clonal angiosperm near its range limit: the case of *Cymodocea nodosa* at the Canary Islands. *Marine Ecology Progress Series*, **309**, 117–129.

Alberto F, Raimondi P, Reed D *et al.* (2011) Isolation by oceanographic distance explains genetic structure for *Macrocystis pyrifera* in the Santa Barbara Channel. *Molecular Ecology*, **20**, 2543–2554.

Alcala N, Goudet J, Vuilleumier S (2014) On the transition of genetic differentiation from isolation to panmixia: What we can learn from GST and D. *Theoretical Population Biology*, **93**, 75–84.

Almany GR, Connolly SR, Heath DD *et al.* (2009) Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. *Coral Reefs*, **28**, 339–351.

Altermatt F, Ebert D (2008) Genetic diversity of *Daphnia magna* populations enhances resistance to parasites. *Ecology Letters*, **11**, 918–928.

Anderson CA, Epperson BK, Fortin MJ *et al.* (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology*, **19**, 3565–3575.

André J, Gyuris E, Lawler IR (2005) Comparison of the diets of sympatric dugongs and green turtles on the Orman Reefs, Torres Strait, Australia. *Wildlife Research*, **32**, 53–62.

Arnaud-Haond S, Duarte CM, Diaz-Almela E *et al.* (2012) Implications of extreme life span in clonal organisms: Millenary clones in meadows of the threatened seagrass Posidonia oceanica. *PLoS ONE*, **7**, e30454.

Arnaud-Haond S, Migliaccio M, Diaz-Almela E *et al.* (2007) Vicariance patterns in the Mediterranean Sea: east-west cleavage and low dispersal in the endemic seagrass *Posidonia oceanica*. *Journal of Biogeography*, **34**, 963–976.

Arnaud-Haond S, Vonau V, Rouxel C *et al.* (2008) Genetic structure at different

spatial scales in the pearl oyster (*Pinctada margaritifera cumingii*) in French Polynesian lagoons: beware of sampling strategy and genetic patchiness. *Marine Biology*, **155**, 147–157.

Arriesgado DM, Kurokochi H, Nakajima Y *et al.* (2015) Genetic diversity and structure of the tropical seagrass *Cymodocea serrulata* spanning its central diversity hotspot and range edge. *Aquatic Ecology*, **49**, 357–372.

Ávila E, Yáñez B, Vazquez-Maldonado LE (2015) Influence of habitat structure and environmental regime on spatial distribution patterns of macroinvertebrate assemblages associated with seagrass beds in a southern Gulf of Mexico coastal lagoon. *Marine Biology Research*, DOI: 10.1080/17451000.2015.1007875.

Ayre DJ, Minchinton TE, Perrin C (2009) Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Molecular Ecology*, **18**, 1887–1903.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Barber PH (2009) The challenge of understanding the Coral Triangle biodiversity hotspot. *Journal of Biogeography*, **36**, 1845–1846.

Barber PH, Cheng SH, Erdmann M V, Tenggardjaja K, Ambariyanto (2011) Evolution and conservation of marine biodiversity in the Coral Triangle. *Crustacean Issues: Phylogeography and Population Genetics in Crustacea*, **19**, 129–156.

Barber PH, Erdmann M V, Palumbi SR (2006) Comparative phylogeography of three codistributed Stomatopods: Origins and timing of regional lineage diversification in the Coral Triangle. *Evolution*, **60**, 1825–1839.

Barber PH, Palumbi SR, Erdmann M., Moosa MK (2000) A marine Wallace's line? *Nature*, **406**, 692–693.

Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A (2004) The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 3747–3752.

Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, **23**, 38–44.

Baums IB, Paris CB, Chérubin LM (2006) A bio-oceanographic filter to larval dispersal in a reef-building coral. *Limnology and Oceanography*, **51**, 1969–1981.

Bay LK, Buechler K, Gagliano M, Caley MJ (2006) Intraspecific variation in the pelagic larval duration of tropical reef fishes. *Journal of Fish Biology*, **68**, 1206–1214.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341–345.

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 of the United States*, **98**, 4563–456.

Beger M, Selkoe KA, Treml E *et al.* (2014) Evolving coral reef conservation with genetic information. *Bulletin of Marine Science*, **90**, 159–185.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations.* Laboratoire Génome, Populations, Interactions, CNRS UMR. Université de Montpellier II, Montpellier, France.

Bernhardt JR, Leslie HM (2013) Resilience to climate change in coastal marine ecosystems. *Annual Review of Marine Science*, **5**, 371–92.

Berry O, England P, Marriott RJ, Burridge CP, Newman SJ (2012) Understanding age-specific dispersal in fishes through hydrodynamic modelling, genetic simulations and microsatellite DNA analysis. *Molecular Ecology*, **21**, 2145–59.

Bierne N, Bonhomme F, David P (2003) Habitat preference and the marine-speciation paradox. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 1399–1406.

de Boer TS, Naguit MRA, Erdmann M V. *et al.* (2014) Concordant phylogenetic patterns inferred from mitochondrial and microsatellite DNA in the giant clam *Tridacna crocea*. *Bulletin of Marine Science*, **90**, 301–329.

de Boer TS, Subia MD, Erdmann M V, Kovitvongsa K, Barber PH (2008) Phylogeography and limited genetic connectivity in the endangered boring giant clam across the Coral Triangle. *Conservation Biology*, **22**, 1255–1266.

Bolliger J, Lander T, Balkenhol N (2014) Landscape genetics since 2003: Status, challenges and future directions. *Landscape Ecology*, **29**, 361–366.

Bongaerts P, Riginos C, Ridgway T *et al.* (2010) Genetic divergence across habitats in the widespread coral *Seriatopora hystrixand* its associated *Symbiodinium*. *PLoS ONE*, **5**, e10871.

Borsa P (2003) Genetic structure of round scad mackerel *Decapterus macrosoma* (Carangidae) in the Indo-Malay archipelago. *Marine Biology*, **142**, 575–581.

Bowen BW, Rocha LA, Toonen RJ, Karl SA (2013) The origins of tropical marine biodiversity. *Trends in Ecology & Evolution*, **28**, 359–66.

Bradbury IR, Laurel B, Snelgrove PVR, Bentzen P, Campana SE (2008) Global patterns in marine dispersal estimates: the influence of geography, taxonomic

category and life history. *Proceedings of the Royal Society B Biological Sciences*, **275**, 1803–9.

Bricker E, Waycott M, Calladine A, Zieman J (2011) High connectivity across environmental gradients and implications for phenotypic plasticity in a marine plant. *Marine Ecology Progress Series*, **423**, 57–67.

Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A practical Information-Theoretic Approach*. Springer-Verlag New York.

Calcagno V, Mazancourt C de (2010) glmulti : An R package for easy automated model selection with (Generalized) Linear Models. *Journal of Statistical Software*, **34**, 1–29.

Campbell ML (2002) Getting the foundation right: A scientifically based management framework to aid in the planning and implementation of seagrass transplant efforts. *Bulletin of Marine Science*, **71**, 1405–1414.

Carpenter KE, Barber PH, Crandall ED *et al.* (2011) Comparative phylogeography of the Coral Triangle and implications for marine management. *Journal of Marine Biology*, **2011**, 1–14.

Charlesworth B, Charlesworth D, Barton NH (2003) The effects of genetic and geographic structure on neutral variation. *Annual Review of Ecology Evolution and Systematics*, **34**, 99–125.

Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. *Nature Reviews. Genetics*, **10**, 783–796.

Ciannelli L, Knutsen H, Olsen EM *et al.* (2010) Small-scale genetic structure in a marine population in relation to water circulation and egg characteristics. *Ecology*, **91**, 2918–2930.

Clark PU, Mix AC (2002) Ice sheets and sea level of the Last Glacial Maximum. *Quaternary Science Reviews*, **21**, 1–7.

Collins LB (2011) Geological setting, marine geomorphology, sediments and oceanic shoals growth history of the Kimberley Region. *Journal of the Royal Society of Western Australia*, **94**, 89–105.

Condie SA, Andrewartha JR (2008) Circulation and connectivity on the Australian North West Shelf. *Continental Shelf Research*, **28**, 1724–1739.

Côté IM, Darling ES (2010) Rethinking ecosystem resilience in the face of climate change. *PLoS Biology*, **8**, e1000438.

Couceiro L, Barreiro R, Ruiz JM, Sotka EE (2007) Genetic isolation by distance among populations of the netted dog whelk *Nassarius reticulatus* (L.) along the European Atlantic coastline. *Journal of Heredity*, **98**, 603–610.

Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, **1**, 443–466.

Coyer J, Diekmann O, Serrão E *et al.* (2004) Population genetics of dwarf eelgrass *Zostera noltii* throughout its biogeographic range. *Marine Ecology Progress Series*, **281**, 51–62.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution*, **15**, 290–295.

Crandall ED, Frey MA, Grosberg RK, Barber PH (2008a) Contrasting demographic history and phylogeographical patterns in two Indo-Pacific gastropods. *Molecular Ecology*, **17**, 611–26.

Crandall ED, Jones ME, Muñoz MM *et al.* (2008b) Comparative phylogeography of two seastars and their ectosymbionts within the Coral Triangle. *Molecular Ecology*, **17**, 5276–90.

Cresswell GR, Badcock KA (2000) Tidal mixing near the Kimberley coast of NW Australia. *Marine and Freshwater Research*, **51**, 641–646.

Cross H, Biffen E, Dijk K van, Lowe A, Waycott M (2016) Effective application of next generation sequencing (NGS) approaches in systematics and population genetics: Case studies in Eucalyptus and Acacia. *Australian Systematic Botany*, Accepted 08 September 2016.

Csardi G, Nepusz T (2015) Package "igraph", Network Analysis and Visualization. version 1.0.1. http://igraph.org/r/.

Cullen-Unsworth L, Unsworth R (2013) Seagrass meadows, ecosystem services, and sustainability. *Environment: Science and Policy for Sustainable Development*, **55**, 14–28.

Dąbrowski MJ, Bornelöv S, Kruczyk M, Baltzer N, Komorowski J (2015) "True" null allele detection in microsatellite loci: a comparison of methods, assessment of difficulties and survey of possible improvements. *Molecular Ecology Resources*, **15**, 477–488.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772–772.

Dharmarajan G, Beatty WS, Rhodes OE (2013) Heterozygote deficiencies caused by a Wahlund effect: Dispelling unfounded expectations. *The Journal of Wildlife Management*, **77**, 226–234.

Diekmann OE, Serrão EA (2012) Range-edge genetic diversity: locally poor extant southern patches maintain a regionally diverse hotspot in the seagrass Zostera marina. *Molecular Ecology*, **21**, 1647–1657.

van Dijk K, Mellors J, Waycott M (2014) Development of multiplex microsatellite PCR panels for the seagrass *Thalassia hemprichii* (Hydrocharitaceae). *Applications in Plant Sciences*, **2**, 1400078.

van Dijk JK, van Tussenbroek BI, Jiménez-Durán K, Márquez-Guzmán GJ, Ouborg J (2009) High levels of gene flow and low population genetic structure related to high dispersal potential of a tropical marine angiosperm. *Marine Ecology Progress Series*, **390**, 67–77.

Doherty P, Planes S, Mather P (1995) Gene flow and larval duration in seven species of fish from the Great Barrier Reef. *Ecology*, **76**, 2373–2391.

Drew J, Barber PH (2009) Sequential cladogenesis of the reef fish *Pomacentrus moluccensis* (Pomacentridae) supports the peripheral origin of marine biodiversity in the Indo-Australian archipelago. *Molecular Phylogenetics and Evolution*, **53**, 335–339.

Du Y, Qu T (2010) Three inflow pathways of the Indonesian throughflow as seen from the simple ocean data assimilation. *Dynamics of Atmospheres and Oceans*, **50**, 233–256.

Duarte C (2002) The future of seagrass meadows. *Environmental Conservation*, **29**, 192–206.

Duarte CM, Kennedy H, Marbà N, Hendriks I (2013) Assessing the capacity of seagrass meadows for carbon burial: Current limitations and future strategies. *Ocean and Coastal Management*, **83**, 32–38.

Duda Jr TF, Palumbi SRR (1999) Population structure of the black tiger prawn, *Penaeus monodon*, among western Indian Ocean and western Pacific populations. *Marine Biology*, **134**, 705–710.

Duffy JE (2006) Biodiversity and the fuctioning of seagrass ecosystems. *Marine Ecology Progress Series*, **233**, 233–250.

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

Ehlers A, Worm B, Reusch TBH (2008) Importance of genetic diversity in eelgrass *Zostera marina* for its resilience to global warming. *Marine Ecology Progress Series*, **355**, 1–7.

Epperson BK (2005) Mutation at high rates reduces spatial structure within populations. *Molecular Ecology*, **14**, 703–710.

Epps CW, Keyghobadi N (2015) Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Molecular Ecology*, 6021–6040.

Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D (2012) qgraph : Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, **48**, 1–18.

Erftemeijer PLA, Robin Lewis RR (2006) Environmental impacts of dredging on seagrasses: A review. *Marine Pollution Bulletin*, **52**, 1553–1572.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Evans SM, Sinclair EA, Poore AGB *et al.* (2014) Genetic diversity in threatened *Posidonia australis* seagrass meadows. *Conservation Genetics*, **15**, 717–728.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Faurby S, Barber PH (2012) Theoretical limits to the correlation between pelagic larval duration and population genetic structure. *Molecular Ecology*, **21**, 3419–32.

Feng M, Slawinski D, Beckley LE, Keesing JK (2010) Retention and dispersal of shelf waters influenced by interactions of ocean boundary current and coastal geography. *Marine and Freshwater Research*, **61**, 1259–1267.

Feutry P, Vergnes A, Broderick D *et al.* (2013) Stretched to the limit; can a short pelagic larval duration connect adult populations of an Indo-Pacific diadromous fish (*Kuhlia rupestris*)? *Molecular Ecology*, **22**, 1518–1530.

Fisher R, Radford BT, Knowlton N *et al.* (2011) Global mismatch between research effort and conservation needs of tropical coral reefs. *Conservation Letters*, **4**, 64–72.

Foggo A, Bilton DT, Rundle SD (2007) Do developmental mode and dispersal shape abundance-occupancy relationships in marine macroinvertebrates? *Journal of Animal Ecology*, **76**, 695–702.

Fourqurean JW, Duarte CM, Kennedy H *et al.* (2012) Seagrass ecosystems as a globally significant carbon stock. *Nature Geoscience*, **5**, 505–509.

Fox J, Weisberg S (2011) *An {R} Companion to Applied Regression*. SAGE Publications, California.

Fox J, Weisberg S, Adler D *et al.* (2015) *Package "car" 2.0-25 : Companion to Applied Regression*. (http://cran.r-project.org/web/packages/car/index.html).

Frankham R (2005) Stress and adaptation in conservation genetics. *Journal of Evolutionary Biology*, **18**, 750–755.

Frankham R, Ballou JD, Briscoe DA (2010) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, U.K.

Frisk MG, Jordaan A, Miller TJ (2014) Moving beyond the current paradigm in marine population connectivity: Are adults the missing link? *Fish and Fisheries*, **15**, 242–254.

Gaggiotti OE, Bekkevold D, Jørgensen HBH *et al.* (2009) Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. *Evolution*, **63**, 2939–2951.

Gaither MR, Bowen BW, Bordenave T-R *et al.* (2011) Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific Barrier with contemporary overlap in The Coral Triangle. *BMC evolutionary biology*, **11**, 189.

Gaither MR, Rocha LA (2013) Origins of species richness in the Indo-Malay-Philippine biodiversity hotspot: evidence for the centre of overlap hypothesis. *Journal of Biogeography*, **40**, 1638–1648.

Galarza J, Carreras-Carbonell J, Macpherson E *et al.* (2009) The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 1473–1478.

Gartner A, Tuya F, Lavery PS, McMahon K (2013) Habitat preferences of macroinvertebrate fauna among seagrasses with varying structural forms. *Journal of Experimental Marine Biology and Ecology*, **439**, 143–151.

Giles EC, Saenz-Agudelo P, Hussey NE, Ravasi T, Berumen ML (2015) Exploring seascape genetics and kinship in the reef sponge *Stylissa carteri* in the Red Sea. *Ecology and Evolution*, **5**, 2487–2502.

Grech A, Chartrand-Miller K, Erftemeijer P *et al.* (2012) A comparison of threats, vulnerabilities and management approaches in global seagrass bioregions. *Environmental Research Letters*, **7**, 24006.

Guillot G, Rousset F (2013) Dismantling the Mantel tests. *Methods in Ecology and Evolution*, **4**, 336–344.

Haasl RJ, Payseur BA (2011) Multi-locus inference of population structure: a comparison between single nucleotide polymorphisms and microsatellites. *Heredity*, **106**, 158–171.

Hall R (2009) Southeast Asia's changing palaeogeography. *Blumea - Biodiversity, Evolution and Biogeography of Plants*, **54**, 148–161.

Halpern BS, Walbridge S, Selkoe KA *et al.* (2008) A global map of human impact on marine ecosystems. *Science*, **319**, 948–952.

Harrell Jr FE (contributions from CD and many others) (2015) *Hmisc: Harrell Miscellaneous, R package version 3.15-0.* (http://cran.r-project.org/web/packages/Hmisc/index.html).

He Z, Feng M, Wang D, Slawinski D (2015) Contribution of the Karimata Strait transport to the Indonesian Throughflow as seen from a data assimilation model. *Continental Shelf Research*, **92**, 16–22.

Hedgecock D, Barber P, Edmands S (2007) Genetic approaches to measuring connectivity. *Oceanography*, **20**, 70–79.

Hedrick P (2012) Population Genetics and Ecology. In: *The Princeton Guide to Ecology* (eds Levin SA, Carpenter SR, Godfray HCJ, et al.), pp. 109–116. Princeton University Press, Princeton, NJ, USA.

Hellberg ME (2009) Gene flow and isolation among populations of marine animals. *Annual Review of Ecology and Systematics*, **40**, 291–310.

Hey J (2006) Recent advances in assessing gene flow between diverging populations and species. *Current Opinion in Genetics & Development*, **16**, 592–596.

Hird SM, Brumfield RT, Carstens BC (2011) PRGmatic: An efficient pipeline for collating genome-enriched second-generation sequencing data using a "provisional-reference genome." *Molecular Ecology Resources*, **11**, 743–748.

Hodoki Y, Ohbayashi K, Tanaka N, Kunii H (2013) Evaluation of genetic diversity in *Zostera japonica* (Aschers. et Graebn.) for seagrass conservation in brackish lower reaches of the Hii river system, Japan. *Estuaries and Coasts*, **36**, 127–134.

