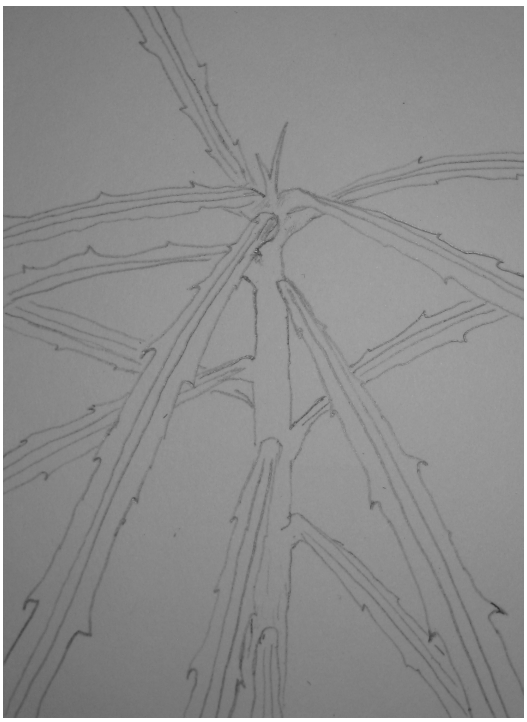
INFERRING THE LATE QUATERNARY PHYLOGEOGRAPHY OF *PSEUDOPANAX CRASSIFOLIUS* USING MICROSATELLITE ANALYSIS



BY

MICHAEL GEMMELL

A thesis submitted to the Victoria University of Wellington
Te Whare Wananga ote Upoko o te Ika a Maui
in fulfilment of the requirement for the degree of Masters in
Ecology and Biodiversity
2013

Abstract

Geologic processes have shaped the New Zealand archipelago throughout its existence. The last major geologic event was the Pleistocene glaciations beginning around 2.5 million years ago. This cold period left its mark in the phylogeography (the geographic distribution of genetic variation) of New Zealand's globally significant biota. Studies into the phylogeography of New Zealand have largely focused on species with limited distributions through rarity or ecological preferences. This study focuses on the ubiquitous species *Pseudopanax crassifolius* (Sol. Ex A. Cunn) K. Koch, also known commonly as Horoeka or Lancewood. This species is widespread and almost continuously distributed throughout New Zealand giving a broad scale look at the patterns and processes that have influenced the formation of New Zealand's natural history.

Seven microsatellite loci and two rps4 chloroplast haplotypes were utilised to study 247 *Pseudopanax crassifolius* and nine *P. chathamicus* individuals sampled from populations from around New Zealand. *Pseudopanax crassifolius* was found to have levels of genetic diversity and overall differentiation consistent with common widespread trees. The genetic structuring suggests *P. crassifolius* is not a single homogenous population across a southern cluster. The geographic structuring of genetic variation within these clusters is poor.

The genetic patterns and the spatial distribution of these patterns may reflect the response of *Pseudopanax crassifolius* to changing environmental conditions during the late Quaternary following the maximum extent of the last glacial maximum (LGM) period. During the maximally cold periods of the LGM, *P. crassifolius* is likely to have been eliminated or at least greatly reduced in the south and west coast of the South Island. In the remainder of the South Island and throughout the North Island it remained widespread. The heterogeneous pattern of genetic variation with little geographic correlation in the northern cluster may reflect either the extent of the historic distribution of the species or the effect of gene flow between populations acting to inhibit population structuring from establishing. The reduction in genetic diversity and the homogeneity of structure in the south indicate a pattern of leading edge re-colonisation into southern areas as conditions became more favourable following the

LGM. The leading edge mode is supported by asymmetric introgression of rps4 haplotype seen between *P. crassifolius* and *P. ferox* along the east coast of the South Island.

This study also investigated levels of differentiation between *Pseudopanax crassifolius* and *P. chathamicus*. There is limited evidence of differentiation based on microsatellite markers. There is therefore no strong genetic evidence for either the support or rejection of the current species delimitation of the crassifolius group of *Pseudopanax* species. The two species are morphologically different and geographically isolated. This, alongside evidence from previous studies suggest that *P. chathamicus* is possibly an example of a group undergoing incipient allopatric speciation. A recent founder event is proposed with enough potential diversity carried in two individual fruit to account for the diversity seen in *P. chathamicus*.

Acknowledgements

Without the help and support of many people this project would have been not have been possible. I would like to take this opportunity to thank those who have helped me. Firstly special thanks to my supervisors Leon Perrie and Joe Zuccarello whose instruction guidance and feedback have been invaluable. Thank you also to Lara Shepherd whose assistance in all areas was of great help

I am thankful to Museum of New Zealand Te Papa Tongarira for granting me their MSc scholarship in molecular systematics.

Thank you to others who contributed to the completion of this project. Fellow students at VUW including the bioinformatics working group and my lab mates in KK612 for help and guidance in bioinformatics software use and advice with lab techniques. NZPCN for allowing me to use their data regarding distribution. Wellington Regional Council, Zealandia and D.O.C for permission to collect samples and to those people who carried out the bulk of sample collecting and DNA extraction. Thank you to those who helped proof read.

Thanks to friends and family and flatmates.

Table of Contents **Page**

Abstract **III**

Acknowledgements **IV**

Table of Contents **VI**

List of Tables and Figures **VIII**

Chapter One General Introduction **1**

1.1 Phylogeography 1

1.2 Population Genetics 1

1.3 Geologic History of New Zealand 3

1.4 Reconstructing the Past 4

1.5 Range Restriction 6

1.6 Small vs. Large Refugia 10

1.7 Repopulation 11

1.8 Distribution Gaps 12

1.9 Microsatellites 13

1.10 Study Aims 14

Chapter 2 Materials And Methods **17**

2.1 Sample Collection 17

2.2 DNA Extraction 18

2.3 Primer Development 19

2.4 PCR 20

2.5 Chloroplast Genotyping 21

2.6 Gel Electrophoresis 22

2.7 Population Delimitation 23

2.8 Microsatellite Validation and Variation 23

2.9 Allele Frequencies 24

2.10 Population Structure 24

2.11 Analysis of Molecular Variance 25

2.11.1 Amova 25

2.11.2 SAMOVA 25

2.12 Bayesian Analysis 26
2.13 Gene Flow and Demography 26
2.14 Bottkeneck 27
2.15 Selfing and Seed Diversity 27
Chapter 3 Results 28
3.1 Microsatellite Validation And Variation 28
3.2 Microsatellite Diversity 29
3.3 Population Structure 31
3.4 Population History 41
3.5 Chloroplast Haplotypes 41
3.6 Parental Analysis 42
Chapter 4 Discussion 43
4.1 Population Genetics 43
4.1.1 Diversity 43
4.1.2 Differentiation 46
4.1.3 Genetic Structure 47
4.2 Distribution 51
4.3 Ecological Tolerances 55
4.4 Chloroplast Haplotype 56
4.5 *Pseudopanax Chathamicus* 59
4.6 Future Directions 60

References 62

Appendices 80
Appendix I Allele frequencies for population clusters 80
Appendix II SAMOVA Voronoi diagram output 83
Appendix III Microsatellite data from parental analysis 84
Appendix IV Microsatellite data included CD

List of Tables

Table 2.1 Microsatellite loci and primer characteristics used in this study. 20

Table 3.1 Genetic characteristics of the seven microsatellite loci used. 29

Table 3.2 Population diversity measures of populations sampled 30

Table 3.3 Pairwise F_{ST} comparisons between populations with five or more samples. 32

Table 3.4 Pairwise D_{EST} comparisons between populations with five or more samples 33

Table 3.5 Average Q per population when K=2 35

List of Figures

Figure 1.1 Glacial and vegetation zones in New Zealand during LGM taken from Alloway et al. (2007). 5

Figure 1.2 Mean annual temperature and total sunshine hours for New Zealand. 9

Figure 1.3 *Pseudopanax crassifolius* distribution. 15

Figure 2.1 Sample locations and population groupings. 18

Figure 3.1 Delta K as estimated by Structure Harvester for *P. crassifolius* 34

Figure 3.2 STRUCTURE Bar graph for all *Pseudopanax crassifolius* samples. 36

Figure 3.3 STRUCTURE Bar graph for *P. crassifolius* and *P. chathamicus*. *37*

Figure 3.4 Delta K as estimated by Structure Harvester for *P. crassifolius, P. chathamicus* and *P. ferox*. 38

Figure 3.5 STRUCTURE Bar graph for *P. crassifolius* and *P. ferox* 39

Figure 3.6 Isolation by distance chart F_{ST} against ln distance 40

Figure 3.7 rps4 haplotype distribution 41

Figure 4.1 Inferred cluster boundaries from STRUCTURE output. 48

Figure 4.3 Expected genetic pattern following leading edge re-colonisation from Trewick *et al* (2011). 54

Pseudopanax crassifolius *growing as a garden plant at Victoria University Wellington.*

Chapter 1: General Introduction

1.1 Phylogeography

The scientific field of phylogeography, initially conceived by Avise *et al.* (1987), is a holistic approach allowing greater understanding of the causes and processes leading to the present genetic patterns in the world's biota and their relationships with the geography of the enviroments they exist in. Research into phylogeography is growing exponentially (Beheregaray 2008 and references therein). As biologists elucidate phylogeographic patterns in more species, we gain a better understanding of processes that have influenced genetic diversity and structure (Marske *et al*. 2012). A more complete understanding of historical processes enables us to predict how these processes will continue to shape the world we live in, as evolution continues to shape the species present and as the species affect the environments they inhabit.

1.2 Population genetics

New Zealand has an international reputation for its exceptional natural history. With high levels of endemicity in both its flora and fauna (Wilton & Breitwieser 2000), it is one of the world's biodiversity hotspots (Myers *et al*. 2000). New Zealand's biota is unique and vulnerable to anthropogenic pressures such as those instigated by human habitation, both at local scales through deforestation and introduction of novel species and global scales through climatic change. It is important to gain insights into the patterns and processes that exist amongst New Zealand biota to be able to understand, predict and respond to future changes.

Genetic diversity is one of the metrics used to measure the health of a species. There is wide consensus that it has an important role in determining the vulnerability of species to extinction risks. High levels of genetic diversity give an indication of a species' robustness and an ability to cope with disease and ecological stressors (Booy *et al*. 2000; Jamieson & Allendorf 2012). Tools such as the Minimum Viable Population (MVP) have been developed to predict a

species extinction risk (Franklin 1980), which is based on underlying principles of genetics and maintaining diversity. Common measures of diversity used allelic richness and proportion of private alleles. Generally large populations harbour more diversity than smaller populations (Frankham 1996) and therefore smaller populations face more risk than larger ones (Ellstrand & Elam 1993).

The processes influencing population genetics are understood from fundamental principles of genetics. A combination of life history traits, selection pressure and stochastic processes all contribute to the evolution of a species. Stochastic processes have a stronger influence on small populations. For example, genetic drift, which leads to the fixation or loss of alleles from a population thereby reducing diversity, is more influential on small populations. In large populations there is a greater probability of mutations occurring and being maintained in the gene pool. This acts to maintain a higher level of genetic diversity. In small populations there is more chance of inbreeding and inbreeding depression, which can contribute to a reduction of variability. The genetic diversity within New Zealand trees is often low (Hawkins & Sweet 1989; Billington 1991; Haase 1992; Barnaud & Houliston 2010), which might indicate population bottlenecks. However some New Zealand trees show greater diversity, a study of *Pseudopanax ferox* Kirk showed that small populations can harbour diversity (Shepherd & Perrie 2011).

How genetic patterns are structured geographically can give valuable insight into the history of a species. How much structuring is present and why the structuring has formed are all part of the puzzle in trying to understand phylogeography. While the general trend is one of weak structure and diversity in New Zealand trees (Young *et al*. 2001; Broadhurst *et al*. 2008; Barnaud & Houliston 2010), *Pseudopanax ferox* shows strong structure (Shepherd & Perrie 2011). This contradictory pattern can be reasonably explained by the disjunct pattern of *P. ferox* populations, which are restricted to small isolated populations throughout New Zealand. Population differentiation and therefore structure is expected to be higher in disjunct populations compared to those species with more continuous distributions (Muller *et al*. 2009)

Structuring is determined by the amount of gene flow. High gene flow inhibits the formation of population structure. Trees are associated with high gene flow compared to other plant forms, in part due to their tall habit and pollen and seed dispersal mechanisms (Petit & Hampe

2006). There are other factors that can also contribute to the lack of genetic structure and differentiation. Barriers to gene flow such as human mitigated deforestation in New Zealand probably developed too recently to be detectable genetically (Barnaud & Houliston 2010). Older barriers such as Southern Alps are readily discerned in the genetic record (Liggins *et al*. 2008). The effects of more recent barriers like the Pleistocene glaciations that resulted in glaciers extending to sea level and outwash plains have also been observed in some species (Trewick & Wallis 2001; Marshall *et al*. 2009; Bunce *et al*. 2009). Detecting recent barriers to gene flow is especially problematic when using slowly evolving genetic markers. In such cases rapidly changing neutral genetic markers should be used, for example single nucleotide polymorphisms (SNPs) or microsatellite markers.

1.3 Geologic History of New Zealand

Since its break off from the Gondwana supercontinent around 80 million years ago (mya) (Wallis & Trewick 2001; Winkworth *et al*. 2002), New Zealand's tumultuous geology is intrinsically linked to the evolution of this archipelagos globally significant biodiversity (Myers *et al*. 2000), These geologic events include the Oligocene inundation (≈30 mya), the building of the alpine regions (≈7 mya) and the Quaternary glacial period beginning (≈2 mya). The legacy of these geologic events is shown in the genetic make up of the extant species.

During the Oligocene inundation, much of New Zealand was submerged as sea levels increased (Cooper & Cooper 1995; Gibbs 2006). Debate continues over the extent of the inundation (Waters & Craw 2006; Knapp *et al*. 2007), but the general consensus is that there was at least an archipelago remaining above sea level. The reduced land area and the restricted habitat types in the first 70 million years of New Zealand's natural history had a significant impact on the resulting biodiversity. For example, the height of the highest peaks would have been around 200 meters, with none of the alpine and subalpine ecosystems seen today (Daugherty *et al*. 1993; Cooper & Cooper 1995; Gibbs 2006; Knapp *et al*. 2007).

The next major influence on the biota was the mountain building period during the Pliocene beginning circa 10 mya. During this uplift the Southern Alps were formed, bisecting the South Island (Cox & Findlay 1995). Mountain ranges also extended into the North Island along the

major fault lines. In contrast to the restriction to the biotas, as happened during the Oligocene drowning, new habitats and niches were created; notably were the alpine and subalpine regions. The formation of these major geographic features also altered weather patterns. The most significant change was the creation of more orographic rain on the west coast and a rain shadow to the east (Winkworth *et al*. 2005; Liggins *et al*. 2008; Meudt *et al*. 2009). The conditions caused by these events in turn led to many speciation events (McGlone 1985; Winkworth *et al*. 2005).

