PATTERNS OF POPULATION GENETIC STRUCTURE AND CONNECTIVITY OF SQUAT LOBSTERS ASSOCIATED WITH VULNERABLE MARINE ECOSYSTEMS

BY

RUO-JIN YAN

A thesis submitted to the Victoria University of Wellington in fulfilment of the requirements for the degree of Doctor of Philosophy in Marine Biology.

Victoria University of Wellington

2019

*Dedicated to the memory of my grandfather **Chun-Fu Yan***

This thesis was conducted under the supervision of:

Prof. Jonathan P. A. Gardner

Victoria University of Wellington

Wellington, New Zealand

&

Prof. Ashley A. Rowden

Victoria University of Wellington

Wellington, New Zealand

National Institute of Water and Atmospheric Research

Wellington, New Zealand

Dr. Kareen E. Schnabel

National Institute of Water and Atmospheric Research

Wellington, New Zealand

Abstract

Vulnerable marine ecosystems (VMEs) are susceptible to the impact of intense or long-term anthropogenic activities (e.g., bottom trawling). Networks of marine protected areas (MPAs) can help facilitate the conservation and restoration of biodiversity and ecosystem function provided by VMEs. An understanding of the connectivity amongst populations of deep-sea organisms is crucial for informing the management of VMEs, by assessing the effectiveness of existing MPAs and informing the placement of new MPAs. Genetic evaluation of population structure is one of the most commonly used indirect approaches for interpreting connectivity. In contrast to corals or sponges, which are typically habitat-forming organisms as VME-indicator taxa, squat lobsters are often found in close association with VMEs and can be considered to be VME-associated taxa. Nowadays, population genetic studies of deep-sea fauna mainly focus on VME-indicator taxa, whilst relatively few studies have focussed on VME-associated taxa, such as squat lobsters, whose distribution is not exclusively limited to VMEs. In this study, three deep-sea squat lobster species, *Munida isos* Ahyong & Poore, 2004, *Munida endeavourae* Ahyong & Poore, 2004 and *Munida gracilis* Henderson, 1885, were selected based on their association with VMEs (e.g., cold-water coral reefs and seamounts), wide distributional ranges across the southwest Pacific Ocean, and sample availability.

The overall aims of this research are to evaluate patterns of population structure and genetic connectivity of three squat lobster taxa in the southwest Pacific Ocean and consider how the acquired genetic information can contribute to the management and conservation of VMEs in the southwest Pacific Ocean. A general introduction of VMEs, MPAs, connectivity of deep-sea fauna, High-Throughput Sequencing (HTS), study area and study taxa are presented in Chapter 1.

To provide background information for this research, a review was conducted of the molecular-based studies of the systematics, taxonomy and phylogenetics of marine squat lobster taxa (Chapter 2). Recent molecular-based studies have dramatically increased our understanding of squat lobster phylogenetics and systematics, and thereby the taxonomy of this diverse and challenging group, which provide a valuable starting point for evaluating hypotheses concerning speciation, biogeography, adaptation and co-evolution (e.g., squat lobsters and corals). Notably, accurate taxonomy is critical for population genetic studies and consequently supports the conservation efforts of VMEs.

A range of molecular genetic markers, including the mitochondrial COI region, nuclear microsatellites and single nucleotide polymorphisms (SNPs), were utilised to evaluate the genetic connectivity amongst populations of three VME-associated taxa (*Munida isos*, *M. endeavourae* and *M. gracilis*). In addition to this, universal invertebrate primers were used to yield partial COI fragments of 649 bp (DNA barcoding) for the three *Munida* species to confirm the taxonomic identity and to exclude the possibility of cryptic species. Due to limited genetic information for the three *Munida* species, novel microsatellite loci were developed for *M. isos* based on the HiSeq 2500 sequencing platform and used for cross-species amplification in *M. endeavourae* and *M. gracilis* (Chapter 3). Additionally, a Genotyping by Sequencing (GBS) protocol and the Universal Network Enabled Analysis Kit (UNEAK) pipeline were employed to develop novel SNPs for *M. isos* samples from the southwest Pacific Ocean (Chapter 5).

A spatially explicit hierarchical testing framework (Northern-Southern biogeographical provinces, North-Central-South regions, and individual geomorphic features) was employed for the evaluation of connectivity amongst populations of the three deep-sea squat lobster taxa across their distributional range in the southwest Pacific Ocean (Chapter 4). The level of genetic diversity was high as revealed by variation at the COI region, and moderate based on microsatellite markers across the three *Munida* species. With more than 96% of the variance being attributed to differences within populations in the three *Munida* species, based on both marker types, no genetic subdivision was detected in *M. endeavourae*, whilst little genetic differentiation was observed in *M. isos* and *M. gracilis* based on microsatellite variation. For *M. isos*, populations from the Tasmanian slope were potentially genetically different from all other populations and may act as source populations, whereas populations from the Kermadec Ridge may be sink populations. Robust evidence of recent demographic expansions was detected in the three *Munida* species, based on COI and microsatellite marker types. The estimated time of demographic expansions for the three *Munida* species was ca. 16.1 kya, 24.4 kya and 21.6 kya for the *M. isos*, *M. endeavourae* and *M. gracilis*, respectively, coinciding with the late Pleistocene. The results are discussed in the context of the distribution of existing MPAs, and contribute new information useful to the management of VMEs within national and international waters in the region.

To further investigate patterns of connectivity in deep-sea squat lobster populations and provide valuable information for the design of management strategies to protect VMEs, newly developed SNPs were utilised (Chapter 5). The results showed that the Tasmanian slope and Macquarie Ridge populations were genetically different from all other populations, both within

New Zealand's exclusive economic zone (EEZ) and the high seas beyond, with little gene flow derived from Tasmanian slope populations to Macquarie Ridge populations. The results are discussed in the context of existing MPAs, and highlight the complexity of the endeavour to maintain population diversity and gene flow across multiple national jurisdictions as well as international waters, all of which employ different spatial protective measures.

The findings of this research are summarised and discussed in relation to the usefulness of genetic studies to provide new and valuable information about the genetic diversity and connectivity of VME-associated species, and to highlight what additional genetic research is needed to assist in the management and conservation of VMEs in the southwest Pacific Ocean (Chapter 6).

IV

Acknowledgements

Publications from this Thesis

Published Articles

Yan R-J, Schnabel KE, Guo X-Z, Gardner JPA (2019) Development and characterization of 20 polymorphic microsatellite loci in the deep-sea squat lobster, *Munida isos* Ahyong and Poore, 2004 and cross-amplification in two congeneric species. *Journal of Genetics* **98**:11

Conference Proceedings

Yan R-J, Schnabel KE, Rowden AA, Guo X-Z, Gardner JPA (2017) Development and characterisation of microsatellite loci in *Munida isos* and cross-amplification in *Munida endeavourae* and *Munida gracilis*. New Zealand Marine Sciences Society 56th Annual Conference. Christchurch, New Zealand.

Yan R-J, Schnabel KE, Rowden AA, Guo X-Z, Gardner JPA (2018) Population genetic structure and connectivity of squat lobsters in vulnerable marine ecosystems. Ninth International Crustacean Congress. Washington, DC, USA.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Vulnerable marine ecosystems

The deep sea is the largest ecosystem on earth that encompasses more than 90% of the global ocean area, which harbours a wide range of marine ecosystem services with unique abiotic and biological characteristics that can support a particularly rich diversity of life (Ramirez-Llodra *et al.* 2010). However, anthropogenic activities and effects such as trawling, mining, oil and gas exploration, disposal of litter and waste, climate change and ocean acidification now, more than ever, threaten the deep sea, leading to a new imperative for national and international organisations, science community and industry to work together and develop management actions leading to the conservation of deep-sea ecosystems (Ramirez-Llodra *et al.* 2011; Levin and Le Bris 2015; Danovaro *et al.* 2017).

Vulnerable marine ecosystems (VMEs), including seamounts, canyons, hydrothermal vent and cold seep habitats, as well as cold water coral reefs and sponge beds, are vulnerable to the impact of intense or long-term anthropogenic activities, primarily to bottom trawling (Figure 1.1) (Food and Agricultural Organisation of the United Nations 2009). VMEs can be functionally significant, structurally complex and/or fragile habitat that contains unique/rare species or species with life-history traits that make recovery from disturbance difficult (Food and Agricultural Organisation of the United Nations 2009). Seamounts, for example, are "undersea mountains" that are elevated from the seafloor, most of which are volcanic in origin and are associated with mid-ocean ridges, intraplate hotspots and island arcs (Yesson *et al.* 2011). Seamounts can support highly productive ecosystems that can provide important feeding and breeding areas for benthic invertebrates, demersal deep-sea fish and pelagic predators, and which potentially support high levels of biodiversity and/or unique biological communities (Clark *et al.* 2010). Particularly, deep-sea seamounts around the New Zealand region are often characterised by the existence of cold-water habitat-forming corals (Tracey *et al.* 2011), which provide habitats for other benthic invertebrates and fish species, yet are vulnerable to the impacts from human activities (Clark and Rowden 2009; Rowden *et al.* 2010a). Nonetheless, the resilience and recovery dynamics of deep water habitats such as VMEs that are impacted by anthropogenic activities (primarily bottom trawling) remain poorly known. The nature and speed of recovery can differ greatly, based on the physical and biological characteristics of the habitat-forming benthic species, the duration and form of human impact, as well as the protection and management measures implemented during the recovery time scales (Lotze *et al.* 2011). Althaus *et al.* (2009) reported no obvious recovery in deep-sea coral fauna on previously trawled seamounts (10 years <5% trawling and 5 years post-

closure) off Tasmania. Similar findings were obtained from seamounts off New Zealand (5 years post-closure) and Australia (10 years post-closure), with no change in the benthic community structure where bottom trawling had ceased (Williams *et al.* 2010). More recently, Clark *et al.* (2019) conducted a fishing impact recovery comparison over a 15-year period on deep-sea habitats (thicket-forming scleractinian corals) from seamounts off eastern New Zealand, and reported low resilience of benthic fauna on the seamounts under conditions of bottom trawling. Collectively, it can be concluded that deep-sea benthic community structure may take decades or longer to restore to a pre-disturbance situation, due to longevity and slow growth rate of the biogenic habitat-forming invertebrates.

To aid in the identification of VMEs, benthic taxa have been identified as indicators of the presence of VMEs for specific ocean regions (e.g., Parker *et al.* (2009) for the South Pacific Ocean, Parker and Bowden (2010) for the Southern Ocean and Antarctica). These VME indicator taxa normally possess characteristics like fragility, longevity, late maturity, slow growth rate, and also include species that could form structurally complex features that are essential for diverse communities (e.g., coral reefs and sponge beds). However, there are other taxa that are often found in association with VMEs that can also act as a proxy for the likely occurrence of a VME (e.g., crinoids or sea lilies, and brisingid seastars that are associated with coral reefs (Parker *et al.* 2009)). In this regard, squat lobsters can also be indicators of VMEs because they may be found in close association with deep-water coral reefs (Baba 1974; Rice and Miller 1991; Kilgour and Shirley 2008; Rowden *et al.* 2010b), a VME that is found on seamounts, steep continental margins and sides of canyons (Rogers *et al.* 2007; Baeza 2011).
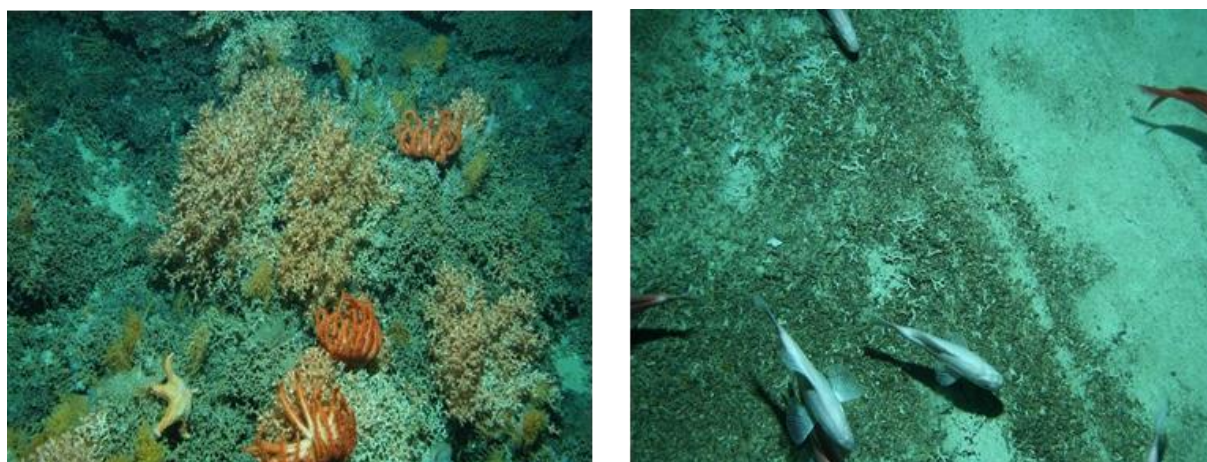
Figure 1.1 Before (left) and after (right) bottom trawling in the deep sea.

Photo credits: NIWA.

1.2 Marine protected areas

Over the last several decades, greater attention has been paid to the ongoing declines in marine seafloor biodiversity, caused by pressures including environmental change, pollution and human exploitation (Thrush and Dayton 2002; Pandolfi *et al.* 2003; Ramirez-Llodra *et al.* 2011; Cordes *et al.* 2016). Consequently, a strong international agreement on the demand for marine protected areas (MPAs) has been reached, that is widely acknowledged to be an important tool to help reduce anthropogenic impacts on biodiversity (Leathwick *et al.* 2008; Gjerde *et al.* 2016). Over the last few years, the number and spatial coverage of MPAs have increased greatly, with a ten-fold increase compared to that has been achieved by the year 2000 (~ 2 million km^2). To date, the number of MPAs has reached 14,882, spanning an area of 27,494,100 km^2 and covering 7.59% of the ocean globally (UNEP-WCMC and IUCN 2019). Most marine conservation has focussed on waters lying within the 200-nautical mile exclusive economic zones (EEZs) of nation states, whilst waters in international waters, i.e., Areas Beyond National Jurisdiction (ABNJ), has attracted less attention (relative to the size of the area) (Baco *et al.* 2016). Specifically, amongst national waters that account for 39% of the global ocean, 17.62% (24,875,947 km^2) are established as MPAs. In contrast, only 1.18% (2,618,153 km^2) of ABNJ, which represents the remaining 61% of the global ocean, has been designated as protected areas (UNEP-WCMC and IUCN 2019). Notably, the design and placement of MPAs in national waters with dedicated legal systems in place is more easily achieved compared to ABNJ, where creating MPAs becomes difficult owing to the complex legal framework in place (Gjerde *et al.* 2016). Nonetheless, in recent years, the implementation of regulatory measures in international waters has also been established to provide more extensive protection for deep-sea biodiversity (Ramirez-Llodra *et al.* 2011).

VMEs have suffered severe damage from anthropogenic activities at a regional or local scale, as well as from physical environmental changes such as increasing ocean temperature and acidification (Thresher *et al.* 2015; Gollner *et al.* 2017). To date, some efforts have been made to lessen human impacts on the VMEs by establishing MPAs, which provide a means to restrict human activity for a conservation purpose to protect critical natural habitats, which can help facilitate the protection and restoration of biodiversity and ecosystem function of VMEs (Agardy 1994; Lubchenco *et al.* 2003; Clark *et al.* 2011)

In 2001, New Zealand implemented a seamount management strategy, with 17 areas (containing 19 seamounts) within New Zealand's EEZ closed to all bottom trawling in order

to protect the benthic habitat and communities on seamounts, including VMEs (Brodie and Clark 2003). Following these seamount closures, in 2007 the New Zealand Government closed 17 benthic protected areas (BPAs) within the EEZ to dredging and bottom trawling (and prohibiting mid-water trawling within 50 m of the seafloor) to protect benthic biodiversity, including VMEs (Helson *et al.* 2010) (Figure 1.2). Since then, 32% of New Zealand's seafloor is protected from dredging and bottom trawling, a total area which comprises roughly four times the landmass of New Zealand.

In 2007, the United Nations General Assembly Resolution 61/105 on sustainable fisheries (United Nations General Assembly 2007) appealed to States and regional fisheries management organisations to protect VMEs from significant adverse impacts caused by destructive fishing activities in international waters. Consequently, in the South Pacific Ocean, interim measures to protect VMEs have been adopted by the South Pacific Regional Fisheries Management Organization (SPRFMO). These measures include 20 min longitude × 20 min latitude areas closed to bottom trawling, as well as areas open to trawling, some of which include the provision of a 'move-on rule' to avoid damage to VMEs (Penney *et al.* 2009).

To date, the suitability and effectiveness of the BPAs and seamount closures within New Zealand's EEZ, and the SPRFMO interim measures, have been assessed using cost-benefit analyses (Leathwick *et al.* 2008; Penney and Guinotte 2013), as well as some limited evaluation using genetic tools and connectivity patterns (e.g., Bors *et al.* (2012); Boschen *et al.* (2016); Holland *et al.* (2018)). For example, pronounced self-recruitment amongst populations within geomorphic features was detected in three deep-sea stony corals, and the Kermadec Ridge was identified as a source of migrants, suggesting protection measures are needed to maintain the genetic diversity of coral populations that achieved by self-recruitment at geomorphic features, particularly Kermadec Ridge (Zeng *et al.* 2017). Given that two BPAs on the Chatham Rise can only protect a very limited depth range along the Rise, additional protected areas on the Rise were proposed, considering the observed high genetic variation in deep-sea sponge populations across the Rise (Zeng *et al.* 2019).

Therefore, an understanding of genetic diversity as well as patterns of connectivity within and amongst populations of deep-sea organisms that characterise or are associated with VMEs, is fundamental and can provide information useful to inform the management of impacts on VMEs, through assessing the effectiveness of current MPAs and informing the design of the spatial arrangement of new MPAs (Palumbi 2003; Coleman *et al.* 2011; Pujolar *et al.* 2013; Santini *et al.* 2016).
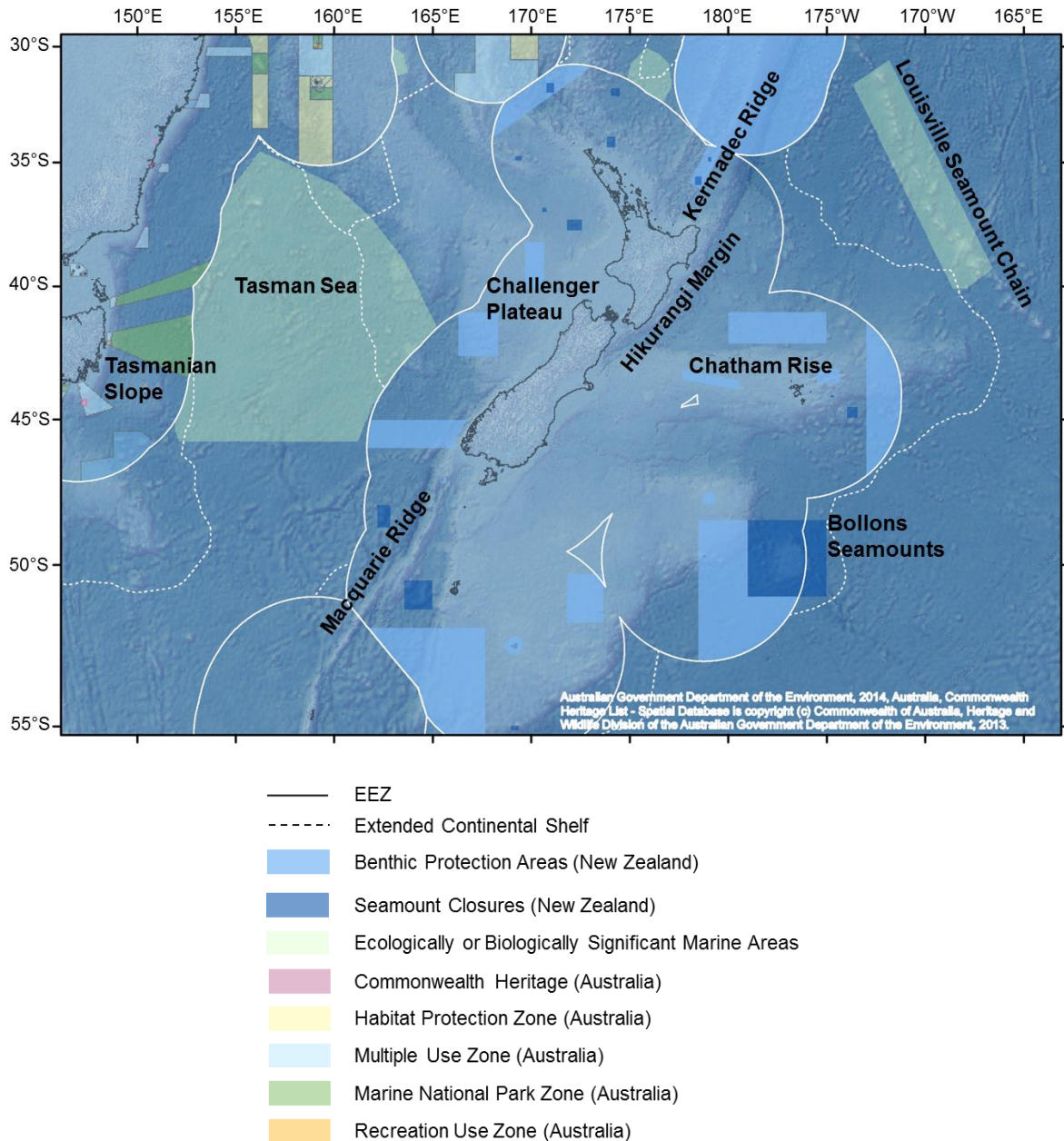
Figure 1.2 Distribution of major geomorphic features and marine protected areas in the southwest Pacific Ocean region.

This map was generated using ArcGIS 10.4.1 software (Mercator projection and the World Geodetic System (WGS 1984) as the reference coordinate system). The base map bathymetry data is derived from the GEBCO_2008 Grid, version 20100927, www.gebco.net; the marine boundaries for the New Zealand region are based on National Institute of Water and Atmospheric Research (NIWA) data; those for the Australian region are credited to the Australian Government Department of the Environment, 2014 and the Australia, Commonwealth Heritage List (Spatial Database is copyright (c) Commonwealth of Australia, Heritage and Wildlife Division of the Australian Government Department of the Environment, 2013).

1.3 Population connectivity in the deep sea

The deep sea is the largest habitat on earth that encompasses roughly 65% of the planet's surface (Danovaro *et al.* 2008). However, ecological information (e.g., biological developmental processes, species' distributions, connectivity patterns, population structure and environmental characteristics) remains poorly known due to the immenseness and remoteness of this habitat (Mengerink *et al.* 2014). To date, although rapid development of technology (e.g., remotely operated vehicles, permanent seafloor observatories) have enhanced the capability and accessibility of biological sources in the deep sea (Ramirez-Llodra *et al.* 2010), less than 1% of the deep seabed has been sampled and studied intensively.

Connectivity, or gene flow, is defined as the dispersal, establishment and reproduction of migrants that contribute to the local gene pool (Hedgecock *et al.* 2007). For many deep-sea taxa, connectivity is maintained through the planktonic dispersal process, yet the accuracy of connectivity estimates remains poorly known in the majority of deep-sea populations, especially when direct estimates of dispersal are often impossible to make (Hilário *et al.* 2015). Nonetheless, an assumption of a link between the pelagic larval duration (PLD) and dispersal distance can help predict population connectivity in poorly studied species (Waples 1998; Bohonak 1999; D'Aloia *et al.* 2015). For species that have a planktonic larval dispersal phase, those that produce lecithotrophic larvae have historically been considered to have a shorter PLD than those that produce planktotrophic larvae, which allows greater potential for dispersal and long-distance connectivity (Taylor and Roterman 2017). However, recent studies have revealed that such an assumption does not apply to all taxa in marine systems, and that dispersal can be influenced by many factors, including larval behaviour, underwater topography and oceanographic features, such as currents that may facilitate dispersal and oceanic fronts that may constrain dispersal (Weersing and Toonen 2009; Selkoe and Toonen 2011; Faurby and Barber 2012; Arellano *et al.* 2014; Mertens *et al.* 2018). Moreover, some species produce both brooding and non-brooding populations (O'Hara *et al.* 2014), and studies have suggested that lower temperatures (e.g., in the deep sea) may extend PLDs to a greater extent than in shallow (warmer) water (Mercier *et al.* 2013), which further complicates the situation.

Direct estimation of larval dispersal in the ocean requires the ability to track microscopic larvae of benthic invertebrates throughout the pelagic dispersal process, from spawning sites to recruitment locations, which is extremely difficult (Levin 2006; Thorrold *et al.* 2007). Given the aforementioned uncertainty around the variables that influence dispersal

and the difficulty of implementing direct estimation methods, ecologists have embraced a number of indirect approaches that have been developed for the evaluation of dispersal and connectivity in marine ecosystems, including trace-element signatures, biophysical models and genetic techniques (Hellberg *et al.* 2002; Cowen and Sponaugle 2009; Lowe and Allendorf 2010).

Elemental fingerprints of calcified structures have been applied to estimate connectivity, primarily in shallow water (Carson *et al.* 2013), whilst a number of biophysical models simulating larval dispersal have been applied in the deep sea to estimate dispersal (Hilário *et al.* 2015). Genetic approaches provide means for deep-sea ecologists to evaluate genetic diversity and to model connectivity patterns, gaining insights into the resilience and dynamics of deep water populations (e.g., Bors *et al.* (2012); Boschen *et al.* (2016); Holland *et al.* (2018)). Over the last several decades, based on a proliferation of genetic techniques and bioinformatic tools, the availability of information about the genetic diversity and population structure of deep-sea invertebrates has greatly increased (Taylor and Roterman 2017). Consequently, this new knowledge is fundamental and essential for the description of conservation units as well as evaluation and placement of MPAs to maintain deep sea biodiversity (Danovaro *et al.* 2008; Wedding *et al.* 2013; Benestan *et al.* 2016).

Given the many technical limitations (e.g., logistics, sampling difficulty, uncertainty in taxonomy), the genetic information of connectivity patterns in marine organisms remains sparse. Baco *et al.* (2016) compiled 51 population genetic studies of deep-water fauna and analysed isolation-by-distance slopes, a proxy for pelagic dispersal distances, to reveal connectivity patterns. As expected, species with a pelagic larval phase were more dispersive than other larval types. The authors also proposed that deep sea MPAs design would need to consider the variety of taxa, life histories, habitat types (e.g., seamount, hydrothermal vent, cold seep) and depth zone, and the dispersal distances for deep-sea organisms might be comparable to or a bit larger than their shallow water counterparts (Baco *et al.* 2016). A critical limitation recognised by the authors was the relatively low number of comparable studies, which hampered the ability of multivariate statistics to distinguish confounding variables. They remarked that more connectivity studies of deep-sea fauna would be needed before greater insights can be achieved (Baco *et al.* 2016). Subsequently, Taylor and Roterman (2017) summarised 77 population genetic studies of benthic invertebrates inhabiting non-chemosynthetic ecosystems on the deep seabed, amongst which, 13 were focussed on VME-indicator taxa (i.e., corals and sponges). Overall, populations sampled spanning a range of

depths exhibited population structure at relatively small scales (100s–1,000s m), whilst those sampled at similar depths were well connected over 100-1000s km. Findings of depth-related genetic structure were in accordance with isolation by adaptation across environmental characteristics (Nosil *et al.* 2009) and/or physical barriers to connectivity with depth (Taylor and Roterman 2017). The latter authors also concluded that demographic instability is typical in the deep sea (for both hydrothermal vent and non-vent fauna), with the majority of studies that utilised single-locus mitochondrial genes/regions showing a common non-neutrality pattern that is consistent with population demographic instability (e.g., expansions or selective sweeps), which render studies using single-locus mitochondrial genes/regions demographically uninformative (Taylor and Roterman 2017).

Collectively, these reviews show that population genetic studies on deep-sea fauna have long been hampered by the extreme difficulties of sampling in the deep sea, which limits the geographical scope, number of organisms, sample sites and sample size. Given the time and cost constraints, future effort should be focussed in areas that have attracted less attention, such as VMEs-indicator and VMEs-associated taxa that are functionally and ecologically important.

Previous population genetic studies of southwest Pacific Ocean VME-indicator taxa (mostly focussed on deep-sea corals) have revealed different patterns of connectivity. For example, Miller *et al.* (2011) assessed connectivity levels of the deep-sea solitary cup coral *Desmophyllum dianthus* (Esper, 1794) in the waters off Australia, New Zealand and Chile, and reported genetic differentiation amongst the three geographic regions using combined assessments of variations in the ribosomal internal transcribed spacer (ITS) (nuclear DNA), 16S and the D-loop (mtDNA). Besides, evidence of strong depth stratification was detected amongst coral populations, indicating limited vertical larval dispersal. Subsequently, Zeng *et al.* (2017) assessed connectivity levels of three stony corals (*Goniocorella dumosa* (Alcock, 1902)*, Madrepora oculata* Linnaeus, 1758 and *Solenosmilia variabilis* Duncan, 1873) based on COI region and microsatellite variation, and predicted four barriers to gene flow within/around New Zealand's EEZ, isolating the populations associated with the geomorphic features of the Kermadec Ridge, Louisville Seamount Chain, Chatham Rise, Macquarie Ridge, and the combined populations of the Campbell Plateau, Bounty Plateau and Bounty Trough. Nonetheless, conflicting estimates of connectivity patterns amongst populations of different VME-indicator taxa within the Australian and New Zealand region have been detected. Evidence of genetic subdivision was observed in three deep-sea coral species, *D. dianthus*, *Antipathes robillardi* (Bell, 1891) and *Stichopathes variabilis* van Pesch, 1914, and low levels

of genetic differentiation was revealed in *S. variablis* and *M. oculata*, whilst no genetic differentiation was detected in *Stephanocyathus spiniger* (Marenzeller, 1888) and *Stichopathes filiformis* Gray, 1868, based on 16S, ITS and D-loop variation (Miller *et al.* 2010). Similarly, different genetic differentiation levels were resolved for two coral species (*D. dianthus* and *S. variabilis*) of the same family spanning the South Tasman Sea and Macquarie Ridge regions utilising microsatellite variation, providing empirical evidence that seamounts can function either as stepping stone or isolated islands for dispersal for different taxa (Miller and Gunasekera 2017). Holland *et al.* (2018) used microsatellite markers and DNA sequences (ITS, 16S, *nd5* and *trnW*) to test patterns of deep-sea connectivity for two scleractinian and two black corals within New Zealand's EEZ and in the surrounding high seas. Regional genetic structure was detected in *D. dianthus,* but little genetic structure was observed in the other three coral species. More recently, deep-sea sponge species, another group of VME indicator taxa, have been shown to exhibit contrasting patterns of genetic subdivision within the New Zealand region, with significant mitochondrial (COI and Cytb) and nuclear DNA (microsatellite variations) genetic differences observed in *Penares* sp. and *Poecillastra laminaris* (Sollas, 1886), but no evidence of genetic structure detected in *Neoaulaxinia persicum* Kelly, 2007 (COI, 12S) and *Pleroma menoui* Lévi & Lévi, 1983 (COI) (Zeng *et al.* 2019).

There is, therefore, a lack of consistent patterns of genetic subdivision for coral and sponge VME-indicator taxa within the South Pacific region, which exhibit a wide range of reproductive strategies (e.g., lecithotrophic and/or planktotrophic) and distributional ranges. It can be concluded that both historical and contemporary processes, including biological factors (larval dispersal strategy), and the influence of large-scale marine current patterns and/or small-scale hydrological features associated with geomorphic features, may contribute in shaping the connectivity patterns. It would, therefore, be beneficial to include both VME-indicator and VME-associated taxa in the connectivity studies to inform the ecosystem-based management of VMEs. However, the latter have so far only received limited attention in genetic connectivity studies.

1.4 High-Throughput Sequencing

Whilst there are a number of studies examining the connectivity of VME indicator species from the southwest Pacific Ocean, relatively few studies have focussed on VME-associated taxa whose distributions are not exclusively limited to VMEs, including deep-sea

squat lobsters (Bors *et al.* 2012; Thaler *et al.* 2014). This limitation has been mainly due to both the difficulty of obtaining a sufficient number of specimens and the paucity of molecular markers that yield results (Shank 2010). Whilst small sample size and limited spatial scale sampling are unavoidable constraints in population genetic studies of many deep-sea organisms (Baco *et al.* 2016), molecular marker technologies have undergone a revolution in the last several decades, shifting from "genotyping-by-assay" to "genotyping-to-sequencing" (Chen *et al.* 2013) and thereby increasing the ease of genetic data acquisition.

The earliest phase of population genetic studies in deep-sea organisms was characterised exclusively by allozyme electrophoresis (e.g., Goocii and Schopf (1972); Costa and Bisol (1978)). During the 1980s and 1990s, technological advances such as DNA Sanger sequencing (Sanger *et al.* 1977) and the discovery of microsatellite markers (Litt and Luty 1989), offered researchers with a variety of tools for population genetic studies. In 1998, the first deep-sea benthic paper using Sanger-sequenced gene (i.e., ribosomal 16S gene in bivalve molluscs) was published (Chase *et al.* 1998). Since then, gene sequence studies, generally utilising mitochondrial COI and/or 16S genes, have become the most prevalent methodology in population genetics (e.g., Samadi *et al.* (2006); Haye *et al.* (2010); Bors *et al.* (2012)). Compared to allozymes, the presence of hidden sequence variations can be revealed by DNA sequence data, which can provide more genetic information (Parker *et al.* 1998).

Advances in High-Throughput Sequencing (HTS) technologies have enabled genome assembly of an increasing number of species and had a tremendous impact on many areas of biology, providing unprecedented discovery and characterisation of molecular markers (O'Leary *et al.* 2018). With the development of HTS technologies, sequencing cost has declined rapidly, facilitating rapid and inexpensive analyses of the genome in both model (organisms with reference genomes) and non-model organisms (those not subject to large-scale genome projects) (Seeb *et al.* 2011). In particular, HTS is now the preferred approach for the development, validation and assessment of genetic markers (e.g., microsatellites, single nucleotide polymorphisms, SNPs) in many non-model species (Abdelkrim *et al.* 2009; Davey *et al.* 2011; Lance *et al.* 2013; Bahassi and Stambrook 2014).

Microsatellites have emerged as one of the most powerful molecular markers in conservation genetic studies due to their abundance, co-dominant mode of inheritance, relative ease and low cost of development, reproducibility and levels of intraspecific variability that are attributable to the high mutation rate of these non-coding regions (Wright and Bentzen 1995; Selkoe and Toonen 2006; Jost *et al.* 2018). Nonetheless, microsatellite development using the

enrichment library method has been relatively problematic in some marine invertebrates (Selkoe and Toonen 2006), which may be due to the existence of inconspicuous cryptic repeats on presumed unique flanking regions in microsatellite sequences (Bailie *et al.* 2010). Despite this potential limitation, HTS technologies have been employed to obtain a large number of microsatellite loci for marine squat lobster species from the southwest Pacific Ocean, North Atlantic Ocean, Southern Ocean and Northwest Pacific Ocean (Boyle *et al.* 2013; Coykendall and Morrison 2013; Roterman *et al.* 2013b; Nakajima *et al.* 2018).

SNPs, as a third-generation molecular marker, have become the marker of choice for many reasons, including the high density of annotated markers, low scoring error rates, calibration amongst laboratories, and the associated ability to assemble combined datasets from multiple resources (Brumfield *et al.* 2003; Morin *et al.* 2004; Helyar *et al.* 2011). HTS offers multiple approaches that simultaneously enables genome-wide SNPs discovery and genotyping in a single step, even in organisms for which we have limited genetic information (Davey *et al.* 2011; Helyar *et al.* 2011). The distributions and frequencies of SNP markers are keys to evaluating diversity amongst closely related populations and species (Tautz *et al.* 2010), something that can now be achieved with unprecedented speed, accuracy and low cost.

1.5 Study area

New Zealand lies in the western South Pacific Ocean and has a large EEZ (approximately 4,270,00 km^2) that contains a variety of marine habitats (Gordon *et al.* 2010; Helson *et al.* 2010; MacDiarmid *et al.* 2013). The seafloor around New Zealand has a diverse array of deep-sea geomorphic features (Figure 1.2). The topography of these features, such as seamounts, rises, ridges, and plateaux, can induce particular current flow patterns and hydrographic features (e.g., Taylor Columns) that may limit the dispersal of larvae amongst populations, even those in close geographic proximity (Mullineau and Mills 1997; Richer de Forges *et al.* 2000; Samadi *et al.* 2006).

The surface ocean current circulation in the New Zealand region is composed of Subtropical Water (STW) that arrives via the East Australian Current and the South Pacific subtropical gyre, Subantarctic Water (SAW) that is driven north by Ekman transport and westerly winds, and Antarctic Surface Water (AASW) that is found south of the Polar Front and extends southwards to the margins of Antarctica. These are separated by three major fronts — the Subtropical Front (STF), Subantarctic Front (SAF) and Polar Front (PF) (Orsi *et al.* 1995; Chiswell *et al.* 2015) (Figure 1.3). Notably, STW has the warmest and most saline surface water, whilst SAW is characterised by relatively low temperature and salinity. A zonal Antarctic Circumpolar Current (ACC) is identified as the water between the SAF and the PF, which is a deep-reaching eastward flow that is driven mainly by the westerly winds, and is the only current to connect major oceans (Carter *et al.* 2008). In the Tasman Sea, the Western South Pacific Central Water (WSPCW) or Tasman Sea Central Water (TSCW) is formed to the east of Tasmania and is characterised by multiple layers, inversions and intrusions, and the STF is density-compensated with little flow, which sits below the TSCW (Szymanska and Tomczak 1994; Hamilton 2006). These large-scale currents, gyres and fronts can be responsible for facilitating the connectivity of populations, or they may present barriers to population connectivity.

Thus, the waters around New Zealand are an ideal place to test hypotheses of how physical oceanographic and environmental processes might influence larval dispersal patterns and thereby population genetic structure and connectivity.
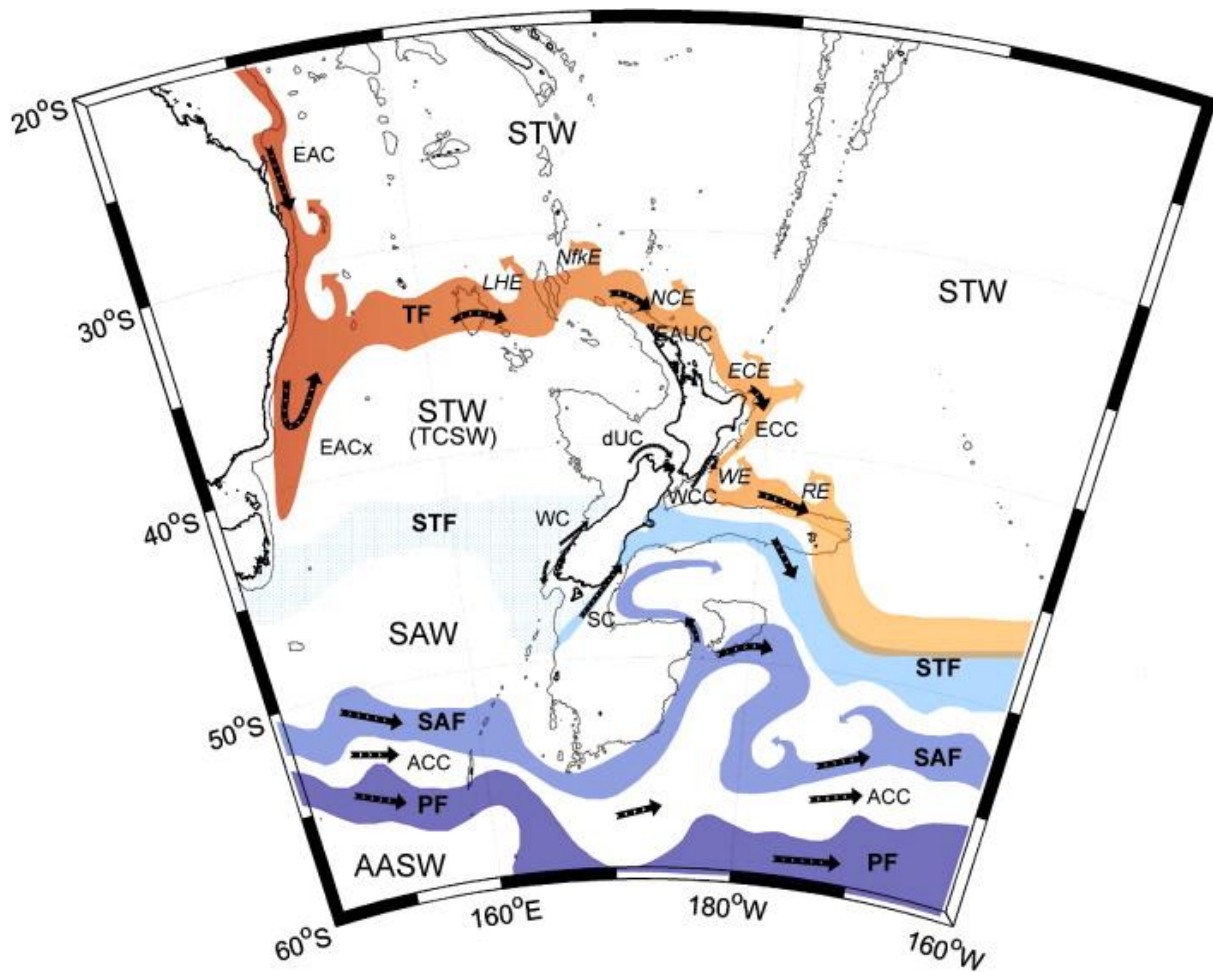
Figure 1.3 Summary of the surface circulation in the southwest Pacific Ocean region.

Subtropical Water (STW), Tasman Sea Central Water (TSCW), Subantarctic Water (SAW) and Antarctic Surface Water (AASW); Tasman Front (TF), Subtropical Front (STF), Subantarctic Front (SAF) and Polar Front (PF); East Australia Current (EAC), East Australia Current extension (EACx), East Auckland Current (EAUC), East Cape Current (ECC), d'Urville Current (dUC), Wairarapa Coastal Current (WCC), Westland Current (WC), Southland Current (SC) and Antarctic Circumpolar Current (ACC). Lord Howe Eddy (LHE), Norfolk Eddy (NfkE), North Cape Eddy (NCE), East Cape Eddy (ECE), Wairarapa Eddy (WE) and Rekohu Eddy (RE). Figure credit: adapted from Chiswell *et al.* (2015).

1.6 Study taxa – squat lobsters

Squat lobsters (Galatheoidea and Chirostyloidea), a diverse group of decapod crustaceans, are dominant and abundant at all depths throughout many of the world's oceans (Forest *et al.* 2000; Baba *et al.* 2008; Macpherson *et al.* 2010; Tracey *et al.* 2011; Coykendall *et al.* 2017). Marine squat lobsters comprise ~1,500 species worldwide, with many not yet formally described (Schnabel *et al.* 2011b; Coykendall *et al.* 2017; WoRMS Editorial Board 2018). As VME-associated taxa, squat lobsters can be found in close association with reef-forming coral species (Baba 1974; Rice and Miller 1991; Kilgour and Shirley 2008; Rowden *et al.* 2010b). For example, representatives of the family Munididae have been reported as engaging in a putative partnership with corals. These include *Munida rugosa* (Fabricius, 1775) in the North East Atlantic that is often found associated with the scleractinian *Lophelia pertusa* (Linnaeus, 1758) (Jonsson *et al.* 2004), and *Sadayoshia acroporae* Baba, 1972 that occurs with Acropora corals of the Indo-West Pacific (Baba 1972, 1979). *Munida*, with ~270 species, is a cosmopolitan and the most diverse genus in the family Munididae and has attracted much attention in recent years, particularly in taxonomic, systematic and phylogenetic studies (Macpherson *et al.* 2010; Macpherson and Baba 2011).

In this study, three deep-sea squat lobsters (*Munida isos* Ahyong & Poore, 2004, *M. endeavourae* Ahyong & Poore, 2004 and *M. gracilis* Henderson, 1885) were selected based on their association with VMEs, their widespread distributions across and beyond New Zealand's EEZ, and the availability of samples in existing natural history collections (Figure 1.4). *Munida isos*, typically associated with seamounts (i.e., VMEs), is widespread in the New Zealand and eastern Australia regions, and has been recorded in water depths of 462–2756 m. The known distribution of *M. endeavourae* includes the southeastern Australia region and the far northern part of the New Zealand region; it is primarily recorded from seamounts in a depth range of 554–2756 m. *M. gracilis* is a New Zealand endemic species and has been reported from both seamount and soft sediment (non-VME) habitats in a depth range of 44–1211 m, distributed across the entire New Zealand continental shelf, excluding the far southern Campbell Plateau and the Tasman Sea (Schnabel *et al.* 2009; Yaldwyn and Webber 2011; Ahyong *et al.* 2015a).

Understanding the connectivity of squat lobster populations amongst the ecosystems they inhabit is useful for informing the management of impacts on VMEs, through assessing the effectiveness of existing MPAs and informing the placement of new MPAs to establish and maintain a connected network of MPAs.
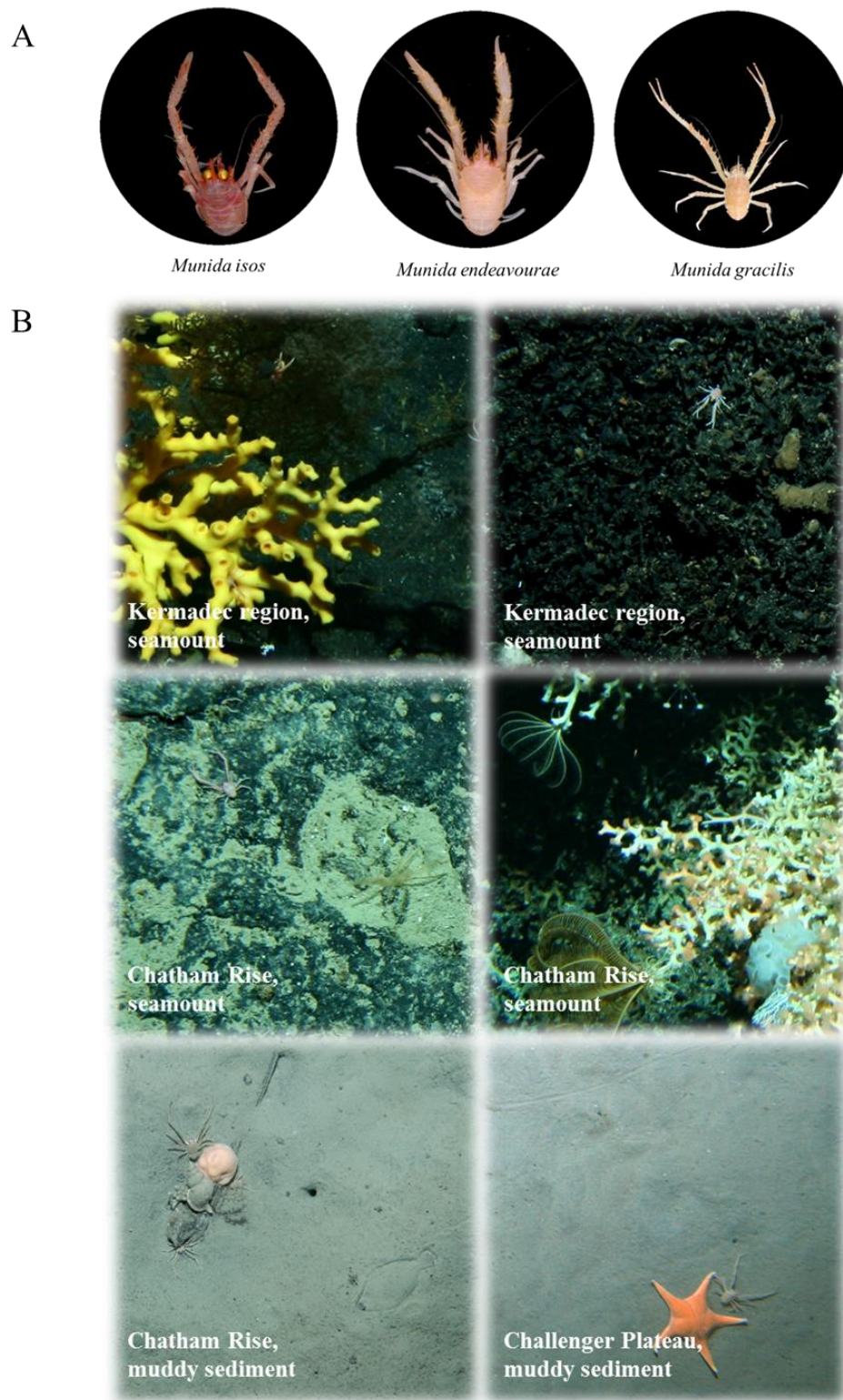
Figure 1.4 Specimens of *M. isos*, *M. endeavourae* and *M. gracilis* (A) and habitats of squat lobsters (B).

Munida isos (body length: 2 cm), *M. endeavourae* (body length: 2 cm) and *M. gracilis* (body length: 11 cm). Photo credits: NIWA.

1.7 Aims and thesis structure

1.7.1 Aims

This PhD research forms part of a larger New Zealand government funded project, led by NIWA and supported by Victoria University of Wellington, to predict the occurrence of VMEs within New Zealand's EEZ, and to better understand how existing VME locations are connected genetically via gene flow. Earlier work has included quantification of population genetic structure and connectivity of a range of VME indicator species, including corals and sponges (Zeng *et al.* 2017; Holland *et al.* 2018; Zeng *et al.* 2019).

The aim of this PhD study is to investigate the patterns of population structure and genetic connectivity of VME-associated squat lobsters. The specific objectives are to:

1. Review molecular-based studies of the systematics, taxonomy and phylogenetics of marine squat lobster taxa.

2. Quantify the genetic diversity of three *Munida* species in the southwest Pacific Ocean.

3. Evaluate patterns of population genetic structure and connectivity of three *Munida* species across their distributional ranges.

4. Explore the drivers of the observed connectivity patterns.

5. Consider how genetic information can contribute to the management and conservation of VMEs in the southwest Pacific Ocean.

1.7.2 Thesis structure

Each data chapter of this thesis has its own objectives and individual sections for Introduction, Methods, Results and Discussion. Chapter 2 presents a review of the taxonomy, systematics and phylogenetics of the marine squat lobsters with a particular emphasis on molecular studies, and provides useful background knowledge for the study taxa. Chapter 3 describes the development of polymorphic microsatellite loci for *M. isos,* and their cross-amplification in two congeneric species, *M. endeavourae* and *M. gracilis* (also reported in a publication in the *Journal of Genetics* (Yan *et al.* 2019). Chapter 4 presents patterns of population structure and genetic connectivity of three *Munida* species in the southwest Pacific Ocean utilising the newly developed nuclear microsatellite markers and the mitochondrial COI

region. In Chapter 5, Genotype-by-sequencing method and UNEAK pipeline are employed to develop novel SNPs, which have been used to further test connectivity patterns amongst populations of *M. isos* in the southwest Pacific Ocean. Finally, Chapter 6 summarises the key findings, indicates future research directions, and provides recommendations for the spatial management and conservation of VMEs in the southwest Pacific Ocean.

Chapter 2. A review of molecular work for marine squat lobsters: taxonomy, systematics and phylogenetics

Abstract

Over the past several decades, squat lobsters have attracted much attention, being one of the most diverse groups in the Decapoda, both morphologically and ecologically. However, there is a lack of clarity around the taxonomy, systematics and phylogenetics of the marine squat lobsters, all of which are reviewed with a particular emphasis on molecular research addressing these topics. Over the years, numerous revisions have been made to squat lobster phylogeny using morphological and/or molecular approaches. Whilst morphological inferences can sometimes be confused by convergent and parallel evolution, molecular information has shed new light on squat lobster phylogeny. To date, a number of marker types, such as mitochondrial DNA, nuclear ribosomal genes and protein-coding genes, have been employed in squat lobster phylogenetic studies at various taxonomic levels. Molecular-based studies have reshaped our understanding of evolutionary and phylogenetic relationships of the squat lobster group, leading to the confirmation of the taxonomic status of many taxa that corroborates traditional morphological views and the reclassifications of major taxonomic groups. New sources of molecular information are essential for future phylogenetic evaluation, the identification of patterns of speciation, genealogy reconstruction, and divergence time estimations of this extraordinarily complicated and challenging group. This review highlights the important contribution that molecular studies can make to accurate assessments of taxonomy and systematics, and the role that correct taxonomy may play in the conservation of vulnerable marine ecosystems.

2.1 Introduction

Squat lobsters, Galatheoidea Samouelle, 1819 and Chirostyloidea Ortmann, 1892, a speciose and highly diverse group of marine decapod crustaceans, are dominant and abundant at all depths throughout the world's oceans (Schnabel *et al.* 2011b) (Figure 2.1). Over the past two decades, squat lobsters have been the focus of intense taxonomic, systematic and phylogenetic studies, which has significantly increased the number of newly recognised taxa, and has seen a series of reorganisations of phylogenetic classification at different taxonomic levels (Ahyong and Poore 2004; Baba 2005; Baba *et al.* 2008; Ahyong *et al.* 2010; Ahyong *et al.* 2011b; Yaldwyn and Webber 2011; Ahyong *et al.* 2015a; Macpherson and Robainas-Barcia 2015). These ongoing studies provide a valuable starting point for evaluating hypotheses concerning speciation, biogeography, adaptation and co-evolution (e.g., squat lobsters and corals), as well as refining the phylogenetic placement of fossil taxa (Ahyong *et al.* 2011b). A robust phylogeny is very important to elucidate the evolution and diversification of squat lobster taxa. Although morphological cladistic approaches have provided many useful insights into the phylogeny of squat lobsters, the high morphological diversity of this group has posed challenges to its phylogenetic study, leaving a number of critical disputes unresolved, especially for deeper taxonomic levels (Chu *et al.* 2015). Therefore, researchers have turned to new sources of information based on nuclear and mitochondrial DNA variation to help resolve squat lobster phylogeny.

With the rapid advances in molecular genetic technologies, barcoding and other DNA-based approaches to taxonomy have become reliable and cost-effective tools for studies of evolutionary ecology and biology (Hebert *et al.* 2003; Costa *et al.* 2007). Specifically, their uses in phylogenetic studies have confirmed the taxonomic status of many taxa based solely on morphological characters (Hendry *et al.* 2000; Lefébure *et al.* 2006). Mitochondrial genes have long been the most popular markers in animal phylogenetic studies, and they now benefit from the availability of a number of universal primers and ease of amplification (Simon *et al.* 1994). To date, most phylogenetic analyses have used rapidly evolving mitochondrial genes, primarily the cytochrome *c* oxidase subunit I (COI) and 16S ribosomal RNA (rRNA), which are commonly employed to infer intergeneric and interspecific relationships (e.g., most recently Coykendall *et al.* (2017); Macpherson *et al.* (2017); Schnabel *et al.* (2017)).
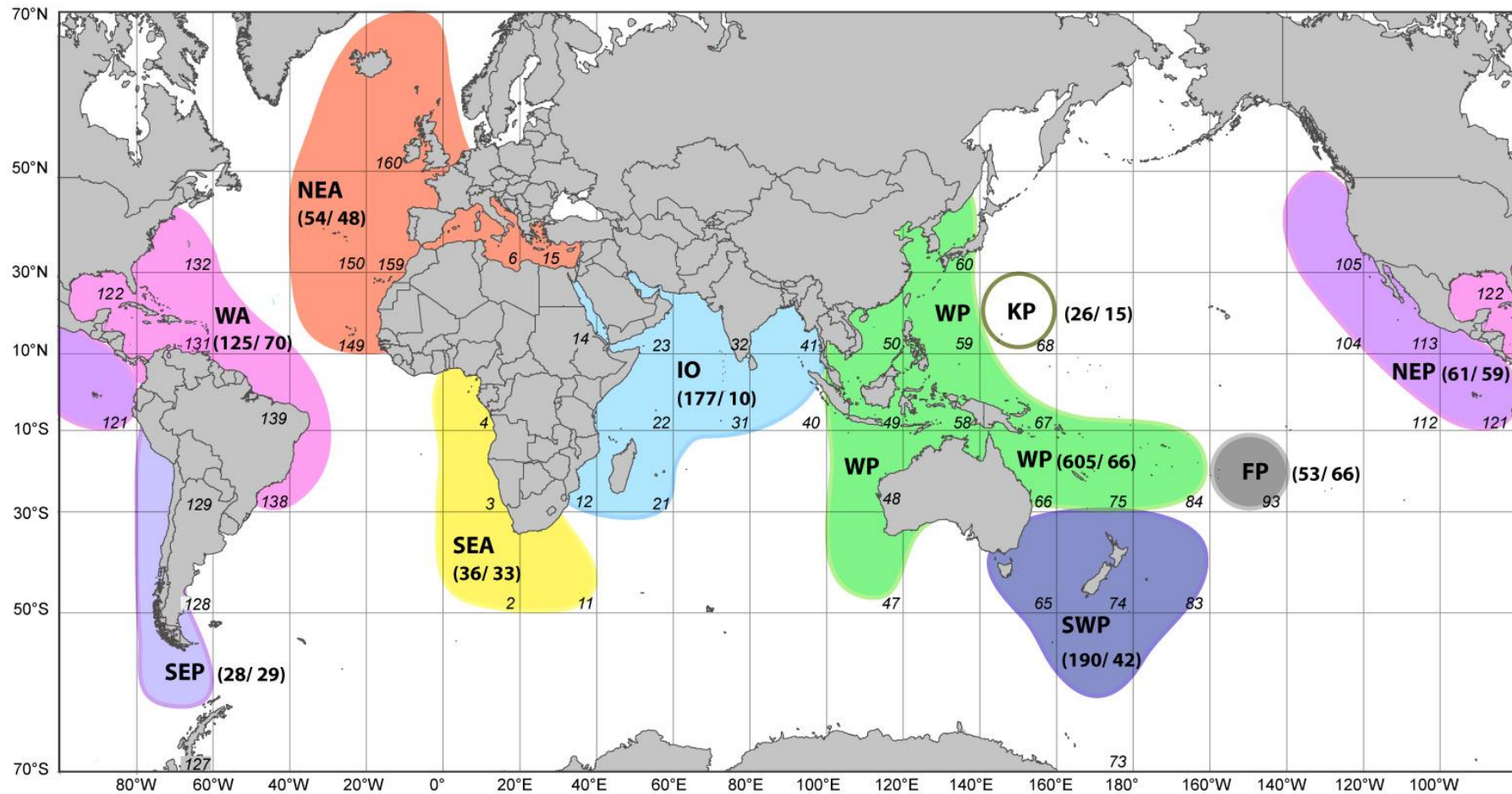
Figure 2.1 Biogeographical distribution of marine squat lobsters in the world oceans.

Bold numbers in parentheses indicate total number of species recorded in the province and percentage endemic. Italic number indicates unique cell identification number. Biogeographic provinces code: NEA = north-eastern Atlantic; SEA = south-eastern Atlantic; WA = western Atlantic; IO = Indian Ocean; WP = central western Pacific; SWP = south-western Pacific; KP = Kyushu-Palau/Bonin; FP = French Polynesia; NEP = north-eastern Pacific; SEP = south-eastern Pacific. Figure credit: adapted from Schnabel *et al.* (2011b).

The possible drawbacks of this mitochondrial DNA-based approach, such as the high mutation rates of these genes and their linked status (the mitochondrial genome, i.e., mitogenome, is a non-recombining circular molecule on which genes are effectively linked), may potentially limit their utility in high-level phylogenetics (Ahyong *et al.* 2011b; Chu *et al.* 2015). Notably, the mitogenome has proved its ability to effectively resolve phylogenetic relationships in several squat lobster species, as well as providing an additional source of phylogenetic information, such as novel gene orders (Shen *et al.* 2013; Tan *et al.* 2018). As a group that lacks genomic resources, squat lobsters benefit from the relative ease of sequencing the mitogenome compared to the whole (nuclear) genome, in light of the large reported genome sizes (1–40 Gbp) in decapod crustaceans (http://www.genomesize.com). To date, only six complete squat lobster mitogenomes are available on the NCBI database, making the squat lobster group one of the most poorly represented amongst major crustacean groups (Yang and Yang 2008; Lee *et al.* 2016a; Zhang *et al.* 2017; Tan *et al.* 2018). However, it's often also true that mitochondrial DNA (mtDNA) history alone is not an ideal marker of evolutionary history (e.g., Westfall and Gardner (2013)). In view of these limitations, researchers have used methods that incorporate nuclear ribosomal genes (e.g., 18S rRNA and 28S rRNA) which evolve at a relatively slower rate and are evolutionarily independent of the mitogenome, to address questions in deeper time for decapod phylogeny (e.g., most recently Rogers *et al.* (2012); Thatje *et al.* (2015); Rodríguez-Flores *et al.* (2018)). Unfortunately, nuclear rRNA genes often suffer from alignment ambiguities which may hamper phylogenetic inferences, especially for phylogenetically diverse taxa (Ahyong *et al.* 2011b). More recently, the nuclear protein-coding genes (e.g., arginine kinase (AK) and phosphoenolpyruvate carboxykinase (PEPCK)), have been applied in decapods and have proven their ability to resolve phylogenetic relationships at different taxonomic levels (e.g., most recently Chu *et al.* (2015); Palero *et al.* (2017); Roterman *et al.* (2018)). These genes are easy to align and have many variable sites, yet, several limitations may restrict the development and utilisation of these markers. For example, the degenerate third codon positions may complicate primer design, and they are more difficult to amplify because there are fewer copies in the genome compared to ribosomal genes (Ahyong *et al.* 2011b; Chu *et al.* 2015). Considering across all molecular approaches, each type of molecular marker has its own strengths and weaknesses, highlighting the need to combine different sources of information in the phylogenetic evaluation, speciation pattern identification, genealogy reconstruction, and divergence time estimations at different taxonomic levels (Liu and Cordes 2004; Pettersson *et al.* 2009).

This review addresses recent molecular studies of the groups of species commonly referred to as squat lobsters, with emphasis on taxonomy, systematics and phylogenetics. A summary of the studies and the gene regions used is provided in Table S2.1. Primary research studies are included and discussed here with the following categories: 1) taxonomic evaluation of squat lobster taxa, 2) molecular phylogeny of the Anomura, 3) molecular phylogeny of squat lobster taxa, and 4) divergence time estimation of squat lobster taxa.

2.2 Taxonomic evaluation of squat lobster taxa: species complexes, cryptic species and new species

The taxonomic challenge posed by cryptic or pseudo-cryptic (i.e., morphologically recognised as cryptic until formally described) sibling species has long been recognised in decapod Crustacea (Knowlton 1986). With rapid development in DNA sequencing technologies, the existence of cryptic species has been resolved in many decapods including squat lobsters, particularly in cosmopolitan species or monospecific genera, on the basis of molecular data (Bickford *et al.* 2007). Such taxa have been recognised as a species complex, which, although morphologically similar, are genetically very distinct.

Currently, there are two themes that explain why morphological change is not correlated with species discrimination: cryptic species appear to be under sufficient evolutionary/selection pressure that maintain morphological stasis (e.g., Vrijenhoek *et al.* (1994)), or are differentiated by nonvisual mating signals (e.g., mating pheromones or mating calls) (e.g., Henry (1994)). One the one hand, extreme environmental conditions may strengthen stabilising selection on morphology, leading to the reduction or elimination of morphological changes that accompany speciation (Schönrogge *et al.* 2002). On the other hand, although cryptic species can be discriminated by differences in mating signals, the morphological machinery that produces different acoustic or olfactory signals need not differ markedly, and thus escape the scrutiny of visually oriental taxonomists (Bickford *et al.* 2007).

Patterns of morphological conservation and convergence are common and frequent in marine squat lobsters, and the real biological diversity of this group is largely underestimated (Baba *et al.* 2008). One of the most probable hypotheses for the evolution of squat lobsters is that a rapid/explosive mode of speciation has been accompanied by stasis/constraints in its morphological evolution (Machordom and Macpherson (2004). The prevalence of apparently

"cryptic speciation" in marine squat lobsters is fascinating: nearly all genera and species examined have revealed a high proportion of genetic diversity with little morphological change.

