A population genetic investigation of genetic diversity and biogeographic patterns in *Piper excelsum subsp. excelsum* (G. Forst) (Piperaceae) on the North Island of New Zealand

A thesis presented in partial fulfilment of the

requirements for the degree of

Master of Science

in

Biological Sciences

At Massey University, [Manawatū],

New Zealand

Anna-Leigh Gillies

2023

Abstract

New Zealand has been exposed to significant geological activity over time, including tectonic movements, land submersion by water, glacial cycles, and volcanic eruptions. These geological events have altered the spatial and genetic distribution of many species in New Zealand and provide an interesting and complex environment to study biogeographic patterns, genetic diversity, and population structure. Population genetics studies can offer meaningful insight when exploring biogeographic patterns, however, there are relatively few population genetic studies on New Zealand native plants.

Piper excelsum subsp. excelsum (G. Forst.) (Piperaceae) is a native, understory tree that is abundant in mixed angiosperm-podocarp New Zealand forest (Burrows, 1995). This species was chosen to study biogeographic patterns in the North Island using a population genetics approach due to its abundance, widespread distribution and combination of life history characteristics. Accordingly, ten consistently amplifying, polymorphic microsatellite markers were developed to characterise the genetic diversity, population structure and to investigate the potential influence of historical geological events on the genetic and spatial distribution of the species.

An east-west pattern was discovered in sampled North Island sites across all analysis methods, potentially a result of repeated volcanic activity across the volcanic plateau, the axial mountain range that runs from Wellington to East Cape, or Last Glacial Maximum (LGM) microrefugia on either side of this mountain range. Low between sampling site divergence was uncovered for *P. excelsum*. A bottleneck event is speculated in this study to have wiped out the frost intolerant *P. excelsum* during the LGM, with persistence proposed only in Northern refugia. Effective pollen and seed dispersal mechanisms among other ecological variables are also thought to be contributing to this finding.

The genetic distribution of *P. excelsum* has been explored across the North Island of New Zealand. The results from this study contribute meaningful insights to understanding widescale biogeographical patterns of native New Zealand species, although a range of taxa need to be investigated to determine whether these patterns are relatively common among native New Zealand plant species.

Acknowledgements

Firstly, I'd like to thank my supervisor Vaughan Symonds for all his help, time, patience, and advice through the writing of my thesis. I feel very grateful to have had such a kind mentor through this process and have enjoyed learning so much from you.

Thank you to Prashant Joshi for your teachings in the laboratory. You have been very patient and taught me all the tricks in the book and I appreciate the time and effort you have put in to help me complete my thesis. A further thank you to the LoST lab members and Boffin lounge crew who have welcomed me in very kindly; the tea, interesting debates and potluck dinners made for a very fun few years.

Thank you to Manaaki Whenua Landcare Research for the sampling permit and the various DOC staff for allowing me to sample across the North Island. A special thanks all those that came with me on sampling adventures and to Marley and Graham for sampling for me in Northland and East Cape. Thank you also to the Famers' Union Scholarship for supplying me some funding. A massive thank you to my family and friends for their unwavering support throughout my project, I am very grateful for the encouragement.

List of Abbreviations

µL	microlitres
AFLP	amplified fragment length polymorphism
BP	before present
bp	base pair
CASS	cheaply amplified size standard
CTAB	cetyltrimethylammonium bromide
DAPC	discriminant analysis of principal components
DNA	deoxyribonucleic acid
e.g.	for example
ENM	Ecological Niche Modelling
F_{IS}	Wright's fixation coefficient
F_{ST}	measure of population differentiation
GBS	genotyping by sequencing
He	expected heterozygosity
Ho	observed heterozygosity
HWE	Hardy Weinberg equilibrium
IBD	isolation by distance
ID	sampling site code
ISSR	inter simple sequence repeat
ka	thousands of years ago
Km	kilometers
LGM	Last Glacial Maximum
PCA	principle component analysis
PCR	polymerase chain reaction
PCoA	principle coordinates analysis
m	meters
Ma	millions of years ago

mM	millimolar
N	number of samples
Na	number of alleles
Ne	effective number of alleles per locus
P	number of private alleles
RADseq	restriction site association sequencing
RAPD	random amplified polymorphic DNA
SE	standard error
SNP	single nucleotide polymorphism
SSR	single sequence repeats
STE	sucrose, tris, EDTA
TBE	tris borate EDTA
TVZ	Taupō Volcanic Zone

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Biogeography

1.1.1 Biogeography vs Phylogeography

Biogeography and phylogeography are two interlinked fields of study that are used to help understand species distributions and the relevant historical processes that underlie observed patterns. Biogeography studies, referred to as pattern and process identification by Myers and Giller (2013), investigate species genetic divergence and distributional patterns over time and aim to decipher the influence of historical, ecological or geological factors (Ebach, 2015; Shelley et al., 2020). Biogeography is multidisciplinary, combining systematics, ecology, phylogenetics, and palaeontology and can utilise species distribution modelling, for example ecological niche modelling (ENM), to address important questions, such as the impact of global climate change on the shift in species ranges (Franklin, 2023; Hu et al., 2009). There are two main approaches to biogeography: historical and ecological. They have similar aims but have contrasting spatial and temporal magnitudes, with historical working at a longer evolutionary timeframe and ecological at a more recent timeframe and smaller spatial scale (Myers & Giller, 2013).

Phylogeography is a sub-discipline of biogeography. It involves determining the phylogenetic and historical processes causing spatial distribution among and within species gene lineages, with this information considered in a broader temporal perspective (Avise, 2000; Summerell et al., 2010). Phylogeography studies often combine many taxa and incorporate many relevant fields, such as phylogenetics, geology and ecology (Wallis & Trewick, 2009). Paleoclimate and ecological niche models, among others, have given phylogeography a framework for testing hypotheses of species distribution, with the aim of offering spatial context for interpreting phylogeographic and biogeographic results (Carstens & Richards, 2007; Hugall et al., 2002). New Zealand is a particularly interesting place to study biogeography and phylogeography, with the distinctness of the biota attributed to the extended isolation from other landmasses (Köhler & Schmitt, 2023; McGlone, 1985).

1.1.2 New Zealand's geological history

New Zealand's geological history has been turbulent. During the Oligocene, evidence suggests that present day New Zealand was partially (Kamp, 1986; King, 2000; McGlone, 1985) or fully submerged by water (Campbell & Landis, 2001; Landis et al., 2008). The extent of the submersion is a topic that is somewhat disputed in the literature (Waters & Craw, 2006), with Mildenhall et al. (2014) suggesting that although total marine submersion is not as likely, it cannot be discounted due to the inaccuracy of contemporary biostratigraphic and radiometric dating methods. New Zealand's geological history is characterised by many historically important events, several of which are a product of tectonic activity. From around 23 Ma, New Zealand has been raised, and northern and southern crustal plates have separated from Zealandia, which in turn produced the Alpine Fault and subduction zones below Fiordland and the eastern North Island (Campbell et al., 2012).

Volcanism and glaciation have also shaped New Zealand's biota significantly (Shepherd et al., 2007). The Taupō Volcanic Zone (TVZ) has been active for the last 2 Ma, is 60km wide, 300km long and contains several active volcanoes, with the central section erupting large amounts of rhyolitic magma over this time (Houghton et al., 1995; Newnham et al., 1999). Since the Last Glacial Maximum (LGM), there have been two notable volcanic eruptions in the TVZ, the Taupō and Oruanui, approximately 1.8 ka and 26 ka respectively (Alloway et al., 2007; Manville & Wilson, 2004). The Oruanui emitted approximately 800km^3 of pyroclastic material (Newnham et al., 1999) and the Taupō eruption is thought to have produced a large ash blanket in a 20,000km^2 blast zone that eliminated vegetation (Manville et al., 2009), potentially halting gene flow between eastern and western areas of the North Island. The Taupō eruption devastated large areas of New Zealand forest but this area is purported to have revegetated within ~200 years post-eruption (Wilmshurst & McGlone, 1996).

During the Pleistocene, there were many alternating interglacial and glacial periods. The fluctuation between warm and cold environments would have affected the distribution and genetic variation of a wide range of biota (Hewitt, 2000). The LGM is thought to have begun anywhere between 34 ka (Alloway et al., 2007) to 23 ka (Hughes et al., 2013; Mix et al., 2001) and ended around 18,000 BP. Temperatures during the LGM are estimated to have ranged from 5-7°C (Eaves & Brook, 2021), 4.1-5.3°C (Newnham et al., 2013) or 5°C (Rojas et al., 2009) colder than the present day, on average, across New Zealand. During the LGM, glaciation was reasonably extensive in the South Island, covering the Southern Alps and pockets of land surrounding the Alps (McGlone, 1985). North Island

glaciers were scarce in comparison, with glaciers reported on Mt Ruapehu, Tongariro, Taranaki and the Tararua Ranges, with other sites, such as the Ruahine and Kaweka Ranges suggested as possible locations of marginal glaciation, although the extent is uncertain (Eaves & Brook, 2021; Eaves et al., 2016). Evidence from pollen, geomorphic, beetle and macrofossil records suggest that during the LGM, Northland was the only area with near continuous podocarp-broadleaf and beech forest, apart from Mt George in the north-western South Island (Moar et al., 2008) and other small forest patches (McGlone, 1985). Dry, open grassland was typical across most of the east coast of both Islands, with shrubland and grassland largely making up the rest of New Zealand's vegetation cover (McGlone et al., 2010; Newnham et al., 2013). The complex geological history of New Zealand, among other factors, has led to several biogeographic patterns within native and endemic species.

1.1.3 Biogeographic boundaries on the North Island of New Zealand

There are several biogeographical 'boundaries' that have been discussed in the context of New Zealand's biota. These include the Kauri, Northland, Taupō and Cockayne Lines (Figure 1.1). Ellis et al. (2015) note that knowing the rough timing of clade splits is critical when attempting to assign biogeographic reasoning to a species or group.

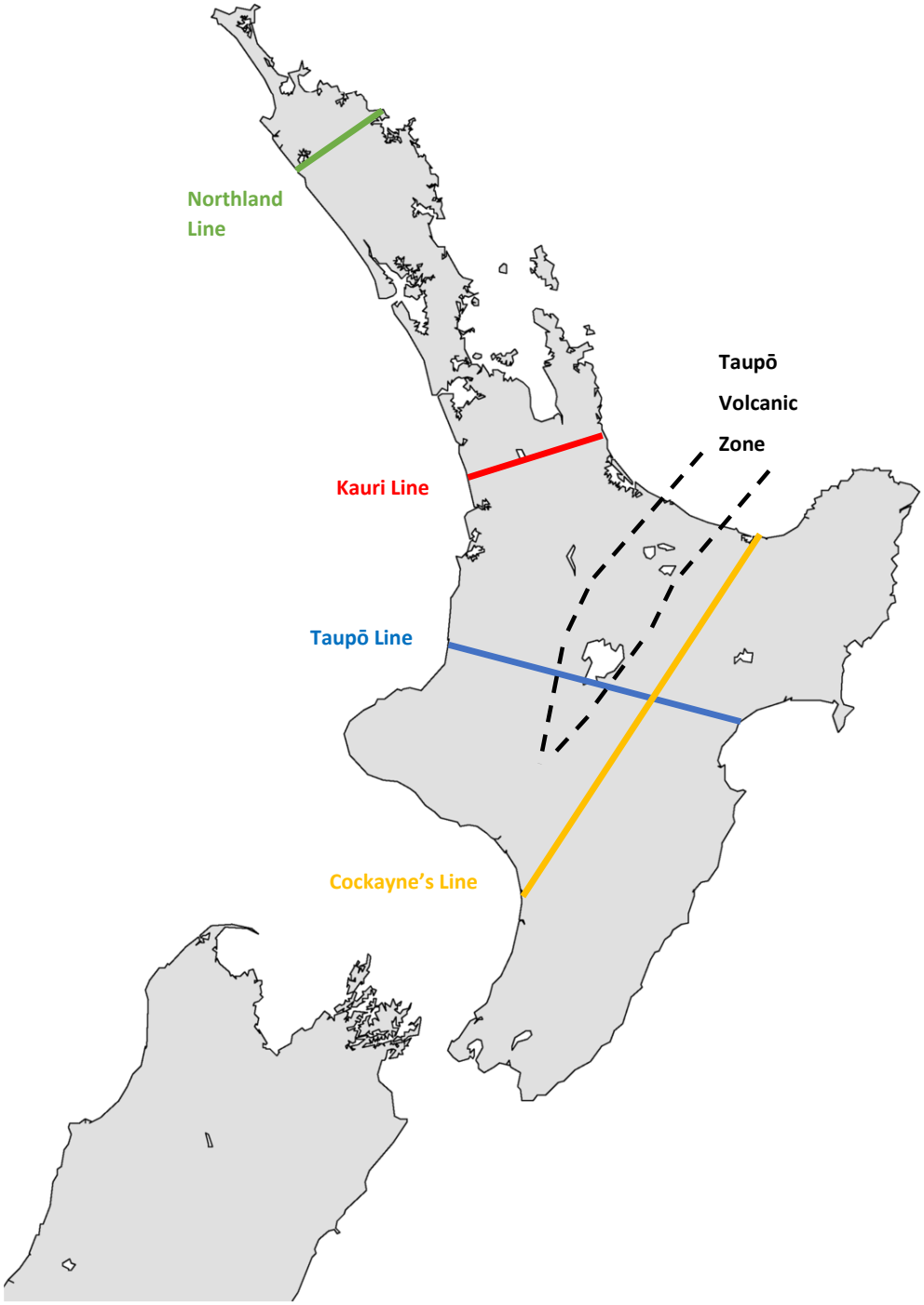


Figure 1.1: A map of the North Island of New Zealand showing the approximate positions of biogeographically significant lines identified in the literature and the Taupō Volcanic Zone.

The Kauri Line is the southern boundary of the Kauri tree, *Agathis australis* (Wagstaff & Clarkson, 2012). A pattern of 'northern richness and southern purity' is found within several endemic species, with diversity levels higher above this line and genetic uniformity below this boundary (Ellis et al.,

2015; Wallis & Trewick, 2009) and this pattern seen in *Lasiorhynchus barbicornis* (Attalabidae) (Painting et al., 2017), *Kikihia ochrina* (Cicadidae) (Ellis et al., 2015), *Clitarchus hookeri* (Phasmatidae) (Buckley et al., 2009; Morgan-Richards et al., 2010), fungus beetles (Zopheridae) (Marske et al., 2011) and *Hemideina thoracica* (Anostostomatidae) (Morgan-Richards et al., 2001), among others. This pattern aligns with the continuous forest during the LGM above this line (Newnham et al., 2013) providing a likely refugial area, and subsequent reestablishment of species down New Zealand post-LGM, however, mtDNA evidence suggests this 'northern richness and southern purity' pattern is likely a result of several processes, not only glacial cycles (Trewick et al., 2011).

North of the Northland Line is characterised by distinct ultramafic soil, which has high magnesium and iron concentrations (Kruckeberg, 2004), leading to the inhabitance of atypical biota (Ellis et al., 2015). During the Pliocene, raised sea levels around 3 Ma led to the formation of several islands in Northland (Trewick & Bland, 2012). High contemporary endemism levels in plants (McGlone, 1985) and insects have been found in this region, with distinct haplotypes noted for the giraffe weevil *Lasiorhynchus barbicornis* (Painting et al., 2017) and in several cicada taxa (Ellis et al., 2015) above the Northland Line.

The Taupō Line was first discussed in terms of endemism, with higher levels of endemic diversity above (125 species) and lower levels below this line (36 species) (McGlone, 1985; Wardle, 1963). The Taupō Line is currently used as a reference point to the more diversity and species found above the line and the lower variation below it. This is demonstrated in several species (Buckley et al., 2010; Buckley et al., 2009; Bulgarella et al., 2014; Heenan et al., 2022; Marske et al., 2011; Painting et al., 2017). Theories for this pattern include Pliocene flooding, tectonic, and volcanic and glacial activity during the Pleistocene (Alloway et al., 2007; Buckley et al., 2010; Bunce et al., 2009; McGlone, 1985; Trewick & Bland, 2012). Ellis et al. (2015) noted that the majority (95 of 125) of the endemic species north of the Taupō Line, were located above the Kauri Line, suggesting that the Kauri Line is more meaningful than the Taupō Line.

The term 'Cockayne's Line' was coined by Ellis et al. (2015), to refer to the previously identified east-west split through the North Island, and refers more specifically to the Wellington-East Cape mountain ranges that uplifted around 2 Ma (Bunce et al., 2009). The higher elevation in the southern end of the ranges is discussed historically by Cockayne to decrease dispersal ability across this boundary. Consequently, an east-west pattern is observed in several species such as *Clitarchus*

hookeri (Phasmatidae) (Buckley et al., 2010), *Pachyornis mappini* (Emeidae) (Baker et al., 2005) and the cicada genus *Notopsalta* (Cicadidae) (Marshall et al., 2012).

1.1.4 Biogeographic and phylogeographic studies in New Zealand

New Zealand is characterised by a high level of endemic species. However, Wallis and Trewick (2009) discuss in a phylogeography review how it is typical to observe low endemism areas and distributional gaps in the New Zealand flora, and both are attributed to a wide range of factors such as tectonic or climatic disruptions. For example, the varying levels of endemism have been identified by McGlone et al. (2001) and Wardle (1988), and are suggested to be a result of variation in terrain, climate and environment that have led to isolation or speciation opportunities (McGlone et al., 2001). Endemism levels are typically lower in areas that have had more geological disturbance, such as extensive glaciation in the central South Island and marine transgression in the lower half of the North Island, and higher in areas that served as glacial refugia through the LGM, such as the northern North Island (McGlone, 1985; Wardle, 1963).

Disjunct spatial and genetic distributions have been observed across many New Zealand species. Several native patchy distributions are thought to be due to the impact of the LGM confining species to refugia, as suggested for *Pseudopanax crassifolius* (Araliaceae) (Gemmell et al., 2022), *Metrosideros* (Myrtaceae) (Drummond et al., 2000; Gardner et al., 2004) and *Nothofagus truncata* (Nothofagaceae) (Haase, 1992). In contrast, some species are thought to have survived relatively widespread during the LGM (Buckley et al., 2010; Shepherd & Perrie, 2011; Shepherd et al., 2007). Determining widescale distribution patterns across species is difficult as there are a variety of adaptations within species, such as behavioural and ecological, that can influence distribution (Shelley et al., 2020). Soltis et al. (2006) stress that a comprehensive study of flora and fauna is required to discover biogeographical trends that may exist, as factors, like habitat preference, can create differentiating phylogenetic and dispersal trends.

Volcanism has also impacted the distribution of New Zealand species. Holzapfel et al. (2002) identified a North Island east-west partitioning of genetic variation in Dactylanthus (Balanophoraceae), thought to be a result of volcanism in the Taupō Volcanic Zone (TVZ). This pattern also has been revealed in native fauna: velvet worms (Peripatopsidae) (Trewick, 2000), short-tailed bat (Mystacinidae) (Lloyd, 2003), North Island brown kiwi (Apterygidae) (Shepherd & Lambert, 2008) and Moa (Baker et al., 2005). However, evidence on this east-west split in native flora is still very limited, with Shepherd et al. (2007) further investigating this pattern in *Asplenium hookerianum*

(Aspleniaceae), and similarly noting that the eruptions likely resulted in a genetic bottleneck in the east and west for *A. hookerianum* and McLay et al. (2022) reinforcing and expanding on the Dactylanthus story. There are many factors that can influence a species' distribution, with the geological history important to consider when investigating distributional patterns.

1.2 Factors influencing species distribution

1.2.1 Climatic and edaphic factors influencing species distribution

Climatic and edaphic factors are the main environmental drivers of plant species' distribution and evolutionary adaptive potential (de la Vega, 1997). Strong associations have been found between plant distribution and the temperature and precipitation factors of climate (Woodward & Williams, 1987), with species distribution models found to be a useful way of investigating climate change impacts (Franklin et al., 2013). Climate change is exerting a strong influence on species ranges, with a study on 1350 European plants predicting more than half may be threatened or at least vulnerable by 2080 (Thuiller et al., 2005). Though climate is a significant external factor, Kruckeberg (2004) coined the term geoedaphic, (an edaphic term that encompasses more broad geological influences, such as topography) where determinants, such as elevation and alpine exposure, among others, are more specifically influencing local tolerance of species ranges. The edaphic factor refers to soil-plant interactions, i.e., the medium the plants grow in and obtain nutrients and water from (Rajakaruna, 2004). The chemical characteristics of soil influence the survival ability of plants, and therefore particular soil preferences impact spatial distribution (Dubuis et al., 2013). Edaphic stresses, such as limited access to oxygen and water, can cause a decrease in plant fitness (Lynch, 2022) and consequently, also limit distribution.

1.2.2 Life history traits influencing species genetic distribution

The genetic structure of a species is formulated through the evolutionary forces of migration, natural selection, random genetic drift, and mutation. These factors are influenced by a species' historical, ecological, and biological background. Therefore, for plants, genetic structure is associated with the spatial structure and distribution of a species due to its restricted distribution abilities (Loveless & Hamrick, 1984).

Many ecological variables and life history traits have been identified as being influential to genetic structure and diversity within and among plant populations. In a review based on isozyme data,

Loveless and Hamrick (1984) consider the influences of several life history characteristics and ecological variables on genetic structure in plants, including breeding system, floral morphology, reproduction process and timing, pollination method, seed dispersal method and dormancy, phenology, life cycle length, successional stage, geographic range, population size, density and spatial distribution. Loveless and Hamrick (1984) discovered that geographic range, breeding system, life form and taxonomic status have a significant impact on within and between plant population genetic composition. The combination of variables producing the highest levels of genetic diversity was found to exist in species that exhibit broad geographic ranges, high reproduction rates, are long lived, outcrossing, and wind-pollinated (Hamrick et al., 1979). Hamrick and Godt (1996) and Gottlieb (1977) also found outcrossing species to be more variable. Hamrick and Godt (1996) found life form and breeding system were the most influential factors on diversity and distribution and that woody species had higher diversity levels than herbaceous species. Conversely, a review deciphering the impact of life history traits on genetic structure of seed plants, found the only associated factor was mating system when using nuclear markers, and seed dispersal or geographic range when using organelle markers (Duminil et al., 2007). Therefore, the interaction between life history traits and genetic structure is complex.

1.3 Population genetics

Population genetics studies provide insights into the level, origin, maintenance, and evolutionary significance of genetic variation within species (Holliday et al., 2018). Population geneticists investigate the genetic structuring and diversity of populations and the gene flow between ecological populations (Balloux & Lugon-Moulin, 2002; Habel et al., 2015). Ecological populations are considered a group of sexually reproducing individuals that inhabit an area (Dobzhansky, 1955) whereas a genetic population involves the assignment of individuals to populations based on genotypic information and is useful for understanding “cryptic” population structure (Pritchard et al., 2000). Populations are subject to several influencing factors, such as migration, natural selection, assortative mating, and mutation (Wigginton et al., 2005). The absence of these factors would lead to the genotype frequencies simply being dependent on allele frequencies and random mating, a phenomenon called Hardy Weinberg equilibrium (HWE) (Chen, 2010). Deviations from HWE could be due to the above influencing factors, or separate causes, such as genotyping errors (Chen et al.,

2017). HWE is considered a valuable tool (Crow, 1988) and has long been used in population genetics studies to make biological inferences.

1.3.1 Molecular markers

Molecular genetic markers are a useful way to study species populations and allow genetic variation to be characterised between and within populations (Altıntaş et al., 2008; Dar et al., 2019). It is imperative to understand that parts of the genome, such as neutral or coding, are subject to differing levels of selection pressures, leading to differences in the way a genetic marker displays genetic variation (Matala et al., 2014) and need to be aligned appropriately to the research questions being posed (Fischer et al., 2017). Using molecular markers to understand species genetic variation can be valuable in many settings; for example: determining conservation status (Ferreira et al., 2020), inbreeding levels (Downey & Baskauf, 2020), relatedness between ecotypes (Hopkins et al., 2021) and interaction between environment and diversity (Ahn et al., 2019; Gasca-Pineda et al., 2020).

The approach to studying the population genetics of natural populations has certainly evolved over time. It started with the introduction of allozymes in the 1960's and utilisation of mitochondrial DNA in the 1970's. Microsatellites then began to be implemented in the 1990's, single nucleotide polymorphisms (SNPs) in the 2000's and more recently, the introduction of the population genomics approach (Allendorf, 2017). Currently, microsatellite loci and SNPs are most commonly used in population genetics studies (Tsykun et al., 2017). Microsatellite loci are made up of one-six base pair DNA repeat regions that are highly polymorphic, co-dominant, mutate at a fast rate, and are useful in plant genetics and breeding (Parida et al., 2009). These consecutive repeat units cause slip-strand mispairing during DNA synthesis which can often result in the insertion or deletion of repeat units (Levinson & Gutman, 1987). 'Alleles' at microsatellite loci largely consist of variation in the number of repeat units (Semagn et al., 2006). These markers are effective at determining genetic variation and population structure (Gebremedhin et al., 2009; Lemopoulos et al., 2019); however, they are also subject to genotyping errors (Van Oosterhout et al., 2004) and homoplasy via mutation. This can confound genetic and allelic variation, resulting in misleading levels of population diversity (Sefc et al., 2007).

SNPs are a single base difference at a certain point in the genome (Zhang et al., 2020). In contrast with microsatellites, they tend to be biallelic and evolve at a more steady pace (Defaveri et al., 2013). Advantageously, they are stable, and found more frequently in the genome than microsatellites (Park et al., 2021). Thus, they are utilised for high-throughput sequencing over microsatellites (Zhang

et al., 2020). High-throughput sequencing technologies, such as Restriction Site Association Sequencing (RADseq) or Genotyping by Sequencing (GBS) are utilised for simultaneously identifying and genotyping SNPs. They are relatively cheaper than assay methods, like fluidigm (Holliday et al., 2018). However, GBS has a random fragmentation nature. This can result in difficulty determining if fragments are neutral or under selection without a reference genome (Holliday et al., 2018). Missing data is also problematic and should be estimated via reference SNPs of the same genome, sequenced via a different method, if possible (Browning & Browning, 2016). Fischer et al. (2017) advises that SNPs sequenced with next generation sequencing techniques, such as RADseq, are the most efficient way to analyse true genetic variation in and between populations and produce meaningful inferences.

There are several issues that can arise when using molecular markers. Missing data and null alleles are problematic for both SNPs and microsatellite markers. Mutations can arise in neighbouring regions of microsatellite priming sites, leading to null alleles that are unable to be amplified as the primer cannot anneal to template DNA (Chapuis & Estoup, 2006). This is problematic as it will artificially lower genetic variation and can lead to an erroneous excess of homozygosity and deviations from HWE (Willyard et al., 2020).

1.3.2 Analyses methods

There are many analysis methods used in study population genetics, such as F-statistics, Bayesian clustering methods, discriminant analysis of principal components (DAPC) and principal component analysis (PCA). Some of the most commonly used analysis methods and software for studying population genetics have been briefly discussed but see Grünwald et al. (2017) for a comprehensive review of these and other frequent practices.

Wright contributed greatly to the field of evolutionary genetics (Wright, 1931), with a notable development of the fixation indices or F-statistics (Wright, 1943, 1951). F-statistics are a widely employed parameter utilised to investigate population diversity and structure (Pearse & Crandall, 2004). F_{ST} describes the between population differentiation, essentially providing a rough estimate of gene flow between populations. F_{ST} values have a numerical range of 0 (no genetic differentiation) to 1 (complete differentiation between populations; i.e., populations are fixed for distinct alleles). F_{IS} examines the relationship between observed and expected heterozygosity within a population. Negative F_{IS} values signify heterozygote excess, potentially suggesting individuals are outcrossing with outside populations, clonal reproduction, or negative assortative mating (da Cunha et al., 2022;

Stoeckel et al., 2006). Positive values indicate homozygote excess, which can imply the presence of within population inbreeding, but can also be attributed to several factors, such as inaccurately defined populations (Wahlund effect), null alleles (Chapuis & Estoup, 2006; De Meeûs, 2017; Montarry et al., 2015) or non-neutral loci. A F_{IS} value of 0 indicates a population is mathematically at HWE.

