The population history and demography of the Long-spined sea urchin (*Centrostephanus rodgersii*) in Aotearoa New Zealand

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science in
Biological Sciences

at Massey University, Albany, New Zealand.

Jenny Ann Sweatman

2021

Abstract

Climate-driven changes in range and abundance can go undetected, particularly in regions like north-east New Zealand (NENZ) where there have not been routine surveys for many species. The Long-spined sea urchin (*Centrostephanus rodgersii*) has extended its range and increased its abundance in Tasmania over the past 40 years, a change that has dramatically impacted the ecosystem and local fisheries. *Centrostephanus rodgersii* is also found in NENZ but we lack systematic survey information to understand whether a similar range extension could have occurred here. Given the similarity of the Tasmanian and NENZ ecosystems, *C. rodgersii* poses a potential threat to New Zealand's (NZ) marine biodiversity and fisheries. The sizes of individuals within populations were analysed and population genomics was used to study the population history, population structure and recruitment dynamics of *C. rodgersii* across its NZ range. Although population size structure revealed no overall signatures of a poleward range extension, the northern part of the NENZ *C. rodgersii* range did have patterns indicative of a poleward range extension. Furthermore, the size structure of populations further south and east suggested that these populations had more regular recruitment than northern populations. Population genomic analysis revealed that Rangitāhua (the Kermadec Islands) populations and populations of NENZ are genetically differentiated, but there is some ongoing migration from Rangitāhua to NENZ. Within NENZ there was no evidence of population genetic structure, however, population graphs revealed that some groups of populations were more similar in genetic composition, and presumably shared higher geneflow. Two demographic groups (younger than 15 years and older than 15 years) were created to examine differences in genetic composition among age classes found within the same populations. This comparison recovered different patterns of connectivity and within-population variance between the two demographic groups that was not present when the groups were combined. My results indicate that the population demography and structure of *C. rodgersii* in NZ is changing and therefore we need to routinely monitor this urchin across its NZ range to prevent damage to our ecosystems and fisheries.

Acknowledgements

I would like to thank my amazing supervisors Libby Liggins and J. David Aguirre. Thank you for being so patient and encouraging. You have taught me so much over the past few years and I am very grateful to have had the opportunity to work with you.

I am grateful for the help of Vanessa Arranz for laboratory assistance, Jose I. (Iggy) Carvajal for bioinformatic help, Craig R. Johnson for providing information about the population size structure data for Tasmania, Adam N. H. Smith for his help with the Bayesian model, and Eric A. Treml for several discussions regarding the biophysical larval dispersal of *C. rodgersii*.

Thank you to all the people and organisations who made this research possible. Expeditions to Rāngitahua (the Kermadec Islands) were made possible by the Ngāti kuri, Blake, the Commanding Officer and ship's company of HMNZS Canterbury, the RV Braveheart crew (Stoney Creek Shipping Company), and the RV Tangaroa with the support of the Auckland Museum Institute, Auckland Museum, the Pew Charitable Trusts, and the School of Natural and Computational Sciences (SNCS) Massey University, the Marine Funding Advisory Research Group, National Institute of Water and Atmospheric Research (Project COBR1705), Te Papa Tongarewa, and the Ministry for the Environment. Permissions to sample within the Kermadec Islands Marine Reserve and Poor Knights Islands Marine Reserve were provided by mana whenua and the Department of Conversation (47976-MAR and 50153-MAR). I thank J. David Aguirre, Tom Trnski, Adam N. H. Smith, David Acuña-Marrero, Emma Betty, Mat Betty, Crispin Middleton, Phil Ross, Sam McCormack, Severine Hannam, Nick Dunn, and Brad Lang for collecting help. I wish to thank and acknowledge mana whenua (the traditional owners), in particular Ngāti Kuri and Ngātiwai for their support. The samples and derived data in this thesis have a Biocultural (BC) Notice attached. The BC Notice is a visible notification that there are accompanying cultural rights and responsibilities that need further attention for any future sharing and use of this material or data. The BC Notice recognises the rights of Indigenous peoples to permission the use of information, collections, data and digital sequence information generated from the biodiversity or genetic resources associated with traditional lands, waters, and territories. The BC Notice may indicate that BC (Biocultural) Labels are in development and their implementation is being negotiated. For more information about the BC Notices, visit https://localcontexts.org/notice/bc-notice/.

I am grateful for the scholarships and funding that supported my thesis. My research was supported by the Massey University Masterate Scholarship, the Massey University Vice Chancellor's Natural Sciences Excellence Award, the Hutton Fund and funding awarded to Libby Liggins as a Rutherford Postdoctoral Fellow (RFT-MAU1502-PD), Rutherford Discovery Fellow (RDF-20-MAU-001), and a Massey University Early Career Fund.

Thank you to my lab group, it has been a privilege to work alongside such amazing people (and Tiki). Thank you to all my friends at uni, I enjoyed all our lunchtime chats. You have all been such an great part of my thesis experience.

Thank you to my Mum and Dad for supporting me throughout my thesis and all of my education, you have helped me a lot. Thank you to all my family and friends for your support and encouragement. A special thank you to Sam, you have given me the best support and encouragement throughout my thesis but especially through finishing it in lockdown.

Table of contents

Chapter One: General introduction

This thesis contributes to the understanding of marine species range and abundance redistributions in response to climate change. The goals of this thesis were to use the individual size structure of populations and population genomics to infer the population history and demography of a renowned range extending urchin, *Centrostephanus rodgersii* in New Zealand. Chapter One is an overview of the climate-driven redistribution of biodiversity, the impacts of the range redistribution of species, the New Zealand marine environment, urchins in north-east New Zealand, the biogeography of *C. rodgersii,* and how we study species redistributions. An overview of subsequent chapters is outlined at the end of Chapter One.

Climate-driven redistribution of biodiversity

Climate-driven changes in species distributions are happening in marine, terrestrial, and freshwater ecosystems, but the fastest changes in distribution are occurring in marine ecosystems (Sorte et al. 2010; Pecl et al. 2017; Pinsky et al. 2020). There are a number of reasons why marine species are more sensitive to environmental changes than terrestrial ones. First, most marine species are ectotherms and thus more sensitive to changes in ocean temperature. Second, marine ectotherms have a smaller thermal tolerance range than terrestrial ectotherms (Pinsky et al. 2019). Third, marine species have a greater ability, than terrestrial species, to colonise new habitats due to their longer larval dispersal phase (Pinsky et al. 2020) and because the ocean has fewer dispersal barriers. Overall, the general trends in marine redistributions are species shifting poleward and into deeper waters (Poloczanska et al. 2016) because of changing ocean temperatures, ocean circulation, oxygen depletion, and increasing acidification.

Increased ocean temperature and marine heatwaves are major climate drivers for range redistributions (Hillebrand et al. 2018). In the South Island of New Zealand, Thomsen et al. (2019) found that the presence of bull kelp was impacted by the 2017/18 austral summer heatwave. The heatwave led to the loss of bull kelp and even the local extinction of the bull kelp species, *Durvillaea poha*, at the warmest location. In contrast, in Japan, four tropical coral species (*Acropora hyacinthus, Acropora muricata, Acropora solitaryensis, Pavona decussata*) have extended their range in a poleward direction due to the increase in ocean temperature (Yamano et al. 2011). As temperatures get warmer the most common response is moving in a poleward direction towards cooler temperatures (Poloczanska et al. 2013; Melbourne-Thomas et al. 2021), but since warming happens unevenly across the ocean, increased temperature will impact species and areas differently.

Climate change is causing oxygen depletion which impacts ocean productivity, nutrient cycling, carbon cycling, and habitats (Keeling et al. 2010). Oxygen depletion can impact the behaviour and physiology of marine species and in extreme cases can cause local extinctions or force redistributions (Townhill et al. 2017). In the northern Gulf of St. Lawrence in Canada, modelling found that decreased dissolved oxygen will impact the distribution of the Greenland halibut (*Reinhardtius hippoglossoides*) and northern shrimp (*Pandalus borealis*), and in combination with warming, the impact of oxygen depletion will be even greater

(Stortini et al. 2017). Overall, increased ocean temperatures and oxygen depletion will make equatorial and shallow regions increasingly hostile (Deutsch et al. 2015).

Ocean acidification impacts the calcification, photosynthesis, nitrogen fixation, and reproduction of marine organisms (Doney et al. 2009). Echinoderms and molluscs are particularly impacted in both their reproduction and calcification (Doney et al. 2009). When the collector urchin (*Tripneustes gratilla*) was exposed to acidification (pH 7.6) in experiments, it produced almost no gonads and grew more slowly compared to at the control ambient pH treatment (Dworjanyn et al. 2018). Another echinoderm, the brittle star (*Amphiura filformis*), moved to shallower depths within the sediment in response to severe short-term acidification in the laboratory (Murray et al. 2013). We are likely to see local extinctions and changes in species distribution as a result of ocean acidification, but responses will vary according to species biology and in interaction with other environmental factors impacted by climate change.

Climate change also impacts ocean circulation, which in turn can lead to changes in dispersal pathways and species distributions (Wilson et al. 2016). Hydrodynamic barriers can limit climate-driven range redistributions (Keith et al. 2011), and coupled with changing environmental factors, hydrodynamic barriers that prevent poleward range shifts could lead to range contractions. However, when a hydrodynamic barrier is relaxed, potentially as a result of climate change, species can disperse and therefore extend their ranges. Additionally, climate change can strengthen or change currents allowing species to extend their ranges in unpredicted directions and/or with unprecedented speed. New dispersal routes will not always lead to a range shift, but if the newly colonised location is habitable, or made habitable by changing climate, then the species can extend its range. In particular, changes in the strength of poleward boundary currents lead to both an increase in poleward dispersal potential and a greater penetrance of warm tropical water to higher latitudes. One of the most striking examples is the strengthening of the East Australian Current (EAC) which caused a climate change cascade in Tasmania (Johnson et al. 2011). The strengthening of the EAC has allowed both the common Sydney octopus (*Octopus tetricus*) (Ramos et al. 2018) and the Long-spined sea urchin (*Centrostephanus rodgersii*) (Ling et al. 2009c) to extend their ranges to Tasmania. As climate change increasingly impacts ocean currents and hydrodynamic barriers we will see more widespread changes to species distributions.

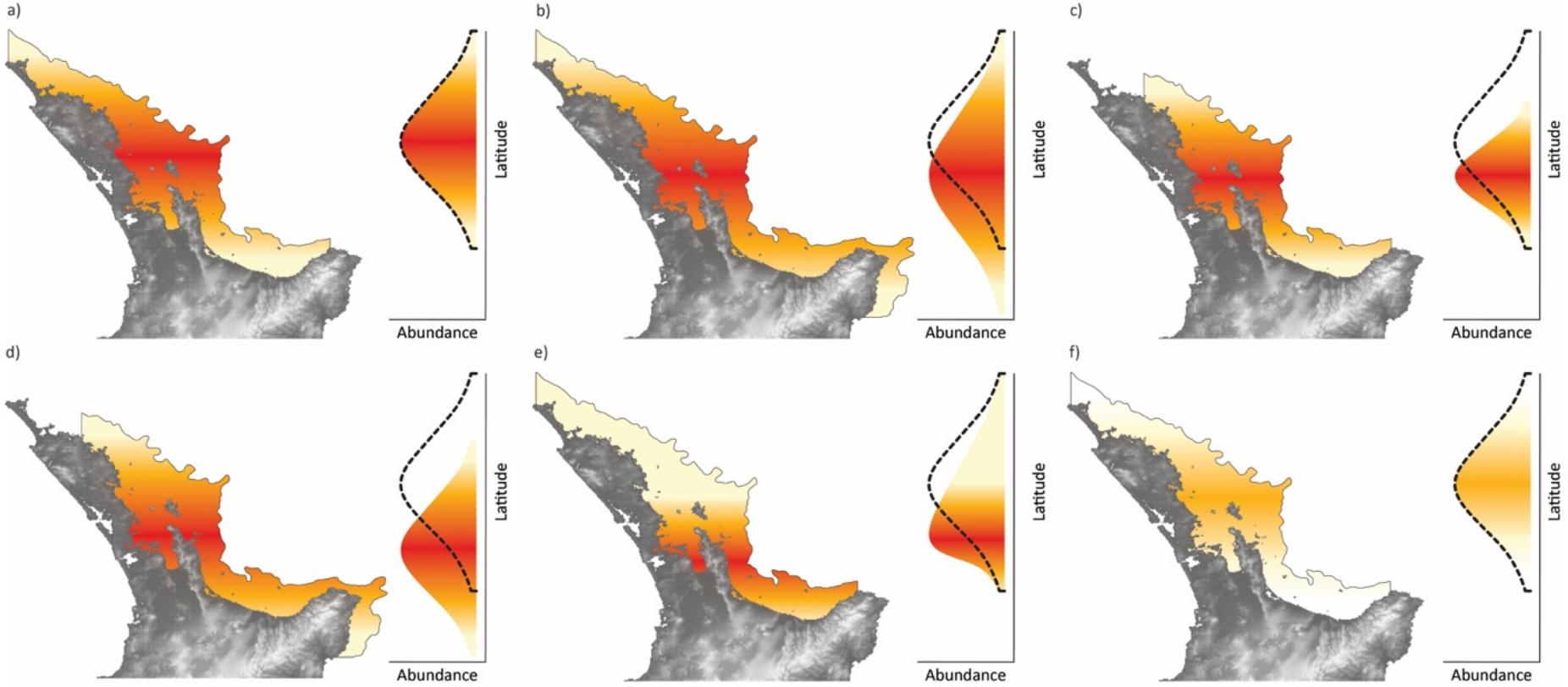

Figure 1: Different hypothetical scenarios for how a species range and abundance may respond to climate change: a) the range and abundance could remain *stable*; b) the range could have an *extension*; c) the range could have a *contraction*; d) the range and centre of abundance could *march*; e) the centre of abundance could *lean*; f) the species abundance could *collapse*.

Climate-driven changes in range and abundance happen in several different ways (see Fig. 1), but the most advantageous for the spread of a species is a range extension (Fig. 1b). In a range extension, the species will travel to a new location, increase in abundance and persist in the new location (Bates et al. 2014). Range extensions have occurred in a wide variety of species in various parts of the world. For example, range extensions have occurred for: mangrove forests on the Florida east coast (Cavanaugh et al. 2014); Mahimahi (*Coryphaena hippurus*) in the North Island of New Zealand (Middleton et al. 2021); four tropical reef corals (*Acropora hyacinthus*, *Acropora muricata*, *Acropora solitaryensis*, *Pavona decussata*) in Japan (Yamano et al. 2011); and Adelie penguins (*Pygoscelis adeliae*) in the Ross Sea, Antarctica (Taylor et al. 1990). Range extensions will likely becoming increasingly common as we continue to see the impacts of climate change.

In range contractions (Fig. 1c), local extinctions occur at the rear edge of the range and therefore a portion of the range is lost. This generally occurs when the environment changes and the species cannot adapt or acclimate, but also cannot shift its range because of some other limiting feature of the environment. For example, in Western Australia, the habitat-forming seaweed species, *Scytothalia dorycarpa*, lost approximately 100km of its range in a marine heatwave as the species could not cope with the increased temperature (Smale et al. 2013; Wernberg et al. 2021). In marine species, those species with ecological specialization, small geographic ranges and endemic species are more vulnerable to marine contractions and extinctions (Dulvy et al. 2003).

In a range march the range size may stay the same but both the leading and trailing edges of the range are shifted in the same direction by a similar amount (Fig. 1d). The kelp species, *Laminaria hyperborea*, has had a poleward range march in Europe (Assis et al. 2016). Increasing temperatures have made the northern waters warm enough for the kelp while the southern waters have become too warm for the kelp, thereby shifting the range north. Increasing ocean temperatures are predicted to cause this range march to continue. Climate change will likely lead to more marine range marches globally.

Climate-driven changes across a species range are more complex than just an increase (range extension), decrease (range contraction), or change in positioning (range march) of a species range. Climate change can also lead to changes in species abundances and densities without a change in range, particularly where there is no suitable habitat to move into. In a range lean, the range extent remains stable but the centre of abundance is shifted towards one edge of the range (Fig. 1e). The worst-case scenario is a collapse in the species abundance, whereby abundance is decreased, causing a contraction of the species from the edges of the range (Fig. 1f). Early detection of climate-driven changes in the range and abundance of species is an important part of mitigating the impacts (Melbourne-Thomas et al. 2021). To detect these changes we need to be aware of all the complex ways ranges and abundance can change in response to climate change.

Impacts of the range redistribution of species

The impacts for humans of the climate-driven redistribution of marine species include loss of food security (e.g. loss of fisheries), decreased human health (e.g. distribution of pathogens), loss of supporting livelihoods (e.g. fisheries providing people's livelihoods),

climate feedbacks (e.g. melting of sea ice leads to the suppression of spring blooms which means less carbon dioxide is absorbed leading to further melting of sea ice), compromised governance (e.g. a fisheries stock shifting to a new geopolitical area creating conflict over who governs it) and cultural values (e.g. Indigenous groups losing their culturally important species) (Pecl et al. 2017; Scheffers et al. 2019; Melbourne-Thomas et al. 2021). These impacts can be directly from the species that has changed its distribution or through an ecosystem function change impacting other species. Being able to predict these impacts and detect them early will help us minimise the effect on human populations.

The effects of range redistribution on ecosystems will vary dramatically depending on the ecological role of the species undergoing a redistribution. Habitat-forming species play a large role in creating the habitat and ecosystem for other species, such as macroalgae (Verges et al. 2016), reef-building corals, and seagrasses, which create habitat for diverse assemblages (Bulleri et al. 2018). Habitat-forming species can also modify the temperature of the environment, sometimes even mitigating latitudinal and elevational trends in temperature (Jurgens et al. 2018). The habitat-forming seaweed, *Scytothalia dorycarpa*, lost 100km of its range in Western Australia (Wernberg et al. 2021) and contracted its range by about 5% globally in the 2011 marine heatwave (Smale et al. 2013). In Western Australia, the loss of this habitat-forming species led to an increase in turf-forming algae and a decrease in encrusting algae and sponges, thereby changing the habitat (Wernberg et al. 2021). Even substitutions of habitat-forming species, such as a different kelp species, can still lead to decreased biodiversity (Teagle et al. 2018).

The range redistribution of species that directly interact with habitat-forming species can also change communities and ecosystems. A common example is urchins, which graze on the habitat-forming macroalgae species, changing the habitat from macroalgae forests to urchin barrens. A range extension or abundance increase of an urchin species can cause the macroalgae to be overgrazed leading to the loss of the macroalgal forest and the community it supports. Poleward range extensions of urchins, likely due to temperature, have happened in several locations such as Japan (*Heliocidaris crassispina* and *Hemicentrotus pulcherrimus*; Agatsuma et al. 2007; Feng et al. 2019), Australia (*Centrostephanus rodgersii*; Ling et al. 2009c), and possibly are beginning in California (*Centrostephanus coronatus*; Freiwald et al. 2016). Likewise, the formation of urchin barrens is happening across the globe (Ling et al. 2015).

The New Zealand marine environment

The New Zealand marine environment is varied, ranging from subtropical in Rangitāhua (the Kermadec archipelago) to subantarctic in the Subantarctic Islands. Within temperate mainland New Zealand there are 11 bioregions, based on macroalgal species presence-absence, within 2 biogeographical provinces: northern and southern (Shears et al. 2008). Ocean warming is already happening in at least some parts of New Zealand (Shears et al. 2017) and we can expect this to increase (Law et al. 2017a) along with ocean acidification (Law et al. 2017b). New Zealand has experienced three recorded marine heatwaves, in the summers of 1934/35, 2017/18, and 2018/19 (Salinger et al. 2019; Salinger et al. 2020). The 2017/18 summer marine heatwave led to the decrease and local extinction of the bull kelp

species, *Durvillaea poha*, in the South Island (Thomsen et al. 2019). Future temperature increases will likely impact species across New Zealand's marine environment.

Currently, there have been few recorded impacts of climate change on the distribution of marine species in New Zealand. New Zealand does not have a strong boundary current, which is a current that follows the coastline and is often implicated in species range extensions. Instead, we have the highly variable East Auckland Current (EAuC) (Stanton et al. 1997), which may cause range changes to be decoupled with latitude along the coastline and therefore difficult to detect. Due to the lack of monitoring, we do not know whether changes in species distributions are happening. However, species like *C. rodgersii* have been recorded to be increasing in abundance at some locations (Balemi et al. 2021). Consistent monitoring will allow us to detect climate-driven changes in species distribution.

Urchins in north-east New Zealand

In the shallow reef ecosystems of north-east New Zealand, there are four large urchin species: the Long-spined sea urchin (*Centrostephanus rodgersii*), kina (*Evechinus chloroticus*), *Heliocidaris tuberculata* and *Tripneustes kermadecensis*. *Evechinus chloroticus* is endemic to New Zealand, a fisheries species, and a taonga, whereas *C. rodgersii* is found in both New Zealand and Australia and does not currently have a fishery in New Zealand. In north-east New Zealand, both these urchin species are found in kelp forests, but *C. rodgersii* has been found to eat a higher proportion of invertebrates than *E. chloroticus* (Balemi et al. 2021).

In areas of New Zealand kelp forests, urchins have formed barrens (Shears et al. 2002). The lack of predators in macroalgal forests, often due to the overfishing of the predator, leads to an explosion of the urchin population. The large urchin population leads to the over-grazing and loss of macroalgae, along with the many species macroalgae support, creating a 'barren' habitat. Even though there is a lack of macroalgae in barrens, the urchins can persist by changing their physiology and behaviour (Johnson et al. 1982) so they can feed on filamentous and coralline algae (Ling et al. 2009a), preventing the macroalgae from returning. In north-east New Zealand, there are barrens formed by both *E. chloroticus* and *C. rodgersii*. Some of the impacts from urchins on the shallow marine ecosystems of north-east New Zealand, particularly by *C. rodgersii*, are expected to be exacerbated by climate change.

The Long-spined sea urchin, *Centrostephanus rodgersii*

Centrostephanus rodgersii has a three-month larval stage (Huggett et al. 2005) which has allowed it to travel long distances; and this, combined with increasing ocean temperatures, has led to range extensions of this urchin. Historically, *C. rodgersii* was common off the coast of New South Wales, Australia, but the species has now extended its range to Tasmania and also occurs in Lord Howe Island, Norfolk Island, north-east New Zealand and Rangitāhua (Byrne et al. 2020). In order to extend its range, the *C. rodgersii* larvae need to be able to reach the new area, that has a minimum winter ocean temperature of 15°C (Pecorino et al. 2013b). In the last few decades, climate change has strengthened the EAC (Oke et al. 2019) which both warmed Tasmanian waters and transported *C. rodgersii* larvae

to Tasmania thereby resulting in a range extension and an increase in abundance of the urchin. The range extension occurred through multiple colonisation events, starting in 1978 (Ling et al. 2009c; Johnson et al. 2011). We may see further similar range extensions in other locations, such as New Zealand, due to climate change's impact on currents and ocean warming.

Barrens have been able to form in much of *C. rodgersii*'s habitat due to the lack of predators (Ling et al. 2009b). *Centrostephanus rodgersii* barrens are found in Tasmania (Johnson et al. 2005), New South Wales (Andrew 1994), Victoria (Cartwright et al. 2019), and New Zealand (Liggins, pers. comm. 2021). Large sized southern rock lobsters (*Jasus edwardsii*, also known as crayfish) and packhorse rock lobsters (*Sagmariasus verreauxi*) are the natural predators of *C. rodgersii*, and therefore can prevent barren formation by controlling urchin populations, but they have been overfished in many of the locations where *C. rodgersii* is found. In Tasmania, barrens have formed where *J. edwardsii* populations have been overfished (Ling et al. 2009b). In New Zealand, the Bay of Plenty *J. edwardsii* stock is overfished (Webber et al. 2018) and in the Hauraki Gulf, *J. edwardsii* are functionally extinct (Hauraki Gulf Forum 2020) leaving these locations vulnerable to the formation of *C. rodgersii* barrens.

In Tasmania, the extensive *C. rodgersii* barrens have negatively impacted biodiversity and fisheries by altering the local ecosystem (Johnson et al. 2005; Ling 2008; Lisson 2018). The formation of barrens in Tasmania has caused macroalgae beds to lose about 150 taxa (Ling 2008). Alongside the loss of biodiversity is the loss of important fisheries species. The blacklip abalone (*Haliotis rubra*) and southern rock lobster (*J. edwardsii*) fisheries are both supported by the macroalgae. The loss of these macroalgae beds due to *C. rodgersii* barrens has resulted in Tasmanian abalone exports decreasing over the last 20 years from 1,500 tonnes in 1998 to 294 tonnes in 2018 representing an annual loss of $49 million (AUD) in export revenue (Lisson 2018). There has been a *C. rodgersii* fishery in Tasmania started to compensate the loss of fisheries, however, the individuals found in barrens are smaller and therefore not desirable for fisheries (Ling et al. 2009a).

Centrostephanus rodgersii has been present in New Zealand longer than Tasmania, however, less is known about the population history and demography. *Centrostephanus rodgersii* was first recorded in New Zealand in 1897 (Farquhar 1897) but was subsequently removed from the faunal list on two occasions for lack of evidence (Fell 1949), potentially indicating a low prevalence. It was not until 1949 when live specimens were collected from several locations in the North Island (the Cavalli Islands, Stephen's Island, Whangaroa, and Little Barrier) that there was conclusive evidence of the species presence (Fell 1949). Population genetic studies confirm that the origin of the New Zealand populations is the East Coast of Australia (Banks et al. 2007), but suggest that the New Zealand populations are no longer reliant on dispersal from Australia and are a self-sustaining meta-population (Thomas et al. 2021). This urchin is increasing in abundance at some locations in north-east New Zealand (Balemi et al. 2021) and is the most abundant urchin species at Rangitāhua (Liggins, pers. comm. 2021), but outside these few locations, there have not been surveys on abundance. Additionally, there are no surveys that would detect a range extension.

Climate change will likely lead to increases in *C. rodgersii*'s range and abundance in New Zealand. Currently, *C. rodgersii*'s range in New Zealand occurs up to its minimum thermal temperature 15°C (Pecorino et al. 2013b). As ocean temperatures increase the 15°C winter isotherm will likely move south allowing *C. rodgersii* populations to extend southwards, similar to the range extension in Tasmania (Ling et al. 2009c). The other impact of climate change is increased ocean acidification. Although decreased pH impacts the jelly coat of *C. rodgersii*'s eggs (Foo et al. 2017) and *C. rodgersii*'s larval development (Doo et al. 2012), the future predictions of pH will likely have minimal impacts on *C. rodgersii* in New Zealand (Pecorino et al. 2013c). Additionally, *C. rodgersii* may be able to adapt to the decreased pH (Foo et al. 2012). Overall climate change is likely to benefit *C. rodgersii* and extend its range in New Zealand (Pecorino et al. 2013c).

Centrostephanus rodgersii poses a threat to New Zealand's ecosystem and fisheries. The dominant macroalgae in north-east New Zealand is *Ecklonia radiata*, which was also the dominant canopy-forming kelp in Tasmania where extensive barrens formed (Ling et al. 2018). In New Zealand, *E. radiata* supports a high-value fishery of crayfish (*J. edwardsii*) which is the same species as the Tasmanian southern rock lobster, pāua (*Haliotis iris*), and kina (*Evechinus chloroticus*). Therefore, New Zealand could have similar losses in biodiversity and fisheries to Tasmania (Johnson et al. 2005; Ling 2008; Lisson 2018). *Centrostephanus rodgersii* needs further study in New Zealand for us to understand and manage the threat it poses to New Zealand's ecosystem and fisheries.

Studying species redistributions

In the absence of time-series survey data, population size structure analysis can be used to examine changes in a species abundance as well as changes in population demography and connectivity to infer range changes. The sizes of individuals in a population indicate the range of ages within a population, and therefore population demography. When we study demography across multiple populations we can detect past range shifts or expansions (Gurevitch et al. 2016). In a range extension (Fig. 1b) of a species with a pelagic larval stage and benthic adult stage, the newly established leading edge will have young individuals that have newly arrived. In contrast, in a range contraction (Fig. 1c) a population may stop recruiting and/or reproducing shifting the demography to older individuals. Measures like the coefficient of variance for individual size, have been used to detect relationships between recruitment and physical ocean features, like upwelling. For example, Ebert et al. (1988) found recruitment of the purple sea urchin (*Strongylocentrotus purpuratus*) was irregular closer to locations of intense upwelling. Similarly, Black et al. 2011 found the coefficient of variance for size in the giant clam (*Tridacna maxima*) increased with increasing latitude, indicative of an increase in recruitment at higher latitudes. Size structure is less time and resource intensive than the ecological time-series surveys used to regularly survey the demography and abundance of a species and therefore is an important tool for detecting differences in demography.

Population genomics can be used to identify the population structure and connectivity of a species. Neutral and adaptive genetic loci can be used to identify the distinct meta-populations. Population graphs give more in-depth genetic structuring including the shared genetic variance between populations, which indicates the geneflow, and the within-

population variance (Dyer et al. 2004). The adaptive genetic variance can inform us how resilient populations may be to future events (Dalongeville et al. 2018b). Using these genomics tools in a species that is potentially changing in abundance and range can inform management decisions.

Currently, there is no monitoring of whether a range extension of *C. rodgersii* could be occurring in New Zealand. In this thesis, I used population size structure analysis to look for signals of recent range extensions in north-east New Zealand. I also studied the population genomics of *C. rodgersii* in north-east New Zealand and Rangitāhua. Although there have been population genetic studies on *C. rodgersii* in New Zealand (Banks et al. 2007; Thomas et al. 2021), this thesis uses the most New Zealand locations and is the first study on *C. rodgersii* to use genome-wide single nucleotide polymorphisms. Using population genomics I examined the patterns of connectivity across the range of *C. rodgersii* and how these patterns differed between demographic groups.

Structure of this thesis

This thesis contains two empirical data chapters (Chapters Two and Three) and a general discussion chapter (Chapter Four). Chapters Two and Three are written in manuscript format as I plan to submit them for peer-review and publication after receiving the examiner's comments. For this reason, there is some repetition of background information and explanation of the study system in these chapters. These chapters will be submitted for peer-review as co-authored manuscripts, and so I use "we" (first-person plural) to acknowledge this; nonetheless, the thesis is my own work, completed under the guidance of my supervisors.

Chapter Two uses population size structure to look for signals of a poleward range extension. This chapter reviews the use of size to infer demography, recruitment and range dynamics. I created a Bayesian model to detect signals of a poleward range extension and confirmed the ability to detect demographic changes using the model in a simulation study. I then used this model firstly, on the known range extension of *C. rodgersii* in Tasmania, and then, on the data from north-east New Zealand. I discuss the use and limitations of this model especially in a changing ecosystem with potentially changing species ranges.

Chapter Three uses population genomics to study New Zealand populations of *C. rodgersii*. I was interested in the patterns of connectivity across north-east New Zealand and Rangitāhua. I was also interested in whether population dynamics were changing over time. I examined this by dividing the individuals within each population into two demographic groups (i.e. up to 15 years old, and over 15 years old) and comparing their genetic composition and patterns of population connectivity. I then discuss what this analysis means for the New Zealand meta-population of *C. rodgersii*.

Chapter Four gives a general overview of Chapters Two and Three and how their findings cross inform each other. I discuss the future directions and applications of this research and the management of range-extending species including future scenarios for New Zealand.

Chapter Two: Using the size structure of populations to infer range extensions and the regularity of recruitment

Abstract

Climate change is causing shifts in the distributional ranges of species. Range shifts of habitat-forming species, or species that influence the predominant habitat-forming species, can have great impacts on ecosystems. *Centrostephanus rodgersii* (the Long-spined urchin) extended its range southward to Tasmania, where it drastically reduced biodiversity and impacted the local fisheries. *Centrostephanus rodgersii* is also found in north-east New Zealand (NENZ), where it could potentially impact subtidal ecosystems similarly to those in Tasmania. Unfortunately, there are no time series data from NENZ from which to infer population or recruitment dynamics. Instead, we inferred population dynamics by analysing the size structure – the mean and standard deviation of individual urchin sizes – of *C. rodgersii* populations across its range using a Bayesian modelling approach. We detected a poleward decrease in mean sizes and a poleward increase in the standard deviations of sizes in *C. rodgersii* in both Tasmania and in northern NENZ, consistent with a poleward range extension, which is well-documented in Tasmania but not yet NENZ. Southern NENZ populations had varied population mean sizes and similar standard deviations of sizes, indicating there has not been a strictly poleward range extension. Across both the Tasmanian and NENZ range of *C. rodgersii*, our model suggested an increase, with latitude, in the regularity of recruitment in higher latitude populations, indicating that the southern range limits may not be recruitment limited. Our study demonstrates that size structure data can be a valuable resource in understanding population histories and recruitment dynamics in the absence of time-series data.

Introduction

Climate-mediated range shifts are altering recipient ecosystems causing widespread socio-economic impacts (Pecl et al. 2017). Range extensions, a common type of range shift, involve a species colonising a new location, followed by population growth through increased immigration or local recruitment (Bates et al. 2014). In the marine environment, we expect to see increased immigration and colonisation of new locations as a consequence of climate-forced changes in ocean circulation altering, and in some cases strengthening, dispersal pathways (Doney et al. 2012, Wilson et al. 2016). Furthermore, localised warming of ocean temperatures due to climate change helps species to survive and reproduce in the new location, promoting population growth. Despite the prevalence of range extensions in the marine environment (Pinsky et al. 2020), determining whether a certain species is undergoing a range extension is often difficult to establish.

Range shifts of habitat-forming species, or species that influence the predominant habitat-forming species, can substantially impact the structure and function of marine communities (Doney et al. 2012, Jurgens & Gaylord 2018). For example, the loss of a habitat-forming species like kelp will dramatically impact the ecosystem (Teagle et al. 2018), and therefore an increase or introduction of herbivores that eat kelp can threaten the ecosystem. One such species is the Long-spined sea urchin (*Centrostephanus rodgersii*), which can create

'barrens' by overgrazing macroalgal kelp forests. The impact of *C. rodgersii* has been most pronounced where the species has extended its range into south-east Australia and Tasmania. The strengthening of the East Australian Current (EAC), due to climate change (Oke et al. 2019), has enabled *C. rodgersii* larvae (capable of dispersing for three to four months; Huggett et al. 2005) to disperse into Tasmanian waters (Ling et al. 2009) during multiple colonisation events (Johnson et al. 2011). The EAC has also increased the sea temperature along the south-east of Australia allowing the persistence of *C. rodgersii* in Tasmania. As a result of the barrens habitats created by *C. rodgersii*, Tasmanian macroalgae forests have lost at least 150 taxa (Ling 2008), and important fisheries such as blacklip abalone (*Haliotis rubra*) and southern rock lobster (*Jasus edwardsii*) have been affected (Johnson et al. 2005, Lisson 2018).

The natural range of *C. rodgersii* includes the east Australian coast, Lord Howe Island, Norfolk Island, Rangitāhua (the Kermadec archipelago), and north-east New Zealand (Byrne & Andrew 2013). However, we know little about the historic range-wide dynamics of the species. Presumably, *C. rodgersii* colonised the Tasman Sea Islands (Lord Howe Island and Norfolk Island) via the eastern flowing Tasman Front, an offshoot of the EAC (Oke et al. 2019), and north-east New Zealand was colonised via the south-easterly flowing East Auckland Current (EAuC) that arises from the Tasman Front. *Centrostephanus rodgersii* was first recorded in New Zealand in 1897 (Farquhar 1897) but was subsequently removed from the faunal list on two occasions for lack of evidence. It was not until 1949, when live specimens were collected from several locations (the Cavalli Islands, Stephen's Island, Whangaroa, and Little Barrier), that there was conclusive evidence of the species presence (Fell 1949). Population genetic studies identified high connectivity between *C. rodgersii* populations in New Zealand and the East Coast of Australia (Banks et al. 2007), and suggest that New Zealand populations have since become a self-sustaining meta-population and are no longer reliant on dispersal from Australia (Thomas et al. 2021). Although *C. rodgersii* is now widespread in New Zealand (ranging from Rangitāhua to Ariel Reef, Gisborne), due to a lack of survey data, we do not know how this increase in range and abundance since 1897 occurred, and whether the species continues to extend its range.

Centrostephanus rodgersii likely poses a threat to New Zealand's coastal ecosystem and fisheries. The dominant macroalgae in north-east New Zealand is *Ecklonia radiata* which was the macroalgal species most impacted by *C. rodgersii* in Tasmania (Ling & Keane, 2018). In New Zealand, *E. radiata* supports a high-value fishery of crayfish (*J. edwardsii*, also known as southern rock lobster), pāua (*Haliotis iris*), and an endemic urchin, kina (*Evechinus chloroticus*). Although ocean temperatures have increased over recent decades, the climate-driven changes to New Zealand's oceanography and ocean environment have not been as dramatic as those seen in south-east Australia and Tasmania (Law et al. 2017). Nonetheless, larval rearing experiments have confirmed that *C. rodgersii* can reproduce and complete the crucial stages of larval development in New Zealand (Pecorino et al. 2013b), suggesting the species can self-recruit under current conditions (Pecorino et al. 2012). Furthermore, monitoring data from 1999 to 2020 shows there has been an increase in the abundance of *C. rodgersii* at the Poor Knights Islands Marine Reserve and the Mokohinau Islands (Shears 2020, Balemi et al. 2021). More generally, however, a thorough understanding of the historical and likely future range dynamics of the species in New Zealand is lacking.

In the absence of time-series survey information, population age-class or size structure information can be used to infer the relative timing of colonisation and frequency of recruitment into populations. Size structure has been used both alone and in combination with abundance and/or density to infer population dynamic for many marine invertebrates, including: the red sea urchin (*Strongylocentrotus franciscanus*;Tegner & Dayton 1981, Botsford et al. 1994, Morgan et al. 2001), the purple sea urchin (*Strongylocentrotus purpuratus*; Ebert & Russell 1988, Ebert et al. 1999, Ebert 2010), and the Kellet's whelk (*Kelletia kelletii*) in California, USA (Zacherl et al. 2003); and the mulberry whelk (*Morula marginalba*), a marine snail (*Afrolittorina pyramidalis*), and the rose barnacle (*Tesseropora rosea*) in south-east Australia (Hidas et al. 2010). The mean size of individuals in a population can give an estimation of how long a population has been established. For instance, if a population has been present for a long time, a large proportion of the individuals will be full-sized adults, so the population will have a large mean size. In contrast, a newly colonised population will have a higher proportion of younger individuals, and therefore a smaller mean size for the population. The standard deviation of sizes in a population also provides information about the regularity of recruitment. For instance, a small standard deviation indicates less variation in sizes and therefore similar ages, so there may have only been one recruitment event leading to one age class. In contrast, a large standard deviation indicates the individuals are spread across all age classes, therefore there would have been regular recruitment.

In combination, these two measures – the mean and standard deviation of sizes – have been used to calculate the coefficient of variation to reveal how recruitment patterns vary across species ranges (Ebert & Russell 1988). A large coefficient of variation indicates the sizes of individuals are highly variable relative to the mean and therefore suggests more regular recruitment (Morgan et al. 2001). In contrast, a small standard deviation relative to the mean, and therefore a small coefficient of variation, may indicate recruitment into the population is episodic as only a few size classes are found. For example, a relationship between the coefficient of variation and the distance from a headland where upwelling occurred for purple sea urchin (*S. purpuratus*) populations sampled from central California to central Oregon (Ebert & Russell, 1988) indicating recruitment was irregular closer to the upwelling. Black et al. (2011) also used the coefficient of variation for small giant clam (*Tridacna maxima*) populations within the Ningaloo Marine Park, Australia, to infer that recruitment increased with decreasing northward latitude.

Although the coefficient of variation has been used more frequently in previous studies than the mean or standard deviation independently, the latter approach may yield more valuable insights in climate impacted oceans. For example, a newly established population at the leading edge of a range extension that has only young individuals from a few recruitment events would have a small standard deviation and small mean size, giving it a medium coefficient of variation. Using the coefficient of variation alone, the dynamics of this population would be indistinguishable from a long-established population with regular recruitment which would have a large mean and large standard deviation also giving the population a medium coefficient of variation. Thus, more nuanced analysis of population size structures – including the mean, standard deviation, and coefficient of variation – can provide a view into both the population histories and ongoing recruitment dynamics of

populations, particularly when appropriate measures to the expected demographic scenarios are used.

Our study uses population size structure to investigate signs for a recent range extension, and to examine patterns of recruitment for *C. rodgersii*. In Tasmania, Ling et al. (2009) examined the abundance and size structure of *C. rodgersii* urchin populations along the range extension axis. The authors found an exponential decline in mean urchin size (and age inferred via growth models) with distance from the EAC and a poleward decrease in size, and therefore age, and abundance, confirming demographic signatures for the documented poleward range extension. Here, we use population size structure data to detect similar signs of recent or varied colonisation timing, or varied recruitment, across the north-east New Zealand range of *C. rodgersii*. Specifically, we model trends in population size structure parameters (mean size, standard deviation of size and coefficient of variation) with latitude where we expect smaller mean sizes in southern populations, indicative of a poleward range extension similar to what was observed in Tasmania. In addition, a high standard deviation in population sizes could indicate where populations are regularly recruiting. To test the performance of our model, we first use simulated data to confirm if trends in both the mean and standard deviation of urchin test diameters could be recovered. Second, we modelled the size structure data from Tasmanian *C. rodgersii* populations to test our inferences in the case of a known range extension. Last, based on our verified approach and assumptions, we use our model to investigate the size structure data of *C. rodgersii* populations in New Zealand to infer the species population history and range dynamics in New Zealand.

Methods

Modelling population size structure trends

Size is typically determined by age and growth rate. For urchins, differences in growth rates can be due to differences in food availability or ambient temperature (Pecorino et al. 2012). In our study, we assume that all north-east New Zealand locations have similar growth rates and so size can be used as a surrogate for age. In support of our assumption, Ling et al. (2009) found that all sampled Tasmanian *C. rodgersii* populations, across a larger latitudinal gradient, had the same growth rates. Furthermore, growth models created for *C. rodgersii* at the Mokohinau Islands in north-east New Zealand were very similar to the models derived for Tasmanian populations (Pecorino et al. 2012). Additionally, the entire north-east New Zealand range of *C. rodgersii* is within one bioregion within which food availability is assumed to be similar (Shears et al. 2008). Nonetheless, our approach will only allow us to see signals of a recent range extension, up to 10-15 years prior to the sampling, as once the urchins reach full size (at 10-15 years; Pecorino et al. 2012) they are indistinguishable.