To date, species complexes and hidden diversity have been demonstrated in many genera: *Eumunida* Smith, 1883 (Puillandre *et al.* 2011), *Heteroptychus* Baba, 2018 (Baba 2018) and *Uroptychus* Henderson, 1888 (Poore and Andreakis 2011) in the Chirostyloidea; *Allogalathea* Baba, 1969 (Cabezas *et al.* 2011b), *Coralliogalathea* Baba & Javed, 1974 (Rodríguez-Flores *et al.* 2018), *Galathea* Fabricius, 1793 (Macpherson and Robainas-Barcia 2015), *Leiogalathea* Baba, 1969 (Rodríguez-Flores *et al.* 2019), *Munida* Leach, 1820 (Machordom and Macpherson 2004; Macpherson and Machordom 2005), *Munidopsis* Whiteaves, 1874 (Jones and Macpherson 2007), and *Paramunida* Baba, 1988 (Cabezas *et al.* 2010) in the Galatheoidea. Molecular tools have also supported the description of new species in many genera, without the inference of a species complex: *Kiwa* Macpherson, Jones & Segonzac, 2005 (Macpherson *et al.* 2005; Thurber *et al.* 2011; Rogers *et al.* 2012; Roterman *et al.* 2013a; Thatje *et al.* 2015; Lee *et al.* 2016b) and *Uroptychus* (Schnabel *et al.* 2017; Baba 2018) in the Chirostyloidea; *Agononida* Baba & de Saint Laurent, 1996 (Cabezas *et al.* 2009b), *Babamunida* Cabezas, Macpherson & Machordom, 2008 (Schnabel *et al.* 2009), *Fennerogalathea* Baba, 1988 (Rodríguez-Flores *et al.* 2017), *Lauriea* Baba, 1971 (Macpherson and Robainas-Barcia 2013), *Munida* (Cabezas *et al.* 2009b; Cabezas *et al.* 2011a), *Paramunida* (Cabezas *et al.* 2009b; Cabezas and Chan 2014; Mccallum *et al.* 2016), *Plesionida* Baba & Saint Laurent, 1996 (Cabezas *et al.* 2009b), *Raymunida* Macpherson & Machordom, 2000 (Macpherson and Machordom 2001) and *Triodonthea* Macpherson & Robainas-Barcia, 2013 (Macpherson and Robainas-Barcia 2013) in the Galatheoidea. The aforementioned molecular-based taxonomic studies show the utility of molecular tools to resolve taxonomic questions in marine squat lobster taxa by uncovering the existence of species complexes, cryptic species and new species, and demonstrate the challenges still faced by taxonomists given that insufficient molecular work has focussed on squat lobsters.

2.3 Phylogenetic studies of the Anomura through time

For a century, phylogenetic relationships amongst the major decapod infraorders have attracted much attention, with hypotheses around the placement of anomuran groups being particularly diverse, likely due to their extraordinary ecological and morphological diversity (Martin and Davis 2001; Dixon *et al.* 2003; Ahyong and O'Meally 2004; Porter *et al.* 2005;

Tsang *et al.* 2008; Bracken-Grissom *et al.* 2009; Shen *et al.* 2013; Chu *et al.* 2015; Toon *et al.* 2015). Early classifications were mainly based on adult morphological characters, and these classification schemes usually differed in higher-level composition, and remained controversial (McLaughlin 1983; Scholtz and Richter 1995; McLaughlin *et al.* 2007). More recently, molecular and/or morphological data have been utilised to re-evaluate anomuran phylogeny (Lemaitre and McLaughlin 2009; Schnabel *et al.* 2011a; Tsang *et al.* 2011; Bracken-Grissom *et al.* 2013; Ahyong *et al.* 2015b; Tan *et al.* 2018). During this time, ongoing cladistic analyses have seen a steady refinement in the exploration of decapod evolution and interrelationships, and have established a universal agreement on the monophyletic sister clades formed by Anomura and Brachyura (Ahyong and O'Meally 2004; Tsang *et al.* 2008; Bracken-Grissom *et al.* 2013; Tan *et al.* 2018). Until around 2008, the relationships amongst groups within Anomura were considered largely resolved, with most classifications recognising three major anomuran superfamilies: Galatheoidea (squat lobsters and porcelain crabs), Hippoidea Latreille, 1825 (mole crabs) and Paguroidea Latreille, 1802 (king and hermit crabs) (Morrison *et al.* 2002; Ahyong and O'Meally 2004; Baba *et al.* 2008; Tsang *et al.* 2008). However, the emphasis of anomuran phylogeny and evolution had focused on the king and hermit crabs, leaving squat lobsters at the periphery of the debate. Since the Census of Marine Life project COMARGE (Continental Margin Ecosystems) brought together squat lobster scientists, squat lobsters have received focussed attention that resulted in a number of reviews of the group's phylogenetic relationships (Ahyong *et al.* 2011b; Schnabel *et al.* 2011a; Ahyong *et al.* 2015b). Building on the efforts undertaken during COMARGE, further reviews of anomuran phylogenetics have recently been produced by Bracken-Grissom *et al.* (2013) and Tan *et al.* (2018). Comparing with previous classifications and hypotheses of anomuran phylogeny, Bracken-Grissom *et al.* (2013) supported the status of most anomuran superfamilies and families as being monophyletic, with three families and eleven genera recovered as para- and polyphyletic, based on morphological and molecular data (Figure 2.2). Tan *et al.* (2018) conducted genome skimming to recover the mitogenome sequences of anomuran species, and revealed strong support for the relationships within families, which were largely consistent with current taxonomic classifications, yet many relationships at higher taxonomic levels remained unsolved.
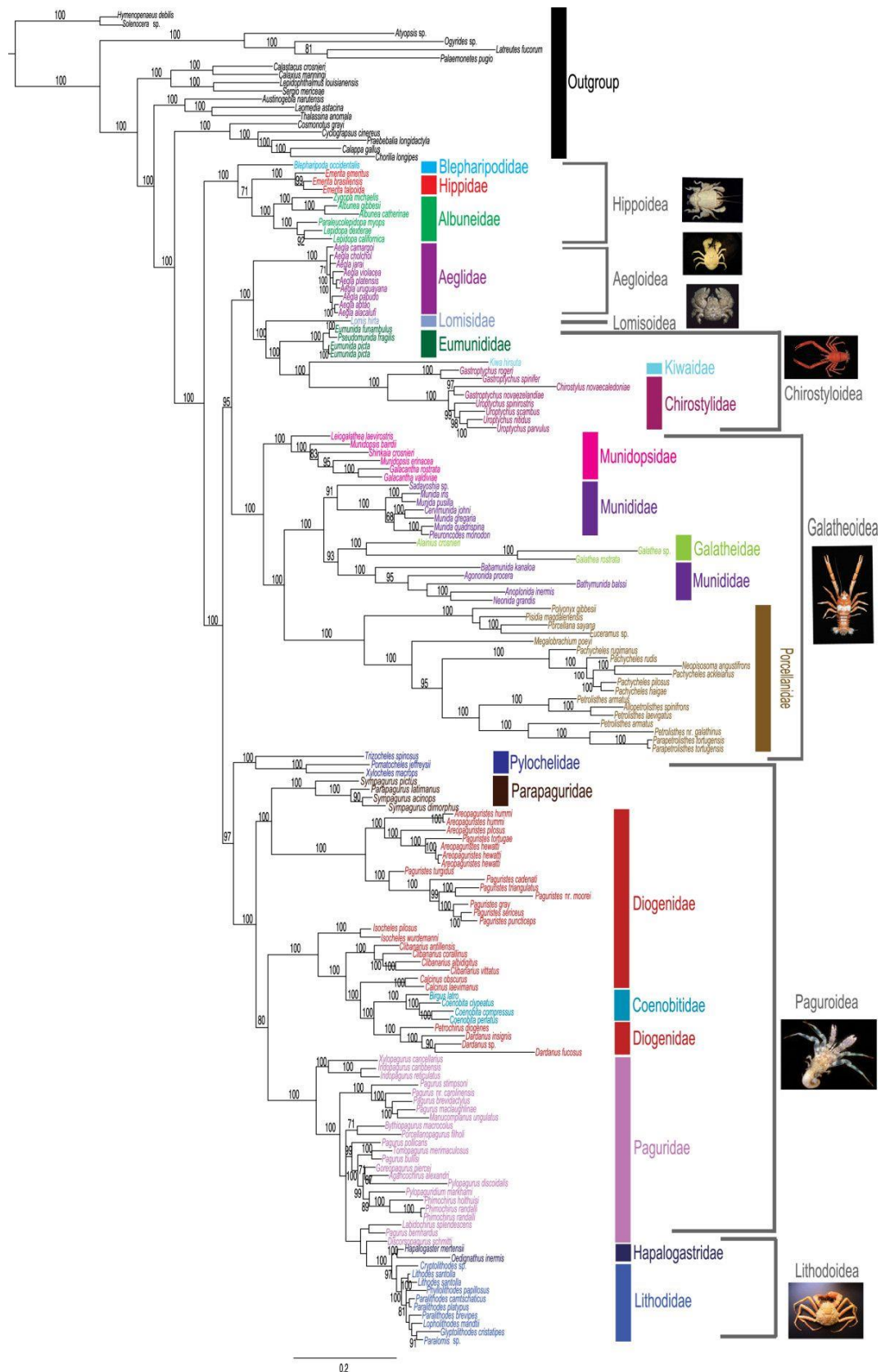
Figure 2.2 Bayesian phylogram of Anomura.

Bayesian posterior probabilities represented as percentages and >70% are noted above or below branches. Figure credit: adapted from Bracken-Grissom *et al.* (2013).

2.4   Phylogenetic   relationships   within   Galatheoidea   and   Chirostyloidea

Galatheoidea *sensu lato* is one of the most speciose and diverse superfamilies of anomuran decapod crustaceans. The phylogenetic position of the superfamily Galatheoidea has long been ambiguous, with several hypotheses proposed regarding the composition and the relationships amongst its families (e.g., Martin and Abele (1986); Scholtz and Richter (1995)). In the early 2000s, the monophyly of the Galatheoidea (including families Aeglidae, Chirostylidae, Galatheidae and Porcellanidae) was commonly accepted (Martin and Davis 2001). With the rapid development of molecular technologies, the systematics and phylogenetics of the superfamily Galatheoidea were re-examined (Morrison *et al.* 2002). According to a series of molecular-based phylogenetic studies, a clear separation was found between the marine galatheoid families and the freshwater squat lobsters Aeglidae Dana, 1852 (Morrison *et al.* 2002; Pérez-Losada *et al.* 2002a; Pérez-Losada *et al.* 2002b; Pérez-Losada *et al.* 2004; Macpherson *et al.* 2005), and the family Aeglidae was formally removed from the superfamily Galatheoidea, and placed in its own superfamily, the Aegloidea, in 2007 (McLaughlin *et al.* 2007). At around the same time, the much-publicised yeti crab, *Kiwa hirsuta* Macpherson, Jones & Segonzac, 2005 was discovered around the hydrothermal vents of the eastern Pacific and was placed in its own family Kiwaidae, within the Galatheoidea, by Macpherson *et al.* (2005).

Relationships within the remaining families of the superfamily Galatheoidea remained unresolved for some time (e.g., both Pérez-Losada *et al.* (2002a) and Macpherson *et al.* (2005) concluded that chirostylids and porcellanids were sister groups based on 18S rRNA data). This conclusion was incongruent with later findings that the family Galatheidae Samouelle, 1819 and the family Porcellanidae Haworth, 1825 were monophyletic and only remotely related to the Chirostylidae Ortmann, 1892 and the Kiwaidae Macpherson, Jones & Segonzac, 2005 (Schnabel *et al.* 2011a; Ahyong *et al.* 2015b). The latter were more closely related to the hermit crabs than they were to the remaining squat lobsters. In light of this research, the chirostylids and kiwaids were suggested to be removed from the superfamily Galatheoidea and formed a separate superfamily, Chirostyloidea (Ahyong *et al.* 2010; Schnabel and Ahyong 2010). At the same time, Schnabel and Ahyong (2010) established the family Eumunididae A. Milne Edwards & Bouvier, 1900 for the two genera *Eumunida* and *Pseudomunida* Haig, 1979, included previously in the Chirostylidae, initially based on morphological characters but since

supported by molecular analyses (Bracken-Grissom *et al.* 2013; Roterman *et al.* 2013a; Roterman *et al.* 2018).

By 2010, it appeared that the higher classification of the superfamilies Chirostyloidea (Chirostylidae, Eumunididae and Kiwaidae) and Galatheoidea *sensu stricto* (Galatheidae, Munididae Ahyong, Baba, Macpherson & Poore, 2010, Munidopsidae Ortmann, 1898 and Porcellanidae) had stabilised (Ahyong *et al.* 2010; Schnabel and Ahyong 2010; Schnabel *et al.* 2011a). The reciprocal monophyly of this two-superfamily phylogenetic classification of the squat lobsters was subsequently supported by Tsang *et al.* (2011) in a cladistic analysis of 14 anomuran families based on five nuclear protein-coding genes, and Bracken-Grissom *et al.* (2013) in a phylogenetic estimation of 19 families using two mitochondrial and three nuclear genes. The latest published phylogeny of the Anomura by Tan *et al.* (2018) is in keeping with findings on the monophyly of mostly anomuran superfamilies and families, as revealed by the aforementioned molecular-based taxonomic studies, utilising mitogenome datasets of twelve anomuran species. Here, the most up-to-date squat lobster classification scheme is followed (Ahyong *et al.* 2010; Schnabel and Ahyong 2010).

2.4.1 Phylogenetics relationships within Chirostyloidea

In 2010, two genera *Eumunida* and *Pseudomunida* were excluded from Chirostylidae, and the new family Eumunididae was established, based on morphological evidence, and a third marine squat lobster family was added to the Chirostyloidea: Chirostylidae, Eumunididae, and the deep-sea chemosynthetic-associated Kiwaidae (yeti crabs) (Macpherson *et al.* 2005; Schnabel and Ahyong 2010; Schnabel *et al.* 2011a). A number of studies supported the reciprocal monophyly of these families, and phylogenetic reconstructions in all cases place the Eumunididae as basal, with the Kiwaidae and Chirostylidae as more derived sister taxa (Schnabel *et al.* 2011a; Tsang *et al.* 2011; Ahyong *et al.* 2015b; Chu *et al.* 2015). Most recently, Baba *et al.* (2018) use morphological characters in combination with published and unpublished molecular evidence to establish a new genus *Sternostylus* Baba, Ahyong & Schnabel, 2018 and family Sternostylidae Baba, Ahyong & Schnabel, 2018. This family includes a group of 12 species previously included in the genus *Gastroptychus* Caullery, 1896, which have consistently been placed remotely in the phylogenetic tree of the Chirostyloidea (both *Gastroptychus rogeri* Baba, 2000 and *G. formosus* (Filhol, 1884) shown in Roterman *et al.* (2018) are now referred to the new family). The authors argue that the Sternostylidae are

placed at an intermediate position between the Kiwaidae and the Chirostylidae based on both molecular and morphological characters.

2.4.1.1 Phylogenetic relationships within Chirostylidae

The family Chirostylidae is currently divided into six genera: *Chirostylus* Ortmann, 1892, *Gastroptychus*, *Hapaloptyx* Stebbing, 1920, *Heteroptychus*, *Uroptychodes* Baba, 2004 and *Uroptychus*.

Heteroptychus is recently described from the Western and Central Pacific, which is established to host nine *Uroptychus* species, on the basis of morphological and molecular data (Baba 2018). Within Chirostylidae, *Gastroptychus* and *Uroptychus* were polyphyletic, based on a nine-gene dataset (Roterman *et al.* 2013a), echoing the findings of (Schnabel *et al.* 2011a). The *Gastroptychus* has been resolved with the formation of *Sternostylus*, but the genus *Uroptychus* remains to be formally addressed and most likely will require the establishment of a number of new genera (Baba *et al.* 2018).

The tropical species complex *Uroptychus naso* Van Dam, 1933 was reviewed by Poore and Andreakis (2011) based upon molecular and morphological data, describing three biologically distinct species, with the inclusion of two cryptic species. The designation of a large number of new species has been supported by molecular data with subtle morphological differences in two genera: *Uroptychus* (Schnabel *et al.* 2017; Baba 2018; Schnabel in press) and *Heteroptychus* (Baba 2018).

2.4.1.2 Phylogenetic relationships within Eumunididae

Eumunididae currently contains two genera, *Eumunida* and *Pseudomunida*, and very few cladistic analyses are available across these. The monophyly of *Eumunida* was supported by Puillandre *et al.* (2011), using a partial mitochondrial gene (COI) and a nuclear gene (28S), in addition to morphological characters, and sequencing type specimens of 13 of the 17 species examined. A putative cryptic species was discovered, and three *Eumunida* species were found to have very limited genetic divergences and were proposed to be formally united. More recently, a new species, *Eumunida spiridonovi* Macpherson, Rodriguez-Flores & Machordom, 2017, was described from the eastern coast of Africa using 16S rRNA and COI (Macpherson *et al.* 2017).

2.4.1.3 Phylogenetic relationships within Kiwaidae

Members of the Kiwaidae are found exclusively in deep-sea chemosynthetic ecosystems and are represented by a single genus, *Kiwa*. Kiwaidae was initially proposed as monotypic, to host the type species *K. hirsuta*, which was discovered on the Pacific-Antarctic Ridge (Macpherson *et al.* 2005). In 2006, a second species, *K. puravida* Thurber, Jones & Schnabel, 2011, was discovered at methane seeps on the Pacific continental slope off Costa Rica. A molecular-based analysis revealed the close affinity between the two *Kiwa* species (Thurber *et al.* 2011). A third species, *K. tyleri* Thatje, 2015, was discovered at two hydrothermal vents on the East Scotia Ridge in the Southern Ocean in 2010 (Rogers *et al.* 2012) and resolved as the sister taxon of *K. hirsuta* on the basis of data from three ribosomal genes (Thatje *et al.* 2015). Shortly after the discovery in the Southern Ocean, a further *Kiwa* species, *K.* sp. SWIR, was spotted on the Southwest Indian ridges in 2011 (Roterman *et al.* 2013a). A concatenated nine-gene sequence dataset indicated a basal split between the seep-endemic *K. puravida* and a vent clade comprising *K. hirsuta*, *K. tyleri* and *K.* sp. SWIR. This pattern was compatible with a seep-to-vent evolutionary trajectory hypothesised for mytilid mussels (Jones *et al.* 2006) and vestimentiferan tubeworms (Halanych 2005). The fifth species of deep-sea yeti crab, *K. araonae* Lee, Lee & Won, 2016 was described from a hydrothermal vent field on the Australian-Antarctic Ridge, being the second known species from the Southern Ocean (Lee *et al.* 2016b). The phylogeny, based on mitochondrial genes, revealed that the vent species *K. araonae* was genetically close to the seep-endemic *K. puravida*, despite the distant geographical locations (~12,000 km distance) between these two species, which was incongruent with the previous seep-to-vent hypothesis as proposed by Roterman *et al.* (2013a). A newly discovered species, *Kiwa* sp. GM, was collected from hydrothermal vents on the Galapagos Microplate and remained to be formally described, but brings the number of known *Kiwa* species to six (Wang *et al.* 2013a). All species have been included in the latest phylogeny of the Kiwaidae by Roterman *et al.* (2018), who have demonstrated a two-clade phylogeny with strong support. The "Bristly" clade comprises the vent species *K. araonae*, *K. puravida* and *Kiwa* sp. GM, whilst the "Plumose" clade consists of both vent and seep species *K. puravida*, *K. tyleri* and *K.* sp. SWIR. This phylogenetic pattern is in agreement with that reported by Lee *et al.* (2016b), suggesting the evolutionary progression of kiwaids may be more complicated than the previous suggestion of seep-to-vent as proposed by Roterman *et al.* (2013a). Limited support for the phylogenetic tree has been observed within the "Bristly" clade, which might be explained by rapid radiation during a cooling transition from a warmer episode.

Roterman *et al.* (2018) also argue that the primary split within this clade is near the equator, followed by a southward migration.

2.4.2 Phylogenetic relationships within Galatheoidea

Galatheoidea has been revised to exclude the freshwater squat lobster Aeglidae, Chirostylidae and Kiwaidae, and comprises three marine squat lobster families (Galatheidae, Munididae and Munidopsidae) and the porcelain crabs (Porcellanidae) (Ahyong *et al.* 2010). The internal galatheoid relationships are still not adequately resolved, particularly when considering the Galatheidae and Munididae, which are the most speciose and diverse groups of the squat lobsters (Ahyong *et al.* 2011b). On the basis of two mitochondrial and three nuclear markers, the Munididae was demonstrated to be possibly paraphyletic with Galatheidae nested inside, whilst Galatheidae and Munidopsidae were recovered as monophyletic with strong support (Bracken-Grissom *et al.* 2013). However, based on phylogenetic analyses that combined both molecular (three ribosomal genes) and morphological characters, Galatheidae is revealed as the sister group of the Munididae (Rodríguez-Flores *et al.* 2018). It appears that the relationships, even at a family level, have not entirely stabilised to date.

2.4.2.1 Phylogenetic relationships within Galatheidae

Species discrimination amongst the squat lobsters of the family Galatheidae is proving extremely difficult due to their many conservative morphological traits (Cabezas *et al.* 2009b). Hence, DNA sequencing is proving to be invaluable in aiding species identifications in a group that is morphologically highly conserved (Macpherson and Robainas-Barcia 2015). The evolution of the family Galatheidae has been marked by rapid speciation accompanied by morphological stasis or some constraints on the morphological evolution, rather than a consequence of adaptation to extreme environments, resulting in the existence of similar morphotypes that are genetically divergent (Machordom and Macpherson 2004; Baba *et al.* 2008; Cabezas *et al.* 2012b; Palero *et al.* 2017).

For example, in the genus *Lauriea*, a combined morphological and molecular analysis revealed the existence of six new species, with subtle morphological characters and distinct molecular differences (Macpherson and Robainas-Barcia 2013). Among these, five species belong to *Lauriea*, whereas the sixth was referred to a new genus, *Triodonthea*. Subsequently, Palero *et al.* (2017) analysed concatenated DNA sequences (two mitochondrial and one nuclear

genes) for the reassessment of the taxonomy of *Lauriea*. The molecular phylogeny showed that *Lauriea siagiani* Baba, 1994 was phylogenetically closer to two monotypic genera, *Triodonthea* and *Macrothea* Macpherson & Cleva, 2010, than to other *Lauriea* species, suggesting the previously described morphospecies still underestimated the real diversity.

The genus *Galathea* is one of the most speciose and diverse groups in the family Galatheidae (Macpherson and Robainas-Barcia 2015). A detailed systematic review, with combined molecular (two mitochondrial genes) and morphological analyses of more than 9000 specimens of 144 species resulted in the description of 92 new species (Macpherson and Robainas-Barcia 2015), confirming the great utility of molecular markers in taxon discrimination.

Clear genetic differences were also found in the mitochondrial gene phylogeny of *Fennerogalathea*, with two major divergent clades and description of three new species of this genus (Rodríguez-Flores *et al.* 2017).

The genus *Coralliogalathea* had long been considered monospecific, comprising *Coralliogalathea humilis* (Nobili, 1905), which lives in close association with corals in the Indian and Pacific Oceans. Using a combined molecular and morphological approach based on a concatenated mitochondrial genes dataset, the validity of two species that are previously synonymised with *C. humilis* and the existence of three new sympatric species was shown. Phylogenetic analyses have revealed the *Coralliogalathea* as a well-supported monophyletic group, and the authors suggest that the genus might be a basal genus of the superfamily, which most likely diversified during the Miocene (Rodríguez-Flores *et al.* 2018).

Another monospecific genus, *Allogalathea*, was examined using a combined molecular and morphological approach (Cabezas *et al.* 2011b). Four deeply divergent clades were revealed, indicating that *Allogalathea elegans* (Adams & White, 1848) was a species complex comprising four species, three of which were described as new.

2.4.2.2 Phylogenetic relationships within Munididae

Munididae, including the highly diverse genera *Agononida* and *Munida*, contains the majority of galatheoids and presents a critical phylogenetic challenge. A number of studies have previously demonstrated that the intrageneric phylogenetic relationships within Munididae could not be deduced with confidence owing to their morphological complexity and cryptic diversity.

Well-supported monophyletic assemblages of smaller genera such as *Babamunida* (Cabezas *et al.* 2008), *Hendersonida* Cabezas & Macpherson, 2014 (Mccallum *et al.* 2016), *Paramunida* (Machordom and Macpherson 2004; Cabezas *et al.* 2009b), *Plesionida* (Cabezas *et al.* 2009b), *Raymunida* (Macpherson and Machordom 2001; Machordom and Macpherson 2004; Cabezas *et al.* 2009b) and *Sadayoshia* Baba, 1969 (Palero *et al.* 2017), had been confirmed by different molecular studies. However, studies that examined the genus *Munida* found it to be polyphyletic, e.g., using a five-gene phylogenetic analysis (Bracken-Grissom *et al.* 2013). The same finding was revealed from an analysis of the genus *Agononida*, which was not recovered as monophyletic, likely due to hidden diversity or species complexes (Machordom and Macpherson 2004; Cabezas *et al.* 2009b).

Showcasing the unlocking of previously hidden diversity within a single widespread species, are the studies of *Agononida incerta* (Henderson, 1888). This species was previously known across the entire Indo-West Pacific, but the use of molecular tools (COI and 16S) revealed a total of eight new species (Macpherson and Baba 2009; Poore and Andreakis 2012, 2014) and helped identify two distinct clades that displayed small but constant differences in morphological characters (facial-spine clade, without facial-spine clade) (Poore and Andreakis 2014).

According to a series of detailed molecular-based phylogenetic and taxonomic studies of the genus *Paramunida*, *P. granulata* (Henderson, 1885) was found to occupy a separate node with the remaining species of *Paramunida* united remotely, indicating this species might be the single representative of a different lineage (Cabezas *et al.* 2010; Cabezas *et al.* 2012b; Cabezas and Chan 2014). Subsequently, a new genus, *Hendersonida*, was proposed to accommodate *P. granulata* (Cabezas and Macpherson 2014). These findings were also supported by a later study employing morphological treatment and molecular evidence (Mccallum *et al.* 2016).

Munida is by far the most speciose and diverse genus within the family Munididae and within the squat lobsters as a whole. Although this group has attracted considerable taxonomic attention during the last decade (Baba 2005; Baba *et al.* 2008), the consensus is that the number of species remains largely underestimated with numerous cryptic species being predicted to exist (Macpherson and Machordom 2005). The genus was previously split, e.g., species previously considered as *Munida* were combined in the genera *Agononida* by Baba and de Saint Laurent (1996), *Crosnierita* Macpherson, 1998, *Raymunida* (Macpherson and Machordom 2000) and *Babamunida* (Cabezas *et al.* 2008). Combined molecular and

morphological phylogenetic analyses strongly supported the taxonomic splitting of the genus *Munida* (Machordom and Macpherson 2004).

The lack of resolution for internal relationships was observed within the genus *Munida* based on two mitochondrial genes, with high levels of genetic divergence, suggesting the existence of rapid diversification events in the evolutionary progression (Machordom and Macpherson 2004; Cabezas *et al.* 2011a). A recent phylogenetic analysis of four *Munida* species collected from cold-water coral and cold seep communities in the northwestern Atlantic Ocean and the Gulf of Mexico, based on combined molecular and morphological data, demonstrated that more recent phylogenetic relationships tend to have high support, but some older relationships remain unresolved (Coykendall *et al.* 2017) (Figure 2.3).

Interestingly, the taxonomic status of two sympatric morphospecies, *Munida gregaria* (Fabricius, 1793) and *Munida subrugosa* Dana, 1852, was investigated by DNA sequence analysis, only to find them identical and invariant, suggesting that the two species represent a single dimorphic species (Pérez-Barros *et al.* 2008). A follow-up study was conducted by Diez *et al.* (2016), who elucidated evidence that *M. gregaria* is the only recorded species with two morphotypes, and hypothesise that transitions between the two morphotypes do not take place, i.e., pelagic individuals of the '*gregaria*' form do not transform to the '*subrugosa*' form, or vice versa. This conflict between genetic and morphological data supported the fact that squat lobster phylogenetic relationships are not easily assessed. Another interesting study was conducted by Pérez-Barros *et al.* (2014), who investigated the genetic cohesion of *M. gregaria* (morphs *M. gregaria* and *M. subrugosa*) across the East Pacific Barrier using mitochondrial markers, yet no signal of genetic differentiation amongst populations was observed, confirming *M. gregaria* constitutes a real trans-Pacific species.

Within the family Munididae, the existence of cryptic or pseudo-cryptic sibling species had been reported within the genera *Munida* (Machordom and Macpherson 2004; Macpherson and Machordom 2005) and *Paramunida* (Cabezas *et al.* 2010). In addition, numerous new species were examined and described in the genera *Agononida* (Cabezas *et al.* 2009b), *Babamunida* (Schnabel *et al.* 2009), *Munida* (Cabezas *et al.* 2009b; Cabezas *et al.* 2011a; Macpherson *et al.* 2017), *Paramunida* (Cabezas *et al.* 2009b; Cabezas and Chan 2014; Mccallum *et al.* 2016), *Plesionida* (Cabezas *et al.* 2009b) and *Raymunida* (Macpherson and Machordom 2001).
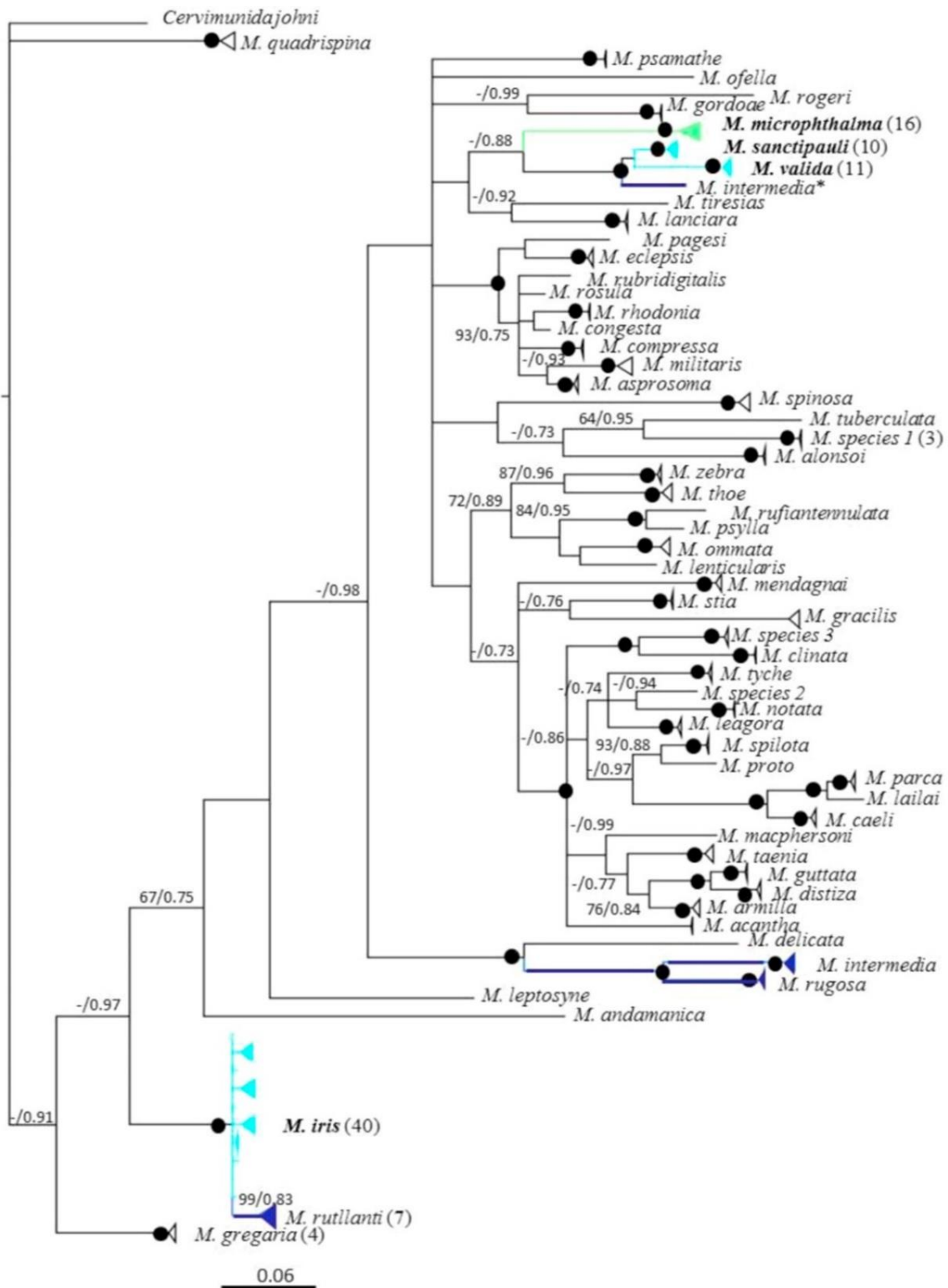
Figure 2.3 Bayesian phylogram of *Munida*.

Bootstrap support value ≥ 60% (maximum likelihood) and posterior probabilities ≥ 0.70 (Bayesian inference) are noted above branches. Scale bar denotes the number of nucleotide substitutions per site. Figure credit: adapted from Coykendall *et al.* (2017).

2.4.2.3 Phylogenetic relationships within Munidopsidae

As early as 2000, Creasey *et al.* (2000) applied ten allozyme markers to evaluate the genetic relationships of five species within the genus *Munidopsis*, and revealed this genus to be putatively monophyletic. This study appeared to be the first piece of recorded molecular-based phylogenetic work on squat lobsters.

Munidopsidae was previously divided into two subfamilies, Munidopsinae and Shinkaiinae. However, according to Ahyong *et al.* (2015b) and (Schnabel *et al.* 2011a), the *Galathea*-like *Leiogalathea* turned out to be a munidopsid, which resulted in Munidopsidae currently comprising four genera: *Galacantha* A. Milne Edwards, 1880, *Leiogalathea*, *Munidopsis* and *Shinkaia* Baba & Williams, 1998. The phylogenetic framework of these four munidopsid genera was recognised based on two mitochondrial genes, rendering the former bi-subfamily classification phylogenetically invalid (Ahyong *et al.* 2011a). Within the Munidopsidae, *Leiogalathea* was sister to the remaining munidopsids, whereas *Shinkaia* and *Galacantha* were nested within *Munidopsis*, making this genus polyphyletic. These findings were supported by mitochondrial genes based phylogenetic analyses (Jones and Macpherson 2007; Ahyong *et al.* 2011a), a five nuclear protein-coding genes dataset (Bracken-Grissom *et al.* 2013), and a combined molecular and morphological dataset (Coykendall *et al.* 2017). In the latter study, similar to *Munida*, nine *Munidopsis* species were sampled from cold-water coral and cold seep communities in the northwestern Atlantic Ocean and the Gulf of Mexico, and assessed in a phylogenetic and population genetic context with regard to habitat and biogeography, with seep taxa scattered across the phylogenetic tree (Coykendall *et al.* 2017).

The existence of pseudo-cryptic sibling species was spotted in the genus *Munidopsis* from the East Pacific Rise, with most *Munidopsis* individuals previously considered as *Munidopsis subsquamosa* Henderson, 1885 from the East Pacific Rise appearing to be members of *M. recta* Baba, 2005 (Jones and Macpherson 2007). Another finding of this study was that *Munidopsis* populations were closely interconnected between distant and discrete localities, supporting the findings that there is no correlation between morphological and genetic divergence (Hendry *et al.* 2000). Additionally, new species were found in the genus *Munidopsis* as reported by several studies (Jones and Macpherson 2007; Coykendall *et al.* 2017; Macpherson *et al.* 2017).

2.5 Divergence time estimation

A few older studies have examined deeper divergence times across the wider order Decapoda or infraorder Anomura (Porter *et al.* 2005; Bracken-Grissom *et al.* 2013), whilst most recent studies have examined more recent divergences within a family or genus (e.g., Palero *et al.* (2017); Roterman *et al.* (2018)).

In 2005, a divergence time estimation for the entire Decapoda was conducted by Porter *et al.* (2005), utilising newly developed Bayesian and likelihood methods based on nuclear (18S rRNA, 28S rRNA and Histone H3 gene) and mitochondrial (16S rRNA) genes, with eight fossil and geological calibrations. Results demonstrated that decapods originated earlier than 437 MYA, followed by rapid speciation within the group. For example, Anomura lineages originated in the Permian-Triassic with all the extant superfamilies (i.e., Galatheoidea, Hippoidea, Lomisoidea and Porcellanidae) radiating between 244 and 309 MYA.

Bracken-Grissom *et al.* (2013) incorporated 31 fossil calibrations to estimate divergence time within anomurans, using a combined molecular and morphological approach based on two mitochondrial (16S rRNA and 12S rRNA) and three nuclear (H3, 18S rRNA and 28S rRNA) genes. The origin of Anomura from the true crabs, Brachyura, was dated to the Late Permian (259 MYA). Two superfamilies, Galatheoidea and Chirostyloidea, were derived from a squat lobster-like ancestor approximately 206 MYA. Galatheoidea split into the Munidopsidae, and the other families at 180 MYA, whilst the other two squat lobster families Munididae and Galatheidae arose soon thereafter at 150 MYA. Within the Chirostyloidea, the three families Chirostylidae, Eumunididae, and Kiwaidae all originated in the Cretaceous (109, 95 and 95 MYA, respectively).

In the same year, and based on a nine-gene dataset with 3 fossil calibrations across 15 chirostyloids, Roterman *et al.* (2013a) arrived at very similar divergence times, with the three Chirostyloidea families originating in the Mid-Cretaceous, i.e., Chirostylidae (106 MYA), Eumunididae (115 MYA), and Kiwaidae (106 MYA), whilst most chirostylids and extant kiwaids radiated from the Eocene onwards. This study was extended by Roterman *et al.* (2018), utilising a longer alignment of the nine-gene dataset with more fossil calibrated nodes, with divergence time estimations broadly similar to the previous study.

The first study using divergence time estimations that considered squat lobsters was that of Machordom and Macpherson (2004) who examined the diversification of the genus *Munida*. These authors resolved a radiation between 7–14 million years ago (MYA), which

corresponds with the Middle or Late Miocene, on the basis of general mean divergence values of two mitochondrial genes (16S rRNA and COI) (Brown *et al.* 1982; Moritz *et al.* 1987; Schubart *et al.* 2000; Stillman and Reeb 2001). The authors pointed out that this group might have experienced rapid radiation, followed by constraints or stasis in its morphological evolution during the Miocene. However, no molecular clock calibration was taken, given that no fossil record was available. Similarly, Cabezas *et al.* (2011b) speculated that the diversification of *Allogalathea* might have occurred during the Late Miocene, based on estimated COI interspecific divergence of this genus (Brown *et al.* 1982; Moritz *et al.* 1987).

With the absence of a fossil record of the deep-sea genus *Paramunida*, Cabezas *et al.* (2012b) reconstructed a molecular phylogeny based on the combination of mitochondrial and nuclear molecular genes. These results provided first indications that both shallow water and deep-sea environments might have been affected by similar evolutionary and ecological processes, that is, a period of accelerated radiation, in the late Oligocene-Miocene, followed by a slower lineage accumulation rate since then. Besides, the southwest Pacific region was identified as a major centre for the diversification of *Paramunida*. The importance of the southwest Pacific region was also highlighted by Macpherson *et al.* (2010), who found that the Coral Sea (Solomon-Vanuatu-New Caledonia islands) and Indo-Malay-Philippines archipelago exhibited the highest biodiversity of squat lobsters in the Pacific Ocean.

More recently, Palero *et al.* (2017) combined two mitochondrial (16S rRNA and COI) genes and one nuclear gene (H3) to estimate the divergence time of two shallow water genera, *Lauriea* and *Sadayoshia*. Results were in agreement with the hypothesis that rapid speciation for shallow water squat lobsters from the Indo-Pacific region arose in the late Paleogene-Neogene transition (< 50 MYA), which might be related to increased tectonic activity and corresponding changes in ocean currents and distribution of habitats during late Paleogene. Similarly, the tropical southwest Pacific region was recognised as a major diversification centre for shallow-water squat lobster taxa, from where species dispersed to Indian and other Pacific Oceans, which was congruent with previous studies (Macpherson *et al.* 2010; Cabezas *et al.* 2012b).

The latest published calibrated phylogenetic tree for the *Coralliogalathea* species and related genera (44 taxa) within Galatheoidea, using multiple calibrations based on fossils and nucleotide substitution rate, was constructed on the basis of two mitochondrial genes, COI and 16S (Rodríguez-Flores *et al.* 2018). Results showed that *Coralliogalathea* emerged earlier than

other genera of Galatheoidea, with a most recent common ancestor (MRCA) at 62 MYA, during the Miocene.

In contrast, very little is known for chirostyloids, which is in part due to the near entire absence of this group in the fossil record (Ahyong *et al.* 2011b). A concatenated mitochondrial (16S rRNA and COI) and nuclear (Histone H3 gene) dataset was applied to approximate divergence time of the genus *Uroptychus*, using a log-normal uncorrelated relaxed molecular clock with two-calibration fossils (Poore and Andreakis 2011). The MRCA was estimated for *Uroptychus* at 66–98 MYA. Given that no substitution rate was available for anomurans for either 18S or 28S, Rogers *et al.* (2012) estimated the divergence date between two *Kiwa* species, *Kiwa tyleri* and *K. hirsuta*, at 12 MYA, in light of the divergence rate of 16S rRNA (Stillman and Reeb 2001). The number of kiwaid species was subsequently increased to four (*K. puravida*, *K. tyleri*, *K. hirsuta* and *K.* sp. SWIR), with findings showing that Kiwaidae radiated at 30.6 MYA (22.7–39.3 MYA), and the split between the Pacific and non-Pacific lineages occurred at 19.1 MYA (13.4–25.9 MYA) (Roterman *et al.* 2013a). Consequently, Roterman *et al.* (2018) incorporate all six extant species of kiwaid squat lobsters and updated the time estimate for the MRCA for the Kiwaidae to 33.2–38.8 MYA, with "Plumose" clade and "Bristly" clade at 17.1–21.71 MYA and 13.3–15.9 MYA, respectively.

2.6 Conclusion

To date, molecular-based studies have dramatically increased our understanding of squat lobster phylogenetics, systematics, and thereby the taxonomy of this extraordinarily complicated and challenging group. Compared with the extensive applications of molecular tools in the aforementioned research fields, less work has been achieved in the population genetics of squat lobsters, and those in the context of vulnerable marine ecosystems (VMEs) are even more limited. The systematic placement of squat lobster taxa has been well studied and defined, making them ideal candidates for population genetic study. Detailed genetic data can be utilised to estimate patterns of connectivity amongst marine populations of some lesser well-studied VME-associated squat lobster taxa, but only once the taxonomy of the study taxa has been resolved. Knowledge of taxonomic, phylogenetic and phylogeographic relationships of squat lobsters is critical for the interpretation of findings as revealed by population genetic studies, in light of their geographical distribution and evolutionary history, and will ultimately contribute to the management and protection of VMEs.

Appendix

Table S2.1 Molecular work on taxonomy, systematics and phylogenetics for the marine squat lobsters.

	Family	*Genus*	Molecular methods	Resources	Reference
Decapoda	Eumunididae, Munididae	*Eumunida, Munida*	16S, 18S, 28S	Sequences from GenBank	Ahyong and O'Meally (2004)
	Chirostylidae, Munididae	*Uroptychus, Munida*	16S, 18S, 28S, H3	Specimens from the collection of Monte L. Bean Life Science Museum (USA)	Porter et al. (2005)
	Munididae	*Munida, Paramunida*	NaK, PEPCK	Taiwan (China)	Tsang et al. (2008)
	Chirostylidae, Eumunididae, Kiwaidae, Munididae, Munidopsidae	*Uroptychus, Eumunida, Kiwa, Munida, Munidopsis*	16S, 18S, 28S, H3	Sequences from GenBank	Bracken et al. (2009)
	Munidopsidae	*Shinkaia*	Mitogenome	Sequences from GenBank	Shen et al. (2013)
	Chirostylidae, Eumunididae, Kiwaidae, Munididae, Munidopsidae	*Uroptychus, Eumunida, Kiwa, Munida, Paramunida, Shinkaia*	NaK, PEPCK	Not stated	Chu et al. (2015)
Anomura	Eumunididae, Munididae	*Eumunida, Munida*	COII, 16S, 18S, 28S	Sequences from GenBank	Morrison et al. (2002)
	Chirostylidae, Munididae,	*Uroptychus, Munida*	18S	Brazil, Chile, USA	Pérez-Losada et al. (2002b)

Chirostylidae, Kiwaidae, Munididae	*Uroptychus, Kiwa, Munida,*	18S	Pacific-Antarctic Ridge	Macpherson et al. (2005)
Chirostylidae, Eumunididae, Kiwaidae, Galatheidae, Munididae, Munidopsidae	*Chirostylus, Gastroptychus, Sternostylus,Uroptychodes, Uroptychus, Eumunida, Pseudomunida, Kiwa, Alainius, Allogalathea, Galathea, Agononida, Anoplonida, Babamunida, Bathymunida, Cervimunida, Heteronida, Munida, Neonida, Pleuroncodes, Sadayoshia, Torbenella, Galacantha, Leiogalathea, Munidopsis, Shinkaia*	16S, 18S, 28S	Sequences from GenBank; Specimens from collections of Natural History Museum of Los Angeles County (US), Natural History Museum of Vienna (Austria), Museum National d'Histoire Naturelle (France), National Taiwan Ocean University (Taiwan, China), National Institute of Water and Atmospheric Research (Wellington, NZ)	Schnabel et al. (2011a)
Chirostylidae, Eumunididae, Kiwaidae, Munididae, Munidopsidae	*Uroptychus, Eumunida, Kiwa, Cervimunida, Munida, Paramunida, Munidopsis, Shinkaia*	AK, En, GAPDH, NaK, PEPCK	Taiwan (China), Okinawa, South East Pacific Rise	Tsang et al. (2011)
Chirostylidae, Eumunididae, Kiwaidae, Galatheidae, Munididae, Munidopsidae	*Chirostylus, Gastroptychus, Sternostylus,Uroptychus, Eumunida, Pseudomunida, Kiwa, Alainius, Galathea, Agononida, Babamunida, Cervimunida, Munida, Neonida, Pleuroncodes, Sadayoshia, Galacantha, Leiogalathea, Munidopsis, Shinkaia*	12S, 16S, 18S, 28S, H3	Sequences from GenBank; Specimens collected on cruise and field expeditions, from collaborators, from the collection of the University of Louisiana at Lafayette (USA)	Bracken-Grissom et al. (2013)

	Chirostylidae, Eumunididae, Kiwaidae, Galatheidae, Munididae, Munidopsidae	*Chirostylus, Gastroptychus, Uroptychus, Eumunida, Pseudomunida, Kiwa, Galathea, Agononida, Allogalathea, Cervimunida, Munida, Pleuroncodes, Sadayoshia, Galacantha, Leiogalathea, Munidopsis, Shinkaia*	16S, 18S, 28S	Sequences from collaborators; Specimens from collections of Museum National d'Histoire Naturelle (France), National Institute of Water and Atmospheric Research (Wellington, New Zealand), National Taiwan Ocean University (Taiwan, China)	Ahyong et al. (2015)
	Chirostylidae, Munididae	*Gastroptychus, Sternostylus, Munida*	Mitogenome	Kulumburu L29 transect, Northwestern Australia, Australia, Southern Ocean, Tasmania, Mongrel Seamount, Dory Hill	Tan et al. (2018)
	Eumunididae, Munididae, Munidopsidae	*Eumunida, Munida, Munidopsis*	COI, 16S	Western Indian Ocean	Macpherson et al. (2017)
Chirostyloidea	Chirostylidae, Eumunididae, Kiwaidae	*Chirostylus, Gastroptychus, Uroptychodes, Uroptychus, Sternostylus, Eumunida, Kiwa*	COI, 16S, 18S, 28S, NaK, En, AK, GAPDH, PEPCK	Sequences from GenBank; Specimens collected on two research cruises on East Scotia Ridge (Scotia Sea), South West Indian Ridge and surrounding seamounts; Specimens from collections of National Institute of Water and Atmospheric Research (Wellington, New Zealand), National Taiwan Ocean University (Taiwan, China), Oxford University Museum of Natural History (UK), Centro de Estudios Avanzados de Blanes (Spain), Leetown Science Center (USA)	Roterman et al. (2013a)

	Chirostylidae	*Uroptychus*	COI, 16S, H3	Indonesia, Philippines, Taiwan (China), Japan, north-western Australia	Poore and Andreakis (2011)
		Uroptychus	COI	Macquarie Ridge (Wellington, New Zealand)	Schnabel et al. (2017)
		Uroptychus, Heteroptychus	Molecular methods	Western and Central Pacific	Baba (2018)
	Eumunididae	*Eumunida*	COI, 28S	South West Pacific and Indian Oceans	Puillandre et al. (2011)
	Kiwaidae	*Kiwa*	18S	Costa Rica	Thurber et al. (2011)
		Kiwa	16S, 18S, 28S	Southern Ocean	Rogers et al. (2012)
		Kiwa	COI, 16S, 18S, 28S	Southern Ocean	Thatje et al. (2015)
		Kiwa	16S	Southern Ocean	Lee et al. (2016b)
		Kiwa	COI, 16S, 18S, 28S, NaK, En, AK, GAPDH, PEPCK	East Pacific Ocean	Roterman et al. (2018)
Galatheoidea	Galatheidae, Munididae	*Lauriea, Macrothea and Triodonthea, Sadayoshia*	COI, 16S, H3	Indo-Pacific Oceans	Palero et al. (2017)
	Munididae, Munidopsidae	*Munida, Munidopsis*	COI	Western North Atlantic Ocean	Coykendall et al. (2017)
	Galatheidae	*Allogalathea*	COI, 16S, PEPCK	Indo-West Pacific Ocean	Cabezas et al. (2011b)
		Coralliogalathea	COI, 16S, 18S, 28S	Indian and Pacific Oceans	Rodríguez-Flores et al. (2018)
		Fennerogalathea	COI, 16S	Pacific Ocean	Rodríguez-Flores et al. (2017)

	Galathea	COI, 16S	Indian and Pacific Oceans	Macpherson and Robainas-Barcia (2015)
	Lauriea, Triodonthea	COI, 16S	Pacific and Indian Oceans	Macpherson and Robainas-Barcia (2013)
Munididae	*Agononida, Munida, Paramunida, Plesionida*	COI, 16S	Solomon Islands, Fiji, New Caledonia	Cabezas et al. (2009b)
	Agononida	COI	Australia, East Africa, Taiwan (China), Philippines and Norfolk Ridge (south-western Pacific)	Poore and Andreakis (2012)
	Agononida	COI, 16S	Madagascar, Vanuatu, Papua New Guinea, Fiji, eastern Australia and French Polynesia	Poore and Andreakis (2014)
	Babamunida	COI, 16S	South West Pacific and Indian Oceans	Cabezas et al. (2008)
	Babamunida	COI, 16S	North-western Hawaiian Islands	Schnabel et al. (2009)
	Munida	COI, 16S	Southwest Pacific	Machordom and Macpherson (2004)
	Munida	COI, 16S	New Caledonia	Macpherson and Machordom (2005)
	Munida	COI, 16S	Taiwan (China)	Cabezas et al. (2011a)

	Munida	COI, ND1	Southeast Pacific, southern tip of South America, southwest Atlantic, southwest Pacific	Pérez-Barros et al. (2014)
	Munida	COI, 16S, ND1, ITS	Southern South America	Perez-Barros et al. (2008)
	Paramunida	16S, ND1	Southwestern Indian Ocean, Philippines, Taiwan (China), Indonesia, Solomon Islands, Vanuatu, New Caledonia, Fiji, Tonga, Wallis, Futuna, French Polynesia	Cabezas et al. (2010)
	Paramunida	COI, 16S, 18S, 28S, ND1	Indo-Pacific Ocean	Cabezas et al. (2012b)
	Paramunida	16S, ND1	Philippines	Cabezas and Chan (2014)
	Paramunida	16S, ND1	North-western Australia	Mccallum et al. (2016)
	Raymunida	COI	Pacific and Indian Oceans	Macpherson and Machordom (2001)
Munidopsidae	*Galacantha, Leiogalathea, Munidopsis, Shinkaia*	COI, 16S	Sequences from GenBank; Specimens from collections of Museum Victoria (Australia), National Institute of Water and Atmospheric Research (Wellington, New Zealand), University of Louisiana at Lafayette (USA), Catholic University of the North (Chile)	Ahyong et al. (2011a)

Leiogalathea	COI, 16S	Atlantic, Indian and Pacific Oceans	Rodríguez-Flores et al. (2019)
Munidopsis	Allozymes	Northeast Atlantic, West Indian, Northeast Pacific Oceans, English Channel, Arabian Sea, Wembury beach (UK)	Creasey et al. (2000)
Munidopsis	COI	East Pacific Ocean	Jones and Macpherson (2007)

Mitochondrial DNA (mtDNA): COI, cytochrome *c* oxidase subunit I; COII cytochrome *c* oxidase subunit II; 12S; 16S. Nuclear DNA (nDNA): H3, histone H3; PEPCK, phosphoenolpyruvate carboxykinase; NaK, sodium potassium ATPase a-subunit; AK, arginine kinase; En, enolase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; ND1, NADH dehydrogenase subunit 1; ITS1, internal transcribed spacer 1; 18S; 28S.

Chapter 3. Development and characterisation of twenty polymorphic microsatellite loci in the deep-sea squat lobster, *Munida isos* Ahyong & Poore, 2004 and cross-amplification in two congeneric species

Abstract

Munida isos is a deep-sea squat lobster species that is widely distributed within the New Zealand and the east Australian regions and is often associated with deep-sea vulnerable marine ecosystems (VMEs). To investigate its population genetic structure and patterns of regional connectivity, microsatellite loci for *M. isos* were developed from two genomic libraries using the Illumina HiSeq 2500 sequencing platform. Twenty-six loci amplified consistently in *M. isos* from the Tasman Sea, amongst which twenty loci were polymorphic and selectively neutral. Evidence of null alleles was observed at eight loci. Most loci exhibited moderate to high levels of polymorphism, with an average Polymorphic Information Content value of 0.482. The mean number of alleles per locus was 7.45, with a mean expected heterozygosity of 0.520. Thirteen loci exhibited significant deviation from Hardy-Weinberg equilibrium, whilst only one locus pair was in linkage disequilibrium after False Discovery Rate correction for multiple testing ($p < 0.05$). Cross-species amplification tests revealed that the transferability of 14 loci (70%) was positive for the two congeners *M. endeavourae* and *M. gracilis*. The accessibility to new polymorphic microsatellite loci will facilitate population genetic studies and aid in developing conservation and management strategies for VMEs.

3.1 Introduction

Microsatellites, also known as simple sequence repeats (SSRs) or short tandem repeats (STRs), have emerged as one of the most versatile and powerful marker types, with applications in many different fields (e.g., genetic breeding, genome mapping, gene location, species identification), and particularly in population genetic studies, owing to their intraspecific variability, high transferability, co-dominant inheritance and reproducibility (Tautz 1989; Wright and Bentzen 1995; Selkoe and Toonen 2006). The widespread use of microsatellites has provided a comparative phylogenetic and population genetic framework for the evaluation of genetic diversity and connectivity in many marine invertebrates, and has contributed to the merging of population genetic and phylogenetic studies (Slatkin and Barton 1989; Frankham *et al.* 2002).

Until recently, the development and optimisation of microsatellite loci involved the screening of genomic libraries obtained by cloning, which was labour-intensive, time-consuming and expensive. Previous attempts in acquiring microsatellite loci from squat lobster species were mainly based on the construction of enriched genomic libraries (Cabezas *et al.* 2009a; Bailie *et al.* 2010; Molecular Ecology Resources Primer Development Consortium *et al.* 2010; Wang and Held 2015) using a range of protocols (Kijas *et al.* 1994; Jones *et al.* 2002; Zane *et al.* 2002; Bloor *et al.* 2006; Selkoe and Toonen 2006; Leese *et al.* 2008). Specifically, Molecular Ecology Resources Primer Development Consortium *et al.* (2010) isolated eight di- and trinucleotide microsatellite loci for *Munida isos* Ahyong & Poore, 2004 collected from submarine mountains and continental slopes in southeastern Australia, using methods described by Jones *et al.* (2002). However, microsatellite loci development using the enrichment library method has been relatively problematic in some marine invertebrates (Selkoe and Toonen 2006), which may be due to the existence of inconspicuous cryptic repeats on presumed unique flanking regions in microsatellite sequences (Bailie *et al.* 2010).

Alternatively, next generation methodologies may pave the way forward for marker development of species characterised by complex genome organisation, particularly in crustaceans (Bailie *et al.* 2010). Currently, based on high throughput parallel and multiplexing genotyping (e.g., Roche 454, Illumina MiSeq/HiSeq, ABI-SOLiD), High-Throughput Sequencing (HTS) technologies have reduced the time and cost associated with the construction of enrichment libraries for microsatellite isolation. To date, this technique has been employed to obtain a large number of microsatellite loci for several squat lobster species

from the southwest Pacific Ocean, North Atlantic Ocean, Southern Ocean and Northwest Pacific Ocean (Boyle *et al.* 2013; Coykendall and Morrison 2013; Roterman *et al.* 2013b; Nakajima *et al.* 2018). Notably, the Illumina HiSeq 2500 paired-end sequencing system, one of the most popular HTS technologies, is capable of identifying numerous microsatellite loci, both quickly and cheaply, and in doing so, it provides a large amount of data.

Munida isos, typically associated with seamounts (i.e., vulnerable marine ecosystems, VMEs), is an abundant squat lobster species that is widespread in New Zealand region, and also in eastern Australia in a depth range of 462–2756 m (Schnabel 2009; Clark *et al.* 2010; Ahyong *et al.* 2015a). Hence, given the potential difficulties of obtaining many microsatellite loci using enrichment methods to construct the genomic library, the Illumina HiSeq 2500 platform was selected for microsatellite development for *M. isos* to increase the number of informative microsatellite loci with varying repeat motifs (di-, tri- and tetranucleotides).

Sourcing suitable microsatellite loci from closely related species, using cross-species amplification, is a cost-effective and time-saving alternative to develop new primers, especially for species with limited available genetic information and when resources are also limited. To date, cross-species amplification transferability has been successful in several squat lobster species (Boyle *et al.* 2013; Wang and Held 2015). For the present study, cross-amplification of microsatellite primers developed for *M. isos* is tested with *Munida endeavourae* Ahyong & Poore, 2004 and *Munida gracilis* Henderson, 1885, two congeneric species in the family *Munida*. *M. endeavourae* primarily inhabits seamounts (554–2756 m) that occur in the northern region of the New Zealand's Exclusive Economic Zone (EEZ) and in the southeastern Australia region, whereas *M. gracilis* is normally found on both seamounts and sediments (i.e., non-VMEs) habitats (44–1211 m) that are distributed across the entire New Zealand continental shelf (Schnabel 2009).

The objective of the study recorded in this chapter is: 1) to develop and characterise a suite of novel microsatellite markers for *M. isos*, and 2) to conduct cross-amplification analysis in two congeners, *M. endeavourae* and *M. gracilis*.

3.2 Material and methods

3.2.1 Sample collection and DNA extraction

Specimens were obtained from the National Institute of Water and Atmospheric Research (NIWA) Invertebrate Collection (Wellington, New Zealand) and Museum Victoria (Melbourne, Australia) from archived specimens collected by mainly scientific sampling expeditions in the southwest Pacific Ocean since the 2000s, considering that old museum specimens might be problematic due to DNA degradation.

Two specimens of *M. isos*, collected from the Chatham Rise, were used for the library development process. Seventy-three specimens of *M. isos* collected from four localities in the Tasmanian slope, were utilised for the subsequent characterisation of the microsatellite loci. Ten specimens of *M. endeavourae* collected from the Kermadec Ridge, and ten specimens of *M. gracilis* collected from the Challenger Plateau were used for the cross-amplification test (Table 3.1). Abdominal muscle or pereopod tissue samples were collected and stored individually in 95% ethanol before being transferred to the laboratory, for further analysis. Whole genomic DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The integrity and molecular weight of the DNA samples were assessed in a 3% agarose gel and with a NanoPhotometer (Implen, Munich, Germany).

Squat lobster samples were previously identified using the methods and keys described in Ahyong and Poore (2004) and Baba (2005). Species identification was further confirmed by genetic screening with GenBank accession numbers MF457406 (*M. isos*) and KJ544249 (*M. gracilis*) using the mitochondrial universal DNA barcoding primers LCO1490/HCO2198 (Folmer *et al.* 1994).

Table 3.1 Summary of collection sites from the southwest Pacific Ocean from which the three *Munida* species, were sampled for microsatellite development and characterisation and cross-amplification analysis.

Species	Geomorphic features	Collection Registration Number	Collection Station Number	Sample Date	Latitude (S)	Longitude (positive = E, negative = W)	Depth (m)	Samples (*N*)
Munida isos	Chatham Rise	NIWA 53802	TAN0905/104	26/06/2009	-44.158	-174.562	681–913	1
	Chatham Rise	NIWA 53915	TAN0905/107	27/06/2009	-44.177	-174.559	760–960	1
	Tasmanian slope	NMV J68478	SS01/2008 25	17/01/2008	-44.277	147.266	900	12
	Tasmanian slope	NMV J68560	SS01/2008 11	14/01/2008	-44.280	147.138	1260	13
	Tasmanian slope	NMV J67888	SS02/2007 14	01/04/2007	-44.321	147.178	1150–1280	16
	Tasmanian slope	NMV J67902	SS02/2007 03	29/03/2007	-43.692	146.976	100–110	32
M. endeavourae	Kermadec Ridge	NIWA 8993	TAN0413/20	08/11/2004	-36.797	177.469	1020–1183	1
	Kermadec Ridge	NIWA 8994	TAN0413/21	08/11/2004	-36.828	177.448	1118–1400	1
	Kermadec Ridge	NIWA 9792	TAN0413/54	11/11/2004	-37.237	177.277	1126–1134	1
	Kermadec Ridge	NIWA 18583	TAN0205/107	29/04/2002	-35.000	178.957	1255–1400	1
	Kermadec Ridge	NIWA 24573	TAN1206/97	24/04/2012	-36.455	177.838	920–950	1
	Kermadec Ridge	NIWA 24575	TAN1206/164	30/04/2012	-37.181	176.983	998–1000	1
	Kermadec Ridge	NIWA 64715	TAN1007/101	06/06/2010	-35.427	178.655	1028–1082	1
	Kermadec Ridge	NIWA 64817	TAN1007/106	06/06/2010	-35.350	178.514	1382–1416	1
	Kermadec Ridge	NIWA 72079	TAN1104/9	03/03/2011	-36.504	177.877	1576–1583	2
M. gracilis	Challenger Plateau	NIWA 33712	TAN0707/30	29/05/2007	-38.623	168.947	479–482	10

3.2.2 DNA sequencing and microsatellite isolation

Two paired-end DNA libraries were constructed from the two DNA samples of *M. isos*, each sequenced in half a lane using an Illumina HiSeq 2500 Genome Analyser (Sangon Biotech Co. Ltd., Shanghai, China). Raw sequences were trimmed to remove adapter sequences (Read1 3': AGATCGGAAGAGCACACGTCTGAAC; Read2 3': AGATCGGAAGAGCGTCGTGTAGGGA) using Cutadapt (https://pypi.python.org/pypi/cutadapt/1.2.1). Reads containing the ambiguous base N and/or with a quality score < 20 were removed using Prinseq (http://prinseq.sourceforge.net/). Clean reads were subject to *de novo* assembly into contigs using SOAPDenovo (http://soap.genomics.org.cn/soapdenovo.html). Contigs were then used as input into MISA (http://pgrc.ipk-gatersleben.de/misa/), specifying a search for di-, tri-, tetra-, and pentanucleotide motifs with a minimum of 6, 5, 5, 5 repeats, respectively. MISA was also employed to design primers with sufficient flanking regions on the resultant contigs, those with PCR product lengths ≥ 80 bp were kept.

3.2.3 PCR amplification and microsatellite genotyping

The experimental screening was conducted with the optimisation of PCR conditions against a subset of *M. isos* ($N = 10$) from New Zealand's EEZ. Primers successfully optimised were subsequently scored for amplification success in a sample of 73 geographically separated individuals of *M. isos* from the Tasmanian slope. Reactions were performed in a 15 μL total volume containing approximately 1.5 μL of genomic DNA (~30 ng/μL), 7.5 μL MyTaq Red Mix (Total Lab Systems Ltd, Auckland, New Zealand), 0.6 μL of each primer (forward and reverse) at concentration of 10 mmol/L, and with the addition of 4.8 μL ddH$_2$O. Single locus polymerase chain reaction (PCR) amplifications for microsatellite genotyping were carried out in a 96-well PCR Thermal Cycler (Applied Biosystems) programmed as follows: denaturation at 94 °C for 4 min, followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at primer-specific annealing temperature for 30 s, extension at 72 °C for 45 s, and a final extension at 72 °C for 10 min. The optimal annealing temperature of primer pairs was decided after testing with gradient PCR. The amplification products were then visualised on a 3% agarose gel stained with RedSafe Nucleic Acid Staining Solution (Ngaio Diagnostics Limited, Nelson, New Zealand) at 80 W for 1 h. Primers that showed unique, clear and reproducible fragments and that successfully amplified across all specimens were retained for further analysis. The

remaining forward primers from each pair were fluorescently end-labelled directly with 6-FAM, HEX or TAMRA and tested for robust amplification and level of polymorphism in all 73 individuals to assess their applicability to population genetics. PCR products were resolved using capillary electrophoresis on an ABI 3730xl DNA Analyzer (Applied Biosystems) with GeneScan 500 LIZ (Applied Biosystems) as the internal size standard. Genotyping was performed using PeakScanner v2.0 (Applied Biosystems).