The Pleistocene glaciations are the most recent major geologic processes to affect New Zealand. During this time the world was in a cyclic pattern of climatic conditions. The climate would alternate between cooler glacial periods lasting in the order of 80,000 years to shorter warmer interglacials of around 10,000 years (Dobrowski 2011). In New Zealand the onset of the Last Glacial Maximum (LGM) was at 29 – 26 kya. The extent of glaciation and different vegetation zones in New Zealand are shown in Fig. 1.1 The LGM lasted until the warming phase, which started at 18 kya. The warming has not been continuous. At 14 – 12.7 kya there was a reversal of the warming that coincided with the Antarctic cold reversal. Warming then continued until another small decline in temperature at 9 – 7 kya. Warming has continued since and does so today (Kaplan *et al*.; Alloway *et al*. 2007; Putnam *et al*. 2010)

1.4 Reconstructing the Past

The fossil record and our understanding of the geologic processes together with the genetic evidence, allow us to gain excellent insight to the history of life on earth. The fossil record provides an invaluable tool in the determination of the how species composition has evolved in a specific area over time. Unfortunately, for a variety of reasons, the fossil record is frequently incomplete and not always truly representative of the historical biota (Anderson *et al*. 2006). In some cases pollen arrives by long distance dispersal, conversely some plants reduce or stop pollen production during stressful conditions. This creates a disparity between the palynological record and the actual species composition at the time (Hicks 2006; Birks & Willis 2008). This is where analysis of the genetics of extant species provides a tool for understanding the processes acting in the past. Encoded in the DNA are clues to population structure and the interconnectivity of historical populations. Whilst we can largely only look at

extant species, we can infer a great deal about what might have influenced the present species composition.

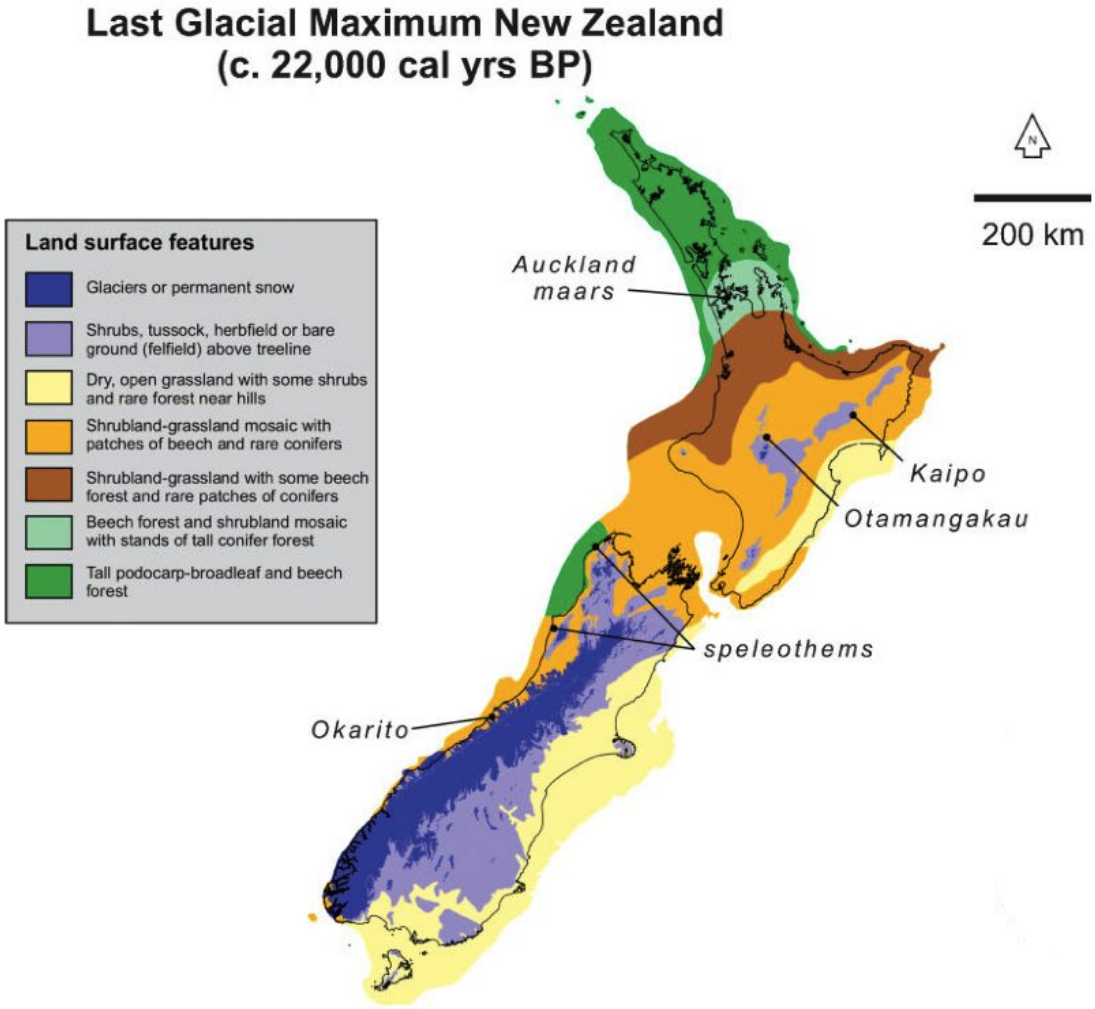


Figure 1.1. Glacial and vegetation zones in New Zealand during LGM inferred from the pollen record. Taken from Alloway et al. (2007). Present coast line shown as black line for comparison to LGM land area.

1.5 Range Restriction

As habitats change so does species composition. Species distributions will change with changes in the suitable range of environmental conditions. This is provided there is space to do so, and the rate of environmental change is not faster than the rate at which the species can disperse or migrate. Extirpation from areas can be through physical means, caused by loss of suitable land due to glaciations, or physiological whereby survival or reproduction are restricted through harsh climatic conditions. During the LGM, temperatures and climate were different to those of today. In New Zealand temperatures were on average depressed 4.5–6°C and the snowline lowered by 800–1000 m (Wilmshurst *et al*. 2007). Models suggest that the east of the South Island underwent the greatest cooling (Drost *et al*. 2007). The differences were more pronounced in the Northern Hemisphere where temperatures may have been lowered by 10–20°C at high latitudes and 2–5°C in low latitudes (Petit *et al*. 1999). This disparity was probably due to the moderating effect of a maritime climate on the islands of New Zealand. During the cooler periods the forests were much more restricted and confined to lower latitudes (Fig.1.1). In the warmer interglacial periods, the woody forests thrived and their distribution pushed into higher latitudes while alpine plants were restricted to refugia (Birks & Willis 2008).

There is uncertainty about the duration of the very coldest temperature periods within the cooler glacial periods. If these coldest periods were of short enough duration, plant populations might have been able to survive *in situ*. Even if they were not able to survive and reproduce normally as a result of the extreme temperatures, populations might have survived in either an un-reproductive or minimally reproductive state. This state can maintain refugia until conditions return to the suitable range where the population can continue as normal (Birks & Willis 2008). Another mechanism by which refugia could be maintained is through microclimates. These are small areas where conditions are more favourable than the predominant conditions. An example of this is a sheltered valley with a favourable aspect, providing, a warmer habitable zone in an otherwise cold climate.

During the LGM in New Zealand, forest was the predominant ecosystem in the northern third of North Island and western Nelson. Elsewhere in the South Island there were only a few

isolated pockets of forest, and the remaining vegetation was dominated by grass and shrubland (Alloway *et al*. 2007; see Fig.1.1). The presence of these large forest refugia is supported by the rates of endemicity. Patterns of endemicity can give an indication as to which regions have had the most ecological stability. Stable environments can allow time for both genetic diversity and specific diversity to accrue (Wardle 1963). The north of the North Island and the two ends of the South Island, have relatively high specific endemicity. The alpine region of the Southern Alps has high specific endemicity which could indicate refugia for cold tolerant species during periods of warm climactic conditions (McGlone 1985).

The ecology of New Zealand in recent prehuman times was one of widespread forestation, although there are residual artefacts of the previous cooler climates in the distribution of different species. North of Taupo endemic species have a high proportion of woody trees and forest dwelling fauna. South of Taupo there is a higher proportion of shrubby and herbaceous species (Gibbs 2006). The cycling of conditions during the Pleistocene would have favoured the survival of species that were reasonably tolerant of a range of conditions (Hewitt 2004). This may explain the extinction of *Acacia*, eucalypts and many Proteaceae species in New Zealand (Lee *et al*. 2001; McGlone *et al*. 2001). It is likely that subtropical species that require warmer temperatures and /or moist conditions did not to survive the coldest and driest periods, while more temperate and drought resistant species did. It is possible that cold adapted species could retreat to high altitude refugia during the warmest periods.

Different climatic agents can act on species at varying levels or in concert (Chapin & Shaver 1985; Stewart *et al*. 2010). Wardle (1991) noted that the natural ranges of many New Zealand species do not extend as far as their physiological tolerances might allow them to. Knowing species' limiting factors are can help us infer what conditions drove species' geographic present boundaries. This could indicate which areas were refugia, and which might become refugia, as the climatic cycles change in the future (Dobrowski 2011). This assumes a continuity of physiology over time, usually a valid assumption for shorter time periods. However as the time frame increases so do the effects of evolutionary pressures, making these assumptions less reliable.

An increase in southerly winds, which brought the cold and dry conditions, could have been

the major factor affecting New Zealand plant communities during the LGM (Drost *et al*. 2007). The frost-sensitive genus *Metrosideros* is an example where temperature can be a limiting factor (Gardner *et al*. 2004). The moderate reduction in temperature during the LGM in New Zealand suggests that temperature was not a significant limiting factor for some cold-tolerant species such as *Pseudopanax ferox* (Shepherd & Perrie 2011). Instead the aridity of the time is likely to have been an important restrictive factor. Analysis of New Zealand's pollen using principal component analysis has shown that precipitation levels are an important factor influencing species composition (Norton *et al*. 1986). This has also been shown to be the case in other regions also such as North America (Stephenson 1990; Lutz *et al*. 2010). That moisture is a major variable influencing New Zealand's plant distribution fits well with the estimations of the climate during the LGM being drier and winder than present conditions, but not much colder (Wilmshurst *et al*. 2007), yet still having wide ranging effects on the distribution of the vegetation in New Zealand.

It remains to be explained how well these factors influence the distribution of *Pseudopanax* K. Koch species which are the focus of this study. Their distribution is only very weakly associated with Mean Annual Temperature (MAT) and October Vapour Pressure Deficit, a measure indicating how much water the air can hold versus what it is holding and can be used as a proxy for moisture levels (Wilmshurst *et al*. 2007). This weak association potentially indicates that the genus has some robust species, tolerant of a range of conditions. It is difficult to distinguish between *Pseudopanax* species in the palynological record meaning it cannot be used to make strong inferences about the historical distribution of the different species. However, given that *Pseudopanax crassifolius* (Sol. ex A. Cunn.) K. Koch is widely distributed latitudinally at present, MAT might not have been a limiting factor to its distribution during LGM. *Pseudopanax* therefore may have been able survive extensively or within numerous refugia during cooler periods as opposed to *Metrosideros* species which have higher negative correlation to MAT, and are more likely to be restricted to moister northern areas as reflected by findings made by Gardner *et al.*(2004).

The consensus is that there was some forest throughout New Zealand during the LGM. More extensive forests were present in the north of the North Island and the north of the South Island. This is shown in the fossil record and genetic signatures of forest and forest obligate

species (Fig. 1.1, McGlone 1985; Gardner *et al*. 2004; Alloway *et al*. 2007; Shepherd *et al*. 2007; Leschen *et al*. 2008; Marshall *et al*. 2009; Shepherd & Perrie 2011). These regions are the warmest in each of New Zealand's two major islands with the highest levels of insolation (Fig. 1.2). Areas outside the northern parts of the two main islands seem to have had at least some level of woody vegetation (Alloway *et al*. 2007; Shepherd *et al*. 2007).

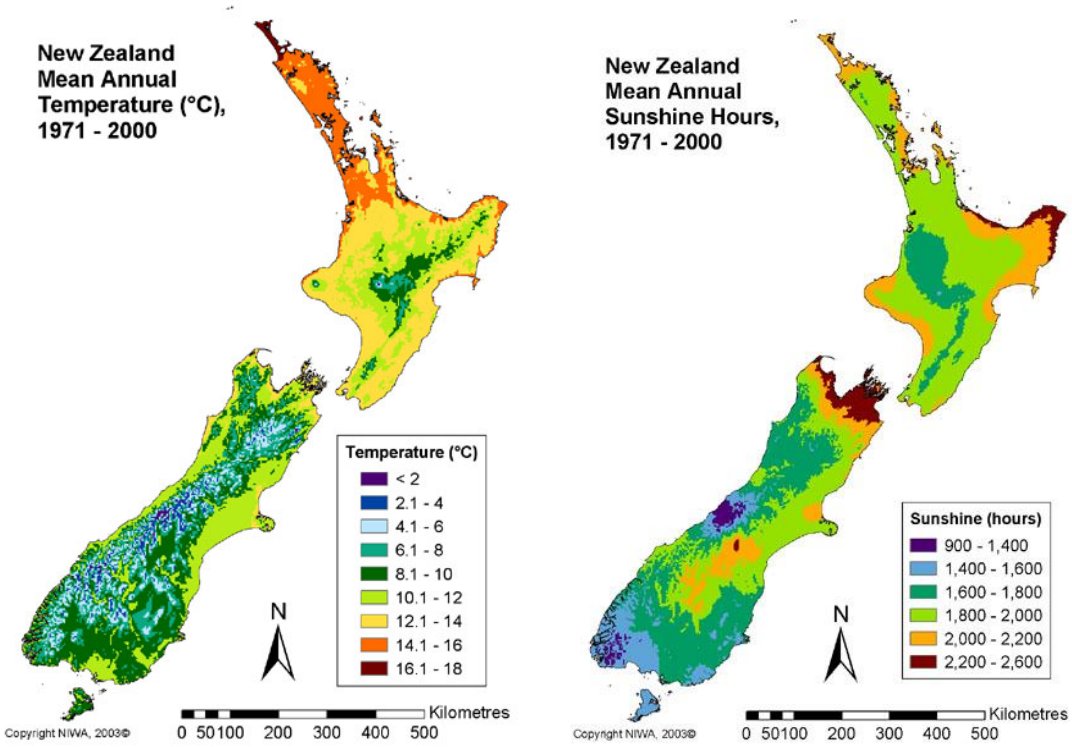


Figure 1.2 Mean annual temperature and total sunshine hours for New Zealand. Image credit NIWA 2003

The classical idea of what happened to species during the glaciations is that biota was pushed into increasingly lower latitudes as the ice sheets and associated adverse conditions advanced. As species were pushed to physical boundaries (such as the ocean for terrestrial species) their ranges became restricted or they became extinct as no suitable habitat remained (Fraser *et al*. 2012). In the Northern Hemisphere where glaciation was more extensive than in the Southern Hemisphere (Suggate & Almond 2005). The European biota was pushed southwards or restricted to refugia in Iberia, Italy and the Caucuses. This is shown by high genetic diversity in those areas. Re-colonising following glacial retreat is shown by decreasing diversity in populations to the north (Hewitt 1999). In New Zealand there is some evidence for a similar pattern in the genus *Metrosideros,* marine algae and the short-tailed bat, which show patterns

of increased diversity in the north compared to the south (Lloyd 2003; Gardner *et al*. 2004; Buchanan & Zuccarello 2012). The haplotype distribution in a New Zealand orb web spider shows a similar north south pattern. Northern most populations show ancestral haplotypes and other haplotypes have developed further south indicative of north to south dispersal pattern (Opell 2006).