Genetic structure can be investigated through several approaches that utilise different methodologies, with the implementation of a combination of these often ideal to provide confidence in each outcome. STRUCTURE (Pritchard et al., 2000) is a commonly used and valuable tool in population genetic studies (Novembre, 2016). The software uses a model-based Bayesian clustering method to deduce the number of genetic groupings (*K*) present (Pritchard et al., 2000). This method uses the assumption that genetic populations are at HWE and markers are unlinked (Falush et al., 2003) to assign individuals to 'ancestral' genetic groups.

DAPC (Jombart et al., 2010) is a multivariate analysis method also used to assess the genetic structure of populations. DAPC involves the conversion of the input data into principal components to confirm the alleles are uncorrelated before proceeding to discriminant analysis (DA). From here, DAPC provides a scatter plot visualisation of clustered individuals of genetic similarity. DAPC has demonstrated the ability to decipher intricate population structures and can be completed with computational ease, in comparison to STRUCTURE (Jombart et al., 2010).

PCA (Novembre & Stephens, 2008) and PCoA are also utilised to study structure. PCA helps to decrease the complexity and dimensions of the data via detection of principal components; i.e., where the most variation is observed in the data (Ringnér, 2008). This is carried out with computational ease (Jombart et al., 2010) and without significant loss of variation. PCoA can be utilised to discover population structure between samples of populations that are not pre-assigned (Mao et al., 2020). A more accessible depiction of the data is produced from the principal components as a colour coordinated graph, to provide a visual representation of population dynamics for both methods.

A Neighbor-net (Bryant & Moulton, 2004) uses a distance-based method to cluster genetically similar groups or individuals, which is an alternative method of depicting genetic structure. It produces a network of weighted splits, using SplitsTree (Huson, 1998) to create a splits graph. The Neighbor-net

demonstrates enhanced resolution, compared to similar phylogenetic methods, while also retaining the ability to display conflicts in the data. All these methods and markers are incredibly useful for studying and understanding population genetics.

1.3.3 New Zealand population genetics studies

There are around 2500 native species and nearly 2000 endemic plant species in New Zealand. Currently, the number of population genetic studies completed on New Zealand flora is limited with a range of marker types used. Population genetic studies in New Zealand are commonly aimed at taxonomic uncertainties (for example (Heenan et al., 2018; Karlin et al., 2008; Smissen & Heenan, 2010)) or conservation issues (Armstrong & De Lange, 2005; Barnaud & Houliston, 2010). However, there are many avenues yet to be investigated in the New Zealand flora; for example, comparative studies on rare and common species population structures or compiling evidence to investigate widescale biogeographical patterns.

Microsatellite markers have been utilised to study several New Zealand species. These include: *Craspedia* (Asteraceae) (Breitwieser et al., 2015), *Sphagnum moss* (Phagnaceae) (Karlin et al., 2008), *Pseudopanax ferox* (Araliaceae) (Shepherd & Perrie, 2011), *Fuscospora* (Nothofagaceae) (Smissen et al., 2014), *Arthropodium cirratum* (Asparagaceae) (Shepherd et al., 2018), *Kowhai* (Fabaceae) (Heenan et al., 2018) and *Kunzea* (Myrtaceae) (Heenan et al., 2022) with microsatellites developed for future population genetic studies for *Gingidia enysii and baxterae* (Umbelliferae) (Heenan et al., 2021).

Many combinations of marker types have been used to study New Zealand native species. Amplified fragment length polymorhpisms (AFLPs) were used to investigate genetic differentiation in *Hebe speciosa* (Plantaginaceae) (Armstrong & De Lange, 2005), *Olearia gardneri* (Asteraceae) (Barnaud & Houliston, 2010), *Metrosideros excelsa* (Myrtaceae) (Broadhurst et al., 2008), *Myosotis* (Boraginaceae) (Meudt et al., 2015), *Metrosideros bartlettii* (Myrtaceae) (Drummond et al., 2000), *Cordyline australis* (Laxmanniaceae) (Beever et al., 2013) and *Austroderia turbaria* (Poaceae) (Houliston et al., 2012). Chloroplast markers have been used to look at *Asplenium hookerianum* (Aspleniaceae) (Shepherd et al., 2007), *Sophora* trees (Fabaceae) (Shepherd et al., 2017), *Entelea arborescens* (Malvaceae) with Single Sequence Repeats (SSRs) also used (Shepherd et al., 2019). Random Amplified Polymorphic DNA (RAPD) and microsatellite markers were chosen to study *Dactylanthus taylorii* (Balanophoraceae) (Holzapfel et al., 2002; McLay et al., 2022), and isozymes for *Nothofagus truncata* (Nothofagaceae) (Haase, 1992), *Pseudopanax crassifolius* (Araliaceae) was

investigated with microsatellites and chloroplast haplotypes (Gemmell et al., 2022), Inter Simple Sequence Repeats (ISSRs) were utilised to study *Pittosporum obcordatum* (Pittosporaceae) (Wright et al., 2017), with some using SSRs and AFLPs to study *Phormium tenax* (Asphodelaceae) (Smissen & Heenan, 2010) or a combination of AFLPs and ISSRs to study *Melicytus ramiflorus* (Violaceae) (Stevens et al., 2015).

There are very few published population genetic studies on New Zealand flora utilising SNPs (Buys et al., 2019) and high throughput technologies. This is contrary to the increased popularity of SNPs used elsewhere, (Allendorf, 2017) and could be due to high initial investment costs for sequencing technology (Willyard et al., 2020), with microsatellites more feasible for smaller scaled studies (Kraus et al., 2015).

1.4 *Piper excelsum*

The Piperaceae family contains economically important plants, such as pepper (*Piper nigrum* L.) and notably in Oceania, kava (*Piper methysticum* Forst) (Cox & O'Rourke, 1987). In New Zealand, there are four subspecies of *Piper excelsum*, including *P. excelsum subsp. delangei, peltatum, psittacorum* and *excelsum*. All subspecies except *P. excelsum subsp*. *excelsum* are naturally uncommon, coastal, and found on offshore islands. *P. excelsum subsp. delangei* is found on the Three Kings Islands (de Lange, 2023a), *P. excelsum subsp. peltatum* is distributed on Poor Knights, Mokohinau, Hen, Great Barrier, Hauturu Islands (de Lange, 2023b) and *P. excelsum subsp*. *psittacorum* is found on Kermadec, Norfolk and Lord Howe Islands (de Lange, 2023c).

In New Zealand, one of the most prominent plant species is *Piper excelsum subsp. excelsum* (G. Forst.) (kawakawa or pepper tree) (Figure 1.2). *P. excelsum* is a native, understory tree that is abundant in New Zealand forests (Butts et al., 2019; Hodge et al., 1998). The species is thought to be relatively long-lived, surviving for at least a few decades (Gardner, 1997) and is typically recognised by the heart shaped leaves with many holes created by the herbivore *Cleora sciptaria* (Geometridae). *P. excelsum* prefers damp soil and is naturally found around the entire coastline of the North Island and on the coastline of the South Island down to Banks Peninsula (de Lange, 2021), as well as on several of New Zealand's offshore islands (Lakeman-Fraser & Ewers, 2013) and is common in lowland mixed angiosperm-podocarp forests (Burrows, 1995) (Figure 1.3).

Figure 1.2: *Piper excelsum* leaves displaying herbivore damage on the Sledge Track, Palmerston North. Taken by Anna-Leigh Gillies (May 2021).

Figure 1.3: *Piper excelsum* in its natural habitat. Taken in Orokawa, Waihi by Anna-Leigh Gillies (May 2022).

The pollination method of *P. excelsum* is somewhat disputed. Although previously regarded as entomophilous (Webb et al., 1999), evidence suggests that *P. excelsum* displays traits typical of anemophilous flora, e.g., high pollen to ovule ratio and lack of biotic pollinators visiting the inflorescences (Merrett et al., 2007). However, Gardner (1997) finds conflicting evidence, noting characteristics that indicate both animal and wind pollination. Although suggested anemophily characteristics were noted, such as non-apical dehiscence slits (Burger, 1972) and smooth surfaced pollen grains, the filaments are described as short, a feature more typical of entomophilous plants (Gardner, 1997).

The fruit are an orange, soft spike made up of many individual drupe-like fruit that each have their own seed, forming an aggregate fruit that can contain as many as 150 seeds (Burrows, 1995). Fruits are typically found from January to March (Wotton, 2016). Several birds have been observed feeding on *P. excelsum* fruit, including kereru (*Hemiphaga novaeseelandiae)*, silvereyes (*Zosterops lateralis*) and blackbirds (*Turdus merula*) (Burrows, 1995). The viability of seeds relies heavily on digestion by the frugivore and the seed coat. Burrows (1995) noted that the seeds are likely to have a reasonably

hard coating that can stay intact during bird digestion and that these seeds are transported internally over a minimum of 10-20 metres. Many species of New Zealand native lizards are described as frugivorous, several of which include *P. excelsum* fruits in their diet. These include *Hoplodactylus duvauceli*, *Hoplodactylus pacificus*, *Cyclodina oliveri*, *Leiolopisma smithi* (Whitaker, 1968), *Cyclodina alani* and *Cyclodina whitakeri* (Southey, 1985). Many of these are rare, endangered, and only found on Northern offshore islands of New Zealand. Most skink and gecko species found in New Zealand require minimal chewing for digestion, allowing seeds to be digested without damage (Whitaker, 1987). Therefore, only *H. pacificus* could play a potential minor dispersal vector role for *P. excelsum*. Furthermore, the germination ability of seeds digested by frugivorous lizards and geckos has been investigated. Wotton (2016) found that the. *P. excelsum* seeds digested by gecko *Woodworthia maculata* have no difference in germination success in comparison to seeds taken directly from a mother plant.

P. excelsum has an interesting phytochemical profile. Several previously unidentified compounds in the fruit (Lei et al., 2015; Richardson, 2015) and leaves (Obst et al., 2017) have been characterised, and recently, a detailed phytochemical profile on dried and fresh leaves revealed 58 novel compounds (Jayaprakash et al., 2022). However, recent studies on *P. excelsum* are limited, with the majority of the literature quite dated and focusing on its herbivore, *Cleora scriptaria* (Beever, 1987; Hodge, Barron, et al., 2000; Hodge et al., 1998; Hodge, Keesing, et al., 2000) and phytochemistry (Nishida et al., 1983; Russell & Fenemore, 1973).

There are many plants that have significance and historical uses to Māori, such as *Harakeke* (Asphodelaceae) (Wehi & Clarkson, 2007), with plants such as karamu (Rubiaceae) and kūmarahou (Rhamnaceae) used for rākau rongoā (plant based remedies) (Koia & Shepherd, 2020). *P. excelsum* also has historic rākau rongoā uses for ailments such as toothache, eczema and kidney problems (Brooker & Cooper, 1961). The leaves are also traditionally used in wreaths at tangihanga (Koia & Shepherd, 2020) and were burnt to act as a pesticide to protect important kumara crops (Lambert & Mark-Shadbolt, 2021). The cultural significance of kawakawa could suggest the movement of this species as Māori groups were established and migrating around New Zealand, although there is nothing in the literature to confirm this.

The commercial applications of *P. excelsum* have been explored in several areas. The ability of *P. excelsum* to help treat type two diabetes has been discussed (Koia & Shepherd, 2020) and a small scale study was undertaken where insulin response was shown to increase with consumption of

kawakawa tea (Ramzan et al., 2022). The commercial safety of *P. excelsum* tea has also been investigated (Butts et al., 2019), as dried leaves are commercially used in tea and seasoning (Jayaprakash et al., 2022) and in several types of beauty products. Jayaprakash et al. (2022) also mentions interest is rising in the therapeutic, food and beverage space, with Alderson et al. (2021) also looking into the sensory profile of kawakawa kombucha.

1.5 Research objectives

Currently, there are limited population genetic studies on native plants in New Zealand, leaving a clear gap around our understanding of population genetic structure and diversity of native species. Given its common geographic distribution, interesting phytochemical profile, cultural importance and combination of life history characteristics, *P. excelsum* is an interesting species for a population genetics investigation. A population genetics study on *P. excelsum* would also provide valuable insight for exploring biogeographic patterns, with a relatively low number of studies investigating biogeographical trends in New Zealand flora. Therefore, the primary aims of this study are to:

1. Develop microsatellite markers suitable to assess the population genetics of *P. excelsum*
2. Determine the genetic structure and variation within and between sampled *P. excelsum* sites in the North Island
3. Discuss potential historical influences that could have influenced the present spatial and genetic distribution of *P. excelsum*

1.6 References

Ahn, J.-Y., Lee, J.-W., Lee, M.-W., & Hong, K.-N. (2019). Genetic diversity and structure of *Carpinus laxiflora* populations in South Korea based on AFLP markers. *Forest Science and Technology*, *15*(4), 192-201.

Alderson, H., Liu, C., Mehta, A., Gala, H. S., Mazive, N. R., Chen, Y., Zhang, Y., Wang, S., & Serventi, L. (2021). Sensory Profile of Kombucha Brewed with New Zealand Ingredients by Focus Group and Word Clouds. *Fermentation*, *7*(100).

Allendorf, F. W. (2017). Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology*, *26*(2), 420-430. https://doi.org/https://doi.org/10.1111/mec.13948

Alloway, B. V., Lowe, D. J., Barrell, D. J. A., Newnham, R. M., Almond, P. C., Augustinus, P. C., Bertler, N. A., Carter, L., Litchfield, N. J., McGlone, M. S., Shulmeister, J., Vandergoes, M. J., & Williams, P. W. (2007). Towards a climate event stratigraphy for New Zealand over the past 30 000 years (NZ-INTIMATE project). *Journal of Quaternary Science*, *22*(1), 9-35.

Altıntaş, S., Toklu, F., Kafkas, S., Kilian, B., Brandolini, A., & Özkan, H. (2008). Estimating genetic diversity in durum and bread wheat cultivars from Turkey using AFLP and SAMPL markers. *Plant breeding*, *127*(1), 9-14.

Armstrong, T. T. J., & De Lange, P. J. (2005). Conservation genetics of Hebe speciosa (Plantaginaceae) an endangered New Zealand shrub. *Botanical Journal of the Linnean Society*, *149*(2), 229-239. https://doi.org/10.1111/j.1095-8339.2005.00437.x

Avise, J. C. (2000). *Phylogeography: the history and formation of species*. Harvard university press.

Baker, A. J., Huynen, L. J., Haddrath, O., Millar, C. D., & Lambert, D. M. (2005). Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: The giant moas of New Zealand. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 8257-8262.

Balloux, F., & Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, *11*(2), 155-165.

Barnaud, A., & Houliston, G. J. (2010). Population genetics of the threatened tree daisy Olearia gardneri (Asteraceae), conservation of a critically endangered species. *Conservation Genetics*, *11*, 1515–1522.

Beever, R., Houliston, G., & Armstrong, T. (2013). Geographic patterns of genetic variation in Cordyline australis; does fingerprinting recover phenotypic and phenological patterns? *New Zealand Journal of Botany*, *51*(1), 3-12.

Beever, R. E. (1987). The holes in the leaves of kawakawa *Macropiper excelsum*. *New Zealand Regional Botanical Society Journals*, *42*(1), 9-11.

Breitwieser, I., Ford, K. A., & Smissen, R. D. (2015). Characterisation of SSR markers for New Zealand Craspedia and their application in Kahurangi National Park. *New Zealand Journal of Botany*, *53*(1), 60-73.

Broadhurst, L. M., Young, A. G., & (2008), B. G. M. (2008). AFLPs reveal an absence of geographical genetic structure among remnant populations of pohutukawa (Metrosideros excelsa, Myrtaceae). *New Zealand Journal of Botany*, *46*(1), 13-21.

Brooker, S. G., & Cooper, R. C. (1961). New Zealand medicinal plants. *Economic Botany*, *15*(1), 1-10.

Browning, B. L., & Browning, S. R. (2016). Genotype imputation with millions of reference samples. *American Journal of Human Genetics*, *98*(1), 116-126.

Bryant, D., & Moulton, V. (2004). Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, *21*(2), 255-265.
Buckley, T. R., Marske, K., & Attanayake, D. (2010). Phylogeography and ecological niche modelling of the New Zealand stick insect Clitarchus hookeri (White) support survival in multiple coastal refugia. *Journal of Biogeography*, *37*(4), 682-695.
Buckley, T. R., Marske, K. A., & Attanayake, D. (2009). Identifying glacial refugia in a geographic parthenogen using palaeoclimate modelling and phylogeography: the New Zealand stick insect Argosarchus horridus (White). *Molecular Ecology*, *18*(22), 4650-4663.
Bulgarella, M., Trewick, S. A., Minards, N. A., Jacobson, M. J., & Morgan-Richards, M. (2014). Shifting ranges of two tree weta species (H emideina spp.): competitive exclusion and changing climate. *Journal of Biogeography*, *41*(3), 524-535.
Bunce, M., Worthy, T. H., Phillips, M. J., Holdaway, R. N., Willerslev, E., Haile, J., Shapiro, B., Scofield, R. P., Drummond, A., & Kamp, P. J. (2009). The evolutionary history of the extinct ratite moa and New Zealand Neogene paleogeography. *Proceedings of the National Academy of Sciences*, *106*(49), 20646-20651.
Burger, W. C. (1972). Evolutionary trends in the Central American species of Piper (Piperaceae). *Brittonia*, *24*(4), 356-362.
Burrows, C. J. (1995). Germination behaviour of seeds of the New Zealand species Fuchsia excorticata, Griselinia littoralis, Macropiper excelsum, and Melicytus ramiflorus. *New Zealand Journal of Botany*, *33*(1), 131-140. https://doi.org/10.1080/0028825x.1995.10412950
Butts, C. A., Klink, J. W. v., Joyce, N. I., Paturi, G., Hedderley, D. I., Martell, S., & Harvey, D. (2019). Composition and safety evaluation of tea from New Zealand kawakawa (Piper excelsum). *Journal of Ethnopharmacology*, *232*, 110-118.
Buys, M. H., Winkworth, R. C., De Lange, P. J., Wilson, P. G., Mitchell, N., Lemmon, A. R., Moriarty Lemmon, E., Holland, S., Cherry, J. R., & Klápště, J. (2019). The phylogenomics of diversification on an island: applying anchored hybrid enrichment to New Zealand Leptospermum scoparium (Myrtaceae). *Botanical Journal of the Linnean Society*, *191*(1), 1-17.
Campbell, H., Malahoff, A., Browne, G., Graham, I., & Sutherland, R. (2012). New Zealand Geology. *Episodes Journal of International Geoscience*, *35*(1), 57-71.
Campbell, H. J., & Landis, C. A. (2001). New Zealand awash. *New Zealand Geographic*, *51*, 6-7.
Carstens, B. C., & Richards, C. L. (2007). Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution*, *61*(6), 1439-1454.
Chapuis, M.-P., & Estoup, A. (2006). Microsatellite Null Alleles and Estimation of Population Differentiation. *Molecular Biology and Evolution*, *24*(3), 621-631. https://doi.org/10.1093/molbev/msl191
Chen, B., Cole, J. W., & Grond-Ginsbach, C. (2017). Departure from Hardy Weinberg equilibrium and genotyping error. *Frontiers in genetics*, *8*, 167.
Chen, J. J. (2010). The Hardy-Weinberg principle and its applications in modern population genetics. *Frontiers in Biology*, *5*, 348-353.
Cox, P. A., & O'Rourke, L. (1987). Kava (Piper methysticum, Piperaceae). *Economic Botany*, 452-454.
Crow, J. F. (1988). Eighty years ago: the beginnings of population genetics. *Genetics*, *119*(3), 473.
da Cunha, N. L., Xue, H., Wright, S. I., & Barrett, S. C. (2022). Genetic variation and clonal diversity in floating aquatic plants: Comparative genomic analysis of water hyacinth species in their native range. *Molecular Ecology*, *31*(20), 5307-5325.

Dar, A. A., Mahajan, R., & Sharma, S. (2019). Molecular markers for characterization and conservation of plant genetic resources. *Indian Journal of Agricultural Sciences*, *89*(11), 1755-1763.

de la Vega, M. P. (1997). Plant genetic adaptedness to climatic and edaphic environment. Adaptation in Plant Breeding: Selected Papers from the XIV EUCARPIA Congress on Adaptation in Plant Breeding held at Jyväskylä, Sweden from July 31 to August 4, 1995,

de Lange, P. J. (2021). *Piper excelsum subsp. excelsum Fact Sheet (content continuously updated)*. New Zealand Plant Conservation Network. Retrieved 01 May from https://www.nzpcn.org.nz/flora/species/piper-excelsum-subsp-excelsum/

de Lange, P. J. (2023a). *Piper excelsum subsp. delangei Fact Sheet (content continuously updated)*. New Zealand Plant Conservation Network. Retrieved 22 June from https://www.nzpcn.org.nz/flora/species/piper-excelsum-subsp-delangei/

de Lange, P. J. (2023b). *Piper excelsum subsp. peltatum Fact Sheet (content continuously updated)*. New Zealand Plant Conservation Network. Retrieved 22 June from https://www.nzpcn.org.nz/flora/species/piper-excelsum-subsp-peltatum/

de Lange, P. J. (2023c). *Piper excelsum subsp. psittacorum Fact Sheet (content continuously updated)*. New Zealand Plant Conservation Network. Retrieved 22 June from https://www.nzpcn.org.nz/flora/species/piper-excelsum-subsp-psittacorum/

De Meeûs, T. (2017). Revisiting FIS, FST, Wahlund Effects, and Null Alleles. *Journal of Heredity*, *109*(4), 446-456. https://doi.org/10.1093/jhered/esx106

Defaveri, J., Viitaniemi, H., Leder, E., & Merila, J. (2013). Characterizing genic and nongenic molecular markers: comparison of microsatellites and SNPs. *Molecular Ecology Resources*, *13*(2), 377–392.

Dobzhansky, T. (1955). A review of some fundamental concepts and problems of population genetics. Cold Spring Harbor symposia on quantitative biology,

Downey, K. M., & Baskauf, C. J. (2020). Population genetics of a rare wetland species, the Tennessee yellow-eyed grass (Xyris tennesseensis, Xyridaceae). *Conservation Genetics*, *21*, 735-746.

Drummond, R. S. M., Keeling, D. J., Richardson, T. E., Gardner, R. C., & Wright, S. D. (2000). Genetic analysis and conservation of 31 surviving individuals of a rare New Zealand tree, Metrosideros bartlettii (Myrtaceae). *Molecular Ecology*, *9*(8), 1149-1157.

Dubuis, A., Giovanettina, S., Pellissier, L., Pottier, J., Vittoz, P., & Guisan, A. (2013). Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables. *Journal of Vegetation Science*, *24*(4), 593-606.

Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007). Can population genetic structure be predicted from life-history traits? *The American Naturalist*, *169*(5), 662-672.

Eaves, S. R., & Brook, M. S. (2021). Glaciers and glaciation of North Island, New Zealand. *New Zealand Journal of Geology and Geophysics*, *64*(1), 1-20. https://doi.org/DOI: 10.1080/00288306.2020.1811354

Eaves, S. R., Mackintosh, A. N., Winckler, G., Schaefer, J. M., Alloway, B. V., & Townsend, D. B. (2016). A cosmogenic 3He chronology of late Quaternary glacier fluctuations in North Island, New Zealand (39 S). *Quaternary Science Reviews*, *132*, 40-56.

Ebach, M. C. (2015). Origins of biogeography. *History, Philosophy and Theory of the Life Sciences*, *13*, 1-20.

Ellis, E. A., Marshall, D. C., Hill, K. B., Owen, C. L., Kamp, P. J., & Simon, C. (2015). Phylogeography of six codistributed New Zealand cicadas and their relationship to multiple biogeographical

boundaries suggest a re-evaluation of the Taupo Line. *Journal of Biogeography*, *42*(9), 1761-1775.
Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics*, *164*(4), 1567-1587. https://doi.org/10.1093/genetics/164.4.1567
Ferreira, D. M. C., Palma-Silva, C., Néri, J., Medeiros, M. C. M. P. d., Pinangé, D. S., Benko-Iseppon, A. M., & Louzada, R. B. (2020). Population genetic structure and species delimitation in the Cryptanthus zonatus complex (Bromeliaceae). *Botanical Journal of the Linnean Society*, 1-18. https://doi.org/10.1093/botlinnean/boaa094
Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., 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*(69).
Franklin, J. (2023). Species distribution modelling supports the study of past, present and future biogeographies. *Journal of Biogeography*.
Franklin, J., Davis, F. W., Ikegami, M., Syphard, A. D., Flint, L. E., Flint, A. L., & Hannah, L. (2013). Modeling plant species distributions under future climates: how fine scale do climate projections need to be? *Global change biology*, *19*(2), 473-483.
Gardner, R. C., De Lange, P. J., Keeling, D. J., Bowala, T., Brown, H. A., & Wright, S. D. (2004). A late Quaternary phylogeography for Metrosideros (Myrtaceae) in New Zealand inferred from chloroplast DNA haplotypes. *Biological Journal of the Linnean Society*, *83*(3), 399-412.
Gardner, R. O. (1997). Macropiper (Piperaceae) in the south-west Pacific. *New Zealand Journal of Botany*, *35*(3), 293-307.
Gasca-Pineda, J., Gutiérrez-Guerrero, Y. T., Aguirre-Planter, E., & Eguiarte, L. E. (2020). The role of environment, local adaptation, and past climate fluctuation on the amount and distribution of genetic diversity in two subspecies of Mexican wild Zea mays. *American Journal of Botany*, *107*(11), 1542-1554.
Gebremedhin, B., Ficetola, G. F., Naderi, S., Rezaei, H. R., Maudet, C., Rioux, D., Luikart, G., Flagstad, Ø., Thuiller, W., & Taberlet, P. (2009). Frontiers in identifying conservation units: From neutral markers to adaptive genetic variation. *Animal Conservation*, *12*(2), 107-109.
Gemmell, M. R., Shepherd, L. D., Zuccarello, G. C., & Perrie, L. R. (2022). Phylogeography of the widespread New Zealand tree lancewood/horoeka (Pseudopanax crassifolius; Araliaceae). *New Zealand Journal of Botany*, 1-17.
Gottlieb, L. (1977). Electrophoretic evidence and plant systematics. *Annals of the Missouri Botanical Garden*, 161-180.
Grünwald, N. J., Everhart, S. E., Knaus, B. J., & Kamvar, Z. N. (2017). Best Practices for Population Genetic Analyses. *Phytopathology®*, *107*(9), 1000-1010. https://doi.org/10.1094/phyto-12-16-0425-rvw
Haase, P. (1992). Isozyme variability and biogeography of Nothofagus truncata (Fagaceae). *New Zealand Journal of Botany*, *30*(3), 315-328.
Habel, J. C., Zachos, F. E., Dapporto, L., Roedder, D., Radespiel, U., Tellier, A., & Schmitt, T. (2015). Population genetics revisited–towards a multidisciplinary research field. *Biological Journal of the Linnean Society*, *115*(1), 1-12.
Hamrick, J., Linhart, Y., & Mitton, J. (1979). Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. *Annual Review of Ecology and Systematics*, *10*(1), 173-200.

Hamrick, J. L., & Godt, M. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1345), 1291-1298.

Heenan, P., Mitchell, C., & Houliston, G. (2018). Genetic Variation and Hybridisation among Eight Species of kōwhai (Sophora: Fabaceae) from New Zealand Revealed by Microsatellite Markers. *Genes*, *9*(2), 111.