3.2.4 Genetic analysis

The genotype data were initially assessed using Lositan (Antao *et al.* 2008) to identify outlier loci that had high or low F_{ST} values compared to neutral expectations, considering the relationship between F_{ST} and H_E (expected heterozygosity) in an island model under the infinite alleles model (IAM) with 50,000 simulations. In addition, Microchecker v2.2.3 (van Oosterhout *et al.* 2004) was employed to detect the presence of null alleles, large allele dropout and scoring errors by stuttering, as well as calculating null allele frequency (r) at 95% confidence intervals by performing 10,000 randomisations using the Brookfield (1996) method. Genetic diversity statistics including the number of alleles (N_A), effective number of alleles (N_E), expected (H_E) and observed heterozygosity (H_O), Polymorphism Information Content (*PIC*) values and inbreeding coefficient (F_{IS}) were calculated for each locus using GenAlEx v6.5 (Peakall and Smouse 2012) and Cervus v3.03 (Kalinowski *et al.* 2007). Exact tests for deviation from Hardy–Weinberg equilibrium (HWE) (Guo and Thompson 1992) and tests for genotypic linkage disequilibrium between pairs of loci were generated with Arlequin v3.5 (Excoffier and Lischer 2010) and Genepop v4.2 (Raymond and Rousset 1995; Rousset 2008). The False Discovery Rate (FDR) correction (Benjamini and Hochberg 1995) was applied when multiple pair-wise tests were performed to assess statistical significance ($p < 0.05$).

3.2.5 Cross-species amplification

The utility of a final set of loci was assessed by means of cross-amplification across two congeneric species, *M. endeavourae* ($N = 10$) and *M. gracilis* ($N = 10$), using the same PCR amplification protocol as described above.

3.3 Results

3.3.1 DNA sequencing and microsatellite isolation

The extractions yielded approximately 100 ng/µL of high-quality genomic DNA (A260/280 > 1.8) with unique and clear lane signals, as shown on an agarose gel (Figure 3.1). Two genomic libraries generated by the HiSeq 2500 sequencing platform produced a total of 156.99 M and 198.40 M raw reads, respectively. Clean reads were obtained after trimming to remove adapter-related, low quality and duplicate reads. Subsequently, *de novo* assembly produced 11.42 M and 12.84 M contigs from the resultant clean reads. The Q30 percentage exceeded 80% with GC contents ranged from 47.24% to 48.49%.
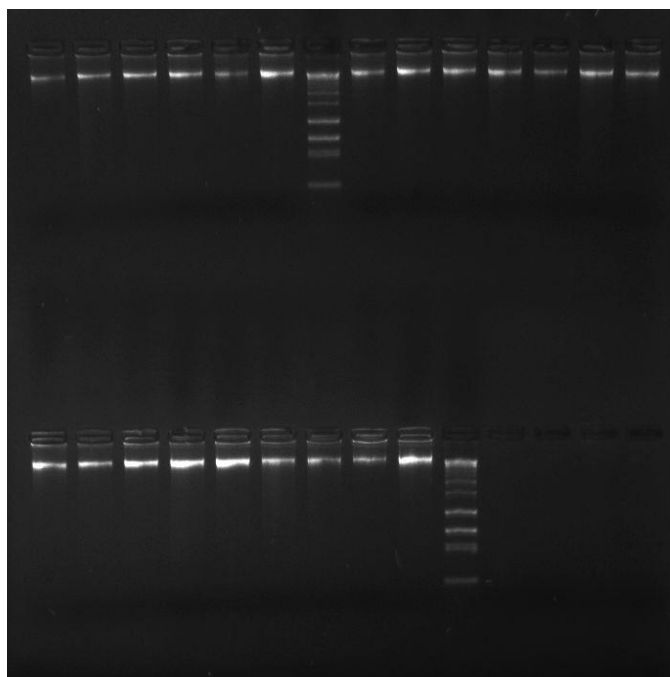
Figure 3.1 Agarose gel electrophoresis of the genomic DNA of *M*. *isos*.

A total of 1.32 M and 1.43 M sequences containing putative microsatellite loci were screened from the two libraries, respectively. Amongst these, dinucleotide repeat motifs (~71%) were the most common, followed by tri- (~18%), tetra- (~9%) and pentanucleotides (~2%). 38, 972 and 56, 165 putative microsatellite sequences were identified after excluding compound and interrupted sequences from the two library datasets, respectively. Sixty-four primer pairs containing di-, tri- and tetranucleotide motifs in length from 18 to 27 bp were selected and tested in ten individuals of *M. isos*. Pentanucleotide motifs were not polymorphic, thus none were developed as microsatellites.

Twenty-six microsatellite loci displayed seemingly successful amplification with targeted repeats on an agarose gel (clear bands, correct placement, and no non-specific amplification) for the initial evaluation. The remaining loci either failed to generate the desired amplification products, and despite retests under modified PCR conditions, or showed weak or inconsistent amplification. Further polymorphism screening on 73 individuals revealed that six of the 26 loci were monomorphic, the remaining 20 loci were polymorphic with optimal annealing temperature from 50 °C to 58 °C. A homology search using the BLAST program showed that none of the 26 sequences was similar to any GenBank sequences (Benson *et al.* 2018). The sequences of the developed microsatellite loci were then deposited in the National Centre for Biotechnology Information (NCBI) Sequence Read Archive under accession numbers MH649295 to MH649320. Descriptive information for the 20 polymorphic loci is reported in Table 3.2.

3.3.2 Characterisation of microsatellite loci

No locus was identified as experiencing positive selection according to Lositan. The number of alleles varied between 3 and 16, with an average of 7.45. The observed and expected heterozygosities ranged from 0.016 to 0.969 and 0.016 to 0.887, with an average of 0.505 and 0.520, respectively. Locus MI_21 had the lowest values for both H_E (0.016) and *PIC* (0.027), whereas the highest values of H_E (0.887) and *PIC* (0.871) were observed for locus MI_07. The detailed characteristics of all selected microsatellite loci are summarised in Table 3.2.

Table 3.2 Characterisation of twenty microsatellite loci from 73 specimens of *M. isos*.

Locus	GenBank Accession No.	Primer sequences (5'–3')	5' Dye	Repeat motif	Size range (bp)	N_A / N_E	H_O / H_E	PIC	F_{IS}	r
MI_01	MH649295	F: AAAACACACACAACAATCACAGAA	FAM	$(CA)_6$	94–112	5/3.372	0.469/0.722	0.65	0.330*	0.141
		R: AAAATGACGCTCATGATCCC								
MI_02	MH649296	F: ACGTATTTGTGCGAGATGCT	HEX	$(TG)_6$	89–121	12/4.095	0.616/0.777	0.753	0.181*	0.092
		R: TCAAACGCACACACAGATGA								
MI_04	MH649297	F: ATACACGATCACCTGCACGA	FAM	$(AC)_8$	132–156	10/4.494	0.725/0.801	0.764	0.062	0.037
		R: TTCATGTGAATGCGTGTGTG								
MI_05	MH649298	F: ATGGCTGTCTGACGCTACCT	HEX	$(CT)_9$	88–132	15/5.46	0.729/0.836	0.810	0.097*	0.058
		R: TTGTAGTGGACTTTTCCGGC								
MI_07	MH649300	F: CCACCACGGAAAGGTTCATA	TAMRA	$(CT)_8$	89–121	16/7.086	0.75/0.887	0.871	0.127*	0.068
		R: GAGATACCGTAGTGGACTGCC								
MI_09	MH649301	F: CGGCACTGAATGCAGTACTTT	TAMRA	$(AC)_9$	97–105	5/1.946	0.519/0.492	0.440	−0.100	−0.016
		R: ATTTAACCCGGGGACGAATC								
MI_10	MH649302	F: GACGTATGTATACGCCACAGAGA	FAM	$(TA)_7$	90–104	6/2.523	0.623/0.608	0.572	−0.073*	−0.018
		R: CGTGTGTAGGCTGTCCCTCT								
MI_14	MH649303	F: GTCAAGGCATTGCAAGTCAA	FAM	$(CA)_9$	102–116	8/2.351	0.579/0.59	0.525	−0.006*	−0.011
		R: TGGTGGTGGTGTGTCTGC								
MI_17	MH649304	F: TCCTTGAGGCTCACTTGGAT	FAM	$(TG)_6$	182–194	4/2.039	0.969/0.526	0.394	−0.904*	−0.296
		R: GGCAGGATCAGATTGGACAT								
MI_19–2	MH649305	F: TAAGTGGGTAAATAAGGACG	HEX	$(ATG)_6$	81–108	10/2.604	0.652/0.635	0.574	−0.067*	0.011
		R: CCATCTACAGAAATATATACCATC								

MI_20	MH649306	F: CAATAATGGTGTGTGAGTTGTGA	FAM	(TGG)$_7$	88–115	8/3.05	0.739/0.69	0.626	−0.104	−0.015
		R: TGATGAGGATGACGAGGTGA								
MI_21	MH649307	F: CAATGTTATTTCAAAGTCTTCTCCAA	HEX	(CAA)$_6$	89–110	3/1.016	0.016/0.016	0.027	−0.024	0.000
		R: TGCGTAAAGCAGGGTTATGA								
MI_23–2	MH649308	F: CTCCTTACATAATTTACTTCCTTCC	TAMRA	(ATT)$_7$	100–121	7/1.621	0.261/0.384	0.385	0.254*	0.093
		R: TTGATGGTTGCTAACAGACTACA								
MI_25	MH649309	F: AAGCAGGGAAGAGCACCAT	FAM	(AACA)$_5$	90–114	4/2.113	0.884/0.543	0.420	−0.684*	−0.248
		R: CTCTCTGTCTGCCCTCCCAG								
MI_29	MH649310	F: ATCACCTCCGTTTAGCGTGT	HEX	(GTCT)$_5$	90–106	5/1.189	0.103/0.157	0.174	0.205*	0.059
		R: TTTGGTAGCAAAAGGTTTAAAATG								
MI_31	MH649311	F: CCTTAAAATCGGGAAAACTCC	FAM	(GATA)$_5$	90–130	9/1.835	0.378/0.437	0.409	0.088	0.050
		R: TGTACGGTAGAGTCTGGAATAAAGG								
MI_39	MH649312	F: AAAATAATTCAGTATCCTGGAGCG	TAMRA	(GTA)$_7$	93–111	5/1.169	0.151/0.143	0.114	−0.074	−0.005
		R: ATAAAGGCAAACACCCCACA								
MI_40	MH649313	F: AACACTTTGTTCTTCGTGTGGA	FAM	(GAG)$_5$	101–107	3/1.086	0.078/0.073	0.053	−0.088	−0.001
		R: TCTCCTCCTCCTACTCCTTTCA								
MI_51	MH649317	F: ATGATCCCTTCGACCCTTG	HEX	(CAT)$_6$	87–108	8/1.771	0.353/0.422	0.445	0.111*	0.064
		R: TGGTTAGACTGAAGAGTTGGCA								
MI_52	MH649318	F: ATTACAGACTCGCCGCTCAC	TAMRA	(TTG)$_5$	110–125	6/2.896	0.515/0.669	0.638	0.205*	0.095
		R: TTCTCACCATCACCACTATTTCA								

N_A number of alleles; N_E number of effective alleles; H_O observed heterozygosity; H_E expected heterozygosity; *PIC* polymorphism information content; F_{IS} inbreeding coefficient; *r* null allele frequency. * Significant departure from Hardy–Weinberg equilibrium, after FDR correction ($p < 0.05$).

Analysis with Microchecker revealed that eight of the 20 microsatellite loci showed evidence of null alleles, with four loci (MI_01, MI_23-2, MI_29 and MI_31) exhibiting stuttering but no large allele drop out. Ten microsatellite loci showed evidence of heterozygote deficiencies as revealed by positive inbreeding coefficients ($F_{IS} > 0$). Amongst which, seven loci had significant homozygote excesses after FDR correction ($p < 0.05$) and were flagged by Microchecker for the presence of null alleles. The maximum null allele frequency observed in this study was relatively low ($r = 0.141$) according to Dakin and Avise (2004) who suggested that null frequencies below 0.2 are acceptable in most microsatellite datasets. All loci were retained for subsequent analysis.

HWE test showed significant deviation at 13 loci after FDR correction ($p < 0.05$). Eight (61.5%) and five (38.5%) of the deviations occurred in the form of deficits and excesses of heterozygotes, respectively. Amongst the 190 locus-pairs tested, only one pair (0.5%) of loci (MI_23-2 and MI_29) showed significant linkage disequilibrium after applying the FDR correction ($p < 0.05$).

3.3.3 Cross-species amplification

To facilitate genetic studies, the cross-amplification of the 20 microsatellite loci from *M. isos* in two congeners, *M. endeavourae* ($N = 10$) and *M. gracilis* ($N = 10$), revealed high amplification success (70%) and stable polymorphism (60%). Amongst the 20 loci, 14 produced expected amplicons, whilst the other six were non-specific. Eight of these 14 loci reliably amplified across both species and could be readily scored. In addition, eleven (55%) microsatellite loci were successfully amplified in *M. endeavourae* and *M. gracilis*, amongst which, two (MI_10 and MI_17) and three loci (MI_10, MI_14 and MI_51) were monomorphic, respectively. Fewer alleles in total were observed compared with the target species *M. isos* (N_A = 149), with only 37 in *M. endeavourae* and 28 in *M. gracilis*. Specifically, loci MI_51 and MI_17 generated the largest number of alleles in *M. endeavourae* (N_A = 6) and *M. gracilis* (N_A = 5), respectively. The allelic size (length in bp) varied between the two species for a given locus, but several alleles were shared by the three species. Descriptive information for Cross-amplification is reported in Table 3.3.

Table 3.3 Cross-amplification of fourteen microsatellite loci in 10 specimens of *M. endeavourae* and 10 specimens of *M. gracilis*.

Locus	*M. endeavourae*			*M. gracilis*		
	Successful amplifications (N)	N_A	Size range (bp)	Successful amplifications (N)	N_A	Size range (bp)
MI_01	10	4	92–98	10	3	90–96
MI_02	10	5	99–113	10	3	87–101
MI_04	—	—	—	10	3	100–106
MI_09	10	3	98–102	—	—	—
MI_10	10	1	93	10	1	93
MI_14	—	—	—	10	1	119
MI_17	10	1	192	10	5	188–200
MI_20	10	3	87–93	10	2	84–87
MI_23–2	10	4	127–151	—	—	—
MI_25	—	—	—	10	4	95–115
MI_29	10	4	122–146	10	3	100–108
MI_39	10	2	99–105	—	—	—
MI_40	10	4	98–107	10	2	101–104
MI_51	10	6	90–108	10	1	99

N_A number of alleles.

3.4 Discussion

3.4.1 Microsatellite isolation and characterisation

Twenty of the 26 microsatellite loci were polymorphic when screened on 73 specimens of *M. isos*, which were collected from four localities in the Tasmanian slope region. However, it is presumed that evaluation in populations from a larger spatial scale across the species' distributional range may reveal additional polymorphic loci. Evaluation of genetic diversity of the 20 microsatellite loci has revealed moderate to high levels of diversity, in general agreement with previous studies based on other HTS methods (Roche 454, Illumina MiSeq) to identify microsatellite loci of other squat lobsters (Molecular Ecology Resources Primer Development Consortium *et al.* 2010; Boyle *et al.* 2013; Coykendall and Morrison 2013; Roterman *et al.* 2013b; Nakajima *et al.* 2018). *PIC* is the most widely applied estimate to help measure the information content of molecular markers. In this study, 16 of 20 microsatellite loci showed moderate to high levels of polymorphism ($PIC > 0.25$) according to the criteria of Botstein *et al.* (1980). Notably, those represented a relatively lower level of polymorphisms were tri- and

tetranucleotide-repeat loci rather than dinucleotide repeat loci, which may be attributed to different mutation rates and/or repeat numbers.

Microsatellite null alleles are very common in marine invertebrates (Weetman *et al.* 2006; Rippe *et al.* 2017; Santos *et al.* 2018) and crustaceans, in particular, exhibit relatively high null allele frequencies (Crivello *et al.* 2005; Boyle *et al.* 2013; Nakajima *et al.* 2018; Santos *et al.* 2018). Dinucleotide-repeat microsatellites tend to produce stutter bands or truncated products, caused by the mispairing of the template and newly synthesised strand during the elongation, which may affect the efficiency and accuracy of allele assignment (Hearne *et al.* 1992). However, in this case, the four loci that exhibited stuttering represent di- (MI_01), tri (MI_23-2) and tetranucleotide (MI_29 and MI_31) repeat motifs.

Deviation from HWE with heterozygote deficiencies can be explained by a variety of different causes, including self-fertilisation, positive assortative mating, null alleles which might have led to the underestimation of heterozygotes even in relatively low frequencies, and the Wahlund effect (Johnson and Black 1984) that is caused by a mixture of individuals from different localities resulting in excess homozygosity and significant F_{IS} estimates (Freeland 2005; Waples 2014). On the other hand, factors that commonly produce heterozygote excesses include population history, selection and sex linkage (Waples 2014). In our study, 13 loci showed significant deviation from HWE after correction for multiple tests ($p < 0.05$), with eight and five of the deviations occurred in the form of deficits and excesses of heterozygotes, respectively. However, all of these described causes of heterozygote deficit/excess should affect all loci, rather than just one or a few (Selkoe and Toonen 2006; Waples 2014). Significant deviations from HWE ($p < 0.05$) have been observed in other squat lobster species (Molecular Ecology Resources Primer Development Consortium *et al.* 2010; Boyle *et al.* 2013; Nakajima *et al.* 2018).

3.4.2 Cross-species amplification

The high cross-amplification percentages reflect the close relatedness of these two species to the target species. Similar high transferability (cross-species amplification rates) was also spotted between other squat lobsters in previous studies (Boyle *et al.* 2013; Wang and Held 2015). Six loci could be amplified in only one species, indicating a differential mutation in the priming (flanking) regions of the microsatellite loci in the two *Munida* species. It is noteworthy that loci exhibiting monomorphism might show polymorphism if larger sample sizes across each species' distributional range were to be included. In addition, fewer alleles were observed in the two congeners compared with the target species *M. isos*. This discrepancy may be due to the small sample sizes highlighting the fact that larger sample sizes are beneficial for obtaining more genetic variety.

3.5 Conclusion

Newly developed microsatellite loci will permit detailed genetic estimates amongst squat lobster populations of three *Munida* species across their distributional ranges in the southwest Pacific Ocean. The availability of new microsatellite loci, in conjunction with other types of molecular markers (e.g., COI, SNPs), will facilitate further evaluation of genetic diversity and connectivity of these invertebrates amongst sites, and offer valuable information to help formulate management and conservation strategies of VMEs within and beyond New Zealand's EEZ.

Chapter 4. Patterns of population structure and genetic connectivity of squat lobsters (*Munida* Leach, 1820) in the southwest Pacific Ocean based on mitochondrial DNA and microsatellite variation

Abstract

Studies of genetic diversity and population genetic structure in deep-sea fauna mainly focus on vulnerable marine ecosystem (VME) indicator taxa, whilst relatively few studies have focussed on VME-associated taxa whose distribution is not exclusively limited to VMEs. Knowledge of the connectivity amongst VME-associated taxa, such as squat lobsters, will contribute to ongoing management decision making related to the protection of VMEs. To better understand the genetic diversity and population genetic structure of squat lobsters (*Munida isos*, *M. endeavourae* and *M. gracilis*) at different spatial scales (Northern-Southern provinces, North-Central-South regions and geomorphic features) within the New Zealand's Exclusive Economic Zone (EEZ) and beyond, mitochondrial COI region and nuclear microsatellite markers were employed. Overall, the level of COI genetic diversity was high and was moderate based on microsatellite markers across all three *Munida* species. With more than 96% of the variance being attributed to differences within populations in all three *Munida* species based on both marker types, no genetic subdivision was detected in *M. endeavourae*, whilst little genetic differentiation was observed in *M. isos* or *M. gracilis* based on microsatellite variation. A low to moderate level of assignment success was achieved for all three *Munida* species, suggesting high levels of gene flow and possible panmixia. Nonetheless, for *Munida isos*, populations from the Tasmanian Slope were potentially genetically different from all other populations and may act as source populations, whereas populations from the Kermadec Ridge region may be sink populations for all three *Munida* species. Moreover, robust evidence of recent demographic expansion was detected for all three *Munida* species, based on both COI and microsatellite variation. These population genetic studies for the three *Munida* species have revealed patterns of genetic diversity and connectivity of squat lobster populations across their distributional range, with results discussed in the context of existing marine protected areas, which can contribute new information useful to the management of VMEs within the southeastern Australia and New Zealand regions.

4.1 Introduction

Vulnerable marine ecosystems (VMEs) in the deep sea, such as cold water coral reefs and seamounts, are vulnerable to the impact of intense or long-term anthropogenic activities (e.g., bottom trawling) (Food and Agricultural Organisation of the United Nations 2009). Marine protected areas (MPAs) provide a means to restrict damaging human activities and thereby allow for the conservation of critical natural habitats such as VMEs, which can help facilitate the maintenance and restoration of biodiversity and ecosystem function (Agardy 1994; Clark *et al.* 2011). Given that population declines can lead to reduced individual fitness and population resilience, and lower mean genetic diversity, an understanding of the population connectivity of VME indicator taxa is useful for informing the design and implementation of MPAs, and assessing the effectiveness of existing MPAs, and thereby improve the management of impacts on VMEs (Palumbi 2003; Miller and Gunasekera 2017).

The development of conservation strategies using networks of MPAs based on population connectivity requires genetic information from rapidly evolving and variable marker types, i.e., those that can indicate contemporary gene flow (Frankham *et al.* 2002; Boschen *et al.* 2016). Over the last few decades, microsatellites have emerged as one of the most powerful molecular markers in population genetic studies (Wright and Bentzen 1995; Selkoe and Toonen 2006), and the combination of mitochondrial and nuclear DNA markers has significantly improved the power of molecular data in conservation genetic, phylogenetic and phylogeographic studies (Toews and Brelsford 2012).

Whilst there are numerous studies examining the population structure of VME indicator taxa from the southwest Pacific Ocean (e.g., Miller *et al.* (2010); Miller *et al.* (2011); Herrera *et al.* (2012); Miller and Gunasekera (2017); Zeng *et al.* (2017); Holland *et al.* (2018); Zeng *et al.* (2019)), relatively few studies have focussed on VME-associated taxa) whose distribution is not exclusively limited to VMEs but are often found together with VMEs (e.g., crinoids, brisingid seastars *sensu* Parker *et al.* (2009). Squat lobsters are highly diverse anomuran groups that are sometimes closely associated with VMEs (e.g., coral thickets and sponge gardens) in the world's oceans (Lumsden *et al.* 2007; Baba *et al.* 2008; Baeza 2011). Recent studies have shown the effectiveness of molecular markers (e.g., mitochondrial COI and microsatellite variation) in population genetic studies of squat lobsters across a number of different families.

For example, cytochrome *c* oxidase subunit I (COI) sequence data revealed no genetic structure amongst populations of two *Munida* species and two *Eumunida* species sampled from

the southwest Pacific Ocean (Samadi *et al.* 2006), and *Pleuroncodes monodon* (H. Milne Edwards, 1837) collected from the southeast Pacific Ocean coast of South America with signal showing a recent demographic expansion (Haye *et al.* 2010). A similar result was observed for *Munida gracilis* Henderson, 1885, where no significant population genetic differentiation was identified in populations of the New Zealand region, based on COI and 16S genetic variation (Bors *et al.* 2012). Additionally, no genetic subdivision was observed for *Kiwa tyleri* Thatje, 2015 across three sites on the East Scotia Ridge (Roterman *et al.* 2016), *Munidopsis polymorpha* Koelbel, 1892 from the east Atlantic region (Cabezas *et al.* 2012a), and *Munida gregaria* (Fabricius, 1793) sampled across its entire South American distribution (Wang *et al.* 2016), based on a combination of the COI region and microsatellite markers. The first species, *K. tyleri*, exhibits extremely abbreviated lecithotrophic larval development that can contribute to a limited dispersal range. Nonetheless, the authors argued that the larval lifespans might be extended by the low temperature-induced metabolic rate reduction in the Southern Ocean, mitigating the impact of larval dispersal strategy in shaping genetic structure. Moreover, evidence showed that *K. tyleri* underwent a demographic expansion ~500 kya (Roterman *et al.* 2016).

Munida gregaria was further examined across its range in the southern Pacific and Atlantic Oceans. Sequence data of mitochondrial COI and NADH dehydrogenase subunit 1 (ND1) genes confirmed that it is a trans-Pacific species, with some degree of genetic subdivision observed between populations of the southwest Pacific Ocean and those off the southern tip of South America, whilst populations of the southwest Atlantic Ocean and southeast Pacific Ocean exhibited no sign of genetic subdivision from the southwest Pacific Ocean. It was argued that the Antarctic Circumpolar Current might contribute in shaping the sub-Antarctic Marine invertebrates' distribution. Besides, evidence of a demographic expansion in southern South America during the Pleistocene was observed, in which the Pleistocene glaciations might play an important role (Pérez-Barros *et al.* 2014). Similarly, for *Shinkaia crosnieri* Baba & Williams, 1998 in the South China Sea, genetic data (COI and nuclear adenine nucleotide translocase (ANT)) suggested that individuals from vent and cold seep environments (also VMEs) represented separate populations but were conspecific, as revealed by relatively large COI sequence divergences but no significant genetic differentiation based on ANT between these two populations. However, it remained to be decided whether this genetic discontinuity was due to geographical separation and/or distinct environment factors (i.e., vent and seep) (Yang *et al.* 2016). Moreover, for a vent-associated squat lobster

species, *Munidopsis lauensis* Baba & de Saint Laurent, 1992, although genetic homogeneity was observed amongst populations sampled from the western Pacific Ocean as suggested by COI sequence data, pronounced genetic differentiation was revealed on the basis of microsatellite markers, which might be attributed to the hydrographic and/or physical barriers (Thaler *et al.* 2014). Taken together, with results spanning from panmixia to pronounced genetic differentiation at different geographic scales, no consistent patterns of connectivity have emerged in shallow water or deep-sea squat lobsters.

Whilst no consistent pattern of population connectivity among squat lobster species has emerged in the southwest Pacific Ocean, detailed assessments of genetic diversity and population genetic structure in squat lobster taxa, based on multiple genetic marker types, are still scarce. As VME-associated taxa, knowledge of connectivity amongst deep-sea squat lobster populations will undoubtedly contribute to ongoing management decision-making about the protection in the region, by assessing the effectiveness of existing MPAs and informing the placement of new MPAs. In light of this, the species-specific patterns of genetic connectivity amongst populations of three locally abundant *Munida* species (*M. isos* Ahyong & Poore, 2004, *M. endeavourae* Ahyong & Poore, 2004 or *M. gracilis* Henderson, 1885) from the southwest Pacific Ocean region are investigated in this study. *Munida* species have free-swimming planktonic larval phases that may drift with ocean currents, but unfortunately, the larval dispersal strategy is not known in detail for these three *Munida* species. However, it is presumed that they have a similar development as other munidids which includes between 4–6 planktotrophic zoeal larval stages and 15–83 days of larval development (under laboratory rearing conditions) (Baba *et al.* 2011). Based on what is known about the dispersal strategies of the three species and their habitat preference, it is expected that: *M. isos* and *M. endeavourae*, which are typically found on physically isolated seamounts, will exhibit population genetic structure within the study area; whereas connectivity amongst populations of *M. gracilis,* which is also found of soft sediments that are generally contiguous over large areas of the deep sea, which will allow greater physical movement of adults and larvae, will exhibit higher levels of population genetic homogeneity. If indeed the connectivity of populations inhabiting seamounts differs from those of contiguous non-seamount habitats, this would have implications for spatial management measures of species across their respective ranges.

The objective of this chapter is: 1) to assess the genetic diversity of the three VMEs-associated squat lobster taxa (*M. isos*, *M. endeavourae* and *M. gracilis*) across their distributional ranges, 2) to evaluate the population structure and genetic connectivity amongst populations of the three *Munida* species in the southwest Pacific Ocean following a spatially explicit hierarchical testing framework, 3) to explore the drivers of the observed connectivity patterns, and 4) to provide information that will contribute to the management of VMEs within the southwest Pacific Ocean.

4.2 Methods

4.2.1 Sample collection and DNA extraction

Specimens were obtained from the National Institute of Water and Atmospheric Research (NIWA) Invertebrate Collection (Wellington, New Zealand) and Museum Victoria (Melbourne, Australia) from archived specimens collected by mainly scientific sampling expeditions in the southwest Pacific Ocean since the 2000s, considering that old museum specimens might be problematic due to DNA degradation. Most specimens were collected within the New Zealand's Exclusive Economic Zone (EEZ), with additional samples from the south Tasman Sea within the Australian EEZ, and from the Louisville Seamount Chain to the northeast of the New Zealand's EEZ (Figure 4.1). The majority of specimens of *M. isos* and *M. endeavourae* were from seamount habitats, whilst the majority of specimens of *M. gracilis* were from soft sediment habitats at relatively shallow depth. All specimens were preserved in > 80% ethanol after collection. Both the mitochondrial DNA and nuclear DNA extraction, and quantification procedures, are the same as described in Chapter 3. Samples exhibiting poor quality DNA were discarded.

The maps were generated using ArcGIS 10.4.1 software (Mercator projection and the World Geodetic System (WGS 1984) as the reference coordinate system). The base map bathymetry data is derived from the GEBCO_2008 Grid, version 20100927, www.gebco.net; the marine boundaries for the New Zealand region are based on NIWA data; those for the Australian region are credited to the Australian Government Department of the Environment, 2014 and the Australia, Commonwealth Heritage List (Spatial Database is copyright (c) Commonwealth of Australia, Heritage and Wildlife Division of the Australian Government Department of the Environment, 2013).
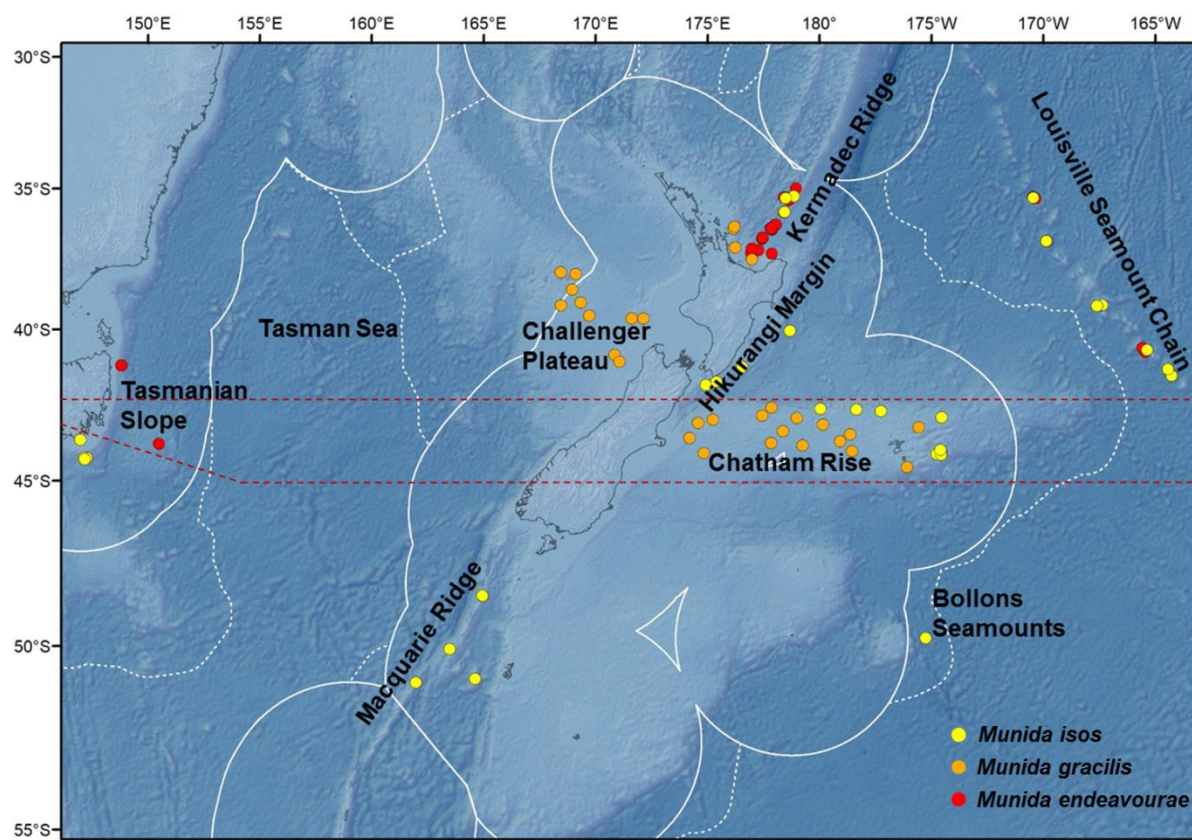
Figure 4.1 Map showing the locations of the samples for the study species.

For *Munida isos*, hierarchical testing framework was employed: (1) red dashed line at approximately 45°S indicates the boundary between Northern province (< 45°S) and Southern province (> 45°S), representing the two biogeographic lower bathyal provinces (BY6 and BY10) of Watling *et al.* (2013) in the New Zealand region; (2) red dashed lines at 42°S and approximately 45°S indicate the boundaries between the North (< 42°S), Central (42°S to 45°S) and south (> 45°S) regions; (3) major geomorphic features are named on the map.

4.2.2 Data testing framework

For *M. isos*, for which there were large samples sizes and good spatial coverage, a spatially explicit hierarchical testing framework was employed to evaluate genetic differentiation amongst populations. This approach follows Zeng *et al.* (2017); Zeng *et al.* (2019) and Holland *et al.* (2018).

First, samples of *M. isos* were assigned to the two deep, bathyal (800–3500 m), ocean floor biogeographic provinces that are currently recognised within the New Zealand region by Watling *et al.* (2013); BY6 or a Northern province, and BY10 or a Southern province, with a boundary at approximately 45°S. The differential water mass characteristics (e.g., salinity, temperature and particulate organic carbon flux) of the two provinces may affect the Northern-Southern population distribution, resulting in a province-scale pattern of genetic differentiation.

Second, samples were assigned to three regions; North (north of Chatham Rise), Central (Chatham Rise), and South (south of Chatham Rise) regions. Subtropical Water (STW) that reaches New Zealand region via the East Australian Current and the South Pacific subtropical gyre, and Subantarctic Water (SAW) that is driven north by Ekman transport and westerly winds, meet along the Chatham Rise forming the Subtropical Front (Chiswell *et al.* 2015). This front may act as a barrier to pelagic larval dispersal, leading to a North-Central-South regional scale pattern of genetic subdivision of populations, and a high level of genetic diversity may be expected in the Central region on the Chatham Rise where North-South population mixing may occur.

Third, samples were assigned to populations according to major geomorphic features of the seafloor: Kermadec Ridge, Louisville Seamount Chain, Hikurangi Margin, Chatham Rise, Macquarie Ridge, and Tasmanian slope. Given the species-specific distribution and habitat requirement, the distinctive topographic and local hydrodynamic characteristics may act as barriers to gene flow. Samples from Bollons Seamount were included in the province and regional spatial scale analyses, but not in geomorphic feature scale analyses, due to the limited sample size.

For *M. endeavourae* and *M. gracilis*, due to patchy sampling sites, samples were only assigned to populations according to major geomorphic features for the assessment of genetic connectivity. For *M. endeavourae*, populations were included from the Kermadec Ridge, Louisville Seamount Chain, and the Tasmanian slope. For *M. gracilis*, populations included those from the Kermadec Ridge, Chatham Rise, and Challenger Plateau. Sample details are summarised in Table 4.1; full details are presented in Tables S4.1–S4.3.

Table 4.1 Summary of sample sizes (*N*, number of specimens) for *M. isos* at three spatial scales and for *M. endeavourae* and *M. gracilis* at the geomorphic feature scale.

	Spatial class	Code	COI	Microsatellite
Munida isos	Northern province	Northern	214	219
	Southern province	Southern	134	136
	Total (provinces)		**348**	**355**
	North region	North	106	109
	Central region	Central	108	110
	South region	South	134	136
	Total (regions)		**348**	**355**
	Kermadec Ridge	KR	13	13
	Louisville Seamount Chain	LS	82	85
	Hikurangi Margin	HM	11	11
	Chatham Rise	CR	108	110
	Macquarie Ridge	MR	59	60
	Tasmanian slope	TA	72	73
	Total (geomorphic features)		**345**	**352**
M. endeavourae	Kermadec Ridge	KR	38	42
	Louisville Seamount Chain	LS	28	28
	Tasmanian slope	TS	23	26
	Total (geomorphic features)		**89**	**96**
M. gracilis	Kermadec Ridge	KR	19	20
	Chatham Rise	CR	154	155
	Challenger Plateau	CP	159	164
	Total (geomorphic features)		**332**	**339**

4.2.3 COI sequencing and data analyses

The mitochondrial DNA protein encoding COI region was amplified for the three *Munida* species using universal invertebrate primers LCO1490 and HCO2198 (Folmer *et al.* 1994), yielding fragments of 649 bp in length. Reactions were performed in a 20 µL total volume containing approximately 1 µL of genomic DNA (~30 ng/µL), 10 µL MyTaq Red Mix (Total Lab Systems Ltd, Auckland, New Zealand), 1 µL of both primers (forward and reverse) at a concentration of 10 mmol/L, and with the addition of 7 µL ddH$_2$O. PCR amplifications for COI sequencing were carried out in a 96-well PCR Thermal Cycler (Applied Biosystems) programmed as follows: denaturation at 95 °C for 5 min, followed by 40 cycles of denaturation at 95 °C for 30 s, 48 °C for 30 s, extension at 72 °C for 45 s, and a final extension at 72 °C for 10 min. PCR products were visualised on a 1.5 % agarose gel and then sequenced on an ABI 3730xl DNA Analyzer. Despite multiple attempts, some COI fragments could not be amplified for several samples. Sequences have been deposited in GenBank (Acc# MK155662–MK156049).

COI sequences were aligned in MEGA v7 (Kumar *et al.* 2016) following manual checking and editing using DNAStar software (Burland 2000). Genetic diversity statistics including number of haplotypes (*h*), average number of nucleotide differences (*K*), number of segregating sites (*S*), haplotype diversity (*H$_d$*) and nucleotide diversity (*π*) were calculated within species using DNASP v6.11.01 (Rozas *et al.* 2017) and Arlequin v.3.5 (Excoffier and Lischer 2010).

To assess hierarchical population genetic differentiation amongst squat lobster populations, the Kimura 2-parameter (K2P) genetic distance values within/between regions were calculated in MEGA v7 (Kumar *et al.* 2016). Arlequin v.3.5 (Excoffier and Lischer 2010) was applied to perform a spatially explicit analysis of molecular variance (AMOVA) and to compute pairwise Φ_{ST} values. All groupings were evaluated with 10,000 permutations. Haplotype networks were constructed using a median-joining method in Popart (Leigh and Bryant 2015). For each species, a maximum likelihood (ML) tree was drawn under the best-fit evolutionary model with the lowest Bayesian information criterion (BIC) score and 500 bootstraps in MEGA v7 (Kumar *et al.* 2016). A congeneric species, *Munida thoe* Macpherson, 1994 (Acc# AY351013) and a yeti crab, *Kiwa tyleri* Thatje, 2015 (Acc# KP763652), were designated as outgroups. Nodes with bootstrap values > 90% were included on the final ML trees.

Demographic history was assessed with Fu's F_S (Fu 1997) tests and mismatch distributions evaluated with the sum of squared deviations (SSD) and Harpending's raggedness index (Hri) (Rogers and Harpending 1992; Harpending 1994) implemented in Arlequin v.3.5 (Excoffier and Lischer 2010). Possible expansions were calculated following the equation: t = $\tau/2\mu k$, where τ (Tau) is the estimated number of generations since the demographic expansions, μ is the mutation rate per nucleotide per generation, k is the sequence length (i.e., 649 bp). A COI substitution rate of 1.015% per lineage per million years was applied (Pérez-Barros *et al.* 2014). It was demonstrated that molecular population-level studies (1–2 mya) could exhibit ten-fold higher mutation rates than the long-term substitution rates suggested by phylogenetic studies (Ho *et al.* 2005; Ho *et al.* 2007; Ho *et al.* 2011). Thus, a ten-fold correction was applied.

All analyses described above were carried out at three spatial scales (Northern-Southern provinces, North-Central-South regions, and geomorphic features) for *M. isos*, and at the geomorphic level only for *M. endeavourae* and *M. gracilis*.

4.2.4 Microsatellite genotyping and data analyses

From the 20 microsatellite markers developed previously (see Chapter 3), 17, 11 and 9 microsatellite primer pairs were selected for *M. isos*, *M. endeavourae* and *M. gracilis*, respectively, based on their scoring reliability. Locus-specific PCR amplification conditions are described in Chapter 3. PCR products were run alongside a LIZ500 internal size standard on an ABI 3730xl DNA Analyzer (Sangon Inc., Shanghai, China). Alleles were scored automatically and reviewed manually using PeakScanner v2.0 (Applied Biosystems).

The genotype data were initially assessed in Lositan (Antao *et al.* 2008) for the neutrality test in an island model under the infinite alleles model (IAM) with 50,000 simulations. Microchecker v2.2.3 (van Oosterhout *et al.* 2004) was subsequently employed to detect the presence of null alleles, large allele dropout and scoring errors caused by stuttering, as well as to calculate null allele frequency (r) using the Brookfield (1996) method for each locus.

To quantify population genetic diversity, summary statistics including the actual number of alleles (N_A), effective number of alleles (N_E), observed (H_O) and expected (H_E) heterozygosity, polymorphism information content (PIC) values, allelic richness (A_R), number of private alleles (N_{AP}) and inbreeding coefficient (F_{IS}) were calculated for each locus using GenAlEx v6.5 (Peakall and Smouse 2012), Cervus v3.03 (Kalinowski *et al.* 2007) and Fstat

v2.9.3.2 (Goudet 2002). Exact tests of deviation from Hardy–Weinberg equilibrium (HWE) (Guo and Thompson 1992) for each locus × population combination and tests for genotypic linkage disequilibrium (LD) between pairs of loci were performed in Arlequin v3.5 (Excoffier and Lischer 2010) and Genepop v4.2 (Raymond and Rousset 1995; Rousset 2008). Significance levels ($p < 0.05$) for multiple pair-wise tests were adjusted using the False Discovery Rate (FDR) correction (Benjamini and Hochberg 1995).

To quantify population genetic differentiation, Nei's genetic distance (D_A) (Nei et $al.$ 1983) and pairwise F_{ST} values were computed using Populations v1.2.31 (available at http://bioinformatics.org/~tryphon/populations/) and Arlequin v.3.5 (Excoffier and Lischer 2010). The latter was used to perform spatially explicit Analysis of Molecular Variance (AMOVA), with all groupings tested with 10,000 permutations, for the described different spatial scales.

General relationships amongst individuals were depicted through a Principal Coordinate Analysis (PCoA) on the basis of a covariance pairwise genetic distance matrix in GenAlEx v6.5 (Peakall and Smouse 2012). Discriminant analysis of principal components (DAPC) (Jombart et $al.$ 2010) was applied to test for genetic structure amongst populations with the "adegenet" package v2.0.1" (Jombart and Ahmed 2011) in the program R v3.4.1 (R Development Core Team, 2017). Notably, for $M.$ $isos$, DAPC was conducted at regional and population scales but not at the province-level due to the limited number of groups. To determine the best K, denoting the most likely number of putative genetic clusters, a Bayesian clustering approach was implemented in Structure v2.3.4 (Pritchard et $al.$ 2000) using an admixture ancestry model with corrected initial ALPHA values (~$1/K$), uncorrelated allele frequencies and prior location information (Hubisz et $al.$ 2009; Wang 2017). Each Markov Chain Monte Carlo (MCMC) simulation ran for 500,000 iterations, with an initial burn-in period of 50,000 iterations. Ten independent runs were performed for each cluster set. The ΔK metric (Evanno et $al.$ 2005) was employed to calculate the best K (average number of nucleotide differences) in Structure Harvest (http://taylor0.biology.ucla.edu/structureHarvester/). Post-processing of Structure results was carried out using CLUMPAK (Kopelman et $al.$ 2015). Another clustering approach was carried out in BAPS v6 (Corander et $al.$ 2008) as a complement to the Structure analyses. Calculations were performed under the mixture model, with the maximum number of populations set as the number of regions, i.e., six, three and three for $M.$ $isos$, $M.$ $endeavourae$ and $M.$ $gracilis$, respectively.

Individual assignment tests and first generation migrant detection with "L_home" likelihood computation were implemented in GeneClass v2 (Piry *et al.* 2004). The contemporary migration rates amongst populations were calculated in BayesAss v3 (Wilson and Rannala 2003), which can produce accurate estimates of migration and self-replenishment when F_{ST} is larger than 0.05 (Faubet *et al.* 2007). Calculation parameters were set as follows: 10,000,000 iterations, of which 3,000,000 were burn-in; sampling frequency 1000; migration rates, allele frequencies and inbreeding coefficients set at 1, 1 and 1, respectively, to make the acceptance rates between 20 and 60%.

Effective population size (*N*e) was evaluated using NeEstimator v2 (Do *et al.* 2014). P_{crit} values were set according to the formula $P_{crit} > 1/2S$, where S denotes the number of samples (Waples and Do 2010). Recent (the past 2*N*e–4*N*e generations) bottleneck events or phases of expansion were tested in Bottleneck v1.2.02 (Piry *et al.* 1999), by assessing heterozygosity excess or deficiency. The two-phase model (TPM) incorporating 95% of the Stepwise Mutation Model (SMM) and 5% of the Infinite Allele Model (IAM), the Wilcoxon sign rank test, and the "mode-shift" indicator were applied.

All analyses described above were carried out at three spatial scales (Northern-Southern provinces, North-Central-South regions, and geomorphic features) for *M. isos*, and at geomorphic feature level for *M. endeavourae* and *M. gracilis*.

4.3 Results

4.3.1 Population genetic structure based on the COI region

4.3.1.1 Genetic diversity

Haplotypic diversity was very high across populations for all provinces, regions and geomorphic features of the three *Munida* species. For *M. isos*, 111 haplotypes, based on a 649-bp fragment of the gene, were obtained from 348 individuals at both the province and regional spatial scale. The haplotype diversity and nucleotide diversity values of populations were slightly greater in populations of the Northern province compared to the Southern province, and both estimates in the North region were higher than the other two regions (detailed results are reported in Table 4.2). At the geomorphic feature scale, a total of 110, 67 and 200 COI haplotypes were obtained from 345, 89 and 332 individuals of *M. isos*, *M. endeavourae* and *M. gracilis*, respectively. Notably, the mean population haplotype diversity and nucleotide

diversity values for *M. isos* (0.764; 0.26%) were lower than for *M. endeavourae* (0.980; 0.59%) and *M. gracilis* (0.973; 0.56%), respectively.

4.3.1.2 Genetic differentiation

As revealed by K2P genetic distance and Φ_{ST} values, genetic differentiation amongst populations of provinces, regions and geomorphic features was low in the three *Munida* species. Specifically, all of the K2P genetic distances were relatively low and sat within the Decapoda intraspecific barcode threshold (0.285–1.375%) (da Silva *et al.* 2011). Significance ($p < 0.05$) was detected for three Φ_{ST} values, all from population pairs at geographical feature scale of *M. isos*, yet none was observed amongst regions or province population pairs of *M. isos*, nor population pairs of geomorphic features for *M. endeavourae* or *M. gracilis* (Table 4.3). This lack of differentiation was supported by findings from AMOVA analyses, with no significant genetic differentiation observed for all *F*-statistics values and over 99% of the variance being attributed to differences within populations (Table 4.4).

The median-joining networks, representing the most parsimonious relationship amongst haplotypes, were illustrated for the three *Munida* species. For *M. isos*, a starburst-like pattern was observed from the 111 haplotypes, but no pattern of genetic structure was detected in populations at either the province (Figure 4.2) or regional spatial scale (Figure 4.3). At the geomorphic feature scale, haplotypes from both *M. isos* and *M. endeavourae* populations exhibited a starburst-like pattern, consistent with the mutational evolution of new haplotypes over time (Figure 4.4). However, the number of equally parsimonious haplotype connections was too great in *M. gracilis*, meaning that the most likely network itself could not be determined. Similarly, there was no evidence of a pattern of genetic structure amongst populations of the three *Munida* species across their distributional ranges (Figure 4.5). The maximum likelihood phylogenetic trees were similar for the three *Munida* species, with all haplotypes occurring in a major non-differentiated cluster with strong support, and clearly differentiated from the two outgroup taxa (Figure 4.6). These analyses indicate that there is no clear evidence of genetic subdivision amongst populations at the geomorphic feature, region, or province scale for the three *Munida* species.

Table 4.2 Population-specific genetic diversity statistics of three *Munida* species based on COI (unshaded) and microsatellite variation (shaded).

		COI							Microsatellite						
		N	h	H_d	π	K	S	Fs	N	N_A	N_E	H_O	H_E	F_{IS}	N_{AP}
Munida isos	Northern province	214	77	0.786	0.27%	1.765	66	**-141.278**	219	8.765	2.317	0.410	0.506	0.256	31
	Southern province	134	49	0.732	0.25%	1.589	44	**-75.803**	136	8.000	2.320	0.413	0.456	0.161	18
	Total/Mean (provinces)	348	111	0.765	0.26%	1.697	81	**-33.746**	355	8.382	2.319	0.412	0.481	0.209	25.000
	North region	106	43	0.807	0.27%	1.743	41	**-33.259**	109	7.353	2.297	0.406	0.501	0.232	10
	Central region	108	43	0.765	0.27%	1.781	40	**-59.638**	110	7.353	2.322	0.415	0.505	0.245	11
	South region	134	49	0.732	0.25%	1.589	44	**-75.803**	136	8.000	2.320	0.413	0.456	0.161	18
	Total/Mean (regions)	348	111	0.765	0.26%	1.697	81	**-33.746**	355	7.569	2.313	0.411	0.487	0.213	13.000
	KR	13	8	0.859	0.33%	2.115	11	**-3.433**	13	4.176	2.263	0.356	0.505	0.293	1
	LS	82	37	0.814	0.26%	1.713	37	**-50.447**	85	6.824	2.343	0.419	0.506	0.216	8
	HM	11	6	0.727	0.25%	1.618	7	**-2.132**	11	3.059	1.859	0.364	0.417	0.056	1
	CR	108	43	0.765	0.27%	1.781	40	**-59.638**	110	7.353	2.322	0.415	0.505	0.245	11
	MR	59	25	0.792	0.28%	1.814	24	**-24.736**	60	6.471	2.149	0.418	0.455	0.129	7
	TA	72	29	0.662	0.21%	1.363	33	**-37.596**	73	6.647	2.356	0.410	0.449	0.069	9
	Total/Mean (geomorphic features)	345	110	0.764	0.26%	1.700	81	**-234.280**	352	5.755	2.215	0.397	0.473	0.168	6.167
M. endeavourae	KR	38	35	0.994	0.65%	4.192	43	**-33.514**	42	4.636	1.860	0.355	0.384	0.073	14
	LS	28	25	0.984	0.63%	4.066	33	**-25.480**	28	3.727	1.737	0.360	0.366	0.004	6
	TA	23	18	0.957	0.45%	2.885	21	**-15.656**	26	3.727	1.840	0.416	0.397	-0.073	6
	Total/Mean (geomorphic features)	89	67	0.980	0.59%	3.805	64	**-34.056**	96	4.030	1.812	0.377	0.382	0.001	8.667
M. gracilis	KR	19	14	0.947	0.52%	3.345	22	**-8.237**	20	4.667	2.408	0.457	0.512	0.142	0
	CR	154	106	0.977	0.60%	3.860	77	**-196.496**	155	7.222	2.410	0.463	0.528	0.167	10
	CP	159	101	0.972	0.54%	3.510	75	**-184.150**	164	6.889	2.442	0.467	0.516	0.122	7
	Total/Mean (geomorphic features)	332	200	0.973	0.56%	3.657	107	**-32.498**	339	6.259	2.420	0.462	0.519	0.144	5.667

For COI, N = sample size; h = number of haplotypes; H_d = haplotype diversity; π = nucleotide diversity; K = average number of nucleotide differences; S = number of segregating sites; F_S = Fu's F_S test statistic ($p < 0.05$, in bold). For microsatellites, N_A = average number of alleles; N_E = average number of effective alleles; H_O = observed heterozygosity; H_E = expected heterozygosity; F_{IS} = inbreeding coefficient; N_{AP} = number of private alleles. Population codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.

Table 4.3 Genetic distance and F-statistics amongst populations of three *Munida* species based on COI (unshaded) and microsatellite variation (shaded).

COI							Microsatellite					
Munida isos	Northern province	Southern province						Northern province	Southern province			
Northern province	**0.003**	0.003					Northern province		0.026			
Southern province	0.000	**0.002**					Southern province	0.013*				
	North region	Central region	South region					North region	Central region	South region		
North region	**0.003**	0.003	0.003				North region		0.025	0.031		
Central region	0.003	**0.003**	0.003				Central region	0.010*		0.032		
South region	0.002	-0.001	**0.002**				South region	0.018*	0.014*			
	KR	LS	HM	CR	MR	TA		KR	LS	HM	CR	MR	TA
KR	**0.003**	0.003	0.003	0.003	0.003	0.003	KR		0.054	0.107	0.064	0.070	0.082
LS	-0.003	**0.003**	0.003	0.003	0.003	0.002	LS	0.001		0.087	0.029	0.044	0.048
HM	-0.028	-0.021	**0.003**	0.003	0.003	0.002	HM	0.033	0.036*		0.084	0.057	0.086
CR	0.005	0.003	-0.022	**0.003**	0.003	0.002	CR	0.011	0.013*	0.021*		0.036	0.049
MR	0.018	0.020*	-0.014	0.003	**0.003**	0.003	MR	0.024*	0.021*	0.004	0.011*		0.044
TA	0.010	0.003	-0.012	0.011*	0.035*	**0.002**	TA	0.036*	0.030*	0.051*	0.026*	0.022*	
M. endeavourae	KR	LS	TA					KR	LS	TA		
KR	**0.007**	0.006	0.005				KR		0.050	0.053		
LS	-0.002	**0.006**	0.005				LS	0.020*		0.035		
TA	-0.011	-0.002	**0.004**				TA	0.023*	0.003			
M. gracilis	KR	CR	CP					KR	CR	CP		
KR	**0.005**	0.006	0.005				KR		0.030	0.026		
CR	-0.014	**0.006**	0.006				CR	0.004		0.015		
CP	-0.013	-0.001	**0.005**				CP	0.003	0.003			

COI (unshaded) Kimura 2-parameter (K2P) pairwise genetic distance within regions (on the diagonal), between regions (above diagonal) and pairwise Φ_{ST} (below diagonal) values, and microsatellite variation (shaded) Nei's genetic distance (D_A) (above diagonal) and pairwise F_{ST} (below diagonal) values amongst regions of three *Munida* species. * $p < 0.05$. Population codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.

Table 4.4 Analysis of molecular variance (AMOVA) amongst squat lobster populations of three *Munida* species based on COI (unshaded) and microsatellite variation (shaded).

Source of variation	Variance components	Percentage of variation	F-statistics	Variance components	Percentage of variation	F-statistics
Munida isos						
A) One group						
Amongst populations	0	0.45	$\Phi_{ST} = 0.005$	0.14	3.39	$F_{ST} = 0.034***$
Within populations	0.84	99.55		4.03	96.61	
B) Two provinces (Northern / Southern)						
Between provinces	0	-0.1	$\Phi_{CT} = 0.001$	0.04	0.99	$F_{CT} = 0.010**$
Amongst provinces within populations	0	0.5	$\Phi_{SC} = 0.004$	0.12	2.88	$F_{SC} = 0.029***$
Within populations	0.84	99.6	$\Phi_{ST} = 0.005$	4.03	96.13	$F_{ST} = 0.039***$
C) Three regions (North / Central / South)						
Amongst regions	0	0.07	$\Phi_{CT} = 0.001$	0.04	1.04	$F_{CT} = 0.010***$
Amongst regions within populations	0	0.4	$\Phi_{SC} = 0.004$	0.11	2.66	$F_{SC} = 0.027***$
Within populations	0.84	99.52	$\Phi_{ST} = 0.005$	4.03	96.31	$F_{ST} = 0.037***$
D) Six geomorphic features (KR / LS / HM / CR / MR / TA)						
Amongst geomorphic features	0.01	0.8	$\Phi_{CT} = 0.008$	0.06	1.49	$F_{CT} = 0.149***$
Amongst geomorphic features within populations	0	-0.15	$\Phi_{SC} = -0.001$	0.09	2.16	$F_{SC} = 0.022***$
Within populations	0.85	99.34	$\Phi_{ST} = 0.007$	4.03	96.34	$F_{ST} = 0.037***$
M. endeavourae						
A) One group						
Amongst populations	0.01	0.33	$\Phi_{ST} = 0.003$	0.07	3.33	$F_{ST} = 0.033$
Within populations	1.9	99.67		2.05	96.67	
B) Three geomorphic features (KR / LS / TA)						
Amongst geomorphic features	-0.01	-0.66	$\Phi_{CT} = -0.007$	0.03	1.28	$F_{CT} = 0.013*$
Amongst geomorphic features within populations	0.02	0.79	$\Phi_{SC} = 0.008$	0.05	2.19	$F_{SC} = 0.022$
Within populations	1.9	99.87	$\Phi_{ST} = 0.001$	2.06	96.53	$F_{ST} = 0.035$

Source of variation	Variance components	Percentage of variation	F-statistics	Variance components	Percentage of variation	F-statistics
M. gracilis						
A) One group						
Amongst populations	-0.01	-0.56	$\Phi_{ST} = -0.006$	0.07	2.83	$F_{ST} = 0.028***$
Within populations	1.84	100.56		2.25	97.17	
B) Three geomorphic features (KR / CP / CR)						
Amongst geomorphic features	-0.01	-0.31	$\Phi_{CT} = -0.003$	0	-0.1	$F_{CT} = 0$
Amongst geomorphic features within populations	-0.01	-0.38	$\Phi_{SC} = -0.004$	0.07	2.9	$F_{SC} = 0.029***$
Within populations	1.84	100.69	$\Phi_{ST} = -0.007$	2.25	97.2	$F_{ST} = 0.028***$

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Population codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.
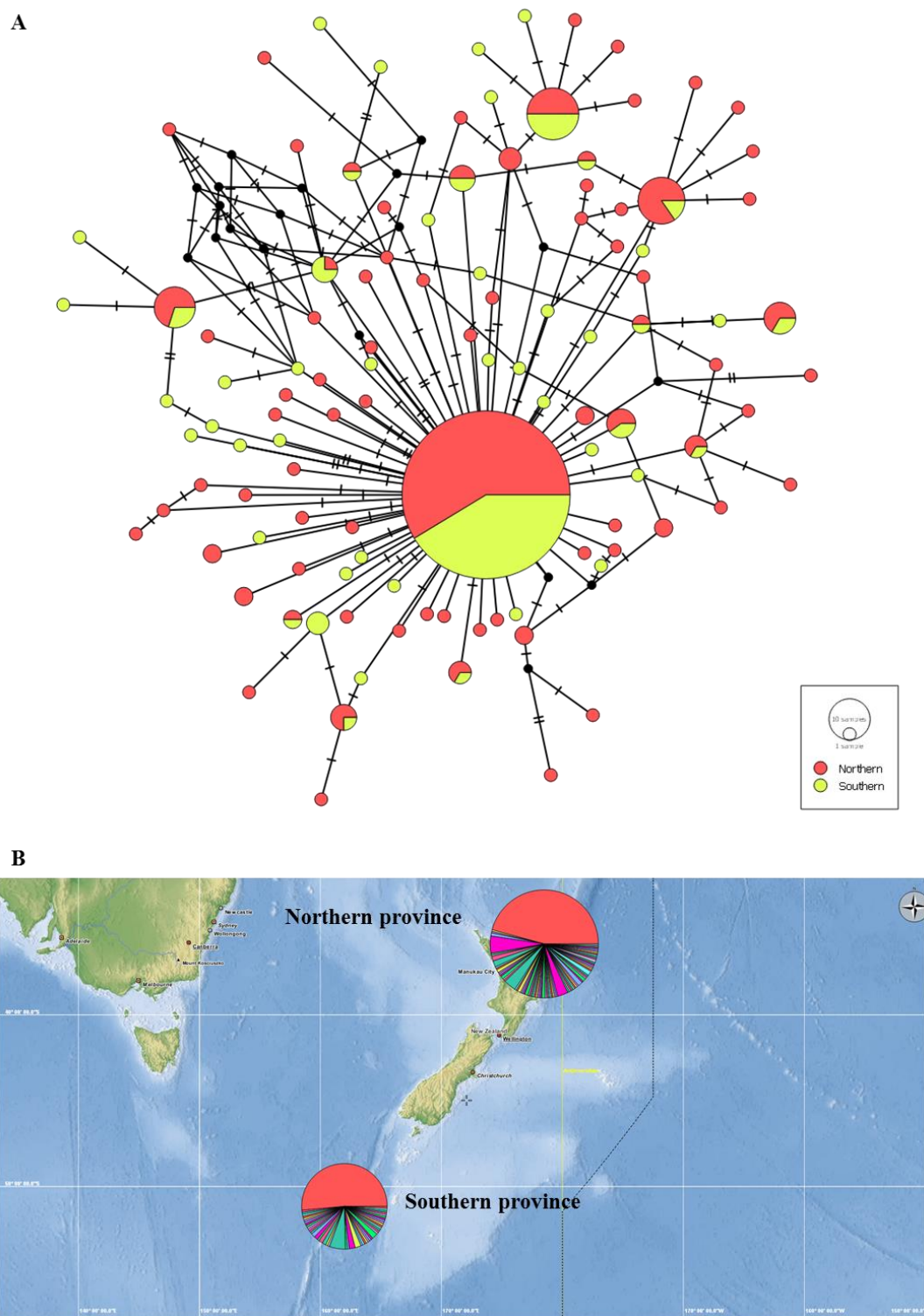
Figure 4.2 Median-joining network (A) and COI haplotype map (B) of *M. isos* based on spatial differentiation of two provinces.

(A) Coloured circles represent haplotypes and circle sizes are proportional to the number of individuals. Black circles represent hypothesised (not observed) haplotypes that are intermediate between observed haplotypes. Lines denote base pair changes; (B) Circles represent province populations and colours denote haplotypes. The location of the circles is arbitrary within the province boundary.

A

B
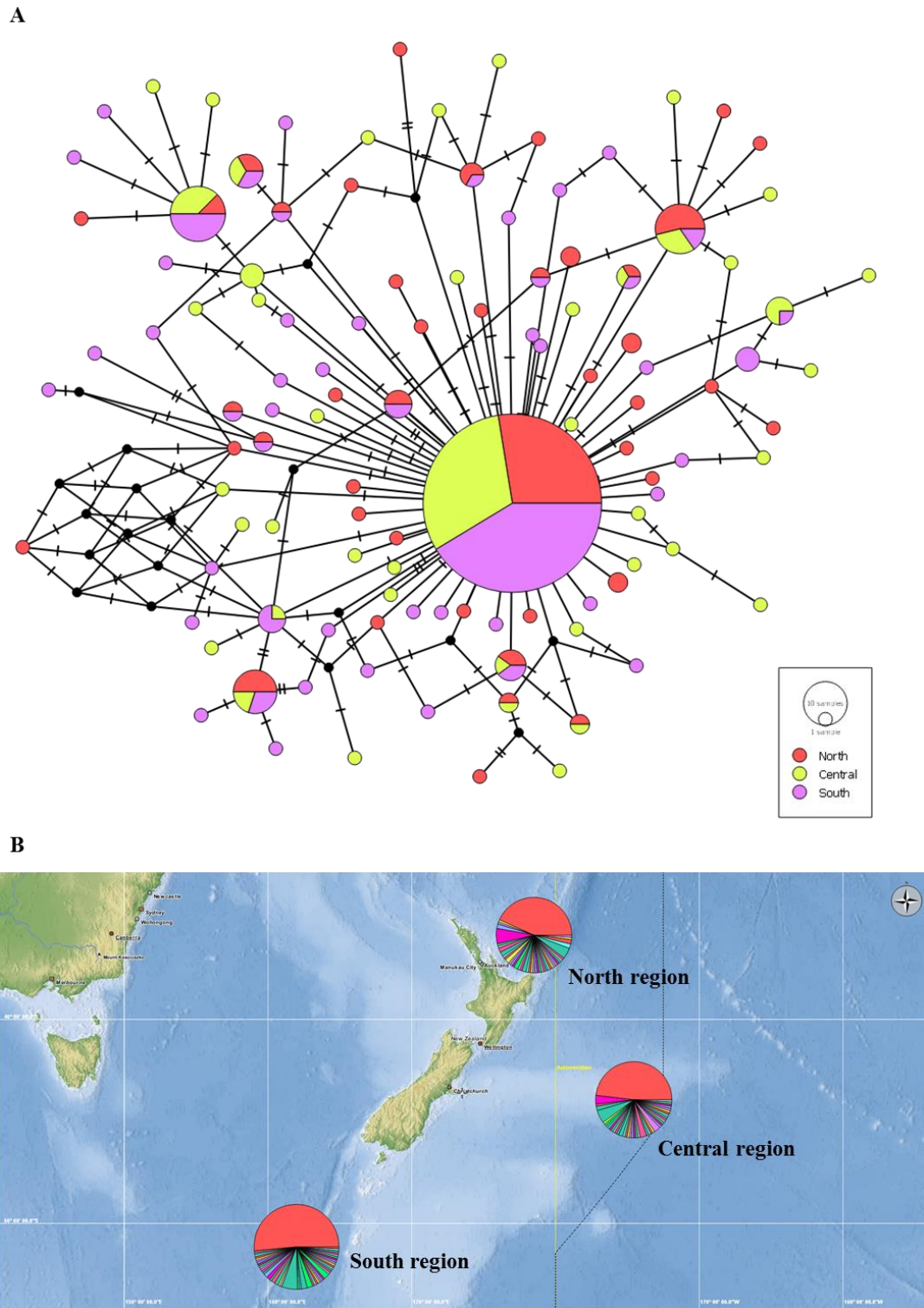
Figure 4.3 Median-joining network (A) and COI haplotype map (B) of *M. isos* based on spatial differentiation of three regions.