1.6 Small vs. Large-scale Refugia

There are two most commonly accepted patterns of species survival during periods of unfavourable conditions. Either large scale survival occurs but where the range of the species has been shifted or restricted to areas of generally favourable conditions, or small scale survival occurs in cryptic or micro refugia, where species survive in small areas of favourable conditions remain within areas that have become uninhabitable. Micro-refugia are increasingly recognised as important in understanding the ebb and flow of species in response to climate change (Birks & Willis 2008). Increasingly evidence suggests that the model of macro-refugia, whilst being a good account of what happened to some species, is perhaps too simplistic or at least inadequate to explain the patterns seen in the species studied so far. There is support for the idea that many species survived in smaller disjunct refugia which were often in small restricted pockets within their former ranges (Drost *et al*. 2007; Birks & Willis 2008; Parducci *et al*. 2012). There also seems to be no strict consensus with different patterns seen amongst the species studied even in close geographic proximity. For example, the pattern of few large refugia in northern localities for *Metrosideros* trees (Gardner *et al*. 2004) seems quite different from the pattern of many complex microrefugia in the fern *Asplenium hookerianum* (Shepherd *et al*. 2007). As *A. hookerianum* requires some level of canopy protection its survival in widespread populations as inferred from genetic analysis would seem to indicate a similar survival of some level of forestation across its distribution (Shepherd *et al*. 2007).

1.7 Repopulation

The conditions during the climatically variable Pleistocene swung from one extreme to the other. As a result the changing distribution of biota was largely due to re-colonisation into newly exposed suitable environments from areas where populations had remained extant, such as refugia or beyond the boundaries of the adverse environments back (Avise 2000; Wallis & Trewick 2001). The rate and pattern of range expansion depends on many factors. Important factors include ecological conditions, dispersal methods, symbiotic relationships, pollination and seed dispersal in some species of plant, diet and habitat specificity in some animals, epiphytes and parasites (Hewitt 1999). Pollen records suggest that most of New Zealand was treeless during periods of glacial maxima and that reforestation happened rapidly on a large scale during the warming following glaciation (McGlone & Bathgate 1983; Alloway *et al*. 2007; Wilmshurst *et al*. 2007). Forest obligate species such as the short tailed bat responded in concert with forest expansion and also rapidly expanded their ranges (Lloyd 2003).

Reforestation seemed to have occurred throughout New Zealand without an obvious north to south or altitudinal pattern. This reforestation was rapid at most scales with little evidence of prolonged successional sequences (McGlone & Bathgate 1983). Pollen records can show the transition of the flora from pioneer species through to climax species (Huntley 1990). The absence of a transition in the record might indicate a rapid expansion where successional species are present only briefly. This supports a scenario where forest species expanded from refugia situated throughout the country (McGlone 1985, 1988; Newnham 2012). This is supported by studies of the short tailed bat (Lloyd 2003), *Aspelenium hookerianum* and *Pseudopanax ferox* studies (Shepherd *et al*. 2007; Shepherd & Perrie 2011). Caution should be taken when inferring reforestation and migration rates from the palynological record as previously unrecognised refugia can make the re-forestation rate appear faster than was actually the case (McLachlan & Clark 2004; McLachlan *et al*. 2005). It is also difficult to distinguish between expansion from multiple refugia or a rapid north-south expansion (Ogden *et al*. 1992). Species with efficient dispersal mechanisms generally have a wider distribution, as they have been able to disperse rapidly from refugia or centres of endemism, whereas those with poorer dispersal mechanisms are often confined to the regions near the

refugia (McGlone *et al*. 2001).

1.8 Distribution Gaps

Disjunctions between populations can occur from localised extinction of a species in intervening areas. This might result from transient events such as fire or severe frosts, or from longer term events such as the disappearance of suitable habitats resulting from tectonic processes or climatic changes. Geological processes can lead to disjunction over longer periods of time through the segregating of habitat. An example of this is the north-west Nelson region and southland Fiordland, which were once adjacent and are now vastly separated through the action of plate tectonics (McGlone 1985). Disjunctions can also arise from lack of re-colonisation. A putative example of this is the "beech gap" whose presence is explained by the effect of the glacial period, but whose persistence is not so easily explained (Hall & McGlone 2006; Leschen *et al*. 2008).

Some species may have had extensive ranges during the warmest periods since the LGM, but with the cooling since 9000 ya may have experienced range restriction (McGlone 1985). Recent range restriction has been proposed as an explanation for the disjunction in *Pseudopanax ferox. Pseudopanax ferox is* thought to have thrived immediately post-glaciation but became restricted as the climate has shifted and soil nutrients have became depleted (Shepherd & Perrie 2011).

The pattern of genetically disjunct populations in *Pseudopanax ferox* (Shepherd & Perrie 2011), is not expected to precisely match the pattern in *P. crassifolius. Pseudopanax crassifolius* is presently a widespread tree, therefore such clearly defined distinct genetic boundaries are not expected. We would expect to see either one population with high diversity with decreasing amounts of variation further from the centre of diversity, or we would expect to see multiple areas of high diversity (Shepherd *et al*. 2007). Possibly regions of hybridisation/introgression could be found where the population boundaries meet or overlap (Hewitt 1999; Gardner *et al*. 2004). Whilst ranges of many species were restricted during the last glacial maxima, this does not preclude their ranges/distribution remaining large but disjunct.

1.9 Microsatellites

Microsatellites are short chain tandem repeat DNA sequences with repeat units 1–6 base pairs long with dinucleotide repeats being the most common. Repeat sequences are commonly between 5 and 40 repeat units in length. They are highly polymorphic and ubiquitous in genomes of most taxa. (Li *et al*. 2002; Ellegren 2004; DeYoung & Honeycutt 2005; Selkoe & Toonen 2006). They provide information regarding genetic structure from between populations down to the level of the individual. They are bi-parentally inherited compared with uniparentally inherited mtDNA or cpDNA and are rapidly mutating. Their high levels of polymorphism provide information that would not be available in more stable parts of the genome (Angers & Bernatchez 1998; Ellegren 2004). Microsatellite markers could be a useful tool to help understand pollination dynamics in wild populations due to the bi-parental inheritance, determining whether plants are selfing or interbreeding with siblings or offspring or outbreeding. Manipulative experiments have been carried out to see it this is a viable investigation (Kelly & Willis 2002).

The rapid mutation of microsatellites, however, brings up some important considerations in their analysis. The mutation process must be taken into account. Two models have been developed to account for this. The Independent Allele Model (IAM) (Kimura & Crow 1964)and the Stepwise Mutation Model (SMM) (Kimura & Ohta 1978). The IAM assumes that each mutation creates a new allele of a size independent of the allele it evolved from. This has been the most commonly used model. The SMM takes into account the size of the source allele. Both models have benefits and drawbacks. The SMM is regarded as most accurately representing the mutation processes of microsatellites whereas the IAM is thought to be more robust to non-stepwise processes in mutations and therefore though to be more reliable (Balloux & Lugon-Moulin 2002; Ellegren 2004; Selkoe & Toonen 2006). Whilst their high mutation rate precludes their use in analysis of deep time events microsatellites are becoming an important tool in understanding phylogeography (Harris & Taylor 2010).

1.10 Study Aims

The study species *Pseudopanax crassifolius* is a member of the Araliaceae that consists of 12 species endemic to New Zealand. The species in this genus cover a diverse range of habitats and habits. There are trees and shrubs that inhabit areas from coastal up to low alpine zones. The lancewood group consisting of the species *P. crassifolius, P. ferox, P. chathamicus* Kirk, and *P. linearis*, is found throughout New Zealand up to sub alpine regions (Perrie & Shepherd 2009). Microsatellite markers have been developed for *Pseudopanax* (Shepherd *et al*. 2008). Shepherd & Perrie (2011) investigated the population genetics of *P. ferox* to determine its phylogeography since the LGM. Shepherd and Perrie (2011) raise the question of whether or not the high genetic diversity in *P. ferox* compared to other tree studies is related to biology or methodology. Other investigations into the phylogeography of New Zealand trees have utilised different markers, for example isozyme analysis chloroplast haplotypes and AFLP (Hawkins & Sweet 1989; Billington 1991; Young *et al*. 2001; Gardner *et al*. 2004; Beever *et al*. 2013). By utilising similar methods to Shepherd & Perrie (2011) this study gives a good control in both methodology and certain aspects of biology to investigate the general patterns of genetic differentiation in New Zealand trees, and to infer the post-glacial history of *P. crassifolius.*

Pseudopanax ferox is a closely related species to *P. crassifolius* (Perrie & Shepherd 2009).There are many similarities between the two species. They have similar reproductive strategies (insect pollinated and frugivore dispersed) and both are somewhat successional species (Dawson & Lucas 2012). Both are found throughout New Zealand. One major difference between *P. crassifolius* and *P. ferox* is the distribution. *Pseudopanax crassifolius* is widespread with an almost continuous distribution (see Fig.1.3) in contrast to *P. ferox* which is uncommon and has very disjunct populations (Shepherd & Perrie 2011). The difference in the distribution of the two species has been attributed to slight differences in ecologic preferences. *Pseudopanax ferox* is more tolerant of low moisture levels and may be more suited to more nutrient rich soils (Shepherd & Perrie 2011).

Figure 1.3 *Pseudopanax crassifolius* distribution. Each point represents an individual *P. crassifolius* record. Data used with permission from NZPCN. Compiled using D.O.C. plant database, NZPCN distribution database and a national plant checklist database developed by Graeme Jane.

This present study will also include analysis of *Pseudopanax chathamicus* Kirk (1899) a species restricted to the Chatham Islands. Genetic analysis using AFLP data and chloroplast DNA sequences of P*seudopanax* have given some indication that the *P. crassifolius* might not be monophyletic with *P. chathamicus* nested within *P. crassifolius* (Perrie & Shepherd 2009). Analysis of microsatellite markers will be performed to clarify the delimitation of these two species.

Understanding the genetic diversity of a species is an important step in the conservation effort (Taylor *et al*. 2010). While *Pseudopanax crassifolius* is not threatened, it is a member of a clade that includes naturally uncommon species *P. linearus* and *P. ferox.* Comprehensive

understanding of the genetics and the interactions within this hybridising genus is important for the conservation of New Zealand's biodiversity, both in efforts to understand and maintain genetic diversity and in understanding what is happening at the species boundaries. As noted in (Liggins *et al*. 2008) most phylogeographic studies in New Zealand tend to focus on species with limited distribution, whether through ecological tolerances or through rarity. *Pseudopanax crassifolius* is widespread and ubiquitous and as such can provide a valuable insight to broader scale processes.

The aims of this study are to gain an understanding of the genetic patterns present in *P. crassifolius.* Genetic diversity will be explored along with the structuring of this genetic diversity to infer as much as possible about the late Quaternary phylogeography of the species.

Chapter 2: Materials and Methods

2.1 Sample collection

The primary focus of this study was *Pseudopanax crassifolius*. The closely related species *P. chathamicus* was included in the study as it has been reported to be nested within *P. crassifolius* (Shepherd *et al*. 2008). Samples of *P. ferox* were also included to aid comparison for some analyses.

Samples had been previously collected from around New Zealand to get samples from populations representative of the natural distribution of *Pseudopanax crassifolius (*see Fig. 2.1 for sample locations) In this study a total of 268 samples from 26 putative populations were used; 247 *P. crassifolius* samples, 9 *P. chathamicus*, 12 *P. ferox.*

A suitable tree would be identified and notes taken of its habit and habitat. Its location was also recorded either by GPS or map reference. Ideally the selected tree was a mature adult. Branch tips with new growth at the apex were chosen for sampling as the softer tissue enabled easier DNA extraction. The branch tip was collected with enough leaves to allow DNA extraction (see extraction section) and to serve as a voucher specimen to be accessioned into the herbarium of the Museum of New Zealand Te Papa Tongarewa (WELT) if deemed necessary. The voucher specimen can be used for future reference, which is important if there is any subsequent doubt about the identification of the sampled tree. There is a large amount of hybridisation within the genus *Pseudopanax* and this is especially common between *P. crassifolius* and *P. lessonii* (Morgan-Richards *et al*. 2009; Perrie & Shepherd 2009) and has been noted between *P. crassifolius* and *P. ferox,* albeit rarely (LR Perrie pers com). The hybrids have a large morphological range, intermediate between the two parent species. In some cases especially in the juvenile stage hybrids can closely resemble one of the parent species. Care was taken in the field to select pure *P. crassifolius* with typical diagnostic features (i.e. unifoliolate leaves and no indication of a joint between the petiole and blade, a feature derived from the *P. lessonii* parent).

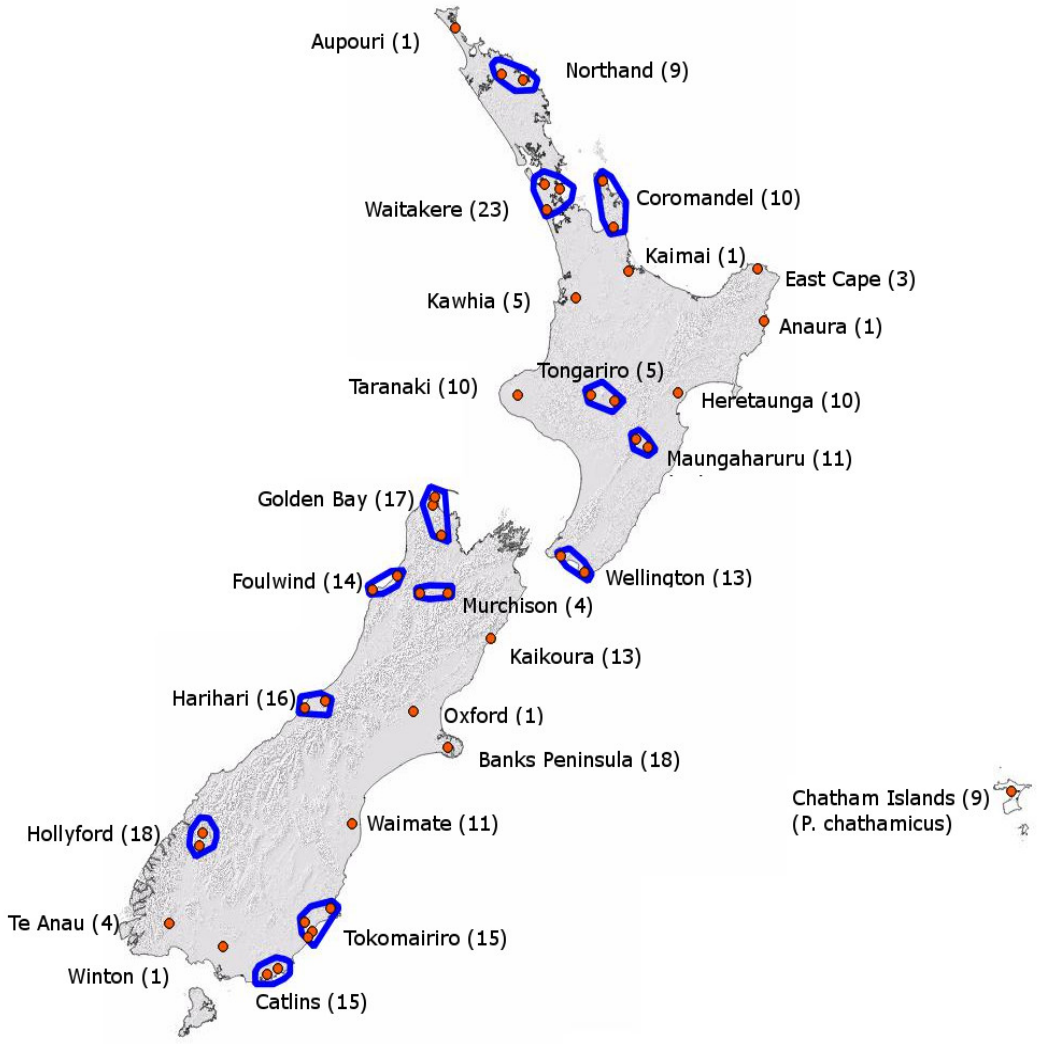


Figure 2.1 Approximate locations of *Pseudopanax crassifolius* populations sampled for this study. Orange points indicate sample population locations Blue lines indicate sample locations grouped together based on proximity for analysis.

