Does marine geospatial/environmental variation affect genetic variation? A meta-analysis

Daniel Alejandro Cárcamo Segovia

A thesis submitted to Victoria University of Wellington

in fulfilment of the requirements for the degree of

Master of Science in Marine Biology

2021

Dedicado con mucho cariño a mis abuelas,

Olga Guzmán, Albertina Vergara y Sally Santa Cruz.

A mi familia,

David Cárcamo, Diana Segovia y Pamela Cárcamo.

Alcanzar y concluir exitosamente el gran desafío de mi vida,

no fue más que el fruto de todo

el amor y apoyo que recibí desde pequeño.

Siempre les estaré agradecido.

Los amo.

Abstract

Genetic information is important to inform management and conservation. However, few studies have tested the relationship between genetic variation and geospatial/environmental variation across marine species. Here, I test two genetics-based ideas in evolutionary theory using data from 55 New Zealand coastal marine taxa. The Core-Periphery Hypothesis (CPH) states that populations at the centre of a species' distribution exhibit greater genetic variability than populations at the periphery (the 'normal' model). Variants of this model include the 'ramped north' (greatest variation in the north), the 'ramped south' (greatest variation in the south), and the 'abundant edge' (greatest variation at the distributional edges, least variation at the centre). The Seascape Genetics Test (SGT) null hypothesis predicts no association between genetic variation and environmental variation. I conducted a meta-analysis of published/unpublished material on population genetic connectivity and diversity and marine environmental data to test both hypotheses. To assess the CPH, genetic data were fitted to four models (Normal, Ramped North, Ramped South, Abundant Edge). I also conducted a descriptive analysis between the genetic outcomes of the CPH and abundance records for a subset of species. The SGT involved GLM analyses using eleven geospatial/environmental variables and species-specific F_{ST}-Φ_{ST} (genetic distance) estimates plus a smaller subset of genetic diversity data. The CPH results showed that 55 of 249 tests (evaluating on average 2.9 ± 1.3 genetic indices in each of the 84 studies) fitted at least one of the four models: Ramped North (10%), Ramped South (8%), Normal (2%) and Abundant Edge (2.4%). Species-specific abundance records followed the same patterns detected by the CPH. These results indicate that edge populations (Ramped North, Ramped South, Abundant Edge) exhibit greater genetic variability than central populations amongst marine taxa from New Zealand, but that most taxa do not conform to any model (~78% of all tests were not statistically significant). For the seascape genetics multi-species analysis (comprising 498 individual tests), the F_{ST}-Φ_{ST} estimates (genetic distance estimates between pairs of populations) were mostly affected by four factors related to sea surface temperature. For genetic diversity indices the most significant predictors were latitude and longitude. Whilst different factors (e.g., physical oceanography, food availability, life-history traits and harvesting), either acting alone or acting synergistically, are likely to be important in explaining patterns of genetic diversity in New Zealand's marine coastal species, my results indicate that variables including SST and to a lesser extent the geospatial variables (latitude and longitude) explain much of the variation in the genetic indices tested here.

Keywords: The Core-Periphery Hypothesis, Seascape genetics, genetic variability, geospatial variation, environmental variation, allozymes, mitochondrial DNA markers, nuclear DNA markers, genetic patterns, life-history traits, southwestern Pacific, New Zealand.

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Jonathan Gardner, for having developed the idea of this research project and for trusting me to conduct the investigative process behind this meta-analysis, here in New Zealand. I am very pleased that he was willing to wait for me (for more than four years), in order to be able to save enough money to pay the university fees and associated costs while I was in Chile, which took a lot of time and personal efforts.

I would like to thank my dad, David Cárcamo, my mom, Diana Segovia and my sister Pamela Cárcamo. My family were critical in the process of making this dream come true, and this dream was partly theirs as well. Special thanks for the emotional and financial support they gave me during these challenging times, without which I would never have been able to succeed. Thank you, also, for providing support and being a shoulder to lean on.

Next, a special greeting to my Chilean undergraduate supervisors. Doris Oliva and L. René Durán have been such fantastic friends and mentors, and a wealth of knowledge and advice. I am extremely thankful for all the opportunities they gave me while I was studying my BSc in Biology at Universidad de Valparaíso, and for the professional experience that I gained, working hand in hand with them, in my first years after graduation. I will be eternally grateful.

The School of Biological Sciences at Victoria University of Wellington has many support and administrative staff without whom life would have been a bit more difficult. A special thank you to Mary Murray, Mark Stephen and Patricia Stein. It goes without saying that from the beginning it was a pleasure to feel part of a School that is so welcoming, with a genuine New Zealand style, proud of being relaxed and happy, but nonetheless highly productive.

Special thanks to my closest friends here in New Zealand, Giulia Trauzzi, Sonali Pawaskar, Lorenzo Ravalo, Jana Dobelmann, Romain Chaput, Laura Somerset and Georgia Kahan.

Lastly and most importantly, I express special gratitude to my Cuban friend Mailiu Díaz Peña. I will be eternally grateful for your unselfish help. Without your statistical advice I cannot even imagine how I would have been able to overcome such a huge barrier with statistics and RStudio.

I truly believe that life is unpredictable. In a year full of challenges, fears, tears, insecurities and uncertainties under COVID-19, I was able to arrive in New Zealand with the perfect timing, just one month before quarantine started. Despite many frustrations that I experienced waiting to reach my dream to come and study in New Zealand, which are not worth mentioning right now, life showed me that this was the perfect moment. Unconsciously (I believe) I decided to take advantage of what fate brought me. As the opening sentence stated by Robin S. Waples in his Food for Thought article goes, "*How much of your scientific career has unfolded as you planned, and how much has been shaped by blind luck?"* (Waples, 2020). This quote makes a lot more sense to me right now than the first time I read it. I feel sure that being in the right place during 2020-2021 was the perfect mixture between one of my deepest wishes and, of course, serendipity.

Table of contents

List of Figures

Chapter 2

Chapter 3

List of Tables

Chapter 2

Chapter 3

List of Formulae

Chapter 2

Chapter 3

Supplementary Material List

Review of phylogeographic genetic studies of marine vertebrates and invertebrates, macroalgae and plants conducted in New Zealand

Chapter 3

Research aims and thesis structure

The overall objective of this Master's thesis was to test two fundamental ideas in evolutionary theory about the relationship between genetic variation and geospatial/environmental variation. To fulfil this objective, data were compiled from studies/projects (the meta-analysis component) and then used to test both hypotheses.

The specific aims of the thesis were:

1. To carry out an exhaustive bibliographic review of population genetics studies, performed on a wide range of organisms that inhabit marine habitats in New Zealand, including endemic species, regionally native species (e.g., those found in New Zealand and Australia), and also Southern hemisphere species (e.g., those found in New Zealand, Chile, Australia, South Africa, etc), but excluding introduced and non-native species.

2. To examine the relationship between population genetic variation and population geographic location, to assess if there is a common or prevalent distribution pattern based on a multi-species genetic dataset. That is, to test the Core-Periphery Hypothesis, which predicts that species-specific genetic variation will be greatest at the centre of the distribution of each species, and lowest at the edges of the distribution.

3. To evaluate the relationship between population genetic variation at a site and the geospatial/environmental variation at that site based on a multi-species genetic dataset. That is, to use seascape genetics to better understand patterns of multi-taxon genetic variation.

The hypotheses to be tested were:

H1: Latitudinal variation across the distribution range of marine species is an important determinant which affects genetic variation.

H2: Variation in geospatial and environmental variables/factors explains genetic variation.

The thesis is divided into four chapters. Chapter 1 is a general introduction to the scientific topic behind this macroecological and evolutionary research, with emphasis on the theoretical framework of abundance and distribution and population genetics: The Abundant-Centre Hypothesis, the Core-Periphery Hypothesis, phylogeography, phylogenetics and seascape genetics. Chapter 2 is a data chapter that describes the results of the testing of the

Core-Periphery Hypothesis across 84 studies (comprising in total 55 species), all of them conducted in New Zealand waters. Chapter 3 is a data chapter and takes a seascape genetics approach to identify associations between geospatial/environmental variables and genetic structure at a phylum level within the 84 studies (comprising in total 55 species). Chapter 4 examines the key findings in a broader context and indicates the limitations and opportunities for further research on these topics. The Supplementary Material section contains a summary of the multi-taxon data set on which this research was based.

Chapters 2 and 3 have been written in a 'ready to publish' scientific article format and as a consequence a certain amount of repetition was unavoidable. However, where entire sections were identical (such as parts of the Materials and methods and Results sections) they have been cross-referenced. Chapter 2 is intended to be submitted to ***Global Ecology and Biogeography*** or ***Ecography*** and Chapter 3 to ***Molecular Ecology.***

It should be clarified that despite the fact that both data chapters are written in a 'ready to publish' paper format, alluding to collaborative work between me and my co-authors, the work presented here was written and developed entirely by me.

CHAPTER 1 General introduction

Life under the ocean implies the need to constantly overcome challenges. Most marine organisms live in a dense and viscous moving fluid, which houses and transports a number of other components (i.e., dissolved nutrients, food, organic material such as detritus and inorganic material such as sediment, among others). Furthermore, individual life forms exhibit species-specific life-history traits determining, among many other things, the presence of many different life stages in the ocean (Riginos & Liggins, 2013). These features generate a complex matrix for the development of diverse structures and biological dynamics in the world's seas. Perhaps because of this complexity and the scale of the oceans the identification of marine patterns of individual- or population-based genetic diversity is still imprecise, even when patterns of biogeography are much better understood.

Main drivers for the establishment of genetic patterns within biogeographic distributions

Wallis & Trewick (2009) summarised molecular phylogenetic analyses conducted in New Zealand within their review entitled: "New Zealand phylogeography: evolution on a small continent". In this article the authors highlight that 10% of the included studies involving approx 100 groups (plants and animals) revealed an archaic origin dating to the vicariant spitting of Zealandia from Gondwana. They also state, based on a compilation covering phylogenetic structuring of terrestrial, freshwater, and marine taxa, that the most common patterns are characterised by east-west splits across both main islands, north-south splits and star phylogenies in mountain systems. Interestingly, the diversity of patterns across taxa suggest that geological processes (i.e., uplifting and sinking activity in the earth's crust, sea level change, erosion, volcanism and glaciation - Suggate *et al.*, 1978; Campbell & Hutching, 2007; Campbell, 2013) may have produced these forms of species structuring within an insular system (Wallis & Trewick, 2009).

Biogeographers largely have attempted to explain the association and/or patterns between the environment and the ecology of organisms (e.g., abundance, biodiversity, richness, etc). Subsequently, three biogeographic models have been proposed for the origins of marine biodiversity worldwide, which are the most accepted and cited across the field (Briggs & Bowen, 2013).

First, the centre of origin, which states that species tend to evolve in or at the centres of their distributions, rather than elsewhere. Interestingly, it is also hypothesised that newly formed species do not have equal dispersal potentials. Those that are produced in the diversity centres by large numbers of individuals with high levels of genetic variation are the ones that disperse outward to keep the system going (Briggs, 2003, 2005).

Second, the centre of accumulation or the 'vortex model' which has been proposed to occur in nature through the transportation of new species from the peripheral areas into the centre of high diversity (Jokiel & Martinelli, 1992).

Third, the centre of overlap between the Pacific and Indian Ocean faunas. This theory was established based on the description of an Indian Ocean centred-distribution (comprising 26 marine tropical Indo-West Pacific fishes), which seems to be the result of an overlap between two faunas (one distributed from Indonesia to the Gulf of Oman and the other from Indonesia to French Polynesia) (Woodland, 1983).

New Zealand's diversity has been influenced mainly by the North Pacific Ocean and Southern Ocean (Briggs & Bowen, 2013). In this way, consecutive migrations have contributed a broad variety of species to the Southern hemisphere (more interhemispheric dispersals have taken place from north to south than vice versa), most of these processes being initiated by upwelling systems (Lindberg, 1991). Furthermore, Griffiths *et al.* (2009) demonstrated that all the species' distributions evaluated (~7,500) within the Southern hemisphere (comprising 29 studied regions) clustered into three main groups: New Zealand-Australia, Antarctica, and southern South America. It is worth highlighting the major outcome in relation to high levels of biological endemism made by Griffiths *et al.* (2009) who concluded that New Zealand rates as one of the provinces with the richest fauna and highest endemism, along with other places such as Tasmania, the Falkland Islands, southern Argentina, southern Chile and Tierra del Fuego.

Researchers have tried to explain how parameters such as a species' abundance may change as a function of position within the species' distribution. For example, the Abundant-Centre Hypothesis (Brown, 1984; Brown *et al.*, 1995) states that populations at the centre of a species' distribution will exhibit greater abundance than populations at the distribution edges. This is due to physiological constrains that limit population size or the abundance of individuals in sub-optimal habitat at the edge of the species' range. This classical theory in evolutionary biology, proposed by Brown (1984), attempts to evaluate the origin of

distribution shapes to better understand the structure of populations at the species level, but on a smaller scale than the biogeographic theories mentioned above. Although the Abundant-Centre Hypothesis has been well-explored in terrestrial systems (Santini *et al.*, 2019), it remains unclear whether such a pattern applies in marine systems.

Just as the Abundant-Centre Hypothesis has been proposed to explain patterns of species-specific abundance as a function of a species' distributional extent, the statement behind the hypothesis may be transformed to reflect, for example, genetic variation as a function of distribution (Carson, 1959). Thus, the Core-Periphery Hypothesis (hereafter, CPH) it has been postulated for the genetic expectations behind this idea in which genetic variation is expected to be greater at the centre of the species' distribution and for it to decline to the distributional edges.

To date, only a few literature reviews covering the genetic issue behind the theory of the CPH have been carried out, but all reviews have been focussed on terrestrial organisms. It is worth noting the final conclusions of Eckert *et al.* (2008), who reported from 134 studies representing 115 species that 64% of them followed the expected decline towards the edges in diversity indices (in this case, expected heterozygosity estimates). There is only one study available in which genetic data has been exclusively applied to the assessment of the CPH in marine organisms. This study, conducted on mussels in South Africa (the native *Perna perna* and the introduced *Mytilus galloprovincialis*), did not find the expected pattern of a centre of high genetic diversity when using estimates of allelic richness for the analysis (Ntuli *et al.*, 2020). This led the authors to reject the genetic diversity variant of the CPH.

Today, we know that many New Zealand coastal marine species have been examined for their population genetic structure (see details below about New Zealand's well-developed set of genetic population studies and its spatially explicit environmental data set), providing a great source of information to test hypotheses. Testing of the CPH could help improve our understanding of macroecological patterns of structure in relation to major biogeographic patterns and boundaries across marine systems.

Seascape genetics: background and general patterns

The multidisciplinary approach behind seascape genetics has grown rapidly over the last two decades (Selkoe *et al.*, 2008). The most remarkable event in this incipient field of genetics is the rate of development of different methodologies and techniques to evaluate the

influence of environmental variables on genetic variation (Selkoe *et al.*, 2008, 2016). In this discipline geneticists try to assess how environmental variation (in a broad sense) influences genetic variation (Riginos & Liggins, 2013). This is one reason why scientists worldwide should pay special attention to the understanding of environmental data and its collection, because this information can give rise to significant predictors for the observed genetic variability. Among other uses, this can provide valuable information for decision making related to marine organisms, i.e., setting of management areas for conservation (Zeng *et al.*, 2020).

Trying to understand how ocean currents affect distribution and connectivity of marine species provides valuable information for the effective conservation management of marine organisms (Foll & Gaggiotti, 2006; Jensen *et al.*, 2020). Despite the presence of spatially replicated gradients that occur in the marine environment (Schmidt *et al.*, 2008), research on seascape genetics is relatively uncommon (Riginos & Liggins, 2013). Conversely, landscape genetics approaches are more frequent and are, as a consequence of this, undergoing greater development and reaching comparatively more robust conclusions (Vucetich & Waite, 2003).

Recently it has been suggested that multiple environmental factors (i.e., geography and oceanography) (Selkoe *et al.*, 2016) and life-history traits (Ross *et al.*, 2009) can interact synergistically to shape genetic patterns (Riginos & Liggins, 2013). In this sense, it is worth noting some examples that have helped further our understanding of changes in allele frequencies and genetic diversity at small spatial scales linked to biogeographic boundaries (e.g., the biogeographic break at Point Conception in south-central California, Cape Agulhas in South Africa and the faunal boundary of Wallace's Line). Key species-specific case studies which have addressed this issue include studies of *Triakis semifasciata* (elasmobranch endemic to the Eastern North Pacific Ocean - Barker *et al.*, 2015), *Clinus cottoides* (fish endemic to the coast of South Africa - Von der Heyden *et al.*, 2008) and *Hippocampus trimaculatus* (a seahorse species from Southeast Asia - Lourie & Vincent, 2004). These case studies all found a pattern of genetic divergence associated with natural barriers.

A recent study, which serves as a good example to show the global predictive power of seascape genetics, was conducted on all six extant rock lobster species (*Jasus* spp.) with a worldwide presence at approximately 40°S. These taxa have populated the continental shelf and seamount habitats in a narrow latitudinal band in the Southern hemisphere across the

Pacific, Atlantic and Indian Oceans (Booth, 2006). In this study, Silva *et al.* (2021) reported that benthic temperature is associated with the differentiation between species/populations from island/seamount and continental shelf habitats. This finding suggests that temperature may limit reproduction between species-specific populations (which are adapted to local benthic temperatures) since this factor has been associated with embryological development in lobsters (Benestan *et al.*, 2016).

New Zealand is both unusual and fortunate in having a well-developed set of marine environmental variables for the whole country, including the Exclusive Economic Zone - Marine Environment Classification (EEZ-MEC) (Snelder *et al.*, 2005) and one specifically developed for estuaries, the Estuarine Environment Classification (EEC) (Hume *et al.*, 2007). Both schemes were produced for the Ministry for the Environment by the National Institute of Water and Atmospheric Research (NIWA). This unprecedented national coverage of marine environmental information makes it possible to plan and execute seascape genetics studies that require site-specific environmental data for their development.

The above makes even more sense if we consider the wide range coverage of population genetic studies conducted across multiple species, i.e., coastal marine taxa, within a country such as New Zealand, which shows habitats shaped by seismic activity and sea level changes (Suggate *et al.*, 1978; Campbell & Hutching, 2007; Campbell, 2013), high levels of endemism (Gordon *et al.*, 2010), isolation (Rouse *et al.*, 2003) and that spans substantial latitudinal and moderate longitudinal variation.

In a review developed by Gardner *et al.* (2010) for the New Zealand Ministry of Fisheries, the authors reported 58 studies that have described population genetic structure and/or genetic indices variation in 42 different taxa (vertebrates, invertebrates, macroalgae and plants) of New Zealand's coastal marine biota. These different studies used a range of marker types including allozymes, dominant and anonymous markers [random amplified polymorphic DNAs (RAPDs), restriction fragment length polymorphisms (RFLPs) and amplified fragment length polymorphisms (AFLPs)], microsatellites and nuclear/mitochondrial DNA sequences. Interestingly, the authors described that among these studies, 36 focussed on rocky intertidal and subtidal taxa, 2 on open coastal, 9 on estuarine, 9 on fjordic, and 2 on soft substrate subtidal taxa. The authors highlight that among all the reviewed taxa, macroinvertebrates were the best represented group (28 of 42 taxa = 67%). In terms of genetic population structure, 16 studies reported no spatial structure, 9 reported an

isolation by distance model, 12 reported divergence within and/or among populations, 20 reported a north-south differentiation and 1 reported an east-west difference pattern. Whilst this report was important for bringing together a large collection of New Zealand marine studies, and also for identifying patterns of genetic diversity across taxa/phyla, it did not consider environmental variation as a driver of genetic diversity.

Some approaches in seascape genetics have been successfully completed in endemic and native species of New Zealand. With a few exceptions, sea surface temperature has been identified as the key factor in explaining genetic variation (using measures of genetic distance among populations). First, genetic variation in the New Zealand blue mussel (*Mytilus galloprovincialis*) was found to be strongly associated with sea surface temperature and winter solar radiation (Westfall, 2010). The case study of the endemic New Zealand greenshell mussel (*Perna canaliculus*) suggested a macrogeographic scale of genetic variation also associated with sea surface temperature variation (Wei *et al.*, 2013b). Genetic variation in two flatfish species (*Rhombosolea leporina* and *Rhombosolea plebeia*) was strongly associated with variation in latitude, sediment type, mean orbital velocity and sea surface temperature (Constable, 2014). At the same time, and including samples from some of the same sites as those sampled by Constable (2014), Hannan (2014) concluded that variation in sea surface temperature, tidal current and bed slope explained significant genetic variation in the wave-exposed sandy beach clam, *Paphies subtriangulata*, whereas variation in latitude, longitude geospatial distance, sea surface temperature and bed slope helped to explain significant genetic variation in the estuarine clam, *Paphies australis*. Subsequently, Silva & Gardner (2016) reported for the New Zealand scallop (*Pecten novaezelandiae*) that variation in population genetic structure was associated with variation in localised freshwater input and suspended particulate matter concentration. Finally, Zeng *et al.* (2020), working with four deep-sea species in the New Zealand region reported a relationship between variation in dissolved oxygen and genetic variation for a sponge (*Poecillastra laminaris*), and between dynamic topography (*Goniocorella dumosa*), sea surface temperature (*Madrepora oculata*) and tidal current (*Solenosmilia variabilis*) for genetic variation in three cold-water corals. These studies provide valuable new insights into the key environmental drivers of genetic differentiation among populations, setting a precedent for future research worldwide.

Intended research: looking for patterns across multiple taxa

The background information presented in this introduction highlights an absence of marine-related studies covering first, the CPH and second, seascape genetics approaches. I will try to fill some of this information gap through the design of the first spatially explicit tests of species-specific genetic variation covering the CPH and seascapes genetics across a marine multi-taxon dataset from New Zealand's marine biota.

CHAPTER 2 Testing the Core-Periphery Hypothesis: a multi-phylum assessment of marine species from New Zealand

Abstract

The Abundant-Centre Hypothesis states that populations at the centre of a species' distribution will exhibit greater abundance than populations at the distributional edges, due to physiological constraints faced by populations at the distributional edges (the 'normal' model). Although the Abundant-Centre Hypothesis has been well-examined for species from terrestrial habitats, it is unclear if this observed pattern on land applies to marine systems (probably due to the paucity of actual/observed abundance data for most marine taxa). It is interesting to dive into the genetic expectations behind this assumption. The Core-Periphery Hypothesis (CPH) assumes that abundance is no longer the response variable, but genetic diversity (in various different forms) is the variable of interest. Variants of the CPH model include the 'ramped north' (greatest variation in the north), the 'ramped south' (greatest variation in the south), and the 'abundant edge' (greatest variation at the distributional edges). To date, surprisingly little information is available about this modified (genetics-focussed) hypothesis, for either terrestrial or marine taxa. What information exists comes from relatively few genetic studies that have shown big differences for the expected genetic variability distribution. These previous studies have attempted to fit into a normal distribution both connectivity metrics (e.g., F_{ST} and Φ_{ST} values) and diversity genetic values (e.g., haplotype diversity, allelic richness), which according to the CPH theory should display greater diversity values at the centre of the distribution compared to the edges. Published reports from elsewhere, employing this approach, have reported that mussels in South Africa did not exhibit higher values of genetic diversity at the centre of the distribution, whereas an endemic flowering plant in the south-eastern United States fitted the expected model. Here, I present the first broad-scale multi-phylum analysis of the CPH, using a large dataset of genetic indices of New Zealand marine biota covering 55 species. I conducted a meta-analysis by gathering published and unpublished material on marine population genetic connectivity and diversity to test the CPH. Based on a selection of 84 studies, across 9 phyla, values of genetic connectivity and diversity indices were fitted to four models of distribution (Normal, Ramped North, Ramped South, Abundant Edge). The degree of fit of each model to the observed and randomised data was evaluated by calculating the sum of squared deviations. In summary, the results of this study revealed ~25% of the tests were statistically significant ($p<0.05$) for the four models: Ramped North (10%), Ramped South (8%), Normal (2%) and Abundant Edge (2.4%). Among the Chordata (Ramped North), Arthropoda (Ramped South) and Mollusca (Ramped North) it was possible to determine a reasonably consistent pattern of genetic variation across all or most species within each phylum. The

remaining taxa fitted multiple models but without any obvious pattern across the phyla. Nonetheless, across all taxa, most tests (~75%) were not statistically significant. In general terms, previous authors have argued for the rejection of the Abundant-Centre Hypothesis and the CPH. The results presented here revealed that only 5 of 29 species (the New Zealand sea lion, the New Zealand sea urchin, the New Zealand dredge oyster, the spotted whelk and the Southern bull kelp) actually fitted the Normal model. In other words, the large majority of species did not fit the Normal model. These results also indicate that edge populations (Ramped North, Ramped South, Abundant Edge) exhibit greater genetic variability than populations at the species' distributional centre amongst marine taxa from New Zealand, but that most taxa (approx 26 of 55 species) do not conform to any model i.e., that genetic variation is independent of geographic location within a species' distribution.

Keywords: The Abundant-Centre Hypothesis, the Core-Peryphery Hypothesis, genetic variability, geospatial variation, allozymes, mitochondrial DNA markers, nuclear DNA markers, pelagic larval duration, life-history traits, New Zealand oceanography, edge populations genetic variability, New Zealand.

Introduction

New Zealand has a rich diversity of marine habitats that include over 15,000 known species and also manages the fourth largest Exclusive Economic Zone in the world (Department of Conservation and Ministry of Fisheries, 2005). This ecosystem status gives biologists an opportunity to compare organismal life-history traits under different geographic, oceanographic, geologic and environmental conditions (Parenti, 2019) with the aim of better understanding which environmental factors contribute to biological variation. In this sense, organisms can interact synergistically with environmental variables/factors (e.g., via migration, occupation of a niche, sympatric speciation), generating new phenotypes driven by positive selection for change (adaptation) (Mustonen & Lässig, 2009). New Zealand's long latitudinal cline from the subtropical north at ~29°S to the subantarctic south at ~53°S is reflected in the large number of endemic species and a wide range of marine habitats (Gordon *et al.*, 2010), making this country an excellent case study location for the study of interactions between the environment and the genome.

Genetic information is crucially important for informing the management and conservation of species, but genetic surveys are time-consuming and costly to conduct meaning that relatively few detailed studies have been carried out. For these reasons, the ability to apply 'general rules' to patterns of observed genetic variation, and the ability to use environmental surrogates for genetic variation is very appealing. The Abundant-Centre Hypothesis states that populations at the centre of a species' distribution will exhibit greater abundance than populations at the distributional edges, due to physiological constrains that limit populations (Brown, 1984, 1995; Brown *et al.*, 1996). Put simply, it is expected that environmental conditions at the centre of the distributional range are more 'optimal' than those at the distributional edges, therefore numbers of individuals are expected to be greater in the centre than at the edges. Although this hypothesis has been well-explored in terrestrial systems (Brown *et al.*, 1996), it remains unclear whether such a pattern applies in marine systems (see below). Just as the Abundant-Centre Hypothesis is proposed to explain patterns of species-specific abundance as a function of distributional extent, the hypothesis may be modified to reflect genetic variation as a function of site-specific location with the species' distribution which is better know as the Core-Periphery Hypothesis (hereafter CPH) (Carson, 1959). The analysis of population genetic data (e.g., connectivity and diversity indices) might however be associated with population size (i.e., abundance) because it is expected that populations at the periphery of the species' range will show lower intra-population diversity

and higher inter-population divergence when compared to central populations (Brown *et al.*, 1995; Gilman, 2006; Levy *et al.*, 2016). Thus, measures of site-specific abundance might not be completely independent of measures of genetic diversity because larger populations are expected to, on average, exhibit greater levels of genetic diversity. Nonetheless, testing of the fit of the hypotheses to both abundance data and to genetic indices data to the same species may permit, for the first time, an improved understanding of how abundance and genetic variation co-vary or vary independently as a function of a species' distributional range.

Previous reviews have, in general, demonstrated a non-consistent relationship between abundance and the centroid of the species distribution; that is, limited or no support for the Abundant-Centre Hypothesis. For example, Sagarin & Gaines (2002b) found that only 39% of individual tests (N_{tests} = 145, comprising 121 species, including plants, insects, marine invertebrates, fishes, birds and mammals mainly distributed in North America and Europe) based on abundance data, supported the hypothesis. Of particular interest is their finding that most of the studies did not sample correctly the species' range of distribution. Not surprisingly, this finding suggests that the design and execution of a neat and detailed sampling regime (representative spatial coverage) is critical in establishing and testing biogeographical patterns. Subsequently, Martinez-Meyer *et al.* (2012) did not find the expected distribution pattern for the abundances of 11 marine and terrestrial species (including birds, mammals and reptiles) in western North America. These authors used an alternative approach to test the Abundant-Centre Hypothesis by placing the distribution centre in the middle of the ecological niche for each species that had been studied (i.e., its habitable conditions but not necessarily its realised niche). More recently, Santini *et al.* (2019) do not report support for the Abundant-Centre Hypothesis for data from 9 bird and 99 mammal species, although they tested 9 models of geographic and ecological centrality/marginality measures. Their analysis showed that variables such as dispersal distance, geographic coverage and environmental coverage of the data appear to be important in explaining the observed variation between different species native to Africa, Europe and Asia. Most recently, although the relative abundance of a small mammal in Central and Southwest China was shown to exhibit a unimodal symmetric pattern of distribution around the centre of the species' distribution, this model appeared only once among the twelve studied elevational gradients for all species (Wen *et al.*, 2020). Interestingly, Wen *et al.* (2020) reported differences in results when testing a data set for small mammals at local and regional scales. The possible explanations for this difference are likely to be topographic

variation, climate and interspecific interaction within the study areas (Wen *et al.*, 2020). Thus, only a very small proportion of the total tested species within the reviews actually followed the expected abundant-centre pattern. These findings are mostly based on studies conducted in the Northern hemisphere, and most often (but not exclusively) for non-marine taxa. Thus, at broad-scale levels of taxonomic (phylum) diversity and environment, further analyses using other geographical areas and habitat types are still necessary before it is possible to draw a reasonably firm conclusion about the Abundant-Centre Hypothesis.

Research across a limited range of marine organisms has reported various outcomes for the Abundant-Centre Hypothesis when tested against abundance data. Tuya *et al.* (2008) found that only two of five endemic demersal fish species (*Chromis klunzingeri* and *Ephinephelides armatus*) from south-western Australia had maximum mean abundances at the centres of their full distributional ranges. Rivadeneira *et al.* (2010) studied five common intertidal porcelain crabs at 13 sites along the Chilean coast (a study performed across 25 degrees of latitude). The relative abundance varied among the different species, showing for only two species (*Allopetrolisthes angulosus* and *Petrolisthes tuberculatus*) a clear trend of greatest densities near the centre of their respective geographic ranges. Hidas *et al.* (2010) tested the Abundant-Centre Hypothesis for three rocky-shore intertidal invertebrates (two gastropods *Morula marginalba* and *Nodilittorina pyramidalis*, and one barnacle, *Tesseropora rosea*) in south-eastern Australia, concluding that at a regional scale (not full distributional range) species' abundances were lower at their range limits. In Atlantic Canada, using as predictor species two mussels and one dogwhelk (*Mytilus edulis*, *Mytilus trossulus* and *Nucella lapillus*, respectively), limited support was reported for the general expectations behind the Abundant-Centre Hypothesis, but based on regional subset of the full distributional ranges of the three species (Tam & Scrosati, 2011). It is interesting to note that the two conspecific mussels, *M. trossulus* and *M. edulis*, show a Ramped South and Abundant Edge distribution pattern, respectively (Tam & Scrosati, 2011). That is, even for two closely related species within the same genus, the observed patterns of abundance are not always the same, at least at the regional scale. This suggests that the aforementioned distributions may be attributed to abrupt changes in habitat properties or environmental conditions (Brown, 1984; Tam & Scrosati, 2011). Among other molluscan species the abundance of the rocky intertidal owl limpet (*Lottia gigantea*) on the Pacific coast of North America was found to provide support for the Abundant-Centre Hypothesis (Fenberg & Rivadeneira, 2011). Subsequently, Baldanzi *et al.* (2013) evaluated the distribution of two

sandhopper species along the Namibian and South African coasts, and provided weak support for the abundance predictions, with *Africorchestia quadrispinosa* being the species that best fits the expected Abundant-Centre Hypothesis model. Soon after, Scrosati & Freeman (2019) evaluated the density of the intertidal barnacle, *Semibalanus balanoides*, across its full elevational range in Nova Scotia, Atlantic Canada (from low to high elevations – not its full distributional range) and reported that this species followed the Abundant-Centre Hypothesis, as barnacle density peaked at the middle of the vertical distribution. Most recently, Shalom *et al.* (2020), using a large data set of underwater visual estimates of fish abundance (1,215 species), with an almost global coverage, found support for a general gradient in abundance across species' ranges, identifying a centre of high abundance and occupancy and lower abundance towards the range limits. In conclusion, the general trend in the results of the studies described above for marine organisms clearly indicates that the Abundant-Centre Hypothesis does not predict the relative abundances of most species across their distributions. Nevertheless, the Abundant-Centre Hypothesis may help to predict the distribution patterns of some species in an accurate way, even if only at the regional scale (e.g., Tam & Scrosati, 2011; Scrosati & Freeman, 2019).

To date, only two literature reviews covering the CPH have been carried out, both with a focus on terrestrial taxa. These reviews have reported in general a medium-high support for the CPH using population-based genetic data. First, Eckert *et al.* (2008) evaluated geographical variation in population studies that assayed spatially explicit variation of nuclear DNA (e.g., allozymes, RAPDs, ISSRs, AFLPs, microsatellites and DNA sequences). The authors found that from 134 studies representing 115 species, 64% of them followed the expected decline towards the edges in diversity indices. Subsequently, Sexton *et al.* (2009) summarised the main outcomes for the association of range limits with variation in population size. Interestingly, they found support for the Abundant-Centre Hypothesis prediction for 5 of 9 studies (56%). In their review they highlight a statement by Goldberg *et al.* (2007), which reads: "Dispersal barriers limit gene flow and allow for local adaptation and spread. If gene flow is restricting a species' range, species' borders should not be associated with dispersal barriers" (Sexton *et al.*, 2009, p419, Table 1). This statement underlines the idea that theoretical models predict many ways for range limits to arise when referring to genetic variation. But it is worth noting that most models, and their underlying assumptions, have not yet been tested.

There are only two studies where genetic data have been applied exclusively to the assessment of the CPH, and only one study for marine taxa. Dixon *et al.* (2013) reported support for the CPH but also rejection of the Abundant-Centre Hypothesis for an endemic flowering plant (*Leavenworthia stylosa*) in the south-eastern United States. In this sense, some authors have suggested that effective population size should decline towards the population distributional limits, with net movement of migrants from the centre to the edges (Vucetich & Waite, 2003). The case study of *L. stylosa* showed that locations farther from the distributional centre exhibited a significant reduction in average allelic diversity and the opposite pattern for F_{ST} estimates (F_{ST} is a measure of differentiation or genetic distance). Thus, it was concluded that the demographic variability did not follow the expected pattern while genetic variability and genetic differentiation showed results consistent with its predictions (Dixon *et al.*, 2013). In contrast, a recent genetic study in South Africa did not find for either studied species of mussels (*Perna perna* and *Mytilus galloprovincialis*) the expected pattern of population-specific allelic richness as a function of distributional range, arguing for rejection of the CPH (Ntuli *et al.*, 2020). These two studies represent very limited testing of the CPH and highlight the need for further research and discussion on the prediction of the CPH when describing genetic variability for marine systems.

Overall, as noted above, there is surprisingly little assessment of the CPH based on genetic variation, although there has been much testing of it based on abundance estimates. Given the multitude of recent marine genetics investigations carried out in New Zealand, mostly on endemic species, the conditions are well-placed to test this hypothesis for genetics data. Building on the data set of Gardner *et al.* (2010) that contained information from 58 studies of population genetic structure across 42 different taxa (including plants, macroalgae, invertebrates and vertebrates) of New Zealand's coastal marine biota, I have added studies published since 2010 to enhance the database. I conducted the first comprehensive test of the CPH in marine systems using a large multispecies New Zealand data set covering endemic and native species, comprising 10 phyla (Chordata, Echinodermata, Arthropoda, Annelida, Mollusca, Brachiopoda, Platyhelminthes, Cnidaria, Ochrophyta and Tracheophyta). The initial assumption, following the CPH, was that a species' genetic variation will be greatest at the centre of its distribution and will be lowest (or at least much reduced) at the periphery (edge of the distribution). If this CPH using genetic indices is true, then it may be possible to predict patterns of genetic diversity or even structure at the phylum level without the time and cost of detailed genetic investigations. If some species show a statistically significant

association under one of the distribution models using genetic diversity and connectivity indices across a geographical area, the management implications may be substantial in terms of population prioritisation, restoration, recolonisation programmes and setting of management units for conservation.

As noted above, the importance behind studying this topic regarding marine biology and distribution lies in the fact that New Zealand has a great latitudinal range in association with varied seafloor relief, which reflects a wide diversity of marine habitats (MacDiarmid *et al.*, 2012). Moreover, these habitats are home to a great diversity of endemic species (Gordon *et al.*, 2010), for many of which it has been corroborated that demographic traits vary along with latitudinal variation (Trip *et al.*, 2014). Due to this, it is possible to take advantage of the country's geography, specifically the long latitudinal gradient. However, to be able to analyse genetic data across a distributional range, extracted from multiple studies, it is recommended to test this hypothesis with a general coverage of sampling points that reflects the actual distributional reality of each species, rather than a regional subset of the full distributional range (Sagarin & Gaines, 2002b).

Interestingly, the relationship between genetic differentiation among populations and the difference in the abundance of populations, has not yet been tested. Therefore, where possible throughout this chapter, I test the idea that the species-specific abundance records follow the same pattern or model as the genetic expectations of the same species for the CPH. Whilst this testing was only possible for a subset of all studies, I seek to link genetics and abundance in ecology, to assess future monitoring efforts which may lead to species-specific effective conservation measures, for the resolution of future conservation challenges.

Materials and methods

To carry out an effective CPH testing procedure I reviewed phylogeographic genetic studies conducted across marine taxa in New Zealand. To test the CPH, a large dataset was compiled from published papers, theses, grey literature (e.g., unpublished reports), etc. This was the meta-analysis component (see details in the Supplementary Material 1-10: "*Review of phylogeographic genetic studies of marine vertebrates and invertebrates, macroalgae and plants conducted in New Zealand*"). The review in this chapter was based on work previously developed for the Ministry of Fisheries of New Zealand (Gardner *et al.*, 2010). This 2010 report involved a review of studies of marine connectivity and population genetics amongst New Zealand coastal organisms to determine the preliminary range of spatially explicit patterns of genetic structure, but not testing of the CPH. Among the studies presented in the report, I selected those that focussed on New Zealand endemic and native species. Subsequently, to cover the period 2010 to present, I searched for and added additional studies. Published papers, theses, report, etc., were downloaded directly from the VUW library e-journal subscription or in the case of items not available an interloan request was lodged via the VUW library.

The raw database containing the genetic attributes included a wide range of information, including: systematics data, species binomial, common name, distributional range, sampling location names (sites), latitude and longitude of each site sampled (if not specified in the publication, the stated site name was used to estimate the coordinates as an "approximate location"), fixation indices also referred to as genetic differentiation or connectivity measures (e.g., F_{ST} and/or Φ_{ST}), measures of genetic diversity (e.g., F_{IS}, Number of haplotypes, Allelic richness, Haplotype richness, Nucleotide diversity, Haplotype diversity, H_O, H_E, among others), life-history characteristics, depth, and the paper's metadata (for tracking purposes).

To assist in CPH testing, I arbitrarily divided New Zealand into three regions based on different latitudes – north, central, south – to try to ensure that data from each study covered all three regions, when appropriate to do so. Data from the studies that were finally tested for the CPH was filtered on a limited number of requirements. These were (1) that all of the sampling points of the study were located throughout the distributional range of the species in New Zealand (with representation of studies that presented distribution in Stewart Island, Chatham Islands and the Subantarctic Islands) and (2) a minimum of 6 sampling points (sites), with at least 2 of them in each region (north, central and south). The final aim was to

obtain representativeness along the entire distributional latitudinal ranges (referring to each study/species) in the testing procedure of the CPH.

The ultimate purpose of testing the CPH was to analyse patterns or trends in the variation of genetic indices along the latitudinal cline of New Zealand. For this, I used a group of genetic studies with multi-phylum representation that took marine species as study individuals, including in my analysis only studies that fulfilled the evaluation criteria among all those incorporated in the review of phylogeographic genetic studies conducted in New Zealand (see footnotes on Supplementary Material 1-10 to take a closer look into the specific studies which were tested).

Following Enquist *et al.* (1995) I used a geographic position parameter that allows comparison to be made between quantitative variables for a species (intraspecific comparisons) and facilitates comparison across different taxa (interspecific comparisons). This was performed separately for each study/species, using the formula proposed by Enquist *et al.* (1995) and subsequently used by other authors:

(Eq. 2.1) $$RI = 2(L - S)/R$$

Equation 2.1. Formula to calculate the range index or the relative position of a site across a latitudinal range.

where *RI* is the range index, or the relative position of each site across the range (varying between −1 and 1, where −1 = northern edge, 0 = centre of range, 1 = southern edge), *L* is the site's latitudinal location, *S* is the latitudinal midpoint of the species' range, and *R* is the latitudinal range (position parameters were measured in decimal degrees of latitude) (Enquist *et al.*, 1995; Sagarin & Gaines, 2002a; Tuya *et al.*, 2008; Rivadeneira *et al.*, 2010; Fenberg & Rivadeneira, 2011). The range index allows for the testing of data from a single species or study in which the full distributional range of the species has not been tested – that is, where only a subset of the full distributional range has been sampled, which is the usual case for most studies. The location of the subset sampled can then be considered in the context of where that subset exists, for example at the northern end, perhaps in the centre, or at the southern end of the full distributional range of the species in question. Latitudinal distribution limits for each species were determined using natural history information stated

in the introduction or discussion sections of the study (if contained) or by searching in literature records in relation to abundance and distributional data. These parameters were set within New Zealand's coastal waters, regardless of whether the species has an endemic or native distribution. For endemic species this is reasonable, and it was assumed that for non-endemic but native species, shared with other countries, that the geographic isolation of New Zealand (i.e., reduced or absent gene flow with other countries) would be sufficient for the CPH to be tested in a meaningful way.

Because most of the marine species in New Zealand have a distribution on both coasts of the mainland islands, the range index was calculated taking as a starting point the midpoint of the distribution for each species. To achieve this, a latitudinal axis was made starting from the distribution midpoint and crossing as far as possible through the centre of the country (adapting the way of calculating the distance index to the distribution centre in an island system). This makes it possible to calculate the latitudinal midpoint of the species' range, and the latitudinal range (the *S* and *R* components of Eq. 2.1). This procedure was also carried out considering the abundance and distribution records stated in the literature for each species (where both were available), to reach the most precise records for subsequent analyses. These points were taken manually using Google Earth Pro v7.3©, to later enter the geographical coordinates in decimal degrees for the respective data sets into Equation 2.1. Most of the studies included in this meta-analysis had sampling points throughout the entire range of the species' distribution in New Zealand (that is, within the three regions previously defined). This coverage of the sampling points in the complete distributional range of the species improves the robustness of statistical testing (Rivadeneira *et al.*, 2010).

After the quality control steps described above, a concise database per study was made with the geographical information of the sampling sites stated in the study, together with the genetic variables to be tested. This included the geographical position data (*RI*) of each sampling point and the measures on genetic connectivity (F_{ST} and/or Φ_{ST}) and diversity (F_{IS}, Number of haplotypes, Allelic richness, Haplotype richness, Nucleotide diversity, Haplotype diversity, H_O, H_E). Of the genetic indices, the F_{ST} and Φ_{ST} values had to be estimated in some cases by calculating an average between the pairwise differences matrix values across all sampling sites. Negative values in the original pairwise matrix were set to zero, in accordance with other studies (Verwimp *et al.*, 2020). The other variables were not modified, but were transcribed according to what was reported in each study.

Broad-scale genetic connectivity and diversity indices were fitted to four hypothetical models, following the procedure developed by Sagarin & Gaines (2002a). These four models of distribution characterise the most common abundance patterns of a species throughout its geographic range (Sagarin & Gaines, 2002b; Gaston, 2004; Sagarin *et al.*, 2006) and have been employed elsewhere (Rivadeneira *et al.*, 2010; Fenberg & Rivadeneira, 2011). The hypothetical models included: Normal (this is the classical CPH), Abundant Edge, Ramped South and Ramped North (Figure 2.1). In this way, not only was the CPH hypothesis tested, other theoretically possible distribution patterns were also tested.

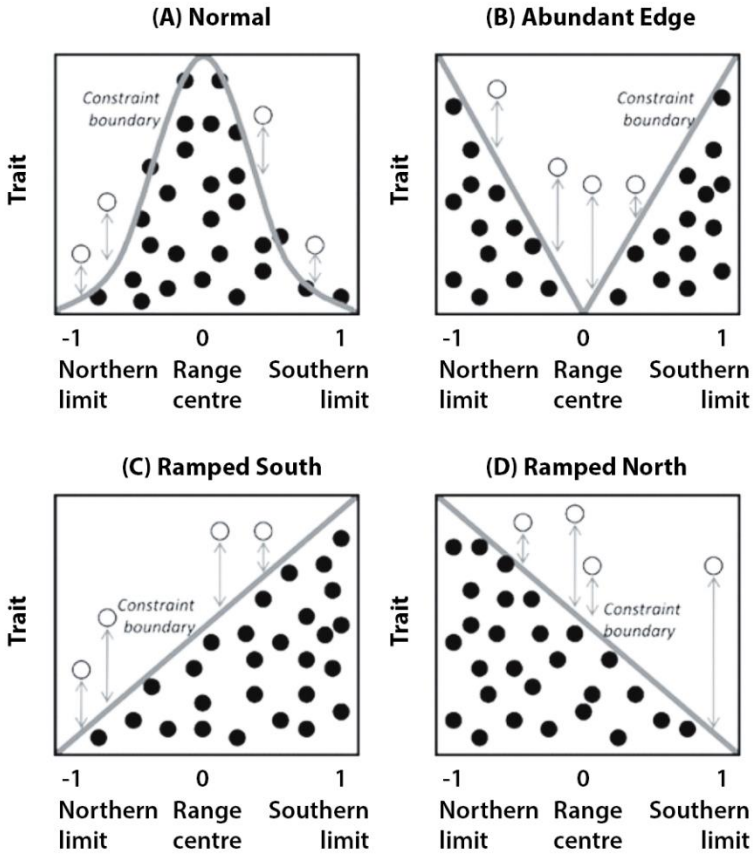


Figure 2.1. Four hypothetical models characterising the genetic diversity indices variation to be tested as a function of the geographic ranges of the species of interest. This does not apply to genetic differentiation indices (as these are expected to have different patterns). (**A**) Normal, (**B**) Abundant Edge, (**C**) Ramped South and (**D**) Ramped North. The goodness of fit of each model to the observed data was evaluated by calculating the sum of squared deviations (SS, deviations indicated by arrows) for sites exceeding the constraint boundary (open dots). Image modified from (Rivadeneira *et al.*, 2010).

This methodological approach only surveys the bounds on the trait that the researcher decides to test in different parts of the species' range. Thus, for example, it does not assume that all sites near the range centre will exhibit great values of the variable under analysis. This recognises the probability that even in the most ideal parts of the species' distribution there may be some sites with unsuitable habitat or other detrimental factors that reduce local population size or effect individual fitness (Sagarin & Gaines, 2002a).

As modified to apply to genetic metrics, in the first model (Normal, Fig. 2.1 A) maximum variation is expected to follow a normal or unimodal symmetrical distribution, with peak values being reached at the centre of the range, decreasing to zero towards the edges. The second model (Abundant Edge, Fig. 2.1 B) assumes the opposite pattern, with maximum variation observed at the edges of the range, and a minimum value in the centre. The two last models (Ramped South Fig. 2.1 C, Ramped North, Fig. 2.1 D) assume that the maximum values of variation decline from one range limit to the other and that intermediate values of variation are reached at the centre of the range.