Heenan, P. B., Houliston, G. J., & Mitchell, C. M. (2021). Microsatellite markers for Gingidia enysii and G. baxterae (Apiaceae) and their taxonomic utility. *New Zealand Journal of Botany*.

Heenan, P. B., McGlone, M. S., Mitchell, C. M., Cheeseman, D. F., & Houliston, G. J. (2022). Genetic variation reveals broad-scale biogeographic patterns and challenges species' classification in the Kunzea ericoides (kānuka; Myrtaceae) complex from New Zealand. *New Zealand Journal of Botany*, *60*(1), 2-26.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), 907-913.

Hodge, S., Barron, M., & Wratten, S. D. (2000). Induced defences in kawakawa (Macropiper excelsum): do caterpillars avoid previous leaf damage? *New Zealand Journal of Ecology*, *24*(1), 91-95.

Hodge, S., Keesing, V., Wratten, S. D., Lovei, G. L., Palmer, J., & Cilgi, T. (1998). Herbivore damage and leaf loss in the New Zealand pepper tree ('kawakawa', Macropiper excelsum, Piperaceae). *New Zealand Journal of Ecology*, *22*(2), 173-180.

Hodge, S., Keesing, V. F., & Wratten, S. D. (2000). Leaf damage does not affect leaf loss or chlorophyll content in the New Zealand pepper tree, kawakawa (Macropiper excelsum). *New Zealand Journal of Ecology*, *24*(1), 87-89.

Holliday, J. A., Hallerman, E. M., & Haak, D. C. (2018). *Genotyping and Sequencing Technologies in Population Genetics and Genomics*. Springer, Cham.

Holzapfel, S., Faville, M. Z., & Gemmill, C. E. C. (2002). Genetic variation of the endangered holoparasite Dactylanthus taylorii (Balanophoraceae) in New Zealand. *Journal of Biogeography*, *29*(5-6), 663-676. https://doi.org/https://doi.org/10.1046/j.1365-2699.2002.00715.x

Hopkins, E., Al-Yahyai, R., Al-Sadi, A. M., & Al-Subhi, A. (2021). Population structure of two morphotypes of *Sideroxylon mascatense* (A.DC.) T.D.Penn. in Oman. *Genetic Resources and Crop Evolution*, *68*, 1299-1308.

Houghton, B., Wilson, C., McWilliams, M., Lanphere, M., Weaver, S., Briggs, R., & Pringle, M. (1995). Chronology and dynamics of a large silicic magmatic system: Central Taupo Volcanic Zone, New Zealand. *Geology*, *23*(1), 13-16.

Houliston, G. J., Dawson, M. I., Lange, P. J. D., & Heenan, P. B. (2012). Using AFLP markers to inform population management of the endemic Chatham Island toetoe, Austroderia turbaria (Poaceae). *Pacific Conservation Biology*, *18*(1), 33-40.

Hu, F. S., Hampe, A., & Petit, R. J. (2009). Paleoecology meets genetics: deciphering past vegetational dynamics. *Frontiers in Ecology and the Environment*, *7*(7), 371-379.

Hugall, A., Moritz, C., Moussalli, A., & Stanisic, J. (2002). Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail Gnarosophia bellendenkerensis (Brazier 1875). *Proceedings of the National Academy of Sciences*, *99*(9), 6112-6117.

Hughes, P. D., Gibbard, P. L., & Ehlers, J. (2013). Timing of glaciation during the last glacial cycle: evaluating the concept of a global 'Last Glacial Maximum'(LGM). *Earth-Science Reviews*, *125*, 171-198.

Huson, D. H. (1998). SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics*, *14*(1), 68-73.
Jayaprakash, R., Ramzan, F., Miles-Chan, J. L., Foster, M., Mithen, R. F., & Pook, C. (2022). Exploring the Chemical Space of Kawakawa Leaf (Piper excelsum). *Nutrients*, *14*(23), 5168.
Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics*, *11*(94).
Kamp, P. J. (1986). The mid-Cenozoic Challenger Rift System of western New Zealand and its implications for the age of Alpine fault inception. *Geological Society of America Bulletin*, *97*(3), 255-281.
Karlin, E. F., Boles, S. B., & Shaw, A. J. (2008). Systematics of Sphagnum section Sphagnum in New Zealand: A microsatellite-based analysis. *New Zealand Journal of Botany*, *46*(2), 105-118.
King, P. R. (2000). Tectonic reconstructions of New Zealand: 40 Ma to the present. *New Zealand Journal of Geology and Geophysics*, *43*(4), 611-638.
Köhler, A., & Schmitt, T. (2023). Northern Richness, Southern Dead End—Origin and Dispersal Events of Pseudolycoriella (Sciaridae, Diptera) between New Zealand's Main Islands. *Insects*, *14*(6), 548.
Koia, J. H., & Shepherd, P. (2020). The Potential of Anti-Diabetic Rakau Rongoa (Maori Herbal Medicine) to Treat Type 2 Diabetes Mellitus (T2DM) Mate Huka: A Review. *Front Pharmacol*, *11*, 935. https://doi.org/10.3389/fphar.2020.00935
Kraus, R. H. S., Holdt, B. v., Cocchiararo, B., Harms, V., Bayerl, H., Kühn, R., Förster, D. W., Fickel, J., Roos, C., & Nowak, C. (2015). A single-nucleotide polymorphism-based approach for rapid and cost-effective genetic wolf monitoring in Europe based on noninvasively collected samples. *Molecular Ecology Resources*, *15*(2).
Kruckeberg, A. R. (2004). *Geology and plant life: the effects of landforms and rock types on plants*. University of Washington Press.
Lakeman-Fraser, P., & Ewers, R. M. (2013). Enemy release promotes range expansion in a host plant. *Oecologia*, *172*(4), 1203-1212. https://doi.org/10.1007/s00442-012-2555-x
Lambert, S. J., & Mark-Shadbolt, M. (2021). Indigenous biosecurity: Past, present and future. In *Routledge handbook of biosecurity and invasive species* (pp. 55-65). Routledge.
Landis, C., Campbell, H., Begg, J., Mildenhall, D., Paterson, A., & Trewick, S. (2008). The Waipounamu Erosion Surface: questioning the antiquity of the New Zealand land surface and terrestrial fauna and flora. *Geological magazine*, *145*(2), 173-197.
Lei, J., Burgess, E. J., Richardson, A. T., Hawkins, B. C., Baird, S. K., Smallfield, B. M., van Klink, J. W., & Perry, N. B. (2015). Cytotoxic Amides from Fruits of Kawakawa, Macropiper excelsum. *Planta Med*, *81*(12-13), 1163-1168. https://doi.org/10.1055/s-0035-1546106
Lemopoulos, A., Prokkola, J. M., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Hyvärinen, P., Koljonen, M. L., Koskiniemi9, J., & Vainikka, A. (2019). Comparing RADseq and microsatellites for estimating genetic diversity and relatedness — Implications for brown trout conservation. *Ecology and Evolution*, *9*(4), 2106–2120.
Levinson, G., & Gutman, G. A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*, *4*(3), 203-221.
Lloyd, B. D. (2003). Intraspecific Phylogeny of the New Zealand short-tailed bat Mystacina tuberculata inferred from multiple mitochondrial gene sequences. *Systematic Biology*, *52*, 460-476.
Loveless, M. D., & Hamrick, J. L. (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, *15*(1), 65-95.

Lynch, J. P. (2022). Edaphic stress interactions: Important yet poorly understood drivers of plant production in future climates. *Field Crops Research*, *283*, 108547.
Manville, V., Segschneider, B., Newton, E., White, J. D. L., Houghton, B. F., & Wilson, C. J. N. (2009). Environmental impact of the 1.8 ka Taupo eruption, New Zealand: Landscape responses to a large-scale explosive rhyolite eruption. *Sedimentary Geology*, *220*(318-336).
Manville, V., & Wilson, C. J. (2004). The 26.5 ka Oruanui eruption, New Zealand: A review of the roles of volcanism and climate in the post-eruptive sedimentary response. *New Zealand Journal of Geology and Geophysics*, *47*(3), 525-547.
Mao, X., Augyte, S., Huang, M., Hare, M. P., Bailey, D., Umanzor, S., Marty-Rivera, M., Robbins, K. R., Yarish, C., & Lindell, S. (2020). Population genetics of sugar kelp throughout the Northeastern United States using genome-wide markers. *Frontiers in Marine Science*, *7*, 694.
Marshall, D. C., Hill, K. B., Marske, K. A., Chambers, C., Buckley, T. R., & Simon, C. (2012). Limited, episodic diversification and contrasting phylogeography in a New Zealand cicada radiation. *BMC Evolutionary Biology*, *12*(1), 1-18.
Marske, K. A., Leschen, R. A., & Buckley, T. R. (2011). Reconciling phylogeography and ecological niche models for New Zealand beetles: looking beyond glacial refugia. *Molecular Phylogenetics and Evolution*, *59*(1), 89-102.
Matala, A. P., Ackerman, M. W., Campbell, M. R., & Narum, S. R. (2014). Relative contributions of neutral and non-neutral genetic differentiation to inform conservation of steelhead trout across highly variable landscapes. *Evolutionary Applications*, *7*(6), 682-701.
McGlone, M., Duncan, R., & Heenan, P. (2001). Endemism, Species Selection and the Origin and Distribution of the Vascular Plant Flora of New Zealand. *Journal of Biogeography*, 199-216.
McGlone, M. S. (1985). Plant biogeography and the late Cenozoic history of New Zealand. *New Zealand Journal of Botany*, *23*(4), 723-749.
McGlone, M. S., Newnham, R. M., & Moar, N. T. (2010). The vegetation cover of New Zealand during the Last Glacial Maximum: do pollen records under-represent woody vegetation. *Terra Australis*, *32*, 49-68.
McLay, T. G., Tate, J. A., Gemmill, C. E., Holzapfel, A. S., & Symonds, V. V. (2022). Inferring the biogeography of New Zealand's only endemic holoparasitic plant, the threatened Dactylanthus taylorii (Mystropetalaceae). *New Zealand Journal of Botany*, 1-23.
Merrett, M. F., Robertson, A. W., & Peterson, P. G. (2007). Pollination performance and vulnerability to pollination breakdown of sixteen native shrub species from New Zealand. *New Zealand Journal of Botany*, *45*(4), 579-591. https://doi.org/https://doi.org/10.1080/00288250709509740
Meudt, H. M., Prebble, J. M., & Lehnebach, C. A. (2015). Native New Zealand forget-me-nots (Myosotis, Boraginaceae) comprise a Pleistocene species radiation with very low genetic divergence. *Plant Systematics and Evolution*, *301*, 1455-1471.
Mildenhall, D., Mortimer, N., Bassett, K., & Kennedy, E. (2014). Oligocene paleogeography of New Zealand: maximum marine transgression. *57*(2), 107-109.
Mix, A. C., Bard, E., & Schneider, R. (2001). Environmental processes of the ice age: land, oceans, glaciers (EPILOG). *Quaternary Science Reviews*, *20*(4), 627-657.
Moar, N., Suggate, R. P., & Burrows, C. (2008). Environments during the Kaihinu Interglacial and Otira Glaciation, coastal north Westland, New Zealand. *New Zealand Journal of Botany*, *46*(1), 49-63.

Montarry, J., Jan, P. L., Gracianne, C., Overall, A. D., Bardou-Valette, S., Olivier, E., Fournet, S., Grenier, E., & Petit, E. J. (2015). Heterozygote deficits in cyst plant-parasitic nematodes: Possible causes and consequences. *Molecular Ecology*, *24*(8), 1654-1677.

Morgan-Richards, M., Trewick, S. A., & Wallis, G. P. (2001). Chromosome races with Pliocene origins: evidence from mtDNA. *Heredity*, *86*(3), 303-312.

Morgan-Richards, M., Trewick, S. A., & Stringer, I. A. (2010). Geographic parthenogenesis and the common tea-tree stick insect of New Zealand. *Molecular Ecology*, *19*(6), 1227-1238.

Myers, A. A., & Giller, P. (2013). *Analytical biogeography: an integrated approach to the study of animal and plant distributions*. Springer Science & Business Media.

Newnham, R., McGlone, M., Moar, N., Wilmshurst, J., & Vandergoes, M. (2013). The vegetation cover of New Zealand at the last glacial maximum. *Quaternary Science Reviews*, *74*, 202-214.

Newnham, R. M., Lowe, D. J., & Williams, P. W. (1999). Quaternary environmental change in New Zealand: a review. *Progress in Physical Geography*, *23*(4), 567-610.

Nishida, R., Bowers, W. S., & Evans, P. H. (1983). Juvadecene: Discovery of a juvenile hormone mimic in the plant, Macropiper Excelsum. *Archives of Insect Biochemistry and Physiology*, *1*(1), 17-24. https://doi.org/10.1002/arch.940010103

Novembre, J. (2016). Pritchard, Stephens, and Donnelly on Population Structure. *Genetics*, *204*(2), 391-393. https://doi.org/10.1534/genetics.116.195164

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

Obst, K., Lieder, B., Reichelt, K. V., Backes, M., Paetz, S., Geissler, K., Krammer, G., Somoza, V., Ley, J. P., & Engel, K. H. (2017). Sensory active piperine analogues from Macropiper excelsum and their effects on intestinal nutrient uptake in Caco-2 cells. *Phytochemistry*, *135*, 181-190. https://doi.org/10.1016/j.phytochem.2016.12.016

Painting, C. J., Myers, S., Holwell, G. I., & Buckley, T. R. (2017). Phylogeography of the New Zealand giraffe weevil Lasiorhynchus barbicornis (Coleoptera: Brentidae): A comparison of biogeographic boundaries. *Biological Journal of the Linnean Society*, *122*(1), 13-28.

Parida, S. K., Kalia, S. K., Kaul, S., Dalal, V., Hemaprabha, G., Selvi, A., Pandit, A., Singh, A., Gaikwad, K., & Sharma, T. R. (2009). Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theoretical and Applied Genetics*, *118*, 327-338.

Park, G., Choi, Y., Jung, J.-K., Shim, E.-J., Kang, M.-y., Sim, S.-C., Chung, S.-M., Lee, G. P., & Park, Y. (2021). Genetic Diversity Assessment and Cultivar Identification of Cucumber (Cucumis sativus L.) Using the Fluidigm Single Nucleotide Polymorphism Assay. *Plants*, *10*(395).

Pearse, D. E., & Crandall, K. A. (2004). Beyond FST: Analysis of population genetic data for conservation. *Conservation Genetics*, *5*, 585-602.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945-959. https://doi.org/10.1093/genetics/155.2.945

Rajakaruna, N. (2004). The edaphic factor in the origin of plant species. *International Geology Review*, *46*(5), 471-478.

Ramzan, F., Jayaprakash, R., Pook, C., Foster, M., Miles-Chan, J. L., & Mithen, R. (2022). Acute Effects of Kawakawa (Piper excelsum) Intake on Postprandial Glycemic and Insulinaemic Response in a Healthy Population. *Nutrients*, *14*(8), 1638.

Richardson, A. T. B. (2015). Hot new chemistry from a kiwi pepper tree. *Chemistry in New Zealand*, *79*, 91-93.

Ringnér, M. (2008). What is principal component analysis? *Nature Biotechnology*, *26*, 303-304.

Rojas, M., Moreno, P., Kageyama, M., Crucifix, M., Hewitt, C., Abe-Ouchi, A., Ohgaito, R., Brady, E. C., & Hope, P. (2009). The Southern Westerlies during the last glacial maximum in PMIP2 simulations. *Climate Dynamics*, *32*(4), 525-548.

Russell, G. B., & Fenemore, P. G. (1973). New lignans from leaves of Macropiper excelsum. *Phytochemistry*, *12*(7), 1799-1803. https://doi.org/10.1016/0031-9422(73)80407-8

Sefc, K. M., Payne, R. B., & Sorenson, M. D. (2007). Genetic differentiation after founder events: an evaluation of F(ST) estimators with empirical and simulated data. *Evolutionary Ecological Research*, *9*, 21-39.

Semagn, K., Bjørnstad, Å., & Ndjiondjop, M. (2006). An overview of molecular marker methods for plants. *African journal of biotechnology*, *5*(25).

Shelley, J. J., David, B. O., Thacker, C. E., Hicks, A. S., Jarvis, M. G., & Unmack, P. J. (2020). Phylogeography of the Cran's bully Gobiomorphus basalis (Gobiiformes: Eleotridae) and an analysis of species boundaries within the New Zealand radiation of Gobiomorphus. *Biological Journal of the Linnean Society*, *130*(2), 365-381. https://doi.org/10.1093/biolinnean/blaa052

Shepherd, L. D., Bulgarella, M., & de Lange, P. J. (2018). Genetic structuring of the coastal herb Arthropodium cirratum (Asparagaceae) is shaped by low gene flow, hybridization and prehistoric translocation. *Plos One*, *13*(10), e0204943.

Shepherd, L. D., de Lange, P. J., Perrie, L. R., & Heenan, P. B. (2017). Chloroplast phylogeography of New Zealand Sophora trees (Fabaceae): extensive hybridization and widespread Last Glacial Maximum survival. *Journal of Biogeography*, *44*(7), 1640-1651.

Shepherd, L. D., Frericks, J., Biggs, P. J., & de Lange, P. J. (2019). Phylogeography of the endemic New Zealand tree Entelea arborescens (whau; Malvaceae). *New Zealand Journal of Botany*, *57*(3), 154-168.

Shepherd, L. D., & Lambert, D. M. (2008). Ancient DNA and conservation: lessons from the endangered kiwi of New Zealand. *Molecular Ecology*, *17*, 2174-2184.

Shepherd, L. D., & Perrie, L. R. (2011). Microsatellite DNA analyses of a highly disjunct New Zealand tree reveal strong differentiation and imply a formerly more continuous distribution. *Molecular Ecology*, *20*(7), 1389-1400. https://doi.org/https://doi.org/10.1111/j.1365-294X.2011.05017.x

Shepherd, L. D., Perrie, L. R., & Brownsey, P. J. (2007). Fire and ice: volcanic and glacial impacts on the phylogeography of the New Zealand forest fern Asplenium hookerianum. *Molecular Ecology*, *16*(21), 4536-4549. https://doi.org/https://doi.org/10.1111/j.1365-294X.2007.03451.x

Smissen, R. D., & Heenan, P. B. (2010). A taxonomic appraisal of the Chatham Islands flax (Phormium tenax) using morphological and DNA fingerprint data. *Australian Systematic Botany*, *23*(5), 371-380. https://doi.org/https://doi.org/10.1071/SB10023

Smissen, R. D., Richardson, S. J., Morse, C. W., & Heenan, P. B. (2014). Relationships, gene flow and species boundaries among New Zealand Fuscospora (Nothofagaceae: southern beech). *The New Zealand Journal of Botany*, *52*(4), 389-406.

Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, *15*(14), 4261-4293.

Southey, I. C. (1985). *The ecology of three rare skinks on Middle Island, Mercury Island* [Unpublished thesis]. University of Auckland.

Stevens, M., Clarke, A., Clarkson, F., Goshorn, M., & Gemmill, C. (2015). Are current ecological restoration practices capturing natural levels of genetic diversity? A New Zealand case

study using AFLP and ISSR data from mahoe (Melicytus ramiflorus). *New Zealand Journal of Ecology*, *39*(2), 190-197.
Stoeckel, S., Grange, J., Fernandez-Manjarres, J. F., Bilger, I., Frascaria-Lacoste, N., & Mariette, S. (2006). Heterozygote excess in a self-incompatible and partially clonal forest tree species—Prunus avium L. *Molecular Ecology*, *15*(8), 2109-2118.
Summerell, B. A., Laurence, M. H., Liew, E. C., & Leslie, J. F. (2010). Biogeography and phylogeography of Fusarium: a review. *Fungal Diversity*, *44*(1), 3-13.
Thuiller, W., Lavorel, S., Araújo, M. B., Sykes, M. T., & Prentice, I. C. (2005). Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences*, *102*(23), 8245-8250.
Trewick, S., & Bland, K. (2012). Fire and slice: palaeogeography for biogeography at New Zealand's North Island/South Island juncture. *Journal of the Royal Society of New Zealand*, *42*(3), 153-183.
Trewick, S. A. (2000). Mitochondrial DNA sequences support allozyme evidence for cryptic radiation of New Zealand Peripatoides (Onychophora). *Molecular Ecology*, *9*, 269-281.
Trewick, S. A., Wallis, G. P., & Morgan-Richards, M. (2011). The invertebrate life of New Zealand: a phylogeographic approach. *Insects*, *2*(3), 297-325.
Tsykun, T., Rellstab, C., Dutech, C., Sipos, G., & Prospero, S. (2017). Comparative assessment of SSR and SNP markers for inferring the population genetic structure of the common fungus Armillaria cepistipes. *Hereditary*, *119*, 371-380.
Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3), 535-538.
Wagstaff, S. J., & Clarkson, B. R. (2012). Systematics and ecology of the Australasian genus Empodisma (Restionaceae) and description of a new species from peatlands in northern New Zealand. *PhytoKeys*(13), 39.
Wallis, G. P., & Trewick, S. A. (2009). New Zealand phylogeography: evolution on a small continent. *Molecular Ecology*, *18*(17), 3548-3580.
Wardle, P. (1963). Evolution and distribution of the New Zealand flora, as affected by Quaternary climates. *New Zealand Journal of Botany*, *1*(1), 3-17.
Wardle, P. (1988). Effects of glacial climates on floristic distribution in New Zealand 1. A review of the evidence. *New Zealand Journal of Botany*, *26*(4), 541-555.
Waters, J. M., & Craw, D. (2006). Goodbye Gondwana? New Zealand biogeography, geology, and the problem of circularity. *Systematic Biology*, *55*(2), 351-356.
Webb, C. J., Lloyd, D. G., & Delph, L. F. (1999). Gender dimorphism in indigenous New Zealand seed plants. *New Zealand Journal of Botany*, *37*(1), 119-130. https://doi.org/https://doi.org/10.1080/0028825X.1999.9512618
Wehi, P. M., & Clarkson, B. D. (2007). Biological flora of New Zealand 10. Phormium tenax, harakeke, New Zealand flax. *New Zealand Journal of Botany*, *45*(4), 521-544.
Whitaker, A. (1968). The lizards of the Poor Knights Islands, New Zealand. *New Zealand journal of science*, *11*, 623-651.
Whitaker, A. H. (1987). The roles of lizards in New Zealand plant reproductive strategies. *New Zealand Journal of Botany*, *25*(2), 315-328. https://doi.org/https://doi.org/10.1080/0028825X.1987.10410078
Wigginton, J. E., Cutler, D. J., & Abecasis, G. R. (2005). A note on exact tests of Hardy-Weinberg equilibrium. *The American Journal of Human Genetics*, *76*(5), 887-893.

Willyard, A., Bower, A., Hipkins, V., Snelling, J., & DeWoody, J. (2020). Genetic diversity and population structure in Chrysolepis chrysophylla (golden chinquapin; Fagaceae): SSRs vs SNPs. *Canadian Journal of Forest Research*, *50*(8), 788-799. https://doi.org/10.1139/cjfr-2020-0009

Wilmshurst, J. M., & McGlone, M. S. (1996). Forest disturbance in the central North Island, New Zealand, following the 1850 BP Taupo eruption. *The holocene*, *6*(4), 399-411.

Woodward, F. I., & Williams, B. (1987). Climate and plant distribution at global and local scales. *Vegetatio*, *69*, 189-197.

Wotton, D. M., Drake, D. R., Powlesland, R. G., & Ladley, J. J. . (2016). The role of lizards as seed dispersers in New Zealand. *Journal of the Royal Society of New Zealand*, *46*(1), 40-65.

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, *16*(2), 97-159.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 139-156.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, *15*(4), 323-354.

Wright, S. A., Hutchison, M., Hale, M. L., Gemmill, C. E. C., Lange, P. J. d., & Pelser, P. B. (2017). A preliminary conservation genetic study of Pittosporum obcordatum (Pittosporaceae), an endemic New Zealand species with a disjunct distribution. *New Zealand Journal of Botany*, *55*(4), 424-438.

Zhang, J., Yang, J., Zhang, L., Luo, J., Zhao, H., Zhang, J., & Wen, C. (2020). A new SNP genotyping technology Target SNP-seq and its application in genetic analysis of cucumber varieties. *Scientific reports*, *10*(1), 5623-5634.

Chapter 2: Microsatellite marker development

2.1 Abstract

Piper excelsum subsp. excelsum (G. Forst.) (Piperaceae) is a native, lowland, and culturally important species that is common in New Zealand forests. The phytochemical profile and herbivore to plant interaction has been examined for *P. excelsum,* but no studies have been completed investigating genetic variation within the species or how that variation is distributed across the landscape. Population genetics studies provide a framework to evaluate genetic diversity and structure within a species using molecular markers. Microsatellite markers are commonly used in population genetics due to their highly variable nature and have been developed in this study using standard methodology to investigate population genetics parameters for *P. excelsum*. Thirty-seven novel primer pairs were developed, and of these, 10 were selected for consistent amplification, scorablility, and polymorphic properties. These 10 markers were trialled on 15 samples from four geographically dispersed sites on the North Island. Allelic richness was found to be relatively low in the trial data set, with 1.83 average alleles per locus per sampling site. The average observed (0.31) and expected (0.26) heterozygosity also indicated potentially low diversity levels compared to average values from a review of over 200 species in the Magnoliid clade, other New Zealand native species, and species with similar life history traits to *P. excelsum*. These polymorphic markers can subsequently be used to examine genetic variation, population structure, and to explore intraspecific biogeographical patterns in *P. excelsum*.

2.2 Introduction

Piper excelsum subsp. *excelsum* (kawakawa or pepper tree) belongs to the Piperaceae family and is a native, understory tree that is abundant in New Zealand forests (Butts et al., 2019; Hodge et al., 1998). *P. excelsum* is common in lowland mixed angiosperm-podocarp forests (Burrows, 1995), has a preference for damp soil and is naturally found around the entire coastline of the North Island and the South Island as far south as Banks Peninsula (de Lange, 2021), as well as on several of New Zealand's offshore islands (Lakeman-Fraser & Ewers, 2013). The species has an extensive phytochemical profile (Jayaprakash et al., 2022; Lei et al., 2015; Obst et al., 2017) and is typically recognised by its heart shaped leaves that often have many distinct holes created by the herbivore *Cleora scriptaria* (Hodge, Barron, et al., 2000; Hodge et al., 1998). Kawakawa is a plant of significance

in traditional Māori culture, and has many historical uses; e.g., rākau rongoā (plant based remedies) for ailments such as toothache, eczema and kidney problems (Brooker & Cooper, 1961). The leaves are also used in wreaths at tangihanga (Koia & Shepherd, 2020). Recently, there has been rising interest in the commercialisation of *P. excelsum,* particularly in medicine, food and beverages areas (Butts et al., 2019; Jayaprakash et al., 2022). Due to the abundance (Hodge et al., 1998) and widespread distribution of *P. excelsum*, a population genetics study provides valuable information for better understanding widescale biogeographic patterns in native New Zealand flora.

Currently, there are a limited number of population genetic and biogeographic studies on native plants in New Zealand. Population genetics studies provide insight into biogeographical studies which examine patterns of species distribution, divergence and investigate the potential of geological or ecological processes (Ebach, 2015; Shelley et al., 2020). This is a particularly difficult concept in New Zealand due to an eventful geological history (Wallis & Trewick, 2009). Studying native species with a widespread distribution is especially valuable for investigating widescale biogeographic patterns, although evidence is limited to a few studies (Beever et al., 2013; Buys et al., 2019; Gemmell et al., 2022; Shepherd et al., 2017; Shepherd et al., 2007). This leaves an investigable gap around understanding population structure, diversity, and biogeographical patterns of native species in New Zealand.