2.2 DNA extraction

DNA extraction was performed following a modified CTAB method based on Doyle & Doyle (1990). Approximately 2cm^2 of leaf tissue was torn into small pieces and placed into a 1.7ml centrifuge tube. The microcentrifuge tube was suspended in liquid nitrogen to make the tissue brittle. The tissue was then crushed into a fine powder using a small pestle. At this stage 600µl CTAB extraction buffer (2% w/v CTAB (cetyl- trimethyl-ammonium bromide, Sigma), 1% w/v

PVP (polyvinylpyrrolidone, Sigma), 1.4 M sodium chloride, 100 mM Tris-HCl pH 8.0, and 20 mM EDTA (ethylenediaminetetra-acetic acid, BDH) was then added. The solution was mixed by vigorous shaking. The tube was then placed into a heating block at 65°C for 40 minutes with occasional agitation. When incubation was completed an equal volume of chloroform was added to the tube. This was mixed vigorously and then centrifuged in a short spin until 10000rpm were reached to separate the solutions. The aqueous phase was then transferred to a new tube. An equal volume of isopropyl alcohol was added to the aqueous solution and the tube was inverted gently to mix the contents. This mixture was then placed on ice for several minutes and if sufficient DNA was present it could be seen to precipitate at the interphase layer. The DNA was removed using another wide mouth pipette tip and transferred to another tube containing 80% ethanol to further remove contaminants. If no DNA was visibly precipitated the interphase layer was removed and put into a tube and centrifuged in a short run up to around 6000 RPM. The DNA would then appear as a smear on the side of the microcentrifuge tube. The DNA was collected and 80% ethanol used to further clean the DNA. The DNA pellet was air dried until no ethanol remained. TE buffer (10 mM Tris-HCl, 1 mM EDTA pH 8.0) was added at a volume based on an visually estimated DNA yield approximately 30-70µl. Samples were stored at -80°C until required.

2.3 Primer Development

Ten microsatellite loci were used for genotyping the samples. Six of these loci were developed for *Pseudopanax* by (Shepherd *et al*. 2008). These six loci were selected as they amplified consistently and were polymorphic in *P. crassifolius*. A further four loci were selected from a pool of 12 candidate sequence derived from unpublished *P. crassifolius* Solexa sequence data of Josie Monaghan (Massey University) and Lara Shepherd (Te Papa). These four were chosen by testing the loci for amplification, at a temperature suitable for future use with M13 tagged fluorescent labels.

Using M13 primers in the tailed Polymerase Chain Reactions (PCR) can reduce costs by removing the necessity of having fluorescent labels attached directly to primers. Different labels can be used for the different primer sets enabling ease of poolplexing as required (Schuelke 2000). Three of the M13 tagged loci were found to have consistently high levels of

null alleles and were discarded from further analysis. The locus that amplified successfully was further amplified using a M13 label protocol for further analysis.

Table 2.1: Microsatellite loci and primer characteristics used in this study.

Locus	GenBank Accession	Repeat motif	Dye	Primer sequence (5′-3′)	T_A (°C)	Size range (bp)	A	H_O	H_E
Psecra2	EU408321	$TA_5GA_3GT_{11}$	HEX	F-GCGGATGGTTAAGAGGTTGA	55	152-182	15	0.695	0.761
				R-GAGAATTCTGCATCACACCCTA					
Psecra3A	EU408322	TG_{13}	6-FAM	F-ATGTTTGTGAGTGTGGC	52	93-113	9	0.680	0.708
				R-CCCCATCTTTTGTCCCTCAT					
Psecra3B	EU408322	TG_{13}	HEX	F-ATGTTTGTGAATTGTGAGTGTGG	50	106-152	25	0.787	0.915
				R-CCCCATCTTTTGTCCCTCAT					
Psecra5	EU408323	TC_7TA_6	NED	F-CCAGCGTCACCTCCATTATT	53	168–216	18	0.538	0.786
				R-TCACAGCCAGCCACTGTATC					
Pseles5	EU408320	GT_8	6-FAM	F-GATCATCGGAAATGGATGCT	52	212-238	12	0.370	0.690
				R-ACATTTTCCCACGTTTGCTC					
Psearb9	HM640018	TC_{22}	HEX	CCAAATGTCCGCACTCACT	50	117-173	22	0.525	0.804
				TCATTCATCAACACAAACGTGA					
M13J	Unpublished	TA10	M13	F-GGAGTTGAAATCACCTGTCCC	55	265-339	17	0.270	0.706
				R-GAGTCGGATTTGCTGACCAC					
Pserps4AF	FJ470223	na	na	F-ATTATTCGCGCCAGACTTAAAC	62				
	FJ470242	na	na	R-CGTACTAACCAGAATAGATCAAATAGC	62				
Pserps4BF	FJ470243	na	na	F-ATTATTCGCGCCAGACTTAAAA	62				
	FJ470243	na	na	R-CGTACTAACCAGAATAGATCAAATAGG	62				

T_A Annealing temperature. A number of alleles. H_O observed heterozygosity. H_E expected heterozygosity.

2.4 PCR Microsatellite amplification.

PCR wasperformed using a Biometra Tgradient PCR machine. Two PCR protocols were followed with each having different master mix recipes and programs. For reactions containing primers with a fluorescent label directly attached to the 5' end (see Table 2.1) master mix recipe and conditions were modified from Shepherd *et al.* (2008). Amplifications were done in 20µl volumes containing 2M Betaine, 500µmol dNTPs, 2µl 10× reaction buffer (New England Biolabs), 10pmol each of the forward and reverse primers, 1U Taq polymerase (New England Biolabs), and approximately 50ng template DNA.

The PCR program for these was; initial denaturation at 94°C for 2 min, followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 50° – 56°C for 1 min, and elongation at 72°C for 50 sec. This was followed by a final elongation step at 72°C for 15min. The optimal

temperature for the annealing step of the reactions was ascertained by carrying out a gradient PCR with annealing temperatures ranging from 48° – 58°C covering the expected optimal ranges for these primers. The optimal temperature was that which produced the brightest bands in the electrophoresis gel (see 2.6) and had no extra bands.

The second protocol was for reactions containing primers with a M13 tag appended to the 5' end and was modified from Schuelke (2000). For these reactions a M13 tag is appended to the 5'end of the microsatellite forward primer and included in the reaction was a fluorescent label attached to the M13 tag. M13 PCR amplifications were done in 20µl volumes containing 2 M Betaine, 500 µmol dNTPs, 2µl 10× reaction buffer (New England Biolabs), 2.5pmol forward primer and 10pmol reverse primer 3.5 pmol M13 label, 1 U Taq polymerase (New England Biolabs), and approximately 50 ng template DNA. The PCR program for this reaction was; initial denaturation at 94°C for 2 min, followed by 30 cycles of denaturation at 94°C for 30 sec, annealing at 56°C for 45 sec, elongation at 72°C for 45 sec. This was followed by the M13 annealing stage with 8 cycles of denaturing at 94°C for 30 sec, annealing at 53°C for 45 sec, elongation 72°C 45 sec, final elongation 72°C 15min.

2.5 Chloroplast Genotyping

Two haplotypes for the *rps*4-*trn*S intergenic spacer, hereafter rps4, have been reported in studies on *Pseudopanax* (Perrie & Shepherd 2009; Shepherd & Perrie 2011). These two haplotypes differ by one transition and one transversion at the (rps4 gene and rps4-trns spacer) locus. Amplification of the rps4 chloroplast haplotypes were performed as per the protocol for labelled microsatellite primers (see above) with the annealing stage temperature increased to 62°C based on the protocol used in (Shepherd & Perrie 2011). Each sample was amplified for both A and B haplotypes using different primer sets. Product was electrophoresed in a 1% agarose gel and scored for either presence or absence of a PCR product for each haplotype. Perrie & Shepherd (2009) recorded two *rps*4 haplotypes in *Pseudopanax ferox*. Each individual assayed was found to only amplify for one of the haplotypes and the haplotypes were separated geographically with no *P. ferox* population containing both haplotypes. Haplotype A was found only in the in the North Island and haplotype B only found in the South Island. To investigate the distribution of rps4 haplotypes

in *P. crassifolius*, samples were assayed from different populations to get a broad geographic spread. Where both haplotypes were found to be present in an area (Fig. 3.7) more individuals were assayed to clarify the relative abundance of the two haplotypes. In the present study a total 80 samples of *P. crassifolius* and nine of *P. chathamicus* were assayed for rps4 haplotypes.

2.6 Gel Electrophoresis

Microsatellite PCR products were electrophoresed in an agarose gel to determine quality of the amplification. 3% agarose gels were made using Seakem Agarose and 1x TAE buffer (40 mM Tris acetate, 1mM EDTA pH 8.0), and electrophoresed within the TAE buffer. PCR product was loaded into gels using loading solution and electrophoresed at constant 95V for 30 min alongside Hyperladder II ladder (Bioline). The gel was then stained with ethidium bromide solution for approximately 15 minutes then fluoresced under a UV transilluminator (wavelength 302 nm, UVP Incorporated) and an image of the gel was captured using a camera and imaging software. This indicated the quality of the amplification and was used as a guide for the amount of product to be used when poolplexing samples to be genotyped. The gel also gave an indication as to whether or not the product was at the expected range when compared to the ladder used. Products of rps4 amplifications were electrophoresed in a 1% agarose gel giving a binary result of presence / absence for either of the rps4 haplotypes.

Prior to genotyping, PCR products were poolplexed together. The poolplex was organised using Multiplex Manager (Holleley & Geerts 2009). Genotyping was performed through the Massey Genome Service using a capillary ABI3730 Genetic Analyser, from Applied Biosystems Inc. The resulting ABI files were examined using GeneMarker v2.4.0. (Holland & Parson 2011). Scoring bins were set by hand for each locus based on an overlay of all peaks across all samples. Each individual was double checked by eye to ensure correct scoring.

2.7 Population Determination

Many of the samples were collected from areas within close geographic proximity such that defining populations was straightforward. In some cases, the number of samples from an area was small and thus necessitated grouping with nearby samples to make a population group of

sufficient size for analysis. (see Fig. 2.1) Some tests are reduced in power when sample sizes are too small e.g. F_{ST} and diversity indices. For F_{ST} tests population with fewer than five samples were omitted. For calculations of allelic richness and private allelic measures, only populations with 8 or more samples were used.

2.8 Microsatellite Validation and Variation

Problems caused by amplification or scoring errors that could result in departures from Hardy-Weinberg Equilibrium (HWE) were tested using Micro-Checker 2.2.3 (Van Oosterhout *et al*. 2004). This was to identify amplification or scoring errors such as null alleles, large allele dropout, or stuttering. 1000 randomisations were used for these tests. To test for adherence to HWE and any linkage disequilibrium (LD) Genepop 3.4 (Raymond & Rousset 1995) was used. 10000 dememorisations and iterations and 1000 batches were used for both HWE and linkage disequilibrium tests. A Bonferroni correction (Rice 1989) was implemented to reduce the chance of type 1 errors. The data were tested for outlier loci using LOSITAN (Beaumont & Nichols 1996; Antao *et al*. 2008) which uses an F_{ST} outlier method. To test if any loci was driving the population differentiation an exact G test was performed in Genepop 3.4 using Markov chain parameters 10000 dememorisations and iterations and 1000. Pairwise tests were performed for each locus and significance levels corrected using the sequential Bonferroni method.

Fstat 2.9.3.2 (Goudet 1995) was used to estimate inbreeding coefficients (F_{IS}) and their significance. Fstat and HP-Rare (Kalinowski 2005) were used to calculate genetic diversity and number of alleles for each population. Allelic richness is calculated based on the smallest population size so populations with small numbers were omitted (see table 2) This left 18 population groups with a minimum number of samples per population group of 8. Bonferroni correction was implemented to reduce the chance of type 1 errors. GenAlEx (Peakall & Smouse 2006) was used to determine expected (H_E) and observed (H_O) heterozygosities of populations and fixation indices.

2.9 Allele Frequencies.

Allelic frequencies were generated for each locus using GenAlEx. These were sorted into groups based on STRUCTURE clusters and species and were as follows North, South and Northwest Nelson. Also included were *P. chathamicus*, *P. ferox* and a final cluster consisting of populations with high frequencies of the rps4 haplotype. This cluster includes the three populations Banks Peninsula, Waimate and Tokomairiro. These charts were generated to look for any differences in allele frequencies between different groups species and to look for signs of correlation between *P. ferox* and *P. crassifolius* in areas where haplotypes are shared indicating putative hybridisation and introgression.

2.10 Population Structure

Differentiation between populations was measured using F_{ST} (Weir & Cockerham 1984). Pairwise comparisons for each population were performed using Fstat and significance levels adjusted using Bonferroni correction. As another comparison for population differentiation SMOGD (Crawford 2010) was used. SMOGD uses D_{EST} (Jost 2008) and G'_{ST} (Hedrick 2005), which are purported to better account for differences in allelic diversity than F_{ST} and G_{ST} especially when within-population diversity is high (Crawford 2010) and references therein). R_{ST} another commonly used measure of differentiation was tested for suitability using an allele permutation test in SPGAGeDI (Hardy & Vekemans 2002). Allele size permutation tests require information about the number of repeat units rather than overall allele length. The genotypes of the data file were coded in such a way that the values represented the number of repeats of the microsatellite repeat unit enabling these calculations. 10000 permutations were performed for comparisons involving populations $n \geq 5$.