Hoegh-Guldberg O (2010) Coral reef ecosystems and anthropogenic climate change. *Regional Environmental Change*, **11**, 215–227.

Hoeksema B (2007) Delineation of the Indo-Malayan centre of maximum marine biodiversity: the Coral Triangle. In: *Biogeography, Time, and Place: Distributions, Barriers, and Islands* (ed Renema W), pp. 117–178. Springer Netherlands.

Hoffman JI, Clarke A, Clark MS, Fretwell P, Peck LS (2012) Unexpected fine-scale population structure in a broadcast-spawning Antarctic marine mollusc. *PLoS ONE*, **7**, e32415.

Hoffman EA, Kolm N, Berglund A, Arguello JR, Jones AG (2005) Genetic structure in the coral-reef-associated Banggai cardinalfish, *Pterapogon kauderni*. *Molecular Ecology*, **14**, 1367–1375.

Hoffmann AA, Blows MW (1994) Species borders: ecological and evolutionary perspectives. *Trends in Ecology & Evolution*, **9**, 223–227.

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 congener. *Molecular Phylogenetics and Evolution*, **49**, 629–638.

Hothorn T, Bretz F, Westfall P (2008) Simultaneous inference in general parametric models. *Biometrical Journal*, **50**, 346–363.

Hu ZM, Zhang J, Lopez-Bautista J, Duan DL (2013) Asymmetric genetic exchange in the brown seaweed *Sargassum fusiforme* (Phaeophyceae) driven by oceanic currents. *Marine Biology*, **160**, 1407–1414.

Hughes AR, Stachowicz JJ (2004) Genetic diversity enhances the resistance of a seagrass ecosystem to disturbance. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 8998–9002.

Hunt HL, Scheibling RE (1997) Role of early post-settlement mortality in recruitment of benthic marine invertebrates. *Marine Ecology Progress Series*, **155**, 269–301.

Hutchinson N, Jenkins GP, Brown A, Smith TM (2014) Variation with depth in temperate seagrass-associated fish assemblages in Southern Victoria, Australia. *Estuaries and Coasts*, **37**, 801–814.

Jackson EL, Rees SE, Wilding C, Attrill MJ (2015) Use of a seagrass residency index to apportion commercial fishery landing values and recreation fisheries expenditure to seagrass habitat service. *Conservation Biology*, **29**, 899–909.

Jahnke M, Christensen A, Micu D *et al.* (2016) Patterns and mechanisms of dispersal in a keystone seagrass species. *Marine Environmental Research*, **117**, 54–62.

Jahnke M, Olsen JL, Procaccini G (2015) A meta-analysis reveals a positive correlation between genetic diversity metrics and environmental status in the long-lived seagrass Posidonia oceanica. *Molecular Ecology*, **24**, 2336–2348.

Jardine DI, Dormontt EE, van Dijk KJ *et al.* (2015) A set of 204 SNP and INDEL markers for Bigleaf maple (*Acer macrophyllum* Pursch). *Conservation Genetics Resources*, **7**, 797–801.

Johansson ML, Alberto F, Reed DC *et al.* (2015) Seascape drivers of *Macrocystis pyrifera* population genetic structure in the northeast Pacific. *Molecular Ecology*, **24**, 4866–4885.

Johnson MS, Black R (1982) Chaotic genetic patchiness in an intertidal limpet, *Siphonaria sp*. *Marine Biology*, **70**, 157–164.

Johnson MS, Black R (2006) Islands increase genetic subdivision and disrupt patterns of connectivity of intertidal snails in a complex archipelago. *Evolution*, **60**, 2498–2506.

Jombart T (2008) adegenet: A R package for the multivariate analysis of genetic

markers. *Bioinformatics*, **24**, 1403–1405.

Jordano P (2010) Pollen, seeds and genes: the movement ecology of plants. *Heredity*, **105**, 329–330.

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

de Juan S, Demestre M, Thrush S (2009) Defining ecological indicators of trawling disturbance when everywhere that can be fished is fished: A Mediterranean case study. *Marine Policy*, **33**, 472–478.

Kalinowski ST (2005) Hp-Rare 1.0: A computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.

Kalinowski ST, Taper ML (2006) Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conservation Genetics*, **7**, 991–995.

Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, **2**, e281.

van Katwijk MM, Bos AR, de Jonge VN *et al.* (2009) Guidelines for seagrass restoration: Importance of habitat selection and donor population, spreading of risks, and ecosystem engineering effects. *Marine Pollution Bulletin*, **58**, 179–188.

van Katwijk MM, Wijgergangs LJM (2004) Effects of locally varying exposure, sediment type and low-tide water cover on *Zostera marina* recruitment from seed. *Aquatic Botany*, **80**, 1–12.

Keenan K, Mcginnity P, Cross TF, Crozier WW, Prodöhl PA (2013) DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, **4**, 782–788.

Kelly RP, Palumbi SR (2010) Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PloS ONE*, **5**, e8594.

Kendrick GA, Orth RJ, Statton J *et al.* (2016) Demographic and genetic connectivity: the role and consequences of reproduction, dispersal and recruitment in seagrasses. *Biological Reviews*, **6**, Early View (Online Version, DOI:10.1111/brv.12261).

Kendrick GA, Waycott M, Carruthers TJB *et al.* (2012) The central role of dispersal in the maintenance and persistence of seagrass populations. *BioScience*, **62**, 56–65.

Kobayashi S, Ota Y, Harada Y *et al.* (2015) The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, **93**, 5–48.

Kochzius M, Nuryanto A (2008) Strong genetic population structure in the boring giant clam, *Tridacna crocea*, across the Indo-Malay Archipelago: Implications related to evolutionary processes and connectivity. *Molecular Ecology*, **17**, 3775–3787.

Koh KS, Matz C, Tan CH *et al.* (2012) Minimal increase in genetic diversity enhances predation resistance. *Molecular Ecology*, **21**, 1741–1753.

Kool JT, Moilanen A, Treml EA (2013) Population connectivity: recent advances and new perspectives. *Landscape Ecology*, **28**, 165–185.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak : a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, **15**, 1179–1191.

Kuhnt W, Holbourn A, Hall R, Zuvela M, Kase R (2004) Neogene history of the Indonesian Throughflow. *Continent-Ocean Interactions Within East Asian Marginal Seas: Geophysical Monograph.*, **149**, 299–320.

Kumar S, Stecher S, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. *Molecular Biology and Evolution*, First published online: March 22, 2016.

Kuo J, Coles R, Long WL, Mellors J (1991) Fruits and seeds of *Thalassia hemprichii* (Hydrocharitaceae) from Queensland, Australia. *Aquatic Botany*, **40**, 165–173.

Kuparinen A, Markkanen T, Riikonen H, Vesala T (2007) Modeling air-mediated dispersal of spores, pollen and seeds in forested areas. *Ecological Modelling*, **208**, 177–188.

de la Torre-Castro M, Di Carlo G, Jiddawi NS (2014) Seagrass importance for a small-scale fishery in the tropics: The need for seascape management. *Marine Pollution Bulletin*, **83**, 398–407.

Lacap CDA, Vermaat JE, Rollon RN, Nacorda HM (2002) Propagule dispersal of the SE Asian seagrasses *Enhalus acoroides* and *Thalassia hemprichii*. *Marine Ecology Progress Series*, **235**, 75–80.

Leache AD, Banbury BL, Felsenstein J, De Oca ANM, Stamatakis A (2015) Short tree, long tree, right tree, wrong tree: New acquisition bias corrections for inferring SNP phylogenies. *Systematic Biology*, **64**, 1032–1047.

Legendre P, Anderson M (1999) Distance-based Redundancy Analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, **69**, 1–24.

Legendre P, Fortin M-J (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, **10**, 831–844.

Legendre P, Legendre L (1998) *Numerical Ecology*. Developments in Environmental Modelling 20. Elsevier.

Lesica P, Allendorf FW (1995) When are peripheral populations valuable for conservation. *Conservation Biology*, **9**, 753–760.

Letunic I, Bork P (2011) Interactive Tree of Life v2: Online annotation and display of phylogenetic trees made easy. *Nucleic Acids Research*, **39**, 1–4.

Lind CE, Evans BS, Taylor JJU, Jerry DR (2007) Population genetics of a marine bivalve, *Pinctada maxima*, throughout the Indo-Australian Archipelago shows differentiation and decreased diversity at range limits. *Molecular Ecology*, **16**, 5193–5203.

Liu N, Chen L, Wang S, Oh C, Zhao H (2005) Comparison of single-nucleotide polymorphisms and microsatellites in inference of population structure. *BMC Genetics*, **6**, S26.

Lohman DJ, de Bruyn M, Page T *et al.* (2011) Biogeography of the Indo-Australian Archipelago. *Annual Review of Ecology, Evolution, and Systematics*, **42**, 205–226.

López-Duarte PC, Carson HS, Cook GS *et al.* (2012) What controls connectivity? An empirical, multi-species approach. *Integrative and Comparative Biology*, **52**, 511–24.

Lourie SA, Vincent ACJ (2004) A marine fish follows Wallace's Line: The phylogeography of the three-spot seahorse (*Hippocampus trimaculatus*, Syngnathidae, Teleostei) in Southeast Asia. *Journal of Biogeography*, **31**, 1975–1985.

Louviere JJ, Hensher DA, Swait JD (2000) *Stated Choice Methods: Analysis and Application*. Cambridge University Press, Cambridge, U.K.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology*, **19**, 3038–51.

Lowe A, Harris S, Ashton P (2004) *Ecological Genetics: Design, Analysis and Application*. Blackwell Publishing, Malden, MA USA.

Lowe RJ, Leon AS, Symonds G, Falter JL, Gruber R (2015) The intertidal hydraulics of tide-dominated reef platforms. *Journal of Geophysical Research: Oceans*, 4845–4868.

Macpherson E, Raventos N (2006) Relationship between pelagic larval duration and geographic distribution of Mediterranean littoral fishes. *Marine Ecology Progress Series*, **327**, 257–265.

Magris RA, Pressey RL, Weeks R, Ban NC (2014) Integrating connectivity and climate change into marine conservation planning. *Biological Conservation*,

170, 207–221.

Maguire J-J, Sissenwine M, Csirke J, Grainger R (2006) The state of the world highly migratory, straddling and other high seas fish stocks, and associated species. *FAO Fisheries Technical Paper*, **495**, 67.

Marko PB, Rogers-Bennett L, Dennis AB (2007) MtDNA population structure and gene flow in lingcod (*Ophiodon elongatus*): Limited connectivity despite long-lived pelagic larvae. *Marine Biology*, **150**, 1301–1311.

Marshall DJ, Monro K, Bode M, Keough MJ, Swearer S (2010) Phenotype-environment mismatches reduce connectivity in the sea. *Ecology Letters*, **13**, 128–140.

Mayr E (1944) Wallace's Line in the light of recent zoogeographic studies. *The Quarterly Review of Biology*, **19**, 1–14.

Mayr E (1963) *Animal Species and Evolution*. Belknap Press, Cambridge, Massachusetts.

McCreadie JW, Adler PH (2006) Ecoregions as predictors of lotic assemblages of blackflies (Diptera: Simuliidae). *Ecography*, **29**, 603–613.

McGlathery KJ, Sundbäck K, Anderson IC (2007) Eutrophication in shallow coastal bays and lagoons: The role of plants in the coastal filter. *Marine Ecology Progress Series*, **348**, 1–18.

McLeod E, Moffitt R, Timmermann A *et al.* (2010) Warming seas in the Coral Triangle: Coral reef vulnerability and management implications. *Coastal Management*, **38**, 518–539.

McMahon K, Dijk K van, Ruiz-Montoya L *et al.* (2014) The movement ecology of seagrasses. *Proceedings of the Royal Society B Biological Sciences*, **281**, 20140878.

de Meester L, Gómez A, Okamura B, Schwenk K (2002) The Monopolization Hypothesis and the dispersal-gene flow paradox in aquatic organisms. *Acta Oecologica*, **23**, 121–135.

Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, **24**, 3223–3231.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F(ST) and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Miller AD, Versace VL, Matthews TG, Montgomery S, Bowie KC (2013) Ocean currents influence the genetic structure of an intertidal mollusc in southeastern Australia - implications for predicting the movement of passive dispersers across a marine biogeographic barrier. *Ecology and Evolution*, **3**, 1248–1261.

Molcard R, Fieux M, Syamsudin F (2001) The throughflow within Ombai Strait. *Deep Sea Research I*, **48**, 1237–1253.

Mori AS, Furukawa T, Sasaki T (2012) Response diversity determines the resilience of ecosystems to environmental change. *Biological Reviews*, **88**, 349–369.

Muldoon KM, Goodman SM (2010) Ecological biogeography of Malagasy non-volant mammals: Community structure is correlated with habitat. *Journal of Biogeography*, **37**, 1144–1159.

Muñiz-Salazar R, Talbot SL, Sage GK, Ward DH, Cabello-Pasini A (2006) Genetic structure of eelgrass *Zostera marina* meadows in an embayment with restricted water flow. *Marine Ecology Progress Series*, **309**, 107–116.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Nadiarty, Riani E, Djuwita I *et al.* (2012) Challenging for seagrass management in Indonesia. *Journal of Coastal Development*, **15**, 234–242.

Nakajima Y, Matsuki Y, Lian C *et al.* (2014) The Kuroshio Current influences genetic diversity and population genetic structure of a tropical seagrass, *Enhalus acoroides*. *Molecular Ecology*, **23**, 6029–6044.

Nanninga GB, Saenz-Agudelo P, Manica A, Berumen ML (2014) Environmental gradients predict the genetic population structure of a coral reef fish in the Red Sea. *Molecular Ecology*, **23**, 591–602.

Nathan R, Getz WM, Revilla E *et al.* (2008) A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 19052–19059.

Neiva J, Pearson GA, Valero M, Serrão EA (2012) Fine-scale genetic breaks driven by historical range dynamics and ongoing density-barrier effects in the estuarine seaweed *Fucus ceranoides* L. *BMC Evolutionary Biology*, **12**, 78.

Nosil P (2008) Speciation with gene flow could be common. *Molecular Ecology*, **17**, 2103–2106.

Nosil P, Vines T, Funk D (2005) Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, **59**, 705–719.

Nuryanto A, Kochzius M (2009) Highly restricted gene flow and deep evolutionary lineages in the giant clam *Tridacna maxima*. *Coral Reefs*, **28**, 607–619.

Oke PR, Sakov P, Cahill ML *et al.* (2013) Towards a dynamically balanced eddy-resolving ocean reanalysis: BRAN3. *Ocean Modelling*, **67**, 52–70.

Oksanen J, Blanchet FG, Kindt R *et al.* (2015) vegan: Community Ecology Package. R package version 2.3-0.

Olsen J, Coyer J, Stam W *et al.* (2013) Eelgrass *Zostera marina* populations in northern Norwegian fjords are genetically isolated and diverse. *Marine Ecology Progress Series*, **486**, 121–132.

Olsen JL, Stam WT, Coyer J a *et al.* (2004) North Atlantic phylogeography and large-scale population differentiation of the seagrass *Zostera marina* L. *Molecular Ecology*, **13**, 1923–41.

Ooi JLS, Kendrick GA, Van Niel KP, Affendi YA (2011) Knowledge gaps in tropical Southeast Asian seagrass systems. *Estuarine, Coastal and Shelf Science*, **92**, 118–131.

Orsini L, Vanoverbeke J, Swillen I, Mergeay J, De Meester L (2013) Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, **22**, 5983–5999.

van Orsouw NJ, Hogers RCJ, Janssen A *et al.* (2007) Complexity reduction of polymorphic sequences (CRoPS^{TM}): A novel approach for large-scale polymorphism discovery in complex genomes. *PLoS ONE*, **2**, e1172.

Orth RJ, Carruthers TIMJB, Dennison WC *et al.* (2006) A global crisis for seagrass ecosystems. *BioScience*, **56**, 987–996.

Paling EI, Keulen M van, Tunbridge DJ (2007) Seagrass transplanting in Cockburn Sound, Western Australia: A comparison of manual transplantation methodology using *Posidonia sinuosa* Cambridge et Kuo. *Restoration Ecology*, **15**, 240–249.

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, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, **17**, 3428–3447.

Palumbi S (1992) Marine speciation on a small planet. *Trends in Ecology & Evolution*, **7**, 114–118.

Palumbi S (2003) Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*, **13**, 146–158.

Pante E, Simon-Bouhet B (2013) marmap: A Package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE*, **8**, e73051.

Parker PG, Snow A a, Schug MD, Booton GC, Fuerst PA (1998) What molecules an tell us about populations: Choosing and using a molecular marker. *Ecology*, **79**, 361–382.

Peakall R, Smouse PE (2012) GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, **28**, 2537–2539.

Peres-Neto PR, Legendre P, Dray S, Borcard D (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*, **87**, 2614–25.

Pringle JM, Wares JP (2007) Going against the flow: maintainance of alongshore variation in allele frequency in a coastal ocean. *Marine Ecology Progress Series*, **335**, 69–84.

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

Procaccini G, Olsen JL, Reusch TBH (2007) Contribution of genetics and genomics to seagrass biology and conservation. *Journal of Experimental Marine Biology and Ecology*, **350**, 234–259.

Quadfasel DR, Cresswell G (1992) A note on the seasonal variability in the South Java Current. *Journal of Geophysical Research*, **97**, 3685–3688.

Raes N, van Welzen PC (2009) The demarcation and internal division of Flora Malesiana: 1857 - present. *Blumea*, **54**, 6–8.

Ranta E, Bach LA, Kaitala V, Bjorklund M, Lundberg P (2009) Competition enhances spatial genetic differentiation. *Evolutionary Ecology Research*, **11**, 1251–1258.

Ravago-Gotanco RG, Juinio-Meñez MA (2010) Phylogeography of the mottled spinefoot *Siganus fuscescens*: Pleistocene divergence and limited genetic connectivity across the Philippine archipelago. *Molecular Ecology*, **19**, 4520–4534.

Ravago-Gotanco RG, Magsino RM, Juinio-Meñez MA (2007) Influence of the North Equatorial Current on the population genetic structure of *Tridacna crocea* (Mollusca: Tridacnidae) along the eastern Philippine seaboard. *Marine Ecology Progress Series*, **336**, 161–168.

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

Renema W, Bellwood D, Braga J (2008) Hopping hotspots: global shifts in marine biodiversity. *Science*, **321**, 654–657.

Reusch T, Ehlers A, Hammerli A, Worm B (2005) Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 2826–2831.