Outliers in each data set were identified when one or more observations behaved outside the interval around the middle 50% of the data (Q1-1.5IQR and Q3+1.5IQR) (following Tukey's rule) (Tukey, 1977). The identified outliers above the interval were removed, but those under the interval were maintained in accordance with the theory behind the ACH analysis (referring to the fact that some sites may be in unsuitable habitats).

The degree of fit of each model to the observed data was evaluated by calculating the sum of squared deviations (SS) for sites exceeding the constraint boundary generated by each model. Values of SS close to zero indicate very good agreement between the model and the observed data. The statistical significance of the observed SS values was evaluated by generating 10^5 randomised values of *RI* and the genetic indices. The fit of the model was considered significant when the observed SS value was lower than the 5^{th} percentile of the randomised distribution (i.e., at alpha less than 0.05).

The degree of support for each model based on the observed data was evaluated by calculating the Akaike's information criterion (AIC) plus Akaike's difference and weight, selecting all models with Akaike weights > 0.25 (Fenberg & Rivadeneira, 2011).

It is important to appreciate that genetic indices data such as allelic richness, haplotype diversity, etc., are expected to have a positive relationship with abundance estimates, whereas index data such as F_{ST} and Φ_{ST} may be characterised as distance estimates

between (pairs of) sites. Thus, patterns of response for F_{ST} and Φ_{ST} as a function of latitude are, in effect, inverse to (mirror images of) those of allelic richness, haplotype diversity, etc., as a function of latitude. In addition, F_{ST} and Φ_{ST} estimates exhibit an asymmetrical pattern of distribution as a function of effective population size (Vucetich & Waite, 2003), meaning that models testing variation of these genetic distance estimates may need to be interpreted differently from models involving the other genetic indices variation.

The analyses were carried out by running the R script of Fenberg & Rivadeneira (2011) using RStudio©. The R script was modified in several sections, to fulfil the distribution limits of New Zealand species and for the identification of outliers and the deletion of missing values in certain rows. The latter modification was required because the analysis does not allow for blank cells, so these lines were removed from the databases (see Supplementary Material 13).

To run the data fitting analysis for each of the four distribution models, using *RI* as an independent variable and the genetic index variation as the dependent variable, I exported each database in .csv file format into RStudio. This process was carried out one study at a time, changing the dependent variable as many times as necessary depending on the total number of variables to be tested for each study.

Finally, I used abundance data from the current IUCN Red List of Threatened Species to track baseline records for taxa that are listed as Endangered and of Least Concern (i.e., considering categories other than those of Not Evaluated and Data Deficient). The purpose of this analysis was to test to determine if the genetic index variation analysed here had a similar behaviour when compared to results of population abundance or population effective size for the same species. That is, I was testing to see if the best fit model for genetic data was the same as, or different to, the best fit model for abundance data, for each species. If genetic index variation and population abundance estimates per species were both associated with some form of the latitudinal models, then the results would have powerful implications in relation to conservation and prioritisation of resources for the most endangered species inhabiting New Zealand waters. The selected data were combined with geographic coordinates to create distribution maps of the sampling points (abundance records) across the latitudinal cline using ArcGIS 10.8©. For visualisation of the dependent variables (genetic index variation) along the latitudinal gradient scatter plots were created in RStudio 1.3.1056© using the library ggplot2©.

Results

In summary, I compiled data from 134 studies that have described population genetic structure and/or genetic index variation in 65 different taxa (vertebrates, invertebrates, macroalgae and plants) of New Zealand's coastal marine biota (see Supplementary Material 12). Of these, 73 focussed on rocky intertidal and subtidal taxa, 27 on open coastal, 12 on estuarine, 11 on fjordic, 10 on soft substrate subtidal and 1 on soft substrate intertidal taxa. Vertebrates and macroinvertebrates were particularly well represented (19 and 43, respectively, out of 65 taxa). In terms of genetic population structure, 42 studies reported no spatial structure, 19 reported an isolation by distance model, 30 reported divergence within and/or among populations, 30 reported a north-south differentiation, 8 reported an east-west difference, and 5 did not specify any trend or pattern. Among the data set, 49 studies were conducted using mitochondrial markers, 33 using microsatellites, 31 using allozymes, 5 using RFLPs, 5 using RAPDs, 4 using SSCPs, 4 using SNPs, 2 using AFLPs and 1 study sequenced the ITS2 region of nuclear ribosomal DNA.

The collected data from the initial 134 studies included in the review of genetic studies comes from species within 10 different phyla (Table 2.1) (see Supplementary Material 1-10). Among these studies the most frequently occurring number of sites (sampling locations) per study (i.e., the mode) was seven, with a mean ± SD of 11 ± 9 (see Supplementary Material 11 A).

After data quality control according to the criteria outlined above (see Material and methods), sufficient information was obtained from the 134 studies for 84 data sets (62.6%) to be analysed (Table 2.1). Nine of the 10 phyla were represented in the selection (the Platyhelminthes could not be included), and at the class level the Polyplacophora (phylum Mollusca) could not be included (Table 2.1). Only 10 of the 65 different taxa were not included in the CPH model testing due to failing the underlying requirements regarding the analysis and testing (number and broad scale regional distribution of sites). The average numbers of 'sampling locations' for each phylum after the quality control, were as follow: Chordata (10±4), Echinodermata (15±9), Arthropoda (11±5), Annelida (8±0), Mollusca (17±9), Brachiopoda (12±7), Cnidaria (20±9), Ochrophyta (33±1) and Tracheophyta (18 – one study only) (see Supplementary Material 11 B-K). Across all studies the number of individual genetic indices ranged from 1 to 5 (out of 10 different indices for which data were compiled) with a mean (± SD) of 2.9 indices ± 1.3. Based on the data set of 84 studies, and because not all nine index values existed for each study, I had a total of 246 individual index-

based tests (Table 2.1). For any genetic index within each data set (i.e., any one species) only one of the four models could be statistically significant, or all four models could be not statistically significant.

Table 2.1. Summary table of the studies and number of executed tests included in the testing of the CPH. The studies were classified first by Phylum and second by Class, including information about the total number of studies, percentages (for Phylum and Class), total number of species (for Class), number of significant tests and total number of tests conducted (by Phylum).

Phylum	**N s**	**%**	**Class**	**N s**	**%**	**N spp.**	**Significant tests (p<0.05)**	**Total tests**
Chordata	26	31.0	Mammalia	9	10.7	3	20	85
			Aves	2	2.4	1		
			Actinopterygii	13	15.5	10		
			Elasmobranchii	1	1.2	1		
			Ascidiacea	1	1.2	1		
Echinodermata	8	9.5	Echinoidea	4	4.8	1	1	18
			Ophiuroidea	1	1.2	1		
			Asteroidea	3	3.6	2		
Arthropoda	9	10.7	Malacostraca	9	10.7	8	9	27
Annelida	2	2.4	Polychaeta	2	2.4	1	0	8
Mollusca	30	35.7	Bivalvia	17	20.2	7	21	89
			Gastropoda	13	15.5	11		
			Polyplacophora	0	0.0	0		
Brachiopoda	3	3.6	Rhynchonellata	3	3.6	2	1	6
Platyhelminthes	0	0.0	Trematoda	0	0.0	0		
Cnidaria	3	3.6	Anthozoa	2	2.4	2	2	8
			Hydrozoa	1	1.2	1		
Ochrophyta	2	2.4	Phaeophyceae	2	2.4	2	1	4
Tracheophyta	1	1.2	Magnoliopsida	1	1.2	1	0	1
Total	84	100		84	100	55	55	246

N s: Number of studies; %: Percentage of representation; N spp.: Number of species. The classes Polyplacophora and Trematoda (Phylum Mollusca and Platyhelminthes, respectively) are included in the table because there was representation of these taxa in the review (see Material and methods). These studies did not meet the selection criteria or simply did not have genetic data available for the hypothesis testing. They are presented here for visualisation of the broad scope of the review.

Approximately three quarters of all data sets tested did not provide a statistically significant result (i.e., 77.6% of all tests did not fit one of the four models). The results for the four hypothetical distribution models across all species revealed that 10% of all tests conducted fitted the Ramped North and 8% the Ramped South models, using a 95% confidence interval (Table 2.2). The numbers of significant test results for models with a peak or decrease of the index value at the centre of the species' distribution were under-represented. The Normal and Abundant Edge models accounted for only 2% and 2.4% of all tests conducted, respectively, of the total cases (Table 2.2). Interestingly, after clustering the results under four confidence intervals (90%, 95%, 99% and 99.9%) the same patterns or trends remained, demonstrating a homogeneity through different settings of statistical confidence parameters (Table 2.2). Each of the ten genetic metrics (i.e., diversity and distance indices) was statistically significant at least once across all taxa (see Supplementary Material 14-15). The two indices that exhibited the greatest number of significant results were number of haplotypes and haplotype diversity (≥10 significant responses each). The percentages of statistically significant results for all tests considering the genetic distance metrics (F_{ST} and/or Φ_{ST}) showed the same tendency compared to those excluding these variables (Table 2.2). That is, the inclusion or exclusion of tests for the genetic distance metrics did not influence the overall pattern of results. A notable pattern in the results was that, for any one data set, if a genetic index such as allelic richness or haplotypic diversity was significant for the Abundant Edge model then a genetic distance-based index was statistically significant for the Normal model, or vice versa. And similarly, if a data set exhibited a significant result for a non-distance-based genetic index for the Ramped North model then a genetic distance-based index was statistically significant for the Ramped South model, or vice versa.

Table 2.2. Summary table with overall percentages including all the tested genetic indices under the CPH, emphasising four models (based on different distributional ranges) of statistical significance. To the left side, including the genetic distance metrics and to the right side, excluding them.

	Indices including F_{ST} and Φ_{ST}				**Indices excluding F_{ST} and Φ_{ST}**			
Model	**NO**	**RN**	**RS**	**AE**	**NO**	**RN**	**RS**	**AE**
p<0.10	**2.8**	**15**	**12**	**3.7**	3.2	19	11	3.2
p<0.05	**2**	**10**	**8**	**2.4**	2.6	13	7	2.6
p<0.01	0.8	6.1	4	0.8	1.1	7.9	3.7	1.1
p<0.001	0.4	1.6	1	0	0.5	2.1	1.1	0

Four hypothetical models. NO: Normal; RN: Ramped North; RS: Ramped South; AE: Abundant Edge.

In total, only 29 of 55 species (~53%) exhibited a significant fit under the genetics expectations of the multi-model CPH testing. The specific results by phylum are detailed below separately (significant responses by individual index-based test, $p<0.05$) (Tables 2.3 - 2.9).

Within the Chordates (Table 2.3) all four models were statistically significant, at least once each. Of the 85 tests carried out, 20 were significant ($p<0.05$). Of the dependent variables four - haplotype diversity, nucleotide diversity, F_{ST} and Φ_{ST} - were the most commonly observed significant responses. Of the 20 significant results, 8 were for the Ramped South, 6 for the Ramped North, 5 for the Abundant Edge, and only 1 for the Normal model. By analysing the data in relation to the genetic markers used in each study for any given single species, it was observed that mitochondrial (13 significant responses) and nuclear markers (7 significant responses) in general exhibited different responses, for example, in the cases of the Hector's dolphin (*Cephalorhynchus hectori*) and the Yellow-eyed penguin (*Megadyptes antipodes*) – both exhibited Abundant Edge for the mtDNA and Ramped South for the nDNA. Mitochondrial markers appear to be the most sensitive across this phylum (mtDNA significant responses 13/43; nDNA significant responses 7/38).

Table 2.3. Phylum Chordata: Statistically significant goodness-of-fit tests for four hypothetical distribution models as a function of genetic index type.

Species	Habitat/Endemism	Index	Model	P-value	Marker	Reference
Cephalorhynchus hectori	Open Coast/NZ	Pi h	Abundant Edge Abundant Edge	***p<0.01*** ***p<0.05***	mtDNA (CR)	(Hamner *et al.*, 2012b)
	Open Coast/NZ	F_{ST}	Ramped South	***p<0.01***	SSR	(Hamner *et al.*, 2012b)
Arctocephalus forsteri	Open Coast/NZ-AUS	F_{ST}	Ramped South	***p<0.05***	SSR	(Dussex *et al.*, 2016)
Phocarctos hookeri	Open Coast/NZ	h	Normal	***p<0.05***	mtDNA (CytB)	(Collins *et al.*, 2016, 2017)
Megadyptes antipodes	Open Coast/NZ	H Pi h	Abundant Edge Abundant Edge Abundant Edge	***p<0.05*** ***p<0.01*** ***p<0.05***	mtDNA (HVI)	(Boessenkool *et al.*, 2009)
	Open Coast/NZ	H_O H_E	Ramped South Ramped South	***p<0.05*** ***p<0.01***	SSR	(Boessenkool *et al.*, 2009)
Bellapiscis medius	Rocky intertidal/NZ	Pi h	Ramped North Ramped North	***p<0.05*** ***p<0.05***	mtDNA (CR)	(Hickey *et al.*, 2009)
Forsterygion lapillum	Rocky subtidal/NZ	Pi h	Ramped North Ramped North	***p<0.001*** ***p<0.001***	mtDNA (CR)	(Hickey *et al.*, 2009)
	Rocky subtidal/NZ	Ar	Ramped North	***p<0.01***	SSR	(Rabone *et al.*, 2015)
Parapercis colias	Rocky intertidal/NZ	Pi	Ramped North	***p<0.01***	mtDNA (CR)	(Smith, 2012; Gebbie, 2014)
Rhombosolea plebeia	Estuarine/NZ	Ar h	Ramped South Ramped South	***p<0.01*** ***p<0.05***	SSR	(Constable, 2014)
Hippocampus abdominalis	Rocky intertidal/NZ-AUS	Φ_{ST} Hr	Ramped South Ramped North	***p<0.05*** ***p<0.05***	mtDNA (CR)	(Ashe & Wilson, 2020)

Model refers to model distribution shapes from Fig. 2.1. Significant ***p-values*** (bold) indicate that the distribution of the genetic indices (connectivity and diversity) fit the distribution shape more closely than a random distribution in 95% or more of the 10^5 randomisation procedures (see Material and methods). Indices. Pi: Nucleotide diversity, h: Haplotype diversity, F_{ST}-Φ_{ST}: Fixation index, H: Number of haplotypes, H_O: Observed heterozygosity, H_E: Expected heterozygosity, Ar: Allelic richness, Hr: Haplotype richness. SSR/µmicrosat = Simple sequence repeats/Microsatellites.

For the phylum Arthropoda there were 9 of 27 significant models ($p<0.05$), of which 3 were for the Ramped North and 6 for the Ramped South models (Table 2.4). The genetic indices F_{ST}-Φ_{ST}, number of haplotypes and expected heterozygosity were represented twice each, whereas all other variables that were statistically significant for a model fit were represented only once. The common rock crab (*Hemigrapsus sexdentatus*) was distinguished by its opposed results for two different diversity indices (H – Ramped South; h – Ramped North). In summary, the marker types allozymes (4 significant results) and mitochondrial COI (5 significant results) accounted for three significant studies each, whilst other marker types were not observed to be significant.

Table 2.4. Phylum Arthropoda: Statistically significant goodness-of-fit tests for four hypothetical distribution models as a function of genetic index type.

Species	**Habitat/Endemism**	**Index**	**Model**	**P-value**	**Marker**	**Reference**
Pinnotheres atrinicola	Rocky intertidal/NZ	F_{ST}	Ramped North	***p<0.05***	Allozymes	(Stevens, 1991)
Paracorophium lucasi	Estuarine/NZ	H_E	Ramped South	***p<0.05***	Allozymes	(Stevens & Hogg, 2004)
Paracorophium excavatum	Estuarine/NZ	H_O H_E	Ramped South Ramped South	***p<0.001*** ***p<0.05***	Allozymes	(Stevens & Hogg, 2004)
Hemigrapsus sexdentatus	Rocky intertidal/NZ	H h	Ramped South Ramped North	***p<0.001*** ***p<0.01***	mtDNA (COI-RFLP)	(Hinnendael, 2008)
Munida gracilis	Soft substrate subtidal/NZ-AUS	H	Ramped South	***p<0.05***	mtDNA (COI)	(Bors *et al.*, 2012)
Metanephrops challengeri	Soft substrate subtidal/NZ	Φ_{ST} Pi	Ramped South Ramped North	***p<0.05*** ***p<0.05***	mtDNA (COI)	(Verry *et al.*, 2020)

Model refers to model distribution shapes from Fig. 2.1. Significant ***p-values*** (bold) indicate that the distribution of the genetic indices (connectivity and diversity) fit the distribution shape more closely than a random distribution in 95% or more of the 10^5 randomisation procedures (see Material and methods). Indices. F_{ST}-Φ_{ST}: Fixation index, H_E: Expected heterozygosity, H_O: Observed heterozygosity, H: Number of haplotypes, h: Haplotype diversity, Pi: Nucleotide diversity.

Within the phylum Mollusca, 21 of 89 tests were statistically significant ($p<0.05$), of which 13 were for Ramped North, 5 for Ramped South and 3 for the Normal model. A significant fit for the Abundant Edge model was not observed (Table 2.5). The genetic indices number of haplotypes and haplotype diversity, both measures of population diversity, were the most sensitive variables for 7 of 9 species. Among the 13 molluscan studies which fitted the Ramped North model, 8 involved mitochondrial markers, 3 involved microsatellites

and 1 involved allozymes. Mitochondrial markers appear to be the most sensitive marker type across the molluscan studies (mtDNA significant responses 13/37; nDNA significant responses 8/45). Whilst model fits were generally consistent within a study, data for the spotted whelk (*Cominella maculosa*), which used the same mitochondrial marker, revealed both a Ramped North and Normal model fit.

Table 2.5. Phylum Mollusca: Statistically significant goodness-of-fit tests for four hypothetical distribution models as a function of genetic index type.

Species	Habitat/Endemism	Index	Model	P-value	Marker	Reference
Paphies subtriangulata	Soft substrate subtidal/NZ	H_E	Ramped North	***p<0.05***	Allozymes	(Smith *et al.*, 1989)
	Soft substrate subtidal/NZ	Ar H_E	Ramped North Ramped North	***p<0.01*** ***p<0.01***	SSR	(Hannan, 2014)
Perna canaliculus	Rocky intertidal/NZ	H h	Ramped South Ramped South	***p<0.01*** ***p<0.01***	mtDNA (NADH4, NADH2, COI - SSCP & RFLP)	(Apte *et al.*, 2003)
Pecten novaezelandiae	Soft substrate subtidal/NZ	F_{ST} F_{IS} Ar	Ramped South Ramped North Ramped North	***p<0.01*** ***p<0.01*** ***p<0.001***	SSR	(Silva & Gardner, 2016)
Ostrea chilensis	Rocky subtidal/NZ	h	Normal	***p<0.05***	mtDNA (COI)	(Guo *et al.*)
Cellana radians	Rocky intertidal/NZ	Pi h	Ramped South Ramped South	***p<0.01*** ***p<0.05***	mtDNA (CytB)	(Goldstien, 2005; Goldstien *et al.*, 2006)
Cellana flava	Rocky intertidal/NZ	Pi h	Ramped North Ramped North	***p<0.05*** ***p<0.05***	mtDNA (CytB)	(Goldstien, 2005; Goldstien *et al.*, 2006)
Haliotis iris	Rocky intertidal/NZ	Ar H_E	Ramped North Ramped North	***p<0.01*** ***p<0.01***	SSR	(Will & Gemmell, 2008; Will *et al.*, 2015)
	Rocky intertidal/NZ	H h	Ramped North Ramped North	***p<0.01*** ***p<0.001***	mtDNA (COI, ATP8-ATP6)	(Will & Gemmell, 2008; Will *et al.*, 2011)
Cominella virgata	Rocky intertidal/NZ	H	Ramped North	***p<0.05***	mtDNA (COI)	(Walton, 2017; Fleming *et al.*, 2018; Walton *et al.*, 2019)
Cominella maculosa	Rocky intertidal/NZ	H	Ramped North	***p<0.05***	mtDNA (COI)	(Walton, 2017; Fleming *et al.*, 2018; Walton *et al.*, 2019)
	Rocky intertidal/NZ	H Pi	Normal Normal	***p<0.01*** ***p<0.05***	mtDNA (COI)	(Dohner *et al.*, 2018)

Model refers to model distribution shapes from Fig. 2.1. Significant ***p-values*** (bold) indicate that the distribution of the genetic indices (connectivity and diversity) fit the distribution shape more closely than a random distribution in 95% or more of the 10^5 randomisation procedures (see Material and methods). Indices. H_E: Expected heterozygosity, Ar: Allelic richness, H: Number of haplotypes, h: Haplotype diversity, F_{ST}: Fixation index, F_{IS}: Inbreeding coefficient, Pi: Nucleotide diversity. SSR/µmicrosat = Simple sequence repeats/Microsatellites.

The results for the remaining six phyla described a wide range of model patterns or more usually the tests were not statistically significant. For the phylum Echinodermata only 1 of 18 tests was statistically significant ($p<0.05$) (Table 2.6), with only the New Zealand sea urchin (*Evechinus chloroticus*) fitting a model (Abundant Edge) for F_{ST} estimates, which is best interpreted as an asymmetrical pattern of genetic diversity.

Table 2.6. Phylum Echinodermata: Statistically significant goodness-of-fit tests for four hypothetical distribution models as a function of genetic index type.

Species	**Habitat/Endemism**	**Index**	**Model**	**P-value**	**Marker**	**Reference**
Evechinus chloroticus	Rocky subtidal/NZ	F_{ST}	Abundant Edge	***p<0.05***	SSR	(Nagel *et al.*, 2015)

Model refers to model distribution shapes from Fig. 2.1. Significant ***p-values*** (bold) indicate that the distribution of the genetic indices (connectivity and diversity) fit the distribution shape more closely than a random distribution in 95% or more of the 10^5 randomisation procedures (see Material and methods). Indices. F_{ST}: Fixation index. SSR/µmicrosat = Simple sequence repeats/Microsatellites.

For the phylum Brachiopoda only 1 of 6 tests was statistically significant ($p<0.05$) (Table 2.7). The red brachiopod (*Terebratella sanguinea*) fitted a Ramped South distribution model using AFLPs as genetic assay under H_E estimates.

Table 2.7. Phylum Brachiopoda: Statistically significant goodness-of-fit tests for four hypothetical distribution models as a function of genetic index type.

Species	**Habitat/Endemism**	**Index**	**Model**	**P-value**	**Marker**	**Reference**
Terebratella sanguinea	Fjord/NZ	H_E	Ramped South	***p<0.05***	AFLP	(Ostrow, 2004)

Model refers to model distribution shapes from Fig. 2.1. Significant ***p-values*** (bold) indicate that the distribution of the genetic indices (connectivity and diversity) fit the distribution shape more closely than a random distribution in 95% or more of the 10^5 randomisation procedures (see Material and methods). Indices. H_E: Expected heterozygosity.

For the phylum Cnidaria only 2 of 8 tests were statistically significant ($p<0.05$) (Table 2.8). The black coral (*Antipathes fiordensis*) and the Waratah anemone (*Actinia tenebrosa*) both showed a decrease of genetic diversity towards the south (Ramped North). These results involved only nuclear DNA markers (allozymes and SSRs).

Table 2.8. Phylum Cnidaria: Statistically significant goodness-of-fit tests for four hypothetical distribution models as a function of genetic index type.

Species	Habitat/Endemism	Index	Model	P-value	Marker	Reference
Antipathes fiordensis	Fjord/NZ	H_O	Ramped North	***p<0.01***	Allozymes	(Miller, 1997)
Actinia tenebrosa	Rocky intertidal/NZ-AUS	H	Ramped North	***p<0.01***	SSR	(Veale, 2007; Veale & Lavery, 2012)

Model refers to model distribution shapes from Fig. 2.1. Significant ***p-values*** (bold) indicate that the distribution of the genetic indices (connectivity and diversity) fit the distribution shape more closely than a random distribution in 95% or more of the 10^5 randomisation procedures (see Material and methods). Indices. H_O: Observed heterozygosity, H: Number of haplotypes. SSR/µmicrosat = Simple sequence repeats/Microsatellites.

Finally, for the phylum Ochrophyta only 1 of 4 tests was statistically significant ($p<0.05$) (Table 2.9). The Southern bull kelp (*Durvillaea antarctica*) exhibited maximum diversity peak at the centre of its distribution based on mitochondrial marker (COI) number of haplotypes.

Table 2.9. Phylum Ochrophyta: Statistically significant goodness-of-fit tests for four hypothetical distribution models as a function of genetic index type.

Species	Habitat/Endemism	Index	Model	P-value	Marker	Reference
Durvillaea antarctica	Rocky intertidal/SH	H	Normal	***p<0.001***	mtDNA (COI)	(Collins *et al.*, 2010)

Model refers to model distribution shapes from Fig. 2.1. Significant ***p-values*** (bold) indicate that the distribution of the genetic indices (connectivity and diversity) fit the distribution shape more closely than a random distribution in 95% or more of the 10^5 randomisation procedures (see Material and methods). Indices. H: Number of haplotypes.

No statistically significant results were observed for data sets from two phyla, the Annelida and the Tracheophyta (8 and 1 total tests, respectively) (see Supplementary Material 14). It is also worth highlighting that only the New Zealand sea lion for haplotype diversity (*Phocarctos hookeri*), the New Zealand sea urchin for F_{ST} (*E. chloroticus*), the New Zealand dredge oyster for haplotype diversity (*Ostrea chilensis*), the spotted whelk for both number of haplotypes and nucleotide diversity (*C. maculosa*) and the Southern bull kelp for number of haplotypes estimates (*D. antarctica*) fulfilled the normal model expectations of the

CPH (i.e., 6 significant responses from 5 species across 249 tests). For further details see Supplementary Material 16.

Nine species in this study are listed as being either Endangered or of Least Concern (Table 2.10). However, abundance data from the IUCN Red List could only be obtained for 5 of these 9 species (Table 2.10). In the case of marine mammals, the observed trend in species-specific abundance corroborates the pattern of genetic index diversity. For both the Hector's dolphin (Figure 2.2) and the New Zealand fur seal (Figure 2.3) abundance decreased towards the south (equivalent F_{ST} recorded pattern - asymmetrical shape). However, pup production for the New Zealand sea lion exhibits a peak in the centre of the distribution range following the same distribution as the haplotype diversity index (Figure 2.4). For the other species the number of breeding pairs of the Yellow-eyed penguin shows a decrease towards the north which coincided with the Ramped South pattern for expected heterozygosity (Figure 2.5). The relative abundance data for blue cod exhibits a decline towards the south similar to the genetic diversity attribute which fitted the Ramped North model (nucleotide diversity) (Figure 2.6). For both *C. hectori* and *M. antipodes* the best fit model shapes for the nuclear DNA marker were different from the models for the mitochondrial DNA marker (Figures 2.2 and 2.5).

Table 2.10. Conservation status for the species that showed significant observed SS values after the randomized procedure for the corresponding genetic indices, under The IUCN Red List of Threatened Species.

Phylum	**Class**	**Species**	**Conservation Status**
Chordata	Mammalia	*Cephalorhynchus hectori*	Endangered
		Arctocephalus forsteri	Least Concern
		Phocarctos hookeri	Endangered
	Aves	*Megadyptes antipodes*	Endangered
	Actinopterygii	*Bellapiscis medius*	Least Concern
		Forsterygion lapillum	Least Concern
		Parapercis colias	Least Concern
		Rhombosolea plebeia	Least Concern
		Hippocampus abdominalis	Least Concern
Echinodermata	Echinoidea	*Evechinus chloroticus*	Not Evaluated
Arthropoda	Malacostraca	*Pinnotheres atrinicola*	Not Evaluated
		Paracorophium lucasi	Not Evaluated
		Paracorophium excavatum	Not Evaluated
		Hemigrapsus sexdentatus	Not Evaluated
		Munida gracilis	Not Evaluated
		Metanephrops challengeri	Not Evaluated
Mollusca	Bivalvia	*Paphies subtriangulata*	Not Evaluated
		Perna canaliculus	Not Evaluated
		Pecten novaezelandiae	Not Evaluated
		Ostrea chilensis	Not Evaluated
	Gastropoda	*Cellana radians*	Not Evaluated
		Cellana flava	Not Evaluated
		Haliotis iris	Not Evaluated
		Cominella virgata	Not Evaluated
		Cominella maculosa	Not Evaluated
Brachiopoda	Rhynchonellata	*Terebratella sanguinea*	Not Evaluated
Cnidaria	Anthozoa	*Antipathes fiordensis*	Not Evaluated
		Actinia tenebrosa	Not Evaluated
Ochrophyta	Phaeophyceae	*Durvillaea antarctica*	Not Evaluated

The IUCN Red List Categories and Criteria are intended to be an easily and widely understood system for classifying species at high risk of global extinction. It divides species into nine categories: Not Evaluated, Data Deficient, Least Concern, Near Threatened, Vulnerable, Endangered, Critically Endangered, Extinct in the Wild and Extinct (IUCN 2020).

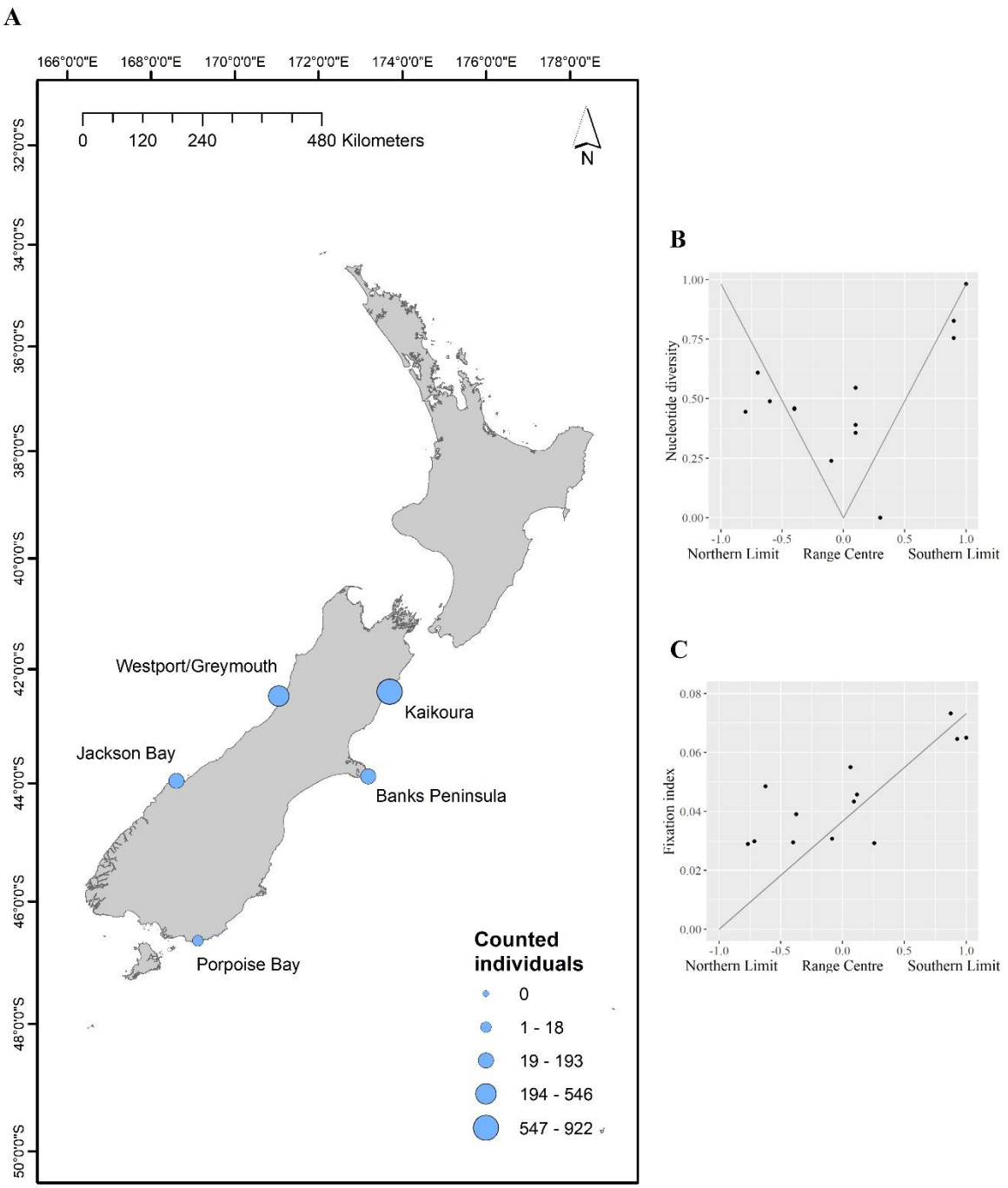


Figure 2.2. Abundance, diversity and connectivity patterns of *Cephalorhynchus hectori*. (**A**) Abundance records, (**B**) mitochondrial DNA Nucleotide diversity (Pi) and (**C**) nuclear DNA Fixation index (F_{ST}) values within the distributional range of the species in New Zealand.

Abundance data came after the review of historical censuses (Brager & Schneider, 1998; Bejder & Dawson, 2001; Dawson *et al.*, 2004; Weir & Sagnol, 2015) while genetic data are sourced from contemporary studies (Hamner *et al.*, 2012b).

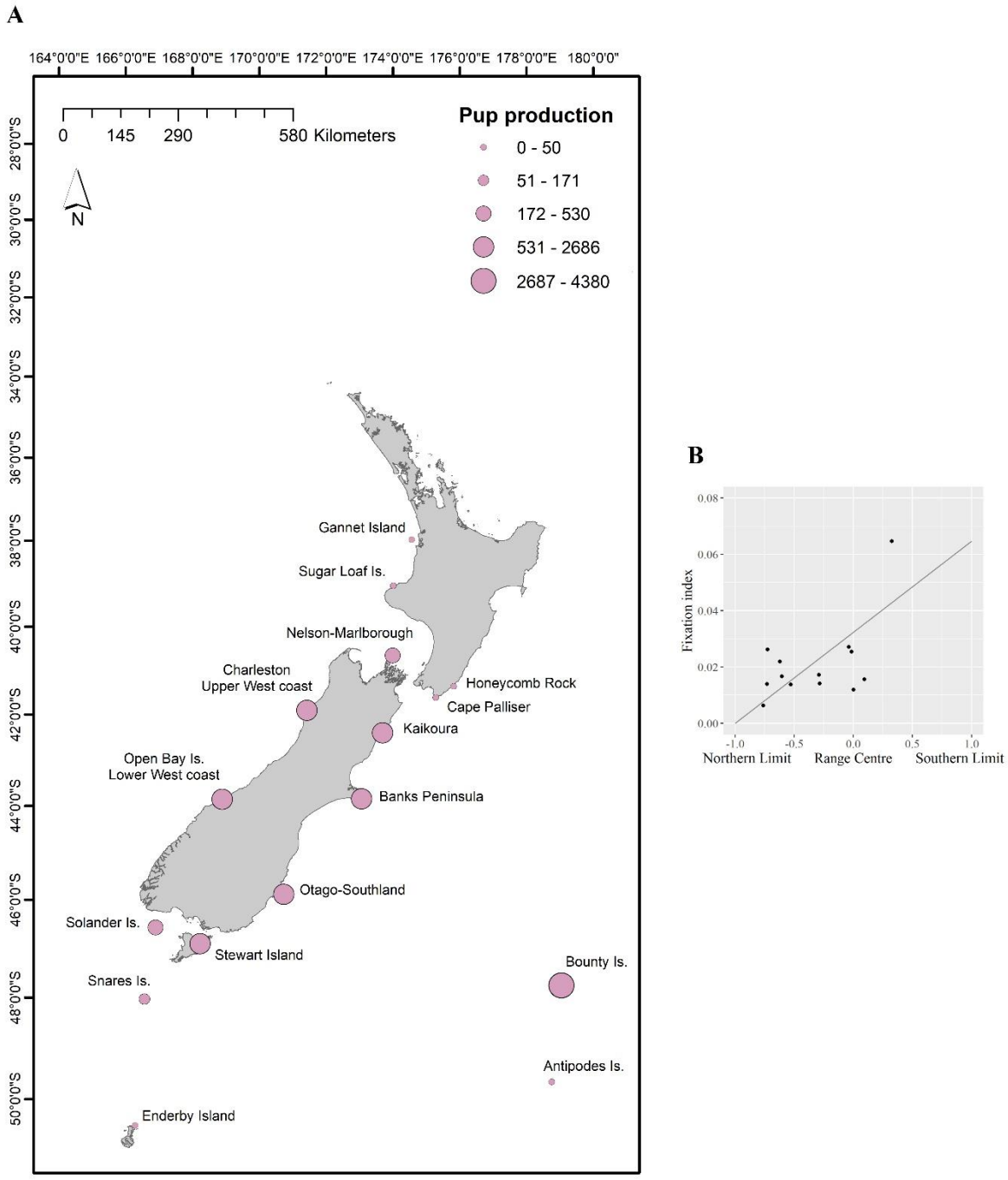


Figure 2.3. Abundance and connectivity patterns of *Arctocephalus forsteri*. (**A**) Pup production records and (**B**) nuclear DNA Fixation index (F_{ST}) values within the distributional range of the species in New Zealand.

Abundance data came after the review of historical censuses (Baird, 2011) while genetic data are sourced from contemporary studies (Dussex *et al.*, 2016).

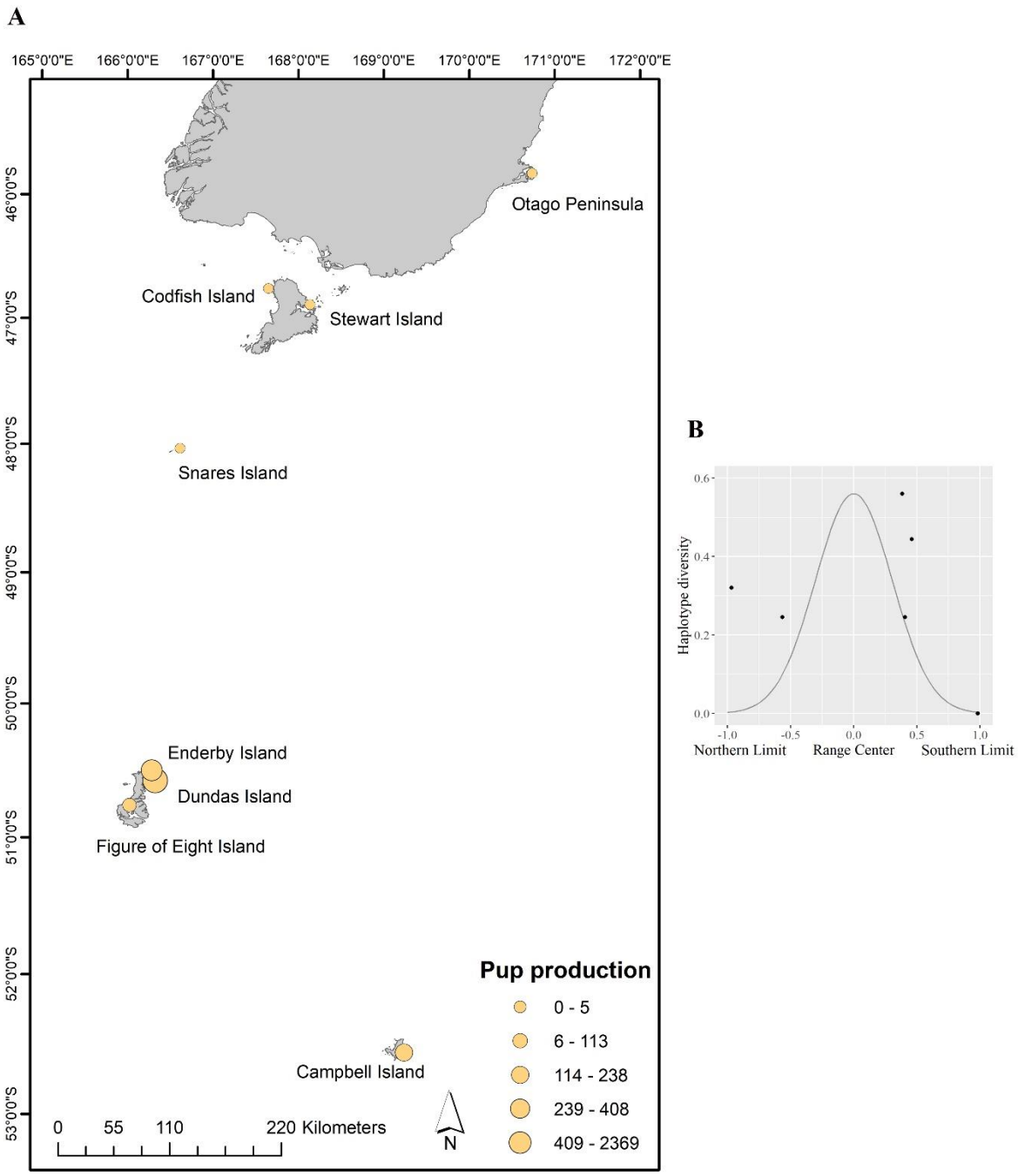


Figure 2.4. Abundance and diversity patterns of *Phocarctos hookeri*. (**A**) Pup production records and (**B**) mitochondrial DNA Haplotype diversity (h) values within the distributional range of the species in New Zealand.

Abundance data came after the review of historical censuses (Cawthorn, 1993; Childerhouse & Gales, 1998; McNally, 2001; Mcconkey *et al.*, 2002; Childerhouse *et al.*, 2005) while genetic data are sourced from contemporary studies (Collins *et al.*, 2016, 2017).

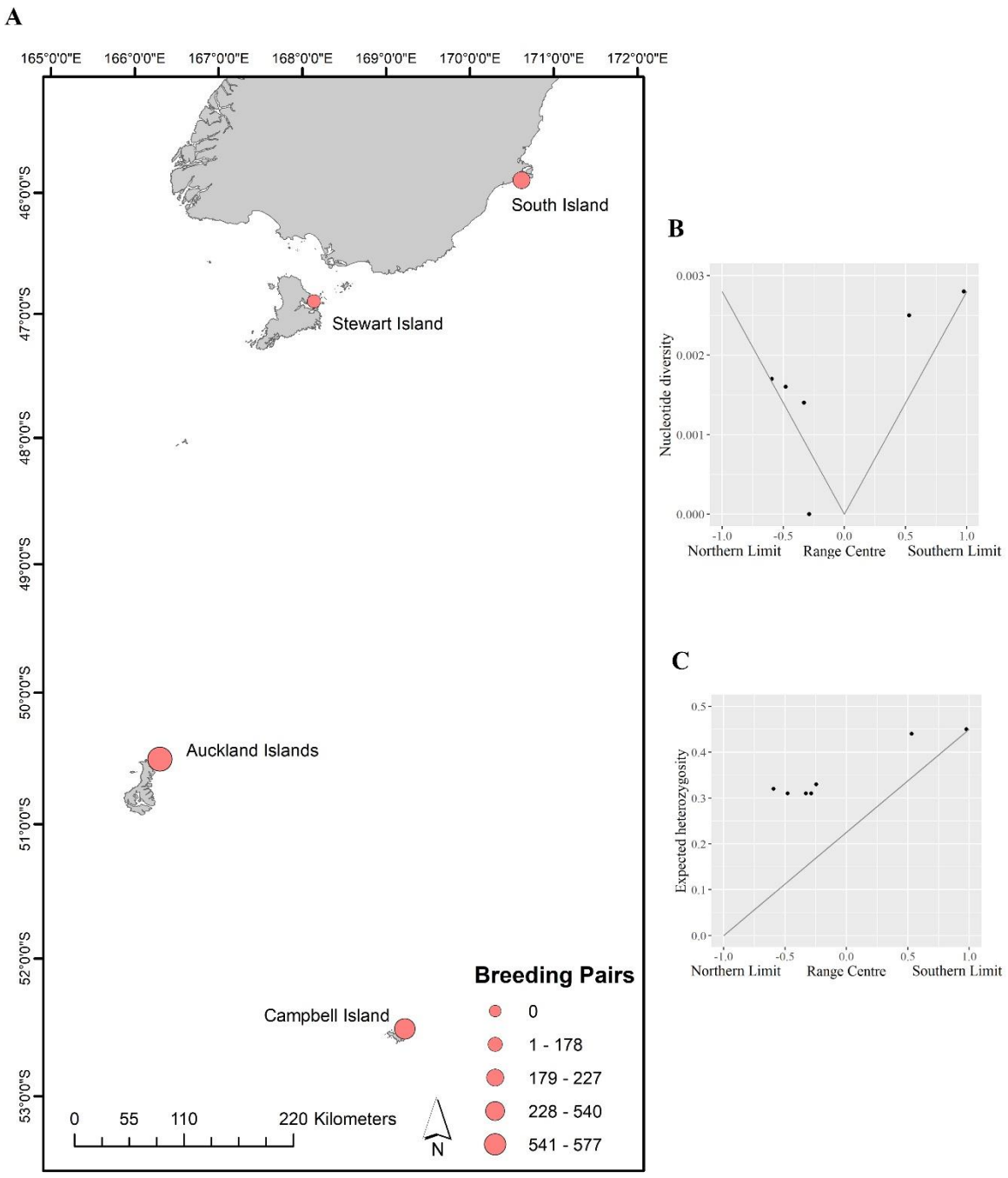


Figure 2.5. Abundance and diversity patterns of *Megadyptes antipodes*. (**A**) Breeding pair records, (**B**) mitochondrial DNA Nucleotide diversity (Pi) and (**C**) nuclear DNA Expected heterozygosity (H_E) values within the distributional range of the species in New Zealand.

Abundance data came after the review of historical censuses (Moore *et al.*, 2001; Massaro & Blair, 2003; Muller *et al.*, 2020) while genetic data are sourced from contemporary studies (Boessenkool *et al.*, 2009).

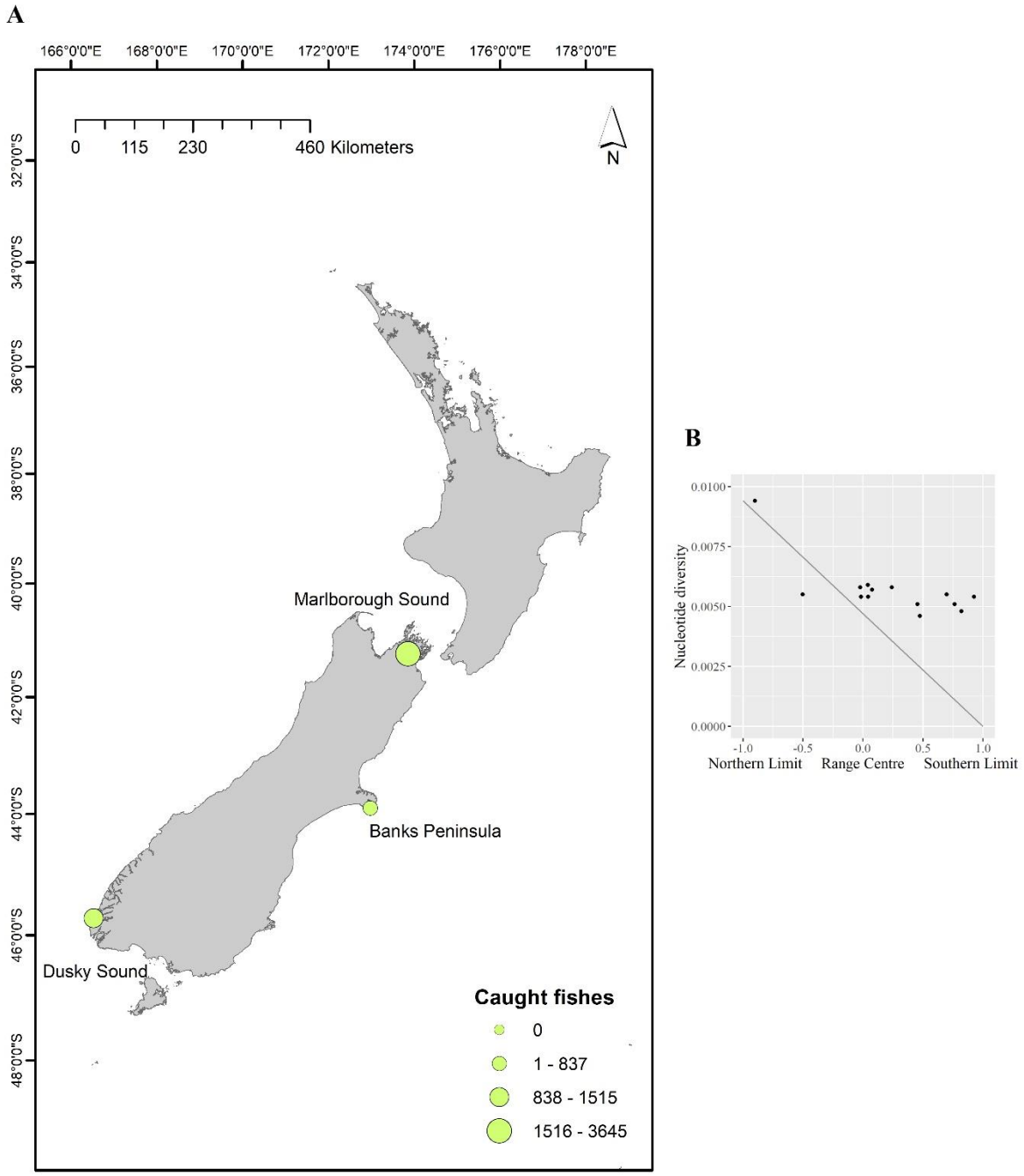


Figure 2.6. Abundance and diversity patterns of *Parapercis colias*. (**A**) Relative abundance records and (**B**) mitochondrial DNA Nucleotide diversity (Pi) values within the distributional range of the species in New Zealand.