Population structure was tested using STRUCTURE (Falush *et al*. 2007), which uses a Bayesian clustering method. STRUCTURE assigns an individual proportionally to a set number (K) of clusters with no prior information. The admixture model was used with Independent Allele Frequency Mutation model (IAM). Which mutation model best represents microsatellite mutations is an issue under much scrutiny and assumptions used in models may not accurately reflect the complex processes involved. While the stepwise mutation model is often regarded as accurately representing the processes in microsatellite evolution the IAM can be

more robust to problems such as imperfect repeats and multiple repeat units (Estoup & Cornuet 1999). The IAM model was used in determining population structure for *Pseudopanax ferox* (Shepherd & Perrie 2011) and was deemed the most suitable model to analyse *P. crassifolius*. K, the number of clusters was set from 2 – 24. 100000 generations of burn-ins were discarded and 100000 generations of sampling were used with 10 simulations for each K. To determine the optimal value of K, Structure Harvester (Pritchard *et al*. 2000; Falush *et al*. 2007; Hubisz *et al*. 2009) was used to analyse the output file from STRUCTURE. Structure Harvester uses the ΔK statistic (Evanno *et al*. 2005) where possible and creates a plot of the mean likelihood values per K and the optimal value for K is determined using log probability of data *(L(K))* and the Delta K estimator.

2.11 Analyses of Molecular Variance

2.11.1 AMOVA

A locus-by-locus AMOVA was performed to determine the level of genetic variation within populations, between populations and within groups. Arlequin 3.5 (Excoffier & Lischer 2010) was used to test the 18 populations with 5 or more individuals sampled. Populations were also grouped based on the STRUCTURE cluster results.

2.11.2 SAMOVA

SAMOVA (Dupanloup *et al*. 2002) can define populations that are geographically proximate but that are genetically separate, and can indicate areas of restricted gene flow using a simulated annealing procedure (Dupanloup *et al*. 2002). Genetic data was analysed alongside spatial coordinates determined using Google Earth version 6.2.2.7373. Data are analysed using an *a priori* number of clusters K. 100 iterations, the suggested number, were used for this analysis. Analysis was performed using a sum of squared size difference microsatellite model. A pairwise DNA model was also run for comparison and returned the same results.

2.12 Bayesian Analysis

Geneclass 2 (Piry *et al*. 2004) was used to perform assignment tests using a Bayesian approach (Rannala & Mountain 1997). To do this the program creates a reference genotype for each population and assigns each sample to a population using a "leave out" method based on probabilities of genotypic match to each of the reference populations (Paetkau *et al*. 2004). When there is high gene flow between populations this method can assign individuals more correctly than STRUCTURE (Waples & Gaggiotti 2006). 10000 simulated individuals were used for probability computations and set to an exclusion threshold of $p<0.05$ Two runs were performed firstly using putative populations based on sample location and secondly using population clusters as inferred from STRUCTURE results.

2.13 Gene Flow and Demography

To test for the importance of genetic drift versus immigration on the population genetics of *Pseudopanax crassifolius,* isolation by distance (IBD) was analysed. This compares the pairwise genetic distance with the pairwise geographic distances. A stepping stone model of population structure is assumed when genetic distances are calculated (Wright 1943; Shepherd & Perrie 2011). To test for IBD straight line distances between populations were calculated using Google Earth version 6.2.2.7373 version and log transformed using LibreOffice 3.5.6.2. The Mantel test was performed using Arlequin 3.5 with 10000 permutations.

Migrate-n (Beerli & Felsenstein 1999) was used to test for gene flow and migration between the sampled populations. To calculate estimates of migration and effective population sizes the program uses a Markov chain Monte Carlo approach with sample genealogies to create maximum likelihood estimates of a matrix for migration with population sizes. Dispersal rates N_EM and effective population size N_E were calculated using a Brownian motion model which is a stepwise mutation model. (The Brownian motion model is much faster than other models and was used despite other tests being done using I_{AM}.) Results are reported as a value M, a mutation scaled migration rate and Theta, the mutation scaled population size. When $M<1$ gene flow is insufficient to prevent loss of homozygosity between populations through loss of alleles (Mills & Allendorf 1996). These two values when multiplied together give an estimated number of migrants per generation.

2.14 Bottleneck analysis

Bottleneck (Cornuet & Luikart 1996)was used to test of evidence of bottle necks in Pseudopanax populations. Tests were performed using an IAM model with 5000 iterations. Significance was adjusted using the bonferoni correction. Bottleneck tests for evidence of bottle necks comparing the allele numbers and heterozygosities at polymorphic loci. Allelic diversity is reduced faster than heterozygosity if a population has been through a bottleneck. A sign test, a standardized differences test, and a Wilcoxon sign-rank test are three tests used by the program to look for evidence of bottleneck (Piry et al 1999)

2.15 Selfing and Seed Diversity

The extent of selfing in plant species has implications as to how well populations conform to the Hardy-Weinberg equilibrium. To get an indication of whether or not selfing was occurring in *Pseudopanax crassifolius*, samples of seeds were collected from two individuals and DNA extracted from a selection of seeds. The extraction included some seed coat tissue, which is from the parent plant, meaning that there is DNA from potentially two genetically different individuals. If the seed was the result of selfing there would be only two possible alleles from the parent plant. If the seed was the result of outcrossing there would be the possibility of three alleles the two maternal (one from the seed's DNA and the homologous allele from the seed coat in the case of heterozygous mothers) and one paternal. In all cases at least one of the seed's alleles should be the same as the maternal plant and if the paternal source has a different allele this should show up.

Chapter 3: Results

3.1 Microsatellite Validation and Variation

Analysis by Microchecker indicated possible amplification errors. The presence of null alleles was detected in all ten loci for at least one population, but was not found across all populations for the most part. However, three of the four loci amplified using the M13 tailed reaction showed evidence for null alleles in most populations and so were excluded from further analysis. This problem has been attributed to problems with the PCR using the M13 protocol. None of the remaining loci showed evidence for null alleles across all populations and no population showed null alleles in all loci. Despite the presence of null alleles, the remaining loci were retained because it has been reported that their impact on the elucidation of genetic differentiation is limited (Carlsson 2008). The presence of null alleles is expected in groups with a large effective population size (Chapuis & Estoup 2007) and *Pseudopanax crassifolius* is at least presently common and widespread. There was no evidence for large allele dropout. One locus, M13j, showed possible evidence for scoring errors due to stuttering but this was confined to four populations only. Populations with possible stuttering were rechecked and stuttering was not found to be present and therefore this is likely to be an artefact of null alleles.

There was varying levels of non-compliance to Hardy-Weinberg Equilibrium (HWE). Four loci differed significantly from HWE. None of these four loci were found to be significantly different from HWE in all populations after Bonferroni correction. This departure from HWE can largely be accounted for by the presence of null alleles as loci found to have null alleles in Microchecker detected were those out of HWE. No significant linkage disequilibrium was detected between any locus pairs.

The expected heterozygosity for all alleles ranged from 0.692 (Pseles5) to 0.915 for (Psecra3b) and the number of alleles per locus ranged from 9 to 25. Average F_{IS} values ranged from -0.06 to 0.556 High levels of F_{IS} indicate some level of nonrandom mating potentially through

selfing. Null alleles may also be driving the high F_{IS} as loci with the most null alleles show the highest F_{IS}.

Lositan showed two loci to be candidates for selection: M13j (positive selection) and Cra3b (balancing selection). Therefore a pairwise exact test was performed by locus. This showed no evidence of unusual patterns or strong indication that a particular locus was driving differentiation.

Table 3.1: Genetic characteristics of the seven microsatellite loci utilised

LOCUS	N_A	Size range	H_O	H_E	F_{IS}
Psecra2	15	152-182	0.695	0.761	0.015
Psecra3a	9	93-113	0.680	0.708	-0.060
Psecra3b	25	106-152	0.787	0.915	0.093
Psecra5	18	168-216	0.538	0.786	0.265
Pseles5	12	212-238	0.370	0.692	0.433
PSearb9	22	117-173	0.525	0.804	0.292
m13j	17	265-339	0.270	0.706	0.556

N_A number of alleles, Range of allele sizes, H_O observed heterozygosity, H_E expected Heterozygosity and F_{IS}

3.2 Microsatellite Diversity

The allelic richness of the individual populations ranged from 2.92 to 4.73 (Table 2). The most diverse populations corresponded to those populations encompassed by STRUCTURE cluster 1 North (see below). Only one population was monomorphic for any loci. This was *Pseudopanax chathamicus*, which was monomorphic for the Pseles5 locus. The private allele indices showed a more varied range from 0.01 to 0.65 with the highest values for the Coromandel, Auckland and Waimate populations. The lowest values corresponded to populations in the south and West Coast of the South Island.

Table 3.2: Population diversity measures.

Population	ID#	N	AR	PAR	PAR cra only	Private alleles	H_O	H_E	F_{IS}
Aupori	1	2	*	*	*	2	0.429	0.429	0.309
Northland	2	9	3.87	0.27	0.27	1	0.571	0.670	0.219
Auckland	3	23	4.38	0.56	0.56	11	0.483	0.742	0.377
Coromandel	4	10	4.73	0.55	0.59	2	0.428	0.769	0.430
Kaimai	5	1	*	*	*	*	*	*	*
Kawhia	6	5	*	*	*	*	0.421	0.673	0.309
East Cape	7	4	*	*	*	*	0.381	0.476	0.227
Anaura	8	1	*	*	*	*	*	*	*
Maungaharuru	9	11	4.02	0.12	0.14	2	0.439	0.714	0.422
Heretaunga	10	10	3.63	0.15	0.16	1	0.453	0.694	0.326
Tongariro	11	5	*	*	*	*	0.536	0.640	0.290
Taranaki	12	10	4.2	0.18	0.23		0.482	0.717	0.248
Wellington	13	13	3.96	0.2	0.25	1	0.438	0.693	0.355
NW Nelson	14	17	4.06	0.19	0.25	2	0.424	0.683	0.427
Murchison	15	4	*	*	*	*	0.560	0.635	0.049
Kaikoura	16	13	4.21	0.16	0.27	1	0.480	0.721	0.341
Oxford	17	1	*	*	*	*	*	*	*
Banks Peninsula	18	15	3.87	0.37	0.42	3	0.378	0.712	0.452
Waimate	19	11	4.22	0.41	0.44	3	0.554	0.736	0.275
Tokomairiro	20	15	3.52	0.15	0.18	1	0.470	0.633	0.267
Catlins	21	15	3.19	0	0.01	0	0.364	0.581	0.394
Winton	22	1	*	*	*	*	*	*	*
Te Anau	23	4	*	*	*	*	0.286	0.442	0.392
Hollyford	24	18	2.92	0.04	0.05	0	0.418	0.524	0.268
Harihari	25	16	3.05	0.07	0.07	0	0.427	0.558	0.214
Foulwind	26	14	3.33	0.21	0.22	3	0.356	0.547	0.247
Chatham	27	9	3.08	0.15	*	*	0.419	0.502	0.189
P. ferox	28	10	3.91	0.97	*	*	0.424	0.623	0.373

N, number of samples in population; AR, allelic richness; PAR, Private allelic richness (calculated for populations with 9 or more samples); HO, Observed Heterozygosity, HE, Expected Heterozygosity, F_{IS} fixation index. See Fig 4.1 for population locations

3.3 Population structure

Allele size permutation tests in Spagedi showed no significant result for the comparison including all populations (p=0.054, 10 000 iterations). This indicates that F_{ST} is a more appropriate than R_{ST} to determine differentiation. The global F_{ST} calculated from all *Pseudopanax crassifolius* populations was 0.082, (p <0.0002) indicating moderate levels of differentiation Wright (1978) defined F_{ST} values that ranged from 0 – 0.05 as being low, 0.05 – 0.15 as being moderate, 0.15 – 0.25 as being high and over 0.25 as being very highly

differentiated. No *P. crassifolius* population was significantly differentiated from all other populations after pairwise comparisons were corrected for multiple tests (Table 3.3). The Catlins population is the most often significantly different population (10 significant comparisons) followed by Northwest Nelson (9) and then Auckland and Banks Peninsula with 7. The range for pairwise F_{ST} is from -0.0164 – 0.2317 and of these the range of significant values is from 0.0309 – 0.2124. None of the pairwise comparisons within the North Island are significant while there are significant F_{ST} between the North Island and South Island and within the South Island. With *P. chathamicus* included in the analysis global F_{ST} rose to 0.093 ($p<0.0002$). Pairwise comparisons using the harmonic mean of D_{EST} as implemented in SMOGD (see Table 3.4) indicated a similar pattern of genetic differences to that shown in pairwise F_{ST} values.

AMOVA results of *P. crassifolius* alone indicate that 87.3% of variation occurs within populations and the remaining 12.7% between the populations. A further breakdown into STRUCTURE cluster groups gives 4.3 % of the variation occurring from among populations within groups (North and South clusters) and 8.4% from among these groups. When *Pseudopanax chathamicus* was included in the analysis the results remain consistent with 86.4% of variation coming from within groups and 13.6% from between the population. Global F_{ST} was calculated for each STRUCTURE cluster separately. The northern cluster has an overall FST of 0.038 ($p<0002$) and the southern cluster 0.048. ($p<0002$)

Table 3.3; Pairwise FST values between populations with five or more sampled individuals

	Northland	Auckland	Coromandel	Kawhia	Maungaharuru	Heretaunga	Tongariro	Taranaki	Wellington	N.W. Nelson	Kaikoura	Banks Peninsula	Waimate	Tokomairiro	Catlins	Hollyford	Harihari	Foulwind
Northland	0																	
Auckland	0.0228	0																
Coromandel	0.0051	0.0128	0															
Kawhia	0.0291	0.0278	-0.0164	0														
Maungaharuru	0.0124	0.0162	0.0272	0.0425	0													
Heretaunga	0.0645	0.0175	0.0189	0.0409	0.0351	0												
Tongariro	0.0179	0.0305	0.0058	0.0364	-0.0043	0.076	0											
Taranaki	0.0272	0.0112	0.0195	0.0072	-0.0078	0.062	-0.0011	0										
Wellington	0.0931	0.0499	0.0574	0.0233	0.052	0.1021	0.0906	0.0271	0									
N.W. Nelson	**0.0518**	**0.0726**	0.026	0.0594	**0.0467**	0.0951	0.0264	0.062	**0.1469**	0								
Kaikoura	0.0475	0.014	0.0314	0.0349	0.0174	0.0464	0.0444	0.0086	0.0238	0.1046	0							
Banks Peninsula	0.0354	**0.0309**	0.0391	0.0784	0.0069	0.0141	0.0184	0.0441	0.1089	**0.0659**	0.0324	0						
Waimate	0.0285	0.0212	0.0275	0.0341	-0.0005	0.0493	0.0229	0.0069	0.0681	0.0371	0.0303	0.0128	0					
Tokomairiro	0.0645	**0.0884**	0.0697	0.1292	0.058	0.1082	0.0506	**0.0862**	**0.1594**	**0.0444**	**0.128**	**0.0893**	0.0699	0				
Catlins	**0.0949**	**0.1237**	0.0783	0.1445	**0.0742**	**0.1565**	**0.0855**	**0.1133**	0.1702	**0.0323**	**0.1608**	**0.1324**	0.0836	0.0065	0			
Hollyford	**0.0974**	**0.1502**	0.1113	0.1746	**0.113**	0.2197	0.0759	**0.1393**	**0.2124**	**0.0557**	**0.1905**	**0.1591**	0.1163	0.0369	0.0193	0		
Harihari	0.1516	**0.1634**	0.1241	0.2056	**0.1241**	0.2317	0.1681	0.1478	0.2089	**0.0757**	0.2108	**0.1818**	0.1065	0.1008	**0.047**	0.075	0	
Foulwind	0.1076	**0.1421**	0.1094	0.193	**0.0783**	0.1772	0.1214	**0.141**	0.2241	0.0384	0.1806	**0.1353**	0.0958	0.0606	0.0228	0.0453	0.0555	0
P. chathamicus	0.2287	0.1749	0.1532	0.2221	0.1619	0.2089	0.1663	0.1593	0.1947	0.1666	0.1616	0.1936	0.1937	0.1878	0.1982	0.2276	0.2626	0.2047

Values underlined and in bold indicate significant values after sequential Bonferroni correction.