Reusch TBH, Hughes RA (2006) The emerging role of genetic diversity for

ecosystem functioning: Estuarine macrophytes as models. *Estuaries and Coasts*, **29**, 159–164.

Reynolds LK, McGlathery KJ, Waycott M (2012) Genetic diversity enhances restoration success by augmenting ecosystem services. *PLoS ONE*, **7**, e38397.

Reynolds LK, Waycott M, McGlathery KJ (2013) Restoration recovers population structure and landscape genetic connectivity in a dispersal-limited ecosystem. *Journal of Ecology*, **101**, 1288–1297.

Rieseberg LH, Burke JM (2001) A genic view of species integration. *Journal of Evolutionary Biology*, **14**, 883–886.

Riginos C, Douglas KE, Jin Y, Shanahan DF, Treml EA (2011) Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, **34**, 566–575.

Riginos C, Liggins L (2013) Seascape genetics: Populations, individuals, and genes marooned and adrift. *Geography Compass*, **7**, 197–216.

Rivers MC, Brummitt NA, Nic Lughadha E, Meagher TR (2014) Do species conservation assessments capture genetic diversity? *Global Ecology and Conservation*, **2**, 81–87.

Rollon R, Vermaat J, Nacorda H (2003) Sexual reproduction in SE Asian seagrasses: the absence of a seed bank in *Thalassia hemprichii*. *Aquatic Botany*, **75**, 181–185.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Ruiz-Montoya L, Lowe R (2012) The role of hydrodynamics on seed dispersal in seagrasses. *Limnology and Oceanography*, **57**, 1257–1265.

Ruiz-Montoya L, Lowe RJ, Kendrick GA (2015) Contemporary connectivity is sustained by wind- and current-driven seed dispersal among seagrass meadows. *Movement Ecology*, **3**, 1–14.

Schlötterer C (2004) The evolution of molecular markers--just a matter of fashion? *Nature Reviews: Genetics*, **5**, 63–69.

Schueler S, Schlunzen KH (2006) Modeling of oak pollen dispersal on the landscape level with a mesoscale atmospheric model. *Environmental Modeling and Assessment*, **11**, 179–194.

Selig ER, Turner WR, Troëng S *et al.* (2014) Global priorities for marine biodiversity conservation. *PLoS ONE*, **9**, e82898.

Selkoe KA, Gaggiotti OE, Bowen BW, Toonen RJ (2014) Emergent patterns of population genetic structure for a coral reef community. *Molecular Ecology*, **23**,

3064–79.

Selkoe K, Toonen R (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series*, **436**, 291–305.

Selkoe KA, Watson JR, White C *et al.* (2010) Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology*, **19**, 3708–3726.

Serra IA, Innocenti AM, Di Maida G *et al.* (2010) Genetic structure in the Mediterranean seagrass Posidonia oceanica: disentangling past vicariance events from contemporary patterns of gene flow. *Molecular Ecology*, **19**, 557–568.

Sgrò CM, Lowe AJ, Hoffmann AA (2011) Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, **4**, 326–337.

Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. *Biological Bulletin*, **216**, 373–385.

Shanks A, Grantham B, Carr M (2003) Propagule dispersal distance and the size and spacing of marine reserves. *Ecological Applications*, **13**, S159–S169.

Sherman CDH, York PH, Smith TM, Macreadie PI (2016) Fine-scale patterns of genetic variation in a widespread clonal seagrass species. *Marine Biology*, **163**, 82 (1-11).

Short FT (1987) Effects of sediment nutrients on seagrasses: Literature review and mesocosm experiment. *Aquatic Botany*, **27**, 41–57.

Short F, Carruthers T, Dennison W, Waycott M (2007) Global seagrass distribution and diversity: A bioregional model. *Journal of Experimental Marine Biology and Ecology*, **350**, 3–20.

Short FT, Coles R, Fortes MD *et al.* (2014) Monitoring in the Western Pacific region shows evidence of seagrass decline in line with global trends. *Marine Pollution Bulletin*, **83**, 408–416.

Siegel D, Kinlan B, Gaylord B, Gaines S (2003) Lagrangian descriptions of marine larval dispersion. *Marine Ecology Progress Series*, **260**, 83–96.

Sinclair EA, Krauss SL, Anthony J, Hovey R, Kendrick GA (2014) The interaction of environment and genetic diversity within meadows of the seagrass *Posidonia australis* (Posidoniaceae). *Marine Ecology Progress Series*, **506**, 87–98.

Sjöqvist C, Godhe A, Jonsson PR, Sundqvist L, Kremp A (2015) Local adaptation and oceanographic connectivity patterns explain genetic differentiation of a

marine diatom across the North Sea-Baltic Sea salinity gradient. *Molecular Ecology*, **24**, 2871–2885.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Spalding MD, Fox HE, Allen GR *et al.* (2007) Marine Ecoregions of the World: A bioregionalization of coastal and shelf areas. *BioScience*, **57**, 573–583.

Sponaugle S, Cowen RK, Shanks A *et al.* (2002) Predicting self-recruitment in marine populations: biophysical correlates and mechanisms. *Bulletin of Marine Science*, **70(1) Supp**, 341–375.

Sprintall J (2009) Indonesian Throughflow. In: *Encyclopedia of Ocean Sciences* (eds Steele JH, Thorpe SA, Thorpe SA, Turekian KK), pp. 237–243. Elsevier, London, UK.

van der Strate HJ, van de Zande L, Stam WT, Haroun RJ, Olsen JL (2003) Within-island differentiation and between island homogeneity: non-equilibrium population structure in the seaweed *Cladophoropsis membranacea* (Chlorophyta) in the Canary Islands. *European Journal of Phycology*, **38**, 15–23.

Sunday JM, Popovic I, Palen WJ, Foreman MGG, Hart MW (2014) Ocean circulation model predicts high genetic structure observed in a long-lived pelagic developer. *Molecular Ecology*, **23**, 5036–5047.

Swindell WR, Bouzat JL (2006) Gene flow and adaptive potential in Drosophila melanogaster. *Conservation Genetics*, **7**, 79–89.

Tanaka N, Demise T, Ishii M, Shoji Y, Nakaoka M (2011) Genetic structure and gene flow of eelgrass *Zostera marina* populations in Tokyo Bay, Japan: Implications for their restoration. *Marine Biology*, **158**, 871–882.

Tanaka Y, Kayanne H (2007) Relationship of species composition of tropical seagrass meadows to multiple physical environmental factors. *Ecological Research*, **22**, 87–96.

Team RDC (2015) R Development Core Team. *R: A Language and Environment for Statistical Computing*.

Temby N, Miller K, Mundy C (2007) Evidence of genetic subdivision among populations of blacklip abalone (*Haliotis rubra* Leach) in Tasmania. *Marine and Freshwater Research*, **58**, 733–742.

Terrados J, Duarte CM, Kamp-Nielsen L *et al.* (1999) Are seagrass growth and survival constrained by the reducing conditions of the sediment? *Aquatic Botany*, **65**, 175–177.

Thiel M, Haye PA (2006) The ecology of rafting in the marine envonment. III.

Biogeographical and evolutionary consequences. In: *Oceanography and Marine Biology: An Annual Review* (eds Gibson RN, Atkinson RJA, Gordon JDM), pp. 323–429. Taylor & France.

Thomas L, Kennington WJ, Stat M *et al.* (2015) Isolation by resistance across a complex coral reef seascape. *Proceedings of the Royal Society B Biological Sciences*, **282**, 20151217.

Tiffin P, Ross-Ibarra J (2014) Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology and Evolution*, **29**, 673–680.

Timm J, Kochzius M (2008) Geological history and oceanography of the Indo-Malay Archipelago shape the genetic population structure in the false clown anemonefish (*Amphiprion ocellaris*). *Molecular Ecology*, **17**, 3999–4014.

van Tussenbroek BI, Vonk JA, Stapel J *et al.* (2006) The biology of *Thalassia*: Paradigms and recent advances in research. *Seagrasses: Biology, Ecology and Conservation*, 409–439.

Tuya F, Haroun R, Espino F (2014) Economic assessment of ecosystem services: Monetary value of seagrass meadows for coastal fisheries. *Ocean & Coastal Management*, **96**, 181–187.

Underwood JN (2009) Genetic diversity and divergence among coastal and offshore reefs in a hard coral depend on geographic discontinuity and oceanic currents. *Evolutionary Applications*, **2**, 222–233.

Underwood JN, Travers MJ, Gilmour JP (2012) Subtle genetic structure reveals restricted connectivity among populations of a coral reef fish inhabiting remote atolls. *Ecology and Evolution*, **2**, 666–679.

UNEP (2008) *National reports on seagrass in the South China Sea*. UNEP/GEF/SCS Technical Publication No. 12.

Vandewoestijne S, Schtickzelle N, Baguette M (2008) Positive correlation between genetic diversity and fitness in a large, well-connected metapopulation. *BMC Biology*, **6**, 46.

Verheij E, Erftemeijer A (1993) Distribution of seagrasses and associated macroalgae in South Sulawesi, Indonesia. *Blumea*, **38**, 45–64.

Vo ST, Pernetta JC, Paterson CJ (2013) Status and trends in coastal habitats of the South China Sea. *Ocean & Coastal Management*, **85**, 153–163.

Vonk JA, Christianen MJA, Stapel J (2008) Redefining the trophic importance of seagrasses for fauna in tropical Indo-Pacific meadows. *Estuarine, Coastal and Shelf Science*, **79**, 653–660.

Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**, 1153–1167.

Vos P, Hogers R, M B, Reijans M, T L (1995) AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, **44**, 388–396.

Wainwright BJ, Arlyza IS, Karl SA (2013) Isolation and characterization of twenty-four polymorphic microsatellite loci for the topical seagrass, *Thalassia hemprichii*. *Conservation Genetics Resources*, **5**, 939–941.

Wang J (2015) Does GST underestimate genetic differentiation from marker data? *Molecular Ecology*, **24**, 3546–3558.

Wang I, Bradburd G (2014) Isolation by environment. *Molecular Ecology*, **23**, 5649–5662.

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.

Watson JR, Mitarai S, Siegel DA *et al.* (2010) Realized and potential larval connectivity in the southern California bight. *Marine Ecology Progress Series*, **401**, 31–48.

Waycott M, Duarte CM, Carruthers TJB *et al.* (2009) Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 12377–12381.

Weersing K, Toonen R (2009) Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, **393**, 1–12.

Wellington GM, Victor BC (1992) Regional differences in duration of the planktonic larval stage of reef fishes in the eastern Pacific Ocean. *Marine Biology*, **113**, 491–498.

Wentworth C (1922) A scale of grade and class terms for clastic sediments. *The Journal of Geology*, **30**, 377–392.

Whitaker K (2004) Non-random mating and population genetic subdivision of two broadcasting corals at Ningaloo Reef, Western Australia. *Marine Biology*, **144**, 593–603.

White C, Selkoe K, Watson J *et al.* (2010) Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B Biological Sciences*, **277**, 1685–1694.

Whitlock MC (2011) *G'ST* and *D* do not replace *FST*. *Molecular Ecology*, **20**, 1083–91.

Willi Y, Van Buskirk J, Hoffmann AA (2006) Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 433–458.

Williams SL (2001) Reduced genetic diversity in eelgrass transplantation affects both population growth and individual fitness. *Ecological Application*, **11**, 1472–1488.

Williams ST, Benzie JAH (1998) Evidence of a phylogeographic break between populations of a high-dispersal starfish: congruent regions within the Indo-West Pacific defined by color morphs, mtDNA and allozyme data. *Evolution*, **52**, 87–99.

Wilson B (2013) *The Biogeography of the Australian North West Shelf: Environmental change and life's response*. Elsevier Inc., San Diego, USA.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97–159.

Wright S (1943) Isolation by Distance. *Genetics*, **28**, 114–138.

Wright S (1951) The genetical structure of natural populations. *Annals of Eugenics*, **15**, 323–354.

Wu K, Chen C-NN, Soong K (2016) Long distance dispersal potential of two seagrasses Thalassia hemprichii and Halophila ovalis. *PLoS ONE*, **11**, e0156585.

Wyrtki K (1961) *Physical oceanography of the Southeast Asian waters. NAGA report*. NAGA Report Vol. 2, The University of California, Scripps Institution of Oceanography, La Jolla, California, USA.

Yaakub SM, McKenzie LJ, Erftemeijer PLA, Bouma T, Todd PA (2014) Courage under fire: Seagrass persistence adjacent to a highly urbanised city-state. *Marine Pollution Bulletin*, **83**, 417–424.

Zardi GI, Nicastro KR, McQuaid CD, Hancke L, Helmuth B (2011) The combination of selection and dispersal helps explain genetic structure in intertidal mussels. *Oecologia*, **165**, 947–58.

Zhang X, Oke P, Feng M *et al.* (2016) A near-global eddy-resolving OGCM for climate studies. *Geoscientific Model Development*, gmd-2016-17.

APPENDIX

Table A2.1. Dataset generated from 55 peer-reviewed publications. Variable "Group" (taxonomic group), Moll=mollusca, Echi=echinoderms, Crust=crustaceans, Fish=fishes, Cor=corals, Plant=marine plant. Variable "Marker type", seq=DNA sequences, allo=allozyme, other=msat and other marker. Variable "AD" (adult life habit), ses=sessile, sed=sedentary, fswim=motile, mig=migratory. Variable "RS" (rep. strategy), bc=broadcaster, br=brooder. Variable "Egg", pel=pelagic egg, bent=benthic egg, dir=direct developer. Variable "PLD" is for pelagic larva duration in hours. Variable "Eco" is number of marine ecoregions represented by each study. Variable "Geo. dist" is maximum geographic distance by sea among sampling sites. Variable "Gen. struct" is number of population clusters. Main sources for variable "AD", "RS", "Egg" and "PLD" are IUCN Redlist, Fishbase, and LarvalBase, or otherwise references stated in the column Notes. References listed Table A2.1

Main Ref	Species	Group	Marker type	AD	RS	Egg	PLD	Eco	Geo. dist	Fst	Gen. struct	Notes
DeBoer et al (2008)	*Tridacna crocea*	Moll	seq	ses	bc	pel	408	9	5773	0.549	3	PLD and Reprod. traits is based on Jameson (1976)
Kochzius & Nuryanto (2008)	*Tridacna crocea*	Moll	seq	ses	bc	pel	408	8	4770	0.28	4	PLD and Reprod. traits is based on Jameson (1976)
DeBoer et al (2014)	*Tridacna crocea*	Moll	seq	ses	bc	pel	408	8	6562	0.57	3	PLD and Reprod. traits is based on Jameson (1976)
Ravago-Gotanco et al (2007)	*Tridacna crocea*	Moll	allo	ses	bc	pel	408	1	1400	0.056	2	PLD and Reprod. traits is based on Jameson (1976)
Junio-Menez et al (2003)	*Tridacna crocea*	Moll	allo	ses	bc	pel	408	2	850	0.066	2	PLD and Reprod. traits is based on Jameson (1976)
DeBoer et al (2014)	*Tridacna crocea*	Moll	other	ses	bc	pel	408	8	6562	0.045	3	PLD and Reprod. traits is based on Jameson (1976)
Nuryanto & Kochzius (2009)	*Tridacna maxima*	Moll	seq	ses	bc	pel	456	7	4770	0.72	3	PLD and Reprod. traits is based on Jameson (1976)
Tisera et al (2012)	*Tridacna maxima*	Moll	seq	ses	bc	pel	456	1	378	0.268	3	PLD and Reprod. traits is based on Jameson (1976)
Kochzius et al (2009)	*Linckia laevigata*	Echi	seq	sed	bc	pel	672	8	6300	0.03	3	Egg type based on *Acanthaster planci*, PLD and Reprod. traits is based on Yamaguchi (1977)
Crandall et al (2008)	*Linckia laevigata*	Echi	seq	sed	bc	pel	672	8	6400	0.08	2	Egg type based on *Acanthaster planci*, PLD and Reprod. traits is based on Yamaguchi (1977)

Junio-Menez et al (2003)	*Linckia laevigata*	Echi	allo	sed	bc	pel	672	2	850	0.049	2	Egg type based on *Acanthaster planci*, PLD and Reprod. traits is based on Yamaguchi (1977)
Crandall et al (2008)	*Protoreaster nodosus*	Echi	seq	sed	bc	pel	720	7	3320	0.17	2	PLD is based on Yamaguchi (1977); Reprod. traits is based on Bos et al (2008) and Scheibling & Metaxas (2008)
Crandall et al (2008)	*Periclimenes soror*	Crust	seq	sed	br	bent	768	7	6400	0.01	1	PLD is based on *P. sagittifer* (Dos Santos et al, 2004); Reprod. traits is based on Wear (1976) and breeding information of *P. indicus* in marinebreeder.org
Barber et al (2006)	*Haptosquilla pulchella*	Crust	seq	sed	br	bent	144	10	3640	0.8	4	Reprod. traits is based on Christy & Salmon (1991); PLD is based on Erdmann (1997)
Barber et al (2006)	*Haptosquilla glyptocercus*	Crust	seq	sed	br	bent	144	10	3240	0.527	2	Reprod. traits is based on Christy & Salmon (1991); PLD is based on Erdmann (1997)
Barber et al (2006)	*Gonodactylellus viridis*	Crust	seq	sed	br	bent	144	10	3770	0.674	3	Reprod. traits is based on Christy & Salmon (1991); PLD is based on Erdmann (1997)
Lourie et al (2005)	*Hippocampus barbouri*	Fish	seq	sed	br	bent	0	5	3190	0.89	2	Reprod. traits is based on Lourie et al (2004)
Lourie et al (2005)	*Hippocampus kuda*	Fish	seq	sed	br	bent	0	12	4700	0.754	2	Reprod. traits is based on Lourie et al (2004)
Lourie et al (2005)	*Hippocampus spinosissimus*	Fish	seq	sed	br	bent	0	9	4700	0.19	3	Reprod. traits is based on Lourie et al (2004)
Lourie et al (2005)	*Hippocampus trimaculatus*	Fish	seq	sed	br	bent	0	10	4700	0.677	2	Reprod. traits is based on Lourie et al (2004)
Magsino et al (2008)	*Siganus argenteus*	Fish	allo	fswim	bc	pel	552	1	1680	0.015	1	PLD is based on *S. fuscescens* (Mellin et al, 2009); Reprod.traits is based on Robinson et al (2004)
Magsino et al (2008)	*Siganus fuscescens*	Fish	allo	fswim	bc	pel	552	1	1680	0.033	1	PLD and reprod.traits is based on Mellin et al (2009)
Borsa et al (2012)	*Neotrygon kuhlii*	Fish	other	sed	br	dir	0	4	3169	0.316	3	Rep. traits is based on Pierce et al (2009)
Borsa et al (2012)	*Himantura gerrardi*	Fish	other	sed	br	dir	0	2	4448	0.159		Rep. traits is based on White & Dharmadi (2007)
Borsa et al (2012)	*Taeniura lymma*	Fish	other	sed	br	dir	0	4	2789	0.074		Rep. traits is based on Pierce et al (2009), Ferreira (2013)
Ovenden et al (2011)	*Rhizoprionodon acutus*	Fish	other	fswim	br	dir	0	2	5130	0.058	2	Reproductive cycle based on fishbase.org and Henderson et al (2006)
Ovenden et al (2011)	*Sphyrna lewini*	Fish	other	mig	br	dir	0	2	5540	0.002	1	Reproductive cycle based on Hazin et al (2001)
Timm & Kochzius (2008)	*Amphiprion ocellaris*	Fish	seq	sed	br	bent	288	8	4260	0.241	4	PLD is based on Timm & Kochzius (2008)
Timm et al (2012)	*Amphiprion ocellaris*	Fish	other	sed	br	bent	288	9	4260	0.048	3	PLD is based on Timm & Kochzius (2008)
Borsa (2003)	*Decapterus macrosoma*	Fish	seq	mig	bc	pel	312	7	3452	0	1	PLD is based on *Decapterus russelli* (plate E III A6) by Allan Connell (http://fisheggs-and-larvae.saiab.ac.za/)
Rohfritsch & Borsa (2005)	*Decapterus russelli*	Fish	seq	mig	bc	pel	312	6	4020	0.37	2	PLD is based on *Decapterus russelli* (plate E III A6) by Allan Connell (http://fisheggs-and-larvae.saiab.ac.za/)
Horne et al (2011)	*Eleutheronema tetradactylum*	Fish	seq	fswim	bc	pel	648	3	4660	0.62	3	PLD is estimated based on Horne et al (2011)