Microsatellite markers are highly polymorphic, co-dominant, and have a high mutation rate (Kalia et al., 2011); these characteristics are beneficial for population genetic studies as they provide insight into the genetic variation and population structure of a species (Gebremedhin et al., 2009; Lemopoulos et al., 2019). The high mutation rate of microsatellites is particularly useful for deciphering variation between individuals within a population (Anne, 2006; Selkoe & Toonen, 2006) over other marker types. For example, Random Amplified Polymorphic DNAs (RAPDs) or Amplified Fragment Length Polymorphisms (AFLPs) are unable to provide an estimation of allele frequencies (Ouborg et al., 1999), the chloroplast genome has a slower mutation rate (Provan et al., 2001) making them less suitable to assess contemporary gene flow and SNPs have high costs for sequencing technology (Willyard et al., 2020) with limited funding and access to the appropriate technology for high-throughput sequencing not available for this study. Here, ten microsatellite markers for *P. excelsum* have been developed for future use to investigate population genetic parameters and explore the geographic distribution of genetic variation. Ultimately, this work will

contribute to a better understanding of the broad impacts of New Zealand's geological history on the distribution of native biota.

2.3 Methods

2.3.1 Sampling and DNA extraction

Fifteen *P. excelsum* samples were collected from four sites across the North Island. At each site, 1-2 leaves from 20 to 25 plants were collected to represent each location and three or four of these samples were randomly chosen for trialling. The leaves were stored in labelled coffee filters and placed in bags of silica gel for desiccation for at least seven days. DNA was extracted following the STE CTAB method (Shepherd & McLay, 2011) with a modification of 1% CTAB instead of 2%. DNA samples were visualized on a 1% TBE agarose gel stained with ethidium bromide.

2.3.2 Microsatellite marker development

To isolate microsatellite loci to develop as markers, DNA from a single *P. excelsum* sample was randomly fragmented to sizes ranging from 200-500bp using a sonicator. Following this, the adapters with individual-specific barcodes and Illumina sequencing sites were ligated to the ends to construct a genome sequence library. In combination with two prepared libraries of other projects, the *P. excelsum* library was pooled in equal amounts and submitted for paired-end (2 x 250bp) sequencing on an Illumina Miseq at the Massey Genome Service facility with ~eight million raw reads (~four million paired reads) produced as a result.

To develop microsatellite markers, Geneious v9.05 software (Kearse et al., 2012) was used to produce merged sequences from the cleaned paired-end sequences. The merged sequence file was converted to FASTA format and uploaded to the Galaxy web platform (Jalili et al., 2020). The STR detection function was used in Galaxy to search for sequences that possessed microsatellite loci that met the following criteria: all di- and trinucleotide repeats, 8-12 uninterrupted repeats, and at least 70 bp of sequence on each side of the repeat region. Those sequences were then re-uploaded to Geneious v9.05 for primer design. Thirty seven primer pairs were developed in Geneious utilising primer3 v2.3.4 (Rozen & Skaletsky, 2000) with the following characteristics: primers >50 bp from repeat region, minimum (150bp) and maximum product size (380bp), primer length of 18-22, optimal GC content of 50% and G/C clamp of 1 bp on the 3' end. For all markers, a 19bp M13

sequence was added to the 5’ end of the forward primer, and a 7bp pigtail sequence was attached to the 5’ end of the reverse primer (Boutin-Ganache et al., 2001).

Twenty-five of these primer pairs were piloted on material collected from the Manawatū area to identify markers that amplified consistently. Further testing was then completed on eight herbarium leaf material samples collected from the Dame Ella Campbell Herbarium (MPN46480, MPN46532, MPN26897, MPN45207, MPN41318, MPN46552, MPN38578, MPN51498) representing a wide geographical range on the North Island in attempt to identify polymorphic markers (fresh widespread field collections had not yet been made). Data is not shown here for these preliminary trials. Following this, nine consistently amplifying markers that presented two or more alleles from the herbarium samples were chosen for a trial on 15 samples collected from four geographically dispersed regions (Figure 2.1). With the aim of developing 10-12 markers, six further primer pairs were chosen comprising a range of product sizes so that several markers could be pool-plexed when genotyping. These primer pairs were also trialled on the same 15 samples.

Figure 2.1: Four sampling sites chosen for initial microsatellite screening are shown here. Site names: TET(14)= Te Tapui, KAR(2)=Kaitake Ranges, TOR(8)=Totara Reserve and AOF(11)=Aorangi Forest Park.

Polymerase Chain Reaction (PCR) was carried out to amplify the microsatellite loci for all trialled DNA samples at a volume of 10 µL. The final reagents and concentrations were ddH_2O, 1x Buffer BD (Solis Biodyne, Tartu Estonia), 2.5mM $MgCL_2$, 20µM dNTP, 200µM forward primer, 450µM reverse primer, 450µM of M13 primer with fluorescent dye (FAM, VIC, or NED) attached, and 0.5U Firepol *Taq* Polymerase (Solis Biodyne, Tartu Estonia), with 1µL of DNA added of each sample. PCR amplification is as follows, three minutes at 95°C, with 30 cycles of 95°C for 30 seconds, 52°C for 30s

and 72°C for 45s followed by a 20-minute hold at 72°C. The PCR products were assessed on a 1.5% agarose gel stained with ethidium bromide.

PCR products from the fifteen trialled markers were pooled, in groups of three or six, and arranged in accordance with non-overlapping expected product size ranges and differentiating fluorescent dyes. 1.25µL of the pooled products were added to 9µL of Hi-Di formamide (Applied Biosystems, Carlsbad, California, USA) and CASS size standard (Symonds & Lloyd, 2004). The products were genotyped on an ABI DNA Analyser 3730 at the Massey Genome Service and the allele sizes were scored in GENEMAPPER (v4.0) (Applied Biosystems) and recorded manually. The number of alleles (Na), observed heterozygosity (Ho) and expected heterozygosity (He) were calculated using GenAlEx v6.5 (Peakall & Smouse, 2006).

2.4 Results

Originally, 37 primer pairs were developed and trialled. Six of the first set of 25 primer pairs amplified well, were easily scorable, and displayed some variation; i.e., two or more alleles. As the aim was to develop 10-12 markers for this project, a second set of 12 primer pairs were considered and a further four loci were found to be polymorphic and amplified consistently. These ten markers were subsequently chosen (Table 2.1) and the trialled results were analysed in GenAlEx v6.5 (Peakall & Smouse, 2006) (Table 2.2). The number of alleles per marker per population ranged from 1 – 3, with an average Na of 1.83. Markers 51, 339876 and 65864 displayed the highest number of alleles in the collection sites of Kaitake Ranges, Aorangi Forest, and Totara Reserve (although, it is acknowledged that the sample sizes for marker development were very small). Observed heterozygosity at several loci across the four sites was 0, with loci 50, 52 and 252751 displaying Ho of 0 for three out of four sites. The highest Ho value was 1 for loci 51 at the Kaitake Ranges site. Expected heterozygosity (He) ranged from 0 at several loci to 0.66 for loci 51 and 339876 at the Kaitake Ranges site. Ho is almost always higher than He in this trial data set, with average Ho (0.31) and average He (0.26), but again, these trial results are based on small sample sizes.

Table 2.1: Microsatellite primer information. Primer sequences, repeat type, product size of the initial sequence, and M13 dye used for amplification.

Locus	Primer sequences	Repeat type	Product size	M13 dye
127770	F: TGCAAGATTCTTCTTCGGG R: CCCCATCATATGTGTCCAC	AG(10)	193	NED
1415	F: CACCCCTGCATAGCCTAAC R: CATGCAAACTGCTAGC	AT(10)	253	VIC
65864	F: TTGTGTTTGACTTGATCGC R: TCCCCACAAACTTTACAGC	CA(21)	286	FAM
339876	F: AGCTTGAAATCGTATGGCC R: TACGTACTTGCCGGTTATG	TA(8)	203	NED
338542	F: AAGCATGAGCCAGAATGTC R: ACCACCAAGAGTCAACAAG	CA(10)	200	FAM
252751	F: TCATCCTGACCCTATCGA R: GCCGTCCTTATGATCTGAC	AC(11)	285	FAM
50	F: TGGTTAATCCTGCAAATGCG R: CACCTTCATCAATGGCTTCC	AT(8)	175	NED
51	F: GCTTGAAATCGTATGGCCTC R: GATGACTTCCAATCACCTGC	TA(8)	329	VIC
52	F: ACGTACTACCCCAGTTGTTC R: TATTGTTCCCAGTCCCATGC	TA(8)	262	FAM
55	F: ACGGTCTTTATTTGTGCGTG R: GCAATCTATGTGGTAGCACC	AT(9)	258	FAM

Table 2.2: Ten selected microsatellite loci across four geographically dispersed *P. excelsum* sampled sites. Sample size (N), number of alleles (Na), observed heterozygosity (Ho) and expected heterozygosity (He). Site codes: Taranaki (TAR), Waikato (WAI), Wellington (WEL) and Manawatū/Whanganui (MNW).

Site	**Kaitake Ranges**				**Te Tapui**				**Aorangi Forest**				**Totara Reserve**			
Code	**(TAR)**				**(WAI)**				**(WEL)**				**(MNW)**			
Locus	**N**	**Na**	**Ho**	**He**	**N**	**Na**	**Ho**	**He**	**N**	**Na**	**Ho**	**He**	**N**	**Na**	**Ho**	**He**
50	4	1	0	0	4	1	0	0	4	2	0.50	0.38	3	1	0	0
51	4	3	1	0.66	4	2	0.25	0.47	4	3	0.75	0.53	3	3	0.67	0.50
52	4	1	0	0	4	2	0.25	0.22	4	1	0	0	3	1	0	0
55	4	2	0.50	0.38	4	2	0.75	0.47	4	2	0.25	0.22	3	1	0	0
252751	4	1	0	0	4	2	0.50	0.38	4	1	0	0	3	1	0	0
1415	4	1	0	0	3	2	0.67	0.44	3	2	0.33	0.28	3	1	0	0
339876	4	3	0.75	0.66	4	2	0.25	0.47	4	3	0.75	0.53	3	3	0.67	0.50
338542	4	2	0.25	0.22	4	2	0.50	0.38	4	2	0.25	0.47	3	1	0	0
65864	4	3	0.50	0.53	4	2	0.25	0.22	4	3	0.50	0.59	3	2	0.33	0.28
127770	4	1	0	0	4	1	0	0	4	2	0.50	0.50	3	2	0.33	0.28

2.5 Discussion

Microsatellite markers are commonly used in population genetics studies as they often mutate at a relatively fast evolutionary rate and are highly polymorphic. Here, ten microsatellite markers were developed that would be useful for undertaking a population genetics study and investigating population structure and genetic diversity in *P. excelsum*. From the 15 trialled samples from four different geographic regions, the highest number of alleles per locus in a sampling site was three, with an average of 1.83 alleles per locus. While these values are low, this may just be a result of the small sample sizes. The within site diversity of Ho (0.31) and He (0.26) for the trialled data was relatively low in comparison to the average Ho (0.58), He (0.49) values for Magnoliids, the clade including the Piperaceae family, from an empirical plant microsatellite review (Merritt et al., 2015). Furthermore, the average expected heterozygosity values for New Zealand native species of genus *Gingidia* (0.522) (Heenan et al., 2021), *Sophora* (0.720) (Van Etten et al., 2014), *Corybas* (0.390) (Van Etten et al., 2018) and other species with similar life history traits to *P. excelsum*, including *Cryptomeria japonica* (Cupressaceae) (0.77) (Takahashi et al., 2005) and *Allocasuarina humilis* (Casuarinaceae) (0.76) (Llorens et al., 2017) were also low in comparison. With this considered, the number of samples trialled here was considerably lower than these studies, as well as other microsatellite marker development studies (Forgiarini et al., 2017; Soares et al., 2012; Thapa et al., 2019) suggesting more samples may be required for a more robust analysis. As these markers will be used for a much larger population genetics study, better comparisons between genetic variation in *P. excelsum* and other taxa will appear in a later chapter. The markers developed here will be suitable for investigating population structure and genetic variation, insight that will be valuable when exploring biogeographical patterns within *P. excelsum*.

2.6 References

Anne, C. (2006). Choosing the right molecular genetic markers for studying biodiversity: from molecular evolution to practical aspects. *Genetica*, *127*(1-3), 101-120.

Beever, R., Houliston, G., & Armstrong, T. (2013). Geographic patterns of genetic variation in Cordyline australis; does fingerprinting recover phenotypic and phenological patterns? *New Zealand Journal of Botany*, *51*(1), 3-12.

Boutin-Ganache, I., Raposo, M., Raymond, M., & Deschepper, C. F. (2001). M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *Biotechniques*, *31*(1), 25-28.

Brooker, S. G., & Cooper, R. C. (1961). New Zealand medicinal plants. *Economic Botany*, *15*(1), 1-10.

Burrows, C. J. (1995). Germination behaviour of seeds of the New Zealand species Fuchsia excorticata, Griselinia littoralis, Macropiper excelsum, and Melicytus ramiflorus. *New Zealand Journal of Botany*, *33*(1), 131-140.

Butts, C. A., Klink, J. W. v., Joyce, N. I., Paturi, G., Hedderley, D. I., Martell, S., & Harvey, D. (2019). Composition and safety evaluation of tea from New Zealand kawakawa (Piper excelsum). *Journal of Ethnopharmacology*, *232*, 110-118.

Buys, M. H., Winkworth, R. C., De Lange, P. J., Wilson, P. G., Mitchell, N., Lemmon, A. R., Moriarty Lemmon, E., Holland, S., Cherry, J. R., & Klápště, J. (2019). The phylogenomics of diversification on an island: applying anchored hybrid enrichment to New Zealand Leptospermum scoparium (Myrtaceae). *Botanical Journal of the Linnean Society*, *191*(1), 1-17.

de Lange, P. J. (2021). *Piper excelsum subsp. excelsum Fact Sheet (content continuously updated)*. New Zealand Plant Conservation Network. Retrieved 01 May from https://www.nzpcn.org.nz/flora/species/piper-excelsum-subsp-excelsum/

Ebach, M. C. (2015). Origins of biogeography. *History, Philosophy and Theory of the Life Sciences*, *13*, 1-20.

Forgiarini, C., Curto, M., Stiehl-Alves, E. M., Bräuchler, C., Kollmann, J., Meimberg, H., & Teixeira de Souza-Chies, T. (2017). Fifteen microsatellite markers for Herbertia zebrina (Iridaceae): An endangered species from South American grasslands. *Applications in Plant Sciences*, *5*(5), 1700035.

Gebremedhin, B., Ficetola, G. F., Naderi, S., Rezaei, H. R., Maudet, C., Rioux, D., Luikart, G., Flagstad, Ø., Thuiller, W., & Taberlet, P. (2009). Frontiers in identifying conservation units: From neutral markers to adaptive genetic variation. *Animal Conservation*, *12*(2), 107-109.

Gemmell, M. R., Shepherd, L. D., Zuccarello, G. C., & Perrie, L. R. (2022). Phylogeography of the widespread New Zealand tree lancewood/horoeka (Pseudopanax crassifolius; Araliaceae). *New Zealand Journal of Botany*, 1-17.

Heenan, P. B., Houliston, G. J., & Mitchell, C. M. (2021). Microsatellite markers for Gingidia enysii and G. baxterae (Apiaceae) and their taxonomic utility. *New Zealand Journal of Botany*.

Hodge, S., Barron, M., & Wratten, S. D. (2000). Induced defences in kawakawa (Macropiper excelsum): do caterpillars avoid previous leaf damage? *New Zealand Journal of Ecology*, *24*(1), 91-95.

Hodge, S., Keesing, V., Wratten, S. D., Lovei, G. L., Palmer, J., & Cilgi, T. (1998). Herbivore damage and leaf loss in the New Zealand pepper tree ('kawakawa', Macropiper excelsum, Piperaceae). *New Zealand Journal of Ecology*, *22*(2), 173-180.

Jalili, V., Afgan, E., Gu, Q., Clements, D., Blankenberg, D., Goecks, J., Taylor, J., & Nekrutenko, A. (2020). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic acids research*, *48*(W1), W395-W402.

Jayaprakash, R., Ramzan, F., Miles-Chan, J. L., Foster, M., Mithen, R. F., & Pook, C. (2022). Exploring the Chemical Space of Kawakawa Leaf (Piper excelsum). *Nutrients*, *14*(23), 5168.

Kalia, R. K., Rai, M. K., Kalia, S., Singh, R., & Dhawan, A. K. (2011). Microsatellite markers: an overview of the recent progress in plants. *Euphytica*, *177*, 309-334.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., & Duran, C. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647-1649.

Koia, J. H., & Shepherd, P. (2020). The Potential of Anti-Diabetic Rakau Rongoa (Maori Herbal Medicine) to Treat Type 2 Diabetes Mellitus (T2DM) Mate Huka: A Review. *Front Pharmacol*, *11*, 935. https://doi.org/10.3389/fphar.2020.00935

Lakeman-Fraser, P., & Ewers, R. M. (2013). Enemy release promotes range expansion in a host plant. *Oecologia*, *172*(4), 1203-1212. https://doi.org/10.1007/s00442-012-2555-x

Lei, J., Burgess, E. J., Richardson, A. T., Hawkins, B. C., Baird, S. K., Smallfield, B. M., van Klink, J. W., & Perry, N. B. (2015). Cytotoxic Amides from Fruits of Kawakawa, Macropiper excelsum. *Planta Med*, *81*(12-13), 1163-1168. https://doi.org/10.1055/s-0035-1546106

Lemopoulos, A., Prokkola, J. M., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Hyvärinen, P., Koljonen, M. L., Koskiniemi9, J., & Vainikka, A. (2019). Comparing RADseq and microsatellites for estimating genetic diversity and relatedness — Implications for brown trout conservation. *Ecology and Evolution*, *9*(4), 2106–2120.

Llorens, T. M., Tapper, S. L., Coates, D. J., McArthur, S., Hankinson, M., & Byrne, M. (2017). Does population distribution matter? Influence of a patchy versus continuous distribution on genetic patterns in a wind-pollinated shrub. *Journal of Biogeography*, *44*(2), 361-374.

Merritt, B. J., Culley, T. M., Avanesyan, A., Stokes, R., & Brzyski, J. (2015). An empirical review: characteristics of plant microsatellite markers that confer higher levels of genetic variation. *Applications in Plant Sciences*, *3*(8), 1500025.

Obst, K., Lieder, B., Reichelt, K. V., Backes, M., Paetz, S., Geissler, K., Krammer, G., Somoza, V., Ley, J. P., & Engel, K. H. (2017). Sensory active piperine analogues from Macropiper excelsum and their effects on intestinal nutrient uptake in Caco-2 cells. *Phytochemistry*, *135*, 181-190. https://doi.org/10.1016/j.phytochem.2016.12.016

Ouborg, N., Piquot, Y., & Van Groenendael, J. (1999). Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology*, *87*(4), 551-568.

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

Provan, J., Powell, W., & Hollingsworth, P. M. (2001). Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution*, *16*(3), 142-147.

Rozen, S., & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods in molecular biology*, *132*, 365-386.

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, *9*(5), 615-629.

Shelley, J. J., David, B. O., Thacker, C. E., Hicks, A. S., Jarvis, M. G., & Unmack, P. J. (2020). Phylogeography of the Cran's bully Gobiomorphus basalis (Gobiiformes: Eleotridae) and an analysis of species boundaries within the New Zealand radiation of Gobiomorphus.

Biological Journal of the Linnean Society, *130*(2), 365-381. https://doi.org/10.1093/biolinnean/blaa052

Shepherd, L. D., de Lange, P. J., Perrie, L. R., & Heenan, P. B. (2017). Chloroplast phylogeography of New Zealand Sophora trees (Fabaceae): extensive hybridization and widespread Last Glacial Maximum survival. *Journal of Biogeography*, *44*(7), 1640-1651.

Shepherd, L. D., & McLay, T. G. (2011). Two micro-scale protocols for the isolation of DNA from polysaccharide-rich plant tissue. *Journal of Plant Research*, *124*(2), 311-314.

Shepherd, L. D., Perrie, L. R., & Brownsey, P. J. (2007). Fire and ice: volcanic and glacial impacts on the phylogeography of the New Zealand forest fern Asplenium hookerianum. *Molecular Ecology*, *16*(21), 4536-4549. https://doi.org/https://doi.org/10.1111/j.1365-294X.2007.03451.x

Soares, T. N., Melo, D. B., Resende, L. V., Vianello, R. P., Chaves, L. J., Collevatti, R. G., & Telles, M. P. d. C. (2012). Development of microsatellite markers for the neotropical tree species Dipteryx alata (Fabaceae). *American Journal of Botany*, *99*(2), e72-e73.

Symonds, V. V., & Lloyd, A. M. (2004). A simple and inexpensive method for producing fluorescently labelled size standard. *Molecular Ecology Notes*, *4*(4), 768-771.

Takahashi, T., Tani, N., Taira, H., & Tsumura, Y. (2005). Microsatellite markers reveal high allelic variation in natural populations of Cryptomeria japonica near refugial areas of the last glacial period. *Journal of Plant Research*, *118*, 83-90.

Thapa, R., Bayer, R. J., & Mandel, J. R. (2019). Development and characterization of microsatellite markers for Antennaria corymbosa (Asteraceae) and close relatives. *Applications in Plant Sciences*, *7*(6), e11268.

Van Etten, M. L., Houliston, G. J., Mitchell, C. M., Heenan, P. B., Robertson, A. W., & Tate, J. A. (2014). Sophora microphylla (Fabaceae) microsatellite markers and their utility across the genus. *Applications in Plant Sciences*, *2*(3), 1300081.

Van Etten, M. L., Lehnebach, C. A., Pearson, S. M., Robertson, A. W., & Tate, J. A. (2018). Microsatellite markers for Corybas (Orchidaceae) species in New Zealand. *Applications in Plant Sciences*, *6*(11), e01192. https://doi.org/https://doi.org/10.1002/aps3.1192

Wallis, G. P., & Trewick, S. A. (2009). New Zealand phylogeography: evolution on a small continent. *Molecular Ecology*, *18*(17), 3548-3580.

Willyard, A., Bower, A., Hipkins, V., Snelling, J., & DeWoody, J. (2020). Genetic diversity and population structure in Chrysolepis chrysophylla (golden chinquapin; Fagaceae): SSRs vs SNPs. *Canadian Journal of Forest Research*, *50*(8), 788-799. https://doi.org/10.1139/cjfr-2020-0009

Chapter 3: Deciphering population structure and biogeographic patterns of *P. excelsum*

3.1 Abstract

The complex geological history of New Zealand provides a fascinating opportunity to investigate biogeographic patterns, genetic diversity, and population structure. To study widescale biogeographic patterns, *Piper excelsum* was selected based on the abundance of the species in mixed angiosperm-podocarp forest in the North Island. Ten microsatellite markers were used to characterise the genetic variation and population structure of 249 plants from 16 sampled sites across the North Island. Although the overall genetic structure was found to be weak, an east-west pattern was uncovered and was hypothesised in this study to be explained by a combination of factors, including LGM microrefugia, repeated volcanic activity across the volcanic plateau, and the axial mountain range running from Wellington to East Cape. Low within site variation was found and is suggested to be a consequence of a mass die-off of frost intolerant *P. excelsum* during the LGM causing a bottleneck with potential persistence only in Northern refugial areas. The finding of low between site variation levels are potentially due to the effective pollen and seed dispersal mechanisms, among other factors, that have contributed to the nearly continuous spatial distribution and promotion of gene flow. This study provides meaningful insight into the genetic diversity and biogeographic patterns within *P. excelsum* and the results are potentially useful for predicting the influence of changing climatic conditions and contributing to the understanding of the broad influence of historic events, such as the LGM on native New Zealand species with similar characteristics to *P. excelsum*, such as lowland forest trees with widespread distribution and frost intolerance.

3.2 Introduction

Biogeography studies aim to understand both within and between species genetic divergence and distributional patterns over time, exploring influential historical, ecological, or geological factors (Ebach, 2015; Shelley et al., 2020). New Zealand has had an eventful geological history, characterised by tectonic activity (Campbell et al., 2012), some degree of submersion by water during the Oligocene (Waters & Craw, 2006), cycles of interglacial and glacial periods during the Pleistocene, the most recent of which being the Last Glacial Maximum (LGM) that occurred approximately 23-19 ka (Hughes et al., 2013) and volcanic activity, particularly in the Taupō Volcanic Zone (TVZ) with the most recent activity being the Taupō eruption 1.8 ka (Alloway et al., 2007). These historic events would have greatly influenced the survival and distribution of native taxa. The LGM rendered continuous podocarp-broadleaf and beech forest only to Northland (Newnham et al., 2013) and small pockets throughout New Zealand (McGlone, 1985), with grassland encompassing the remainder of New Zealand's vegetation at this time (McGlone et al., 2010; Newnham et al., 2013). The Taupō eruption was significant, destroying around 20,00km^2 of native forest (Manville et al., 2009).

Glaciation and volcanic disruptions have also influenced the distribution of genetic variation. The environmental conditions during the LGM were severe, with only the most cold resistant plants likely to have persisted through the frost, apart from on offshore islands, coastal areas, and northern areas (Newnham et al., 2013; Wardle, 1963; Wardle, 1988). Studies of New Zealand species have identified several refugial areas during the LGM, one of which is the upper North Island (Buckley et al., 2010; Buckley et al., 2009; Gardner et al., 2004; Gemmell et al., 2022). Species that persisted through the LGM in northern refugia may have recolonised rapidly throughout the North Island as the climate became less severe (McGlone et al., 2001), with higher levels of genetic diversity generally found in the upper north and less in the south (Buckley et al., 2010). Repeated volcanic activity in the TVZ is often cited as a barrier leading to an east-west pattern of diversification in some North Island species (Holzapfel et al., 2002; McLay et al., 2022; Shepherd et al., 2007). A north-south pattern of decreasing diversity has been noted in native flora in the North Island, suggesting the southern populations have a more recent origin than northern populations or have recolonised the lower North Island from refugial areas following glacial periods (Armstrong & De Lange, 2005; Beever et al., 2013). Therefore, investigating biogeographic patterns is complex in New Zealand as there are many environmental factors that will have influenced spatial and genetic distributions, with recolonisation making these patterns especially difficult to decipher (Wallis & Trewick, 2009).

Climatic, geological and ecological factors all need to be considered when interpreting spatial distribution as a combination of these is likely influencing or has historically influenced species distributions (McGlone et al., 2001). Studying a diverse range of taxa is also imperative to understanding historical environmental influence on species distribution patterns as each species often has a unique combination of life history traits, colonisation and dispersal capability (Petit et al., 2003).

Population genetics studies provide insights into biogeographic patterns. Such studies look at genetic diversity, structure and gene flow (Balloux & Lugon-Moulin, 2002; Habel et al., 2015), where gene flow is driven by pollen and seed dispersal among plant populations (Ennos, 1994). The limited population genetic studies published on New Zealand native species have tended to focus on taxonomic uncertainties or to inform conservation efforts for endangered species. There are very few papers on native widespread species (Beever et al., 2013; Gemmell et al., 2022; Shepherd et al., 2007) which should be particularly useful for investigating widescale biogeographic patterns.