Table 3.4: Genetic distance estimates using harmonic mean of DEST by SMOGD

--	Northland	Auckland	Coromandel	Kawhia	Maungaharuru	Heretaunga	Tongariro	Taranaki	Wellington	N.W. Nelson	Kaikoura	Banks Peninsula	Waimate	Tokomairiro	Catlins	Hollyford	Harihari	Foulwind
Northland	0																	
Auckland	0.019	0																
Coromandel	0.028	0.026	0															
Kawhia	0.020	0.035	0.001	0														
Maungaharuru	0.001	0.045	0.074	0.049	0													
Heretaunga	0.033	0.008	0.017	0.099	0.021	0												
Tongariro	0.003	0.029	0.004	0.061	0.001	0.024	0											
Taranaki	0.057	0.043	0.021	0.025	0.000	0.052	0.000	0										
Wellington	0.159	0.086	0.089	0.055	0.054	0.131	0.147	0.015	0									
N.W. Nelson	0.114	0.164	0.062	0.099	0.118	0.169	0.045	0.109	0.309	0								
Kaikoura	0.038	0.053	0.067	0.039	0.001	0.086	0.062	0.002	0.012	0.160	0							
Banks Peninsula	0.060	0.104	0.134	0.114	0.043	0.003	0.034	0.100	0.176	0.152	0.070	0						
Waimate	0.018	0.085	0.092	0.031	0.013	0.089	0.005	0.021	0.097	0.092	0.007	0.021	0					
Tokomairiro	0.075	0.160	0.135	0.200	0.078	0.120	0.012	0.166	0.243	0.078	0.179	0.131	0.072	0				
Catlins	0.142	0.266	0.156	0.196	0.123	0.300	0.111	0.228	0.308	0.052	0.260	0.261	0.166	0.009	0			
Hollyford	0.141	0.303	0.159	0.226	0.171	0.366	0.102	0.212	0.341	0.049	0.286	0.280	0.222	0.042	0.015	0		
Harihari	0.186	0.268	0.234	0.237	0.162	0.273	0.086	0.201	0.271	0.077	0.267	0.266	0.139	0.050	0.029	0.060	0	
Foulwind	0.191	0.260	0.156	0.209	0.144	0.284	0.177	0.223	0.300	0.058	0.228	0.252	0.137	0.046	0.009	0.039	0.025	0

The results from STRUCTURE were evaluated using the Evanno method as implemented by Structure Harvester (Evanno *et al*. 2005; Earl & vonHoldt 2012). This method indicated that K=2 was the most appropriate for the *P. crassifolius* data set. Delta K dropped sharply from K=2 to K=3 a slight rise at K=4 and then further drop off to close to zero at K=6 and for all higher K (Fig.3.1).

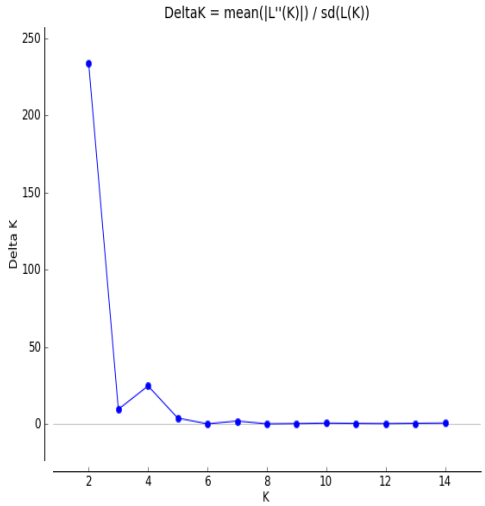


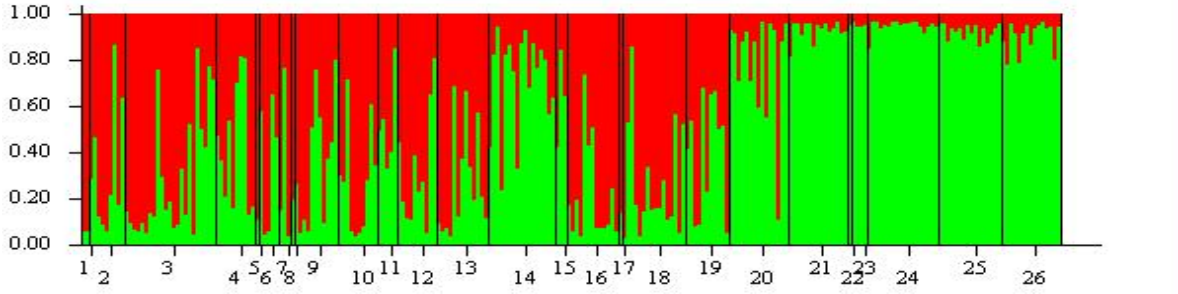
Figure 3.1 Delta K as estimated by Structure Harvester

At K=2 the inferred clusters appeared to be geographically aligned. This pattern persists at higher K but becomes less well defined. The first cluster included all North Island populations and populations on the east coast of the South Island extending to and including the Waimate population. These populations proportionally assign to this cluster with average Q values ranging from 0.551 – 0.814. These values do not generally show strong support for cluster allocation. The second cluster includes the west coast of the South Island and the southern populations of the South Island. These populations assign to this cluster with average Q values ranging from 0.776 – 0.968, which is reasonably robust. There is some ambiguity in allocation of the Northwest Nelson group into the two clusters. It does align more readily with the southern group but not to the same extent as other South Island populations, with average Q 0.549. This may be indicating a boundary region between the two clusters. This group become more differentiated from the southern group at K=4 and 5.

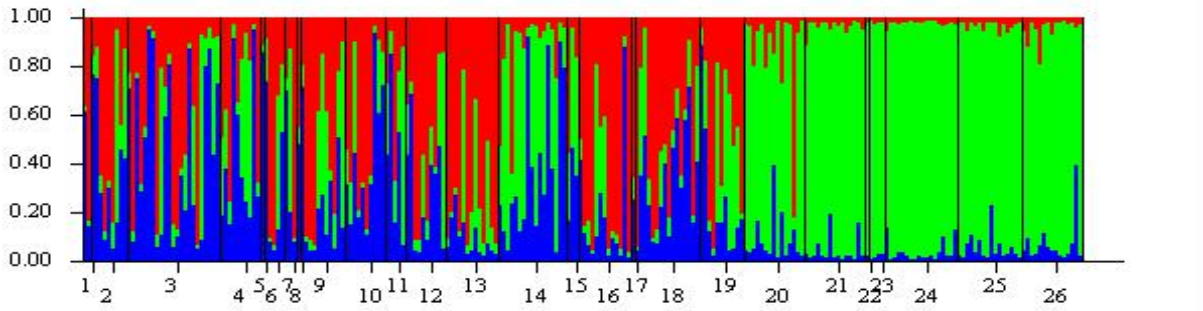
Table 3.5 Average Q for each population into K= two clusters from STRUCTURE

Population	Cluster 1 (north)	Cluster 2 (south)
Aupouri	0.929	0.071
Northland	0.687	0.313
Auckland	0.749	0.251
Coromandel	0.714	0.286
Kaimai	0.919	0.081
Kawhia	0.696	0.304
East cape	0.710	0.290
Anaura	0.813	0.187
Maungaharuru	0.633	0.367
Heretaunga	0.732	0.268
Tongariro	0.598	0.402
Taranaki	0.737	0.263
Wellington	0.736	0.264
N.W. Nelson	0.451	0.549
Murchison	0.442	0.558
Kaikoura	0.783	0.217
Oxford	0.821	0.179
Banks Peninsula	0.706	0.294
Waimate	0.661	0.339
Tokomairiro	0.277	0.723
Catlins	0.119	0.881
Winton	0.079	0.921
Te Anau	0.074	0.926
Hollyford	0.086	0.914
Harihari	0.116	0.884
Foulwind	0.185	0.815

A



B



C

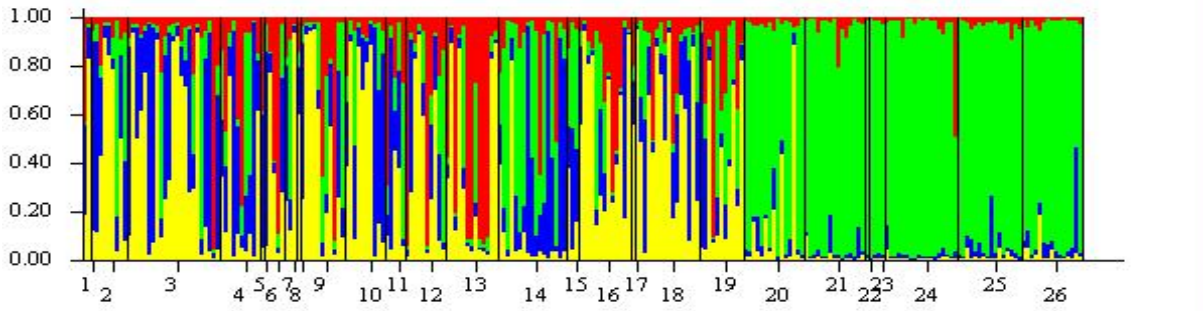
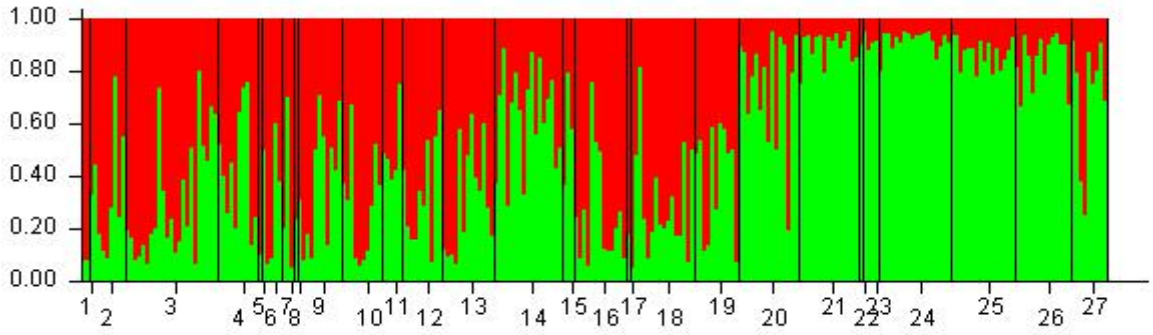


Fig. 3.2. STRUCTURE output with the number of clusters K set to two (A) three (B) and four (C). Each bar indicates an individual sample. The vertical partitioning into the different colours represents the estimated fraction of an individual allocation into the K clusters Population numbers 1 – 13 are from the North Island. 14 – 15 Northwest Nelson and Murchison. 16 – 19 East coast of the South Island from Kaikoura to Waimate. 20 – 22 Tokomairiro, Catlins and Winton. 23 – 24 Te Anau and Hollyford Valley. 24 – 26. West Coast of the South Island.

A



B

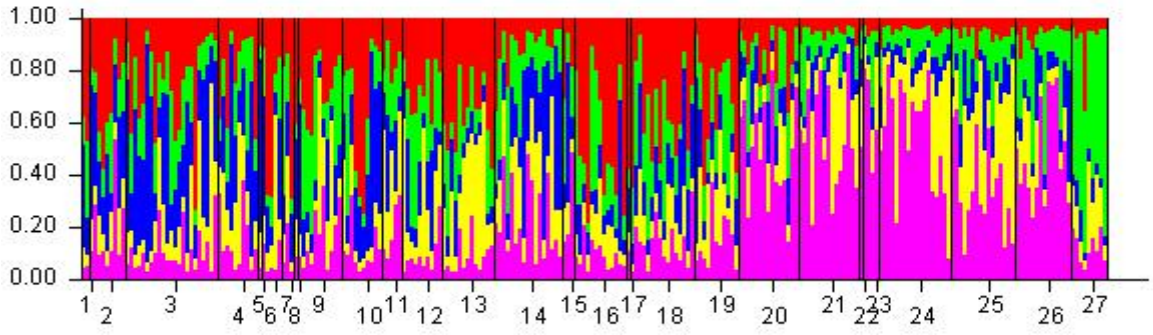


Figure 3.3 STRUCTURE output with *Pseudopanax chathamicus* included in the analysis the number of clusters K set to two (A) and five (B) clusters. Population numbers 1-13 are from the North Island. 14 -15 Northwest Nelson and Murchison. 16-19 East coast of the South Island Kaikoura to Waimate. 20-22 Tokomairiro, Catlins and Winton. 23-24 Te Anau and Hollyford Valley. 24-26 West Coast of the South Island. 27 *P. chathamicus*.

The inclusion of *Pseudopanax ferox* in the analysis alters the results only moderately. K=2 is still evaluated as the best by Structure Harvester with only a small drop to K=3 before getting close to zero at K=5 (Fig. 3.4). The bar graph shows *P. ferox* as a distinct group at K=3 (average Q=0.875) and above (Fig. 3.5). There is little impact on the assignment of other populations indicating that *P. ferox* is reasonably distinct and is not greatly influencing the allocation of *P. crassifolius* and *P. chathamicus* populations. There is one individual in the Banks Peninsula *P. crassifolius* population that assigns more closely to the *P. ferox* samples.