Horne et al (2011)	*Eleutheronema tetradactylum*	Fish	other	fswim	bc	pel	648	3	4660	0.073	3	PLD is estimated based on Horne et al (2011)
Ackiss et al (2013)	*Caesio cuning*	Fish	seq	fswim	bc	pel	1128	11	5146	0.142	3	Rep. traits is based on Yokoyama et al (1994), PLD is based on Ackiss et al (2013) (Doherty et al., 1995)
Lind et al (2012)	*Pinctada maxima*	Moll	seq	ses	bc	pel	744	8	6817	0.372	1	Rep. traits is based on Rose et al (1990), PLD is based on Rose & Baker (1994)
Lind et al (2007)	*Pinctada maxima*	Moll	other	ses	bc	pel	744	8	6817	0.027		Rep. traits is based on Rose et al (1990), PLD is based on Rose & Baker (1994)
Ovenden et al (2009)	*Carcharinus sorrah*	Fish	seq	fswim	br	dir	0	5	2439	0.903	2	Rep. traits is based on Stevens & Wiley (1986)
Ovenden et al (2009)	*Sphyrna lewini*	Fish	seq	mig	br	dir	0	5	2439	na	1	Rep. traits based on Hazin et al (2001)
Phillips et al (2011)	*Pristis zijsron*	Fish	seq	fswim	br	dir	0	4	4209	0.202	1	Biological traits is based on Peverell (2005)
Phillips et al (2011)	*Pristis clavata*	Fish	seq	fswim	br	dir	0	3	2981	0.419	1	Biological traits is based on Peverell (2005)
Phillips et al (2011)	*Pristis microdon*	Fish	seq	fswim	br	dir	0	4	5526	0.811	2	Biological traits is based on Peverell (2005)
Starger et al (2013)	*Seriatopora hystrix*	Cor	other	ses	br	pel	100	1	1519	0.357	1	PLD is based on Figueiredo et al (2013); Reprod. traits is based on Lin et al (2013) and Baird (2001)
Starger et al (2013)	*Pocillopora damicornis*	Cor	other	ses	br	pel	2400	1	1519	0.139	1	PLD is based on Harii et al (2002); Reprod. traits is based on Scmidt-Roach et al (2012)
Raynal et al (2014)	*Dascyllus aruanus*	Fish	seq	fswim	br	bent	576	6	4035	0.016	3	PLD is based on wellington & Victor (1989); Rep. traits is based on Pillai et al (1985)
Knitweis et al (2009)	*Heliofungia actiniformes*	Cor	seq	ses	bc	pel	72	6	3143	0.026	2	PLD is based on Rinchmond & Hunter (1990); Sampling region is Indo-Maya Archipelago
Knitweis et al (2009)	*Heliofungia actiniformes*	Cor	seq	ses	bc	pel	72	6	63	0.093	1	PLD is based on Rinchmond & Hunter (1990); Sampling region is Spermonde Islands
Mirams et al (2011)	*Apogon doederleini*	Fish	seq	fswim	br	bent	648	2	5557	0.905	2	PLD is based on Mirams et al (2011); Reprod.traits is based on Takeyama et al (2002)
Mirams et al (2011)	*Pomacentrus coelestis*	Fish	seq	fswim	bc	bent	480	2	5557	0.064	1	PLD is based on Mirams et al (2011); Reprod. traits is based on Jan (1997)
Mirams et al (2011)	*Dascyllus trimaculatus*	Fish	seq	fswim	bc	bent	624	2	5557	0.078	1	PLD is based on Mirams et al (2011)
Mirams et al (2011)	*Acanthurus triostegus*	Fish	seq	fswim	bc	pel	1440	2	5557	0.059	1	PLD is based on McCormick (1999); Reprod. traits is based on Craig (1998)
Underwood et al (2007)	*Seriatopora hystrix*	Cor	other	ses	br	pel	100	1	69	0.095		PLD is based on Figueiredo et al (2013); Rep. traits is based on Lin et al (2013)
Ovenden et al (2004)	*Pristipomoides multidens*	Fish	seq	fswim	bc	pel	960	3	2415	0.055	2	PLD is based on Ovenden et al (2004)
Bernardi et al (2004)	*Pterapogon kauderni*	Fish	seq	fswim	br	bent	0	1	222	na	2	The fish is mouthbrooder and doesn't have pelagic stage
Jackson et al (2014)	*Auxis thazard*	Fish	seq	mig	bc	pel	228	6	5600	0.007	2	PLD data is based on Houde & Zastrow (1993)
Jackson et al (2014)	*Euthynnus affinis*	Fish	seq	mig	bc	pel	585.6	7	5796	0.007	2	PLD is based on *Euthynnus alletteratus* in Houde & Zastrow (1993)
Jackson et al (2014)	*Katsuwonus pelamis*	Fish	seq	mig	bc	pel	288	5	5991	0.031	1	PLD is based on Houde & Zastrow (1993) and Tanabe (2002)

Jackson et al (2014)	*Rastrelliger kanagurta*	Fish	seq	mig	bc	pel	240	6	5499	0.077	2	PLD data is based on *R. niphonius* (Shoji & Tanaka, 2006)
Jackson et al (2014)	*Scomberomorus commerson*	Fish	seq	mig	bc	pel	206	7	5238	0.023	2	PLD data is based on *S. maculatus* (Houde & Zastrow, 1993)
Casilagan et al (2013)	*Tripneustes gratilla*	Echi	seq	sed	bc	pel	1248	2	1557	0	1	PLD data is based on Casilagan et al (2013)
Casilagan et al (2013)	*Tripneustes gratilla*	Echi	other	sed	bc	pel	1248	2	1557	0.001		PLD data is based on Casilagan et al (2013)
Ravago-Gotanco & Juinio-Menez (2010)	*Siganus fuscescens*	Fish	seq	fswim	bc	pel	552	3	2733	0.23	2	PLD and reprod.traits is based on (Mellin et al, 2009)
Underwood (2009)	*Acropora tenuis*	Cor	other	ses	bc	pel	1656	2	1415	0.062		PLD is based on Nishikawa et al (2003)
Donrung et al (2011)	*Paphia undulata*	Moll	other	ses	bc	pel	360	2	2341	0.315	5	Reprod.traits is based on Nabuab et al (2010), PLD is based on Del Narto-Campos et al (2010)
Wanna et al (2004)	*Penaeus merguiensis*	Crust	other	fswim	bc	bent	240	2	2203	0.203	2	PLD is based on Zacharia & Kakati (2004), Reprod. traits is based on Hoang et al (2002)
Sugama et al (2002)	*Penaeus monodon*	Crust	allo	fswim	bc	bent	360	6	3933	0.013	3	PLD is based on Motoh (1984), Reprod. traits is based *P. marguiensis* Hoang et al (2002)
Imron et al (2007)	*Haliotis asinina*	Moll	seq	sed	bc	pel	96	9	7498	0.812	4	PLD is based on Counihan (2002), Reprod. traits is based on Counihan et al (2001)
Yahya et al (2014)	*Rhizopora apiculata*	Plant	other	ses	br	pel	2136	7	5157	0.381	4	PLD is estimated based on Drexler (2001), Reprod. traits is based on Clarke et al (2001) & Christensen & Wium-Andersen (1977)
Benzie & Smith-Kane (2006)	*Pinctada maxima*	Moll	other	ses	bc	pel	744	4	1888	0.009		Reproduction timing is estimated based on Rose et al (1990), PLD is based on Rose & Baker (1994)
Wee et al (2014)	*Rhizopora mucronata*	Plant	other	ses	br	pel	2136	4	3189	0.625	2	PLD is estimated based on Drexler (2001), Reprod. traits is based on Clarke et al (2001) & Christensen & Wium-Andersen (1977)
Nakajima et al (2014)	*Enhalus acoroides*	Plant	other	ses	br	bent	120	2	2200	0.323	2	PLD is estimated based on Lacap et al (2002)
Aziz et al (2011)	*Penaeus monodon*	Crust	other	fswim	bc	bent	360	3	2087	0.631	2	PLD is based on Motoh (1984), Reprod. traits is based *P. marguiensis* Hoang et al (2002)
Xu et al (2001)	*Penaeus monodon*	Crust	other	fswim	bc	bent	360	2	598	0.013	2	PLD is based on Motoh (1984), Reprod. traits is based *P. marguiensis* Hoang et al (2002)
Lim et al (2014)	*Nemipterus japonicus*	Fish	seq	fswim	bc	pel	456	2	1511	0.07	2	PLD is based on Nemipteridae Fisher (2005)
Yap et al (2011)	*Carcinoscopius rotundicauda*	Crust	allo	sed	br	bent	576	1	700	0.092	2	Reproductive traits is based on Srijaya et al (2014); PLD is based on the first instar of *Limulus polyhemus* (Sekiguchi et al. 1988)
Lin et al (2012)	*Perna viridis*	Moll	other	ses	bc	pel	480	1	230	0.039		Reproductive traits and PLD is based on Nair & Appukuttan (2003)
Yap et al (2013)	*Perna viridis*	Moll	other	ses	bc	pel	480	2	500	0.077	2	Reproductive traits and PLD is based on Nair & Appukuttan (2003)
Wang et al (2011)	*Epinephelus coioides*	Fish	other	fswim	bc	pel	600	2	2312	na	1	Reproductive traits is based on Grandcourt et al (2005); PLD is based on *E. marginatus* Cunha et al (2013)
Antoro et al (2006)	*Epinephelus coioides*	Fish	other	fswim	bc	pel	600	5	2980	0.074	2	Reproductive traits is based on Grandcourt et al (2005); PLD is based on *E. marginatus* Cunha et al (2013)

Abidin et al (2014)	*Crassostrea iredalei*	Moll	seq	ses	bc	pel	480	3	2885	na	2	Reproductive traits is based on Angell (1986); PLD is based on Ver (1986)
Phinchongsakuldit et al (2013)	*Rachycentron canadum*	Fish	other	fswim	br	pel	648	2	2773	na	1	PLD is based on Salze et al (2011)
Hobbs et al (2013)	*Centropyge joculator*	Fish	seq	fswim	bc	pel	792	1	1194	0.011	2	PLD is estimated based on *C. bispinosus* Brothers & Thresher (1985).
Kochzius et al (2009)	*Thyca crystallina*	Moll	seq	sed	br	pel		7	3380	0.005	1	Reprod. traits is based on other parasitic eulimid *Hypermastus tokunagai* Matsuda et al (2013)
Crandall et al (2008)	*Thyca crystallina*	Moll	seq	sed	br	pel		6	3880	0.01	2	Reprod. traits is based on other parasitic eulimid *Hypermastus tokunagai* Matsuda et al (2013)
Bastidas et al (2002)	*Clavularia koellikeri*	Cor	allo	ses	br	pel		1	1104	0.134	1	

References listed in Table A2.1

Abidin, D.H.Z., Mustaffa, S., Rahim, M.A., Nair, D.M., Naim, D.M, & Nor, S.A.M. (2014) Population genetics of the black scar oyster, *Crassostrea iredalei*: repercussion of anthropogenic interference. *Mitochondrial DNA*, (Early publication), doi: 10.3109/19401736.2014.913137.

Ackiss, A.S., Pardede, S., Crandall, E.D., Ablan-Lagman, M.C., Ambariyanto, Romena, N., Barber, P.H., & Carpenter, K.E. (2013) Pronounced genetic structure in a highly mobile coral reef fish, *Caesio cuning*, in the Coral Triangle. *Marine Ecology Progress Series*, **480**, 185-197.

Angell, C.L. (1986) *The Biology and Culture of Tropical Oysters*. ICLARM Studies and Reviews, Manila.

Antoro, S., Na-Nakorn, U., & Koedprang, W. (2006) Study of genetic diversity of orange-spotted grouper, *Epinephelus coioides*, from Thailand and Indonesia using microsatellite markers. *Marine Biotechnology*, **8**, 17-26.

Aziz, D., Siraj, S.S., Daud, S.K., Panandam, J.M., & Othman, M.F. (2011) Genetic diversity of wild and cultured populations of *Penaeus monodon* using microsatellite markers. *Journal of Fisheries and Aquatic Science*, **6**, 614-623.

Baird, A.H. (2001) *The ecology of coral larvae: settlement patterns, habitat selection and the length of the larval phase*. PhD Dissertation, James Cook University, Australia.

Barber, P. H., Erdmann, M.V., & Palumbi, S. R. (2006) Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the Coral Triangle. *Evolution*, **60**, 1825-1839.

Bastidas, C., Benzie, J.A.H, & Fabricius, K.E. (2002) Genetic differentiation among populations of the brooding soft coral *Clavularia koellikeri* on the Great Barrier Reef. *Coral Reefs*, **21**, 233-241.

Benzie, J.A.H., & & Smith-Kane, C. (2006) Microsatellite variation in Australian and Indonesian pearl oyster *Pinctada maxima* populations. *Marine Ecology Progress Series*, **314**, 197-211.

Bernardi, G. & Vagelli, A. (2004) Population structure in Banggai cardinalfish, *Pterapogon kauderni*, a coral reef species lacking a pelagic larval phase. *Marine Biology*, **145**, 803-810.

Borsa, P. (2003) Genetic structure of round scad mackerel *Decapterus macrosoma* (Carangidae) in the Indo-Malay archipelago. *Marine Biology*, **142**, 575-581.

Borsa, P., Arlyza, I.S., Laporte, M., & Berrebi, P. (2012). Population genetic structure of blue-spotted maskray *Neotrygon kuhlii* and two other Indo-West Pacific stingray species (Myliobatiformes: Dasyatidae), inferred from size-polymorphic intron markers. *Journal of Experimental Marine Biology and Ecology*, **438**, 32-40.

Bos, A.R., Gumanao, G.S., Alipoyo, J.C.E., & Cordona, L.T. (2008) Population dynamics, reproduction and growth of the Indo-Pacific horned sea star, *Protoreaster nodosus* (Echinodermata; Asteroidea). *Marine Biology*, **156**, 55-63.

Brothers, E.B., & Thresher. R.E. (1985) Pelagic duration, dispersal and the distribution of Indo-Pacific coral-reef fishes. *The Ecology of Coral Reefs. Symposia Series for Undersea Research. NOAA Undersea Research Program* **3**(1): 53-69.

Casilagan, I.L.N., Juinio-Menez, M.A., & Crandall, E.D. (2013) Genetic diversity, population structure, and demographic history of exploited sea urchin populations (*Tripneustes gratilla*) in the Philippines. *Journal of Experimental Marine Biology and Ecology*, **449**, 284-293.

Christensen, B., & Wium-Andersen, S. (1977) Seasonal growth of mangrove trees in Southern Thailand. I. The phenology and *Rhizopora apiculata. Aquatic Botany*, **3**, 281-286.

Christy, J.H., & Salmon, M. (1991) Comparative studies of reproductive behaviour in mantis shrimps and fiddle crabs. *American Zoology*, **31**, 329-337.

Clarke, P.J., Kerrigan, R.E., & Westphal, C.J. (2001) Dispersal potential and early growth in 14 tropical mangroves: do early life history traits correlate with patterns of adult distribution? *Journal of Ecology*, **89**, 648-659.

Counihan, R.T, McNamara, D.C., Souter, D.C., Jebreen, E.J., Preston, N.P., Johnson, C.R., & Degnan, B. M. (2001) Pattern, synchrony and predictability of spawning of the tropical abalone *Haliotis asinina* from Heron Reef, Australia. *Marine Ecology Progress Series*, **213**, 193-202.

Counihan, R.T. (2002). *Early life history of* Haliotis asinina*: A model for Haliotid aquaculture and molluscan research*. PhD Dissertation, University of Queensland, Australia.

Craig, P.C. (1998) Temporal spawning patterns of several surgeonfishes and wrasses in American Samoa. *Pacific Science*, **52**, 35-39.

Crandall, E.D., Jones, M.E., Muñoz, M.M., Akinronbi, B., Erdmann, M.V., & Barber, P.H. (2008) Comparative phylogeography of two seastars and their ectosymbionts within the Coral Triangle. *Molecular Ecology*, **17**, 5276-5290.

Cunha, M.E., Re, P., Quental-Ferreira, H., Gavaia, P.J., & Pouso-Ferreira, P. (2013) Larval and juvenile development of dusky grouper *Epinephelus marginatus* reared in mesocosms. *Journal of Fish Biology*, **83**, 448-465.

DeBoer, T.M., Naguit, M.A., Erdmann, M.V., Ablan-Lagman, M.C., Ambariyanto, Carpenter, K.E., Toha, A.H.A, & Barber, P.H. (2014) Concordant phylogenetic patterns inferred from mitochondrial and microsatellite DNA in the giant clam *Tridacna crocea*. *Bulletin of Marine Science*, **90**, 301-329.