Piper excelsum subsp. *excelsum* (kawakawa or pepper tree) belongs to the Piperaceae family and is a native, and lowland, understory tree that is common in New Zealand forests (Butts et al., 2019; Hodge et al., 1998). *P. excelsum* is found in lowland mixed angiosperm-podocarp forests (Burrows, 1995), is frost intolerant, prefers damp soil and a shaded habitat (Hall, 2020; Koia & Shepherd, 2020) and is naturally found around the coastline of the North Island and on the South Island as far south as the Banks Peninsula (de Lange, 2021), as well as on several of New Zealand's offshore islands (Lakeman-Fraser & Ewers, 2013). The species has interesting phytochemical properties (Jayaprakash et al., 2022; Lei et al., 2015; Obst et al., 2017) and many historical and cultural uses by Māori; e.g., rākau rongoā (plant based remedies) for ailments such as toothache, eczema and kidney problems (Brooker & Cooper, 1961). Additionally, the leaves have been used in wreaths at tangihanga (Koia & Shepherd, 2020). The heart shaped leaves are frequently covered in holes made by *Cleora scriptaria* (Hodge, Barron, et al., 2000; Hodge et al., 1998). Although evidence of characteristics typical of entomophily have been noted, such as short filaments (Gardner, 1997), there is more support towards anemophily for *P. excelsum*, including a high pollen to ovule ratio, smooth surfaced pollen grains, non-apical dehiscence slits and absence of biotic inflorescence visitors (Burger, 1972; Merrett et al., 2007). The seed dispersal mechanism for *P. excelsum* is indicated as animal-ingested, with kereru, silvereyes, blackbirds and bellbirds observed feeding on *P. excelsum* fruit (Burrows, 1995;

Wyman, 2013). The commonness of *P. excelsum* makes it a potentially useful species for contributing knowledge to widescale biogeographic patterns.

This study involves the use of microsatellite markers to investigate the following objectives: (1) evaluate the genetic variation within and between sampled *P. excelsum* sites (2) investigate the population structure of sampled *P. excelsum* sites in the North Island (3) discuss potential explanations, such as historic events, for any observed biogeographic patterns. The results suggest relatively weak genetic structuring and low levels of within site population diversity and among site differentiation. Potential ecological and environmental influences on the spatial distribution of genetic variation are also discussed.

3.3 Methods

3.3.1 Sampling and DNA extraction

Four hundred and fifty-five *P. excelsum* samples were collected from 21 sites across the North Island. Permission for collection was granted through Manaaki Whenua/Landcare Research Limited's Concession (CA-31615-OTH), the Department of Conservation and local Iwi. At each site, 1-2 leaves from each of 20 to 25 plants were collected, notations of coexisting plant species, estimated levels of *C. scriptaria* herbivory, site characteristics, a scaled photograph, and GPS coordinates (Supplementary Table 1) for one individual were noted, where possible. As *P. excelsum* was plentiful at most sites, the plants were collected along two transects to best represent the site. The leaves were stored in labelled coffee filters and placed in sealed bags with silica gel for desiccation for at least seven days. Each site was assigned a number and each individual a separate ID (e.g., 4-3) to reflect the individual identification. DNA was extracted following the STE CTAB method (Shepherd & McLay, 2011) as stated in Chapter 2.

3.3.2 Microsatellite markers and genotyping

Of the 21 sites, two were excluded due to an insufficient number of individuals sampled and three further sites were not used as they were thought not to be relictual. For the remaining 16 sites (Table 3.1), which cover a reasonable geographic range on the North Island (Figure 3.1), 249 *P. excelsum* samples were chosen for microsatellite analyses with approximately 16 samples selected per site; initial analyses indicted that there was relatively little genetic variation within sites, so it was decided to maximize the number of sites. Ten microsatellite loci were chosen to investigate these samples

utilising the PCR and genotyping methodology described in Chapter 2. As the PCR products were found to be consistently strong and of high concentration for all markers, the PCR products were diluted 1/50 prior to pooling. The 10 loci were split into two pools, with a group of six (markers 50, 51, 52, 65864, 1415 and 339876) and four (markers 338542, 55, 252751 and 127770) arranged to avoid expected size overlap and differentiating fluorescent dyes were used. 1.25µL of the pooled products were added to 9µL of Hi-Di formamide (Applied Biosystems, Carlsbad, California, USA) and CASS size standard (Symonds & Lloyd, 2004). The products were genotyped on an ABI DNA Analyser 3730 at the Massey Genome Service and the allele sizes were scored in GENEMAPPER (v4.0) (Applied Biosystems), manually recorded in Microsoft Excel, and processed in GenAlEx v6.5 (Peakall & Smouse, 2006). Individuals with missing data for three or more loci were removed from the data set. Microchecker x2.3.3 (Van Oosterhout et al., 2004) was utilised to test for the presence of null alleles.

Table 3.1: Collection site information for 16 sampled *P. excelsum* locations.

Site number	Name	ID	Region	Code
1	Meeting of the Waters	MEW	Taranaki	TAR
2	Kaitake Ranges	KAR	Taranaki	TAR
3	Orokawa	ORK	Bay of Plenty	BOP
4	Aongatete	AOT	Bay of Plenty	BOP
7	Bruce Park	BRP	Manawatu-Wanganui	MNW
8	Totara Reserve	TOR	Manawatu-Wanganui	MNW
11	Aorangi Forest	AOF	Wellington	WEL
12	Remutaka Forest	REM	Wellington	WEL
14	Te Tapui	TET	Waikato	WAI
16	Atene Skyline	ATS	Manawatu-Wanganui	MNW
18	Wairere Boulders	WAB	Northland	NOR
19	Clints, Oue	CLT	Northland	NOR
20	Maungaroa	MGA	East Cape	E.C
21	Motu River Delta	MRD	East Cape	E.C
22	White Pine	WHP	Hawkes Bay	HB
23	Shine Falls	SHF	Hawkes Bay	HB

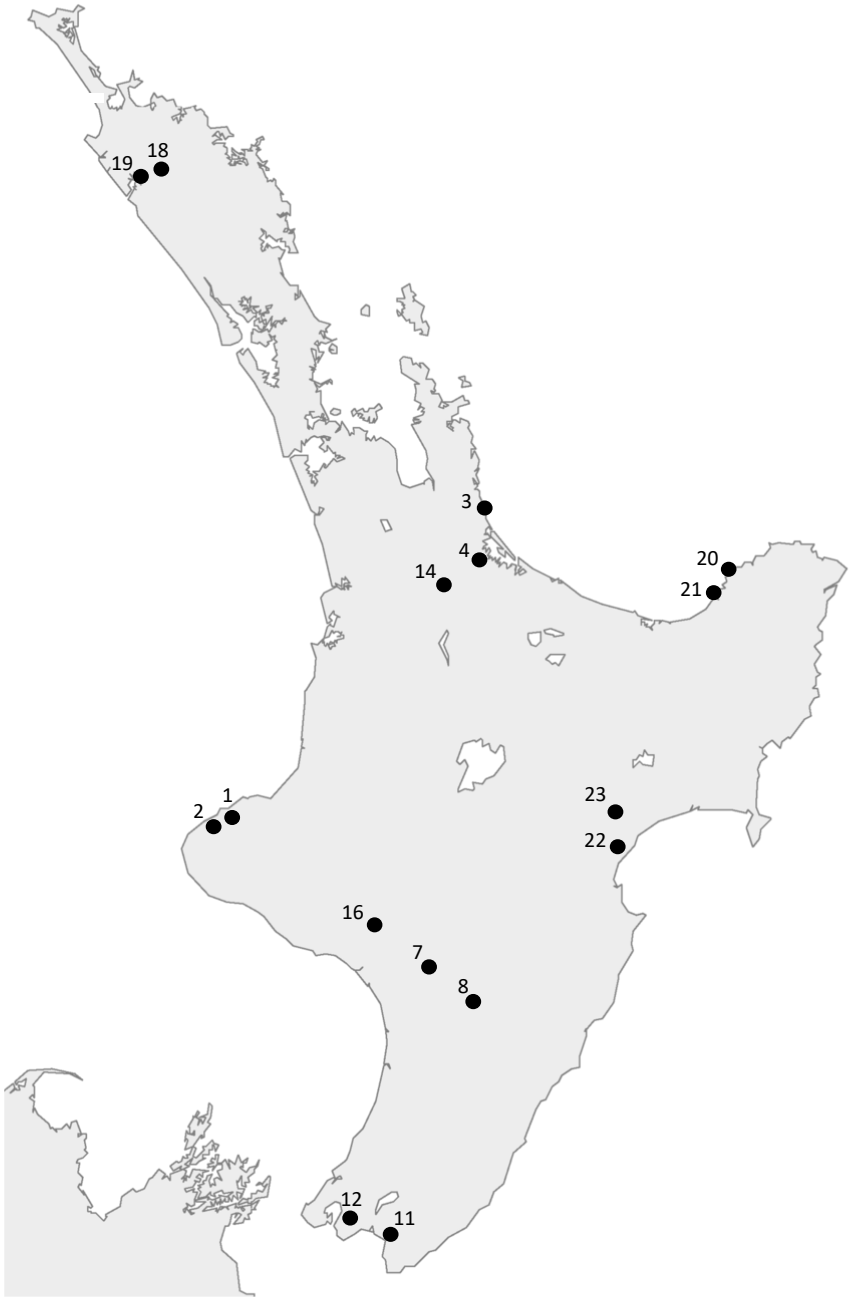


Figure 3.1: Map of the North Island of New Zealand displaying the 16 sampled *P. excelsum* collection sites, as indicated in Table 3.1.

3.3.3 Determination of genetic structure and variation of sampled *P. excelsum* sites

The Bayesian software STRUCTURE v2.3.3 (Pritchard et al., 2000) was utilised to identify genetic groupings. The genetic groupings (i.e., *K* values) were set from 1-8 with 11 replicates for each *K* value. It became clear that a full analysis was not required above *K*=5, with the final run consisting of *K*=1-5 with 11 replicates for each *K* value. The number of burn-in iterations was set at 100,000 and the number of data collection iterations was set to one million using the admixture model with

independent allele frequencies. The STRUCTURE results were exported to CLUMPAK (Kopelman et al., 2015), which uses a Markov clustering algorithm called CLUMPP (Jakobsson & Rosenberg, 2007). For each *K* value, CLUMPP takes the replicate runs and produces a most likely output. The STRUCTURE results were also exported to STRUCTURE HARVESTER v0.6.94 (Earl & vonHoldt, 2012) which uses the Evanno method to generate the mean likelihood values of *K* and estimate the number of genetic groupings (Evanno et al., 2005).

The final data set was imported into RStudio (RStudio-Team, 2020) and used to examine among individual genetic variation in R package *poppr* (Kamvar et al., 2014) where a distance matrix was produced at the individual level using provesti.dist. The R package *ape* (Paradis et al., 2004) was utilised to produce PCoA output at individual and population level using the matrix. The colours used in PCoA graphs represent the STRUCTURE outputs. Genetic structure was also explored using a distance matrix through population level pairwise genetic differences to create a Neighbor-net splits-graph in SPLITSTREE v.4.14.6 (Huson & Bryant, 2006).

GenAlEx v6.5 (Peakall & Smouse, 2006) was used to undertake a codominant Mantel test with 999 permutations to assess isolation by distance in *P. excelsum* sites. GenAlEx was also utilised to calculate the number of alleles per locus, effective number of alleles per locus, number of private alleles, observed heterozygosity per locus, expected heterozygosity per locus, Wright's fixation index for each location, a mean F_{ST} value over all loci and populations, and a pairwise F_{ST} matrix.

3.4 Results

3.4.1 Sampling collection and data handling

Two hundred and forty-nine *P. excelsum* plants from 16 locations were genotyped at ten microsatellite loci, resulting in only 2.7% missing data. Forty-two alleles were discovered in total, ranging from 2-10 per locus with an average of 4.2 (data not shown). Microchecker results suggested potential evidence for a null allele at loci 127770, 51, 339876, 1415, and 65864 based on excess homozygosity. However, there are many explanations for excess homozygosity, so no changes were made to the data set before analysis. The mantel test was completed to test IBD and a weak positive correlation between geographic and genetic distance that is statistically significant was found (R^2=0.245, p = 0.001) (Figure 3.2).

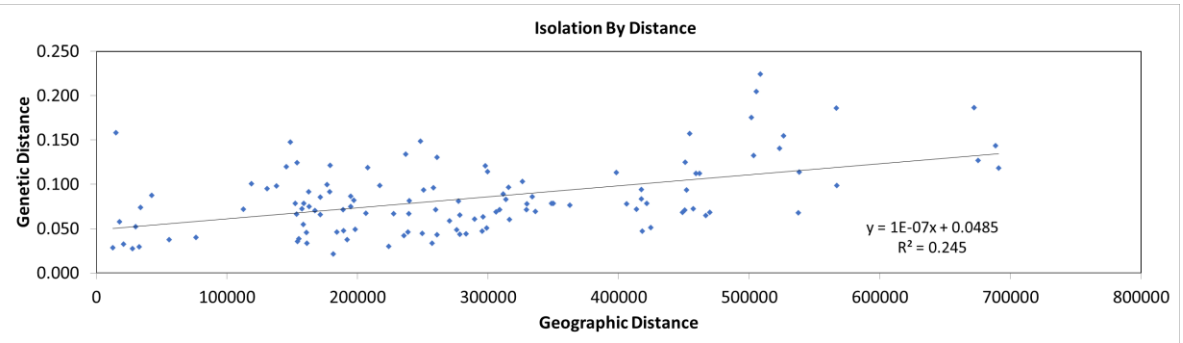


Figure 3.2: Mantel test showing isolation by distance computed in GenAlEx. A weak positive relationship between geographic distance and genetic distance was found (R^2=0.245, p=0.001).

3.4.2 Genetic structure of *P. excelsum* sites

STRUCTURE was used to identify genetic clusters of *P. excelsum* (Figure 3.3). STRUCTURE HARVESTER found the greatest support for *K*=2, with a Δ*K* value of 488.921 (Figure 3.4, Table 3.2). The Δ*K* value for *K*=3 (0.09701) and *K*=4 (0.70690) was very low in comparison. The STRUCTURE mean likelihood outputs for the 11 replicates at *K*=3 and *K*=4 produced considerable variation. This led to high standard deviation values at *K*=3 (46.7614) and *K*=4 (51.3379), and ultimately low Δ*K* values. This tends to be a drawback of the Evanno method, where run-to-run variation can lead to lower support for a given *K* value, potentially diminishing important biological patterns. The STRUCTURE outputs for *K*=2 and *K*=3 are plotted on a map of the North Island of New Zealand for visualisation (Figure 3.5). The pie charts were created using the output generated in STRUCTURE labelled "proportion of membership of each pre-defined population to each of the two clusters" for *K*=2 and three clusters for *K*=3 and represent the proportion of genotypes within a sample site that were assigned to a particular genetic cluster.

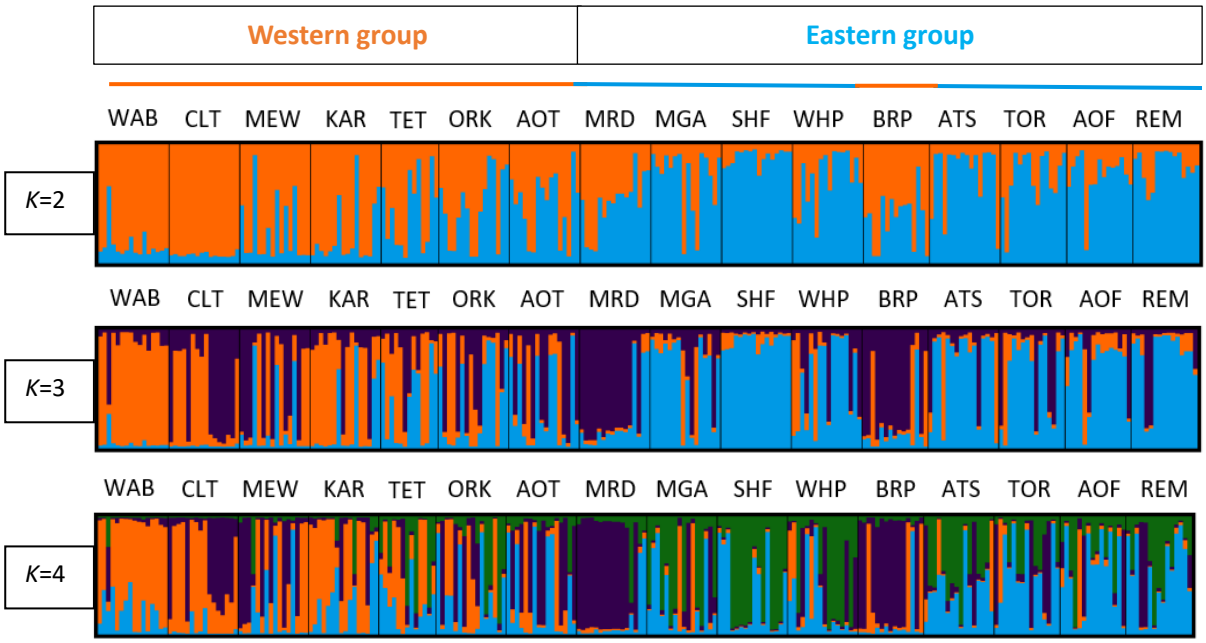


Figure 3.3: STRUCTURE outputs for the full *P. excelsum* data set displaying assignment of individuals to genetic clusters at *K*=2, *K*=3 and *K*=4. Collection sites are delineated by black vertical bars, with the site ID taken from Table 3.1. At *K*=2, individuals mostly split into western and eastern groups. At *K*=3, individuals from site CLT, MRD and BRP form a separate genetic grouping. At *K*=4, individuals in the eastern group most obviously further subdivide into another genetic cluster. The colour scheme utilised will be used in successive figures.

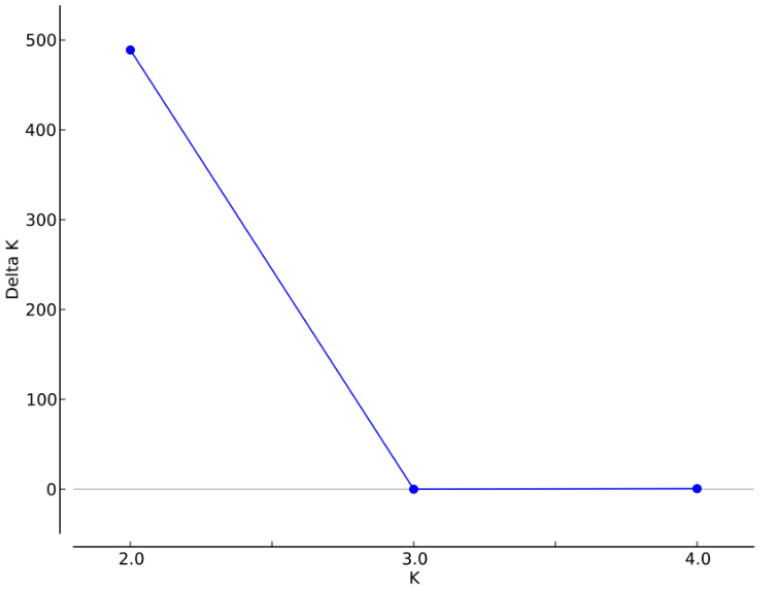


Figure 3.4: Δ*K* output from the Evanno method (Evanno et al., 2005) vs *K* STRUCTURE outputs based on 11 replicates at *K*=2, *K*=3 and *K*=4. This method suggests that *K*=2 best fits this data set.

Table 3.2: Evanno method output from *K*=1-5 STRUCTURE runs. Δ*K* shows the most support for *K*=2.

K	Reps	Mean LnP(*K*)	Stdev LnP(*K*)	Ln'(*K*)	|Ln''(*K*)|	Δ*K*
1	11	-3331.69	0.03015	-	-	-
2	11	-3115.34	0.31292	216.355	153.482	488.921
3	11	-3052.46	46.7614	62.8727	4.53636	0.09701
4	11	-2985.05	51.3379	67.4091	36.2909	0.70690
5	11	-2953.94	2.26640	31.1182	-	-

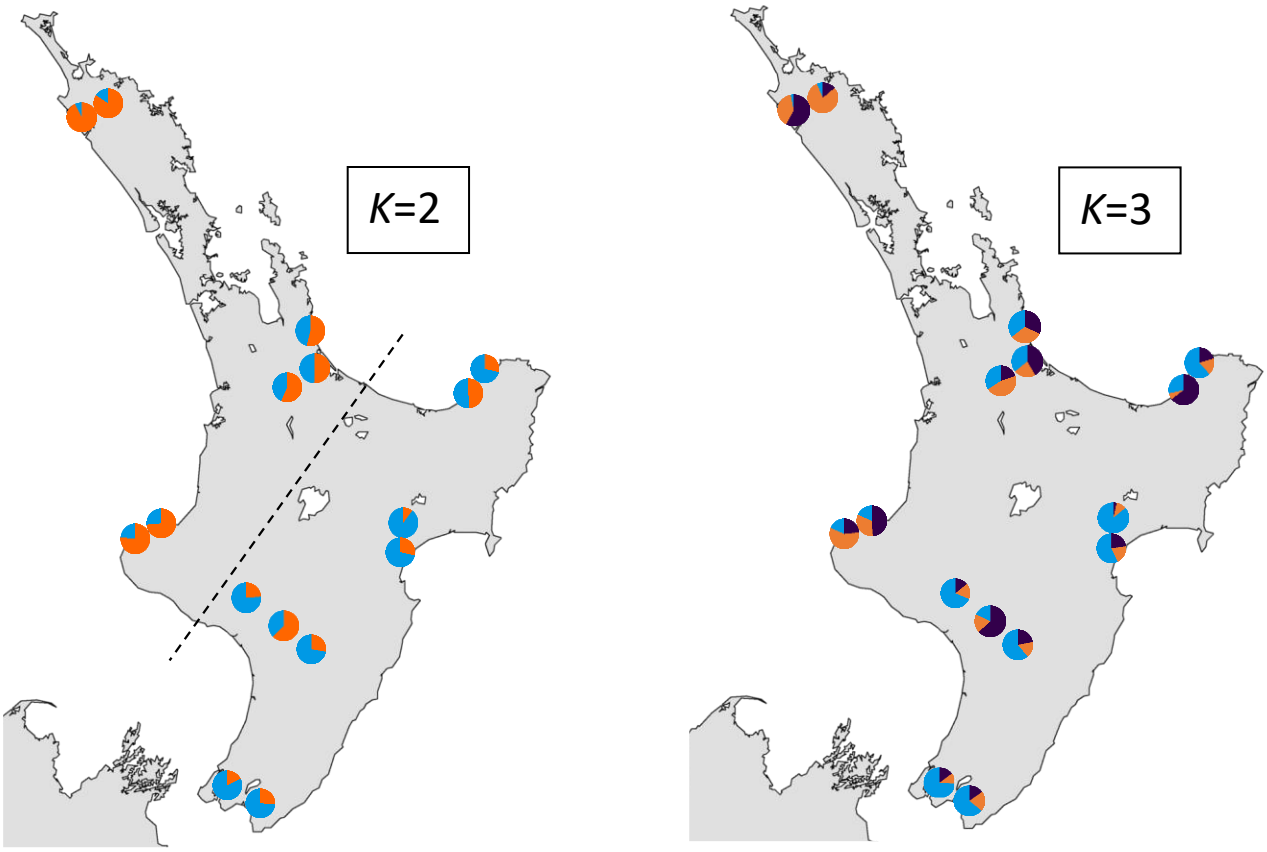


Figure 3.5: Geographical representation of *P. excelsum* STRUCTURE outputs at *K*=2 and *K*=3 for the 16 sampled sites on a map of the North Island of New Zealand. The dotted line in the *K*=2 output represents the east/west split observed. The colours utilised for the pie charts are the same as those in Figure 3.3, representing the proportion of genotypes within a collection site that were assigned to a genetic cluster identified by STRUCTURE.

STRUCTURE, the Neighbor-net and PCoA demonstrate several similar patterns. The most prominent is the splitting of the data into an eastern and western group. The STRUCTURE analyses show that at *K*=2, the sampled sites effectively split into two genetic groups that are defined by an east-west distribution (see dotted line in Figure 3.5). BRP(7) is the only site that does not fit this pattern, situated in the east but having a higher proportion of genotypes belonging to the western group. The main split of the Neighbor-net (Figure 3.6) also denotes the general east-west pattern. This split

separated sites into the proposed STRUCTURE western and eastern groups. However, sites BRP(7) and MRD(21) are geographically positioned in the eastern group in STRUCTURE but are clustered with the sites in the western group in the Neighbor-net. Furthermore, the population-level PCoA using STRUCTURE *K*=2 outputs (Figure 3.7) depict this east-west pattern (see dotted line), with BRP(7) again positioned with sites in the western group.

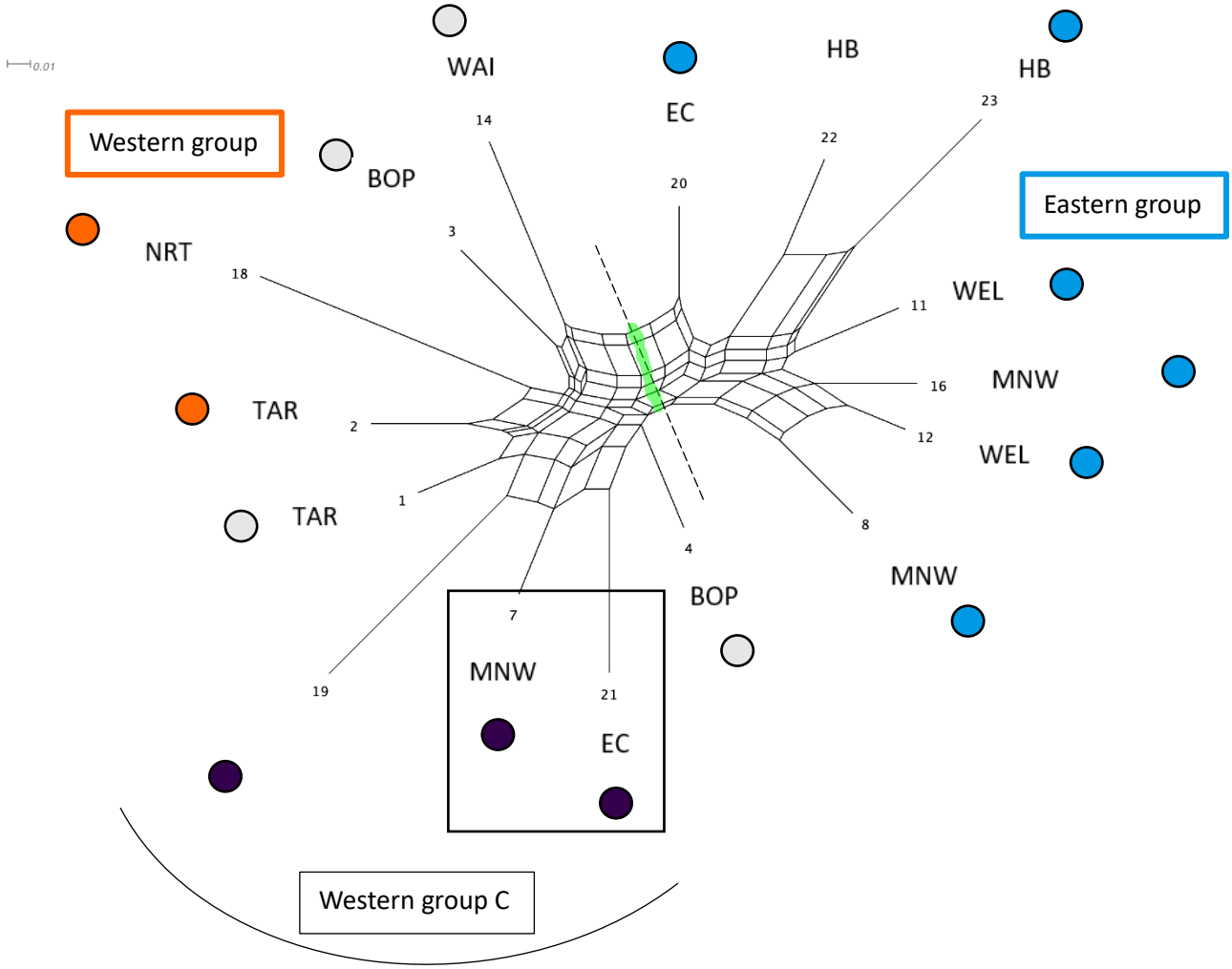


Figure 3.6: The Neighbor-net network for full *P. excelsum* data set using population pairwise genetic distances. The coloured circles identify where more than 50% of genotypes within a population are assigned to a single genetic group in reference to *K*=3. The grey circle indicates where no genetic grouping within a population has more than 50% of individuals assigned to it. The dotted line represents the east-west geographical split observed at *K*=2 in Figure 3.5 and is highlighted as it is a genetically important split. The two boxed populations (7, 21) are geographically positioned in the proposed eastern group but are genetically more similar to the western group, according to this analysis.