(A) Coloured circles represent haplotypes and circle sizes are proportional to the number of individuals. Black circles represent hypothesised (not observed) haplotypes that are intermediate between observed haplotypes. Lines denote base pair changes; (B) Circles represent region populations and colours denote haplotypes. The location of the circles is arbitrary within the region boundary.
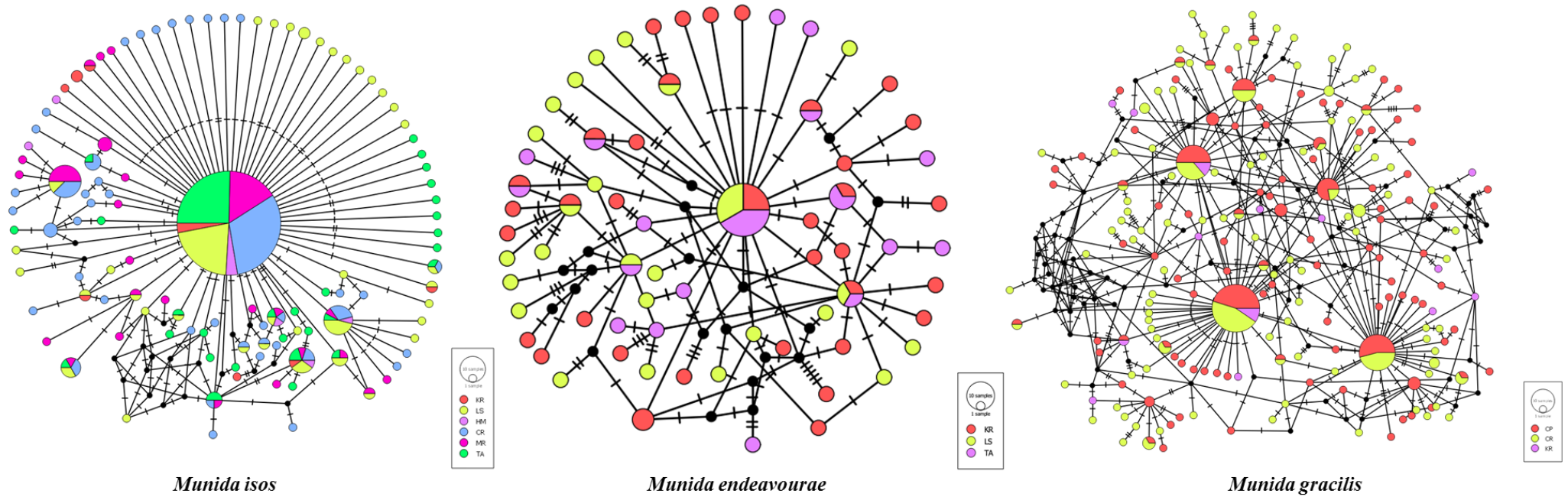
Figure 4.4 Median-joining network of COI haplotypes of three *Munida* species for geomorphic features.

Coloured circles represent haplotypes and circle sizes are proportional to the number of individuals. Black circles represent hypothesised (not observed) haplotypes that are intermediate between observed haplotypes. Lines denote base pair changes. Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.
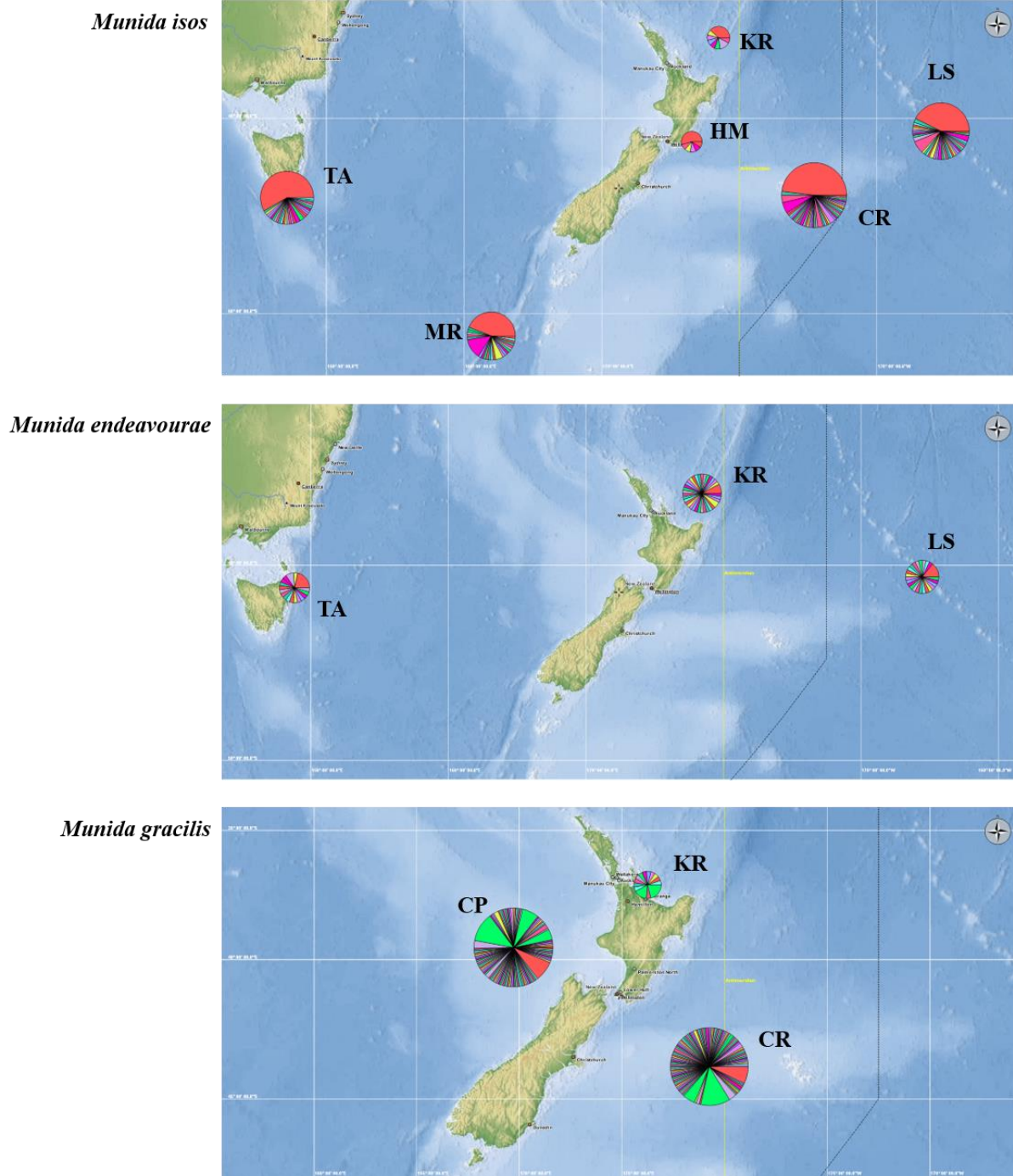
Figure 4.5 COI haplotype maps of three *Munida* species and geomorphic features.

Circles represent populations, colours denote haplotypes, and circle sizes represent relative samples sizes. Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau. The location of the circles is representative only of the location of the geomorphic feature.
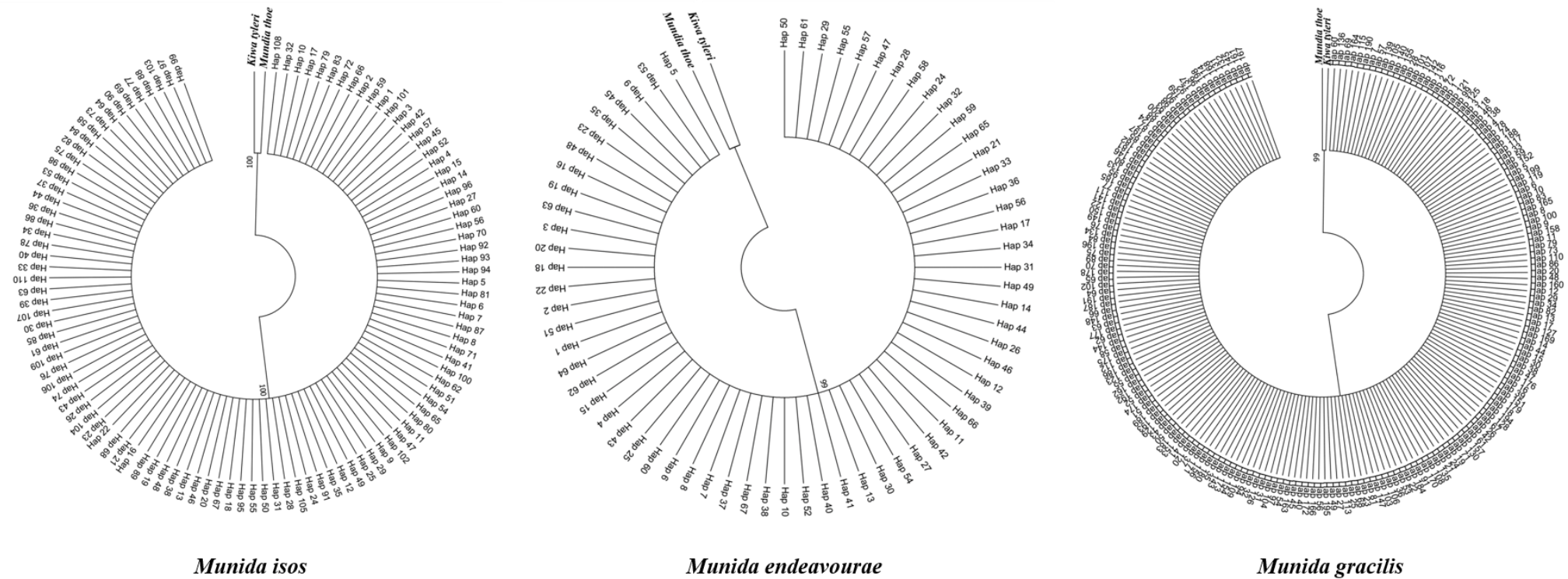
Figure 4.6 Maximum likelihood (ML) phylogenetic trees of three *Munida* species based on COI variation.

The numbers above the branches correspond to bootstrap support > 90% obtained in the ML analyses.

4.3.1.3 Demography

The Fu's F_S value was significantly ($p < 0.02$) negative in populations of both provinces and regions for *M. isos*, as well as in populations of geomorphic features for the three *Munida* species (Table 4.2). Mismatch distributions were all unimodal across all analyses (Figure 4.7–4.8). Additionally, *Hri* scores were non-significant for all analyses in three *Munida* species. Significantly negative Fu's F_S, unimodal mismatch distributions and non-significant *Hri* scores were consistent across the three *Munida* species, suggesting a departure from neutral expectations and exponential demographic expansions cannot be rejected. The onset of potential population expansions in southwest Pacific were estimated in ca. 161 kya ($\tau = 2.115$), 244 kya ($\tau = 3.213$) and 216 kya ($\tau = 2.842$) for the *M. isos*, *M. endeavourae* and *M. gracilis*, respectively, coinciding with the Middle Pleistocene. After applying the ten-fold correction, these estimates become 16.1 kya, 24.4 kya and 21.6 kya, respectively, which correspond to Late Pleistocene.
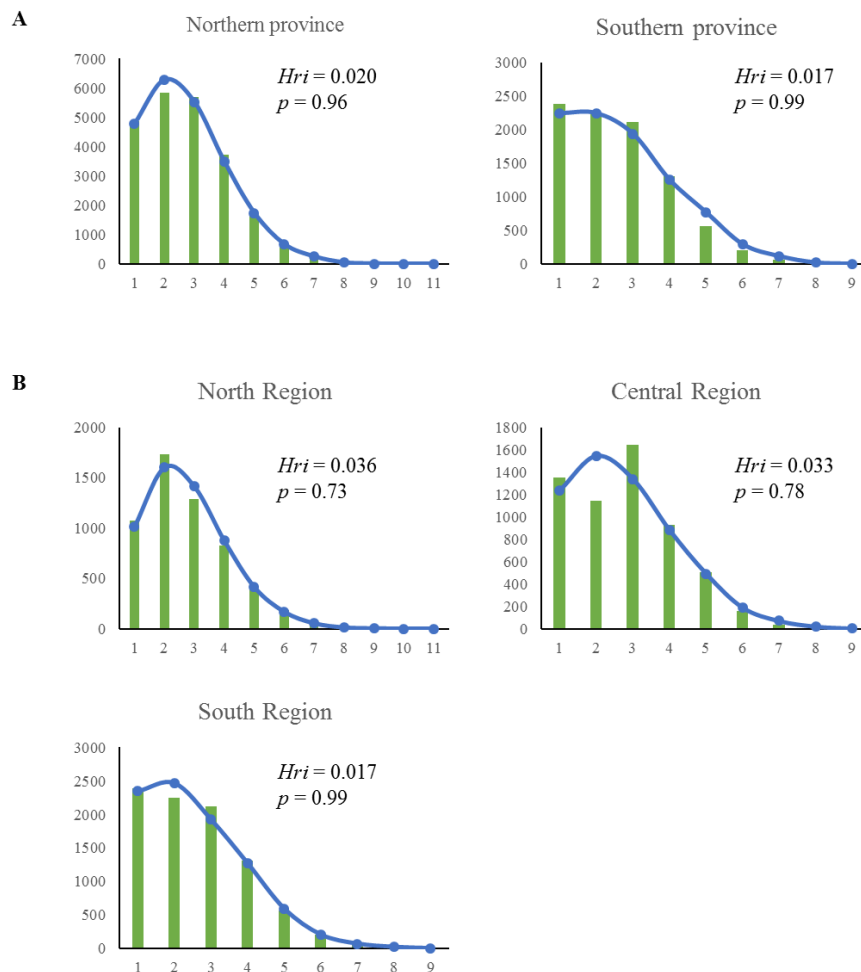
Figure 4.7 Mismatch distributions of COI nucleotide pairwise differences of *M. isos* based on spatial differentiation of two provinces (A) and three regions (B).
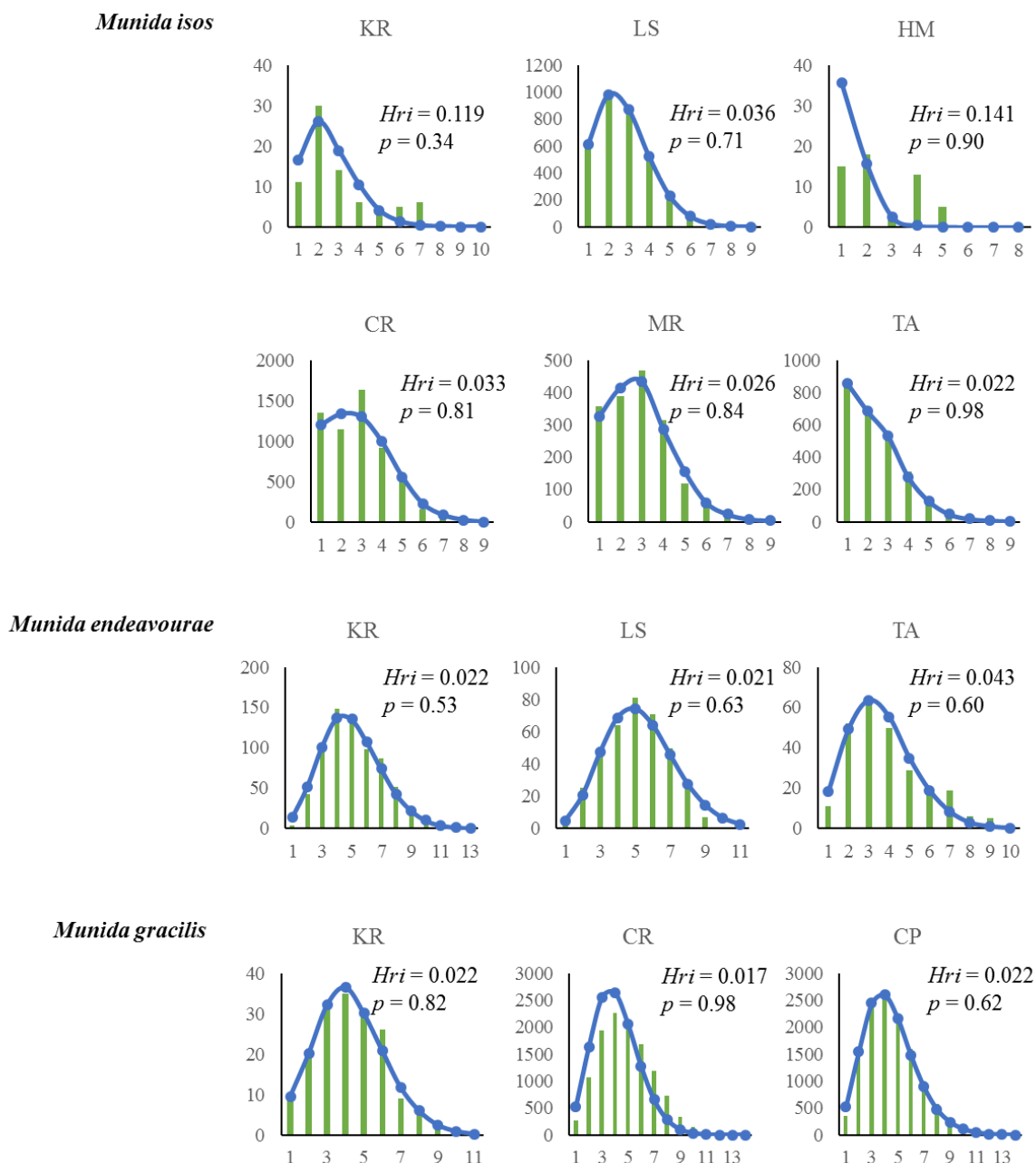
Figure 4.8 Observed and expected mismatch distributions of COI nucleotide pairwise differences for three *Munida* species for geomorphic feature populations.

Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.

4.3.2 Population genetic structure based on microsatellite variation

4.3.2.1 Genetic diversity

According to Lositan analysis, no microsatellite locus was identified as likely to be experiencing selection in any species. The presence of null alleles ($r > 0.2$) was detected at loci MI_06 (*M. isos*) and MI_02 (*M. gracilis*), whilst none was observed at these same loci in the two other species. Evidence of homozygote excess and stuttering was observed, but these were randomly distributed across loci in the three *Munida* species. No large allele dropout was indicated in any locus.

In total, 167, 63 and 74 alleles with size ranges of 78–134 bp, 83–172 bp and 82–200 bp were amplified at seventeen, eleven and nine microsatellite loci in *M. isos*, *M. endeavourae* and *M. gracilis*, respectively. In general, allelic variation (e.g., H_E and *PIC*) of tested loci was only moderately high. At locus level, four loci (MI_21, MI_29, MI_39 and MI_40) in *M. isos*, three (MI_06, MI_10 and MI_39) in *M. endeavourae* and one (MI_14) in *M. gracilis* exhibited a relatively low level of polymorphism compared to other loci (details are summarised in Table 4.5). For *M. isos*, the Northern province exhibited a higher level of genetic diversity than the Southern province. Additionally, the average number of effective alleles and expected heterozygosity were greatest in populations of the Central region, but the largest number of private alleles was found in the South region. At the geomorphic feature level, the observed and expected heterozygosities in the populations were reasonably consistent in all features for the three species (detailed results are reported in Table 4.2). At the geomorphic feature scale, populations of *M. isos* and *M. gracilis* exhibited moderate levels of genetic diversity whilst *M. endeavourae* showed slightly lower genetic diversity. A number of private alleles were identified in nearly all populations of geomorphic features across the three *Munida* species, with the highest numbers being observed for the Chatham Rise (*M. isos* and *M. gracilis*) and Kermadec Ridge (*M. endeavourae*) populations, whilst none were detected for the Kermadec Ridge population of *M. gracilis*. Average F_{IS} values were positive across populations of geomorphic features, with the only exception being for the population of *M. endeavourae* from the Tasmanian slope, which exhibited a deficit of heterozygotes.

For *M. isos*, after FDR correction for multiple comparisons, 28 (82.35%) and 37 (72.55%) tests were significantly different from HWE ($p < 0.05$) at pairwise combinations of loci × populations at the province and regional spatial scale, respectively. For geomorphic features, 38 (37.25%), 5 (15.15%) and 14 (51.85%) tests were significantly different from

HWE ($p < 0.05$) in *M. isos*, *M. endeavourae* and *M. gracilis*, respectively, with the majority exhibiting heterozygote deficiencies. In particular, loci MI_25 and MI_02 demonstrated significant departures from HWE within all populations at geomorphic feature scale of *M. isos* and *M. gracilis*, respectively.

For *M. isos*, only four out of 136 locus-pairs showed significant LD after adjusting *p*-values using FDR correction ($p < 0.05$). No evidence of LD was found in *M. endeavourae* and *M. gracilis*.

Table 4.5 Summary statistics of locus-specific genetic variation of three *Munida* species based on microsatellite variation.

Locus	*Munida isos*							*M. endeavourae*							*M. gracilis*						
	N_A	N_E	H_O	H_E	PIC	A_R	F_{IS}	N_A	N_E	H_O	H_E	PIC	A_R	F_{IS}	N_A	N_E	H_O	H_E	PIC	A_R	F_{IS}
MI_01	6	3.178	0.477	0.696	0.644	4.003	0.304	5	2.668	0.813	0.603	0.597	4.945	-0.397	11	3.028	0.813	0.676	0.622	5.881	-0.215
MI_02	16	3.463	0.569	0.708	0.709	6.203	0.187	10	2.024	0.545	0.512	0.487	6.681	-0.089	8	1.804	0.046	0.446	0.418	3.814	0.893
MI_04															13	2.570	0.441	0.616	0.560	6.146	0.278
MI_05	24	4.089	0.743	0.750	0.762	7.529	-0.008														
MI_06	6	1.551	0.002	0.332	0.303	3.337	0.996	2	1.082	0.078	0.076	0.068	1.895	-0.041							
MI_09	6	2.056	0.455	0.519	0.458	3.380	0.097	4	1.812	0.542	0.454	0.399	3.462	-0.213							
MI_10								2	1.145	0.132	0.122	0.102	1.972	-0.073							
MI_14	12	2.450	0.534	0.594	0.575	4.842	0.079								7	1.042	0.030	0.040	0.055	1.950	0.223
MI_17															9	4.707	0.961	0.795	0.759	5.803	-0.222
MI_19-2	12	2.796	0.526	0.634	0.627	5.516	0.151														
MI_20	12	3.220	0.677	0.701	0.640	4.589	0.014	4	1.454	0.282	0.312	0.279	3.084	0.077	3	1.879	0.391	0.466	0.388	2.407	0.157
MI_21	8	1.271	0.044	0.197	0.196	2.335	0.529														
MI_23-2	11	1.646	0.303	0.383	0.414	4.243	0.154	13	2.612	0.491	0.602	0.555	7.715	0.172							
MI_25	4	2.015	0.929	0.513	0.388	2.210	-0.846								6	2.281	0.693	0.567	0.487	3.957	-0.240
MI_29	6	1.230	0.054	0.174	0.217	2.852	0.538	7	1.649	0.343	0.399	0.359	4.262	0.127	11	2.762	0.422	0.643	0.604	6.197	0.336
MI_31	12	1.863	0.390	0.465	0.408	4.345	0.129														
MI_39	9	1.410	0.152	0.280	0.237	2.862	0.357	5	1.087	0.063	0.076	0.090	2.969	0.102							
MI_40	5	1.083	0.049	0.075	0.092	1.913	0.214	4	1.686	0.241	0.407	0.370	3.560	0.389	6	1.708	0.364	0.418	0.353	3.357	0.109
MI_51	10	2.233	0.442	0.538	0.547	4.484	0.127	7	2.718	0.617	0.641	0.574	5.792	0.018							
MI_52	8	2.102	0.401	0.481	0.529	4.636	0.097														
Mean	9.824	2.215	0.397	0.473	0.456	4.075	0.183	5.727	1.812	0.377	0.382	0.353	4.212	0.007	8.222	2.420	0.462	0.519	0.472	4.390	0.147

N_A = average number of alleles; N_E = average number of effective alleles; H_O = observed heterozygosity; H_E = expected heterozygosity; PIC = polymorphism information content; A_R= allelic richness; F_{IS} = inbreeding coefficient.

4.3.2.2 Genetic differentiation

Genetic differentiation amongst populations was low for the three *Munida* species, as revealed by Nei's genetic distance (D_A) and F_{ST} values. For *M. isos*, F_{ST} values amongst populations of the two provinces and three regions remained low but significant ($p < 0.05$). For *M. isos* at the geomorphic feature scale, only the pairwise comparison between populations of the Hikurangi Margin and Tasmanian slope ($F_{ST} = 0.051$) exhibited a moderate level of differentiation (Table 4.3). For the other two species' population pairs, significance ($p < 0.05$) was detected for thirteen F_{ST} values after correction, of which eleven and two were from *M. isos* and *M. endeavourae*, respectively.

AMOVA revealed significant genetic differentiation at different spatial levels: within all samples of *M. isos* and *M. gracilis*; amongst the populations of the two provinces, three regions, and six geomorphic features for *M. isos*; and amongst populations of the three geomorphic features for *M. endeavourae* and *M. gracilis* (Table 4.4). Nonetheless, AMOVA partitioned more than 96% of the genetic variation within samples for the three *Munida* species, leaving less than 4% variance attributable to the differences amongst populations and amongst populations within groups.

4.3.2.3 Genetic structure

PCoA, with relatively low cumulative variation represented by the first two factors, revealed no evidence of genetic differentiation across populations of provinces or regions for *M. isos*, and amongst populations of geomorphic features for the three *Munida* species (Figure 4.9–4.10), whilst Structure showed the highest support for two clusters ($\Delta K = 2$) for all datasets. However, given that the ΔK approach is unable to detect the best K when $K = 1$, post-processing of Structure bar plots was conducted in CLUMPAK. As a result, $K = 1$ and $K = 2$ could not be successfully differentiated between for *M. isos* (10/10; 10/10) at three spatial scales (provinces, regions and geomorphic features) and *M. gracilis* (10/10; 10/10) at geomorphic feature level, whereas $K = 2$ (5/10) was rejected in favour of $K = 1$ for *M. endeavourae* (Figure 4.11–4.12). For *M. isos*, when $K = 2$, individuals from the Southern province and South region were mostly assigned to the blue cluster based on provinces and region scales. At the geomorphic feature level, the majority of individuals from the Hikurangi Margin, Macquarie Ridge and Tasmanian slope populations were assigned to the blue cluster, whilst populations from other features showed a more complex genetic structure with a shared membership of both the blue and

orange clusters. This finding was supported by the fact that ten out of eleven significant pairwise F_{ST} values were detected in populations from these three geomorphic features (Hikurangi Margin, Macquarie Ridge and Tasmanian slope). For *M. endeavourae*, no pattern was revealed, with all individuals assigned equally to both clusters, indicating no structure. For *M. gracilis*, individuals from populations of all three geomorphic features were randomly assigned to two clusters, echoing the AMOVA results as major variance was attributed to the differences within geomorphic features. In the BAPS analyses, the inferred number of populations was one ($K = 1$) for all datasets of three *Munida* species, indicating no genetic structure.

A

B
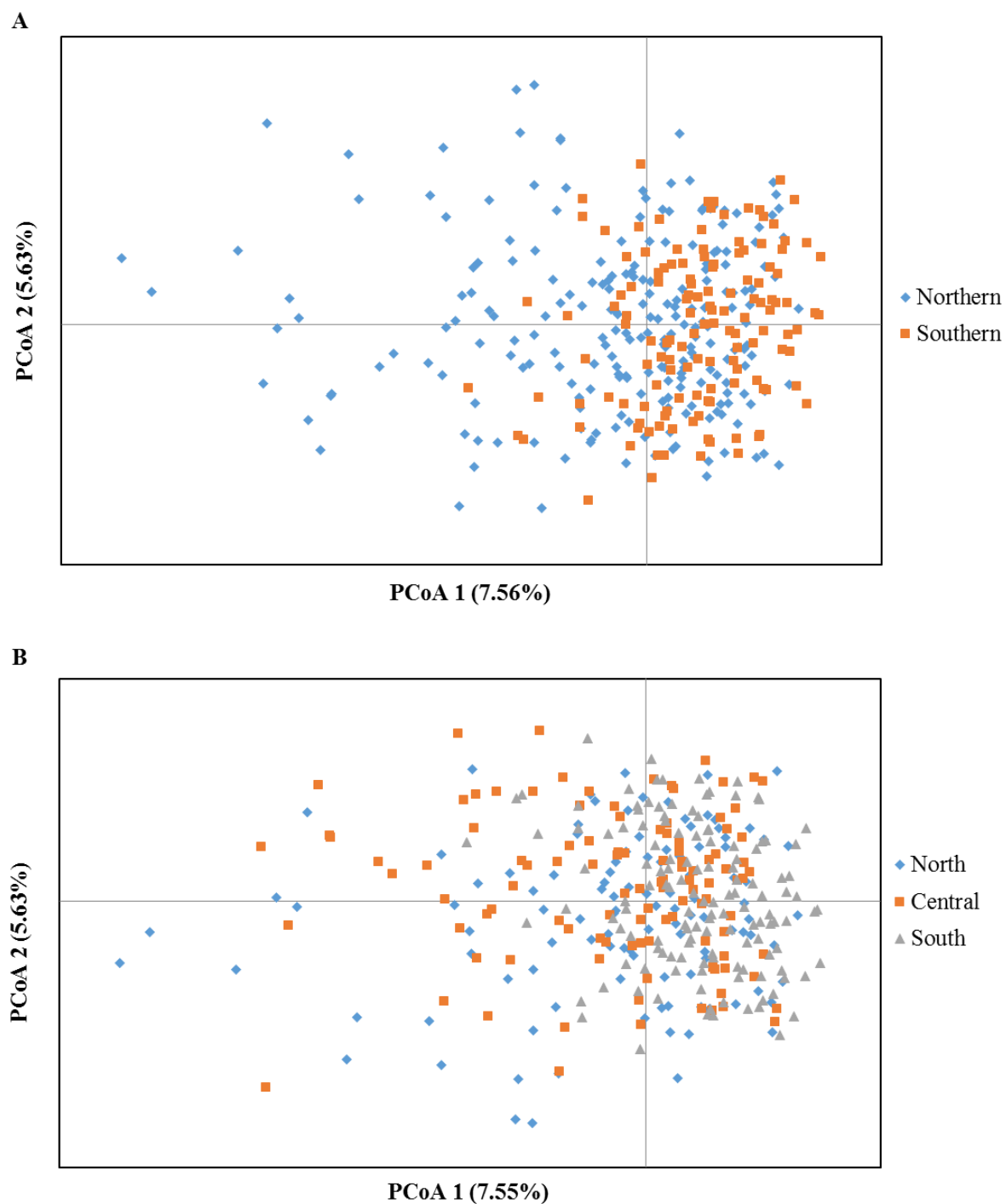
Figure 4.9 Principal coordinates analyses (PCoA) of by microsatellite variation exhibited by *M*. *isos* based on spatial differentiation of two provinces (A) and three regions (B).

Munida isos

Munida endeavourae

Munida gracilis
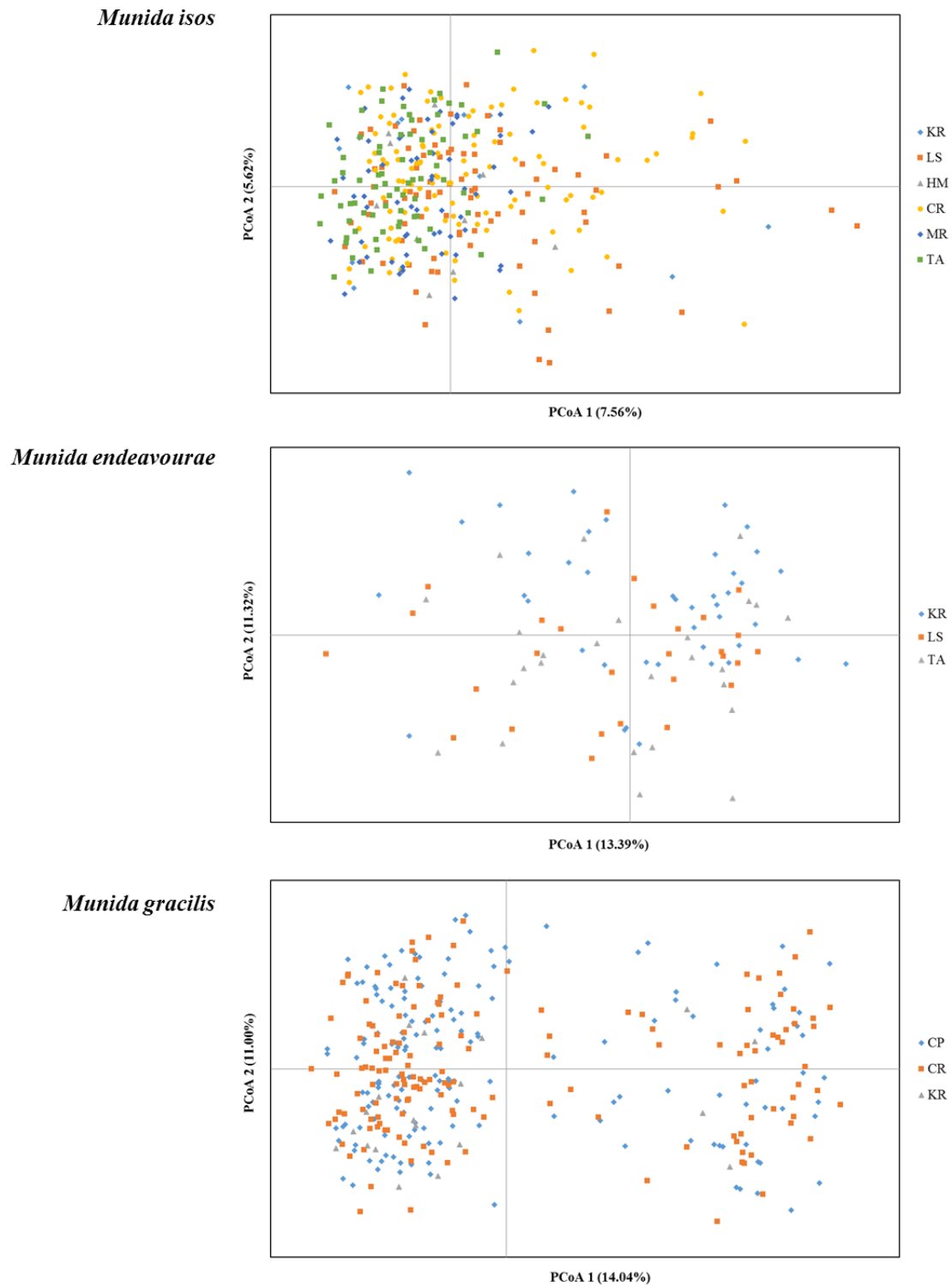
Figure 4.10 Principal coordinates analyses (PCoA) of microsatellite variation of three *Munida* species amongst populations of geomorphic features.

Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.

A   K=1 (10/10)

Northern	Southern

K=2 (10/10)

Northern	Southern

B   K=1 (10/10)

North	Central	South

K=2 (10/10)

North	Central	South

K=3 (4/10)

North	Central	South

Figure 4.11 Structure cluster analyses of *M. isos* based on spatial differentiation of two provinces (A) and three regions (B) as revealed by microsatellite variation.
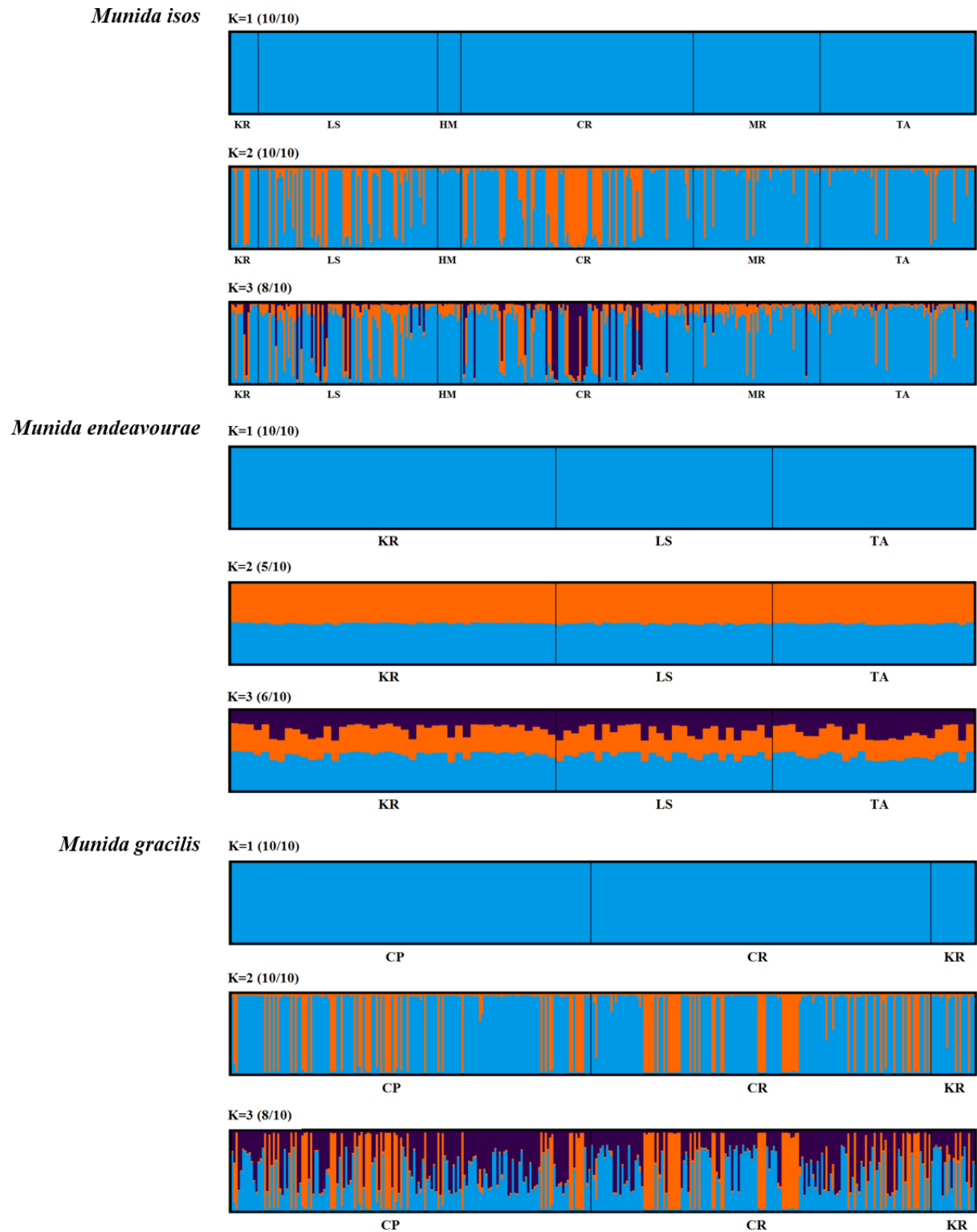
Figure 4.12 Structure cluster analyses of three *Munida* species based on microsatellite variation amongst populations on geomorphic features.

Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.

Results from DAPC, an analysis that can maximise genetic variation amongst groups whilst minimising separation within groups, indicated three genetic clusters corresponding to the three regions as defined by the first two principal components (PC) for *M. isos* (Figure 4.13). At the geomorphic feature level, 2, 3 and 3 major genetic clusters were calculated for *M. isos*, *M. endeavourae* and *M. gracilis*, respectively (Figure 4.14). For *M. isos*, PC1 separated the Tasmanian slope population from the populations of all other geomorphic features. The major cluster was further divided by PC2 into individuals from populations of the Kermadec Ridge and Louisville Seamount Chain, Chatham Rise, and Macquarie Ridge and the Hikurangi Margin, which was congruent with the F_{ST} and Structure results. For *M. endeavourae* and *M. gracilis*, three distinct clusters corresponding to the populations from the three geomorphic features were revealed, with clusters overlapping to different degrees for the two species.
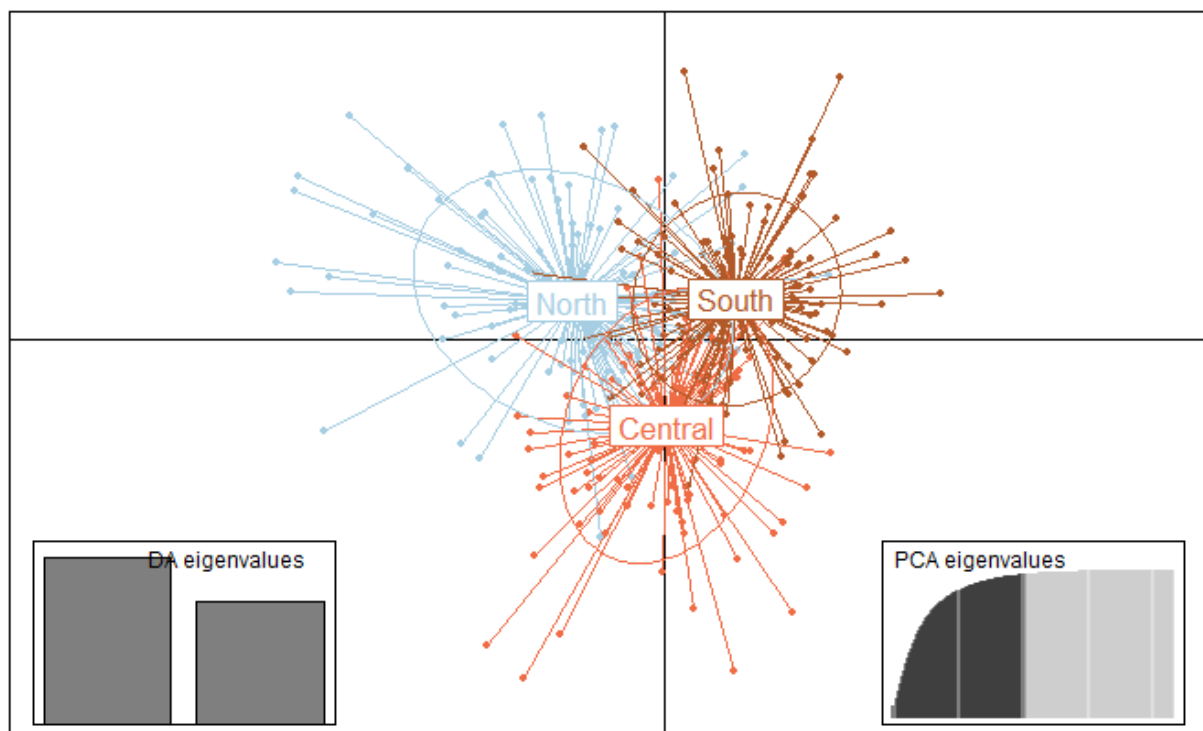
Figure 4.13 Scatterplot generated by the discriminant analysis of principal components (DAPC) of *M. isos* based on three regions as revealed by microsatellite variation.

Munida isos

Munida endeavourae

Munida gracilis
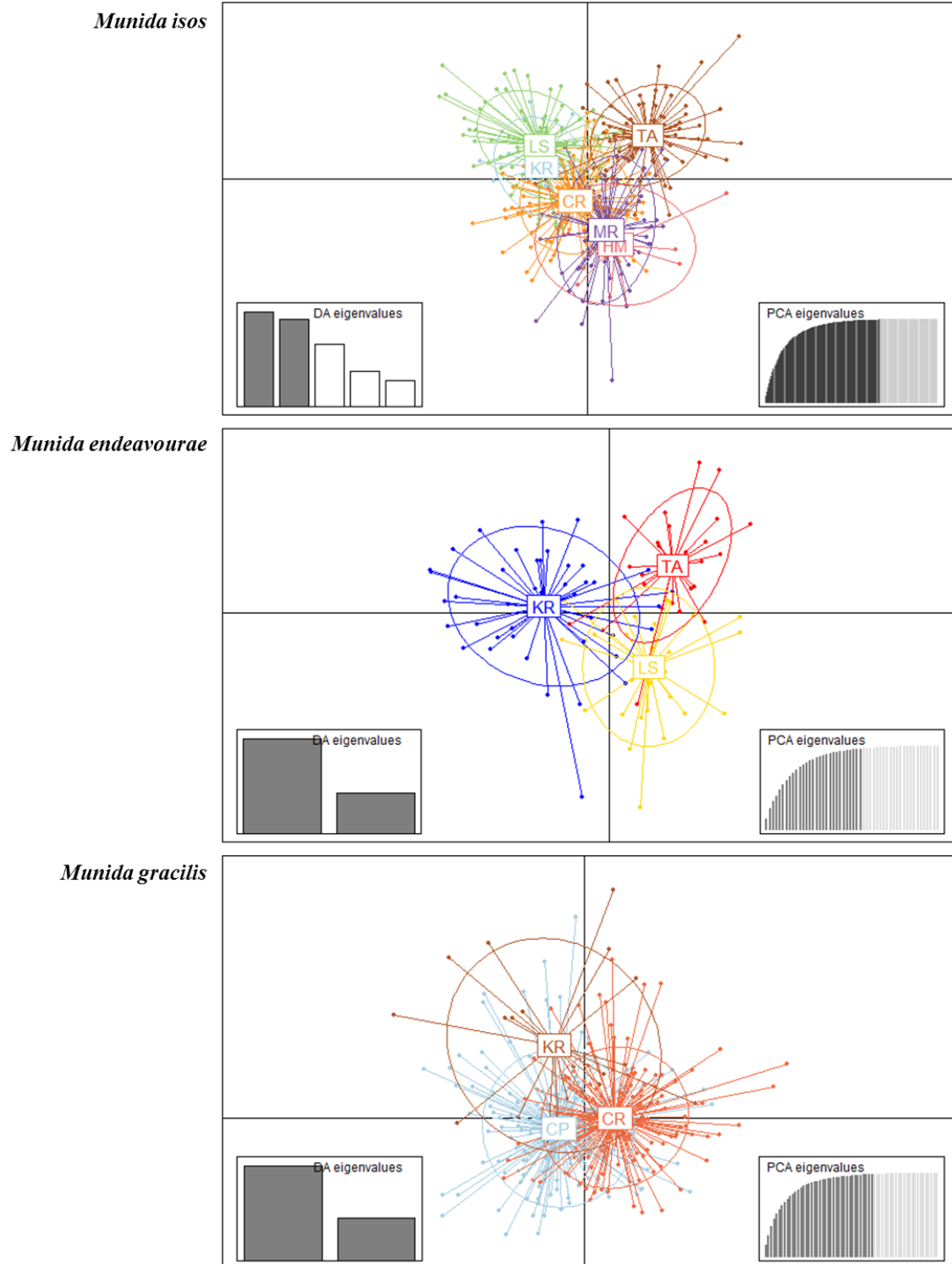
Figure 4.14 Scatterplot generated by the discriminant analysis of principal components (DAPC) of three *Munida* species based on microsatellite variation among populations of geomorphic features.

Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.

4.3.2.4 Gene flow and connectivity

The assignment-based tests performed in GeneClass elucidated a low to moderate level of classification success. For *M. isos*, 231 (65.1%) and 179 (50.4%) individuals were correctly assigned to their own province and region populations, respectively. At the geomorphic feature level, a total of 143 (40.6%), 53 (55.2%) and 156 (40.0%) individuals were correctly assigned to their own geomorphic feature populations for *M. isos*, *M. endeavourae* and *M. gracilis*, respectively. The "L_home" statistics demonstrated some evidence of first-generation migrants amongst sampled province/region/geomorphic feature populations. In total, 19, 2 and 8 individuals were below the threshold of the assignment tests ($\alpha = 0.01$) for *M. isos*, *M. endeavourae* and *M. gracilis*, respectively, increasing the possibility that these individuals may be first generation immigrants that have been introduced from an unknown source.

According to BayesAss, for *M. isos*, the acceptance rates of migration rates, allele frequencies and inbreeding coefficients were unable to reach the standard of 20–60% in the province and regional spatial scale datasets, thus excluded from the analyses, and only results for geomorphic features are presented. At the geomorphic feature level, estimates of contemporary migrants between populations were low, and the proportions of self-recruitment were high to very high (Table 4.6). For the *M. isos*, the Macquarie Ridge (81.22%) and Tasmanian slope (96.43%) populations exhibited strong signals of genetic isolation with high proportions of contribution from within each population (self-recruitment). Similar patterns were observed for individuals from the Louisville Seamount Chain for *M. endeavourae* (95.47%) and from the Challenger Plateau for *M. gracilis* (88.26%). Notably, the Kermadec Ridge population was identified as a sink population (only receiving migrants from other populations) in all *Munida* species. For *M. isos*, the Tasmanian slope population was identified as a source population (providing migrants to other populations while receiving no or few migrants). This was incongruent with findings from *M. endeavourae*, for which the Tasmanian slope population (68.01%) acted as a sink population, receiving migrants (30.27%) from the Louisville Seamount Chain population. For *M. gracilis*, populations from both the Chatham Rise and Challenger Plateau were both source and sink populations.

Table 4.6 Estimated migration rates from source populations (top) to recipient populations (side) for three *Munida* species based on microsatellite variation.

Geomorphic features	Source					
Munida isos	KR	LS	HM	CR	MR	TA
KR	**68.41%**	6.59%				16.35%
LS		**72.23%**				20.31%
HM			**68.59%**		19.99%	5.02%
CR				**78.74%**	6.22%	14.02%
MR					**81.22%**	14.87%
TA						**96.43%**
M. endeavourae	KR	LS	TA			
KR	**79.65%**	19.45%				
LS		**95.47%**				
TA		30.27%	**68.01%**			
M. gracilis	KR	CR	CP			
KR	**68.13%**	17.86%	14.00%			
CR		**76.81%**	22.98%			
CP		11.54%	**88.26%**			

Empty cells denote migration rates < 0.05, estimates of self-recruitment in bold (on the diagonal). Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.

4.3.2.5 Demography

For *M. isos*, populations from the Southern province and from the South region showed the largest effective contemporary population sizes at these two different spatial scales (Table 4.7). At the geomorphic feature scale, the contemporary effective population sizes for geomorphic features of the three *Munida* species were either infinite or in the range of 19.5–378.3. Populations of *M. isos* from the Kermadec Ridge and the Louisville Seamount Chain, as well as from the Chatham Rise for both *M. isos* and *M. gracilis*, exhibited comparable values of *Ne*. The *M. isos* population from the Tasmanian slope (*Ne* ± CI, 193.0 ± 115.5–480.6) and the *M. gracilis* population from the Challenger Plateau (378.3 ± 198.9–1624.6) exhibited large, but not infinitely large *Ne* values. The *Ne* estimates of all populations at geomorphic feature scale of *M. endeavourae* were infinite, with lower CI values of 40.6 to 696.8.

Table 4.7 Estimates of effective population size (*N*e) and associated confidence intervals for three *Munida* species based on microsatellite variation.

		N	P_{crit}	*N*e	lower CI	upper CI
Munida isos	Northern province	219	0.01	131.9	116.4	150.7
	Southern province	136	0.01	765.2	406.7	4212.5
	North region	109	0.01	91.6	76.9	111.3
	Central region	110	0.01	127.7	104.3	161.1
	South region	136	0.01	765.2	406.7	4212.5
	KR	13	0.05	19.5	9.9	66.6
	LS	85	0.02	87.5	69.2	115.2
	HM	11	0.06	∞	15.3	∞
	CR	110	0.01	127.7	104.3	161.1
	MR	60	0.02	∞	666.3	∞
	TA	73	0.02	193.0	115.5	480.6
M. endeavourae	KR	42	0.02	∞	696.8	∞
	LS	28	0.03	∞	40.6	∞
	TA	26	0.03	∞	69.1	∞
M. gracilis	KR	20	0.04	∞	776.0	∞
	CR	155	0.01	199.3	126.0	392.1
	CP	164	0.01	378.3	198.9	1624.6

Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau.

Based on microsatellite variation, no bottleneck signal was detected in populations in any province, region or geomorphic feature according to the Wilcoxon sign rank test and the mode-shift approach, signifying no genetic bottleneck during the past few generations (Table 4.8). Besides, a significant ($p < 0.05$) pattern of heterozygote deficiency was observed under both SMM and TPM in populations in all provinces, regions and geomorphic features in the three *Munida* species, with only one exception (the Hikurangi Margin population of *M. isos*), suggesting recent population expansions, which was consistent with findings as revealed by variation in the COI region.

Table 4.8 Heterozygosity deficiency probabilities of Wilcoxon sign rank tests for three *Munida* species based on microsatellite variation.

		SMM	TPM 95%	Mode-shift
Munida isos	Northern province	**0.00002**	**0.00002**	
	Southern province	**0.00001**	**0.00002**	
	North region	**0.00003**	**0.00003**	
	Central region	**0.00002**	**0.00002**	
	South region	**0.00001**	**0.00002**	
	KR	**0.00232**	**0.01008**	
	LS	**0.00005**	**0.00034**	
	HM	0.10388	0.2106	
	CR	**0.00002**	**0.00005**	L-shaped
	MR	**0.00004**	**0.00013**	
	TA	**0.00008**	**0.00011**	
M. endeavourae	KR	**0.00049**	**0.00073**	
	LS	**0.00122**	**0.00464**	
	TA	**0.00684**	**0.00928**	
M. gracilis	KR	**0.02734**	**0.03711**	
	CR	**0.00488**	**0.00488**	
	CP	**0.00293**	**0.00488**	

Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope; CP = Challenger Plateau. $p < 0.05$, in bold.
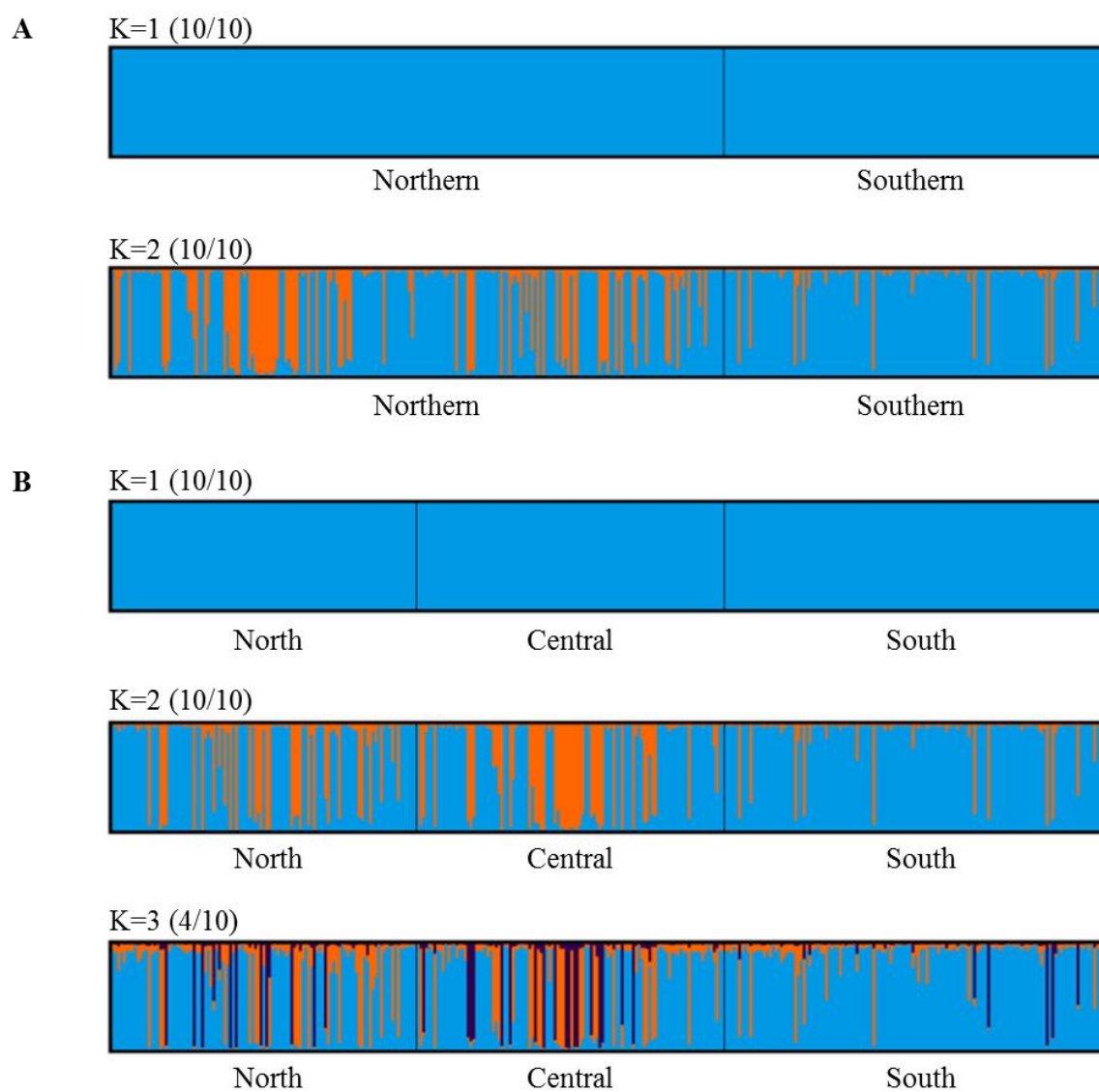
4.4 Discussion

Three squat lobsters, *Munida isos*, *M. endeavourae* and *M. gracilis*, have been tested for levels of genetic diversity and patterns of genetic connectivity in the southwest Pacific Ocean region based on mitochondrial COI region and nuclear microsatellite variation. Overall, the level of genetic diversity was high as revealed by the COI region, and moderate as revealed by microsatellite markers, across the three *Munida* species. There was no evidence of genetic subdivision amongst populations of the three *Munida* species based on the COI region data, and only limited genetic differentiation was observed in *M. isos* and *M. gracilis* based on microsatellite variation. Notably, both marker types demonstrated the occurrence of recent demographic expansions in the three species.

4.4.1 Genetic diversity

Genetic diversity in three *Munida* species was randomly distributed across the study area, as indicated by both the COI region and microsatellite markers. The mitochondrial COI region, in particular, exhibited a high level of genetic diversity for haplotypes across the three *Munida* species. Notably, within the genus *Munida*, high levels of polymorphisms in COI can be indicative of a relatively fast mutation rate, or because the taxa have been isolated for extended periods, allowing for mutations to accumulate (Coykendall *et al.* 2017). The observation of high haplotype diversity was consistent with findings from other deep-sea squat lobsters, such as *K. tyleri* (Roterman *et al.* 2016) and *P. monodon* (Haye *et al.* 2010). Nonetheless, *M. isos* exhibited lower genetic diversity level compared to *M. endeavourae* and *M. gracilis*, as revealed by other genetic diversity estimates (e.g., nucleotide diversity and average number of nucleotide differences) based on COI region, whilst *M. endeavourae* showed the lowest genetic diversity on the basis of microsatellite markers with respect to expected heterozygosity and effective number of alleles that could reflect genetic drift (Hedrick 1999). Given that sample size is uneven when comparing populations of *M. isos* and *M. gracilis*, allelic richness, which is independent on sample size, is included as an index to evaluate the level of genetic diversity for the newly developed microsatellite markers. Consequently, these microsatellite markers performed reasonably well, exhibited a moderate level of genetic diversity and provided valuable genetic information for the three *Munida* species.

Significant deviations from HWE ($p < 0.05$) were observed in three *Munida* species, which is common in other squat lobster species (Molecular Ecology Resources Primer

Development Consortium *et al.* 2010; Boyle *et al.* 2013; Nakajima *et al.* 2018). Deviations from HWE could be attributed to a variety of different causes such as null alleles and the Wahlund effect (Johnson and Black 1984). The presence of null alleles may result in mistyping of homozygotes and heterozygotes, leading to an over-estimation of genetic differentiation (DeWoody *et al.* 2006). However, null alleles are commonly observed in a variety of marine invertebrate species (e.g., Dailianis *et al.* (2011); Thaler *et al.* (2014); Roterman *et al.* (2016)) and in population genetic studies, null allele frequencies below 0.2 are commonly accepted when using microsatellite markers (Dakin and Avise 2004). In this study, although low-to-moderate frequencies of null alleles were detected in a number of loci across the three *Munida* species, only two loci in *M. isos* (MI_06) and *M. gracilis* (MI_02) exhibited null allele frequencies above 0.2. Given that little genetic differentiation was observed amongst populations of *M. isos* and *M. gracilis*, the null alleles at these two loci might not influence the overall interpretation of the results. The Wahlund effect is caused by individuals from different localities that are characterised by different genotype proportion, being mixed during sampling or during data analysis, resulting in excess homozygosity and significant F_{IS} estimates (Freeland 2005; Waples 2014). The highly significant deficit of heterozygote genotypes, as revealed by positive inbreeding coefficients ($F_{IS} > 0$), was obtained across all populations in three *Munida* species with only one exception (i.e., the Tasmanian slope population of *M. endeavourae*), indicating the presence of inbreeding. Moreover, HWE deviations occurred in the form of homozygote excesses, which might also result from small sample size (e.g., populations from Kermadec Ridge and Hikurangi Margin in *M. isos*), given the potential for extensive gene flow amongst populations, and population demographic history (e.g., expansions or selective sweeps). However, small sample size and limited spatial scale sampling are unavoidable constraints in population genetic studies of deep-sea organisms (Baco *et al.* 2016). For microsatellite-based population genetic studies, 25–30 individuals per population is considered sufficient to provide accurate estimations of genetic diversity and gene flow (Hale *et al.* 2012), and the majority of the populations tested in the three *Munida* species meet this criterion.

4.4.2 Genetic differentiation and connectivity

Genetic differentiation at different geographic scales was evaluated based on mitochondrial COI region and nuclear microsatellite markers in the three *Munida* species. For the COI region, according to the Decapoda intraspecific barcode threshold (0.285–1.375%) (da

Silva *et al.* 2011), the maximum pairwise K2P genetic distances (0.3% for *M. isos* and 0.6% for *M. endeavourae* and *M. gracilis*) amongst populations did not suggest the presence of cryptic species. These results support previous findings for *M. gracilis* in the New Zealand region (0–1.9%) (Bors *et al.* 2012) and align with observations in other squat lobsters, including *Fennerogalathea chacei* Baba, 1988 (0.13–0.61%) (Rodríguez-Flores *et al.* 2017). Despite the large geographic coverage of sampling sites, non-significantly low global Φ_{ST} values were observed for the three *Munida* species, suggesting genetic similarity in the squat lobster populations within the southwest Pacific Ocean study area. In contrast, microsatellite markers revealed some limited evidence of genetic differentiation amongst populations of *M. isos* and *M. gracilis*, but not *M. endeavourae*. Although nuclear microsatellite markers showed signs of low levels of genetic differentiation with significant global F_{ST} values in *M. isos* (0.039, 0.037 and 0.034 when defined at province, region and geomorphic feature scales, respectively) and *M. gracilis* (0.028), more than 96% of the variance was attributed to differences within populations, again suggesting high levels of gene flow leading to panmixia. Collectively, little genetic differentiation was observed amongst populations of three *Munida* species in the southwest Pacific Ocean, either from seamounts or soft sediments. The absence of genetic differentiation is usually consistent with either panmixia or recent demographic expansions that have left little or no signature in the genome. This lack of genetic differentiation has been observed in a number of squat lobster species (e.g., Samadi *et al.* (2006); Haye *et al.* (2010); Bors *et al.* (2012); Cabezas *et al.* (2012a); Wang *et al.* (2016)). For example, little or no evidence of differentiation has been observed previously amongst populations of *M. gracilis* based on the variation of the COI and 16S rRNA genes in the New Zealand region (Bors *et al.* 2012). Similarly, no significant structure was observed across the province, regional and geomorphic feature spatial scales for a deep-sea sponge species (Zeng *et al.* 2019). However, this pattern of genetic similarity is incongruent with a variety of deep-sea marine species, such as corals and sponges (e.g., Miller *et al.* (2011); Herrera *et al.* (2012); Miller and Gunasekera (2017)). For example, a North-Central-South regional pattern of the structure has been detected in a cup coral utilising ITS variation (Miller *et al.* 2010), and a province-scale differentiation has been reported in branching stony corals using a combination of D-loop, ITS and microsatellite markers in the New Zealand region (Zeng *et al.* 2017). Genetic differentiation has also been detected in a sponge species based on province, regional and geomorphic feature spatial scales as revealed by COI, Cytb and microsatellite variations (Zeng *et al.* 2019).