DeBoer, T.S., Subia, M.D., Ambariyanto, Erdmann, M.V., Kovitvongso, K., & Barber, P.L. (2008) Phylogeography and limited genetic connectivity in the endangered boring giant clam across the Coral Triangle. *Conservation Biology*, **22**, 1255-1266.

del Narto-Campos, A., Nabuab, F.N., Palla, R.Q., & Burlas, M.R. (2010) The early development of the short-necked clam, *Paphia undulata* (Born 1778) (Mollusca, Pelecypoda: Veneridae) in the laboratory. *Science Diliman*, **22**, 13-20.

Donrung, P., Tunkijjanukij, S., Jarayabhand, P., & Poompuang, S. (2011) Spatial genetic structure of the surf clam *Paphia undulata* in Thailand waters. *Zoological Studies*, **50**, 211-219.

Dos Santos, A., Calado, R., Bartilotti, C., & Narciso, L. (2004) The larval development of the partner shrimp *Periclimenes sagittifer* (Norman, 1861) (Decapoda: Caridea: Palaemonidae: Pontoniinae) described from laboratory-reared material, with a note on chemical settlement cues. *Helgoland Marine Research*, **58**, 129-139

Drexler, J.Z. (2001) Maximum longevities of *Rhizophora apiculata* and *R. mucronata* propagules. *Pacific Science*, **55**, 17-22.

Erdmann, M.V.N. (1997) *The ecology, distribution, and bioindicator potential of Indonesian coral reef stomatopod communities*. PhD Dissertation, University of California, Berkeley.

Ferreira, A. S. (2013) *Breeding and juvenile growth of the ribbontail stingray* Taeniura lymma, *at Oceanario de Lisboa*. PhD Dissertation, Universidade deLisboa.

Figueiredo, J., Baird, A.H., & Connolly, S.R. (2013) Synthesizing larval competence dynamics and reef-scale retention reveals a high potential for self-recruitment in corals. *Ecology*, **94**, 650-659.

Fisher, R. (2005) Swimming speeds of larval coral reef fishes: impacts on self-recruitment and dispersal. *Marine Ecology Progress Series*, **285**, 223-232.

Grandcourt, E.M., Al-Abdessalam, T.Z., Francis, F., & Al-Shamsi, A.T. (2005) Population biology and assessment of the orange-spotted grouper, *Epinephelus coioides* (Hamilton, 1822), in the southern Arabian Gulf. *Fisheries Research*, **74**, 55-68.

Harii, S., Kayanne, H., Takigawa, H., Hayashibara, T., & Yamamoto, M. (2002) Larval survivorship, competency periods and settlement of two brooding corals, *Heliopora coerulea* and *Pocillopora damicornis*. *Marine Biology*, **141**, 39-46.

Hazin, F., Fischer, A., & Broadhurst, M. (2001). Aspects of reproductive biology of the scalloped hammerhead shark, *Sphyrna lewini*, off northeastern Brazil. *Environmental Biology of Fishes*, **61**, 151-159.

Henderson, A.C., McIlwain, J.L., Al-Oufi, H.S., & Ambu-Ali, A. (2006) Reproductive biology of the milk shark *Rhizoprionodon acutus* and the bigeye houndshark *Iago omanensis* in the coastal waters of Oman. *Journal of Fish Biology*, **68**, 1662-1678.

Hoang, T., Lee, S.Y., Keenan, C.P., & Marsden, G.E. (2002) Spawning behaviour of *Penaeus (Fenneropennaeus) merguiensis* de Man and the effect of light intensity on spawning. *Aquaculture Research*, **33**, 351-57.

Hobbs, J.P.A., Herwerden, L.V., Jerry, D.R., Jones, G.P., & Munday, P.L. (2013) High genetic diversity in geographically remote populations of endemic and widespread coral reef angelfishes (genus: *Centropyge*). *Diversity*, **5**, 39-50.

Horne, J.B., Momigliano, P., Welch, D.J., Newman, S.J., & 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.

Houde, E.D. & Zastrow, C.E. (1993) Ecosystem- and taxon-specific dynamic and energetics properties of larval fish assemblages. *Bulletin of Marine Science*, **53**, 290-335.

Imron, Jeffrey, B., Hale, P., Degnan, B., & Degnan, S.M. (2007) Pleistocene isolation and recent gene flow in *Haliotis asinina*, an Indo-Pacific vetigastropod with limited dispersal capacity. *Molecular Ecology*, **16**, 289-304.

Jackson, A.M., Ambariyanto, Erdmann, M.V., Toha, A.H.A., Stevens, L.A., Barber, P.H. (2014) Phylogeography of commercial tuna and mackerel in the Indonesian Archipelago. *Bulletin of Marine Science*, **90**, 471-492.

Jameson, S.C. (1976) Early life history of the giant clams *Tridacna crocea* Lamarck, *Tridacna maxima*, (Roding), and *Hippopus hippopus* (Linnaeus). *Pasific Science*, **30**, 219-233.

Jan, R.Q. (1997) Sympatric spawning of the damselfishes *Chromis fumea* and *Pomacentrus coelestis* on the northern coast of Taiwan. *Zoological Studies*, **36**, 26-32.

Juinio-Menez, M.A., Magsino, R.M., Ravago-Gotanco, R., & Yu, E.T. (2003) Genetic structure of *Linckia laevigata* and *Tridacna crocea* populations in the Palawan shelf and shoal reefs. *Marine Biology*, **142**, 717-726.

Knittweis, L., Kraemer, W.E., Timm, J., & Kochzius, M. (2009) Genetic structure of *Heliofungia actiniformis* (Scleractinia: Fungiidae) populations in the Indo-Malay Archipelago: implications for live coral trade management efforts. *Conservation Genetics*, **10**, 241-249.

Kochzius, M. & Nuryanto, A. (2008) Strong genetic population structure in the boring giant clam, *Tridacna crocea*, across the Indo-Malay Archipelago: implications related to evolutionary processes and connectivity. *Molecular Ecology*, **17**, 3775-3787.

Kochzius, M., Seidel, C., Hauschild, J., Kirchhoff, S., Mester, P., Meyer-Wachsmuth, I., Nuryanto, A., & Timm, J. (2009) Genetic population structure of the blue starfish *Linckia laevigata* and its gastropod ectoparasite *Thyca crystallina. Marine Ecology Progress Series*, **396**, 211-219.

Lacap, C.D.A. Vermaat, J.E., Rollon, R.N., & Nacorda, H.M. (2002) Propagule dispersal of the SE Asian seagrasses *Enhalus acoroides* and *Thalassia hemprichii*. *Marine Ecology Progress Series*, **235**, 75-80.

Lim, H.C. Ahmad, A.B., Nuruddin, A.A., & Nor, S.A.M. (2014) Cytochrome *b* gene reveals panmixia among Japanese Threadfin Bream, *Nemipterus japonicus* (Bloch, 1791) populations along the coasts of Peninsular Malaysia and provides evidence of a cryptic species. *Mitochondrial DNA*, (Early publication), doi:10.3109/19401736.2014.908354.

Lin, C.H., Soong, K., & Fan, T.Y. (2013) Hourglass mechanism with temperature compensation in the diel periodicity of planulation of the coral, *Seriatopora hystrix*. *PloS ONE*, **8**, 5, e64584.

Lin, G., Lo, L.C., & Yue, G.H. (2012) Genetic variations in populations from farms and natural habitats of Asian green mussel, *Perna viridis*, in Singapore inferred from nine microsatellite markers. *Journal of the World Aquaculture Society*, **43**, 270-277.

Lind, C.E., Evans, B.S., Elphinstone, M.S., Taylor, J.J.U., & Jerry, D.R. (2012) Phylogeography of a pearl oyster (*Pinctada maxima*) across the Indo-Australian Archipelago: evidence of strong regional structure and population expansions but no phylogenetic breaks. *Biological Journal of the Linnaean Society*, **107**, 632-646.

Lind, C.E., Evans, B.S., Taylor, J.J.U., & Jerry, D.R. (2007) Population genetics of a marine bivalve, *Pinctada maxima*, throughout the Indo-Australian Archipelago shows differentiation and decreased diversity at range limits. *Molecular Ecology*, **16**, 5193-5203.

Lourie, S.A., Foster, S.J., Cooper, E.W.T., & Vincent, A.C.J. (2004) *A Guide to the Identification of Seahorses*. Project Seahorse and TRAFFIC North America. University of British Columbia and World Wildlife Fund, Washington D.C.

Lourie, S.A., Green, D.M., & Vincent, A.C.J. (2005) Dispersal, habitat differences, and comparative phylogeography of Southeast Asian seahorses (Syngnathidae: *Hippocampus*). *Molecular Ecology*, **14**, 1073-1094.

Magsino, R.M. & Juinio-Menez, M.A. (2008) The influence of contrasting life history traits and oceanic processes on genetic structuring of rabbitfish populations *Siganus argenteus* and *Siganus fuscescens* along the eastern Philippine coasts. *Marine Biology*, **154**, 519-532.

Matsuda, H., Hamano, T., & Nagasawa, K. (2013) Growth and reproductive cycle of *Hypermastus tokunagai* (Caenogastropoda: Eulimidae), an ectoparasite of the sand dollar *Scaphechinus mirabilis* (Clypeasteroida: Scutellidae) in the Seto Inland Sea, Japan. *Journal of the Marine Biological Association of the United Kingdom*, **93**, 1041-1051.

McCormick, M.I. (1999) Delayed metamorphosis of a tropical reef fish *(Acanthurus triostegus):* a field experiment. *Marine Ecology Progress Series*, **176**, 25-38.

Mellin, C., Galzin, R. Ponton, D., & Vigliola, L. (2009) Detecting age-structured effects in growth performance of coral reef fish juveniles. *Aquatic Biology*, **6**, 31-39.

Mirams, A.G.K., Treml, E.A., Shields, J.L., Liggins, L., & Riginos, C. (2011) Vicariance and dispersal across an intermittent barrier: population genetic structure of marine animals across the Torres Strait land bridge. *Coral Reefs*, **30**, 937-949.

Motoh, H. (1984) Biology and ecology of Penaeus monodon. *In*: Taki Y., Primavera J. H. & Llobrera J. A. (*Eds*.). *Proceedings of the First International Conference on the Culture of Penaeid Prawns/Shrimps*, 4-7 December 1984, Iloilo City, Philippines. pp. 27-36.

Nabuab, F.N., Ledesma-Fernandez, L., & del Norte-Campos, A. (2010). Reproductive biology of the short-necked clam, *Paphia undulata* (Born 1778) from Southern Negros Occidental, Central Philippines. *Science Diliman*, **22**, 31-40.

Nair, M. & Appukuttan, K.K. (2003) Effect of temperature on the development, growth, survival and settlement of green mussel *Perna viridis* (Linnaeus, 1758). *Aquaculture Research*, **34**, 1037-1045.

Nakajima, Y., Matsuki, Y., Lian, C., Fortes, M.D., Uy, W. D., Campos, W. L., Nakaoka, M., & Nadaoka, K. (2014) The Kuroshio Current influences genetic diversity and population genetic structure of a tropical seagrass, *Enhalus acoroides*. *Molecular Ecology*, **23**, 6029-6044.

Nishikawa, A., Katoh, M., & Sakai, K. (2003) Larval settlement rates and gene flow of broadcast-spawning (*Acropora tenuis*) and planula-brooding (*Stylophora pistillata*) corals. *Marine Ecology Progress Series*, **256**, 87-97.

Nuryanto, A. & Kochzius, M. (2009) Highly restricted gene flow and deep evolutionary lineages in the giant clam *Tridacna maxima*. *Coral Reefs*, **28**, 607-619.

Ovenden, J.R., Kashiwagi, T., Broderick, D. Giles, J., & Salini, J. (2009) The extent of population genetic subdivision differs among four co-distributed shark species in the Indo-Australian archipelago. *BMC Evolutionary Biology*, **9**, 40.

Ovenden, J.R., Morgan, J.A.T., Street, R., Tobin, A., Simpfendorfer, C., Macbeth, W., & Welch, D. (2011). Negligible evidence for regional genetic population structure for two shark species *Rhizoprionodon acutus* (Ruppell, 1837) and *Sphyrna lewini* (Griffith & Smith, 1834) with contrasting biology. *Marine Biology*, **158**, 1497-1509.

Ovenden, J.R., Salini, J., O'Connor, S., & Street, R. (2004) Pronounced genetic population structure in a potentially vagile fish species (*Pristipomoides multidens*, Teleostei; Perciformes; Lutjanidae) from the East Indies triangle. *Molecular Ecology*, **13**, 1991-1999.

Peverell, S.C. (2005) Distribution of sawfishes (Pristidae) in the Queensland Gulf of Carpentaria, Australia, with notes on sawfish ecology. *Environmental Biology of Fishes*, **73**, 391-402.

Phillips, N.M., Chaplin, J.A., Morgan, D.L., & Peverell, S.C. (2011) Population genetic structure and genetic diversity of three critically endangered *Pristis* sawfishes in Australian waters. *Marine Biology*, **158**, 903-915.

Phinchongsakuldit, J., Chaipakdee, P., Collins, J.F., Jaroensutasinee, M., & Brookfield, J.F.Y. (2013) Population genetics of cobia (*Rachycentron canadum*) in the Gulf of Thailand and Andaman Sea: fisheries management implications. *Aquaculture International*, **21**, 197-217.

Pierce, S.J., Pardo, S.A., & Bennet, M.B. (2009). Reproduction of the blue-spotted maskray *Neotrygon kuhlii* (Myliobatoidei: Dasyatidae) in south-east Queensland, Australia. *Journal of Fish Biology*, **74**, 1291-1301.

Pillai, C.S.G., Mohan, M, & Koya, K.K. (1985) Ecology and biology of the white-tailed humbug *Dascyllus aruanus* (Pomacentriade, Pisces) from Minicoy Atoll. *Journal of Marine Biology Association of India*, **27**, 113-123.

Ravago-Gotanco, R.G., & Juinio-Menez, M.A. (2010) Phylogeography of the mottled spinefoot *Siganus fuscescens*: Pleistocene divergence and limited genetic connectivity across the Philippine archipelago. *Molecular Ecology,* **19**, 4520-4534.

Ravago-Gotanco, R.G., Magsino, R.M., & Juinio-Meñez, M.A. (2007) Influence of the North Equatorial Current on the population genetic structure of *Tridacna crocea* (Mollusca: Tridacnidae) along the eastern Philippine seaboard. *Marine Ecology Progress Series*, **336**, 161-168.

Raynal, M., Crandall, E.D., Barber, P.H., Mahardika, G.N., Lagman, M.C., & Carpenter, K.E. (2014) Basin isolation and oceanographic features influencing lineage divergence in the humbug damselfih *(Dascyllus aruanus)* in the Coral Triangle. *Bulletin of Marine Science*, **90**, 513-532.

Richmond, R.H., & Hunter, C. L. (1990) Reproduction and recruitment of corals: comparisons among the Caribbean, the Tropical Pacific, and the Red Sea. *Marine Ecology Progress Series*, **60**, 185-203.

Robinson1, J., Isidore1, M., Marguerite1, M.A., Öhman, M.C., & Payet, R.J. (2004). Spatial and temporal distribution of reef fish spawning aggregations in the Seychelles – An interview-based survey of artisanal fishers. *Western Indian Ocean Journal of Marine Science*, **3**, 63-69.

Rohfritsch, A., & Borsa, P. (2005) Genetic structure of Indian scad mackerel *Decapterus russelli*: Pleistocene vicariance and secondary contact in the Central Indo-West Pacific Seas. *Heredity*, **95**, 315-326.

Rose, R.A., & Baker, S.B. (1994) Larval and spat culture of the Western Australian silverlip pearl oyster, *Pinctada maxima* (Jameson) (Molusca: Pteriiidae). *Aquaculture*, **126**, 35-50.

Rose, R.A., Dybdahl, R.E., & Harders, S. (1990). Reproductive cycle of the Western Australian silverlip pearl oyster, *Pinctada maxima* (Jameson) (Molusca: Pteriiidae). *Journal of Shellfish Research*, **9**, 261-272.

Salze, G., Craig, S.R., Smiths, B.H., Smith, E.P., & McLean, E. (2011) Morphological development of larval cobia *Rachycentron canadum* and the influence of dietary taurine supplementation. *Journal of Fish Biology,* **78**, 1470-1491.

Scheibling, R. E. & Metaxas, A. (2008) Abundance, spatial distribution, and size structure of the sea star *Protoreaster nodosus* in Palau, with notes on feeding and reproduction. *Bulletin of Marine Science,* **82**, 221-235

Schmidt-Roach, S., Miller, K.J., Woolsey, E., Gerlach, G., & Baird, A.H. (2012) Broadcast spawning by *Pocillopora* species on the Great Barrier Reef. *PloS ONE,* **7**, 12, e50847.

Sekiguchi, K., Seshimo, H., & Sugita, H. (1988) Post-embryonic development of the horseshoe crab. *Biological Bulletin*, **174**, 337-345.

Shoji, J. & Tanaka, M. (2006) Growth-selective survival in piscivorous larvae of Japanese Spanish mackerel *Scomberomorus niphonius*: early selection and significance of ichthyoplankton prey supply. *Marine Ecology Progress Series,* **321**, 245-254.

Srijaya, T.C., Pradeep, P.J., Hassan, A., Chaterji, A., Shaharom, F., & Jeffs, A. (2014) Oxygen consumption in trilobite larvae of the mangrove horseshoe crab (*Carcinoscorpius rotundicauda*; Latreille, 1802): effect of temperature, salinity, pH, and light–dark cycle. *International Aquatic Research*, **6**, 60.

Starger, C.J., Barber, P.H., Erdmann, M.V.N, Toha, A.H.A, & Baker, A.C. (2013). Strong genetic structure among coral populations within a conservation priority region, the Bird's Head Seascape (Papua and West Papua, Indonesia). *PeerJ PrePrints*, **1**, e25v1.

Stevens, J.D., & Wiley, P.D. (1986) Biology of two commercially important carcharhinid sharks from Northern Australia. *Australian Journal of Marine and Freshwater Research*, **37**, 671-688.

Sugama, K., Haryanti, Benzie, J.A.H., & Ballmen, E. (2002) Genetic variation and population structure of the giant tiger prawn, *Penaeus monodon*, in Indonesia. *Aquaculture*, **205**, 37-48.