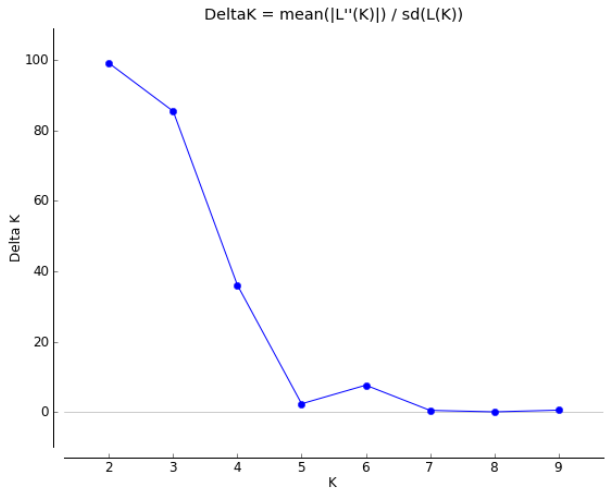
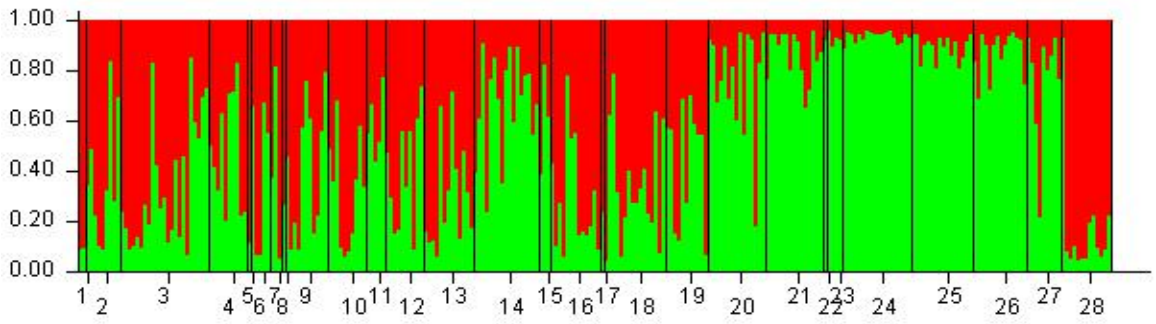


Figure 3.4 Delta K inferred by Structure Harvester with *Pseudopanax chathamicus* and *P. ferox* included in the analysis alongside *P. crassifolius*

The bar graph output file from structure for K=2 shows that the populations 16 – 23, (Dunedin through to Foulwind), are reasonably strongly assigned to the southern (green) cluster, corresponding to average Q 0.776 – 0.968 (see Table 4). Populations 14 (NW Nelson) and 15 (Murchison), also assign more strongly to the southern cluster but with lower average Q, 0.549 and 0.558. The remaining populations, the northern cluster, are variable in their assignment levels. All assign more strongly to the North cluster with average Q values of 0.551 – 0.814. *Pseudopanax chathamicus* assigns more closely to the southern group 0.685. *P. ferox* is the only population assigning closely to the third cluster. 0.875. This might indicate a more genetic homogeneity in the southern cluster and variability in the north.

A



B

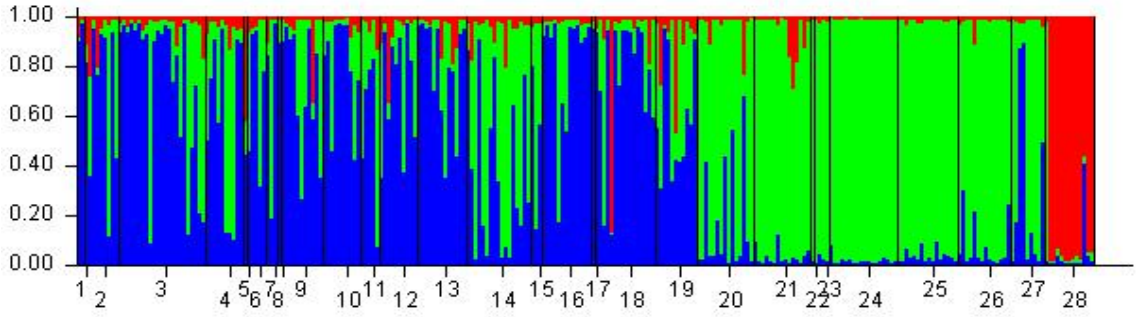


Figure 3.5 STRUCTURE output with *Pseudopanax chathamicus* and *P. ferox* included in the analysis the number of clusters K set to two (top) and three (Bottom). Population numbers 1-13 are from the North Island. 14 -15 Northwest Nelson and Murchison. 16-19 East coast of the South Island Kaikoura to Waimate. 20-22 Tokomairiro, Catlins and Winton. 23-24 Te Anau and Hollyford Valley. 24-26 West Coast of the South Island 27. *P. chathamicus*. 28 *P. ferox*.

Population allocation in the data set including *Pseudopanax crassifolius* and *P. chathamicus* using Bayesian analysis reiterates the weak structuring. Geneclass was only able to correctly assign individuals to their correct population source 20.6% (55 individuals) of the time giving a quality index of 10.69. When the populations were aggregated into 3 groups (North South and Chatham) as inferred from STRUCTURE results there is a much higher success rate of correct assignment 64.6% (191 individuals) with a quality index of 71.5%.

A Spatial AMOVA (SAMOVA) was implemented using a sum of squares microsatellite analysis. At K=2 the populations are clustered along species lines, showing a combination of both geographic and genetic distance. At K=3 *P. chathamicus* remains separate, and *P.*

crassifolius is divided into a northern group including all of the North Island and the northwest and east coast south to Waimate (see appendix II for the Voronoi diagram for K=3).

3.4 Population Structure

A Mantel test showed some indication of Isolation by Distance (IBD) (Fig. 6). This was not a strong effect but it was significant (Regression coefficient: 0.0465, Correlation coefficient: 0.383, Determination of Y by X1: 0.147, p=0.00017)

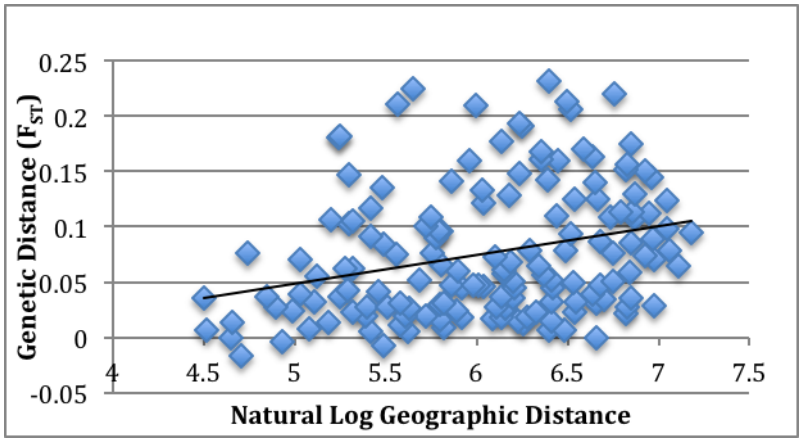


Figure 3.6 Isolation by distance for *Pseudopanax crassifolius*. Genetic distance (F_{ST}) plotted against the natural log of straight line geographic distances.

All population pairs were tested for gene flow and migration using migrate-n. In all cases M>1 indicating high levels of gene flow. M, a mutation scaled migration rate ranged from 1.46 – 23.38. The number of migrants however was low only 18 population pairs (2%) of tests indicating a migration of greater than one migrant per generation. The high gene flow and poor structuring suggest that accurate detection of gene flow and migration may be problematic. There was no geographic pattern to the migration results. Given the lack of strong structure or clear patterns of differentiation these results should be interpreted with caution.

Bottleneck found no population showed a significant indication of a genetic bottleneck after sequential Bonferroni correction.

3.5 Chloroplast Haplotypes.

Pseudopanax crassifolius was tested for the presence of two rps4 haplotypes following the analysis done in Shepherd & Perrie (2011). Each sample would only amplify one rps4 haplotype (A or B). Both A and B haplotypes were found to be present in *P. crassifolius* as with *P. ferox*. Haplotype A was found in all 12 populations tested. The North Island populations exclusively had the A haplotype. Haplotype A was the most widely distributed haplotype in the South Island. Haplotype B was only detected in populations from the south east of the South Island from Banks Peninsula to Catlins. These populations had both haplotypes present.

Figure 3.7 Distribution of rps4 haplotypes found in *Pseudopanax crassifolius* and *P. chathamicus* with number individuals from populations sampled. The North Island contains only A haplotype. The South Island contains both A and B and the distribution of the two haplotypes are shown.

3.6 Parental analysis

Microsatellite analysis of seeds from two parent trees showed the presence of up to 3 alleles per sample per locus. This indicates the occurrence of outcrossing in *P. crassifolius*. The detection

of 3 alleles in a sample indicated 2 alleles from the seed DNA and the presence of parental DNA from the seed coat. Evidence for selfing is difficult to establish because even a single allele found in analysis of seeds could be from two parents sharing the same allele. 10 of 14 seed samples showed evidence for outcrossing in at least one locus by the presence of an allele not found in the parent sample. It is worth noting that absence of a different allele to those of the parent does not preclude outcrossing as the two parents may share the same alleles.

Chapter 4: Discussion

4.1 Population Genetics

The genetic differentiation and the distribution of genetic variation exhibited is enough to conclude that *Pseudopanax crassifolius* does not consist of a single homogenous population. *Pseudopanax crassifolius* shows moderate genetic differentiation and can be divided into two clusters based on the STRUCTURE analysis. Two measures of diversity, allelic richness (AR) and private alleles (PA) are lower in the southern cluster relative to the northern cluster. Private alleles are present in both of the clusters. They are common in populations from the northern cluster and rare in populations in the southern cluster, restricted to only the Foulwind and Tokomairiro populations. Like many measures of diversity the presence of private alleles is strongly influenced by sample size. If only few individuals are sampled from a population a true representation of the alleles present will not be discovered lending to the possibility that false positive results for private alleles will be shown. As no population is markedly genetically depauperate, or can be regarded as a subset of any other single population, there is no strong indication that any population is the product of recent long distance dispersal. The geographic pattern of genetic diversity and differentiation does, however, indicate the possibility that the southern cluster has a more recent origin than the northern cluster.

4.1.1 Diversity

The allelic richness of *Pseudopanax crassifolius* populations ranges from 2.92 to 4.73. A range of genetic diversity has been found in New Zealand from high in *P. ferox* (Shepherd & Perrie 2011) and *Metrosideros* (Young *et al*. 2001), to low in rimu (*Dacrydium cupressinum*) (Hawkins & Sweet 1989) and *Nothofagus* (Haase 1992). The levels of allelic richness in *P. crassifolius* and of expected heterozygosity (H_E 0.429 – 0.769) compares closely to that of *P. ferox* indicating high diverstity (Shepherd & Perrie 2011). The life histories of species have associations with their genetic diversity. Outcrossing, being woody and having animal dispersal vectors are all

associated with high genetic diversity and low differentiation (Hamrick *et al*. 1992; Hamrick & Godt 1996; Broadhurst *et al*. 2008). *Pseudopanax crassifolius* displays all of these characteristics to at least some extent, although the extent of outcrossing and selfing rates has not yet been determined. This combination of factors goes at least someway to explaining the high genetic diversity and levels of differentiation found in *P. crassifolius*.

In *Pseudopanax crassifolius,* population genetic richness measures are higher in the north and reduced in the south. All populations tested from the northern STRUCTURE cluster show higher allelic richness than those from the southern STRUCTURE cluster. High genetic diversity is associated with areas of long-term inhabitation, which can be used to infer refugial populations when considering the effects of the Pleistocene glaciations. Numerous studies have inferred this association (for example(Hewitt 1999; Widmer & Lexer 2001; Gardner *et al*. 2004; Shepherd *et al*. 2007; Marshall *et al*. 2009; Barnaud & Houliston 2010; Shepherd & Perrie 2011). The highest value 4.73 comes from the Coromandel and the lowest value from the Hollyford Valley 2.92. The gradient in genetic diversity seen in some cases of low latitude richness and high latitude paucity lead to the term "Southern Richness Northern Purity" (Hewitt 1999) when considering Northern Hemisphere species and vice versa in the Southern Hemisphere.

This gradient as seen in *P. crassifolius* suggests a longer history in northern areas compared to the south during the LGM. This low latitude richness and high latitude purity pattern is not consistent across all species. Species like *P. ferox* and *Agyrtodes labralis* (a new Zealand fungus beetle) show evidence of different patterns indicating alternative patterns of survival and directional expansion (Marske *et al*. 2009; Shepherd & Perrie 2011). In *P. ferox*, the highest levels of richness are found in the south and south east of the South Island with the Taieri/Waipouri population having the highest at 4.17. The Auckland population has the lowest with 2.16 (Shepherd & Perrie 2011). This reversal of patterns between *P. crassifolius* and *P. ferox* is also seen in the STRUCTURE results (see Fig. 3.2) where the northern populations of *P. crassifolius* are unable to be separated from the north eastern South Island populations while in *P. ferox* the southern South Island were indistinguishable from the eastern South Island. This might indicate contrasting patterns of survival in the north and rarity in the south for *P. crassifolius* with a more abundant southern survival and reduced northern populations in

P. ferox. *Pseudopanax ferox* has been suggested to have had a widespread distribution but richness and connectivity in the north have been eroded through subsequent reductions in population sizes and distribution (Shepherd & Perrie 2011).

The levels of heterozygosity found in *P. crassifolius* are comparable to those found in *P. ferox* but indicate levels of diversity higher than that found in other tree species from New Zealand (Hawkins & Sweet 1989; Haase 1992; Drummond *et al*. 2000). There are many factors that can influence the levels of heterozygosity. Population size, breeding systems, null alleles, genetic drift, and small sample size can all be of some influence. Small sample size can show the Wahlund effect, where by sampled individuals may not accurately represent the whole population. Genetic drift, which has been proposed for the lack of diversity found in rimu (Hawkins & Sweet 1989), can be an important influence on small populations. The distribution of *P. crassifolius* (see Fig. 1.2) would suggest that it has a large effective population size, at least until recent habitat fragmentation. Given the large effective populations size of *P. crassifolius* genetic drift would not be expected to be a significant factor and this is supported by Bottleneck results finding no significant signal of recent bottlenecks.

Pseudopanax crassifolius has reasonably low levels of observed heterozygosity in relation to the expected heterozygosity (H_O 0.286 – 0.571). The microsatellite data did show the presence of null alleles, which would reduce the apparent diversity. It has been recognised that in groups with high effective population size, high levels of null alleles are to be expected (Chapuis & Estoup 2007) Given the distribution of *P. crassifolius,* it is likely that it has a large effective population size. This could be the reason for the high levels of null alleles seen in some loci. It is also possible that, despite the presence of out crossing, a species may self fertilise with this recorded in pohutukawa (Schmidt-Adam *et al*. 2000). *Pseudopanax crassifolius* has hermaphrodite flowers and other members of its genus like *P. lessonii* are presumed to be able to self-pollinate (Perrie & Shepherd 2009).

4.1.2 Differentiation

Pseudopanax crassifolius has an overall F_{ST} of 0.083 very close to the average value for trees (0.084) found by (Hamrick *et al*. 1992). This places it in the mid to low range of F_{ST} values for tree species as reported by (Muller *et al*. 2009) and indicates moderate genetic differentiation amongst populations. Wright (1978) suggested that values from 0.05 to 0.15 indicate moderate differentiation although this rule of thumb was envisaged for Bi-allelic loci and may not be entirely appropriate for highly variable markers such as microsatellites. This level of differentiation places *P. crassifolius* in the mid range of New Zealand trees. Some such as rimu and *Nothofagus* and an alpine conifer have lower differentiation (Hawkins & Sweet 1989; Billington 1991; Haase 1992) while pohutukawa and *P. ferox* have higher, with *P. ferox* having one of the highest recorded for trees (Young *et al*. 2001; Shepherd & Perrie 2011). The contrasting geographic distribution patterns of *P. crassifolius* and *P. ferox* account for the differences in F_{ST} values. *Pseudopanax crassifolius* has an almost continuous distribution and therefore low population differentiation is expected. The current disjunct pattern of *P. ferox* accounts for its high F_{ST}, being at levels comparable to species inhabiting archipelagos (Muller *et al*. 2009; Shepherd & Perrie 2011).