Based on mitochondrial COI sequence data, an overall lack of population genetic structure was detected for the three *Munida* species accordingly to maximum likelihood tree and median-joining network. Similar haplotype networks have been seen in other squat lobster taxa, such as *K. tyleri* (Roterman *et al.* 2016), *Munida intermedia* var. sarsi Huus, 1935 (García-Merchán *et al.* 2012), *M. gracilis* (Bors *et al.* 2012), *M. gregaria* (Pérez-Barros *et al.* 2014) and *P. monodon* (Haye *et al.* 2010). For nuclear microsatellite data, given that the differentiation amongst populations was weak for the three *Munida* species, it was difficult to infer population genetic structure. Notably, the Bayesian clustering software programmes Structure and BAPS can perform reasonably well at low levels of genetic differentiation (F_{ST} = 0.02–0.03) and are able to correctly identify the number of genetically distinct groups (Latch *et al.* 2006). In this study, considering the low F_{ST} values (*M. isos*: 0.001–0.05, *M. endeavourae*: 0.003–0.023 and *M. gracilis*: 0.003–0.004) amongst populations of three *Munida* species, Structure and BAPS were employed to calculate the best *K* (average number of nucleotide differences). Little or no structure was proposed for three *Munida* species in both approaches. Nonetheless, evidence showed that populations from the Tasmanian slope were potentially genetically different from the rest of *M. isos*, populations as revealed by results from the combination of Structure and DAPC analyses. The potentially genetic isolation of populations from Tasmanian slope may attribute to the hydrographic barriers to gene flow that limit migration or colonisation of populations to the east. The Subtropical Front (STF) may act as a barrier to pelagic larval dispersal, leading to a pattern of genetic differentiation between populations of Tasmanian slope and all other populations in New Zealand's EEZ.

The three *Munida* species exhibited low to moderate levels of assignment success, which was most probably derived from the high levels of connectivity and which lead to panmixia across the study area. Interestingly, the estimations of migration rate present were not always consistent for populations from the Tasmanian slope, which acted as a source population for *M. isos* and as a sink population for *M. endeavourae*. Moreover, populations from the Kermadec Ridge acted as sink populations in the three *Munida* species, which was inconsistent with a previous study of two stony corals where the Kermadec Ridge is a migrant source (Zeng *et al.* 2017). Given the fact that BayesAss can only produce accurate estimates of migration and self-replenishment when F_{ST} is greater than 0.05 (Faubet *et al.* 2007), all estimates of migration rates amongst populations as revealed by BayesAss should be cautiously interpreted and are best viewed in combination with other analyses. For *M. isos*, only one pairwise comparison exhibited a moderate level of differentiation ($F_{ST} > 0.05$), whilst none

was larger than 0.05 in *M. endeavourae* and *M. gracilis*. The hypothesis that populations from the Tasmanian slope being predicted as source populations was incongruent with the results showing them as being genetically different from all other populations. Nonetheless, if this hypothesis was true, it might attribute to the major surface current circulations – Subtropical Water (STW) and Subantarctic Water (SAW), both flowing eastward before meeting along the Chatham Rise (Chiswell *et al.* 2015), which might increase gene flow to all other populations within New Zealand region.

Collectively, patterns of connectivity vary across the range of deep-sea organisms with different ecological characteristics at different spatial scales, highlighting the need to expand the sampling scale with the inclusion of more representatives, as well as to test larger numbers of higher-resolution markers, to more robustly elucidate the potential genetic differentiation amongst squat lobster populations. Given that samples of the three *Munida* species, as VME-associated taxa, were collected across their distributional range in the southwest Pacific Ocean, it would be reasonably difficult to call for the sampling of more sites, making this data the best that is presently available.

Larval dispersal is crucial for connectivity and colonisation through migration in deep-sea environments. Larvae of *Munida* species are planktotrophic, but relatively little is known about the larval development of *M. isos*, *M. endeavourae* and *M. gracilis*. However, larval dispersal was studied for a congeneric species, *M. gregaria*, with a PLD that may reach 83 days at a rearing temperature of 8 °C (Pérez-Barros *et al.* 2007). The potentially long PLD of the three *Munida* species that lasts for perhaps several months is but one of various possible explanations for the panmixia observed in the study region. The larval lifespans may be extended by low temperature-induced metabolic rate reduction in a low-temperature environment, lessening the impact of dispersal strategy on patterns of genetic differentiation (Roterman *et al.* 2016). In addition, the density of seamounts, as potential stepping stones, may be higher than is presently thought, facilitating dispersal, and deep water currents may act as a driving force for the enhancement of connectivity as well (Handayani *et al.* 2014). Interestingly, the estimates of genetic homogeneity in marine organisms are often found to be inconsistent with expectations from inferred PLD, which may be due to oversimplification of PLD estimation (Weersing and Toonen 2009; Selkoe and Toonen 2011; Faurby and Barber 2012). Developmental biology studies of the three *Munida* species in future work should test the potential relationship between PLD and connectivity in squat lobster populations across their distributional range. Understanding this relationship will help inform the construction of

biophysical models that are also used to inform the design of MPA networks (Hilário *et al.* 2015).

In summary, the population genetics findings from three *Munida* species provide evidence for the rationality of the testing framework at the province, region, and geomorphic feature scale. However, little support for genetic differentiation has been observed, which suggests that oceanic currents, fronts, and geomorphic feature-forced currents and structures are not a barrier to gene flow for *Munida* taxa with relatively long PLDs.

4.4.3 Effective population size and demography

Based on nuclear microsatellite data, estimates of *Ne* varied amongst populations for the three *Munida* species, from relatively small effective contemporary population sizes to functionally infinite effective population sizes, which might be attributed to genetic drift and genetic sweeps. Similar estimations have been conducted in other squat lobster species, such as *M. lauensis*, with all populations estimated to be functionally infinite (Thaler *et al.* 2014), and *Munidopsis polymorpha*, exhibiting extremely low effective population sizes across all populations (Cabezas *et al.* 2012a). Recent bottleneck events or phases of expansion were tested with both the COI and microsatellite datasets.

For COI sequence data, neutrality tests revealed evidence of recent population expansions or selective sweeps across populations for all provinces, regions and geomorphic features of the three *Munida* species, and the subsequent mismatch-analyses could not reject the hypotheses of demographic expansions. Given that selective sweep would have reduced the haplotype diversity which is incongruent with our results, it would be indicative of recent demographic expansions in the three *Munida* species. This hypothesis of population expansions was further supported by microsatellite variations, with a significant ($p < 0.05$) pattern of heterozygote deficiency observed under both SMM and TPM in populations in all provinces, regions and geomorphic features in the three *Munida* species. However, for microsatellites, at least 20 individuals per populations and at least 4 polymorphic microsatellites are required to achieve reliable results in the software Bottleneck (Piry *et al.* 1999). In view of the Hikurangi Margin populations in *M. isos* showed inconsistent results with evidence of deviation from the model of migration-drift equilibrium (based on microsatellite variations) but conformed to the assumption of mismatch analysis of panmixia (based on COI), the observed non-significant pattern of heterozygote deficiency as revealed by microsatellite variations may attribute to the

small sample size ($N = 11$). Overall, robust consistency was observed in the demographic evaluation based on both mitochondrial and nuclear marker types for the three *Munida* species, indicating the likelihood of recent demographic expansions. Evidence of a recent population expansion following bottlenecks was reported for a number of squat lobster species, such as *P. monodon* (Haye *et al.* 2010), *M. gregaria* (Pérez-Barros *et al.* 2008; Pérez-Barros *et al.* 2014), *M. intermedia* (García-Merchán *et al.* 2012) and *K. tyleri* (Roterman *et al.* 2016).

In this study, the estimated time of demographic expansions for the three *Munida* species was ca. 161 kya, 244 kya and 216 kya for the *M. isos*, *M. endeavourae* and *M. gracilis*, respectively, coinciding with the Middle Pleistocene. However, with the ten-fold corrections of the mutation rate (intraspecies vs interspecies), the estimations (16.1 kya, 24.4 kya and 21.6 kya, respectively) would fit into Late Pleistocene. The demographic expansions inferred from rapid mutation rates (i.e., ten-fold correction) could be more realistic, in keeping with the lack of genetic structure in the three *Munida* species, than the relatively old phylogenetic substitution rates, as suggested by Ho *et al.* (2005; 2007; 2011). Notably, the estimated time of demographic expansion for *M. gregaria* (220–160 kya) from the southern Pacific and Atlantic Oceans (Pérez-Barros *et al.* 2014), *M. intermedia* (125–102 kya) from the Atlantic–Mediterranean transition area (García-Merchán *et al.* 2012) and *K. tyleri* (500 kya) from the Southern Ocean (Roterman *et al.* 2016) all occurred during the middle to late Pleistocene, which is consistent with estimations in the three southwest Pacific *Mundia* species. The Pleistocene was characterised by repeated cycles of glaciation, which have influenced the genetic composition of many organisms (Petit *et al.* 2003; Fraser *et al.* 2012). Distributional changes of deep-sea organisms, in particular, may be attributed to the severe climatic fluctuations as produced by glaciation events (Hewitt 2000). It was hypothesised that their distributions might have been restricted to lower latitudes during the periods of glaciation, followed by population expansion and recolonisation with the increase of sea level and seawater temperature after the glaciations (Pérez-Barros *et al.* 2008).

4.4.4 Implications for conservation and management

This study found that the population of *M. isos* from the Tasmanian slope was potentially genetically distinct (Figure 4.15). Populations from the Tasmanian slope are located in the 9991 km^2 Huon Commonwealth Marine Reserve (part of the South-east Commonwealth Marine Reserve Network) within the Australian EEZ. Three of the sample sites are within the

Habitat Protection Zone (IUCN IV), with mining prohibition and commercial fishing if authorised, whilst the fourth sample site for this population is situated in the Huon-Multiple Use Zone (IUCN VI), where activities such as mining and fishing can occur if authorised (Iftekhar *et al.* 2014). Moreover, the population of *M. isos* in the Huon Marine Park is a potential source population for populations within New Zealand's EEZ, and therefore its protection is helpful for the conservation of the genetic diversity of this species in the New Zealand region. Thus, the results of this study not only demonstrate the usefulness of the Huon Marine Park as an MPA to potentially protect VMEs in Australian waters, but also illustrates the importance of considering the international cross-boundary implications of conservation measures when trying to design effective MPA networks (e.g., (Miller *et al.* 2018; van Dover *et al.* 2018).
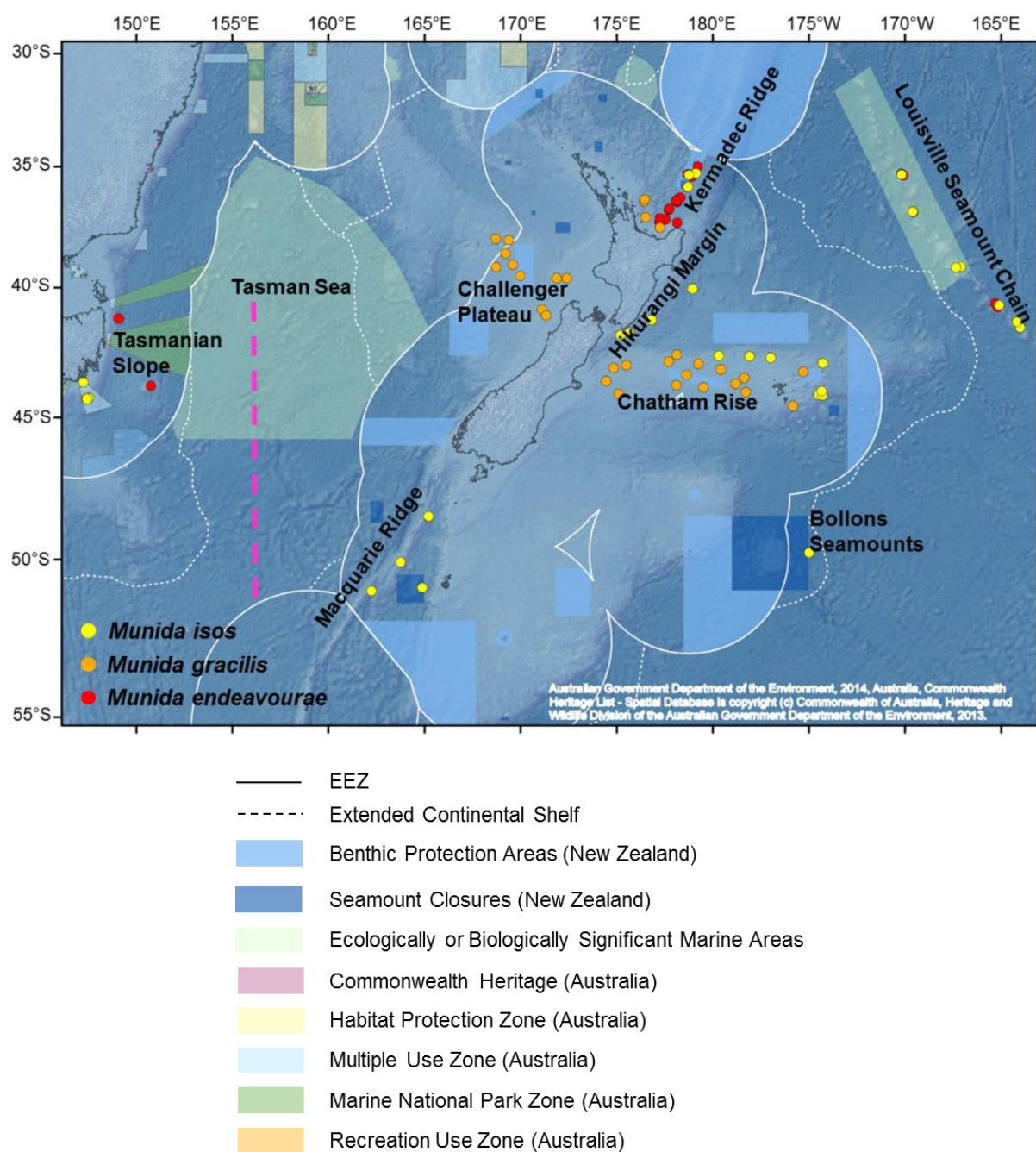
Figure 4.15 Map showing the distribution of marine protected areas, major geomorphic features and locations of the samples in the southwest Pacific Ocean region.

The dashed line represents the potential genetic break amongst populations of *M. isos* within the study area.

4.5 Conclusion

Overall, based on mitochondrial COI region and nuclear microsatellite markers, the present population genetic studies of the three *Munida* species (as VME-associated taxa) represent moderate to high level of genetic diversity but little genetic differentiation at different spatial scales. This lack of genetic subdivision pattern amongst squat lobster populations at different spatial scales within and beyond New Zealand's EEZ might be considered surprising given the coverage of the entire known range of three *Munida* species. Nonetheless, further genetic assessment of squat lobster populations, particularly the Tasmanian slope population which is recognised as being potentially genetically distinct population and as a potential source, can be expected with more extensive sampling and utilising more sensitive and higher-resolution markers (e.g., SNPs), to robustly detail the pattern of connectivity amongst the geographical distributional range of the three *Munida* species, and to contribute new information useful for assessing the effectiveness of existing protection measures and the design of new protected area networks in the study area.

Appendix

Table S4.1 Summary of collection sites and sample sizes (*N*, number of specimens) from the southwest Pacific Ocean from which the squat lobster, *M. isos*, was sampled for genomic analysis.

Geomorphic features	Code	Collection Registration Number	Collection Station Number	Sample Date	Latitude (S)	Longitude (positive = E, negative = W)	Depth (m)	COI	microsatellite
Kermadec Ridge	KR	NIWA 65036	TAN1007/123	08/06/2010	-35.284	178.863	1244–1276	2	2
Kermadec Ridge	KR	NIWA 72936	TAN1104/124	19/03/2011	-35.857	178.448	1237–1460	8	8
Kermadec Ridge	KR	NIWA 85237	TAN1104/33	07/03/2011	-35.360	178.506	1290–1395	3	3
Louisville Seamount Chain	LS	NIWA 94100	TAN1402/15	09/02/2014	-35.337	-170.443	1075–1100	1	1
Louisville Seamount Chain	LS	NIWA 94257	TAN1402/57	15/02/2014	-36.908	-169.846	1010–1013	10	10
Louisville Seamount Chain	LS	NIWA 94395	TAN1402/91	20/02/2014	-39.164	-167.350	910–934	12	12
Louisville Seamount Chain	LS	NIWA 94404	TAN1402/97	21/02/2014	-39.196	-167.590	1082–1090	2	2
Louisville Seamount Chain	LS	NIWA 94414	TAN1402/97	21/02/2014	-39.196	-167.590	1082–1090	1	1
Louisville Seamount Chain	LS	NIWA 94437	TAN1402/111	23/02/2014	-40.717	-165.343	814–912	6	7
Louisville Seamount Chain	LS	NIWA 94457	TAN1402/112	23/02/2014	-40.718	-165.343	768–913	2	2
Louisville Seamount Chain	LS	NIWA 94541	TAN1402/137	28/02/2014	-41.566	-164.261	1060–1061	20	20
Louisville Seamount Chain	LS	NIWA 94560	TAN1402/138	28/02/2014	-41.582	-164.255	1223–1241	8	10
Louisville Seamount Chain	LS	NIWA 94609	TAN1402/156	02/03/2014	-41.365	-164.419	1220–1250	20	20
Hikurangi Margin	HM	NIWA 31950	SO191-3/316	18/03/2007	-41.291	176.551	728–762	2	2
Hikurangi Margin	HM	NIWA 32003	SO191-2/138	18/02/2007	-41.780	175.401	1043–1059	6	6
Hikurangi Margin	HM	NIWA 32016	SO191-3/218	04/03/2007	-40.055	178.682	1111–1120	2	2
Hikurangi Margin	HM	NIWA 64201	TAN1004/134	27/04/2010	-41.896	174.927	913–948	1	1
Chatham Rise	CR	NIWA 33689	TAN0705/189	19/04/2007	-42.705	-178.340	997–1042	3	3
Chatham Rise	CR	NIWA 33690	TAN0705/189	19/04/2007	-42.705	-178.340	997–1042	2	2
Chatham Rise	CR	NIWA 53012	TAN0905/6	14/06/2009	-42.674	-179.955	150–1030	30	30
Chatham Rise	CR	NIWA 53802	TAN0905/104	26/06/2009	-44.158	-174.562	681–913	13	13
Chatham Rise	CR	NIWA 54079	TAN0905/113	27/06/2009	-44.150	-174.757	519–609	10	11
Chatham Rise	CR	NIWA 54178	TAN0905/116	27/06/2009	-44.175	-174.552	716–745	30	30

Geomorphic features	Code	Collection Registration Number	Collection Station Number	Sample Date	Latitude (S)	Longitude (positive = E, negative = W)	Depth (m)	COI	microsatellite
Chatham Rise	CR	NIWA 54367	TAN0905/121	28/06/2009	-44.028	-174.591	801–823	18	19
Chatham Rise	CR	NIWA 62805	TRIP2862/113	07/06/2009	-42.767	-177.252	965–1018	1	1
Chatham Rise	CR	NIWA 65961	TRIP3126/11	11/05/2010	-42.967	-174.533	1054	1	1
Macquarie Ridge	MR	NIWA 39574	TAN0803/17	30/03/2008	-48.549	164.963	1318–1327	4	4
Macquarie Ridge	MR	NIWA 39597	TAN0803/19	30/03/2008	-48.532	164.948	1060–1112	4	4
Macquarie Ridge	MR	NIWA 39762	TAN0803/33	01/04/2008	-50.091	163.482	1077–1408	22	22
Macquarie Ridge	MR	NIWA 39983	TAN0803/52	04/04/2008	-51.043	161.981	506–560	5	5
Macquarie Ridge	MR	NIWA 46520	TAN0306/6	14/04/2003	-50.943	164.609	1105–1140	24	25
Bollons Seamount	BP	NIWA 115316	TAN0307/83	02/05/2003	-49.768	-175.242	1261–1278	3	3
Tasmanian slope	TA	NMV J68478	SS01/2008 25	17/01/2008	-44.277	147.266	900	12	12
Tasmanian slope	TA	NMV J68560	SS01/2008 11	14/01/2008	-44.280	147.138	1260	13	13
Tasmanian slope	TA	NMV J67888	SS02/2007 14	01/04/2007	-44.321	147.178	1150–1280	16	16
Tasmanian slope	TA	NMV J67902	SS02/2007 03	29/03/2007	-43.692	146.976	100–110	31	32

Table S4.2 Summary of collection sites and sample sizes (*N*, number of specimens) from the southwest Pacific Ocean from which the squat lobster, *M. endeavourae*, was sampled for genomic analysis.

Geomorphic features	Code	Collection Registration Number	Collection Station Number	Sample Date	Latitude (S)	Longitude (positive = E, negative = W)	Depth (m)	COI	micros atellite
Kermadec Ridge	KR	NIWA 8993	TAN0413/20	08/11/2004	-36.797	177.469	1020–1183	1	1
Kermadec Ridge	KR	NIWA 8994	TAN0413/21	08/11/2004	-36.828	177.448	1118–1400	1	1
Kermadec Ridge	KR	NIWA 9792	TAN0413/54	11/11/2004	-37.237	177.277	1126–1134	1	2
Kermadec Ridge	KR	NIWA 18583	TAN0205/107	29/04/2002	-35.000	178.957	1255–1400	1	1
Kermadec Ridge	KR	NIWA 24573	TAN1206/97	24/04/2012	-36.455	177.838	920–950	1	1
Kermadec Ridge	KR	NIWA 24575	TAN1206/164	30/04/2012	-37.181	176.983	998–1000	1	1
Kermadec Ridge	KR	NIWA 64715	TAN1007/101	06/06/2010	-35.427	178.655	1028–1082	1	1
Kermadec Ridge	KR	NIWA 64817	TAN1007/106	06/06/2010	-35.350	178.514	1382–1416	1	1
Kermadec Ridge	KR	NIWA 72079	TAN1104/9	03/03/2011	-36.504	177.877	1576–1583	2	2
Kermadec Ridge	KR	NIWA 72135	TAN1104/11	03/03/2011	-36.459	177.847	1101–1172	2	2
Kermadec Ridge	KR	NIWA 72163	TAN1104/13	03/03/2011	-36.449	177.840	835–877	1	1
Kermadec Ridge	KR	NIWA 72282	TAN1104/20	03/03/2011	-36.484	177.887	1272–1328	1	1
Kermadec Ridge	KR	NIWA 72316	TAN1104/31	07/03/2011	-35.354	178.525	1150–1400	1	1
Kermadec Ridge	KR	NIWA 72356	TAN1104/33	07/03/2011	-35.360	178.506	1290–1395	1	1
Kermadec Ridge	KR	NIWA 72417	TAN1104/39	08/03/2011	-35.334	178.429	1520–1550		1
Kermadec Ridge	KR	NIWA 72440	TAN1104/48	09/03/2011	-35.348	178.543	1180–1306	1	1
Kermadec Ridge	KR	NIWA 72868	TAN1104/122	19/03/2011	-35.870	178.439	1235–1485	1	1
Kermadec Ridge	KR	NIWA 82145	TAN1206/21	17/04/2012	-36.331	178.026	1124–1127		1
Kermadec Ridge	KR	NIWA 82186	TAN1206/27	17/04/2012	-36.323	178.039	996–1035	1	1
Kermadec Ridge	KR	NIWA 82206	TAN1206/29	17/04/2012	-36.320	178.047	1126–1148	1	1
Kermadec Ridge	KR	NIWA 82686	TAN1206/77	22/04/2012	-36.811	177.465	878–911	1	1
Kermadec Ridge	KR	NIWA 82952	TAN1206/101	24/04/2012	-36.790	177.456	1150–1210	1	1
Kermadec Ridge	KR	NIWA 83045	TAN1206/120	26/04/2012	-37.335	176.960	1078–1083	2	2
Kermadec Ridge	KR	NIWA 83329	TAN1206/166	30/04/2012	-37.184	176.984	928–928	1	2
Kermadec Ridge	KR	NIWA 85225	TAN1206/34	18/04/2012	-36.447	177.839	850–980	1	1
Kermadec Ridge	KR	NIWA 82567	TAN1206/68	21/04/2012	-37.365	177.879	1229–1250	12	12
Louisville Seamount Chain	LS	NIWA 94045	TAN1402/10	08/02/2014	-35.368	-170.346	1175–1280	2	2
Louisville Seamount Chain	LS	NIWA 94191	TAN1402/31	11/02/2014	-35.317	-170.452	1205–1600	1	1
Louisville Seamount Chain	LS	NIWA 94274	TAN1402/59	15/02/2014	-36.924	-169.836	1147–1400	3	3
Louisville Seamount Chain	LS	NIWA 94475	TAN1402/124	25/02/2014	-40.778	-165.419	1230–1260	1	1
Louisville Seamount Chain	LS	NIWA 94505	TAN1402/134	26/02/2014	-40.639	-165.553	1370–1448	1	1
Louisville Seamount Chain	LS	NIWA 96367	TAN1402/14	09/02/2014	-35.326	-170.439	1137–1154	2	2
Louisville Seamount Chain	LS	NIWA 96465	TAN1402/111	23/02/2014	-40.717	-165.343	814–912	1	1

Geomorphic features	Code	Collection Registration Number	Collection Station Number	Sample Date	Latitude (S)	Longitude (positive = E, negative = W)	Depth (m)	COI	micros atellite
Louisville Seamount Chain	LS	NIWA 96467	TAN1402/57	15/02/2014	-36.908	-169.846	1010–1013	1	1
Louisville Seamount Chain	LS	NIWA 96514	TAN1402/137	28/02/2014	-41.566	-164.261	1060–1061	11	11
Louisville Seamount Chain	LS	NIWA 96515	TAN1402/156	02/03/2014	-41.365	-164.419	1220–1250	5	5
Tasmanian slope	TA	NMV J59291	TN228 J2-389-004	01/01/2009	-41.239	148.821	1374	1	2
Tasmanian slope	TA	NMV J59309	TN228 J2-389-002	01/01/2009	-41.240	148.821	1356	1	1
Tasmanian slope	TA	NMV J59310	TN228 J2-389-006	01/01/2009	-41.239	148.822	1309	2	2
Tasmanian slope	TA	NMV J59312	TN228 J2-390-015	04/01/2009	-43.828	150.500	1061	2	2
Tasmanian slope	TA	NMV J68481	SS01/2008 25	17/01/2008	-44.277	147.266	900	5	6
Tasmanian slope	TA	NMV J68559	SS01/2008 11	14/01/2008	-44.280	147.138	1260	12	13

Table S4.3 Summary of collection sites and sample sizes (*N*, number of specimens) from the southwest Pacific Ocean from which the squat lobster, *M. gracilis*, was sampled for genomic analysis.

Geomorphic features	Co de	Collection Registration Number	Collection Station Number	Sample Date	Latitude (S)	Longitude (positive = E, negative = W)	Depth (m)	COI	microsatellite
Kermadec Ridge	KR	NIWA 9884	TAN0413/119	13/11/2004	-37.560	176.985	313–314	5	6
Kermadec Ridge	KR	NIWA 9819	TAN0413/119	13/11/2004	-37.560	176.985	313–314	1	1
Kermadec Ridge	KR	NIWA 85529	KAH1205/70	24/03/2012	-37.149	176.242	351–353	3	3
Kermadec Ridge	KR	NIWA 85527	KAH1205/63	22/03/2012	-36.473	176.172	411–410	2	2
Kermadec Ridge	KR	NIWA 85538	KAH1205/64	22/03/2012	-36.395	176.202	460–465	8	8
Chatham Rise	CR	NIWA 33660	TAN0705/25	04/04/2007	-44.131	174.844	516–518	4	4
Chatham Rise	CR	NIWA 33669	TAN0705/111	11/04/2007	-44.576	-176.079	415–265	20	20
Chatham Rise	CR	NIWA 33677	TAN0705/138	14/04/2007	-43.291	-175.562	638–644	2	2
Chatham Rise	CR	NIWA 33683	TAN0705/179	18/04/2007	-43.520	-178.619	423–425	16	16
Chatham Rise	CR	NIWA 33685	TAN0705/181	18/04/2007	-44.068	-178.549	451–452	20	20
Chatham Rise	CR	NIWA 33698	TAN0705/252	24/04/2007	-42.993	178.999	534–539	12	13
Chatham Rise	CR	NIWA 33704	TAN0705/260	25/04/2007	-42.645	177.869	1211–1215	1	1
Chatham Rise	CR	NIWA 33705	TAN0705/281	27/04/2007	-43.037	175.238	309–310	3	3
Chatham Rise	CR	NIWA 53604	TAN0905/97	26/06/2009	-44.147	-174.690	440–600	8	8
Chatham Rise	CR	NIWA 54278	TAN0905/119	28/06/2009	-44.158	-174.555	487–616	11	11
Chatham Rise	CR	NIWA 78110	TAN1116/3	03/11/2011	-43.145	174.589	530–530	2	2
Chatham Rise	CR	NIWA 78242	TAN1116/19	05/11/2011	-43.639	174.199	498–497	2	2
Chatham Rise	CR	NIWA 78645	TAN1116/78	12/11/2011	-43.876	179.253	502–500	4	4
Chatham Rise	CR	NIWA 78695	TAN1116/84	13/11/2011	-43.793	177.854	491–490	4	4
Chatham Rise	CR	NIWA 78966	TAN1116/103	15/11/2011	-42.904	177.454	414–414	2	2
Chatham Rise	CR	NIWA 79157	TAN1116/127	18/11/2011	-43.411	178.377	349–347	6	6
Chatham Rise	CR	NIWA 85652	TAN1208/70	26/06/2012	-43.753	-179.077	397–397	17	17
Chatham Rise	CR	NIWA 85790	TAN1208/36	19/06/2012	-43.202	-179.841	518–516	20	20
Challenger Plateau	CP	NIWA 33712	TAN0707/30	29/05/2007	-38.623	168.947	479–482	11	11
Challenger Plateau	CP	NIWA 33715	TAN0707/34	29/05/2007	-38.024	168.447	570–575	30	30
Challenger Plateau	CP	NIWA 33722	TAN0707/80	02/06/2007	-39.183	168.469	554–557	23	27
Challenger Plateau	CP	NIWA 33727	TAN0707/86	03/06/2007	-38.083	169.115	565–570	13	13
Challenger Plateau	CP	NIWA 33731	TAN0707/91	04/06/2007	-39.085	169.342	523–526	17	17

Geomorphic features	Code	Collection Registration Number	Collection Station Number	Sample Date	Latitude (S)	Longitude (positive = E, negative = W)	Depth (m)	COI	microsatellite
Challenger Plateau	CP	NIWA 33735	TAN0707/93	04/06/2007	-39.544	169.715	634–636	30	30
Challenger Plateau	CP	NIWA 33739	TAN0707/120	06/06/2007	-40.875	170.846	534–536	29	30
Challenger Plateau	CP	NIWA 33741	TAN0707/126	06/06/2007	-41.112	171.071	288–298	3	3
Challenger Plateau	CP	NIWA 33742	TAN0707/130	07/06/2007	-39.637	171.624	473–491	1	1
Challenger Plateau	CP	NIWA 33744	TAN0707/140	07/06/2007	-39.642	172.148	266–267	2	2

Chapter 5. Population genetics assessment of *Munida isos* in the southwest Pacific Ocean: the application of SNPs

Abstract

Advances in High-Throughput Sequencing (HTS) technologies make Genotyping by Sequencing (GBS) feasible for large genome non-model species. To facilitate genetics research in deep-sea squat lobster populations and to provide information for the design of management strategies, GBS was conducted for the vulnerable marine ecosystem (VME) associated taxon *Munida isos* from the New Zealand and part of the Australian exclusive economic zones (EEZ) utilising a *PstI-MspI* double restriction-enzyme digestion. After stringent filtering in a *de novo* bioinformatics pipeline (Universal Network Enabled Analysis Kit (UNEAK)), a total of 32,681 single nucleotide polymorphisms (SNPs) were retained with a sample depth of 1.82. Relatedness and population genetic differentiation were assessed among *M. isos* populations, using two methods: Genomic Relationship Matrix (GRM) analysis and Principle Component Analysis (PCA). Evidence of genetic discontinuity was observed amongst regions of the Tasmanian slope, Macquarie Ridge and the eastern New Zealand region, whilst no genetic differentiation was detected within the eastern New Zealand region. These genetic patterns amongst squat lobster populations may be attributed to local hydrodynamic regimes and life-history strategy (putative dispersal patterns). Nonetheless, genetic data were relevant to inform management of VMEs at a variety of spatial scales. Genetically distinct regions (Tasmanian slope in the Australian EEZ and Macquarie Ridge in New Zealand's EEZ) need to be managed separately. Squat lobsters from populations in eastern New Zealand's EEZ region that have high levels of gene flow highlight the importance of recognising and protecting genetic diversity and routes of gene flow between/amongst populations. This study presents the first attempt to use genome-wide SNP analysis in squat lobsters and will facilitate population genomics studies for the VME-associated species *M. isos* in the future. The results provide information to aid the New Zealand Government's development of conservation and management strategies for VMEs.

5.1 Introduction

Recent advances in High-Throughput Sequencing (HTS) technologies have enabled genome assembly of an increasing number of species and made a tremendous impact on many areas of biology, providing unprecedented discovery and characterisation of molecular markers (O'Leary *et al.* 2018). The revolution in genetic marker discovery relying on HTS has facilitated research in population genetics, genome-wide association studies (GWAS), phylogeography, quantitative trait loci mapping and selective breeding for any organism, and genome-wide studies on wild populations, all of which will contribute to conservation genetics and ecology (Luikart *et al.* 2003; Davey *et al.* 2011; McCormack *et al.* 2013). As a third-generation molecular marker, single nucleotide polymorphisms (SNPs) have become the marker of choice for many reasons, such as the high density of annotated markers and the associated ability to assemble combined datasets from multiple resources (Brumfield *et al.* 2003; Morin *et al.* 2004; Helyar *et al.* 2011). HTS offers multiple approaches that simultaneously enable genome-wide SNPs discovery and genotyping in a single step, even in organisms for which we have limited genetic information (Davey *et al.* 2011; Helyar *et al.* 2011). The distributions and frequencies of SNP markers are key to evaluating diversity amongst closely related populations and species (Tautz *et al.* 2010), something that can now be achieved with unprecedented speed, accuracy and low cost.

Genomic data can be sourced from a variety of available sequencing platforms, including whole genome, transcriptome, targeted gene capture, and reduced representative libraries. The selection of approach mainly depends on the availability of existing genomic information, the study organism's genomic characteristics, the questions being addressed and funding (Ahrens *et al.* 2018). In recent years, reduced-representation sequencing has emerged as an ideal choice for genetic studies due to its relatively low cost, and the fact that many research questions can be answered with a smaller set of markers (e.g., SNPs). Notably, for non-model organisms, a genome reduction step is needed to acquire deep assemblage of redundant contigs for SNP calling (Slate *et al.* 2009; Tautz *et al.* 2010).

The use of genome complexity reduction with restriction enzymes and multiplex HTS technologies for high-density SNP discovery and genotyping was initially demonstrated through restriction site-associated DNA sequencing (RADseq) by Baird *et al.* (2008). Cost benefits and increased efficiency were achieved by combining a multiplex sequencing strategy that applied an inexpensive barcoding system. Three years after the study by Baird *et al.* (2008),

Elshire *et al.* (2011) proposed a less complicated and even more cost-effective genotyping procedure (i.e., genotyping by sequencing or GBS), based on HTS technologies. GBS approaches are simple, fast, affordable, extremely specific and highly robust when based on parallel high-throughput sequencing of genomic subsets targeted by restriction enzymes. Both RADseq and GBS approaches couple HTS technologies with genome complexity reduction techniques to simultaneously identify and genotype SNPs, yet GBS library development is greatly simplified. In addition to its simplicity of less DNA required, no fragment size selection, and fewer enzymatic and purification steps required, GBS is cost and time efficient through its completion of only two steps for enzyme digestion and adapter ligation on plates prior to PCR amplification of the pooled library. Another advantage of GBS is that it may cover important genome regions that are inaccessible to DNA target-enrichment approaches (e.g., Mamanova *et al.* (2010). Moreover, low per sample cost can be achieved by multiplexing samples simultaneously through the use of barcodes ligated to each sample prior to the sequencing (Huang *et al.* 2014). The cost will also be affected by the choice of single-end or paired-end sequencing on any HTS platform. Nonetheless, one drawback of GBS is that there is a possibility that only one of the two alleles can be seen, leading to an overestimation of homozygotes. Also, GBS lacks specificity to the genic regions of the genome, resulting in a significant proportion of sequences originating from repetitive and non-informative regions. As such, if the coverage of the sequenced region is low, there will be a significant loss of genotypes. However, this scenario can be avoided, and lower copy regions can be targeted with up to threefold higher efficiency by choosing appropriate restriction enzymes using GBS (Elshire *et al.* 2011). Enzyme selection in GBS library preparation is critical, and depends on the number of markers required and the desired level of multiplexing. The original GBS protocol only applies a single restriction enzyme for genome complexity reduction, but Poland *et al.* (2012b) extended the GBS approach to a two-enzyme system that includes a "rare-cutter" and a second "common-cutter". This two-enzyme approach greatly simplifies quantification by generating a suitable and uniform complexity reduction of the library prior to sequencing with a range of different restriction enzymes to produce a higher or lower complexity reduction of the genome. Last but not least, the term "genotyping-by-sequencing" (GBS) generically refers to any reduced-representation genotyping method (Davey *et al.* 2011).

The success of GBS library preparation is directly determined by the initial quality of DNA samples and the genome size of the studied organism (Elshire *et al.* 2011). The performance of GBS in very large genome organisms (e.g., non-model marine invertebrates)

with high genetic diversity can be very challenging due to their genome complexity, which may result in a high proportion of non-homologous DNA fragments with lower coverage. For large genome-size organisms, different sets of restriction enzymes or enzyme combinations can be selected to optimise the number of DNA fragments, capturing a certain proportion of the target genome. If possible, additional sequencing is recommended to reduce missingness and increase sequencing depth in large-genome taxa (Chen *et al.* 2013). To date, only three genome size records for squat lobsters are available in the Animal Genome Size Database (Gregory 2018), with the largest 1C genome yet reported for the vent squat lobster *Munidopsis recta* Baba, 2005 (15.56 pg), and two *Galathea* species (*G. squamifera* Leach, 1814 and *G. strigosa* (Linnaeus, 1761)) of 8.46 and 6.99, respectively (Bonnivard *et al.* 2009). The target species here, *Munida isos* Ahyong & Poore, 2004 belongs to a different family of squat lobster (the Munididae). However, it might reasonably be expected to have a similarly large genome since it belongs in the same superfamily Galatheoidea.

Whatever library preparation approach is chosen, bioinformatics must be applied to filter out useful information on genotypes from the vast amount of raw sequence reads (Davey *et al.* 2011; McCormack *et al.* 2013). GBS requires intense bioinformatics analyses for SNPs calling and data filtering, which can be complex and challenging (Torkamaneh *et al.* 2016). A GBS pipeline must include steps to filter out poor-quality reads, classify reads based on barcodes, either align reads to the reference genome or identify alleles and loci *de novo* for polymorphism discovery, and score genotypes for every individual (Torkamaneh *et al.* 2016). SNP calling can be done using a variety of bioinformatics pipelines including reference-based pipelines (e.g., TASSEL-GBS v2, Stacks, see below) when a reference genome is available and *de novo*-based pipelines (e.g., UNEAK, Stacks, see below) when a reference genome is lacking. When a fully sequenced or a reference genome is available, SNPs discovery can be easily performed for whole-genome re-sequencing projects by aligning reads to the physical map, minimising the impact of short read length and high error rates in HTS data (Nielsen *et al.* 2011). However, challenges arise when a reference genome sequence is not available. For a study of wild populations in the absence of a reference genome, many markers scored accurately in the majority of individuals are required to ensure population parameters are precisely estimated (Davey *et al.* 2011). The identification of high-density markers in species that lack a reference genome will subsequently facilitate the development of the reference genome sequences.

The most commonly applied pipelines for a *de novo*-based approach are UNEAK (Lu *et al.* 2013) and Stacks (Catchen *et al.* 2013). The general design of UNEAK including three steps: 1) reads trimmed to 64 bp, 2) tags collapsed from identical 64-bp reads, and 3) SNPs identified from tag pairs with 1 bp mismatch. The UNEAK pipeline enforces a strictly bi-allelic criterion for the detection of putative SNPs, in line with the evidence that most SNPs are bi-allelic (Krawczak 1999; Morin *et al.* 2004), yet tri- or tetra- allelic SNPs are common in many non-model marine species as well (Gardner *et al.* 2016; Larraín *et al.* 2018). Complexities that hinder SNPs discovery, such as a mixture of repeats, paralogs and sequencing errors, are easily resolved with UNEAK. Additionally, a "network filter" is designed for basic data filtering, generating a collection of reciprocal SNPs. Stacks can identify SNPs via alignment against a reference genome or *de novo* from non-model organisms. The pipeline is described as follows: 1) reads demultiplexed and cleaned, 2) polymorphic nucleotide sites identified, 3) loci grouped together, 4) allelic state determined in each individual at each locus, and 5) allelic data output into population genetic statistics or mappable genotypes. With regard to the number of SNPs, UNEAK called approximately 2-fold more SNPs than Stacks and was ~3 times faster. Among which, ~67% of SNPs called by Stacks were also detected in UNEAK data (Torkamaneh *et al.* 2016). Although with different algorithms, both pipelines are relatively conservative concerning SNP calling, resulting in high-quality datasets. However, there appears to be not much difference between the two pipelines with regard to the proportion of missing data and genotypic accuracy. Thus, for deep-sea marine invertebrates that lack a reference genome (e.g., squat lobsters), both *de novo* pipelines can perform extremely well in terms of accuracy, yet UNEAK can yield nearly two-folds more SNPs (Oetjen and Reusch 2007). The "HapMap count" file that generated by the UNEAK pipeline can be directly applied in the Kinship (genetic relatedness) using GBS with Depth adjustment (KGD) method, which is an unbiased estimation of relatedness and genetic differentiation for GBS data (Dodds *et al.* 2015). One advantage of the KGD is that it has a SNP calling step that can further filter for excess heterozygosity and very low MAF from the SNPs data, which may improve the estimation of the downstream analyses. Besides, an optimal depth of 1–5 is suggested for between-individual relationship evaluation, making it more feasible if the total sequencing effort is fixed, especially for organisms with a large genome (a trade-off between coverage depth and density of SNPs) (Dodds *et al.* 2015).

In the marine environment, patterns of gene flow can be influenced by oceanographic features such as currents and fronts that can facilitate or prevent diffusion and mixing of pelagic

larvae, as well as biogeographical boundaries (Wares *et al.* 2001; Weersing and Toonen 2009; White *et al.* 2010). For example, two adjacent sites may rarely exchange genetic information if located on different sides of a front (Gilg and Hilbish 2003), and two distant sites may be well connected driven by a strong current between them (Mitarai *et al.* 2009). Understanding the degree to which populations rely on self-recruitment or receive a subsidy from other populations through dispersal is of fundamental importance in ecology and evolution, and will aid in the evaluation of species' colonisation and persistence under contemporary anthropogenic stresses (Cowen and Sponaugle 2009; Adams *et al.* 2012; Clobert *et al.* 2012). Therefore, dispersal has a crucial impact upon management and conservation that can provide useful information in designing marine protected area (MPA) networks worldwide (Etter and Bower 2015; Ross *et al.* 2016). However, direct estimation of larval dispersal in the ocean can be very challenging, which requires the ability to track microscopic larvae of benthic invertebrates throughout the pelagic dispersal process (Levin 2006; Thorrold *et al.* 2007). Alternatively, genetic techniques provide an indirect means to evaluate dispersal and connectivity in marine ecosystems (Cowen and Sponaugle 2009; Lowe and Allendorf 2010).

Presently, utilising HTS technologies, marine population genetics stands at the forefront of evaluating genetic diversity and connectivity patterns at the genomic level (Luikart *et al.* 2003; Reitzel *et al.* 2013). HTS technologies provide means to increase the ease of genetic data acquisition (i.e., smaller sample sizes with an abundance of markers), and more importantly, the higher resolution genome-wide data allows fine-scale connectivity patterns to be evaluated (Taylor and Roterman 2017). Limitations of conducting population genetic studies in the deep sea have been mainly due to both the difficulty of obtaining a sufficient number of specimens and the paucity of molecular markers that yield results (Shank 2010; Baco *et al.* 2016). With the huge increase in abundance of SNPs offered by population genomics, 10-20 specimens might be enough to obtain useful genetic information of populations (Willing *et al.* 2012). Moreover, unlike microsatellite data that needs to be subjectively genotyped for every study, HTS data can be easily used in conjunction with HTS data from another study in building genome assemblies and/or larger genomic datasets (Taylor and Roterman 2017).

To date, however, relatively few population genomic studies have applied SNPs generated by HTS technologies to benthic marine species for connectivity pattern assessment (e.g., Van Wyngaarden *et al.* (2017); Al-Breiki *et al.* (2018); Silliman (in press)), and only rarely to deep-sea organisms (e.g., Everett *et al.* (2016); Xu *et al.* (2018)).

GBS is now starting to prove its efficiency in SNPs discovery in a variety of non-model marine invertebrates. For instance, Wells and Dale (2018) provided the first population genomic studies of a coastal marine species in New Zealand. The authors assessed the utility of GBS in examining gene flow patterns in an endemic marine isopod *Isocladus armatus*, and suggested that this technique can be an ideal alternative to traditional markers (e.g., microsatellites). Genetic variation of 4,000 SNPs revealed strong genetic subdivision between North and South Island populations, whilst no genetic disjunction was observed between two populations from North Island with only 8 km distance. Interestingly, two outlier SNPs represented strong association with the expression of a striped morph, as genetic basis support to colouration.

Up to now, only a handful of population genomic studies have been conducted in deep-sea organisms utilising SNPs. The first deep sea population genomic study was published by Everett *et al.* (2016). The authors utilised RAD-tag sequencing for the assessment of putative population structure of the deep-sea octocoral *Swiftia simplex* sampled from the Eastern Pacific Ocean along the west coast of the United States. However, no genetic subdivision was observed across all geographic areas, and no isolation by distance was detected, indicating a high level of connectivity and potential panmixia. Another example was conducted on a deep-sea mussel *Bathymodiolus platifrons* in the Northwest Pacific Ocean (Xu *et al.* 2018). Contrasting results were revealed in populations from four hydrothermal vents and two methane seeps by two types of molecular markers, with no genetic differentiation as revealed by three mitochondrial genes (i.e., atp6, cox1, and nad4) but pronounced genetic subdivision being detected based on SNPs generated by the type IIB restriction site-associated DNA (2b-RAD) approach.

The deep sea encompasses 95% of the ocean's volume and represents the largest and least explored biome of the Earth's Biosphere (Danovaro *et al.* 2017). Deep sea ecosystems harbour many different habitats with contrasting environmental conditions, including the functionally and ecologically important vulnerable marine ecosystems (VMEs). Therefore, whole and reduced representation genomes studies are needed, to achieve greater insights into the fine-scale connectivity patterns in the deep sea, with particular focus on both VME-indicator and VME-associated taxa, aiding in the establishment of informed management plans and the designation of new MPAs in New Zealand's EEZ.

Previously, mitochondrial COI region and nuclear microsatellite markers were employed for population genetic analyses of the deep-sea squat lobster *M. isos* Ahyong & Poore, 2004 (see Chapter 3). In both approaches, the marker types failed to detect pronounced

genetic structure amongst squat lobster populations. It was argued that *M. isos* might exhibit a high level of connectivity within the study area, owing to its larval dispersal and recruitment pattern (e.g., potentially long pelagic larval duration (PLD)) and the oceanographic features (e.g., currents). Nonetheless, more information can be expected to detail the potential genetic differentiation amongst squat lobster populations in the southwest Pacific Ocean based on a higher-resolution nuclear marker type, SNPs. With regard to the absence of a reference genome and limited sampling of *M. isos*, a larger number of SNPs with sufficient accuracy are needed, making UNEAK an ideal pipeline for the GBS library construction.

The objective of the study recorded in this chapter is: 1) to evaluate the effectiveness of the GBS approach and UNEAK SNP calling pipeline for a non-model marine organism, *M. isos*, 2) to estimate the connectivity patterns amongst populations of *M. isos* in the southwest Pacific Ocean; and 3) to provide information that will contribute to the management of VMEs within the southwest Pacific Ocean.

5.2 Methods

5.2.1 Sample collection and DNA extraction

Specimens were obtained from the National Institute of Water and Atmospheric Research (NIWA) Invertebrate Collection (Wellington, New Zealand) and Museum Victoria (Melbourne, Australia) from archived specimens collected by mainly scientific sampling expeditions in the southwest Pacific Ocean since the 2000s, considering that old museum specimens might be problematic due to DNA degradation. Two hundred and seventy-two specimens of *M. isos* collected from six geomorphic regions in the southwest Pacific Ocean, i.e., Kermadec Ridge, Louisville Seamount Chain, Hikurangi Margin, Chatham Rise, Macquarie Ridge and the Tasmanian slope, were used for the preparation of three GBS libraries with the following codes/names: SQ0503-8 (272 individuals), SQ0503-1 (272 individuals) and SQ0642-7 (48 individuals) (Figure 5.1; Table 5.1 and Table S5.1). As no genetic structure was detected amongst squat lobster populations of *M. isos* utilising the spatially explicit hierarchical testing framework (e.g., Northern-Southern biogeographical provinces, North-Central-South regions) by COI region and microsatellite variations (see Chapter 3), the individual geomorphic features were chosen as the testing framework for the SNPs work. Whole genomic DNA was extracted and examined as per Chapter 3. The Quant-iT Picogreen dsDNA Assay Kit (Life

Technologies, Carlsbad, United States) was then used to accurately quantify the DNA samples before being diluted to a working stock of 20 ng/µL. DNA extractions as described above were sent to AgResearch Ltd. (Invermay) for GBS processing (see subsequent sections)

The maps were generated using ArcGIS 10.4.1 software (Mercator projection and the World Geodetic System (WGS 1984) as the reference coordinate system). The base map bathymetry data is derived from the GEBCO_2008 Grid, version 20100927, www.gebco.net; the marine boundaries for the New Zealand region are based on NIWA data; those for the Australian region are credited to the Australian Government Department of the Environment, 2014 and Australia, Commonwealth Heritage List (Spatial Database is copyright (c) Commonwealth of Australia, Heritage and Wildlife Division of the Australian Government Department of the Environment, 2013).

Table 5.1 Summary of sample sizes (*N*, number of specimens) for *M. isos* from the southwest Pacific Ocean, for GBS library construction.

Geomorphic features	Code	SQ0503-8 / SQ0503-1	SQ0642-7
Kermadec Ridge	KR	10	9
Louisville Seamount Chain	LS	73	8
Hikurangi Margin	HM	5	5
Chatham Rise	CR	103	9
Macquarie Ridge	MR	30	9
Tasmanian slope	TA	51	8
Total		272	48
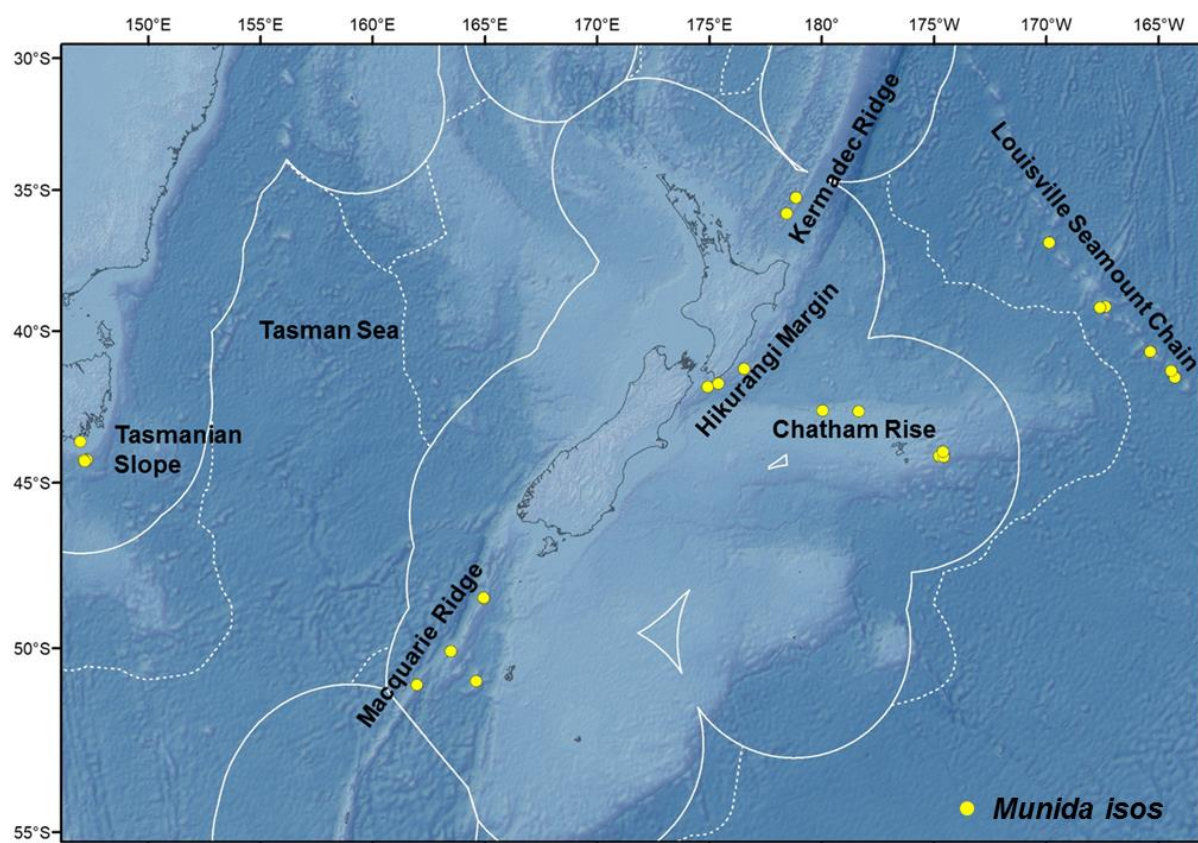
Figure 5.1 Map showing the locations of the samples for *M. isos*.

5.2.2 Enzyme selection and adapter titration

Prior to library preparation, several DNA samples (2 μL each) were pooled for an optimal enzyme or double-enzyme digest check and adapter titration. Selection of enzymes that leave 2~3 bp overhangs, and do not cut either too frequently or infrequently in the major repetitive fraction of the genome, is critical and extremely useful in promoting efficient adapter ligation to insert DNA. In this study, restriction digest buffer (*NEB Buffer 4*) and three restriction endonucleases were chosen: *ApeKI* (recognition at site G↓CWGC) and *PstI* (CTGCA↓G), both as rare-cutting restriction enzyme, and a third enzyme, *MspI* (C↓CGG), which has a more common recognition site. Four combinations of the three enzymes were evaluated for their effectiveness in genome complexity reduction and enrichment of genic SNPs. Genome sequences were digested into fragments at the point of the recognition site of the restriction enzyme(s). In double-digest samples (*ApeKI-MspI* and *PstI-MspI*), *ApeKI* was digested first in the absence of *MspI*, whereas *PstI* and *MspI* were digested together at the same time. Adapter titrations were then performed to minimise adapter dimers (i.e., barcodes ligating

to each other) for the optimal enzyme combination (i.e., *PstI-MspI*). Only *PstI* adapters were used for the titration because *PstI* and *MspI* adapters cannot ligate to each other due to their different cut sites. However, *PstI-MspI* was still checked in a single reaction to ensure the ongoing libraries had a good range of fragments. Details of temperature cycling programmes for digestion, ligation and PCR are provided in the next section "GBS Library Construction". Laboratory work described in this section was conducted by AgResearch Ltd.

5.2.3 GBS library construction and sequencing

Genotyping was performed using the GBS method with the double restriction enzyme combination of *PstI-MspI* according to the protocol described by Elshire *et al.* (2011) and Poland *et al.* (2012b) with minor modification. A single-end (1 × 100 bp reads) GBS library (SQ0503-8) was constructed in 272-plex from the genomic DNA samples of *M. isos* from six geomorphic regions in the southwest Pacific Ocean, using an Illumina HiSeq 2500 Genome Analyser at AgResearch Animal Genomics laboratory (Invermay, New Zealand) in August 2017. Re-sequencing was conducted with SQ0503-1 (272-plex) in September 2017 and SQ0642-7 (48-plex: a subset of samples from six localities) in March 2018 (Table 5.1). *PstI-MspI* barcodes were plated out at a concentration of 8 μL (i.e., 4.8 ng per well) before drying down for 25 min with medium heat. 100 ng of squat lobster DNA samples were then added to the matching well position on the barcode plate followed by 25 min with medium heat to dry down. Double-enzyme *PstI-MspI* digestion (Table 5.2) was performed in PCR Thermal Cycler (Applied Biosystems) programmed as follows: 37 °C for 2 h, followed by 65 °C for 30 min, and then held at 4 °C. Adapter ligation of Adapter 1 (unique for each sample) and Adapter 2 (common adapter) (Table 5.3) was run as follows: 22 °C for 1 h, followed by 65 °C for 30 min, and then held at 4 °C. Five (5) μL of each sample were pooled into a single Eppendorf tube, purified using a QIAquick PCR Purification Kit (Qiagen, Valencia, CA, US) and re-suspended in 100 μL of EB Buffer. Four PCR tubes were set up with 4 μL of the elution in each tube and combined with 46 μL of the PCR cocktail (Table 5.4). Temperature cycling consisted of 72 °C for 5 min, then 98 °C for 30 s, followed by 18 cycles of denaturation at 98°C for 30 s, annealing at 65 °C for 10 s, extension at 72 °C for 30 s, and a final *Taq* extension at 72 °C for 5 min, before holding at 4 °C. All four samples were pooled together and purified using a QIAquick PCR Purification Kit (Qiagen, Valencia, CA, US), and re-suspended in 35 μL of EB Buffer. Thirty (30) μL of the amplified library was mixed with 10 μL of Buffer L and run on Pippin

Prep (SAGE Science, Beverly, MA, US) to capture library fragments without adapter dimers that were between the sizes of 193–318 bp (adapter included). Both the pre-pippin and post-pippin libraries were measured on a 2100 BioAnalyser High Sensitivity DNA kit (Agilent Technologies, Santa Clara, CA, US) for the range of digested DNA fragments and the average size of the libraries. The concentration was further determined by quantifying 1 μL of the library on a Qubit Fluorometer (Life Technologies, Carlsbad, CA, US). Once the appropriate quantity of adapters was empirically determined, single-end sequencing libraries (SQ0503-8, SQ0503-1, SQ0642-7) were performed in three lanes of separate flow cells on an Illumina HiSeq 2500 Genome Analyser using a single-end 100 bp run with version 4 chemistry at AgResearch Animal Genomics laboratory (Invermay, New Zealand). Laboratory work described in this section was conducted by AgResearch Ltd.

Table 5.2 Double-enzyme *PstI-MspI* digestion for GBS library preparation.

	1 x reaction (μL)	Cocktail (μL)
NEB Buffer 4	1.5	450
PstI	1	300
MspI	1	300
ddH$_2$O	11.5	3450
Total	15	4500

Table 5.3 Adapter ligation for GBS library preparation.

	1 x reaction (μL)	Cocktail (μL)
NEB Buffer 4	2.5	750
T4 DNA ligase	1.6	480
ddH$_2$O	5.9	1770
Total	10	3000

Table 5.4 PCR protocol for GBS library construction.

	1 x reaction (μL)	Cocktail (μL)
PCR Primer 1	1	5
PCR Primer 2	1	5
Taq	25	125
ddH$_2$O	19	95
Total	46	230

5.2.4 Sequencing reads and taxonomy check

The fastq file generated by the Illumina HiSeq 2500 Genome Analyser combines the base calls for the sequences that were generated with the encoded quality value for each base, showing the confidence level of the sequencing. Quality control was conducted for raw fastq files generated from three libraries (SQ0503-8, SQ0503-1, SQ0642-7) using FastQC v0.11.7 (Andrews 2010). The taxonomic identification was verified and to rule out contamination of the sample in a random subsample of 10,000 reads for each lane by genetic screening through BLAST searches against the Nucleotide database in GenBank (Benson *et al.* 2018). This work was carried out by AgResearch Ltd in consultation with me. Work described subsequently was carried out by me, initially at AgResearch and in consultation with their staff, then subsequently by me in Wellington, again in consultation with AgResearch staff when necessary.

5.2.5 SNPs calling and initial filtering

To ensure that the SNPs were of high quality, stringent filtering was performed in UNEAK for *de novo* clustering of sequences by trimming adapters and discarding low-quality reads. UNEAK (Lu *et al.* 2013) is a non-reference GBS SNP calling pipeline, which is an extension of the Java program of TASSEL v3.0.174 (Bradbury *et al.* 2007; Glaubitz *et al.* 2014), running as TASSEL v3.0 plugins via command line. SNP calling was applied globally to all available sequencing data (SQ0503-8, SQ0503-1, SQ0642-7) and the "merged library" (i.e., combining sequencing data from three libraries into a single analysis). UNEAK searched for reads that begin with a barcode followed by the expected cut site remnant (CTGCA for *PstI* and CGG for *MspI*) and trimmed them to 64 bases (keeping the cut site remnant but removing the barcode) because the sequencing errors are normally enriched at the ends of the reads. If a read contained either a full cut site (CTGCAG for *PstI* and CCGG for *MspI*) or the beginning of the common adapter within the first 64 bases, it was truncated appropriately and filled to 64 bases with polyA. Sequences containing N were rejected. The 64-base sequence was defined as a "tag". Tag pairs were then identified, and those with only 1 bp mismatch were considered as candidate SNPs. The criteria applied to filter out low-quality data were on a per-SNP basis. Error tolerance rate (ETR) was set at 0.03, slightly above the Illumina sequencing error rate (0.1%). Minimum allele frequency (MAF) was defined as the frequency of the less common allele at a variable site. In this study, minimum and maximum MAF parameters were set at 0.03 and 0.5, respectively. Call rate (CR) denotes a proportion of tags with how many

individuals are covered by at least one tag. The CR was set to a minimum of 0.1 and a maximum of 1. The minimum CR of 0.1 indicates that a particular SNP which covered at least 10% of the individuals would be retained.

5.2.6 SNPs filtering and relatedness analysis

The second phase of SNP filtering was performed using KGD (Dodds *et al.* 2015) in the program R v3.4.1 (R Development Core Team, 2017). The R code is available at https://github.com/AgResearch/KGD. The KGD method is an unbiased estimation of relatedness for GBS data (SNPs with genotype calls through all individuals), using theoretical calculation and simulation. G5, a method with correction for sample depth, was referred to in self-relatedness estimation because it gave results that were very close to true values in relationship evaluation (Dodds *et al.* 2015). A "Fin plot", named from the shape of its boundaries, is a simple graphical method that illustrates Hardy-Weinberg Disequilibrium (HWD, i.e., the proportion of reference allele homozygotes minus its expected proportion) against MAF, using a colour gradient to illustrate SNP average depth. Genomic Relationship Matrices (GRM) and Principle Component Analysis (PCA) plots were constructed as follows: GRM was depicted in the form of a heat map with relatedness between individuals on the off-diagonal and self-relatedness on the diagonal, which was achieved by using the observed calls for an individual and the shared calls between individuals for the estimation of self-relatedness and individual to individual relatedness. PCA, a powerful dimensionality reduction technique, has been long used to assess population differentiation without *a priori* specification of subpopulations that distil the complex multi-locus datasets into two or more dimensional scatter plots (Jackson 1991; Novembre and Stephens 2008). Initially, samples with maximum read depth = 1 and/or mean read depth < 0.01 were discarded. All SNPs were additionally filtered to discard those with MAF = 0 or depth < 0.01, with no data through all six regions, with excess heterozygosity (bottom right corner of the fin plot, HWD < -0.05). Subsequently, retained SNPs were filtered for four libraries to test for the extent of SNP-specific genetic differentiation using an index that is equivalent to the traditional F_{ST} index of population differentiation (i.e., $r\chi^2/n(r-1)$, r - the number of populations, n - the number of genotypes (2 × number of individuals for diploid data)). This gave a significant ($p < 0.05$) value of 0.025 for 269-plex merged library (two samples dropped) and 271-plex SQ0503-8 and SQ0503-1 libraries (one sample discarded), as well as a significant ($p < 0.05$) value of 0.144 for the 46-

plex SQ0642-7 library (two samples excluded). A plot of the genetic differentiation index against its frequency was constructed. Among which, the top 5%, 10%, 15% and 20% SNPs (i.e., potential outliers), were removed to check for potential changes of the genetic patterns, respectively.

5.3 Results

5.3.1 Library fragment size distribution

The extractions yielded approximately 100 ng/µL of high-quality genomic DNA (ratio of absorbance A260/280 > 1.8 and 260/230 ratio > 1.6), as shown on an agarose gel (Figure 5. 2A). The genomic DNA was fragmented by enzymatic digestion with four combinations of three enzymes (*ApeKI*, *PstI* and *MspI*) including two controls, and were evaluated for their effectiveness in genome complexity reduction and enrichment of genic SNPs (Figure 5.2B). The use of all enzyme combinations effectively reduced genome complexity. In contrast, the library was well digested by double-enzyme digestion *ApeKI-MspI* and *PstI-MspI* in comparison with single-enzyme *ApeKI* and *PstI*. Specifically, the double digest *PstI-MspI* worked slightly better than *ApeKI-MspI* and was subsequently selected for downstream titration.