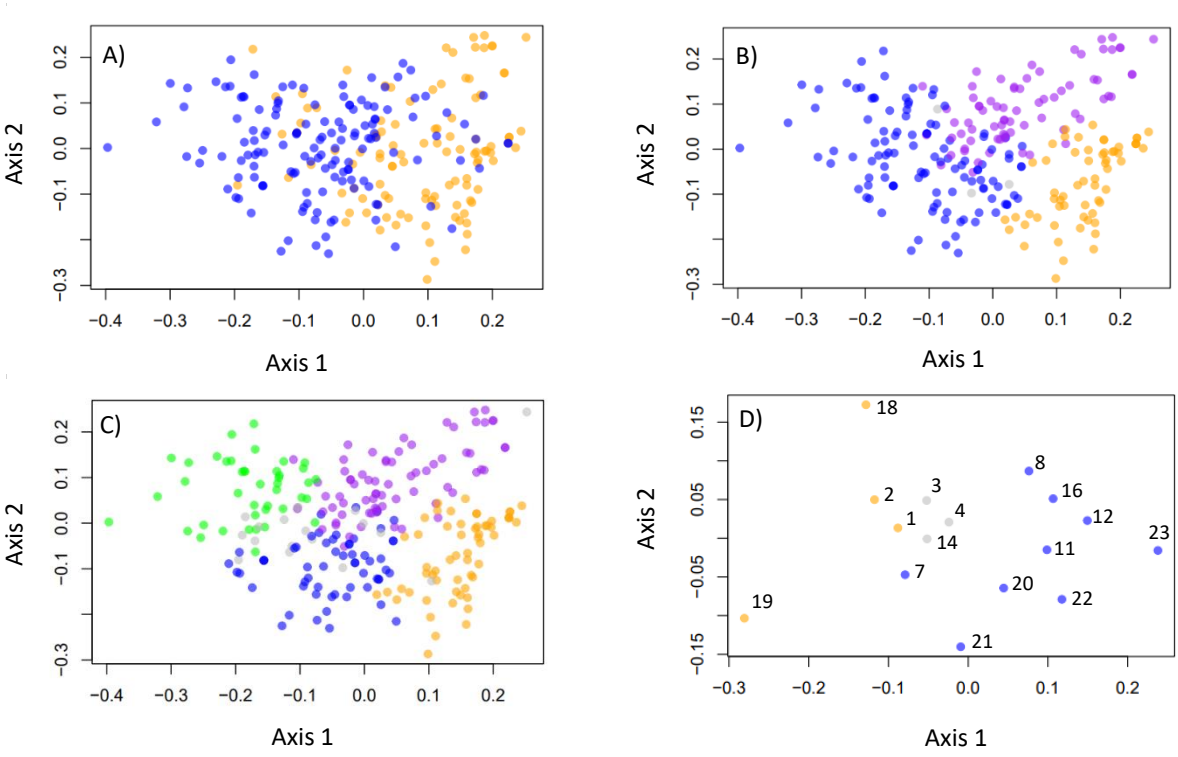


Figure 3.7: Individual and population level PCoA outputs utilising colours to represent the STRUCTURE genetic groupings. Grey dots represent individuals or sampling sites where >50% of genotypes at a collection site are not assigned to a single genetic grouping. Individual level: A) *K*=2, B) *K*=3, C) *K*=4, D) population level PCoA with colours associated with STRUCTURE *K*=2 output.

Another common pattern seen among different analyses is the grouping of individuals from seven sites. At *K*=2, there is a collection of sites, MGA(20), WHP(22), SHF(23), AOF(11), ATS(16), REM(12) and TOR(8), that have more than two thirds of individuals, within each sampling site, assigned to the eastern genetic group. At *K*=3, this collection of seven sites all have at least 50% of individuals, within each site, assigned to the eastern grouping in STRUCTURE. This grouping is depicted clearly in the Neighbor-net, with all seven sites placed adjacent to each other to the right of the main split. The population level PCoA similarly depicts this grouping with all seven collection sites somewhat close together on the right side of the PCoA.

The STRUCTURE and the Neighbor-net analyses demonstrate the grouping of collection sites 19(CLT), 7(BRP) and 21(MRD). At *K*=3, the STRUCTURE output suggests sites 19(CLT), 7(BRP) and 21(MRD) have similar genetic make-up. Similarly, the Neighbor-net placed these three sites together in the network, labelled Western Group C (WGC). The grouping of these three sites is not as obvious in the population level PCoA analysis. They are positioned in the bottom lefthand side but are not particularly close together. The remaining western sampled sites of 1, 3, 4 and 14 do not have more

than 50% of individuals from each site assigned to a genetic grouping (represented by the grey circles in the Neighbor-net), demonstrating a higher rate of admixture in these sites.

The individual PCoA figures are color-coded to reflect the assignment of individuals to STRUCTURE genetic groupings from *K*=2 to *K*=4. At *K*=2, the eastern and western groups are somewhat distinct, but show some overlap. At *K*=3, the individual level PCoA depicts three adjacent and only somewhat overlapping groupings. Here, the eastern group mainly persists and the western group splits into two groups. The western group is now reasonably defined to the bottom right corner and WGC displays some overlap with the eastern group. The grey individuals are all in the centre as they have an approximately equal proportionate assignment of individuals to all three genetic groupings. At *K*=4, the STRUCTURE and individual level PCoA plot show the western group staying relatively unchanged but the STRUCTURE output demonstrates sites in the eastern group becoming somewhat stratified, with Hawkes Bay sites SHF(23) and WHP(22) encompassing the largest proportion of this new genetic group.

Interestingly, where several sites within the same region were collected, they were not always clustered together. For example, the Neighbor-net places the two BOP and two East Cape sites, on opposite sides of the network. The two Northland sites were also positioned on either side of the two Taranaki sites, displaying quite contrasting assignment of individuals that are geographically relatively close, to different genetic groupings.

3.4.3 Population genetic metrics for *P. excelsum*

The population genetic metrics for *P. excelsum* are reported in Table 3.3. The average number of alleles per locus per population (Na) is 2.34, ranging from 1.70 for CLT(19) to 2.90 for ORK(3). The average effective number of alleles per locus (Ne) per population is 1.64, again the lowest Ne is 1.41 for CLT(19) and highest being 1.93 in ORK(3). Ten private alleles were detected, averaging 0.25 per population, with three of these found in WHP(22). WHP(22) was geographically positioned on a valley floor and was surrounded by steep hillside. The average observed heterozygosity (Ho) per population was 0.33, ranging from 0.24 in TOR(8) to 0.43 at WHP(22), and average expected heterozygosity (He) was also 0.33, with the lowest, again, 0.24 in TOR(8) and highest 0.42 at WHP(22). F_{IS} values ranged from -0.17 to 0.09, with nine populations having a slight heterozygote excess, three populations at Hardy Weinberg Equilibrium and four populations having a slight homozygote excess. The pairwise F_{ST} values ranged from 0.022 for populations ATS(16) and REM(12)

to 0.224 for populations CLT(19) and SHF(23) and the average F_{ST} value over all loci and populations is 0.161 (Table 3.4).

Table 3.3: Population genetic diversity metrics for the 16 sampled *P. excelsum* sites.

Site	ID	N		Na	Ne	P	Ho	He	F_{IS}
1	MEW	16	Mean	2.50	1.60	0	0.29	0.30	-0.01
			SE	0.31	0.19		0.06	0.07	0.07
2	KAR	16	Mean	2.40	1.66	0	0.30	0.28	-0.07
			SE	0.34	0.26		0.09	0.08	0.03
3	ORK	16	Mean	2.90	1.93	2	0.33	0.39	0.08
			SE	0.41	0.26		0.06	0.08	0.05
4	AOT	16	Mean	2.40	1.76	0	0.38	0.37	0.03
			SE	0.22	0.20		0.08	0.06	0.09
7	BRP	15	Mean	2.00	1.54	0	0.34	0.28	-0.07
			SE	0.21	0.17		0.10	0.07	0.11
8	TOR	15	Mean	2.20	1.44	0	0.24	0.24	-0.04
			SE	0.29	0.15		0.08	0.07	0.06
11	AOF	15	Mean	2.40	1.79	0	0.39	0.38	-0.03
			SE	0.22	0.18		0.07	0.07	0.07
12	REM	15	Mean	2.20	1.52	1	0.36	0.31	-0.17
			SE	0.20	0.11		0.06	0.05	0.05
14	TET	13	Mean	2.70	1.83	1	0.39	0.41	0.00
			SE	0.30	0.18		0.05	0.06	0.06
16	ATS	16	Mean	2.50	1.50	0	0.27	0.30	0.09
			SE	0.27	0.11		0.06	0.05	0.09
18	WAB	16	Mean	2.20	1.62	1	0.31	0.30	-0.03
			SE	0.33	0.20		0.09	0.08	0.05
19	CLT	16	Mean	1.70	1.41	1	0.27	0.25	0.00
			SE	0.15	0.11		0.08	0.06	0.17
20	MGA	16	Mean	2.30	1.59	0	0.34	0.33	-0.05
			SE	0.15	0.13		0.06	0.06	0.05
21	MRD	16	Mean	2.20	1.51	0	0.29	0.28	0.01
			SE	0.20	0.15		0.09	0.07	0.13
22	WHP	16	Mean	2.60	1.79	3	0.43	0.42	0.00
			SE	0.27	0.12		0.05	0.04	0.10
23	SHF	16	Mean	2.30	1.69	1	0.42	0.38	-0.12
			SE	0.26	0.11		0.07	0.05	0.10
Total			Mean	2.34	1.64	0.625	0.33	0.33	-0.02
			SE	0.07	0.04		0.02	0.02	0.02

Notes: Site=sample site number, ID=site code, N=number of samples, Na=average number of alleles per locus, Ne=effective number of alleles per locus, P=number of private alleles, Ho=average observed heterozygosity per locus, He=average expected heterozygosity per locus, F_{IS}=Wright's fixation index, SE=standard error.

Table 3.4: Pairwise F_{ST} values between sites produced in GenAlEx. Average F_{ST} is reported in the outlined box.

SITE	18	19	1	2	14	3	4	21	20	23	22	7	16	8	11	12
18																
19	0.158															
1	0.094	0.072														
2	0.079	0.084	0.028													
14	0.103	0.078	0.087	0.068												
3	0.069	0.089	0.045	0.043	0.038											
4	0.072	0.086	0.030	0.042	0.052	0.030										
21	0.157	0.112	0.069	0.079	0.100	0.079	0.078									
20	0.125	0.112	0.079	0.076	0.046	0.034	0.055	0.058								
23	0.205	0.224	0.134	0.149	0.121	0.119	0.092	0.125	0.086							
22	0.141	0.155	0.082	0.094	0.082	0.067	0.049	0.066	0.048	0.033						
7	0.114	0.068	0.039	0.046	0.067	0.061	0.034	0.064	0.083	0.148	0.098					
16	0.133	0.175	0.072	0.101	0.099	0.059	0.046	0.114	0.060	0.075	0.073	0.088				
8	0.099	0.186	0.071	0.075	0.130	0.072	0.049	0.121	0.097	0.120	0.095	0.074	0.40			
11	0.119	0.144	0.065	0.081	0.078	0.072	0.051	0.068	0.065	0.047	0.044	0.070	0.038	0.066		
12	0.127	0.186	0.071	0.096	0.113	0.071	0.047	0.094	0.068	0.051	0.044	0.092	0.022	0.036	0.027	

Average F_{ST} value over all loci and populations calculated in GenAlEx: 0.161.

3.5 Discussion

3.5.1 Biogeography and genetic structure

3.5.1.1 East-west patterning in New Zealand biota

The genetic structuring in *P. excelsum* is relatively weak, with low levels of genetic differentiation found among sampled locations. However, all genetic structure analyses methods utilised in this study revealed a consistent east-west pattern of partitioned genetic variation. A similar pattern has been identified in species of native New Zealand plants, including *Dactylanthus taylorii* (Holzapfel et al., 2002; McLay et al., 2022) and *Asplenium hookerianum* (Shepherd et al., 2007), and animals (Baker et al., 2005; Lloyd, 2003; Shepherd & Lambert, 2008; Trewick, 2000) and is suggested to be a result of repeated volcanic activity in the TVZ. Holzapfel et al. (2002) extended this interpretation with a discussion about this potentially causing a bottleneck event and fixation of genetic variation to either side of the TVZ. An east-west pattern is also observed in *Entelea arborescens* (Malvaceae) though this pattern is not thought to be a result of geological events (Shepherd et al., 2019). Genetic diversity levels are higher in the western group for *D. taylorii* (Holzapfel et al., 2002; McLay et al., 2022) and higher in the eastern group for *A. hookerianum* (Shepherd et al., 2007), with genetic diversity equally low across the North Island for *P. excelsum* in this study.

The orientation of the east-west split for *P. excelsum* is roughly in line with that proposed for *D. taylorii*, but both are further west than Cockayne's Line (Ellis et al., 2015). It is possible that the processes that led to the positioning of Cockayne's Line are similar to those that have produced the east-west split for *P. excelsum,* but they are not orientated in the same position. It is not expected that this biogeographically significant line would necessarily align with patterns identified in this study as there are many different factors that impact species-specific distributions. Cockayne's Line has been described from interspecific distribution patterns of a diverse range of taxa, mostly animals, and from factors that may be more relevant over a relatively longer evolutionary timescale. Repeated volcanic activity in the TVZ or the axial mountain ranges disrupting gene flow between eastern and western North Island over time (i.e., Cockayne's Line) or a combination of the two could explain the east-west pattern seen in *P. excelsum* as the TVZ and mountain range align approximately with this east-west split. Furthermore, pollen evidence suggests the presence of patchy forest and woody vegetation throughout most of New Zealand during the LGM (Newnham et al., 2013). Therefore, it is also possible that there were microrefugial sites on either side of this axial mountain

range where *P. excelsum* may have persisted, harbouring genetic variation in small forest pockets to the east and west. However, this is speculative, with McLay et al. (2022) suggesting further study on distribution patterns of native flora to investigate a potentially widescale influence of volcanism on genetic diversity would be informative.

This east-west pattern is observed across a range of taxa that could potentially share a common vicariance event (Shepherd et al., 2007). Although, when common distributional patterns are seen across diverse species, these patterns can be a result of varying influences occurring at different temporal scales (Soltis et al., 2006). Ellis et al. (2015) emphasises the importance of dating divergence times when investigating phylogeographic patterns; however, dating clade splits only provides a very rough temporal approximation and cannot be completed with microsatellite markers. It can be achieved using other methods, for example, Shepherd et al. (2019) utilised molecular clock analyses on ITS sequences of individuals from putative eastern and western groups and suggested divergence during the Pleistocene, although this period spans over 2 Ma so is still a broad estimate. A similar method could be applied to *P. excelsum*, among other native biota, to determine the approximate east-west pattern establishment, although considering how widescale and disruptive events during the Pleistocene were, it is likely many different intraspecific patterns were established during this period.

3.5.1.2 Further findings

The mean F_{ST} value for all sites and loci was 0.161 and the majority of the pairwise F_{ST} values were below 0.10. This suggests a reasonably high level of gene flow between sampled sites of *P. excelsum*. Pairwise F_{ST} values for *P. excelsum* are expected to be generally lower than a disjunct species distribution, as the species has a nearly continuous distribution across the lowlands of the North Island. The two Northland (18, 19) and Hawkes Bay (23) sites exhibit the most differentiation from each other, with pairwise F_{ST} of 0.205 and 0.224 respectively. This relationship could be explained by IBD which showed a significant but weak positive relationship between genetic and geographic distance.

Understanding population structure can inform the degree of gene flow across the distribution of a species, with species homogenisation implying weak structure and divergence indicating strong structure or speciation (Duminil et al., 2007). A New Zealand native species with a somewhat similar distribution to *P. excelsum* is *Pseudopanax crassifolius* (Araliaceae), a lowland widespread forest tree, that was also found to have weak genetic structure (Gemmell et al., 2022). However, Gemmell

et al. (2022) included South Island samples and found the North Island and eastern South Island populations had higher diversity than the western and lower South Island, whereas *P. excelsum* had more consistent genetic variation across the North Island. Future work should include sampling sites on the South Island for *P. excelsum* to compare the genetic diversity and structure to the North Island and other native species with similar distributions.

The east-west patterning for *P. excelsum* is a finding that is concordant among analysis methods and has potential biogeographic explanations. WGC revealed by STRUCTURE at *K*=3 is also concordant among analysis methods but has no obvious link between genetic grouping and geographic positioning. There are several instances where collection sites within a geographic region have reasonably differentiated genetic makeup, i.e., the largest proportion of individuals within sites of the same region are not assigned to the same genetic group in *K*=3 STRUCTURE, e.g., see Northland, East Cape, and Manawatu-Wanganui. These regions each include a site with a high proportion of individuals assigned to WGC: CLT(19), BRP(7) and MRD(21). WGC is a very unusual finding, with STRUCTURE grouping these sites with no obvious geographic link between them. The geographic orientation of this group has not been observed in any other New Zealand species to our knowledge. There are however, other native species where no apparent link between genetic structure and geographic positioning has been found, such as for pōhutukawa (Myrtaceae) (Broadhurst et al., 2008). The WGC goes against the expectations of IBD, although the relationship between genetic and geographic distance for *P. excelsum* was significant but only weakly positive. The sites in WGC do, however, have real genetic similarities. At two loci (51 and 339876), WGC sites display the highest frequency of an allele which is distinct from the alleles at greatest frequency in other sites (data not shown). Potential explanations for this unexpected finding are unsubstantiated, for example, linkage to conserved loci in these sites, incomplete lineage sorting, human distribution, or nursery stocks. Further sampling would be useful to determine if this grouping is present elsewhere in New Zealand.

Collection sites in BOP and Waikato show the highest level of admixture, suggesting these sites are regularly exchanging genetic variation with other sites currently or in the recent past. KAR(2) displays less admixture than MEW(1) which could be due to recurrent volcanic activity at Mt Taranaki over the last 30 ka, with some block-and-ashflows reaching as far as 13.5km to the west coast along the Hangatahua river (Cronin et al., 2021; Platz et al., 2007), potentially disturbing vegetation at KAR(2).

3.5.2 Disruption of species distribution

3.5.2.1 Environmental influence on species distribution

Climate and edaphic factors are important environmental determinants in a species' spatial and genetic distribution. *P. excelsum* has a preference for moist soil (de Lange, 2021; Hall, 2020) but the widespread distribution suggests it is somewhat resilient to soil type. At present, *P. excelsum* is found only up to 500m above sea level (de Lange, 2021) and is intolerant to frost (Koia & Shepherd, 2020), suggesting climate is more influential to survival. During the LGM, glacial modelling found air temperature was predicted to range from 4.1-7°C below present on average across New Zealand (Eaves & Brook, 2021; Newnham et al., 2013; Rojas et al., 2009). The present day mean annual temperature in the North Island is 16°C (NIWA, Accessed 2023), with mean temperatures potentially dropping to 9°C or lower during the LGM. The LGM markedly decreased temperatures and resulted in more regions across New Zealand experiencing frost, likely influencing the distribution and survival of biota across New Zealand. Species exhibiting characteristics such as frost hardiness are thought to have survived relatively widespread during the LGM (Buckley et al., 2010; Shepherd & Perrie, 2011; Shepherd et al., 2007). However, the prevalent frost during the LGM likely would have caused a mass die-off of *P. excelsum*, resulting in a bottleneck event. If this were the case, *P. excelsum* would likely have been restricted to the putative LGM refugial areas of northern North Island, as suggested for several other New Zealand species (Buckley et al., 2010; Buckley et al., 2009; Drummond et al., 2000; Gardner et al., 2004; Gemmell et al., 2022). This area is characterised by near continuous forest (Newnham et al., 2013), providing more habitable conditions for many species compared to the rest of the North Island during the LGM.

3.5.2.2 Human redistribution or removal

It is likely that the distribution and natural state of *P. excelsum* genetic variation has been altered over time by human redistribution, cultivation, or removal. Forest disturbance in New Zealand due to human activity has been prevalent over the last 700 years (Newnham et al., 2013). A period of Polynesian deforestation has been described by Wilmshurst and McGlone (1996) around 600-800 YBP. The arrival of European colonists to New Zealand from the 1800's led to further, extensive forest clearances which did not slow in the upper North Island until the late 1960's (Drummond et al., 2000). Māori introduced important food crops to New Zealand, such as the sweet potato, yam, taro and gourd (Best, 1930) and cultivated these crops during establishment. Similarly, due to the importance of kawakawa to Māori culture and medicine (Brooker & Cooper, 1961), it is possible that

cuttings were manually planted and moved around New Zealand as groups migrated, potentially homogenising variation if the same cuttings were used each time they relocated.

3.5.3 Distribution of genetic variation and structure in sampled *P. excelsum* sites

3.5.3.1 Within population genetic diversity

Most sampled *P. excelsum* sites have a F_{IS} value close to 0, ranging from -0.17 to 0.09. This indicates they are reasonably close to Hardy Weinberg equilibrium (HWE), with sites CLT(19), TET(14) and WHP(22) at HWE. Nine of the 16 sites display heterozygote excess, i.e., deviation from HWE, suggesting that overdominance or negative assortative mating are at play (da Cunha et al., 2022; Stoeckel et al., 2006); however, the deviations from HWE are slight. Sites ORK(3), AOT(4), MRD(21) and ATS(16) exhibit homozygote excess, suggesting these sites may be subject to Wahlund effect, i.e., subpopulation structure, inbreeding, potential inaccurate sampling or presence of null alleles (Chapuis & Estoup, 2006; De Meeûs, 2017; Stoeckel et al., 2006). Although, as all sampled *P. excelsum* sites are so close to HWE, it is unlikely any of these factors have a strong influence on the sites.

Within population diversity was measured by number of alleles (Na), expected (He), and observed (Ho) heterozygosity. He was plotted geographically (Figure 3.8) with the only geographic pattern observed being the sites assigned to WGC (7, 21, 19) generally displaying lower He than the other sites. The average Na was 2.34 and the average He and Ho being 0.33 for *P. excelsum* sites. These values are all reasonably lower than those stated in a review on over 6000 microsatellite studies that found the average number of alleles for Magnoliids (the clade that encompasses Piperaceae) was 5.68, average He was 0.485 and average Ho was 0.579 (Merritt et al., 2015). As the Magnoliid clade contains over 10,000 woody and herbaceous species, these values are a very rough estimate for standard population genetic metrics within this clade. Hence, reported *P. excelsum* values have been compared to other New Zealand species and those with similar life history traits for a closer comparison.

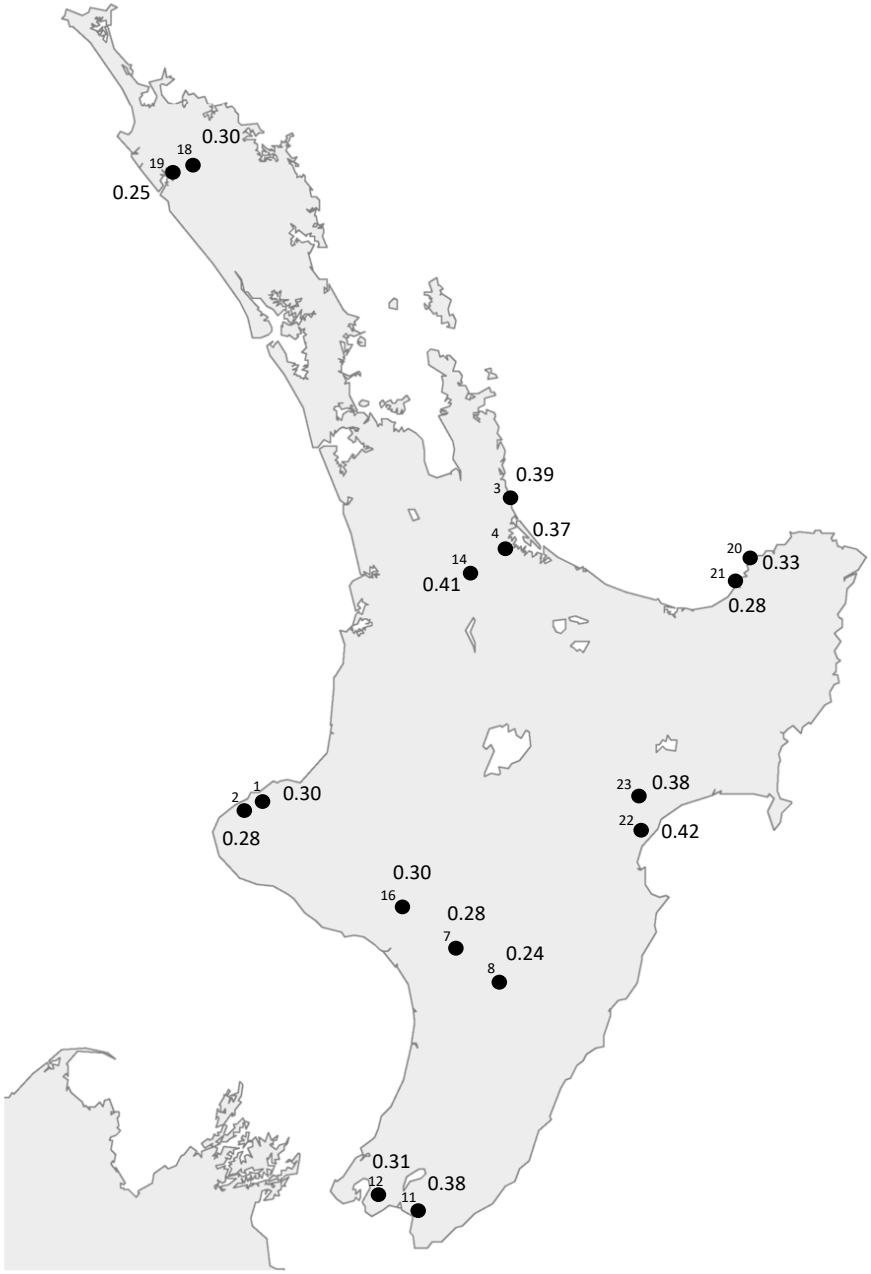


Figure 3.8: Expected heterozygosity values for sampled *P. excelsum* sites.

For *P. excelsum*, the average number of alleles per locus, per population (Na) was 2.34. In other population genetics studies completed on New Zealand flora (Table 3.5), Na ranged from 1.32 (Haase, 1992) to 16.4 for *Sophora microphylla* (Fabaceae) (Heenan et al., 2018), with the most similar value to *P. excelsum* being 1.93 for *Arthropodium cirratum* (Asparagaceae) (Shepherd et al., 2018). In contrast with plant species found outside of New Zealand with similar life histories (Table 3.6), Na varied from 4.41 (Saro et al., 2015) – 37.3 (Zhu et al., 2016), with 4.41 for a palm *Phoenix*

canariensis (Arecaceae) closest to *P. excelsum* Na. Therefore, the average Na for *P. excelsum* is quite low in comparison with other New Zealand species and those with similar life history traits elsewhere.