Using a Bayesian modelling approach fit with the R package Rethinking version 2.01 (McElreath 2016, 2020), we created a model to detect trends in the mean size of individuals in a location, and the standard deviation of size in a location, across latitude. First (2.2), we tested the model's performance using simulated size structure data with several trends in different directions and of different magnitudes. Second (2.3), we used the data from Tasmania's verified poleward range extension of *C. rodgersii* to test whether we could detect the expected poleward trend in size structure. Last (2.4), once we had verified that

the model would pick up trends, we used it to infer trends for the New Zealand populations of *C. rodgersii*. The test diameter (size) is based on a normal distribution where both the mean and standard deviation are modelled. Both the mean and standard deviation have a global mean/standard deviation, a slope for the mean/standard deviation and an error term for the mean/standard deviation (full details in Appendix 1). The mean was estimated as 90mm based on the mean of the observed data and since this is close to the growth asymptote for both New Zealand and Tasmania (Pecorino et al. 2012). The standard deviation of 15 was based on the observed data. Latitude was standardised and centred in the model (see Appendix 1).

(I) Simulation study

In the simulation study, we tested whether the model could recover trends for changes in test diameter with respect to latitudes that were of different magnitudes and directions for both the location means and location standard deviations. For each scenario, we simulated two data sets with two different sample sizes: five locations and fifteen locations (where each location represents a population). Within each location, we simulated test diameter values for 30 individuals across all simulations.

We simulated nine scenarios for a combination of neutral, positive and negative relationships between test diameter and latitude for both the location means and location standard deviations. Next, we explored if our model would be sensitive to differences in the magnitude of the regression parameters as well as differences in direction. Hence, we simulated six additional datasets; three scenarios for a change in the magnitude of the regression parameter for location means (b_L) and three scenarios for a change in the magnitude of the regression parameter for location standard deviations (c_L). Lastly, we examined if non-linear patterns in the location means or standard deviations across latitude influenced our estimates. These two simulations used convex relationships between latitude and either location means or location standard deviations, and no relationship with latitude and either location means or locations standard deviations, respectively. Full details of the simulation study are in Appendix 2.

(II) Modelling population size structure for *Centrostephanus rodgersii* in Tasmania

The range extension of *C. rodgersii* in Tasmania has been well documented and therefore created an opportunity to test the model's ability to detect a known poleward range extension using size structure data. Ling et al. collected urchins from five locations/populations in 2004 and 2005 (Fig. 2a) on the east coast of Tasmania covering the latitude range -43.48° to -41.34° ("Growth and age across range extension region of *Centrostephanus rodgersii* in eastern Tasmania and morphometric comparison of urchins inhabiting kelp versus barrens habitats: dataset 3 - Allometry for conversion between jaw length and test diameter" accessed from the Australian Ocean Data Network). The urchin test diameters ranged from 33 to 133mm. Two of the locations/populations (St Helens Island and Elephant Rock) had a site for both barrens habitat and kelp habitat which were combined so there was only one dataset per location, each with 600 individuals sampled.

Cape Tourville and Mistaken Cape both had 300 individuals sampled and Fortescue had 282 individuals sampled.

(III) Using population size structure for *Centrostephanus rodgersii* in north-east New Zealand to infer population history and recruitment dynamics

We collected test diameter measurements (mm) of 647 *C. rodgersii* individuals from 14 locations along the north-east coast of the North Island of New Zealand (Fig. 3a). Individuals ranged from 32 to 130mm and were collected between 2015 and 2018. The majority of locations were sampled between late 2015 and early 2016, with a few from August and September 2017 and January 2018.

Finally, we used our model to assess patterns across north-east New Zealand's *C. rogersii* locations. We were interested in the trends between latitude and the location means and location standard deviations as well as the location coefficient of variation.

Results

(I) Simulation study

Overall, we found close agreement between the simulated parameters and the model estimates of these parameters across a range of realistic scenarios (Fig. 1, further details in Appendix 2). Though a more thorough sensitivity analysis is required, there was an indication that models slightly underestimated the strength of the relationship between latitude and the summary statistics when strong linear relationships were simulated (Fig. A2, Appendix 2). When a non-linear, convex relationship between latitude and the summary statistics was simulated, and a linear model fit to the simulated data, the relationship and the location estimates for the simulated mean test diameter were recovered, but the location estimates for the simulated standard deviations were not reliably recovered (Fig. A5). Thus, the model can reliably recover a linear relationship with locations means and/or location standard deviations.

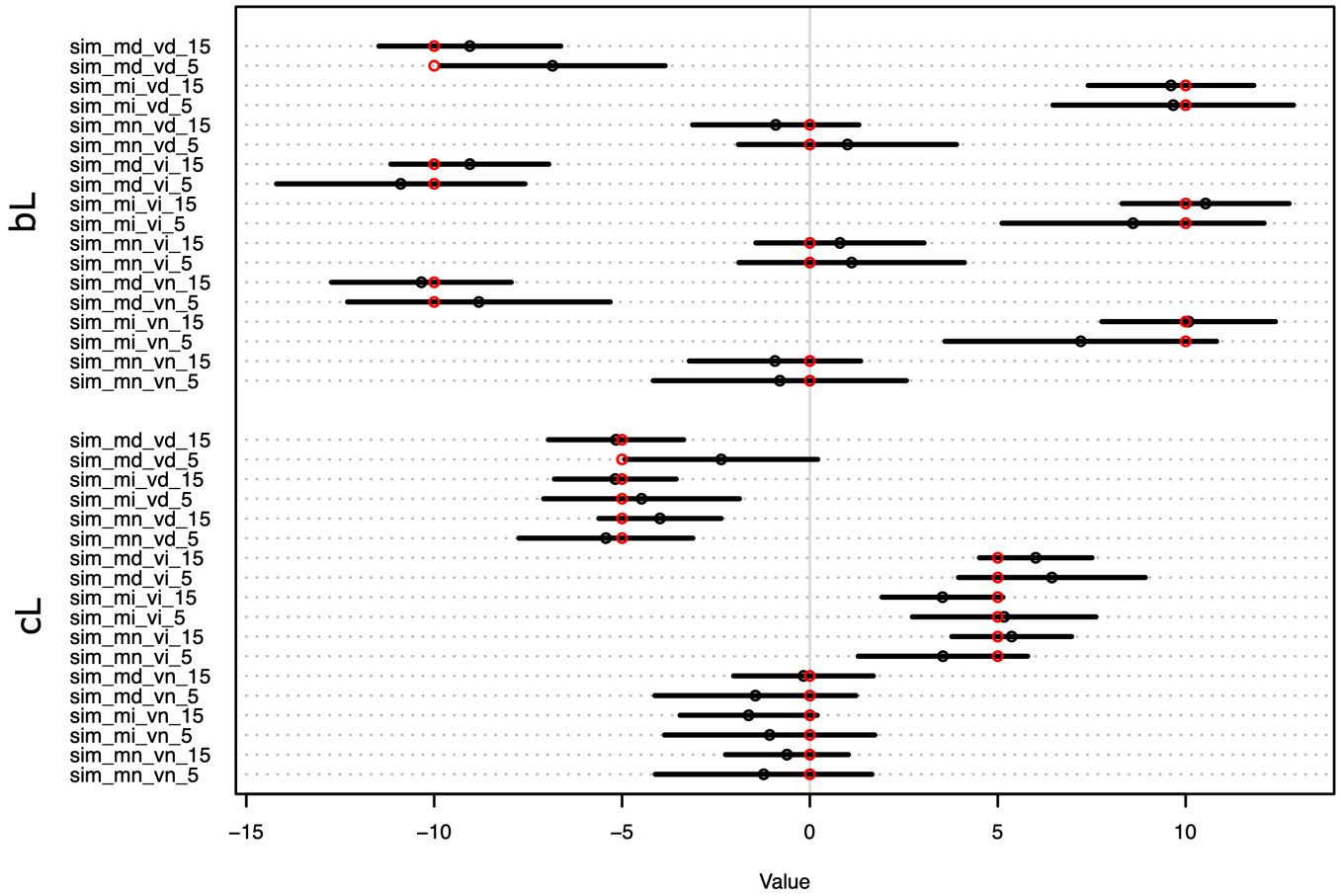

Figure 1: Plot of the regression coefficients bL (regression coefficient between latitude and the mean of test diameters of locations) and cL (regression coefficient between latitude and the variation of test diameters of locations). These were plotted for each of the nine simulated scenarios for both 15 location and 5 location datasets. The models are named based on the relationship between the location means and latitude ("md": bL = -10 ; "mi": bL = 10; "mn": bL = 0;) and the relationship between location standard deviations and latitude ("vd": cL = -5; "vi": cL = 5; "vn": cL = 0). The red circles are the simulated value for each of the parameters in each simulated scenario, the black circles are the model estimates and the lines are the credible intervals.

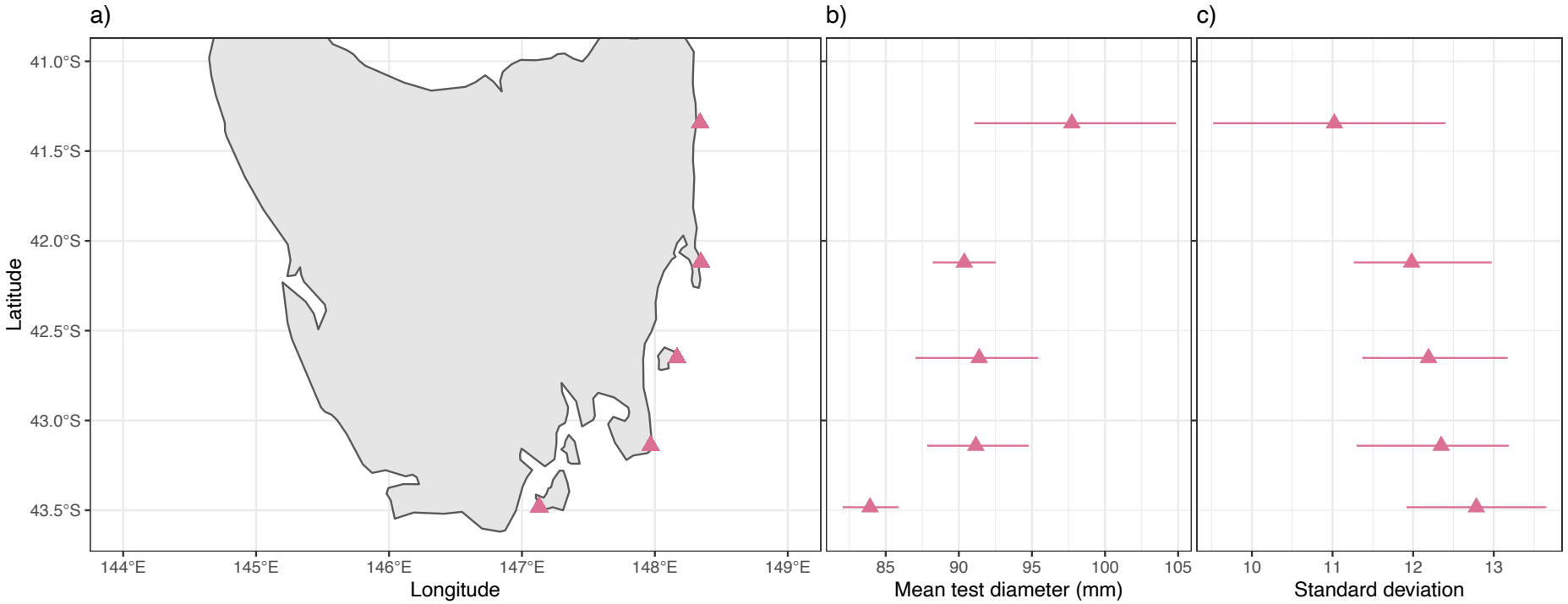

Figure 2: Results of the model for Tasmanian *Centrostephanus rodgersii* populations. a) Map of the locations sampled for *C. rodgersii* on the east coast of Tasmania; b) Observed mean test diameter for each location; c) Observed standard deviation of test diameter for each location. Lines indicate 95% credible intervals.

(II) Modelling population size structure for *Centrostephanus rodgersii* in Tasmania

Using the dataset for Tasmanian *C. rodgersii* populations, there was a positive trend between the location means and latitude (Fig. 2b). Although this trend was not significant (bL: mean= 1.77, 95% credible interval: -1.07 to 4.47) the result supports the ability of our model to infer signatures of the poleward range extension that has occurred. Furthermore, the trend between the location standard deviations and latitude was negative and significant (Fig. 2c; cL: mean= -0.75, 95% credible interval: -1.36 to -0.12), suggesting that regular recruitment is occurring at the high-latitude, southern locations.

(III) Using population size structure for *Centrostephanus rodgersii* in north-east New Zealand to infer population history and recruitment dynamics

In north-east New Zealand there was no trend between the location means and latitude overall (Fig. 3b; bL: mean= -1.79, 95% credible interval: -3.49 to 0.95). The pattern of location means was more varied in New Zealand than in Tasmania. At the northern locations, the means decreased with latitude indicative of a poleward range extension (to -36°), but at the southern latitudes there were varied location means without a trend. Similar to Tasmania, we found a significant negative relationship between the location standard deviations and latitude (Fig. 3c; cL: mean= -1.79, 95% credible interval: -3.09 to -0.48), suggesting southern locations could be recruiting more regularly.

Coefficient of variation

In both New Zealand and Tasmania, there was a negative trend between the coefficient of variation and latitude (Fig. 4). New Zealand had a steeper trend than Tasmania but it comprises more sampled locations and covers a greater absolute range in latitude. These results suggest that there is more regular recruitment occurring at more southern, high latitude locations for both Tasmania and New Zealand.

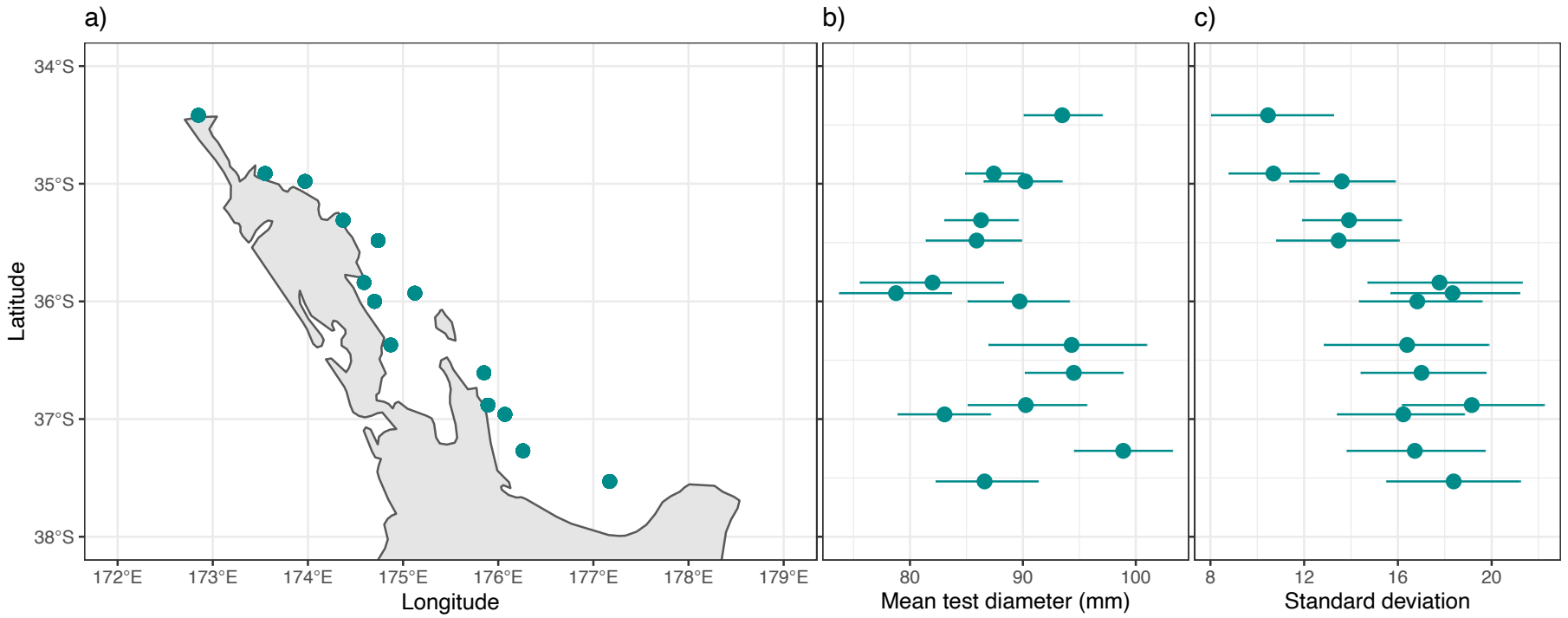

Figure 3: Results of the model for New Zealand *Centrostephanus rodgersii* populations. a) Map of the locations sampled for *C. rodgersii* in north-east New Zealand; b) Observed mean test diameter for each location; c) Observed standard deviation of test diameters for each location. Lines indicate 95% credible intervals.

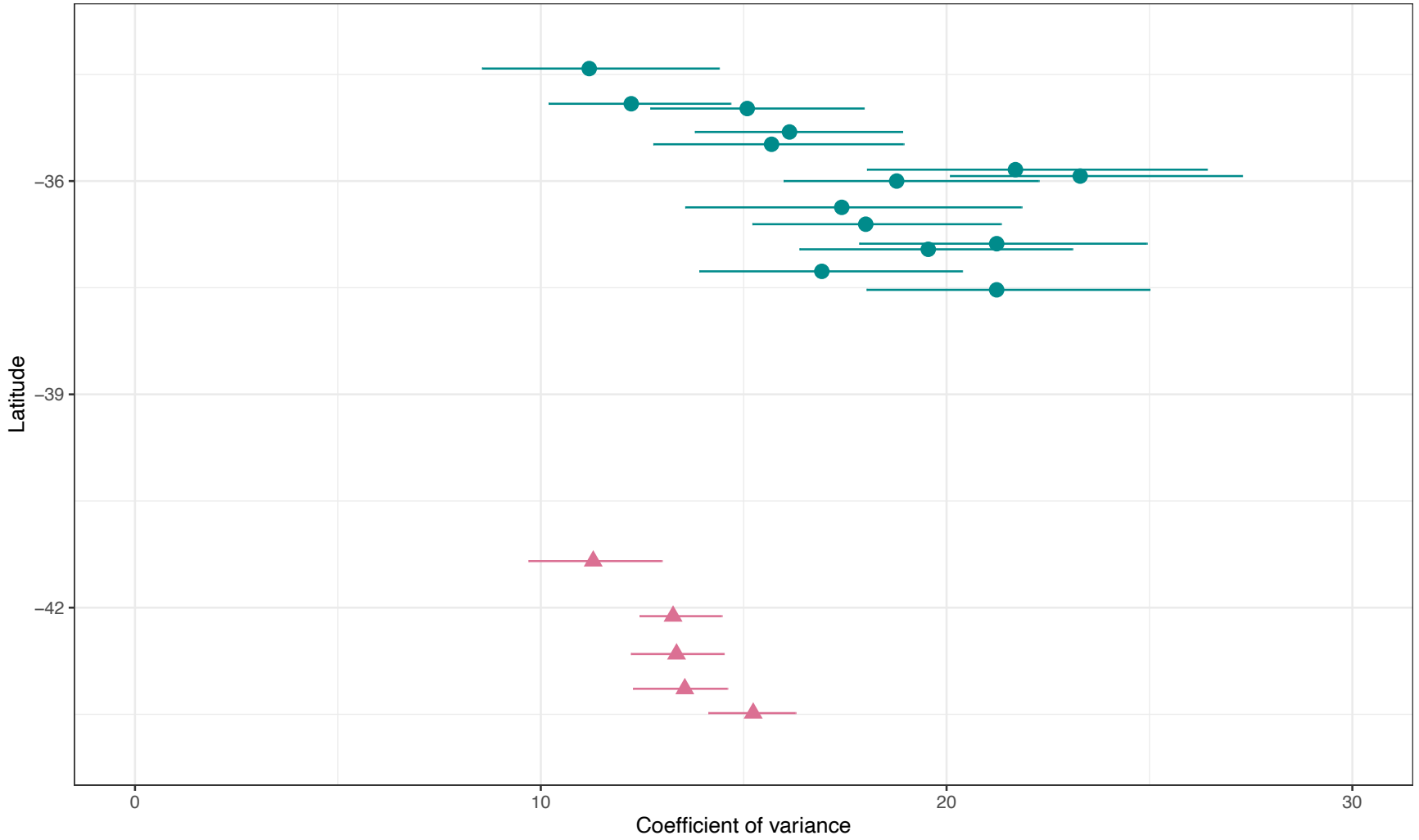

Figure 4: Coefficients of variation for each location plotted by latitude. Green is the north-east New Zealand locations and pink is the Tasmania locations. Lines indicate 95% credible intervals.

Discussion

Our study demonstrates that recent range history and regularity of recruitment into populations can be examined using size structure data when time-series data is not available. Our Bayesian modelling approach found patterns consistent with a range extension of *C. rodgersii* in New Zealand, based on measured size structure data across the species north-east New Zealand range. We first validated our model approach using simulated size structure data and verified its performance in recovering signatures of recent range extensions using size structure data from Tasmania, where *C. rodgersii* is documented to have undergone a range extension. Our results infer previously undescribed patterns of increased regularity of recruitment in a poleward direction in both New Zealand and Tasmania and no overall trend in the population means in New Zealand. We suggest that *C. rodgersii* in New Zealand has not had the same strictly poleward range extension in the same timeframe as Tasmania. Below we discuss the trends detected for *C. rodgersii* in north-east New Zealand, and the likely drivers underlying differences in the population histories of Tasmania and north-east New Zealand populations. We also discuss our approach with reference to other population size structure analyses, and limitations in such an approach.

The patterns in the mean sizes of individuals across populations demonstrate some differences in recent range shifts/colonisation of *C. rodgersii* between Tasmania and New Zealand. Although not statistically significant, we found that the mean sizes of individuals in Tasmanian populations decreased in a poleward direction which fits the documented population history of this species (Ling et al. 2009). The underlying cause of the pattern in mean sizes is likely due to changes in the EAC. In Tasmania, Ling et al. (2009) described an exponential decline in mean urchin age with distance from the western edge of EAC. The EAC travels in a poleward direction and therefore the range extension in Tasmania is associated with latitude.

In contrast to Tasmania, we did not find an overall trend in the mean sizes of individuals across populations associated with latitude in New Zealand. In the north of *C. rodgersii*'s north-east New Zealand range, the mean sizes decreased with latitude, but from approximately -36° to East Cape, there was no trend with latitude. It is important to note that the number of locations and the breadth of latitudes sampled in New Zealand was greater than in Tasmania, and our findings likely correspond to the different physical environments found across latitude in north-east New Zealand. Latitude itself does not drive biological processes; rather, it is the variables that are often correlated with latitude, such as temperature and ocean currents. In this case, unlike the EAC in Australia, the EAuC in north-east New Zealand does not track the coastline, but stays offshore, and more frequently meets the East Cape than more northerly locations (Stanton et al. 1997, Sutton & Bowen 2019). Therefore, a range extension influenced by ocean currents in north-east New Zealand may not be strictly poleward, linear, and associated with latitude. In southern regions of north-east New Zealand, the range extension of *C. rodgersii* may not be recent enough to detect using an analysis of population size structure. *Centrostephanus rodgersii* has been present in New Zealand a lot longer than in Tasmania (Farquhar 1897; Johnson et al. 2005). The size of *C. rodgersii* is likely to reflect their age up to 15 years old, after which their growth slows (estimated from Jolicoeur tag–recapture in Pecorino et al. 2012).

Therefore, all the individuals that are over approximately 15 years old (arrived before approximately 2001) are indifferentiable using size so range extensions and/or recruitment patterns before 2001 cannot be detected using this approach.

For both Tasmania and New Zealand, the standard deviations of sizes within populations increased in a poleward direction, suggesting that recruitment is more regular in higher latitude populations. In Tasmania, *C. rodgersii* has more recently colonised southern locations; therefore, we expected southern locations to have smaller standard deviations of sizes due to having fewer older individuals but, instead, they had a greater standard deviation of sizes. This counterintuitive result may be due to currents and marine heatwave events affecting recruitment patterns. In Tasmania, the annual length of marine heatwaves (periods of abnormally high sea temperature) varies in two ways (Oliver et al. 2018). First, annual variation is associated with southward along-shore circulation which comes from the EAC, predominantly impacting northern populations. Second, there is a less frequent multi-year variability which mainly affects south and south-eastern Tasmania. This multi-year variability is associated with weak currents over the eastern continental shelf which could in turn impact recruitment at the multi-year scale. When the studied urchins (measured in 2004 and 2005) were recruiting between 1993 and 2000 there was a multi-year increase in marine heatwave events (Oliver et al. 2018), which could have led to regular recruitment in the southern populations, leading to larger standard deviations of sizes. If the population was to be resurveyed now, the patterns in standard deviation may be quite different due to the heatwave in 2015/16 from the EAC that likely impacted northern populations (Oliver et al. 2017).

Similar to Tasmania, in New Zealand, the pattern in standard deviations may be due to the variability in the EAuC affecting recruitment of *C. rodgersii*. The EAuC varies from year to year and regularly will reach the more southern locations but the northern locations are not always reached or only reached later in the season (Stanton et al. 1997). Therefore, urchins will likely be recruited to southern populations each year but recruitment may not always reach the northern populations.

The coefficient of variation of *C. rodgersii* populations in both Tasmania and north-east New Zealand was low compared to other studies. In a study of the purple sea urchin (*S. purpuratus*) along the coast of California and Oregon, the coefficients of variation ranged from 18.5% to 46.7% and were considered “consistently poor” and “consistently good”, respectively (Ebert & Russell 1988). Similarly, a study on the small giant clam (*T. maxima*) in Western Australia recovered coefficients of variation between 23.54% and 44.45% (Black et al. 2011). In contrast, the coefficients of variation for New Zealand populations of *C. rodgersii* ranged from 11.19% to 23.38%, and between 11.28% and 15.23% for Tasmanian populations. It may be that the low coefficients of variance we found were due to the difficulty in surveying the very small urchins in the population. Very young urchins tend to have more cryptic behaviour, such as hiding in crevices away from predators (Ling & Johnson 2012, Byrne & Andrew 2020) and these small urchins could have been missed in the surveys. Hence, our population mean sizes could be overestimated, our population standard deviations of size may be underestimated, and our population coefficient of variation underestimated. Although such a survey bias compromises

our ability to compare our recovered coefficients of variation to other studies, it does not impact our study as this bias would have been consistent across all studied locations within New Zealand, and likely within Tasmania.

Although we recovered an increase in the coefficients of variation for populations in the poleward direction, for both Tasmania and New Zealand, this provided little more information than provided by the pattern in standard deviations of sizes. In the case of Tasmania, if only the coefficients of variation were used in our analysis, important demographic insights may have been missed, such as the trend in the mean size of individuals in Tasmania indicating a poleward range extension. Thus, the separation of the regularity of recruitment (i.e., standard deviation in sizes) and the length of populations presence (i.e., mean size) may be important when addressing demographic scenarios that potentially include range extensions (also suggested by Zacherl et al. 2003).

Although a range extension of *C. rodgersii* in north-east New Zealand may be unrelated to latitude, but associated with a more heterogeneous variable (such as sea surface temperature or physical oceanography), it is also plausible that no recent range extension has occurred in north-east New Zealand. *Centrostephanus rodgersii* has been present some locations of north-east New Zealand for at least 120 years (Farquhar 1897) which is much longer than the 40 years since they have been recorded in Tasmania (Johnson et al. 2005). Based on historical records in New Zealand (Farquhar 1897, Fell 1949) and sex ratios at the Poor Knights Islands (Pecorino et al. 2013c), it is possible that the species extended its range throughout many of the studied locations before 2001 (i.e. prior to the timeframe we can analyse in this study). Regardless, further range extension is possible in future. In New Zealand, the distribution of *C. rodgersii* indicates that the species is limited by winter sea surface temperatures of 15°C (Pecorino et al., 2013c). Sea surface temperatures around New Zealand are predicted to increase (Law et al. 2017), which could promote a future range extension of the species (Pecorino et al. 2013a). For instance, Tasmania is at a higher latitude than the New Zealand locations sampled, but is a renowned global warming hotspot, warming at a much greater rate than north-east New Zealand (Shears & Bowen 2017, Sutton & Bowen 2019). Therefore, it is likely the range shifts of *C. rodgersii* differ between Tasmania and New Zealand owing to the different time periods studied, as well as the velocity of ocean climate change in each region.

We have demonstrated that modelling the mean sizes and standard deviations in sizes of individuals in a population can be used to study population demography and recruitment, particularly when time-series data are not available. In the context of the climate crisis, it is important to have a range of methods to understand the range history of species. *Centrostephanus rodgersii* has the potential to detrimentally impact New Zealand's biodiversity and fisheries due to our coastal marine ecosystem being similar to Tasmania (Johnson et al. 2005, Ling et al. 2008, Ling et al. 2009, Lisson 2018). Formal time-series surveys across *C. rodgersii*'s range would give us the fullest and most certain understanding of any future range extension, as well as increases in abundance and density in New Zealand over time, beyond the 15 year post-range extension limit of our approach. However, in many cases, time and/or resources to undertake structured time-series surveys are limited, especially as ocean climate

changes are already impacting our ecosystems; in these situations, the method presented here provides a more immediate and less time-intensive way to infer recent demographic histories, recruitment patterns, and to inform impact assessments and management decisions.

Chapter Two Appendices:

Appendix 1: Supplementary information and tables

Bayesian model written in equation form: (parameters explained in Table A1)

$$TD_i \sim Normal\ (\mu_i,\ \sigma_i)$$
$$\mu_i = \mu_a + \beta_{L\,[i]} sL + z[P]\,\sigma_a$$
$$\mu_a \sim Normal\ (90{,}10)$$
$$\beta_L \sim Normal(0, 5)$$
$$z[P] \sim Normal\ (0, 1)$$
$$\sigma_a \sim Exp(2)$$
$$\sigma_i = \mu_c + c_{L\,[i]} sL + z2[P]\,\sigma_c$$
$$\mu_c \sim Normal\ (15{,}5)$$
$$c_L \sim Normal(0, 5)$$
$$z2[P] \sim Normal\ (0, 1)$$
$$\sigma_c \sim Exp(2)$$

Parameter	Explanation	Prior/value
TD	Test diameter of urchins in mm (individual size)	
sL	Standardised latitude	
P	Population/location	
μ	Mean test diameter of a location	
μ_a	Global mean test diameter	Mean: 90; standard deviation: 10
β_L	Slope of the relationship between standardised latitude and the location mean of test diameter	Mean: 0; standard deviation: 5
σ_a and z[P]	Error term for the mean test diameter of a location using non-centered parameterisation. The z[P] allows the error to differ for each location	z: Mean: 0; standard deviation: 1 σ_a : Exp(2)
σ	Standard deviation of test diameters of a location	
μ_c	Global standard deviation of test diameter	Mean: 15; standard deviation: 5
c_L	Slope of the relationship between standardised latitude and the location standard deviation of test diameter	Mean: 0; standard deviation: 5
σ_c and z2[P]:	Error term for the standard deviation of test diameters of a location using non-centered parameterisation. The z2[P] allows the error to differ for each location	z2: Mean: 0; standard deviation: 1 σ_c : Exp(2)

Table A1: Explanation of each of the parameters in the model.

Appendix 2: Simulation study examining relationships between latitude and location means as well as location standard deviations of sea urchin test diameters.

Background: The Long-spined sea urchin (*Centrostephanus rodgersii*) is found in both Tasmania and north-east New Zealand. In Tasmania, the Long-spined sea urchin extended its range poleward causing dramatic ecosystem change and socio-economic challenges associated with the collapse of lobster and abalone fisheries. Researchers in Tasmania observed that more southerly populations of the range extending *C. rodgersii* had smaller mean test diameters (size) and a younger mean age than more northerly populations.

It is currently unknown whether *C. rodgersii* is undergoing a similar range extension in north-east New Zealand. We used test diameter measurements from populations of the Long-spined sea urchin distributed along a 7° latitudinal gradient from Spirits Bay to White Island (Whakaari) to investigate if we would also observe a decline in mean test diameter with increased southerly latitude. In particular, we were interested in examining if latitude is associated with a change in the mean test diameters as well as a change in the standard deviation of test diameters.

We felt that flexibility afforded by the R package Rethinking (McElreath 2020) would allow us to estimate the location means and standard deviations of test diameters and regress these estimates on latitude within a single model. We were unable to identify a suitable example of a model parameterization that would allow us to run these models in the literature. Therefore, we performed a simulation study to confirm that our model parameterization was appropriate for testing our hypotheses relating to relationships between location specific summary statistics and latitude.

Summary: Overall, we found close agreement between the simulated location parameters and the model estimates of these parameters across a range of realistic scenarios. Though a more thorough sensitivity analysis is required, there was an indication that models slightly underestimated the strength of the relationship between latitude and the summary statistics when strong linear relationships were simulated. When a non-linear, convex, relationship between latitude and the summary statistics was simulated, and a linear model fit to these data, the relationship and the location estimates for the simulated mean test diameter were recovered, but the location estimates for the simulated standard deviations were not reliably recovered.

Simulation study: For each scenario below we simulated two data sets with two different sample sizes: five locations (ending in "_5") and fifteen locations (ending in "_15"). Within each location we simulated test diameter values for 30 individuals across all simulations.

Our nine different scenarios were:

sim_mn_vn_5/sim_mn_vn_15: no relationship between location means and latitude ("mn": bL = 0) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_mi_vn_5/sim_mi_vn_15: positive relationship between location means and latitude ("mi": bL = 10) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_md_vn_5/sim_md_vn_15: negative relationship between location means and latitude ("md": bL = -10) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_mn_vi_5/sim_mn_vi_15: no relationship between location means and latitude ("mn": bL = 0) and positive relationship between location standard deviations and latitude ("vi": cL = 5)

sim_mi_vi_5/sim_mi_vi_15: positive relationship between location means and latitude ("mi": bL = 10) and positive relationship between location standard deviations and latitude ("vi": cL = 5)

sim_md_vi_5/sim_md_vi_15: negative relationship between location means and latitude ("md": bL = -10) and positive relationship between location standard deviations and latitude ("vi": cL = 5)

sim_mn_vd_5/sim_mn_vd_15: no relationship between location means and latitude ("mn": bL = 0) and negative relationship between location standard deviations and latitude ("vd": cL = -5)

sim_mi_vd_5/sim_mi_vd_15: positive relationship between location means and latitude ("mi": bL = 10) and negative relationship between location standard deviations and latitude ("vd": cL = -5)

sim_md_vd_5/sim_md_vd_15: negative relationship between location means and latitude ("md": bL = -10) and negative relationship between location standard deviations and latitude ("vd": cL = -5)

For each simulated dataset we then fit a model with the following parameters:

TD: test diameter of urchins in mm

sL: standardised latitude

P: Urchin location ID

mu: mean test diameter of a location

mu_a: global mean test diameter

bL: slope of the relationship between standardised latitude and the location mean of test diameter

sigma_a and z[P]: error term for the mean test diameter of a location using non-centered parameterisation. The z[P] allows the error to differ for each location

sigma: standard deviation of test diameters of a location

mu_c: global standard deviation of test diameter

cL: slope of the relationship between standardised latitude and the location standard deviation of test diameter

signma_c and z2[P]: error term for the standard deviation of test diameters of a location using non-centered parameterisation. The z2[P] allows the error to differ for each location

Data simulation and model function

The function below was written using the ulam function from the rethinking package.

```
data.sim_out <- function(pop, ind, lat, mu, sig){

  diam_m.v <- cbind(seq(mu[1], mu[2], length = pop), seq(sig[1], sig[2], leng
th = pop))

  dat <- list(sL = rep(seq(lat[1], lat[2], length = pop), each = ind), TD = c
(apply(diam_m.v, 1, function(x) rnorm(ind, x[1], (x[2])))), P = rep(seq(1, po
p), each = ind))

  out <- ulam(

    alist(

      TD ~ dnorm(mu , sigma),

      mu <- mu_a + bL*sL + z[P]*sigma_a,

      z[P] ~ dnorm( 0, 1),

      sigma_a ~ dexp(2),

      mu_a ~ dnorm(90, 10),

      bL ~ dnorm( 0 , 5 ),

      sigma <- mu_c + cL*sL + z2[P]*sigma_c,

      cL ~ dnorm( 0 , 5 ),

      mu_c ~ dnorm(15, 5),

      z2[P] ~ dnorm( 0, 1),

      sigma_c ~ dexp(2)

    ), data= dat , chains=4, cores=1 , control=list(adapt_delta=0.99, max_tre
edepth = 15), log_lik = TRUE, iter=3000, constraints=list( mu_c="lower=0", mu
_a="lower=0"))

  out

}
```

Simulating data and generating model output for each of the 9 scenarios for both 5 and 15 locations:

```
sim_mn_vn_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(90,90), sig = c(15,15))
sim_mi_vn_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(80,100), sig = c(15,15))
sim_md_vn_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(100,80), sig = c(15,15))
sim_mn_vi_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(90,90), sig = c(10,20))
sim_mi_vi_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(80,100), sig = c(10,20))
sim_md_vi_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(100,80), sig = c(10,20))
sim_mn_vd_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(90,90), sig = c(20,10))
sim_mi_vd_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(80,100), sig = c(20,10))
sim_md_vd_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(100,80), sig = c(20,10))

sim_mn_vn_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(90,90), sig = c(15,15))
sim_mi_vn_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(80,100), sig = c(15,15))
sim_md_vn_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(100,80), sig = c(15,15))
sim_mn_vi_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(90,90), sig = c(10,20))
sim_mi_vi_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(80,100), sig = c(10,20))
sim_md_vi_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(100,80), sig = c(10,20))
sim_mn_vd_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(90,90), sig = c(20,10))
sim_mi_vd_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(80,100), sig = c(20,10))
sim_md_vd_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(100,80), sig = c(20,10))
```

Function to plot regression coefficients

The following function based of the coeftab_plot function plots the regression coefficients (± CI) for the relationship between latitude and the location mean (bL) or location standard deviation (cL) in each scenario (rows) as well as the simulated bL an cL

```
coeftab_plot_with_sim_estimates <- function (x, pars, sim.est, col.ci = "blac
k", by.model = FALSE, prob = 0.95, xlab = "Value", cex) {

  xse <- x@se

  x <- x@coefs

  if (!missing(pars)) {

    x <- x[pars, ]

    xse <- xse[pars, ]

    sim.est <- sim.est[pars, ]

  }

  if (by.model == FALSE) {

    xse <- t(xse)

    x <- t(x)

    sim.est <- t(sim.est)

  }

  z <- qnorm(1 - (1 - prob)/2)

  left <- x

  right <- x

  for (k in 1:nrow(x)) {

    for (m in 1:ncol(x)) {

      ci <- x[k, m] + c(-1, 1) * z * xse[k, m]

      left[k, m] <- ci[1]

      right[k, m] <- ci[2]

    }
```

```
  }

  llim <- min(left, na.rm = TRUE)

  rlim <- max(right, na.rm = TRUE)

  dotchart(x, xlab = xlab, xlim = c(llim, rlim), cex=cex)

  for (k in 1:nrow(x)) {

    for (m in 1:ncol(x)) {

      if (!is.na(left[k, m])) {

        kn <- nrow(x)

        ytop <- ncol(x) * (kn + 2) - 1

        ypos <- ytop - (m - 1) * (kn + 2) - (kn - k + 1)

        lines(c(left[k, m], right[k, m]), c(ypos, ypos), lwd = 2, col = col.c
i)

        points(sim.est[k,m], ypos, cex = cex, pch = 1, col = "red")

      }

    }

  }

  abline(v = 0, lty = 1, col = col.alpha("black", 0.15))

}
```

Plotting the nine scenarios for both 5 and 15 locations

```
md <- -10
mn <- -0
mi <- 10
vd <- -5
vn <- 0
vi <- 5

x_sim.est <- matrix(1:36,2,18)
x_sim.est[1,]<-rep(c(mn,mn,mi,mi,md,md), times = 3)
x_sim.est[2,]<-rep(c(vn,vi,vd), each = 6)
rownames(x_sim.est) <- paste(c("bL","cL"))
```

```
coeftab_plot_with_sim_estimates(coeftab(sim_mn_vn_5, sim_mn_vn_15, sim_mi_vn_5, sim_mi_vn_15, sim_md_vn_5, sim_md_vn_15, sim_mn_vi_5, sim_mn_vi_15, sim_mi_vi_5, sim_mi_vi_15, sim_md_vi_5, sim_md_vi_15, sim_mn_vd_5, sim_mn_vd_15, sim_mi_vd_5, sim_mi_vd_15, sim_md_vd_5, sim_md_vd_15), pars = c("bL", "cL"), sim.est=x_sim.est, cex=0.5)
```

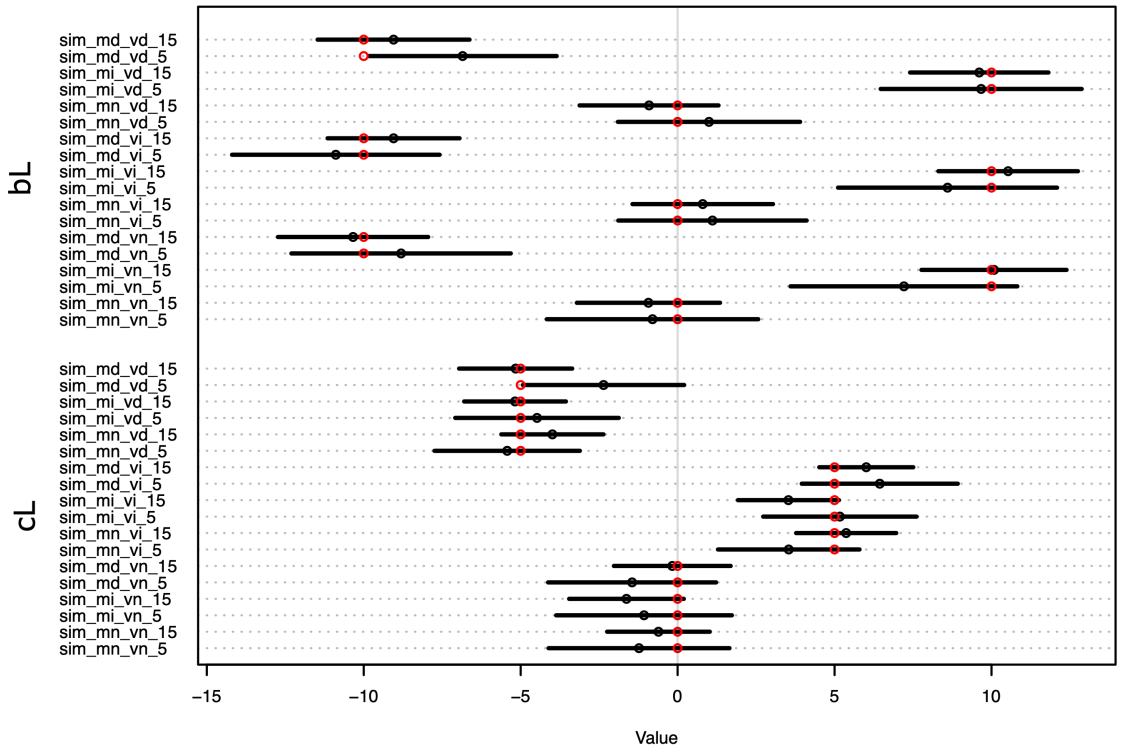

Figure A1: Plot of the regression coefficients bL (regression coefficient between latitude and the mean of test diameters of locations) and cL (regression coefficient between latitude and the variance of test diameters of locations). These were plotted for each of the nine simulated scenarios for both 15 location and 5 location datasets. The red circles are the simulated value for each of the parameters in each simulated senario, the black circles are the model estimates and the lines are the credible intervals. Note: this is the same figure as Fig. 1 Chapter Two.