Satisfactory performances of titration were observed, suggesting high ligation efficiency and no potential technical issues. Both the pre-pippin and post-pippin libraries were checked, and the range of size-selection (bp) for digested DNA fragments was evaluated. The titration results for *PstI* and *PstI-MspI* with a range of barcode volumes (4, 6, 8, 10 µL) are shown in Figure 5.3. The fragment distribution of squat lobster DNA generated by the *PstI* endonuclease displayed some discrete, and repetitive DNA peaks, generally considered as repetitive DNA fragments. A similar observation was made for the *PstI-MspI* library as well. The presence of excess PCR primers (tall peak at ~72 bp) and adapter-dimers (small peak at ~124 bp) were observed in all *PstI* libraries, whereas no signal of adapter-dimer was detected in the *PstI-MspI* library. In addition to the six-base cutter *PstI* (CTGCAG), the *PstI-MspI* library was much larger, showing an increased number of both small and large fragments with the inclusion of the four-base cutter *MspI* (CCGG). Therefore, 8 µL barcode adapter with a concentration of 4.8 ng for double-enzyme *PstI-MspI* was chosen for subsequent GBS library construction, i.e., SQ0503-8, SQ0503-1 and SQ0642-7. The amplified fragment size distributions of SQ0503-8 constructed from the 272 DNA samples of *M*. *isos* from six regions

digested with *PstI-MspI* are shown in Figure 5.4 (results of SQ0503-1 and SQ0642-7 not shown). Although the fragments size distribution curve still contained a relatively high proportion of repetitive DNA in both pre-pippin and post-pippin libraries, a remarkable degree of enrichment was gained when the double-enzyme combination of *PstI-MspI* was employed. Fragments without adapter dimers between the sizes of 193–318 bp (adapter included) were captured. The spruce post-pippin library was very strong, indicating a large number of genomic restriction sites in *M. isos*.
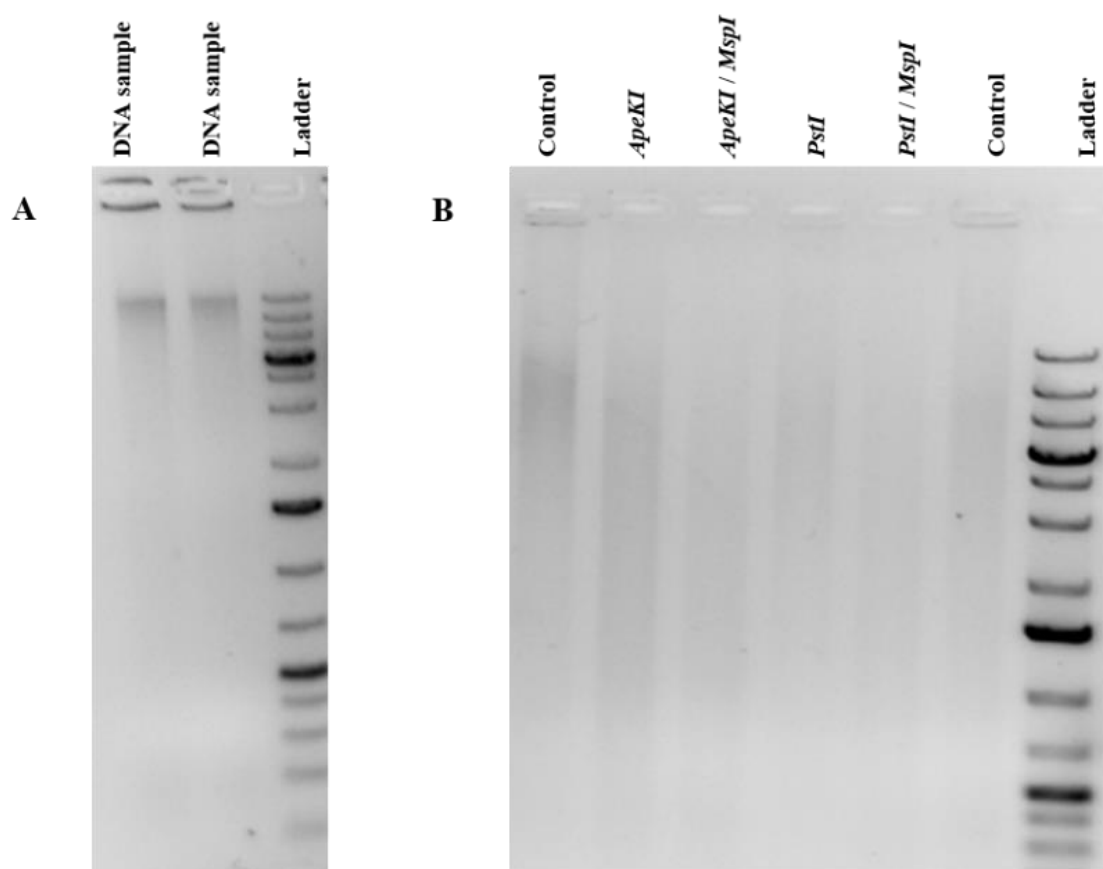
Figure 5.2 (A) Two genomic DNA samples of *M. isos* and (B) fragmented DNA samples of *M. isos* generated using three enzymes.
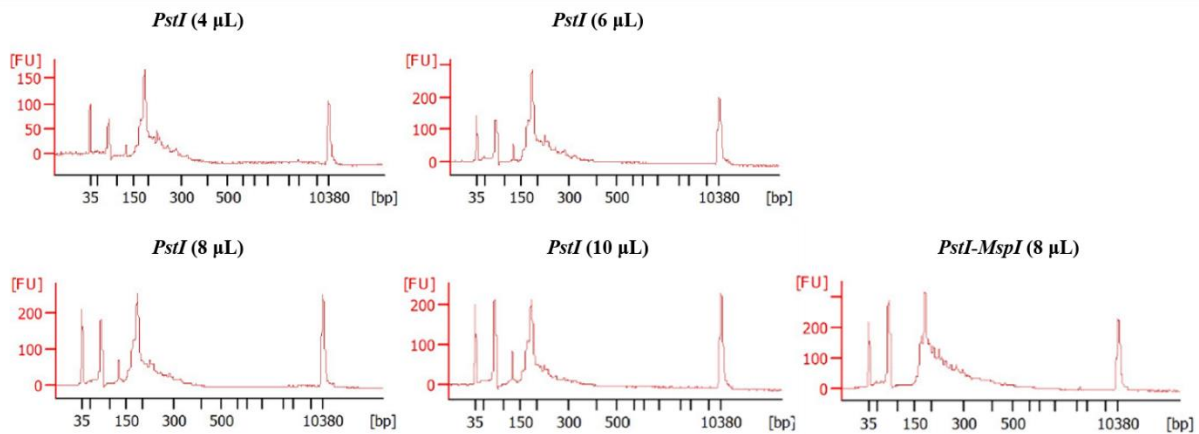
Figure 5.3 Fragment size distributions of GBS library constructed with several DNA samples of *M. isos*.

Peaks at 35 bp and 10,308 bp are size standards, ~72 bp are excess PCR primers, and ~124 bp are adapter-dimers.
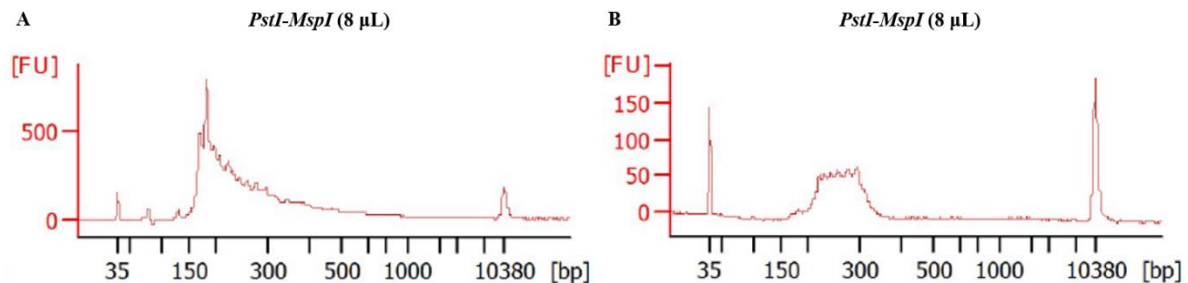
Figure 5.4 Fragment size distributions of pre-pippin (A) and post-pippin (B) GBS libraries of SQ0503-8 with 272 DNA samples of *M. isos* from six regions.

Peaks at 35 bp and 10,308 bp are size standards, ~72 bp are excess primers, and ~124 bp are adapter-dimers.

5.3.2 Sequencing of GBS libraries

Sequences from three libraries (SQ0503-8, SQ0503-1, SQ0642-7) of the squat lobster *M. isos* DNA were collected in three lanes of separate flow cells (Table 5.5). Lane-to-lane sequencing run variation of the read counts was moderately high, yielding ca. 257.01 M, 190.08 M and 260.44 M GBS raw reads (35–101 bp), respectively. Amongst these, 91.97% (ca. 236.37 M), 91.24% (ca. 173.44 M) and 95.54% (ca. 248.82 M) passed the initial filtering criteria (i.e., contained the barcode followed by the expected cut site remnant and no "Ns", that is, "good barcode reads"), respectively. On a per-sample basis, the average numbers of "good barcode reads" of 48 barcoded genotypes (i.e., SQ0642-7) was over 5.18 M reads, whereas the two 272-plex libraries gave fewer reads with 637 K for SQ0503-1 and 869 K for SQ0503-8. No sequence was flagged as poor quality. No contamination was observed, and the negative controls (no DNA samples) proved to be negative.

Table 5.5 Summary number of individuals, raw sequences, good barcoded reads and good barcoded reads per individual analysed by the FastQC.

	SQ0503-8	SQ0503-1	SQ0642-7
Number of individuals	272	272	48
Raw reads	257,005,421	190,083,092	260,435,568
Good barcoded reads	236,370,370	173,438,885	248,819,819
Good barcoded reads per individual	869,009	637,643	5,183,746

Good barcoded reads: 64 bp reads with barcode, cut-site remnant and no "Ns".

5.3.3 SNPs identification and selection

The UNEAK *de novo* bioinformatics pipeline was used to detect variants and report alternative allele counts at variant sites for four datasets (merged, SQ0503-8, SQ0503-1 and SQ0642-7 libraries). Initially, a total of ca. 4.75M, 2.13 M, 1.69 M and 2.57 M sequencing tags were identified from the merged and the three single libraries, respectively, representing *M. isos* across six localities in the South Pacific Ocean (Table 5.6). Amongst these, ca. 176,688 (merged library), 110,256 (SQ0503-8), 93,485 (SQ0503-1) and 112,819 (SQ0642-7) SNPs that were composed of 64 bp biallelic tag pairs with 1 bp mismatch were filtered with the filtration threshold of sequencing ETR lower than 0.03. The subsequent thresholds for MAF and CR were applied to remove poorly genotyped SNPs. Additional filters on tags presented in at least 10% of the individuals (CR $\geq$ 0.1) and MAF $\geq$ 0.03 were applied, resulting in three single libraries with putative SNP numbers ranging from 36,408 (SQ0503-1) to 63,676 (SQ0642-7), averaged at 46,166, which was similar to that from the merged library (42,880). It was apparent that a much higher number of SNPs per sample was called by the 48-plex library SQ0642-7 compared with the other three 272-plex libraries.

Table 5.6 Summary number of raw tags and the corresponding SNPs retained following filtering steps analysed by the UNEAK pipeline.

	Merged library	SQ0503-8	SQ0503-1	SQ0642-7
Number of individuals	272	272	272	48
Raw tags	4,751,005	2,127,326	1,694,801	2,565,299
Tags (ETR $\leq$ 0.03)	353,376	220,512	186,970	225,638
Tag pairs / SNPs (ETR $\leq$ 0.03)	176,688	110,256	93,485	112,819
SNPs (MAF = 0.03–0.5; CR = 0.1–1)	42,880	38,415	36,408	63,676

ETR, error tolerance rate; MAF, minor allele frequency; CR, call rate.

5.3.4 Characteristics of filtered SNPs and relatedness analyses

Initial filtering yielded a total number of 181,379 SNPs from four datasets of *M. isos* across all six regions in UNEAK before the data set was exported for KGD analysis. Amongst the three single libraries, SNPs were highest for the SQ0642-7 (62,709) and lowest for the SQ0503-1 (36,408) libraries. Prior to the analyses, two samples (Lob120 and Lob185) were dropped from the merged and the SQ0642-7 libraries due to uncertainty about their identity; one sample (Lob196) was discarded from the merged, SQ0503-8 and SQ0503-1 libraries with the filtration threshold of maximum sample depth $= 1$ and/or mean sample depth < 0.3. As a result, 269 (merged library), 271 (SQ0503-8), 271 (SQ0503-1) and 46 (SQ0642-7) samples of *M. isos* from the six regions were retained. Through subsequent pruning, approximately 6% of the filtered SNPs were excluded due to MAF $= 0$ or read depth < 0.01, leaving a total number of 170,213 SNPs from the four libraries. Notably, the mean sample depth ranged from 1.10 (SQ0503-1) to 4.31 (SQ0642-7) in three single libraries whilst a mean sample depth of 1.82 was observed in the merged library (Table 5.7).

Table 5.7 Summary statistics following filtering steps analysed by the KGD.

	Merged library	SQ0503-8	SQ0503-1	SQ0642-7
Number of individuals	269	271	271	46
SNPs (MAF $\neq 0$ or depth ≥ 0.01)	32,681	38,415	36,408	62,709
Mean sample depth	1.82	1.36	1.10	4.31
Proportion of missing genotypes	0.54	0.55	0.58	0.50
CR	0.46	0.45	0.42	0.50
SNPs (present across all regions)	30,844	38,415	36,408	62,709
Mean sample depth	4.15	2.00	1.75	4.87
SNPs (HWD $\geq$ -0.05)	30,158	37,808	35,974	61,127
Mean sample depth	3.05	1.67	1.50	4.07
SNPs (genetic differentiation index)	19,692	15,017	14,576	13,711
Mean sample depth	2.08	1.69	1.52	3.26

MAF: minor allele frequency; CR: call rate; Genetic differentiation index: 0.025 for the merged, SQ0503-8 and SQ0503-1 libraries and 0.144 for the SQ0642-7 library.

The relationships between average sample SNP read depth and sample SNP CR for four datasets of *M. isos* are depicted in Figure 5.5. This curvilinear pattern is commonly observed in GBS data, whereby samples with a higher depth tend to have higher call rates. The SQ0642-7 library, with only 46 samples, exhibited a relatively scattered distribution compared to the other three libraries. Nonetheless, it was apparent that all *M. isos* samples were clustered as only one group in the merged library, indicating that the differences of genotyping in the three different flow cells were low. Figure 5.6 depicts the relationships between average SNP depth and SNP CR for four libraries of *M. isos*. All datasets demonstrated a sigmoidal relationship between depth and log call rate with the only exception of the SQ0642-7 library which displayed a fatter "S" pattern due to the small sample size and high depth. The average call rates (0.46) and the average proportion of missing genotypes (0.54) were similar between pairs of sets. Overall, KGD captures the relationship between SNP depth and SNP CR appropriately.
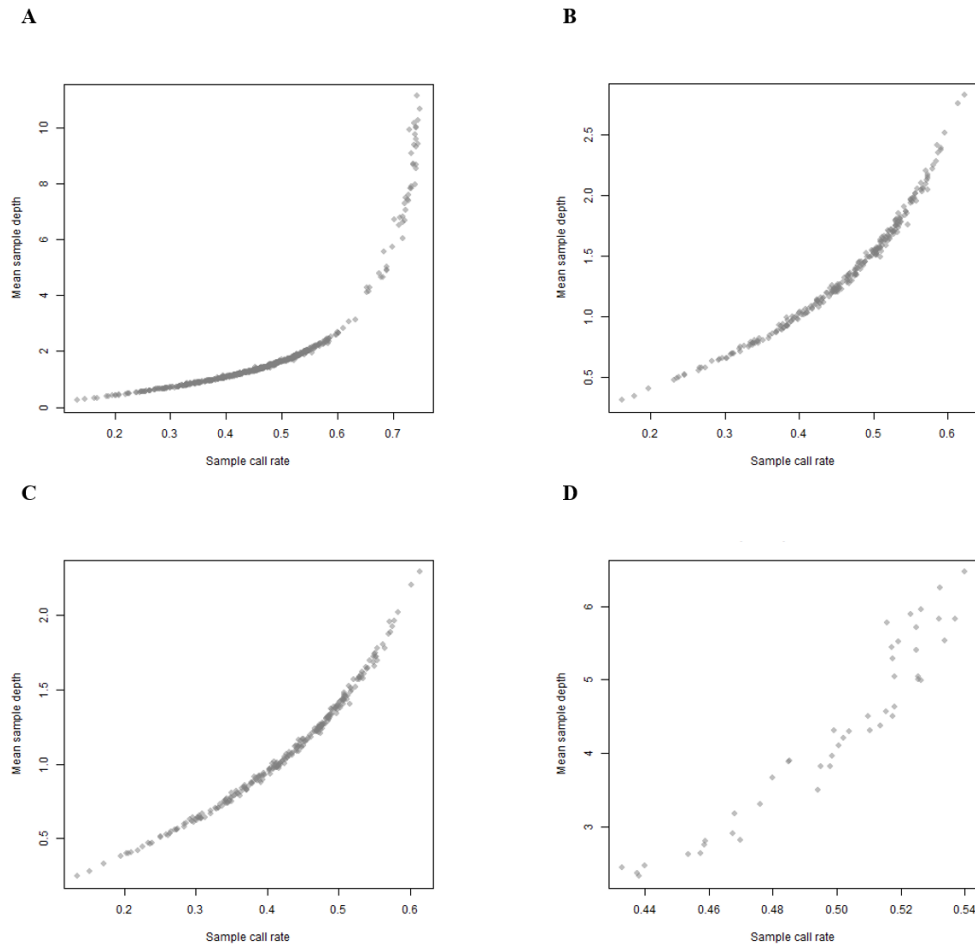
Figure 5.5 Relationship between average sample SNP read depth and sample SNP call rate.

Merged (A), SQ0503-8 (B), SQ0503-1 (C) and SQ0642-7 (D) libraries with 269, 271, 271 and 46 individuals of *M. isos* from six regions, respectively.

A

B

C

D
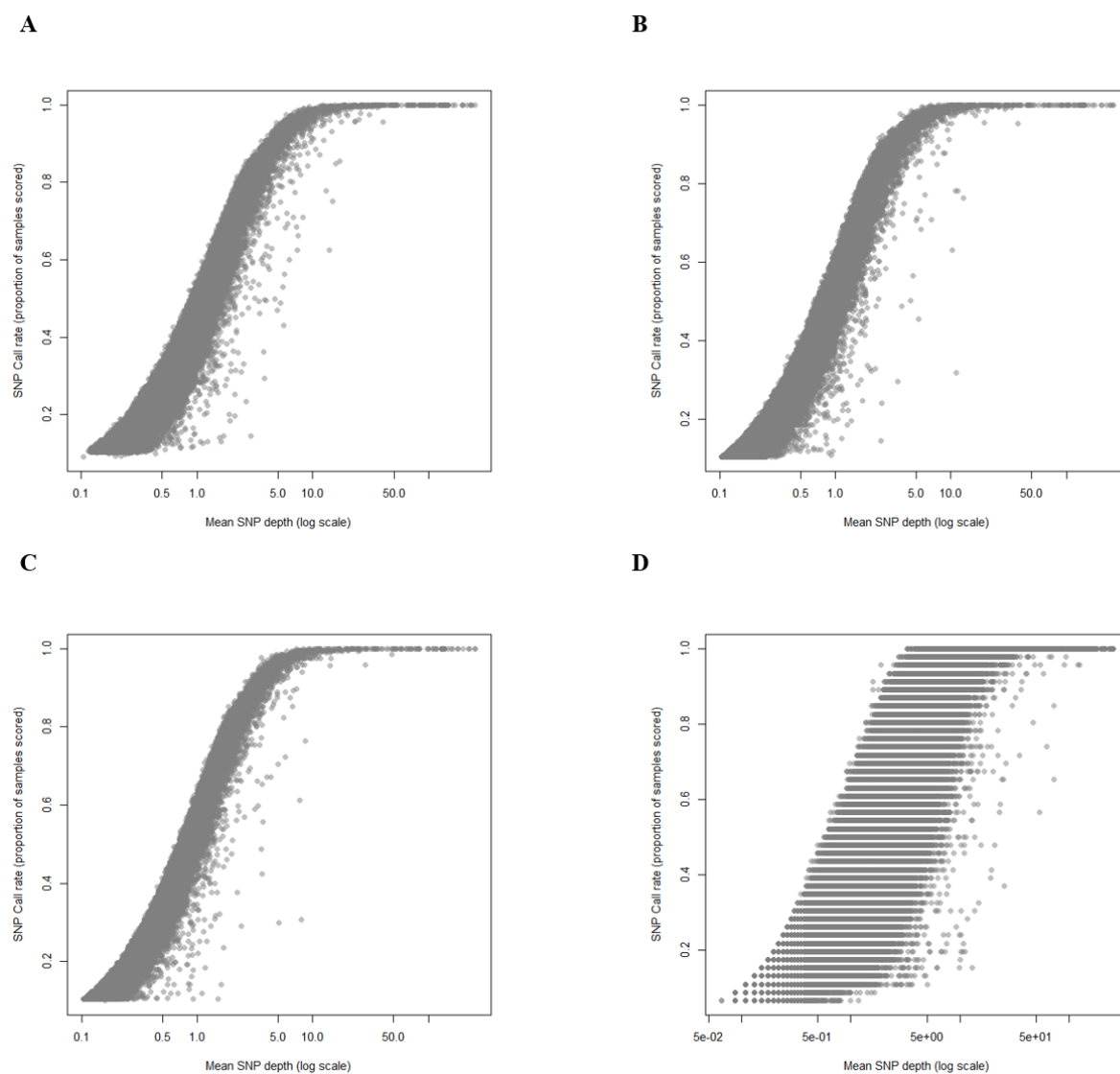
Figure 5.6 Relationships between average SNP read depth and SNP call rate.

Merged (A), SQ0503-8 (B), SQ0503-1 (C) and SQ0642-7 (D) libraries with 269, 271, 271 and 46 individuals of *M. isos* from six regions, respectively.

Finplots were obtained for the merged and the three single libraries (Figure 5.7). SNPs from the upper edge of the fin plot were of a high density with low depth (coloured grey), those with a depth of 1 (corresponding genotype shown as a homozygote) appeared near this region. It is very much likely that SNPs located in the bottom right corner, with relatively high depth (coloured blue) and excess heterozygosity (at the given MAF), represent calls from duplicated regions in the genome where the depth of SNPs was at least double that from other regions.
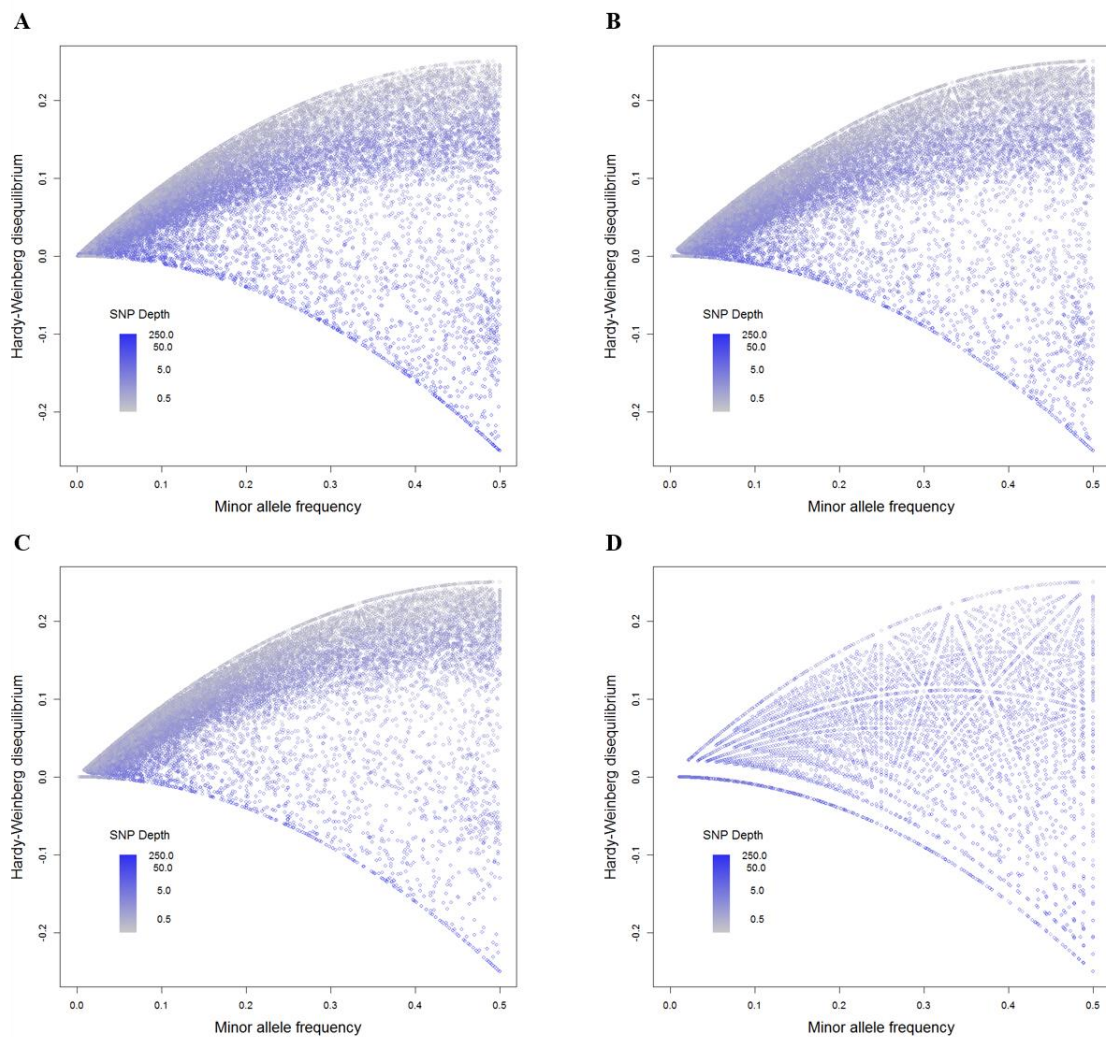
Figure 5.7 Finplots for SNPs.

Merged (A), SQ0503-8 (B), SQ0503-1 (C) and SQ0642-7 (D) libraries with 269, 271, 271 and 46 individuals of *M. isos* from six regions, respectively.

Panels of SNPs were additionally filtered: 1) with data for all six regions, 2) with HWD $\geq$ -0.05, and 3) with a significant level of genetic differentiation index. The overall MAF distributions of SNPs from the three filters are provided in Figures 5.8–5.10. It can be argued that little was lost through the first removal step of the merged library, with only 1,837 (5.6%) SNPs dropped from the panel due to the sample merging process. In general, the three filters resulted in a reduction of MAF frequencies, as well as a drop in the number of SNPs, especially for those with large MAF values which often represent duplicated regions of the genome. Only small differences were observed amongst the merged and the two 271-plex libraries. However, it was of interest to note that the 46-plex SQ0642-7 library experienced a substantial loss of SNPs after utilising the genetic differentiation index $\geq$ 0.144 (46 samples from 6 regions) filter, showing a more evenly distributed MAF. As a result, for the SQ0642-7 library, a total of 13,711 SNPs were retrieved with a mean sample depth of 3.26, which took up 22.43% of the SNPs from the previous filter of HWD $\geq$ -0.05.
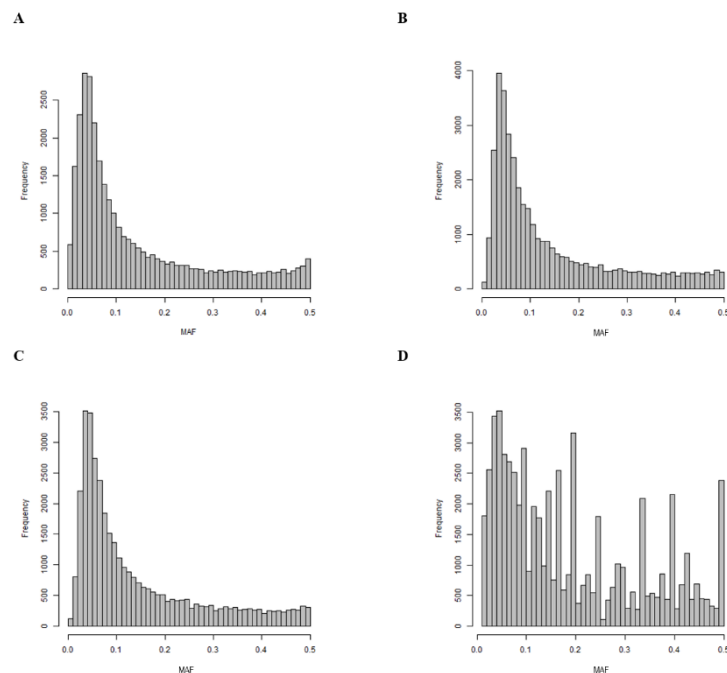
Figure 5.8 Distribution of minor allele frequency (MAF) values for SNPs.

Merged (A), SQ0503-8 (B), SQ0503-1 (C) and SQ0642-7 (D) libraries with 269, 271, 271 and 46 individuals of *M. isos* from six regions, respectively.
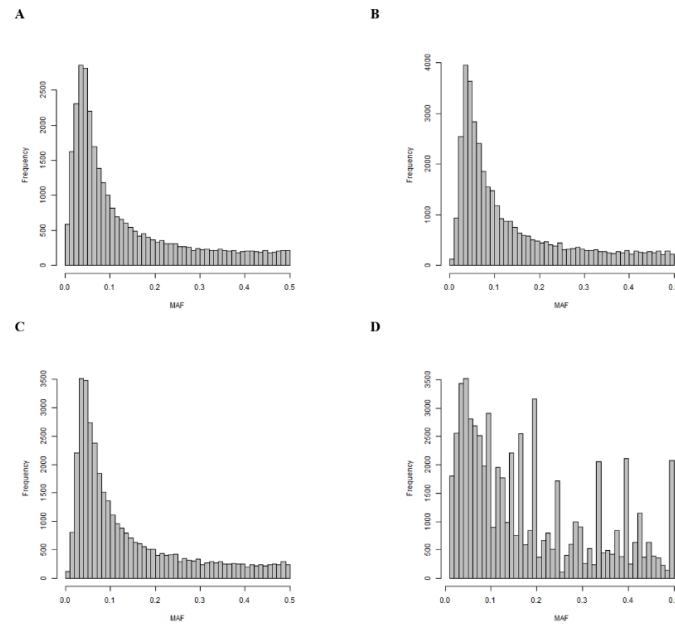
Figure 5.9 Distribution of minor allele frequency (MAF) values for SNPs with HWD filter of ≥ -0.05.

Merged (A), SQ0503-8 (B), SQ0503-1 (C) and SQ0642-7 (D) libraries with 269, 271, 271 and 46 individuals of *M. isos* from six regions, respectively.
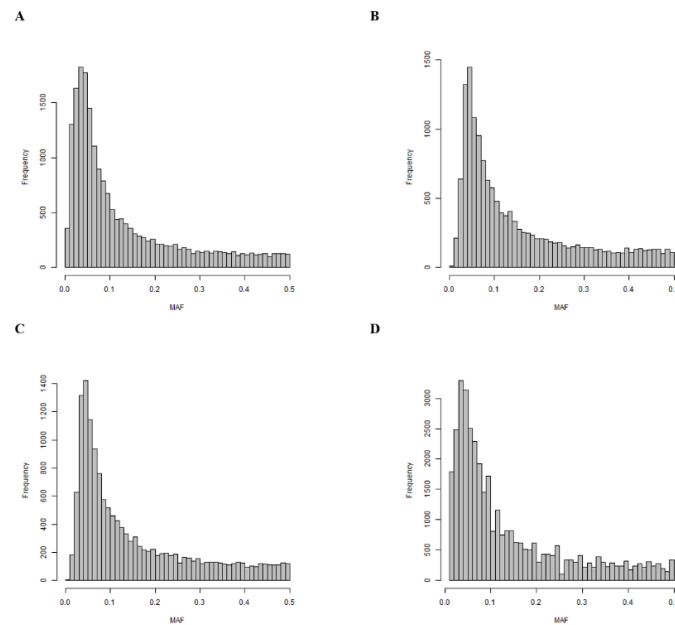
Figure 5.10 Distribution of minor allele frequency (MAF) values for SNPs with HWD filter of ≥ -0.05, and had a significant level of genetic differentiation index.

A significant level of genetic differentiation index of 0.025 for the merged (A), SQ0503-8 (B), and SQ0503-1 (C) libraries, and 0.144 for the SQ0642-7 (D) library with 269, 271, 271 and 46 individuals of *M. isos* from six regions, respectively.

GRMs of dimension 269 × 269, 271 × 271, 271 × 271 and 46 × 46 were constructed using the KGD model for the merged, SQ0503-8, SQ0503-1 and SQ0642-7 libraries, respectively. The GRM was depicted in the heat map with the relatedness between individuals on off-diagonal and the self-relatedness on the diagonal, with a colour scale of red (related) to white (unrelated). SNPs with a filter of HWD ≥ -0.05 and an additional filter of a significant level of genetic differentiation index were used accordingly. After utilising the filter of HWD ≥ -0.05 in four datasets, a panel of 165,067 SNPs allowed for differentiation of populations across six regions in four datasets, showing two main clusters with the majority of individuals from the Macquarie Ridge population being grouped with the Tasmanian slope population, and a second cluster formed by individuals from all other populations, including the Louisville Seamount Chain (to the northeast of the New Zealand's EEZ) (Figure 5.11). Notably, as revealed by the merged dataset, four individuals (Lob97, Lob111, Lob122, Lob129) from Macquarie Ridge were found to be present in the Tasmanian slope cluster (Figure 5.12). Amongst these, three were from the population NIWA39983 with the depth of 506–560 m, representing the shallowest sampling site of *M. isos* within New Zealand's EEZ. High levels of relatedness were observed amongst individuals from populations of the Kermadec Ridge, the Hikurangi Margin, the Chatham Rise and the Louisville Seamount Chain. In contrast, individuals from the Tasmanian slope and the Macquarie Ridge exhibited high levels of relatedness within populations from each geomorphic region.

Figure 5.11 Genomic relatedness matrix of the merged SNPs library and three individual SNPs libraries with HWD filter of ≥ -0.05.

Analyses were done on individuals from all six regions for the merged (A), SQ0503-8 (B), SQ0503-1(C), and SQ0642-7 (D) libraries with 269, 271, 271 and 46 individuals, respectively. Colour scale is from red (related) to white (unrelated). Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope.

Figure 5.12 Genomic relatedness matrix of the merged SNPs library with HWD filter of ≥ -0.05.

Colour scale is from red (related) to white (unrelated). Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope.

The GRMs shown in Figure 5.13 were composed of the SNPs that passed the HWD filter > -0.05 and had a significant level of genetic differentiation index of 0.025 for the merged, SQ0503-8 and SQ0503-1 libraries ($N = 272$) and a significant level of genetic differentiation index of 0.144 for the SQ0642-7 library ($N = 46$). Although slightly different groupings were detected compared with the previous filter (i.e., HWD > -0.05), the retained SNPs with a filter of genetic differentiation index of 0.025 were still able to accurately cluster the majority of individuals from Tasmanian slope and Macquarie Ridge populations according to their localities in the merged library (Figure 5.14). A more detailed structure was revealed for squat lobsters from the Hikurangi Margin and Kermadec Ridge populations, with four out of five individuals from Hikurangi Margin being clustered together in the merged and SQ0503-8 libraries, whilst all five individuals formed one group in SQ0503-1, and the Kermadec Ridge samples were grouped together in SQ0503-8. In contrast, there was a more prominent variability observed in the SQ0642-7 library which had a relatively conservative genetic differentiation index (0.025) due to its sample size (46 individuals). The GRM showed six unique clusters based on different localities, with a similar break between Tasmanian slope, Macquarie Ridge and populations in other regions (Figure 5.12D). After excluding potential outlier SNPs by removing the largest 5%, 10%, 15% and 20% of the genetic differentiation index, patterns of GRMs remained the same for all four datasets.

Figure 5.13 Genomic relatedness matrix of the merged SNPs library and three individual SNPs libraries with HWD filter of ≥ -0.05, and had a significant level of genetic differentiation index.

A significant level of genetic differentiation index of 0.025 for the merged (A), SQ0503-8 (B), and SQ0503-1 (C) libraries, and 0.144 for the SQ0642-7 (D) library with 269, 271, 271 and 46 individuals, respectively. Colour scale is from red (related) to white (unrelated). Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope.

Figure 5.14 Genomic relatedness matrix of the merged SNPs library with HWD filter of ≥ -0.05, and had a significant level of genetic differentiation index.

A significant level of genetic differentiation index of 0.025 for the merged library with 269 individuals. Colour scale is from red (related) to white (unrelated). Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope.

The PCA plots roughly grouped individuals into three main clusters according to their geographical locations, consistent with the results of the GRM, with three major genetic clusters defined by the first four principal components (PCs). After utilising the filter of HWD $\geq$ -0.05 in the four datasets, the first two PC axes cumulatively accounted for 13.3%, 11.6%, 11.2% and 9.1% of the total variation for the merged, SQ0503-8, SQ0503-1 and SQ0642-7 libraries, respectively (Figure 5.15). Notably, PC1 only showed the ability to separate cluster one (Tasmanian slope), cluster two (Macquarie Ridge) and cluster three (Kermadec Ridge, Louisville Seamount Chain, Hikurangi Margin, Chatham Rise), which account for 11.9%, 10.3%, 9.9% and 6.2% of the total variation for the merged, SQ0503-8, SQ0503-1 and SQ0642-7 libraries, respectively. The Tasmanian slope population was highly differentiated and did not overlap with any other populations, with the exception of four Macquarie Ridge individuals (consistent with results from the GRM) (Figure 5.16). Compared with the Tasmanian slope, the Macquarie Ridge cluster was situated closer to the cluster of the other four regions, which exhibited some overlap and, therefore, a relatively low level of population genetic differentiation.
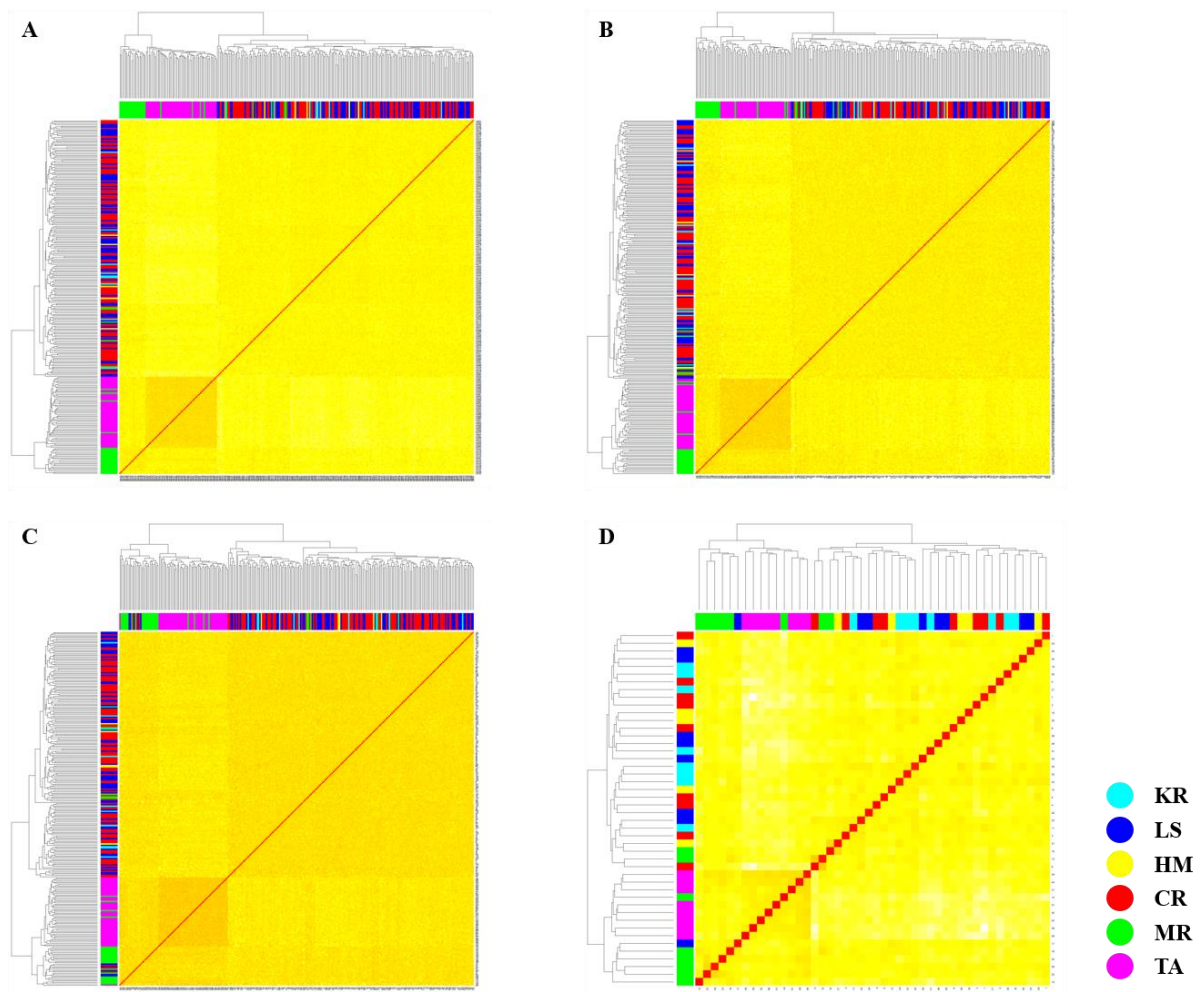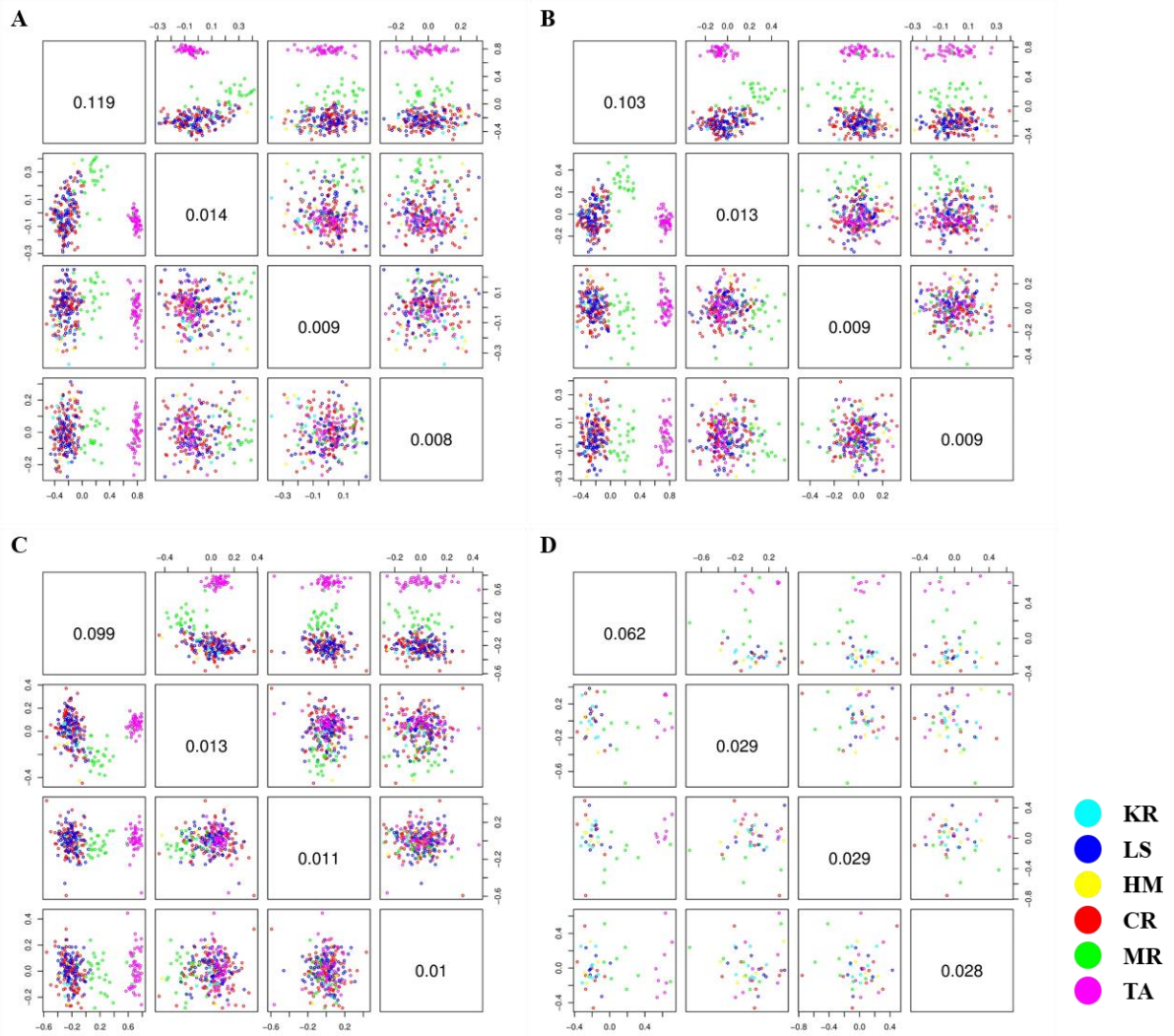
Figure 5.15 PCA plots of the merged SNPs library and three individual SNPs libraries with HWD filter of ≥ -0.05.

Analyses were done on individuals from all six regions for the merged (A), SQ0503-8 (B), SQ0503-1(C), and SQ0642-7 (D) libraries with 269, 271, 271 and 46 individuals, respectively. First four principal components were plotted on separate axes. Each point represents an individual colour coded according to its region. Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope.
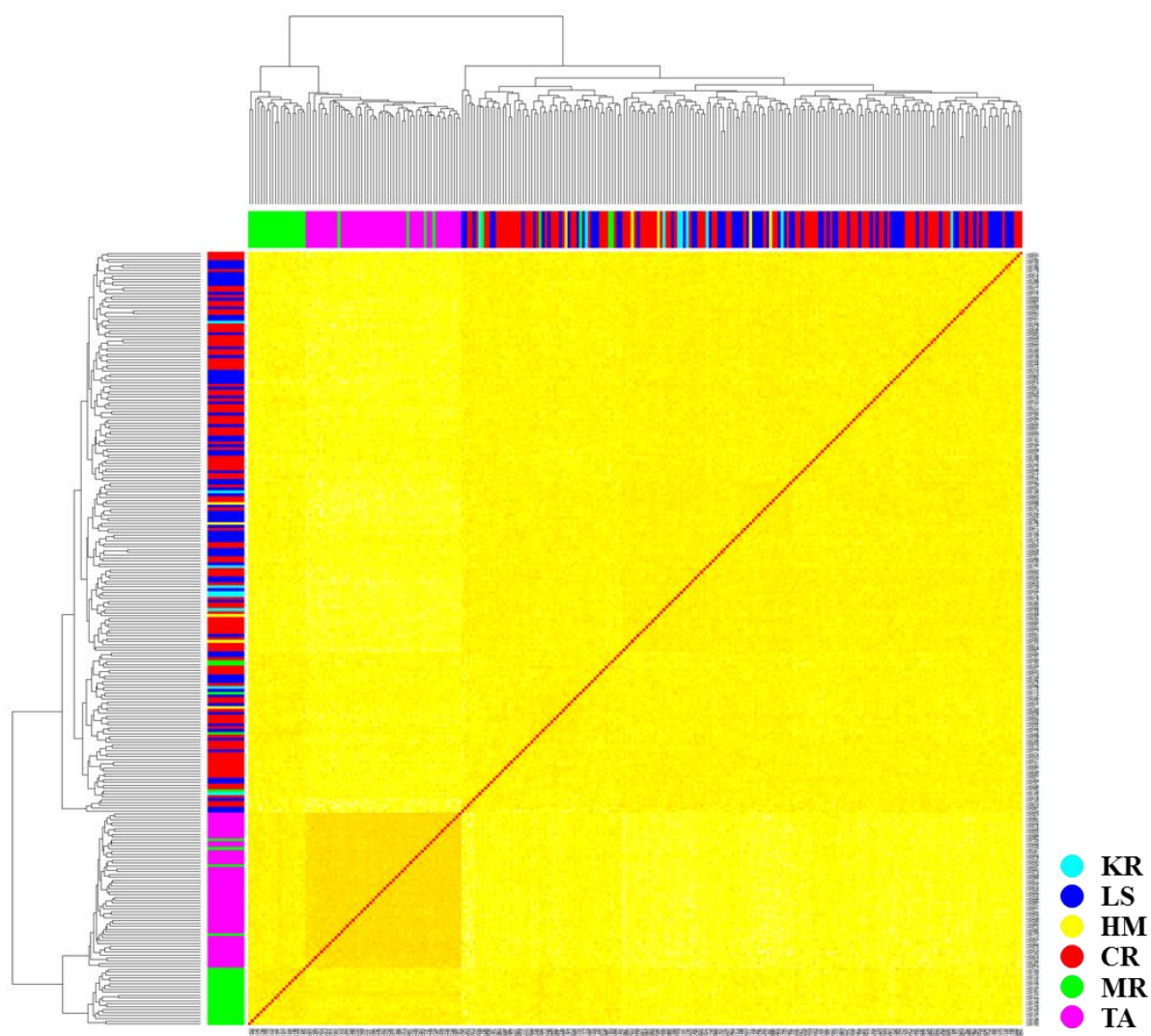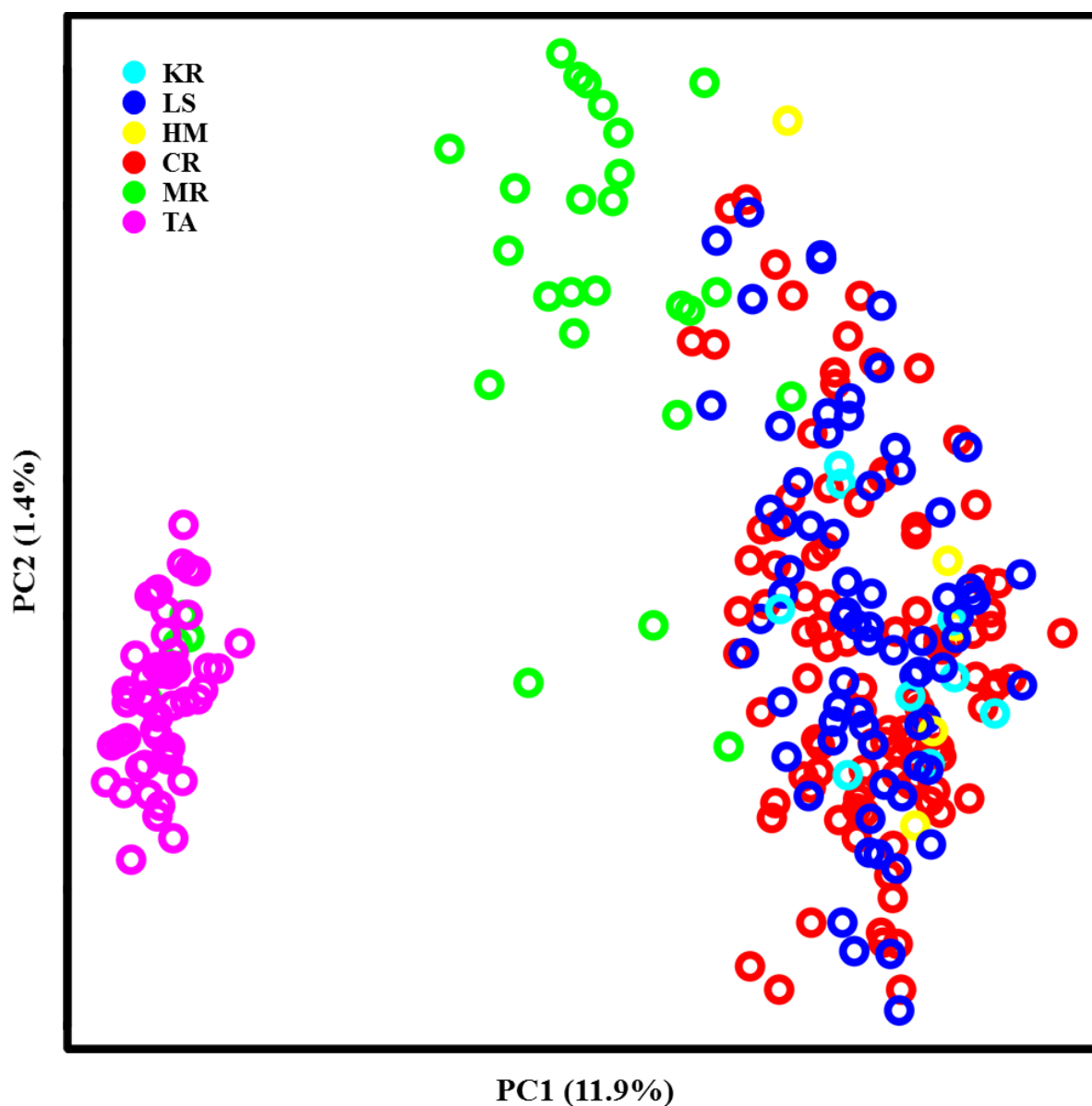
Figure 5.16 PCA plot of the merged SNPs library with HWD filter of ≥ -0.05.

Each point represents an individual colour coded according to its region. Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope.

PCA results (Figure 5.17) were computed by the SNPs that passed the HWD filter > -0.05 and had a significant level of genetic differentiation index of 0.025 for the merged, SQ0503-8 and SQ0503-1 libraries ($N = 272$) and a significant level of genetic differentiation index of 0.144 for the SQ0642-7 library ($N = 46$). PC1 and PC2 combined explained 32.9%, 28.2%, 25.7% and 27.4% of the total variation for the merged (Figure 5.18), SQ0503-8, SQ0503-1 and SQ0642-7 libraries, respectively. Patterns observed from PCA spatial data were consistent with previous findings, showing three main clusters. However, an interesting exception occurred in the SQ0642-7 dataset which demonstrated a relatively clear differentiation amongst samples from the six regions (Figure 5.17D). PC1 and PC2 successfully differentiated cluster one (Tasmanian slope), cluster two (Macquarie Ridge) and cluster three (Kermadec Ridge, Louisville Seamount Chain, Hikurangi Margin, Chatham Rise), whilst the third and fourth PCs mainly separated individuals from Hikurangi Margin, Chatham Rise and Kermadec Ridge with 37.4% of the total variation. These results are consistent with those obtained from the GRM. After excluding potential outlier SNPs by removing the largest 5%, 10%, 15% and 20% of the genetic differentiation index, patterns of PCA plots remained similar for all four datasets.
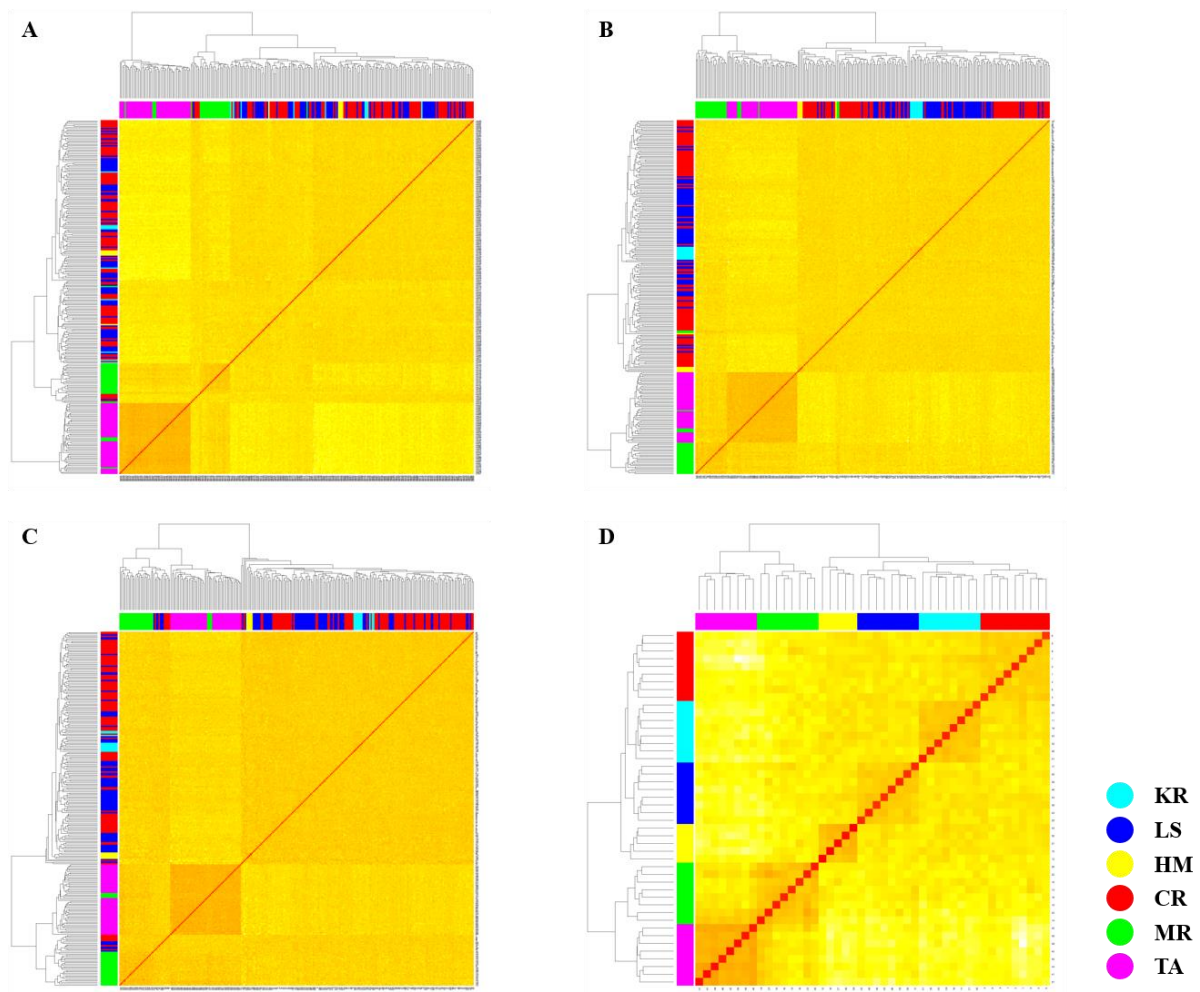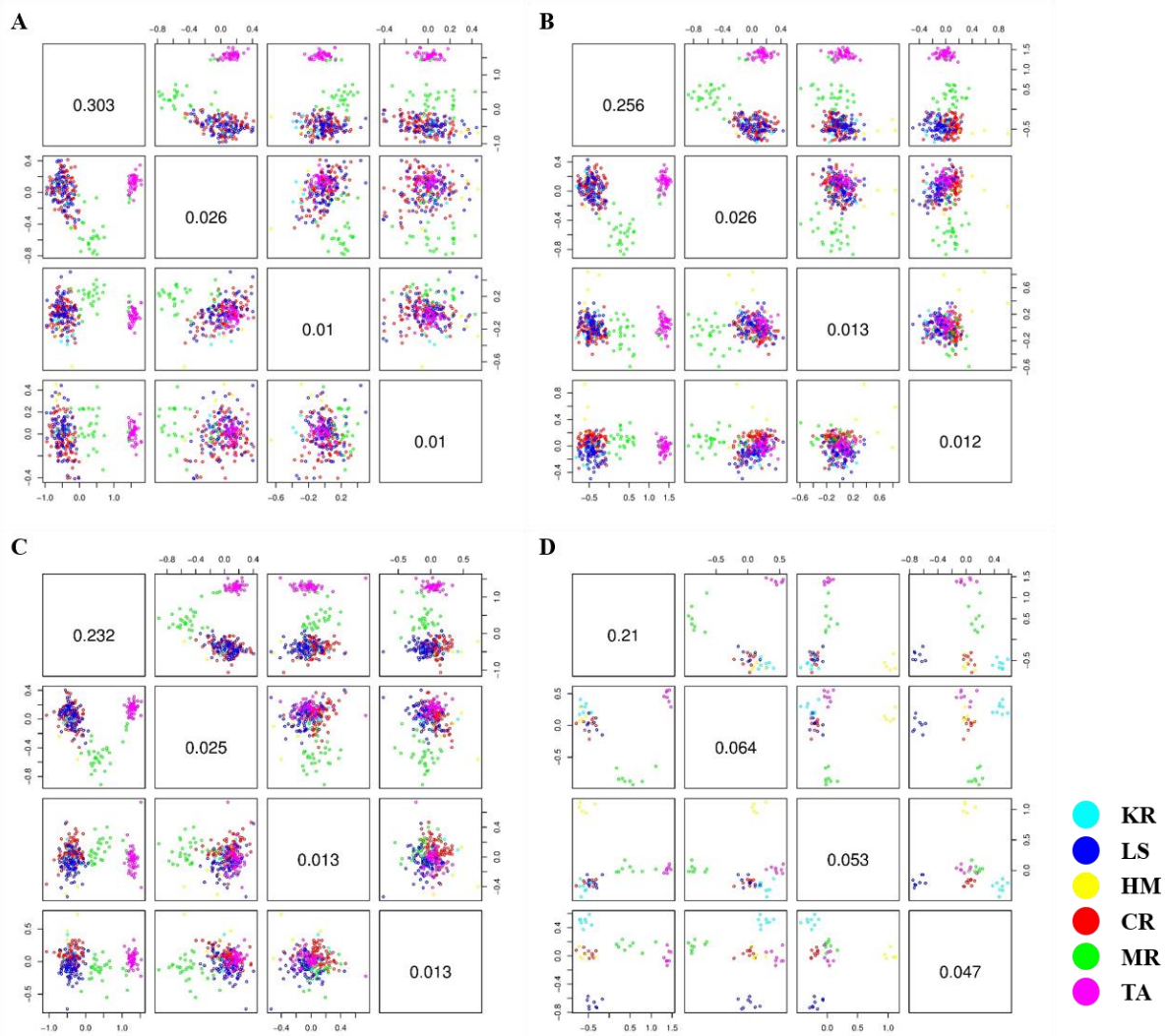
Figure 5.17 PCA plots of the merged SNPs library and three individual SNPs libraries with HWD filter of ≥ -0.05, and had a significant level of genetic differentiation index.