The average expected (He) and observed (Ho) heterozygosity across the collection sites for *P. excelsum* is the same at 0.33. He is the expected heterozygosity under random mating and *P. excelsum* appears to be effectively undergoing random mating as it is wind pollinated and the seed appears to be dispersed effectively, meaning similar He and Ho values is expected. For New Zealand species, He varies from 0.051 (Haase, 1992) – 0.66 for *Sophora microphylla* (Heenan et al., 2018) and for similar life history species, ranges from 0.44 (Zhu et al., 2016) – 0.79 (Pazouki et al., 2016). The most similar He value to New Zealand species is 0.39 for *Sphagnum* moss (Phagnaceae) (Karlin et al., 2008) and 0.44 for similar life history species, the tree, *Lindera glauca* (Lauraceae) (Zhu et al., 2016). For New Zealand studies, Ho varies from 0.051 (Haase, 1992) to 0.66 (Heenan et al., 2018) and for similar life history species, Ho ranges from 0.47 (Dev et al., 2011) to 0.95 (Pazouki et al., 2016). The most similar Ho value to New Zealand species is 0.39 for *D. taylorii* (McLay et al., 2022) and 0.47 for *Ficus hispida* (Moraceae), a species with similar life history traits. Therefore, in comparison with New Zealand species and those with similar life histories, the within population diversity for *P. excelsum* is relatively low.

Table 3.5: A summary of New Zealand population genetics studies on native plant species. Studies that reported standard population genetic diversity metrics were used for comparison with *P. excelsum*.

Species	Characteristics	Distribution	Marker type	Average Na	Average Ho	Average He	Reference
Sphagnum **(Phagnaceae)**	Moss, perennial, reproduce with spore filled capsules	*S. cristatum*: common, widespread	SSRs	4.25*	0.60*	0.39*	(Karlin et al., 2008)
Phormium tenax **(Asphodelaceae)**	Perennial, predominantly outcrossing	Across New Zealand and Norfolk Island	AFLPS, SSRs	N/a			(Smissen & Heenan, 2010)
Pseudopanax ferox **(Araliaceae)**	Small lowland tree, entomophilous	Found across New Zealand but particularly disjunct	SSRs, chloroplasts	8.67*	0.44*	0.66*	(Shepherd & Perrie, 2011)
Craspedia **(Asteraceae)**	Herbacious, outcrossing, perennial	Mostly located in Kahurangi National Park	SSRs	7.92	0.41	0.62	(Breitwieser et al., 2015)
Fuscospora ** **(Nothofagaceae)**	Wind pollinated, perennial tree	*F. truncata*: lowland North Island and northern South Island.	SSRs	4	0.46	0.48	(Smissen et al., 2014)
Pittosporum obcordatum **(Pittosporaceae)**	long-lived, shrub, dioecious	Uncommon – 14 known populations across New Zealand	ISSRs			0.163	(Wright et al., 2017)
Melicytus ramiflorus **(Violaceae)**	Dioecious, tree, entomophilous	Common in coastal, lowland, and low elevation montane forest	AFLPs, ISSRs	N/a			(Stevens et al., 2015)
Sophora ** **(Fabaceae)**	Small tree, ornithophilous	*S. microphylla:* across New Zealand	SSRs	16.4	0.66	0.79	(Heenan et al., 2018)
Hebe speciosa **(Plantaginaceae)**	Threatened shrub, entomophilous	Disjunct populations, coastal populations	AFLP	N/a			(Armstrong & De Lange, 2005)

		between Malborough Sounds and Northland					
***Olearia gardneri* (Asteraceae)**	Long lived tree, hermaphroditic, outcrossing	Very rare, North Island only between Rangitikei and Wairarapa	AFLPs	N/a			(Barnaud & Houliston, 2010)
***Myosotis* forget-me-nots (Boraginaceae)**	Many species threatened, self-compatible, perennial	Varying between species	Chloroplast, AFLPs	N/a			(Meudt et al., 2015)
***Metrosideros excelsa* (Myrtaceae)**	Long lived, outcrossing tree	Naturally upper North Island (North of Taupō)	AFLPs	N/a			(Broadhurst et al., 2008)
***Metrosideros bartlettii* (Myrtaceae)**	Endangered tree, long lived, entomophilous	Rare, top of North Island	Chloroplasts, AFLPs			0.18	(Drummond et al., 2000)
***Metrosideros* (Myrtaceae)**	Frost sensitive, seeds wind dispersed	The five species cover New Zealand	Chloroplasts	N/a			(Gardner et al., 2004)
***Dactylanthus taylorii* (Balanophoraceae)**	Endangered root-holoparasite, dioecious, chiropterophilious	North Island, Little Barrier Island	RAPDs, SSRs	3.10	0.39	0.46	(McLay et al., 2022)
***Asplenium hookerianum* (Aspleniaceae)**	Forest fern, spores wind dispersed	Across New Zealand	Chloroplasts				(Shepherd et al., 2007)
***Cordyline australis* (Laxmanniaceae)**	Monocot tree, entomophilous	Widespread	AFLPs			0.11	(Beever et al., 2013)
***Entelea arborescens* (Malvaceae)**	Coastal tree, entomophilous,	Coastal North Island and top of South Island	Chloroplasts, SSRs				(Shepherd et al., 2019)

	short lived, frost intolerant						
***Pseudopanax crassifolious* (Araliaceae)**	Lowland forest tree, entomophilous	Widespread across New Zealand	SSRs and chloroplasts	N/a			(Gemmell et al., 2022)
Kunzea ericoides ** **(Myrtaceae)**	Entomophilous, wind and water dispersed seeds	*K. amathicola:* North and South Island	SSRs	10.81	0.57	0.71	(Heenan et al., 2022)
***Nothofagus truncate* (Nothofagaceae)**	Monoecious, wind pollinated tree	North and upper South Island	Isozymes	1.32	0.051	0.051	(Haase, 1992)
***Sophora* (Fabaceae)**	Tree, ornithophilous	Vary among species	Chloroplasts	N/a			(Shepherd et al., 2017)
***Arthropodium cirratum* (Asparagaceae)**	Coastal herb, perennial	Upwards of Rotorua	SSRs	1.93*	0.12	0.60	(Shepherd et al., 2018)
***Leptospermum scoparium* (Myrtaceae)**	Evergreen, hardy shrub	Mainland NZ and Chatham	SNPs	N/a			(Buys et al., 2019)

*Average calculated from population data. **Where more than one subspecies was studied, the subspecies with the most similar life history characteristics to *P. excelsum* was chosen and information on the chosen subspecies distribution, average Na, He and Ho (if applicable) was included in this table.

Table 3.6: A collection of eight studies of species with similar life history traits to *P. excelsum*. The basic population genetic diversity metrics were used for comparison with *P. excelsum*.

Species	Dioecious	Wind pollinated	Long-lived	Tree	Woody	Na	Ho	He	Reference
Allocasuarina humilis **(Carsuarinaceae)**	Y	Y	Y	Y	Y		0.64	0.76	(Llorens et al., 2017)
Lindera glauca **(Lauraceae)**	Y			Y		37.3		0.44	(Zhu et al., 2016)
Phoenix canariensis **(Arecaceae)**	Y		Y			4.41	0.51	0.50	(Saro et al., 2015)
Ficus hispida and Ficus exasperata* **(Moraceae)**	Y			Y	Y	10.57	0.47	0.78	(Dev et al., 2011)
Milicia excelsa **(Moraceae)**	Y	Y		Y		8.38		0.55	(Bizoux et al., 2009)
Populus simonii **(Salicaceae)**	Y	Y	Y	Y		6.93	0.51	0.68	(Wei et al., 2013)
Cryptomeria japonica **(Cupressaceae)**	Y	Y	Y	Y	Y	24	0.65	0.77	(Takahashi et al., 2005)
Pinus sylvestris **(Pinaceae)**	Y	Y	Y	Y		18.5	0.95	0.95	(Pazouki et al., 2016)

**(only F. exasperata values reported here)*

Low diversity levels in *P. excelsum* could potentially be due to a bottleneck event during the LGM, that would have reduced the pool of genetic variation. It is likely that *P. excelsum* re-established rapidly and effectively following the LGM, with high levels of gene flow likely due to the effective pollen and seed dispersal mechanisms, indicative by its current nearly continuous lowland distribution. A bottleneck event often results in the preservation of diversity in refugial populations during ice ages (Hewitt, 1996). However, this is not always the case, with the *P. excelsum* Northland sites (18, 19) found not to be more diverse than the rest of the North Island sites. Broadhurst et al. (2008) also found diversity was not higher in putative refugial areas for pōhutukawa (Myrtaceae). Further sampling in the Northland area would be beneficial to assess the genetic diversity in more depth and discover if there are areas within this region that are more genetically diverse, such as presumed sites of refugia for other species. It is possible that this potential bottleneck event for *P. excelsum* was severe, with only a small group of individuals encompassing low levels of diversity persisting through the LGM. Furthermore, male *P. excelsum* plants have been found to be more abundant than female, with Merrett et al. (2007) reporting that female plants made up only 25-30% of sampled populations. This would increase the severity of a bottleneck event, lowering the effective population size so that potentially only one quarter of the population would have contributed 50% of the alleles to the next generation. Alternatively, the cultural and medicinal significance of *P. excelsum* to Māori may have led to widespread, manual movement of cuttings by Māori, potentially explaining the low genetic variation across the North Island, although evidence to support this speculative theory is difficult to obtain.

3.5.3.2 Ecological factors impacting the spatial and genetic distribution of *P. excelsum*

There are several ecological factors that could be contributing to the observed low levels of among site differentiation of *P. excelsum*, including pollination mechanism and seed dispersal, breeding system, spatial distribution, and life cycle.

Population spatial distribution is known to influence the distribution of genetic variation (Levy et al., 2016), with species displaying continuous populations often presenting lower levels of interpopulation divergence (Barbará et al., 2007). Broadhurst et al. (2017) investigated several Australian plant species and found mean F_{ST} values for semi-continuous distributions (0.09) to be lower than patchily distributed species (0.16). The distribution of *P. excelsum* varies somewhat around the North Island. The species has a semi-continuous distribution in some areas and a patchy distribution in others, but areas most commonly demonstrate nearly continuous distribution

(Supplementary Figure 1). The overall F_{ST} value for *P. excelsum* is 0.161 which suggests a reasonably high level of inter population gene flow and interestingly aligns with the patchily distributed mean F_{ST} value in the Broadhurst et al. (2017) study. However, the definition for semi-continuous and patchy distribution is not provided in that study, so it is difficult to directly compare with *P. excelsum*. It is likely the few areas of patchy *P. excelsum* distribution has driven this value upward. The relatively low pairwise F_{ST} values between the majority of *P. excelsum* sites aligns with the general expectations of low interpopulation divergence within nearly continuously distributed species.

Free flowing pollen can lead to among population genetic homogeneity (Dyer & Sork, 2001) and the halting of pollen movement can increase IBD (Wright, 1943). The relatively low levels of within population genetic diversity, illustrated by comparatively low He and Na values, and population differentiation shown by relatively low pairwise F_{ST} values, overall F_{ST} value and weak genetic structure, suggest that the anemophilous pollination strategy is allowing frequent pollen-mediated gene flow within and between *P. excelsum* sites. The seed dispersal distances of birds will also influence the distribution and gene flow of a species, with *P. excelsum* seed dispersed by several common bird species such as kererū and blackbird. Frugivores often disperse seed less than 100m from the parent plant; see Debussche et al. (1982) for examples, including blackbirds (Williams, 2006). Kererū are reported to disperse seeds distances of upwards of 30km (Powlesland et al., 2011) but more frequently travel shorter distances (mean of 19 minutes at a time, travelling roughly 85m) (Wotton & Kelly, 2012). It appears the reasonably low-medium dispersal distance range of all the seed dispersing birds for *P. excelsum* does not limit their ability to disperse seed effectively. As these seed dispersing birds are common in New Zealand forests, they generally should not have to travel far to disperse seed to nearby populations. This is reflected by the nearly continuous distribution and the low levels of sampled inter-site differentiation (Table 3.4), suggesting gene flow among sampled sites is frequent in the present or in the recent past.

Long lived species, i.e., perennials, typically lower the influence of genetic drift, heighten migration frequency and lower divergence between populations (Loveless & Hamrick, 1984). *P. excelsum* is a perennial, long-lived, dioecious, species, with these characteristics all potentially contributing to the low between site diversity levels.

3.6 Summary

Overall genetic structure for *P. excelsum* was found to be weak, but an east-west pattern was observed. This pattern could potentially be explained by repeated volcanic activity, the axial mountain range running from Wellington to East Cape, LGM microrefugia, or a combination of these factors. A compilation of many studies on a diverse range of taxa investigating species spatial and genetic distribution, in combination with further pollen and fossil evidence is required to provide a better understanding of the response of species to geologically significant events in New Zealand. Low within site diversity was noted for *P. excelsum* compared to other New Zealand natives and species with similar life history traits. This may be a result of frost during the LGM causing a bottleneck in *P. excelsum* and retreatment of the species to Northern refugia. As the sampled sites in Northland were not found to be more diverse than the other sampled sites, as expected if the species was restricted to Northland in the LGM, more sampling is required in Northland to further investigate the diversity in this region; however, the establishment of multiple refugia and subsequent dispersal could also explain this finding. Low between site differentiation was also found for *P. excelsum*, with a combination of ecological variables, such as dioecious, long lived, effective pollen and seed dispersal mechanisms may have led to the nearly continuous spatial distribution. These factors are also suggested to be promoting gene flow and homogenisation of *P. excelsum* in this study.

3.7 References

Alloway, B. V., Lowe, D. J., Barrell, D. J. A., Newnham, R. M., Almond, P. C., Augustinus, P. C., Bertler, N. A., Carter, L., Litchfield, N. J., McGlone, M. S., Shulmeister, J., Vandergoes, M. J., & Williams, P. W. (2007). Towards a climate event stratigraphy for New Zealand over the past 30 000 years (NZ-INTIMATE project). *Journal of Quaternary Science*, *22*(1), 9-35.

Armstrong, T. T. J., & De Lange, P. J. (2005). Conservation genetics of Hebe speciosa (Plantaginaceae) an endangered New Zealand shrub. *Botanical Journal of the Linnean Society*, *149*(2), 229-239. https://doi.org/10.1111/j.1095-8339.2005.00437.x

Baker, A. J., Huynen, L. J., Haddrath, O., Millar, C. D., & Lambert, D. M. (2005). Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: The giant moas of New Zealand. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 8257-8262.

Balloux, F., & Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, *11*(2), 155-165.

Barbará, T., Martinelli, G., Fay, M., Mayo, S., & Lexer, C. (2007). Population differentiation and species cohesion in two closely related plants adapted to neotropical high-altitude 'inselbergs', Alcantarea imperialis and Alcantarea geniculata (Bromeliaceae). *Molecular Ecology*, *16*(10), 1981-1992.

Barnaud, A., & Houliston, G. J. (2010). Population genetics of the threatened tree daisy Olearia gardneri (Asteraceae), conservation of a critically endangered species. *Conservation Genetics*, *11*, 1515–1522.

Beever, R., Houliston, G., & Armstrong, T. (2013). Geographic patterns of genetic variation in Cordyline australis; does fingerprinting recover phenotypic and phenological patterns? *New Zealand Journal of Botany*, *51*(1), 3-12.

Best, E. (1930). Maori agriculture. Cultivated food-plants of the Maori and native methods of agriculture. *The Journal of the Polynesian Society*, *39*(4 (156), 346-380.

Bizoux, J. P., Daïnou, K., Bourland, N., Hardy, O. J., Heuertz, M., Mahy, G., & Doucet, J. L. (2009). Spatial genetic structure in Milicia excelsa (Moraceae) indicates extensive gene dispersal in a low-density wind-pollinated tropical tree. *Molecular Ecology*, *18*(21), 4398-4408.

Breitwieser, I., Ford, K. A., & Smissen, R. D. (2015). Characterisation of SSR markers for New Zealand Craspedia and their application in Kahurangi National Park. *New Zealand Journal of Botany*, *53*(1), 60-73.

Broadhurst, L., Breed, M., Lowe, A., Bragg, J., Catullo, R., Coates, D., Encinas-Viso, F., Gellie, N., James, E., & Krauss, S. (2017). Genetic diversity and structure of the Australian flora. *Diversity and Distributions*, *23*(1), 41-52.

Broadhurst, L. M., Young, A. G., & (2008), B. G. M. (2008). AFLPs reveal an absence of geographical genetic structure among remnant populations of pohutukawa (Metrosideros excelsa, Myrtaceae). *New Zealand Journal of Botany*, *46*(1), 13-21.

Brooker, S. G., & Cooper, R. C. (1961). New Zealand medicinal plants. *Economic Botany*, *15*(1), 1-10.

Buckley, T. R., Marske, K., & Attanayake, D. (2010). Phylogeography and ecological niche modelling of the New Zealand stick insect Clitarchus hookeri (White) support survival in multiple coastal refugia. *Journal of Biogeography*, *37*(4), 682-695.

Buckley, T. R., Marske, K. A., & Attanayake, D. (2009). Identifying glacial refugia in a geographic parthenogen using palaeoclimate modelling and phylogeography: the New Zealand stick insect Argosarchus horridus (White). *Molecular Ecology*, *18*(22), 4650-4663.

Burger, W. C. (1972). Evolutionary trends in the Central American species of Piper (Piperaceae). *Brittonia*, *24*(4), 356-362.

Burrows, C. J. (1995). Germination behaviour of seeds of the New Zealand species Fuchsia excorticata, Griselinia littoralis, Macropiper excelsum, and Melicytus ramiflorus. *New Zealand Journal of Botany*, *33*(1), 131-140.

Butts, C. A., Klink, J. W. v., Joyce, N. I., Paturi, G., Hedderley, D. I., Martell, S., & Harvey, D. (2019). Composition and safety evaluation of tea from New Zealand kawakawa (Piper excelsum). *Journal of Ethnopharmacology*, *232*, 110-118.

Buys, M. H., Winkworth, R. C., De Lange, P. J., Wilson, P. G., Mitchell, N., Lemmon, A. R., Moriarty Lemmon, E., Holland, S., Cherry, J. R., & Klápště, J. (2019). The phylogenomics of diversification on an island: applying anchored hybrid enrichment to New Zealand Leptospermum scoparium (Myrtaceae). *Botanical Journal of the Linnean Society*, *191*(1), 1-17.

Campbell, H., Malahoff, A., Browne, G., Graham, I., & Sutherland, R. (2012). New Zealand Geology. *Episodes Journal of International Geoscience*, *35*(1), 57-71.

Chapuis, M.-P., & Estoup, A. (2006). Microsatellite Null Alleles and Estimation of Population Differentiation. *Molecular Biology and Evolution*, *24*(3), 621-631. https://doi.org/10.1093/molbev/msl191

Cronin, S. J., Zernack, A. V., Ukstins, I. A., Turner, M. B., Torres-Orozco, R., Stewart, R. B., Smith, I. E., Procter, J. N., Price, R., & Platz, T. (2021). The geological history and hazards of a long-lived stratovolcano, Mt. Taranaki, New Zealand. *New Zealand Journal of Geology and Geophysics*, *64*(2-3), 456-478.

da Cunha, N. L., Xue, H., Wright, S. I., & Barrett, S. C. (2022). Genetic variation and clonal diversity in floating aquatic plants: Comparative genomic analysis of water hyacinth species in their native range. *Molecular Ecology*, *31*(20), 5307-5325.

de Lange, P. J. (2021). *Piper excelsum subsp. excelsum Fact Sheet (content continuously updated)*. New Zealand Plant Conservation Network. Retrieved 01 May from https://www.nzpcn.org.nz/flora/species/piper-excelsum-subsp-excelsum/

De Meeûs, T. (2017). Revisiting FIS, FST, Wahlund Effects, and Null Alleles. *Journal of Heredity*, *109*(4), 446-456. https://doi.org/10.1093/jhered/esx106

Debussche, M., Escarré, J., & Lepart, J. (1982). Ornithochory and plant succession in Mediterranean abandoned orchards. *Vegetatio*, *48*, 255-266.

Dev, S. A., Kjellberg, F., Hossaert-McKey, M., & Borges, R. M. (2011). Fine-scale population genetic structure of two dioecious Indian keystone species, Ficus hispida and Ficus exasperata (Moraceae). *Biotropica*, *43*(3), 309-316.

Drummond, R. S. M., Keeling, D. J., Richardson, T. E., Gardner, R. C., & Wright, S. D. (2000). Genetic analysis and conservation of 31 surviving individuals of a rare New Zealand tree, Metrosideros bartlettii (Myrtaceae). *Molecular Ecology*, *9*(8), 1149-1157.

Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007). Can population genetic structure be predicted from life-history traits? *The American Naturalist*, *169*(5), 662-672.

Dyer, R., & Sork, V. (2001). Pollen pool heterogeneity in shortleaf pine, Pinus echinata Mill. *Molecular Ecology*, *10*(4), 859-866.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359-361. https://doi.org/doi: 10.1007/s12686-011-9548-7

Eaves, S. R., & Brook, M. S. (2021). Glaciers and glaciation of North Island, New Zealand. *New Zealand Journal of Geology and Geophysics*, *64*(1), 1-20. https://doi.org/DOI: 10.1080/00288306.2020.1811354

Ebach, M. C. (2015). Origins of biogeography. *History, Philosophy and Theory of the Life Sciences*, *13*, 1-20.

Ellis, E. A., Marshall, D. C., Hill, K. B., Owen, C. L., Kamp, P. J., & Simon, C. (2015). Phylogeography of six codistributed New Zealand cicadas and their relationship to multiple biogeographical boundaries suggest a re-evaluation of the Taupo Line. *Journal of Biogeography*, *42*(9), 1761-1775.

Ennos, R. (1994). Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*, *72*(3), 250-259.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*(8), 2611-2620.

Gardner, R. C., De Lange, P. J., Keeling, D. J., Bowala, T., Brown, H. A., & Wright, S. D. (2004). A late Quaternary phylogeography for Metrosideros (Myrtaceae) in New Zealand inferred from chloroplast DNA haplotypes. *Biological Journal of the Linnean Society*, *83*(3), 399-412.

Gardner, R. O. (1997). Macropiper (Piperaceae) in the south-west Pacific. *New Zealand Journal of Botany*, *35*(3), 293-307.

Gemmell, M. R., Shepherd, L. D., Zuccarello, G. C., & Perrie, L. R. (2022). Phylogeography of the widespread New Zealand tree lancewood/horoeka (Pseudopanax crassifolius; Araliaceae). *New Zealand Journal of Botany*, 1-17.

Haase, P. (1992). Isozyme variability and biogeography of Nothofagus truncata (Fagaceae). *New Zealand Journal of Botany*, *30*(3), 315-328.

Habel, J. C., Zachos, F. E., Dapporto, L., Roedder, D., Radespiel, U., Tellier, A., & Schmitt, T. (2015). Population genetics revisited–towards a multidisciplinary research field. *Biological Journal of the Linnean Society*, *115*(1), 1-12.

Hall, M. M. (2020). *Enrichment planting of late-successional plant species within restored urban forests* Waikato University].

Heenan, P., Mitchell, C., & Houliston, G. (2018). Genetic Variation and Hybridisation among Eight Species of kōwhai (Sophora: Fabaceae) from New Zealand Revealed by Microsatellite Markers. *Genes*, *9*(2), 111.

Heenan, P. B., McGlone, M. S., Mitchell, C. M., Cheeseman, D. F., & Houliston, G. J. (2022). Genetic variation reveals broad-scale biogeographic patterns and challenges species' classification in the Kunzea ericoides (kānuka; Myrtaceae) complex from New Zealand. *New Zealand Journal of Botany*, *60*(1), 2-26.

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, *58*(3), 247-276.

Hodge, S., Barron, M., & Wratten, S. D. (2000). Induced defences in kawakawa (Macropiper excelsum): do caterpillars avoid previous leaf damage? *New Zealand Journal of Ecology*, *24*(1), 91-95.

Hodge, S., Keesing, V., Wratten, S. D., Lovei, G. L., Palmer, J., & Cilgi, T. (1998). Herbivore damage and leaf loss in the New Zealand pepper tree ('kawakawa', Macropiper excelsum, Piperaceae). *New Zealand Journal of Ecology*, *22*(2), 173-180.

Holzapfel, S., Faville, M. Z., & Gemmill, C. E. C. (2002). Genetic variation of the endangered holoparasite Dactylanthus taylorii (Balanophoraceae) in New Zealand. *Journal of Biogeography*, *29*(5-6), 663-676. https://doi.org/https://doi.org/10.1046/j.1365-2699.2002.00715.x

Hughes, P. D., Gibbard, P. L., & Ehlers, J. (2013). Timing of glaciation during the last glacial cycle: evaluating the concept of a global 'Last Glacial Maximum'(LGM). *Earth-Science Reviews*, *125*, 171-198.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*(2), 254-267.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801-1806.

Jayaprakash, R., Ramzan, F., Miles-Chan, J. L., Foster, M., Mithen, R. F., & Pook, C. (2022). Exploring the Chemical Space of Kawakawa Leaf (Piper excelsum). *Nutrients*, *14*(23), 5168.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Karlin, E. F., Boles, S. B., & Shaw, A. J. (2008). Systematics of Sphagnum section Sphagnum in New Zealand: A microsatellite-based analysis. *New Zealand Journal of Botany*, *46*(2), 105-118.

Koia, J. H., & Shepherd, P. (2020). The Potential of Anti-Diabetic Rakau Rongoa (Maori Herbal Medicine) to Treat Type 2 Diabetes Mellitus (T2DM) Mate Huka: A Review. *Front Pharmacol*, *11*, 935. https://doi.org/10.3389/fphar.2020.00935

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179-1191.

Lakeman-Fraser, P., & Ewers, R. M. (2013). Enemy release promotes range expansion in a host plant. *Oecologia*, *172*(4), 1203-1212. https://doi.org/10.1007/s00442-012-2555-x

Lei, J., Burgess, E. J., Richardson, A. T., Hawkins, B. C., Baird, S. K., Smallfield, B. M., van Klink, J. W., & Perry, N. B. (2015). Cytotoxic Amides from Fruits of Kawakawa, Macropiper excelsum. *Planta Med*, *81*(12-13), 1163-1168. https://doi.org/10.1055/s-0035-1546106

Levy, E., Byrne, M., Coates, D., Macdonald, B., McArthur, S., & Van Leeuwen, S. (2016). Contrasting influences of geographic range and distribution of populations on patterns of genetic diversity in two sympatric Pilbara acacias. *Plos One*, *11*(10), e0163995.

Llorens, T. M., Tapper, S. L., Coates, D. J., McArthur, S., Hankinson, M., & Byrne, M. (2017). Does population distribution matter? Influence of a patchy versus continuous distribution on genetic patterns in a wind-pollinated shrub. *Journal of Biogeography*, *44*(2), 361-374.

Lloyd, B. D. (2003). Intraspecific Phylogeny of the New Zealand short-tailed bat Mystacina tuberculata inferred from multiple mitochondrial gene sequences. *Systematic Biology*, *52*, 460-476.

Loveless, M. D., & Hamrick, J. L. (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, *15*(1), 65-95.

Manville, V., Segschneider, B., Newton, E., White, J. D. L., Houghton, B. F., & Wilson, C. J. N. (2009). Environmental impact of the 1.8 ka Taupo eruption, New Zealand: Landscape responses to a large-scale explosive rhyolite eruption. *Sedimentary Geology*, *220*(318-336).

McGlone, M., Duncan, R., & Heenan, P. (2001). Endemism, Species Selection and the Origin and Distribution of the Vascular Plant Flora of New Zealand. *Journal of Biogeography*, 199-216.

McGlone, M. S. (1985). Plant biogeography and the late Cenozoic history of New Zealand. *New Zealand Journal of Botany*, *23*(4), 723-749.

McGlone, M. S., Newnham, R. M., & Moar, N. T. (2010). The vegetation cover of New Zealand during the Last Glacial Maximum: do pollen records under-represent woody vegetation. *Terra Australis*, *32*, 49-68.