In summary, Figure 1 shows there was good agreement between the simulation parameters and the estimated coefficients and that accuracy and precision of our estimated coefficients improves with greater samples sizes (i.e. comparing the accuracy of the 5 and 15 location datasets for the same scenario).

Testing different magnitudes of the regression parameters

Next, we explored if our model would be sensitive to differences in magnitude of the regression parameters as well as differences in direction. Hence, we simulated six additional datasets under the following scenarios for both 5 (ending in _5) and 15 locations (ending in _15):

Changing the magnitude of bL:

sim_mi_low_vn_5/sim_mi_low_vn_15: low positive relationship between location means and increasing latitude ("mi_low": bL = 5) and no relationship between location variances latitude ("vn": cL = 0)

sim_mi_med_vn_5/sim_mi_med_vn_15: medium positive relationship between location means and increasing latitude ("mi_med": bL = 10) and no relationship between location variances latitude ("vn": cL = 0)

sim_mi_hig_vn_5/sim_mi_hig_vn_15: high positive relationship between location means and increasing latitude ("mi_hig": bL = 20) and no relationship between location variances latitude ("vn": cL = 0)

Changing the magnitude of cL:

sim_mn_vi_low_5/sim_mn_vi_low_15: no relationship between location means and latitude ("mn": bL = 0) and low positive relationship between location standard deviations and latitude ("vi_low": cL = 2)

sim_mn_vi_med_5/sim_mn_vi_med_15: no relationship between location means and latitude ("mn": bL = 0) and medium positive relationship between location standard deviations and latitude ("vi_med": cL = 5)

sim_mn_vi_hig_5/sim_mn_vi_hig_15: no relationship between location means and latitude ("mn": bL = 0) and high positive relationship between location standard deviations and latitude ("vi_hig": cL = 10)

Simulating the scenarios for changes in magnitudes of the regression parameters

```
sim_mi_low_vn_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(85,
95), sig = c(15,15))
sim_mi_med_vn_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(80,
100), sig = c(15,15))
sim_mi_hig_vn_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(70,
110), sig = c(15,15))

sim_mn_vi_low_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(90,
90), sig = c(13,17))
sim_mn_vi_med_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(90,
90), sig = c(10,20))
sim_mn_vi_hig_5 <- data.sim_out(pop = 5, ind = 30, lat = c(-1, 1), mu = c(90,
90), sig = c(5,25))

sim_mi_low_vn_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(8
5,95), sig = c(15,15))
sim_mi_med_vn_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(8
0,100), sig = c(15,15))
sim_mi_hig_vn_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(7
0,110), sig = c(15,15))

sim_mn_vi_low_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(9
0,90), sig = c(13,17))
sim_mn_vi_med_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(9
0,90), sig = c(10,20))
sim_mn_vi_hig_15 <- data.sim_out(pop = 15, ind = 30, lat = c(-1, 1), mu = c(9
0,90), sig = c(5,25))
```

Plotting regression coefficients (± CI) for the relationship between latitude and the location mean (bL) or location standard deviation (cL) for scenarios with different magnitudes of the regression coefficient (rows)

```
x_sim.est <- matrix(1:24,2,12)
x_sim.est[1,]<-c(0,0,0,0,0,0,5,5,10,10,20,20)
x_sim.est[2,]<-c(2,2,5,5,10,10,0,0,0,0,0,0)
rownames(x_sim.est) <- paste(c("bL","cL"))
colnames(x_sim.est) <- letters[1:12]

coeftab_plot_with_sim_estimates(coeftab(sim_mn_vi_low_5, sim_mn_vi_low_15, si
m_mn_vi_med_5, sim_mn_vi_med_15, sim_mn_vi_hig_5, sim_mn_vi_hig_15, sim_mi_lo
w_vn_5, sim_mi_low_vn_15, sim_mi_med_vn_5, sim_mi_med_vn_15, sim_mi_hig_vn_5,
sim_mi_hig_vn_15), pars = c("bL", "cL"), sim.est=x_sim.est, cex=0.5)
```

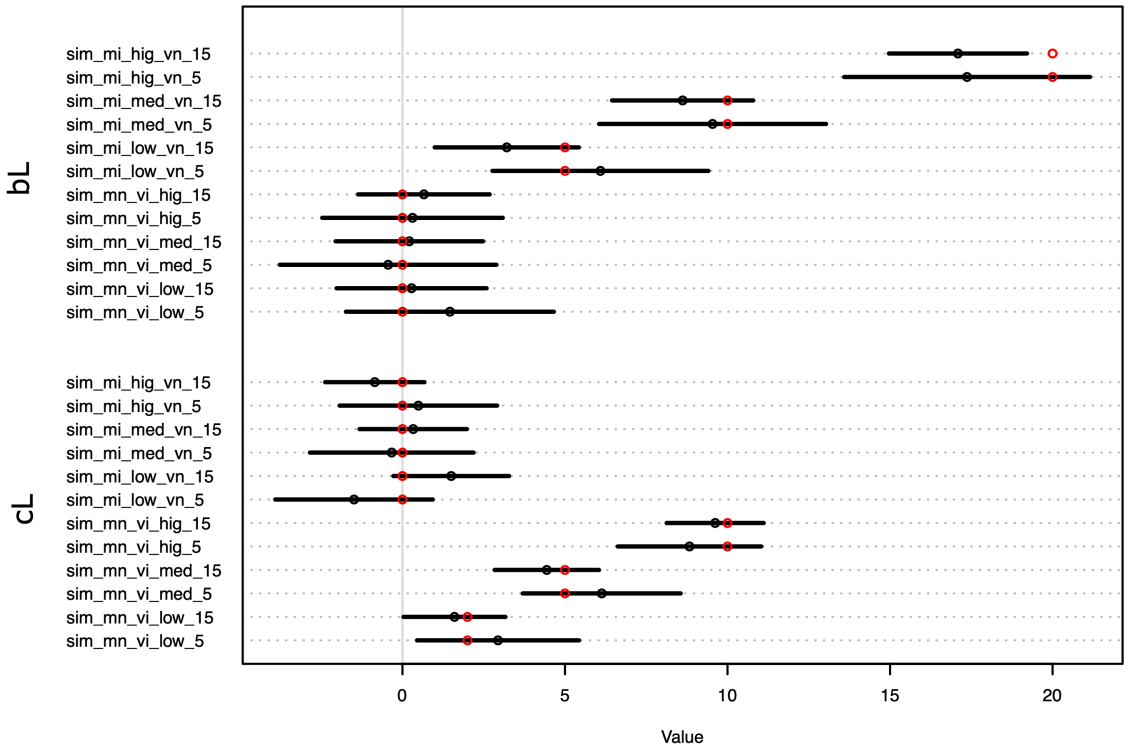

Figure A2: Plot of the regression coefficients bL (regression coefficient between latitude and the mean of test diameters of locations) and cL (regression coefficient between latitude and the variance of test diameters of locations). These were plotted for each of the 3 simulated scenarios for the change in the magnitude of bL and the change in the magnitude of cL for both 15 location and 5 location datasets. The red circles are the simulated values for each parameter in each simulated scenario, the black circles are the model estimates and the lines are the credible intervals.

In summary, Figure 2 shows there was good agreement between the simulation parameters and the estimated coefficients for all models except for the estimate of bL in model sim_mi_hig_vn_15 suggesting that that for steeper slopes the model estimates tend to underestimate the magnitude of the relationship. We will perform a follow up sensitivity analysis to examine the generality of this result. Nevertheless, although the sim_mi_hig_vn_15 credible interval of bL did not include the simulated bL it did have a greater magnitude of bL than sim_mi_med_vn_15 so the estimates showed an increase in magnitude.

Recovering location means and location standard deviations

Next, examined if the estimates of the location means and location standard deviations were reliably recovered.

Functions to extract the location means and location standard deviation from the model output:

```
all_model_samples<- function(modelname){
  p<-extract.samples(modelname)
  mu_out <- matrix(nrow= nrow(p$z), ncol = ncol(p$z))
  sigma_out <- matrix(nrow= nrow(p$z), ncol = ncol(p$z))
  for (i in 1:nrow(p$z)){
    mu_out[i,] <- p$mu_a[i] + p$bL[i]*unique(modelname@data$sL) +  p$z[i,]*as
.numeric(p$sigma_a[i])
    sigma_out[i,] <- p$mu_c[i] + p$cL[i]*unique(modelname@data$sL) +  p$z2[i,
]*as.numeric(p$sigma_c[i])
  }
  list(mu_out,sigma_out)
}

means_TD<-function(modelname){
  mu<-apply(all_model_samples(modelname)[[1]],2,mean)
  sd<-apply(all_model_samples(modelname)[[2]],2,mean)
  list(mu,sd)
}

model_credible_intervals_means<-function(modelname){
  samples_mu<-all_model_samples(modelname)[[1]]
  samples_sd<-all_model_samples(modelname)[[2]]
  ci_out_mu<-matrix(nrow=2,ncol=length(unique(modelname@data$P)))
  ci_out_sd<-matrix(nrow=2,ncol=length(unique(modelname@data$P)))
  for (i in 1:length(unique(modelname@data$P))){
    ci_out_mu[1,i]<-PI(samples_mu[,i], prob=0.95)[1]
   ci_out_mu[2,i]<-PI(samples_mu[,i], prob=0.95)[2]
   ci_out_sd[1,i]<-PI(samples_sd[,i], prob=0.95)[1]
   ci_out_sd[2,i]<-PI(samples_sd[,i], prob=0.95)[2]
  }
  list(ci_out_mu, ci_out_sd)
}
```

Plotting the estimated location means and location standard deviations for different directional relationships between the location means and latitude:

Given that we had good agreement between the simulated parameters and the model regression parameters, we explored the accuracy of the model estimates of the location means and standard deviations.

First, we plotted the estimated location means and location standard deviations for simulated datasets with 15 populations but different relationships between the latitude and location

means (positive, negative and none) and no relationship between latitude and locations standard deviations. The following models were plotted:

sim_mn_vn_15: no relationship between location means and latitude ("mn": bL = 0) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_mi_vn_15: positive relationship between location means and latitude ("mi": bL = 10) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_md_vn_15: negative relationship between location means and latitude ("md": bL = -10) and no relationship between location standard deviations and latitude ("vn": cL = 0)

Then, we plotted the estimated location means and locations standard deviations for simulated datasets with 15 populations but different relationships between the latitude and location standard deviations (postive, negative and none) and no relationship between latitude and locations means.

sim_mn_vn_15: no relationship between location means and latitude ("mn": bL = 0) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_mn_vi_15: no relationship between location means and latitude ("mn": bL = 0) and positive relationship between location standard deviations and latitude ("vi": cL = 5)

sim_mn_vd_15: no relationship between location means and latitude ("mn": bL = 0) and negative relationship between location standard deviations and latitude ("vd": cL = -5)

```
plot_location_means_and_sd<-function(modelname, start, end, colour, m_sd){poi
nts(unique(modelname@data$sL), means_TD(modelname)[[m_sd]], col=colour, pch=1
6)
arrows(unique(modelname@data$sL),model_credible_intervals_means(modelname)[[m
_sd]][1,],y1=model_credible_intervals_means(modelname)[[m_sd]][2,],length = 0
, col=colour)
points(seq(-1,1,length=length(unique(modelname@data$P))),seq(start, end,lengt
h=length(unique(modelname@data$P))), col=colour, pch=1)
}

#m_sd=1 for location means and m_sd=2 for location standard deviations

par(mfrow = c(2, 2), mar = c(4,4,2,2))

plot(NULL, xlim=c(-1,1), ylim=c(75,105), xlab="latitude", ylab="test diameter
", cex.axis=1.1, cex.lab=1.1, main="a", cex.main=1.3)
plot_location_means_and_sd(modelname=sim_mn_vn_15, start=90, end=90, colour="
blue", m_sd=1)
plot_location_means_and_sd(modelname=sim_mi_vn_15, start=80, end=100, colour=
"magenta", m_sd=1)
plot_location_means_and_sd(modelname=sim_md_vn_15, start=100, end=80, colour=
"green", m_sd=1)

plot(NULL, xlim=c(-1,1), ylim=c(5,25), xlab="latitude", ylab="standard deviat
```

```
ion", cex.axis=1.1, cex.lab=1.1, main="b", cex.main=1.3)
plot_location_means_and_sd(modelname=sim_mn_vn_15, start=15, end=15, colour="
blue", m_sd=2)
plot_location_means_and_sd(modelname=sim_mi_vn_15, start=15, end=15, colour="
magenta", m_sd=2)
plot_location_means_and_sd(modelname=sim_md_vn_15, start=15, end=15, colour="
green", m_sd=2)

plot(NULL, xlim=c(-1,1), ylim=c(75,105), xlab="latitude", ylab="test diameter
", cex.axis=1.1, cex.lab=1.1, main="c", cex.main=1.3)
plot_location_means_and_sd(modelname=sim_mn_vn_15, start=90, end=90, colour="
blue", m_sd=1)
plot_location_means_and_sd(modelname=sim_mn_vi_15, start=90, end=90, colour="
magenta", m_sd=1)
plot_location_means_and_sd(modelname=sim_mn_vd_15, start=90, end=90, colour="
green", m_sd=1)

plot(NULL, xlim=c(-1,1), ylim=c(5,25), xlab="latitude", ylab="standard deviat
ion", cex.axis=1.1, cex.lab=1.1, main="d", cex.main=1.3)
plot_location_means_and_sd(modelname=sim_mn_vn_15, start=15, end=15, colour="
blue", m_sd=2)
plot_location_means_and_sd(modelname=sim_mn_vi_15, start=10, end=20, colour="
magenta", m_sd=2)
plot_location_means_and_sd(modelname=sim_mn_vd_15, start=20, end=10, colour="
green", m_sd=2)
```

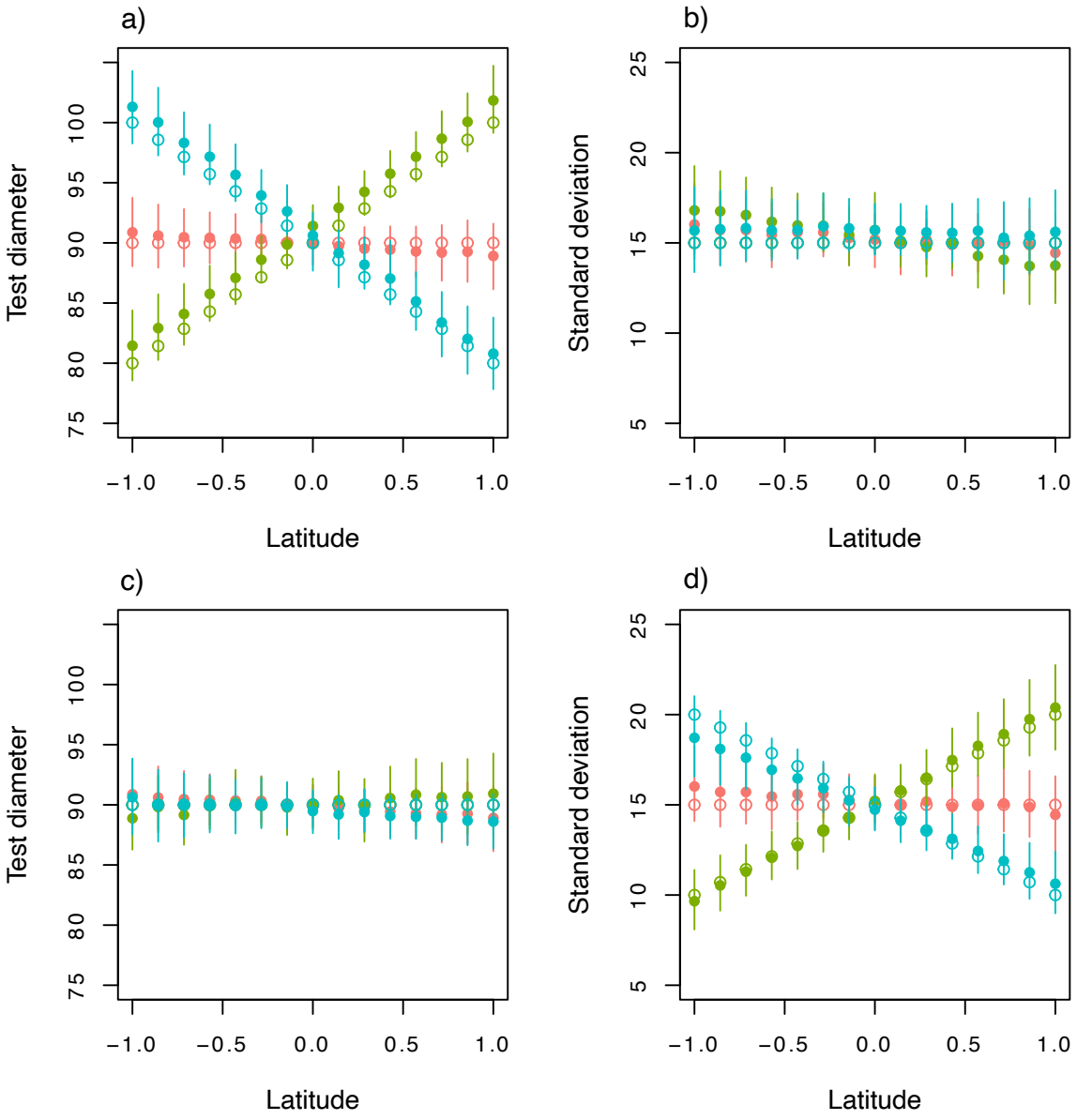

Figure A3: Plot of estimated and simulated location means (Figure 3a,c) and location standard deviations (Figure 3b,d) for the following models: (3a,b) sim_mn_vn_15: orange (no linear relationship between latitude and locations means and location standard deviations), sim_mi_vn_15: green (a positive relationship between locations means and latitude and no relationship between locations standard deviations and latitude), and sim_md_vn_15: blue (a negative relationship between locations means and latitude and no relationship between locations standard deviations and latitude); (3c,d) sim_mn_vn_15: orange (no linear relationship between latitude and locations means and location standard deviations), sim_mn_vi_15: green (no relationship between locations means and latitude and a positive relationship between locations standard deviations and latitude), sim_mn_vd_15: blue (no relationship between locations means and latitude and a negative relationship between locations standard deviations and latitude). Closed symbols are the estimated location mean and standard deviations from the model and open symbols are the simulated location mean and standard deviations. The error bars are the 0.95 Highest Posterior Density Intervals (HPDI). Note, that in panels b and c all models had the same simulated value i.e. the simulated value is equal to zero.

The simulated locations means were within the 95% HPDI for the estimated location means for all 5 models. Therefore the model was predicting them reasonably well for all 5 models. For both models sim_mi_vn_15 and sim_md_vn_15 the estimates of the location means were all slightly higher than was simulated (Fig. 3a).

The simulated locations standard deviations were within the 95% HPDI for the estimated location standard deviations for all 5 models. Although within the 95% HPDI, there was a slight negative trend in location standard deviations for models sim_mi_vn_15 and sim_mn_vn_15 even when there was no simulated slope for location standard deviations. However, the simulated location standard deviations were still within the 95% HPDI (Fig. 3b). Similarly, although within the 95% HPDI, for model sim_md_vn_15 the locations standard deviations were all slightly higher than was simulated (Fig. 3b).

Testing if non-linear relationships between latitude and the location summary statistics influence parameter estimates

Lastly, we examined if non-linear patterns in the location means influenced our estimates for the linear relationships between latitude and the location means, and for no relationship between latitude and location standard deviations. These simulations used a convex relationship between latitude and either location means or standard deviations.

sim_mc_vn_5/sim_mc_vn_15: this model has a convex relationship between location means and latitude ("mc": bL = 0) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_mn_vc_5/sim_mn_vc_15: this model has no relationship between location means and latitude ("mn": bL = 0) and a convex relationship between location standard deviations and latitude ("vc": cL = 0)

Function to create non-linear relationships between latitude and location means or standard deviations using the ulam bayesian model from the rethinking package.

```
data.sim_out.convex <- function(pop, ind, lat, mu, sig){

  diam_m.v <- cbind(mu, sig)

  dat <- list(sL = rep(seq(lat[1], lat[2], length = pop), each = ind), TD = c
(apply(diam_m.v, 1, function(x) rnorm(ind, x[1], (x[2])))), P = rep(seq(1, po
p), each = ind))

  out <- ulam(

    alist(

      TD ~ dnorm(mu , sigma),

      mu <- mu_a + bL*sL + z[P]*sigma_a,
```

```
      z[P] ~ dnorm( 0, 1),

      sigma_a ~ dexp(2),

      mu_a ~ dnorm(90, 10),

      bL ~ dnorm( 0 , 5 ),

      sigma <- mu_c + cL*sL + z2[P]*sigma_c,

      cL ~ dnorm( 0 , 5 ),

      mu_c ~ dnorm(15, 5),

      z2[P] ~ dnorm( 0, 1),

      sigma_c ~ dexp(2)

    ), data= dat , chains=4, cores=1 , control=list(adapt_delta=0.99, max_tre
edepth = 15), log_lik = TRUE, iter=3000, constraints=list( mu_c="lower=0", mu
_a="lower=0"))

  out

}
```

Simulating data and generating model output for the 2 convex scenarios for both 5 and 15 locations:

```
sim_mc_vn_5<-data.sim_out.convex(pop = 5, ind = 30, lat = c(-1, 1), mu = c(10
0,90,80,90,100), sig = c(rep(15,5)))
sim_mn_vc_5<-data.sim_out.convex(pop = 5, ind = 30, lat = c(-1, 1), mu = c(re
p(90,5)), sig = c(20,15,10,15,20))

sim_mn_vc_15<-data.sim_out.convex(pop = 15, ind = 30, lat = c(-1, 1), mu = c(
rep(90,15)), sig = c(20,19,18,17,16,15,14,13,14,15,16,17,18,19,20))
sim_mc_vn_15<-data.sim_out.convex(pop = 15, ind = 30, lat = c(-1, 1), mu = c(
101,98,95,92,89,86,83,80,83,86,89,92,95,98,101), sig = c(rep(15,15)))
```

Plotting regression coefficients (± CI) for the relationship between latitude and the location mean (bL) or location standard deviation (cL) for senario with convex relationship between latitude and the location means

```
x_sim.est <- matrix(1:8,2,4)
x_sim.est[1,]<-c(0,0,0,0)
x_sim.est[2,]<-c(0,0,0,0)
rownames(x_sim.est) <- paste(c("bL","cL"))

coeftab_plot_with_sim_estimates(coeftab(sim_mc_vn_5, sim_mc_vn_15, sim_mn_vc_
5, sim_mn_vc_15), pars = c("bL", "cL"), sim.est=x_sim.est, cex=1)
```

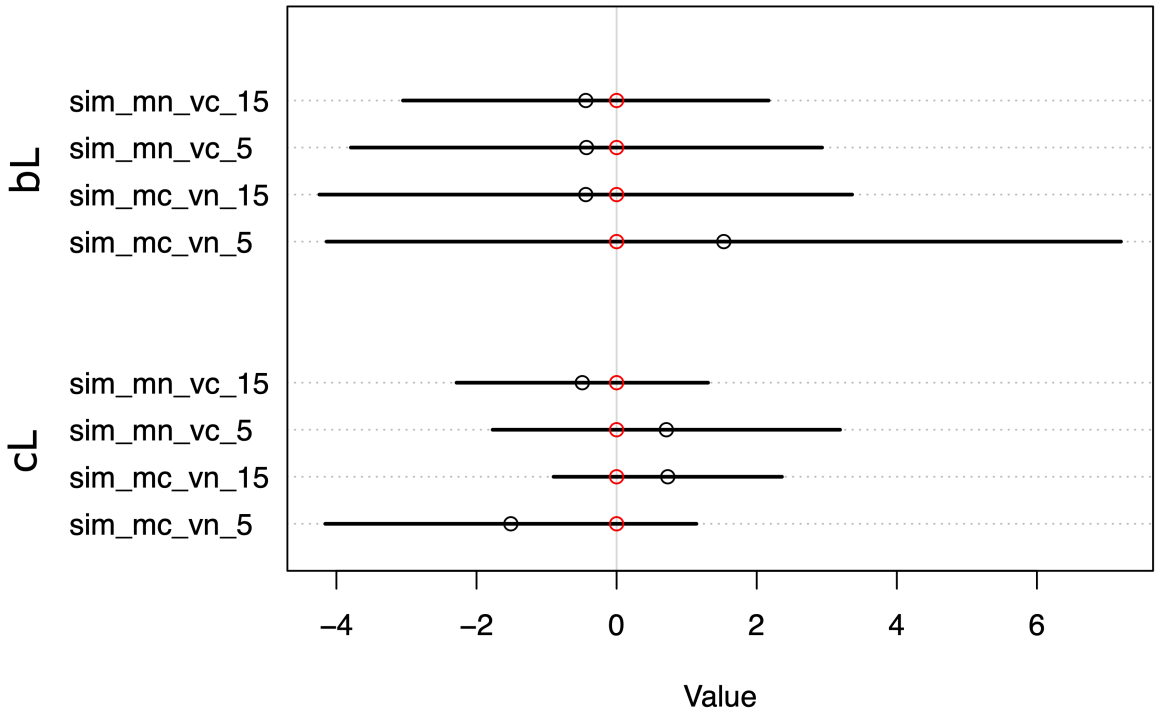

Figure A4: Plot of the regression coefficients bL (regression coefficient between latitude and the mean of test diameters of locations) and cL (regression coefficient between latitude and the variance of test diameters of locations). These were plotted for models: sim_mc_vn_5/sim_mc_vn_15 (a convex relationship between the location means and latitude and no relationship between latitude and the location standard deviations) and sim_mn_vc_5/sim_mn_vc_15 (no relationship between the location means and latitude and a convex relationship between latitude and the location standard deviations). The red circles are simulated values for each of the parameters in each of the simulated scenarios, the black circles are the model estimates and the lines are the 0.95 credible intervals.

In summary, Figure 4 shows there was good agreement between the simulation parameters and the estimated coefficients. The convex relationship gave an estimate of 0 for bL in models sim_mc_vn_5 and sim_mc_vn_15 and an estimate of 0 for cL in models sim_mn_vc_5 and sim_mn_vc_15 as was expected.

Plotting the estimated location means and location standard deviations for a convex relationship between the location means and latitude but no relationship between latitude and the location standard deviations

```
plot_location_means_and_sd_2<-function(modelname, seq, colour, m_sd){points(u
nique(modelname@data$sL), means_TD(modelname)[[m_sd]], col=colour, pch=16)
arrows(unique(modelname@data$sL),model_credible_intervals_means(modelname)[[m
_sd]][1,],y1=model_credible_intervals_means(modelname)[[m_sd]][2,],length = 0
, col=colour)
points(seq(-1,1,length=length(unique(modelname@data$P))),seq, col=colour, pch
=1)
}

#m_sd=1 for location means and m_sd=2 for location standard deviations

convex.means_5<-c(100,90,80,90,100)
convex.means_15<-c(101,98,95,92,89,86,83,80,83,86,89,92,95,98,101)
convex.sd_5<-c(20,15,10,15,20)
convex.sd_15<-c(20,19,18,17,16,15,14,13,14,15,16,17,18,19,20)

par(mfrow = c(2, 2), mar = c(4,4,2,2))

plot(NULL, xlim=c(-1,1), ylim=c(75,105), xlab="latitude", ylab="test diameter
", cex.axis=1.1, cex.lab=1.1, main="a", cex.main=1.3)
plot_location_means_and_sd_2(modelname=sim_mc_vn_5, seq=convex.means_5, colou
r="purple", m_sd=1)
plot_location_means_and_sd(modelname=sim_mn_vc_5, start=90, end=90, colour="#
e34234", m_sd=1)

plot(NULL, xlim=c(-1,1), ylim=c(5,25), xlab="latitude", ylab="standard deviat
ion", cex.axis=1.1, cex.lab=1.1, main="b", cex.main=1.3)
plot_location_means_and_sd(modelname=sim_mc_vn_5, start=15, end=15, colour="p
urple", m_sd=2)
plot_location_means_and_sd_2(modelname=sim_mn_vc_5, seq=convex.sd_5, colour="
#e34234", m_sd=2)

plot(NULL, xlim=c(-1,1), ylim=c(75,105), xlab="latitude", ylab="test diameter
", cex.axis=1.1, cex.lab=1.1, main="a", cex.main=1.3)
plot_location_means_and_sd_2(modelname=sim_mc_vn_15, seq=convex.means_15, col
our="purple", m_sd=1)
plot_location_means_and_sd(modelname=sim_mn_vc_15, start=90, end=90, colour="
#e34234", m_sd=1)

plot(NULL, xlim=c(-1,1), ylim=c(5,25), xlab="latitude", ylab="standard deviat
ion", cex.axis=1.1, cex.lab=1.1, main="b", cex.main=1.3)
plot_location_means_and_sd(modelname=sim_mc_vn_15, start=15, end=15, colour="
purple", m_sd=2)
plot_location_means_and_sd_2(modelname=sim_mn_vc_15, seq=convex.sd_15, colour
="#e34234", m_sd=2)
```

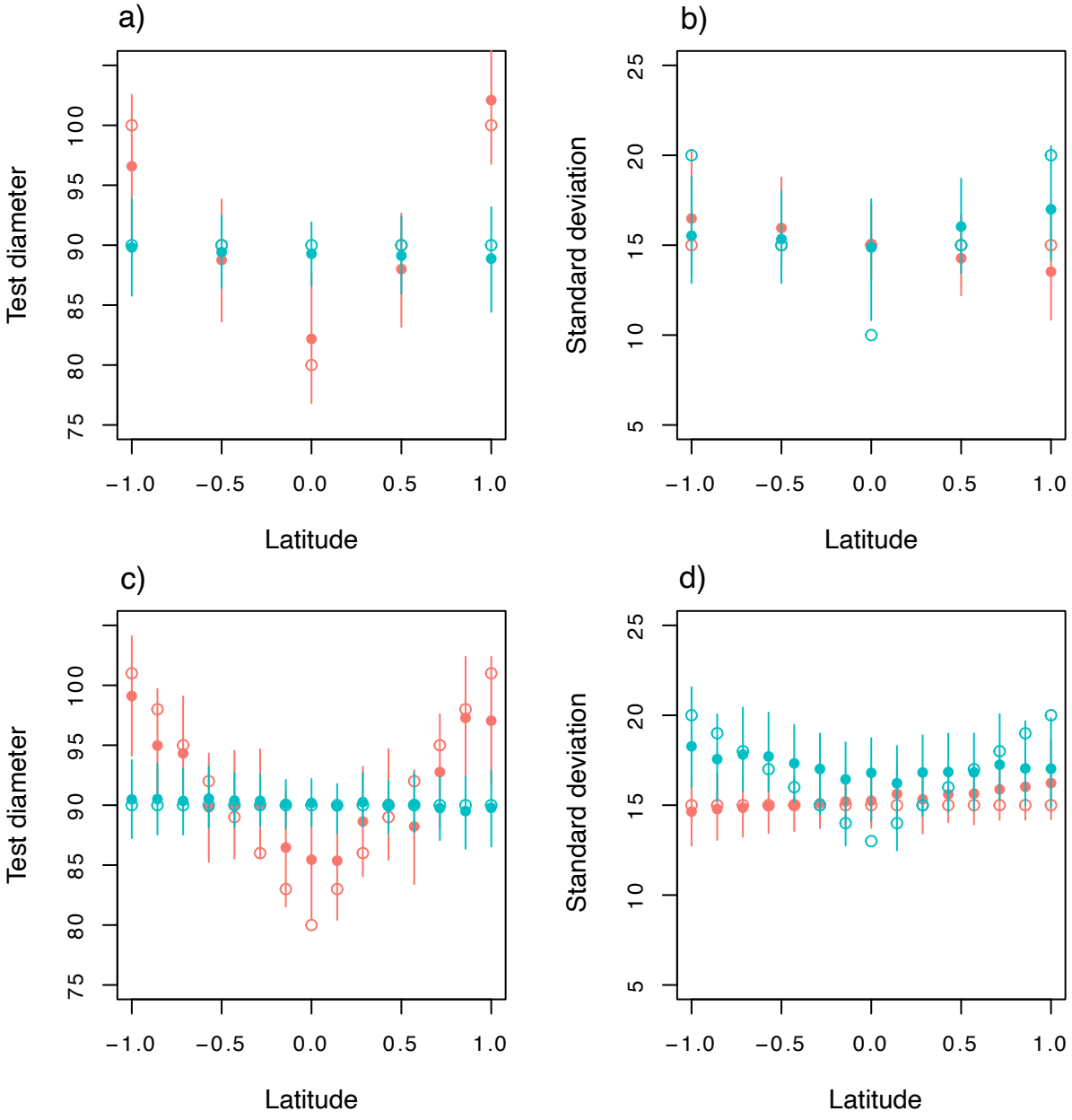

Figure A5: Plotting the estimated and simulated location means (a,c) and location standard deviations (b,d) for models: sim_mc_vn_5 (organge; a,b)/sim_mc_vn_15 (orange; c,d) (convex relationship between the location means and latitude and no relationship between latitude and the location standard deviations) and sim_mn_vc_5 (blue; a,b)/sim_mn_vc_15 (blue; c,d) (no relationship between the location means and latitude and a convex relationship between latitude and the location standard deviations). Closed symbols are the estimated location mean and standard deviations from the model and open symbols are the simulated location mean and standard deviations. The error bars are the 0.95 HPDI.

For both the 15 location and 5 location simulations, the convex relationship between the location means and latitude was recovered (Fig. 5a,c). However, the extent of the convex relationship was less extreme for the estimated location means than the simulated location means. For model sim_mc_vn_15, the lowest simulated location mean at latitude 0 was below the 0.95 HPDI (Fig. 5c). The rest of the location means had a simulated location mean within the estimated 0.95 HPDI.

For both the 15 location and 5 location simulations, the convex relationship between the location standard deviations and latitude was not recovered well (Fig. 5b,d). For model sim_mn_vc_5, the simulated location standard deviation at latitude -1 was above the 0.95 HPDI and the simulated location standard deviation at latitude 0 was below the 0.95 HPDI. Therefore the convex relationship was not recovered and location standard deviation estimates were not accurate. For model sim_mc_vn_5, the slope for the location standard deviations was negative when there was no simulated slope but the simulated location standard deviations were within the estimated HDPI errors. Model sim_mc_vn_5 gave a positive slope for the location standard deviations when there was no simulated slope but the simulated location standard deviations were within the estimated HDPI errors. For model sim_mn_vc_15, the convex relationship was partially recovered. Although the points did not follow the simulated trend, all the location standard deviations were within the HDPI error except the location at latitude 1 for which the simulated location standard deviation was above the estimated HDPI.

Supplementary tables: Outputs from models

These tables show the estimates for each of the parameters in the model: bL, cL, sigma_a, sigma_c, mu_a, mu_c, and z[1] to z[5] (for 5 populations) or z[1] to z[15] (for 15 populations).

Table A2: Nine different scenarios.

The estimates for the following 9 scenarios are shown for both 5 and 15 populations:

sim_mn_vn_5/sim_mn_vn_15: no relationship between location means and latitude ("mn": bL = 0) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_mi_vn_5/sim_mi_vn_15: positive relationship between location means and latitude ("mi": bL = 10) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_md_vn_5/sim_md_vn_15: negative relationship between location means and latitude ("md": bL = -10) and no relationship between location standard deviations and latitude ("vn": cL = 0)

sim_mn_vi_5/sim_mn_vi_15: no relationship between location means and latitude ("mn": bL = 0) and positive relationship between location standard deviations and latitude ("vi": cL = 5)

sim_mi_vi_5/sim_mi_vi_15: positive relationship between location means and latitude ("mi": bL = 10) and positive relationship between location standard deviations and latitude ("vi": cL = 5)

sim_md_vi_5/sim_md_vi_15: negative relationship between location means and latitude ("md": bL = -10) and positive relationship between location standard deviations and latitude ("vi": cL = 5)

sim_mn_vd_5/sim_mn_vd_15: no relationship between location means and latitude ("mn": bL = 0) and negative relationship between location standard deviations and latitude ("vd": cL = -5)

sim_mi_vd_5/sim_mi_vd_15: positive relationship between location means and latitude ("mi": bL = 10) and negative relationship between location standard deviations and latitude ("vd": cL = -5)

sim_md_vd_5/sim_md_vd_15: negative relationship between location means and latitude ("md": bL = -10) and negative relationship between location standard deviations and latitude ("vd": cL = -5)

```
coeftab(sim_mn_vn_5, sim_mn_vn_15, sim_mi_vn_5, sim_mi_vn_15, sim_md_vn_5, sim_md_vn_15, sim_mn_vi_5, sim_mn_vi_15, sim_mi_vi_5, sim_mi_vi_15, sim_
md_vi_5, sim_md_vi_15, sim_mn_vd_5, sim_mn_vd_15, sim_mi_vd_5, sim_mi_vd_15, sim_md_vd_5, sim_md_vd_15)

##         sim_mn_vn_5 sim_mn_vn_15 sim_mi_vn_5 sim_mi_vn_15 sim_md_vn_5 sim_md_vn_15 sim_mn_vi_5 sim_mn_vi_15 sim_mi_vi_5 sim_mi_vi_15
## z[1]       0.01        0.08        -0.27        0.10         0.03        -0.05        -0.02       -0.45        -0.25        0.13
## z[2]       0.01       -0.12         0.10        0.13         0.03         0.13         0.04        0.37         0.18        0.00
## z[3]      -0.08       -0.11         0.30       -0.19         0.00        -0.08        -0.05       -0.39        -0.10       -0.05
## z[4]       0.05       -0.01        -0.13        0.07         0.03         0.24         0.08        0.20         0.31        0.25
## z[5]      -0.01        0.10        -0.01       -0.04        -0.11         0.19         0.00        0.39        -0.13       -0.53
## sigma_a    0.45        0.42         0.57        0.46         0.45         0.62         0.43        0.67         0.58        0.78
## mu_a      89.29       89.88        89.82       91.46        91.72        91.02        90.24       90.12        89.29       89.54
## bL        -0.80       -0.93         7.21       10.08        -8.81       -10.34         1.11        0.80         8.60       10.53
## cL        -1.23       -0.61        -1.07       -1.63        -1.45        -0.17         3.54        5.37         5.17        3.53
## mu_c      16.08       15.26        16.26       15.34        15.70        15.65        13.88       15.10        16.00       14.54
## z2[1]     -0.12        0.21         0.01       -0.20        -0.12        -0.19         0.07       -0.13         0.08       -0.12
## z2[2]      0.05       -0.10        -0.11        0.01         0.25        -0.07        -0.33        0.07        -0.16        0.04
## z2[3]      0.12        0.02         0.16        0.05        -0.19         0.06         0.05        0.07        -0.07        0.10
## z2[4]      0.25       -0.22         0.11       -0.13         0.06        -0.08         0.28        0.05         0.11       -0.03
## z2[5]     -0.25        0.08        -0.14       -0.09        -0.03        -0.03        -0.12       -0.14         0.04       -0.06
## sigma_c    0.50        0.44         0.45        0.49         0.47         0.42         0.49        0.35         0.44        0.33
## z[6]         NA        0.25           NA        0.03           NA        -0.04           NA        0.16           NA        0.26
## z[7]         NA       -0.02           NA       -0.19           NA         0.12           NA       -0.21           NA        0.17
## z[8]         NA        0.05           NA       -0.06           NA        -0.37           NA       -0.01           NA       -0.56
## z[9]         NA       -0.05           NA        0.02           NA        -0.39           NA        0.14           NA       -0.09
## z[10]        NA       -0.14           NA       -0.12           NA         0.12           NA       -0.41           NA       -0.35
## z[11]        NA       -0.06           NA       -0.04           NA         0.43           NA        0.09           NA        0.08
## z[12]        NA       -0.09           NA       -0.08           NA         0.04           NA        0.24           NA        0.49
## z[13]        NA       -0.04           NA       -0.01           NA        -0.24           NA       -0.06           NA       -0.22
## z[14]        NA        0.23           NA       -0.05           NA        -0.12           NA       -0.12           NA        0.31
## z[15]        NA       -0.05           NA        0.39           NA         0.12           NA        0.00           NA        0.12
## z2[6]        NA        0.21           NA        0.19           NA         0.31           NA        0.01           NA        0.02
## z2[7]        NA       -0.09           NA       -0.13           NA         0.17           NA       -0.03           NA        0.08
## z2[8]        NA       -0.10           NA        0.50           NA         0.04           NA        0.20           NA        0.01
## z2[9]        NA       -0.24           NA        0.02           NA         0.05           NA       -0.11           NA       -0.09
## z2[10]       NA        0.12           NA       -0.15           NA        -0.03           NA       -0.20           NA       -0.22
## z2[11]       NA       -0.16           NA        0.45           NA        -0.04           NA        0.14           NA        0.13
## z2[12]       NA        0.03           NA       -0.17           NA         0.16           NA        0.17           NA        0.19
## z2[13]       NA        0.37           NA       -0.15           NA        -0.34           NA       -0.02           NA       -0.02
## z2[14]       NA        0.19           NA       -0.28           NA        -0.15           NA        0.08           NA        0.09
## z2[15]       NA       -0.28           NA        0.03           NA         0.17           NA       -0.13           NA       -0.12
## nobs        150         450          150         450          150         450          150         450          150         450
##         sim_md_vi_5 sim_md_vi_15 sim_mn_vd_5 sim_mn_vd_15 sim_mi_vd_5 sim_mi_vd_15 sim_md_vd_5 sim_md_vd_15
## z[1]       0.10        0.04        -0.04        0.22        -0.14        -0.09         0.01        0.35
## z[2]      -0.03       -0.25         0.10       -0.29         0.01        -0.09        -0.05       -0.21
## z[3]       0.00        0.11        -0.01       -0.14         0.04         0.14         0.23        0.03
## z[4]       0.09        0.25        -0.06        0.00         0.22        -0.06        -0.17        0.40
## z[5]      -0.16       -0.04         0.04       -0.03        -0.09        -0.12        -0.08        0.09
## sigma_a    0.47        0.41         0.43        0.47         0.49         0.39         0.49        0.97
## mu_a      91.61       89.89        90.77       89.53        90.72        90.08        89.42       90.16
```

```
## bL       -10.89       -9.05         1.00       -0.91         9.67         9.61        -6.85        -9.05
## cL         6.44        6.01        -5.43       -3.99        -4.48        -5.18        -2.36        -5.16
## mu_c      15.82       14.55        13.63       14.72        14.87        15.10        14.12        15.23
## z2[1]     -0.02       -0.08         0.15        0.00         0.05        -0.12        -0.15        -0.08
## z2[2]     -0.07        0.12        -0.30       -0.10        -0.26         0.10         0.16         0.14
## z2[3]      0.08        0.07         0.23        0.08         0.42         0.32         0.10        -0.50
## z2[4]      0.00        0.08         0.02       -0.11        -0.03        -0.05         0.11        -0.12
## z2[5]      0.02       -0.21        -0.10        0.05        -0.18         0.03        -0.17         0.32
## sigma_c    0.41        0.33         0.50        0.32         0.54         0.53         0.45         0.67
## z[6]         NA        0.25           NA        0.26           NA         0.05           NA        -0.34
## z[7]         NA       -0.12           NA        0.23           NA         0.11           NA        -0.06
## z[8]         NA       -0.03           NA       -0.06           NA        -0.05           NA         0.02
## z[9]         NA       -0.04           NA       -0.23           NA        -0.05           NA        -0.14
## z[10]        NA       -0.06           NA        0.16           NA         0.10           NA         0.10
## z[11]        NA       -0.06           NA       -0.09           NA         0.04           NA         0.03
## z[12]        NA        0.13           NA       -0.01           NA         0.03           NA        -0.67
## z[13]        NA       -0.15           NA        0.07           NA         0.15           NA         0.67
## z[14]        NA       -0.05           NA       -0.07           NA        -0.04           NA        -0.71
## z[15]        NA        0.04           NA       -0.01           NA        -0.11           NA         0.47
## z2[6]        NA       -0.09           NA        0.12           NA        -0.05           NA         0.54
## z2[7]        NA        0.01           NA       -0.08           NA         0.34           NA         0.26
## z2[8]        NA        0.09           NA        0.03           NA        -0.42           NA        -0.42
## z2[9]        NA        0.01           NA       -0.06           NA        -0.27           NA         0.08
## z2[10]       NA       -0.05           NA        0.07           NA        -0.27           NA         0.14
## z2[11]       NA       -0.04           NA        0.21           NA         0.33           NA        -0.12
## z2[12]       NA        0.08           NA        0.02           NA        -0.40           NA         0.21
## z2[13]       NA       -0.23           NA        0.05           NA         0.24           NA        -0.56
## z2[14]       NA        0.15           NA       -0.07           NA        -0.03           NA        -0.01
## z2[15]       NA        0.12           NA       -0.18           NA         0.21           NA         0.11
## nobs        150         450          150         450          150         450          150         450
```

Table A3: Testing different magnitudes of the regression parameters.