Abundance data came after the review of historical censuses (Beentjes & Carbines, 2005; Blackwell, 2006) while genetic data are sourced from contemporary studies (Smith, 2012; Gebbie, 2014).

Discussion

Species range boundaries are the expression of a species' ecological niche in space (Sexton *et al.*, 2009). Multiple explanations to address the range limits of species have been proposed, however, none of these has prevail over the others. Among the conclusions and the more interesting insights gained by researchers in the last decades on this matter, we can mention at least three. Natural barriers raised from past geological processes in marine ecosystems (Chiswell & Sutton, 2015), extent of dispersal of pelagic larvae in marine organisms (Ross *et al.*, 2009) and life-history traits which may promote or prevent colonisation (Veale, 2007; Veale & Lavery, 2012). As is frequently the case in nature, these factors interact in an antagonistic or synergistic way to explain the observed species-specific patterns of distribution.

Because the New Zealand coast spans a latitudinal range that encompasses subtropical to subantarctic regions, it shows a great number of habitats for coastal marine flora and fauna (Rouse *et al.*, 2003; Gordon *et al.*, 2010). In this sense, the first approach in trying to explain species' distributions is the one related to geological factors. In evolutionary biology, the adaptive radiation of a species is a process in which organisms diversify rapidly from an ancestral taxon into a multitude of new forms (Grant & Grant, 2008). As can be predicted from New Zealand's geologic history and period of isolation, New Zealand subtropical-related species tend to be distributed towards the north (nearer the equator), whereas cooler water species tend to be distributed towards the south (nearer the pole). Interestingly, the fact that New Zealand land masses in past geological periods were largely submerged and then subsequently exposed, after uplifting activity in the earth's crust pair with dropping ocean water levels (Suggate *et al.*, 1978; Campbell & Hutching, 2007; Campbell, 2013), may also raise the possibility of other forms of species distributional patterns or models being possible within an insular system.

My results on the CPH revealed that only 6 of 246 tests (~2.5%) carried out fitted the peak of the genetic variation within the centre of the distribution for 55 species under examination, whereas the other 49 significant responses (~20%) fitted one of the three alternative models. When compared to other similar reviews the percentage of significant results reported here is comparatively low. For example, in support of the Abundant-Centre Hypothesis, Sagarin & Gaines (2002b) reported 39%, Martinez-Meyer *et al.* (2012) reported 36%, and Santini *et al.* (2019) reported 38% of individual test results. It should be noted, however, that in all three reviews the authors tested the Abundant-Centre Hypothesis based

on abundance data using mostly terrestrial species (including plants, insects, birds and mammals). The only two available reviews of genetic data are the ones conducted by Eckert *et al.* (2008) and Sexton *et al.* (2009). In the first paper they performed 134 tests representing 115 species, from which 64% of them followed the expected decline towards the edges of the distribution, using expected heterozygosity (H_E) as the dependent factor. H_E is an expected (not observed) estimate of genetic diversity, and whilst it is an important metric for several different reasons (e.g., testing of Hardy Weinberg equilibrium) it is frequently the case that observed heterozygosity (H_O) is significantly less than H_E for many species (Wei *et al.*, 2013b; Zeng *et al.*, 2017, 2019). In the second study, the authors summarised the main outcomes for the association of range limits with variation in population size, reporting that edges popultions are characterized by increased genetic isolation, genetic differentiation and variability. Both reviews differing substantially from my review due to the studied species (terrestrial animals and plants), number of executed tests and genetic indices used. Previous testing of marine species-specific data sets suggests that abundance data from fishes follow other types of distribution rather than the Normal (Tuya *et al.*, 2008). The same situation is almost always true for abundance data for intertidal invertebrates (Hidas *et al.*, 2010; Rivadeneira *et al.*, 2010), although it is noted that three porcelain crab species - *A. angulosus*, *P. tuberculatus* and *L. gigantea* - fitted the Abundant-Centre Hypothesis predictions (Rivadeneira *et al.*, 2010; Fenberg & Rivadeneira, 2011). For bivalve molluscs I reported greater genetic diversity towards the distributional edges, similar to the patterns that have been previously been reported in the case of the Ramped South model for *M. trossulus* and the Abundant Edge model for *M. edulis* (Tam & Scrosati, 2011). Interestingly, Shalom *et al.* (2020) reported a general gradient amongst global fish species' ranges that highlights a decrease in abundance towards the edges, but which contradicts my fish-specific results. It seems that based on this comparison the CPH is not really a good model to describe patterns of genetic diversity in marine taxa, and that the Ramped North or Ramped South models are more prevalent or important in the marine realm, although overall no model fit was the most frequently observed result.

The relationship between early life-history characteristics and genetic structure varies significantly among taxa. Previous research reported a negative correlation between Pelagic Larval Duration (hereafter PLD) and genetic differentiation, suggesting that increased larval duration is associated with decreases in genetic structure or regional differentiation (Bohonak, 1999; Bradbury *et al.*, 2008; Ross *et al.*, 2009). For example, species with longer

or extended PLDs appear to be good dispersers often with low levels of differentiation (e.g., *Antipathes fiordensis*). Conversely, for taxa with shorter PLDs, differentiation amongst populations is generally greater but highly variable (Bradbury *et al.*, 2008; Ross *et al.*, 2009). With this relationship between PLD and spatial genetic differentiation in mind, the models tested here that represent greater variability within the range of distributions of marine species are the Normal and Abundant Edge because first they assume a marked division between north and south, and second because both describe a dramatic decline (Normal) or increase (Abundant Edge) from the centre of the distribution to its edges. In this study, I report that those species which fitted the Normal model have, in general, short dispersal phases for their larvae or spores (the New Zealand dredge oyster, the spotted whelk, and the Southern bull kelp) or showed other reproductive strategies such as live birth or egg laying (the Hector's dolphin, the New Zealand sea lion, the yellow-eyed penguin), which do not promote widespread dispersal (Supplementary Material 12). The only exception here in my results is the New Zealand sea urchin, which has a PLD of approx 56 days (Supplementary Material 12). Interpretation of the results for the other species that fitted the Ramped North and Ramped South models (including The New Zealand sea urchin), will be best explained by other biological (e.g., larval behaviour and post-settlement dispersal) and physical factors (e.g., oceanographic, geological and/or climatic changes, isolation within unique systems such as the fjords). These factors, acting in combination, will determine the final extent of the species-specific dispersal process and the resulting pattern of population genetic structure (Wallis & Trewick, 2009).

Among the physical factors that may be influencing the spatially explicit patterns of genetic variation (and therefore the prevalence of the model results) for New Zealand marine species, one may highlight as the most important the oceanographic features and currents in the study area. To the south of New Zealand, the Subantarctic Front together with the related Antarctic Circumpolar Current, which surround both the Campbell Plateau and Chatham Rise, lead to complex and dynamic oceanographic features in the Southern Ocean (Heath, 1981). The Subantarctic Front represents a major natural barrier to movement (gene flow or adult migration) as a consequence of the current's circular (clockwise) pattern of flow. Northward of this proposed limit, flow is differentiated between the surface and 1000 m depth (i.e., in the area covered by the studies here tested) while further south flow tends to isolate regions inside (i.e., the Antarctic region) from regions outside (i.e., the Southern Ocean) (Chiswell & Sutton, 2015). The coastal and offshore currents around 40°S are

relevant to understand the processes shaping the distributions and population diversity at broad spatial extents (Laing & Chiswell, 2003; Will & Gemmell, 2008; Garden *et al.*, 2014). This correlates strongly with the 49 studies (~40% of all the studies included in the review) which reported an isolation by distance model or a north-south differentiation (19 and 30 studies, respectively). I suggest that the higher proportion of best fits for the ramped models are reflecting the same overall pattern suggested by the main outcomes behind this review (see Results section), as well as the results previously reported by other authors (Gardner *et al.*, 2010). Most of the studies included in my work argued for a genetic discontinuity between north and south around ~42°S, a fact that confirms a difference between the northern versus southern edge populations of many/most New Zealand marine species. I hypothesise that the aforementioned oceanographic events may play a deterministic role for the observed patterns of distribution since they have been reported extensively in the literature as being influential in terms of larval dispersal (mainly coastal currents and surface waters) and/or may act as natural barriers to larval transport (Sharples & Greig, 1998; Adams & Flierl, 2010; Chiswell *et al.*, 2017; Limer *et al.*, 2020). Nevertheless, the differences that can be attributed to a specific taxon must be analysed taking into consideration oceanographic attributes linked to inherent life-history characteristics.

In this study, data from mitochondrial and nuclear markers revealed different model fits for some marine species. It is worth highlighting the case of *C. hectori* and *M. antipodes* for which the abundant-centre model fit of abundance data does not correspond with model fit for mitochondrial DNA (nucleotide diversity). It has been widely stated and validated that genetic markers can help to understand and predict ancient (mitochondrial DNA) and recent (nuclear DNA) information (e.g., evolutionary history) about the status of populations. For example, the case study of the South American fur seal (*Arctocephalus australis*) along its full range of distribution shows that Pleistocene glaciation events may have played a deterministic role in the living conditions of these marine mammals (revealing two haplogroups corresponding to the main breeding areas in Peru and Uruguay), suggesting a long period of isolation between the Pacific and Atlantic populations (Túnez *et al.*, 2013). In addition, de Oliveira *et al.* (2009), using microsatellite markers, identified a recent genetic bottleneck in the Peruvian population of this species. These authors reported significant deviations from neutrality-equilibrium, suggesting that the Pacific population has undergone different demographic events from the Atlantic population, probably influenced by strong ENSO events (de Oliveira *et al.*, 2009). This supports the hypothesis that it is highly likely

that my results reflect changes at the population level based on the sensitivity of each marker under evaluation: mtDNA better reflects ancient (100s of generations ago) events whereas nDNA better reflects recent or contemporary events. In interpreting model fits to both mtDNA and nDNA data sets it is important to appreciate the different signals (ancient versus contemporary) that both marker types may represent.

Multiple gaps are and will continue to be opened in relation with patterns in ecology. Here, I tested a genetics-based idea in evolutionary theory about species-specific patterns of genetic variation across a latitudinal cline. Even though my results showed support for greater genetic variability at the edges of the distribution for coastal marine taxa, the significant responses in my analysis can have multiple interpretations. Moreover, my meta-analysis suggests that on the whole, different measures of diversity and connectivity are exhibiting parallel patterns of latitudinal variation, but not all of them are equally informative, as shown by the greater number of significant results for indices such as number of haplotypes and haplotype diversity compared to all other indices. Perhaps, based on these results, it is too simplistic to think that the Normal, Ramped North, Ramped South or Abundant Edge models are common when so many different factors are at play (e.g., PLD, physical oceanography, habitat, evolutionary factors influencing contemporary population genetic variation, life-history traits, etc). Nevertheless, this study has, for the first time, tested the CPH with a focus on marine coastal taxa and I report that in most cases it is not possible to fit a statistically significant model of any form to the data.

Because my meta-analysis is based on surveys that are essentially a 'snapshot' of genetic variation within populations at a particular time, it is uncertain whether the distribution patterns are representative of long-term dynamics. Very few population genetics studies examine or test temporal variation (but see Gardner & Kathiravetpillai, 1997; Gardner & Palmer, 1998; Oliver *et al.*, 2009), whereas the vast majority focus on spatial differences. Continued monitoring efforts of these sites and species would test both the robustness in time of the patterns found in this study and will provide information on how species' distributions respond to environmental change.

The ability to test for or perhaps demonstrate the relationship between species-specific abundance and genetic variability is feasible, but in-depth studies need to be conducted taking these results as a starting point to assess a statistical comparison among life-history traits and genetic attributes. The results described here indicate that it is possible to determine a

reasonably consistent pattern of genetic variation across all/most species within a phylum (Chordata: Ramped North, Arthropoda: Ramped South and Mollusca: Ramped North). Nevertheless, more sampling effort must be conducted on phyla with poor representation – that is, on the so-called 'lesser' phyla, all of which are important and represent, by definition, unique evolutionary life forms (see Table 2.1). Clearly, the results presented here make it impossible to predict patterns of latitudinal genetic variation at the taxon or phylum level for New Zealand's marine biota. Further research on this topic might perhaps predict specific pathways of genetic differentiation and diversity across multi-species by adding first, more distribution models (i.e., different from the four tested in this study) and second, more species/phyla within the analysis. To be able to use such genetic metrics as surrogates of population attributes for a given species for management purposes without the timely and consuming delays of additional new genetic investigations would be beneficial, but is not presently a realistic approach. Meanwhile, to monitor populations of key species is still relevant. Such monitoring will, it is hoped, increase the number of studies for under-represented phyla (e.g., Echinodermata, Arthropoda, Annelida, Brachiopoda, Platyhelminthes, Cnidaria, Ochrophyta and Tracheophyta) and indeed, other contributions to the biological knowledge of species and the conformation of their populations in New Zealand waters.

In conclusion, whether the CPH itself, plus its other variants (Ramped North, Ramped South, Abundant Edge), are a simplistic or a useful approach to understanding patterns of abundance and/or genetic variation, will depend on the questions that other researchers may choose to address. In my opinion, 55 of 249 significant responses (22.4%) for a pool of 55 species (29 of them fitting a model) is low. This suggests that these models are not common, at least across New Zealand's coastal biota, and that they therefore do not help us to any great extent to better understand the relationship (if one exists) between distributional range and genetic index variation. Indeed, if one thinks about the data sets which fitted the Normal model (a total of only 6 significant responses) it becomes more apparent still that the CPH does not really help explain much at all. This opens a window to propose other models of genetic distributions, here not tested, that may explain more or novel patterns in the data set. In conclusion, based on my results taking 84 studies on genetic population structure of New Zealand coastal marine taxa, I reject the CPH based on a multi-phylum marine assessment.

CHAPTER 3 Determining which environmental variables have the greatest influence on multi-species genetic variation: a seascape genetics meta-analysis of studies from New Zealand

Abstract

Genetic information is important to inform management and conservation. However, few studies have tested the relationship between genetic variation and geospatial/environmental variation across a wide range of species. To test this genetics-based idea in evolutionary theory I based my work on the approaches used previously within the discipline of seascape genetics, which in general terms, test the null hypothesis of no association between genetic variation and environmental variation. My overall aim was to test for the presence of consistent patterns of explanatory (independent) variables across species/taxa and across phyla. Here, I evaluated the role of geospatial and environmental factors as retarding or promoting gene flow (connectivity) and population genetic diversity. I conducted a meta-analysis of published/unpublished material of genetic connectivity (e.g., F_{ST} and Φ_{ST} values) and diversity indices (e.g., haplotype diversity, allelic richness) of New Zealand marine biota and the associated spatially-explicit geospatial and marine environmental data. The environmental variables were extracted from raster files of the New Zealand Exclusive Economic Zone – Marine Environmental Classification scheme (NZ EEZ-MEC) developed for the Ministry for the Environment by the National Institute of Water and Atmospheric Research (NIWA). The seascape genetics analyses involved the execution of multiple regressions using eleven geospatial and environmental variables (Latitude, Longitude, Depth, Annual mean solar radiation, Wintertime sea surface temperature, Annual amplitude of sea surface temperature, Spatial gradient annual mean sea surface temperature, Summertime sea surface temperature anomaly, Mean orbital velocity, Tidal current and Bed slope) and species-specific genetic variation. Two approaches were conducted, a general linear model (GLM, entering all independent variables simultaneously) and a methodological extension better-known as generalized linear model (GLZ stepwise regression - backward elimination, excluding non-significant independent variables sequentially) using 84 datasets of New Zealand coastal marine taxa examined for their population genetic structure. My results showed that four factors related to sea surface temperature (SSTwint, SSTanamp, SSTgrad and SSTanom) were identified as explaining significant variation in F_{ST}-Φ_{ST} (genetic distance between pairs of populations within a single study). In contrast, for genetic diversity indices, the main factors explaining variation in Allelic richness and Nucleotide diversity were latitude and longitude (but with fewer significant responses than for the F_{ST}-Φ_{ST} analyses). The results for the two phyla with greatest representation of datasets (Chordata and Mollusca) showed similar patterns. Among the connectivity indices the most important explanatory variables were wintertime sea surface temperature and summertime sea surface temperature,

whereas for the diversity indices the most important explanatory variables were latitude, longitude and mean orbital velocity. Special attention must be given to inference statistical analyses based on stepwise model selection due to its inflated Type I error rates and the greater number of statistically significant terms compared to a model building without a step-by-step iterative construction (all variables entered at the same time), which has lower Type I error rates and lower number of statistically significant terms. Whilst a range of different factors (e.g., physical oceanography (localised currents), food availability, life-history traits and harvesting), either acting alone or acting synergistically, are likely to be important in explaining patterns of genetic diversity in New Zealand's marine coastal species, my results indicate that variables including SST (SSTwint, SSTanamp, SSTgrad and SSTanom) and to a lesser extent the geospatial variables (latitude and longitude) explain much of the variation in the genetic indices tested here. The identification of such patterns across multiple phyla and species is helpful in the broader context of better understanding how environmental variation - past, present and future - contributes to genetic diversity.

Keywords: genetic variability, geospatial variation, environmental variation, allozymes, mitochondrial DNA markers, nuclear DNA markers, seascape genetics, southwestern Pacific, New Zealand.

Introduction

New Zealand has a rich diversity of marine habitats that include over 15,000 known species (Department of Conservation and Ministry of Fisheries, 2005). This status, combined with numerous recent published studies, provides an opportunity to compare the genetic diversity of organisms under different environmental and oceanographic conditions. Within the last two decades, researchers across the globe have increased our understanding of the explanatory power of genetic data by combining it with environmental data (Selkoe *et al.*, 2008). In simple terms, the discipline of seascape genetics aims to predict if environmental variation can explain (is associated with) genetic variation across the marine realm, taking as a starting point the variability in baseline environmental conditions (Liggins *et al.*, 2013). Typically, neutral genetic markers are used to understand the associations between these factors (Wei *et al.*, 2013b; Liggins *et al.*, 2019). However, the researcher must assume that the genetic marker being studied is representative of population processes throughout the genome, such as gene flow, genetic drift, and mutation (Liggins *et al.*, 2013). Equally important to be aware of is the fact that genetic differentiation between populations may increase with differences in the environment independent of the geographic distance – the so-called isolation by environment effect (Wang & Bradburd, 2014).

New Zealand's latitudinal cline (~13 degrees of latitude) and its moderate longitudinal variation (~8.5 degrees of longitude) makes the country an excellent case study location for the study of environmental features that may retard or promote gene flow between or amongst populations. Furthermore, an additional influencing factor is the north to south shoreline of the country which determines, for example, the temperature ranges and consequently the weather on both islands (warm subtropical and cool temperate rainy in the north and south, respectively) (Wei *et al.*, 2013b; Chiswell & Sutton, 2015). Because many marine species are distributed along the entire coastline of New Zealand's different habitats (Gordon *et al.*, 2010) there exists an opportunity to examine large-scale seascape genetics across multiple different taxa to test the effect of marine environmental variation on species-specific genetic structuring.

It is well-known that some parameters in nature can facilitate or restrict connections between populations or individuals. In this sense, Selkoe *et al.* (2016) conducted an insightful review of the developments in seascape genetics over the previous decade to inform methodological approaches, to identify spatial genetic patterns, and to provide best practice recommendations for applications in conservation. The authors noted that temperature,

oceanography, and geography showed an equal contribution of influence on genetic variation. Furthermore, they argued that multiple other seascape factors may be impacting connectivity at distinct spatio-temporal scales. Overall, these results suggest that connectivity in the oceanic environment is a cause-effect relationship of many interacting variables (Selkoe *et al.*, 2016). Elsewhere, the complex patterns of oceanic circulation, climatic events, pelagic larval duration, larval behaviour, food availability and species-specific evolutionary histories have been identified as key factors in explaining rates and routes of species-specific population connectivity (Barber *et al.*, 2000, 2002; Westfall, 2010; Wei *et al.*, 2013b; Silva & Gardner, 2016). The lack of any predictable order among or ranking of how these variables affect larval transport suggests at least some component of random or stochastic recruitment and dispersal events depending on each marine system under study (Siegel *et al.*, 2008).

Within New Zealand, the study of the blue mussel (*Mytilus galloprovincialis*) is one example of how the discipline of seascape genetics has effectively been used to relate environmental variables to genetic variability. In this case, annual mean sea surface temperature and winter solar radiation were significantly associated with multi-locus microsatellite allele frequencies (Westfall, 2010). The case study of the endemic greenshell mussel (*Perna canaliculus*) highlights a macrogeographic scale of genetic variation associated with sea surface temperature over its full distributional range (~11.5° of latitude) (Wei *et al.*, 2013b). Elsewhere, the contribution of other factors such as latitude, sediment type, mean orbital velocity (water speed of movement) and sea surface temperature were significant in explaining genetic variation for two New Zealand-endemic flatfish species, the open coast *Rhombosolea leporina* and the estuarine *Rhombosolea plebeia*, and in addition for the sand flounder, *R. plebeia*, both longitude and width of the estuary mouth were also significant in explaining population genetic variation (Constable, 2014). Hannan (2014), using distance estimates among populations (F_{ST}) found that environmental factors (annual amplitude of sea surface temperature, tidal and bed slope) rather than geospatial ones were better able to explain the genetic variation of open coast surf clams, the tuatua (*Paphies subtriangulata*). In addition, a mixture of geospatial and environmental predictors (latitude, longitude, geospatial distance, annual amplitude of sea surface temperature and spatial gradient annual mean sea surface temperature) contributed to explaining the genetic variation for estuarine clams, the pipi (*Paphis australis*) using the same aforementioned measure of differentiation among populations. Subsequently, Silva & Gardner (2016) reported that the scallop (*Pecten novaezelandiae*), across its full distributional range in New Zealand, shows

population genetic structuring that can be attributed to a combination of both environmental and geospatial variables, highlighting two environmental variables - freshwater input and suspended particulate matter - as being particularly important. Most recently, four deep-sea cold water taxa (one sponge, three corals) have been reported to exhibit differences among environmental variables which may be shaping the genetic structure of their populations. The observed microsatellite allele patterns exhibited a causality in dissolved oxygen for the sponge, *Poecillastra laminaris*, while dynamic topography (*Goniocorella dumosa*), sea surface temperature and surface water primary productivity (*Madrepora oculata*) and tidal current speed (*Solenosmilia variabilis*) were important for explaining genetic variation in the corals (Zeng *et al.*, 2020). These examples show how multiple case studies from New Zealand have been used in the discipline of seascape genetics, and how most results to date are species-specific, with no obvious patterns yet being apparent.

Today, a total of 42 different New Zealand coastal marine taxa have been examined for population genetic structure, with results that came from a literature review made to interpret and understand genetic structuring and connectivity in New Zealand (Gardner *et al.*, 2010). These studies, combined with work published since the 2010 report, provide the opportunity to perform a complete set of seascapes genetic analyses amongst the same species tested for the Core-Periphery Hypothesis (as explored in Chapter 2) with the aim of identifying common patterns or species-specific patterns of environmental variables that help to explain multi-taxon genetic variation.

To date several classification schemes within New Zealand's marine habitats have been carried out with the aim of characterising the geography and oceanography of key locations for management purposes (Snelder *et al.*, 2005; Hume *et al.*, 2007; Shears *et al.*, 2008; Walls, 2009). The schemes with greatest coverage and the best developed set of marine environmental variables for the entire Exclusive Economic Zone (EEZ) are the Marine Environment Classification (MEC) and the Estuarine Environment Classification (EEC), the latter specifically developed for estuaries, both produced for the Ministry for the Environment by the National Institute of Water and Atmospheric Research (NIWA) (Snelder *et al.*, 2005; Hume *et al.*, 2007). Using this unprecedented national coverage of environmental information, I aim to carry out the first spatially explicit, multispecies test of seascape genetics to explore if and how well environmental variation predicts genetic variation across a wide range of New Zealand coastal marine species. Most studies of New

Zealand seascape genetics to date have used either indices of genetic differentiation (e.g., F_{ST} or Φ_{ST}) or have used locus-specific allele frequencies. The underlying assumption in both cases has generally been that these markers are neutral, that is, not under direct selection. By working with genetic response variables such as *F*-statistics and estimates of gene flow and allelic/haplotypic/nucleotide diversity, my aim is to study how seascape variation explains genetic variability (Holsinger & Weir, 2009; Selkoe *et al.*, 2016) in the coastal waters of New Zealand.

If a model for estimating genetic variation based on environmental information using genetic connectivity and diversity indices is feasible, then I will be able to apply this methodological approach to a multispecies dataset. This multidisciplinary analysis will identify patterns of taxon-specific genetic structuring, as well as the influence of key environmental factors shaping marine genetic structure in New Zealand's coastal realm (Wei *et al.*, 2013b; Silva & Gardner, 2016; Zeng *et al.*, 2020). In this way, it will be possible to predict patterns of environmental and genetic co-variation at the phylum as well as the species level, and in the near future it may be possible to use such patterns (the genetic attributes of populations) for management purposes without the timely and costly delays of additional/new genetic investigations. The management implications are profound in terms of population prioritisation, captive breeding, restoration, recolonisation programmes and setting of management units for commercial interest species. This is the first study to explore the relationship between oceanscape and the genetics of New Zealand marine biota through the execution of a detailed meta-analysis.

Materials and methods

I carried out a broad scale assessment of seascape genetics based on phylogeographic genetic studies conducted in New Zealand during the last 40 years. To test the null hypothesis of no association between environmental and genetic variation a large dataset was compiled from published papers, theses, grey literature (e.g., unpublished reports), etc. This was the meta-analysis component (see details in the Supplementary Material 1-10: "*Review of phylogeographic genetic studies of marine vertebrates and invertebrates, macroalgae and plants conducted in New Zealand*"). The review presented in my study was based on work previously develop for the New Zealand Ministry of Fisheries. This report involved a review of studies of marine connectivity and population genetics in New Zealand coastal organisms to determine the preliminary range of patterns observed. Among the studies presented in this report I selected those that focussed on endemic and native species to New Zealand. Subsequently, I searched for new studies that post-date the original report (Gardner *et al.*, 2010) to build up a data base that spans approximately 40 years of research. Published papers, theses, report, etc., were downloaded directly from the VUW library e-journal subscription or in case the items were not available an interloan request was lodged via the VUW library system.

The raw database containing the genetic attributes included a wide range of information, including: taxonomic data, species name (binomial), common name, distributional range, sampling location names, latitude and longitude of each population sampled (if they were not specified in the publication, the stated site name was used to estimate the coordinates as an "approximate location"), fixation indices also referred to as genetic differentiation or connectivity measures (e.g., F_{ST} and/or Φ_{ST}), measures of genetic diversity (e.g., F_{IS}, Number of haplotypes, Allelic richness, Haplotype richness, Nucleotide diversity, Haplotype diversity, H_O, H_E), life-history characteristics, sample depth, paper metadata (for tracking purposes), among others.

The independent variables tested in the analysis came from three main sources. Geospatial variables come from each paper's metadata or, otherwise, from Google Earth Pro 7.3©. Environmental variables were extracted from raster files of the New Zealand Exclusive Economic Zone – Marine Environmental Classification scheme (EEZ-MEC). I did not use for this analysis the categorical variables sediment type and freshwater fraction due to concerns regarding their reliability (Snelder *et al.*, 2005). Following Leathwick *et al.* (2004) I used the angle of the slope of the seabed (bed slope) as the best single independent variable for seabed

form. The two extracted geospatial variables were: Latitude (Lat) and Longitude (Lon). The eleven extracted environmental variables were: Depth (Depth), Annual mean solar radiation (Rad_mean), Winter solar radiation (Rad_wint), Wintertime sea surface temperature (SSTwint), Annual amplitude of sea surface temperature (SSTanamp), Spatial gradient annual mean sea surface temperature (SSTgrad), Summertime sea surface temperature anomaly (SSTanom), Mean orbital velocity (Orb_v_mean), Extreme orbital velocity (Orv_v_95), Tidal current (Tidal) and Bed slope (Bed_slope) (see Supplementary Material 17 and 18 for further details on the EEZ-MEC environmental variables and extraction methodology using ArcGIS v10.8).

Correlation analyses and a Principal Component Analysis (PCA) were performed using the software package STATISTICA v.10© and Minitab17©, respectively, to identify highly correlated environmental variables. Pearson correlation coefficients between all pairs of environmental variables across all studies were calculated for the 11 environmental data variables collected from EEZ-MEC. To overcome colinearity, one of the predicator variables was eliminated from the data set when bivariate correlations were >0.95 (e.g. Silva & Gardner, 2016). The PCA was conducted after elimination of correlated variables to corroborate that the final variables to be incorporated in the seascape genetics analysis did not show varimax factor (VF) values > 0.75, because they are considered to be strong factor loadings (second test to verify that there were no high correlations among the selected variables suitable to performed the seascapes genetics approach) (Azid *et al.*, 2015). After this process, I used for the analyses 2 geospatial and 9 environmental variables (for more details on the specific independent variables used see Results section).

The selection of the genetics datasets to be used in the seascape analysis was based on the total number of sampled populations in each study. In strict terms, to run multiple regressions it is necessary to evaluate the relationship between the number of predictor variables (maximum) and the number of observations for each dependent variable under evaluation (the value that will define the extension of the model) (Eq. 3.1) (Lindsey, 1997; Dobson & Barnett, 2008).

(Eq. 3.1) $$N\ rows\ (sites\ obs.) = N\ ind.variables + intercept + 1$$

Equation 3.1. Explanatory formula for the required number of observations under general multivariate regression models.

From Eq. 3.1 it follows that each study requires 11 (independent variables) + 1 (intercept) + 1 (constant) = 13 site-specific observations. This meant that I had a problem caused by the limited number of sites sampled by most studies and greater number of geospatial and environmental variables available to be included in the models (n=11). I therefore decided to set the limit of entry to testing at ≥ 6 sampling locations (sites) to ensure a fair degree of representativeness among the reviewed studies (see Supplementary Material 11 A). Therefore, I split the 11 geospatial/environmental variables (after elimination of highly correlated variables) to be tested into groups, for the case of studies with 6-7 (Eq. 3.2) and ≥8 (Eq. 3.3) sites, respectively. As detailed below:

(Eq. 3.2) $$Group\ 1\ (\ 6-7\ site\ obs.\ ) = (Lat, Lon, Tidal, Depth) + (Bed.slope, Orb.v.mean, Rad.mean) + (SST.anom, SST.wint, SST.grad, SST.anamp)$$

Equation 3.2. Explanatory formula for the geospatial/environmental variables grouping for the studies with 6 or 7 site observations.

(Eq. 3.3) $$Group\ 2\ (\geq 8\ site\ obs.\ ) = (Lat, Lon, Tidal, Depth, Bed.slope, Orb.v.mean) + (SST.anom, SST.wint, SST.grad, SST.anamp, Rad.mean)$$

Equation 3.3. Explanatory formula for the geospatial/environmental variables grouping for the studies with ≥8 site observations.

A database entry per study was made with the dependent variables (genetic indices) together with the 11 geospatial/environmental variables (i.e., after correlation testing and PCA) intended to be evaluated in each routine. Among the measures of genetic variability, I included measures of genetic differentiation (F_{ST} and/or Φ_{ST}) and diversity indices (F_{IS}, Number of haplotypes, Allelic richness, Haplotype richness, Nucleotide diversity, Haplotype diversity, H_O, H_E). Among all the genetic indices, only the genetic differentiation values were obtained by calculating an average between the pairwise differences matrix values across sampling locations. Negative F_{ST} or Φ_{ST} values were set to zero in the pairwise matrix, following Silva & Gardner (2016) and Verwimp *et al.*, (2020). The other variables were not modified but were transcribed according to what was reported in each study.

In the first instance a general linear model (GLM), which is a multiple regression analysis between several independent variables and a dependent variable, was conducted for each dataset in the database. For the GLM, I used a procedure for variable selection in which all variables in a block are entered in a single step. Secondarily, a generalized linear model (GLZ), which is an extension of the GLM that allows residuals to have distributions other than normal, was conducted for each dataset to evaluate the same hypothesis. For the GLZ, I used a backward elimination stepwise regression approach, which it is commonly used to remove those independent variables that do not make a significant contribution to explaining variation in the dependent variable (Lindsey, 1997; Dobson & Barnett, 2008). GLM and GLZ analyses were performed separately on data from each study. All analyses were conducted using the package StepReg© in RStudio 1.3.1056© assuming a normal distribution for continuous dependent variables but a poisson distribution for discrete dependent variables (Lindsey, 1997; Dobson & Barnett, 2008) (see Supplementary Material 19). It is worth highlighting that after the extraction of the variables from the EEZ-MEC, the process returned missing values among the independent variables for many sampled locations. To help address this unbalanced data set I used a missing values imputation procedure, for which the mean value for the variable was calculated from all other site-specific values (i.e., mean imputation) and this value was used in place of the site-specific missing value (Little & Rubin, 2020). After this process, some studies showed identical values for certain environmental variables across the sampling sites (i.e., no variation). Because of this, some GLM/GLZ analyses were conducted with fewer independent variables than the full 11 geospatial/environmental variables.

I exported each database in .csv file format to RStudio. This process was carried out one study at a time, changing the dependent variable as many times as necessary depending on the total number of variables to be tested for each study.

To summarise the outcomes of both GLM and GLZ analyses, heatmaps were created using the frequencies of occurrences for each independent variable that was significant (p-values <0.05) for the testing of each genetic index in RStudio using the library ggplot2© and tidyr©. Because inference statistical analyses based on stepwise model selection are expected to exhibit inflated Type I error rates than a model building without a step-by-stepiterative constuction (Mundry & Nunn, 2009), I focus on the results from the GLMs, but include the GLZ results in the Supplementary Materials section for completeness.

The main purpose behind this methodology was to look for patterns across multiple taxa/studies and thereby to see if it may be possible to identify if any of the independent variables here tested here consistently help explain variation in the dependent variables that inform genetic attributes of differentiation and diversity across multiple taxa. In particular, my aim was to elucidate if some independent variables are much more important than others in terms of explaining genetic variation across phyla and amongst species.

Results

Studies spanning 10 different phyla were identified, from which data to be used in the seascape analysis could be compiled (see Supplementary Material 1-10). Among these studies the mode value of the variable 'sampling locations' was 7 sites with a mean ± SD of 11 ± 9 (see Supplementary Material 11 A). In summary, I identified 134 studies which have described population genetic structure in 65 different taxa (vertebrates, invertebrates, macroalgae and plants) of New Zealand's coastal marine biota (see Supplementary Material 12). Of the 134 studies, 73 focussed on rocky intertidal and subtidal taxa, followed by open coastal (27), estuarine (12), fjord (11), soft substrate subtidal (10) and soft substrate intertidal (1) taxa. Vertebrates and macroinvertebrates are particularly well represented (19 and 43, respectively, out of 65 taxa). In terms of genetic population structure, 42 studies reported no structure, 19 reported an isolation by distance model, 30 reported divergence within and/or among populations, 30 reported a north-south differentiation, 8 reported an east-west difference, and 5 did not specify any trend. Among the studies, 49 were conducted using mitochondrial markers, 33 microsatellites, 31 allozymes, 5 RFLPs, 5 RAPDs, 4 SSCPs, 4 SNPs, 2 AFLPs and 1 study sequenced the ITS2 region of nuclear ribosomal DNA.

Approximately 63% of the total studies (i.e., 84 of 134 studies) in the review were selected for analysis (Table 3.1) based on the sole entry criterion of a minimum number of sites ($n \geq 6$) per study (the 50 studies that were not included covered 10 species, two additional classes - the Trematoda and the Polyplacophora - and one additional phylum, the Platyhelminthes).

The mean ± SD numbers of 'sampling locations' for each phylum after the quality control, were as follow: Chordata (10±4), Echinodermata (15±9), Arthropoda (11±5), Annelida (8±0), Mollusca (17±9), Brachiopoda (12±7), Cnidaria (20±9), Ochrophyta (33±1) and Tracheophyta (18 – one study only) (see Supplementary Material 11 B-K). Across all studies the number of individual genetic indices ranged from 1 to 5 (out of 10 different indices) with a mean ± SD of 2.9 ± 1.3. Based on the data set of 84 studies, and because not all ten index values existed for each study, I had a total of 246 individual index-based analyses (multiplied by 2 if I consider both the GLM and the GLZ analyses) (Table 3.1).

Table 3.1. Summary table of the studies which met the basic requirements to test the seascape genetics approach and the total number of tests (GLM and GLZ) executed. The studies were classified first by Phylum and then by Class, and finally by species.

Phylum	**N s**	**%**	**Class**	**N s**	**%**	**N spp.**	**GLM and GLZ tests**
Chordata	26	31.0	Mammalia	9	10.7	3	170
			Aves	2	2.4	1	
			Actinopterygii	13	15.5	10	
			Elasmobranchii	1	1.2	1	
			Ascidiacea	1	1.2	1	
Echinodermata	8	9.5	Echinoidea	4	4.8	1	36
			Ophiuroidea	1	1.2	1	
			Asteroidea	3	3.6	2	
Arthropoda	9	10.7	Malacostraca	9	10.7	8	54
Annelida	2	2.4	Polychaeta	2	2.4	1	16
Mollusca	30	35.7	Bivalvia	17	20.2	7	178
			Gastropoda	13	15.5	11	
			Polyplacophora	0	0.0	0	
Brachiopoda	3	3.6	Rhynchonellata	3	3.6	2	12
Platyhelminthes	0	0.0	Trematoda	0	0.0	0	
Cnidaria	3	3.6	Anthozoa	2	2.4	2	16
			Hydrozoa	1	1.2	1	
Ochrophyta	2	2.4	Phaeophyceae	2	2.4	2	8
Tracheophyta	1	1.2	Magnoliopsida	1	1.2	1	2
Total	84	100		84	100	55	492

N s: Number of studies; %: Percentage of representation; N spp.: Number of species. The classes Polyplacophora and Trematoda (Phylum Mollusca and Platyhelminthes, respectively) are included in the table because there was representation of these taxa in the review (see Material and methods). These studies did not meet the selection criteria or simply did not have genetic data available for the hypothesis testing. They are presented here for visualization of the broad scope of the review. GLM and GLZ tests conducted by phylum were multiplied by 2 (because the same variables were used for both routines).

The analysis of bivariate (pairwise) correlations showed that 4 of 11 environmental variables extracted from the EEZ-MEC were significantly correlated across all the studies included in the review. The pairs of variables which exhibited a high correlation value were annual mean solar radiation - winter solar radiation (r^2= 0.99) and mean orbital velocity - extreme orbital velocity (r^2= 0.95). Consequently, annual mean solar radiation and mean orbital velocity were retained in the analyses and chosen to represent the variables solar

radiation and orbital velocity, respectively (i.e., including, by association, variation in the two dropped variables). This process resulted in 9 environmental variables plus 2 geospatial variables (latitude and longitude) to be tested as independent variables for each site.

The output of the initial explorative analysis of the data (PCA) showed that the first six components (axes) explained 81% of the cumulative variance out of a total of 11 components. The first component explained 24% of the variance and the major contributing factors on this axis where latitude (0.57), wintertime sea surface temperature (0.54), mean orbital velocity (-0.31) and tidal (-0.22). The second component explained 16% of the variance. The major contributing factors on this axis were bed slope (-0.54), depth (-0.51), annual amplitude of sea surface temperature (0.43) and longitude (0.38). The following four components explained 12%, 11%, 9% and 8% of the variance, respectively. There was no evident segregation when coding the data points by phylum. VF scores were < 0.75, so that the 11 variables could be entered into the GLM/GLZ analyses.

Among the 10 connectivity/diversity genetic indices selected as dependent variables, 9 were represented across multiple studies, and one - haplotype richness - was represented in only one study (Tables 3.2-3.10). This variable was therefore not included in further analyses. The greatest number of individual studies was observed for the Chordata (26 of 84 studies, ~31%) and the Mollusca (30 of 84 studies, ~36%) (Table 3.1). My results focus on the 'all phyla' testing and then secondarily on the Chordata and the Mollusca because these two phyla have enough studies to allow for phylum-specific interrogation of patterns in the results, whereas the other phyla have too few studies for this.

Across all studies regardless of phylum, there were no pronounced patterns indicating how variation in the 11 independent variables contributed to variation in the nine genetic indices (Figure 3.1). One variable (Tidal) did not explain any statistically significant variation in the genetic indices, and two variables (Bedslope and Radmean) explained only a low level of significant variation in the genetic metrics. For the genetic distance metrics (F_{ST} and/or Φ_{ST}) variation was best explained by the four sea surface temperature variables (SSTanamp = SSTgrad > SSTanom > SSTwint) (Figure 3.1). Elsewhere, whilst the other independent variables did exhibit low level statistical significance, the only notable results were for Latitude, which exhibited several statistically significant associations with the genetic metrics, but not with F_{ST} or Φ_{ST} (Figure 3.1). The most pronounced statistically significant

result was noted for variation in Allelic Richness (Ar) as a function of variation in Latitude (Figure 3.1).

Across all studies in the analysis for the phylum Chordata there were no pronounced patterns of results (Table 3.2), consistent with the case presented for 'all phyla' (Figure 3.1). Two variables (Tidal and SSTanom) did not explain any statistically significant variation in the genetic indices, and based on the observed proportion of significant responses three variables (Bedslope, Orbvmean and Radmean) explained only a low level of significant variation in the genetic metrics. For the genetic distance metrics (F_{ST} and/or Φ_{ST}) variation was best explained by two sea surface temperature variables (SSTwint > SSTanamp) and less by three other variables (Bedslope = Orbvmean = Radmean) (Figure 3.2). While the other independent variables did exhibit low/middle level statistical significance, the only notable results were for Latitude, Longitude, Depth and Spatial gradient annual mean sea surface temperature, all of which exhibited statistically significant associations with the genetic metrics, but not with F_{ST} or Φ_{ST} (Figure 3.2). The most pronounced statistically significant result was noted for variation in Number of haplotypes (H) as a function of variation in Longitude, Inbreeding coefficient (F_{IS}) with Depth, Expected heterozygosity (H_E) with Spatial gradient annual mean sea surface temperature and Nucleotide diversity (Pi) with Latitude (Figure 3.2).

Across the studies for the phylum Mollusca, there were no pronounced patterns in the results (Figure 3.3), consistent with the results for 'all phyla' and also for the Chordata (Figure 3.1-3.2). Three variables (Depth, Bedslope and Radmean) did not explain statistically significant variation in the genetic indices, and interestingly, one variable (Tidal) explained for the genetic index Observed heterozygosity (H_O) a greater proportion of significant responses (but not significant variation for the other genetic metrics). For the genetic distance metrics (F_{ST} and/or Φ_{ST}) variation was best explained by three sea surface temperature variables (SSTanom > SSTgrad > SSTanamp) (Figure 3.3). Elsewhere, whilst the other independent variables did exhibit low/middle level statistical significance, the only notable results were for Latitude, Longitude and Mean orbital velocity which exhibited several statistically significant associations with the genetic metrics, but not for F_{ST} or Φ_{ST} (Figure 3.3). The most pronounced statistically significant result was noted for variation in Allelic Richness (Ar) as a function of variation in Latitude, Inbreeding coefficient (F_{IS}) with Longitude, and Observed heterozygosity and also Inbreeding coefficient with Mean orbital velocity (Figure 3.3).

Interestingly, for the Chordata, variation in Allelic richness (Ar) was related to variation in four predictors (Lon = Depth = Bedslope = Radmean), whereas variation in Nucleotide diversity (Pi) was related to variation in two (Lat = Lon) (Figure 3.2). In contrast, variation in Ar for the Mollusca was associated with variation in 7 predictors (Lat > Long = Orbvmean = SSTanom = SSTwint = SSTgrad = SSTanamp) and for Pi it was associated with just one (SSTwint) (Figure 3.3).

In general terms, the results presented here using the GLM approach showed the same key predictor variables as the results from the GLZ analyses (see Supplementary Material 20 and 21). The only observed difference was the greater number of significant responses of independent variables for the GLZ, the GLZ being more sensitive than the GLM approach, but also more subject to increased Type I error rates, due to the greater number of tests carried out in the stepwise regression (see Supplementary Material 21).