A significant level of genetic differentiation index of 0.025 for the merged (A), SQ0503-8 (B), and SQ0503-1 (C) libraries, and 0.144 for the SQ0642-7 (D) library with 269, 271, 271 and 46 individuals, respectively. First four principal components were plotted on separate axes. Each point represents an individual coloured according to its region. Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope.
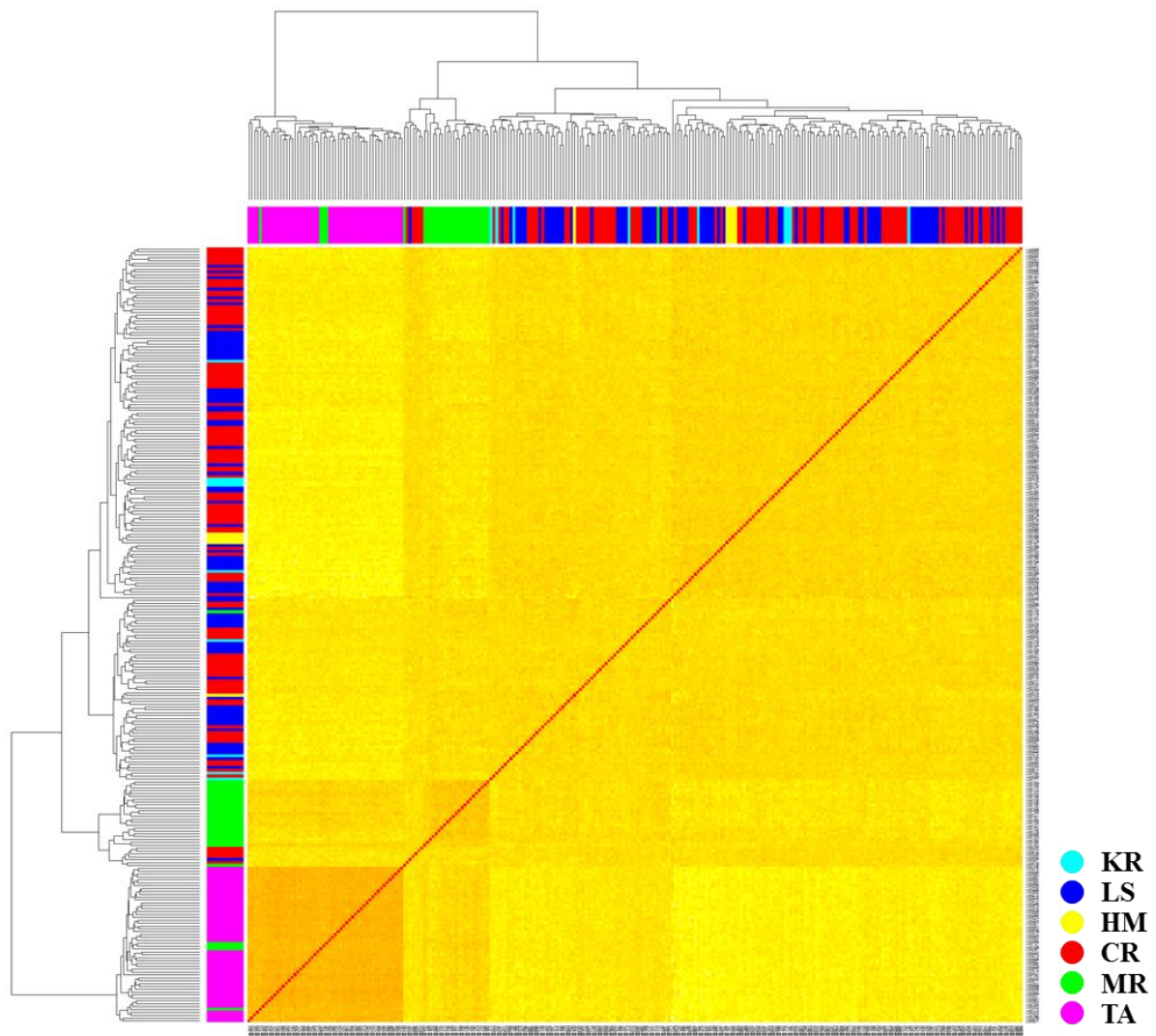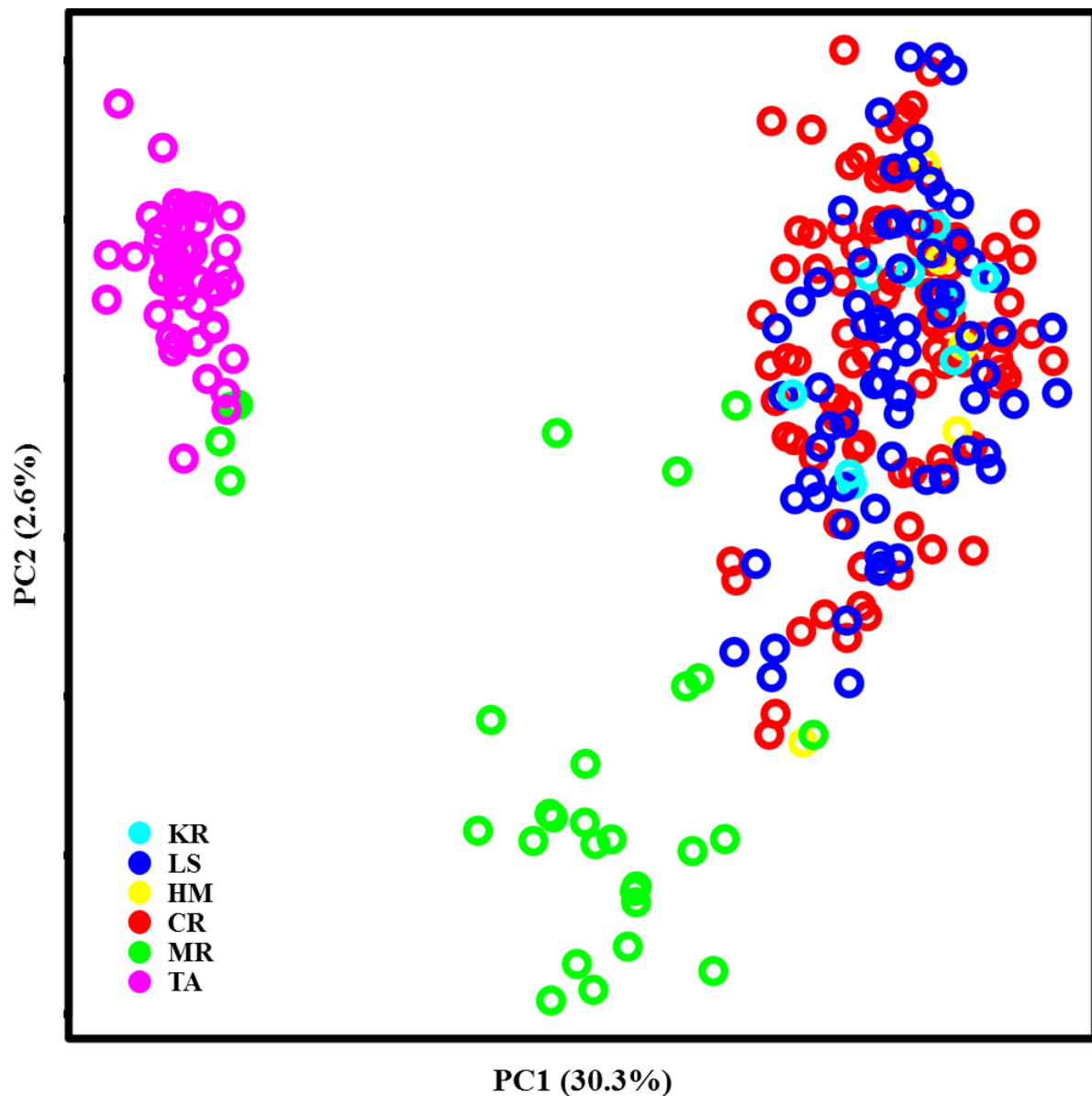
Figure 5.18 PCA plot of the merged SNPs library with HWD filter of ≥ -0.05, and had a significant level of genetic differentiation index.

A significant level of genetic differentiation index of 0.025 for the merged library with 269 individuals. Each point represents an individual coloured according to its region. Geomorphic feature codes: KR = Kermadec Ridge; LS = Louisville Seamount Chain; HM = Hikurangi Margin; CR = Chatham Rise; MR = Macquarie Ridge; TA = Tasmanian slope.

5.4 Discussion

Three GBS libraries of the deep-sea squat lobster *M. isos* were successfully constructed utilising the double-enzyme *PstI-MspI* approach. The degree of genetic relatedness and connectivity of squat lobster populations in the southwest Pacific Ocean was assessed with filtered SNPs using the UNEAK bioinformatics pipeline, for the purpose of aiding the development of conservation and management strategies for VMEs. The results showed that the Tasmanian slope and Macquarie Ridge populations were genetically different from other populations to the east, both within New Zealand's EEZ and beyond. This finding of regional genetic differences amongst the squat lobster populations is in contrast to the results from the microsatellite and mtDNA sequence analyses, and has new and important implications for the management of VMEs.

5.4.1 GBS efficiency in library construction

The completeness of GBS SNPs calling and the number of SNPs filtered is known to be influenced by several factors including 1) the number of fragments produced by enzyme digestion, 2) the sequencing depth, determined by the number of reads (e.g., HTS efficiency), the number of multiplexed samples, and the genome size of the target organism, and 3) the density of SNPs, which is related to the diversity and structure of the populations (Poland and Rife 2012). Here, knowing that some limitations do exist in the methodological approach applied, the following precautions were employed aiming at increasing GBS efficiency for squat lobster populations: 1) sample bias due to unequal starting DNA volume was minimised by equimolar pooling of individual DNA samples through careful evaluation of DNA quality and quantity, 2) digestion enzymes were carefully selected and evaluated, 3) multiple PCR reactions with fewer cycles were performed to minimise potential PCR bias, 4) the bioinformatics filters were relatively conservative, which effectively classified the allele counts into genotype calls, and 5) additional GBS runs and lower multiplex level were applied aiming at increasing sequencing depth.

The double-enzyme digestion in the GBS library construction enables high-throughput genome-wide genotyping with HTS technology, leading to more accurate detection of SNPs. Choosing an appropriate combination of enzymes for GBS library construction is a critical factor that needs to be considered carefully before laboratory application depending on specific experimental requirements (Hamblin and Rabbi 2014). Firstly, a sufficient number of different

restriction fragments with appropriate size need to be produced. Secondly, the genome needs to be cut randomly without apparent over-representation of duplicate DNA sequences (Gore *et al.* 2007). Here, *ApeKI* and *MspI* were selected as candidate enzymes. Recognition sites of *MspI* are mostly located in high-CG regions (Gu *et al.* 2010), the *MspI*-digested DNA fragments will be further digested by adding a rare-cutting enzyme such as *ApeKI* and *PstI* to create a more representative coverage in low-CG regions (Wang *et al.* 2013b). Notably, *ApeKI* and *PstI* were previously demonstrated to give the most random fragment distribution for the maize (Elshire *et al.* 2011), wheat (Poland *et al.* 2012a; Poland *et al.* 2012b) and cattle (De Donato *et al.* 2013) genomes, making these two enzymes ideal for large genome complexity reduction. In this study, the double digest *PstI-MspI* worked slightly better than *ApeKI-MspI* and hence was selected for the subsequent GBS procedure.

The sequencing depth required to obtain a complete dataset for all SNP-containing fragments is currently very expensive and impractical (Huang *et al.* 2014). As such, there is always a trade-off between coverage depth and density of SNPs. For complex genomes with a high level of genetic diversity, it may be better to increase sequencing depth per SNP locus for more reliable heterozygote calling, which can be achieved by adding a rare-cutter (Chen *et al.* 2013). However, for applications such as genome-wide association studies (GWAS) that require a large number of SNPs, it is advisable to select a common-cutter for more reads and increase the read depth by performing additional sequencing runs, or if the budget is tight, by sequencing at lower multiplex levels (Chen *et al.* 2013). In this study, a 272-plex library (SQ0503-8) was initially constructed with 38,415 SNPs and a mean sample depth of 1.36. The coverage depth and density of SNPs were further improved by performing two additional sequencing runs (SQ0503-1 and SQ0642-7), with the latter consisting of a lower multiplex sample size (48 individuals). The three libraries successfully generated relatively large and uniform fragment size distributions and produced an average of ca. 2,129,142 unique sequence polymorphisms with an average depth of 2.25.

5.4.2 SNPs filtering

As a powerful non-reference-genome bioinformatics pipeline, UNEAK was initially applied to perform SNP calling, giving clear and satisfying results and passing different filtering criteria. However, SNPs identified by GBS can vary considerably with different pipelines (Mascher *et al.* 2013; Torkamaneh *et al.* 2016), and there is a possibility that the additional SNPs called by a particular pipeline may contain important information which would otherwise be discarded by some other pipelines. Although differences were observed for retained SNPs through different calling pipelines due to initial filtering parameters, there is high concordance on genotype calls for common SNPs (Mascher *et al.* 2013). It will be of interest to compare the results obtained using UNEAK with other *de novo* pipelines (e.g., Stacks) in future studies.

KGD was utilised for the second phase filtering to obtain an unbiased relatedness estimation using theoretical calculation and simulation. The number of detected tags and depth of coverage can affect both the number and quality of generated variants (Elshire *et al.* 2011; Brouard *et al.* 2017). Moreover, appropriate filtering is likely to improve genetic estimates, and the optimal filters commonly depend on the GBS approaches and the properties of the species in question (e.g., genome size) (Dodds *et al.* 2015). Therefore, SNPs with a sufficient level of diversity can be filtered by variants (e.g., MAF, depth) to suit actual study requirements.

SNPs with a low MAF value are more susceptible to sequencing error due to a smaller number of observations in the rare homozygote and heterogeneity regions which may result in the inconstant estimation of population structure (Mathieson and McVean 2012; De la Cruz and Raska 2014). As a consequence, many researchers have chosen to remove markers with a small MAF prior to analysis, although pre-selection of high-frequency SNPs may sometimes result in bias in population genetic studies. However, SNPs with low MAF values may be useful for identifying selection in non-model organisms with sufficiently large sample sizes (Ahrens *et al.* 2018). MAF thresholds have often been set to "default" values (i.e., 0.05 or 0.10), and most studies fail to provide justifications for their chosen MAF threshold. In general, SNPs with MAF values > 5% are included in GWAS, and typically, those with MAF ≥ 1% (i.e., as a "polymorphism") are retained for regular SNP calling approaches (Bodmer and Bonilla 2008). In this study, because the relatedness estimation is commonly contributed by SNPs with a higher MAF, removal of SNPs with MAF ≤ 0.03 was likely to be beneficial. With regard to CR, given the fact that cut-offs as high as 95% or an even higher threshold of 99% are often

employed, SNPs with a CR less than 10% were initially removed using UNEAK, followed by another filter of "SNPs calling across all regions" achieved in KGD. In addition, genomic prediction accuracy is likely to benefit from including more individuals with many SNPs at relatively low depth rather than a limited number individuals with high depth (Zhan *et al.* 2011; Gorjanc *et al.* 2015). This approach is supported by Dodds *et al.* (2015) who argued that SNPs with a relativity low sequencing depth had satisfactory quality for relatedness analyses with the optimal (smoothed) depth as 2–4, yet a lower sequencing depth (1–5) was also desirable.

5.4.3 Comparison of libraries

Overall, the total number of filtered bi-allelic SNPs per dataset differed depending on the sequencing depth, which was the highest for SQ0642-7 (48-plex) and lowest for SQ0503-1 (272-plex). Genotype disconcordance was strongly dependant on the unevenness of the sample size differences amongst the squat lobster populations. Although SQ0503-8 and SQ0503-1 were constructed with the same number of individuals, SQ0503-8 performed much better than SQ0503-1 concerning the number of sequences and SNPs obtained. One possible explanation for the inconsistency may be attributed to DNA degradation. GBS for SQ0503-1 was accomplished almost two months after SQ0503-8 was performed, suggesting that temporal change in quality and quantity for some DNA samples may fail to meet the GBS standard. High-quality genomic DNA that is free of contamination (e.g., RNA) is crucial to the success of GBS, given that efficiency of digestion, ligation and amplification can vary in the final marker set (Davey *et al.* 2011). Elshire *et al.* (2011) pointed out that for any multiplex sequencing protocol, accurate quantification of high molecular weight DNA samples is the most likely source of sample-to-sample variation in sequence coverage. Another reason for the lack of congruence may be due to inconsistent pipetting during robotic liquid handling. The subset of 48-plex (SQ0642-7) was additionally genotyped in order to obtain greater mean sample depth. Regardless of the disproportionate sample representation, a merged library was created containing sequencing data from three single libraries. As a result, considerable improvement in read counts and SNPs were gained. Interestingly, these values were still lower than anticipated, perhaps depicting the fact that the squat lobster *M. isos* is likely to have a large genome. This interpretation is in line with other studies in squat lobster genome size prediction that species from the superfamily Galatheoidea have relatively large genomes compared to other anomurans (Bonnivard *et al.* 2009).

5.4.4 Population genetic pattern and connectivity

This study is the first to add SNPs to examine the widespread marine invertebrate squat lobster *M. isos* within the southwest Pacific Ocean. Even in low-depth sequencing, GBS still successfully generated accurate information for application in relatedness and genetic differentiation analyses. By using large SNP datasets and sampling of populations of the squat lobster *M. isos* across the southwest Pacific Ocean, population differentiation was successfully identified (Figure 5.15). As revealed by GRM, panels of retained SNPs from different filters in four datasets (merged, SQ0503-8, SQ0503-1 and SQ0642-7 libraries) suggested similar genetic patterns which were roughly clustered with the geographical localities examined. Two main clusters were detected with populations from the Tasmanian slope and Macquarie Ridge being grouped together, compared to a cluster of populations from all others in the eastern New Zealand region (Kermadec Ridge, Hikurangi Margin, Chatham Rise, and Louisville Seamount Chain). To confirm the clustering identified by GRM, PCA was utilised. As expected, a similar pattern of genetic differentiation amongst squat lobster populations was revealed by the PCA-based approach. Although this study was conducted based on geomorphic feature scale, a Northern-Southern biogeographical provinces differentiation was observed. Similarly, a province-scale differentiation has been reported in branching stony corals using a combination of D-loop, ITS and microsatellite markers (Zeng *et al.* 2017) and in a sponge species as revealed by COI, Cytb and microsatellite variations within the New Zealand region (Zeng *et al.* 2019).

Notably, four individuals from Macquarie Ridge were observed to be present in the Tasmanian slope cluster, and amongst which, three were collected from the Eltanin Seamount at a relatively low depth of 506–560 m, suggesting potential gene flow derived from the west (i.e., Tasmanian slope) to the east (i.e., Macquarie Ridge) and implying that there are barriers to gene flow at deeper depths. Given that populations from both Tasmanian slope and Macquarie Ridge were genetically different from all others in the eastern New Zealand region, the Macquarie Ridge may act as a southern geographical barrier to gene flow from the southwest to the eastern/northeast New Zealand region. Within the study region, the Subantarctic Water (SAW) has a roughly northeastwards flow across the south Tasman Sea, passing over the Macquarie Ridge, before meeting the Subtropical Water (STW) along the Chatham Rise, forming the Subtropical Front (STF) (Chiswell *et al.* 2015), which is consistent with the estimated gene flow from the Tasmanian slope to the Macquarie Ridge. Moreover, the STF may act as a barrier to pelagic larval dispersal northwards of the Chatham Rise, leading to

a pattern of genetic subdivision amongst populations of the Tasmanian slope and Macquarie Ridge from all other populations to the east.

In contrast, no evidence of genetic differentiation was observed amongst squat lobster populations in the eastern New Zealand region, which was similar to previous results based on the COI region and microsatellite variations (see Chapter 4) and other comparable studies (e.g., Bors *et al.* (2012); Zeng *et al.* (2019)). This lack of genetic differentiation pattern amongst squat lobster populations within and beyond New Zealand's EEZ might be considered surprising given the large geographic distances between the sampling sites in this study. Nonetheless, the potentially long PLD of the *M. isos* that lasts for perhaps several months is but one of various possible explanations for the panmixia observed in the eastern New Zealand region.
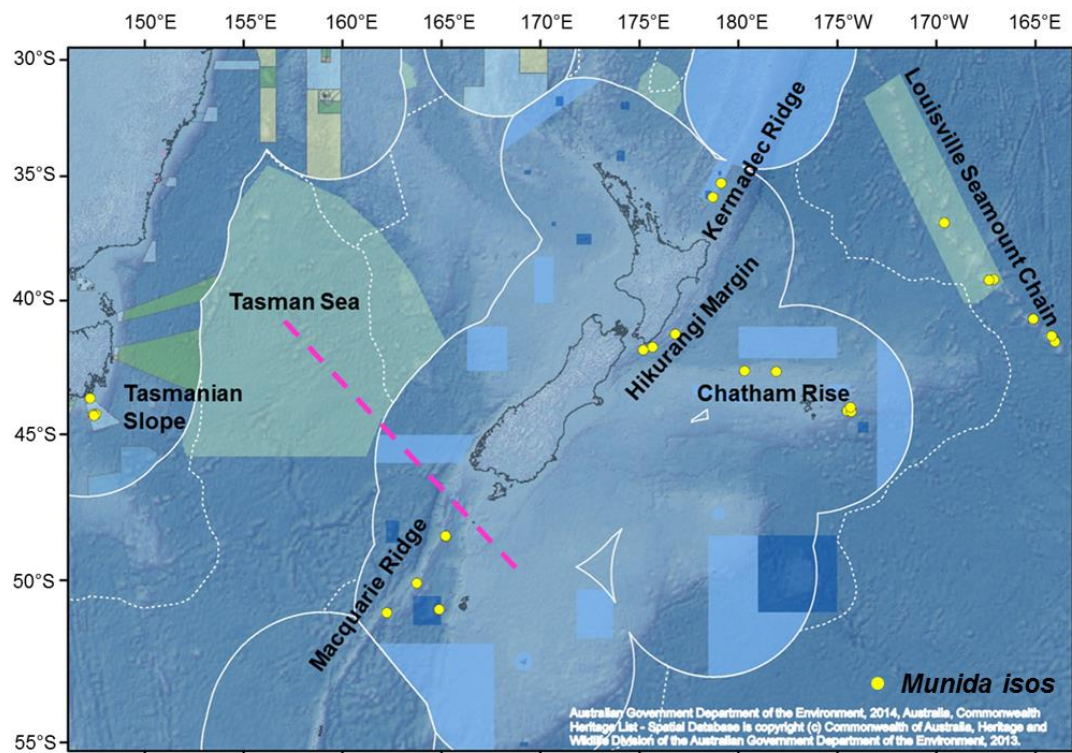
Figure 5.19 Map showing the distribution of marine protected areas, major geomorphic features and locations of the *M. isos* samples in the southwest Pacific Ocean region.

The dashed line represents the potential genetic break amongst populations of *M. isos* within the study area.

5.4.5 Molecular marker performance

In this study, SNPs outperformed COI and microsatellites to reveal a genetic pattern for squat lobster populations, indicating increased genetic differentiation in the SNP panels for genomic studies. Assessments drawn from mtDNA sequences are limited by the nature of maternal inheritance, whereas microsatellites suffer from variable mutation patterns and null alleles that increase ambiguity in analyses (Morin *et al.* 2004). The statistical power to detect population differentiation is related to the total number of alleles. Kalinowski (2002) pointed out that the discriminatory power of ~100 neutral SNPs is thought to be very roughly equivalent to 10–20 microsatellites. Whilst Morin *et al.* (2009) believed that ~30 SNPs would be sufficient for moderate differentiation detection, ≥ 80 SNPs and large sample size are needed for assessing demographic independence. Whilst mutation rates for SNPs are in general considerably lower than microsatellites, the multi-locus genotypes at these loci might differ in their utility in genetics research (Foll and Gaggiotti 2008). A number of findings suggest that SNPs tend to achieve better insight into kinship assignment and population clustering compared to microsatellites (e.g., Liu *et al.* (2005); Smith and Seeb (2008); Glover *et al.* (2010); Hauser *et al.* (2011)), which is coincident with that revealed by the squat lobster populations. However, the most informative SNPs may sometimes rival (or even exceed) the average information content of microsatellites in assessing broad-scale population differentiation and assignment tests, yet microsatellites still perform better than SNPs with fine-scale structure estimation (Liu *et al.* 2005; Rengmark *et al.* 2006; Smith *et al.* 2007; Narum *et al.* 2008; Hess *et al.* 2011; DeFaveri *et al.* 2013). This conclusion was supported by Haasl and Payseur (2011) with an added emphasis on the statement that fewer unlinked microsatellites are needed for the evaluation of recent divergence times. As such, the value of microsatellites remains apparent in relatedness estimation and fine-scale population differentiation, owing to their high level of polymorphism and fast mutation rate, which is in contrast to the slower evolving biallelic SNPs (DeFaveri *et al.* 2013). In addition, as genetic markers are generally derived from transcribed elements, temporal and spatial patterns may not reflect neutral evolutionary history but are likely to have been shaped by natural selection (Oetjen and Reusch 2007). Generally, stronger divergence, lower variability and higher frequency were observed for molecular markers in genic regions compared with non-genic regions. Therefore, seeking SNPs that are under selection through neutrality testing is necessary to create a comparison with the patterns revealed by the combination of non-genic and genic SNPs in future research.

In this study, with the exclusion of potential outlier SNPs, genetic relatedness and differentiation patterns remain the same for all datasets. Whilst the possibility of some SNPs being under selection cannot be ruled out, this scenario is unlikely due to the inclusion of SNP markers that are under selection. Besides, non-neutral markers can be useful for individual assignment and are robust for distinguishing groups of conservation concern (Seeb *et al.* 2011; Wei *et al.* 2013a; Gagnaire *et al.* 2015). Given that not all genotyped loci are necessary for contributing assignment power, future studies should focus on the exploration of a smaller set of PCA-correlated SNPs that explicitly capture the observed regional genetic structure by assigning individuals into particular clusters. Identifying such markers will provide insight into genetic loci that have undergone selective pressure, and will also reduce genotyping costs (Paschou *et al.* 2007).

5.4.6 Implications for conservation and management

In this study, the assessment of relatedness and population connectivity using GBS SNPs data provides new insights into the connectivity patterns of squat lobster populations in the southwest Pacific Ocean. The results showed that the Tasmanian slope and Macquarie Ridge populations are genetically different from all other populations to the east, both within New Zealand's EEZ and beyond. Thus, both regions (i.e., Tasmanian slope and Macquarie Ridge) would ideally be managed separately, especially for the Tasmanian slope population, which might be source populations to those on the Macquarie Ridge. Implications for conservation and management with regard to the Tasmanian slope have been discussed in Chapter 4. In this study, apart from the Tasmanian slope, populations from the Macquarie Ridge were found to be genetically distinct as well. Although no seamount closure or BPA is established on the Macquarie Ridge, there is one BPA to the north (Puysegur, BPA8) with a bottom dredging/trawling prohibition, and two seamount closures to the west (Seamount FMA5) and east (Seamount FMA6/2) also with trawling prohibitions, all of which might contain a representative pool for the protection of Macquarie Ridge populations. Particularly, squat lobsters from the rest of the New Zealand region and beyond that have high levels of gene flow highlight the importance to maintain evolutionary diversity (through self-recruitment) across eastern New Zealand's EEZ and beyond. Thus it is also essential to have MPAs distributed throughout these regions. This situation highlights the complexity of trying to protect population diversity and gene flow when not all areas fall within a single national

jurisdiction, and cross over into international waters in which there may or may not be spatial protective measures (e.g., Miller *et al.* (2018); van Dover *et al.* (2018)).

5.5 Conclusion

The standard GBS protocol and UNEAK pipeline have been employed to develop a large number of SNPs for *M. isos* samples from six regions in the southwest Pacific Ocean, at relatively low cost and in the absence of a reference genome. SNPs developed in this study will facilitate genetics and genomics studies for the VME-associated species *M. isos* in the future. Furthermore, the genetic data may be relevant to inform management of VMEs at a variety of spatial scales.

Appendix

Table S5.1 Summary of collection sites from the southwest Pacific Ocean from which the squat lobster, *M. isos*, was collected for genomic analysis.

Geomorphic features	Code	Collection Registration Number	Collection Station Number	Sample Date	Latitude (S)	Longitude (positive = E, negative = W)	Depth (m)	SQ0503-8 / SQ0503-1	SQ0642-7
Kermadec Ridge	KR	NIWA 65036	TAN1007/123	8/06/2010	-35.284	178.863	1244–1276	2	2
Kermadec Ridge	KR	NIWA 72936	TAN1104/124	19/03/2011	-35.857	178.448	1237–1460	8	7
Louisville Seamount Chain	LS	NIWA 94257	TAN1402/57	15/02/2014	-36.908	-169.846	1010–1013	7	2
Louisville Seamount Chain	LS	NIWA 94395	TAN1402/91	20/02/2014	-39.164	-167.35	910–934	8	1
Louisville Seamount Chain	LS	NIWA 94404	TAN1402/97	21/02/2014	-39.196	-167.59	1082–1090	2	1
Louisville Seamount Chain	LS	NIWA 94414	TAN1402/97	21/02/2014	-39.196	-167.59	1082–1090	1	1
Louisville Seamount Chain	LS	NIWA 94437	TAN1402/111	23/02/2014	-40.717	-165.343	814–912	6	1
Louisville Seamount Chain	LS	NIWA 94457	TAN1402/112	23/02/2014	-40.718	-165.343	768–913	2	
Louisville Seamount Chain	LS	NIWA 94541	TAN1402/137	28/02/2014	-41.566	-164.261	1060–1061	20	
Louisville Seamount Chain	LS	NIWA 94560	TAN1402/138	28/02/2014	-41.582	-164.255	1223–1241	7	1
Louisville Seamount Chain	LS	NIWA 94609	TAN1402/156	2/03/2014	-41.365	-164.419	1220–1250	20	1
Hikurangi Margin	HM	NIWA 31950	SO191-3/316	18/03/2007	-41.291	176.551	728–762	2	2
Hikurangi Margin	HM	NIWA 32003	SO191-2/138	18/02/2007	-41.78	175.401	1043–1059	2	2
Hikurangi Margin	HM	NIWA 64201	TAN1004/134	27/04/2010	-41.896	174.927	913–948	1	1
Chatham Rise	CR	NIWA 33690	TAN0705/189	19/04/2007	-42.705	-178.34	997–1042	1	1
Chatham Rise	CR	NIWA 53012	TAN0905/6	14/06/2009	-42.674	-179.955	150–1030	30	2
Chatham Rise	CR	NIWA 53802	TAN0905/104	26/06/2009	-44.158	-174.562	681–913	13	2
Chatham Rise	CR	NIWA 54079	TAN0905/113	27/06/2009	-44.15	-174.757	519–609	11	2
Chatham Rise	CR	NIWA 54178	TAN0905/116	27/06/2009	-44.175	-174.552	716–745	30	1
Chatham Rise	CR	NIWA 54367	TAN0905/121	28/06/2009	-44.028	-174.591	801–823	18	1

Geomorphic features	Code	Collection Registration Number	Collection Station Number	Sample Date	Latitude (S)	Longitude (positive = E, negative = W)	Depth (m)	SQ0503-8 / SQ0503-1	SQ0642-7
Macquarie Ridge	MR	NIWA 39574	TAN0803/17	30/03/2008	-48.549	164.963	1318–1327	2	2
Macquarie Ridge	MR	NIWA 39597	TAN0803/19	30/03/2008	-48.532	164.948	1060–1112	3	2
Macquarie Ridge	MR	NIWA 39762	TAN0803/33	1/04/2008	-50.091	163.482	1077–1408	7	2
Macquarie Ridge	MR	NIWA 39983	TAN0803/52	4/04/2008	-51.043	161.981	506–560	5	2
Macquarie Ridge	MR	NIWA 46520	TAN0306/6	14/04/2003	-50.943	164.609	1105–1140	13	1
Tasmanian slope	TA	NMV J68478	SS01/2008 25	17/01/2008	-44.277	147.266	900	7	2
Tasmanian slope	TA	NMV J68560	SS01/2008 11	14/01/2008	-44.28	147.138	1260	10	2
Tasmanian slope	TA	NMV J67888	SS02/2007 14	1/04/2007	-44.321	147.178	1150–1280	12	2
Tasmanian slope	TA	NMV J67902	SS02/2007 03	29/03/2007	-43.692	146.976	100–110	22	2

Chapter 6. General Discussion

The primary objectives of my research study were to investigate patterns of population structure and genetic connectivity of three deep-sea squat lobster species (*Munida isos* Ahyong & Poore, 2004, *Munida endeavourae* Ahyong & Poore, 2004 and *Munida gracilis* Henderson, 1885) in the southwest Pacific Ocean. The former two species are nearly exclusively found on seamounts, while the latter is found mostly on the soft-sediments of the New Zealand plateaux and rises, but the three species can be associated with vulnerable marine ecosystem (VME) habitats. Comparing molecular patterns across their distributional ranges and habitats will contribute to the management and conservation of VMEs in the study area. A spatially explicit hierarchical testing framework and three genetic marker types (most of them developed in this study) were employed for the investigation. This study presents the first attempt of genome-wide single nucleotide polymorphism (SNP) analysis in squat lobsters (i.e., *M. isos*), contributes a broader understanding of patterns of connectivity in the southwest Pacific Ocean, and provides new information useful to the management of VMEs.

In this chapter, major findings of this study are summarised and reviewed, and considered in light of their contributions to the knowledge of deep-sea population genetics in the southwest Pacific Ocean. Management implications are highlighted, and future research directions are presented.

6.1 Major findings

6.1.1 Molecular marker performance

In recent years, with the development of sequencing technologies, the costs of Sanger sequencing, microsatellite development, and genotyping have decreased substantially, permitting a move from the analysis of single-locus datasets towards multilocus datasets (generally with DNA sequence data and microsatellite markers) (Thaler *et al.* 2014; Zeng *et al.* 2017). Presently, population genetic studies of marine taxa have gradually shifted from population genotyping to genome typing (Luikart *et al.* 2003), utilising High-Throughput Sequencing (HTS) technologies to develop large numbers of SNPs in a variety of model and non-model organisms (e.g., Wells and Dale (2018)). Unlike microsatellite data which needs to be subjectively genotyped for comparison between individual studies, HTS data can be applied in conjunction with other HTS data across studies for future use (e.g., genome assembly), similar to one of the benefits of DNA sequence data (Taylor and Roterman 2017).

In this study, three molecular genetic marker types, including mitochondrial COI region, nuclear microsatellites and SNPs, were utilised to evaluate the genetic connectivity amongst populations of three *Munida* species (*M. isos*, *M. endeavourae* and *M. gracilis*). Universal invertebrate primers (Folmer *et al.* 1994) were used to yield partial COI fragments for the three *Munida* species, whilst microsatellites and SNPs (developed by Genotyping by Sequencing (GBS)) were newly developed for *M. isos*, both based on an Illumina HiSeq 2500 sequencing platform.

In many respects, the species is the fundamental unit of biological science, but despite this, there is still much debate what exactly constitutes a species (Ghiselin 2002). Because accurate taxonomy is critical for population genetic studies, particularly when used to support management decisions, the mitochondrial COI marker was utilised for species delimitation in the three *Munida* species (see below). As a result, a number of misidentified specimens were identified from each *Munida* species and excluded from the study, highlighting the utility of barcoding for identification and taxonomy before population genetic analyses. The existence of cryptic species has been resolved in many squat lobster taxa, particularly in 'cosmopolitan' species or monospecific genera, based on molecular data (e.g., Machordom and Macpherson (2004); Macpherson and Machordom (2005)). Such taxa have been recognised as constituting a species complex, which are morphologically similar but genetically very distinct. Thus, further investigations are required to confirm the taxonomy of the individuals excluded from the analysis.

From a population genetics perspective, high haplotypic diversity and moderate nucleotide diversity were revealed by COI sequence data, yet there was no evidence of genetic subdivision amongst populations of the three *Munida* species. This pattern of high genetic diversity, paired with no genetic differentiation, was consistent with findings from some other deep-sea squat lobsters, such as *Munida gregaria* (Fabricius, 1793) (Pérez-Barros *et al.* 2014), *Kiwa tyleri* Thatje, 2015 (Roterman *et al.* 2016) and *Pleuroncodes monodon* (H. Milne Edwards, 1837) (Haye *et al.* 2010). High levels of haplotypic diversity paired with relatively lower levels of nucleotide diversity can be indicative of a relatively fast mutation rate, isolation of the taxa for extended periods allowing for mutations to accumulate (Coykendall *et al.* 2017), or a history of recent population expansion following bottlenecks (Grant and Bowen 1998).

Microsatellite locus development using the enrichment library method has been problematic in some marine invertebrates (Selkoe and Toonen 2006; Bailie *et al.* 2010), whilst HTS technologies have enabled marker development for species characterised by complex

genomes, such as crustaceans (Bailie *et al.* 2010). Cross-species amplification is a cost-effective and time-saving alternative to the development of new primers, especially for species with limited genetic information (e.g., Boyle *et al.* (2013); Wang and Held (2015)). In the present study, novel microsatellites were developed for *M. isos* and cross-amplifications were conducted in two congeners. Despite moderate levels of genetic diversity, little genetic differentiation was observed in populations of *M. isos* and *M. gracilis* based on microsatellite variation, whilst no genetic subdivision was detected in *M. endeavourae*. In contrast, for *M. isos*, populations from the Tasmanian slope were potentially genetically different from all other populations. Collectively, *M. isos* populations exhibited relatively lower genetic diversity compared to *M. endeavourae* and *M. gracilis* populations based on mitochondrial COI region data, whilst *M. endeavourae* populations showed the lowest genetic diversity according to nuclear microsatellite variation, which might be due to the smaller sample sizes for this species.

Genomic data can be sourced from a variety of available sequencing platforms, and GBS is now starting to prove its efficiency in SNPs discovery in a variety of non-model marine invertebrates (e.g., Wells and Dale (2018)). Here, three GBS libraries were successfully constructed for *M. isos* utilising the double enzyme *PstI-MspI* approach, followed by the use of a powerful non-reference-genome bioinformatics pipeline for SNPs filtering. Pronounced genetic differentiation was observed between populations from the Tasmanian slope and Macquarie Ridge, and all other populations to the east of these features. This finding was in contrast to that based on the COI region data for which no population genetic differentiation was observed, and for microsatellite variation for which little population genetic subdivision was detected. Thus, SNPs data, with higher resolution than either the microsatellite or the COI sequence data sets, enabled fine-scale patterns of connectivity to be observed in the non-model deep-sea species (*M. isos*) in the southwest Pacific Ocean. Future analyses of genetic statistic estimation, outlier SNPs characterisation and genetic differentiation evaluation based on obtained SNPs data are expected.

In short, although mitochondrial COI region data failed to reveal any genetic pattern, it did permit confirmation of the taxonomic status of all individuals for each *Munida* species, which was critical for the population genetic studies. With an increase in marker resolution, microsatellite variation started to show some signal of regional level divergence. Subsequently, with higher quality data and broader genome coverage, SNPs showed evidence of pronounced genetic structure amongst squat lobster populations. As demonstrated in my study, SNPs have the potential for higher genome-wide coverage, genotyping efficiency, data quality and

analytical simplicity and have become the marker of choice in population ecology, evolution and conservation genetics (Morin *et al.* 2004).

6.1.2 Patterns of population structure and connectivity in the southwest Pacific Ocean

To date, the few studies exploring the pattern of connectivity amongst populations of deep-sea organisms (e.g., corals and sponges) in the southwest Pacific Ocean have revealed sharply different patterns of genetic differentiation on the basis of different marker types (e.g., Herrera *et al.* (2012); Miller and Gunasekera (2017); Zeng *et al.* (2017)). In this study, population genetic information based on three marker types (i.e. COI, microsatellites and SNPs) has broadened our knowledge of connectivity patterns of squat lobster populations for three *Munida* species across their distributional range in the southwest Pacific Ocean. Based on COI sequence data, despite the sizeable geographic coverage of sampling sites, genetic homogeneity was observed in the squat lobster populations of the three *Munida* species within the southwest Pacific Ocean. Similarly, although the microsatellite markers possessed sufficient variation, limited genetic differentiation was observed amongst populations of either *M. isos* or *M. gracilis*, whilst no genetic subdivision was detected amongst populations of *M. endeavourae*. In addition to this, more than 96% of the variance was attributed to differences within populations based on microsatellite variation, again suggesting high levels of genetic homogeneity. Notably, the homogenising process is usually consistent with either panmixia (i.e., gene flow), recent demographic expansions or selective sweeps, or lack of time for the accumulation of new mutations after founder events. This pattern of no genetic differentiation observed in COI region sequences, paired with some genetic subdivision detected through microsatellite variations, was consistent with a previous study on the squat lobster *Munidopsis lauensis* from southwest Pacific back-arc basins using a similar approach (i.e., COI and microsatellites) (Thaler *et al.* 2014). In contrast, SNPs revealed pronounced genetic structure amongst squat lobster populations in the study area (Figure 6.1).
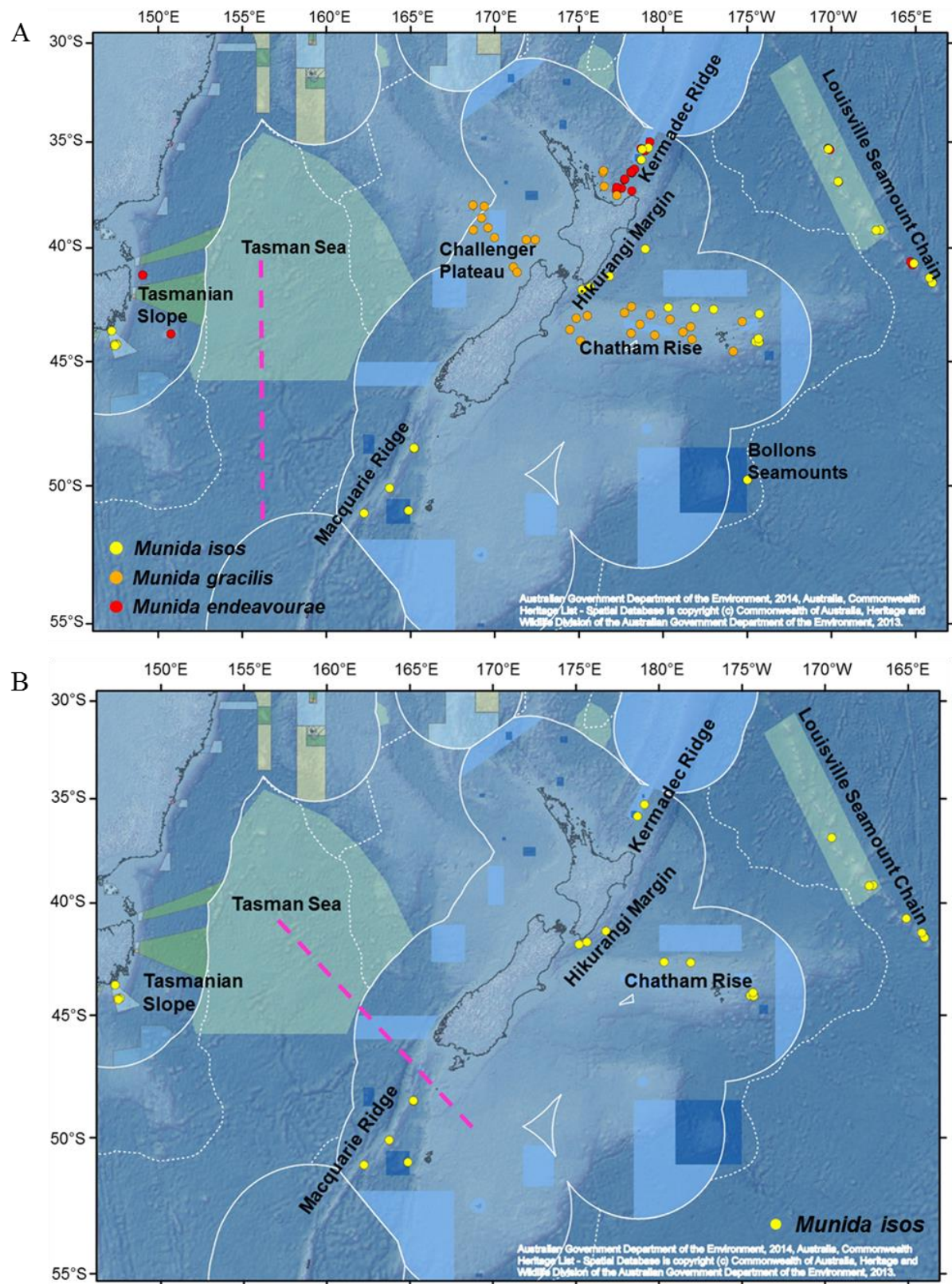
Figure 6.1 Maps showing the distribution of marine protected areas, major geomorphic features and locations of the samples in the southwest Pacific Ocean region.

The dashed line represents the potential genetic break amongst populations of *M. isos* based on microsatellite variations (A) and SNPs (B).

Given that limited information on genetic differentiation was observed in *M. endeavourae* and *M. gracilis* based on both marker types (i.e., COI region and microsatellite variations), the following discussion regarding genetic structure and connectivity is only focussed on *M. isos*, for which I have the largest data set (number of individuals) and also the best genetic variation data. For *M. isos*, microsatellite variation revealed that populations from the Tasmanian slope were potentially genetically different from all other populations within the New Zealand region (including the Macquarie Ridge population). However, according to the SNPs data, strong evidence was observed that populations from the Tasmanian slope and the Macquarie Ridge were differentiated and genetically isolated from all other populations, and that populations from the east of New Zealand region (both inside and outside the EEZ) were genetically undifferentiated. These results highlight the power and potential utility of SNPs, in that they revealed a previously unseen pattern of genetic structure and differentiation for *M. isos* in the southwest Pacific Ocean. The divergent estimates from different marker types highlight one of the major challenges in conservation genetics that no single gene or a suite of genetic loci can provide a complete picture of population genetic structure (Avise 1994). Ultimately, results of this type reported by scientists are only as good as the markers employed, and this as a major limitation has to be recognised.

Patterns of genetic differentiation (Tasmanian slope and Macquarie Ridge) and homogeneity (eastern New Zealand region) inferred from different marker types may have multiple explanations, but which are likely to a broad extent controlled by the planktonic duration of the larval dispersal phase and the influence of large-scale oceanographic features (e.g., major currents and fronts) and/or smaller scale topographically-forced hydrological features associated with geomorphic features (e.g., Taylor Columns, upwelling and down-welling) which may facilitate or impede dispersal (Bossart and Pashley Prowell 1998). *Munida* species have free-swimming planktonic larval phases that may drift with ocean currents, but the larval dispersal strategies of the three *Munida* species studied here are not known in detail and are inferred from laboratory rearing studies of related species. It is presumed that they have a similar development to other munidids, which includes 4–6 planktotrophic zoeal larval stages and 15–83 days of larval development during which larvae may disperse on ocean currents (Baba *et al.* 2011). SNPs data revealed that four individuals from the Macquarie Ridge population were present in the Tasmanian slope population cluster, suggesting potential gene flow derived from the west (i.e., Tasmanian slope) to the east (e.g., Macquarie Ridge). Given that populations from both geomorphic features were genetically different from all other

populations in the eastern New Zealand region, hydrographic barriers to gene flow are likely to be limiting migration or colonisation of populations to the east. Within the study region, two major surface current circulations, Subtropical Water (STW) and Subantarctic Water (SAW), are separated by the Subtropical Front (STF) (Chiswell *et al.* 2015). Notably, the SAW has a roughly easterly flow across the south Tasman Sea, passing over the Macquarie Ridge, which is consistent with the estimated gene flow from the Tasmanian slope to the Macquarie Ridge, dispersing the larvae eastward. After passing over the Macquarie Ridge, the deep-moving SAW is diverted southwards by the Campbell Plateau before it continues northwards and eastwards as it moves around the plateau. Therefore, the ability of this current to disperse larvae to sampled populations east of New Zealand is limited. In addition, while the more shallow-moving STW could potentially transport larvae eastwards from the Tasman Sea, the STF may act as a barrier to larval dispersal northwards of the Chatham Rise, and also contribute to the pattern of genetic subdivision amongst populations of the Tasmanian slope and Macquarie Ridge from all other populations to the east. Collectively, it can be concluded that both historical and contemporary processes, including life-history traits, marine current patterns and bathymetry, may contribute in shaping the patterns genetic differentiation amongst squat lobster populations in the southwest Pacific Ocean, and examining whether these patterns apply more broadly across the other species of *Munida* and, indeed, other VME-associated taxa is necessary.

6.1.3 Genetic diversity and demographic expansion

Demographic history was assessed for three *Munida* species based on the variation of the mitochondrial COI region and nuclear microsatellites, with sound evidence of recent demographic expansions detected in the three *Munida* species, based on both marker types, indicating the likelihood of recent demographic expansions. For the COI region, significantly negative Fu's F_S, unimodal mismatch distributions and non-significant *Hri* scores were consistent across the three *Munida* species, suggesting a departure from neutral expectations, and that exponential demographic expansions or selective sweeps could not be rejected. However, given that a selective sweep is expected to reduce haplotype diversity, something that was not observed, this is indicative of recent demographic expansions in the three *Munida* species. This hypothesis of population expansions was further supported by microsatellite variation, with a statistically significant) pattern of heterozygote deficiency observed under

both SMM and TPM in populations in all provinces, regions and geomorphic features in the three *Munida* species (the only exception was the Hikurangi Margin population, which might be biased due to its small sample size). To date, evidence of recent population expansions following bottlenecks have been reported for a number of squat lobster species, such as *P. monodon* (Haye *et al.* 2010), *Munida intermedia* A. Milne Edwards & Bouvier, 1899 (García-Merchán *et al.* 2012), *M. gregaria* (Pérez-Barros *et al.* 2008; Pérez-Barros *et al.* 2014) and *K. tyleri* (Roterman *et al.* 2016).

In this study, the estimated time of demographic expansions for the three *Munida* species from the southwest Pacific Ocean was congruent with estimations for *M. gregaria* from the southern Pacific and Atlantic Oceans (Pérez-Barros *et al.* 2014), *M. intermedia* from the Atlantic–Mediterranean transition area (García-Merchán *et al.* 2012) and *K. tyleri* from the Southern Ocean (Roterman *et al.* 2016) that occurred during the middle to late Pleistocene. The Pleistocene was characterised by repeated cycles of glaciation, which have influenced the genetic composition of many organisms (Petit *et al.* 2003; Fraser *et al.* 2012). Distributional changes of deep-sea organisms, in particular, may be attributed to the severe climatic fluctuations as produced by glaciation events (Hewitt 2000). It was hypothesised that their distributions might have been restricted to lower latitudes during the periods of glaciation, followed by population expansion and recolonisation with the increase of sea level and seawater temperature after the glaciations (Pérez-Barros *et al.* 2008).

6.1.4 Study limitations

Museums and natural history collections worldwide house millions of specimens that can be used for population genetic studies. For example, the major marine collections in New Zealand are housed at the Museum of New Zealand Te Papa Tongarewa and the NIWA Invertebrate Collection, both in Wellington. With the rapid development of molecular approaches, these collections have become sources of many conservation genetic studies (Wandeler *et al.* 2007). Nonetheless, one practical limitation of deep-sea population genetic studies is that relatively few samples have been historically preserved in an appropriate way for genetic analyses. For example, many deep-sea specimens are initially preserved in formalin before transferring to ethanol for long-term fixation (Boschen *et al.* 2015; Taylor and Roterman 2017). Formalin can degrade DNA into small fragments (< 200 bp) and only short reads can be generated from sequencing. This limitation has hampered population genetic studies in

many deep-sea taxa, because high molecular weight DNA with limited degradation is a prerequisite, especially for HTS technologies (Elshire *et al.* 2011; Etter *et al.* 2011). Generally, high-quality genome sequencing is required in genome-level studies to obtain large numbers of *de novo* SNPs and high-quality data. Therefore, genome-wide SNPs analysis may not be applicable for archived samples, whilst microsatellites are more appropriate to obtain high-quality genetic data using preserved specimens. Nonetheless, with correct preservation methods and high-quality DNA extractions, large amounts of genetic data can be generated through HTS technology, even in non-model organisms (Helyar *et al.* 2011). In this study, considering that old museum specimens might be problematic due to DNA degradation, only those collected since the year 2000 were considered for analysis. This approach has been successful, and good-quality DNA was extracted from the majority of squat lobster samples, and high-quality genomic DNA samples were obtained for the downstream population genetic studies.

The deep-sea is an environment where traditional ecological data collection approaches are extremely difficult, and the collection of sufficient numbers of specimens for robust comparative analyses is the primary challenge (Taylor and Roterman 2017). The huge cost and the logistical difficulty of sampling have resulted in relatively low sampling effort in the deep sea compared to terrestrial and shallow-water fauna, hampering the production of high-resolution genetic data sets to aid in resource management and conservation endeavours (Taylor and Roterman 2017). Thus, small sample size and limited spatial scale sampling are unavoidable constraints in population genetic studies of deep-sea organisms (Baco *et al.* 2016; Boschen *et al.* 2016). Generally, the minimum number of samples required for statistically robust analyses in population genetic studies is determined by a variety of factors, such as the number of loci and their level of diversity (Landguth *et al.* 2012). Given the limitations of specimen collection in the deep sea, it has been challenging to obtain even a modest number of samples (~30) from many collection sites or regions. For microsatellite-based population genetic studies, 25–30 individuals per population is considered sufficient to provide sufficient coverage (Hale *et al.* 2012). Nonetheless, with a significant increase in the numbers of markers that are now offered by population genomics (i.e., 1000s SNP loci rather than 12–20 microsatellite loci), 10–20 individuals per population has been proposed to be sufficient to obtain useful information from a population, even though the use of more individuals is advised (Willing *et al.* 2012). In this study, for instance, *M. isos* populations from the Hikurangi Margin have most probably produced biased results in several analyses based on microsatellite

variation because of its small sample size ($N = 11$), whilst SNPs data have produced robust results regardless of the small sample size. Thus, the HTS technologies can be applied for SNPs development in population studies of deep-sea organisms, particularly in those with insufficient sample sizes for statistically robust analyses utilising other marker types (e.g., DNA sequence data and microsatellites) (Reitzel *et al.* 2013; Taylor and Roterman 2017). As noted earlier, the application of HTS technology such as GBS-derived SNPs, highlights the power and utility of a combined marker approach to encompass disciplines such as taxonomy, population genetics and connectivity. With new collection conducted with the consideration of molecular technologies, more specimens of the species considered here, and other relevant VME-associated taxa will be available to expand this study.

Squat lobsters, as a morphologically and ecologically diverse group, have been the focus of intense taxonomic, systematic and phylogenetic studies (Baba 2005; Ahyong *et al.* 2011b; Ahyong *et al.* 2015b). *Munida* is by far the most speciose genus within the squat lobster group and has captivated considerable attention (Baba 2005; Baba *et al.* 2008). Nonetheless, the internal relationships within *Munida* remain largely unresolved (Machordom and Macpherson 2004; Cabezas *et al.* 2011a; Coykendall *et al.* 2017). In this study, three VME-associated taxa (*M. isos*, *M. endeavourae* and *M. gracilis*) were selected to evaluate the connectivity patterns across their distributional ranges in the southwest Pacific Ocean. To date, limited phylogenetic and/or phylogeographic information is available for the three *Munida* species. For example, Coykendall *et al.* (2017) conducted a phylogenetic analysis of 59 *Munida* species, including *M. gracilis* sampled from the northern Atlantic Ocean that nested with other *Mundia* species, based on a combined molecular and morphological dataset. Additionally, *M. isos* was selected as a representative of the family Munididae in constructing a phylogenetic tree of infraorders Anomura and Brachyura, and was clustered with the other congener *M. gregaria* with maximum support (BPP = 1.00, SH = 100%, UFBoot = 100%) (Tan *et al.* 2018). Given the relatively wide distributional ranges of the three *Munida* species, and the fact that New Zealand and southern Australian squat lobster fauna represents the only distinctly temperate biogeographic region globally (Schnabel *et al.* 2011b), it is feasible and necessary to investigate the phylogenetic and phylogeographic status of the three *Munida* species in a broader context (e.g., *Munida* phylogram with species worldwide) in the near future, utilising the obtained genetic data in this study. More information refers to 6.3.2.

6.2 Management implications

In recent years, population genetic studies have been increasingly undertaken for VME-indicator taxa (Miller *et al.* 2010; Miller *et al.* 2011; Herrera *et al.* 2012; Miller and Gunasekera 2017; Zeng *et al.* 2017; Holland *et al.* 2018; Zeng *et al.* 2019) to inform management and conservation strategies (Palumbi 2003; Hilário *et al.* 2015), and to help facilitate the protection and restoration of biodiversity and ecosystem function (Lubchenco *et al.* 2003; Clark *et al.* 2011). However, relatively few studies have focussed on VME-associated taxa such as deep-sea squat lobsters, whose distribution is not exclusively limited to VMEs. In this study, the assessment of genetic diversity and population connectivity using different marker types provides new insights into the connectivity patterns of VME-associated taxa in the southwest Pacific Ocean, which can be useful for assessing the potential effectiveness of current MPAs and the design of future MPAs alongside assessments based on VME-indicator taxa.

In New Zealand's EEZ, seventeen seamount closure areas (encompassing 19 seamounts) and 17 benthic protection areas (BPAs) were established in 2001 and 2007, respectively, to protect benthic fauna (including VMEs) from anthropogenic activities (Brodie and Clark 2003; Helson *et al.* 2010). In this study, most of the geomorphic features used in the spatial hierarchical analyses currently contain one or more seamount closures or BPAs. Although no genetic differentiation was observed amongst populations of *M. isos* from eastern New Zealand's EEZ, as well as all the populations of *M. endeavourae* and *M. gracilis* within New Zealand's EEZ, it is still essential to have MPAs distributed throughout New Zealand's EEZ to recognise and protect the high levels of gene flow amongst populations and thus maintain the evolutionary diversity of species associated with VMEs. For *M. isos*, populations from the Macquarie Ridge were genetically distinct from all other populations to the east within New Zealand's EEZ, and would therefore ideally be managed separately. Although no seamount closure or BPA is established on the Macquarie Ridge, there is one BPA to the north (Puysegur, BPA8) with a bottom dredging/trawling prohibition, and two seamount closures to the west (Seamount FMA5) and east (Seamount FMA6/2) also with trawling prohibitions, all of which might contain a representative pool for the protection of Macquarie Ridge populations.

The Tasmanian slope populations are located within the Australian EEZ. In this study, populations of *M. isos* from the Tasmanian slope were found to be genetically isolated from all other populations to the east within New Zealand's EEZ and might be source populations to those on the Macquarie Ridge. In addition to this, potential gene flow was observed from the

west (i.e., Tasmanian slope) to the east (e.g., Macquarie Ridge). Therefore, the populations of the Tasmanian slope should be treated as a separate biological unit, and in turn, its protection would be helpful for the conservation of the genetic diversity of this species on the Macquarie Ridge. Populations from the Tasmanian slope are located in the 9991 km^2 Huon Commonwealth Marine Reserve (part of Australia's South-east Commonwealth Marine Reserve Network). Three of the sample sites are within the Habitat Protection Zone (IUCN IV), with mining prohibition and commercial fishing if authorised, whilst the fourth sample site for this population is situated in the Huon-Multiple Use Zone (IUCN VI), where activities such as mining and fishing can occur if authorised (Iftekhar *et al.* 2014). Thus, the results of this study not only demonstrate the usefulness of the Huon Marine Park as an MPA to potentially protect VMEs in Australian waters, but also illustrates the importance of considering the international cross-boundary implications of conservation measures when trying to design effective MPA networks (van Dover *et al.* 2018).

The Louisville Seamount Chain, located within the South Pacific Regional Fisheries Management Organization (SPRFMO) Convention Area to the northeast of the New Zealand's EEZ, is an important area for New Zealand commercial fisheries (e.g., orange roughy) and is known and predicted to host VME-indicator taxa and VME habitat (Clark *et al.* 2014; Anderson *et al.* 2016; Rowden *et al.* 2017). Currently, New Zealand has established spatial closures of 40% of the total bottom trawl fishing footprint (distributed among 20-minute blocks of closed areas) along the Louisville Seamount Chain to protect VMEs (Penney *et al.* 2009; Penney and Guinotte 2013). In this study, no genetic differentiation was observed between the *Munida* populations of the Louisville Seamount Chain and those within New Zealand's EEZ. However, it is necessary to separately protect VMEs on the Louisville Seamount Chain, because the impacts of bottom trawling on VMEs in the SPRFMO Convention Area are managed separately from the management of similar impacts in New Zealand's EEZ. Furthermore, the currently closed areas on the Louisville Seamount Chain are part of a broader protection measure that not only distributes VME protection elsewhere in the SPRFMO Convention Area, but also includes limits on VME indicator taxa bycatch (including VME habitat associated taxa). Ongoing monitoring of genetic diversity and patterns of connectivity within and amongst populations of deep-sea fauna before and after bottom trawling would allow researchers to evaluate the impact of this anthropogenic activity on population differentiation and connectivity, and provide valuable information to inform best practices in management and conservation.

The aim of this research is to produce scientific information that can assist New Zealand, Australia and SPRFMO initiatives on spatial management in the southwest Pacific Ocean, and particularly New Zealand's EEZ. The results of the study highlight the complexity of trying to protect population genetic diversity and gene flow when not all areas fall within single national jurisdiction, but cross over into international waters. This, in turn, highlights the need for a flexible spatial management system to maintain functional connectivity for various deep-sea taxa at different spatial scales, when new information comes to light (e.g., Miller *et al.* (2018); van Dover *et al.* (2018)).

6.3 Future research directions

6.3.1 Understanding the drivers of population connectivity

In the marine environment, genetic connectivity is often maintained through the planktonic dispersal process, and patterns of gene flow can be influenced by oceanographic features that can facilitate or prevent diffusion and mixing of pelagic larvae, and biogeographical boundaries that may act as barriers to larval dispersal (Wares *et al.* 2001; Weersing and Toonen 2009; White *et al.* 2010). Dispersal of individuals amongst marine populations is critical with respect to population persistence, metapopulation dynamics and species expansion, and an understanding of connectivity amongst distant populations has a crucial impact upon management and conservation (Ross *et al.* 2016). One of the most commonly used indirect approaches for interpreting connectivity is the assessment of population genetic structure (Hellberg *et al.* 2002), which has been implemented in the present study. In addition to this, biophysical dispersal models, incorporating biological characteristics of larval dispersal together with dynamic physical oceanographic data, can be used to simulate spatial patterns of larval dispersal at different spatial scales (e.g., (Treml *et al.* 2008; Ross *et al.* 2016; Mertens *et al.* 2018). Integrated interdisciplinary approaches that pair high-resolution biophysical modelling and empirical data are likely to provide a greater understanding of the observed dispersal patterns. The use of "network modelling", on the basis of three connectivity typologies (i.e., structural, functional and realised), has been proposed as a means to predict marine connectivity using a simple, yet flexible and robust framework for investigating the flow of individuals amongst habitat patches or population (Treml and Kool 2018). Future studies assessing the connectivity of squat lobster populations, should use hydrodynamic

connectivity models along with network models that are informed by the results of genetic connectivity analyses (e.g., molecular markers) as produced by the likes of the present study.

As a relatively recent method in marine ecology, seascape genetics/genomics is the integration of genetic analyses, biological parameters of dispersal and oceanographic information to assess environmental drivers of genetic differentiation patterns (Hansen and Hemmer-Hansen 2007). To date, the first wave of seascape genetics has broadened our conceptual understanding of a variety of historical and contemporary processes that contribute to the dispersal and gene flow in marine systems (Selkoe *et al.* 2016). Notably, amongst a number of seascape factors, oceanography, temperature and geography exhibit equal prevalence in shaping the spatial connectivity patterns (Selkoe *et al.* 2016). New Zealand contains a diverse seafloor bathymetry and topography as well as complex oceanic dynamics, making it an ideal region to test the potential effect of environmental factors on deep-sea connectivity. Within New Zealand, seascape genetic studies have been conducted on a number of shallow-water marine taxa, such as greenshell mussels (Wei *et al.* 2013b) and scallops (Silva and Gardner 2015), and on deep-sea taxa such as corals and sponges (Zeng PhD thesis). Squat lobsters have distributions that encompass a range of environments and depths and would be good model taxa to examine the effects of environmental variations on population genetic structure in the future. Seascape genetics has the power to answer questions that we have not previously been able to answer, and just as importantly, to be used in a predictive modelling framework (Selkoe *et al.* 2016; Taylor and Roterman 2017).

6.3.2 Understanding the phylogeny and phylogeography of squat lobsters

In the present study, novel SNPs were developed for *M. isos* and utilised in relatedness and population genetic differentiation evaluation. The obtained SNPs data not only provide valuable insights into the management and conservation of VMEs in the southwest Pacific Ocean but can facilitate future genetics research in deep-sea squat lobsters.

Nowadays, the boundaries amongst population genetics, phylogenetics and phylogeography are increasingly obscure (Knowles and Maddison 2002). Being one of the most morphologically and ecologically diverse groups in the Decapoda, squat lobsters have attracted much attention over the past several decades. Molecular-based studies have reshaped our understanding of evolutionary and phylogenetic relationships of the squat lobster group, leading to the confirmation of the taxonomic status of many taxa that both corroborate and defy

traditional morphological views and has resulted in the reclassifications of major taxonomic groups (Ahyong *et al.* 2011b). In this study, a number of misidentified specimens, and possibly two species requiring taxonomic validation, were detected after utilising mitochondrial COI marker for species delimitation, emphasising the utility of this approach to ongoing work on the inventory of squat lobsters in this region. The New Zealand and southern Australian squat lobster fauna is unique in that it represents the only distinctly temperate biogeographic region globally (Schnabel *et al.* 2011b). Using the population genetics approach here, this study lends support to a united trans-Tasman Sea fauna (*M. isos*), but further studies will be required to examine this hypothesis more broadly, in particular considering other VME-associated species with different larval or other biological characteristics. These might include e.g., members of the Chirostylidae such as *Uroptychus* that have an abbreviated larval development and are typically commensals of gorgonian and hard corals, or members of the Munidopsidae such as *Munidopsis*, which also have an abbreviated larval development and can sometimes be found in large numbers on hydrothermal vents, cold seeps and whale and wood falls, habitats also considered as VME. As discussed under Limitations, the numbers of specimens are often not available for a robust comparison across the species' distribution range, but future collections across the region will allow for some of these questions to be addressed in more detail.

A robust taxonomy, accurate knowledge of the geographic distributions and their relationships in a wider phylogenic framework is very important to elucidate the evolution and diversification of any organism. For squat lobsters, the genetic information of three *Munida* species obtained in this study will contribute to future phylogenetic evaluations, the identification of patterns of speciation, genealogy reconstruction, and divergence time estimation in the New Zealand region and within the context of the surrounding temperate and tropical regions.

6.4 Conclusion

The use of new markers such as SNPs and the application of new technologies such as whole-genome sequencing for many different deep-sea taxa will soon be possible, both technically and economically, and when applied appropriately may well provide entirely new insights into patterns of genetic structure and connectivity (as in the use of SNPs for *M. isos*), and will thereby aid new management strategies to better protect the VMEs in the southwest Pacific Ocean region.