The estimates for the following 6 scenarios are shown for both 5 and 15 populations:

change in bL magnitude:

sim_mi_low_vn_5/sim_mi_low_vn_15: low positive relationship between location means and increasing latitude ("mi_low": bL = 5) and no relationship between location variances latitude ("vn": cL = 0)

sim_mi_med_vn_5/sim_mi_med_vn_15: medium positive relationship between location means and increasing latitude ("mi_med": bL = 10) and no relationship between location variances latitude ("vn": cL = 0)

sim_mi_hig_vn_5/sim_mi_hig_vn_15: high positive relationship between location means and increasing latitude ("mi_hig": bL = 20) and no relationship between location variances latitude ("vn": cL = 0)

change in cL magnitude:

sim_mn_vi_low_5/sim_mn_vi_low_15: no relationship between location means and latitude ("mn": bL = 0) and low positive relationship between location standard deviations and latitude ("vi_low": cL = 2)

sim_mn_vi_med_5/sim_mn_vi_med_15: no relationship between location means and latitude ("mn": bL = 0) and medium positive relationship between location standard deviations and latitude ("vi_med": cL = 5)

sim_mn_vi_hig_5/sim_mn_vi_hig_15: no relationship between location means and latitude ("mn": bL = 0) and high positive relationship between location standard deviations and latitude ("vi_hig": cL = 10)

```
coeftab(sim_mn_vi_low_5, sim_mn_vi_low_15, sim_mn_vi_med_5, sim_mn_vi_med_15, sim_mn_vi_hig_5, sim_mn_vi_hig_15, sim_mi_low_vn_5, sim_mi_low_vn_15,
sim_mi_med_vn_5, sim_mi_med_vn_15, sim_mi_hig_vn_5, sim_mi_hig_vn_15)
```

```
##         sim_mn_vi_low_5 sim_mn_vi_low_15 sim_mn_vi_med_5 sim_mn_vi_med_15 sim_mn_vi_hig_5 sim_mn_vi_hig_15 sim_mi_low_vn_5
## z[1]      -0.01           -0.15            -0.02            0.07             0.05            0.18             -0.05
## z[2]       0.06            0.09            -0.01            0.00            -0.13            0.01              0.00
## z[3]      -0.17           -0.03             0.05           -0.20             0.06           -0.14              0.05
## z[4]       0.17           -0.08            -0.06           -0.63             0.11           -0.21             -0.11
## z[5]      -0.03            0.09             0.03            0.36            -0.08           -0.02              0.12
## sigma_a    0.48            0.67             0.44            0.79             0.44            0.39              0.45
## mu_a      91.51           90.17            92.13           89.45            88.52           89.34             90.25
## bL         1.46            0.28            -0.44            0.21             0.31            0.66              6.09
## cL         2.94            1.60             6.13            4.44             8.83            9.62             -1.49
## mu_c      14.81           14.77            15.79           14.84            13.82           15.37             15.25
## z2[1]      0.00            0.19            -0.03            0.20            -0.27            0.04              0.02
## z2[2]      0.05           -0.17             0.13           -0.29             0.33           -0.14              0.10
## z2[3]     -0.28            0.02            -0.11            0.04            -0.21           -0.12             -0.07
## z2[4]      0.27            0.01            -0.15           -0.24             0.09            0.05             -0.09
## z2[5]     -0.07           -0.13             0.18            0.02             0.02            0.06              0.08
## sigma_c    0.49            0.33             0.46            0.42             0.49            0.31              0.43
## z[6]         NA           -0.40               NA            0.16               NA            0.11                NA
## z[7]         NA            0.08               NA            0.43               NA            0.19                NA
## z[8]         NA            0.29               NA            0.20               NA           -0.15                NA
## z[9]         NA            0.10               NA            0.11               NA            0.21                NA
## z[10]        NA           -0.36               NA           -0.12               NA           -0.09                NA
## z[11]        NA            0.54               NA            0.10               NA           -0.03                NA
## z[12]        NA            0.26               NA           -0.33               NA           -0.05                NA
## z[13]        NA           -0.13               NA           -0.44               NA           -0.09                NA
## z[14]        NA           -0.14               NA           -0.12               NA            0.10                NA
## z[15]        NA           -0.25               NA            0.43               NA            0.01                NA
## z2[6]        NA            0.05               NA            0.28               NA            0.06                NA
## z2[7]        NA           -0.08               NA           -0.09               NA            0.21                NA
## z2[8]        NA            0.06               NA           -0.24               NA           -0.06                NA
## z2[9]        NA           -0.09               NA           -0.04               NA            0.03                NA
## z2[10]       NA            0.03               NA            0.34               NA           -0.10                NA
## z2[11]       NA            0.08               NA            0.11               NA           -0.15                NA
## z2[12]       NA           -0.15               NA            0.02               NA           -0.04                NA
## z2[13]       NA            0.10               NA            0.01               NA            0.07                NA
## z2[14]       NA           -0.02               NA           -0.06               NA           -0.09                NA
## z2[15]       NA            0.07               NA           -0.11               NA            0.20                NA
## nobs        150             450              150             450              150             450               150
##         sim_mi_low_vn_15 sim_mi_med_vn_5 sim_mi_med_vn_15 sim_mi_hig_vn_5 sim_mi_hig_vn_15
## z[1]       0.06            -0.16            -0.08            -0.56            0.12
## z[2]       0.09             0.07            -0.10             0.18            0.02
## z[3]       0.02            -0.06            -0.11             0.41           -0.16
## z[4]      -0.02             0.13             0.13            -0.24           -0.17
## z[5]      -0.24             0.00            -0.08             0.21            0.01
## sigma_a    0.42             0.48             0.45             0.82            0.40
## mu_a      90.02            90.59            90.03            88.96           90.49
```

```
## bL         3.21            9.54           8.62           17.37          17.09
## cL         1.50           -0.33           0.33            0.49          -0.85
## mu_c      15.14           15.90          14.96           15.64          14.10
## z2[1]     -0.81            0.41          -0.07            0.21          -0.04
## z2[2]      0.39           -0.43           0.02           -0.21          -0.06
## z2[3]      0.51            0.03           0.11           -0.08           0.25
## z2[4]      0.12           -0.38           0.08            0.00          -0.03
## z2[5]      0.13            0.38          -0.10            0.09          -0.09
## sigma_c    0.80            0.76           0.38            0.47           0.33
## z[6]      -0.14              NA           0.00              NA          -0.12
## z[7]       0.08              NA           0.15              NA          -0.03
## z[8]      -0.17              NA           0.24              NA           0.13
## z[9]       0.23              NA          -0.01              NA           0.15
## z[10]      0.11              NA          -0.03              NA          -0.08
## z[11]     -0.05              NA          -0.18              NA          -0.14
## z[12]      0.10              NA           0.04              NA           0.07
## z[13]      0.02              NA           0.17              NA          -0.02
## z[14]      0.02              NA           0.17              NA           0.08
## z[15]     -0.08              NA          -0.22              NA           0.16
## z2[6]     -0.11              NA           0.14              NA           0.08
## z2[7]     -0.17              NA          -0.04              NA          -0.11
## z2[8]      0.21              NA          -0.28              NA          -0.06
## z2[9]     -0.47              NA          -0.06              NA           0.08
## z2[10]    -0.25              NA          -0.06              NA          -0.09
## z2[11]    -0.13              NA           0.12              NA          -0.11
## z2[12]     0.39              NA          -0.08              NA          -0.01
## z2[13]     0.37              NA           0.29              NA           0.27
## z2[14]     0.15              NA           0.14              NA          -0.07
## z2[15]    -0.39              NA          -0.21              NA          -0.04
## nobs        450             150            450             150            450
```

Table A4: Testing if non-linear relationships between latitude and the location summary statistics influence parameter estimates.

The estimates for the following model is shown for 5 populations.

sim_mc_vn_5: convex relationship between location means and latitude ("mc": bL = 0) and no relationship between location standard deviations and latitude ("vn": cL = 0)

```
coeftab(sim_mc_vn_5)
```

```
##         sim_mc_vn_5
## z[1]       1.44
## z[2]      -0.42
## z[3]         -2
## z[4]      -0.92
## z[5]       1.96
## sigma_a    4.74
## mu_a      91.45
## bL         1.53
## cL        -1.51
## mu_c      15.06
## z2[1]     -0.11
## z2[2]       0.2
## z2[3]      0.01
## z2[4]     -0.05
## z2[5]     -0.03
## sigma_c    0.44
## nobs        150
```

Chapter Three: Long-distance dispersal and changing meta-population dynamics shape neutral and adaptive genomic variation across the New Zealand range of the Long-spined sea urchin (*Centrostephanus rodgersii*)

Abstract

Climate-driven changes in species range and abundance can influence the genetic connectivity between populations and the genetic composition of populations. The Long-spined sea urchin (*Centrostephanus rodgersii*) has extended in range and increased in abundance in Tasmania impacting both the local fisheries and ecosystem. *Centrostephanus rodgersii* is also found in New Zealand, however, we know little about the its population history in its New Zealand range including Rangitāhua (the Kermadec Islands). We used population genomics to study the population structure and connectivity of *C. rodgersii* across its New Zealand range. We found that Rangitāhua populations and north-east New Zealand populations are genetically differentiated, but there is some ongoing migration from Rangitāhua to north-east New Zealand. *Centrostephanus rodgersii* populations in north-east New Zealand did not form distinct genetic clusters indicating one meta-population. However, modularity analysis of the population graphs, based on genetic covariances, showed some groups of populations were more similar to each other and presumably share a common colonisation history or higher geneflow. To reveal whether there has been a change in the genetic composition and connectivity of populations in north-east New Zealand over the last few decades, we compared the genetic composition of the younger (smaller) and older (larger) individuals within populations. This comparison recovered different patterns of connectivity and within-population variance between the two demographic groups that were not present when the groups were combined. The changing connectivity and population structure of *C. rodgersii* in New Zealand along with the impact the urchin had on fisheries and the ecosystem in Tasmania indicates that we should continue to monitor *C. rodgersii* within New Zealand.

Introduction

Climate change is impacting the marine environment, through changing ocean temperatures, chemistry, oxygen, and circulation. These impacts affect the distribution of species, their local abundance, and population connectivity (Wilson et al. 2016). For instance, a changing environment leads to local increases or decreases in the abundance of species (Wassmann et al. 2011; Poloczanska et al. 2016), in extreme cases causing the spatial distribution of species to shift. The most evident climate-driven changes in species ranges are away from warming low-latitude regions toward the poles (i.e. poleward range shifts; Sorte et al. 2010), but the direction of range shifts may vary according to the geography of local environmental changes. As a consequence, subtle changes in the abundance and range of species in response to climate change may be difficult to detect. Nonetheless, it is important to monitor these early changes, particularly for species predicted to have widespread impacts on local ecosystems, food security, and human livelihoods (Pecl et al. 2017; Melbourne-Thomas et al. 2021).

Changes in a species abundance and range are not independent of changes in the connectivity among populations and their genetic composition. Climate-driven range extensions can impact the genetic variance of populations and their connectivity, and climate-induced population decline can erode the genetic diversity of populations (Coleman et al. 2020). Such population genetic inferences have long been used to infer demographic and spatial expansion (Excoffier 2004). Seascape genomics has more recently helped to characterise physical and environmental influences on the distribution of neutral and adaptive genetic variance (Dalongeville et al. 2018a; Dalongeville et al. 2018b). Therefore, when a species climate-mediated change in range and abundance is subtle, or still in the very early stages, population genomics may be particularly useful in detecting changes in the genetic variance and connectivity of populations.

Several high latitude temperate reef areas have suffered climate change impacts, such as Western Australia (Smale et al. 2013; Wernberg et al. 2021), Japan (Agatsuma et al. 2007; Feng et al. 2019), and Tasmania (Ling et al. 2009c). These locations have all been described as climate change hotspots where several species have gone locally extinct as the local environment has changed and new species have extended their ranges into the area. Although New Zealand is similar in latitude, and its ocean climate has measurably changed over recent decades (Law et al. 2017b), species responses to these local environmental changes have been less evident. In many climate change hotspots, strong and sometimes strengthening boundary currents facilitate poleward range extensions by driving larval movement to, and warming, poleward locations (Ling et al. 2009c). In contrast, New Zealand's poleward current, the East Auckland Current (EAuC), is a weak offshoot of the East Australian Current (EAC), that travels across the Tasman Sea from Australia as the Tasman Front. Along the geographically complex East Coast of north-east New Zealand the EAuC has an unpredictable flow, forming many eddies, and often sits far offshore (Stanton et al. 1997). As a consequence, although there are a few accounts of poleward range shifts in this region (Middleton et al. 2021), biodiversity responses to climate change in north-east New Zealand may be more subtle than those observed in other temperate reef areas.

Around the world, the impacts of species shifting their ranges cause societal, cultural, ecosystem, governance, and climate feedback impacts (Pecl et al. 2017), some of which we may soon see in New Zealand. One of the most notorious examples is the Long-spined sea urchin, *Centrostephanus rodgersii*. This urchin is native to New South Wales, but in the 1980s began to extend its range south toward Tasmania facilitated by the strengthening and extension of the EAC, and has been increasing in abundance and range since (Ling et al. 2009c). The extension of *C. rodgersii* into Tasmania has dramatically impacted the ecosystem and local fisheries of abalone (*Haliotis rubra*) and rock lobster (*Jasus edwardsii*) (Johnson et al. 2005; Ling 2008; Lisson 2018). *Centrostephanus rodgersii* also occurs in New Zealand, along the north-east coast of the North Island, including several nearshore and offshore islands, and at Rangitāhua (the Kermadec archipelago, Fig. 1). North-east New Zealand has a similar ecosystem to Tasmania. For instance, the dominant macroalgae species in north-east New Zealand is *Ecklonia radiata* (the most impacted macroalgae species in Tasmania; Ling et al. 2018). North-east New Zealand has high-value fisheries, similar to Tasmania, of crayfish (*J. edwardsii,* known as rock lobster in Tasmania), pāua (*Haliotis iris,* an endemic abalone species), and kina (*Evechinus chloroticus,* an endemic urchin). Based on the biodiversity and fisheries impact of *C. rodgersii* in Tasmania (Johnson

et al. 2005; Ling 2008; Lisson 2018), *C. rodgersii* poses a considerable threat to New Zealand's coastal ecosystem and fisheries.

Although there are no time-series surveys of *C. rodgersii* throughout its New Zealand range, available research suggests the species has increased in abundance in certain locations (Shears 2020; Balemi et al. 2021), and likely in range extent. *Centrostephanus rodgersii* was first recorded in New Zealand in 1897 (Farquhar 1897) but was subsequently removed from the faunal list on two occasions for lack of evidence (Fell 1949), indicating it was not abundant. It was not until 1949 when live specimens were collected from several locations in the North Island (the Cavalli Islands, Stephen's Island, Whangaroa, and Little Barrier) that there was conclusive evidence of the species presence (Fell 1949). The species is most abundant in Rangitāhua (L. Liggins, pers. comm. 2021), but overall New Zealand's population densities appear to be lower than observed in Australia (Pecorino et al. 2013b; Edgar et al. 2017). Population genetic studies identified high connectivity between *C. rodgersii* populations in New Zealand and the East Coast of Australia (Banks et al. 2007), but suggest that New Zealand populations are no longer reliant on dispersal from Australia and are a self-sustaining meta-population (Thomas et al. 2021). Based on population size structure analysis across the species New Zealand range, we expect there has been a poleward range extension occurring in the northern part of north-east New Zealand and that there is more regular recruitment in southern locations of north-east New Zealand (Chapter Two).

Here we studied the population genomics of *C. rodgersii* across its' New Zealand range, including Rangitāhua. We were interested in patterns of neutral genetic structure and covariance across the New Zealand range of *C. rodgersii*, and whether there was any indication of changing population connectivity or genetic composition. To identify changes in population connectivity and genetic composition over time, we analysed two different age classes within each population – those up to 15 years old and those over 15 years old – based on size-at-age relationships determined for the species (Pecorino et al. 2012). We additionally considered patterns of adaptive genetic variance and covariance among populations. Putatively adaptive loci were identified by several methods that fundamentally aim to identify loci that depart from the prevailing pattern of genetic population structure. We analysed population structure at this set of loci to further explore the qualitative and quantitative differences in patterns of population structure at the putatively adaptive set relative to the presumed neutral set of loci to detect signatures of local adaptation, infer the adaptive potential of populations, and identify any conspicuous patterns.

Methods

Sampling locations and collection

Centrostephanus rodgersii individuals were collected throughout the New Zealand range of the species, including the north-east coast of the North Island and Rangitāhua (Fig. 1). Gonad tissue was collected from each individual and stored in 95% ethanol. Most populations were sampled in late 2015 to early 2016, with a few sampled in late 2017 and January 2018. Twenty-nine to thirty-one individuals were used for genomic analysis from 15 locations including Raoul Island, Macauley Island, L'Esperance Rock, Spirits Bay, Berghan Point, Cavalli Islands, Home Point, Poor Knights Islands, Mokohinau Islands, Sail Rock, Mercury Islands, Castle Rock, Alderman Islands, Mayor Island, White Island and then 14 individuals from Takatu Point, where population density was lower.

We measured the size (test diameter in mm) of each individual urchin as a proxy for age. As urchin size is determined by both age and the growth rate, which is impacted by food availability or ambient temperature (Pecorino et al. 2012), the size-based analyses excluded Rangitāhua as the food availability and ambient temperature there is not consistent with north-east New Zealand. North-east New Zealand is considered as a single bioregion supporting our assumption that food availability and ambient temperature were the same for our sites in this region (Shears et al. 2008). We created two size cohorts: below 90mm, and 90mm and above. The cut-off point of 90mm was used because this is approximately the point where the urchin's growth slows as they reach their full adult size at 15 years (Pecorino et al. 2012). Therefore beyond this size, the urchins are indistinguishable by size and therefore difficult to age. These two age/size groups were represented across all north-east New Zealand locations.

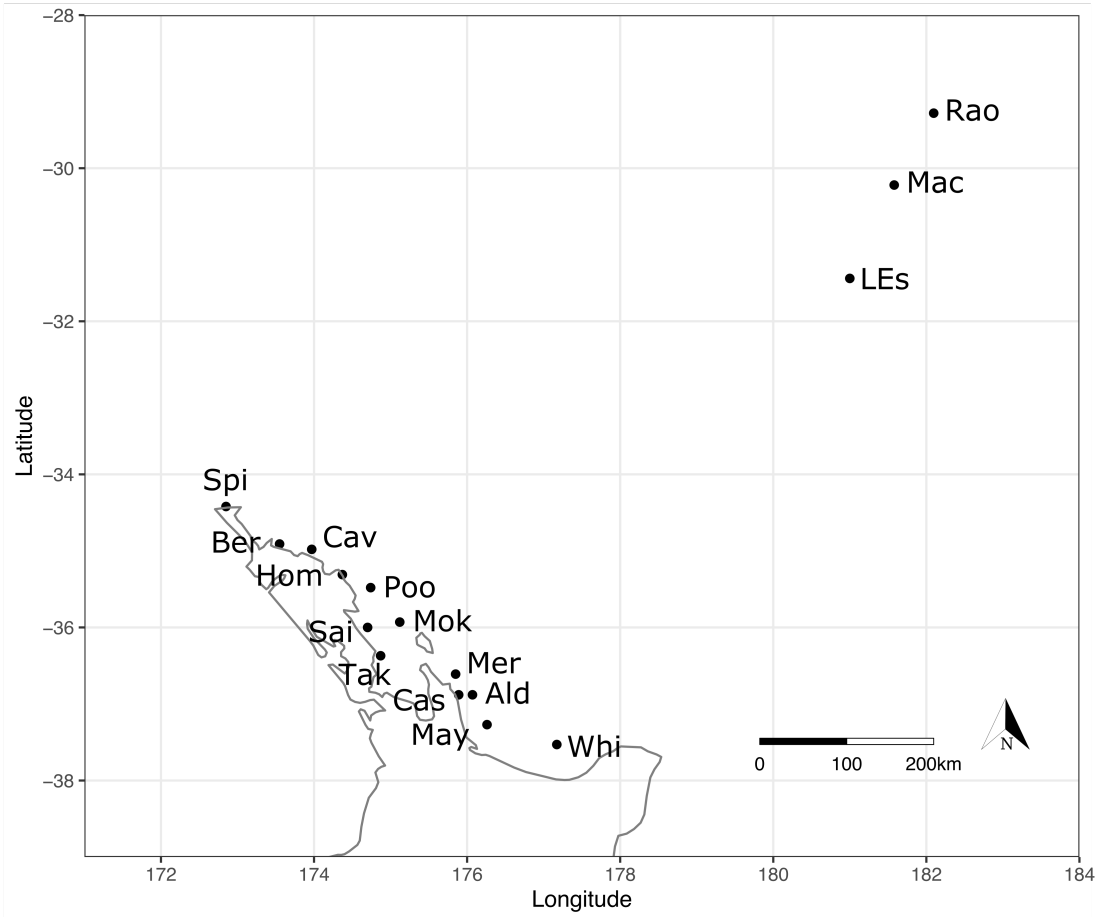

Figure 1: Map of the sampling locations of *C. rodgersii* across its New Zealand range using abbreviated names as follows: Alderman Islands: Ald, Berghan Point: Ber, Castle Rock: Cas, Cavalli Islands: Cav, Home Point: Hom, L'Esperance Rock: LEs, Macauley Island: Mac, Mayor Island: May, Mercury Islands: Mer, Mokohinau Islands: Mok, Poor Knights Islands: Poo, Raoul Island: Rao, Sail Rock: Sai, Spirits Bay: Spi, Takatu Point: Tak, White Island: Whi.

DNA extractions and Genotype-By-Sequencing

Genomic DNA was extracted from the gonad tissue using the Qiagen DNeasy Blood & Tissue Kit following the manufacturer's protocol for animal tissue with the following modifications: the lysis step was conducted overnight at 56°C with 20µl Proteinase K; after digestion, 3µl of RNAse A (Monarch) was added and samples were incubated for 20 minutes at 37°C; and to maximize DNA yield, the elution step using UltraPure DNase/RNase-Free Distilled Water (Invitrogen) was repeated, 50µl each, for a total of 100µl. Extractions were initially assessed for quality by running 2µL of DNA on a 1% agarose gel. A selection of the extractions were quantified using a Qubit 2.0 Fluorometer (Fisher Scientific), which were then used to estimate the concentration for all extractions.

Genotyping-by-Sequencing (GBS) was done by Diversity Arrays Technology Pty Ltd (DArTseq; Australia) using the DArT genome complexity reduction methods followed by 100 cycles on an Illumina Hiseq2500 as described in Kilian et al. (2012). After testing several combinations of restriction enzymes for the best complexity reduction method, PstI and NlaIII were selected to digest DNA prior to library preparation. Library preparation was conducted for 468 DNA extractions (four of these were individual replicates to investigate genotyping and/or variant calling errors).

Single nucleotide polymorphism calling and filtering

Single nucleotide polymorphisms (SNPs) were generated from the Illumina reads using proprietary DArT pipelines which demultiplex the reads to each sample using a strict barcode mismatch criteria and remove low-quality sequences. The provided DArT SNP file was a matrix where 102,052 SNP loci were coded as 0 for homozygotes, 1 for heterozygotes and 2 for homozygotes, for the alternative allele, and contained various QC statistics for each SNP.

The R dartR version 1.9.4 software package (Gruber et al. 2018) in RStudio Version 1.2.1335 (RStudio Team 2020) was used to convert the DArT SNP matrix into a genlight object (adegenet; Jombart 2008), generate QC reports, and perform several filtering steps. Based on the QC reports generated by dartR, loci were removed based on the following criteria: minor allele frequencies less than 0.05, more than 80% missing data, read depths less than 5, DArT replicability score below 0.97, and any monomorphic loci. We pruned the data, retaining only one SNP per locus using the filter.secondaries function in dartR, to minimise linkage among our SNP loci (Gruber et al. 2018).

Four individuals were sequenced twice. Any loci that provided different genotypes for the same individuals were removed from the dataset. Then for these four individuals, the replicate with the most data was retained (Table A1).

Detection of loci putatively under selection

Loci that are under selection, or linked to loci that are under selection, can influence the inference of population demography. To identify loci that may be under divergent selection indicative of local adaptation we used two outlier detection methods: the R packages

pcadapt version 4.3.3 (Luu et al. 2017) and Outflank version 0.2 (Whitlock et al. 2015) in RStudio Version 1.2.1335 (RStudio Team 2020). Pcadapt identifies outlier loci by their divergence from the overall population structure within a Principal Components Analysis (PCA) using the method described by Luu et al. (2017). This method does not require *a priori* grouping of individuals and has been shown to perform well in demographic scenarios potentially relevant to *C. rodgersii* in New Zealand, such as continuous population structure, hierarchical population structure, the presence of admixed individuals, and range expansion (Luu et al. 2017). Using Cattel's rule, we determined the number of principal components for the PCA from the scree plot (Fig. A1) and used the Mahalanobis distance as our test statistic where loci with a *q*-value less than 0.05 were considered to be potentially under selection.

Although pcadapt is less sensitive than Outflank to demographic effects that may cause co-dependence among populations (Lotterhos et al. 2014), it does suffer from low power if the spatial pattern of selection does not correlate with spatial genetic structure (Capblancq et al. 2018). To detect any loci that pcadapt may have missed for this reason, and to increase our confidence in the neutral dataset, we used the R package Outflank (Whitlock et al. 2015) as well as pcadapt. OutFlank detects outlier loci based on the distribution of loci-specific F_{ST} values, where loci with values in the centre of the distribution are assumed to be neutral. Loci with the highest (and lowest) F_{ST} are identified as outliers. Unlike pcadapt, Outflank requires the *a priori* grouping of individuals into populations and is generally more conservative returning few false positives across many demographic scenarios (Whitlock et al. 2015; Luu et al. 2017). OutFlank was run using default parameters, and a *q*-value threshold of 0.05 was used to identify loci with the highest F_{ST} values which are likely under divergent selection.

We created two datasets from the outlier detection methods to be used in subsequent analyses. First, the 'neutral' dataset included only loci not detected as outliers using either Outflank or pcadapt; and an 'adaptive' dataset including all loci detected as outliers across both Outflank and pcadapt.

Population genetic summary statistics

Several summary statistics were calculated for each of the sampled populations based on the neutral loci. The population summary statistics included the number of loci, observed heterozygosity, and expected heterozygosity using dartR version 1.9.4 (Gruber et al. 2018); the Shannon-Weiner Diversity index and Simpson's index using poppr version 2.9.1 (Kamvar et al. 2014, 2015); and the mean allelic richness using popgenreport version 3.0.4 (Adamack et al. 2014; Gruber et al. 2019). Nei's (1987) pairwise F_{ST} was also calculated using hierfstat version 0.5-7 (Goudet et al. 2020).

Population genetic structure

To examine the neutral population genetic structure, we performed an analysis in STRUCTURE (Pritchard et al. 2000). STRUCTURE uses a Bayesian iterative algorithm to distinguish groups with similar patterns of variation and assigns individuals to groups based on likelihood (Pritchard et al. 2000; Porras-Hurtado et al. 2013). STRUCTURE analyses were only used to infer genetic clusters based on the neutral loci, at the recommended value of K (clusters), for each of the following data subsets: (1) all populations, (2) north-east New Zealand populations only, (3) the large size class across north-east New Zealand populations, and (4) the small size class across north-east New Zealand populations. StrAuto (Chhatre et al. 2017) was used to automate the STRUCTURE runs without any priors. Ten replicate runs, each with 10,000 MCMC chains (after 1,500 burn-in chains were discarded), were completed for each data subset. For the (3) and (4) data subsets, 1-9 clusters (k) were tested; for (1) and (2), 1-7 clusters were tested. The STRUCTURE results were uploaded to Structure Harvester (Earl 2011) to select the best number of clusters for each data subset, CLUMPP (Jakobsson 2009) was used to standardize the cluster designations across the 10 replicate runs, and ggplot2 version 3.3.3 (Wickham et al. 2020) was used to plot the results.

For each of the data subsets, we further investigated genetic structuring based on both neutral and adaptive loci using a Discriminatory Analysis of Principal Components (DAPC) using the adegenet package version 2.1.3 (Jombart 2008) in RStudio Version 1.2.1335 (RStudio Team 2020). DAPC is a multivariate method for clustering genetically related individuals and can handle a larger number of loci than STRUCTURE, also providing the option for significance testing according to *a priori* groupings (Jombart et al. 2010). In the DAPC method, first, a PCA is performed and then a Discriminatory Analysis (DA) that maximises the between-group variance and minimises the within-group variance. We used DAPC to test for significant genetic structuring naively (using the recommended value of k) for data subset (1) and (2), and according to our *a priori* population designations for (1), (2), (3) and (4). The number of principal components used was determined with cross-validation. We also investigated whether the genetic composition of the population size classes within north-east New Zealand populations differ, by using *a priori* above (large) and below (small) 90mm size groups (excluding Rangitāhua). Each of these DAPCs was performed using both the adaptive and neutral loci, separately.

Meta-population structure and population genetic covariance

Population graphs and their underlying covariance measures are a model-free way to examine meta-population structure. Using the methods of Dyer et al. (2004) the covariance measure is derived from an individual pairwise genetic distance calculated across all loci. Population graphs are a network created using graph theory with nodes, representing populations, connected by edges, which represent the genetic covariance between populations (Dyer et al. 2004). They differ from traditional population genomic analyses because the relationships among populations are studied simultaneously rather than through several pairwise relationships. The edges represent conditional covariance and the overall population graph will only include the edges that have significant power (using significance of 0.05) in explaining the topology of the population graph (Dyer 2015). In our case, we anticipated that only populations that have had recent gene-flow, and/or

represent source-sink relationships would retain edges between them. We used the R packages ‘gstudio’ version 1.5.2 (Dyer 2012) and ‘popgraph’ version 1.5.1 (Dyer 2017) in RStudio Version 1.2.1335 (RStudio Team 2020) to create the population graphs. We created population graphs for data subsets (1) - (4) based using both the adaptive and neutral loci, separately. Then, we created population graphs, excluding Rangitāhua, for both the neutral and adaptive datasets. The edge lengths in the population graphs are based on the between-population variance (for edges that are included) and the node size is based on the within-population variance (both displayed in three-dimensions). These variances are the same as the within and between population variances in AMOVA (Dyer 2015).

Using the R package “graph4lg” (Savary et al. 2020) we analysed the population graphs. Using the “compute_node_metric” function we calculated three measures of centrality for each node. The “degree centrality” is the number of edges from the node to other nodes and shows how many populations that population has significant genetic covariance with (Cross et al. 2018). The “closeness centrality” index measures the average shortest path from a node to all other nodes and tells us on average how many steps are required for that node to share genomic variance with the other nodes (Cross et al. 2018). Lastly, the “between centrality” index measures how many shortest paths the node is found on and is useful for identifying the importance of a particular node for the genetic exchange in the meta-population (Cross et al. 2018). Next, the nodes were grouped into modules using the “compute_graph_modul” function. Modules are groups of nodes that are made by maximising the modularity, which in our case are groups of populations that are demographically interdependent.

To compare the modularity of two graphs constructed based on different data subsets, we used the “graph_modul_compar” function of the graph4lg R package (Savary et al. 2020). This function calculates the Adjusted Rand Index reflecting how often pairs of nodes from the same module in one graph are also in the same module in the second graph. We also used the “graph_node_compar” function to assess whether the connectivity of each node is similar between the graphs. This function calculates Spearman’s correlation coefficient between the graph-theoretic metric values.

Results

Single nucleotide polymorphism dataset

The final dataset, after all the filtering steps, had 6,629 binary SNPs. There were 226 outlier SNPs identified using pcadapt (based on two PCs, see Appendix, Fig. A1) and 208 identified using Outflank. Between both methods, there was a total of 253 outliers (159 of which were detected by both Outflank and pcadapt). These loci were used to inform the creation of two datasets: adaptive (all outliers from both pcadapt and Outflank, 326 SNPs) and neutral (all except adaptive, 6,303 SNPs).

Population genetic summary statistics

All populations had a similar Shannon-Weiner Diversity index (3.367 to 3.434) and Simpson's index (0.966 to 0.968) except Takatu Point (Shannon-Weiner Diversity index 2.639; Simpson's index 0.929; Table 1). The Rangitāhua populations had lower observed heterozygosity (Raoul Island 0.155; L'Esperance Rock 0.156, Macauley Island: 0.159), expected heterozygosity (Raoul Island 0.250; L'Esperance Rock 0.250, Macauley Island: 0.252), and mean allelic richness (Raoul Island 1.697; L'Esperance Rock 1.696, Macauley Island: 1.701) than the rest of the New Zealand populations that we examined (observed heterozygosity 0.168 to 0.188; expected heterozygosity 0.262 to 0.269; mean allelic richness 1.721 to 1.743; Table 1). Pairwise F_{ST}-values were highest between Rangitāhua and north-east New Zealand populations (Table A2). Based on pairwise F_{ST}, Takatu Point was the most genetically differentiated of the north-east New Zealand populations (pairwise F_{ST}'s 0.0010 to 0.0033; Table A2). Of the Rangitāhua populations, L'Esperance Rock and Macauley Island were the least genetically differentiated based on pairwise F_{ST} (0.0005).

Population	Number of individuals	Number of individuals <90mm	Number of individuals ≥90mm	Number of loci	Observed Heterozygosity	Expected Heterozygosity	Shannon-Weiner Diversity index	Simpson's index	Mean allelic richness
Raoul Island (Rao)	30	NA	NA	1864	0.155	0.250	3.401	0.967	1.697
Macauley Island (Mac)	30	NA	NA	1946	0.159	0.252	3.401	0.967	1.701
L'Esperance Rock (LEs)	30	NA	NA	1875	0.156	0.250	3.401	0.967	1.696
Spirits Bay (Spi)	31	7	24	1975	0.171	0.266	3.434	0.968	1.736
Berghan Point (Ber)	29	17	12	1758	0.168	0.266	3.367	0.966	1.734
Cavalli Islands (Cav)	31	13	18	1999	0.172	0.268	3.434	0.968	1.739
Home Point (Hom)	30	19	11	728	0.172	0.265	3.401	0.967	1.734
Poor Knights Islands (Poo)	29	14	15	2051	0.173	0.267	3.367	0.966	1.738
Mokohinau Islands (Mok)	30	22	8	2067	0.179	0.268	3.401	0.967	1.741
Sail Rock (Sai)	30	12	18	2058	0.180	0.266	3.401	0.967	1.737
Takatu Point (Tak)	14	3	11	3236	0.188	0.262	2.639	0.929	1.721
Mercury Islands (Mer)	30	14	16	1977	0.184	0.269	3.401	0.967	1.741
Alderman Islands (Ald)	30	25	5	2032	0.179	0.268	3.401	0.967	1.740
Castle Rock (Cas)	30	12	18	1928	0.176	0.267	3.401	0.967	1.739
Mayor Island (May)	30	5	25	1904	0.187	0.269	3.401	0.967	1.743
White Island (Whi)	30	19	11	2062	0.179	0.269	3.401	0.967	1.743

Table 1: Population genetic summary statistics for each sampled population of *Centrostephanus rodgersii* based on the neutral loci dataset. The number of individuals in each of the size categories (<90mm and ≥ 90mm) are not reported for populations of Rangitāhua as they were not included in any size-based analyses. The evenness was the same for all populations: 1. Populations are ordered from north to south.

Population genetic structure

Naïve clustering of individuals based on the neutral loci using both STRUCTURE and DAPC found two significant clusters (K = 2), roughly equating to populations from north-east New Zealand and populations from Rangitāhua. In the STRUCTURE analysis, four north-east New Zealand individuals (one each from Mercury Islands, Castle Rock, Mayor Island, and White Island) were assigned to the "Rangitāhua" cluster with almost 100% probability, and several others from these populations as well as the Mokohinau Islands were assigned to this group with between 50-80% probability (Table A3). In the DAPC without priors, only one discriminant function was retained, which distinguished the north-east New Zealand and Rangitāhua clusters (Fig. A2). However, between one and three individuals from each of Castle Rock, Mayor Island, Mercury Islands, Mokohinau Islands, and White Island were placed in the Rangitāhua cluster (Fig. 2b). When we used the k-means clusters on these two groups, no further subclusters were found.

In the DAPC that assigned individuals based on population (Fig. 2c; using the first 152 PCs and the proportion of conserved variance was 0.46), the first discriminant function corresponded to the separation between the same north-east New Zealand and Rangitāhua clusters found with the naïve clustering (Fig. A2). The individuals from north-east New Zealand that were found clustered with the Rangitāhua group ("Rangitāhua-like" individuals) in the naïve DAPC (Fig. 2b, A2) were likely those individuals found between the two clusters along the first axis (Fig. 2c). On the second discriminatory axis, only the population of Takatu Point was somewhat separated from the other north-east New Zealand populations and the population of L'Esperance Rock from the other Rangitāhua populations, though to a lesser extent (Fig. 2c).

DAPCs for just the north-east New Zealand populations revealed no further significant clustering (either with, Fig. A3, or without *a priori* population groupings which showed no clustering) based on the neutral loci. The DAPC for the north-east New Zealand populations using *a priori* population groupings again showed the population of Takatu Point was slightly separated (Fig. A3). The DAPC for the Rangitāhua populations using *a priori* groupings revealed each population centroid differed, but their dispersions overlapped considerably (Fig. A4).

Although the STRUCTURE analyses suggested k =2 within the north-east New Zealand samples, there was no genetic structuring according to geography (Fig. A5). Instead, the same ten individuals (except one), from north-east New Zealand offshore island populations, previously assigned to the Rangitāhua cluster (based on neutral SNPs for all populations) were all assigned in the same STRUCTURE grouping, with over 100 further individuals (Table A3).