McLay, T. G., Tate, J. A., Gemmill, C. E., Holzapfel, A. S., & Symonds, V. V. (2022). Inferring the biogeography of New Zealand's only endemic holoparasitic plant, the threatened Dactylanthus taylorii (Mystropetalaceae). *New Zealand Journal of Botany*, 1-23.

Merrett, M. F., Robertson, A. W., & Peterson, P. G. (2007). Pollination performance and vulnerability to pollination breakdown of sixteen native shrub species from New Zealand. *New Zealand Journal of Botany*, *45*(4), 579-591. https://doi.org/https://doi.org/10.1080/00288250709509740

Merritt, B. J., Culley, T. M., Avanesyan, A., Stokes, R., & Brzyski, J. (2015). An empirical review: characteristics of plant microsatellite markers that confer higher levels of genetic variation. *Applications in Plant Sciences*, *3*(8), 1500025.

Meudt, H. M., Prebble, J. M., & Lehnebach, C. A. (2015). Native New Zealand forget-me-nots (Myosotis, Boraginaceae) comprise a Pleistocene species radiation with very low genetic divergence. *Plant Systematics and Evolution*, *301*, 1455-1471.

Newnham, R., McGlone, M., Moar, N., Wilmshurst, J., & Vandergoes, M. (2013). The vegetation cover of New Zealand at the last glacial maximum. *Quaternary Science Reviews*, *74*, 202-214.

NIWA. (Accessed 2023). *Overview of New Zealand's climate*.

Obst, K., Lieder, B., Reichelt, K. V., Backes, M., Paetz, S., Geissler, K., Krammer, G., Somoza, V., Ley, J. P., & Engel, K. H. (2017). Sensory active piperine analogues from Macropiper excelsum and their effects on intestinal nutrient uptake in Caco-2 cells. *Phytochemistry*, *135*, 181-190. https://doi.org/10.1016/j.phytochem.2016.12.016

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language V3.3. *Bioinformatics*, *20*, 289-290.

Pazouki, L., Shanjani, P. S., Fields, P. D., Martins, K., Suhhorutšenko, M., Viinalass, H., & Niinemets, Ü. (2016). Large within-population genetic diversity of the widespread conifer Pinus sylvestris at its soil fertility limit characterized by nuclear and chloroplast microsatellite markers. *European journal of forest research*, *135*, 161-177.

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

Petit, R. J., Aguinagalde, I., de Beaulieu, J., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Muller-Starck, G., Demesure-Musch, B., Palme, A., Martın, J. P., Rendell, S., & Vendramin, G. G. (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, *300*(5625), 1563–1565.

Platz, T., Cronin, S. J., Cashman, K. V., Stewart, R. B., & Smith, I. E. (2007). Transition from effusive to explosive phases in andesite eruptions—a case-study from the AD1655 eruption of Mt. Taranaki, New Zealand. *Journal of Volcanology and Geothermal Research*, *161*(1-2), 15-34.

Powlesland, R. G., Moran, L. R., & Wotton, D. M. (2011). Satellite tracking of kereru (Hemiphaga novaeseelandiae) in Southland, New Zealand: impacts, movements and home range. *New Zealand Journal of Ecology*, 229-235.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945-959. https://doi.org/10.1093/genetics/155.2.945

Rojas, M., Moreno, P., Kageyama, M., Crucifix, M., Hewitt, C., Abe-Ouchi, A., Ohgaito, R., Brady, E. C., & Hope, P. (2009). The Southern Westerlies during the last glacial maximum in PMIP2 simulations. *Climate Dynamics*, *32*(4), 525-548.

RStudio-Team. (2020). *RStudio: Integrated Development for R.* In *RStudio, PBC, Boston, MA.*

Saro, I., González-Pérez, M. A., García-Verdugo, C., & Sosa, P. A. (2015). Patterns of genetic diversity in Phoenix canariensis, a widespread oceanic palm (species) endemic from the Canarian archipelago. *Tree Genetics & Genomes*, *11*, 1-13.

Shelley, J. J., David, B. O., Thacker, C. E., Hicks, A. S., Jarvis, M. G., & Unmack, P. J. (2020). Phylogeography of the Cran's bully Gobiomorphus basalis (Gobiiformes: Eleotridae) and an analysis of species boundaries within the New Zealand radiation of Gobiomorphus. *Biological Journal of the Linnean Society*, *130*(2), 365-381. https://doi.org/10.1093/biolinnean/blaa052

Shepherd, L. D., Bulgarella, M., & de Lange, P. J. (2018). Genetic structuring of the coastal herb Arthropodium cirratum (Asparagaceae) is shaped by low gene flow, hybridization and prehistoric translocation. *Plos One*, *13*(10), e0204943.

Shepherd, L. D., de Lange, P. J., Perrie, L. R., & Heenan, P. B. (2017). Chloroplast phylogeography of New Zealand Sophora trees (Fabaceae): extensive hybridization and widespread Last Glacial Maximum survival. *Journal of Biogeography*, *44*(7), 1640-1651.

Shepherd, L. D., Frericks, J., Biggs, P. J., & de Lange, P. J. (2019). Phylogeography of the endemic New Zealand tree Entelea arborescens (whau; Malvaceae). *New Zealand Journal of Botany*, *57*(3), 154-168.

Shepherd, L. D., & Lambert, D. M. (2008). Ancient DNA and conservation: lessons from the endangered kiwi of New Zealand. *Molecular Ecology*, *17*, 2174-2184.

Shepherd, L. D., & McLay, T. G. (2011). Two micro-scale protocols for the isolation of DNA from polysaccharide-rich plant tissue. *Journal of Plant Research*, *124*(2), 311-314.

Shepherd, L. D., & Perrie, L. R. (2011). Microsatellite DNA analyses of a highly disjunct New Zealand tree reveal strong differentiation and imply a formerly more continuous distribution. *Molecular Ecology*, *20*(7), 1389-1400. https://doi.org/https://doi.org/10.1111/j.1365-294X.2011.05017.x

Shepherd, L. D., Perrie, L. R., & Brownsey, P. J. (2007). Fire and ice: volcanic and glacial impacts on the phylogeography of the New Zealand forest fern Asplenium hookerianum. *Molecular Ecology*, *16*(21), 4536-4549. https://doi.org/https://doi.org/10.1111/j.1365-294X.2007.03451.x

Smissen, R. D., & Heenan, P. B. (2010). A taxonomic appraisal of the Chatham Islands flax (Phormium tenax) using morphological and DNA fingerprint data. *Australian Systematic Botany*, *23*(5), 371-380. https://doi.org/https://doi.org/10.1071/SB10023

Smissen, R. D., Richardson, S. J., Morse, C. W., & Heenan, P. B. (2014). Relationships, gene flow and species boundaries among New Zealand Fuscospora (Nothofagaceae: southern beech). *The New Zealand Journal of Botany*, *52*(4), 389-406.

Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, *15*(14), 4261-4293.

Stevens, M., Clarke, A., Clarkson, F., Goshorn, M., & Gemmill, C. (2015). Are current ecological restoration practices capturing natural levels of genetic diversity? A New Zealand case study using AFLP and ISSR data from mahoe (Melicytus ramiflorus). *New Zealand Journal of Ecology*, *39*(2), 190-197.

Stoeckel, S., Grange, J., Fernandez-Manjarres, J. F., Bilger, I., Frascaria-Lacoste, N., & Mariette, S. (2006). Heterozygote excess in a self-incompatible and partially clonal forest tree species—Prunus avium L. *Molecular Ecology*, *15*(8), 2109-2118.

Symonds, V. V., & Lloyd, A. M. (2004). A simple and inexpensive method for producing fluorescently labelled size standard. *Molecular Ecology Notes*, *4*(4), 768-771.

Takahashi, T., Tani, N., Taira, H., & Tsumura, Y. (2005). Microsatellite markers reveal high allelic variation in natural populations of Cryptomeria japonica near refugial areas of the last glacial period. *Journal of Plant Research*, *118*, 83-90.

Trewick, S. A. (2000). Mitochondrial DNA sequences support allozyme evidence for cryptic radiation of New Zealand Peripatoides (Onychophora). *Molecular Ecology*, *9*, 269-281.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3), 535-538.

Wallis, G. P., & Trewick, S. A. (2009). New Zealand phylogeography: evolution on a small continent. *Molecular Ecology*, *18*(17), 3548-3580.

Wardle, P. (1963). Evolution and distribution of the New Zealand flora, as affected by Quaternary climates. *New Zealand Journal of Botany*, *1*(1), 3-17.

Wardle, P. (1988). Effects of glacial climates on floristic distribution in New Zealand 1. A review of the evidence. *New Zealand Journal of Botany*, *26*(4), 541-555.

Waters, J. M., & Craw, D. (2006). Goodbye Gondwana? New Zealand biogeography, geology, and the problem of circularity. *Systematic Biology*, *55*(2), 351-356.

Wei, Z., Du, Q., Zhang, J., Li, B., & Zhang, D. (2013). Genetic diversity and population structure in Chinese indigenous poplar (Populus simonii) populations using microsatellite markers. *Plant Molecular Biology Reporter*, *31*(3), 620-632.

Williams, P. A. (2006). The role of blackbirds (Turdus merula) in weed invasion in New Zealand. *New Zealand Journal of Ecology*, 285-291.

Wilmshurst, J. M., & McGlone, M. S. (1996). Forest disturbance in the central North Island, New Zealand, following the 1850 BP Taupo eruption. *The holocene*, *6*(4), 399-411.

Wotton, D. M., & Kelly, D. (2012). Do larger frugivores move seeds further? Body size, seed dispersal distance, and a case study of a large, sedentary pigeon. *Journal of Biogeography*, *39*(11), 1973-1983.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 139-156.

Wright, S. A., Hutchison, M., Hale, M. L., Gemmill, C. E. C., Lange, P. J. d., & Pelser, P. B. (2017). A preliminary conservation genetic study of Pittosporum obcordatum (Pittosporaceae), an endemic New Zealand species with a disjunct distribution. *New Zealand Journal of Botany*, *55*(4), 424-438.

Wyman, T. E. (2013). *Consequences of reduced bird densities for seed dispersal* University of Canterbury].

Zhu, S., Ding, Y., Yap, Z., & Qiu, Y. (2016). De novo assembly and characterization of the floral transcriptome of an economically important tree species, Lindera glauca (Lauraceae), including the development of EST-SSR markers for population genetics. *Molecular biology reports*, *43*(11), 1243-1250.

Chapter 4: Summary and future work

4.1 Introduction

New Zealand is considered an ideal location to create and test biogeographic theories, due to a large number of endemic species, climatic conditions and isolation from other landmasses (Köhler & Schmitt, 2023; McGlone, 1985). New Zealand has also been subject to many geological influences over time, including glaciation, volcanism, tectonic activity, and inundation, which have affected the spatial and genetic distribution of species. Species distribution is also influenced by climate and edaphic factors (de la Vega, 1997) and species-specific life history traits (Gottlieb, 1977; Hamrick et al., 1979; Loveless & Hamrick, 1984). Biogeographic patterns of *P. excelsum* were explored in this study and the geographic history of New Zealand and other native species patterns were considered when discussing the potential explanations for the patterns observed. Population genetic studies look at the structure, genetic variation and gene flow of species (Balloux & Lugon-Moulin, 2002) and provide meaningful insight for studying biogeographic patterns, such as those found in native New Zealand species. Therefore, a population genetics approach was used to examine the genetic variation and population structure of sampled *P. excelsum* sites in the North Island.

P. excelsum (Piperaceae) is a native, understory tree that is common in lowland mixed angiosperm-podocarp forests (Burrows, 1995; Butts et al., 2019; Hodge et al., 1998). *P. excelsum* is naturally positioned around the coastline of the North and South Island as far as the Banks Peninsula (de Lange, 2021) and on New Zealand's offshore islands (Lakeman-Fraser & Ewers, 2013). The geographic distribution and life history characteristics of *P. excelsum* made it a useful species for studying population genetics parameters and contributing knowledge to understanding the geological widescale biogeographic patterns.

The objectives for this study were to:

1. Develop microsatellite markers suitable to assess the population genetics of *P. excelsum*
2. Determine the genetic structure and variation within and between sampled *P. excelsum* sites in the North Island
3. Discuss potential historical influences that could have influenced the present spatial and genetic distribution of *P. excelsum*

4.2 Findings

Ten microsatellite markers were developed to assess the population genetics of *P. excelsum*. The findings were discussed in chapter 2 and 3 and are summarised here.

4.2.1 Objective one

Thirty-seven novel primer pairs were developed. These primers were trialled on *P. excelsum* samples collected in the Manawatū area, on leaf specimens retrieved from the Dame Ella Campbell Herbarium and on 15 samples collected from across four geographically dispersed regions. Ten of these primer pairs were chosen for this study based on their consistent amplification, easily scorable, and polymorphic properties. These ten markers were genotyped and analysed and a relatively low average number of alleles per locus found in the small trial sample (1.83). Furthermore, within population diversity was also relatively low when compared to other plant species in the Magnoliid clade, other New Zealand species and other species with similar life history traits to *P. excelsum*, although this study had a small number of samples compared with the other studies.

4.2.2 Objectives two and three

A general pattern of east-west genetic partitioning was identified in sites across several analysis methods. Potential explanations for this pattern are suggested as: the axial mountain range that runs from Wellington to East Cape interrupting gene flow between the east and west currently or in the recent past, repeated volcanic activity in the Taupō Volcanic Zone or the persistence of *P. excelsum* in microrefugia forest pockets during the LGM.

Genetic structure was found to be relatively weak across the North Island sites of *P. excelsum*. This suggests the level of gene flow between sampled sites has been relatively frequent and is reflected in the overall F_{ST} value of 0.161. Furthermore, within population diversity is low with the average Na of 2.34 and average He and Ho being 0.33 across all sampled *P. excelsum* sites in comparison to other New Zealand natives and other species with similar life history traits. This is suggested to be a result of colder conditions during the LGM wiping out majority of *P. excelsum*, causing a bottleneck event and retreatment of the species to Northern refugia. If this were the case, the effective pollen, and seed dispersal, when considered with the perennial, dioecious characteristics of *P. excelsum*, is likely to have promoted the nearly continuous distribution and increased species homogenisation during the re-establishment following the LGM. An alternate theory regarding the movement of *P. excelsum* cuttings around the North Island by Māori is postulated to have contributed to species homogenisation but evidence for this is difficult to obtain.

4.3 Limitations

One of the biggest limitations for this project is the difficulty in determining if sampled sites are relictual or replanted. While relictual sites were sought, it is not always possible to determine the

history of a location. The documentation on replanting schemes is not available to the public and the detail of these schemes is unknown. Cost was another major project restriction. The cost of genotyping is significant and limited the number of samples for microsatellite screening to 15 total from four sites and approximately 16 per site for this study. The limited number of samples per collection site potentially lowered the accuracy in the depiction of genetic variation and structure that exists for *P. excelsum*. There is a chance that the samples collected from each site are not representative of the entire population which could have skewed my findings, although with low within population diversity this seems unlikely. There were also some geographic gaps in sampling, particularly in Auckland, the South Island and around the east and west coastline. Further sampling across the range would allow for a more complete assessment of genetic variation and population structure in *P. excelsum.*

4.4 Future work

This study has uncovered the general state of genetic diversity in *P. excelsum*, with the findings creating several pathways for future work. Further sampling across the whole range of *P. excelsum* would strengthen and validate my findings. More sampling in the upper North Island would be useful for investigating the presence of potential refugial populations harbouring more diversity than those in this study, as it is suggested *P. excelsum* only survived in Northland during the LGM due to frost intolerance. Sampling across the species range would provide more thorough investigation into the unusual grouping of geographically separated sites found at *K*=3. Sampling in the South Island would be interesting for further study on the influence of the LGM on *P. excelsum* distribution.

This population genetics study on *P. excelsum* can provide important insight to a range of areas. As it is not possible to undertake a population genetics study on all species, it would be worthwhile to complete population genetics studies on several species with a range of life history traits. This would hopefully lead to the identification of broad patterns of species spatial and genetic distribution, and consequently allow a more accurate prediction of species ability to cope in a changing environment. This is important as species can become vulnerable for various reasons, with climate change and introduction of pests and diseases likely to impact the New Zealand flora and fauna. Here, this study can contribute evidence to understanding the broad impact of historical events, such the LGM on frost intolerant, lowland forest trees and can help predict the influence of a changing climate on New Zealand native plants, particularly those with similar characteristics to *P. excelsum*, such as frost

intolerance or widespread spatial distribution. A study by Lakeman-Fraser and Ewers (2013) demonstrated that *P. excelsum* can survive further south than its current southern range limit of Banks Peninsula and grow to a larger size than naturally occurring *P. excelsum,* likely due to the absence of *C. scriptaria*, suggesting some ability to tolerate colder conditions.

Furthermore, the molecular markers used to characterise genetic variation in population genetics studies have varying benefits and disadvantages. To further investigate and provide a broader understanding of the genetic variation within *P. excelsum*, it would be beneficial to utilise an additional marker type. For example, SNPs are stable, and more common in the genome than microsatellites (Park et al., 2021). Therefore, they are more suitable for Next Generation Sequencing (NGS), such as Genotyping By Sequencing (GBS) and Restriction Site Association Sequencing (Rad-Seq), which are used to identify and sequence SNPs (Zhang et al., 2020). GBS and RADseq are ideal methods of sequencing SNPs as a reference genome is not needed (Al Salameen et al., 2022; Andrews et al., 2016). These technologies have increased the convenience and lowered the costs associated with nucleotide sequencing (Yang et al., 2020) and are considered an efficient method for examining true genetic variation within and between populations (Fischer et al., 2017). Furthermore, comparing the nuclear evidence from microsatellites to markers obtained from the mitochondrial or plastid genome would be beneficial as the genomes vary in how they display genetic variation (Birky Jr, 1988). Chloroplast markers have a lower mutation rate and can be useful for investigating longer scale historical processes (Wang, 2011). The differentiation found using microsatellite markers in this study could be compared to variation found for chloroplast markers, with the often uniparental inheritance providing an opportunity to compare the direct contribution that either pollen or seeds have to overall gene flow (Ouborg et al., 1999).

Knowledge of species genetic diversity, often provided by population genetics studies, is an important aspect of restoration. Restoration and conservation of native forests is essential in New Zealand due to historical obliteration of indigenous forests (Leathwick et al., 2003). It is paramount for the appropriate plants to be eco-sourced, i.e. genetically diverse, and local ecological interactions considered so that the natural integrity of populations can be restored (Reay & Norton, 1999). Although genetic diversity of *P. excelsum* discovered in this study was found to be relatively low, this information should be useful for conservation scientists and land managers and can be used to help eco-source plants for restoration or conservation from sites with slightly higher within population diversity, such as WHP (22) or TET (14) which would also be valuable if the species were to become

threatened by a pest or disease. These more genetically diverse sites would also be an interesting place to study the potential resistance to the species herbivore *C. scriptaria* as well as building on the work of Jayaprakash et al. (2022) who found geographical and seasonal differences in the phytochemical profile of kawakawa.

4.5 References

Al Salameen, F., Habibi, N., Al Amad, S., & Al Doaij, B. (2022). Genetic diversity of Rhanterium eppaposum Oliv. Populations in Kuwait as revealed by GBS. *Plants*, *11*(11), 1435.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81-92.

Balloux, F., & Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, *11*(2), 155-165.

Birky Jr, C. W. (1988). Evolution and variation in plant chloroplast and mitochondrial genomes. In *Plant evolutionary biology* (pp. 23-53). Springer.

Burrows, C. J. (1995). Germination behaviour of seeds of the New Zealand species Fuchsia excorticata, Griselinia littoralis, Macropiper excelsum, and Melicytus ramiflorus. *New Zealand Journal of Botany*, *33*(1), 131-140. https://doi.org/10.1080/0028825x.1995.10412950

Butts, C. A., Klink, J. W. v., Joyce, N. I., Paturi, G., Hedderley, D. I., Martell, S., & Harvey, D. (2019). Composition and safety evaluation of tea from New Zealand kawakawa (Piper excelsum). *Journal of Ethnopharmacology*, *232*, 110-118.

de la Vega, M. P. (1997). Plant genetic adaptedness to climatic and edaphic environment. Adaptation in Plant Breeding: Selected Papers from the XIV EUCARPIA Congress on Adaptation in Plant Breeding held at Jyväskylä, Sweden from July 31 to August 4, 1995,

de Lange, P. J. (2021). *Piper excelsum subsp. excelsum Fact Sheet (content continuously updated)*. New Zealand Plant Conservation Network. Retrieved 01 May from https://www.nzpcn.org.nz/flora/species/piper-excelsum-subsp-excelsum/

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., 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*(69).

Gottlieb, L. (1977). Electrophoretic evidence and plant systematics. *Annals of the Missouri Botanical Garden*, 161-180.

Hamrick, J., Linhart, Y., & Mitton, J. (1979). Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. *Annual Review of Ecology and Systematics*, *10*(1), 173-200.

Hodge, S., Keesing, V., Wratten, S. D., Lovei, G. L., Palmer, J., & Cilgi, T. (1998). Herbivore damage and leaf loss in the New Zealand pepper tree ('kawakawa', Macropiper excelsum, Piperaceae). *New Zealand Journal of Ecology*, *22*(2), 173-180.

Jayaprakash, R., Ramzan, F., Miles-Chan, J. L., Foster, M., Mithen, R. F., & Pook, C. (2022). Exploring the Chemical Space of Kawakawa Leaf (Piper excelsum). *Nutrients*, *14*(23), 5168.

Köhler, A., & Schmitt, T. (2023). Northern Richness, Southern Dead End—Origin and Dispersal Events of Pseudolycoriella (Sciaridae, Diptera) between New Zealand's Main Islands. *Insects*, *14*(6), 548.

Lakeman-Fraser, P., & Ewers, R. M. (2013). Enemy release promotes range expansion in a host plant. *Oecologia*, *172*(4), 1203-1212. https://doi.org/10.1007/s00442-012-2555-x

Leathwick, J. R., Overton, J. M., & McLeod, M. (2003). An environmental domain classification of New Zealand and its use as a tool for biodiversity management. *Conservation Biology*, *17*, 1612-1623.

Loveless, M. D., & Hamrick, J. L. (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, *15*(1), 65-95.

McGlone, M. S. (1985). Plant biogeography and the late Cenozoic history of New Zealand. *New Zealand Journal of Botany*, *23*(4), 723-749.

Ouborg, N., Piquot, Y., & Van Groenendael, J. (1999). Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology*, *87*(4), 551-568.

Park, G., Choi, Y., Jung, J.-K., Shim, E.-J., Kang, M.-y., Sim, S.-C., Chung, S.-M., Lee, G. P., & Park, Y. (2021). Genetic Diversity Assessment and Cultivar Identification of Cucumber (Cucumis sativus L.) Using the Fluidigm Single Nucleotide Polymorphism Assay. *Plants*, *10*(395).

Reay, S. D., & Norton, D. A. (1999). Assessing the Success of Restoration Plantings in a Temperate New Zealand Forest. *Restoration Ecology 7*(3), 298-308.

Wang, I. J. (2011). Choosing appropriate genetic markers and analytical methods for testing landscape genetic hypotheses. *Molecular Ecology*, *20*, 2480–2482.

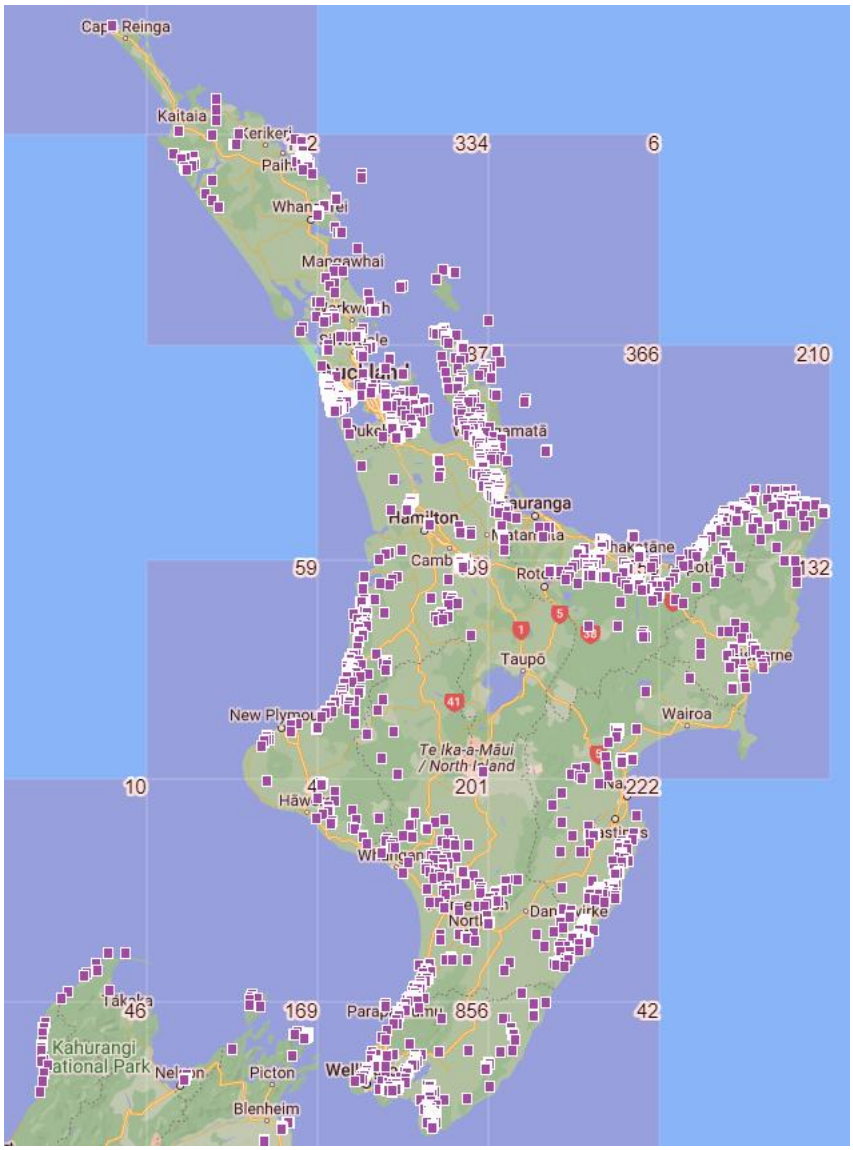
Yang, T.-Y., Gao, T.-X., Meng, W., & Jiang, Y.-L. (2020). Genome-wide population structure and genetic diversity of Japanese whiting (Sillago japonica) inferred from genotyping-by-sequencing (GBS): Implications for fisheries management. *Fisheries Research*, *225*, 105501.

Zhang, J., Yang, J., Zhang, L., Luo, J., Zhao, H., Zhang, J., & Wen, C. (2020). A new SNP genotyping technology Target SNP-seq and its application in genetic analysis of cucumber varieties. *Scientific reports*, *10*(1), 5623-5634.

Supplementary information

Supplementary Table 1: GPS coordinates for the *P. excelsum* collection sites.

Site number	Site ID	Latitude	Longitude
1	MEW	-39.100649	174.11146
2	KAR	-39.148228	173.979941
3	ORK	-37.378079	175.936982
4	AOT	-37.672548	175.916488
7	BRP	-39.961197	175.530687
8	TOR	-40.149626	175.843417
11	AOF	-41.449152	175.235152
12	REM	-41.349319	174.931121
14	TET	-37.813516	175.624054
16	ATS	-39.726801	175.139285
18	WAB	-35.374509	173.599353
19	CLT	-35.423968	173.445512
20	MGA	-37.71611	177.714964
21	MRD	-37.86243	177.630351
22	WHP	-39.286167	176.879336
23	SHF	-39.101142	176.854182



Supplementary Figure 2: New Zealand National Vegetation Survey Databank vegetation survey plots for *P. excelsum* in the North Island. This map approximates the distribution of P. excelsum in the North Island.