Takeyama, T., Okuda, N., & Yanagisawa, Y. (2002) Seasonal pattern of filial cannibalism by *Apogon doederleini* mouthbrooding males. *Journal of Fish Biology*, **61**, 633-644.

Tanabe, T. (2002) Studies on the early life ecology of skipjack tuna, *Katsuwonus pelamis*, in the tropical western-north Pacific. *Bulletin of Fisheries Research Agency*, **3**, 63-132.

Timm, J., & Kochzius, M. (2008) Geological history and oceanography of the Indo-Malay Archipelago shape the genetic population structure in the false clown anemonefish (*Amphiprion ocellaris*). *Molecular Ecology*, **17**, 3999-4014

Timm, J., Planes, S., & Kochzius, M. (2012) High similarity of genetic population structure in the false clown anemonefish (*Amphiprion ocellaris*) found in microsatellite and mitochondrial control region analysis. *Conservation Genetics*, **13**, 693-706.

Tisera, W.L., Naguit, M.R.A., Rehatta, B.M., & Calumpong, H. (2012) Ecology and genetic structure of giant clams around Savu Sea, East Nusa Tenggara Province, Indonesia. *Asian Journal of Biodiversity*, **3**, 174-194.

Underwood, J.N. (2009) Genetic diversity and divergence among coastal and offshore reefs in a hard coral depend on geographic discontinuity and oceanic currents. *Evolutionary Applications*, **2**, 222-233.

Underwood, J.N., Smith, L.D., van Open, J.H., & Gilmour, J.P. (2007) Multiple scales of genetic connectivity in a brooding coral on isolated reefs following catastrophic bleaching. *Molecular Ecology*, **16**, 771-784.

Ver, L.M.M. (1986). Early development of *Crassostrea iredalei* (Faustino, 1932) (Bivalvia, Osteridae), with notes of the structure of the larval hinge. *The Veliger*, **29**, 78-85.

Wang, L., Meng, Z., Liu, X., Zhang, Y., & Lin, H. (2011) Genetic diversity and differentiation of the orange-spotted grouper (*Epinephelus coioides*) between and within cultured stocks and wild populations inferred from microsatellite DNA analysis. *International Journal of Molecular Sciences*, **12**, 4378-4394.

Wanna, W., Rolland, J.L., Bonhomme, F., & Phongdara, A. (2004) Population genetic structure of *Penaeus merguiensis* in Thailand based on nuclear DNA variation. *Journal of Experimental Marine Biology and Ecology*, **311**, 63-78.

Wear, R. G. (1976) Larva of the commensal shrimp *Periclimenes (Periclimenes) soror* Nobili, 1904 (Crustacea: Decapoda: Pontoniinae) from Fiji, New Zealand. *Journal of Marine and Freshwater Research*, **10**, 527-532.

Wee, A.K.S., Takayama, K., Asakawa, T., Thompson , B., Onrizal, Sungkaew, S., Tung, N.Z., Nazre, M., Soe, K.K., Tan, H.T.W., Watano, Y., & Baba, S. Kajita, T., & Webb, E.L. (2014) Oceanic currents, not land masses, maintain the genetic structure of the mangrove *Rhizophora mucronata* Lam. (Rhizophoraceae) in Southeast Asia. *Journal of Biogeography*, **41**, 954-964.

Wellington, G.M. & Victor, B.C. (1989) Planktonic larval duration of one hundred species of Pacific and Atlantic damselfishes (Pomacentridae). *Marine Biology*, **101**, 557-567.

White, W.T. & Dharmadi (2007). Species and size compositions and reproductive biology of rays (Chondrichthyes, Batoidea) caught in target and non-target fisheries in eastern Indonesia. *Journal of Fish Biology*, **70**, 1809-1837.

Xu, Z., Primavera, J.H., Pena, L.D.D.L., Pettit, P., Bleak, J., Alcivar-Warren, A. (2001) Genetic diversity of wild and cultured Black Tiger Shrimp (*Penaeus monodon)* in the Philippines using microsatellites. *Aquaculture*, **199**, 13-40.

Yahya, A.F., Hyun, J.O., Lee, J.H., Kim, Y.Y., Lee, K.M., Hong, K.N., & Kim S.C. (2014) Genetic variation and population genetic structure of *Rhizophora apiculate* (Rhizophoraceae) in the greater Sunda Islands, Indonesia using microsatellite markers. *Journal of Plant Research*, **127**, 287-297.

Yamaguchi, M. (1977) Population structure, spawning, and growth of the coral reef asteroid *Linckia laevigata* (Linnaeus), *Pacific Science*, **31**, 13-30.

Yap, C.K., Cheng, W.H., Ong, C.C., & Tan, S.G. (2013) Heavy metal contamination and physical barrier are main causal agents for the genetic differentiation of *Perna viridis* populations in Peninsular Malaysia. *Sains Malaysiana*, **42**, 1557-1564.

Yap, C.K., Chong, C.M., & Tan, S.G (2011) Allozyme polymorphisms in horseshoe crabs, *Carcinoscorpius rotundicauda*, collected from polluted and unpolluted intertidal areas in Peninsular Malaysia. *Environmental Monitoring Assessment*, **174**, 389-400.

Yokoyama, K., Kamei, Y., Toda, M., Hirano, K., & Iwatsuki, Y. (1994) Reproductive behaviour, eggs, and larvae of a Caesionine fish, *Pterocaesio digramma,* observed in an aquarium. *Japanese Journal of Ichthyology*, **41**, 261-274.

Zacharia, S. & Kakati, V.S. (2004) Optimal salinity and temperature for early developmental stages of *Penaeus merguiensis* de Man. *Aquaculture*, **232**, 373-382.

Table A3.1. Likelihood and AIC scores of 24 nucleotide substitution models computed using jModelTest 2.1.7 (Darriba *et al.* 2012). -lnL: negative log likelihood; K: number of estimated parameters, AIC: Akaike Information Criterion; delta: AIC difference, weight: AIC weight, cumWeight: cumulative AIC weight.

Model	-lnL	K	AIC	delta	weight	cumWeight
HKY+I+G	4399.76	312	9423.52	0	0.9997450	0.999745
GTR+I+G	4404.04	316	9440.07	16.55	0.0002550	0.999999
SYM+I+G	4413.06	313	9452.12	28.60	0.0000006	1
K80+I+G	4421.06	309	9460.12	36.59	0	1
HKY+I	4425.83	311	9473.66	50.14	0	1
SYM+I	4429.12	312	9482.24	58.71	0	1
GTR+I	4430.42	315	9490.84	67.31	0	1
K80+I	4439.49	308	9494.99	71.46	0	1
SYM+G	4543.37	312	9710.74	287.22	0	1
GTR+G	4540.86	315	9711.71	288.19	0	1
HKY+G	4546.94	311	9715.88	292.35	0	1
K80+G	4550.37	308	9716.75	293.22	0	1
JC+I+G	4693.13	308	10002.27	578.74	0	1
F81+I+G	4695.89	311	10013.79	590.26	0	1
F81+I	4712.47	310	10044.93	621.41	0	1
JC+I	4717.11	307	10048.22	624.69	0	1
F81+G	4831.42	310	10282.85	859.32	0	1
JC+G	4840.80	307	10295.61	872.08	0	1
HKY	5190.05	310	11000.10	1576.58	0	1
GTR	5186.35	314	11000.70	1577.18	0	1
K80	5210.40	307	11034.79	1611.27	0	1
SYM	5206.92	311	11035.84	1612.32	0	1
F81	5441.64	309	11501.29	2077.76	0	1
JC	5462.82	306	11537.65	2114.12	0	1

Table A3.2. Network parameters constructed using the package *igraph* from the oceanographic connectivity between the sampling sites. Definition of the network parameters can be found in the Materials and Methods. Note that, for calculating *closeness* and *betweenness*, the package treats the connection weights as 'cost' instead of 'connection strength', thus the parameter values represent the cost needed to connect nodes.

Site ID	Site	Strength	Closeness	Transitivity	Betweenness
1	Biak	4767	0.008	0.78	61
2	Tual	15196	0.011	0.73	38
3	Ambon	26016	0.009	0.69	41
4	Kendari	22664	0.009	1.02	31
5	Bitung	27696	0.006	0.97	11
6	Palu	28279	0.006	1.01	0
7	Jepara	29961	0.006	0.89	0
8	Pari Is.	32488	0.008	0.87	71
9	Bangka	20520	0.005	0.78	15
10	Natuna	14758	0.006	0.72	14
11	Kupang	10923	0.011	0.98	16
12	Lombok	23356	0.011	1.03	0
13	Drini	18029	0.014	0.96	8
14	Padang	2262	0.012	0.69	63
15	Cocos Keeling	4086	0.004	1.56	3
16	Kimberley	11242	0.008	1.11	92
17	Exmouth	11021	0.008	0.95	82

Table A3.3. Migration estimates (number of migrants per generation) between groups of populations calculated in MIGRATE (Beerli and Felsenstein, 2001) using SNP data. 'Sunda' corresponds to pooled samples in the Sunda shelf clade, 'East' corresponds to the East clade, while 'Indian Ocean' corresponds to the Indian Ocean clade on the phylogenetic tree. Estimates are with confidence intervals.

From	To	2.5%	25%	75%	97.5%	Median	Mean
East	Sunda	22.67	51.33	92.67	143.33	77.67	79.88
East	Indian Ocean	0	21.33	58	109.33	47	50.83
Sunda	Indian Ocean	0	0	12.67	52.67	13	16.41
Sunda	East	0	2.67	17.33	36	15	12.57
Indian Ocean	Sunda	0	0	9.33	32.67	9.67	9.58
Indian Ocean	East	0	0	8	24.67	8.33	6.65

Table A3.4. Migration estimates (number of migrants per generation) between groups of populations calculated in MIGRATE (Beerli and Felsenstein, 2001) using microsatellite data. Samples were pooled based on five genetic clusters of populations from STRUCTURE analysis. Cluster A consists of samples from site 1-6; cluster B from site 8-10; cluster C from site 11; 12, & 14; cluster D from site 13 & 15; and cluster E from site 16-17. Estimates are with confidence intervals.

From	To	2.5%	25%	75%	97.5%	Median	Mean
Eastern Indonesia	Sunda Shelf	34.07	36.07	37.17	38.13	40.67	36.71
Eastern Indonesia	Lesser Sunda	25.93	29.8	31.37	32.53	39.53	32.37
Eastern Indonesia	Indian Ocean	24.33	25.73	26.63	27.47	28.87	26.66
Eastern Indonesia	Western Australia	18.53	19.2	20.5	21.8	22.6	25.6
Sunda Shelf	Eastern Indonesia	3.33	5.13	6.3	7.27	16.13	16.32
Sunda Shelf	Lesser Sunda	6.47	7.73	8.57	9.33	10.53	8.74
Sunda Shelf	Indian Ocean	10.67	12	12.97	13.87	15.13	14.01
Sunda Shelf	Western Australia	12.93	14.47	15.3	16.13	17.73	15.07
Lesser Sunda	Eastern Indonesia	1.87	2.8	4.03	5.2	12	6.12
Lesser Sunda	Sunda Shelf	14.67	16.13	17.17	18.13	22	17.52
Lesser Sunda	Indian Ocean	12.8	14.53	15.5	16.33	18.53	15.88
Lesser Sunda	Western Australia	2	4	4.9	5.73	7.73	5.99
Indian Ocean	Eastern Indonesia	2.2	3.13	3.97	4.67	5.6	3.97
Indian Ocean	Sunda Shelf	16.47	18.67	20.5	22.2	33.47	19.07
Indian Ocean	Lesser Sunda	0	0.2	1.17	2.07	3.13	4.64
Indian Ocean	Western Australia	6.93	7.87	9.23	10.6	13.33	13.22
Western Australia	Eastern Indonesia	2.6	3.53	4.37	5.07	6.07	4.35
Western Australia	Sunda Shelf	18.67	21.8	23.17	24.13	26.27	21.41
Western Australia	Lesser Sunda	6.8	8.8	9.83	10.8	12.93	11.05
Western Australia	Indian Ocean	2.33	3.8	4.63	5.4	6.87	5.31

Table A3.5. AMOVA result using SNPs and microsatellite dataset. For SNPs, sites were pooled based on the phylogenetic clades from MEGA analysis. For microsatellite data, sites were pooled according to 5 clusters from STRUCTURE analysis.

AMOVA	SNPs	Microsatellite
Variance (%)		
- *Among groups*	31.25**	10.08**
- *Among sites within groups*	55.97**	24.32**
- *Within sites*	12.78**	65.58**

**= p value <0.01

Table A3.6. Test of association between genetic differentiation (F_{ST}) with oceanographic connectivity (OC) and historical processes (HP) using distance-based redundancy analysis. Explained variation is based on adjusted R-squared.

Response variable	Predictors	Explained variation (%)	p-value
F_{ST}	HP	20.30	< 0.001
	OC	8.44	0.008

*db-RDA output from the package *vegan* in R

```
# dbRDA with fst.IAA as response variable
# biogeo3 is the HP (historical processes)
# oc.var is the OC (oceanographic connectivity)

> fst0 <- capscale(fst.IAA ~ 1, sqrt.dist=TRUE)
> fst1 <- capscale(fst.IAA ~ 1 + biogeo3 + oc.var, sqrt.dist=TRUE)';
> ordiR2step (fst0, fst1)

Step: R2.adj= 0
Call: fst.IAA ~ 1

                 R2.adjusted
+ biogeo3         0.20304169
<All variables>   0.12113303
+ oc.var          0.08442828
<none>            0.00000000

Call: capscale(formula = fst.IAA ~ 1, sqrt.dist = TRUE)

              Inertia Rank
Total            2.82
Unconstrained    2.82   16
Inertia is Unknown distance

Eigenvalues for unconstrained axes:
  MDS1   MDS2   MDS3   MDS4   MDS5   MDS6   MDS7   MDS8   MDS9
MDS10  MDS11  MDS12  MDS13
0.4850 0.3145 0.2909 0.2557 0.2349 0.2144 0.1855 0.1601 0.1322
0.1140 0.1081 0.1003 0.0862
 MDS14  MDS15  MDS16
0.0579 0.0547 0.0259

> ordistep(fst0, scope = formula (fst1), direction = 'forward',
permutations=how(nperm = 9999))

Start: fst.IAA ~ 1

          Df    AIC      F Pr(>F)
+ biogeo3  7 18.957 1.5823 0.0001 ***
+ oc.var   4 20.206 1.3689 0.0083 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Step: fst.IAA ~ biogeo3

         Df    AIC      F Pr(>F)
+ oc.var  4 18.627 0.7903 0.7994
```

```
Call: capscale(formula = fst.IAA ~ biogeo3, sqrt.dist = TRUE)

               Inertia Proportion Rank
Total           2.8204     1.0000
Constrained     1.5560     0.5517    7
Unconstrained   1.2643     0.4483    9
Inertia is Unknown distance

Eigenvalues for constrained axes:
  CAP1   CAP2   CAP3   CAP4   CAP5   CAP6   CAP7
0.4520 0.2892 0.2333 0.2019 0.1664 0.1190 0.0943

Eigenvalues for unconstrained axes:
MDS1    MDS2    MDS3    MDS4    MDS5    MDS6    MDS7    MDS8    MDS9
0.26340 0.21698 0.18609 0.15341 0.13389 0.10410 0.08836 0.07418
0.04391
```

Table A4.1. Environmental parameters recorded in each sampling site. Depth (m) was measured from the mean sea level. All site depths, except Bathurst and Longitude, are above the mean sea level. Sediment type was based on Wenworth (1922).

Site ID	Site	Depth (m)	Sediment type	Habitat type	Presence of coral	Number of other seagrass species
1	Bathurst Is.	0	Granule	Reef terrace	No	0
2	Longitude Is.	0	Sand-granule	Reef lagoon	No	0
3	Bedford Is. –South	4	Sand-granule	Reef terrace	Yes	2
4	Bedford Is. – North	1	Sand-granule	Reef terrace	Yes	2
5	Riptide Is.	4	Sand	Reef lagoon	Yes	1
6	Mermaid Is.	5	Sand-granule	Reef lagoon	Yes	2
7	Sunday Is. –South	1	Sand-granule	Reef terrace	Yes	2
8	Sunday Is. –North	3	Sand-granule	Reef terrace	Yes	4
9	Halls Pool	1	Sand-granule	Reef terrace	Yes	4
10	Talon Is.	3	Sand	Reef lagoon	Yes	2
11	Jackson Is.	1	Sand-granule	Reef terrace	Yes	2
12	Noyon	1	Sand	Reef lagoon	Yes	1
13	Shenton Bluff	2	Sand-granule	Reef lagoon	Yes	1

Table A4.2. Network parameters constructed using the package *igraph* from the oceanographic connectivity between the sampling sites. Definition of the network parameters can be found in the Materials and Methods. Note that, the package treats the connection weights as 'cost' instead of 'connection strength' for calculating *closeness* and *betweenness*, thus the parameter values represent the cost needed to connect nodes.