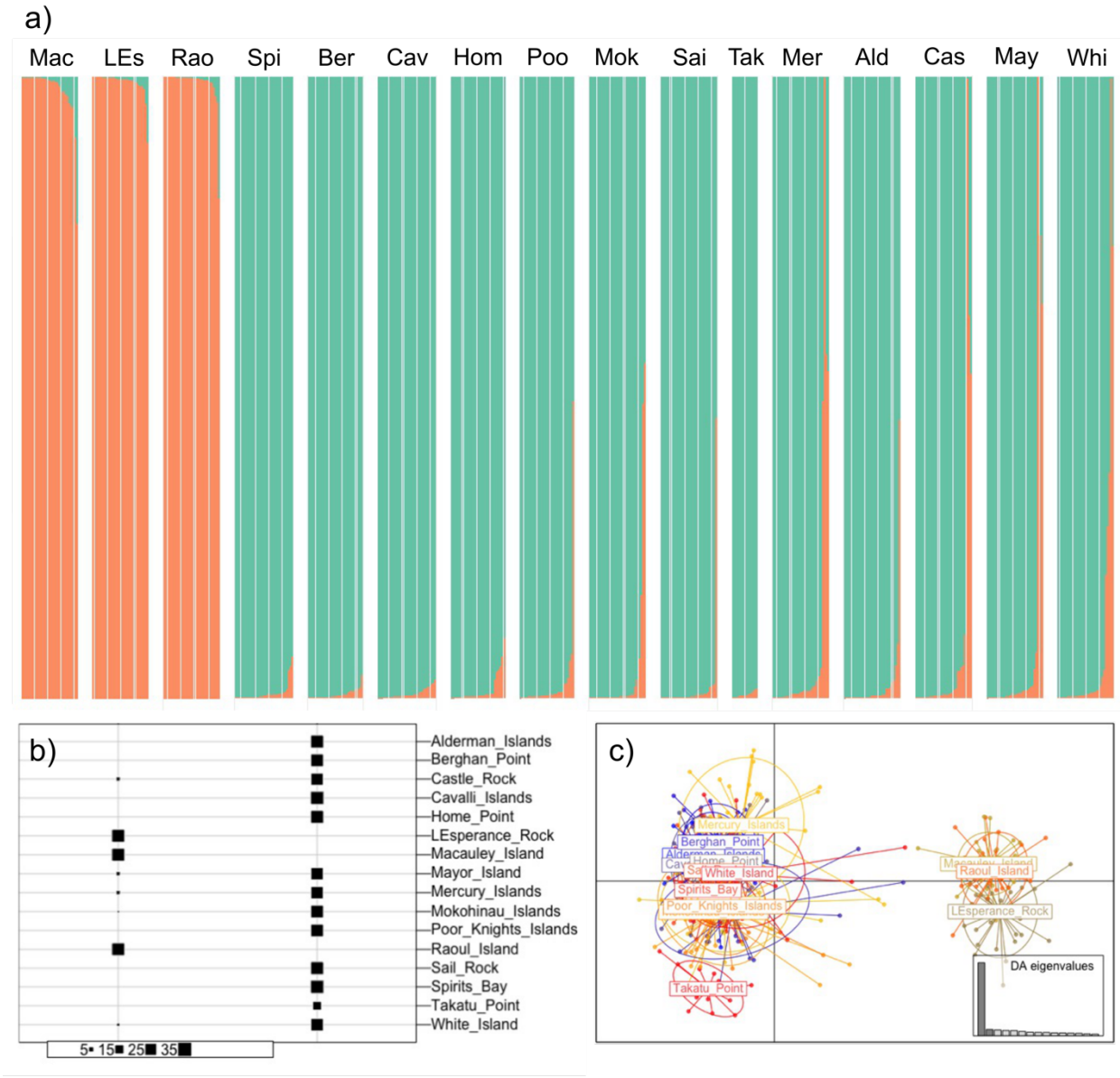

Figure 2: Clustering of *Centrostephanus rodgersii* individuals from north-east New Zealand and Rangitāhua based on the neutral loci. a) STRUCTURE analysis result with colours representing the two clusters found; b) DAPC clustering result without *a priori* grouping of individuals into populations showing two clusters (final DAPC in Figure A2); c) DAPC clustering result with *a priori* population groupings, colours represent the different populations and points represent individuals. Full population names are given in Table A11.

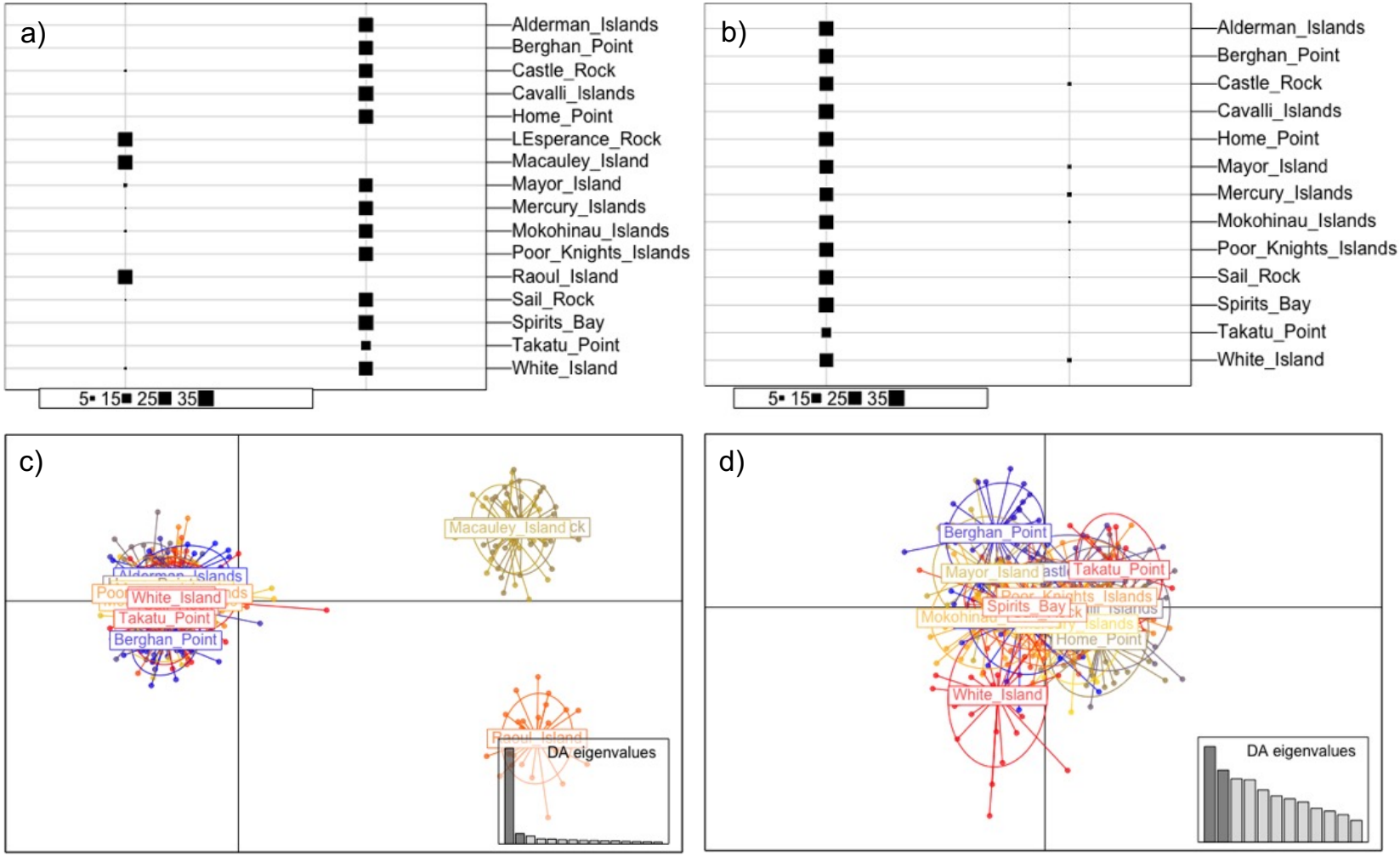

Figure 3: Clustering of *Centrostephanus rodgersii* individuals from north-east New Zealand and Rangitāhua based on the putatively adaptive loci. a) Naïve clustering for north-east New Zealand and Rangitāhua, final DAPC in Fig. A9); b) Naïve clustering for north-east New Zealand, final DAPC in Fig. A10; c) DAPC with *a priori* population groupings for north-east New Zealand and Rangitāhua, colours represent the different populations and points represent individuals. d) DAPC with *a priori* population groupings for the north-east New Zealand, colours represent the different populations and points represent individuals.

Naïve clustering based on the adaptive loci only (using DAPC) recovered similar north-east New Zealand and predominantly Rangitāhua clusters as when using the neutral loci for all populations (Table A3; Fig. 3a; Fig. A9). For the north-east New Zealand populations, two clusters were found based on the adaptive loci (Fig. 3b; Fig. A10). The smaller cluster contained individuals that had a Rangitāhua-like genotype (based on adaptive and/or neutral loci as shown in Table A3), along with six individuals from the Poor Knight Islands (one), White Island (two), the Mercury Islands (two), and the Alderman Islands (one) (Table A3).

The DAPC using *a priori* population groupings and based on the putatively adaptive loci showed the same separation of the Rangitāhua populations and north-east New Zealand populations, similar to the neutral loci, however, the Raoul Island population was distinct from the rest of the other Rangitāhua populations (Fig. 3c; based on 220 PCs and the proportion of conserved variance was 0.966). In the DAPC for north-east New Zealand populations using *a priori* defined populations, all populations were in one clustered group, although the White Island population was slightly differentiated (Fig. 3d; based on first 144 PCs and the proportion of conserved variance was 0.885).

DAPC suggested a significant difference in the genetic composition of the two size classes sampled across populations of north-east New Zealand based on the neutral loci (Fig. 4a; using the first 112 PCs and the proportion of conserved variance was 0.41). Within each of the size classes, there were no significant genetic clusters according to DAPC (naively or using *a priori* population groupings) based on the neutral loci. Even so, when visualised with *a priori* population groupings, the small size class of the Mercury Islands population was somewhat separated from the other populations (Fig. 4b; using the first 54 PCs and the proportion of conserved variance was 0.374), and for the large size class, the White Island and Takatu Point populations were also somewhat differentiated from the other populations (Fig. 4c; using the first 30 PCs and the proportion of conserved variance was 0.21).

STRUCTURE analyses suggested were either K = 2 or K = 7 for the large size class (Fig. A6, A7). There was no evident structuring according to populations for K = 7, but in K = 2, a few individuals from each of the Mercury Islands, Castle Rock, Mayor Island and White Island were being assigned to a second group. Similarly, for the small size class, STRUCTURE analyses suggested K = 2, wherein a few individuals from these populations had a high probability of assignment to the second cluster, albeit not as high as for the large size class (Fig. A8). These were some of the same individuals assigned to the “Rangitahua-like” cluster in previous analyses.

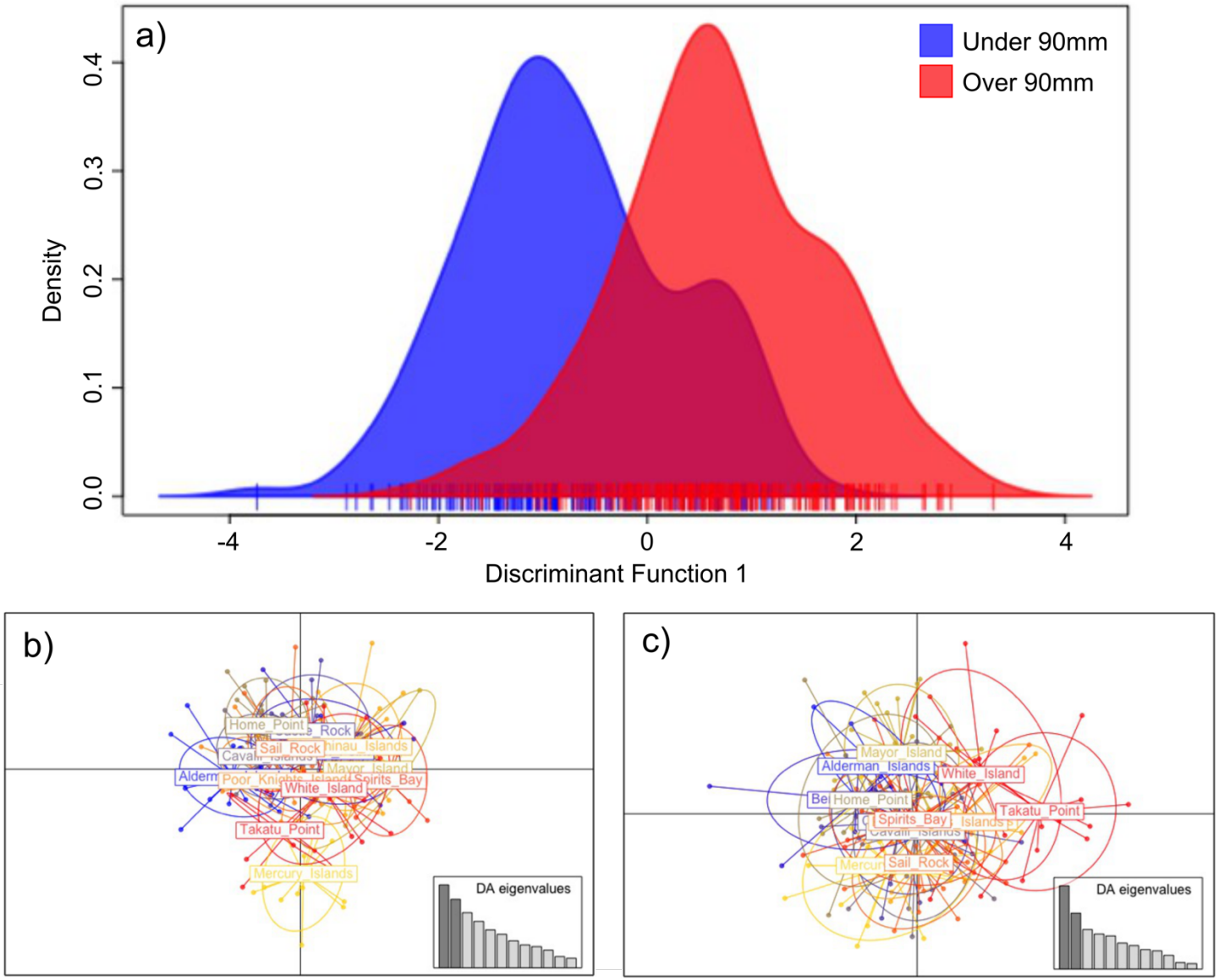

Figure 4: DAPCs of *Centrostephanus rodgersii* individuals from north-east New Zealand distinguished into size classes, and based on neutral loci. a) DAPC using *a priori* size grouping result with colours representing the two clusters found; b) DAPC using *a priori* population groupings using small individuals (under 90mm), colours represent the different populations and points represent individuals. c) DAPC using *a priori* population groupings using large individuals (90mm and over), colours represent the different populations and points represent individuals.

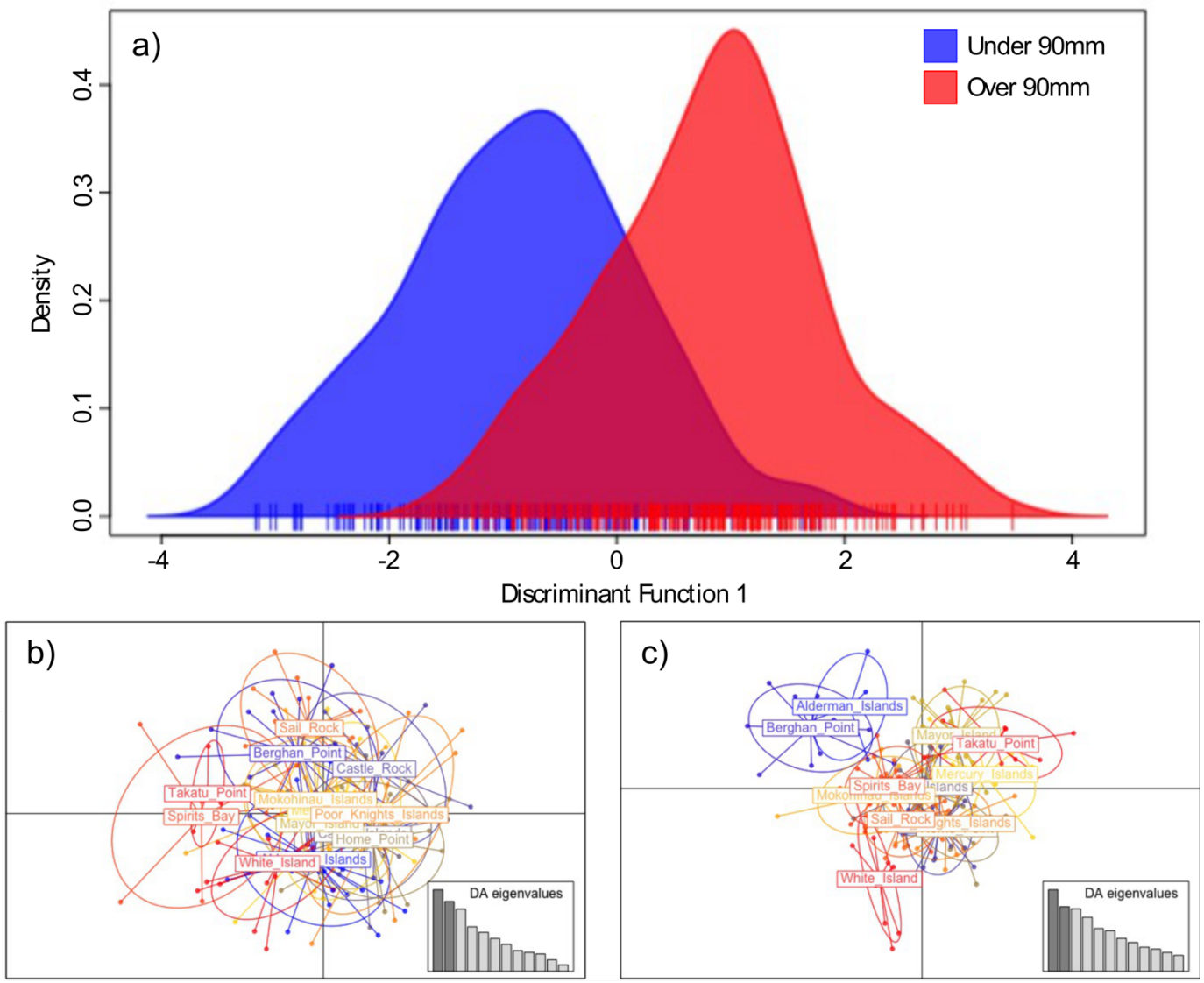

Figure 5: DAPCs of *Centrostephanus rodgersii* individuals from north-east New Zealand distinguished into size classes, and based on putatively adaptive loci. a) DAPC using *a priori* size grouping result with colours representing the two clusters found based on putatively adaptive loci; b) DAPC using *a priori* population groupings using small individuals (under 90mm) based on putatively adaptive loci, colours represent the different populations and points represent individuals. c) DAPC using *a priori* population groupings using large individuals (90mm and over) based on putatively adaptive loci, colours represent the different populations and points represent individuals.

Similar to patterns revealed for the neutral loci, the DAPC based on putatively adaptive loci indicated a difference in the genetic composition between the size classes sampled across populations of north-east New Zealand (Fig. 5a; using the first 144 PCs and the proportion of conserved variance was 0.885). With *a priori* population groupings within the small size class, we found no clusters based on the adaptive loci, but there were two clusters found for the large size class (Fig. A12). Within the small size class, the Takatu Point and Spirits Bay populations were slightly separated from the other populations (Figure 5b; based on the first 42 PCs and the proportion of conserved variance was 0.565). Within the large size class, the White Island, Berghan Point, and the Alderman Islands populations were slightly separated from the other populations (Figure 5c; based on the first 96 PCs and the proportion of conserved variance was 0.862).

All the individuals with a Rangitāhua-like genotype based on the neutral loci were from the populations: Castle Rock, White Island, Mayor Island, and the Mokohinau Islands and they were mostly between 60 and 70 mm, or over 100mm in test diameter (Table A3). Based on putatively adaptive loci, the individuals that have a Rangitāhua-like genotype differed slightly (eight individuals were detected based on both the neutral and adaptive loci, two individuals only based on the neutral loci, and three individuals only based on the adaptive loci; Table A3). The individuals with the Rangitāhua-like genotype based on adaptive loci came from the same populations as those detected based on neutral loci, but additionally included the Sail Rock population, and ranged in size from 46 to 125mm (Table A3).

Meta-population structure and population genetic covariance

The population graph including all populations, based on neutral loci, was split into two disconnected, distinct subgraphs: Rangitāhua and north-east New Zealand (Fig. A13a,b). The north-east New Zealand subgraph had 13 nodes (populations), and 32 edges were retained between populations that had significant conditional genetic covariance. The Rangitāhua subgraph had three nodes and three edges that had significant conditional genetic covariance. The population graph based on putatively adaptive loci for the same populations (Fig. A13c,d) had similar links (correlation 0.506; Table A10) and metric values (correlation 0.730; Table A10), comprising 30 edges, and Rangitāhua was a separate subgraph.

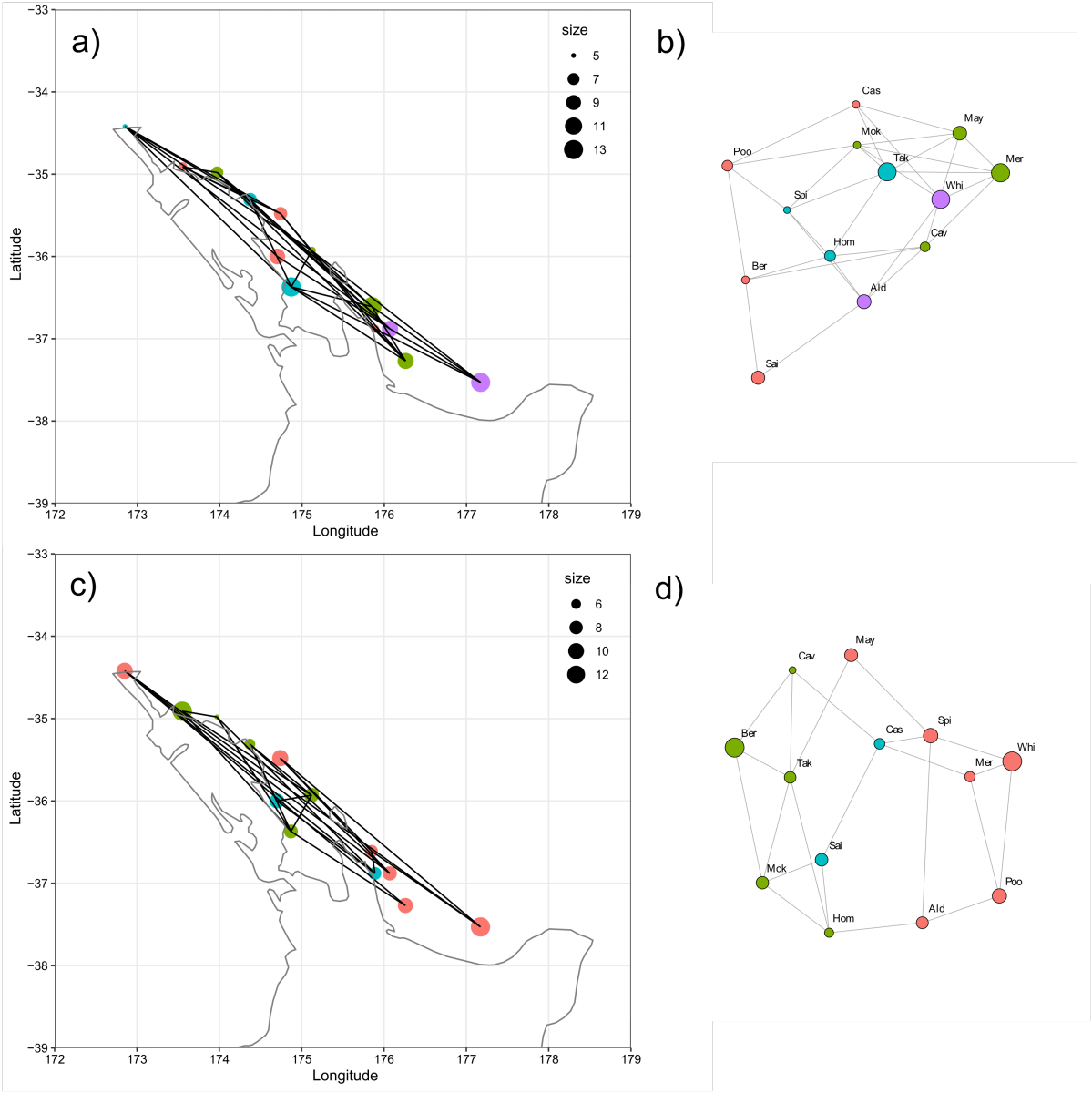

Figure 6: Population graphs of north-east New Zealand: a) based on the neutral loci displayed geographically, and b) based on the neutral loci displayed aspatially; and c) based on the adaptive loci displayed geographically, and d) based on the adaptive loci displayed aspatially. Each node represents a sampled population and edges that significantly contribute to the overall topology of the graph are retained. The colours represent the different modules within the graph as shown in Tables A4 and A5; the size of the nodes represent within-population variance (values in Tables A4 and A5), and the length of the edges represent the between-population covariance (Tables A12 and A13). Note the aspatial population graphs are three dimensional but are only displayed in two dimensions; accurate edge lengths are in Tables A12 and A13. Abbreviated names are as follows: Alderman Islands: Ald, Berghan Point: Ber, Castle Rock: Cas, Cavalli Islands: Cav, Home Point: Hom, L'Esperance Rock: LEs, Macauley Island: Mac, Mayor Island: May, Mercury Islands: Mer, Mokohinau Islands: Mok, Poor Knights Islands: Poo, Raoul Island: Rao, Sail Rock: Sai, Spirits Bay: Spi, Takatu Point: Tak, White Island: Whi.

The population graph for north-east New Zealand populations based on neutral loci was quite connected (Fig. 6b). The four modules detected within the graph did not correspond to the geographic proximity of the populations (Fig. 6a). Takatu Point (13.240), White Island (13.110), and the Mercury Islands (13.328) populations had the greatest within-population neutral genetic variance, whereas the Spirits Bay population had the least (5.000; Table A4). Takatu Point and the Mokohinau Islands were the most connected populations (six degrees; Table A4) and the Sail Rock population was the least (two degrees; Table A4). Along with having the most edges, the Takatu Point population had the shortest edge lengths (9.38 to 10.82; Table A12) and therefore has the most shared variance with other populations. The Sail Rock population had the longest edge lengths (14.65 and 15.02; Table A12) and so is the most genetically distinct population. The shortest edge length was between the Takatu Point and Mercury Islands populations. The Alderman Islands population had the highest closeness centrality index (0.713) and the other populations ranged from 0.512 to 0.697 (Table A4). The Alderman Islands also had the highest betweenness centrality index (12) while the others range from zero to eight (Table A4).

The population graph for north-east New Zealand populations based on putatively adaptive loci had three modules which also did not correspond to the geographic proximity of the populations (Fig. 6c). The White Island (13.645) and Berghan Point (13.717) populations had the most within-population adaptive genetic variance and the Cavalli Islands population had the least (5.000; Table A5). The Takatu Point and Home Point population shared the most genetic variance (edge length: 11.27; Table A13). The Takatu Point population was the most connected (five degrees) and had the shortest edge lengths (11.27 to 12.69; Table A13), meaning it shared the highest amount of variance with other populations. All other edge lengths ranged from (15.03 to 17.22; Table A13). The Mayor Island population was the least connected population (two degrees; Table A5). The Castle Rock population had the highest closeness centrality index (0.779) and the Spirits Bay (16) and Castle Rock (15) populations had the highest betweenness centrality indices (Table A5).

The two population graphs for north-east New Zealand populations based on neutral and adaptive loci differed in their number of edges: the neutral graph had 30 whereas the adaptive only had 22. The topography based on edges of the two population graphs for north-east New Zealand populations based on the neutral and adaptive loci were not correlated (correlation -0.085; Table A10). The degrees of populations between the two population graphs based on the neutral and adaptive loci were correlated both with and without the Rangitāhua populations (0.426 correlation without Rangitāhua populations; 0.730 correlation with Rangitāhua populations; Table A10). The modularity between the two population graphs for north-east New Zealand populations based on the neutral and adaptive loci also differed (Adjusted Rand Index 0.229; Table A10), the neutral graph had four modules (Fig. 6a) whereas the adaptive had three modules (Fig. 6c).

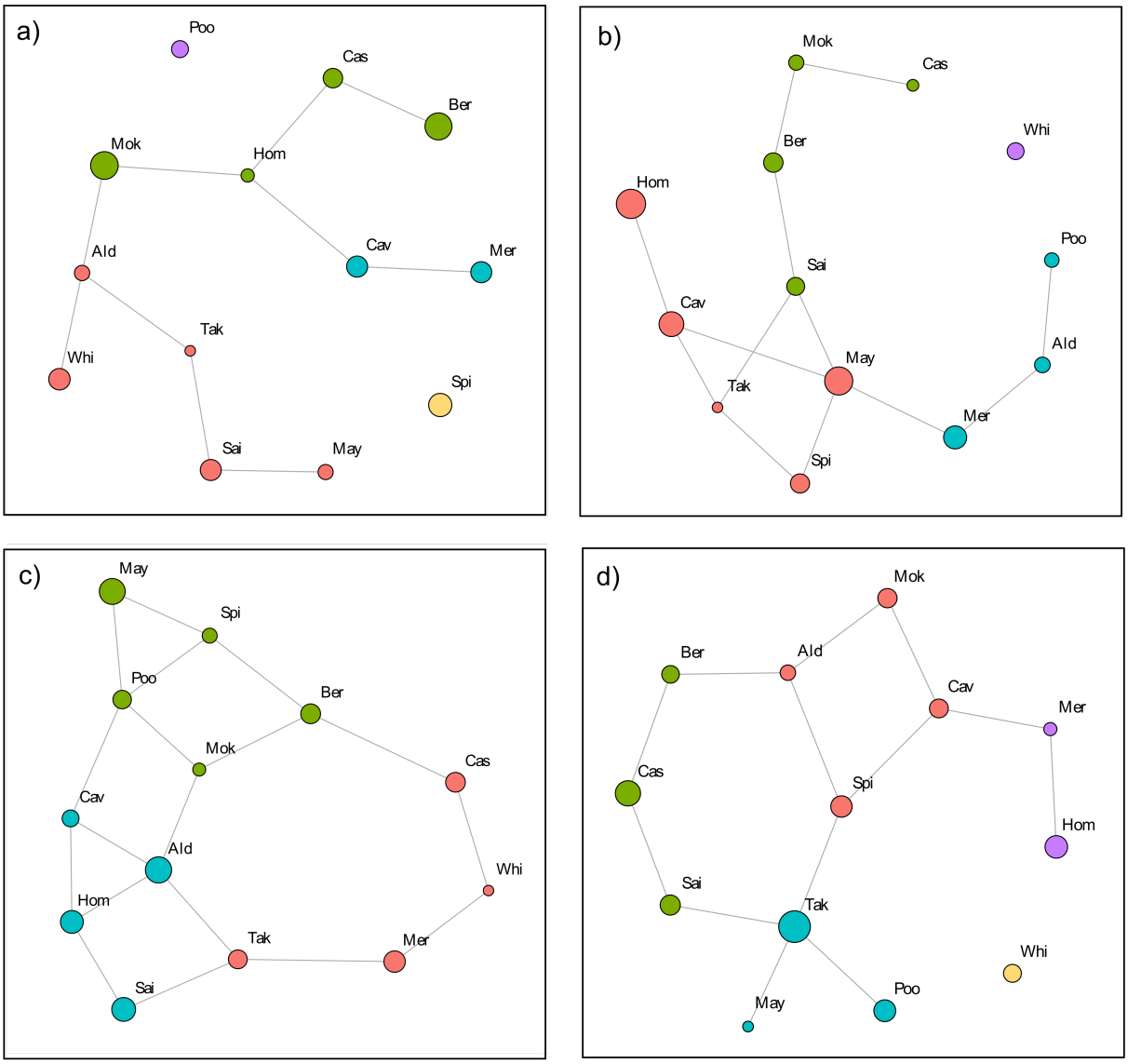

Figure 7: Aspatial population graphs based on small and large individuals in north-east New Zealand: a) based on neutral loci from large individuals in north-east New Zealand; b) based on neutral loci from small individuals in north-east New Zealand; c) based on putatively adaptive loci from large individuals in north-east New Zealand; d) based on putatively adaptive loci from small individuals in north-east New Zealand. There is a node for each population and an edge between the nodes when that edge significantly contributes to the overall topology of the graph. The colours represent the different modules as shown in Tables A6-A9, the sizes of the nodes are set by the within-population variance in three dimensions (numbers in Table A6-A9), and the length of the edges are set by the between-population covariance (Tabless A14-A17). Note the aspatial population graphs are three dimensional but are only displayed in two dimensions; accurate edge lengths are in Tables A14-A17. Abbreviated names as follows: Alderman Islands: Ald, Berghan Point: Ber, Castle Rock: Cas, Cavalli Islands: Cav, Home Point: Hom, L'Esperance Rock: Les, Macauley Island: Mac, Mayor Island: May, Mercury Islands: Mer, Mokohinau Islands: Mok, Poor Knights Islands: Poo, Raoul Island: Rao, Sail Rock: Sai, Spirits Bay: Spi, Takatu Point: Tak, White Island: Whi.

The population graph for small individuals based on neutral loci (Fig. 7b) differed from the population graph for large individuals based on neutral loci (Fig. 7a) in the edges (0.137 correlation; Table A10), slightly in the number of node degrees (0.096 correlation; Table A10), and in the modules (Adjusted Rand Index 0.026; Table A10); suggesting that there is a difference in patterns of population connectivity for individuals up to 15 years old and those 15 years and older.

In the population graph for large individuals based on the neutral loci, there were five modules (Table A7) and the Spirits Bay and the Poor Knight Islands populations were disconnected from the graph (Fig. 7a). The Mokohinau Islands population had the largest within-population variance (13.086; Table A7) and Takatu Point population had the smallest within-population variance (5.000; Table A7). The highest number of degrees was three (Alderman Islands and Home Point; Table A7). All populations had a similar closeness centrality index (0.032 to 0.034; Table A7) except the disconnected populations (Spirits Bay and the Poor Knight Islands). The Alderman Islands (27), Home Point (28), and the Mokohinau Islands (25) populations had the highest betweenness centrality indices (Table A7). The shortest edge length was between the Alderman Islands and the Mokohinau Islands (5.53) populations and all the other shorter edge lengths were between the Alderman Islands population and other populations (5.53 to 6.26; Table A14). All remaining edge lengths were between 7.31 to 9.74 (Table A14).

In contrast to the population graph for large individuals based on neutral loci, the population graph for small individuals based on neutral loci had only the White Island population disconnected and there were four modules (Fig. 7b). The Mayor Island population had the lowest within-population variance (5.000), and Takatu Point population had the highest within-population variance (14.768, in contrast to Takatu Point having the lowest within-population variance for the population graph for large individuals based on neutral loci; Table A8). The Takatu Point population also had the most edges (four degrees; Table A8) and the highest betweenness centrality index (19 along with the Alderman Islands; Table A8). The closeness centrality indices ranged from 0.049 (Mayor Island) to 0.060 (the Alderman Islands and Spirits Bay), except for the White Island population (0.006; Table A8) which was disconnected. The smallest edge length was between the Takatu Point and Spirits Bay populations (3.84) and all the shortest edge lengths were between the Takatu Point population and other populations (3.84 to 4.48; Table A15).

The population graph based on putatively adaptive loci for small individuals differed from the population graph based on putatively adaptive loci for large individuals as their edges were not correlated (Mathew's correlation of 0; Table A10), the number of node degrees were only slightly correlated (Spearman's correlations 0.342; Table A10), and the modules had an Adjusted Rand Index of 0.062 (Table A10).

The population graph based on putatively adaptive loci for large individuals had four modules (Table A9; Fig. 7c). The Alderman Islands (12.529) and Mayor Island (12.376) populations had the largest within-population variance and the White Island population had the smallest within-population variance (5.000; Table A9). The highest number of degrees was four (Alderman Islands and the Poor Knight Islands; Table A9). The closeness centrality index range from 0.214 to 0.268 and the Cavalli Islands population had the highest

betweenness centrality index (16; Table A9). When the population graph based on putatively adaptive loci for large individuals was compared to the population graph based on neutral loci for the same large individuals, the modularity has an Adjusted Rand Index of 0.081, the number of node degrees was 0.120, and the edges were correlated (correlation of 0.245; Table A10). The edge lengths for population graph based on putatively adaptive loci for large individuals ranged from 4.62 to 9.02 (Table A16) and the smallest edge length was between the Mokohinau Islands and the Alderman Islands populations (4.62; Table A16).

In the population graph based on putatively adaptive loci for small individuals, there were five modules and like the graph based on neutral loci for the same individuals, the White Island population was disconnected (Table A8; Fig. 7d). The Takatu Point population had the largest within-population variance (14.768 the opposite of the graph based on neutral loci for the same individuals) and the Mayor Island population had the smallest within-population variance (5.000 which was one of the highest for the graph based on neutral loci for the same individuals; Table A8). The highest number of degrees was four for the Takatu Point population (Table A8) and all the shortest edge lengths were between this population and other populations (3.41 to 4.05; Table A17), indicating that individuals sampled at Takatu Point share a lot of genetic variance with other populations. All other edge lengths ranged from 6.41 to 10.14 (Table A17). The closeness centrality index ranged from 0.049 to 0.060 except for the White Island population (0.006) which was disconnected. The Alderman Islands and Takatu Point populations had the highest betweenness centrality index (19; Table A8). Comparing the two population graphs for small individuals for neutral and putatively adaptive loci, the modularity had an Adjusted Rand Index of 0.165 and the number of node degrees was not very correlated (Spearman's correlation 0.082, Mathew's correlation -0.015, Table A10).

Discussion

A change in the distribution and/or abundance of a species within a region is associated with changes in the genetic composition of its populations and the distribution of genetic variation across its range. In this study, we uncovered likely changes in population connectivity and the distribution of the neutral and putatively adaptive genetic variance in *Centrostephanus rodgersii* populations over recent decades in New Zealand. We found that individuals from Rangitāhua and north-east New Zealand differ genetically but there is some ongoing recruitment from Rangitāhua to north-east New Zealand. Through examining the population structure and covariance of differently sized cohorts within north-east New Zealand, we gained further insights into how the demography of this species has changed over the last decade or so. Here we discuss the implications of our results for understanding the population history and potential future of *C. rodgersii* in New Zealand. We also discuss the broader contributions of our study to understanding population responses to climate change.

Our study revealed that populations of *C. rodgersii* in Rangitāhua had a different genomic composition from populations of *C. rodgersii* in north-east New Zealand. Our findings contrast with a recent study based on microsatellites that suggested that there was little genetic differentiation between populations of *C. rodgersii* in Rangitāhua and north-east

New Zealand (Thomas et al. 2021). Nonetheless, our study includes more New Zealand *C. rodgersii* populations than previous studies (Banks et al. 2007; Thomas et al. 2021) and uses SNPs, which can reveal finer-scale population genetic differentiation than microsatellites (e.g. Candy et al. 2015; Zimmerman et al. 2020). Although our results indicate that there is low migration between Rangitāhua and north-east New Zealand populations of *C. rodgersii*, we did recover some individuals from north-east New Zealand that were consistently assigned to a cluster dominated by individuals from Rangitāhua, indicating they had a Rangitāhua-like genotype. Dispersal between the Rangitāhua archipelago and north-east New Zealand is known to occur based on species occurrences. For example, two fish species endemic to Rangitāhua, the Kermadec scalyfin (*Parma kermadecensis*) and the Kermadec demoiselle (*Chrysiptera rapanui*), are reported in north-east New Zealand periodically indicating there is larval dispersal from Rangitāhua to north-east New Zealand (Francis et al. 1999; Liggins et al. 2021). The *C. rodgersii* individuals with the Rangitāhua-like genotype, based on neutral loci, were found in Castle Rock, Mayor Island, the Mokohinau Islands, and White Island – which are all offshore islands and in the southern part of the distribution of the *C. rodgersii* in New Zealand. Furthermore, they had a range of sizes from 67 to 125 so they likely span multiple migration events. Based on the assignment values of these individuals recovered in the STRUCTURE analysis, some are likely to be first-generation immigrants from Rangitāhua that have dispersed to north-east New Zealand as larvae, and some are likely to be second-generation migrants that are as a result of admixture between the first-generation migrants from Rangitāhua and local north-east New Zealand individuals.

In north-east New Zealand, we found that there were no distinct genetic clusters according to geography, that genetic differentiation among populations was low, and that geographically proximal populations of *C. rodgersii* did not necessarily have high genetic covariance. The population graphs revealed that the north-east New Zealand populations of *C. rodgersii* were a large meta-population, within which some populations uniquely shared genetic variance with many other populations (e.g. over eight of the populations had a degree of five or six; Fig. 6a,b; Table A4) and some that shared genetic variance with relatively few populations (e.g. Sail Rock had only two degrees; Fig.6a,b; Table A4). Modularity analysis allows us to split meta-populations into sub-populations that are anticipated to share gene flow and therefore have a more similar genetic make-up. Based on the population graph topology, four modules were distinguished that comprised of geographically distant populations (Fig. 6a). Previous studies have found that although modules are often geographically close, such as for the red abalone (*Haliotis rufescens*) in the Southern California Bight (Peña et al. 2017), there are cases where modularity is not predictable according to the geographic arrangement of populations (e.g. Fletcher et al. 2013) and can be the result of dispersal being unrelated to geographic distance (Gilarranz 2020). *Centrostephanus rodgersii* has a long pelagic larval stage (Huggett et al. 2005) and the oceanography and currents around north-east New Zealand are complex (Stanton et al. 1997), potentially leading to counterintuitive dispersal among populations. For instance, even the *C. rodgersii* populations of White Island and Spirits Bay, at the geographic extremes of our sampled extent, were still relatively connected to all the other populations, have low pairwise F_{ST} values, and are genetically similar to each other (pairwise F_{ST} of 0 between White Island and Spirits Bay, Table A2). White Island and Spirits Bay could have been colonised at a similar time from the same source, leading to their similarity, and/or they could be receiving immigrating larvae from the same source. These results indicate that

the leading and trailing edges of the range of *C. rodgersii* in New Zealand are likely well connected by dispersal to other populations.