Table 3.2. Genetic differentiation values (F_{ST}-Φ_{ST}: genetic distance between/amongst populations) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLM analysis across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

Index	Ind. Var	All phyla	Chordata (17)	Echinodermata (6)	Arthropoda (7)	Annelida (2)	Mollusca (20)	Brachiopoda (2)	Cnidaria (1)	Ochrophyta (0)	Tracheophyta (1)
F_{ST}-Φ_{ST} (GLM)	Lat	4	0	2	1	1	0	0	0	-	0
	Lon	6	2	1	0	0	3	0	0	-	0
	Tidal	2	1	1	0	0	0	0	0	-	0
	Depth	3	1	0	0	0	2	0	0	-	0
	Bedslope	4	3	0	0	1	0	0	0	-	0
	Orbvmean	5	3	0	0	0	2	0	0	-	0
	SSTanom	9	1	1	0	0	6	1	0	-	0
	SSTwint	8	4	1	0	0	3	0	0	-	0
	SSTgrad	10	1	2	0	0	5	1	1	-	0
	SSTanamp	10	3	1	1	0	4	1	0	-	0
	Radmean	6	3	1	0	0	1	0	1	-	0

Table 3.3. Index of genetic diversity (F_{IS}: Inbreeding coefficient) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLM analysis across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

Index	Ind. Var	All phyla	Chordata (6)	Echinodermata (1)	Arthropoda (1)	Annelida (0)	Mollusca (4)	Brachiopoda (0)	Cnidaria (0)	Ochrophyta (0)	Tracheophyta (0)
F_{IS} (GLM)	Lat	1	1	0	0	-	0	-	-	-	-
	Lon	1	0	0	0	-	1	-	-	-	-
	Tidal	0	0	0	0	-	0	-	-	-	-
	Depth	2	2	0	0	-	0	-	-	-	-
	Bedslope	0	0	0	0	-	0	-	-	-	-
	Orbvmean	2	1	0	0	-	1	-	-	-	-
	SSTanom	0	0	0	0	-	0	-	-	-	-
	SSTwint	0	0	0	0	-	0	-	-	-	-
	SSTgrad	1	0	0	0	-	1	-	-	-	-
	SSTanamp	0	0	0	0	-	0	-	-	-	-
	Radmean	0	0	0	0	-	0	-	-	-	-

Table 3.4. Index of genetic diversity (H: Number of haplotypes) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLM analysis across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

Index	Ind. Var	All phyla	Chordata (8)	Echinodermata (2)	Arthropoda (3)	Annelida (2)	Mollusca (12)	Brachiopoda (0)	Cnidaria (2)	Ochrophyta (2)	Tracheophyta (0)
H (GLM)	Lat	3	0	0	1	0	1	-	1	0	-
	Lon	3	2	0	1	0	0	-	0	0	-
	Tidal	2	1	0	0	0	0	-	0	1	-
	Depth	1	0	0	0	0	1	-	0	0	-
	Bedslope	0	0	0	0	0	0	-	0	0	-
	Orbvmean	1	0	0	0	0	0	-	0	1	-
	SSTanom	1	1	0	0	0	0	-	0	0	-
	SSTwint	1	0	1	0	0	0	-	0	0	-
	SSTgrad	2	1	0	0	0	0	-	1	0	-
	SSTanamp	1	0	0	0	0	0	-	1	0	-
	Radmean	0	0	0	0	0	0	-	0	0	-

Table 3.5. Index of genetic diversity (Ar: Allelic richness) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLM analysis across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

Index	Ind. Var	All phyla	Chordata (6)	Echinodermata (0)	Arthropoda (0)	Annelida (0)	Mollusca (6)	Brachiopoda (0)	Cnidaria (0)	Ochrophyta (0)	Tracheophyta (0)
Ar (GLM)	Lat	3	1	-	-	-	2	-	-	-	-
	Lon	1	0	-	-	-	1	-	-	-	-
	Tidal	0	0	-	-	-	0	-	-	-	-
	Depth	1	1	-	-	-	0	-	-	-	-
	Bedslope	1	1	-	-	-	0	-	-	-	-
	Orbvmean	1	0	-	-	-	1	-	-	-	-
	SSTanom	1	0	-	-	-	1	-	-	-	-
	SSTwint	1	0	-	-	-	1	-	-	-	-
	SSTgrad	1	0	-	-	-	1	-	-	-	-
	SSTanamp	1	0	-	-	-	1	-	-	-	-
	Radmean	1	1	-	-	-	0	-	-	-	-

Table 3.6. Index of genetic diversity (Hr: Haplotype richness) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLM analysis across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

Index	Ind. Var	All phyla	Chordata (1)	Echinodermata (0)	Arthropoda (0)	Annelida (0)	Mollusca (0)	Brachiopoda (0)	Cnidaria (0)	Ochrophyta (0)	Tracheophyta (0)
Hr (GLM)	Lat	0	0	-	-	-	-	-	-	-	-
	Lon	0	0	-	-	-	-	-	-	-	-
	Tidal	0	0	-	-	-	-	-	-	-	-
	Depth	0	0	-	-	-	-	-	-	-	-
	Bedslope	0	0	-	-	-	-	-	-	-	-
	Orbvmean	0	0	-	-	-	-	-	-	-	-
	SSTanom	0	0	-	-	-	-	-	-	-	-
	SSTwint	0	0	-	-	-	-	-	-	-	-
	SSTgrad	0	0	-	-	-	-	-	-	-	-
	SSTanamp	0	0	-	-	-	-	-	-	-	-
	Radmean	0	0	-	-	-	-	-	-	-	-

Table 3.7. Index of genetic diversity (Pi: Nucleotide diversity) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLM analysis across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

Index	Ind. Var	All phyla	Chordata (14)	Echinodermata (1)	Arthropoda (3)	Annelida (2)	Mollusca (10)	Brachiopoda (0)	Cnidaria (0)	Ochrophyta (1)	Tracheophyta (0)
Pi (GLM)	Lat	5	3	0	1	0	1	-	-	0	-
	Lon	4	3	0	1	0	0	-	-	0	-
	Tidal	2	1	0	1	0	0	-	-	0	-
	Depth	3	2	0	1	0	0	-	-	0	-
	Bedslope	1	1	0	0	0	0	-	-	0	-
	Orbvmean	3	2	0	0	0	1	-	-	0	-
	SSTanom	2	0	1	0	0	1	-	-	0	-
	SSTwint	5	2	1	0	0	2	-	-	0	-
	SSTgrad	1	0	0	0	0	1	-	-	0	-
	SSTanamp	3	1	1	0	0	1	-	-	0	-
	Radmean	3	1	1	0	0	1	-	-	0	-

Table 3.8. Index of genetic diversity (h: Haplotype diversity) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLM analysis across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

Index	Ind. Var	All phyla	Chordata (14)	Echinodermata (1)	Arthropoda (3)	Annelida (2)	Mollusca (13)	Brachiopoda (0)	Cnidaria (1)	Ochrophyta (1)	Tracheophyta (0)
h (GLM)	Lat	4	1	0	1	0	2	-	0	0	-
	Lon	4	3	1	0	0	0	-	0	0	-
	Tidal	1	1	0	0	0	0	-	0	0	-
	Depth	1	1	0	0	0	0	-	0	0	-
	Bedslope	4	2	1	0	0	1	-	0	0	-
	Orbvmean	2	2	0	0	0	0	-	0	0	-
	SSTanom	4	2	1	0	0	1	-	0	0	-
	SSTwint	3	2	1	0	0	0	-	0	0	-
	SSTgrad	0	0	0	0	0	0	-	0	0	-
	SSTanamp	2	1	1	0	0	0	-	0	0	-
	Radmean	5	2	1	1	0	1	-	0	0	-

Table 3.9. Index of genetic diversity (H_O: Observed heterozygosity) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLM analysis across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

Index	Ind. Var	All phyla	Chordata (10)	Echinodermata (3)	Arthropoda (5)	Annelida (0)	Mollusca (12)	Brachiopoda (1)	Cnidaria (2)	Ochrophyta (0)	Tracheophyta (0)
H_O (GLM)	Lat	4	0	0	2	-	1	0	1	-	-
	Lon	2	0	0	0	-	2	0	0	-	-
	Tidal	3	0	0	0	-	3	0	0	-	-
	Depth	3	1	0	1	-	0	1	0	-	-
	Bedslope	3	0	1	0	-	1	1	0	-	-
	Orbvmean	5	0	0	0	-	3	1	1	-	-
	SSTanom	2	1	0	0	-	1	0	0	-	-
	SSTwint	1	0	0	0	-	1	0	0	-	-
	SSTgrad	5	2	0	0	-	2	0	1	-	-
	SSTanamp	4	2	0	0	-	1	0	1	-	-
	Radmean	2	0	0	0	-	1	0	1	-	-

Table 3.10. Index of genetic diversity (H_E: Expected heterozygosity) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLM analysis across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

Index	Ind. Var	All phyla	Chordata (9)	Echinodermata (4)	Arthropoda (5)	Annelida (0)	Mollusca (12)	Brachiopoda (3)	Cnidaria (2)	Ochrophyta (0)	Tracheophyta (0)
H_E (GLM)	Lat	3	0	0	0	-	2	1	0	-	-
	Lon	5	1	0	1	-	2	1	0	-	-
	Tidal	2	0	0	1	-	1	0	0	-	-
	Depth	3	0	1	1	-	0	1	0	-	-
	Bedslope	1	0	0	0	-	0	1	0	-	-
	Orbvmean	2	0	0	0	-	2	0	0	-	-
	SSTanom	4	1	1	0	-	1	1	0	-	-
	SSTwint	5	1	1	0	-	2	1	0	-	-
	SSTgrad	4	2	1	0	-	1	0	0	-	-
	SSTanamp	4	1	1	0	-	2	0	0	-	-
	Radmean	2	0	1	0	-	1	0	0	-	-

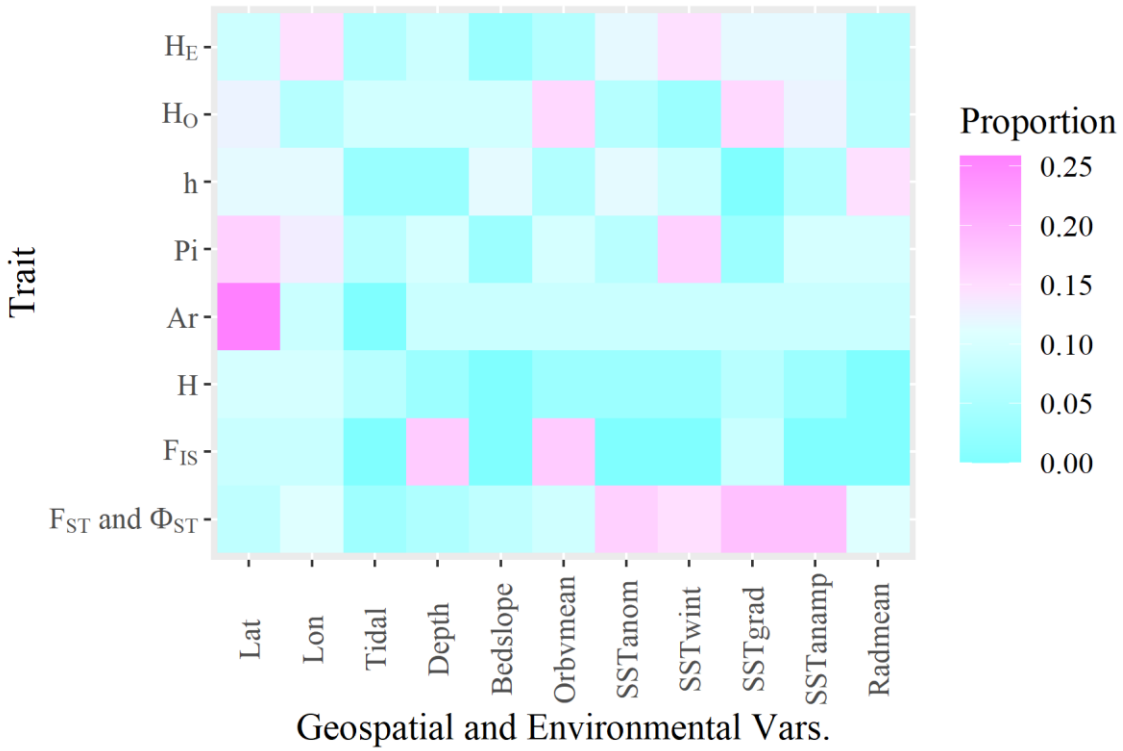


Figure 3.1. Contribution summary heatmap for geospatial/environmental variables in explaining genetic variability across all phyla within New Zealand using the GLM routine.

Genetic indices. H_E: Expected heterozygosity, H_O: Observed heterozygosity, h: Haplotype diversity, Pi: Nucleotide diversity, Ar: Allelic richness, H: Number of haplotypes, F_{IS}: Inbreeding coefficient, F_{ST}-Φ_{ST}: Fixation index. Geospatial variables. Lat: Latitude, Lon: Longitude. Environmental variables. Tidal: Tidal current, Depth: Depth, Bedslope: Seabed curvature, Orbvmean: Mean orbital velocity, SSTanom: Summertime sea surface temperature anomaly, SSTwint: Wintertime sea surface temperature, SSTgrad: Spatial gradient annual mean sea surface temperature, SSTanamp: Annual amplitude of sea surface temperature, Radmean: Annual mean solar radiation.

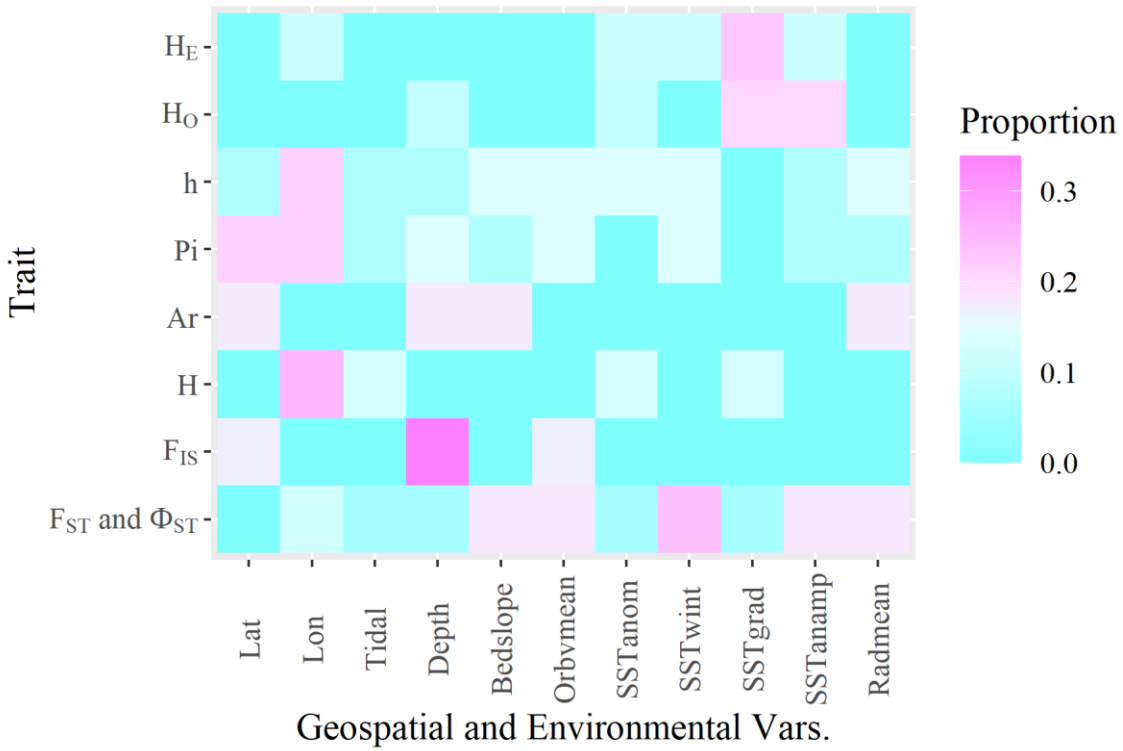


Figure 3.2. Contribution summary heatmap for geospatial/environmental variables in explaining genetic variability across the phylum Chordata within New Zealand using the GLM routine.

Genetic indices. H_E: Expected heterozygosity, H_O: Observed heterozygosity, h: Haplotype diversity, Pi: Nucleotide diversity, Ar: Allelic richness, H: Number of haplotypes, F_{IS}: Inbreeding coefficient, F_{ST}-Φ_{ST}: Fixation index. Geospatial variables. Lat: Latitude, Lon: Longitude. Environmental variables. Tidal: Tidal current, Depth: Depth, Bedslope: Seabed curvature, Orbvmean: Mean orbital velocity, SSTanom: Summertime sea surface temperature anomaly, SSTwint: Wintertime sea surface temperature, SSTgrad: Spatial gradient annual mean sea surface temperature, SSTanamp: Annual amplitude of sea surface temperature, Radmean: Annual mean solar radiation.

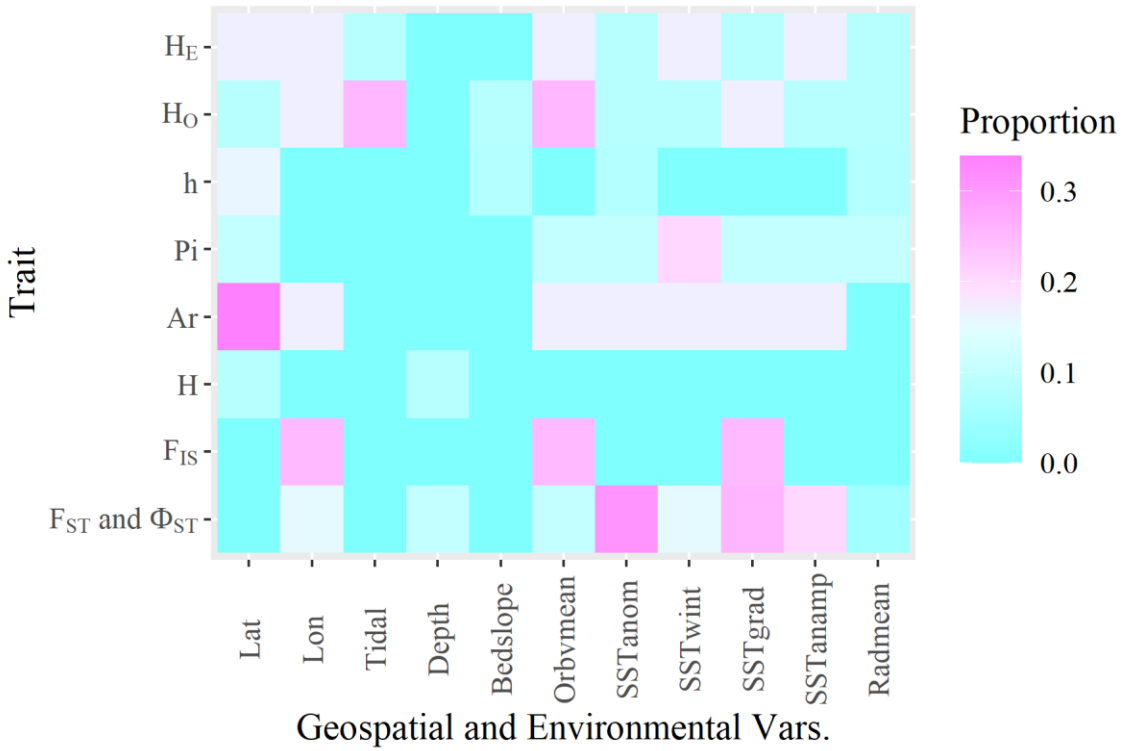


Figure 3.3. Contribution summary heatmap for geospatial/environmental variables in explaining genetic variability across the phylum Mollusca within New Zealand using the GLM routine.

Genetic indices. H_E: Expected heterozygosity, H_O: Observed heterozygosity, h: Haplotype diversity, Pi: Nucleotide diversity, Ar: Allelic richness, H: Number of haplotypes, F_{IS}: Inbreeding coefficient, F_{ST}-Φ_{ST}: Fixation index. Geospatial variables. Lat: Latitude, Lon: Longitude. Environmental variables. Tidal: Tidal current, Depth: Depth, Bedslope: Seabed curvature, Orbvmean: Mean orbital velocity, SSTanom: Summertime sea surface temperature anomaly, SSTwint: Wintertime sea surface temperature, SSTgrad: Spatial gradient annual mean sea surface temperature, SSTanamp: Annual amplitude of sea surface temperature, Radmean: Annual mean solar radiation.

Discussion

Trying to find statistical certainty in ecology is and will continue to be challenging. Here, the seascape genetics null hypothesis of no association between geospatial/environmental variation and genetic variation has been tested by executing two related analytical approaches (GLM and GLZ analyses) to try to determine if one or more environmental variables has a consistent multi-taxon and across phylum affect on variation in genetic diversity indices of New Zealand's coastal biota.

Two-coupled measures of genetic differentiation (F_{ST} and Φ_{ST}) and eight indices of genetic diversity (F_{IS}, Number of haplotypes, Allelic richness, Haplotype richness, Nucleotide diversity, Haplotype diversity, H_O, H_E) were used to generate new understanding about how environmental variation might be influencing genetic variation in New Zealand's marine biota. In general terms, at the species level (across the 84 studies including 55 taxa) no strong patterns were observed. The same conclusion was reached when data were analysed at the level of phylum: Chordata (across 26 of 84 studies including 16 species) and Mollusca (across 30 of 84 studies including 18 species). These two phyla were tested in depth because they presented enough studies to allow for phylum-specific interrogation of patterns, whereas all other phyla had too few studies for this approach.

Overall, many significant associations were found at the phylum level between some genetic metrics and environmental variables. These results therefore provide support for rejecting the null hypothesis of no association between genetic variation and geospatial/environmental variation. Across the genetic indices tested here, I can highlight measures of genetic differentiation (i.e., F_{ST}-Φ_{ST}) and diversity indices (i.e., Allelic richness) as the genetic variables that were most often linked to environmental variation. Four environmental variables were important in explaining genetic distance among populations (i.e., SSTanamp, SSTgrad, SSTanom and SSTwint), whilst one geospatial variable was important in explaining genetic diversity within populations (i.e., Latitude).

At the individual phylum level some significant associations were found between certain genetic metrics and geospatial/environmental variables. It is worth mentioning, across the genetic indices tested for Chordata, the case of measures of genetic

differentiation (i.e., F_{ST}-Φ_{ST}) which were mainly explained by only one sea surface temperature variable (i.e., SSTwint). Also, for the Chordata, the diversity metrics were best explained by four predictors (Latitude, Longitude, Depth and Spatial gradient annual mean sea surface). In contrast, studies from the phylum Mollusca exhibited greater proportions of significant responses for genetic differentiation measures (i.e., F_{ST}-Φ_{ST}) that were mainly explained by one only sea surface temperature (i.e., SSTanom), whereas the diversity metrics were better explained by three predictors (Latitude, Longitude and Mean orbital velocity).

In summary, based on the frequencies of occurrences from the GLM analyses, sea surface temperature variables were the most common predictors of variation for measures of population genetic differentiation, whereas latitude was the most frequently included variable for the genetic diversity variability models of all examined studies, followed by longitude.

According to Wei *et al.* (2013b) it is reasonable to expect that F_{ST}-Φ_{ST} indices are often significant (but other metrics not) for sea surface temperature values, whereas most other metrics (i.e., diversity indices) are significant for latitude, longitude or other environmental variables such as depth (see Supplementary Material 17). Based on my results, I suggest that differentiation among populations (note that increased genetic distance is generally inversely related to gene flow) may be of relevance and is tightly connected with variation in environmental anomalies. For example, my F_{ST}-Φ_{ST} (distance metric) results for the Chordata and the Mollusca were best explained by SSTwint and SSTanom, respectively. My findings are similar to those reported for greenlipped mussels by Wei *et al.* (2013b), suggesting that localised environmental conditions (SSTamon, SSTwint, and SSTgrad) may be of great importance in affecting gene flow amongst population. In addition, in my study, the genetic diversity metrics (which have not been studied before) showed a tendency to be better explained by geospatial variables or variables of terrain.

When contrasting the patterns of the predictor's occurrence among the geospatial variables by phylum, the Chordata showed higher proportions of studies for which longitude was statistically significant, whereas for the Mollusca there was greater

representation for latitude. The relative importance or contribution of latitude to significant models has also been documented in previous seascape genetics studies (e.g., Constable, 2014; Hannan, 2014; Silva & Gardner, 2016) and it seems to be based on the principle that this variable is a surrogate for many important environmental variables that may directly or indirectly contribute to organismal life-history trait variation (Trip *et al.*, 2014), which will in turn be modelling the genetic distances amongst populations. However, the identification of longitude as a factor which explains genetic index variability has not a clear explanation yet. Conversely, some authors have suggested that longitude is an important predictor factor for gene flow because this variable may be playing a deterministic role for genetic connectivity surrogates. Thus, it has been suggested that the west to east direction of the ocean currents in New Zealand may reflect variation of connectivity at large spatial scales in open water systems (Silva & Gardner, 2016). Whilst previous seascape genetics work on New Zealand marine taxa has focussed on F_{ST}-Φ_{ST} indices (Westfall, 2010; Wei *et al.*, 2013b; Constable, 2014; Hannan, 2014; Silva & Gardner, 2016; Zeng *et al.*, 2020), I have compiled data of additional forms of genetic variation that add new insight into the relationships of genetic index variability with geospatial predictors different from the conclusions about environmental variation and genetic connectivity made by other authors. On the basis of the foregoing, the fact that previous research on seascape genetics has not tested genetic diversity metrics makes this multi-species assessment a new step within the field. But the underlying reason why geospatial variables explain variation in the genetic diversity indices is unknown and cannot, for the present, be addressed. Nevertheless, it is important to note that most seascape genetics approaches have found more powerful explanations for genetic distance variation based on environmental predictors rather than on geospatial drivers (i.e., environmental variables > geospatial variables e.g., Wei *et al.*, 2013b; Silva & Gardner, 2016).

The Subantarctic Front represents a major natural boundary or barrier within New Zealand waters (Chiswell & Sutton, 2015). Elsewhere, coastal currents have been shown to play a key role in explaining seascape genetics (Galindo *et al.*, 2010; Selkoe *et al.*, 2010). In this study, I report a low-to-absent influence of currents (i.e., mean orbital velocity), as shown in the case of the ‘all phyla’ testing and secondarily for the Chordata and the Mollusca. The significant responses which involved mean orbital velocity as a predictor of

genetic variation showed in general lower proportions, being greater for the diversity metrics than for distance estimates. It is worth highlighting the case study of *Onchidella* (Mollusca: Gastropoda), for which Cumming *et al.* (2014) tested for transoceanic dispersal in direct-developing species, as these slugs are ecologically associated with buoyant bull kelp (Durvillaea), and are known to raft. They found that the studied populations belong to a single taxon, whose distribution is widespread. The analyses revealed high mtDNA differentiation between mainland New Zealand populations, apparently consistent with low levels of dispersal. However, subantarctic samples show many shared mitochondrial haplotypes and high levels of AFLP similarity between distant populations. These findings suggest that currents, specifically the Subantarctic Front, may be preventing gene flow between regions within New Zealand waters, but promoting at the same time a clockwise direction for dispersal to distant areas. As reported for numerous other studies, these findings suggest that diversity and connectivity are closely related and that both depend on gene flow.

Based on the main outcomes for the distance metrics (i.e., F_{ST}-Φ_{ST}), I decided to conduct a brief approach to better understand the positive or negative contribution of the sea surface temperature predictors on the effects for gene flow across populations. I can report that in terms of a simple linear correlation, there was no statistical significance fit for any of the four variables which were important in explaining genetic distance among populations (i.e., SSTanamp, SSTgrad, SSTanom and SSTwint) against F_{ST}-Φ_{ST}. The apparent absence of relationships here act to preventing our understanding of the real contribution of these factors on the genetic distance and connectivity among populations.

As my seascape genetics analysis cannot give a direct cause and effect linkage between the independent and dependent variables tested here, one possible explanation for relationships would be a specific model of selection based on thermal stress, specifically, acting directly on heat shock proteins (hereafter, HSP) (Buckley *et al.*, 2001; Dutton & Hofmann, 2009; Wei *et al.*, 2013b). The microsatellite markers themselves are thought to be selectively neutral and are therefore unlikely to respond directly to selection, but microsatellite variation may indirectly reflect patterns of selection via hitch-hiking (Chambers & MacAvoy, 2000; England *et al.*, 2003; Holderegger & Wagner, 2006) and as

a consequence may provide a window into a temperature-driven selection response. For example, Smith (1988) suggested more than three decades ago that there may be warm-adapted and cold-adapted populations of the greenshell mussel (*Perna canaliculus*) in New Zealand based on his allozyme-based work. The general pattern of genetic variation observed in this study is consistent with the statement made by Smith (1988), suggesting that adapted populations (tightly linked with sea surface temperature) may not be a feature exclusive to molluscs, but might also be reflected in other coastal taxa. If so, further examination of this matter will advance our understanding of the patterns and processes of genetic structuring in coastal marine biota. In addition, this statement should be also tested in the future for chordates, as was suggested by Constable (2014) for the case of the common flatfish, *R. plebeia.* Constable (2014) suggested that local adaptation to thermal tolerance in the expression of HSPs may be occurring for the New Zealand species, just as earlier research showed for the European flounder (*Platyichthys flesus*) (Hemmer-Hansen *et al.*, 2007).

Wolff *et al.* (2014) suggested that mitonuclear allelic interactions are evolutionarily significant modulators of the expression of specific life-history phenotypes, across and between species (intra- and interpopulational). In this sense, I observed for the Chordata and the Mollusca different significant responses (driven factors) between nuclear (i.e., Allelic richness) and mitochondrial markers (i.e., Nucleotide diversity). This result suggests that different groups of genes would be causing the appearance of specific phenotypes within different species and across multiple phyla. Perhaps the observed patterns across the studies included in this analysis are nothing more than environmental variables exerting modeling on different groups of genes (i.e., mitonuclear interactions) to define species-specific life-history traits and consequently adaptations within the environment (Morales *et al.*, 2015). However, this hypothesis must be widely tested and debated.

While the second routine here tested (GLZ) may be more sensitive than the GLM due to stepwise removal of variables with the smallest partial correlation with the dependent variable, the truth is that the stepwise approach is not as reliable as the GLM performed here. The evidence on this matter has clearly demonstrated that the probability of making a Type I error increases while using stepwise procedures (Derksen & Keselman,

1992). An analysis conducted by Mundry & Nunn (2009) was intended to evaluate this particular issue for different numbers of entry variables, using for this purpose random data for which the null hypothesis held true. The results of this integrated approximation suggest that the inflated Type I error is largely due to multiple testing (i.e., the stepwise procedure itself). That is, the GLZ routine involves multiple different tests before arriving at a final 'best' solution, whereas the GLM involves only one such test (this is the problem of multiple testing giving rise to elevated Type I errors rates in the GLZ). Mundry & Nunn (2009) concluded that the only valid option for employing stepwise procedures would be to adjust error levels for the number of variables evaluated at each step. Nevertheless, statistical inference based on the sequential addition or removal of the expected predictors has been largely used by researchers across ecology and evolution-related fields. Based on this, it is proposed that the methods in future seascape genetics research are standardised, to achieve successful and meaningful comparisons across studies.

Identifying cause and effects is very hard, if not impossible, for much of the testing within the discipline of seascape genetics. We can start to think about causation (e.g., heat shock proteins), but a full mechanistic understanding of the process or processes that underpin seascape genetics is beyond us right now. Sometimes the independent variables versus the dependent variables relationships seems to point to barriers to gene flow, other times the relationships point to the promotion of gene flow, which in turn will influence diversity metrics such as Allelic richness, Haplotype richness, Haplotype diversity, and Nucleotide diversity.

Clearly, the results presented here make it impossible to predict patterns of environmental-genetic variation at the taxon or phylum level for New Zealand's marine biota. Further research on this discipline might perhaps predict specific pathways of genetic differentiation and diversity across multi-species by the development of new tools and models. To be able to use such metrics (the genetic attributes of populations) for management purposes without the timely and consuming delays of additional new genetic investigations would be beneficial, but is not presently a realistic approach. Meanwhile, to monitor populations of key species is still relevant. Such monitoring will, it is hoped, increase the number of studies for under-represented phyla (e.g., the Echinodermata,

Arthropoda, Annelida, Brachiopoda, Platyhelminthes, Cnidaria, Ochrophyta and Tracheophyta) and indeed, other contributions to the biological knowledge of species and the conformation of their populations in New Zealand waters. The incorporation of additional environmental variables into model-testing may also help explain sources of significant genetic variation, as corroborated recently in a cold-water sponge (e.g., silicate) and cold-water corals (e.g., calcite) by Zeng *et al.* (2020).

The availability of an enhanced or broader spatially explicit set of environmental variables is only as good as the environmental monitoring and mapping that occurs. Greater emphasis is being placed on developing a new understanding of fine scale spatial (and temporal) variation in environmental variations across New Zealand's marine realm, but until such high resolution data are available the seascape genetics approach will not advance far beyond its present state. It is important in the future to be able to achieve predictions on a global scale, as some authors have recently suggested. An example is the outcome reported by Silva *et al.* (2021), who suggested that benthic temperature is associated with the differentiation between species/populations across rock lobsters, *Jasus* spp., along a narrow latitudinal band in the Southern hemisphere. This is undoubtedly the limitation that the discipline of seascape genetics is facing - to be able in the future to inform management and conservation - but more efforts and collaborative networks are necessary to achieve such macroecological goals.

CHAPTER 4 General discussion

In this thesis, I describe how I tested genetic variation as a function of geospatial patterns along a latitudinal cline (~13 degrees of latitude) using the Core-Periphery Hypothesis (CPH). In the following chapter, I evaluated the contribution of 11 geospatial and environmental variables to explaining genetic variation by conducting two seascape genetics analyses, testing the null hypothesis of no relationship. The Core-Periphery and the seascape genetics hypotheses were tested using a large data set of genetic index variation for a wide range of marine organisms, with a focus on studies addressing the genetic structure of populations (all conducted in New Zealand waters), across 10 phyla. This macroecological research uses a predictive framework to identify patterns (CPH models) of genetic diversity and to identify variables that help explain such patterns (seascape genetics). Additionally, in a broader sense, it helps to further our understanding of the interactions between the genome and the marine environment. Broad scale testing procedures for these topics have been suggested within the field of marine evolutionary biology for several years (Chave & Bascompte, 2013), but until recently there has been limited opportunity to test such hypotheses, because too few genetic studies had been conducted and in particular because the necessary site-specific environmental data sets did not exist (Snelder *et al.*, 2005). Nevertheless, hypotheses or proposed associations (direct or indirect) are constantly referred to when it is necessary to describe the relationship between genetic variation and environmental variation. The work presented in this thesis is a first approach intending to identify factors which may be producing geospatial patterns of multi-taxon genetic variation and raising predictors of marine genetic variation by gathering and combining huge amounts of data from pre-existing research projects in New Zealand.

The genetic expectations behind the Core-Periphery Hypothesis (CPH) taking as a proxy New Zealand marine biota

The Abundant-Centre and Core-Periphery hypotheses are clearly not a general rule which can be broadly applied to predict that species are most abundant at their range centre, as demographic and genetics approaches have already demonstrated in terrestrial systems (Sagarin & Gaines, 2002b; Eckert *et al.*, 2008; Santini *et al.*, 2019). This suggests that

range limits may be highly mobile and dynamic. The overall results, achieved using a large dataset of New Zealand's coastal marine biota to test this hypothesis, showed that by assessing connectivity and diversity indices under the expectations behind the theory, the edge populations exhibit greater genetic variability and genetic differentiation than those in the centre.

The requirements here, taken into consideration to perform the analysis of the CPH, showed that the selection of studies based on a meta-analysis resulted in an effective methodological approach ideal for replication. My analyses emphasise that the selection of studies to test the hypothesis must be focussed on those that sample populations across the entire distribution range for the interest species, taking at least 6 sampling points in total (two per macrozone or region in the case of New Zealand). This criterion for inclusion in the testing framework provided realistic and robust results, as far as it is possible to judge. I concluded that ~25% of all the possible arrangements which were tested using ten genetic indices, fitted at least one of the four distribution models (for 29 out of 55 species being represented). Moreover, I also reported higher level support for the Ramped North and Ramped South models which accounted for ~20% of the total cases. However, overall, most test results (194 of 249 tests = 78%) showed that none of the four models being tested could be fitted to the data. As the first multi-taxon test of the CPH of this kind I conclude that the CPH as originally defined, and when applied to indices of genetic variation, does not apply to New Zealand's marine biota. Further investigation is required to determine why, of all the significant models that were identified, the Ramped North and the Ramped South models (i.e., positive and negative relationships with latitude) were the most frequently observed (these two models were approximately three-times more frequent that the Normal or Abundant Edge models). One plausible explanation could be that warm water species should tend to distribute towards the North, whereas cold water species towards the South (as suggested in Chapter 2). But this hypothesis remains to be tested.

Here, I suggest that depending on the marker used to perform the genetic analysis in each of the studies included in the CPH testing procedure, it would be possible to determine modern or historical issues closely related to genetic connectivity or diversity indices. This enables me to propose that nuclear markers will agree with current abundance records or

effective population sizes, whereas mitochondrial markers tend to exhibit past tendencies (historical information) (de Oliveira *et al.*, 2009; Túnez *et al.*, 2013). Because this relationship between abundance and genetic data was executed through a descriptive analysis with available abundance records for only some species, it is necessary to make some recommendations mainly in relation to updating baseline information regarding abundance and distribution of marine invertebrates in New Zealand waters. In this way, it would be possible to design a statistical assessment to determine significance levels in relation to this idea and test even more species to establish if this pattern can be extended across phyla. Nevertheless, more studies that integrate genetic data across multiple life-history traits must be conducted to reach more exact and widely applicable conclusions in marine evolutionary ecology and conservation.

In summary, my study reveals low-to-absent support for the CPH. The normal distribution fitted only five species: the New Zealand sea lion (*Phocarctos hookeri*), the New Zealand sea urchin (*Evechinus chloroticus*), the New Zealand dredge oyster (*Ostrea chilensis*), the spotted whelk (*Cominella maculosa*) and the Southern bull kelp (*Durvillaea antarctica*). These species are members of five families (Otariidae, Echinometridae, Ostreidae, Buccinidae and Durvillaeaceae) that share some peculiarities. These include the absence of a pelagic larval phase for the case of large vertebrates and macroalgae and a (very) short pelagic larval duration for the case of the two molluscs species (the oyster is a brooding species and the whelk lays eggs from which live individuals emerge). The case of the New Zealand sea urchin separates itself from the others (average PLD), and is the reason why it requires more examination. Overall, while the CPH may be an oversimplifying concept, these results focus attention on the population genetics and range dynamics, which may be more variable at the distribution periphery in New Zealand, but the reliability of this pattern, and its effect on creating limits, is still unknown and requires further testing and improvements.

A seascape genetics approach: trying to find marine environmental drivers to explain genetic variation in New Zealand coastal species

The seascape genetics analyses, based on population genetics studies that fulfilled the basic requirement for entry (≥ 6 sampling locations per study), were intended to reveal the relationship between genetic differentiation/diversity variability and environmental influences on a macroecological scale (i.e., across New Zealand's coastal realm). It is worth highlighting the value of using an integrative biological, ecological and physical oceanographic approach to investigate indirect patterns of population genetic variation such as the seascape genetics analysis here conducted. By using standardised analytical methods across species, it was possible to determine both the common and species-specific factors that may be responsible for shaping genetic variation of coastal marine populations in New Zealand. Here, I tested a large multi-phylum dataset of genetic indices for 55 species to try to identify common geospatial and/or environmental variables that may explain variation in the indices.

This is the first comprehensive seascape genetics analysis covering a wide range of marine organisms (from mammals to macroalgae) and that attempts to identify common patterns of taxon genetic structuring with key marine environmental drivers. My results show that both geospatial and environmental variables are significantly associated with the population genetic structure across phyla, and at the phylum scale (in the case of the Chordata and Mollusca).

At phylum level, the geospatial variable latitude was an important factor in explaining the genetic diversity variation among populations, followed by longitude. A mixture of environmental variables contributes to the observed genetic index variability, but wintertime sea surface temperature and mean orbital velocity contributed the most. On the other hand, variation in genetic differentiation (i.e., F_{ST} and Φ_{ST}) amongst populations was best explained by four sea surface temperature variables (summertime, wintertime, spatial gradient annual mean and annual amplitude). These predictor factors were the most significant amongst the tested studies using the GLM procedure, following the recommendations made by Mundry & Nunn (2009), who found that stepwise approaches (e.g., the GLZ) significantly raise the probability of erroneously rejecting the null

hypothesis (i.e., the GLZ raises the Type I error). Whilst I employed both the GLM and the GLZ approaches in my testing, I have followed the recommendations of Mundry & Nunn (2009) and have focussed on the results from the GLM, but for the sake of completeness I note that the patterns of results for the GLZ followed those of the GLM (as expected), but that there were more statistically significant results reported from the GLZ than from the GLM testing (again, as expected).

The discipline of population genetics attempts to explain the genetic diversity in contemporary populations and the changes in allele and genotype frequencies in space, and occasionally in time (Hedrick, 2005). In this sense, genetic surveys and monitoring efforts are crucial to aid our understanding (and testing of hypotheses) in relation to effective population sizes, population bottlenecks, dispersal patterns and diversity (Veale, 2007; de Oliveira *et al.*, 2009; Túnez *et al.*, 2013). The focus of my study was in evaluating patterns of genetic variation as a function of the environmental variation at the studies' sampling sites, using for this purpose genetic attributes from published studies. Although this proposes obtaining macro-scale results, it also confronts us with other challenges that must be solved. For example, it was found to be necessary to address and cover the gaps in relation to interpolated data, imputation of missing values after extraction of environmental variables, and mean orbital velocity zero values that are far from the reality (three aspects which concern the EEZ-MEC scheme - NIWA). By addressing these limitations, coupled with access to individual-specific allelic frequencies, it will be possible to firstly, obtain accurate and reliable environmental data, and secondly, perform other routines which strengthen or complement the current analysis (e.g., BEST and DistLM routines) (Silva & Gardner, 2016; Zeng *et al.*, 2020). For example, the use of different analytical methods can increase confidence that the environmental variables are truly associated with the genetic variation (Silva & Gardner, 2015). Consequently, access to complete data sets will promote and facilitate the analysis and comparison of species-specific data (e.g., the Ira Moana (genes of the sea) initiative – https://sites.massey.ac.nz/iramoana/).

A major limitation that I faced for the seascape genetics meta-analysis was the low number of sampling sites per study (on average) across New Zealand coastal marine taxa. I set the threshold at six different sites per study as the minimum number to perform the

routines (in this case I still had to split the 11 geospatial/environmental variables when executing the GLM/GLZ analyses). Every meta-analysis is only as good as the data that the researcher can pull from the individual studies. Consequently, limitations faced by the original researchers then become limitations faced by the meta-analysis researcher too. Beyond this, the seascape genetics approach is only as good as the environmental variables entered into the model. In this sense, the case of New Zealand stands out because of the quality and spatial coverage of site-specific data (Snelder *et al.*, 2005). Nevertheless, in the EEZ-MEC there is a limited number of environmental variables. This means that inevitably I will miss the signal from other important variables that may be influencing gene flow or genetic diversity in some way, if those variables are not in the data set. Even taking into consideration the above-mentioned issues, undoubtedly, this meta-analysis is the best approach we have available right now. In this sense, the discipline of seascape genetics is still really in its infancy, and as such, still faces a number of limitations that need to be acknowledged. Nonetheless, despite the limited number of samples across the meta-analysis and a lack of comparative information about the relationships between environmental factors and genetic differentiation, the results generated here contribute to a more comprehensive understanding of environmental and genetic interactions at both phylum and taxon scales.

Even though the outputs of the seascape genetics analyses can be considered to be informative, various different ways of partitioning the data still remain to be tested. Rearrangements according to the genetic marker (e.g., mitochondrial DNA versus nuclear DNA) used in each study may result in different outcomes or conclusions for each species or phylum. Future different arrangements of the analysis presented here may address this methodological issue, which could highlight a suitable marker among a series of genetic tools specific for seascape genetics analyses. Nevertheless, here I intended to describe general patterns in the multi-species meta-analysis, rather than to identify individual study-based explanations of genetic diversity or variation.

Whether there is a direct selection-based association between genetic and environmental variation, or whether the environmental variables simply act as barriers to gene flow or promote gene flow (a neutral rather than selection-based explanation) remains

to be determined. However, because population genetics is almost exclusively based on neutral markers (it is usual that outlier loci or loci that appear to be under selection are removed from data sets before publication) it is reasonable to assume that the vast majority of the genetic data sets included in my meta-analysis are neutral (or very nearly so) and therefore are not expected to be subject to direct selection activity. Thus, the new discipline of seascape genetics (Selkoe *et al.,* 2016) is a valuable tool for increasing our understanding of genetic structure at broad taxonomic levels as demonstrated by this meta-analysis (phylum-specific and across all phyla) and which is also most widely applicable at the species level.

A greater understanding of the predictions and scopes of the selective pressures influencing genetic variability at a population, community and ecosystem level is still imperative to being able to predict adaptation to novel environmental conditions. With the recent development and application of single nucleotide polymorphism (SNP)-based and also structural variant-based markers, and the ability to identify very large datasets of both neutral and selected loci across the genome, it is now possible to believe that future studies may be able to test for cause and effect when looking to understand how genetic diversity is influenced by and responds to environmental variation.

Supplementary Material

Supplementary Material 1. Chordata – an overview of genetic studies conducted in New Zealand.

Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Chordata - Mammalia	Open Coast	*Cephalorhynchus hectori*	Hector's Dolphin	Endemic NZ	NZ wide (3)	mtDNA (CR)	Three regions (North Island, West and East of South Island)	IBD. Low rate of female dispersal	(Pichler et al., 1998)
Chordata - Mammalia	Open Coast	*Cephalorhynchus hectori*	Hector's Dolphin	Endemic NZ	NZ wide (3)	mtDNA (CR)	Four regions (North Island, West, East and South of South Island)	IBD	(Pichler & Baker, 2000)
Chordata - Mammalia †	Open Coast	*Cephalorhynchus hectori*	Hector's Dolphin	Endemic NZ	NZ wide (13) (4)	mtDNA (CR) SSRs	Four regions (North Island, West, East and South of South Island)	IBD	(Pichler, 2002)
Chordata - Mammalia †	Open Coast	*Cephalorhynchus hectori*	Hector's Dolphin	Endemic NZ	South Island (13) (13)	mtDNA (CR) SSRs	Three regions (West, East and South of South Island)	IBD	(Hamner et al., 2012b)
Chordata - Mammalia	Open Coast	*Cephalorhynchus hectori maui*	Maui's Dolphin	Endemic NZ	North Island (1) (1) Two years monitoring	mtDNA (CR) SSRs	No structure	Only one haplotype	(Hamner et al., 2012a)
Chordata - Mammalia	Open Coast	*Cephalorhynchus hectori maui*	Maui's Dolphin	Endemic NZ	North Island (1) (1)	mtDNA (CR) SSRs	No structure	Only one haplotype	(Hamner et al., 2012b)
Chordata - Mammalia	Open Coast	*Cephalorhynchus hectori maui*	Maui's Dolphin	Endemic NZ	North Island (1) (1)	mtDNA (CR) SSRs	No structure	Only one haplotype	(Baker et al., 2013)
Chordata - Mammalia	Open Coast	*Cephalorhynchus hectori maui*	Maui's Dolphin	Endemic NZ	North Island (1)	SSRs	No structure	Assignable differentiation of microsatellite genotypes	(Baker et al., 2016)

Chordata - Mammalia	Open Coast	*Arctocephalus forsteri*	New Zealand fur seal	Endemic AUS-NZ	South Island (4)	mtDNA (CytB)	No structure	No population structure among colonies around the NZ South Island, and a deep phylogenetic divergence between fur seals from Western Australia	(Lento et al., 1994)
Chordata - Mammalia †	Open Coast	*Arctocephalus forsteri*	New Zealand fur seal	Endemic AUS-NZ	South Island (7)	SSRs	South Island West Coast colonies slightly different from those of the East Coast	IBD	(Robertson & Gemmell, 2005)
Chordata - Mammalia †	Open Coast	*Arctocephalus forsteri*	New Zealand fur seal	Endemic AUS-NZ	South Island and Subantarctic islands (13)	SSRs	No structure	There was no significant pattern of isolation by distance among the NZ colonies	(Dussex et al., 2016)
Chordata - Mammalia †	Open Coast	*Arctocephalus forsteri*	New Zealand fur seal	Endemic AUS-NZ	South Island (8)	SNPs	South Island West Coast colonies different from those of the South Eastern Coast	IBD	(Stovall, 2016)
Chordata - Mammalia	Open Coast	*Phocarctos hookeri*	New Zealand sea lion	Endemic NZ	Auckland Islands (1)	SSRs	No structure	High levels of genetic variation (due to males)	(Acevedo-Whitehouse et al., 2009)

Chordata - Mammalia	Open Coast	*Phocarctos hookeri*	New Zealand sea lion	Endemic NZ	Auckland Islands (4)	SSRs	No structure	Population bottleneck and nonrandom mating	(Osborne et al., 2016)
Chordata - Mammalia †	Open Coast	*Phocarctos hookeri*	New Zealand sea lion	Endemic NZ	South Island and Subantarctic Islands (7) (7) (7)	mtDNA (D-loop) mtDNA (CytB) SSRs	No structure	Female philopatry in combination with distant colonies	(Collins et al., 2016, 2017)
Chordata - Aves	Open Coast	*Megadyptes antipodes*	Yellow-eyed penguin	Endemic NZ	South Island and Subantarctic Islands (4)	Allozymes	Subantarctic different from South Island	Very low immigration to the South Island from Subantartic islands	(Triggs & Darby, 1989)
Chordata – Aves †	Open Coast	*Megadyptes antipodes*	Yellow-eyed penguin	Endemic NZ	South Island and Subantarctic Islands (6) (7)	mtDNA (HVI) SSRs	Subantarctic different from South Island	Subtropical convergence	(Boessenkool et al., 2009)
Chordata - Aves	Open Coast	*Megadyptes antipodes*	Yellow-eyed penguin	Endemic NZ	South Island and Subantarctic Islands (3)	SSRs	Subantarctic different from South Island	Very low immigration to the South Island from Subantartic islands	(Boessenkool et al., 2010)
Chordata – Actinopterygii †	Estuarine	*Grahamina nigripenne*	Estuarine triplefin	Endemic NZ	NZ wide (11)	mtDNA (CR)	North-south differentiation at ~44°S	IBD	(Hickey et al., 2009)
Chordata – Actinopterygii †	Estuarine	*Rhombosolea plebeia*	New Zealand sand flounder	Endemic NZ	NZ wide (19)	SSRs	North-south differentiation	IBD	(Constable, 2014)
Chordata - Actinopterygii	Rocky intertidal	*Bellapiscis lesleyae*	Mottled twister	Endemic NZ	NZ wide (5)	mtDNA (CR)	Differentiation between	IBD	(Hickey et al., 2009)

							populations		
Chordata – Actinopterygii †	Rocky intertidal	*Bellapiscis medius*	Twister	Endemic NZ	NZ wide (8)	mtDNA (CR)	Differentiation between populations	IBD	(Hickey et al., 2009)
Chordata – Actinopterygii †	Rocky intertidal	*Parapercis colias*	New Zealand blue cod	Endemic NZ	NZ wide and Chatham Islands (14) (7)	mtDNA (CR) SSRs	Chatham Islands different from mainland NZ sites	IBD. Adult migration highly infrequent	(Smith, 2012; Gebbie, 2014)
Chordata – Actinopterygii †	Rocky intertidal	*Hippocampus abdominalis*	The pot-bellied seahorse	Endemic AUS-NZ	NZ wide (6) (6)	mtDNA (CR) SSRs	Adelaide different from the rest of Australia and Australia different from NZ	Colonization event to NZ	(Ashe & Wilson, 2020)
Chordata – Actinopterygii †	Rocky subtidal	*Forsterygion varium*	Variable triplefin	Endemic NZ	NZ wide (8)	mtDNA (CR)	No structure	High gene flow within NZ but Three Kings Islands isolated	(Hickey et al., 2009)
Chordata – Actinopterygii †	Rocky subtidal	*Grahamina gymnota*	Tasmanian robust triplefin	Endemic NZ	NZ wide (9)	mtDNA (CR)	No structure	High gene flow within NZ but Three Kings Islands isolated	(Hickey et al., 2009)
Chordata - Actinopterygii	Rocky subtidal	*Ruanoho whero*	Spectacled triplefin	Endemic NZ	NZ wide (5)	mtDNA (CR)	No structure	High gene flow within NZ but Three Kings Islands isolated	(Hickey et al., 2009)
Chordata – Actinopterygii †	Rocky subtidal	*Forsterygion lapillum*	Common triplefin	Endemic NZ	NZ wide (12)	mtDNA (CR)	Differentiation between populations	IBD	(Hickey et al., 2009)
Chordata – Actinopterygii †	Rocky subtidal	*Forsterygion lapillum*	Common triplefin	Endemic NZ	NZ wide (7)	SSRs	North-south differentiation	IBD, stepping-stone dispersal and	(Rabone et al., 2015)

								contiguous gene flow	
Chordata – Actinopterygii †	Rocky subtidal	*Grahamina capito*	Spotted robust triplefin	Endemic NZ	NZ wide, Chatham Islands and Subantarctic islands (15)	mtDNA (CR)	Three distinct geographically restricted lineages	IBD	(Hickey et al., 2009)
Chordata - Actinopterygii	Rocky subtidal	*Forsterygion capito*	Spotted robust triplefin	Endemic NZ	NZ wide (5)	SSRs	North-south differentiation	IBD, stepping-stone dispersal and contiguous gene flow	(Rabone et al., 2015)
Chordata – Actinopterygii †	Soft substrate subtidal	*Rhombosolea leporina*	New Zealand yellowbelly flounder	Endemic NZ	NZ wide (10)	SSRs	East to west and north to south differentiation	There is gene flow among most locations with some level of structure	(Constable, 2014)
Chordata – Elasmobranchii †	Estuarine	*Mustelus lenticulatus*	The spotted estuary smooth-hound or rig	Endemic NZ	NZ wide (9) (7)	Allozymes mtDNA (RFLPs)	No structure	Very low levels of genetic variation	(Hendry, 2004)
Chordata – Ascidiacea †	Rocky intertidal	*Cnemidocarpa nisiotis*	New Zealand sea tunicate	Endemic NZ	NZ wide (18)	mtDNA (COI)	No structure	A single panmictic population	(del Mundo, 2009)

NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV; NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

Supplementary Material 2. Echinodermata – an overview of genetic studies conducted in New Zealand.

Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Echinodermata - Echinoidea †	Fjord	*Evechinus chloroticus*	New Zealand sea urchin	Endemic NZ	Fiordland plus 2 South Island and 2 North Island sites (34)	SSRs	Differentiation within and among fjords	Dispersal is restricted between the islands, at the scale of >1000km, and in and out of the fjords at the scale of <100km	(Perrin & Roy, 2000; Perrin, 2002)
Echinodermata - Echinoidea †	Rocky subtidal	*Evechinus chloroticus*	New Zealand sea urchin	Endemic NZ	NZ wide (6)	Allozymes	No structure (except Doubtful Sound distinct from all other populations)	Longlived planktotrophic larva and oceanographic features	(Mladenov et al., 1997)
Echinodermata - Echinoidea †	Rocky subtidal	*Evechinus chloroticus*	New Zealand sea urchin	Endemic NZ	NZ wide (11) (11)	mtDNA (COI) SSRs	North-south differentiation at ~42°S	Significant IBD correlation for both the mitochondrial and microsatellite datasets	(Nagel et al., 2015)
Echinodermata - Ophiuroidea †	Fjord	*Astrobrachion constrictum*	Snake star	Endemic NZ-AUS	Fiorland (6) (7)	Allozymes mtDNA (SSCPs-COI)	No structure	There is some evidence for subdivision at the level of haplotype frequency	(Steel, 1999; Steel et al., 2015)
Echinodermata - Ophiuroidea	Rocky intertidal	*Ophiomyxa brevirima*	Brittle star	Endemic NZ	South Island and Stewart Islands (4)	Allozymes	Differentiation between populations	Brooding habit and potential	(Garrett, 1994; Garrett et al., 1997)

								dependence on availability of rafting materials	
Echinodermata - Asteroidea †	Fjord	*Coscinasterias muricata*	Eleven armed sea star	Endemic NZ-AUS	Fiordland plus 2 South Island and 1 North Island sites (17)	mtDNA (SSCPs-D-loop)	Differentiation between North and South Islands. Differentiation among fjords and between fjords and open ocean	IBD amongst northern fjords, restricted gene flow between southern fjords	(Perrin, 2002; Perrin et al., 2004)
Echinodermata - Asteroidea †	Fjord	*Coscinasterias muricata*	Eleven armed sea star	Endemic NZ-AUS	Fiordland plus 2 South Island and 1 North Island sites (16)	Allozymes	Differentiation among fjords	Large scale IBD, no IBD on small, relatively isolated hydrography of each fjord	(Sköld et al., 2003)
Echinodermata - Asteroidea	Rocky intertidal	*Patiriella regularis*	Cushion star	Endemic NZ-AUS	NZ wide (22)	mtDNA (CR)	North-south differentiation at ~42°S	IBD across NZ as a whole and across northern NZ	(Waters & Roy, 2004)
Echinodermata - Asteroidea †	Rocky intertidal	*Patiriella regularis*	Cushion star	Endemic NZ-AUS	NZ wide (22)	mtDNA (CR)	North-south differentiation at ~42°S	Upwelling zones south of Cook Strait	(Ayers & Waters, 2005)
Echinodermata - Asteroidea	Rocky subtidal	Coscinasterias muricata	Eleven armed sea star	Endemic NZ-AUS	NZ wide (4) (2) (1)	mtDNA (COI) mtDNA (COI+CR) nDNA (ITS2)	No structure	One haplotype related to southern Australia	(Waters & Roy, 2003)

NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV; NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ

wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

Supplementary Material 3. Arthropoda – an overview of genetic studies conducted in New Zealand.

Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Arthropoda - Malacostraca	Estuarine	*Paracorophium excavatum*	New Zealand amphipod	Endemic NZ	NZ east coast (4)	Allozymes	North-south differentiation	Ocean currents and overland dispersal routes. The species could potentially comprise a complex of at least three subspecies	(Schnabel, 1998; Schnabel et al., 2000)
Arthropoda - Malacostraca †	Estuarine	*Paracorophium excavatum*	New Zealand amphipod	Endemic NZ	NZ wide (21)	Allozymes	North-south differentiation at ~39°S	IBD. Effects of sea-level and landmass changes that have occurred throughout the Plio-Pleistocene	(Stevens & Hogg, 2004)
Arthropoda - Malacostraca †	Estuarine	*Paracorophium lucasi*	New Zealand amphipod	Endemic NZ	Central NZ (10)	Allozymes	East-west and north-south differentiation	Ocean currents and overland dispersal routes. Low levels of gene flow and fixed differences between the two groups of the "eastern" and "western"	(Schnabel, 1998; Schnabel et al., 2000)

Arthropoda - Malacostraca †	Estuarine	*Paracorophium lucasi*	New Zealand amphipod	Endemic NZ	NZ wide (18)	Allozymes	East-west and north-south differentiation	IBD. Effects of sea-level and landmass changes that have occurred throughout the Plio-Pleistocene	(Stevens & Hogg, 2004)
Arthropoda - Malacostraca †	Rocky intertidal	*Pinnotheres novaezelandiae*	Pea crab	Endemic NZ	NZ North Island (7)	Allozymes	Differentiation within and among locations	The variance among host-associated populations within a locality was greater than that between localities	(Stevens, 1990)
Arthropoda - Malacostraca †	Rocky intertidal	*Pinnotheres atrinicola*	Pea crab	Endemic NZ	NZ North Island (7)	Allozymes	Latitudinal cline differentiation	Life-history attributes and current movements	(Stevens, 1991)
Arthropoda - Malacostraca †	Rocky intertidal	*Hemigrapsus sexdentatus*	Common rock crab	Endemic NZ	NZ wide (11)	mtDNA (RFLPs-COI)	Shallow north-south differentiation	IBD. Gene flow via larval dispersal can be considered the primary force	(Hinnendael, 2008)
Arthropoda - Malacostraca	Rocky intertidal	*Isocladus armatus*	Isopod	Endemic NZ	NZ wide (3)	SNPs	North-south differentiation	It appears that NZ isopods are capable of substantial levels of dispersal along coastlines, despite their sedentary reproductive	(Wells & Dale, 2018)

								histories	
Arthropoda - Malacostraca †	Rocky intertidal	*Isocladus armatus*	Isopod	Endemic NZ	NZ wide (9)	SNPs	Strong north-south differentiation	IBD	(Pearman et al., 2020)
Arthropoda - Malacostraca	Rocky subtidal	*Jasus edwardsii*	Red rock lobster	Endemic NZ-AUS	NZ wide (3)	Allozymes	No structure	No genetic differences between the three samples of *J. edwardsii* collected from widely separated areas of NZ	(Smith et al., 1980)
Arthropoda - Malacostraca	Rocky subtidal	*Jasus edwardsii*	Red rock lobster	Endemic NZ-AUS	NZ east coast (2)	mtDNA (RFLPs)	No structure	Lobster haplotypes were shown not to be geographically partitioned. Shared haplotypes	(Ovenden et al., 1992)
Arthropoda - Malacostraca	Rocky subtidal	*Jasus verreauxi*	Packhorse rock lobster	Endemic NZ-AUS	NZ northeast coast (2)	mtDNA (RFLPs)	No structure	Three of the six haplotypes found only in NZ constitute a branch on the phenetic tree	(Brasher et al., 1992)
Arthropoda - Malacostraca	Soft substrate subtidal	*Metanephrops challengeri*	New Zealand scampi	Endemic NZ	NZ wide and Auckland Islands (8)	Allozymes	Auckland Islands different from mainland NZ	Female scampi mature at a greater size around the Auckland Islans than mainland NZ	(Smith, 1999)

								populations	
Arthropoda - Malacostraca †	Soft substrate subtidal	*Metanephrops challengeri*	New Zealand scampi	Endemic NZ	NZ wide and Auckland Islands (6)	mtDNA (COI)	Auckland Islands different from mainland NZ	IBD. The expansion may have been caused by the glacial-interglacial cycles of the Pleistocene, which similarly impacted lobster populations worldwide	(Verry et al., 2020)
Arthropoda - Malacostraca †	Soft substrate subtidal	*Munida gracilis*	Squat lobster	Endemic NZ-AUS	NZ wide (7)	mtDNA (COI)	No structure	Single population with high levels of mixing not impeded by geographic distance or current patterns	(Bors et al., 2012)

NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV; NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

Supplementary Material 4. Annelida – an overview of genetic studies conducted in New Zealand.

Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Annelida - Polychaeta †	Soft substrate subtidal	*Hyalinoecia longibranchiata*	Quill-worm	Endemic NZ	NZ wide (8) (8)	mtDNA (16S) mtDNA (COI)	Northeast to southwest differentiation	The large-scale genetic differences can be explained partly by geographic distribution and partly by currents	(Bors et al., 2012)

NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV; NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

Supplementary Material 5. Mollusca – an overview of genetic studies conducted in New Zealand.

Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Mollusca - Bivalvia †	Estuarine	*Austrovenus stutchburyi*	New Zealand little neck clam	Endemic NZ	NZ wide (10)	Allozymes	No structure	The differences within the sites were possibly the result of the population consisting of a mixture of varying annual cohorts	(Lidgard, 2001)
Mollusca - Bivalvia †	Estuarine	*Austrovenus stutchburyi*	New Zealand little neck clam	Endemic NZ	NZ wide, Chatham Islands and Subantarctic Islands (29) (8)	mtDNA (COI) SSRs	North-south differentiation	IBD. Large and persistent eddies located to the north and south of East Cape, historic dispersal barriers that no longer exist and regional environmental differences	(Ross, 2011; Ross et al., 2012)
Mollusca - Bivalvia	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	NZ wide (12)	Allozymes	North-south and spring-autumn differentiation	Currents movement of East Auckland, Tasman and Southland currents. Partial assortative mating	(Smith, 1988)
Mollusca - Bivalvia	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	South Island (3)	Allozymes	Significant differences in allele frequencies were observed	The differences between the three populations may also reflect	(Sin et al., 1990)

							between the littoral and sublittoral populations and between the northern and southern littoral populations	geographic differences between the north and south, and between intertidal and sublittoral habitats	
Mollusca - Bivalvia	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	Wellington Harbour and Beatrix Bay (2)	Allozymes	No structure	Four of the seven polymorphic loci exhibited significant allelic heterogeneity between the wild and cultured populations	(Gardner et al., 1996a)
Mollusca - Bivalvia †	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	NZ wide (10)	Allozymes	North-south differentiation	IBD	(Gardner et al., 1996b)
Mollusca - Bivalvia	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	Wellington Harbour (3)	Allozymes	Differentiation between geographically proximate (~10 km) locations	The extent of gene flow among the populations will depend to some extent upon their spatial separation as well as upon local hydrographic conditions	(Gardner & Kathiravetpillai, 1997)
Mollusca - Bivalvia	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	NZ wide (35)	Allozymes	No structure	Effect of gene flow is higher relative to that of genetic drift	(Apte & Gardner, 2001)
Mollusca - Bivalvia	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	NZ wide (2)	mtDNA (NADH4-SSCPs)	North-south differentiation at ~42°S	Gene flow was obstructed by the opening of Cook Strait ~15,000-16,000 years ago	(Apte & Gardner, 2002)

								and the subsequent establishment of present-day surface water circulation patterns in Greater Cook Strait	
Mollusca - Bivalvia †	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	NZ wide (35) (21) (19)	Allozymes mtDNA (NADH4, NADH2, COI-SSCPs & RFLPs) RAPDs	North-south differentiation at ~42°S	Genetic diversity was high across the seven allozyme loci. Hydrographic, climatic, and topographic characteristics may have led to adaptations and population genetic differentiation	(Apte et al., 2003)
Mollusca - Bivalvia †	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	NZ wide (19)	RAPDs	North-south differentiation at ~42°S	Temperature would seem to be the most likely cause for the genetic discontinuity	(Star et al., 2003)
Mollusca - Bivalvia †	Rocky intertidal	*Perna canaliculus*	Greenshell mussel	Endemic NZ	NZ wide (14)	SSRs	No structure	Sea surface temperature (SST) explains far more genetic variation (among populations and individuals) than any other variable	(Wei et al., 2013a,b)
Mollusca - Bivalvia	Rocky intertidal	*Mytilus galloprovincialis*	Mediterranean mussel	Endemic NZ	Wellington Harbour (3)	Allozymes	Differentiation between geographically	The extent of gene flow among the populations	(Gardner & Kathiravetpillai, 1997)

							proximate (~10 km) locations	will depend to some extent upon their spatial separation as well as upon local hydrographic conditions	
Mollusca - Bivalvia †	Rocky intertidal	*Mytilus galloprovincialis*	Mediterranean mussel	Endemic NZ	NZ wide, Chatham Islands and Subantartic Islands (13)	SSRs	Northwest population different from the other panmictic population	IBD. Dispersal direction indicated by genetic homogeneity between sample sites and oceanographic currents supports larval transport through lower layers in the water column	(Westfall, 2010; Gardner et al., 2016)
Mollusca - Bivalvia	Rocky subtidal	*Tiostrea lutaria*	New Zealand dredge oyster	Endemic NZ	Faveaux Strait (1)	Allozymes	No structure	They concluded that *T. chilensis* and *T. lutaria* are geographic populations of one species based on our information and that on grounds of priority this species should be *T. chilensis*	(Buroker et al., 1983)
Mollusca - Bivalvia †	Rocky subtidal	*Ostrea chilensis*	New Zealand dredge oyster	Endemic NZ	NZ wide and Chatham Islands (14) (15) (17)	mtDNA (COI) mtDNA (CytB) SSRs	North-south differentiation at ~42°S, three main regional genetic clusters (Hauraki Gulf, northern,	The Cook Strait itself is not a barrier to gene flow, and provides further evidence that an	(Guo et al.)

							southern) were observed	oceanographic barrier of some sort, perhaps coastal upwelling, exists at ~42 °S on both coasts of the South Island that prevents extensive gene flow between the regions	
Mollusca - Bivalvia	Rocky subtidal/Deep thermal vents	*Bathymodiolus spp.*	Hydrothermal vent mussels	Endemic NZ	North Island (4) (4)	Allozymes mtDNA (COI)	Three lineages distinct from other mussel species	Transient gene frequency shifts on local scales. Taylor columns and anti-cyclonic gyres are known to develop around seamounts	(Smith et al., 2004)
Mollusca - Bivalvia †	Soft substrate subtidal	*Paphies subtriangulata*	Tuatua	Endemic NZ	NZ wide and Chatham Islands (13)	Allozymes	North-south differentiation; Chatham Islands differentiation	Influenced by the East Auckland, East Cape, Tasman, D´Urville and Southland currents movements	(Smith et al., 1989)
Mollusca - Bivalvia †	Soft substrate subtidal	*Paphies subtriangulata*	Tuatua	Endemic NZ	NZ wide (10)	SSRs	Mainland New Zealand different from Chatham Island	IBD. Pelagic larval duration	(Hannan, 2014)
Mollusca - Bivalvia †	Soft substrate subtidal	*Paphies australis*	Pipi	Endemic NZ	NZ wide (13)	SSRs	Northern, South Eastern and South Western	IBD. Oceanographic processes act in unison with habitat availability to drive genetic	(Hannan, 2014)

								population structure and connectivity in this species	
Mollusca - Bivalvia †	Soft substrate subtidal	*Pecten novaezelandiae*	New Zealand scallop	Endemic NZ	NZ wide and Chatham Islands (14)	SSRs	North Island separately from central NZ. Fiordland, Stewart Islands and Chatham Islands were plotted separately	IBD. Different combinations of environmental variables were correlated with the genetic variation	(Silva & Gardner, 2016)
Mollusca - Gastropoda	Rocky intertidal	*Haliotis iris*	New Zealand black-foot abalone	Endemic NZ	NZ wide and Chatham Islands (3)	Allozymes	No structure	Pgm electromorph frequencies appeared to be useful in distinguishing between Chatham and mainland samples	(Frusin, 1982)
Mollusca - Gastropoda	Rocky intertidal	*Haliotis iris*	New Zealand black-foot abalone	Endemic NZ	North Island, Stewart Islands and Chatham Islands (4) (4)	mtDNA (ATP8-ATP6) SSRs	Differentiation between sites	Chatham Islands are beyond the limits of typical larval dispersal. Only weak genetic structure was observed in *H. iris* with one mtDNA marker	(Smith & McVeagh, 2006)
Mollusca - Gastropoda †	Rocky intertidal	*Haliotis iris*	New Zealand black-foot abalone	Endemic NZ	NZ wide (28)	mtDNA (COI, ATP8-ATP6)	Four phylogeographic breaks, western Cook Strait region, southeast coast of the South Island, East Cape in the North Island and	IBD	(Will & Gemmell, 2008; Will et al., 2011)

							Chatham Islands		
Mollusca - Gastropoda †	Rocky intertidal	*Haliotis iris*	New Zealand black-foot abalone	Endemic NZ	NZ wide and Stewart Islands (27)	SSRs	North-south differentiation, Chatham Islands different from mainland NZ	IBD. Differetiation of North Island samples corresponds with major components of NZ oceanography, Cook Strait and the East Cape	(Will & Gemmell, 2008; Will et al., 2015)
Mollusca - Gastropoda	Rocky intertidal	*Nerita atramentosa*	Black nerite	Endemic NZ-AUS	North Island and Three Kings Islands (10)	mtDNA (COI)	No structure	The Tasman Sea is not a significant phylogeographical barrier for eastern *N. atramentosa*, likely reflecting their extended planktotrophic larval phase	(Waters et al., 2005)
Mollusca - Gastropoda †	Rocky intertidal	*Cellana radians*	Radiate limpet	Endemic NZ	NZ wide (28)	mtDNA (CytB)	North-south differentiation at ~42°S	Contemporary North and South Island lineages diverged from their respective most recent common ancestor approximately 200,000 to 300,000 years before present	(Goldstien, 2005; Goldstien et al., 2006)
Mollusca - Gastropoda †	Rocky intertidal	*Cellana ornata*	Ornate limpet	Endemic NZ	NZ wide (19)	mtDNA (CytB)	North-south differentiation at ~42°S	Contemporary North and South Island lineages diverged from their respective	(Goldstien, 2005; Goldstien et al., 2006)

								most recent common ancestor approximately 200,000 to 300,000 years before present	
Mollusca - Gastropoda †	Rocky intertidal	*Cellana flava*	Golden limpet	Endemic NZ	NZ wide (6)	mtDNA (CytB)	North-south differentiation	Contemporary North and South Island lineages diverged from their respective most recent common ancestor approximately 200,000 to 300,000 years before present	(Goldstien, 2005; Goldstien et al., 2006)
Mollusca - Gastropoda †	Rocky intertidal	*Scutellastra kermadecensis*	Kermadec Island giant limpet	Endemic NZ	Kermadec Islands (11)	RAPDs	Differentiation between sites	Local hydrographic features, contribute to the observed patterns of small-scale population genetic structuring	(Wood & Gardner, 2007)
Mollusca - Gastropoda †	Rocky intertidal	*Siphonaria raoulensis*	Marine snail	Endemic NZ	Kermadec Islands (6)	RAPDs	Differentiation between sites	Local hydrographic features, contribute to the observed patterns of small-scale population genetic structuring	(Wood & Gardner, 2007)
Mollusca - Gastropoda †	Rocky intertidal	*Scutus breviculus*	Duck's bill limpet, Rori	Endemic NZ	NZ wide (17)	mtDNA (COI)	No structure	The substantial trans-Tasman differentiation observed for Scutus therefore	(Waters et al., 2007)

								suggests that their planktotrophic larval phase is insufficiently long to facilitate dispersal across the Tasman Sea	
Mollusca - Gastropoda †	Rocky intertidal	*Austrolittorina antipodum*	Blue-banded Periwinkle, Ngaeti	Endemic NZ	NZ wide (8)	mtDNA (COI)	No structure	The substantial trans-Tasman differentiation observed for *A. antipodum* therefore suggests that their planktotrophic larval phase is insufficiently long to facilitate dispersal across the Tasman Sea. Blue *Austrolittorina* species [*A. unifasciata* (Australia), *A. antipodum* (NZ), and *A. fernandezensis* (Juan Fernández)] represent a paraphyletic assemblage	(Waters et al., 2007)
Mollusca - Gastropoda †	Rocky intertidal	*Cominella virgata*	Red-mouthed whelk	Endemic NZ	NZ wide (22)	mtDNA (COI)	North-south differentiation	The South Island population of *C. virgata* are therefore likely to have originated through	(Walton, 2017; Fleming et al., 2018; Walton et al., 2019)

								unintentional human-mediated translocations	
Mollusca - Gastropoda †	Rocky intertidal	*Cominella maculosa*	Spotted whelk	Endemic NZ	NZ wide and Chatham Islands (31)	mtDNA (COI)	Two phylogeographic breaks, between Marokopa and New Plymouth on the west coast and between Castle Point and Cape Palliser on the east coast. Chatham Islands different from mainland NZ	Due to a refugium during the LGM at or to the north of New Plymouth. The east coast break may result from the confluence of two linages	(Walton, 2017; Fleming et al., 2018; Walton et al., 2019)
Mollusca - Gastropoda †	Rocky intertidal	*Cominella maculosa*	Spotted whelk	Endemic NZ	North Island (7)	mtDNA (COI)	Four northern sites were genetically differentiated from each other and from the group of three southern sites	Sporadic long-distance dispersal events occur in this species	(Dohner et al., 2018)
Mollusca - Gastropoda †	Rocky intertidal	*Buccinulum vittatum*	Marine snail	Endemic NZ	NZ wide (6)	nDNA (SNPs)	Northern different from eastern and southern	IBD. The nuclear markers suggest that gene flow between populations from within the range of *B. v. vittatum* (north) and B. colensoi (east) is ongoing	(Gemmell et al., 2018)
Mollusca - Polyplacophora	Rocky intertidal	*Sypharochiton pelliserpentis*	Snakeskin chiton	Endemic NZ-AUS	NZ wide and Chatham Islands (29)	mtDNA (COI) mtDNA (COI-RFLPs)	North-south differentiation, using the RFLPs data, the North	The location of the boundaries shown in *S. pelliserpentis* are	(Veale, 2007; Veale & Lavery, 2011)

					(29)		Island populations show a moderate disjunction between the east and west coasts	exactly the locations of upwelling and associated offshore currents believed to be the cause of population disjunction	

NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV; NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

Supplementary Material 6. Brachiopoda – an overview of genetic studies conducted in New Zealand.

Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Brachiopoda - Rhynchonellata †	Fjord	*Terebratella sanguinea*	Red brachiopod	Endemic NZ	Fiordland and Stewart Island (10) (20)	Allozymes AFLPs	Some differentiation among fjords	IBD	(Ostrow, 2004)
Brachiopoda - Rhynchonellata †	Fjord	*Liothyrella neozelanica*	White brachiopod	Endemic NZ	Fiordland (6)	AFLPs	Differentiation within and among fjords	IBD	(Ostrow, 2004)

NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV; NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

Supplementary Material 7. Platyhelminthes – an overview of genetic studies conducted in New Zealand.

Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Platyhelminthes - Trematoda	Estuarine	*Maritrema novaezealandensis*	Marine trematode parasite	Endemic NZ	Otago Harbour (3)	SSRs	No structure	High gene flow due to the bird definitive hosts (mobility)	(Keeney et al., 2008)

NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV; NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

Supplementary Material 8. Cnidaria – an overview of genetic studies conducted in New Zealand

Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Cnidaria - Anthozoa †	Fjord	*Antipathes fiordensis*	Black coral	Endemic NZ	Fiordland and Stewart Island (28)	Allozymes	Divergence within and among fjords	This pattern is suggestive of non-equilibrium populations and is likely to be related to the geologically young age of these populations	(Miller, 1997)
Cnidaria - Anthozoa	Fjord	*Antipathes fiordensis*	Black coral	Endemic NZ	Doubtful Sound, Fiordland (3)	Allozymes	Divergence within fjord	Decrease in larval dispersal associated with reduce water movement	(Miller, 1998)
Cnidaria - Anthozoa †	Rocky intertidal	*Actinia tenebrosa*	Waratah anemone	Endemic NZ-AUS	NZ wide (23)	SSRs	Differentiation between populations	IBD	(Veale, 2007; Veale & Lavery, 2012)
Cnidaria - Hydrozoa †	Fjord	*Errina novaezealandiae*	Red coral	Endemic NZ	Fiordland (9)	Allozymes	Divergence within and among fjords	The scale of larval dispersal in this species may be relatively small	(Miller et al., 2004)

NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV; NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

Supplementary Material 9. Ochrophyta – an overview of genetic studies conducted in New Zealand.

Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Ochrophyta - Phaeophyceae †	Rocky intertidal	*Durvillaea antarctica*	Southern bull-kelp	Endemic SH New Zealand, Gough Island (Atlantic), Argentina, Chile and Australia (Norfolk Island and Macquarie Island) and Antarctic	NZ wide (34)	mtDNA (COI)	North-south differentiation	Long-distance drifting of buoyant bull-kelp provides a powerful mechanism for colonization, but only a poor one for maintaining gene flow	(Collins et al., 2010)
Ochrophyta - Phaeophyceae †	Rocky subtidal	*Carpophyllum maschalocarpum*	Common flapjack	Endemic NZ	NZ wide (32)	mtDNA (23S-tRNA Lys)	North-south differentiation	These patterns are consistent with fucalean morphological and life-history characteristics that limit dispersal at the gamete/zygote stage, but facilitate occasional long-distance dispersal.	(Buchanan, 2011; Buchanan & Zuccarello, 2012)

NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV;

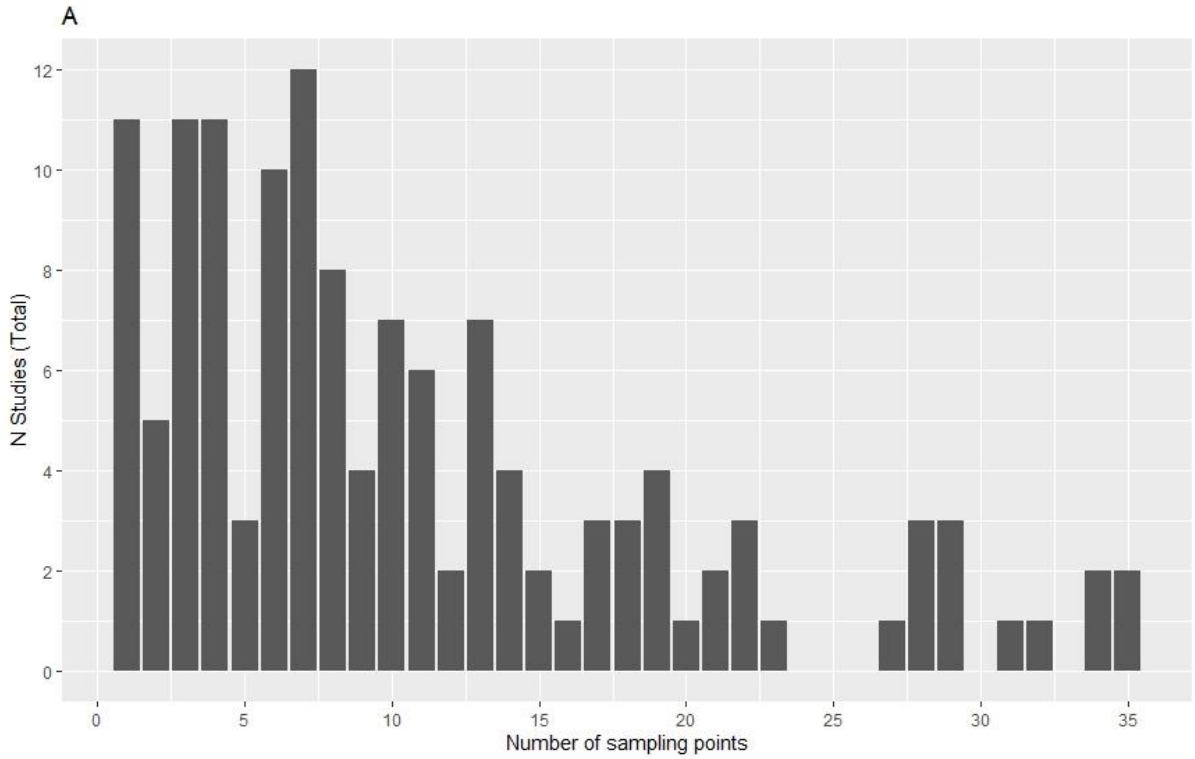
NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

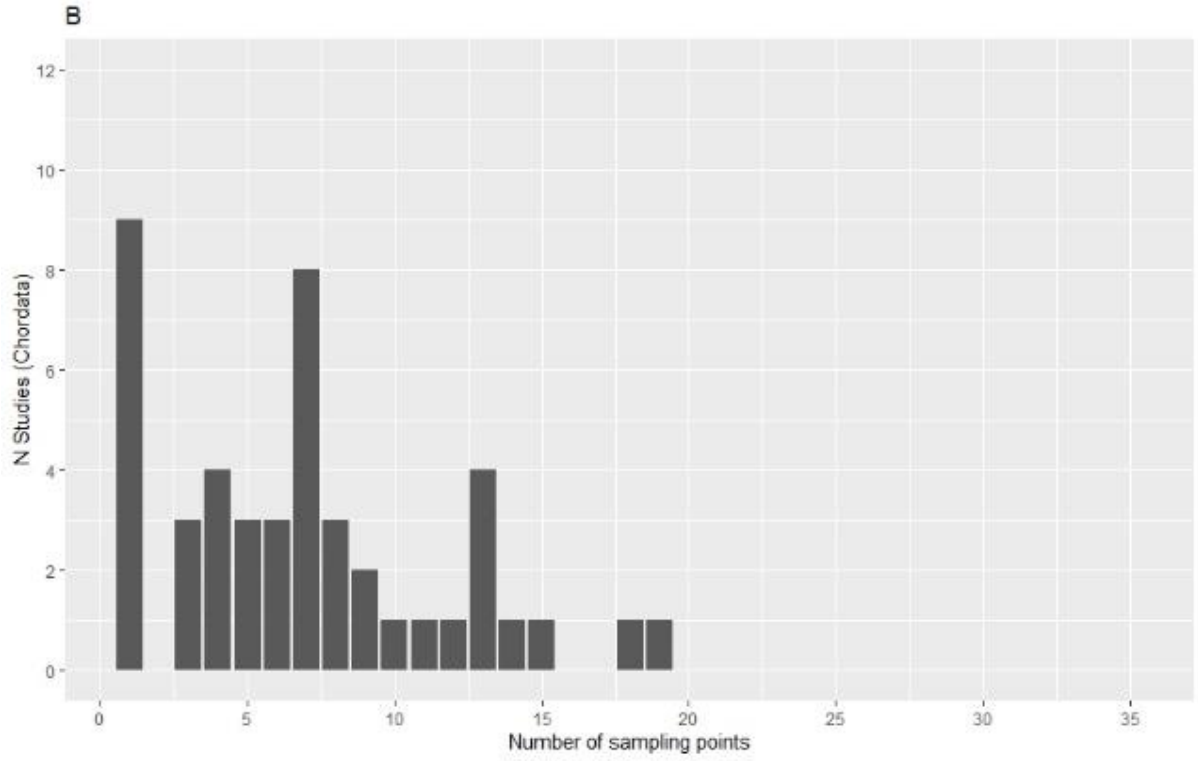
Supplementary Material 10. Tracheophyta – an overview of genetic studies conducted in New Zealand.

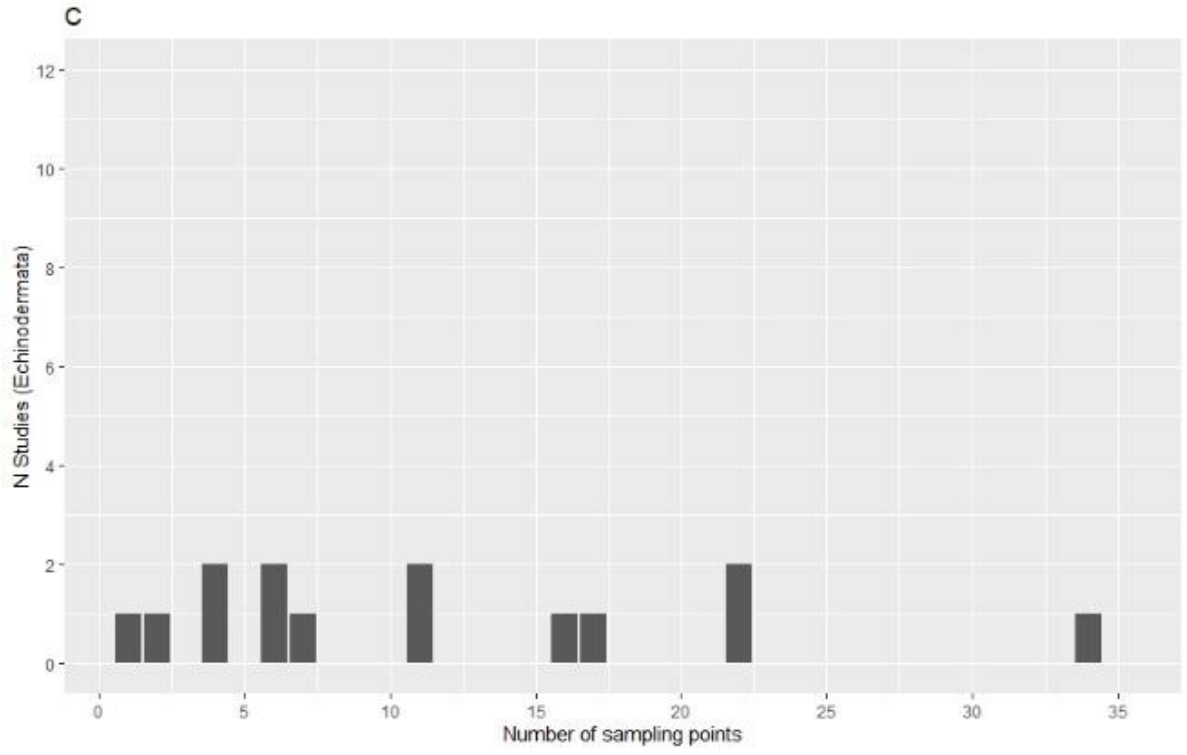
Classification (phylum - class)	Habitat type	Species	Common name	Endemic to NZ / AUS-NZ / SH	N° Sampling points	Marker types	Genetic structure	Factors	Reference
Tracheophyta - Magnoliopsida †	Soft substrate intertidal	*Zostera muelleri*	Seagrass	Endemic NZ-AUS	NZ wide (18)	nDNA (ITS-RAPDs)	East-west (for fine-scale genetic variation within two estuaries: Raglan Harbour and Whangapoua Harbour) & east-west and north-south differentiation (for almost the entire latitudinal gradient of the North and South Islands)	IBD. Gene flow between estuaries is shaped, both directed and in some cases limited, by the coastal currents circulating NZ shores	(Jones, 2004; Jones et al., 2008)

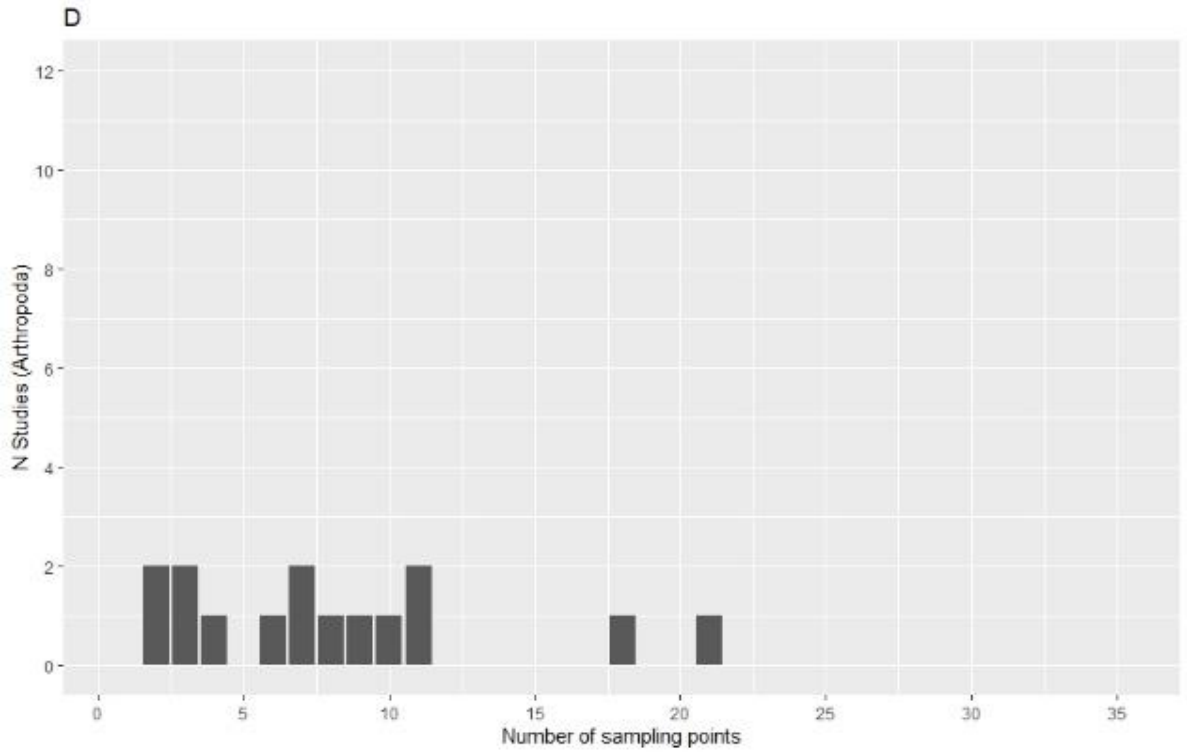
NZ, New Zealand; mtDNA, mitochondrial DNA; nDNA, nuclear DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSCP, single-strand conformation polymorphism; RAPD, random amplification of polymorphic DNA; ITS, internal transcribed spacer; ITS2, internal transcribed spacer 2; CR, control region; COI, cytochrome c oxidase subunit I; D-loop, displacement loop; NADH4, NADH dehydrogenase subunit IV; NADH2, NADH dehydrogenase subunit II; HVI, hyper variable region I; CytB, cytochrome b; 16S, 16S ribosomal RNA; 23S, 23S ribosomal RNA; ATP6, ATPase synthase 6; ATP8, ATP synthase 8; SSR, simple sequence repeat; SNP, single nucleotide polymorphisms; IBD, isolation by distance. Number in parentheses indicates populations sampled (total sampling points across New Zealand territory also including the Chatham Islands and subantartic islands). NZ wide: It refers to the fact that the sampling locations were mainly on the North Island, South Island and Stewart Islands of New Zealand. † The symbol indicates the studies incorporated in the Abundant-Centre hypothesis and Seascapes Genetics analyses.

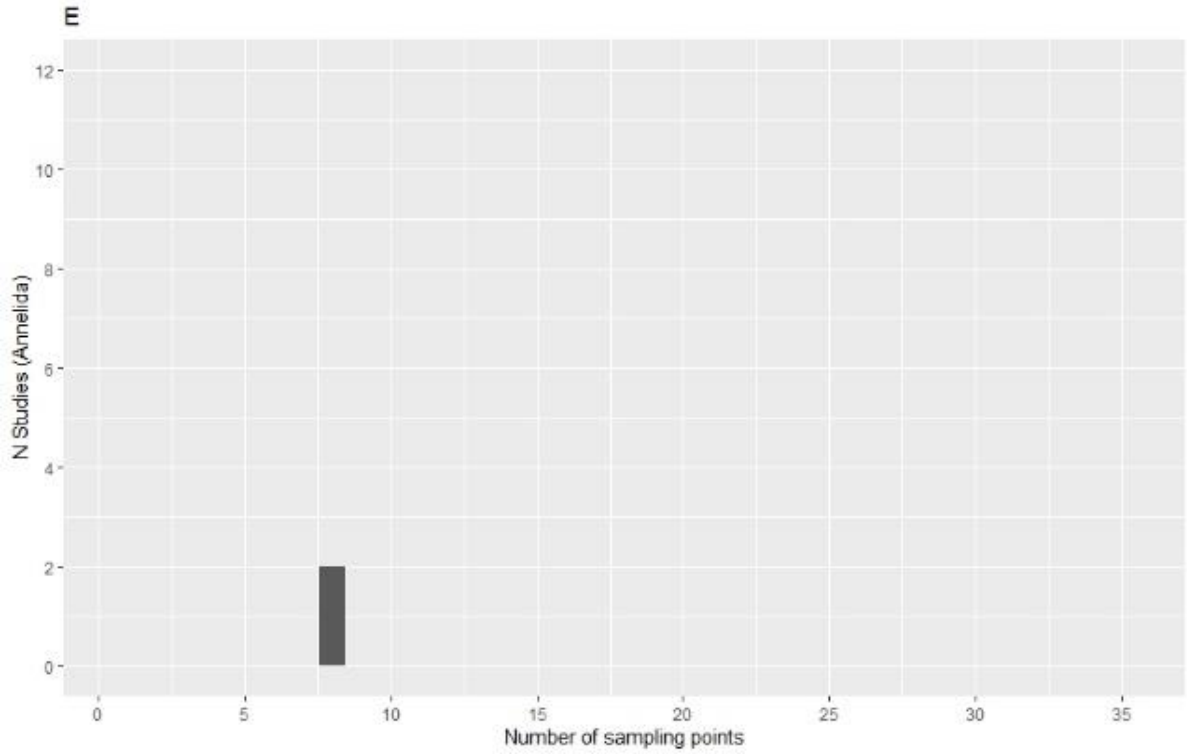
Supplementary Material 11. Summary of studies incorporated in the review of phylogeographic genetic studies of marine vertebrates and invertebrates, macroalgae and plants conducted in New Zealand. Based on the number of sampling points performed in each study. (**A**) Total reviewed studies, and subsequently separated by phylum (**B**) Chordata, (**C**) Echinodermata, (**D**) Arthropoda, (**E**) Annelida, (**F**) Mollusca, (**G**) Brachiopoda, (**H**) Platyhelminthes, (**I**) Cnidaria, (**J**) Ochrophyta and (**K**) Tracheophyta.