Site ID	Site	Strength	Closeness	Transitivity	Betweenness
1	Bathurst Is.	4.70	0.74	1.31	50
2	Longitude Is.	10.26	0.71	0.70	33
3	Bedford Is. –South	15.96	0.78	0.95	1
4	Bedford Is. – North	19.68	0.35	1.12	0
5	Riptide Is.	20.19	0.70	1.06	15
6	Mermaid Is.	16.53	0.66	0.96	0
7	Sunday Is. –South	20.32	0.37	1.03	1
8	Sunday Is. –North	18.46	0.34	0.83	11
9	Halls Pool	27.29	0.59	0.93	21
10	Talon Is.	25.31	0.61	0.87	12
11	Jackson Is.	22.04	0.42	0.80	4
12	Noyon	17.13	0.36	1.07	2
13	Shenton Bluff^{++}	18.81	0.52	0.76	0

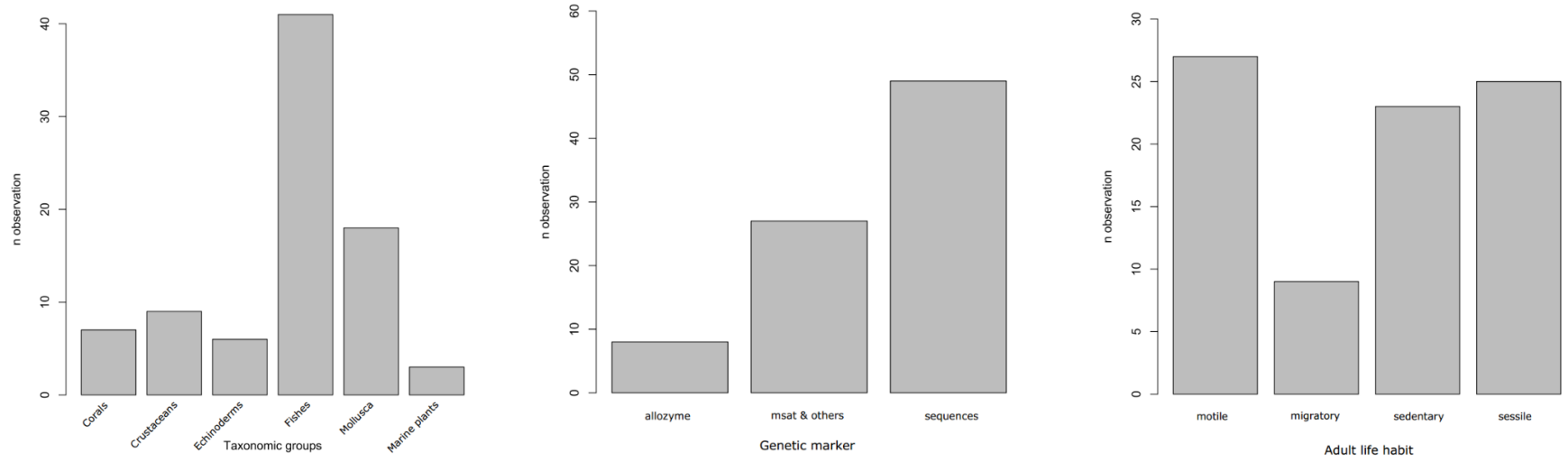

Figure A2.1. Number of data records based on categories in taxonomic groups (left panel), type of genetic markers (middle panel) and adult life habit (right panel).

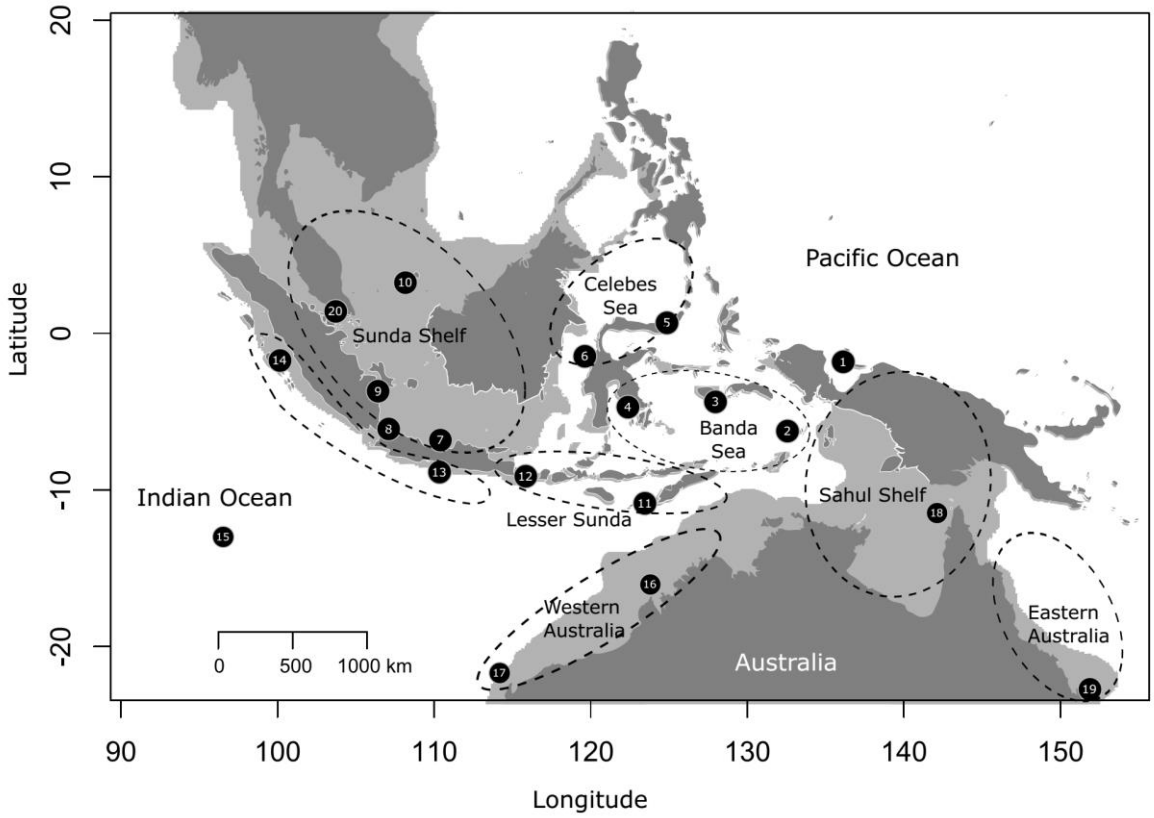

Figure A3.1. Map of the Indo-Australian Archipelago showing sampling sites within distinct regions (dashed circles) representing the impact of lowered sea levels on marine habitat and population connectivity. Dark grey areas indicate present-day island configuration and light grey areas indicate landmass configuration during the Pleistocene low sea level stands (100 m below present-day level) (Voris, 2000; Carpenter et al 2011; Lohman *et al*, 2011)

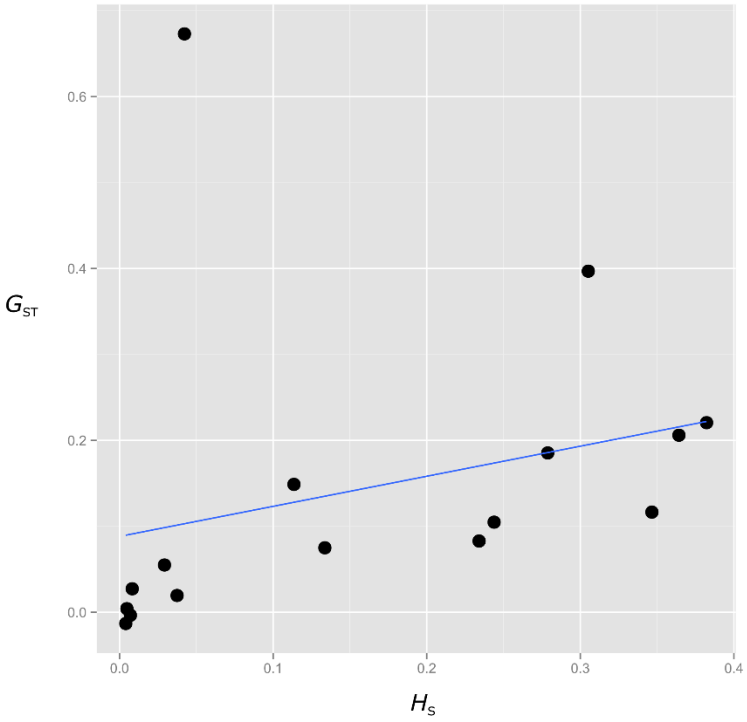

Figure A4.1. Main output of CoDiDi analysis, performed with 10000 permutation, to examine the mutation effect on genetic differentiation. The correlation coefficient (r_{GH}) between G_{ST} and H_S is 0.2878 (p values = 0.279, calculated from the Student's t test). The fitted linear regression (y = a +b.x) coefficient a= 0.0884 and b=0.3500.

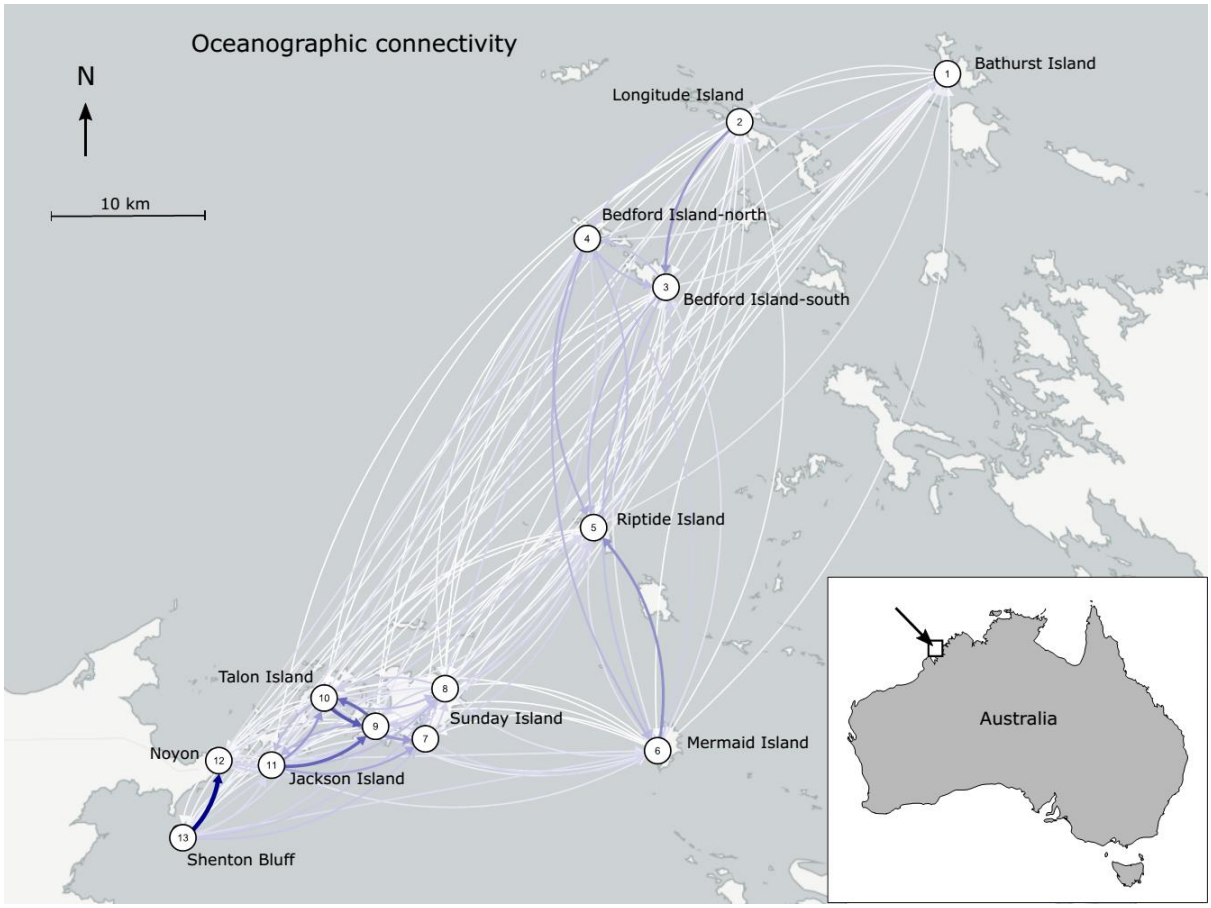

Figure A4.2. Oceanographic connectivity as the average number of particles released from site *i* that were tracked to be in site *j*, ranging from 0 to 7.49 particles/release period. Sampling sites (populations) were represented by numbers within circles (referred to Table 1). Number of particles were represented by curved lines. The thicker the lines, the higher level of connectivity between populations. The base map was obtained from OPENSTREETMAP contributors (https://www.openstreetmap.org/copyright).

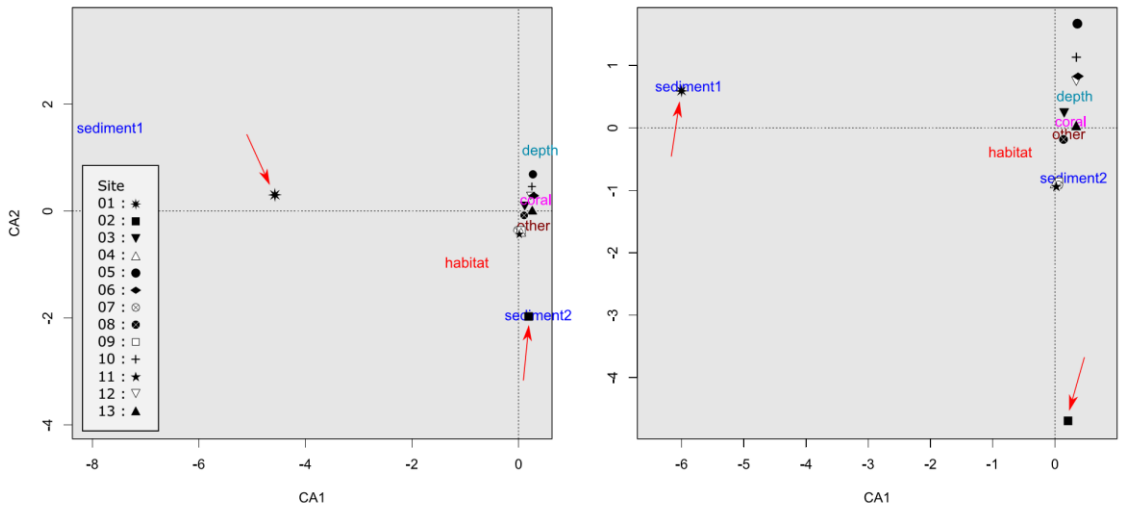

Figure A4.3. Ordination plot for the first two CA axes from the correspondence analysis (CCA) on the environmental variables using scaling method=1 (sites are scaled proportional by eigenvalues; left hand panel) and scaling method=2 (environmental variables are scaled by eigenvalues; right hand panel). Description about the sampling site ID and environmental variables can be found in the Supporting information Table 1. Red arrows points to Site 1 and Site 2 that are farther apart from the other sites. Sites that are closer together in the plots have more similar environmental characteristics, while sites that are farther apart have greater differences in environmental characteristics.

Appendix A3.1. Procedure for SNP analysis in Chapter 3.

This protocol for SNP analysis was outlined by Cross *et al* (2016). The analysis consisted of 5 main steps: 1) restriction/ligation digest, 2) pre-selective amplification, 3) selective amplification, 4) purification and size selection, and 5) SNP screening.

1). Restriction/ligation digest
DNA was digested with the restriction enzymes EcoRI HF and MseI in a 20 uL reaction for three to four hours at 37°C, followed by 65°C for 20 minutes to deactivate the enzyme. Then, a 20-uL mixture of T4 DNA ligase and the annealed adapters (EcoRI: forward CTCGTATACTGCGTACC, reverse AATTGGTACGCAGTA, and MseI: forward GACGATGAGTCCTGAG, reverse TACTCAGGACTCATC) was added into each sample, The reaction mixes were incubated overnight at 16°C or for four hours at room temperature.

2). Pre-selective amplification
The restriction/ligation reactions were diluted 1:10 and used as a template for an initial PCR. The amplification consisted of a 25 uL reaction for each sample containing DyNAzyme buffer, dNTPs, $MgCl_2$, adapter primers (Table S3, S4), DyNAzyme polymerase, and diluted restriction-ligation reaction, run on a thermalcycler for 95°C for five minutes, then 20-30 cycles at 95°C for 30 seconds, 56°C for 30 seconds, and 72°C for one (to two) minute; followed by an extension step of ten minutes at 60°C. The preselective PCR products were then diluted 1:20 in water.

3. Selective amplification
The diluted preselective PCR products were amplified again using fusion primers. In this second PCR, the same preselective primers were used, but in addition to the adapter sequence, a 7 bp barcode sequence followed by the Primer key A (on the *Eco*R1 adapter) and a six bp barcode followed by the Primer key P1 (on the *Mse*I adapter) were added on the 5′ end. To reduce complexity further, primers had an additional 2-4 bases added at the 5′ or 3′ end. Selective amplification was performed with a short thermal cycling run: 95°C for 9 minutes, followed by 5 cycles of 95°C for 30 seconds, 56°C for 30 seconds, and 72°C for 45 seconds; then 15 cycles of 95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 30 seconds; ending with a 72°C extension for 7 minutes.

4. Purification and size selection
The selective amplification products were then pooled (5-10 μl each) and purified using either AMPure XP (Agencourt, Beckman Coulter, Inc., Brea, California). The pooled reaction mixture was then size-selected using an E-Gel (Life Technologies, Carlsbad, California), in order to optimize sequencing by selecting the products at the appropriate size range for Ion Torrent sequencing: 150-250 bp for 200 bp chemistry, ~250-450 for 400 bp chemistry.

The size-selected amplification pool was then purified up to three times using AMPure XP, and then quantified using a 2200 TapeStation (Agilent, Santa Clara,

California) with a high-sensitivity ScreenTape. The resulting quantification was then used to calculate the dilution of the library, which was in the range of 9 - 14 pmol l^{-1}. Emulsion PCR and enrichment were then conducted on the diluted mixture using the Ion OneTouch 2 System (Life Technologies) according to manufacturer's specifications, and the enrichment was checked using the IonSphere Quality Control Kit (Life Technologies). Next-generation sequencing on the Ion Torrent PGM or Proton sequencing machines followed.

5. SNP screening

Sequencing reads were analysed (de-multiplexing, trimming, and assembling) using CLC-Genomic Workbench (Qiagen, Venlo, The Netherlands) to generate a 'provisional reference genome'. The assembly included only full-length reads (where both barcodes were present). Each individual's reads were mapped onto this reference and the consensus sequences were extracted. 530 contigs that had at least 130 reads were selected. Ambiguous bases were called if the least frequent base was present in at least 25% of the reads, using a minimal read depth of 20 reads. SNP loci were manually selected in Geneious R7 (Biomatters, Auckland, New Zealand) and identified 113 contigs with SNP loci. The selected contigs were trimmed, such that each contig contained a single SNP locus that was flanked with invariant sites of up to 15 bp length. The contigs (including the invariant sites) were concatenated and the final SNP dataset consisted of 154 samples, each consisting of a 1435-bp nucleotide sequence.

Reference

Cross H, Biffen E, Dijk K van, Lowe A, Waycott M (2016) Effective application of next generation sequencing (NGS) approaches in systematics and population genetics: Case studies in Eucalyptus and Acacia. *Australian Systematic Botany*, Accepted 08 September 2016.

Appendix A3.2. P-sex values per site considering F_{IS} calculated using Rclone (Bailleul et al, 2016).