Our study found lower genetic diversity of *C. rodgersii* populations in Rangitāhua than in north-east New Zealand (Table 1). This result contrasts with previous findings based on microsatellites, where populations of Rangitāhua were noted to have higher genetic diversity than populations of north-east New Zealand (Thomas et al. 2021). The smaller number of individuals sampled from Rangitāhua, relative to mainland New Zealand, could have contributed to the lower genetic variance observed in our study due to ascertainment bias in bioinformatically selecting polymorphic loci. Nonetheless, Rangitāhua is geographically isolated with a small habitat area available to shallow reef organisms and is at the edge of the *C. rodgersii*'s range, that could also lead to the lower genetic variance in these populations (as also found in other echinoderms species, e.g. Liggins et al. 2014). In contrast to Rangitāhua, north-east New Zealand is influenced by the EAC via the Tasman Front, thus over time it has likely had immigration from parts of Australia (Banks et al. 2007; Thomas et al. 2021) as well as receiving migrants from Rangitāhua, increasing exchange and therefore genetic variance. Within north-east New Zealand, the northern-most populations of *C. rodgersii* had the lowest neutral genetic variance (Fig. 6a). This pattern could indicate that these populations are newer as the founder effect can cause new populations to have smaller genetic variance (Sirkkamaa 1983). However, size structure analysis suggests there has been a poleward range extension of *C. rodgersii* from North Cape (Spirit's Bay) to the Mokohinau Islands and that southern populations are recruiting more regularly (Chapter Two). Therefore, it is more likely that southern populations have higher neutral genetic variance because they have a larger population size and have more regular recruitment from a diverse range of sources.

Often patterns of spatial genetic differentiation based on putatively adaptive loci correlate with those based on neutral loci for the same individuals, but provide greater precision (e.g. Candy et al. 2015; Liggins et al. 2019) that can be useful for informing stock assessments (Mariani et al. 2013) and conservation. In our analyses, Rangitāhua populations were more genetically differentiated from each other based on the putatively adaptive loci compared to neutral loci. Additionally, north-east New Zealand populations of *C. rodgersii* did form two clusters based on the putatively adaptive loci (Fig. 3b), as compared to only one based on the neutral loci (without population groupings imposed *a priori*). The clusters corresponded to those that had a Rangitāhua-like genotype and those that did not. Additionally, the White Island population was more genetically distinct from the other north-east New Zealand populations with the DAPC based on putatively adaptive loci (Fig. 3d) than the DAPC based on neutral loci (Fig. A3). The population graphs provided a view as to why patterns of population genetic variation differed for the neutral versus putatively adaptive loci. In contrast to the poleward increase of within-population variance detected using neutral loci, within-population variance informed by the putatively adaptive loci varied among populations in an irregular spatial pattern. Based on these patterns, populations with higher putatively adaptive genetic variance such as White Island and Berghan Point may have higher putatively adaptive potential, and relaxed selection at those putatively adaptive loci (Table A5). In contrast, populations with low putatively adaptive genetic variance, such as the Cavalli Islands and Home Point may have low putatively adaptive potential, as a consequence of a selection bottleneck or directional selection acting on those

particular loci. Both inferences regarding the putatively adaptive potential within a population, as well as which populations are under selection within a seascape (and according to what selection pressures) are informative for understanding the future potential for populations to thrive under future climate scenarios.

When we examined the two size cohorts of *C. rodgersii* sampled across north-east New Zealand separately, the distribution of putatively adaptive and neutral genetic variance differed and we recovered patterns that were not present when these size cohorts were combined. For instance, when comparing patterns of putatively adaptive genetic variance across the sampled size cohorts, the Mayor Island population had high within-population putatively adaptive genetic variance in the large/older group (12.376; Table A9) and then small within-population putatively adaptive genetic variance in the smaller/younger group (5.000; Table A8), potentially indicating a selection bottleneck of these younger cohorts. Through comparing the population graphs for the older, younger, and overall population samples, we can see how the distribution of genetic variation has shifted over time. For instance, in the population graphs for the older/larger group, the populations of Spirits Bay and the Poor Knights Islands are disconnected from the graph (Fig. 7a), whereas in the younger/smaller group they are connected (Fig.7b). Overall (with both age groups), these two populations are connected (Fig. 6b) and have quite low pairwise F_{ST} (Spirits Bay -0.0002 to 0.0016, Poor Knight Islands 0 to 0.0014; Table A2, excluding Rangitāhua) indicating they are genetically similar to the other populations overall. The difference in the population graphs indicates that the older individuals in each population may have come from separate colonisation events, or different genotypes have persisted within each population, since which time these individuals have contributed to gene-flow among populations. Alternatively, colonisation by the younger/smaller individuals may have been more geographically widespread. This could be because of a potential recent range extension of *C. rodgersii* in northern north-east New Zealand (Chapter Two) that has led to the younger/smaller individuals being very similar. Furthermore, based on putatively adaptive loci, the larger/older individuals of the Spirits Bay population and the Poor Knights Islands population are connected in the graph (Fig. 7c). Such a pattern could be due to similar environmental pressures (i.e. selection pressures) at these locations, meaning that the putatively adaptive variance that is maintained may beneficial for the survival of individuals at both locations.

Using a combination of methods to understand patterns in the distribution of genetic variation across the *C. rodgersii* meta-population allowed us to have further confidence in our inferences as each analysis had different assumptions, for example, population graphs do not assume an underlying model (Dyer 2015). Additionally, each method drew from different information within the genotype dataset. For instance, conditional covariances among populations based on shared alleles are used in population graphs, whereas the discrimination of individuals into clusters based on the decomposition of the overall sampled variance is used in DAPCs. The combination of multiple genetic measures, the multivariate discriminant analyses, and analyses of genetic covariance and meta-population graph topology, revealed some *C. rodgersii* populations of particular interest. First, Takatu Point's population of *C. rodgersii* had the lowest genetic diversity (mean allelic richness 1.721; Simpson's index 0.929; Table 1; smallest within-population variation for both neutral small and neutral large 5.000; Table A6 and Table A7), and was the most genetically

differentiated population (F_{ST} = 0.0010 to 0.0033; Table A2; and DAPCs in Fig. 2c and 4c) within north-east New Zealand. Although the smaller number of sampled individuals (14) in the Takatu Point population means population genetic diversity estimates could be unreliable (Konopinski 2020), pairwise genetic differentiation measures are more robust to sample size differences. Furthermore, the Takatu Point population also had the highest observed heterozygosity (0.188) but the lowest expected heterozygosity (0.262) within north-east New Zealand (Table 1). These results may indicate that the Takatu Point population is the result of only a few immigration events, potentially from formerly genetically distinct populations, that are now admixed. Although Takatu Point's population of *C. rodgersii* was very connected in the population graph (6 degrees; Table A4), and had high within-population genetic variance across both size cohorts (13.24), it had a low between centrality index (1) meaning it is not important to the overall graph connectivity and is likely a demographic sink. Second, our sampling design including two size cohorts, and our analysis of both neutral and putatively adaptive loci, revealed White Island's population of *C. rodgersii* may be self-recruiting more in recent years than previously. Based on the neutral loci, White Island's population of *C. rodgersii* was connected in the older/larger group but not the younger/smaller group (Fig. 7a,b). Although White Island has two individuals with a high probability of the Rangitāhua-like genotype (Table A3) this is likely not the cause of the disconnected populations as these individuals are large. Instead, it may be because the population had been largely self-recruiting in recent years (also suggested by Chapter Two). Furthermore, based on putatively adaptive loci, White Island is disconnected in the older/large group population graph (Fig. 7c) and appears different to other populations in the DAPC (Fig. 3d, 5c) indicating divergent selection at these loci, and local adaptation of the White Island population of *C. rodgersii*.

Overall, we found that population structure and connectivity differed for the two sampled size cohorts of *C. rodgersii* in north-east New Zealand; and that Rangitāhua populations have a different genomic composition to those of north-east New Zealand, but some migrants from Rangitāhua are received into southern offshore islands. Our analyses depended on the sampling of representative size cohorts within populations. In general, small individuals are difficult to find as they are able to hide more effectively within interstitial spaces (Ling et al. 2012; Byrne et al. 2020), and populations naturally have different size landscapes due to their demography and recruitment. Nonetheless, for every population of north-east New Zealand, we were able to sample two representative size cohorts, relevant to the time scales over which the ocean environment of New Zealand has changed (Shears et al. 2017; Salinger et al. 2020). Although the methods we used to detect putatively adaptive loci are robust to population genetic structure and admixture (Whitlock et al. 2015; Luu et al. 2017; Capblancq et al. 2018), without relating these loci to a known divergent selection pressure (i.e. genotype-environment-associations with a seascape genomic approach) and a mechanistic link between gene function and the selection pressure, interpretation of the patterns in putatively adaptive genetic variance should be cautious. Therefore it is important to retain these time-stamped, georeferenced genotypes and related measures as a baseline to inform further analyses and future genetic monitoring. *Centrostephanus rodgersii* has an actively changing population structure and since it has the potential to dramatically impact our biodiversity and fisheries in New Zealand, we should monitor this species and consider ways we can manage the species like has been done in Tasmania (Johnson et al. 2013).

Chapter Three Appendix: supplementary methods and results

Single nucleotide polymorphism calling and filtering

Individual	Match	Mismatch	NA
T8276	6355	30	508
T7284	6300	48	545
T7234	6284	35	574
T7290	6340	87	466

Table A1: Table of the replicate individuals that were sequenced. Match = the number of loci that were called the same, Mismatch = the number of loci that were called differently, NA = the number of loci where the call was missing in at least one of the replicates.

Detection of loci putatively under selection

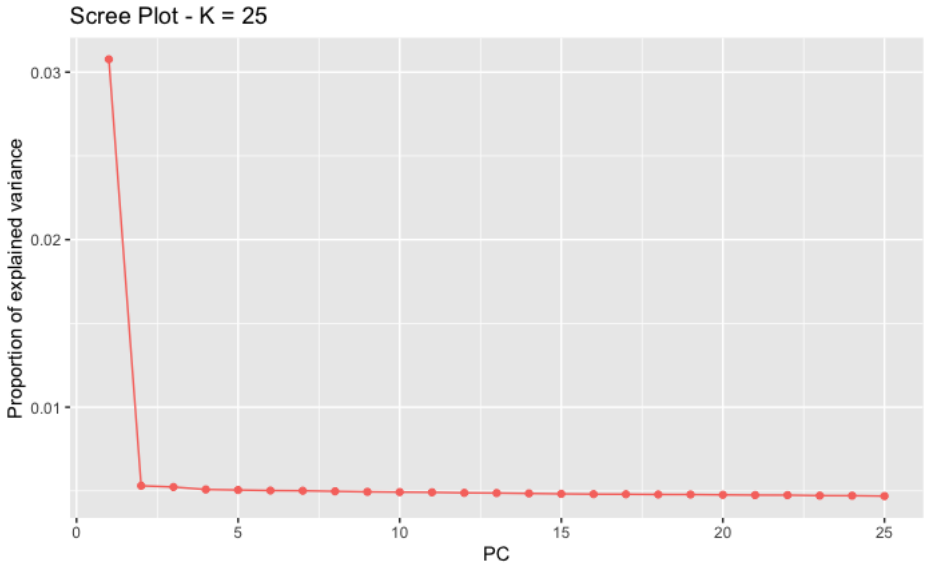

Figure A1: Scree plot used to select the value of K used in the outlier analysis of pcadapt.

Population genetic summary statistics

	Ald	Ber	Cas	Cav	Hom	LEs	Mac	May	Mer	Mok	Poo	Rao	Sai	Spi	Tak
Ber	0.0008														
Cas	0.0012	0.0002													
Cav	0.0007	-0.0005	0.0004												
Hom	0.0016	0.0004	0.0018	0.0009											
LEs	0.0305	0.0309	0.0259	0.0305	0.0310										
Mac	0.0275	0.0279	0.0238	0.0292	0.0281	0.0005									
May	0.0025	0.0013	0.0016	0.0019	0.0013	0.0269	0.0252								
Mer	0.0020	0.0011	0.0016	0.0017	0.0024	0.0257	0.0232	0.0019							
Mok	0.0015	0.0006	0.0004	0.0008	0.0008	0.0297	0.0274	0.0013	0.0017						
Poo	0.0002	0.0001	0	0.0012	0.0004	0.0286	0.0257	0.0012	0.0014	0.0007					
Rao	0.0301	0.0306	0.0266	0.0303	0.0305	0.0012	0.0008	0.0265	0.0254	0.0306	0.0288				
Sai	0.0004	0.0006	0.0011	0.0006	0.0017	0.0311	0.0289	0.0020	0.0016	0.0007	0.0003	0.0307			
Spi	0.0006	0.0002	-0.0002	0.0005	0.0004	0.0313	0.0285	0.0006	0.0010	0.0004	0	0.0311	-0.0001		
Tak	0.0025	0.0026	0.0012	0.0010	0.0031	0.0305	0.0292	0.0016	0.0033	0.0012	0.0012	0.0306	0.0017	0.0016	
Whi	0.0009	0	0.0005	0.0001	0.0010	0.0253	0.0226	0.0005	0.0008	-0.0002	0	0.0256	0.0008	0	0.0020

Table A2: Pairwise F_{ST} matrix based on the neutral SNPs. See Table A11 for full location names for abbreviated terms.

Population genetic structure

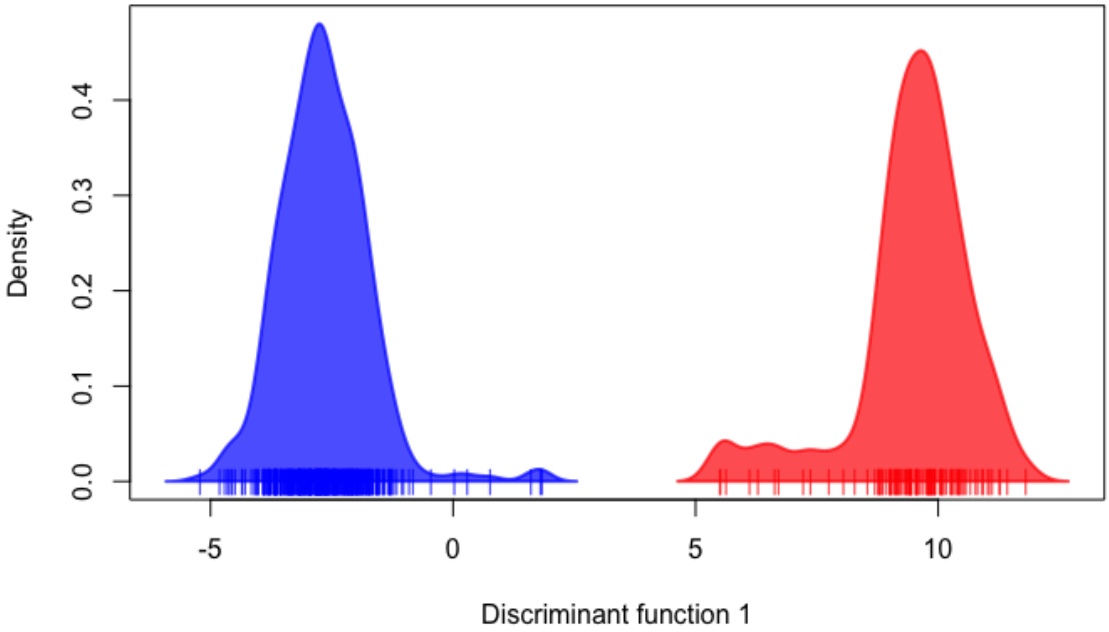

Figure A2: The DAPC with no priors based on the neutral SNPs for all New Zealand populations (including Rangitāhua). The DAPC is based on the first 111 PCs and the proportion of conserved variance was 0.356.

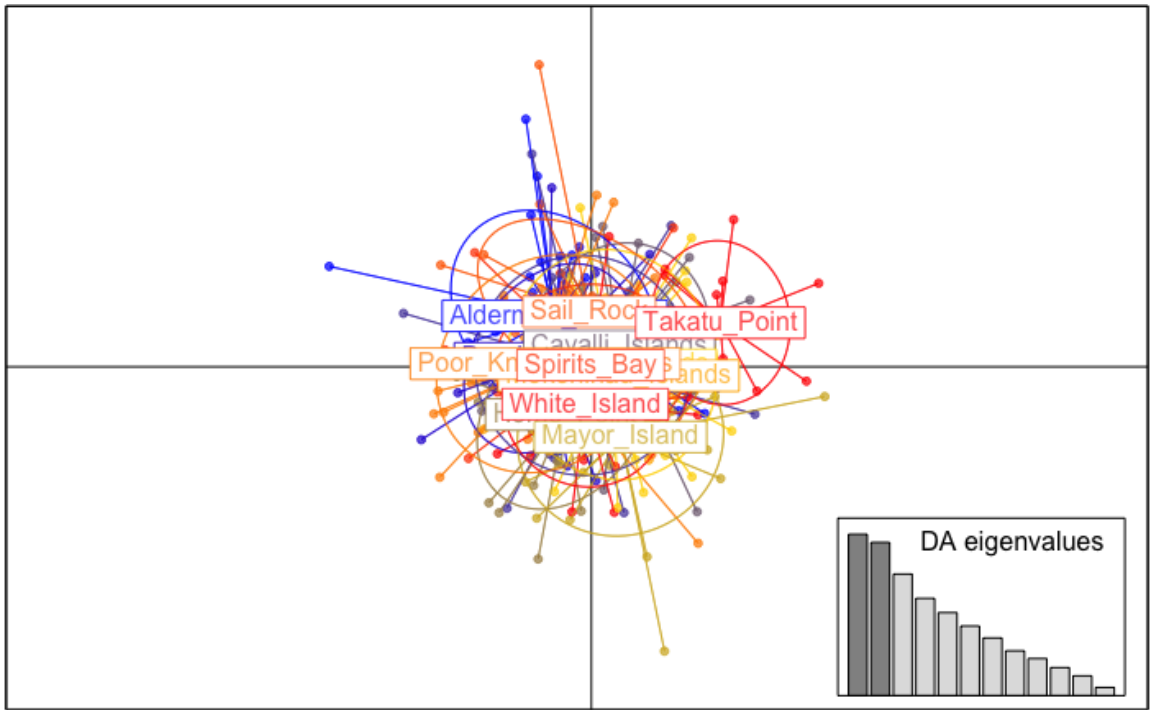

Figure A3: The DAPC with population priors based on the neutral SNPs for north-east New Zealand populations. The DAPC is based on the first 33 PCs and the proportion of conserved variance was 0.137.

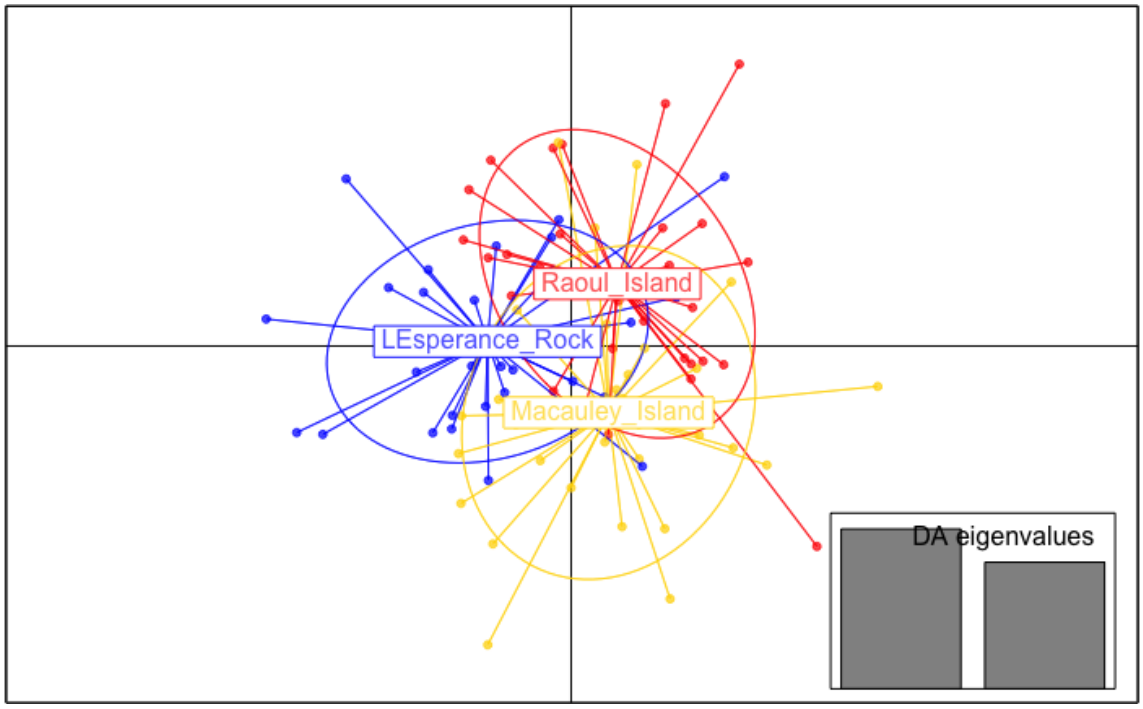

Figure A4: The DAPC with population priors based on neutral SNPs for Rangitāhua populations. The DAPC is based on the first 10 PCs and the proportion of conserved variance was 0.139.

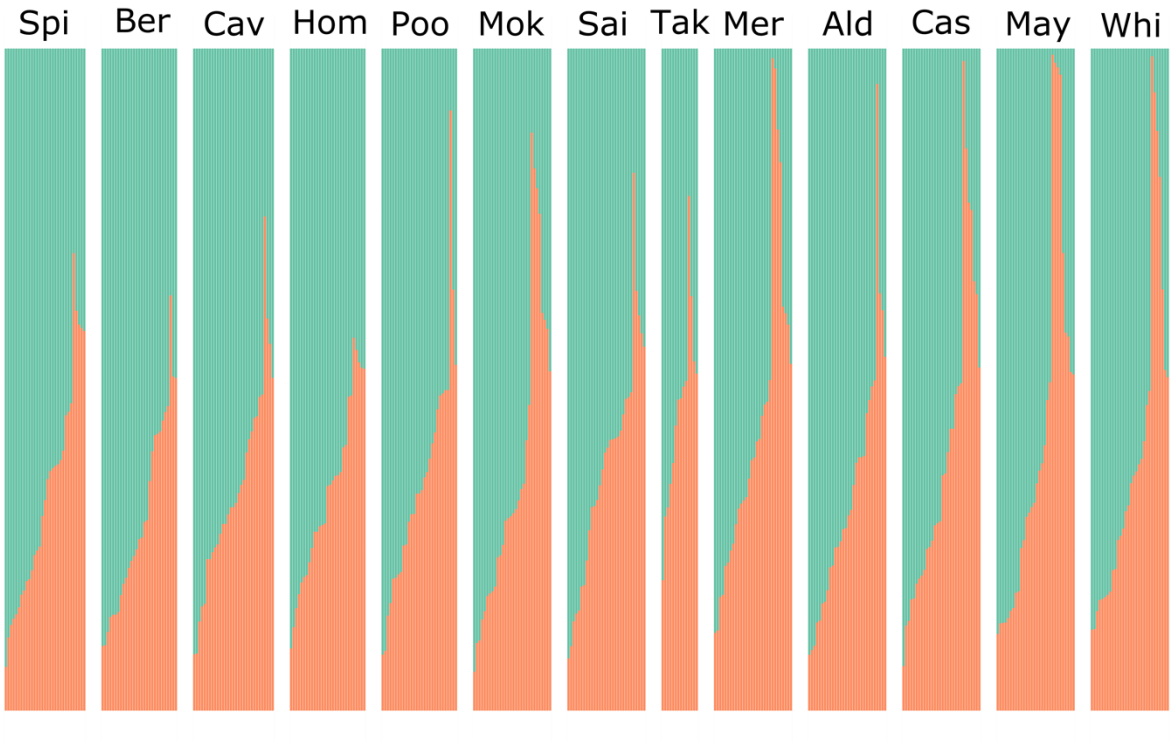

Figure A5: STRUCTURE plot based on neutral SNPs for north-east New Zealand populations, k = 2. Full population names are given in Table A11.

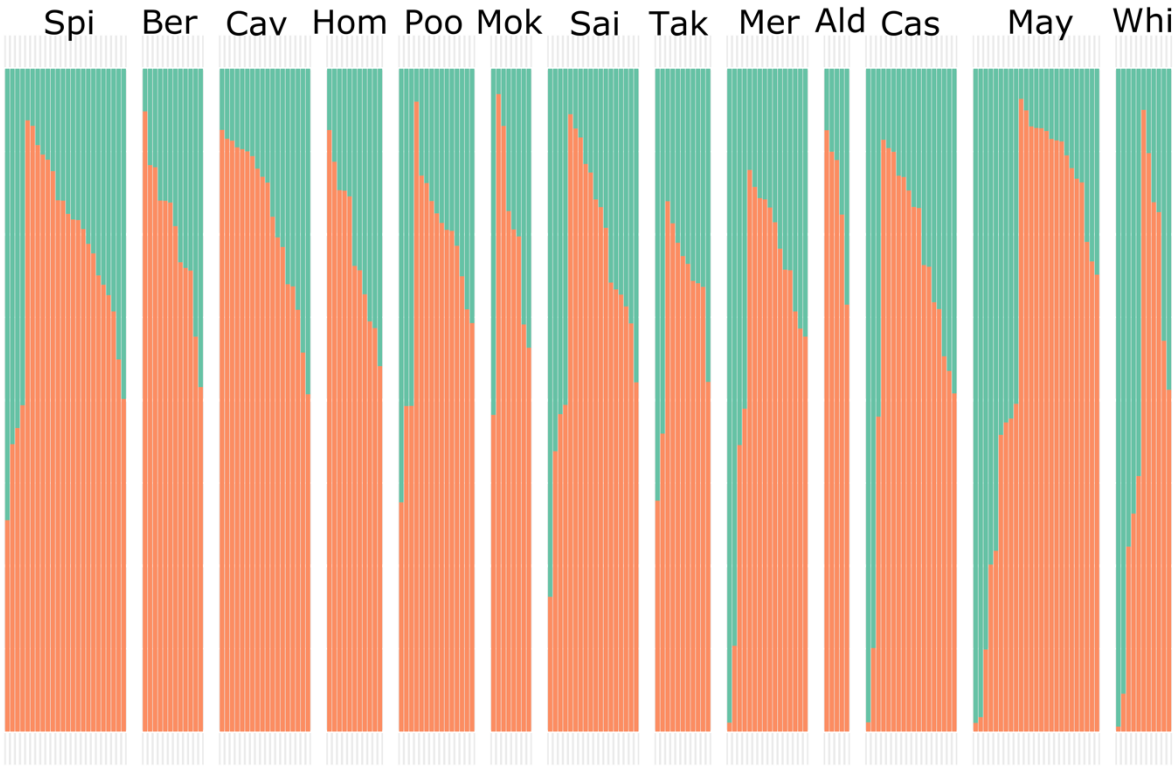

Figure A6: STRUCTURE plot based on neutral SNPs for large individuals in north-east New Zealand populations, k = 2. Full population names are given in Table A11.

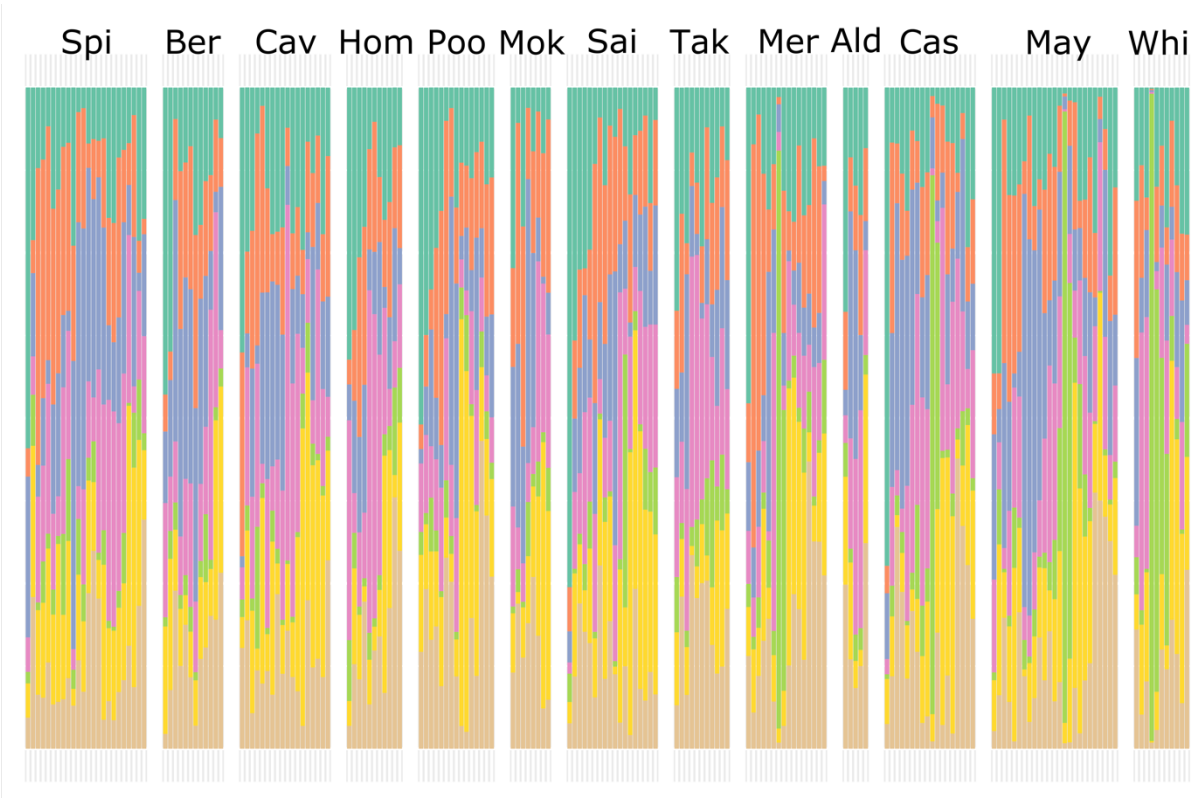

Figure A7: STRUCTURE plot based on neutral SNPs for large individuals in north-east New Zealand populations, k = 7. Full population names are given in Table A11.

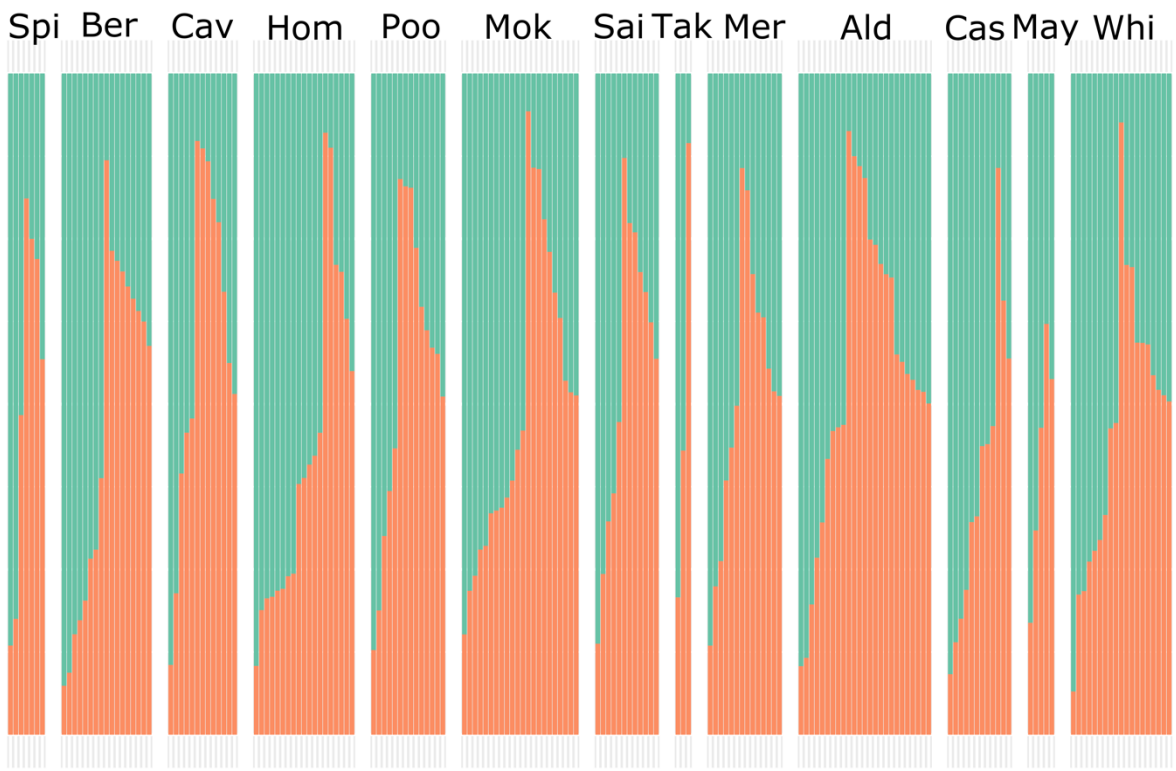

Figure A8: STRUCTURE plot based on neutral SNPs for small individuals in north-east New Zealand populations, k = 2. Full population names are given in Table A11.

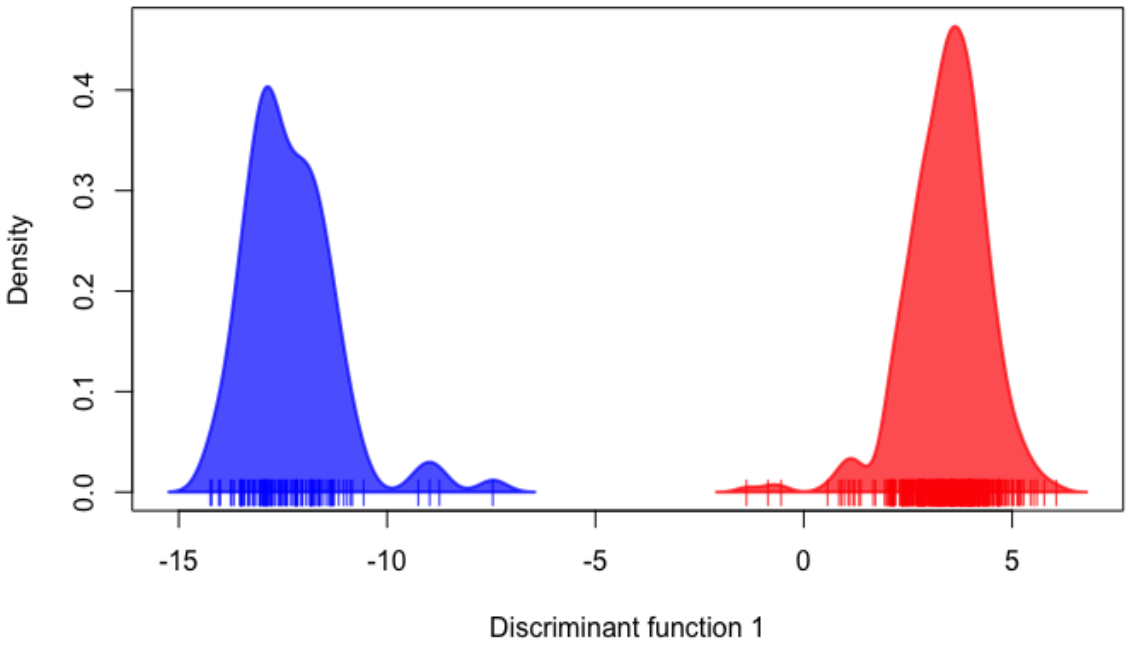

Figure A9: The DAPC with no priors based on the adaptive SNPs for all New Zealand populations (including Rangitāhua). The DAPC is based on the first 200 PCs and the proportion of conserved variance was 0.948.

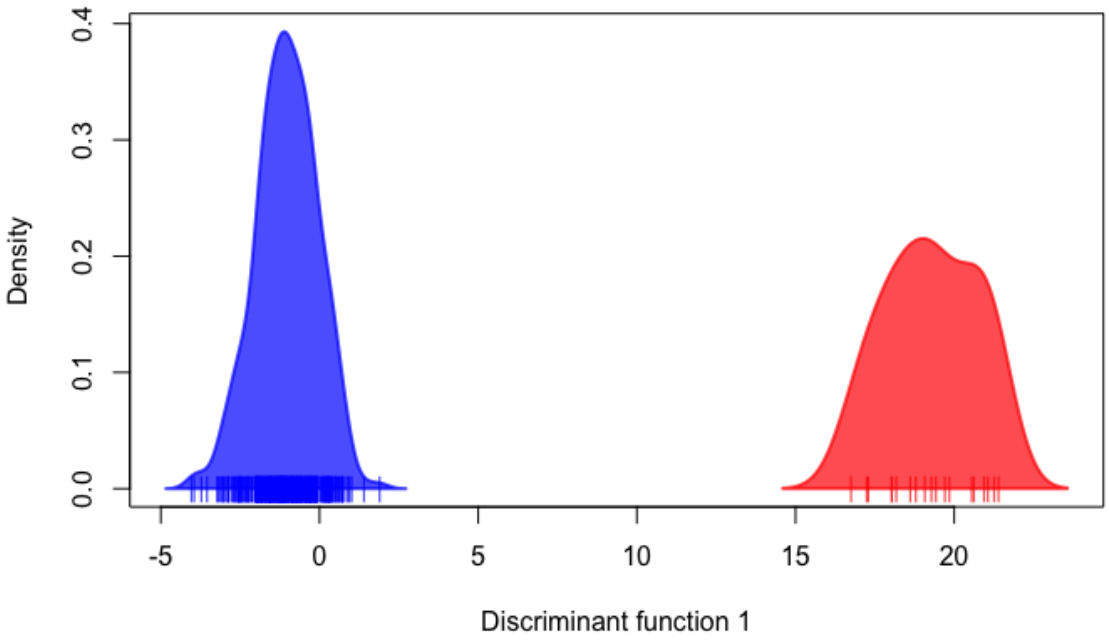

Figure A10: The DAPC with no priors based on the adaptive SNPs for north-east New Zealand populations. The DAPC is based on the first 247 PCs and the proportion of conserved variance was 0.989.

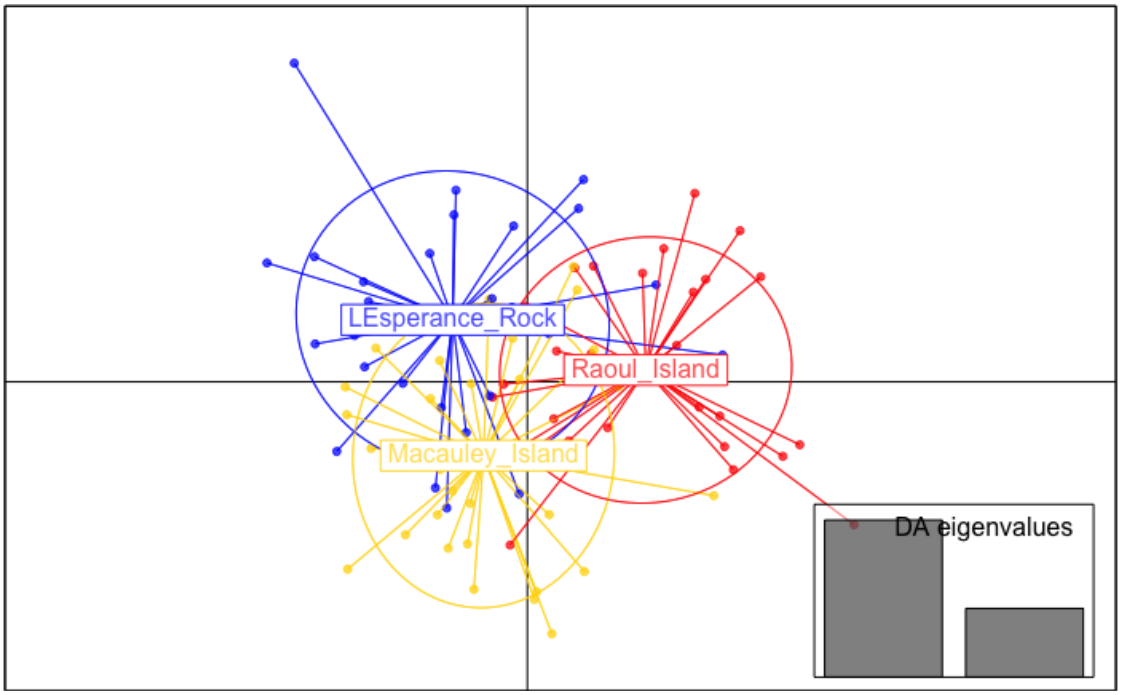

Figure A11: The DAPC with population priors based on the adaptive SNPs for Rangitāhua populations. The DAPC is based on the first 23 PCs and the proportion of conserved variance was 0.471.

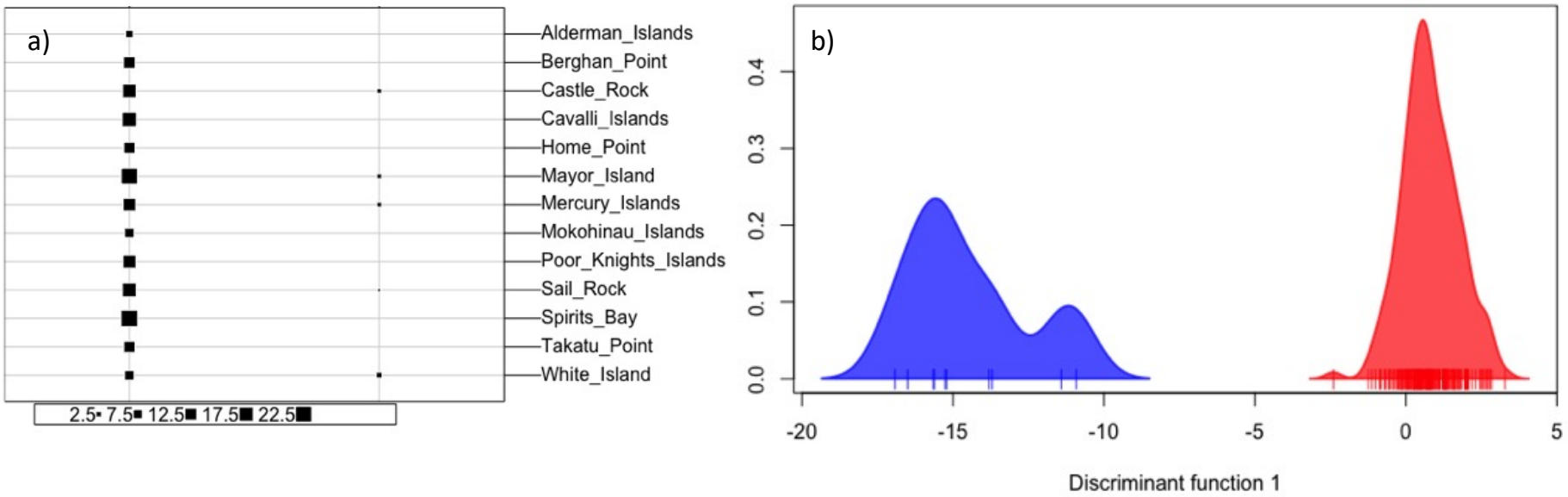

Figure A12: a) clustering based on the adaptive SNPs for large individuals in north-east New Zealand populations which is then used for the DAPC; b) the DAPC with no priors based on the adaptive SNPs for large individuals in north-east New Zealand populations. The DAPC is based on the first 75 PCs and the proportion of conserved variance was 0.777.