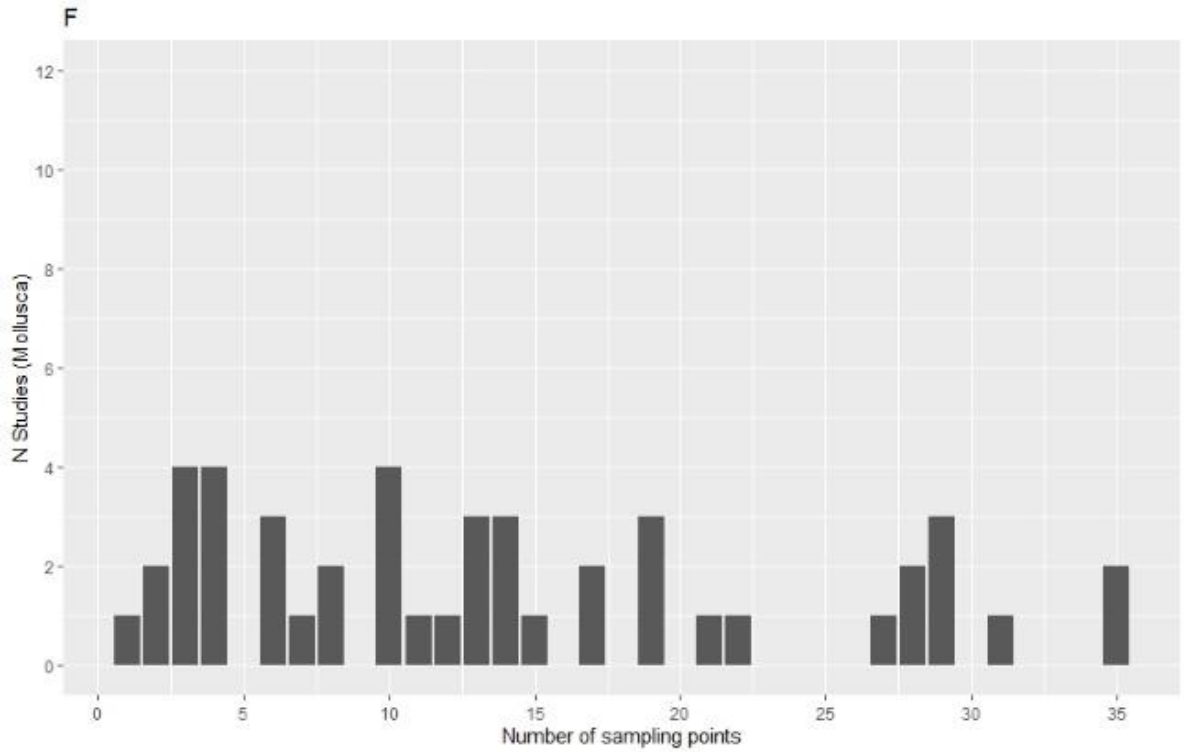


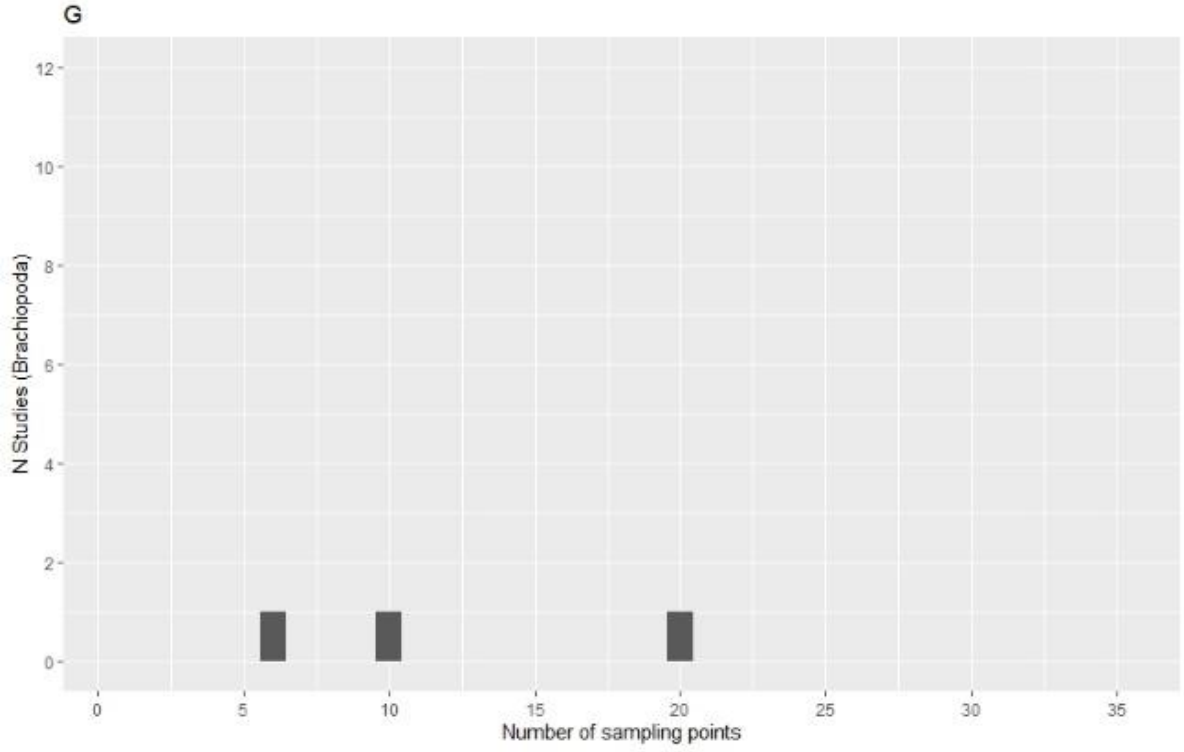

B
N Studies (Chordata)
Number of sampling points
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35

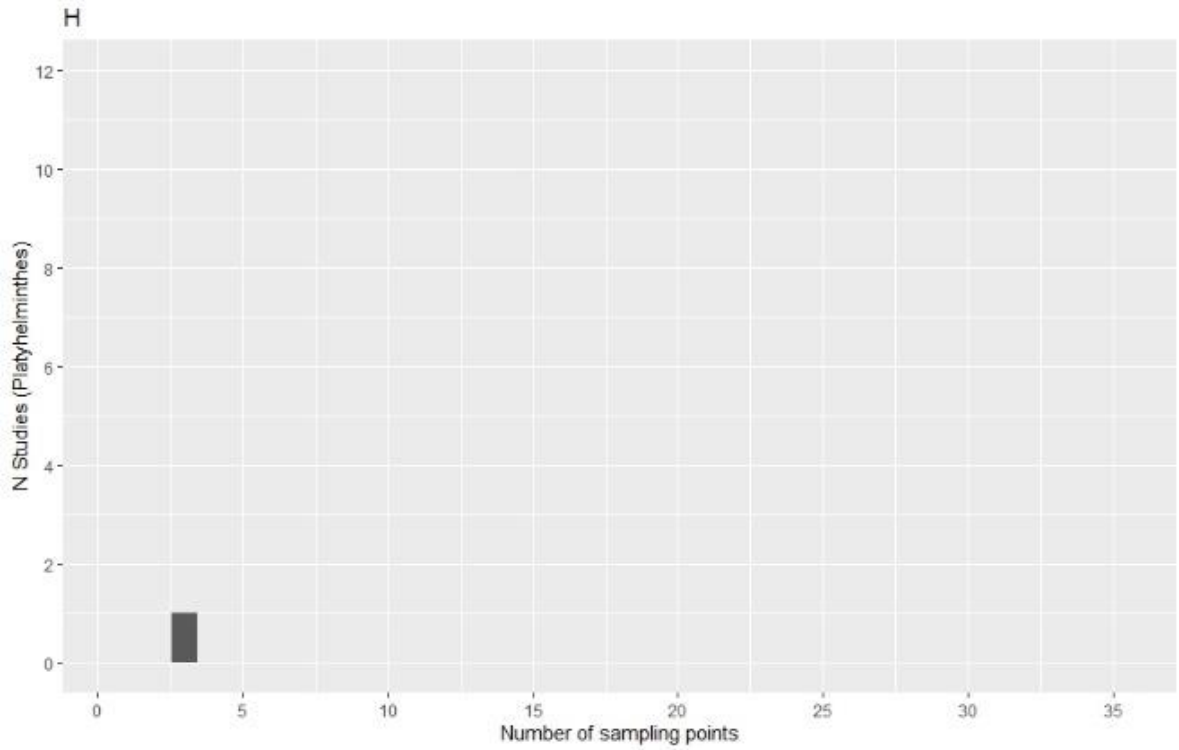

C
N Studies (Echinodermata)
Number of sampling points
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35

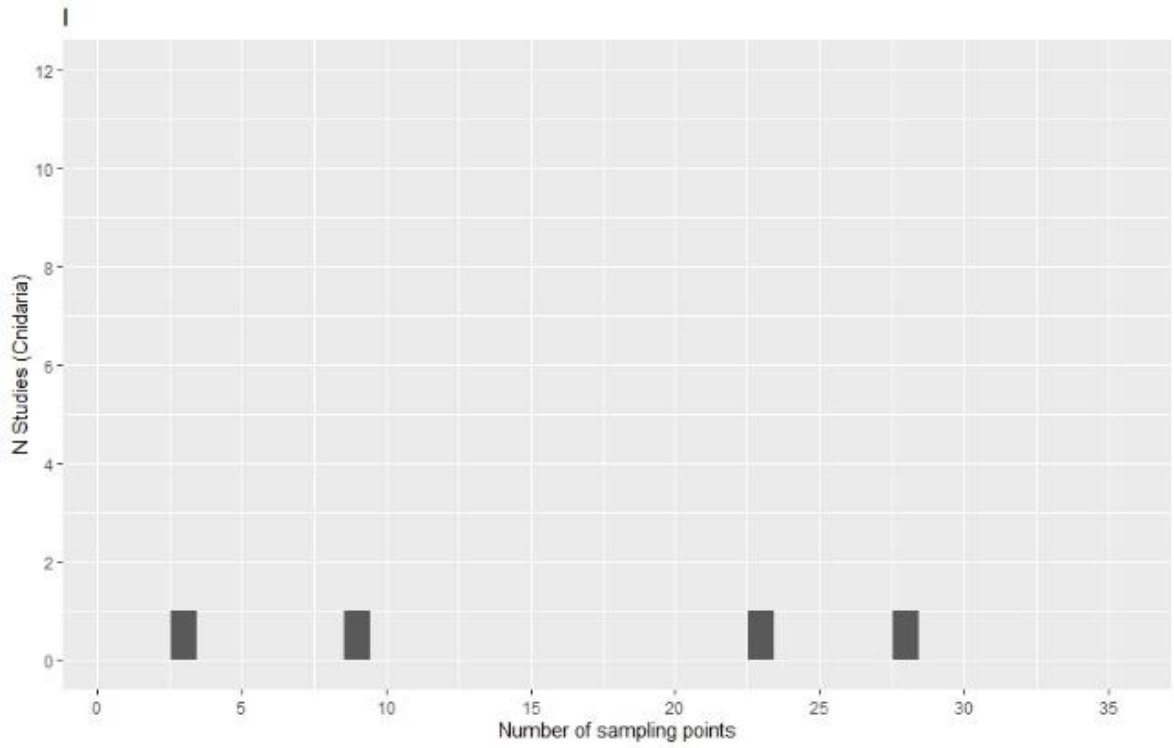

D
N Studies (Arthropoda)
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35
Number of sampling points

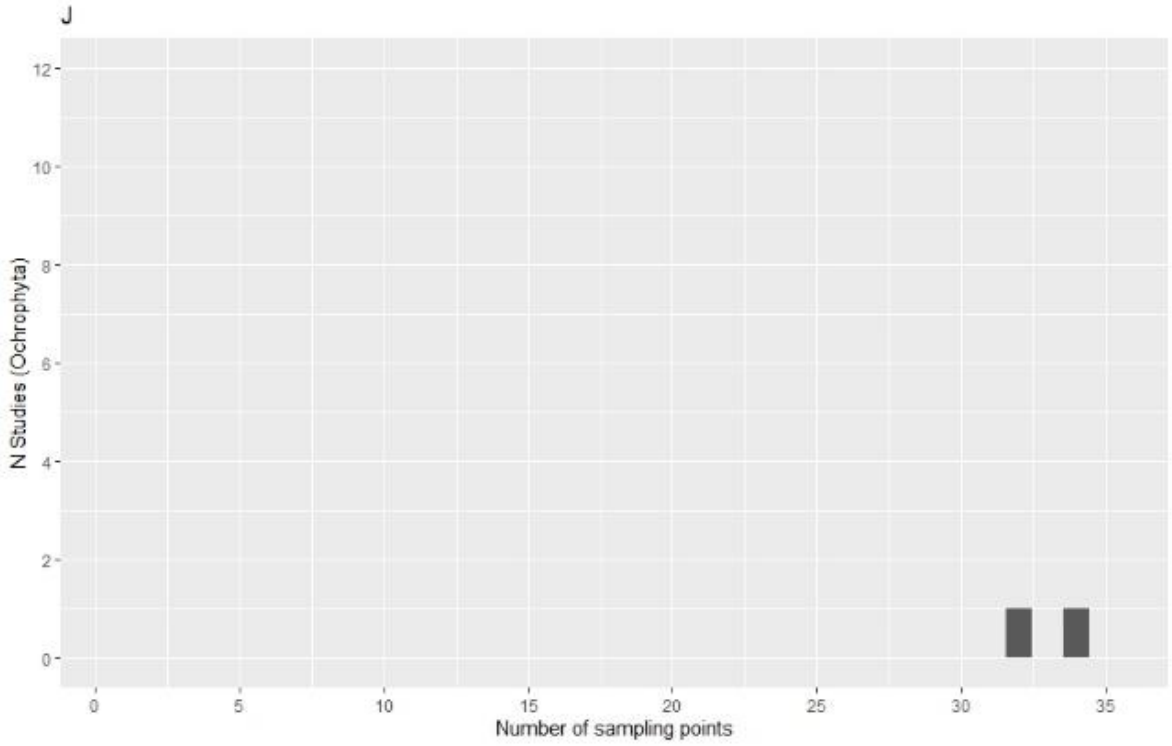

E
N Studies (Annelida)
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35
Number of sampling points

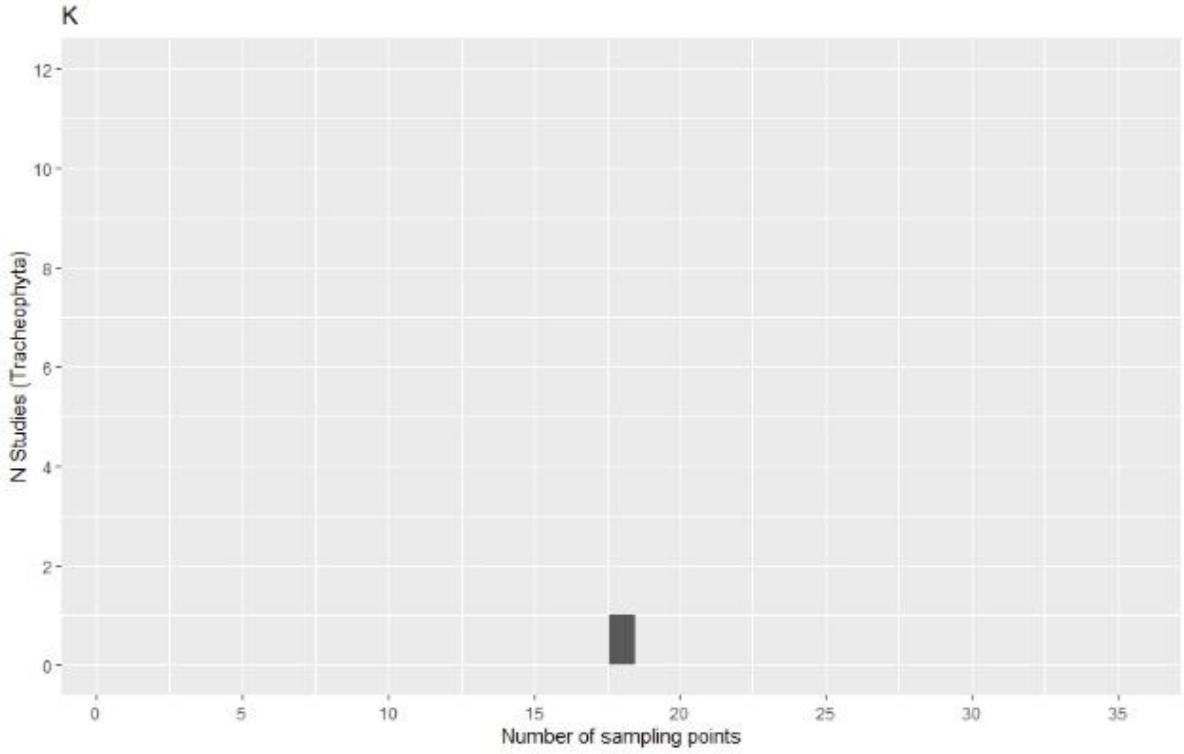

F
N Studies (Mollusca)
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35
Number of sampling points


G
N Studies (Brachiopoda)
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35
Number of sampling points


H
N Studies (Platyhelminthes)
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35
Number of sampling points


I
N Studies (Cnidaria)
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35
Number of sampling points


J
N Studies (Ochrophyta)
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35
Number of sampling points


K
N Studies (Tracheophyta)
0
2
4
6
8
10
12
0
5
10
15
20
25
30
35
Number of sampling points

Supplementary Material 12. Summary of species included in the review of phylogeographic genetic studies of marine vertebrates and invertebrates, macroalgae and plants conducted in New Zealand. Emphasis has been placed on the presence or absence of the pelagic larval stage and its duration for each species.

Taxonomic group	**Common name**	**Species**	**Pelagic Larval Duration (PLD)**	**Reference**
Brown Macroalgae-Ochrophyta	Common flapjack	*Carpophyllum maschalocarpum*	-	(Dunmore, 2006)
Brown Macroalgae-Ochrophyta	Southern bull kelp	*Durvillaea antarctica*	-	(Dunmore, 2006)
Higher Plant-Tracheophyta	Seagrass	*Zostera muelleri*	-	(Orth et al., 1994)
Invertebrate-Cnidaria	Waratah anemone	*Actinia tenebrosa*	20	(Ottaway, 1979)
Invertebrate-Cnidaria	Red coral	*Errina novaezealandiae*	10	(Miller & Mundy, 2003)
Invertebrate-Cnidaria	Black coral	*Antipathes fiordensis*	10	(Miller & Mundy, 2003)
Invertebrate-Platyhelminthes	Marine trematode parasite	*Maritrema novaezealandensis*	-	(Martorelli et al., 2004)
Invertebrate-Brachiopoda	White brachiopod	*Liothyrella neozelanica*	Unknown	(Ostrow, 2004)
Invertebrate-Brachiopoda	Red brachiopod	*Terebratella sanguinea*	Unknown	(Ostrow, 2004)
Invertebrate-Mollusca	Snakeskin chiton	*Sypharochiton pelliserpentis*	4	(Pearse, 1979)
Invertebrate-Mollusca	Marine snail	*Buccinulum vittatum*	-	(Ponder, 1971)
Invertebrate-Mollusca	Spotted whelk	*Cominella maculosa*	-	(Fleming et al., 2018)
Invertebrate-Mollusca	Red-mouthed whelk	*Cominella virgata*	-	(Fleming et al., 2018)
Invertebrate-Mollusca	Blue-banded Periwinkle, Ngaeti	*Austrolittorina antipodum*	28	(Williams et al., 2003)
Invertebrate-Mollusca	Duck's bill limpet, Rori	*Scutus breviculus*	Unknown	(Waters et al., 2007)
Invertebrate-Mollusca	Marine snail	*Siphonaria raoulensis*	70	(Hodgson, 1999)
Invertebrate-Mollusca	Kermadec Island giant limpet	*Scutellastra kermadecensis*	10	(Blackmore, 1969)
Invertebrate-Mollusca	Golden limpet	*Cellana flava*	11	(Dodd, 1955; Dunmore & Schiel, 2000)
Invertebrate-Mollusca	Ornate limpet	*Cellana ornata*	11	(Dodd, 1955; Dunmore & Schiel, 2000)
Invertebrate-Mollusca	Radiate limpet	*Cellana radians*	11	(Dodd, 1955; Dunmore & Schiel, 2000)
Invertebrate-	Black nerite	*Nerita atramentosa*	168	(Underwood, 1975)

Mollusca				
Invertebrate-Mollusca	New Zealand black-foot abalone	*Haliotis iris*	9	(Tong et al., 1992)
Invertebrate-Mollusca	New Zealand scallop	*Pecten novaezelandiae*	21	(Shumway & Parsons, 2006)
Invertebrate-Mollusca	Pipi	*Paphies australis*	21	(Redfearn, 1987)
Invertebrate-Mollusca	Tuatua	*Paphies subtriangulata*	21	(Redfearn, 1987)
Invertebrate-Mollusca	Hydrothermal vent mussels	*Bathymodiolus spp.*	28	(Lutz et al., 1980)
Invertebrate-Mollusca	New Zealand dredge oyster	*Ostrea chilensis*	1	(Millar & Hollis, 1963)
Invertebrate-Mollusca	Mediterranean mussel	*Mytilus galloprovincialis*	49	(Bayne, 1966)
Invertebrate-Mollusca	Greenshell mussel	*Perna canaliculus*	28	(Gardner et al., 1996b)
Invertebrate-Mollusca	New Zealand little neck clam	*Austrovenus stutchburyi*	21	(Lidgard, 2001)
Invertebrate-Annelida	Quill-worm	*Hyalinoecia longibranchiata*	-	(Paxton, 1986)
Invertebrate-Arthropoda	Squat lobster	*Munida gracilis*	83	(Pérez-Barros et al., 2007; Baba et al., 2011)
Invertebrate-Arthropoda	New Zealand scampi	*Metanephrops challengeri*	5	(Wear, 1976)
Invertebrate-Arthropoda	Packhorse rock lobster	*Jasus verreauxi*	336	(Smith et al., 1980)
Invertebrate-Arthropoda	Red rock lobster	*Jasus edwardsii*	336	(Smith et al., 1980)
Invertebrate-Arthropoda	Isopod	*Isocladus armatus*	-	(Dexter, 1977)
Invertebrate-Arthropoda	Common rock crab	*Hemigrapsus sexdentatus*	49	(Wear, 1970)
Invertebrate-Arthropoda	Pea crab	*Pinnotheres atrinicola*	25	(Goodbody, 1960)
Invertebrate-Arthropoda	Pea crab	*Pinnotheres novaezelandiae*	25	(Goodbody, 1960)
Invertebrate-Arthropoda	New Zealand amphipod	*Paracorophium lucasi*	-	(Stevens & Hogg, 2004)
Invertebrate-Arthropoda	New Zealand amphipod	*Paracorophium excavatum*	-	(Stevens & Hogg, 2004)
Invertebrate-Echinodermata	New Zealand common cushion star	*Patiriella regularis*	56	(Byrne & Barker, 1991)
Invertebrate-Echinodermata	Eleven-armed sea star	*Coscinasterias muricata*	28	(Barker, 1977)
Invertebrate-Echinodermata	Brittle star	*Ophiomyxa brevirima*	-	(Fell, 1945)
Invertebrate-Echinodermata	Snake star	*Astrobrachion constrictum*	-	(Fell, 1945)
Invertebrate-Echinodermata	New Zealand sea urchin	*Evechinus chloroticus*	56	(Dix, 1969; Walker, 1984)
Invertebrate-Chordata	New Zealand sea tunicate	*Cnemidocarpa nisiotis*	1	(Svane & Young, 1989)

Vertebrate- Chordata	The spotted estuary smooth-hound or rig	*Mustelus lenticulatus*	-	(Francis, 1988)
Vertebrate- Chordata	New Zealand yellowbelly flounder	*Rhombosolea leporina*	70	(Crawford, 1984)
Vertebrate- Chordata	Spotted robust triplefin	*Forsterygion capito*	90	(Kohn & Clements, 2011)
Vertebrate- Chordata	Common triplefin	*Forsterygion lapillum*	90	(Kohn & Clements, 2011)
Vertebrate- Chordata	Spectacled triplefin	*Ruanoho whero*	90	(Kohn & Clements, 2011)
Vertebrate- Chordata	Tasmanian robust triplefin	*Grahamina gymnota*	90	(Kohn & Clements, 2011)
Vertebrate- Chordata	Variable triplefin	*Forsterygion varium*	90	(Kohn & Clements, 2011)
Vertebrate- Chordata	The pot-bellied seahorse	*Hippocampus abdominalis*	14	(Woods, 2000)
Vertebrate- Chordata	New Zealand blue cod	*Parapercis colias*	10	(Henderson, 2009)
Vertebrate- Chordata	Twister	*Bellapiscis medius*	90	(Kohn & Clements, 2011)
Vertebrate- Chordata	Mottled twister	*Bellapiscis lesleyae*	90	(Kohn & Clements, 2011)
Vertebrate- Chordata	New Zealand sand flounder	*Rhombosolea plebeia*	70	(Crawford, 1984)
Vertebrate- Chordata	Estuarine triplefin	*Grahamina nigripenne*	90	(Kohn & Clements, 2011)
Vertebrate- Chordata	Yellow-eyed penguin	*Megadyptes antipodes*	-	(Efford et al., 1996)
Vertebrate- Chordata	New Zealand sea lion	*Phocarctos hookeri*	-	(Chilvers et al., 2007)
Vertebrate- Chordata	New Zealand fur seal	*Arctocephalus forsteri*	-	(Dickie & Dawson, 2003)
Vertebrate- Chordata	Maui´s Dolphin	*Cephalorhynchus hectori maui*	-	(Dawson, 2009)
Vertebrate- Chordata	Hector´s Dolphin	*Cephalorhynchus hectori*	-	(Dawson, 2009)

Most of the presented PLD stages were consulted/extracted from the review entitled “Phylogeography of New Zealand’s coastal benthos” (Ross et al., 2009).

Supplementary Material 13. R-script to fit spatial patterns of connectivity/diversity to four hypothetical models of distribution.

```
## Script modified from the Original Article entitled "Range limits and geographic patterns of abundance of the rocky intertidal owl limpet, Lottia gigantea"  (Fenberg & Rivadeneira, 2011)
## R-code to evaluate the Abundant-centre hypothesis (ACH) across a latitudinal gradient
## The data set must be composed of at least five-column text file delimited by comas (.csv)
## Column 1 should contain the Range Index of each site
## Column 2-5 should contain the indices (any quantitative variable. i.e. FST, Number of haplotypes, Allelic richness, Haplotype diversity, Total abundance, etc). In this study, genetic indices were tested
## Genetic data was freely available, and it was downloaded from articles, reports or theses (from the main text or supplementary material). Due to most of the studies did not cover the entire latitudinal range limits for each species, range limits were set manually (by searching the records in distribution and abundance papers for each species). Subsequently, the Range Index was calculated for each sampling point.
## No blanks spaces allowed
## Highlighted parts must be changed depending on the dataset

## R-Code

## Step 1: Select directory

setwd("C:/Users/Daniel/Desktop/ach")

## Step 2: Read a data file

data_55<-read.table("63_Ostrea chilensis_ACH.csv",sep=",",header = T)
dim(data_55)
head(data_55)
str(data_55)

## Step 3: Preliminary analysis of the data

summ_d55<-data.frame(t(apply(data_55[,2:5],c(2),summary)))
colnames(summ_d55)<-c("Min.","Q1","Median","Mean","Q3","Max.")

out55<-apply(data_55[,2:5], c(2), function(x) boxplot.stats(x)$out)
out55pos<-apply(data_55[,2:5], c(2), function(x) which(x %in% c(boxplot.stats(x)$out)))
summ_d55$Outlier<-c(as.character(format(out55,digits = 2)))
summ_d55$Outlier_pos<-c(as.character(out55pos))

print(summ_d55,digits = 2)

## Step 4: In case you need to delete variable
## In this case row 17 was eliminated

data55n<-data_55[-17,]

## Step 5: Building the four models

all.models<-function(ri){
  model.normal<-((1/(0.303859*(2*pi^0.5)))*exp(-0.5*(ri/0.303859)^2))
```

```
  model.no<-model.normal/max(model.normal)
  model.rn<-(0.5+(ri*-0.5))
  model.rs<-(0.5+(ri*0.5))
  model.ae<-ifelse(ri<=0,-ri,ri)
  return(cbind(model.no,model.rn,model.rs,model.ae))
}

## Step 6: Defining variables (hereafter the input data must be the Range Index and one dependent variable)
## Repeat the same procedure multiple times, depending in the total number of dependent variables to be
tested
## Remove outliers. Rows specified in brackets can be removed (outlier positions)

ri<-data_55$Range.Index[-c(1,2,7)]
var1<-data_55$Fst[-c(1,2,7)]/max(data_55$Fst[-c(1,2,7)])

## Step 7: Adjust models

models<-all.models(ri)
ss<-colSums(ifelse(var1>models,(var1-models)^2,0))
ss

## Step 8: The randomizing procedure

start_time <- Sys.time()
runs<-1000000
a<-matrix(ncol=4,nrow=runs)
for (k in 1:runs) {
  obs.null<-sample(var1,replace=FALSE)
  nrss<-colSums(ifelse(obs.null>models,(obs.null-models)^2,0))
  a[k,]<- rbind(nrss)
  next
}
end_time <- Sys.time()
end_time - start_time

## The observed RSS contrasted against the percentiles 0.1, 1, 5 and 10% of the null model
## Model selection according the Akaike Information Criterion (AIC)

rss<-rbind(ss,apply(a,2,quantile,probs=c(0.001,0.01,0.05,0.1),na.rm=T))
dim(rss)
aics<-matrix(1,4)
aics[1,1]<-(2*2)+(25*(log(rss[1,1]/25)))
aics[2,1]<-(2*2)+(25*(log(rss[1,2]/25)))
aics[3,1]<-(2*2)+(25*(log(rss[1,3]/25)))
aics[4,1]<-(2*4)+(25*(log(rss[1,4]/25)))
daics<-aics-min(aics)
aicsw<-exp(-0.5*daics)
AICw<-aicsw/sum(aicsw)
aics<-as.vector(aics)
daics<-as.vector(daics)
AICw<-as.vector(AICw)
table.results<-rbind(rss,aics,daics,AICw)
```

table.results

Supplementary Material 14. Representation percentages for each distribution model ($p<0.05$). Representation percentages for each distribution model per genetic index and phylum (considering the studies with a probability of $p<0.05$ in the fitting of one of the four models). In addition, a summary of the total representation by index and phylum (with and without F_{ST} and Φ_{ST} as variable) are also showed. Numbers highlighted indicated the total percentages accounted for each model across all phyla.

	F_{ST} and/or Φ_{ST}				**F_{IS}**				**H**				**Ar**				**Hr**			
	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE
Chordata	0	0	18	0	0	0	0	0	0	0	0	13	0	17	17	0	0	100	0	0
Echinodermata	0	0	0	17	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-
Arthropoda	0	14	14	0	0	0	0	0	0	0	67	0	-	-	-	-	-	-	-	-
Annelida	0	0	0	0	-	-	-	-	0	0	0	0	-	-	-	-	-	-	-	-
Mollusca	0	0	5	0	0	25	0	0	8.3	25	8	0	0	50	0	0	-	-	-	-
Brachiopoda	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cnidaria	0	0	0	0	-	-	-	-	0	50	0	0	-	-	-	-	-	-	-	-
Ochrophyta	-	-	-	-	-	-	-	-	50	0	0	0	-	-	-	-	-	-	-	-
Tracheophyta	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Total	0	1.8	9	1.8	0	8.3	0	0	6.5	13	10	3.2	0	33	8	0	0	100	0	0

Supplementary Material 14 continued.

Pi				**h**				H_O				H_E				**Total w/ F_{ST} and/or Φ_{ST}**				**Total w/o F_{ST} and/or Φ_{ST}**			
NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE
0	21	0	14	7.1	14	7	14	0	0	10	0	0	0	11	0	1.2	8.2	8	5.9	1.5	10	6	7.4
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.6	0	0	0	0
0	33	0	0	0	33	0	0	0	0	20	0	0	0	40	0	0	11	22	0	0	10	25	0
0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0
10	10	10	0	7.7	15	15	0	0	0	0	0	0	25	0	0	3.4	15	6	0	4.3	19	6	0
-	-	-	-	-	-	-	-	0	0	0	0	0	0	33	0	0	0	17	0	0	0	25	0
-	-	-	-	0	0	0	0	0	50	0	0	0	0	0	0	0	25	0	0	0	29	0	0
0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	25	0	0	0	25	0	0	0
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	-	-	-	-
3.2	16	3	6.5	5.7	14	9	5.7	0	3	6	0	0	8.6	11	0	2	10	8	2.4	**2.6**	**13**	**7**	**2.6**

Supplementary Material 15. Representation percentages for each distribution model ($p<0.1$). Representation percentages for each distribution model per genetic index and phylum (considering the studies with a probability of $p<0.1$ in the fitting of one of the four models). In addition, a summary of the total representation by index and phylum (with and without F_{ST} and Φ_{ST} as variable) are also showed. Numbers highlighted indicated the total percentages accounted for each model across all phyla.

	F_{ST} and Φ_{ST}				**F_{IS}**				**H**				**Ar**				**Hr**			
	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE
Chordata	5.9	0	29	0	0	33	0	0	0	25	0	13	0	17	17	0	0	100	0	0
Echinodermata	0	0	0	50	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-
Arthropoda	0	29	14	0	0	0	0	0	0	0	67	0	-	-	-	-	-	-	-	-
Annelida	0	0	0	0	-	-	-	-	0	0	0	0	-	-	-	-	-	-	-	-
Mollusca	0	0	10	0	0	25	0	0	8.3	25	17	0	0	50	0	0	-	-	-	-
Brachiopoda	0	0	50	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cnidaria	0	0	0	0	-	-	-	-	0	50	0	0	-	-	-	-	-	-	-	-
Ochrophyta	-	-	-	-	-	-	-	-	50	0	0	0	-	-	-	-	-	-	-	-
Tracheophyta	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Total	1.8	3.6	16	5.4	0	25	0	0	6.5	19	13	3.2	0	33	8	0	0	100	0	0

Supplementary Material 15 continued.

Pi				**h**				***H*$_O$**				***H*$_E$**				**Total w/ *F*$_{ST}$ and *Φ*$_{ST}$**				**Total w/o *F*$_{ST}$ and *Φ*$_{ST}$**			
NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE	NO	RN	RS	AE
0	36	0	14	7.1	14	7	21	0	0	20	0	0	0	11	0	2.4	15	12	7.1	1.5	19	7	8.8
0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	6	17	0	0	8	0
0	33	0	0	0	33	33	0	0	0	40	0	0	0	40	0	0	15	30	0	0	10	35	0
0	0	0	0	0	0	50	0	-	-	-	-	-	-	-	-	0	0	13	0	0	0	17	0
20	30	10	0	7.7	23	15	0	0	17	0	0	0	33	0	0	4.5	21	8	0	5.8	28	7	0
-	-	-	-	-	-	-	-	0	0	0	0	0	0	33	0	0	0	33	0	0	0	25	0
-	-	-	-	0	0	0	0	0	50	0	0	0	0	0	0	0	25	0	0	0	29	0	0
0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	25	0	0	0	25	0	0	0
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	-	-	-	-
6.5	29	3	6.5	5.7	17	17	8.6	0	9.1	12	0	0	11	11	0	2.8	15	12	3.7	**3.2**	**19**	**11**	**3.2**

Supplementary Material 16. Best fitting constrains shape for goodness-of-fit tests for distributions of different genetic indices (p<0.1). Biogeographic studies for which significant results were registered for one of the four hypothetical distribution models are shown. (**A**) Phylum Chordata, (**B**) Phylum Echinodermata, (**C**) Phylum Arthropoda, (**D**) Phylum Annelida, (**E**) Phylum Molluscs, (**F**) Phylum Brachiopoda, (**G**) Phylum Cnidaria and (**H**) Phylum Ochrophyta. Shape refers to distribution shapes from Fig. 1. Significant ***P-values*** (bold) indicate that the distribution of the genetic indices (connectivity and diversity) fit the distribution shape more closely than a random distribution in 95% or more of the 10^5 randomization procedures (see Material and Methods, Chapter 1). Genetic indices. F_{ST} and Φ_{ST}: Fixation index, F_{IS}: Inbreeding coefficient, H: Haplotypes, Ar: Allelic richness, Hr: Haplotype richness, Pi: Nucleotide diversity, h: Haplotype diversity, H_O: Observed heterozygosity, H_E: Expected heterozygosity.

(**A**) Phylum Chordata

Species	**Habitat/Endemism**	**Index**	**Shape**	**P-value**	**Marker**	**Reference**
Cephalorhynchus hectori	Open Coast/NZ	H Pi	Ramped North Ramped North	p<0.1 p<0.1	mtDNA (CR)	(Pichler, 2002)
	Open Coast/NZ	F_{ST} H Pi h	Ramped South Ramped North Abundant Edge Abundant Edge	p<0.1 p<0.1 ***p<0.01*** ***p<0.05***	mtDNA (CR)	(Hamner *et al.*, 2012b)
	Open Coast/NZ	F_{ST}	Ramped South	***p<0.01***	SSR	(Hamner *et al.*, 2012b)
Arctocephalus forsteri	Open Coast/NZ-AUS	F_{IS}	Ramped North	p<0.1	SSR	(Robertson & Gemmell, 2005)
	Open Coast/NZ-AUS	F_{ST}	Ramped South	***p<0.05***	SSR	(Dussex *et al.*, 2016)
	Open Coast/NZ-AUS	F_{IS}	Ramped North	p<0.1	SNP	(Russell, 2016)
Phocarctos hookeri	Open Coast/NZ	h	Abundant Edge	p<0.1	mtDNA (D-loop)	(Collins *et al.*, 2016, 2017)
	Open Coast/NZ	h	Normal	***p<0.05***	mtDNA (CytB)	(Collins *et al.*, 2016, 2017)
Megadyptes antipodes	Open Coast/NZ	F_{ST} H Pi h	Normal Abundant Edge Abundant Edge Abundant Edge	p<0.1 ***p<0.05*** ***p<0.01*** ***p<0.05***	mtDNA (HVI)	(Boessenkool *et al.*, 2009)
	Open Coast/NZ	H_O H_E	Ramped South Ramped South	***p<0.05*** ***p<0.01***	SSR	(Boessenkool *et al.*, 2009)
Bellapiscis medius	Rocky intertidal/NZ	Pi h	Ramped North Ramped North	***p<0.05*** ***p<0.05***	mtDNA (CR)	(Hickey *et al.*, 2009)

Forsterygion lapillum	Rocky subtidal/NZ	Pi h	Ramped North Ramped North	***p<0.001*** ***p<0.001***	mtDNA (CR)	(Hickey *et al.*, 2009)
	Rocky subtidal/NZ	Ar	Ramped North	***p<0.01***	SSR	(Rabone *et al.*, 2015)
Grahamina nigripenne	Estuarine/NZ	Pi	Ramped North	p<0.1	mtDNA (CR)	(Hickey et al., 2009)
Parapercis colias	Rocky intertidal/NZ	Pi	Ramped North	***p<0.01***	mtDNA (CR)	(Smith, 2012; Gebbie, 2014)
Rhombosolea plebeia	Estuarine/NZ	Ar H H_O	Ramped South Ramped South Ramped South	***p<0.01*** ***p<0.05*** p<0.1	SSR	(Constable, 2014)
Hippocampus abdominalis	Rocky intertidal/NZ-AUS	Φ_{ST} Hr	Ramped South Ramped North	***p<0.05*** ***p<0.05***	mtDNA (CR)	(Ashe & Wilson, 2020)
	Rocky intertidal/NZ-AUS	F_{ST}	Ramped South	p<0.1	SSR	(Ashe & Wilson, 2020)

(**B**) Phylum Echinodermata

Species	**Habitat/Endemism**	**Index**	**Shape**	**P-value**	**Marker**	**Reference**
Evechinus chloroticus	Fjord/NZ	F_{ST}	Abundant Edge	p<0.1	SSR	(Perrin & Roy, 2000; Perrin, 2002)
	Rocky subtidal/NZ	F_{ST}	Abundant Edge	p<0.1	mtDNA (COI)	(Nagel *et al.*, 2015)
	Rocky subtidal/NZ	F_{ST}	Abundant Edge	***p<0.05***	SSR	(Nagel *et al.*, 2015)
Coscinasterias muricata	Fjord/NZ-AUS	h	Ramped South	p<0.1	Allozymes	(Sköld et al., 2003)

(**C**) Phylum Arthropoda

Species	**Habitat/Endemism**	**Index**	**Shape**	**P-value**	**Marker**	**Reference**
Pinnotheres atrinicola	Rocky intertidal/NZ	F_{ST}	Ramped North	***p<0.05***	Allozymes	(Stevens, 1991)
Paracorophium lucasi	Estuarine/NZ	H_O H_E	Ramped South Ramped South	p<0.1 ***p<0.05***	Allozymes	(Stevens & Hogg, 2004)
Paracorophium excavatum	Estuarine/NZ	H_O H_E	Ramped South Ramped South	***p<0.001*** ***p<0.05***	Allozymes	(Stevens & Hogg, 2004)
Hemigrapsus sexdentatus	Rocky intertidal/NZ	F_{ST} H h	Ramped North Ramped South Ramped North	p<0.1 ***p<0.001*** ***p<0.01***	mtDNA (COI-RFLP)	(Hinnendael, 2008)
Munida gracilis	Soft substrate subtidal/NZ-AUS	H h	Ramped South Ramped South	***p<0.05*** p<0.1	mtDNA (COI)	(Bors *et al.*, 2012)
Metanephrops challengeri	Soft substrate subtidal/NZ	Φ_{ST} Pi	Ramped South Ramped North	***p<0.05*** ***p<0.05***	mtDNA (COI)	(Verry *et al.*, 2020)

(**D**) Phylum Annelida

Species	**Habitat/Endemism**	**Index**	**Shape**	**P-value**	**Marker**	**Reference**
Hyalinoecia	Soft substrate	h	Ramped	p<0.1	mtDNA	(Bors et al.,

longibranchiata	subtidal		South		(COI)	2012)

(**E**) Phylum Molluscs

Species	**Habitat/Endemism**	**Index**	**Shape**	**P-value**	**Marker**	**Reference**
Paphies subtriangulata	Soft substrate subtidal/NZ	H_O H_E	Ramped North Ramped North	p<0.1 ***p<0.05***	Allozymes	(Smith *et al.*, 1989)
	Soft substrate subtidal/NZ	F_{ST} Ar H_E	Ramped South Ramped North Ramped North	p<0.1 ***p<0.01*** ***p<0.01***	SSR	(Hannan, 2014)
Perna canaliculus	Rocky intertidal/NZ	H h	Ramped South Ramped South	***p<0.01*** ***p<0.01***	mtDNA (NADH4, NADH2, COI - SSCP & RFLP)	(Apte *et al.*, 2003)
Pecten novaezelandiae	Soft substrate subtidal/NZ	F_{ST} F_{IS} Ar H_E	Ramped South Ramped North Ramped North Ramped North	***p<0.01*** ***p<0.01*** ***p<0.001*** p<0.1	SSR	(Silva & Gardner, 2016)
Ostrea chilensis	Rocky subtidal/NZ	Pi h	Normal Normal	p<0.1 ***p<0.05***	mtDNA (COI)	(Guo *et al.*)
Cellana radians	Rocky intertidal/NZ	Pi h	Ramped South Ramped South	***p<0.01*** ***p<0.05***	mtDNA (CytB)	(Goldstien, 2005; Goldstien *et al.*, 2006)
Cellana ornata	Rocky intertidal/NZ	Pi	Ramped North	p<0.1	mtDNA (CytB)	(Goldstien, 2005; Goldstien et al., 2006)
Cellana flava	Rocky intertidal/NZ	Pi h	Ramped North Ramped North	***p<0.05*** ***p<0.05***	mtDNA (CytB)	(Goldstien, 2005; Goldstien *et al.*, 2006)
Austrolittorina antipodum	Rocky intertidal/NZ	H	Ramped South	p<0.1	mtDNA (COI)	(Waters et al., 2007)
Haliotis iris	Rocky intertidal/NZ	Ar H_O H_E	Ramped North Ramped North Ramped North	***p<0.01*** p<0.1 ***p<0.01***	SSR	(Will & Gemmell, 2008; Will *et al.*, 2015)
	Rocky intertidal/NZ	H Pi h	Ramped North Ramped North Ramped North	***p<0.01*** p<0.1 ***p<0.001***	mtDNA (COI, ATP8-ATP6)	(Will & Gemmell, 2008; Will *et al.*, 2011)
Cominella virgata	Rocky intertidal/NZ	H h	Ramped North Ramped North	***p<0.05*** p<0.1	mtDNA (COI)	(Walton, 2017; Fleming *et al.*, 2018; Walton *et al.*, 2019)
Cominella maculosa	Rocky intertidal/NZ	H	Ramped North	***p<0.05***	mtDNA (COI)	(Walton, 2017; Fleming *et al.*, 2018; Walton *et al.*, 2019)
	Rocky intertidal/NZ	H Pi	Normal Normal	***p<0.01*** ***p<0.05***	mtDNA (COI)	(Dohner *et al.*, 2018)

(**F**) Phylum Brachiopoda

Species	**Habitat/Endemism**	**Index**	**Shape**	**P-value**	**Marker**	**Reference**
Terebratella sanguinea	Fjord/NZ	H_E	Ramped South	***p<0.05***	AFLP	(Ostrow, 2004)
Liothyrella neozelanica	Fjord/NZ	F_{ST}	Ramped South	p<0.1	AFLP	(Ostrow, 2004)

(**G**) Phylum Cnidaria

Species	**Habitat/Endemism**	**Index**	**Shape**	**P-value**	**Marker**	**Reference**
Antipathes fiordensis	Fjord/NZ	H_O	Ramped North	***p<0.01***	Allozymes	(Miller, 1997)
Actinia tenebrosa	Rocky intertidal/NZ-AUS	H	Ramped North	***p<0.01***	SSR	(Veale, 2007; Veale & Lavery, 2012)

(**H**) Phylum Ochrophyta

Species	**Habitat/Endemism**	**Index**	**Shape**	**P-value**	**Marker**	**Reference**
Durvillaea antarctica	Rocky intertidal/SH	H	Normal	***p<0.001***	mtDNA (COI)	(Collins *et al.*, 2010)

Supplementary Material 17. List of independent environmental variables derived from the Exclusive Economic Zone New Zealand Marine Environment Classification (EEZ-MEC).

Environmental variable	**Abbreviation**	**Description**	**Units**
Depth †	Depth	Bathymetry grid (1km resolution)	m
Annual mean solar radiation †	Rad_mean	Mean extra atmospheric solar radiation modified by mean annual cloud cover	Wm^{-2}
Winter solar radiation	Rad_wint	Extra atmospheric solar radiation in June, modified by mean annual cloud cover	Wm^{-2}
Wintertime sea surface temperature †	SSTwint	Mean of daily data from early September when SST is typically lowest	°C
Annual amplitude of sea surface temperature †	SSTanamp	Smoothed annual amplitude of SST	°C
Spatial gradient annual mean sea surface temperature †	SSTgrad	Smoothed magnitude of the spatial gradient of annual mean SST	°C km^{-1}
Summertime sea surface temperature anomaly †	SSTanom	Spatial anomalies with scales between 20 and 450 km in late February when SST is typically highest	°C
Mean orbital velocity †	Orb_v_mean	Orbital velocity at the bed for the mean significant wave height calculated from a 20-year wave hindcast	m/s
Extreme orbital velocity	Orb_v_95	Orbital velocity at the bed for the 95^{th} percentile significant wave height calculated from a 20-year wave hindcast	m/s
Tidal current †	Tidal	Depth averaged maximum tidal current	m/s
Sediment type (categorical variable)	Sed	Sediment type as a categorical variable	na
Seabed rate of change of slope (profile)	Bed_prof	The rate of change of slope for each cell	$0.01m^{-1}$
Seabed curvature	Bed_curv	Curvature of the surface surrounding each grid cell	$0.01m^{-1}$
Seabed planform curvature	Bed_plan	Curvature of the surface perpendicular to the slope direction	$0.01m^{-1}$
Bed slope †	Bed_slope	Slope as a measure of bed shape over the three bed variables	$0.01m^{-1}$
Freshwater fraction	FW	Proportion of fresh water based on river inputs	proportion

The presented details can be further revised from the final report submitted to the Ministry for the Environment (Snelder et al., 2005). † The symbol indicates the variables incorporated in the Seascapes Genetics analysis.

Supplementary Material 18. Description of the method for the extraction of environmental variables using the Exclusive Economic Zone New Zealand Marine Environment Classification (EEZ-MEC).