References

Abdelkrim J, Robertson BC, Stanton JAL, Gemmell NJ (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *Biotechniques* **46**:185–191

Adams DK, Arellano SM, Govenar B (2012) Larval dispersal: vent life in the water column. *Oceanography* **25**:256–268

Agardy MT (1994) Advances in marine conservation: the role of marine protected areas. *Trends in Ecology & Evolution* **9**:267–270

Ahrens CW, Rymer PD, Stow A, Bragg J, Dillon S, Umbers KDL, Dudaniec RY (2018) The search for loci under selection: trends, biases and progress. *Molecular Ecology* **27**:1342–1356

Ahyong ST, O'Meally D (2004) Phylogeny of the Decapoda Reptantia: Resolution using three molecular loci and morphology. *Raffles Bulletin of Zoology* **52**:673–693

Ahyong ST, Poore GCB (2004) Deep-water Galatheidae (Crustacea: Decapoda: Anomura) from southern and eastern Australia. *Zootaxa* **472**:1–76

Ahyong ST, Baba K, Macpherson E, Poore GCB (2010) A new classification of the Galatheoidea (Crustacea: Decapoda: Anomura). *Zootaxa* **2676**:57–68

Ahyong ST, Andreakis N, Taylor J (2011a) Mitochondrial phylogeny of the deep-sea squat lobsters, Munidopsidae (Galatheoidea). *Zoologischer Anzeiger* **250**:367–377

Ahyong ST, Schnabel KE, Macpherson E (2011b) Phylogeny and fossil record of marine squat lobsters. In: Poore GCB, Ahyong ST, Taylor J (eds.), The Biology of Squat Lobsters. CSIRO Publishing, Melbourne, pp 73–104

Ahyong ST, Schnabel KE, Baba K (2015a) Southern high latitude squat lobsters: Galatheoidea and Chirostyloidea from Macquarie Ridge with description of a new species of *Uroptychus. Records of the Australian Museum* **67**:109–128

Ahyong ST, Schnabel KE, Maas EW (2015b) Anomuran phylogeny: new insights from molecular data. In: Martin JW, Crandall KA, Felder DL (eds.), Decapod Crustacean Phylogenetics. CRC Press, Boca Raton, Florida, pp 399–414

Al-Breiki RD et al. (2018) Genome-wide SNP analyses reveal high gene flow and signatures of local adaptation among the scalloped spiny lobster (*Panulirus homarus*) along the Omani coastline. *BMC Genomics* **19**:690

Althaus F et al. (2009) Impacts of bottom trawling on deep-coral ecosystems of seamounts are long-lasting. *Marine Ecology Progress Series* **397**:279–294

Anderson OF, Guinotte JM, Rowden AA, Clark MR, Mormede S, Davies AJ, Bowden DA (2016) Field validation of habitat suitability models for vulnerable marine ecosystems in the South Pacific Ocean: Implications for the use of broad-scale models in fisheries management. *Ocean & Coastal Management* **120**:110–126

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 4 December 2018

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a F_{ST}-outlier method. *BMC Bioinformatics* **9**:323

Arellano SM, Van Gaest AL, Johnson SB, Vrijenhoek RC, Young CM (2014) Larvae from deep-sea methane seeps disperse in surface waters. *Proceedings of the Royal Society B* **281**:20133276

Avise JC (1994) Molecular Markers, Natural History and Evolution. Chapman & Hall, New York

Baba K (1972) A new Species of the Galatheidean Crustacea from the Ryukyu Islands (Decapoda, Anomura). *Memoirs of the Faculty of Education, Kumamoto University, Section 1 (Natural Science)* **20**:42–48

Baba K (1974) Four new species of galatheidean Crustacea from New Zealand waters. *Journal of the Royal Society of New Zealand* **4**:381–393

Baba K (1979) Expédition Rumphius II (1975) Crustacés parasites, commensaux, etc. (Th. Monod et R. Sèrene, éd.) VII. Galatheid crustaceans (Decapoda, Anomura). *Bulletin du Muséum national d'Histoire naturelle, Paris (4e série) Section A* **1**:643–657

Baba K, de Saint Laurent M (1996) Crustacea Decapoda: Revision of the genus *Bathymunida* Balss, 1914, and description of six new related genera (Galatheidae). In: Crosnier A (ed.), Résultats des Campagnes MUSORSTOM. Mémoires du Muséum National d'Histoire Naturelle, Paris, pp 433–502

Baba K (2005) Deep-sea chirostylid and galatheid crustaceans (Decapoda: Anomura) from the Indo-Pacific, with a list of species. *Galathea Report* **20**:1–317

Baba K et al. (2008) Catalogue of squat lobsters of the world (Crustacea: Decapoda: Anomura—families Chirostylidae, Galatheidae and Kiwaidae). *Zootaxa* **1905**:1–220

Baba K, Fujita Y, Wehrtmann IS, Scholtz G (2011) Developmental biology of squat lobsters. In: Poore GCB, Ahyong ST, Taylor J (eds.), The Biology of Squat Lobsters. CSIRO Publishing, Melbourne, pp 105–148

Baba K (2018) Chirostylidae of the Western and Central Pacific: *Uroptychus* and a new genus (Crustacea: Decapoda: Anomura). *Tropical Deep-Sea Benthos* **30**:1–612

Baba K, Ahyong ST, Schnabel KE (2018) Rediagnosis of the squat lobster genus *Gastroptychus* Caullery, 1896, with a new genus *Sternostylus* and a new family Sternostylidae (Crustacea: Decapoda: Anomura: Chirostyloidea). *Zootaxa* **4524**:77–86

Baco AR, Etter RJ, Ribeiro PA, von der Heyden S, Beerli P, Kinlan BP (2016) A synthesis of genetic connectivity in deep-sea fauna and implications for marine reserve design. *Molecular Ecology* **25**:3276–3298

Baeza JA (2011) Squat lobsters as symbionts and in chemo-autotrophic environments. In: Poore GCB, Ahyong ST, Taylor J (eds.), The Biology of Squat Lobsters. CSIRO Publishing, Melbourne, pp 249–270

Bahassi EM, Stambrook PJ (2014) Next-generation sequencing technologies: breaking the sound barrier of human genetics. *Mutagenesis* **29**:303–310

Bailie DA, Fletcher H, Prodöhl PA (2010) High incidence of cryptic repeated elements in microsatellite flanking regions of galatheid genomes and its practical implications for molecular marker development. *Journal of Crustacean Biology* **30**:664–672

Baird NA et al. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One* **3**:e3376

Benestan LM et al. (2016) Conservation genomics of natural and managed populations: building a conceptual and practical framework. *Molecular Ecology* **25**:2967–2977

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

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Ostell J, Pruitt KD, Sayers EW (2018) GenBank. *Nucleic Acids Research* **46**:D41–D47

Bickford D et al. (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution* **22**:148–155

Bloor P, de Laguna IHB, Kemp SJ (2006) Highly polymorphic tetranucleotide microsatellite loci for the eastern Canary Island lizard, *Gallotia atlantica*. *Molecular Ecology Resources* **6**:737–739

Bodmer W, Bonilla C (2008) Common and rare variants in multifactorial susceptibility to common diseases. *Nature Genetics* **40**:695–701

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology* **74**:21–45

Bonnivard E, Catrice O, Ravaux J, Brown SC, Higuet D (2009) Survey of genome size in 28 hydrothermal vent species covering 10 families. *Genome* **52**:524–536

Bors EK, Rowden AA, Maas EW, Clark MR, Shank TM (2012) Patterns of deep-sea genetic connectivity in the New Zealand region: implications for management of benthic ecosystems. *PloS One* **7**:e49474

Boschen RE, Rowden AA, Clark MR, Gardner JPA (2015) Limitations in the use of archived vent mussel samples to assess genetic connectivity among seafloor massive sulfide deposits: A case study with implications for environmental management. *Frontiers in Marine Science* **2**:105

Boschen RE et al. (2016) A primer for use of genetic tools in selecting and testing the suitability of set-aside sites protected from deep-sea seafloor massive sulfide mining activities. *Ocean & Coastal Management* **122**:37–48

Bossart JL, Pashley Prowell D (1998) Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends in Ecology & Evolution* **13**:202–206

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**:314–331

Boyle EA, Thaler AD, Jacobson A, Plouviez S, Van Dover CL (2013) Characterization of 10 polymorphic microsatellite loci in *Munidopsis lauensis*, a squat-lobster from the southwestern Pacific. *Conservation Genetics Resources* **5**:647–649

Bracken-Grissom HD et al. (2009) The decapod tree of life: compiling the data and moving toward a consensus of decapod evolution. *Arthropod Systematics & Phylogeny* **67**:99–116

Bracken-Grissom HD et al. (2013) A comprehensive and integrative reconstruction of evolutionary history for Anomura (Crustacea: Decapoda). *BMC Evolutionary Biology* **13**:128

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**:2633–2635

Brodie S, Clark MR (2003) The New Zealand seamount management strategy–steps towards conserving offshore marine habitat. In: Beumer JP, Grant A, Smith DC (eds.), Aquatic Protected Areas: What Works Best and How Do We Know? Proceedings of the World Congress on Aquatic Protected Areas, Cairns, Australia, August 2002. Australian Society of Fish Biology, Cairns, Australia, pp 664–673

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* **5**:453–455

Brouard J-S, Boyle B, Ibeagha-Awemu EM, Bissonnette N (2017) Low-depth genotyping-by-sequencing (GBS) in a bovine population: strategies to maximize the selection of high quality genotypes and the accuracy of imputation. *BMC Genetics* **18**:32

Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *Journal of Molecular Evolution* **18**:225–239

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

Burland TG (2000) DNASTAR's lasergene sequence analysis software. In: Misener S, Krawetz SA (eds.), Bioinformatics Methods and Protocols. Humana Press, Totowa, NJ, pp 71–91

Cabezas P, Macpherson E, Machordom A (2008) A new genus of squat lobster (Decapoda: Anomura: Galatheidae) from the South West Pacific and Indian Ocean inferred from morphological and molecular evidence. *Journal of Crustacean Biology* **28**:68–75

Cabezas P, Bloor P, Acevedo I, Toledo C, Calvo M, Macpherson E, Machordom A (2009a) Development and characterization of microsatellite markers for the endangered anchialine squat lobster *Munidopsis polymorpha*. *Conservation Genetics* **10**:673–676

Cabezas P, Macpherson E, Machordom A (2009b) Morphological and molecular Descriptionon of new species of squat lobster (Crustacea: Decapoda: Galatheidae) from the Solomon and Fiji Islands (South-West Pacific). *Zoological Journal of the Linnean Society* **156**:465–493

Cabezas P, Macpherson E, Machordom A (2010) Taxonomic revision of the genus *Paramunida* Baba, 1988 (Crustacea: Decapoda: Galatheidae): a morphological and molecular approach. *Zootaxa* **2712**:1–60

Cabezas P, Lin C-W, Chan T-Y (2011a) Two new species of the deep-sea squat lobster genus *Munida* Leach, 1820 (Crustacea: Decapoda: Munididae) from Taiwan: morphological and molecular evidence. *Zootaxa* **3036**:26–38

Cabezas P, Macpherson E, Machordom A (2011b) *Allogalathea* (Decapoda: Galatheidae): a monospecific genus of squat lobster? *Zoological Journal of the Linnean Society* **162**:245–270

Cabezas P, Alda F, Macpherson E, Machordom A (2012a) Genetic characterization of the endangered and endemic anchialine squat lobster *Munidopsis polymorpha* from

Lanzarote (Canary Islands): management implications. *ICES Journal of Marine Science* **69**:1030–1037

Cabezas P, Sanmartín I, Paulay G, Macpherson E, Machordom A (2012b) Deep under the sea: unraveling the evolutionary history of the deep-sea squat lobster *Paramunida* (Decapoda, Munididae). *Evolution: International Journal of Organic Evolution* **66**:1878–1896

Cabezas P, Chan T-Y (2014) Deep-sea squat lobsters of the genus *Paramunida* Baba, 1988 (Crustacea: Decapoda: Munididae) from the Philippines Panglao 2004, Panglao 2005 and Aurora 2007 expeditions, with the description of three new species. *Raffles Bulletin of Zoology* **62**:302–316

Cabezas P, Macpherson E (2014) A new species of *Paramunida* Baba, 1988 from the Central Pacific Ocean and a new genus to accommodate *P. granulata* (Henderson, 1885). *ZooKeys* **425**:15–32

Carson HS, López-Duarte PC, Cook GS, Fodrie FJ, Becker BJ, DiBacco C, Levin LA (2013) Temporal, spatial, and interspecific variation in geochemical signatures within fish otoliths, bivalve larval shells, and crustacean larvae. *Marine Ecology Progress Series* **473**:133–148

Carter L, McCave IN, Williams MJM (2008) Circulation and water masses of the Southern Ocean: a review. *Developments in Earth and Environmental Sciences* **8**:85–114

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

Chase MR, Etter RJ, Rex MA, Quattro JM (1998) Bathymetric patterns of genetic variation in a deep-sea protobranch bivalve, *Deminucula atacellana*. *Marine Biology* **131**:301–308

Chen C, Mitchell SE, Elshire RJ, Buckler ES, El-Kassaby YA (2013) Mining conifers' mega-genome using rapid and efficient multiplexed high-throughput genotyping-by-sequencing (GBS) SNP discovery platform. *Tree Genetics & Genomes* **9**:1537–1544

Chiswell SM, Bostock HC, Sutton PJH, Williams MJM (2015) Physical oceanography of the deep seas around New Zealand: a review. *New Zealand Journal of Marine and Freshwater Research* **49**:286–317

Chu K-H, Tsang L-M, Ma K-Y, Chan T-Y, Ng PKL (2015) Decapod phylogeny: what can protein-coding genes tell us? In: Martin JW, Crandall KA, Felder DL (eds.), Decapod Crustacean Phylogenetics. CRC Press, Boca Raton, Florida, pp 89–100

Clark MR, Rowden AA (2009) Effect of deepwater trawling on the macro-invertebrate assemblages of seamounts on the Chatham Rise, New Zealand. *Deep Sea Research Part I: Oceanographic Research Papers* **56**:1540–1554

Clark MR et al. (2010) The ecology of seamounts: structure, function, and human impacts. *Annual Review of Marine Science* **2**:253–278

Clark MR, Watling L, Rowden AA, Guinotte JM, Smith CR (2011) A global seamount classification to aid the scientific design of marine protected area networks. *Ocean & Coastal Management* **54**:19–36

Clark MR et al. (2014) Voyage report of a survey of deep-sea habitats of the Louisville Seamount Chain (TAN1402). *NIWA Preliminary Project Report* **VMES133**:1–84

Clark MR, Bowden DA, Rowden AA, Stewart R (2019) Little evidence of benthic community resilience to bottom trawling on seamounts after 15 years. *Frontiers in Marine Science* **6**:63

Clobert J, Baguette M, Benton TG, Bullock JM (2012) Dispersal ecology and evolution. Oxford University Press, Oxford

Coleman MA, Chambers J, Knott NA, Malcolm HA, Harasti D, Jordan A, Kelaher BP (2011) Connectivity within and among a network of temperate marine reserves. *PloS One* **6**:e20168

Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9**:539

Cordes EE et al. (2016) Environmental impacts of the deep-water oil and gas industry: a review to guide management strategies. *Frontiers in Environmental Science* **4**:58

Costa FO, deWaard JR, Boutillier J, Ratnasingham S, Dooh RT, Hajibabaei M, Hebert PDN (2007) Biological identifications through DNA barcodes: the case of the Crustacea. *Canadian Journal of Fisheries and Aquatic Sciences* **64**:272–295

Costa R, Bisol PM (1978) Genetic variability in deep-sea organisms. *The Biological Bulletin* **155**:125–133

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

Coykendall DK, Morrison CL (2013) Polymorphic microsatellite markers developed from 454 pyro-sequencing in the cold water coral-associated squat lobster species, *Eumunida picta* (Chirostylidae: Eumunididae). *Conservation Genetics Resources* **5**:495–498

Coykendall DK, Nizinski MS, Morrison CL (2017) A phylogenetic perspective on diversity of Galatheoidea (*Munida*, *Munidopsis*) from cold-water coral and cold seep communities

in the western North Atlantic Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* **137**:258–272

Creasey S, Rogers A, Tyler P, Gage J, Jollivet D (2000) Genetic and morphometric comparisons of squat lobster, *Munidopsis scobina* (Decapoda: Anomura: Galatheidae) populations, with notes on the phylogeny of the genus *Munidopsis*. *Deep Sea Research Part II: Topical Studies in Oceanography* **47**:87–118

Crivello JF, Landers Jr DF, Keser M (2005) The genetic stock structure of the American lobster (*Homarus americanus*) in Long Island Sound and the Hudson Canyon. *Journal of Shellfish Research* **24**:841–848

D'Aloia CC, Bogdanowicz SM, Francis RK, Majoris JE, Harrison RG, Buston PM (2015) Patterns, causes, and consequences of marine larval dispersal. *Proceedings of the National Academy of Sciences* **112**:13940–13945

da Silva JM, Creer S, dos Santos A, Costa AC, Cunha MR, Costa FO, Carvalho GR (2011) Systematic and evolutionary insights derived from mtDNA COI barcode diversity in the Decapoda (Crustacea: Malacostraca). *PloS One* **6**:e19449

Dailianis T, Tsigenopoulos C, S Dounas C, Voultsiadou E (2011) Genetic diversity of the imperilled bath sponge *Spongia officinalis* Linnaeus, 1759 across the Mediterranean Sea: patterns of population differentiation and implications for taxonomy and conservation. *Molecular Ecology* **20**:3757–3772

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity* **93**:504–509

Danovaro R et al. (2008) Exponential decline of deep-sea ecosystem functioning linked to benthic biodiversity loss. *Current Biology* **18**:1–8

Danovaro R, Corinaldesi C, Dell'Anno A, Snelgrove PV (2017) The deep-sea under global change. *Current Biology* **27**:R461–R465

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* **12**:499–510

De Donato M, Peters SO, Mitchell SE, Hussain T, Imumorin IG (2013) Genotyping-by-sequencing (GBS): a novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. *PloS One* **8**:e62137

De la Cruz O, Raska P (2014) Population structure at different minor allele frequency levels. *BMC Proceedings* **8**:S55

DeFaveri J, Viitaniemi H, Leder E, Merilä J (2013) Characterizing genic and nongenic molecular markers: comparison of microsatellites and SNPs. *Molecular Ecology Resources* **13**:377–392

DeWoody J, Nason JD, Hipkins VD (2006) Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes* **6**:951–957

Diez MJ, Rojas-Quiroga ML, Pérez-Barros P, Lezcano A, Florentín O, Lovrich GA (2016) "Gregaria" to "Subrugosa," that is the Question: Shape Changes Under Laboratory Conditions in the Pelagic Morphotype of the Squat Lobster *Munida Gregaria* (Decapoda: Anomura: Munididae). *Journal of Crustacean Biology* **36**:530–537

Dixon CJ, Ahyong ST, Schram FR (2003) A new hypothesis of decapod phylogeny. *Crustaceana* **76**:935–975

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

Dodds KG, McEwan JC, Brauning R, Anderson RM, Stijn TC, Kristjánsson T, Clarke SM (2015) Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genomics* **16**:1047

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One* **6**:e19379

Etter RJ, Boyle EE, Glazier A, Jennings RM, Dutra E, Chase MR (2011) Phylogeography of a pan-Atlantic abyssal protobranch bivalve: implications for evolution in the Deep Atlantic. *Molecular Ecology* **20**:829–843

Etter RJ, Bower AS (2015) Dispersal and population connectivity in the deep North Atlantic estimated from physical transport processes. *Deep Sea Research Part I: Oceanographic Research Papers* **104**:159–172

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

Everett MV, Park LK, Berntson EA, Elz AE, Whitmire CE, Keller AA, Clarke ME (2016) Large-scale genotyping-by-sequencing indicates high levels of gene flow in the deep-sea octocoral *Swiftia simplex* (Nutting 1909) on the west coast of the United States. *PloS One* **11**:e0165279

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

Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Molecular Ecology* **16**:1149–1166

Faurby S, Barber PH (2012) Theoretical limits to the correlation between pelagic larval duration and population genetic structure. *Molecular Ecology* **21**:3419–3432

Foll M, Gaggiotti OE (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**:977–993

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoes R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**:294–299

Food and Agricultural Organisation of the United Nations (2009) International guidelines for the management of deep-sea fisheries in the high seas. FAO, Rome

Forest J, de Saint Laurent M, McLaughlin PA, Lemaitre R (2000) The marine fauna of New Zealand: Paguridea (Decapoda: Anomura) exclusive of the Lithodidae. *NIWA Biodiversity Memoirs* **114**:1–250

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation genetics. Cambridge University Press, Cambridge

Fraser CI, Nikula R, Ruzzante DE, Waters JM (2012) Poleward bound: biological impacts of Southern Hemisphere glaciation. *Trends in Ecology & Evolution* **27**:462–471

Freeland JR (2005) Molecular markers in ecology. In: Kirk H (ed.), Molecular Ecology. Wiley-Blackwell, West Sussex, pp 31–62

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

Gagnaire PA et al. (2015) Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications* **8**:769–786

García-Merchán VH et al. (2012) Phylogeographic patterns of decapod crustaceans at the Atlantic–Mediterranean transition. *Molecular Phylogenetics and Evolution* **62**:664–672

Gardner JPA, Zbawicka M, Westfall KM, Wenne R (2016) Invasive blue mussels threaten regional scale genetic diversity in mainland and remote offshore locations: the need for

baseline data and enhanced protection in the Southern Ocean. *Global Change Biology* **22**:3182–3195

Ghiselin MT (2002) Species concepts: the basis for controversy and reconciliation. *Fish and Fisheries* **3**:151–160

Gilg MR, Hilbish TJ (2003) The geography of marine larval dispersal: coupling genetics with fine-scale physical oceanography. *Ecology* **84**:2989–2998

Gjerde KM et al. (2016) Protecting Earth's last conservation frontier: scientific, management and legal priorities for MPAs beyond national boundaries. *Aquatic Conservation: Marine and Freshwater Ecosystems* **26**:45–60

Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PloS One* **9**:e90346

Glover KA, Hansen MM, Lien S, Als TD, Høyheim B, Skaala Ø (2010) A comparison of SNP and STR loci for delineating population structure and performing individual genetic assignment. *BMC Genetics* **11**:2

Gollner S et al. (2017) Resilience of benthic deep-sea fauna to mining activities. *Marine Environmental Research* **129**:76–101

Goocii JL, Schopf TJM (1972) Genetic variability in the deep sea: relation to environmental variability. *Evolution* **26**:545–552

Gordon DP, Beaumont J, MacDiarmid A, Robertson DA, Ahyong ST (2010) Marine biodiversity of *Aotearoa* New Zealand. *PloS One* **5**:e10905

Gore M et al. (2007) Evaluation of target preparation methods for single-feature polymorphism detection in large complex plant genomes. *Crop Science* **47**:S-135–S-148

Gorjanc G, Cleveland MA, Houston RD, Hickey JM (2015) Potential of genotyping-by-sequencing for genomic selection in livestock populations. *Genetics Selection Evolution* **47**:12

Goudet J (2002) FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3. https://www2.unil.ch/popgen/softwares/fstat.htm. Accessed 4 December 2018

Grant WAS, Bowen BW (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity* **89**:415–426

Gregory TR (2018) Animal Genome Size Database. http://www.genomesize.com/. Accessed 4 December 2018

Gu H et al. (2010) Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nature Methods* **7**:133–136

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**:361–372

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

Halanych KM (2005) Molecular phylogeny of siboglinid annelids (a.k.a. pogonophorans): a review. *Hydrobiologia* **535**:297

Hale ML, Burg TM, Steeves TE (2012) Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. *PloS One* **7**:e45170

Hamblin MT, Rabbi IY (2014) The effects of restriction-enzyme choice on properties of genotyping-by-sequencing libraries: a study in cassava (*Manihot esculenta*). *Crop Science* **54**:2603–2608

Hamilton LJ (2006) Structure of the subtropical front in the Tasman Sea. *Deep Sea Research Part I: Oceanographic Research Papers* **53**:1989–2009

Handayani M, Anggoro S, Widowati I, Hideyuki I (2014) Molecular ecology comparison of blue leg hermit crab (*Calcinus elegans*) based on spatial factor in south coast of Java Island. *International Journal of Marine and Aquatic Resource Conservation and Co-existence* **1**:12–18

Hansen MM, Hemmer-Hansen J (2007) Landscape genetics goes to sea. *Journal of Biology* **6**:6

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* **66**:591–600

Hauser L, Baird M, Hilborn R, Seeb LW, Seeb JE (2011) An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. *Molecular Ecology Resources* **11**:150–161

Haye PA, Salinas P, Acuna E, Poulin E (2010) Heterochronic phenotypic plasticity with lack of genetic differentiation in the southeastern Pacific squat lobster *Pleuroncodes monodon*. *Evolution & Development* **12**:628–634

Hearne CM, Ghosh S, Todd JA (1992) Microsatellites for linkage analysis of genetic traits. *Trends in Genetics* **8**:288–294

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences* **270**:313–321

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

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**:313–318

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

Helson J, Leslie S, Clement G, Wells R, Wood R (2010) Private rights, public benefits: Industry-driven seabed protection. *Marine Policy* **34**:557–566

Helyar SJ et al. (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources* **11**:123–136

Hendry AP, Vamosi SM, Latham SJ, Heilbuth JC, Day T (2000) Questioning species realities. *Conservation Genetics* **1**:67–76

Henry CS (1994) Singing and cryptic speciation insects. *Trends in Ecology & Evolution* **9**:388–392

Herrera S, Shank TM, Sánchez JA (2012) Spatial and temporal patterns of genetic variation in the widespread antitropical deep-sea coral *Paragorgia arborea*. *Molecular Ecology* **21**:6053–6067

Hess JE, Matala AP, Narum SR (2011) Comparison of SNPs and microsatellites for fine-scale application of genetic stock identification of Chinook salmon in the Columbia River Basin. *Molecular Ecology Resources* **11**:137–149

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

Hilário A et al. (2015) Estimating dispersal distance in the deep sea: challenges and applications to marine reserves. *Frontiers in Marine Science* **2**:6

Ho SYW, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution* **22**:1561–1568

Ho SYW, Shapiro B, Phillips MJ, Cooper A, Drummond AJ (2007) Evidence for time dependency of molecular rate estimates. *Systematic Biology* **56**:515–522

Ho SYW, Lanfear R, Bromham L, Phillips MJ, Soubrier J, Rodrigo AG, Cooper A (2011) Time‐dependent rates of molecular evolution. *Molecular Ecology* **20**:3087‐3101

Holland LP, Rowden AA, Hamilton JZ, Clark MR, Chiswell SM, Gardner JPA (2018) Genetic connectivity of deep-sea corals in the New Zealand region. Ministry for Primary Industries, Wellington

Huang YF, Poland JA, Wight CP, Jackson EW, Tinker NA (2014) Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. *PloS One* **9**:e102448

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

Iftekhar MS, Tisdell JG, Gilfedder L (2014) Private lands for biodiversity conservation: Review of conservation covenanting programs in Tasmania, Australia. *Biological Conservation* **169**:176–184

Jackson JE (1991) A user's guide to principal components. Wiley-Interscience, New York

Johnson MS, Black R (1984) The Wahlund effect and the geographical scale of variation in the intertidal limpet *Siphonaria* sp.. *Marine Biology* **79**:295–302

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

Jombart T, Ahmed I (2011) *adegenet 1.3-1*: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**:3070–3071

Jones KC, Levine KF, Banks JD (2002) Characterization of 11 polymorphic tetranucleotide microsatellites for forensic applications in California elk (*Cervus elaphus canadensis*). *Molecular Ecology Resources* **2**:425–427

Jones WJ, Won YJ, Maas PAY, Smith PJ, Lutz RA, Vrijenhoek RC (2006) Evolution of habitat use by deep-sea mussels. *Marine Biology* **148**:841–851

Jones WJ, Macpherson E (2007) Molecular phylogeny of the East Pacific squat lobsters of the genus *Munidopsis* (Decapoda: Galatheidae) with the descriptions of seven new species. *Journal of Crustacean Biology* **27**:477–501

Jonsson LG, Nilsson PG, Floruta F, Lundälv T (2004) Distributional patterns of macro-and megafauna associated with a reef of the cold-water coral *Lophelia pertusa* on the Swedish west coast. *Marine Ecology Progress Series* **284**:163–171

Jost L, Archer F, Flanagan S, Gaggiotti O, Hoban S, Latch E (2018) Differentiation measures for conservation genetics. *Evolutionary Applications* **11**:1139–1148

Kalinowski ST (2002) How many alleles per locus should be used to estimate genetic distances? *Heredity* **88**:62–65

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

Kijas JM, Fowler JC, Garbett CA, Thomas MR (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *Biotechniques* **16**:656–660, 662

Kilgour MJ, Shirley TC (2008) *Eumunida picta* SI Smith, 1883, and *Lophelia pertusa* (Linnaeus, 1758): a relationship or just good friends? *Crustaceana* **81**:587–593

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology* **11**:2623–2635

Knowlton N (1986) Cryptic and sibling species among the Decapod Crustacea. *Journal of Crustacean Biology* **6**:356–363

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

Krawczak M (1999) Informativity assessment for biallelic single nucleotide polymorphisms. *Electrophoresis* **20**:1676–1681

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

Lance SL, Love CN, Nunziata SO, O'Bryhim JR, Scott DE, Flynn RW, Jones KL (2013) 32 species validation of a new Illumina paired-end approach for the development of microsatellites. *PloS One* **8**:e81853

Landguth EL et al. (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources* **12**:276–284

Larraín MA, Zbawicka M, Araneda C, Gardner JPA, Wenne R (2018) Native and invasive taxa on the Pacific coast of South America: Impacts on aquaculture, traceability and biodiversity of blue mussels (*Mytilus* spp.). *Evolutionary Applications* **11**:298–311

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics* **7**:295–302

Leathwick J et al. (2008) Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters* **1**:91–102

Lee CW, Song JH, Min GS, Kim S (2016a) The complete mitochondrial genome of squat lobster, *Munida gregaria* (Anomura, Galatheoidea, Munididae). *Mitochondrial DNA Part B* **1**:204–206

Lee S-H, Lee W-K, Won Y-J (2016b) A New Species of Yeti Crab, Genus *Kiwa* (Decapoda: Anomura: Kiwaidae), from a Hydrothermal Vent on the Australian-Antarctic Ridge. *Journal of Crustacean Biology* **36**:238–247

Leese F, Mayer C, Held C (2008) Isolation of microsatellites from unknown genomes using known genomes as enrichment templates. *Limnology and Oceanography: Methods* **6**:412–426

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

Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6**:1110–1116

Lemaitre R, McLaughlin PA (2009) Recent advances and conflicts in concepts of anomuran phylogeny (Crustacea: Malacostraca). *Arthropod Systematics & Phylogeny* **67**:119–135

Levin LA (2006) Recent progress in understanding larval dispersal: new directions and digressions. *Integrative and Comparative Biology* **46**:282–297

Levin LA, Le Bris N (2015) The deep ocean under climate change. *Science* **350**:766–768

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* **44**:397–401

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

Liu ZJ, Cordes JF (2004) DNA marker technologies and their applications in aquaculture genetics. *Aquaculture (Amsterdam, Netherlands)* **238**:1–37

Lotze HK, Coll M, Magera AM, Ward-Paige C, Airoldi L (2011) Recovery of marine animal populations and ecosystems. *Trends in Ecology & Evolution* **26**:595–605

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology* **19**:3038–3051

Lu F et al. (2013) Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *Plos Genetics* **9**:e1003215

Lubchenco J, Palumbi SR, Gaines SD, Andelman S (2003) Plugging a hole in the ocean: the emerging science of marine reserves. *Ecological Applications* **13**:3–7

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics* **4**:981–994

Lumsden SE, Hourigan TF, Bruckner AW, Dorr G (2007) The State of Deep Coral Ecosystems of the United States. NOA Technical Memorandum CRCP-3, Silver Spring, Maryland

MacDiarmid AB, Law CS, Pinkerton M, Zeldis J (2013) New Zealand marine ecosystem services. In: Dymond JR (ed.), Ecosystem services in New Zealand–Conditions and trends. Manaaki Whenua Press, Lincoln, New Zealand, pp 238–253

Machordom A, Macpherson E (2004) Rapid radiation and cryptic speciation in squat lobsters of the genus *Munida* (Crustacea, Decapoda) and related genera in the South West Pacific: molecular and morphological evidence. *Molecular Phylogenetics and Evolution* **33**:259–279

Macpherson E, Machordom A (2000) *Raymunida*, new genus (Decapoda: Anomura: Galatheidae) from the Indian and Pacific oceans. *Journal of Crustacean Biology* **20**:253–258

Macpherson E, Machordom A (2001) Phylogenetic relationships of species of *Raymunida* (Decapoda: Galatheidae) based on morphology and mitochondrial cytochrome oxidase sequences, with the recognition of four new species. *Journal of Crustacean Biology* **21**:696–714

Macpherson E, Jones W, Segonzac M (2005) A new squat lobster family of Galatheoidea (Crustacea, Decapoda: Anomura) from the hydrothermal vents of the Pacific-Antarctic Ridge. *Zoosystema* **27**:709–723

Macpherson E, Machordom A (2005) Use of morphological and molecular data to identify three new sibling species of the genus *Munida* Leach, 1820 (Crustacea, Decapoda, Galatheidae) from New Caledonia. *Journal of Natural History* **39**:819–834

Macpherson E, Baba K (2009) New species of squat lobsters of the genera *Agononida* and *Paramunida* (Crustacea: Decapoda: Anomura: Galatheidae) from the western Pacific. *Zootaxa* **2024**:56–68

Macpherson E, de Forges BR, Schnabel K, Samadi S, Boisselier M-C, Garcia-Rubies A (2010) Biogeography of the deep-sea galatheid squat lobsters of the Pacific Ocean. *Deep Sea Research Part I: Oceanographic Research Papers* **57**:228–238

Macpherson E, Baba K (2011) Taxonomy of squat lobsters. In: Poore GCB, Ahyong ST, Taylor J (eds.), The Biology of Squat Lobsters. CSIRO Publishing, Melbourne, pp 39–72

Macpherson E, Robainas-Barcia A (2013) A new genus and some new species of the genus *Lauriea* Baba, 1971 (Crustacea, Decapoda, Galatheidae) from the Pacific and Indian Oceans, using molecular and morphological characters. *Zootaxa* **3599**:136–160

Macpherson E, Robainas-Barcia A (2015) Species of the genus *Galathea* Fabricius, 1793 (Crustacea, Decapoda, Galatheidae) from the Indian and Pacific Oceans, with descriptions of 92 new species. *Zootaxa* **3913**:1–335

Macpherson E, Rodríguez-Flores PC, Machordom A (2017) New sibling species and new occurrences of squat lobsters (Crustacea, Decapoda) from the western Indian Ocean. *European Journal of Taxonomy* **34**:1–61

Mamanova L et al. (2010) Target-enrichment strategies for next-generation sequencing. *Nature Methods* **7**:111–118

Martin JW, Abele LG (1986) Phylogenetic relationships of the genus *Aegla* (Decapoda: Anomura: Aeglidae), with comments on anomuran phylogeny. *Journal of Crustacean Biology* **6**:576–612

Martin JW, Davis GE (2001) An updated classification of the recent Crustacea. Natural History Museum of Los Angeles County, Los Angeles, California

Mascher M, Wu S, Amand PS, Stein N, Poland J (2013) Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. *PloS One* **8**:e76925

Mathieson I, McVean G (2012) Differential confounding of rare and common variants in spatially structured populations. *Nature Genetics* **44**:243–246

Mccallum AW, Cabezas P, Andreakis N (2016) Deep-sea squat lobsters of the genus *Paramunida* Baba, 1988 (Crustacea: Decapoda: Munididae) from north-western Australia: new records and description of three new species. *Zootaxa* **4173**:201–224

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

McLaughlin PA (1983) A review of the phylogenetic position of the Lomidae (Crustacea: Decapoda: Anomala). *Journal of Crustacean Biology* **3**:431–437

McLaughlin PA, Lemaitre R, Sorhannus U (2007) Hermit crab phylogeny: a reappraisal and its "fall-out". *Journal of Crustacean Biology* **27**:97–115

Mengerink KJ et al. (2014) A call for deep-ocean stewardship. *Science* **344**:696–698

Mercier A, Sewell MA, Hamel JF (2013) Pelagic propagule duration and developmental mode: reassessment of a fading link. *Global Ecology and Biogeography* **22**:517–530

Mertens L, Treml EA, von Der Heyden S (2018) Genetic and biophysical models help define marine conservation focus areas. *Frontiers in Marine Science* **5**:268

Miller KA, Thompson KF, Johnston P, Santillo D (2018) An overview of seabed mining including the current state of development, environmental impacts and knowledge gaps. *Frontiers in Marine Science* **4**:418

Miller KJ, Williams A, Rowden AA, Knowles C, Dunshea G (2010) Conflicting estimates of connectivity among deep-sea coral populations. *Marine Ecology* **31**:144–157

Miller KJ, Rowden AA, Williams A, Häussermann V (2011) Out of their depth? Isolated deep populations of the cosmopolitan coral *Desmophyllum dianthus* may be highly vulnerable to environmental change. *PloS One* **6**:e19004

Miller KJ, Gunasekera RM (2017) A comparison of genetic connectivity in two deep sea corals to examine whether seamounts are isolated islands or stepping stones for dispersal. *Scientific Reports* **7**:46103

Mitarai S, Siegel DA, Watson JR, Dong C, McWilliams JC (2009) Quantifying connectivity in the coastal ocean with application to the Southern California Bight. *Journal of Geophysical Research: Oceans* **114**:C10026

Molecular Ecology Resources Primer Development Consortium et al. (2010) Permanent genetic resources added to Molecular Ecology Resources Database 1 June 2010–31 July 2010. *Molecular Ecology Resources* **10**:1106–1108

Morin PA, Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution* **19**:208–216

Morin PA, Martien KK, Taylor BL (2009) Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources* **9**:66–73

Moritz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics* **18**:269–292

Morrison CL, Harvey AW, Lavery S, Tieu K, Huang Y, Cunningham CW (2002) Mitochondrial gene rearrangements confirm the parallel evolution of the crab-like form. *Proceedings of the Royal Society of London B: Biological Sciences* **269**:345–350

Mullineau LS, Mills SW (1997) A test of the larval retention hypothesis in seamount-generated flows. *Deep Sea Research Part I: Oceanographic Research Papers* **44**:745–770

Nakajima Y et al. (2018) Isolation and characterization of novel polymorphic microsatellite loci for the deep-sea hydrothermal vent limpet, *Lepetodrilus nux*, and the vent-associated squat lobster, *Shinkaia crosnieri*. *Marine Biodiversity* **48**:677–684

Narum SR et al. (2008) Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. *Molecular Ecology* **17**:3464–3477

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**:153–170

Nielsen R, Paul JS, Albrechtsen A, Song YS (2011) Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics* **12**:443–451

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* **18**:375–402

Novembre J, Stephens M (2008) Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics* **40**:646–649

O'Hara TD, England PR, Gunasekera RM, Naughton KM (2014) Limited phylogeographic structure for five bathyal ophiuroids at continental scales. *Deep Sea Research Part I: Oceanographic Research Papers* **84**:18–28

O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS (2018) These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology* **27**:3193–3206

Oetjen K, Reusch TBH (2007) Genome scans detect consistent divergent selection among subtidal vs. intertidal populations of the marine angiosperm *Zostera marina*. *Molecular Ecology* **16**:5156–5157

Orsi AH, Whitworth III T, Nowlin Jr WD (1995) On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers* **42**:641–673

Palero F, Robainas-Barcia A, Corbari L, Macpherson E (2017) Phylogeny and evolution of shallow-water squat lobsters (Decapoda, Galatheoidea) from the Indo‐Pacific. *Zoologica Scripta* **46**:584‐595

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

Pandolfi JM et al. (2003) Global trajectories of the long-term decline of coral reef ecosystems. *Science* **301**:955–958

Parker PG, Snow AA, Schug MD, Booton GC, Fuerst PA (1998) What molecules can tell us about populations: choosing andusing a molecular marker. *Ecology* **79**:361–382

Parker SJ, Penney AJ, Clark MR (2009) Detection criteria for managing trawl impacts on vulnerable marine ecosystems in high seas fisheries of the South Pacific Ocean. *Marine Ecology Progress Series* **397**:309–317

Parker SJ, Bowden DA (2010) Identifying taxonomic groups vulnerable to bottom longline fishing gear in the Ross Sea region. *CCAMLR Science* **17**:105–127

Paschou P, Ziv E, Burchard EG, Choudhry S, Rodriguez-Cintron W, Mahoney MW, Drineas P (2007) PCA-correlated SNPs for structure identification in worldwide human populations. *Plos Genetics* **3**:e160

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

Penney AJ, Parker SJ, Brown JH (2009) Protection measures implemented by New Zealand for vulnerable marine ecosystems in the South Pacific Ocean. *Marine Ecology Progress Series* **397**:341–354

Penney AJ, Guinotte JM (2013) Evaluation of New Zealand's high-seas bottom trawl closures using predictive habitat models and quantitative risk assessment. *PloS One* **8**:e82273

Pérez-Barros P, Thatje S, Calcagno JA, Lovrich GA (2007) Larval development of the subantarctic squat lobster *Munida subrugosa* (White, 1847) (Anomura: Galatheidae), reared in the laboratory. *Journal of Experimental Marine Biology and Ecology* **352**:35–41

Pérez-Barros P, d'Amato ME, Guzman NV, Lovrich GA (2008) Taxonomic status of two South American sympatric squat lobsters, *Munida gregaria* and *Munida subrugosa* (Crustacea: Decapoda: Galatheidae), challenged by DNA sequence information. *Biological Journal of the Linnean Society* **94**:421–434

Pérez-Barros P, Lovrich GA, Calcagno JA, Confalonieri VA (2014) Is *Munida gregaria* (Crustacea: Decapoda: Munididae) a truly transpacific species? *Polar Biology* **37**:1413–1420

Pérez-Losada M, Jara CG, Bond-Buckup G, Crandall KA (2002a) Phylogenetic relationships among the species of *Aegla* (Anomura: Aeglidae) freshwater crabs from Chile. *Journal of Crustacean Biology* **22**:304–313

Pérez-Losada M, Jara CG, Bond-Buckup G, Porter ML, Crandall KA (2002b) Phylogenetic position of the freshwater anomuran family Aeglidae. *Journal of Crustacean Biology* **22**:670–676

Pérez-Losada M, Bond-Buckup G, Jara CG, Crandall KA (2004) Molecular systematics and biogeography of the southern South American freshwater "crabs" *Aegla* (Decapoda:

Anomura: Aeglidae) using multiple heuristic tree search approaches. *Systematic Biology* **53**:767–780

Petit RJ et al. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**:1563–1565

Pettersson E, Lundeberg J, Ahmadian A (2009) Generations of sequencing technologies. *Genomics* **93**:105–111

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**:502–503

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**:536–539

Poland J et al. (2012a) Genomic selection in wheat breeding using genotyping-by-sequencing. *The Plant Genome* **5**:103–113

Poland JA, Brown PJ, Sorrells ME, Jannink J-L (2012b) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PloS One* **7**:e32253

Poland JA, Rife TW (2012) Genotyping-by-sequencing for plant breeding and genetics. *The Plant Genome* **5**:92–102

Poore GCB, Andreakis N (2011) Morphological, molecular and biogeographic evidence support two new species in the *Uroptychus naso* complex (Crustacea: Decapoda: Chirostylidae). *Molecular Phylogenetics and Evolution* **60**:152–169

Poore GCB, Andreakis N (2012) The *Agononida incerta* species complex unravelled (Crustacea: Decapoda: Anomura: Munididae). *Zootaxa* **3492**:1–29

Poore GCB, Andreakis N (2014) More species of the *Agononida incerta* complex revealed by molecules and morphology (Crustacea: Decapoda: Anomura: Munididae). *Zootaxa* **3860**:201–225

Porter ML, Pérez-Losada M, Crandall KA (2005) Model-based multi-locus estimation of decapod phylogeny and divergence times. *Molecular Phylogenetics and Evolution* **37**:355–369

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

Puillandre N, Macpherson E, Lambourdière J, Cruaud C, Boisselier-Dubayle M-C, Samadi S (2011) Barcoding type specimens helps to identify synonyms and an unnamed new

species in *Eumunida Smith*, 1883 (Decapoda: Eumunididae). *Invertebrate Systematics* **25**:322–333

Pujolar JM et al. (2013) Understanding the effectiveness of marine protected areas using genetic connectivity patterns and Lagrangian simulations. *Diversity and Distributions* **19**:1531–1542

Ramirez-Llodra E et al. (2011) Man and the last great wilderness: human impact on the deep sea. *PloS One* **6**:e22588

Ramirez-Llodra EZ et al. (2010) Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. In: Gattuso JP, Kesselmeier J (eds.), Biogeosciences (Online). Copernicus Publications, Göttingen, pp 2851–2899

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution* **49**:1280–1283

Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology* **22**:2953–2970

Rengmark AH, Slettan A, Skaala Ø, Lie Ø, Lingaas F (2006) Genetic variability in wild and farmed Atlantic salmon (*Salmo salar*) strains estimated by SNP and microsatellites. *Aquaculture (Amsterdam, Netherlands)* **253**:229–237

Rice AL, Miller JE (1991) Chirostylid and galatheid crustacean associates of coelenterates and echinoderms collected from the *Johnson-Sea-Link* submersible, including a new species of *Gastroptychus*. *Proceedings of the Biological Society of Washington* **104**:299–308

Richer de Forges B, Koslow JA, Poore GCB (2000) Diversity and endemism of the benthic seamount fauna in the southwest Pacific. *Nature* **405**:944–947

Rippe JP et al. (2017) Population structure and connectivity of the mountainous star coral, *Orbicella faveolata*, throughout the wider Caribbean region. *Ecology and Evolution* **7**:9234–9246

Rodríguez-Flores PC, Machordom A, Macpherson E (2017) Three new species of squat lobsters of the genus *Fennerogalathea* Baba, 1988 (Decapoda: Galatheidae) from the Pacific Ocean. *Zootaxa* **4276**:46–60

Rodríguez-Flores PC, Macpherson E, Buckley D, Machordom A (2018) High morphological similarity coupled with high genetic differentiation in new sympatric species of coral-

reef squat lobsters (Crustacea: Decapoda: Galatheidae). *Zoological Journal of the Linnean Society* **185**:984–1017

Rodríguez-Flores PC, Macpherson E, Machordom A (2019) Revision of the squat lobsters of the genus *Leiogalathea* Baba, 1969 (Crustacea, Decapoda, Munidopsidae) with the description of 15 new species. *Zootaxa* **4560**:201–256

Rogers AD, Baco A, Griffiths H, Hall-Spencer JM (2007) Corals on seamounts. In: Morato T, Hart PJB, Clark MR, Haggan N, Santos RS (eds.), Seamounts: Ecology Fisheries and Conservation. Blackwell Scientific, Oxford, UK, pp 141–169

Rogers AD et al. (2012) The discovery of new deep-sea hydrothermal vent communities in the Southern Ocean and implications for biogeography. *PLoS Biology* **10**:e1001234

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**:552–569

Ross RE, Nimmo-Smith WAM, Howell KL (2016) Increasing the depth of current understanding: Sensitivity testing of deep-sea larval dispersal models for ecologists. *PloS One* **11**:e0161220

Roterman CN, Copley JT, Linse KT, Tyler PA, Rogers AD (2013a) The biogeography of the yeti crabs (Kiwaidae) with notes on the phylogeny of the Chirostyloidea (Decapoda: Anomura). *Proceedings of the Royal Society B* **280**:20130718

Roterman CN, Copley JT, Linse KT, Tyler PA, Rogers AD (2013b) Development of polymorphic microsatellite loci for three species of vent-endemic megafauna from deep-sea hydrothermal vents in the Scotia Sea, Southern Ocean. *Conservation Genetics Resources* **5**:835–839

Roterman CN, Copley JT, Linse KT, Tyler PA, Rogers AD (2016) Connectivity in the cold: the comparative population genetics of vent-endemic fauna in the Scotia Sea, Southern Ocean. *Molecular Ecology* **25**:1073–1088

Roterman CN, Lee W-K, Liu X, Lin R, Li X, Won Y-J (2018) A new yeti crab phylogeny: Vent origins with indications of regional extinction in the East Pacific. *PloS One* **13**:e0194696

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

Rowden AA et al. (2010a) A test of the seamount oasis hypothesis: seamounts support higher epibenthic megafaunal biomass than adjacent slopes. *Marine Ecology* **31**:95–106

Rowden AA, Schnabel KE, Schlacher TA, Macpherson E, Ahyong ST, de Forges BR (2010b) Squat lobster assemblages on seamounts differ from some, but not all, deep-sea habitats of comparable depth. *Marine Ecology* **31**:63–83

Rowden AA, Anderson OF, Georgian SE, Bowden DA, Clark MR, Pallentin A, Miller A (2017) High-resolution habitat suitability models for the conservation and management of vulnerable marine ecosystems on the Louisville Seamount Chain, South Pacific Ocean. *Frontiers in Marine Science* **4**:335

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34**:3299–3302

Samadi S, Bottan L, Macpherson E, De Forges BR, Boisselier M-C (2006) Seamount endemism questioned by the geographic distribution and population genetic structure of marine invertebrates. *Marine Biology* **149**:1463–1475

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**:5463–5467

Santini L, Saura S, Rondinini C (2016) Connectivity of the global network of protected areas. *Diversity and Distributions* **22**:199–211

Santos MF, Souza IGB, Gomes SO, Silva GR, Bentzen P, Diniz FM (2018) Isolation and characterization of microsatellite markers in the spiny lobster, *Panulirus echinatus* Smith, 1869 (Decapoda: Palinuridae) by Illumina MiSeq sequencing. *Journal of Genetics* **97**:e25–e30

Schnabel KE (2009) Squat lobsters (Crustacea: Anomura: Galatheidae and Chirostylidae) of New Zealand: diversity, phylogeny and biogeography. University of Otago, Dunedin, New Zealand

Schnabel KE, Martin JW, Moffitt RB (2009) Additions to the decapod crustacean fauna of the Hawaiian Islands, III. A new species of the genus *Babamunida* (Crustacea: Galatheidae) from Hawaii based on morphological and molecular evidence. *Zootaxa* **2130**:21–30

Schnabel KE, Ahyong ST (2010) A new classification of the Chirostyloidea (Crustacea: Decapoda: Anomura). *Zootaxa* **2687**:56–64

Schnabel KE, Ahyong ST, Maas E (2011a) Galatheoidea are not monophyletic–molecular and morphological phylogeny of the squat lobsters (Decapoda: Anomura) with recognition of a new superfamily. *Molecular Phylogenetics and Evolution* **58**:157–168

Schnabel KE, Cabezas P, McCallum A, Macpherson E, Ahyong ST, Baba K (2011b) Worldwide distribution patterns of squat lobsters. In: Poore GCB, Ahyong ST, Taylor J (eds.), The Biology of Squat Lobsters. CSIRO Publishing, Melbourne, pp 149–182

Schnabel KE, Burghardt I, Ahyong ST (2017) Southern high latitude squat lobsters II: description of *Uroptychus macquariae* sp. nov. from Macquarie Ridge. *Zootaxa* **4353**:327–338

Schnabel KE (in press) The Marine Fauna of New Zealand. Squat lobsters (Crustacea: Decapoda: Chirostyloidea). *NIWA Biodiversity Memoir* **132**

Scholtz G, Richter S (1995) Phylogenetic systematics of the reptantian Decapoda (Crustacea, Malacostraca). *Zoological Journal of the Linnean Society* **113**:289–328

Schönrogge K et al. (2002) When rare species become endangered: cryptic speciation in myrmecophilous hoverflies. *Biological Journal of the Linnean Society* **75**:291–300

Schubart CD, Neigel JE, Felder DL (2000) Use of mitochondrial 16S rRNA gene for phylogenetic and population studies in Crustacea. *Crustacean Issues* **12**:817–830

Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources* **11**:1–8

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

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

Selkoe KA et al. (2016) A decade of seascape genetics: contributions to basic and applied marine connectivity. *Marine Ecology Progress Series* **554**:1–19

Shank TM (2010) Seamounts: deep-ocean laboratories of faunal connectivity, evolution, and endemism. *Oceanography* **23**:108–122

Shen H, Braband A, Scholtz G (2013) Mitogenomic analysis of decapod crustacean phylogeny corroborates traditional views on their relationships. *Molecular Phylogenetics and Evolution* **66**:776–789

Silliman K (in press) Population structure, genetic connectivity, and adaptation in the Olympia oyster (*Ostrea lurida*) along the west coast of North America. *Evolutionary Applications*

Silva CNS, Gardner JPA (2015) Identifying environmental factors associated with the genetic structure of the New Zealand scallop: linking seascape genetics and ecophysiological tolerance. *ICES Journal of Marine Science* **73**:1925–1934

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87**:651–701

Slate J, Gratten J, Beraldi D, Stapley J, Hale M, Pemberton JM (2009) Gene mapping in the wild with SNPs: guidelines and future directions. *Genetica* **136**:97–107

Slatkin M, Barton NH (1989) A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* **43**:1349–1368

Smith CT, Antonovich A, Templin WD, Elfstrom CM, Narum SR, Seeb LW (2007) Impacts of marker class bias relative to locus-specific variability on population inferences in Chinook salmon: a comparison of single-nucleotide polymorphisms with short tandem repeats and allozymes. *Transactions of the American Fisheries Society* **136**:1674–1687

Smith CT, Seeb LW (2008) Number of alleles as a predictor of the relative assignment accuracy of short tandem repeat (STR) and single-nucleotide-polymorphism (SNP) baselines for chum salmon. *Transactions of the American Fisheries Society* **137**:751–762

Stillman JH, Reeb CA (2001) Molecular phylogeny of eastern Pacific porcelain crabs, genera *Petrolisthes* and *Pachycheles*, based on the mtDNA 16S rDNA sequence: phylogeographic and systematic implications. *Molecular Phylogenetics and Evolution* **19**:236–245

Szymanska K, Tomczak M (1994) Subduction of central water near the subtropical front in the southern Tasman Sea. *Deep Sea Research Part I: Oceanographic Research Papers* **41**:1373–1386

Tan MH et al. (2018) ORDER within the chaos: Insights into phylogenetic relationships within the Anomura (Crustacea: Decapoda) from mitochondrial sequences and gene order rearrangements. *Molecular Phylogenetics and Evolution* **127**:320–331

Tautz D (1989) Hypervariabflity of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**:6463–6471

Tautz D, Ellegren H, Weigel D (2010) Next generation molecular ecology. *Molecular Ecology* **19**:1–3

Taylor ML, Roterman CN (2017) Invertebrate population genetics across Earth's largest habitat: The deep-sea floor. *Molecular Ecology* **26**:4872–4896

Thaler AD et al. (2014) Comparative population structure of two deep-sea Hydrothermal-Vent-Associated decapods (*Chorocaris* sp. 2 and *Munidopsis lauensis*) from southwestern Pacific back-arc basins. *PloS One* **9**:e101345

Thatje S, Marsh L, Roterman CN, Mavrogordato MN, Linse K (2015) Adaptations to hydrothermal vent life in *Kiwa tyleri*, a new species of yeti crab from the East Scotia Ridge, Antarctica. *PloS One* **10**:e0127621

Thorrold SR, Zacherl DC, Levin LA (2007) Population connectivity and larval dispersal: using geochemical signatures in calcified structures. *Oceanography* **20**:80–89

Thresher RE, Guinotte JM, Matear RJ, Hobday AJ (2015) Options for managing impacts of climate change on a deep-sea community. *Nature Climate Change* **5**:635–639

Thrush SF, Dayton PK (2002) Disturbance to marine benthic habitats by trawling and dredging: implications for marine biodiversity. *Annual Review of Ecology and Systematics* **33**:449–473

Thurber AR, Jones WJ, Schnabel KE (2011) Dancing for food in the deep sea: bacterial farming by a new species of yeti crab. *PloS One* **6**:e26243

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21**:3907–3930

Toon A, Finley M, Staples J, Crandall KA (2015) Decapod phylogenetics and molecular evolution. In: Martin JW, Crandall KA, Felder DL (eds.), Decapod Crustacean Phylogenetics. CRC Press, Boca Raton, Florida, pp 15–30

Torkamaneh D, Laroche J, Belzile F (2016) Genome-wide SNP calling from genotyping by sequencing (GBS) data: a comparison of seven pipelines and two sequencing technologies. *PloS One* **11**:e0161333

Tracey DM, Rowden AA, Mackay KA, Compton T (2011) Habitat-forming cold-water corals show affinity for seamounts in the New Zealand region. *Marine Ecology Progress Series* **430**:1–22

Treml EA, Halpin PN, Urban DL, Pratson LF (2008) Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landscape Ecology* **23**:19–36

Treml EA, Kool J (2018) Networks for Quantifying and Analysing Seascape Connectivity. In: Pittman SJ (ed.), Seascape Ecology. John Wiley & Sons, Hoboken, New Jersey, pp 293–318

Tsang LM, Ma KY, Ahyong ST, Chan T-Y, Chu KH (2008) Phylogeny of Decapoda using two nuclear protein-coding genes: Origin and evolution of the Reptantia. *Molecular Phylogenetics and Evolution* **48**:359–368

Tsang LM, Chan T-Y, Ahyong ST, Chu KH (2011) Hermit to king, or hermit to all: multiple transitions to crab-like forms from hermit crab ancestors. *Systematic Biology* **60**:616–629

UNEP-WCMC, IUCN (2019) Marine Protected Planet. UNEP-WCMC and IUCN. www.protectedplanet.net. Accessed 24 April 2019

United Nations General Assembly (2007) Resolution 61/105. Sustainable fisheries, including through the 1995 Agreement for the Implementation of the Provisions of the United Nations Convention on the Law of the Sea of 10 December 1982 relating to the Conservation and Management of Straddling Fish Stocks and Highly Migratory Fish Stocks, and related instruments. UNGA A/RES/61/105. www.un.org/Depts/los/general_assembly/general_assembly_resolutions.htm. Accessed 4 December 2018

van Dover CL et al. (2018) Scientific rationale and international obligations for protection of active hydrothermal vent ecosystems from deep-sea mining. *Marine Policy* **90**:20–28

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Resources* **4**:535–538

Van Wyngaarden M et al. (2017) Identifying patterns of dispersal, connectivity and selection in the sea scallop, *Placopecten magellanicus*, using RAD seq‐derived SNPs. *Evolutionary Applications* **10**:102‐117

Vrijenhoek RC, Schutz SJ, Gustafson RG, Lutz RA (1994) Cryptic species of deep-sea clams (Mollusca: Bivalvia: Vesicomyidae) from hydrothermal vent and cold-water seep environments. *Deep Sea Research Part I: Oceanographic Research Papers* **41**:1171–1189

Wandeler P, Hoeck PEA, Keller LF (2007) Back to the future: museum specimens in population genetics. *Trends in Ecology & Evolution* **22**:634–642

Wang C, Held C (2015) Isolation and characterization of 11 microsatellite markers from the squat lobster *Munida gregaria* (Decapoda: Galatheidae) around the Falkland Islands/Islas Malvinas. *Conservation Genetics Resources* **7**:147–149

Wang C, Agrawal S, Laudien J, Häussermann V, Held C (2016) Discrete phenotypes are not underpinned by genome-wide genetic differentiation in the squat lobster *Munida gregaria* (Crustacea: Decapoda: Munididae): a multi-marker study covering the Patagonian shelf. *BMC Evolutionary Biology* **16**:258

Wang J, Lin R, Bamber RN, Huang D (2013a) Two new species of *Sericosura* Fry & Hedgpeth, 1969 (Arthropoda: Pycnogonida: Ammotheidae) from a hydrothermal vent on the East Pacific Rise. *Zootaxa* **3669**:165–171

Wang J et al. (2013b) Double restriction-enzyme digestion improves the coverage and accuracy of genome-wide CpG methylation profiling by reduced representation bisulfite sequencing. *BMC Genomics* **14**:11

Wang J (2017) The computer program STRUCTURE for assigning individuals to populations: easy to use but easier to misuse. *Molecular Ecology Resources* **17**:981–990

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity* **89**:438–450

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary *N*e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications* **3**:244–262

Waples RS (2014) Testing for Hardy–Weinberg proportions: have we lost the plot? *Journal of Heredity* **106**:1–19

Wares JP, Gaines S, Cunningham CW (2001) A comparative study of asymmetric migration events across a marine biogeographic boundary. *Evolution* **55**:295–306

Watling L, Guinotte J, Clark MR, Smith CR (2013) A proposed biogeography of the deep ocean floor. *Progress in Oceanography* **111**:91–112

Wedding L et al. (2013) From principles to practice: a spatial approach to systematic conservation planning in the deep sea. *Proceedings of the Royal Society B: Biological Sciences* **280**:20131684

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

Weetman D, Hauser L, Bayes MK, Ellis JR, Shaw PW (2006) Genetic population structure across a range of geographic scales in the commercially exploited marine gastropod *Buccinum undatum*. *Marine Ecology Progress Series* **317**:157–169

Wei K, Wood AR, Gardner JPA (2013a) Population genetic variation in the New Zealand greenshell mussel: locus-dependent conflicting signals of weak structure and high gene flow balanced against pronounced structure and high self-recruitment. *Marine Biology* **160**:931–949

Wei K, Wood AR, Gardner JPA (2013b) Seascape genetics of the New Zealand greenshell mussel: sea surface temperature explains macrogeographic scale genetic variation. *Marine Ecology Progress Series* **477**:107–121

Wells SJ, Dale J (2018) Contrasting gene flow at different spatial scales revealed by genotyping-by-sequencing in *Isocladus armatus*, a massively colour polymorphic New Zealand marine isopod. *PeerJ* **6**:e5462

Westfall KM, Gardner JPA (2013) Interlineage *Mytilus galloprovincialis* Lmk. 1819 hybridization yields inconsistent genetic outcomes in the Southern hemisphere. *Biological Invasions* **15**:1493–1506

White C, Selkoe KA, Watson J, Siegel DA, Zacherl DC, Toonen RJ (2010) Ocean currents help explain population genetic structure. *Proceedings of the Royal Society of London B: Biological Sciences* **277**:1685–1694

Williams A et al. (2010) Seamount megabenthic assemblages fail to recover from trawling impacts. *Marine Ecology* **31**:183–199

Willing E-M, Dreyer C, van Oosterhout C (2012) Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PloS One* **7**:e42649

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

WoRMS Editorial Board (2018) World register of marine species. http://www.marinespecies.org. Accessed 4 December 2018

Wright JM, Bentzen P (1995) Microsatellites: genetic markers for the future. In: Carvalho GR, Pitcher TJ (eds.), Molecular Genetics in Fisheries. Springer, Dordrecht, pp 117–121

Xu T et al. (2018) Population genetic structure of the deep‐sea mussel *Bathymodiolus platifrons* (Bivalvia: Mytilidae) in the Northwest Pacific. *Evolutionary Applications* **11**:1915–1930

Yaldwyn JC, Webber WR (2011) Annotated checklist of New Zealand Decapoda (Arthropoda: Crustacea). *Tuhinga* **22**:171–272

Yan R-J, Schnabel KE, Guo X-Z, Gardner JPA (2019) Development and characterization of 20 polymorphic microsatellite loci in the deep sea squat lobster, *Munida isos* Ahyong and Poore, 2004 and cross-amplification in two congeneric species. *Journal of Genetics* **98**:11

Yang C-H, Tsuchida S, Fujikura K, Fujiwara Y, Kawato M, Chan T-Y (2016) Connectivity of the squat lobsters *Shinkaia crosnieri* (Crustacea: Decapoda: Galatheidae) between cold seep and hydrothermal vent habitats. *Bulletin of Marine Science* **92**:17–31

Yang J-S, Yang W-J (2008) The complete mitochondrial genome sequence of the hydrothermal vent galatheid crab *Shinkaia crosnieri* (Crustacea: Decapoda: Anomura): a novel arrangement and incomplete tRNA suite. *BMC Genomics* **9**:257

Yesson C, Clark MR, Taylor ML, Rogers AD (2011) The global distribution of seamounts based on 30 arc seconds bathymetry data. *Deep Sea Research Part I: Oceanographic Research Papers* **58**:442–453

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**:1–16

Zeng C, Rowden AA, Clark MR, Gardner JPA (2017) Population genetic structure and connectivity of deep-sea stony corals (Order Scleractinia) in the New Zealand region: Implications for the conservation and management of vulnerable marine ecosystems. *Evolutionary Applications* **10**:1040–1054

Zeng C, Clark MR, Rowden AA, Kelly M, Gardner JPA (2019) Patterns of genetic connectivity amongst four deep-sea demosponges in the New Zealand region: implications for the protection of Vulnerable Marine Ecosystems. *Scientific Reports* **9**:5482

Zhan B, Fadista J, Thomsen B, Hedegaard J, Panitz F, Bendixen C (2011) Global assessment of genomic variation in cattle by genome resequencing and high-throughput genotyping. *BMC Genomics* **12**:557

Zhang D, Zhou Y, Cheng H, Wang C (2017) The complete mitochondrial genome of a yeti crab *Kiwa tyleri* Thatje, 2015 (Crustacea: Decapod: Anomura: Kiwaidae) from deep-sea hydrothermal vent. *Mitochondrial DNA Part B* **2**:141–142