Individual ID	Population	Size (mm)	Mainland adaptive DAPC	All adaptive DAPC	All neutral DAPC	All neutral STRUCTURE
T7793	Alderman Islands	85	Y			0.4475
T7773	Castle Rock	67	Y		Y	**0.5209**
T7743	Castle Rock	109	Y	Y	Y	**0.6151**
T7768	Castle Rock	125	Y	Y	Y	**0.9971**
T8212	Mayor Island	69	Y	Y	Y	**0.6334**
T8210	Mayor Island	112	Y	Y	Y	**0.9988**
T8229	Mayor Island	122	Y	Y		**0.7438**
T7286	Mercury Islands	130	Y			**0.5522**
T7296	Mercury Islands	67	Y		Y	**0.5255**
T7313	Mercury Islands	90	Y			0.4779
T7306	Mercury Islands	104	Y	Y	Y	**0.9982**
T7188	Mokohinau Islands	46	Y	Y		**0.5376**
T7207	Mokohinau Islands	78	Y	Y	Y	0.4733
T8297	Poor Knight Islands	54	Y			0.4779
T7364	Sail Rock	98	Y	Y		0.4503
T7254	White Island	91	Y	Y	Y	**0.9970**
T7263	White Island	95	Y			0.3628
T7260	White Island	99	Y			0.2766
T7233	White Island	100	Y	Y	Y	**0.7269**

Table A3: Table of *Centrostephanus rodgersii* individuals sampled from north-east New Zealand that clustered with Rangitāhua populations in the DAPC analyses without priors, including their assignment to the cluster according to their test size, which DAPC analysis clustered them with the Rangitāhua populations (indicated by Y), and STRUCTURE. Mainland adaptive – one of the DAPC clusters from the analysis based on the putatively adaptive SNPs including populations sampled from north-east New Zealand only, which recovered a small number of individuals clustering separately; All adaptive – the Rangitāhua-like DAPC cluster for the analysis based on the putatively adaptive SNPs including all sampled populations; All neutral DAPC – the Rangitāhua-like DAPC cluster for the analysis based on the neutral SNPs including all sampled populations; All neutral STRUCTURE – the probability that the individual was in the Rangitāhua-like cluster from the STRUCTURE analysis based on the neutral SNPs including all sampled populations (k=2), bold over 0.5 probability. (Note: the DAPC analysis based on the neutral SNPs including populations sampled from north-east New Zealand did not recover a Rangitāhua-like cluster).

Meta-population structure and population genetic covariance

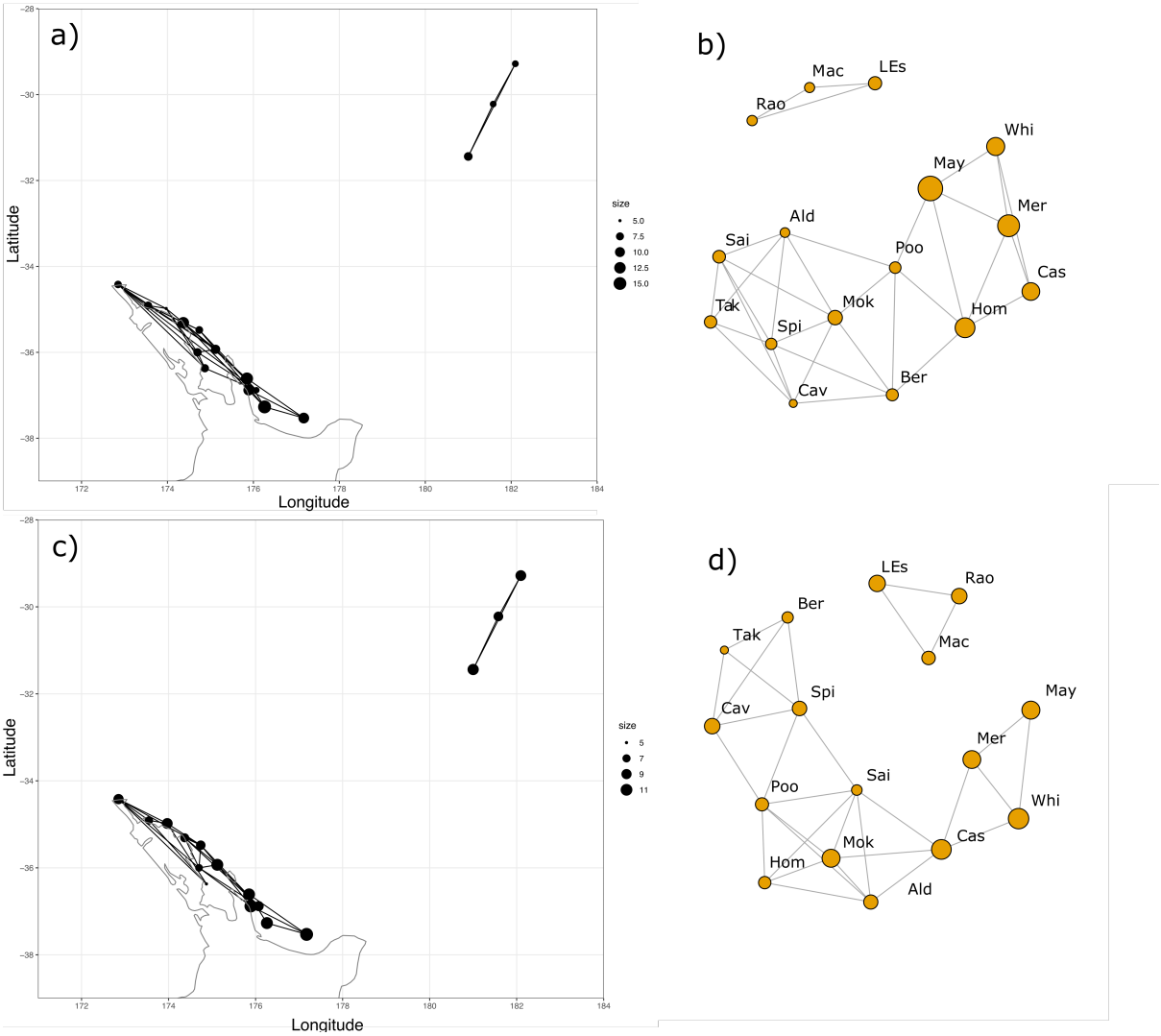

Figure A13: Population graphs for all of New Zealand (including Rangitāhua). a) Population graph based on the neutral SNPs displayed geographically; b) Population graph based on the neutral SNPs displayed aspatially; c) Population graph based on the putatively adaptive SNPs displayed geographically; d) Population graph based on the putatively adaptive SNPs displayed aspatially.

Population	deg	close	btw	str	siw	miw	modules	pop var
Alderman Islands	5	0.713	12	0.363	68.965	13.793	d	10.171
Berghan Point	4	0.624	4	0.305	52.921	13.230	a	5.762
Castle Rock	4	0.613	4	0.321	50.694	12.674	a	5.454
Cavalli Islands	5	0.680	6	0.385	64.931	12.986	b	7.156
Home Point	5	0.653	2	0.415	60.632	12.126	c	8.011
Mayor Island	5	0.643	2	0.417	60.924	12.185	b	10.073
Mercury Islands	5	0.642	0	0.422	60.111	12.022	b	13.328
Mokohinau Islands	6	0.686	3	0.483	75.351	12.558	b	5.398
Poor Knights Islands	4	0.650	4	0.308	51.966	12.991	a	7.852
Sail Rock	2	0.512	2	0.135	29.665	14.833	a	9.659
Spirits Bay	5	0.697	8	0.394	63.882	12.776	c	5.000
Takatu Point	6	0.593	1	0.598	60.356	10.059	c	13.240
White Island	4	0.644	6	0.298	53.815	13.454	d	13.110

Table A4: Measures from the population graph in Fig. 6a based on neutral loci of individuals from north-east New Zealand. Deg = number of degrees/connections to other populations; close = closeness centrality index; btw = betweenness centrality index; str = strength (sum of the weights of the links connected to a node); siw = sum of the inverse weights of the links connected to a node; miw = mean of the inverse weights of the links connected to a node; modules = each letter refers to a unique module; pop var = the within population variance.

Population	deg	close	btw	str	siw	miw	modules	pop var
Alderman Islands	3	0.687	12	0.187	48.205	16.068	a	8.607
Berghan Point	3	0.551	1	0.204	44.582	14.861	b	13.717
Castle Rock	4	0.779	15	0.259	61.807	15.452	c	7.659
Cavalli Islands	3	0.611	6	0.210	43.263	14.421	b	5.000
Home Point	4	0.670	10	0.276	59.507	14.877	b	6.578
Mayor Island	2	0.587	2	0.140	29.493	14.746	a	9.390
Mercury Islands	3	0.595	2	0.194	46.451	15.484	a	7.426
Mokohinau Islands	4	0.622	5	0.262	62.123	15.531	b	8.981
Poor Knights Islands	3	0.562	2	0.189	47.622	15.874	a	10.31
Sail Rock	3	0.656	3	0.185	48.647	16.216	c	9.117
Spirits Bay	4	0.719	16	0.243	65.921	16.48	a	10.394
Takatu Point	5	0.617	5	0.411	60.871	12.174	b	8.384
White Island	3	0.549	2	0.191	47.128	15.709	a	13.645

Table A5: Measures from the populations graph Fig. 6c based on putatively adaptive loci of individuals from north-east New Zealand. Deg = number of degrees/connections to other populations; close = closeness centrality index; btw = betweenness centrality index; str = strength (sum of the weights of the links connected to a node); siw = sum of the inverse weights of the links connected to a node; miw = mean of the inverse weights of the links connected to a node; modules = each letter refers to a unique module; pop var = the within population variance.

Population	deg	close	btw	str	siw	miw	modules	pop var
Alderman Islands	2	0.054	10	0.204	19.677	9.838	c	7.413
Berghan Point	2	0.056	18	0.243	16.581	8.290	b	9.227
Castle Rock	1	0.050	0	0.113	8.826	8.826	b	5.577
Cavalli Islands	3	0.057	14	0.541	17.461	5.820	a	11.764
Home Point	1	0.053	0	0.131	7.636	7.636	a	13.871
Mayor Island	4	0.060	38	0.795	20.236	5.059	a	13.36
Mercury Islands	2	0.057	18	0.289	14.963	7.482	c	10.937
Mokohinau Islands	2	0.054	10	0.224	17.879	8.94	b	7.195
Poor Knights Islands	1	0.051	0	0.098	10.159	10.159	c	6.866
Sail Rock	3	0.059	24	0.571	16.684	5.561	b	8.501
Spirits Bay	2	0.054	0	0.478	8.441	4.220	a	9.096
Takatu Point	3	0.056	0	0.716	12.631	4.210	a	5.000
White Island	0	0.006	0	0	0	NA	d	8.010

Table A6: Measures from the populations graph Fig. 7b based on neutral loci of small individuals from north-east New Zealand. Deg = number of degrees/connections to other populations; close = closeness centrality index; btw = betweenness centrality index; str = strength (sum of the weights of the links connected to a node); siw = sum of the inverse weights of the links connected to a node; miw = mean of the inverse weights of the links connected to a node; modules = each letter refers to a unique module; pop var = the within population variance.

Population	deg	close	btw	str	siw	miw	modules	pop var
Alderman Islands	3	0.034	27	0.513	17.589	5.863	a	7.24
Berghan Point	1	0.032	0	0.127	7.845	7.845	b	12.79
Castle Rock	2	0.033	9	0.253	15.786	7.893	b	9.221
Cavalli Islands	2	0.033	9	0.22	18.17	9.085	c	9.967
Home Point	3	0.034	28	0.374	24.259	8.086	b	6.27
Mayor Island	1	0.031	0	0.103	9.741	9.741	a	7.169
Mercury Islands	1	0.032	0	0.109	9.161	9.161	c	9.927
Mokohinau Islands	2	0.034	25	0.318	12.841	6.421	b	13.086
Poor Knights Islands	0	0.006	0	0	0	NA	d	8.027
Sail Rock	2	0.032	9	0.219	18.338	9.169	a	9.937
Spirits Bay	0	0.006	0	0	0	NA	e	10.983
Takatu Point	2	0.033	16	0.276	14.858	7.429	a	5.000
White Island	1	0.032	0	0.173	5.795	5.795	a	10.149

Table A7: Measures from the populations graph Fig. 7a based on neutral loci of large individuals from north-east New Zealand. Deg = number of degrees/connections to other populations; close = closeness centrality index; btw = betweenness centrality index; str = strength (sum of the weights of the links connected to a node); siw = sum of the inverse weights of the links connected to a node; miw = mean of the inverse weights of the links connected to a node; modules = each letter refers to a unique module; pop var = the within population variance.

Population	deg	close	btw	str	siw	miw	modules	pop var
Alderman Islands	3	0.060	19	0.35	26.761	8.92	a	7.207
Berghan Point	2	0.058	12	0.21	19.099	9.549	b	8.186
Castle Rock	2	0.057	10	0.244	16.492	8.246	b	11.767
Cavalli Islands	3	0.059	18	0.405	22.452	7.484	a	8.779
Home Point	1	0.052	0	0.112	8.955	8.955	d	10.521
Mayor Island	1	0.049	0	0.28	3.575	3.575	c	5.000
Mercury Islands	2	0.055	10	0.237	16.926	8.463	d	6.109
Mokohinau Islands	2	0.058	12	0.224	18.073	9.036	a	8.975
Poor Knights Islands	1	0.050	0	0.246	4.058	4.058	c	10.191
Sail Rock	2	0.055	6	0.389	11.433	5.716	b	9.38
Spirits Bay	3	0.060	15	0.601	16.435	5.478	a	9.971
Takatu Point	4	0.057	19	1.076	14.938	3.735	c	14.768
White Island	0	0.006	0	0	0	NA	e	8.336

Table A8: Measures from the populations graph Fig. 7d based on putatively adaptive of small individuals from north-east New Zealand. Deg = number of degrees/connections to other populations; close = closeness centrality index; btw = betweenness centrality index; str = strength (sum of the weights of the links connected to a node); siw = sum of the inverse weights of the links connected to a node; miw = mean of the inverse weights of the links connected to a node; modules = each letter refers to a unique module; pop var = the within population variance.

Population	deg	close	btw	str	siw	miw	modules	pop var
Alderman Islands	4	0.265	10	0.732	22.178	5.545	c	12.529
Berghan Point	3	0.250	14	0.416	22.203	7.401	b	9.451
Castle Rock	2	0.232	10	0.236	16.959	8.479	a	9.421
Cavalli Islands	3	0.268	16	0.394	23.364	7.788	c	8.053
Home Point	3	0.250	7	0.409	22.501	7.5	c	10.924
Mayor Island	2	0.238	0	0.218	18.391	9.195	b	12.376
Mercury Islands	2	0.214	7	0.283	14.206	7.103	a	10.252
Mokohinau Islands	3	0.254	2	0.552	16.559	5.52	b	6.205
Poor Knights Islands	4	0.278	19	0.517	31.790	7.948	b	8.83
Sail Rock	2	0.224	3	0.251	15.927	7.963	c	11.379
Spirits Bay	3	0.257	13	0.372	24.322	8.107	b	7.152
Takatu Point	3	0.233	10	0.447	20.837	6.946	a	9.014
White Island	2	0.218	7	0.273	14.812	7.406	a	5.000

Table A9: Measures from the populations graph Fig. 7c based on putatively adaptive loci of large individuals from north-east New Zealand. Deg = number of degrees/connections to other populations; close = closeness centrality index; btw = betweenness centrality index; str = strength (sum of the weights of the links connected to a node); siw = sum of the inverse weights of the links connected to a node; miw = mean of the inverse weights of the links connected to a node; modules = each letter refers to a unique module; pop var = the within population variance.

Graph 1	Graph 2	graph_modul_compar (Adjusted Rand Index)	graph_node_compar (Spearman's correlation)	graph_topo_compar (Mathew's correlation)
Small|Adp|NENZ	Large|Adp|NENZ	0.062	0.342 (0.253)	0
Small|Neu|NENZ	Large|Neu|NENZ	0.026	0.096 (0.756)	0.137
All|Neu|All	All|Adp|All	0.509	0.730 (0.001)	0.506
Large|Neu|NENZ	Large|Adp|NENZ	0.081	0.120 (0.695)	0.245
Small|Neu|NENZ	Small|Adp|NENZ	0.165	0.500 (0.082)	-0.015
All|Neu|NENZ	All|Adp|NENZ	0.229	0.426 (0.147)	-0.085

Table A10: Comparisons of populations graphs. Graph 1 and Graph 2: the population graphs which are being compared; labelled as: Size class (Small: small individuals, Large: large individuals, All: both small and large individuals)|SNPs (Neu: neutral SNPs, Adp: putatively adaptive SNPs)|Populations (All: all populations, NENZ: only north-east New Zealand populations). graph_modul_compar uses the Adjusted Rand Index to compare if the modules are similarly classified across the graphs; graph_node_compar uses Spearman's correlation coefficient to compare the degrees (chosen parameter) of the nodes/populations between the graphs using a sample size of 13, number in brackets is the p-value; graph_topo_compar uses Mathew's correlation coefficient to compare the topology/connections between the populations.

Population	Abbreviation
Alderman Islands	Ald
Berghan Point	Ber
Castle Rock	Cas
Cavalli Islands	Cav
Home Point	Hom
L'Esperance Rock	LEs
Macauley Island	Mac
Mayor Island	May
Mercury Islands	Mer
Mokohinau Islands	Mok
Raoul Island	Poo
Poor Knights Islands	Rao
Sail Rock	Sai
Spirits Bay	Spi
Takatu Point	Tak
White Island	Whi

Table A11: Abbreviations of the population names.

	Ald	Ber	Cas	Cav	Hom	May	Mer	Mok	Poo	Sai	Spi	Tak	Whi
Ald	0	0	0	0.07	0.08	0	0	0	0	0.07	0.07	0	0.07
Ber		0	0	0.07	0.08	0	0	0	0.08	0.07	0	0	0
Cas			0	0	0	0.08	0	0	0.07	0	0	0.1	0.07
Cav	13.62	13.43		0	0.08	0.08	0.08	0	0	0	0	0	0
Hom	12.9	11.95		12.59	0	0	0	0	0	0	0.08	0.1	0
May			13.16	12.58		0	0.08	0.07	0	0	0	0.1	0
Mer				12.72		12.26	0	0.08	0	0	0	0.11	0.08
Mok						13.38	12.56	0	0.08	0	0.08	0.1	0.08
Poo		12.53	13.5					13.04	0	0	0.08	0	0
Sai	14.65	15.02								0	0	0	0
Spi	14.12				12.73			13.31	12.9		0	0.09	0
Tak			10.13		10.46	9.55	9.38	10.01			10.82	0	0
Whi	13.68		13.91				13.19	13.04					0

Table A12: Edge lengths of the population graph based on the neutral loci of individuals in north-east New Zealand (Fig. 6b). Above the diagonal is the between-population covariance with zeros where the populations are not connected. Below the diagonal is the inverse of the between-population covariance which corresponds to the lengths of the edges displayed in the population graph.

	Ald	Ber	Cas	Cav	Hom	May	Mer	Mok	Poo	Sai	Spi	Tak	Whi
Ald	0	0	0	0	0.06	0	0	0	0.06	0	0.06	0	0
Ber		0	0	0.06	0	0	0	0.06	0	0	0	0.08	0
Cas			0	0.07	0	0	0.07	0	0	0.06	0.06	0	0
Cav		15.85	15.03	0	0	0	0	0	0	0	0	0.08	0
Hom	15.42				0	0	0	0.06	0	0.06	0	0.09	0
May						0	0	0	0	0	0.06	0.08	0
Mer			15.28				0	0	0.06	0	0	0	0.07
Mok		16.05			16.94			0	0	0.06	0	0.08	0
Poo	16.09						15.98		0	0	0	0	0.06
Sai			15.88		15.88			16.88		0	0	0	0
Spi	16.69		15.62			17.22					0	0	0.06
Tak		12.69		12.39	11.27	12.27		12.25				0	0
Whi							15.19		15.55		16.39		0

Table A13: Edge lengths of the population graph based on the putatively adaptive loci of individuals in north-east New Zealand (Fig. 6d). Above the diagonal is the between-population covariance with zeros where the populations are not connected. Below the diagonal is the inverse of the between-population covariance which corresponds to the lengths of the edges displayed in the population graph.

	Ald	Ber	Cas	Cav	Hom	May	Mer	Mok	Poo	Sai	Spi	Tak	Whi
Ald	0	0	0	0	0	0	0	0.18	0	0	0	0.16	0.17
Ber		0	0.13	0	0	0	0	0	0	0	0	0	0
Cas		7.84	0	0	0.13	0	0	0	0	0	0	0	0
Cav				0	0.11	0	0.11	0	0	0	0	0	0
Hom			7.94	9.01	0	0	0	0.14	0	0	0	0	0
May						0	0	0	0	0.1	0	0	0
Mer				9.16			0	0	0	0	0	0	0
Mok	5.53				7.31			0	0	0	0	0	0
Poo									0	0	0	0	0
Sai						9.74				0	0	0.12	0
Spi											0	0	0
Tak	6.26									8.6		0	0
Whi	5.8												0

Table A14: Edge lengths of the population graph based on the neutral loci of large individuals in north-east New Zealand (Fig. 7a). Above the diagonal is the between-population covariance with zeros where the populations are not connected. Below the diagonal is the inverse of the between-population covariance which corresponds to the lengths of the edges displayed in the population graph.

	Ald	Ber	Cas	Cav	Hom	May	Mer	Mok	Poo	Sai	Spi	Tak	Whi
Ald	0	0	0	0	0	0	0.11	0	0.1	0	0	0	0
Ber		0	0	0	0	0	0	0.11	0	0.13	0	0	0
Cas			0	0	0	0	0	0.11	0	0	0	0	0
Cav				0	0.13	0.19	0	0	0	0	0	0.22	0
Hom				7.64	0	0	0	0	0	0	0	0	0
May				5.35		0	0.18	0	0	0.21	0.22	0	0
Mer	9.52					5.45	0	0	0	0	0	0	0
Mok		9.05	8.83					0	0	0	0	0	0
Poo	10.16								0	0	0	0	0
Sai		7.53				4.85				0	0	0.23	0
Spi						4.6					0	0.26	0
Tak				4.48						4.31	3.84	0	0
Whi													0

Figure A15: Edge lengths of the population graph based on the neutral loci of small individuals in north-east New Zealand (Fig. 7b). Above the diagonal is the between-population covariance with zeros where the populations are not connected. Below the diagonal is the inverse of the between-population covariance which corresponds to the lengths of the edges displayed in the population graph.

	Ald	Ber	Cas	Cav	Hom	May	Mer	Mok	Poo	Sai	Spi	Tak	Whi
Ald	0	0	0	0.16	0.17	0	0	0.22	0	0	0	0.19	0
Ber		0	0.11	0	0	0	0	0.17	0	0	0.13	0	0
Cas		8.84	0	0	0	0	0	0	0	0	0	0	0.12
Cav	6.25			0	0.12	0	0	0	0.12	0	0	0	0
Hom	6.06			8.58	0	0	0	0	0	0.13	0	0	0
May						0	0	0	0.11	0	0.11	0	0
Mer							0	0	0	0	0	0.13	0.15
Mok	4.62	5.93						0	0.17	0	0	0	0
Poo				8.53		9.37		6.02	0	0	0.13	0	0
Sai					7.86					0	0	0.12	0
Spi		7.44				9.02			7.87		0	0	0
Tak	5.25						7.51			8.07		0	0
Whi			8.12				6.69						0

Table A16: Edge lengths of the population graph based on the putatively adaptive loci of large individuals in north-east New Zealand (Fig. 7c). Above the diagonal is the between-population covariance with zeros where the populations are not connected. Below the diagonal is the inverse of the between-population covariance which corresponds to the lengths of the edges displayed in the population graph.

	Ald	Ber	Cas	Cav	Hom	May	Mer	Mok	Poo	Sai	Spi	Tak	Whi
Ald	0	0.1	0	0	0	0	0	0.1	0	0	0.15	0	0
Ber	10.14	0	0.11	0	0	0	0	0	0	0	0	0	0
Cas		8.96	0	0	0	0	0	0	0	0.13	0	0	0
Cav				0	0	0	0.13	0.12	0	0	0.16	0	0
Hom					0	0	0.11	0	0	0	0	0	0
May						0	0	0	0	0	0	0.28	0
Mer				7.97	8.96		0	0	0	0	0	0	0
Mok	10.01			8.07				0	0	0	0	0	0
Poo									0	0	0	0.25	0
Sai			7.53							0	0	0.26	0
Spi	6.62			6.41							0	0.29	0
Tak						3.57			4.06	3.9	3.41	0	0
Whi													0

Table A17: Edge lengths of the population graph based on the putatively adaptive loci of small individuals in north-east New Zealand (Fig. 7d). Above the diagonal is the between-population covariance with zeros where the populations are not connected. Below the diagonal is the inverse of the between-population covariance which corresponds to the lengths of the edges displayed in the population graph.

Chapter Four: General discussion

In this thesis, I used population size structure analysis and population genomics to examine the population history and demography of the Long-spined sea urchin, *Centrostephanus rodgersii,* in New Zealand. Specifically, I used these methods to assess whether there has been a range extension or detectable change in the population demography of the species over recent decades. Using size structure analysis I found that there has been a possible range extension of *C. rodgersii* in the northern part of north-east New Zealand, and that southern locations may have more regular recruitment (Chapter Two). Based on population genomic analysis, I found that *C. rodgersii* populations of the Rangitāhua archipelago are genetically differentiated from populations of north-east New Zealand, but that there is continuing immigration of *C. rodgersii* from Rangitāhua to some offshore islands of north-east New Zealand. Focussing on north-east New Zealand, population genomic analysis of different size classes revealed that the distribution of neutral and adaptive genomic variation has changed over recent decades, likely associated with demographic changes (Chapter Three). In this general discussion, I summarise the main results of my thesis, the implications of these findings, and discuss caveats and future directions for research looking to understand the management and monitoring of range-extending species, or species undergoing demographic changes in response to local and global changes such as climate change.

Main findings

The range extension and associated impacts of the Long-spined sea urchin, *Centrostephanus rodgersii,* on the marine ecosystems and fisheries of south-east Australia are well known (Johnson et al. 2005; Ling 2008; Lisson 2018). My thesis contributes new knowledge about the population history and demography of *C. rodgersii* in New Zealand. Previous research on *C. rodgersii* showed that the species has been in New Zealand since at least 1897 (Farquhar 1897), that New Zealand's population is genetically distinct from Australia (Thomas et al. 2021), is likely self-recruiting (Pecorino et al. 2013a), and is possibly increasing in abundance at some locations (Balemi et al. 2021). In Chapter Two, I found an indication that *C. rodgersii* had undergone a range extension in the northern part of north-east New Zealand (from Spirits Bay to the Mokohinau Islands), but not across its whole New Zealand range. Additionally, I found that the populations in the southern part of the range are likely recruiting more regularly. Chapter Three indicated that there have been demographic and connectivity changes to *C. rodgersii* in north-east New Zealand. Based on the combination of these results it is likely that there was poleward range extension of *C. rodgersii* in the northern part of north-east New Zealand and that recruitment levels vary across the range.

One of the most evident signs of a change in a species' range and demography is a range extension. Range extensions in response to climate change have been recorded for marine organisms in many parts of the world, including mangrove forests on the Florida east coast (Cavanaugh et al. 2014); four tropical reef corals (*Acropora hyacinthus*, *Acropora muricata*, *Acropora solitaryensis*, *Pavona decussata*) in Japan (Yamano et al. 2011); and Adelie penguins (*Pygoscelis adeliae*) in the Ross Sea, Antarctica (Taylor et al. 1990). New Zealand has few recorded range extensions currently, however, this could be due to a lack of monitoring. For instance, recent monitoring from a citizen science project has recorded the

range extension of mahimahi (*Coryphaena hippurus*), in the North Island of New Zealand (Middleton et al. 2021). In Chapter Two, I used size structure analysis to reveal evidence of a potential past range extension of *C. rodgersii* in the northern part of the species' north-east New Zealand range. In contrast, in the southern part of *C. rodgersii*'s north-east New Zealand range, I did not detect evidence for a potential poleward range extension. The dominant current system in north-east New Zealand, the East Auckland Current (EAuC), is very variable (Stanton et al. 1997), and could drive a non-linear non-poleward range extension which would not have been picked up the methods that I used in Chapter Two. To detect such current-driven non-linear range extensions, the methods used in Chapter Two could be extended to incorporate non-linear relationships and/or colonisation/connectivity predictions among locations based on a biophysical model of larval dispersal for the species.

This thesis contributed to the understanding of the neutral population genetic structuring of *C. rodgersii* in New Zealand, by using more locations than previous studies and the use of single nucleotide polymorphisms (SNPs). Using microsatellite markers, Banks et al. (2007) studied only one population from New Zealand (and 15 from Australia) and found little genetic differentiation between Australian and New Zealand populations. The study of Thomas et. al (2021) included seven New Zealand populations (and seven from Australia) and found that Australian and New Zealand populations were substantially differentiated, but found little differentiation between Rangitāhua and north-east New Zealand. I used 16 populations from New Zealand (and none from Australia) and found that Rangitāhua is differentiated from north-east New Zealand. The differentiation between Rangitāhua and north-east New Zealand found in this study but not in Thomas et al. (2021) may have been due to the increased number of locations but could also be due to the use of SNPs rather than microsatellites. SNPs can reveal finer-scale population genetic differentiation than microsatellites (e.g. Candy et al. 2015; Zimmerman et al. 2020) and have been recommended to replace the use of microsatellites (Fischer et al. 2017). Therefore, due to the chosen markers and the number of locations, this thesis gives the most robust understanding of neutral population genetic structuring of *C. rodgersii* in New Zealand to date.

There is likely some ongoing migration of *C. rodgersii* from Rangitāhua to north-east New Zealand. The clustering of individuals based on genotype in Chapter Three showed that some individuals clustered with the Rangitāhua populations but were found in north-east New Zealand. Migration from the Rangitāhua populations is likely possible due to the Antarctic Intermediate Water that flows south along the Kermadec Ridge (where Rangitāhua is located) to north-east New Zealand (Chiswell et al. 2015). There are records of other sub-tropical species found at Rangitāhua being sighted in north-east New Zealand (Middleton et al. 2021), including the fishes the Kermadec scalyfin (*Parma kermadecensis*) and the Kermadec demoiselle (*Chrysiptera Rapanui*) (Francis et al. 1999; Liggins et al. 2021). *Centrostephanus rodgersii*'s long larval dispersal stage means that there would be enough time for an individual larva to travel from Rangitāhua to north-east New Zealand (Huggett et al. 2005). The north-east New Zealand *C. rodgersii* individuals with the "Rangitāhua-like" genotype were found in populations in the southern part of north-east New Zealand, likely because the Antarctic Intermediate Water flows from the Kermadec Ridge to central north-east New Zealand, rather than North Cape, Spirit's Bay (Chiswell et al. 2015). Based on their genotypic composition, I found that the "Rangitāhua-like" individuals in north-east New

Zealand are likely to be a mixture of direct migrants, second generation, and third-generation migrants from Rangitāhua that have interbred with the north-east New Zealand populations.

Populations across the New Zealand range of *C. rodgersii* may be undergoing different demographic changes. The more southern populations in north-east New Zealand are likely to be recruiting more regularly than the northern populations, based on their greater variance in sizes (Chapter Two). Further, I found that in the same populations, younger cohorts and older cohorts had different between-population genomic connectivity. For instance, the White Island population was connected to other north-east New Zealand populations in the younger size cohort, yet not in the older size cohort; and the Poor Knight Islands and Spirits Bay populations were disconnected from other populations in the older size cohort, but connected in the younger size cohort (Chapter Three). Furthermore, based on putatively adaptive loci, the Berghan Point and White Island populations had much higher adaptive potential than other populations such as the Cavalli Islands and Home Point populations. The analysis based on the putatively adaptive SNPs suggest signs of local adaptation. The Cavalli Islands had the lowest adaptive genetic variance indicating that they could have recently undergone a selective sweep. These patterns in adaptive genetic variance were different when the individuals were split into different demographic groups, indicating that recruiting larvae and adults had distinct genomic compositions, potentially due to adaptation. Based on these results, the meta-population of *C. rodgersii* in north-east New Zealand is likely undergoing changes, and these changes are likely to continue in the future, so ongoing monitoring is important.

My thesis confirms, along with previous studies, that self-recruitment of *C. rodgersii* is likely happening in north-east New Zealand. The environmental conditions in north-east New Zealand allow *C. rodgersii* to reproduce (Pecorino et al. 2013a). Additionally, Thomas et al. (2021) found genetic differentiation between Australia and New Zealand suggesting there is reproduction occurring in New Zealand. Chapter Two suggests that recruitment is occurring in north-east New Zealand from the variation of sizes present in populations and that recruitment is higher in southern populations. Chapter Three suggests that there is recruitment occurring in both Rangitāhua and north-east New Zealand meta-populations, due to the genetic differentiation between them. The population graphs from the two demographic groups in Chapter Three indicate that the recruitment and population connectivity may be changing for populations in north-east New Zealand. White Island is likely self-recruiting, particularly more recently, as it had a high standard deviation of sizes (Chapter Two), high within-population genetic variance, and is disconnected in the population graph of small individuals (Chapter Three). The changes happening in recruitment must be monitored as this species has the potential to dramatically impact our ecosystem and fisheries.

Using population genomics and size structure to infer population dynamics and detect range extensions

To manage the climate-driven redistribution of species Melbourne-Thomas et al. (2021) suggests we need to take actions to increase monitoring and detection, harmonize scales of management, increase jurisdictional cooperation, enhance adaptation, and support Indigenous and traditional rights. My thesis is focused on the first action, of increased monitoring and detection. I studied how the analysis of population size structure and population genomics can be used to help detect past range extensions, and changes in the population connectivity and the recruitment of populations over time. Such methods may be particularly important when there have been no previous distributional or abundance baselines defined, or when a less resource-intensive option is required rather than resource-intensive multi-year surveys.

Size structure analysis can be used to look for signals of recent range extensions which can inform how much of a priority further monitoring is. Most previous studies that used size structure, used it in combination with abundance or density (e.g. the Kellet's whelk, *Kelletia kelletii*, Zacherl et al. 2003; and the red sea urchin, *Strongylocentrotus franciscanus*, Morgan et al. 2001; in California USA), and although it is beneficial to have abundance or density information my thesis showed it is not always necessary. Furthermore, size has often been analysed using the coefficient of variance (e.g. the purple sea urchin, *Strongylocentrotus purpuratus*, in California USA; Ebert et al. 1988), however, in Chapter Two I show that separately analysing the mean and standard deviation can be more informative. The size structure analysis I used in Chapter Two detected a likely range extension in the northern north-east New Zealand range of *C. rodgersii*. A survey that only measured presence and abundance, but not size, would not have detected such a range extension until comparisons from further surveys had been done. This study showed the value of measuring size whilst doing surveys.

Using population genomics to analyse two demographic groups, in Chapter Three, allowed me to detect potential changes in the *Centrostephanus rodgersii* meta-population over time. This approach may be useful for understanding populations and demographic responses to a changing climate/environment. Although comparing surveys over multiple years would likely give more accurate data, comparing demographic groups provides a less resource-intensive and quicker way to look at changes in demography and connectivity over time. Using individuals below and above 15 years (estimated from size), I was able to see that the older cohort from the Spirits Bay and the Poor Knights Islands populations shared no genetic variance with other sampled populations. These findings are lost when the age groups are combined. I was able to show that the meta-population structure of the younger cohort differed from that of the older cohort. While the best picture of meta-population structure can be attained from the comprehensive sampling of the population, I demonstrate age cohorts can be used to infer demographic change and whether certain populations may be becoming sinks or sources within the meta-population. These findings are important for understanding responses to local ocean climate changes as well as fisheries and pest management. Furthermore, detecting the age-class specific signature of selection can be informative for how specific populations may be responding to climate change pressures.

The methods and approaches used in this thesis had some caveats. For instance, it was difficult to achieve even sample sizes of age-class cohorts across populations as smaller individuals were harder to find. In the future, also recording density and/or abundance for each sampled population would help provide further context for the analysis of population trends in *C. rodgersii*. Nonetheless, the data provide a baseline of population genomic compositions and size-structures. Resampling of these same populations over time would help to reveal changes in the relative connectivity of north-east New Zealand populations and Rangitāhua as well as changes to the genetic variance and covariance of populations.

Future research directions

For species with a pelagic dispersal phase, such as *C. rodgersii*, biologically informed larval dispersal models can be informative about population connectivity and demography. "Biophysical models" are models that combine physical factors like oceanography with the focal species traits, like pelagic larval duration and larval mortality, which are used to simulate the movement of the focal species in the ocean (Liggins et al. 2013; Treml et al. 2015). Combining biophysical models with other analyses allows the detection of more complex patterns.

Chapter Two was limited to detecting linear poleward range extensions, but a biophysical model could facilitate the detection of some complex non-linear range extensions. The EAuC is variable in both strength and position (Stanton et al. 1997), therefore, the current patterns in north-east New Zealand may not be correlated with latitude contrasting to the consistent poleward flowing East Australian Current (EAC). In the range extension of *C. rodgersii* in Tasmania, the mean age of individuals was related to the distance from the EAC, which was correlated with latitude (Ling et al. 2009c). My model only looked at latitude, and not currents, therefore if a non-linear non-poleward current-driven range extension of *C. rodgersii* was occurring in New Zealand (if currents and latitude are not correlated), Chapter Two would not have detected it. However, in the future, I could add the biophysical model output as another predictor variable with latitude to the Bayesian model in Chapter Two to detect any range extension that may be correlated with the biophysical model output and/or create a non-linear model.

Biophysical models could also be used to give further confidence and reasoning behind results found in Chapter Three, such as why populations are genetically similar and the directions of the gene flow. A study in the Western Mediterranean on the common sea urchin (*Paracentrotus lividus*) used genomics (SNPs) and a biophysical model to study the connectivity between populations (Paterno et al. 2017). The biophysical model explained why there was genetic homogeneity in the populations – because of the high gene flow. A limitation of the methods used in Chapter Three is that the edges of the population graphs were not directional, so we do not know the direction of the gene flow. A biophysical model could be used in combination with the genomics, in Chapter Three, to explain the direction of gene flow and help in detecting source and sink populations. Source populations have a net migration of individuals out of the populations, whereas sink populations have a net migration of individuals into the population. Detecting the reasons for genetic similarity and source and sink populations is important for management decisions as decreasing the size

of a source population will be more effective at decreasing the whole meta-population than decreasing the size of a sink population.

Putatively adaptive SNPs, alongside environmental variables, can be used to look for signals of adaptation. For instance, a redundancy analysis can be done with SNPs and environmental variables to look for correlations between certain SNPs and variables as was done in Xuereb et al. (2018). Xuereb et al. (2018) ran a redundancy analysis on SNPs from the giant California sea cucumber (*Parastichopus californicus*) and 11 bioclimatic variables. They found 59 candidate SNPs, some of which were correlated with mean bottom temperature, surface salinity, and bottom current velocity. A similar analysis could be done using the SNPs from Chapter Three and environmental variables, such as minimum sea surface temperature and current velocity, to find if local adaptation could be occurring with one of these variables. Knowing which environmental variables a species is responding to allows predictions about how a species may adapt to future projections of those environmental variables, such as increased ocean temperature from climate change.

Future research applications

Centrostephanus rodgersii has been very harmful to Tasmania's fisheries causing a need for management. The formation of barrens from the urchin grazing has led to the loss of about 150 taxa (Ling 2008). Additionally, there has been a financial loss from the blacklip abalone (*Haliotis rubra*) and southern rock lobster (*Jasus edwardsii*) fisheries which are supported by macroalgae beds (Johnson et al. 2005). It became necessary for Tasmania to take actions to control *C. rodgersii*'s abundance. A wide range of strategies have been trialled to control the spread of *C. rodgersii*, including culling the urchins, management and translocations of the rock lobsters, and a fishery on *C. rodgersii* (Johnson et al. 2013; Cartwright et al. 2018, 2019). Depending on the trajectory of *C. rodgersii* in New Zealand, similar measures may be required in the future.

Monitoring from 2001 in Tasmania has allowed the documentation of a 75% increase in *C. rodgersii*'s density in eastern Tasmania (excluding the southern sites where they are rarer) in 4m to 18m depths (Ling et al. 2018). Most surveys are done by divers but other methods like remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) have been used to monitor the species (Perkins et al. 2015; Sward et al. 2021). In Tasmania, as part of their monitoring, barrens are classified as incipient barrens: small patches of barrens within the macroalgae, and extensive barrens: large barrens without macroalgae (Johnson et al. 2013). Unlike in Tasmania, here in New Zealand, there is little knowledge about the abundance, spread, or impacts of *C. rodgersii*. There has been monitoring at the Poor Knights Islands and Mokohinau Islands which indicates an increase from 1999 to 2021 (Balemi et al. 2021), however other locations have not been monitored and there is no barren classification system. As discussed above the methods in this thesis can be used for monitoring purposes and this thesis provides a starting point for the state of *C. rodgersii* in New Zealand.

Unlike in Tasmania, in New Zealand, we have not yet taken actions to control *C. rodgersii* even though we know it has the potential to have similar impacts as it has in Tasmania. We know that *C. rodgersii* is increasing in abundance in at least a few locations in New Zealand

(Balemi et al. 2021), that there has likely been a recent range extension in northern north-east New Zealand (Chapter Two), and that connectivity and demography in New Zealand is changing (Chapter Three). According to the future ocean temperature predictions, *C. rodgersii* is likely to become more common (Pecorino et al. 2013c). Some locations in north-east New Zealand have a high adaptive variance and so could be genetically resilient to future events (Chapter Three). Accordingly, we may need to employ some of the management methods, such as those being used and trialled in Tasmania.