Step 1: Studies' Geographic Information System

1. The interest sampling locations (GPS coordinates) from all the genetic studies included in the review, were registered if the paper/report/thesis contained them. If the coordinates were provided in sexagesimal degrees were transform to decimal degrees using Google Earth Pro 7.3.
2. If a study did not contain GPS coordinates information for the sampling locations, these were obtained by searching each of them in Google Earth Pro v7.3. This were done by placing the cursor in the middle of the bay or geographical area described in the study (also recorded in decimal degrees).
3. The geographical information was organized and stored in a database created in Microsoft Excel (Microsoft Office 365 ProPlus).

Step 2: Geographic Information System Projection

1. The GPS coordinates in their majority were obtained in Google Earth, for this reason the default Geographic Information System Projection was WGS1984.
2. For the studies in which the authors provided the information, we assumed that this data was obtained by GPS devices and managed/exported to Google Earth.

Step 3: EEZ-MEC (NIWA) Geographic Information System Projection

1. The raster's files Geographic Information System Projection was Mercator 1SP (as stated in the files specifications), which uses the spheroid Clarke 1866.

Step 4: Export GPS coordinates in decimal degrees to ArcGIS

1. The GPS coordinates were ordered in Excel for each sampling location (decimal degrees). Assigning to each of them an ID number, name, latitude and longitude (four columns). The Excel spreadsheet must be store, when ready to export, as an excel version 1997, to later being correctly identified by ArcGIS v10.8. For matching purposes, the headers names for latitude and longitude in the excel spreadsheet must be, Y and X, respectively.
2. It is highly recommended to create a folder for each excel spreadsheet, for being able later to track the path to the executable file.
3. Once ready, open the excel sheet ("Add Data") with the GPS coordinates in decimal degrees as a new layer in ArcGIS.
4. Right click button in the created layer ("sheet1$") to display XY data.
5. X and Y fields must recognize your longitude (X) and latitude (Y) (by default).
6. After this, press the edit tab, to specify the Geographic Coordinate System of your metadata contained in the excel sheet. Select the plus icon besides the folder named as "Geographic Coordinate System". Select the same plus icon in the folder named "World", search for the system "WGS1984" and select.
7. Given all this, you should be able to see your points in the workspace in the specified geographic coordinate system.
8. Right click button in the new created layer named as "sheet1$Events" and select "Data", subsequently "Export Data".
9. A popup window will appear to save the new shapefile in a folder. I recommend saving it in the same folder where you store the initial excel spreadsheet, by copy and paste the path to the executable file.
10. Once created the shapefile, close everything and open it again (now only the shapefile layer).

Step 5: Overlapping GPS points as shapefile with raster's files

1. Once open the shapefile layer with your points in your table of contents, select the icon "Add Data".
2. Open the folder which contains the "New Zealand Marine Environment Classification". Select "EEZ". Subsequently, select "Env_vars" and finally "esri".
3. Open both folders containing the raster's files named "final" and "extra".
4. Automatically will appear a window called "Geographic Coordinate Systems Warning", showing you the Geographic Coordinate System specified for the rasters (GCS_Clarke_1866).
5. Select "Transformations".

6. In the next window named “Geographic Coordinate System Transformations”, by default must appear the selection of the correct Geographic Coordinate System from which you want to convert (“GCS_Clarke_1866”) your raster’s files into the new one (“GCS_WGS_1984”). Press OK. Finally close the warning window.
7. Points will overlap with rasters.

Step 6: Extract multi values to points (Environmental data)

1. Before the extraction of the environmental values, make sure to select the extensions you want to use, in this case Spatial Analyst (Customize).
2. Select “ArcToolbox” → “Spatial Analyst Tools” → “Extraction” → “Extract Multi Values to Points”.
3. Select input point features (points shapefile layer) and inputs rasters (select interest rasters) in the popup window. Press OK.
4. The data will be exported to the attribute table of your points shapefile layer.
5. Check for the environmental data by selecting right click button in the points shapefile layer and press “Open Attribute Table”.

Download Attribute Table with Environmental Data

1. Select “ArcToolbox” → “Conversion Tools” → “Excel” → “Table To Excel”.
2. Select input table (points shapefile layer) and output excel file (path to the final store folder) in the popup window. Press OK.

Supplementary Material 19. R-script to execute the GLM and GLZ analyses (conventional and backward removal).

R-code to execute the GLM procedure (conventional and backward removal).
The data set must be composed of dependent and independent variables in a text file delimited by comas (.csv).
In this case, the first columns (1-4) are genetic indices (dependent variables; i.e. Nucleotide diversity, Haplotype diversity, Haplotypes, F_{ST}, etc).
The rest of the columns (5-16) contain the environmental variables (independent variables; i.e. Latitude, Longitude, Summertime sea surface temperature anomaly, Tidal current, Wintertime sea surface temperature, Spatial gradient annual mean sea surface temperature, Annual amplitude of sea surface temperature, Annual mean solar radiation, Mean orbital velocity, Depth, Bed slope, etc).
Genetic data was freely available, and it was downloaded from articles, reports or theses (from the main text or supplementary material).
##The environmental variables were download from the New Zealand Exclusive Economic Zone - Marine Environment Classification (EEZ-MEC).
Normal and Poisson distributions must be manually set in the script for continuous or discrete variables, respectively (in this case, only Haplotypes).
No blanks spaces allowed in the dependent variables.
Blank spaces within the environmental variables were replaced by average imputation (mean value arise from the existing data in each column where missing values were encountered).
Highlighted parts must be changed depending on the dataset.
Due to the requirement (to execute a GLM) to have at least, number of rows + intercept + 1 columns of independent variables; the analysis was divided into 2, mainly because of the size of the matrices due to low numbers of sampling points in each study.

```
## R-Code

## Step 1: Set empty environment in RStudio

rm(list=ls())

## Step 2: Select directory

setwd("~/Documents/Daniel/glz")

library(StepReg)

## Step 3: Read a data file

mydata = read.csv("ID5.csv", header= TRUE)

## Step 4: Show columns names

colnames(mydata)

## Step 5: Replace missing values (N/A) by the average value calculated for each column

sum(is.na(mydata))

for(i in 1:ncol(mydata)){

  mydata[is.na(mydata[,i]), i] <- mean(mydata[,i],na.rm = T)

}

mydata$rad_mean

## Step 6: Select dependent and independents variables

varDep <- list("Nucleotide.diversity", "Haplotype.diversity","Haplotypes","Fst")
```

```
varsInd <-
list("Latitude+Longitude+sstanom+tidal+sstwint+sstgrad","sstanamp+rad_mean+orb_v_mean+depth_band+bed
_slope")

## Step 7: GLM conventional procedure

summaryGLM_allmod<-NULL

summGLMback_allmod<-NULL

set.seed(123)

for (i in 1:length(varDep)){

  for (j in 1:length(varsInd)){

    model<-NULL

    model<- glm(paste(varDep[i],"~",varsInd[j]), family = "gaussian", data= mydata)# by default na.omit

    if (paste(varDep[i]) == "Haplotypes"){

     model<- glm(paste(varDep[i],"~",varsInd[j]), family = "poisson", data= mydata)

    }

    summod<-NULL

    summod<-data.frame(summary(model)$coefficients)# summary table with P-value

    rownames(summod)<-NULL

    colnames(summod)<-c("Estimate","Std.Error","t_or_z-value","P-value")

    summod$VarDep<-rep(paste(varDep[i]),nrow(summod))

    summod$VarInd<-rownames(summary(model)$coefficients)

    summod=summod[,c(5,6,1,2,3,4)]

    varindmod<-c(summod$VarInd)

    summaryGLM_allmod<-rbind(summaryGLM_allmod,summod)

 ## Step 8: Backward removal procedure

    summGLMback<-matrix(NA,nrow = nrow(summod),ncol = (ncol(summod)+3),

                dimnames=list(summod$VarInd,c(colnames(summod),"score","pvalue","No.It")))

    backwGLM<-NULL

    backwGLM <- step(model, direction="backward",trace = 0) #trace=1 show the executed steps

    summGLMback[,1:2]<-as.matrix(summod[,1:2])

    summGLMback[rownames(summary(backwGLM)$coefficients),3:6]<-summary(backwGLM)$coefficients

    selcol<-c(rownames(summary(backwGLM)$coefficients))

    if (length(varindmod)>length(selcol)){

    datan<-NULL

    datan<-mydata[,as.character(c(paste(varDep[i]),varindmod[-1]))]

    finalmod<-glm(backwGLM$formula,family = "gaussian", data=mydata)
```

```
  if (paste(varDep[i]) == "Haplotypes"){
    finalmod<- glm(backwGLM$formula, family = "poisson", data= mydata) # by default na.omit
  }
  summGLMback[varindmod[-which(varindmod%in%selcol)],"score"]<-
    scoretest(finalmod, as.matrix(datan[,-which(colnames(datan)%in%c(paste(varDep[i]),selcol[-1]))]))$score
  summGLMback[varindmod[-which(varindmod%in%selcol)],"pvalue"]<-
    scoretest(finalmod, as.matrix(datan[,-which(colnames(datan)%in%c(paste(varDep[i]),selcol[-1]))]))$pvalue
  }
  summGLMback[,"No.It"]<-rep(summary(backwGLM)$iter,nrow(summod))
  rownames(summGLMback)<-NULL
  summGLMback<-data.frame(summGLMback) # summary table with P-value
  colnames(summGLMback)<-c(colnames(summod),"Score","P-value","No.It")
  summGLMback_allmod<-rbind(summGLMback_allmod,summGLMback)

 }
}
## Step 9: Export summary table of the GLM conventional procedure into excel
summaryGLM_allmod
write.table(summaryGLM_allmod,file = "summaryGLM_allmod_st1.csv",sep=",",row.names = FALSE)
## Step 10: Export summary table of the Backward removal procedure into excel
summGLMback_allmod
write.table(summGLMback_allmod,file = "summGLMback_allmod_st1.csv",sep=",",row.names = FALSE)
```

Supplementary Material 20. Genetic differentiation (F_{ST} and Φ_{ST}: genetic distance between/amongst populations, **A**) and genetic diversity indices (F_{IS}: Inbreeding coefficient, **B**; H: Number of haplotypes, **C**; Ar: Allelic richness, **D**; Hr: Haplotype richness, **E**; Pi: Nucleotide diversity, **F**; h: Haplotype diversity, **G**; H_O: Observed heterozygosity, **H**; H_E: Expected heterozygosity, **I**) tested against 11 independent variables (2 geospatial and 9 environmental) using the GLZ (with stepwise backward removal) routine across phyla. In parentheses, maximum number of studies tested for the genetic index using the 11 variables for each analysis. Numbers specified per variable correspond to the number of times that the variable was statistically significant (p-value < 0.05) out of the total number of studies per phylum/phyla.

(**A**)

Index	Ind. Var	All phyla	Chordata (17)	Echinodermata (6)	Arthropoda (7)	Annelida (2)	Mollusca (20)	Brachiopoda (2)	Cnidaria (1)	Ochrophyta (0)	Tracheophyta (1)
F_{ST}-Φ_{ST} (GLZ)	Lat	16	4	2	3	2	4	0	0	-	1
	Lon	13	4	2	1	0	5	0	1	-	0
	Tidal	5	2	1	0	0	1	0	1	-	0
	Depth	9	2	2	1	0	3	0	1	-	0
	Bedslope	4	3	0	0	1	0	0	0	-	0
	Orbvmean	11	6	2	0	0	2	0	0	-	1
	SSTanom	14	1	3	1	0	7	2	0	-	0
	SSTwint	15	5	3	1	1	5	0	0	-	0
	SSTgrad	12	3	2	0	0	5	1	1	-	0
	SSTanamp	15	4	2	1	0	6	2	0	-	0
	Radmean	11	6	2	1	0	1	0	1	-	0

(**B**)

Index	Ind. Var	All phyla	Chordates (6)	Echinodermata (1)	Arthropoda (1)	Annelida (0)	Mollusca (4)	Brachiopoda (0)	Cnidaria (0)	Ochrophyta (0)	Tracheophyta (0)
F_{IS} (GLZ)	Lat	3	2	0	0	-	1	-	-	-	-
	Lon	2	1	0	0	-	1	-	-	-	-
	Tidal	0	0	0	0	-	0	-	-	-	-
	Depth	2	2	0	0	-	0	-	-	-	-
	Bedslope	0	0	0	0	-	0	-	-	-	-
	Orbvmean	2	1	0	0	-	1	-	-	-	-
	SSTanom	1	0	0	0	-	1	-	-	-	-
	SSTwint	2	1	0	0	-	1	-	-	-	-
	SSTgrad	2	0	0	0	-	2	-	-	-	-
	SSTanamp	1	1	0	0	-	0	-	-	-	-
	Radmean	0	0	0	0	-	0	-	-	-	-

(**C**)

Index	Ind. Var	All phyla	Chordates (8)	Echinodermata (2)	Arthropoda (3)	Annelida (2)	Mollusca (12)	Brachiopoda (0)	Cnidaria (2)	Ochrophyta (2)	Tracheophyta (0)
H (GLZ)	Lat	3	0	0	1	0	1	-	1	0	-
	Lon	4	2	0	1	1	0	-	0	0	-
	Tidal	4	1	1	0	0	1	-	0	1	-
	Depth	2	0	0	0	0	2	-	0	0	-
	Bedslope	0	0	0	0	0	0	-	0	0	-
	Orbvmean	3	2	0	0	0	0	-	0	1	-
	SSTanom	2	2	0	0	0	0	-	0	0	-
	SSTwint	4	1	1	0	0	0	-	1	1	-
	SSTgrad	2	1	0	0	0	0	-	1	0	-
	SSTanamp	2	0	0	0	0	0	-	1	1	-
	Radmean	3	1	0	1	0	1	-	0	0	-

(**D**)

Index	Ind. Var	All phyla	Chordates (6)	Echinodermata (0)	Arthropoda (0)	Annelida (0)	Mollusca (6)	Brachiopoda (0)	Cnidaria (0)	Ochrophyta (0)	Tracheophyta (0)
Ar (GLZ)	Lat	5	3	-	-	-	2	-	-	-	-
	Lon	2	1	-	-	-	1	-	-	-	-
	Tidal	1	1	-	-	-	0	-	-	-	-
	Depth	1	1	-	-	-	0	-	-	-	-
	Bedslope	2	1	-	-	-	1	-	-	-	-
	Orbvmean	2	0	-	-	-	2	-	-	-	-
	SSTanom	1	0	-	-	-	1	-	-	-	-
	SSTwint	3	2	-	-	-	1	-	-	-	-
	SSTgrad	1	0	-	-	-	1	-	-	-	-
	SSTanamp	1	0	-	-	-	1	-	-	-	-
	Radmean	3	1	-	-	-	2	-	-	-	-

(E)

Index	Ind. Var	All phyla	Chordates (1)	Echinodermata (0)	Arthropoda (0)	Annelida (0)	Mollusca (0)	Brachiopoda (0)	Cnidaria (0)	Ochrophyta (0)	Tracheophyta (0)
Hr (GLZ)	Lat	0	0	-	-	-	-	-	-	-	-
	Lon	0	0	-	-	-	-	-	-	-	-
	Tidal	0	0	-	-	-	-	-	-	-	-
	Depth	0	0	-	-	-	-	-	-	-	-
	Bedslope	0	0	-	-	-	-	-	-	-	-
	Orbvmean	1	1	-	-	-	-	-	-	-	-
	SSTanom	0	0	-	-	-	-	-	-	-	-
	SSTwint	0	0	-	-	-	-	-	-	-	-
	SSTgrad	0	0	-	-	-	-	-	-	-	-
	SSTanamp	0	0	-	-	-	-	-	-	-	-
	Radmean	0	0	-	-	-	-	-	-	-	-

(F)

Index	Ind. Var	All phyla	Chordates (14)	Echinodermata (1)	Arthropoda (3)	Annelida (2)	Mollusca (10)	Brachiopoda (0)	Cnidaria (0)	Ochrophyta (1)	Tracheophyta (0)
Pi (GLZ)	Lat	8	5	1	1	0	1	-	-	0	-
	Lon	8	4	0	1	1	2	-	-	0	-
	Tidal	8	5	1	1	1	0	-	-	0	-
	Depth	4	2	0	1	0	1	-	-	0	-
	Bedslope	2	2	0	0	0	0	-	-	0	-
	Orbvmean	4	3	0	0	0	1	-	-	0	-
	SSTanom	5	2	1	0	0	2	-	-	0	-
	SSTwint	11	6	1	0	0	3	-	-	1	-
	SSTgrad	3	1	0	0	0	2	-	-	0	-
	SSTanamp	7	3	1	1	0	2	-	-	0	-
	Radmean	10	4	1	1	0	3	-	-	1	-

(**G**)

Index	Ind. Var	All phyla	Chordates (14)	Echinodermata (1)	Arthropoda (3)	Annelida (2)	Mollusca (13)	Brachiopoda (0)	Cnidaria (1)	Ochrophyta (1)	Tracheophyta (0)
h (GLZ)	Lat	12	3	0	2	0	7	-	0	0	-
	Lon	9	5	1	0	0	3	-	0	0	-
	Tidal	3	2	0	0	0	1	-	0	0	-
	Depth	3	3	0	0	0	0	-	0	0	-
	Bedslope	4	2	1	0	0	1	-	0	0	-
	Orbvmean	4	3	0	0	0	1	-	0	0	-
	SSTanom	9	7	1	0	0	1	-	0	0	-
	SSTwint	8	5	1	0	0	1	-	0	1	-
	SSTgrad	4	0	0	1	0	3	-	0	0	-
	SSTanamp	6	3	1	2	0	0	-	0	0	-
	Radmean	11	5	1	1	0	4	-	0	0	-

(**H**)

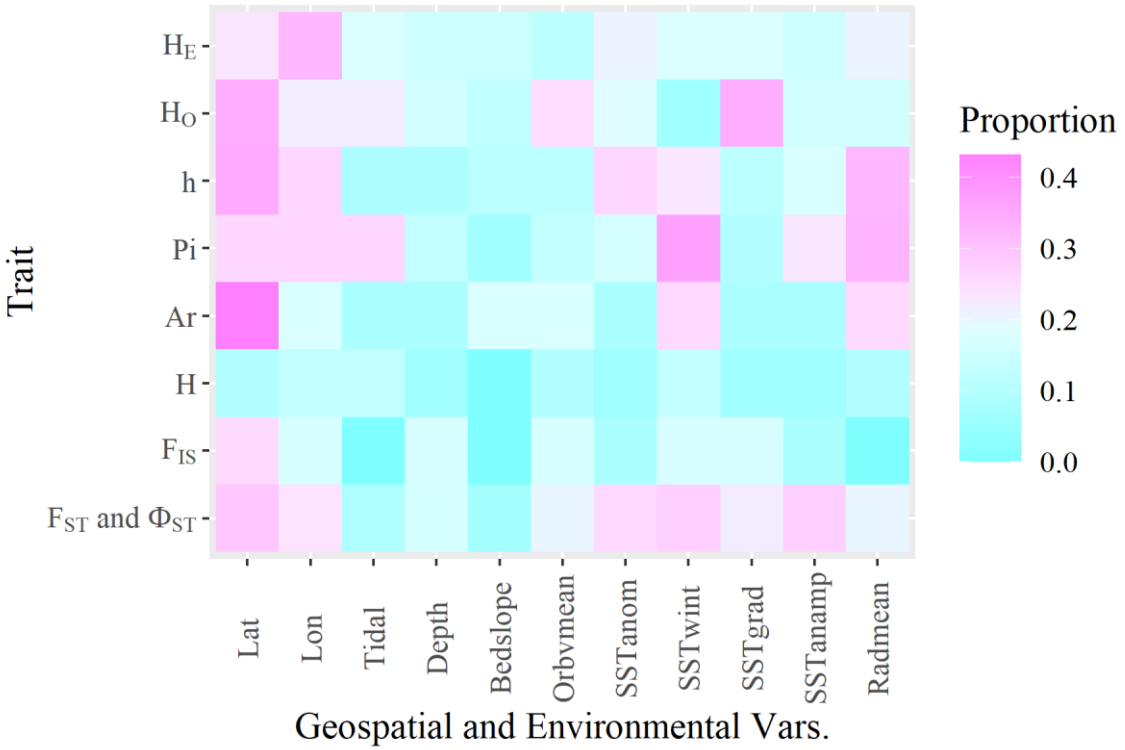
Index	Ind. Var	All phyla	Chordates (10)	Echinodermata (3)	Arthropoda (5)	Annelida (0)	Mollusca (12)	Brachiopoda (1)	Cnidaria (2)	Ochrophyta (0)	Tracheophyta (0)
H_O (GLZ)	Lat	11	4	1	2	-	3	0	1	-	-
	Lon	7	1	0	3	-	3	0	0	-	-
	Tidal	7	0	0	3	-	4	0	0	-	-
	Depth	5	1	1	2	-	0	1	0	-	-
	Bedslope	4	0	1	1	-	1	1	0	-	-
	Orbvmean	8	1	0	1	-	4	1	1	-	-
	SSTanom	6	2	1	0	-	2	1	0	-	-
	SSTwint	2	1	0	0	-	1	0	0	-	-
	SSTgrad	11	4	0	1	-	5	0	1	-	-
	SSTanamp	5	2	0	1	-	1	0	1	-	-
	Radmean	5	0	0	0	-	4	0	1	-	-

(I)

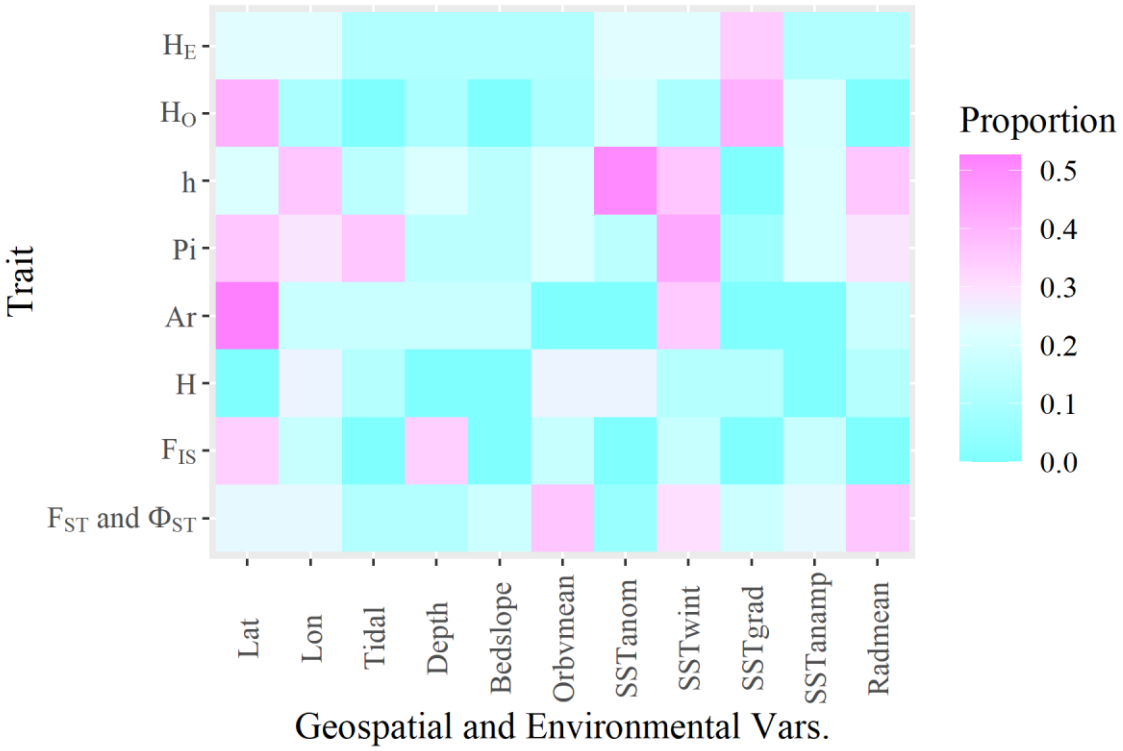
Index	Ind. Var	All phyla	Chordates (9)	Echinodermata (4)	Arthropoda (5)	Annelida (0)	Mollusca (12)	Brachiopoda (3)	Cnidaria (2)	Ochrophyta (0)	Tracheophyta (0)
H_E (GLZ)	Lat	8	2	1	0	-	3	2	0	-	-
	Lon	11	2	3	1	-	4	1	0	-	-
	Tidal	6	1	1	1	-	3	0	0	-	-
	Depth	5	1	1	1	-	1	1	0	-	-
	Bedslope	5	1	0	1	-	2	1	0	-	-
	Orbvmean	4	1	0	1	-	2	0	0	-	-
	SSTanom	7	2	2	1	-	1	1	0	-	-
	SSTwint	6	2	1	0	-	2	1	0	-	-
	SSTgrad	6	3	1	0	-	2	0	0	-	-
	SSTanamp	5	1	1	1	-	2	0	0	-	-
	Radmean	7	1	2	0	-	3	1	0	-	-

Supplementary Material 21. Contribution summary heatmaps for geospatial/environmental variables in explaining genetic variability across all phyla (**A**), Chordata (**B**) and Mollusca (**C**) within New Zealand. Using the GLZ backward removal (stepwise regression) routine.

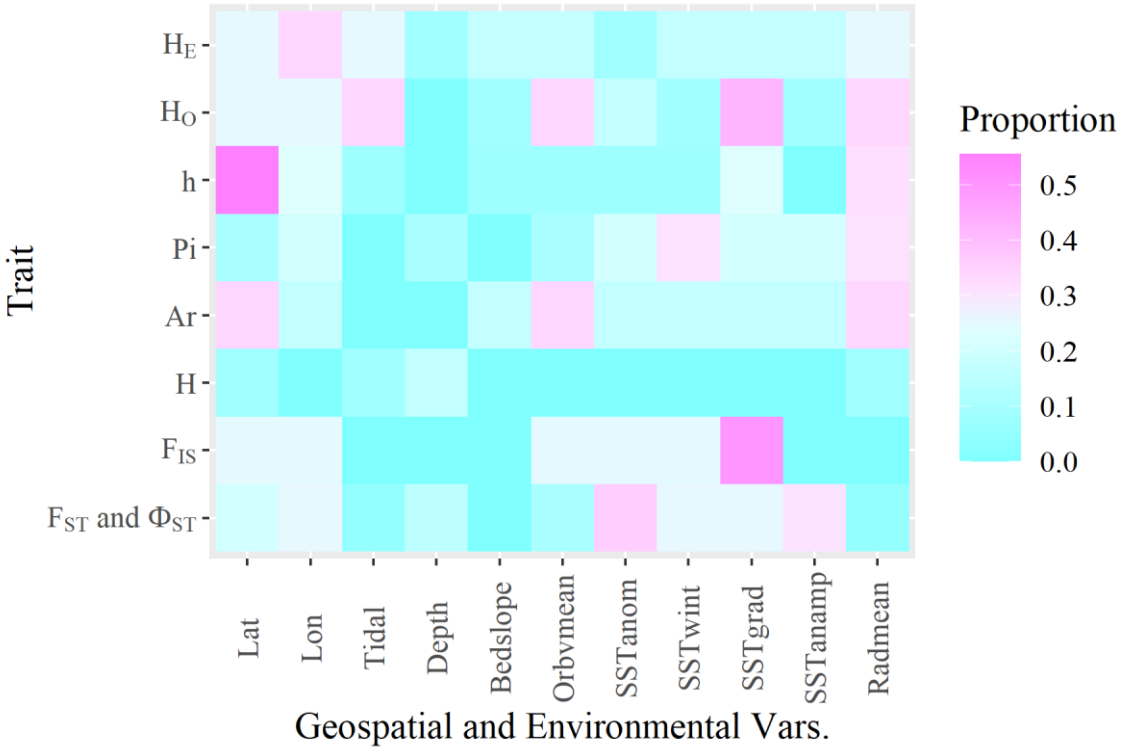
(**A**)



(B)


Trait
H_E
H_O
h
Pi
Ar
H
F_{IS}
F_{ST} and Φ_{ST}
Lat
Lon
Tidal
Depth
Bedslope
Orbvmean
SSTanom
SSTwint
SSTgrad
SSTanamp
Radmean
Geospatial and Environmental Vars.
Proportion
0.5
0.4
0.3
0.2
0.1
0.0

(C)


Trait
H_E
H_O
h
Pi
Ar
H
F_{IS}
F_{ST} and Φ_{ST}
Lat
Lon
Tidal
Depth
Bedslope
Orbvmean
SSTanom
SSTwint
SSTgrad
SSTanamp
Radmean
Geospatial and Environmental Vars.
Proportion
0.5
0.4
0.3
0.2
0.1
0.0

References

Acevedo-Whitehouse, K., Petetti, L., Duignan, P., & Castinel, A. (2009). Hookworm infection, anaemia and genetic variability of the New Zealand sea lion. *Proceedings of the Royal Society B: Biological Sciences*, *276*, 3523–3529. doi: 10.1098/rspb.2009.1001

Adams, D. K., & Flierl, G. R. (2010). Modeled interactions of mesoscale eddies with the East Pacific Rise: Implications for larval dispersal. *Deep Sea Research Part I: Oceanographic Research Papers*, *57*, 1163–1176. doi: 10.1016/j.dsr.2010.06.009

Apte, S., & Gardner, J. P. A. (2001). Absence of population genetic differentiation in the New Zealand greenshell mussel *Perna canaliculus* (Gmelin 1791) as assessed by allozyme variation. Journal of Experimental *Marine Biology and Ecology*, *258*, 173–194. doi: 10.1016/S0022-0981(01)00218-0

Apte, S., & Gardner, J. P. A. (2002). Population genetic subdivision in the New Zealand greenshell mussel (*Perna canaliculus*) inferred from single-strand conformation polymorphism analysis of mitochondrial DNA. *Molecular Ecology*, *11*, 1617–1628. doi: 10.1046/j.1365-294X.2002.01554.x

Apte, S., Star, B., & Gardner, J. P. A. (2003). A comparison of genetic diversity between cultured and wild populations, and a test for genetic introgression in the New Zealand greenshell mussel *Perna canaliculus* (Gmelin 1791). *Aquaculture*, *219*, 193–220. doi: 10.1016/S0044-8486(03)00003-6

Ashe, J. L., & Wilson, A. B. (2020). Navigating the southern seas with small fins: Genetic connectivity of seahorses (*Hippocampus abdominalis*) across the Tasman Sea. *Journal of Biogeography*, *47*, 207–219. doi: 10.1111/jbi.13733

Ayers, K. L., & Waters, J. M. (2005). Marine biogeographic disjunction in central New Zealand. *Marine Biology*, *147*, 1045–1052. doi: 10.1007/s00227-005-1632-7

Azid, A., Juahir, H., Amran, M. A., Suhaili, Z., Osman, M. R., Muhamad, A., Muhamad, A., Abidin, I. Z., Sulaiman, N. H., & Saudi, A. S. M. (2015). Spatial Air Quality Modelling Using Chemometrics Techniques: A Case Study in Peninsular Malaysia. *Malaysian Journal of Analytical Sciences*, *19*, 1415–1430.

Baba, K., Fujita, Y., Wehrtmann, I. S., & Scholtz, G. (Ed.). (2011). *Chapter 4. Developmental biology of squat lobsters. The Biology of Squat Lobsters*. Victoria, Australia: Museum Victoria, Australian Museum,

Baird, S. J. (2011). New Zealand fur seals – summary of current knowledge. New Zealand Aquatic Environment and Biodiversity Report No. 72, Ministry of Fisheries, Wellington, New Zealand.

Baker, C. S., Hamner, R. M., Cooke, J., Heimeier, D., Vant, M., Steel, D., & Constantine, R. (2013). Low abundance and probable decline of the critically endangered Maui's dolphin estimated by genotype capture-recapture. *Animal Conservation, 16*, 224–233.

Baker, C. S., Steel, D., Hamner, R. M., Hickman, G., Boren, L., Arlidge, W., & Constantine, R. (2016). *Estimating the abundance and effective population size of Māui dolphins using microsatellite genotypes in 2015-16, with retrospective matching to 2001-16*. Auckland, New Zealand: Department of Conservation.

Baldanzi, S., McQuaid, C. D., Cannicci, S., & Porri, F. (2013). Environmental Domains and Range-Limiting Mechanisms: Testing the Abundant Centre Hypothesis Using Southern African Sandhoppers. (Research Article). *PLoS ONE, 8*, e54598. doi: 10.1371/journal.pone.0054598

Barber, P. H., Palumbi, S. R., Erdmann, M. V., & Moosa, M. K. (2000). A marine Wallace's line?. *Nature*, *406*, 692–693. doi: 10.1038/35021135

Barber, P. H., Palumbi, S. R., Erdmann, M. V., & Moosa, M. K. (2002). Sharp genetic breaks among populations of *Haptosquilla pulchella* (Stomatopoda) indicate limits to larval transport: patterns, causes, and consequences. *Molecular ecology*, *11*, 659–674. doi: 10.1046/j.1365-294X.2002.01468.x

Barker, A. M., Nosal, A. P., Lewallen, E. A., & Burton, R. S. (2015). Genetic structure of leopard shark (*Triakis semifasciata*) populations along the Pacific coast of North America. *Journal of experimental marine biology and ecology*, *472*, 151–157. doi: 10.1016/j.jembe.2015.06.020

Barker, M. F. (1977). Observations on the settlement of the brachiolaria larvae of *Stichaster australis* (Verill) and *Coscinasterias calamaria* (Gray) (Echinodermata: Asteroidea) in the laboratory and on the shore. *Journal of experimental marine biology and ecology*, *30*, 95–108. doi: 10.1016/0022-0981(77)90030-2

Bayne, B. L. (1966). Growth and the delay of metamorphosis of the larvae of *Mytilus edulis* (L.): Bayne B. L., 1965. Ophelia, Univ. Copenhagen, 2, 1–47. *Deep-sea research and oceanographic abstracts*, *13*, 117–118.

Beentjes, M. P., & Carbines, G. D. (2005). Population structure and relative abundance of blue cod (*Parapercis colias*) off Banks Peninsula and in Dusky Sound, New Zealand. *New Zealand journal of marine and freshwater research*, *39*, 77–90. doi: 10.1080/00288330.2005.9517293

Bejder, L. & Dawson, S. (2001) Abundance, residency, and habitat utilisation of Hector's dolphins (*Cephalorhynchus hectori*) in Porpoise Bay, New Zealand. *New Zealand journal of marine and freshwater research*, *35*, 277–287. doi: 10.1080/00288330.2001.9516998

Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., Rochette, R., & Bernatchez, L. (2016). Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular ecology*, *25*, 5073–5092. doi: 10.1111/mec.13811

Blackmore, D. T. (1969). Studies of Patella vulgata L. I. Growth, reproduction and zonal distribution. *Journal of experimental marine biology and ecology*, *3*, 200–213. doi: 10.1016/0022-0981(69)90018-5

Blackwell, R. G. (2006). Abundance and size composition of blue cod in the Marlborough Sounds and Tasman Bay, September - October 2004, National Institute of Water and Atmospheric Research. Final Research Report BCO 2033-01. Wellington, New Zealand.

Boessenkool, S., Star, B., Waters, J. M., & Seddon, P. J. (2009). Multilocus assignment analyses reveal multiple units and rare migration events in the recently expanded yellow-eyed penguin (*Megadyptes antipodes*). *Molecular Ecology*, *18*, 2390–2400. doi: 10.1111/j.1365-294X.2009.04203.x

Boessenkool, S., Star, B., Seddon, P. J., & Waters, J. M. (2010). Temporal genetic samples indicate small effective population size of the endangered yellow-eyed penguin. *Conservation Genetics*, *11*, 539–546. doi: 10.1007/s10592-009-9988-8

Bohonak, A. J. (1999). Dispersal, Gene Flow, and Population Structure. *The Quarterly review of biology*, *74*, 21–45.

Booth, J.D. (2006) Jasus species. In B. F. Phillips (Ed.) *Lobsters: Biology, Management, Aquaculture and Fisheries* (pp. 340-358). Oxford, UK: Blackwell Scientific Publications.

Bors, E. K., Rowden, A. A., Maas, E. W., Clark, M. R., & Shank, T. M. (2012). Patterns of Deep-Sea Genetic Connectivity in the New Zealand Region: Implications for Management of Benthic Ecosystems (New Zealand Deep-Sea Genetic Connectivity). *Plos One*, *7*, e49474. doi: 10.1371/journal.pone.0049474

Bradbury, I. R., Laurel, B., Snelgrove, P. V. R., Bentzen, P., & Campana, S. E. (2008). Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. *Proceedings of the Royal Society. B, Biological sciences*, *275*, 1803–1809. doi: 10.1098/rspb.2008.0216

Brager, S., & Schneider, K. (1998). Near-shore distribution and abundance of dolphins along the West Coast of the South Island, New Zealand. *New Zealand journal of marine and freshwater research*, *32*, 105–112. doi: 10.1080/00288330.1998.9516809

Brasher, D. J., Ovenden, J. R., Booth, J. D., & White, R. W. G. (1992). Genetic subdivision of Australian and New Zealand populations of *Jasus verreauxi* (Decapoda: Palinuridae)-preliminary evidence from the mitochondrial genome. *New Zealand Journal of Marine and Freshwater Research*, *26*, 53–58. doi: 10.1080/00288330.1992.9516499

Briggs, J. C. (2003). Marine centres of origin as evolutionary engines. *Journal of biogeography*, *30*, 1–18.

Briggs, J. C. (2005). The Marine East Indies: Diversity and Speciation. *Journal of biogeography*, *32*, 1517–1522. doi: 10.1111/j.1365-2699.2005.01266.x

Briggs, J. C., & Bowen, B.W. (2013). Marine shelf habitat: biogeography and evolution. *Journal of biogeography*, *40*, 1023–1035. doi: 10.1111/jbi.12082

Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, *124*, 255–279.

Brown, J. H. (1995). *Macroecology*. Chicago, IL: University of Chicago Press.

Brown, J. H., Mehlman, D. W., & Stevens, G. C. (1995). Spatial Variation in Abundance. *Ecology (Durham)*, *76*, 2028–2043. doi: 10.2307/1941678

Brown, J. H., Stevens, G. C., & Kaufman, D. M. (1996). THE GEOGRAPHIC RANGE: Size, Shape, Boundaries, and Internal Structure. *Annual Review of Ecology and Systematics*, *27*, 597–623. doi: 1 0.1146/annurev.ecolsys.27.1.597

Buchanan, J. (2011). *Molecular Studies of New Zealand Fucales: Phylogeography, Phylogeny and Taxonomy in Carpophyllum and Cystophora (Phaeophyceae)* (PhD thesis). Victoria University of Wellington, Wellington, New Zealand.

Buchanan, J., & Zuccarello, G. C. (2012). Decoupling of short- and long-distance dispersal pathways in the endemic New Zealand seaweed *Carpophyllum maschalocarpum* (Phaeophyceae, Fucales). *Journal of Phycology, 48*, 518–529. doi: 10.1111/j.1529-8817.2012.01167.x

Buckley, B. A., Owen, M. E., & Hofmann, G. E. (2001). Adjusting the thermostat: the threshold induction temperature for the heatshock response in intertidal mussels (genus *Mytilus*) changes as a function of thermal history. *Journal of experimental biology, 204*, 3571–3579. doi: 10.1242/jeb.204.20.3571

Buroker, N., Chanley, P., Cranfield, H., & Dinamani, P. (1983). Systematic status of two oyster populations of the genus Tiostrea from New Zealand and Chile. *Marine Biology, 77*, 191–200. doi: 10.1007/BF00396317

Byrne, M., & Barker, M. F. (1991). Embryogenesis and Larval Development of the Asteroid Patiriella regularis Viewed by Light and Scanning Electron Microscopy. *The Biological Bulletin*, 180, 332–345.

Campbell, H., & Hutching, G. (2007). *In search of ancient New Zealand*. North Shore, NZ: Penguin Books.

Campbell, H. (2013). *The Zealandia Drowning Debate Did New Zealand Sink Beneath the Waves?* Wellington, NZ: Bridget Williams Books.

Carson, H. L. (1959). Genetic conditions which promote or retard the formation of species. *Cold Spring Harbour Symposia on Quantitative Biology, 24*, 87–105. doi: 10.1101/SQB.1959.024.01.010

Cawthorn, M. W., & Martin W. (1993). Census and population estimation of Hooker's sea lion at the Auckland Islands, December 1992 - February 1993. Department of Conservation Technical Series No.2. Wellington, New Zealand.

Chambers, G. K., & MacAvoy, E. S. (2000). Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology, Part B, 126*, 455–476. doi: 10.1016/S0305-0491(00)00233-9

Chave, J., & Bascompte, J. (2013). The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecology letters*, *16*, 4–16. doi: 10.1111/ele.12048

Childerhouse, S. & Gales, N. (1998). Historical and modern distribution and abundance of the New Zealand sea lion *Phocarctos hookeri*. *New Zealand journal of zoology*, *25*, 1–16. doi: 10.1080/03014223.1998.9518131

Childerhouse, S., Gibbs, N., McAlister, G., McConkey, S., McConnell, H., McNally, N., & Sutherland, D. (2005). Distribution, abundance and growth of New Zealand sea lion *Phocarctos hookeri* pups on Campbell Island. *New Zealand journal of marine and freshwater research*, *39*, 889–898. doi: 10.1080/00288330.2005.9517360

Chilvers, B., Robertson, B., Wilkinson, I., & Duignan, P. (2007). Growth and survival of New Zealand sea lions, *Phocarctos hookeri*: birth to 3 months. *Polar Biology*, *30*, 459–469. doi: 10.1007/s00300-006-0203-9

Chiswell, S. M., & Sutton, P. J. H. (2015). Drifter- and float-derived mean circulation at the surface and 1000 m in the New Zealand region. *New Zealand journal of marine and freshwater research*, *49*, 259–277. doi: 10.1080/00288330.2015.1008522

Chiswell, S. M., Zeldis, J. R., Hadfield, M. G., & Pinkerton, M. H. (2017). Wind-driven upwelling and surface chlorophyll blooms in Greater Cook Strait. New Zealand journal of marine and freshwater research, 51, 465–489. doi: 10.1080/00288330.2016.1260606

Collins, C. J., Fraser, C. I., Ashcroft, A., & Waters, J. M. (2010). Asymmetric dispersal of southern bull-kelp (*Durvillaea antarctica*) adults in coastal New Zealand: testing an oceanographic hypothesis. *Molecular Ecology, 19*, 4572–4580. doi: 10.1111/j.1365-294X.2010.04842.x

Collins, C. J., Chilvers, B. L., Taylor, M., & Robertson, B. C. (2016). Historical population size of the threatened New Zealand sea lion *Phocarctos hookeri*. *Journal of Mammalogy*, *97*, 436–443. doi:10.1093/jmammal/gyv187

Collins, C. J., Chilvers, B. L., Osborne, A., Taylor, M., & Robertson, B. C. (2017). Unique and isolated: population structure has implications for management of the endangered New Zealand sea lion. *Conservation Genetics*, *18*, 1177–1189. doi: 10.1007/s10592-017-0969-z

Constable, H. B. (2014). *Population structure, temporal stability and seascape genetics of two endemic New Zealand Pleuronectiformes, Rhombosolea plebeia (sand flounder) and R.*

leporina (yellowbelly flounder) (PhD thesis). Victoria University of Wellington, Wellington, New Zealand.

Crawford, C. M. (1984). Preliminary results of experiments on the rearing of Tasmanian flounders, *Rhombosolea tapirina* and *Ammotretis rostratus*. *Aquaculture, 42*, 75–81. doi: 10.1016/0044-8486(84)90315-6

Cumming, R. A., Nikula, R., Spencer, H. G., & Waters, J. M. (2014). Transoceanic genetic similarities of kelp-associated sea slug populations: long-distance dispersal via rafting? *Journal of biogeography, 41*, 2357–2370. doi: 10.1111/jbi.12376

Dawson, S. M. (Ed.). (2009). *Cephalorhynchus dolphins C. heavisidii, C. eutropia, C. hectori, and C. commersonii. Encyclopedia of marine mammals*. London, UK: Academic Press.

Dawson, S. M., Slooten, E., DuFresne, S., Wade, P., & Clement, D. (2004). Small-boat surveys for coastal dolphins: line-transect surveys for Hector's dolphins (*Cephalorhynchus hectori*). *Fishery bulletin, 102*, 441–451.

Department of Conservation and Ministry of Fisheries (2005). Marine Protected Areas: Policy and Implementation Plan. Wellington, New Zealand.

Derksen, S., & Keselman, H. J. (1992). Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology, 45*, 265–282. doi: 10.1111/j.2044-8317.1992.tb00992.x

Dexter, D. M. (1977). Natural history of the Pan-American sand beach isopod *Excirolana braziliensis* (Crustacea: Malacostraca). *Journal of Zoology, 183*, 103–109. doi: 10.1111/j.1469-7998.1977.tb04175.x

Dickie, G., & Dawson, S.M. (2003). Age, growth, and reproduction in new zealand fur seals. *Marine mammal science, 19*, 173–185. doi: 10.1111/j.1748-7692.2003.tb01100.x

Dix, T. G. (1969). Larval life span of the *echinoid evechinus chloroticus* (val.). *New Zealand journal of marine and freshwater research, 3*, 13–16. doi: 10.1080/00288330.1969.9515273

Dixon, A. L., Herlihy, C. R., & Busch, J. W. (2013). Demographic and population-genetic tests provide mixed support for the abundant centre hypothesis in the endemic plant *Leavenworthia stylosa*. *Molecular ecology, 22*, 1777–1791. doi: 10.1111/mec.12207

Dobson, A. J., & Barnett, A. G. (2008). *An introduction to generalized linear models* (3rd ed.). Boca Raton, FL: CRC Press.

Dodd, J. M. (1955). Artificial fertilisation, larval development and metamorphosis in *Patella vulgata* L. and *Patella coerulea* L. *Pubblicazioni della Stazione Zoologica di Napoli*, *29*, 172–185.

Dohner, M., Phillips, N. E., & Ritchie, P. A. (2018). Fine-scale genetic structure across a New Zealand disjunction for the direct-developing intertidal whelk *Cominella maculosa* (Gastropoda: Buccinidae). *Biological Journal of the Linnean Society*, *123*, 593–602. doi: 10.1093/biolinnean/blx167

Dunmore, R. A. (2006). *Demography of early life stages of habitat-forming intertidal fucoid algae* (PhD thesis). University of Canterbury, Christchurch, New Zealand.

Dunmore, R. A., & Schiel, D. R. (2000). Reproduction in the intertidal limpet *Cellana ornata* in southern New Zealand. *New Zealand journal of marine and freshwater research*, *34*, 653–660. doi: 10.1080/00288330.2000.9516966

Dussex, N., Robertson, B. C., Salis, A. T., Kalinin, A., Best, H., & Gemmell, N. J. (2016). Low Spatial Genetic Differentiation Associated with Rapid Recolonization in the New Zealand Fur Seal *Arctocephalus forsteri. The Journal of heredity*, *107*, 581–592. doi:10.1093/jhered/esw056

Dutton, J. M., & Hofmann, G. E. (2009). Biogeographic variation in Mytilus galloprovincialis heat shock gene expression across the eastern Pacific range. *Journal of experimental marine biology and ecology*, *376*, 37–42. doi: 10.1016/j.jembe.2009.06.001

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular ecology*, *17*, 1170–1188. doi: 10.1111/j.1365-294X.2007.03659.x

Efford, M., Darby, J., & Spencer, N. (1996). *Population studies of yellow-eyed penguins. 1993-94 Progress Report. Science for Conservation: 22.* Wellington, New Zealand: Department of Conservation.

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

Enquist, B., Jordan, M., & Brown, J. (1995). Connections between ecology, biogeography, and paleobiology: Relationship between local abundance and geographic distribution in fossil and recent molluscs. *Evolutionary Ecology*, *9*, 586–604. doi: 10.1007/BF01237657

Fell, H. B. (1945). A Revision of the Current Theory of Echinoderm Embryology. *Transactions of the Royal Society of New Zealand*, *75*, 73–101.