Site 1: Biak
All samples are distinct genets

==========

Site 2: Tual

	genet	psexFis
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40	39	9.24058444600597e-09
41		
42		
43		
44		
45		
46		
47		
48		

==========

Site 3: Ambon

genet	psexFis
1	
2	
3	
4	
5	4 1.83883729328096e-07
6	
7	
8	
9	
10	3 4.50359447134744e-06
11	
12	11 4.70505561781088e-07
13	11 1.08332734467752e-13
14	
15	
16	14 2.79351382888123e-05
17	14 3.81894814769598e-10
18	
19	18 6.79218831389802e-05
20	18 2.25776592724388e-09
21	14 3.4048667536502e-15
22	14 2.22616784830622e-20
23	
24	18 4.89453607839558e-14
25	
26	
27	
28	
29	
30	
31	30 0.00549694875552209
32	
33	30 1.48687253520776e-05
34	30 2.6229451784512e-08
35	32 0.00130303319800964
36	30 3.39317386925797e-11
37	30 3.43185381538437e-14
38	
39	30 2.8252119548669e-17
40	30 1.94608169247325e-20
41	
42	38
43	30 1.1443405473067e-23
44	30 5.83178627172542e-27
45	30 2.60621002449561e-30
46	
47	

==========

Site 4: Kendari

genet	psexFis
1	
2	
3	
4	2 0.000154576410058764
5	
6	2 1.1694551042169e-08

7		
8	3	5.22356861053121e-05
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25	21	0.000185689023052364
26	22	5.38895204201929e-05
27		
28	21	1.68765278046086e-08
29		
30		
31		
32		
33		
34	20	0.000116505098474151
35		
36		
37		
38		
39		
40		
41		
42		
43	42	5.37962074911193e-06
44	21	1.00033022272079e-12
45		
46		
47		

==========

Site 5: Bitung

All samples are distinct genets

==========

Site 6: Palu

	genet	psexFis
1		
2		
3		
4		
5		
6		
7		
8		
9		

10		
11	10	7.79637127371257e-05
12	10	2.97608624055445e-09
13		
14		
15		
16		
17		
18		
19		
20	17	2.35954215001963e-07
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32	31	1.38833354316886e-06
33		
34	29	0.00115191661496907
35		
36	29	6.50383555867962e-07
37		
38		
39		
40		
41		
42		
43	25	0.000295061225335999
44		
45	7	3.0394212428419e-05
46	7	4.52294820367773e-10
47		
48		

==========

Site 7: Jepara

genet		psexFis
1		
2	1	0.00300041838602622
3	1	4.41667326409798e-06
4	1	4.23796758062258e-09
5	1	2.98055424673367e-12
6		
7		
8		
9		
10	6	0.00113019645493398
11	6	6.25495162970066e-07
12		
13	12	0.00250702488470776
14	12	3.08200595204228e-06
15	12	2.4697739471256e-09
16		
17		
18		
19		
20		

```
21     20         3.25202091879363e-05
22
23
24
25
26
27
28
29
30
31     28         5.24209370179957e-09
32
33
34
35
36
37      9
38
39      6         2.25653915711027e-10
40      6         5.96676364749754e-14
41      6         1.23283688066159e-17
42
43      6         2.07217161825059e-21
44
45      1         1.63797493852386e-15
46
```

==========

Site 8: Pari Island

```
genet           psexFis
1
2
3
4
5
6
7
8
9       6         5.16434629528455e-08
10      6         1.30574195796231e-15
11
12
13
14      4         1.69549329241946e-06
15
16
17
18
19
20     15         1.25915258393092e-08
21
22
23
24
25
26
27
28     27         2.56919791082093e-06
29
30
31
32
33
```

```
34
35
36
37     36         2.99440354989516e-05
38
39     38         1.35903602937912e-07
40     38         9.04250223620836e-15
41
42
43     36         4.38995721665321e-10
44
45     36         4.19931991155284e-15
46
47     36         2.94722695245122e-20
48     46         1.29133032819177e-07
```

==========

Site 9: Bangka

```
genet           psexFis
1
2       1         1.52250681251997e-20
3       1         6.58544475305125e-19
4       1         1.85857328999974e-17
5       1         3.84848871852956e-16
6       1         6.23348464200043e-15
7       1         8.22254204452578e-14
8
9       8         0.0504834301827729
10      8         0.117924537200774
11      8         0.179733163984266
12      8         0.200987109387011
13      8         0.175807880600839
14      1         9.08061697037789e-13
15      1         8.56578620689977e-12
16      1         7.00717100703414e-11
17      1         5.02996921761737e-10
18      1         3.19826350619207e-09
19      1         1.81506726565026e-08
20      1         9.25145150728886e-08
21      1         4.25704155229611e-07
22      1         1.7760433058264e-06
23      1         6.74226024979907e-06
24      1         2.33595554180527e-05
25      1         7.40477327587992e-05
26      1         0.00021519759981846
27      1         0.000574332548490027
28      1         0.00140948406309756
29      1         0.00318389881194334
30
31     30         0.352987944462074
32      1         0.00662464747630388
33
34      1         0.0127013400517567
35      1         0.0224428899765484
36      1         0.0365419124106024
37      1         0.0548034870550455
38      1         0.0756532274163385
39
40      1         0.0960323290102239
41      1         0.111945801919132
42      1         0.119640061293594
43      1         0.116985914336819
44      1         0.104399308093703
```

```
45      1         0.0847750327551693
46      1         0.0624145669252765
47      1         0.0414844003497627
48      1         0.0247640882293947
```

==========

Site 10: Natuna

```
genet           psexFis
1
2
3
4
5       4         2.84086788153841e-05
6
7
8
9
10
11      6         0.000109082142790278
12      8         9.86097031859839e-09
13
14
15
16
17
18
19
20
21     18         1.2951690587374e-06
22     20         1.45688174052538e-06
23
24
25
26
27
28
29
30
31
32
33
34
35     34         4.88385622494812e-07
36
37     25         2.92743538873748e-08
38
39
40     25         4.19377018596881e-16
41
42
43
44      2         2.61620008732933e-05
45
46     45         7.94331106444096e-07
47
```

==========

Site 11: Kupang

```
genet           psexFis
```

```
1
2       1        3.49348681879857e-05
3       1        5.97530413749759e-10
4
5
6
7       4        1.11598705241058e-05
8
9       4        6.09747156340238e-11
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30     29        1.85678790027278e-06
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
```

==========

Site 12: Lombok

```
   genet       psexFis
1
2
3
4
5
6
7
8
9
10
11
```

```
12     11        7.67539303834948e-06
13
14
15
16     14        9.44169950285318e-06
17
18     17        4.14197024211811e-05
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
```

==========

Site 13: Drini

```
   genet        psexFis
1
2      1        0.0300773724193837
3      1        0.000456852465251604
4
5      1        4.52773256542905e-06
6      1        3.29231595459202e-08
7
8      1        1.87263180816108e-10
9      1        8.67436751487075e-13
10     1        3.36401006051884e-15
11
12     1        1.11434424081144e-17
13
14
15     1        3.20114297631791e-20
16
17
18
19
20
21     17       0.0136773013783765
22
23
```

```
24    23        0.0287839112191742
25
26
27
28
29
30    29        0.0113146444686211
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
```

==========

Site 14: Padang

```
   genet         psexFis
1
2
3
4
5
6
7
8
9      5        0.0016566472612002
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27    19        0.0140196551799368
28    20        0.0206109259273891
29    19        9.75617441021191e-05
30    20        0.000212307100146124
31    23        0.0325723067224044
32     3        0.0617054999741369
33
34
```

```
35
36
37
38
39
40
41
42
43
44
45
46
47
```

==========

Site 15: Cocos Keeling

```
   genet        psexFis
1
2
3
4
5
6
7
8
9       7        0.0183605839408152
10
11
12
13     12        0.00457403290372556
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35     34        0.0191084782025655
36
37
38
39
40
41     24        0.270245039265622
42
43
44
45
46
```

```
47
```

==========

Site 16: Kimberley

```
   genet       psexFis
1
2
3
4
5
6       5       0.0503171726064935
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28     27       0.0491641426657541
29
30
31
32
33     32       0.330362094365338
34     18       0.304898818010204
35
36
37
38
39
40
41
42
43
44
45
46
47
48
```

==========

Site 17: Exmouth

```
   genet       psexFis
1
2
```

```
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25      5         0.074841502966956
26      5         0.00297116962057343
27      5         7.68884834547267e-05
28
29
30
31      7         0.00974549077531879
32      6         0.010238899515563
33
34
35
36
37
38     20         0.270975067469644
39     13         0.129226348721793
40
41
42
43
44
45
46
```

Appendix A4.1. P-sex values per site considering F_{IS} calculated using Rclone (Bailleul et al, 2016).

```
Bathurst Is.

   genet       psexFis
1
2
3
4
5
6       4        0.205536703516333
7
8       2        0.0921203526850636
9       5        0.22186971277447
10      4        0.26641551478484
11      4        0.22227898886532
12
13      5        0.272310469774344
14
15
16
17      5        0.215129234573599
18     14        0.350313785939881
19      5        0.122914047735021
20     14        0.24831868733486
21     14        0.113300115741084
22
23     22        0.342388705412075
24     12        0.373175211328671
25
26      4        0.134123324216544
27      4        0.0623461694871239
28     14        0.0373868004034803
29
30
=============

Longitude Is.
   genet       psexFis
1
2
3
4       1        0.00417895036981363
5
6
7
8       3        0.0597814214236594
9
10
11     10         0.0788093679408484
12
13     12         0.298031355784719
14
15     12        0.0701748927921248
16      9        0.0198450301798994
17
18     17         0.183158878042981
19     17        0.0206706679943432
20
21
22
23     17        0.00152212305190144
```

```
24
25     17        8.22358073322403e-05
26     17        3.47538007769453e-06
27
28     27        0.363513613948272
29
30
31     27        0.14562620254712
32     30        0.151476676338894
33
34     27        0.0380651095313724
35     27        0.00730012990486062
36     27        0.00109512638528754
37     27        0.000133792703685004
38
39     30        0.0134656075820429
40     27        1.36846805101044e-05
41     27        1.19558240343302e-06
42     27        9.05832666885192e-08
43     27        6.02231638597489e-09
44
45     27        3.54654489043445e-10
46
47
48     17        1.19612946054609e-07


==============
Bedford Is.-South
   genet       psexFis
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18     17        0.263849617212677
19
20
21
22
23
24
25
26
27
28
29
30
31     30        0.235876552415548
32
33     32        0.290857183989338
34
35     34        0.278372014858292
36
37     30        0.0377198707602828
```

```
38     36        0.326086962372233
39     30        0.00393572128294006
40     28        0.37169786873576
41     36        0.0927384726887454
42
43
44     34        0.276247943321268
45     42        0.267610392037565
46
47
48


=============
Bedford Is.-North
   genet       psexFis
1
2       1        0.279920888453883
3
4
5       4        0.269921623693441
6
7       6        0.203184677325755
8       6        0.0263130701487624
9       6        0.0022234164176297
10
11
12     11        0.369953678403679
13     11        0.167152708981296
14     11        0.0492774474800559
15     11        0.0106585697979822
16
17
18
19     18        0.0929084097130379
20
21     17        0.295174180401435
22
23     17        0.274136487403538
24     17        0.166120808494359
25
26     25        0.370543867812566
27     25        0.170540983026394
28
29     28        0.328332780440611
30     28        0.0949104968411244
31
32     31        0.229789597357859
33     17        0.073857937207115
34
35     34        0.0403880838388765
36
37
38
39     10        0.318675610006982
40     10        0.266721458875978
41
42     17        0.0256862355929993
43
44
45
46     17        0.00727508852333863
47     38        0.33043542784608
48


=============
```

```
Riptide Is.
   genet        psexFis
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15     13        0.000343533723701224
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30     24        0.144936361956302
31
32
33
34
35
36
37     36        0.00643715915725498
38
39
40
41
42
43
44      4        0.16694092611612
45
46
47


==============
Mermaid Is.
   genet        psexFis
1
2
3
4
5
6       5        0.0503171726064935
7
8
9
10
11
12
13
```

```
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28     27        0.0491641426657541
29
30
31
32
33     32        0.330362094365338
34     18        0.304898818010204
35
36
37
38
39
40
41
42
43
44
45
46
47
48


==============
Sunday Is.-South
   genet       psexFis
1
2
3       1        0.279629887911877
4
5       4        0.142266016681224
6
7       1        0.0585770122262625
8
9       8        0.262536476706997
10
11      8        0.276501026105846
12      8        0.189824710269123
13      8        0.0955181245078357
14     10        0.239517690459116
15      8        0.0375568923370473
16
17
18      4        0.0117106353210418
19
20     19        0.182318877908043
21     19        0.251753659714714
22
23
24     22        0.0913998022216022
25     22        0.00452043416064378
26     16        0.0234172258196526
27     16        0.0672286455918518
```

28		
29		
30		
31		
32		
33		
34	33	0.238608503559374
35	10	0.273439436233714
36		
37		
38	10	0.203485565875237
39	10	0.110989810269808
40	10	0.0473046062143355
41	28	0.193040637914753
42	28	0.0233176593406201
43	19	0.226604666700801
44		
45	19	0.149499227617787
46	19	0.0770690259510372

==========

Sunday Is. -North

	genet	psexFis
1		
2	1	0.3054264382729
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13	12	0.267229247795763
14		
15	12	0.276430694637857
16	11	0.117838451712043
17		
18		
19	11	0.203073705297989
20	12	0.186576634804712
21	12	0.0923940898480529
22		
23	22	0.0970881688093894
24		
25		
26		
27	11	0.228343720628296
28		
29		
30	22	0.181344607205139
31		
32		
33	31	0.223558442715188
34	11	0.188382431762127
35		
36		
37		
38	5	0.370850718745682
39	5	0.172633357089
40	36	0.243237467410292
41	35	0.3527777228549
42	35	0.240731458693163

```
43     36        0.273902635361287
44
45     44        0.0518555736801186
46
47     46        0.0878361020870784
48      5        0.0524347282740952


==============
Halls Pool

   genet        psexFis
1
2
3       2        0.265626600784242
4
5
6
7
8
9
10
11      2        0.276383093458308
12
13      1        0.277436426817102
14
15
16
17
18      2        0.187637698278805
19
20
21
22     21        0.370494962719947
23
24
25     23        0.357421408528579
26     23        0.234384862895295
27
28      2        0.0934640650253111
29     19        0.00766374725808238
30      2        0.0364165990449515
31
32
33
34
35
36     35        0.191164304278173
37     35        0.256105377599512
38
39     38        0.308344879151253
40     38        0.270710593710339
41
42
43
44
45     38        0.155075245404881
46      1        0.276298799338928
47
48


==============
Talon Is.
   genet        psexFis
1
2       1        0.00719486208443701
```

```
3       1        0.0263249968241694
4
5
6
7
8
9       1        0.0628467604482503
10      1        0.110081264641328
11
12
13
14     11        0.0508078137810718
15
16
17      1        0.150825257003921
18      1        0.168294247252104
19      1        0.157216640919267
20     12        0.28261030106003
21     12        0.0603742522281371
22
23
24
25      1        0.125449915847561
26      1        0.0868092187749026
27      1        0.0527118728040149
28      1        0.0283515854231582
29      1        0.0136105545145975
30
31     30        0.227119118703294
32
33     30        0.0343659558494518
34     22        0.340035859579964
35     23        0.238831431305308
36     23        0.038905923224724
37     23        0.00413531583301445
38
39     38        0.188315538229287
40     38        0.0220366824169918
41     23        0.000322490992217716
42     22        0.253568764468387
43     22        0.123377368880728
44
45      1        0.00586830880603148
46      1        0.00228418334114529
47      1        0.000806114270562384
48     22        0.0440444508608379


============
Jackson Is.
   genet        psexFis
1
2
3
4
5
6
7
8
9
10
11     10        0.114694841824451
12
13     12        0.327690922388124
14     12        0.262106133740713
15
16
```

```
17
18
19
20     17         0.368239832228302
21
22
23
24
25     21         0.141429470089109
26
27
28     27         0.360698197436042
29
30
31
32     31         0.00887036259224075
33
34
35     24         0.292333027977395
36
37      4         0.363917886250809
38
39
40     39         0.0514850073446829
41
42     41        0.327300820435
43
44     33         0.267075226419027
45     31         3.88683395532614e-05
46     33         0.0518211545731621
47
48     21         0.223922685508452


============
Noyon
   genet        psexFis
1
2
3
4
5
6
7       6         8.94223320529311e-06
8       3         0.00087922972775287
9       1         0.356173278244628
10
11
12      6         7.38139921637657e-05
13      6         0.00039755764725933
14      6         0.00157100466506888
15      6         0.00485607054578928
16
17      1         0.236214830706712
18
19      6         0.0122243853823934
20     10         0.026384416546259
21      1         0.102216664142774
22
23      1         0.032452801352397
24
25     24         0.0937462728458237
26     16         0.3420709303167
27      6         0.0257633968752041
28      6         0.046379048430041
29
30     24         0.177517915722783
```

```
31      6        0.0724042661614549
32     24        0.219330528342814
33
34      6        0.0991867370870361
35      6        0.120356446153723
36
37     10        0.0734360971505551
38      6        0.130351092071287
39
40     36        0.198527304469287
41     36        0.259601099775344
42      5        0.34353033471222
43
44      6        0.126793988337407
45     16        0.251849175564186
46      4        0.369149712107767
47      6        0.111343152121153
48      6        0.0886494428434392


=============
Shenton Bluff
   genet        psexFis
1
2       1        2.5294160027697e-05
3       1        0.00019004006741409
4       1        0.000931620002250906
5
6       1        0.00335079901074117
7       1        0.00942731681755357
8
9       8        0.354389344421293
10      1        0.0216004281345804
11      1        0.041435316574032
12
13      1        0.0678924658775133
14      1        0.0964708524773903
15      1        0.120286745567291
16      1        0.132851278452728
17      1        0.130961387527764
18      1        0.115946987091288
19      1        0.0926737054700909
20      1        0.0671585387159829
21      1        0.0442845511051518
22      1        0.0266508064573075
23      8        0.238664976791675
24      8        0.104873423290399
25      1        0.0146742730456886
26      1        0.00740766444356874
27      1        0.00343404795113783
28      1        0.00146386876222253
29      1        0.000574381281692159
30      1        0.000207588285500231
31      1        6.91335185160344e-05
32      1        2.121861072561e-05
33      1        6.00106713150716e-06
34      1        1.56330765108877e-06
35      1        3.74854534259025e-07
36
37     36        0.198751809136312
38     36        0.259702661132847
39     36        0.22141672551154
40     36        0.138503386797188
41     36        0.0677704723195125
42     36        0.0270056829473402
43     36        0.00900956581266697
44     36        0.00256740970651547
```

```
45    36        0.000634468562126307
46    36        0.000137585340111544
47    36        2.64277701050699e-05
48    36        4.53083707150435e-06
```