The potential to monitor and manage range-extending species

Population graphs based on population genetic data have been used to inform conservation and the management of invasive species. For instance, in North Carolina Piedmont, USA, population graphs were used to inform the conservation of songbirds (Minor et al. 2008). Minor et al. (2008) compared graph measures between a graph based on real geographic distances and simulated graphs. Measures indicated that the rate of movement through the graph was slow, and therefore, although there was still enough connectivity for dispersal and gene flow, the spread of disease would be slow. Further measures, like the node-degree distribution, indicated the resilience of the meta-population. Similar graph measures could be used for monitoring and management the *C. rodgersii* meta-population in New Zealand. Population graphs (or networks) have also been used to inform the management plans for the coralivorous crown-of-thorns starfish (*Acanthaster planci*) on the Great Barrier Reef (Hock et al. 2014; Hock et al. 2016). Hock et al. (2016) used state-based models where the edge weights represented the probability of colonization at a discrete time step. The goal was to prevent the spread of the pest species to ecologically or economically important locations. They applied the state-based model to the crown-of-thorns starfish dispersal network from Hock et al. (2014) and ran simulations to protect the most important reefs by intervening with culling at source patches. This model informed which populations would be more effective to cull. We could use a similar model to simulate what the intervention of *C. rodgersii* could look like.

Particular populations (or patches) within population graphs (or networks) are generally either important for the within-module connectivity or the between-module connectivity (Fletcher et al. 2013). Further investigating the importance of the modules identified in Chapter Three would provide information that could be used in management (Peterman et al. 2016). If we want a sustained meta-population, for example, in conservation or a fishery, then it would be important to make sure the populations that account for between-module connectivity are protected or sustainably fished. However, if we wanted to decrease that meta-population, for example, through culling to control a species, then it may be important to target the populations that are important to between-module connectivity. Alternatively, if we want to sustain a module or patch, for example for a fisheries stock, we would be interested in those populations that account for within-module connectivity.

Barrens are very difficult to reverse once they are formed (Johnson et al. 2013; Filbee-Dexter et al. 2014); therefore, it is important that we start preventative actions on *C. rodgersii* in New Zealand before they become an extensive issue. Between kelp forest and urchin barrens there is a discontinuous phase shift meaning that to reverse an urchin barren back into a kelp forest, the density of urchins needs to be much less than the amount that

originally leads to the formation of the urchin barren (Filbee-Dexter et al. 2014; Ling et al. 2015). For *C. rodgersii*, in Sydney, Australia, only a third of the density found in barrens was needed to maintain the barrens (Hill et al. 2003). In Tasmania, they have had far more success with preventing incipient barrens from becoming extensive barrens, and preventing incipient barrens from forming, than restoring extensive barrens (Johnson et al. 2013). The restoration of extensive barrens is likely to be very expensive and take at least 30 years (Johnson et al. 2013). Due to what has happened in Tasmania and the findings of this thesis, we have an opportunity to start prevention while it is less cost-intensive and before we witness the same level of impacts seen in Tasmania here in New Zealand.

Options for active management of *Centrostephanus rodgersii* in New Zealand

Control of *Centrostephanus rodgersii* using culling

Culling of *C. rodgersii* can be done either to prevent urchin barrens or to restore urchin barrens to macroalgae beds, although the prevention is far more efficient. Systematic culling by divers has been effective at reducing the density of urchins in incipient barrens on the east coast of Tasmania (Tracey 2014). To reduce the cost of culling there have been trials to get the abalone divers to cull urchins whilst fishing abalone. Although, many urchins were culled there was no detected difference in barrens between the trial and control reefs (Sanderson et al. 2016). The lack of detected difference is likely because divers only covered small areas. Although abalone divers culling urchins cannot be the only solution, it could still be a part of the solution. In the deeper areas (over 20m) it is harder for divers to cull urchins, due to dive times, and therefore the cost (Tracey 2014). As an alternative to divers, there is a trial using Autonomous Underwater Vehicles (AUVs) to cull the urchins by punching a hole in them (Cartwright et al. 2018). Culling has a cost and therefore we must prioritise certain locations, using methods like identifying source populations as discussed above.

We should be starting to think about whether culling is an option we want to pursue in New Zealand. Currently, there is some culling occurring on kina (*Evechinus chloroticus*), however, it is important that iwi are involved in this process as kina are a taonga species and this could extend to *C. rodgersii*. Due to both the cost and cultural considerations, we may only want very limited culling, and the research done in this thesis could help inform decisions of where that culling should occur. Chapter Two suggests that locations with a smaller coefficient of variance (or large standard deviations and small means) are of greater concern because either they have recently arrived, so have the potential to cause unknown harm, and/or they are recruiting regularly so the population is likely to grow more rapidly. The populations in north-east New Zealand that had the highest coefficients of variance were White Island, Castle Rock, the Alderman Islands, the Mokohinau Islands, and Sail Rock. In Chapter Three, populations with high connectivity and low within-population variance could be source populations and therefore more efficient to cull, such as the Mokohinau Islands, Spirits Bay, Castle Rock, and Berghan Point. The Mokohinau Islands and Castle Rock are likely good populations to cull according to both Chapters Two and Three. To further indicate which populations would be best for culling, we could simulate what would happen to the meta-population in a culling event, as discussed in the future directions.

Control of *Centrostephanus rodgersii* using natural predators

An alternative to humans culling *C. rodgersii* is controlling their abundance through natural predators. Large rock lobsters (*Jasus edwardsii*) can eat *C. rodgersii* and therefore a healthy population of rock lobsters can prevent the formation of barrens. Unfortunately, the overfished population of rock lobsters in Tasmania, has reduced the resilience of kelp beds to the formation of barrens (Ling et al. 2009b). In Tasmania, rebuilding the large rock lobster population (individuals over 140mm) on incipient barrens was successful at recovering algal cover and preventing barrens, but this was not successful on extensive barrens (Johnson et al. 2013). Robinson et al. (2020) proposed eight management strategies for *C. rodgersii* which included, a rock lobster fisheries cap that limits the annual catch on the east coast of Tasmania, maximum legal size limits on rock lobsters, and translocations of rock lobsters. A set of stakeholders assessed the cost-effectiveness of the proposed management strategies and found that, out of the options not including urchin removal, translocation strategies were considered the most cost-effective and closing the rock lobster fishery was considered the least cost effective. However, overall, the most cost-effective strategies were the ones that included both urchin removal (though either culling or fishery harvesting) and a rock lobster fisheries cap (Robinson et al. 2020).

In north-east New Zealand, we have the same species of rock lobster as Tasmania (*Jasus edwardsii*, in New Zealand called crayfish), and like Tasmania, our population is overfished. Stocks were rebuilt since the 1990s when rock lobsters were overexploited (Breen et al. 2016), but currently, the Bay of Plenty rock lobster stock is overfished (Webber et al. 2018) and in the Hauraki Gulf, rock lobsters are functionally extinct (Hauraki Gulf Forum 2020). This means that our ecosystems in north-east New Zealand are vulnerable to the formation of *C. rodgersii* urchin barrens. Therefore it is essential both for the ecosystem and for resilience to barrens that we manage and increase our rock lobster populations.

Control of *Centrostephanus rodgersii* using fisheries

Urchin roe is a delicacy throughout countries in Asia, the Mediterranean, the Caribbean, and South America (Rahman et al. 2014) and there is a market for *Centrostephanus rodgersii* roe in China, Hong Kong, Japan, and Singapore. In Tasmania, Victoria, and New South Wales fisheries on *C. rodgersii* have been formed to help with the control of their numbers. Urchins found within barrens are not suitable for harvest as the urchins are smaller with slower growth rates and little roe (Ling et al. 2009a). Therefore, it is in the best interests of *C. rodgersii* fisheries to prevent barrens. However, although urchin barrens are not ideal for fisheries, a Norway based company called Urchinomics (Urchinomics 2020) removes the ‘starved’ urchins from barrens and then feeds the urchins in land-based facilities so they are suitable for consumption. In Tasmania, the fishery of *C. rodgersii* helps keep the *C. rodgersii* populations under control thereby preventing barrens from forming. Additionally it has created an alternative commercial fishery to the abalone and rock lobster fisheries, that have been impacted by the *C. rodgersii* barrens.

There is likely a viable *C. rodgersii* fishery in New Zealand that is not currently being utilised. Although *C. rodgersii* is included in the fisheries stock (under the broad description of ‘kina’) it is not actively being sold and the current fishery in New Zealand is based primarily on *E.*

chloroticus. The primary method of collection for the commercial fishery of *E. chloroticus* is hand-gathering while free diving, though there have been some small dredge fisheries in Nelson/Marlborough and the Hauraki Gulf (Miller et al. 2011). In 2019, the overseas market for *E. chloroticus* urchin roe was mainly Australia with a small proportion to Canada (Seafood New Zealand 2019). The 2019 review on the sustainability of the north-east New Zealand stocks determined whether there should be an increase to the stock, but, the decision was made to maintain the stock amount (Fisheries New Zealand 2019). From this thesis and the existing literature, it is likely the population of *C. rodgersii* in New Zealand is robust enough to start a fishery. Chapter Three and Thomas et al. (2021) suggest that *C. rodgersii* is self-sustaining in north-east New Zealand. To have an effectively managed fishery we need to differentiate between *E. chloroticus* and *C. rodgersii* in the fishing stocks, particularly if the fishery is to be used as a method to control the population and prevent barrens. Fisheries will be most viable in areas where barrens have not formed, therefore it will be important to control barren formation.

The modules of the population graph described in Chapter Three could be used to inform the management of a *C. rodgersii* fishery in New Zealand. However, since the modules were not geographically contiguous (also seen in Fletcher et al. 2013; Peterman et al. 2016), this may be impractical. Instead, modularity analysis could be used to simulate the impact of fishing on the meta-population as was done in Southern California on abalone and sea urchins (Peña et al. 2017). Modularity analysis on matrices based on yearly larval connectivity found that the urchins had weak spatial structure and changed from year to year, similar to how the younger and older groups of *C. rodgersii* in Chapter Three had different modules. Peña et al. (2017)'s simulations indicated that the urchins would have an abrupt collapse under high fishing pressure, whilst the abalone would have a more stepwise path to extinction. A similar analysis using the modules in Chapter Three could indicate how resilient a fishery on *C. rodgersii* in New Zealand could be.

Conclusions

My thesis highlights several actions that should be taken in north-east New Zealand to prevent devastation of our ecosystems and fisheries as was seen in Tasmania (Johnson et al. 2005; Ling 2008; Lisson 2018). Firstly we need to start a monitoring program that will detect range shifts and increases in abundance. If genomic and size data is collected, it can be compared to the data in this thesis to detect changes in recruitment and connectivity. Secondly, we need to look into prevention measures to stop the formation of extensive barrens, such as increasing the crayfish stocks and culling. The restoration methods discussed do not work well on extensive barrens (Johnson et al. 2013) therefore we must act now. Thirdly, we should create a fisheries stock system for *C. rodgersii* separate from *E. chloroticus* so if a fishery is started it can be controlled appropriately. Genomics from Chapter Three could be used to inform this stock system.

Bibliography

Adamack AT, Gruber B 2014. PopGenReport: Simplifying basic population genetic analyses in R. Methods in Ecology and Evolution 5: 384-387.

Agatsuma Y, Hoshikawa H 2007. Northward extension of geographical range of the sea urchin *Hemicentrotus pulcherrimus* in Hokkaido, Japan. Journal of Shellfish Research 26: 629–635.

Andrew NL 1994. Survival of kelp adjacent to areas grazed by sea urchins in New South Wales, Australia. Australian Journal of Ecology 19: 466-472.

Assis J, Lucas AV, Barbara I, Serrao EA 2016. Future climate change is predicted to shift long-term persistence zones in the cold-temperate kelp *Laminaria hyperborea*. Marine Environmental Research 113: 174-82.

Balemi C, Shears N, Taylor R 2021. Subtropical sea urchin takes hold in northern New Zealand marine reserve. New Zealand Marine Sciences Society Conference. Tauranga.

Banks S, Piggott L, Williamson J, Bové U, Holbrook N, Beheregaray L 2007. Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. Ecology 88: 3055-3064.

Bates AE, Pecl GT, Frusher Set al. 2014. Defining and observing stages of climate-mediated range shifts in marine systems. Global Environmental Change 26: 27-38.

Black R, Johnson MS, Prince J, Brearley A, Bond T 2011. Evidence of large, local variations in recruitment and mortality in the small giant clam, *Tridacna maxima*, at Ningaloo Marine Park, Western Australia. Marine and Freshwater Research 62. 1318–1326

Botsford LW, Smith BD, Quinn JF 1994. Bimodality in size distributions: The red sea urchin *Strongylocentrotus franciscanus* as an example. Ecological Applications 4: 42-50.

Breen PA, Branson AR, Bentley Net al. 2016. Stakeholder management of the New Zealand red rock lobster (*Jasus edwardsii*) fishery. Fisheries Research 183: 530-538.

Bulleri F, Eriksson BK, Queiros A, Airoldi L, Arenas F, Arvanitidis C, Bouma T J, Crowe T P, Davoult D, Guizien K, et al. 2018. Harnessing positive species interactions as a tool against climate-driven loss of coastal biodiversity. PLoS Biology 16: 1-19.

Byrne M, Andrew N 2013. *Centrostephanus rodgersii*. In: Lawrence JM ed. Sea Urchins: Biology and Ecology. Elsevier. p. 243-256.

Byrne M, Andrew NL 2020. *Centrostephanus rodgersii* and *Centrostephanus tenuispinus*. In: Lawrence JM ed. Sea Urchins: Biology and Ecology. Elsevier. p. 379-396.

Candy JR, Campbell NR, Grinnell MH, Beacham TD, Larson WA, Narum SR 2015. Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), an anadromous Pacific smelt. Molecular Ecology Resources 15: 1421-1434.

Capblancq T, Luu K, Blum MGB, Bazin E 2018. Evaluation of redundancy analysis to identify signatures of local adaptation. Molecular Ecology Resources 18: 1223-1233.

Cartwright I, Richardson K, Dutton I, Waters V 2018. Proceedings of the 2018 *Centrostephanus* forum. Tasmania: Department of Primary Industries, Parks, Water and Environment.

Cartwright I, Richardson K, Dutton I, Waters V 2019. Proceedings of the 2019 *Centrostephanus* forum #2. Tasmania: Department of Primary Industries, Parks, Water and Environment.

Cavanaugh KC, Kellner JR, Forde AJet al. 2014. Poleward expansion of mangroves is a threshold response to decreased frequency of extreme cold events. Proceedings of the National Academy of Sciences of the United States of America 111: 723-727.
Chhatre VE, Emerson KJ 2017. StrAuto: automation and parallelization of STRUCTURE analysis. BMC Bioinformatics 18: 1-5.
Chiswell SM, Bostock HC, Sutton PJH, Williams MJM 2015. Physical oceanography of the deep seas around New Zealand: A review. New Zealand Journal of Marine and Freshwater Research 49: 286-317.
Coleman MA, Minne AJP, Vranken S, Wernberg T 2020. Genetic tropicalisation following a marine heatwave. Scientific Reports 10.
Cross TB, Schwartz MK, Naugle DE, Fedy BC, Row JR, Oyler-McCance SJ 2018. The genetic network of greater sage-grouse: Range-wide identification of keystone hubs of connectivity. Ecology and Evolution 8: 5394-5412.
Dalongeville A, Andrello M, Mouillot Det al. 2018a. Geographic isolation and larval dispersal shape seascape genetic patterns differently according to spatial scale. Evolutionary Applications 11: 1437-1447.
Dalongeville A, Benestan L, Mouillot D, Lobreaux S, Manel S 2018b. Combining six genome scan methods to detect candidate genes to salinity in the Mediterranean striped red mullet (*Mullus surmuletus*). BMC Genomics 19.
Deutsch C, Ferrel A, Seibel B, Pörtner H-O, Huey RB 2015. Climate change tightens a metabolic constraint on marine habitats. Science 348: 1132-1135.
Doney SC, Fabry VJ, Feely RA, Kleypas JA 2009. Ocean acidification: The other CO2 problem. Annual Review of Marine Science 1: 169-192.
Doney SC, Ruckelshaus M, Duffy JEet al. 2012. Climate change impacts on marine ecosystems. Annual Review of Marine Science 4: 11-37.
Doo SS, Dworjanyn SA, Foo SA, Soars NA, Byrne M 2012. Impacts of ocean acidification on development of the meroplanktonic larval stage of the sea urchin *Centrostephanus rodgersii*. ICES Journal of Marine Science 69: 460-464.
Dulvy NK, Sadovy Y, Reynolds JD 2003. Extinction vulnerability in marine populations. Fish and Fisheries 4: 25-64.
Dworjanyn SA, Byrne M 2018. Impacts of ocean acidification on sea urchin growth across the juvenile to mature adult life-stage transition is mitigated by warming. Proceedings of the Royal Society B: Biological Sciences 285.
Dyer RJ 2012. Gstudio: Tools Related to the Spatial Analysis of Genetic Marker Data. Version 1.5.3.
Dyer RJ 2015. Population graphs and landscape genetics. Annual Review of Ecology, Evolution, and Systematics 46: 327-342.
Dyer RJ 2017. Popgraph: This is an R package that constructs and manipulates population graphs. Version 1.5.1.
Dyer RJ, Nason JD 2004. Population graphs: The graph theoretic shape of genetic structure. Molecular Ecology 13: 1713-1727.
Earl DA, vonHoldt BM 2011. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4: 359-361.
Ebert TA 2010. Demographic patterns of the purple sea urchin *Strongylocentrotus purpuratus* along a latitudinal gradient, 1985–1987. Marine Ecology Progress Series 406: 105-120.

Ebert TA, Dixon JD, Schroeter SCet al. 1999. Growth and mortality of red sea urchins *Strongylocentrotus franciscanus* across a latitudinal gradient. Marine Ecology Progress Series 190: 189-209.
Ebert TA, Russell MP 1988. Latitudinal variation in size structure of the west coast purple sea urchin: A correlation with headlands. Limnology and Oceanography 33: 286-294.
Edgar GJ, Stuart-Smith RD, Thomson RJ, Freeman DJ 2017. Consistent multi-level trophic effects of marine reserve protection across northern New Zealand. PLoS One 12.
Excoffier L 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: Lessons from the infinite-island model. Molecular Ecology 13: 853-864.
Farquhar H 1897. A contribution to the history of New Zealand echinoderms. Zoological Journal of the Linnean Society 26: 186–198.
Fell HB 1949. The occurrence of Australian echinoids in New Zealand waters. Records of the Auckland Institute and Museum 3: 343-346.
Feng W, Nakabayashi N, Narita K, Inomata E, Aoki MN, Agatsuma Y 2019. Reproduction and population structure of the sea urchin *Heliocidaris crassispina* in its newly extended range: The Oga Peninsula in the Sea of Japan, northeastern Japan. PLoS One 14.
Filbee-Dexter K, Scheibling RE 2014. Sea urchin barrens as alternative stable states of collapsed kelp ecosystems. Marine Ecology Progress Series 495: 1-25.
Fischer MC, Rellstab C, Leuzinger M, Roumet M, Gugerli F, Shimizu KK, Holderegger R, Widmer A 2017. Estimating genomic diversity and population differentiation - an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. BMC Genomics 18.
Fisheries New Zealand 2019. Review of sustainability measures for kina (SUR 1A, SUR 1B) for 2019/20. Wellington.
Fletcher RJ, Jr., Revell A, Reichert BE, Kitchens WM, Dixon JD, Austin JD 2013. Network modularity reveals critical scales for connectivity in ecology and evolution. Nature Communications 4.
Foo SA, Byrne M 2017. Marine gametes in a changing ocean: Impacts of climate change stressors on fecundity and the egg. Marine Environmental Research 128: 12-24.
Foo SA, Dworjanyn SA, Poore AG, Byrne M 2012. Adaptive capacity of the habitat modifying sea urchin *Centrostephanus rodgersii* to ocean warming and ocean acidification: Performance of early embryos. PLoS One 7.
Francis MP, Worthington CJ, Saul P, Clements KD 1999. New and rare tropical and subtropical fishes from northern New Zealand. New Zealand Journal of Marine and Freshwater Research 33: 571-586.
Freiwald J, Wisniewski CJ, Abbott D 2016. Northward range extension of the crowned sea urchin (*Centrostephanus coronatus*) to Monterey Bay, California. California Fish and Game 102: 37-40.
Gilarranz LJ 2020. Generic emergence of modularity in spatial networks. Scientific Reports 10.
Goudet J, Jombart T, Kamvar ZN, Archer E, Hardy O 2020. hierfstat: estimation and tests of hierarchical F-statistics. Version 0.5-7.
Gruber B, Adamack AT 2015. Landgenreport: A new R function to simplify landscape genetic analysis using resistance surface layers. Molecular Ecology Resources 15: 1172-1178.
Gruber B, Unmack PJ, Berry OF, Georges A 2018. DartR: An R package to facilitate analysis of SNP datagenerated from reduced representation genome sequencing. Molecular Ecology Resources 18: 691–699.

Gurevitch J, Fox GA, Fowler NL, Graham CH 2016. Landscape demography: Population change and its drivers across spatial scales. The Quarterly Review of Biology 91: 459-485.

Hauraki Gulf Forum 2020. State of our gulf 2020. Auckland.

Hidas EZ, Ayre DJ, Minchinton TE 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.

Hill NA, Blount C, Poore AGB, Worthington D, Steinberg PD 2003. Grazing effects of the sea urchin *Centrostephanus rodgersii* in two contrasting rocky reef habitats: Effects of urchin density and its implications for the fishery. Marine and Freshwater Research 54. 691–700.

Hillebrand H, Brey T, Gutt Jet al. 2018. Climate change: Warming impacts on marine biodiversity. In: Salomon M, Markus T eds. Handbook on Marine Environment Protection. Springer, Cham. p. 353-373.

Hock K, Wolff NH, Beeden Ret al. 2016. Controlling range expansion in habitat networks by adaptively targeting source populations. Conservation Biology 30: 856-866.

Hock K, Wolff NH, Condie SA, Anthony KRN, Mumby PJ, Paynter Q 2014. Connectivity networks reveal the risks of crown-of-thorns starfish outbreaks on the Great Barrier Reef. Journal of Applied Ecology 51: 1188-1196.

Huggett MJ, King CK, Williamson JE, Steinberg PD 2005. Larval development and metamorphosis of the Australian diadematid sea urchin *Centrostephanus rodgersii*. Invertebrate Reproduction & Development 47: 197-204.

Jakobsson M 2009. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Version 1.1.2. Rosenberg lab at Stanford University.

Johnson CR, Banks SC, Barrett NS, Cazassus F, Dunstan PK, Edgar GJ, Frusher SD, Gardner C, Haddon M, Helidoniotis F, et al. 2011. Climate change cascades: Shifts in oceanography, species' ranges and subtidal marine community dynamics in eastern Tasmania. Journal of Experimental Marine Biology and Ecology 400: 17-32.

Johnson CR, Ling SD, Ross J, Shepherd S, Miller K 2005. Establishment of the Long-spined sea urchin (*Centrostephanus rodgersii*) in Tasmania: First assessment of potential threats to fisheries. Tasmania: Tasmanian Aquaculture and Fisheries Institute.

Johnson CR, Ling SD, Sanderson JC, Dominguez JGS, Flukes E, Frusher S, Gardner C, Hartmann K, Jarman S, Little R, et al. 2013. Rebuilding ecosystem resilience: Assessment of management options to minimise formation of ‘barrens’ habitat by the long- spined sea urchin (*Centrostephanus rodgersii*) in Tasmania. Tasmania: Institute for Marine and Antartic Studies.

Johnson CR, Mann KH 1982. Adaptations of *Strongylocentrotus droebachiensis* for survival on barren grounds in Nova Scotia. In: Lawrence JM ed. International Echinoderm Conference. Tampa Bay. p. 277-283.

Jombart T 2008. Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403-1405.

Jombart T, Devillard Sb, Balloux Fo 2010. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. BMC Genetics 11.

Jurgens LJ, Gaylord B 2018. Physical effects of habitat-forming species override latitudinal trends in temperature. Ecology Letters 21: 190-196.

Kamvar ZN, Brooks JC, Grunwald NJ 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Frontiers in Genetics 6.
Kamvar ZN, Tabima JF, Grünwald NJ 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2.
Keeling RE, Kortzinger A, Gruber N 2010. Ocean deoxygenation in a warming world. Annual Review of Marine Science 2: 199-229.
Keith SA, Herbert RJH, Norton PA, Hawkins SJ, Newton AC 2011. Individualistic species limitations of climate-induced range expansions generated by meso-scale dispersal barriers. Diversity and Distributions 17: 275-286.
Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, Caig V, Heller-Uszynska K, Jaccoud D, Hopper C, et al. 2012. Diversity arrays technology: A generic genome profiling technology on open platforms. In: Pompanon F, Bonin A ed. Data Production and Analysis in Population Genomics. Totowa, NJ, Humana Press. p. 67-89.
Konopinski MK 2020. Shannon diversity index: A call to replace the original Shannon's formula with unbiased estimator in the population genetics studies. PeerJ 8.
Law CS, Bell JJ, Bostock HCet al. 2017b. Ocean acidification in New Zealand waters: Trends and impacts. New Zealand Journal of Marine and Freshwater Research 52: 155-195.
Law CS, Rickard GJ, Mikaloff-Fletcher SEet al. 2017a. Climate change projections for the surface ocean around New Zealand. New Zealand Journal of Marine and Freshwater Research 52: 309-335.
Liggins L, Gleeson L, Riginos C 2014. Evaluating edge-of-range genetic patterns for tropical echinoderms, *Acanthaster planci* and *Tripneustes gratilla*, of the Kermadec Islands, southwest Pacific. Bulletin of Marine Science 90: 379-397.
Liggins L, Kilduff L, Trnski Tet al. 2021. Morphological and genetic divergence supports peripheral endemism and a recent evolutionary history of Chrysiptera demoiselles in the subtropical South Pacific. Coral Reefs.
Liggins L, Treml EA, Riginos C 2013. Taking the plunge: An introduction to undertaking seascape genetic studies and using biophysical models. Geography Compass 7: 173-196.
Liggins L, Treml EA, Riginos C 2019. Seascape genomics: Contextualizing adaptive and neutral genomic variation in the ocean environment. In: Oleksiak M, Rajora O eds. Population Genomics: Marine Organisms. Springer, Cham. p. 171-218.
Ling SD 2008. Range expansion of a habitat-modifying species leads to loss of taxonomic diversity: A new and impoverished reef state. Oecologia 156: 883-894.
Ling SD, Johnson CR 2008a. Growth and age across range extension region of Centrostephanus rodgersii in eastern Tasmania and morphometric comparison of urchins inhabiting kelp versus barrens habitats: dataset 3 - Allometery for conversion between jaw length and test diameter. Australian Ocean Data Network.
Ling SD, Johnson CR 2009a. Population dynamics of an ecologically important range-extender: Kelp beds versus sea urchin barrens. Marine Ecology Progress Series 374: 113-125.
Ling SD, Johnson CR 2012. Marine reserves reduce risk of climate-driven phase shift by reinstating size - and habitat - specific trophic interactions. Ecological Applications 22: 1232-1245.
Ling SD, Johnson CR, Frusher S, King CK 2008b. Reproductive potential of a marine ecosystem engineer at the edge of a newly expanded range. Global Change Biology 14: 907-915.

Ling SD, Johnson CR, Frusher SD, Ridgway KR 2009b. Overfishing reduces resilience of kelp beds to climate-driven catastrophic phase shift. Proceedings of the National Academy of Sciences of the United States of America 106: 22341–22345.
Ling SD, Johnson CR, Ridgway K, Hobday AJ, Haddon M 2009c. Climate-driven range extension of a sea urchin: Inferring future trends by analysis of recent population dynamics. Global Change Biology 15: 719-731.
Ling SD, Keane JP 2018. Resurvey of the Long-spined sea urchin (*Centrostephanus rodgersii*) and associated barren reef in Tasmania. Tasmania: Institute for Marine and Antarctic Studies, University of Tasmania.
Ling SD, Scheibling RE, Rassweiler A, Johnson CR, Shears N, Connell SD, Salomon AK, Norderhaug KM, Pérez-Matus A, Hernández JC, et al. 2015. Global regime shift dynamics of catastrophic sea urchin overgrazing. Philosophical Transactions of the Royal Society B: Biological Sciences 370.
Lisson D 2018. Maintaining healthy abalone reef systems on Tasmania's East Coast. Tasmania: Tasmanian Abalone Council Ltd.
Lotterhos KE, Whitlock MC 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. Molecular Ecology 23: 2178-2192.
Luu K, Bazin E, Blum MG 2017. Pcadapt: An R package to perform genome scans for selection based on principal component analysis. Molecular Ecology Resources 17: 67-77.
Mariani S, Bekkevold D 2013. The nuclear genome: Neutral and adaptive markers in fisheries science. In: Cadrin SX, Kerr L, Mariani S ed. Stock Identification Methods: Applications in Fishery Science. 2 ed, Academic Press. p. 297-327.
McElreath R 2016. Statistical Rethinking: A bayesian course with examples in R and stan, CRC Press.
McElreath R 2020. Rethinking: Statistical Rethinking book package. Version 2.01.
Melbourne-Thomas J, Audzijonyte A, Brasier MJ, Cresswell KA, Fogarty HE, Haward M, Hobday AJ, Hunt HL, Ling SD, McCormack PC, et al. 2021. Poleward bound: adapting to climate-driven species redistribution. Reviews in Fish Biology and Fisheries.
Middleton I, Aguirre JD, Trnski T, Francis M, Duffy C, Liggins L, Leroy B, et al. 2021. Introduced alien, range extension or just visiting? Combining citizen science observations and expert knowledge to classify range dynamics of marine fishes. Diversity and Distributions 27: 1278-1293.
Miller SL, Abraham ER 2011. Characterisation of New Zealand kina fisheries. Wellington: Ministry of Fisheries.
Minor ES, Urban DL 2008. A graph-theory framework for evaluating landscape connectivity and conservation planning. Conservation Biology 22: 297-307.
Morgan LE, Wing SR, Botsford LW, Lundquist CJ, Diehl JM 2001. Spatial variability in red sea urchin (*Strongylocentrotus franciscanus*) recruitment in northern California. Fisheries Oceanography 9: 83-98.
Murray F, Widdicombe S, McNeill CL, Solan M 2013. Consequences of a simulated rapid ocean acidification event for benthic ecosystem processes and functions. Marine Pollution Bulletin 73: 435-442.
Oke PR, Roughan M, Cetina-Heredia P, Pilo GS, Ridgway KR, Rykova T, Archer MR, Coleman RC, Kerry CG, Rocha C, et al. 2019. Revisiting the circulation of the East Australian Current: Its path, separation, and eddy field. Progress in Oceanography 176.

Oliver ECJ, Benthuysen JA, Bindoff NL, Hobday AJ, Holbrook NJ, Mundy CN, Perkins-Kirkpatrick SE 2017. The unprecedented 2015/16 Tasman Sea marine heatwave. Nature Communications 8.

Oliver ECJ, Lago V, Hobday AJ, Holbrook NJ, Ling SD, Mundy CN 2018. Marine heatwaves off eastern Tasmania: Trends, interannual variability, and predictability. Progress in Oceanography 161: 116-130.

Paterno M, Schiavina M, Aglieri G, Ben Souissi J, Boscari E, Casagrandi R, Chassanite A, Chiantore M, Congiu L, Guarnieri G, et al. 2017. Population genomics meet Lagrangian simulations: Oceanographic patterns and long larval duration ensure connectivity among *Paracentrotus lividus* populations in the Adriatic and Ionian seas. Ecology and Evolution 7: 2463-2479.

Pecl GT, Araújo MB, Bell JD, Blanchard J, Bonebrake TC, Chen IC, Clark TD, Colwell RK, Danielsen F, Evengård B, et al. 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science 355.

Pecorino D, Lamare MD, Barker MF 2012. Growth, morphometrics and size structure of the Diadematidae sea urchin *Centrostephanus rodgersii* in northern New Zealand. Marine and Freshwater Research 63: 624-634.

Pecorino D, Lamare MD, Barker MF 2013a. Reproduction of the Diadematidae sea urchin *Centrostephanus rodgersii* in a recently colonized area of northern New Zealand. Marine Biology Research 9: 157-168.

Pecorino D, Lamare MD, Barker MF, Byrne M 2013b. How does embryonic and larval thermal tolerance contribute to the distribution of the sea urchin *Centrostephanus rodgersii* (Diadematidae) in New Zealand? Journal of Experimental Marine Biology and Ecology 445: 120-128.

Pecorino D, Barker MF, Dworjanyn SA, Byrne M, Lamare MD 2013c. Impacts of near future sea surface pH and temperature conditions on fertilisation and embryonic development in *Centrostephanus rodgersii* from northern New Zealand and northern New South Wales, Australia. Marine Biology 161: 101-110.

Peña TS, Watson JR, González-Guzmán LI, Keitt TH 2017. Step-wise drops in modularity and the fragmentation of exploited marine metapopulations. Landscape Ecology 32: 1643-1656.

Perkins NR, Hill NA, Foster SD, Barrett NS 2015. Altered niche of an ecologically significant urchin species, *Centrostephanus rodgersii*, in its extended range revealed using an Autonomous Underwater Vehicle. Estuarine, Coastal and Shelf Science 155: 56-65.

Peterman WE, Ousterhout BH, Anderson TL, Drake DL, Semlitsch RD, Eggert LS 2016. Assessing modularity in genetic networks to manage spatially structured metapopulations. Ecosphere 7: 1-16.

Pinsky ML, Eikeset AM, McCauley DJ, Payne JL, Sunday JM 2019. Greater vulnerability to warming of marine versus terrestrial ectotherms. Nature 569: 108-111.

Pinsky ML, Selden RL, Kitchel ZJ 2020. Climate-driven shifts in marine species ranges: Scaling from organisms to communities. Annual Review Marine Science 12: 153-179.

Poloczanska ES, Brown CJ, Sydeman WJet al. 2013. Global imprint of climate change on marine life. Nature Climate Change 3: 919-925.

Poloczanska ES, Burrows MT, Brown CJet al. 2016. Responses of marine organisms to climate change across oceans. Frontiers in Marine Science 3.

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo A, Lareu MV 2013. An overview of STRUCTURE: Applications, parameter settings, and supporting software. Frontiers in Genetics 4.
Pritchard JK, Stephens M, Donnelly P 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.
Rahman MA, Arshad A, Yusoff FM 2014. Sea urchins (Echinodermata: Echinoidea): Their biology, culture and bioactive compounds. International Conference on Agricultural, Ecological and Medical Sciences. London. p. 39-48.
Ramos JE, Pecl GT, Moltschaniwskyj NA, Semmens JM, Souza CA, Strugnell JM 2018. Population genetic signatures of a climate change driven marine range extension. Scientific Reports 8.
Robinson LM, Marzloff MP, van Putten I, Pecl G, Jennings S, Nicol S, Hobday AJ, Tracey S, Hartmann K, Haward M, et al. 2020. Decision support for the Ecosystem-Based Management of a Range-Extending Species in a Global Marine Hotspot Presents Effective Strategies and Challenges. Ecosystems.
RStudio Team 2020. RStudio: Integrated development for R. RStudio. Version 1.2.1335. Boston, MA.
Salinger MJ, Diamond HJ, Behrens E, Fernandez D, Fitzharris BB, Herold N, Johnstone P, Kerckhoffs H, Mullan AB, Parker AK, et al. 2020. Unparalleled coupled ocean-atmosphere summer heatwaves in the New Zealand region: Drivers, mechanisms and impacts. Climatic Change 162: 485-506.
Salinger MJ, Renwick J, Behrens E, Mullan AB, Diamond HJ, Sirguey P, Smith Robert O, Trought Michael CT, Alexander L, Cullen NJ, et al. 2019. The unprecedented coupled ocean-atmosphere summer heatwave in the New Zealand region 2017/18: drivers, mechanisms and impacts. Environmental Research Letters 14.
Sanderson JC, Ling SD, Dominguez JG, Johnson CR 2016. Limited effectiveness of divers to mitigate 'barrens' formation by culling sea urchins while fishing for abalone. Marine and Freshwater Research 67: 84–95.
Savary P, Foltête JC, Moal H, Vuidel G, Garnier S, Gaggiotti O 2020. Graph4lg: A package for constructing and analysing graphs for landscape genetics in R. Methods in Ecology and Evolution 12: 539-547.
Scheffers BR, Pecl G 2019. Persecuting, protecting or ignoring biodiversity under climate change. Nature Climate Change 9: 581-586.
Seafood New Zealand 2019. New Zealand seafood exports 2019.
Shears N 2020. Changing of the guard: Subtropical sea urchin taking hold in an iconic New Zealand marine reserve. New Zealand Marine Sciences Society Webinar Series 2020. Auckland.
Shears NT, Babcock RC 2002. Marine reserves demonstrate top-down control of community structure on temperate reefs. Oecologia 132: 131-142.
Shears NT, Bowen MM 2017. Half a century of coastal temperature records reveal complex warming trends in western boundary currents. Scientific Reports 7.
Shears NT, Smith F, Babcock RC, Duffy CA, Villouta E 2008. Evaluation of biogeographic classification schemes for conservation planning: Application to New Zealand's coastal marine environment. Conservation Biology 22: 467-81.
Sirkkamaa S 1983. Calculations on the decrease of genetic variation due to the founder effect. Hereditas 99: 11-20.

Smale DA, Wernberg T 2013. Extreme climatic event drives range contraction of a habitat-forming species. Proceedings of the Royal Society B: Biological Sciences 280.
Sorte CJB, Williams SL, Carlton JT 2010. Marine range shifts and species introductions: Comparative spread rates and community impacts. Global Ecology and Biogeography 19: 303-316.
Stanton BR, Sutton PJH, Chiswell SM 1997. The East Auckland Current, 1994–95. New Zealand Journal of Marine and Freshwater Research 31: 537-549.
Stortini CH, Chabot D, Shackell NL 2017. Marine species in ambient low-oxygen regions subject to double jeopardy impacts of climate change. Global Change Biology 23: 2284-2296.
Sutton PJH, Bowen M 2019. Ocean temperature change around New Zealand over the last 36 years. New Zealand Journal of Marine and Freshwater Research 53: 305-326.
Sward D, Monk J, Barrett NS, Pettorelli N, Rowlands G 2021. Regional estimates of a range-extending ecosystem engineer using stereo-imagery from ROV transects collected with an efficient, spatially balanced design. Remote Sensing in Ecology and Conservation.
Taylor RH, Wilson PR 1990. Recent increase and southern expansion of Adelie penguin populations in the Ross Sea, Antarctica, related to climatic warming. New Zealand Journal of Ecology 14: 25-29.
Teagle H, Smale DA, Schoeman D 2018. Climate-driven substitution of habitat-forming species leads to reduced biodiversity within a temperate marine community. Diversity and Distributions 24: 1367-1380.
Tegner MJ, Dayton PK 1981. Population structure, recruitment and mortality of two sea urchins (*Strongylocentrotus franciscanus* and *S. purpuratus*) in a kelp forest. Marine Ecology Progress Series 5: 255-268.
Thomas LJ, Liggins L, Banks SC, Carter L, Byrne M, Cumming RA, Beheregaray LB, Lamare MD, et al. 2021. The population genetic structure of the urchin *Centrostephanus rodgersii* in New Zealand. Marine Biology 168.
Thomsen MS, Mondardini L, Alestra T, Gerrity S, Tait L, South PM, Lilley SA, Schiel DR, et al. 2019. Local extinction of bull kelp (*Durvillaea spp.*) due to a marine heatwave. Frontiers in Marine Science 6.
Townhill BL, Pinnegar JK, Righton DA, Metcalfe JD 2017. Fisheries, low oxygen and climate change: How much do we really know? Journal of Fish Biology 90: 723-750.
Tracey S, Mundy C, Baulch T, Marzloff M, Hartmann K, Ling S, Tisdell J 2014. Trial of an industry implemented, spatially discrete eradication/control program for *Centrostephanus rodgersii* in Tasmania. Tasmania: Institute of Marine and Antartic Studies.
Treml EA, Ford JR, Black KP, Swearer SE 2015. Identifying the key biophysical drivers, connectivity outcomes, and metapopulation consequences of larval dispersal in the sea. Movement Ecology 3.
Urchinomics 2020. Urchinomics global challenge https://www.urchinomics.com/global-challenge (accessed September 2021).
Verges A, Doropoulos C, Malcolm HAet al. 2016. Long-term empirical evidence of ocean warming leading to tropicalization of fish communities, increased herbivory, and loss of kelp. Proceedings of the National Academy of Sciences of the United States of America 113: 13791-13796.

Wassmann P, Duarte CM, AgustÍ S, Sejr MK 2011. Footprints of climate change in the Arctic marine ecosystem. Global Change Biology 17: 1235-1249.
Webber DN, Starr PJ, Haist V, Rudd MB, Edwards CTT 2018. The 2017 stock assessment and management procedure evaluation for rock lobsters (*Jasus edwardsii*) in CRA 2. Wellington: Ministry for Primary Industries.
Wernberg T, Bennett S, Babcock RCet al. 2021. Climate-driven regime shift of a temperate marine ecosystem. Science 353: 169-172.
Whitlock MC, Lotterhos KE 2015. Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of F(ST). The American Naturalist 186 Suppl 1: S24-36.
Wickham H, Chang W, Henry Let al. 2020. Ggplot2: Create elegant data visualisations using the grammar of graphics. Version: 3.3.5.
Wilson LJ, Fulton CJ, Hogg AM, Joyce KE, Radford BTM, Fraser CI 2016. Climate-driven changes to ocean circulation and their inferred impacts on marine dispersal patterns. Global Ecology and Biogeography 25: 923-939.
Xuereb A, Kimber CM, Curtis JMR, Bernatchez L, Fortin MJ 2018. Putatively adaptive genetic variation in the giant California sea cucumber (*Parastichopus californicus*) as revealed by environmental association analysis of restriction-site associated DNA sequencing data. Molecular Ecology 27: 5035-5048.
Yamano H, Sugihara K, Nomura K 2011. Rapid poleward range expansion of tropical reef corals in response to rising sea surface temperatures. Geophysical Research Letters 38: 1-6.
Zacherl D, Gaines SD, Lonhart SI 2003. The limits to biogeographical distributions: insights from the northward range extension of the marine snail, *Kelletia kelletii* (Forbes, 1852). Journal of Biogeography 30: 913–924.
Zimmerman SJ, Aldridge CL, Oyler-McCance SJ 2020. An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. BMC Genomics 21.