Fenberg, P. B., & Rivadeneira, M. M. (2011). Range limits and geographic patterns of abundance of the rocky intertidal owl limpet, *Lottia gigantea*. *Journal of Biogeography*, *38*, 2286–2298. doi: 10.1111/j.1365-2699.2011.02572.x

Fleming, A. M., Dohner, M. M., Phillips, N. E., & Ritchie, P. A. (2018). Genetic connectivity among populations of two congeneric direct-developing whelks varies across spatial scales. *New Zealand Journal of Marine and Freshwater Research*, *52*, 100–117. doi: 10.1080/00288330.2017.1340899

Foll, M., & Gaggiotti, O. (2006). Identifying the environmental factors that determine the genetic structure of populations. *Genetics*, *174*, 875–891. doi: 10.1534/genetics.106.059451

Francis, M P. (1988). Movement patterns of rig (*Mustelus lenticulatus*) tagged in southern New Zealand. *New Zealand Journal of Marine and Freshwater Research*, *22*, 259–272. doi: 10.1080/00288330.1988.9516298

Frusin, A. (1982). *Electrophoretic study of some paua (Haliotis iris) proteins* (MSc thesis). Victoria University of Wellington, Wellington, New Zealand.

Galindo, H. M., Pfeiffer-Herbert, A. S., McManus, M. A., Chao, Y. I., Chai, F. E. I., & Palumbi, S. R. (2010). Seascape genetics along a steep cline: using genetic patterns to test predictions of marine larval dispersal. *Molecular ecology*, *19*, 3692–3707. doi: 10.1111/j.1365-294X.2010.04694.x

Garden, C. J., Currie, K., Fraser, C. I., & Waters, J. M. (2014). Rafting dispersal constrained by an oceanographic boundary. *Marine ecology. Progress series*, *501*, 297–302. doi: 10.3354/meps10675

Gardner, J. P. A., Bell, J. J., Constable, H. B., Hannan, D., Ritchie, P. A., & Zuccarello, G. C. (2010). Multi-species coastal marine connectivity: a literature review with recommendations for further research. Ministry of Fisheries. Wellington, New Zealand.

Gardner, J. P. A., Eyles, R. F., & Pande, A. (1996a). Biochemical-genetic variation in a wild and a cultured population of the greenshell mussel, *Perna canaliculus*. *New Zealand Journal of Marine and Freshwater Research*, *30*, 435–441. doi: 10.1080/00288330.1996.9516731

Gardner, J. P. A., Pande, A., Eyles, R. F., & Wear, R. G. (1996b). Biochemical genetic variation among populations of the greenshell mussel, *Perna canaliculus*, from New Zealand: preliminary findings. *Biochemical Systematics and Ecology, 24*, 763–774. doi: 10.1016/S0305-1978(96)00059-2

Gardner, J. P. A., & Kathiravetpillai, G. (1997). Biochemical genetic variation at a leucine aminopeptidase (LAP) locus in blue (*Mytilus galloprovincialis*) and greenshell (*Perna canaliculus*) mussel populations along a salinity gradient. *Marine Biology*, *128*, 619–625. doi: 10.1007/s002270050128

Gardner, J. P. A., & Palmer, N. L. (1998). Size-dependent, spatial and temporal genetic variation at a leucine aminopeptidase (LAP) locus among blue mussel (*Mytilus galloprovincialis*) populations along a salinity gradient. *Marine biology*, *132*, 275–281. doi: 10.1007/s002270050393

Gardner, J. P. A., Zbawicka, M., Westfall, K. M., & 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. doi: 10.1111/gcb.13332

Garrett, F. K. (1994). *Reproduction and population genetics of the New Zealand brooding brittle star Ophiomyxa brevirima*. (MSc thesis). University of Otago, Dunedin, New Zealand.

Garrett, F. K., Mladenov, P. V., & Wallis, G.P. (1997). Evidence of amictic reproduction in the brittle-star *Ophiomyxa brevirima. Marine Biology*, *129*, 169–174.

Gaston, K. J. (2004). The Structure and Dynamics of Geographic Ranges. Oxford Series in Ecology and Evolution (Book Review). *The Quarterly Review of Biology*, *79*, 100–101.

Gebbie, C. L. (2014). *Population genetic structure of New Zealand blue cod (Parapercis colias) based on mitochondrial and microsatellite DNA markers* (MSc thesis). Victoria University of Wellington, Wellington, New Zealand.

Gemmell, M., Trewick, S. A., Crampton, J., Vaux, F., Hills, S., Daly, E. E., Marshall, B. A., Beu, A. G., & Morgan-Richards, M. (2018). Genetic structure and shell shape variation within a

rocky shore whelk suggest both diverging and constraining selection with gene flow. *Biological Journal Of The Linnean Society*, *125*, 827–843. doi: 10.1093/biolinnean/bly142

Gilman, S. E. (2006). The Northern Geographic Range Limit of the Intertidal Limpet Collisella scabra: A Test of Performance, Recruitment, and Temperature Hypotheses. *Ecography*, *29*, 709–720.

Goldberg, E. E., Lande, R., Gomulkiewicz, A. E. R., & Whitlock, E. M. C. (2007). Species' Borders and Dispersal Barriers. *The American naturalist*, *170*, 297–304.

Goldstien, S. J. (2005). *Phylogeography of the Cellana limpets of New Zealand: Investigating Barriers to Marine Dispersal and Historical Biogeography* (PhD thesis). University of Canterbury, Christchurch, New Zealand.

Goldstien, S. J., Schiel, D. R., & Gemmell, N. J. (2006). Comparative phylogeography of coastal limpets across a marine disjunction in New Zealand. Molecular *Ecology*, *15*, 3259–3268. doi: 10.1111/j.1365-294X.2006.02977.x

Goodbody, I. (1960). Abbreviated development in a pinnotherid crab. *Nature*, 185, 704–705.

Gordon, D. P., Beaumont, J., Macdiarmid, A., Robertson, D. A., & Ahyong, S. T. (2010). Marine biodiversity of Aotearoa New Zealand *PLoS ONE*, *5*, e10905. doi: 0.1371/journal.pone.0010905

Grant, P. R., & Grant, B. R. (2008). *How and why species multiply: the radiation of Darwin's finches*. Princeton, NJ: Princeton University Press.

Griffiths, H. J., Barnes, D. K. A., Linse, K., & Lambshead, J. (2009). Towards a Generalized Biogeography of the Southern Ocean Benthos. *Journal of biogeography*, *36*, 162–177. doi: 10.1111/j.1365-2699.2008.01979.x

Guo, X. Z., Wei, K. J., Yan, R. J., & Gardner, J. Pronounced mitochondrial DNA and subtle nuclear DNA populations genetic structure in a brooding coastal marine invertebrate (paper in review).

Hamner, R. M., Oremus, M., Stanley, M., Brown, P., Constantine, R., & Scott Baker, C. (2012a). *Estimating the abundance and effective population size of Maui's dolphins using microsatellite genotypes in 2010-11, with retrospective matching to 2001-07*. Auckland, New Zealand: Department of Conservation.

Hamner, R. M., Pichler, F. B., Heimeier, D., Constantine, R., & Baker, C. S. (2012b). Genetic differentiation and limited gene flow among fragmented populations of New Zealand endemic Hector's and Maui's dolphins. *Conservation Genetics*, *13*, 987–1002. doi:10.1007/s10592-012-0347-9

Hannan, D. A. (2014). *Population genetics and connectivity in Paphies subtriangulata and Paphies australis (Bivalvia: Mesodesmatidae)* (PhD thesis). Victoria University of Wellington, Wellington, New Zealand.

Heath, R. A. (1981). Oceanic fronts around southern New Zealand. Deep-sea research. Part A. *Oceanographic research papers*, *28*, 547–560. doi: 10.1016/0198-0149(81)90116-3

Hedrick, P. W. (2005). *Genetics of populations* (3rd ed.). Boston, MA: Jones and Bartlett Publishers.

Hemmer-Hansen, J., Nielsen, E. E., Frydenberg, J., & Loeschcke, V. (2007). Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus* L.). *Heredity*, *99*, 592–600. doi: 10.1038/sj.hdy.6801055

Henderson, I. M. (2009). Optimising recreational harvests of blue cod: The effects of catch-and-release mortality and size selectivity. *Fisheries research*, *99*, 184–195. doi: 10.1016/j.fishres.2009.06.003

Hendry, R. T. (2004). *An assessment of the spatial extent and relative importance of nurseries, and of the genetic structure among nurseries of rig (Mustelus lenticulatus), an endemic New Zealand shark* (MSc thesis). Victoria University of Wellington, Wellington, New Zealand.

Hickey, A. J. R., Lavery, S. D., Hannan, D. A., Baker, C. S., & Clements, K. D. (2009). New Zealand triplefin fishes (family Tripterygiidae): contrasting population structure and mtDNA diversity within a marine species flock. *Molecular Ecology*, *18*, 680–696. doi: 10.1111/j.1365-294X.2008.04052.x

Hidas, E. Z., Ayre, D. J., & Minchinton, T. E. (2010). Patterns of demography for rocky-shore, intertidal invertebrates approaching their geographical range limits: tests of the abundant-centre hypothesis in south-eastern Australia. *Marine and freshwater research, 61*, 1243–1251. doi: 10.1071/MF09317

Hill, G. E. (2019). *Mitonuclear ecology*. Oxford, UK: Oxford University Press.

Hinnendael, F. J. (2008). *Phylogeny of Worldwide Varunidae and Phylogeography of New Zealand Hemigrapsus* (MSc thesis). University of Auckland, Auckland, New Zealand.

Hodgson, A. N. (1999). The Biology of Siphonariid Limpets (Gastropoda: Pulmonata). *Oceanography and Marine Biology: An Annual Review, 27*, 245–314.

Holderegger, R., & Wagner, H. H. (2006). A brief guide to Landscape Genetics. *Landscape ecology, 21*, 793–796. doi: 10.1007/s10980-005-6058-6

Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nature Reviews Genetics, 10,* 639. doi: 10.1038/nrg2611

Hume, T. M., Snelder, T., Weatherhead, M., & Liefting, R. (2007). A controlling factor approach to estuary classification. *Ocean & coastal management, 50*, 905–929. doi: 10.1016/j.ocecoaman.2007.05.009

IUCN 2020. The IUCN Red List of Threatened Species. Version 2020-3. https://www.iucnredlist.org. Downloaded on [14.01.2021].

Jones, T. C. (2004). *Population Genetics of Zostera capricorni (a seagrass) as estimated by* RAPDs (MSc thesis). University of Waikato, Hamilton, New Zealand.

Jones, T. C., Gemmill, C., & Pilditch, C. (2008). Genetic variability of New Zealand seagrass (*Zostera muelleri*) assessed at multiple spatial scales. *Aquatic Botany, 88*, 39–46. doi:10.1016/j.aquabot.2007.08.017

Jensen, M. P., Dalleau, M., Gaspar, P., Lalire, M., Jean, C., Ciccione, S., Mortimer, J. A., Quillard, M., Taquet, C., Wamukota, A., Leroux, G., & Bourjea, J. (2020). Seascape Genetics and the Spatial Ecology of Juvenile Green Turtles. *Genes, 11*, 278. doi: 10.3390/genes11030278

Jokiel, P., & Martinelli, F. J. (1992). The Vortex Model of Coral Reef Biogeography. *Journal of biogeography, 19*, 449–458. doi: 10.2307/2845572

Keeney, D. B., Bryan-Walker, K., King, T. M., & Poulin, R. (2008). Local variation of within-host clonal diversity coupled with genetic homogeneity in a marine trematode. *Marine Biology, 154*, 183–190. doi: 10.1007/s00227-008-0914-2

Kohn, Y., & Clements, K. (2011). Pelagic larval duration and population connectivity in New Zealand triplefin fishes (Tripterygiidae). *Environmental Biology of Fishes, 91*, 275–286. doi: 10.1007/s10641-011-9777-3

Laing, A., & Chiswell, S. (2003). The Ocean Medium. In N. Andrew & M. Francis (Eds.), *The living reef: the ecology of New Zealand's rocky reefs* (pp. 24–31). Nelson, NZ: Craig Potton Publishing.

Leathwick, J., Image, K., Snelder, T., Weatherhead, M., & Wild, M. (2004). Definition and testing of Marine Environment classifications of New Zealand's Exclusive Economic Zone and the Hauraki Gulf. National Institute of Water and Atmospheric Research Report CHC2004. Wellington, New Zealand.

Lento, G. M., Mattlin, R. H., Chambers, G. K., & Baker, C. S. (1994). Geographic distribution of mitochondrial cytochrome b DNA haplotypes in New Zealand fur seals (*Arctocephalus forsteri*). *Canadian Journal of Zoology, 72*, 293–299.

Levy, E., Byrne, M., Coates, D. J., Macdonald, B. M., McArthur, S., & van Leeuwen, S. (2016). Contrasting Influences of Geographic Range and Distribution of Populations on Patterns of Genetic Diversity in Two Sympatric *Pilbara Acacias*. *PloS one*, *11*, e0163995–e0163995. doi: 10.1371/journal.pone.0163995

Lidgard, C. W. (2001). *Evaluating the population genetic structure of the New Zealand cockle Austrovenus stutchburyi using allozyme electrophoresis* (MSc thesis). University of Waikato, Hamilton, New Zealand.

Liggins, L., Treml, E. A., & Riginos, C. (2013). Taking the Plunge: An Introduction to Undertaking Seascape Genetic Studies and using Biophysical Models. *Geography compass*, *7*, 173–196. doi: 10.1111/gec3.12031

Liggins, L., Treml, E. A., & Riginos, C. (2019). Seascape Genomics: Contextualizing Adaptive and Neutral Genomic Variation in the Ocean Environment. In M. Oleksiak & O. Rajora (Eds.), *Population Genomics: Marine Organisms. Population Genomics*. Cham: Springer International Publishing. doi: /10.1007/13836_2019_68

Lindberg, D. R. (1991). Marine Biotic Interchange Between the Northern and Southern Hemispheres. *Paleobiology, 17*, 308–324.

Lindsey, J. K. (1997). *Applying Generalized Linear Models* (1st ed.). New York, NY: Springer New York.

Little, R. J. A., & Rubin, D. B. (2020). *Statistical analysis with missing data* (3rd ed.). Hoboken, NJ: Wiley.

Lourie, S. A., & Vincent, A. C. J. (2004). A Marine Fish Follows Wallace's Line: The Phylogeography of the Three-Spot Seahorse (*Hippocampus trimaculatus*, Syngnathidae, Teleostei) in Southeast Asia. Journal of biogeography, 31, 1975–1985. doi: 10.1111/j.1365-2699.2004.01153.x

Lutz, R., Jablonski, D., Rhods, D., & Turner, R. (1980). Larval Dispersal of a Deep-Sea Hydrothermal Vent Bivalve From the Galapagos *Rift. Marine biology*. *Berlin, Heidelberg*, *57*, 127–133. doi: 10.1007/BF00387378

Martorelli, S. R., Fredensborg, B. L., Mouritsen, K. N., & Poulin, R. (2004). Description and Proposed Life Cycle of *Maritrema novaezealandensis* n. sp. (Microphallidae) Parasitic in Red-Billed Gulls, *Larus novaehollandiae scopulinus*, from Otago Harbor, South Island, New Zealand. *The Journal of Parasitology*, *90*, 272–277. doi: 10.1645/GE-3254

MacDiarmid, A., McKenzie, A., Sturman, J., Beaumont, J., Mikaloff-Fletcher, S., & Dunne, J. (2012). Assessment of anthropogenic threats to New Zealand marine habitats. New Zealand Aquatic Environment and Biodiversity Report No. 93. Ministry of Agriculture and Forestry. Wellington, New Zealand.

Martinez-Meyer, E., Diaz-Porras, D., Peterson, A. T., & Yanez-Arenas, C. (2012). Ecological niche structure and rangewide abundance patterns of species. *Biology Letters*, *9*, 20120637. doi:10.1098/rsbl.2012.0637

Massaro, M., & Blair, D. (2003). Comparison of population numbers of yellow-eyed penguins, *Megadyptes antipodes*, on Stewart Island and on adjacent cat-free islands. *New Zealand journal of ecology*, *27*, 107–113.

Mcconkey, S.D., Mcconnell, H., Lalas, C., Heinrich, S., Mcnally, A., Parker, E., Borofsky, C., Schimanski, K., Mcintosh, G., & Gludmerer, N. (2002). A Northward Spread In The Breeding Distribution Of The New Zealand Sea Lion (*Phocartos Hooeri*). *Australian mammalogy*, *24*, 97–106. doi: 10.1071/AM02097

McNally, N. (2001). *New Zealand sea lion abundance, demographics and movements in southern New Zealand* (MSc thesis). University of Otago, Dunedin, New Zealand.

Millar, R. H., & Hollis, P. J. (1963). Abbreviated Pelagic Life of Chilean and New Zealand Oysters. *Nature*, *197*, 512.

Miller, K. J. (1997). Genetic structure of black coral populations in New Zealand's fiords. *Marine Ecology Progress Series, 161*, 123–132. doi:10.3354/meps161123

Miller, K. J. (1998). Short-distance dispersal of black coral larvae: inference from spatial analysis of colony genotypes. *Marine ecology. Progress series (Halstenbek), 163*, 225–233. doi: 10.3354/meps171225

Miller, K. J., & Mundy, C. (2003). Rapid settlement in broadcast spawning corals: implications for larval dispersal. *Coral Reefs, 22*, 99–106. doi: 10.1007/s00338-003-0290-9

Miller, K. J., Mundy, C. N., & Chadderton, W. L. (2004). Ecological and genetic evidence of the vulnerability of shallow-water populations of the stylasterid hydrocoral *Errina novaezelandiae* in New Zealand's fiords. *Aquatic Conservation: Marine and Freshwater Ecosystems, 14*, 75–94. doi: 10.1002/aqc.597

Mladenov, P. V, Allibone, R. M., & Wallis, G. P. (1997). Genetic differentiation in the New Zealand sea urchin *Evechinus chloroticus* (Echinodermata: Echinoidea). *New Zealand Journal of Marine and Freshwater Research, 31*, 261–269. doi: 10.1080/00288330.1997.9516763

Moore, P. J., Fletcher, D., & Amey, J. (2001). Population estimates of Yellow-eyed Penguins, *Megadyptes antipode*s, on Campbell Island, 1987-98. *Emu, 101*, 225–236. doi: 10.1071/MU00037

Morales, H. E., Pavlova, A., Joseph, L., & Sunnucks, P. (2015). Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. *Molecular ecology, 24*, 2820–2837. doi: 10.1111/mec.13203

Muller, C. G., Chilvers, B. L., French, R. K., Hiscock, J. A., & Battley, P. F. (2020). Population estimate for yellow-eyed penguins (*Megadyptes antipodes*) in the subantarctic Auckland Islands, New Zealand. *Notornis, 67*, 299–319.

del Mundo, G. M. (2009). *Phylogenetic and Phylogeographic Study of the New Zealand Endemic Sea Tunicate Cnemidocarpa nisiotis* (MSc thesis). University of Canterbury, Christchurch, New Zealand.

Mundry, R., & Nunn, C. L. (2009). Stepwise model fitting and statistical inference: turning noise into signal pollution. *The American naturalist, 173*, 119–123.

Mustonen, V., & Lässig, M. (2009). From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends in genetics, 25*, 111–119. doi: 10.1016/j.tig.2009.01.002

Nagel, M., Sewell, M., & Lavery, S. (2015). Differences in population connectivity of a benthic marine invertebrate *Evechinus chloroticus* (Echinodermata: Echinoidea) across large and small spatial scales. *Conservation Genetics, 16,* 965–978. doi: 10.1007/s10592-015-0716-2

Ntuli, N. N., Nicastro, K. R., Zardi, G. I., Assis, J., Mcquaid, C. D., & Teske, P. R. (2020). Rejection of the genetic implications of the "Abundant Centre Hypothesis" in marine mussels. *Scientific reports, 10,* 604. doi: 10.1038/s41598-020-57474-0

de Oliveira, L. R., Meyer, D., Hoffman, J., Majluf, P., & Morgante, J. S. (2009). Evidence of a genetic bottleneck in an El Niño affected population of South American fur seals, *Arctocephalus australis*. *Journal of the Marine Biological Association of the United Kingdom, 89,* 1717–1725. doi:10.1017/S0025315409000162

Oliver, M. K., Lambin, X., Cornulier, T., & Piertney, S. B. (2009). Spatio-temporal variation in the strength and mode of selection acting on major histocompatibility complex diversity in water vole (*Arvicola terrestris*) metapopulations. *Molecular ecology, 18,* 80–92. doi: 10.1111/j.1365-294X.2008.04015.x

Orth, R. J., Luckenbach, M., & Moore, K. A. (1994). Seed Dispersal in a Marine Macrophyte: Implications for Colonization and Restoration. *Ecology, 75,* 1927–1939. doi: 10.2307/1941597

Osborne, A. J., Negro, S. S., Chilvers, B. L., Robertson, B. C., Kennedy, M. A., & Gemmell, N. J. (2016). Genetic Evidence of a Population Bottleneck and Inbreeding in the Endangered New Zealand Sea Lion, *Phocarctos hookeri. Journal of Heredity, 107,* 392–402. doi:10.1093/jhered/esw015

Ostrow, D. G. (2004). *Larval dispersal and population genetic structure of brachiopods in the New Zealand fiords* (PhD thesis). University of Otago, Dunedin, New Zealand.

Ottaway, J. R. (1979). Population ecology of the Intertidal Anemone *Actinia tenebrosa* II Geographical distribution, synonymy, reproductive cycle and fecundity. *Australian Journal of Zoology, 27,* 273–290. doi: 10.1071/ZO9790273

Ovenden, J., Brasher, D., & White, R. (1992). Mitochondrial DNA analyses of the red rock lobster *Jasus edwardsii* supports an apparent absence of population subdivision throughout Australasia. *Marine Biology, 112,* 319–326. doi: 10.1007/BF00702478

Parenti, L. R. (2019). The shape of biogeography: Endemism, maps, and classification of fish distributions in the western Pacific. *Journal of biogeography*, *46*, 1841–1856. doi: 10.1111/jbi.13551

Paxton, H. (1986). Generic revision and relationships of the family Onuphidae (Annelida: Polychaeta). *Records of the Australian Museum*, *38*, 1–74. doi:10.3853/j.0067-1975.38.1986.175

Pearman, W., Wells, S., Silander, O., Freed, N., & Dale, J. (2020). Population structure and dispersal across small and large spatial scales in a direct developing marine isopod. *BioRxiv*, in press.

Pearse, J. S. (Ed.). (1979). *Polyplacophora. Reproduction of Marine Invertebrates V5*. Oxford, UK: Elsevier Science.

Pérez-Barros, P., Thatje, S., Calcagno, J. A., & Lovrich, G. A. (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. doi: 10.1016/j.jembe.2007.06.035

Perrin, C. (2002). *The effects of fiord hydrography and environment on the population genetic structures of the sea urchin Evechinus chloroticus and the sea star Coscinasterias muricata in New Zealand* (PhD thesis). University of Otago, Dunedin, New Zealand.

Perrin, C., & Roy, M. S. (2000). Rapid and efficient identification of microsatellite loci from the sea urchin, *Evechinus chloroticus. Molecular Ecology*, *9*, 2220–2222. doi: 10.1046/j.1365-294X.2000.105335.x

Perrin, C., Wing, S. R., & Roy, M. S. (2004). Effects of hydrographic barriers on population genetic structure of the sea star *Coscinasterias muricata* (Echinodermata, Asteroidea) in the New Zealand fiords. *Molecular Ecology*, *13*, 2183–2195. doi: 10.1111/j.1365-294X.2004.02241.x

Pichler, F. B. (2002). *Genetic assessment of population boundaries and gene exchange in Hector's dolphin*. Wellington, New Zealand: Department of Conservation Science Internal Serie.

Pichler, F. B., & Baker, C. S. (2000). Loss of Genetic Diversity in the Endemic Hector's Dolphin Due to Fisheries- Related Mortality. *Proceedings: Biological Sciences*, *267*, 97–102.

Pichler, F. B., Dawson, S. M., Slooten, E., & Baker, C. S. (1998). Geographic Isolation of Hector's Dolphin Populations Described by Mitochondrial DNA Sequences. *Conservation Biology*, *12*, 676–682.

Ponder, W. F. (1971). A review of the New Zealand recent and fossil species of Buccinulum deshayes (Mollusca: Neogastropoda: Buccinidae). *Journal of the Royal Society of New Zealand, 1*, 231–283. doi: 10.1080/03036758.1971.10423314

Rabone, M., Lavery, S. D., Little, A., & Clements, K. D. (2015). Discordance between nuclear and mitochondrial DNA analyses of population structure in closely related triplefin fishes (*Forsterygion lapillum and F. capito, F. Tripterygiidae*) supports speciation with gene flow. *Marine Biology, 162*, 1611–1624. doi: 10.1007/s00227-015-2697-6

Redfearn, P. (1987). Larval shell development of the northern tuatua, *Paphies subtriangulata* (Bivalvia, Mesodesmatidae). *New Zealand journal of marine and freshwater research, 21*, 65–70. doi: 10.1080/00288330.1987.9516201

Riginos, C., & Liggins, L. (2013). Seascape Genetics: Populations, Individuals, and Genes Marooned and Adrift: Seascape Genetics. *Geography compass, 7*, 197–216. doi: 10.1111/gec3.12032

Rivadeneira, M. M., Hernáez, P., Antonio Baeza, J., Boltaña, S., Cifuentes, M., Correa, C., Cuevas, A., Del Valle, E., Hinojosa, I., Ulrich, N., Valdivia, N., Vásquez, N., Zander, A., & Thiel, M. (2010). Testing the abundant-centre hypothesis using intertidal porcelain crabs along the Chilean coast: linking abundance and life-history variation. *Journal of Biogeography, 37*, 486–498. doi: 10.1111/j.1365-2699.2009.02224.x

Robertson, B. C., & Gemmell, N. J. (2005). *Microsatellite DNA markers for the study of population structure in the New Zealand fur seal Arctocephalus forsteri.* Wellington, New Zealand: Department of Conservation Science Internal Serie.

Ross, P. M. (2011). *The genetic structure of New Zealand's coastal benthos: using the estuarine clam Austrovenus stutchburyi, to determine rates of gene flow and population connectivity* (PhD thesis). University of Waikato, Hamilton, New Zealand.

Ross, P. M., Hogg, I. D., Pilditch, C. A., & Lundquist, C. J. (2009). Phylogeography of New Zealand's coastal benthos. *New Zealand journal of marine and freshwater research, 43*, 1009–1027. doi: 10.1080/00288330.2009.9626525

Ross, P. M., Hogg, I., Pilditch, C., Lundquist, C., & Wilkins, R. (2012). Population Genetic Structure of the New Zealand Estuarine Clam *Austrovenus stutchburyi* (Bivalvia: Veneridae)

Reveals Population Subdivision and Partial Congruence with Biogeographic Boundaries. *Estuaries and Coasts*, *35*, 143–154. doi: 10.1007/s12237-0

Rouse, H., Nichol, S. & Goff, J. (2003). Introduction to the New Zealand coast. In J. R. Goff, S. Nichol and H. L. Rouse (Eds.), *The New Zealand coast, Te Tai o Aotearoa* (pp. 9-23). Palmerston North, New Zealand: Dunmore Press.

Russell, W. (2016). *Population genetics of New Zealand fur seals (Arctocephalus forsteri): Genomic tools for research and management* (MSc thesis). University of Otago, Dunedin, New Zealand.

Sagarin, R. D., & Gaines, S. D. (2002a). Geographical abundance distributions of coastal invertebrates: using one-dimensional ranges to test biogeographic hypotheses. *Journal of Biogeography*, *29*, 985–997. doi: 10.1046/j.1365-2699.2002.00705.x

Sagarin, R. D., & Gaines, S. D. (2002b). The 'abundant centre' distribution: to what extent is it a biogeographical rule? *Ecology Letters, 5*, 137–147. doi: 10.1046/j.1461-0248.2002.00297.x

Sagarin, R. D., Gaines, S. D., & Gaylord, B. (2006). Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in ecology & evolution*, *21*, 524–530. doi: 10.1016/j.tree.2006.06.008

Santini, L., Pironon, S., Maiorano, L., & Thuiller, W. (2019). Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses. *Ecography (Copenhagen*), *42*, 696–705. doi: 10.1111/ecog.04027

Schmidt, P. S., Serrão, E. A., Pearson, G. A., Riginos, C., Rawson, P. D., Hilbish, T. J., Brawley, S. H., Trussell, G. C., Carrington, E., Wethey, D. S., Grahame, J. W., Bonhomme, F., & Rand, D. M. (2008). Ecological Genetics in the North Atlantic: Environmental Gradients and Adaptation at Specific Loci. *Ecology (Durham)*, *89*, S91–S107. doi: 10.1890/07-1162.1

Schnabel, K. E. (1998). *Genotypic and Phenotypic Variation in New Zealand Estuarine Corophiid Amphipods* (MSc thesis). University of Waikato, Hamilton, New Zealand.

Schnabel, K. E., Hogg, I. D., & Chapman, M. A. (2000). Population genetic structures of two New Zealand corophiid amphipods and the presence of morphologically cryptic species: Implications for the conservation of diversity. *New Zealand Journal of Marine and Freshwater Research*, *34*, 637–644. doi: 10.1080/00288330.2000.9516964

Scrosati, R. A., & Freeman, M. J. (2019). Density of intertidal barnacles along their full elevational range of distribution conforms to the abundant-centre hypothesis. *PeerJ*, *7*, e6719–e6719. doi: 10.7717/peerj.6719

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

Selkoe, K. A., Watson, J. R., White, C., Horin, T., Iacchei, M., Mitarai, S., Siegel, D. A., Gaines, S. D., & Toonen, R. J. (2010). Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular ecology*, *19*, 3708–3726. doi: 10.1111/j.1365-294X.2010.04658.x

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

Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and Ecology of Species Range Limits. *Annual review of ecology, evolution, and systematics*, *40*, 415–436. doi: 10.1146/annurev.ecolsys.110308.120317

Shalom, H. Y., Granot, I., Blowes, S. A., Friedlander, A., Mellin, C., Leite Ferreira, C. E., Arias-González, J. E., Kulbicki, M., Floeter, Sergio R Chabanet, Pascale Parravicini, & Valeriano Belmaker, J. (2020) A closer examination of the 'abundant centre' hypothesis for reef fishes. *Journal of Biogeography*, *47*, 2194–2209. doi: 10.1111/jbi.13920

Sharples, J., & Greig, M. J. N. (1998). Tidal currents, mean flows, and upwelling on the north-east shelf of New Zealand. *New Zealand journal of marine and freshwater research*, *32*, 215–231. doi: 10.1080/00288330.1998.9516821

Shears, N. T., Smith, F., Babcock, R. C., Duffy, C. A. J., & Villouta, E. (2008). Evaluation of Biogeographic Classification Schemes for Conservation Planning: Application to New Zealand's Coastal Marine Environment. *Conservation biology*, *22*, 467–481. doi: 10.1111/j.1523-1739.2008.00882.x

Shumway, S., & Parsons, J. (2006). *Scallops: Biology, Ecology and Aquaculture*. Developments in Aquaculture and Fisheries Science, Elsevier.

Siegel, D. A., Mitarai, S., Costello, C. J., Gaines, S. D., Kendall, B. E., Warner, R. R., & Winters, K. B. (2008). The Stochastic Nature of Larval Connectivity among Nearshore Marine Populations. *Proceedings of the National Academy of Sciences - PNAS*, *105*, 8974–8979. doi: 10.1073/pnas.0802544105

Silva, C. N. S., & Gardner, J. P. A. (2016). 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. doi:10.1093/icesjms/fsv240

Silva, C. N. S., Murphy, N. P., Bell, J. J., Green, B. S., Duhamel, G., Cockcroft, A. C., Hernández, C. E., & Strugnell, J.M. (2021). Global drivers of recent diversification in a marine species complex. *Molecular ecology*, *30*, 1223–1236. doi: 10.1111/mec.15780

Sin, F. Y. T., Khoo, J., & Parker, G. R. (1990). Enzyme variation between littoral and sublittoral populations of the green-lipped mussel *Perna canaliculus*. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, *95*, 419–422. doi: 10.1016/0305-0491(90)90096-C

Sköld, M., Wing, S. R., & Mladenov, P. V (2003). Genetic subdivision of a sea star with high dispersal capability in relation to physical barriers in a fjordic seascape. *Marine Ecology Progress Series*, *250*, 163–174.

Smith, H. M. (2012). *Characterisation of the Mitochondrial Genome and the Phylogeographic Structure of Blue Cod (Parapercis colias)* (MSc thesis). Victoria University of Wellington, Wellington, New Zealand.

Smith, P. J. (1988). Biochemical-genetic variation in the green-lipped mussel *Perna canaliculus* around New Zealand and possible implications for mussel farming. *New Zealand Journal of Marine and Freshwater Research*, *22*, 85–90. doi:10.1080/00288330.1988.9516281

Smith, P. J. (1999). Allozyme variation in scampi (Metanephrops challengeri) fisheries around New Zealand. *New Zealand Journal of Marine and Freshwater Research*, *33*, 491–497. doi: 10.1080/00288330.1999.9516894

Smith, P. J., & McVeagh, S. M. (2006). *Genetic population structure of blackfoot paua*. Final Research Report for Ministry of Fisheries Project SAP2005-01. Unpublished report held by Ministry of Fisheries., Wellington, New Zealand.

Smith, P. J., McKoy, J. L., & Machin, P. J. (1980). Genetic variation in the rock lobsters *jasus edwardsii* and *jasus novaehollandiae*. *New Zealand Journal of Marine and Freshwater Research, 14*, 55–63. doi: 10.1080/00288330.1980.9515843

Smith, P. J., Macarthur, G. J., & Michael, K. P. (1989). Regional variation in electromorph frequencies in the tuatua, Paphies subtriangulata, around New Zealand. *New Zealand Journal of Marine and Freshwater Research, 23*, 27–33. doi: 10.1080/00288330.1989.9516337

Smith, P. J., McVeagh, S. M., Won, Y., & Vrijenhoek, R. C. (2004). Genetic heterogeneity among New Zealand species of hydrothermal vent mussels (Mytilidae: Bathymodiolus). *Marine Biology, 144*, 537–545. doi: 10.1007/s00227-003-1207-4.1080/00288330.1980.9515843

Snelder, T., Leathwick, J., Dey, K., Weatherhead, M., Fenwick, G., Francis, M., Gorman, R., Grieve, J., Hadfield, M., Hewitt, J., Hume, T., Richardson, K., Rowden, A., Uddstrom, M., Wild, M., & Zeldis, J. (2005). *The New Zealand Marine Environment Classification.* Wellington, New Zealand: Ministry for the Environment.

Star, B., Apte, S., & Gardner, J. P. A. (2003). Genetic structuring among populations of the greenshell mussel *Perna canaliculus* revealed by analysis of randomly amplified polymorphic DNA. *Marine Ecology Progress Series, 249*, 171–182. doi:10.3354/meps249171

Steel, D. (1999). *Genetics of the snake star Astrobrachion constrictum (Ophiuroidea: Asteroschematida) in Fiorland, New Zealand* (MSc thesis). University of Otago, Dunedin, New Zealand.

Steel, D., Stewart, B., Trewick, S. A., & Wallis, G. P. (2015). Fiord populations of *Astrobrachion constrictum* (Ophiuroidea: Asteroschematidae) show little genetic differentiation for mitochondrial DNA. *New Zealand Journal of Zoology, 42*, 165–172. doi: 10.1080/03014223.2015.1059863

Stevens, P. M. (1990). A genetic analysis of the pea crabs (Decapoda: Pinnotheridae) of New Zealand. I. Patterns of spatial and host-associated genetic structuring in Pinnotheres novaezelandiae Filhol. Journal of Experimental Marine Biology and Ecology, *141*, 195–212. doi: 10.1016/0022-0981(90)90224-Z

Stevens, P. M. (1991). A genetic analysis of the pea crabs (Decapoda: Pinnotheridae) of New Zealand. *Marine Biology, 108*, 403–410. doi: 10.1007/BF01313649

Stevens, M. I., & Hogg, I. D. (2004). Population genetic structure of New Zealand's endemic corophiid amphipods: evidence for allopatric speciation. *Biological Journal of the Linnean Society, 81*, 119–133. doi: 10.1111/j.1095-8312.2004.00270.x

Stovall, W. R. (2016). *Population genetics of New Zealand fur seals (Arctocephalus forsteri): Genomic tools for research and management* (MSc thesis). University of Otago, Dunedin, New Zealand.

Suggate, R. P., Stevens, G.R., & Te Punga, M. (1978). *The Geology of New Zealand*. Wellington, NZ: E.C. Keating, Government Printer.

Svane, I. B., & Young, C. M. (1989). The ecology and behaviour of ascidian larvae. *Oceanography and Marine Biology: An Annual Review, 27*, 45–90.

Tam, J. C., & Scrosati, R. A. (2011). Mussel and dogwhelk distribution along the north-west Atlantic coast: testing predictions derived from the abundant-centre model. *Journal of biogeography, 38*, 1536–1545. doi: 10.1111/j.1365-2699.2011.02498.x

Tong, L. J., Moss, G. A., Redfearn, P., & Illingworth, J. (1992). *A manual of techniques for culturing paua, Haliotis iris, through to the early juvenile stage. New Zealand Fisheries. Technical Report No. 31*, Wellington, New Zealand.

Triggs, S., & Darby, J. (1989). *Genetics and conservation of yellow-eyed penguin: an interim report. Science and Research Internal Report No. 43*, Wellington, New Zealand.

Trip, E. D. L., Clements, K. D., Raubenheimer, D., & Choat, J. H. (2014). Temperature-related variation in growth rate, size, maturation and life span in a marine herbivorous fish over a latitudinal gradient. *The Journal of animal ecology, 83*, 866–875. doi: 10.1111/1365-2656.12183

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, Mass: Addison-Wesley Publishing Company.

Túnez, J. I., Cappozzo, H. L., Pavés, H., Albareda, D. A., & Cassini, M. H. (2013). The role of Pleistocene glaciations in shaping the genetic structure of South American fur seals (*Arctocephalus australis*). *New Zealand journal of marine and freshwater research, 47*, 139–152. doi: 10.1080/00288330.2012.753463

Tuya, F., Wernberg, T., & Thomsen, M. S. (2008). Testing the 'abundant centre' hypothesis on endemic reef fishes in south-western Australia. *Marine Ecology Progress Series*, *372*, 225–230. doi: 10.3354/meps07718

Underwood, A. J. (1975). Comparative studies on the biology of *Nerita atramentosa* Reeve, *Bembicium nanum* (Lamarck) and *Cellana tramoserica* (Sowerby) (gastropoda: Prosobranchia) in S.E. Australia. *Journal of experimental marine biology and ecology, 18*, 153–172. doi: 10.1016/0022-0981(75)90071-4

Veale, A. J. (2007). *Phylogeography of two intertidal benthic marine invertebrates around New Zealand: the Waratah Anemone (Actinia tenebrosa) & the Snakeskin Chiton (Sypharochiton pelliserpentis)* (MSc thesis). University of Auckland, Auckland, New Zealand.

Veale, A. J., & Lavery, S. D. (2011). Phylogeography of the snakeskin chiton *Sypharochiton pelliserpentis* (Mollusca: Polyplacophora) around New Zealand: are seasonal near-shore upwelling events a dynamic barrier to gene flow? *Biological Journal of the Linnean Society, 104*, 552–563. doi: 10.1111/j.1095-8312.2011.01743.x

Veale, A. J., & Lavery, S. D. (2012). The population genetic structure of the waratah anemone (*Actinia tenebrosa*) around New Zealand. *New Zealand Journal of Marine and Freshwater Research, 46*, 523–536. doi: 10.1080/00288330.2012.730053

Verry, A. J. F., Walton, K., Tuck, I. D., & Ritchie, P. A. (2020). Genetic structure and recent population expansion in the commercially harvested deep-sea decapod, *Metanephrops challengeri* (Crustacea: Decapoda). *New Zealand Journal of Marine and Freshwater Research, 54*, 251-270. doi: 10.1080/00288330.2019.1707696

Verwimp, C., Vansteenbrugge, L., Derycke, S., Kerkhove, T., Muylle, H., Honnay, O., Ruttink, T., Roldán - Ruiz, I., & Hostens, K. (2020). Population genomic structure of the gelatinous zooplankton species *Mnemiopsis leidyi* in its nonindigenous range in the North Sea. *Ecology and Evolution, 10*, 11-25. doi: 10.1002/ece3.5468.

Von der Heyden, S., Prochazka, K., & Bowie, R. C. K. (2008). Significant population structure and asymmetric gene flow patterns amidst expanding populations of *Clinus cottoides* (Perciformes, Clinidae): application of molecular data to marine conservation planning in South Africa. *Molecular ecology*, *17*, 4812–4826. doi: 10.1111/j.1365-294X.2008.03959.x

Vucetich, J. A., & Waite, T. A. (2003). Spatial patterns of demography and genetic processes across the species' range: Null hypotheses for landscape conservation genetics. *Conservation genetics, 4*, 639–645. doi: 10.1023/A:1025671831349

Walker, M. M. (1984). Larval life span, larval settlement, and early growth of *Evechinus chloroticus* (Valenciennes). *New Zealand journal of marine and freshwater research, 18*, 393–397. doi: 10.1080/00288330.1984.9516060

Walls, K. (2009). Nearshore marine classification and inventory – a planning tool to help identify marine protected areas for the nearshore of New Zealand. Department of Conservation. Wellington, New Zealand.

Wallis, G.P., & Trewick, S.A. (2009). New Zealand phylogeography: evolution on a small continent. *Molecular ecology, 18*, 3548–3580. doi: 10.1111/j.1365-294X.2009.04294.x

Walton, K. (2017) *Phylogeography of the New Zealand whelks Cominella maculosa and C. virgata* (MSc thesis). Victoria University of Wellington, Wellington, New Zealand.

Walton, K., Marshall, B. A., Phillips, N. E., Verry, A. J. F., & Ritchie, P. A. (2019). Phylogeography of the New Zealand whelks *Cominella maculosa* and *C. virgata* (Gastropoda: Neogastropoda: Buccinoidea: Buccinidae). *Biological Journal of the Linnean Society*, *126*, 178–202. doi: 10.1093/biolinnean/bly174

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular ecology, 23*, 5649–5662. doi: 10.1111/mec.12938

Waples, R. S. (2020). Serendipity and me. *ICES Journal of Marine Science, 77*, 1658–1665. doi: 10.1093/icesjms/fsaa061

Waters, J. M., & Roy, M. S. (2003). Marine biogeography of southern Australia: phylogeographical structure in a temperate sea-star. *Journal of Biogeography, 30*, 1787–1796. doi: 10.1046/j.0305-0270.2003.00978.x

Waters, J.M., & Roy, M.S. (2004) Phylogeography of a high-dispersal New Zealand sea-star: does upwelling block gene-flow? *Molecular Ecology, 13*, 2797–2806. doi: 10.1111/j.1365-294X.2004.02282.x

Waters, J. M., King, T. M., O' Loughlin, P .M., & Spencer, H. G. (2005). Phylogeographical disjunction in abundant high-dispersal littoral gastropods. *Molecular Ecology, 14*, 2789–2802. doi: 10.1111/j.1365-294X.2005.02635.x

Waters, J. M., Mcculloch, G. A., & Eason, J. A. (2007). Marine biogeographical structure in two highly dispersive gastropods: implications for trans-Tasman dispersal. *Journal of Biogeography, 34*, 678–687. doi:10.1111/j.1365-2699.2006.01615.x

Wear, R. G. (1970). Life-history studies on New Zealand Brachyura: 4. Zoea larvae hatched from crabs of the family grapsidae. *New Zealand journal of marine and freshwater research*, *4*, 3–35. doi: 10.1080/00288330.1970.9515324

Wear, R. G. (1976). Studies On the Larval Development of Met Anephrops Challengeri (Balss, 1914) (Decapoda, Nephropidae). *Crustaceana*, *30*, 113–122. doi: 10.1163/156854076X00521

Wei, K., Wood, A. R., & Gardner, J. P. A. (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. doi: 10.1007/s00227-012-2145-9

Wei, K., Wood, A. R., & Gardner, J. P. A. (2013b). Seascape genetics of the New Zealand greenshell mussel: Sea surface temperature explains macrogeographic scale genetic variation. *Marine Ecology Progress Series*, *477*, 107–121. doi: 10.3354/meps10158

Weir, J. S. & Sagnol, O. (2015). Distribution and abundance of Hector's dolphins (*Cephalorhynchus hectori*) off Kaikoura, New Zealand. *New Zealand journal of marine and freshwater research*, *49*, 376–389. doi: 10.1080/00288330.2015.1020502

Wells, S. J., & 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. doi: 10.7717/peerj.5462

Wen, Z., Ge, D., Feijó, A., Du, Y., Cheng, J., Sun, J., Wang, Y., Xia, L., & Yang, Q. (2020). Varying support for abundance-centre and congeneric-competition hypotheses along elevational transects of mammals. *Journal of biogeography*, *48*, 616-627. doi: 10.1111/jbi.14025

Westfall, K. M. (2010). *Molecular ecology and systematics of blue mussels (genus Mytilus) (Mytilidae; Bivalvia; Mollusca) in the Southern Hemisphere* (PhD thesis). Victoria University of Wellington, Wellington, New Zealand.

Will, M., & Gemmell, N. J. (2008). *Genetic Population Structure of Black Foot Paua*. Final Research Report for Ministry of Fisheries Project GEN2007A. Unpublished report held by Ministry of Fisheries., Wellington, New Zealand.

Will, M., Hale, M. L., Schiel, D. R., & Gemmell, N. J. (2011). Low to moderate levels of genetic differentiation detected across the distribution of the New Zealand abalone, *Haliotis iris*. *Marine Biology*, *158*, 1417–1429. doi: 10.1007/s00227-011-1659-x

Will, M., McCowan, T., & Gemmell, N. J. (2015). Broad-scale genetic patterns of New Zealand abalone, *Haliotis iris*, across a distribution spanning 13° latitude and major oceanic water masses. *Genetica*, *143*, 487–500. doi: 10.1007/s10709-015-9847-0

Williams, S. T., Reid, D. G., & Littlewood, D. T. J. (2003). A molecular phylogeny of the Littorininae (Gastropoda: Littorinidae): unequal evolutionary rates, morphological parallelism, and biogeography of the Southern Ocean. *Molecular phylogenetics and evolution*, *28*, 60–86. doi: 10.1016/S1055-7903(03)00038-1

Wolff, J. N., Ladoukakis, E. D., Enríquez, J. A., & Dowling, D. K. (2014). Mitonuclear interactions: evolutionary consequences over multiple biological scales. *Philosophical transactions. Biological sciences*, *369*, 20130443. doi: 10.1098/rstb.2013.0443

Wood, A., & Gardner, J. P. A. (2007). Small spatial scale population genetic structure in two limpet species endemic to the Kermadec Islands, New Zealand. *Marine Ecology Progress Series*, *349*, 159–170. doi: 10.3354/meps07110

Woods, C. M. C. (2000). Preliminary observations on breeding and rearing the seahorse *Hippocampus abdominalis* (Teleostei: Syngnathidae) in captivity. *New Zealand journal of marine and freshwater research*, *34*, 475–485. doi: 10.1080/00288330.2000.9516950

Woodland, D. J. (1983). Zoogeography of the Siganidae (Pisces): An Interpretation of Distribution and Richness Patterns. *Bulletin of marine science*, *33*, 713–717.

Zeng, C., Rowden, A. A., Clark, M. R., & Gardner, J. P. A. (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. doi: 10.1111/eva.12509

Zeng, C., Clark, M. R., Rowden, A. A., Kelly, M., & Gardner, J. P. A. (2019). The use of spatially explicit genetic variation data from four deep-sea sponges to inform the protection of

Vulnerable Marine Ecosystems. *Scientific Reports, 9,* 5482. doi: 10.1038/s41598-019-41877-9

Zeng, C., Rowden, A. A., Clark, M. R., & Gardner, J. P. A. (2020). Species-specific genetic variation in response to deep-sea environmental variation amongst Vulnerable Marine Ecosystem indicator taxa. Scientific reports, 10, 2844. doi: 10.1038/s41598-020-59210-0