There is large variation in the range of F_{ST} values between *Pseudopanax crassifolius* population pairs, of which 25% were significant. There is some geographic correlation with the pairwise F_{ST} values. Pairwise F_{ST} values between populations from different STRUCTURE clusters (with K=2) were more likely to be significant than tests within clusters. 7% of the within cluster comparisons were significant, whereas 43% of between-cluster comparisons were significant. A mantel test shows a significant but small effect of isolation-by-distance. This contrasts with the lack of IBD found in *P. ferox*. The contrasting distribution of the two species is again likely to be the reason for this. With a continuous distribution a gradient in differentiation can occur, whereas the isolation of *P. ferox* populations has enabled enough separate evolution to accumulate making differentiation non-directional. In New Zealand isolation by straight line distance may not be entirely appropriate. New Zealand is a long narrow country with alpine regions throughout its centre, with the Southern Alps especially prominent in the South Island. Populations either side of the central mountain chain whilst being geographically isolated are relatively close in a straight-line distance thereby reducing the power of an IBD

test.

4.1.3 Genetic Structure

Genetic structuring is evident amongst the populations of *Pseudopanax crassifolius*. Analysis of the data with STRUCTURE shows that the populations can be divided into two clusters (see Fig. 3.2), and that these clusters have a geographic association. As plants are largely stationary, any genetic structure almost certainly follows a geographic pattern (Loveless & Hamrick 1984). The Evanno method of analysing the STRUCTURE results indicates that the best fit for the number of genetic clusters is two. One cluster contains individuals from the north, including all the North Island and the northeast and central east of the South Island (Fig.4.1), while the second cluster encompasses the south and west coast of the South Island. The inferred boundary between the clusters is along the centre of the South Island, and the boundary between the Otago and Canterbury regions. The Southern Alps dominate the centre of the South Island and are well recognised as a barrier to gene flow for lowland species and are likely to have been so since their uplifting in the Pliocene (Trewick & Wallis 2001; Heenan & Mitchell 2003; Haase *et al*. 2007; Chapple *et al*. 2009; Marske *et al*. 2009, 2011). The Waitaki river marks the boundary region of the central and southern biogeographic regions of the South Island as delineated in other phylogeographic studies (Emerson *et al*. 1997; Trewick *et al*. 2000; Wallis *et al*. 2001; Buckley *et al*. 2001; Greaves *et al*. 2007; O'Neill *et al*. 2008). The Waitaki River itself is unlikely to represent a strong barrier to an animal dispersed species such as *P. crassifolius*. A more likely boundary is the abrupt landscape change and associated climatic conditions at transition from the Canterbury Plains to the Otago ranges. It is in this area that there are ranges that extend almost to the coast through Hawkdun, Ida and Horse Ranges, and the Kakanui Mountains.

Figure 4.1 Inferred cluster boundaries for *Pseudopanax crassifolius* from STRUCTURE results. The two major clusters north and south and a possible third cluster including populations from NW Nelson and Murchison. See table 3.2 for information about populations.

In present conditions, there is scope for gene flow across the boundary between the two clusters and this is reflected in the low pairwise F_{ST} value between Waimate and the Tokomairiro populations. However during the LGM the harsher conditions in this boundary zone may have been enough to inhibit gene flow or represent the southern limit to survival. As apparent in Fig. 1.2, this area marks a change to lower temperature averages. If the temperature

depression was great enough during the LGM, this area could mark a transition from a habitable climate to areas inhospitable further south and thus mark the southern limit of *Pseudopanax crassifolius* during the LGM. Another possibility is that this area represents a break in connectivity between two surviving populations. *Pseudopanax ferox* is thought to have survived south of this area (Shepherd & Perrie 2011). This must raise the possibility that *P. crassifolius* was also able survive in a refugial population in areas south of this boundary. Gene flow connecting these geographically isolated populations would have been possible if there were vectors for this. The animal vectors of *P. crassifolius* raise this possibility to at least some extent. If the intermediate area between populations at Waimate and Tokomairiro were deforested because of the conditions, then the vectors for migration and gene flow are also likely to be absent further reinforcing genetic isolation and maintaining this area as a genetic boundary zone.

Individuals from northern populations show mixed affinity to the two genetic clusters. Individuals from the southern cluster show strong affinity to their cluster as seen in the average Q values for populations (see Table 3.5). When K values increase, the southern group of populations remain largely assigned to the same single cluster, and populations in the northern group retain mixed affinity across several clusters. Affinity of the southern group to one cluster at different K indicates robustness in the assignment of the southern cluster and support for the lack of differentiation within the northern cluster. This reflects the pattern seen in the pairwise F_{ST} values between and amongst populations in the two groups. The mixed affinity of the northern cluster is possibly an artefact of the higher diversity and lack of differentiation and shows that this group has or at least had a large effective population size and this has been present for some time. The homogeneity of the southern group can be interpreted as resulting from its more recent establishment showing in its reduced diversity. At the boundaries of the clusters the assignment of the populations becomes less clear. The average Q for the populations is reduced at the edge of the southern cluster, Tokomairiro and Northwest Nelson (Table 3.5). This could indicate areas of mixing between the two clusters or mark the beginnings of reducing diversity leading to weaker differentiation. Mixing of the two clusters is difficult to infer due to the lack of clear assignment in the northern cluster.

Little differentiation and structuring have formed in the northern STRUCTURE cluster

despite an inferred long history of *Pseudopanax crassifolius* in the area. When analysed separately the northern cluster has a low F_{ST} 0.038 likewise the southern cluster with 0.048. These low values are consistent with the lack of genetic structure found within populations of these areas (Fig. 3.2.). These low values indicate that the overall F_{ST} in *P. crassifolius* across New Zealand is driven by the differentiation between the two clusters. Little structure and differentiation would be expected if there was widespread survival enabling continuous gene flow, which acts to reduce differentiation (Loveless & Hamrick 1984; Slatkin 1987). It is possible that in the future differentiation and structuring will develop as the distribution of forests around New Zealand change. In pre-human times there were periodic deforestations through fire ignited by natural causes such as lightning and volcanism, but these were intermittent enough that forest regenerated between events. Since the arrival of humans, around 1200 years ago, deforestation has been more permanent and has resulted in very fragmented forests (Ogden *et al*. 1998; Ewers *et al*. 2006; McWethy *et al*. 2010).

If *P. crassifolius* was reduced to isolated refugia during the LGM genetic structure could have developed. Later as populations expanded to the widespread distribution seen presently then it is likely that gene flow between newly connected populations occurred eroding any genetic structure. There is no clear and consistent signal that can differentiate between widespread survival during the LGM or one of a rapid expansion immediately post as has been suggested for some species (McGlone & Bathgate 1983; Shepherd & Perrie 2011). Lack of structuring could also be an artefact of the genetic diversity seen. If there are high levels of variability in markers given enough time homoplasy can develop which would act to mask structure and differentiation.

4.2 Distribution

The current distribution of *Pseudopanax crassifolius* is widespread and almost continuous (Perrie & Shepherd 2010; Dawson & Lucas 2012; Fig.1.2) with oceanic and altitudinal boundaries being the main constraints. There is enough evidence in the genetic structuring data to suggest that there has been a change in the distribution of *P. crassifolius* in the late Quaternary. It appears that *P. crassifolius* was only moderately restricted during the LGM remaining largely widespread in the north. This is consistent with other studies that show a

similar inferred widespread survival in forest obligate species (Marshall *et al*. 2009; Marske *et al*. 2009). The genetic patterns seen in the southern cluster indicate a reduction in distribution. The reduced diversity in the southern group could be caused by population restriction or, if reduction was to an extent that extirpation occurred, recent colonisation of the area. The apparent homogeneity of the southern STRUCTURE cluster can be interpreted as there not having been enough time for diversity to accumulate since a recent colonisation.

The lack of signal representing a refugial population in the south suggests absence or severe restriction in the southern South Island and west coast of the South Island in response to the conditions present during the LGM. Some of the restriction in the south must be due to the extensive glaciation in the area covering suitable habitat and some likely to be due to the harsh conditions. The simplest interpretation is that *P. crassifolius* survived during the LGM in populations widespread throughout the North Island and the east of the South Island including Canterbury and has subsequently recolonised the far south and west coast of the South Island following the LGM.

A northern survival during the LGM has previously been inferred from the genetic studies of other New Zealand species (Morgan-Richards *et al*. 2001; Neiman & Lively 2004; Gardner *et al*. 2004; Opell 2006; Chapple *et al*. 2008; Hill *et al*. 2009; Marshall *et al*. 2009; Buckley *et al*. 2009, 2010). This pattern of low latitudinal refugia is reflected through out the world (eg. Hewitt 1999, 2000; Beatty & Provan 2010, 2013). The exact extent of the range encompassed by *P. crassifolius* during the Pleistocene glaciations is not clear. There are three possibilities. *P. crassifolius*, widely distributed with limited restriction, or it occurred in many widespread refugia or it was in few very restricted refugia. Disentangling widely distributed and widespread refugia is difficult. The ill-defined population structure and higher levels of genetic richness of the northern part of New Zealand do suggest that this widespread distribution has been around for sometime with associated gene flow rather than many isolated refugia that have since become connected during subsequent range expansion. A stronger level of structuring would be expected if there had been isolated refugia in the north of New Zealand.

There is no compelling evidence that any particular population was an isolated refugial population during the LGM. Neither the northern cluster nor southern clusters show strong differentiation or markedly higher genetic richness, or have a super or sub sets of alleles compared to nearby populations. The Northland and Coromandel populations have high levels of allelic richness and private alleles, which are indicators of isolation. However, any suggestion that they harboured the only refugia is inconsistent with the lack of a strong gradient of richness measures radiating away from these areas.

The strongest evidence for a possible isolated refugial population in New Zealand comes from the area encompassing the Northwest Nelson and Murchison populations. This area is geographically isolated surrounded by mountain ranges and the ocean. This area is recognised as a forest refugia during the LGM on the basis of the palynological record (Alloway *et al*. 2007), and this is supported by genetic data from Neiman & Lively (2004), Gardner *et al.* (2004) and Marske *et al* (2009). The Northwest Nelson and Murchison populations sit at the edge of the two inferred STRUCTURE clusters and are somewhat problematic in their assignation. There is not enough genetic differentiation to conclusively indicate that they are a separate cluster from the north and south clusters (see Fig. 3.2 population numbers 14, 15). There are lines of evidence suggesting some differentiation, namely the presence of private alleles and the pattern of clustering of these populations in STRUCTURE at higher K values. STRUCTURE analysis indicates that the Northwest Nelson and Murchison populations are closer to the southern group whereas pairwise F_{ST} comparisons show similar differentiation from both groups (see Table 3.3) and SAMOVA groups it with the northern populations. The lack of strong differentiation from the north may reflect the genetic diversity and lack of genetic structure of the northern group, which make differentiation difficult to detect. The high diversity and private alleles found in these populations can be interpreted as support for an isolated remnant population and one that was not severely restricted in population size. The presence of high private allelic richness alone has not previously been interpreted as evidence for isolation for other populations in this study, but has been used to suggest a long-term presence. More corroborating evidence is needed such as high differentiation or strong

STRUCTURE clustering to interpret these populations as definite refugia.

If as suggested *Pseudopanax* crassifolius was absent from the far south and West Coast it has successfully re-established itself. The recolonisation of this area could have come through three potential avenues. Firstly expansion from small refugial populations. While this is possible, the results suggest not, as there is no obvious candidate for a refugia, neither through higher genetic diversity and private alleles nor a pattern of radiating reduction in diversity from any nearby population. The possibility that refugia occurred in the coldest parts of New Zealand cannot be discounted as there is evidence for some species surviving much closer to glacial edges that once thought (Shepherd *et al*. 2007; Marra & Thackray 2010; Parducci *et al*. 2012). This is consistent with more evidence of species surviving in micro refugia. However, that the populations in the south show only moderately reduced diversity and are not greatly differentiated and lack private alleles stands against them being long term refugia. If small populations were to remain isolated for long periods an even stronger reduction in allelic richness would be expected. This could manifest in fixed alleles or low levels of polymorphism through genetic drift and inbreeding. Private alleles would be expected to be more common as separate evolution occurs. Secondly long distance dispersal would be characterised by populations being a subset of diversity from founding populations with a radiating pattern of reducing diversity. Long distance dispersal into an area, and population restrictions are often seen through reduced diversity from bottleneck or founder effects (Widmer & Lexer 2001). The lack of evidence for a bottleneck somewhat refutes these populations having gone through a severe restriction or being the result of a founder event. The third possibility is leading edge recolonisation, which would show with populations being a subset of diversity seen in the northern cluster and a reducing diversity to the most isolated areas. Both patterns are seen with Pseudopanax crassifolius and in the rapid re-colonisation in post-LGM Europe (Hewitt 1999). The contrasting homogeneous and heterogeneous patterns of the two clusters and the higher diversity of the north and the reduced diversity in the south fits with the inferences made by Morgan-Richards *et al.* (2001) and Trewick *et al* (2011) where they postulated that the retreat of glaciers enabled leading edge recolonisation of forest in previously uninhabited areas as conditions became more suitable and habitat available (Fig. 4.3).

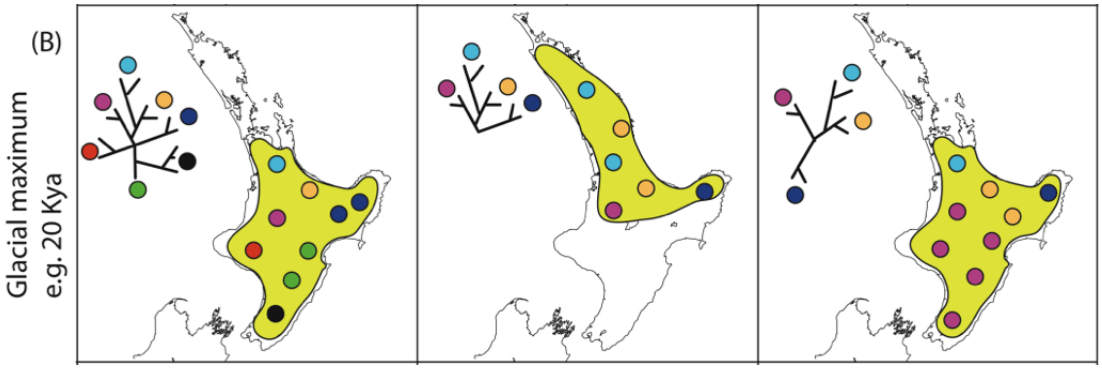


Figure 4.3. An example of the expected genetic pattern associated with distribution reduction and later leading edge re-expansion in the North Island. Taken from Trewick et al. (2011).

The alleles in the locus Psecra3b suggest that the leading edge established in the east coast of the South Island near the boundary of the inferred STRUCTURE clusters. Almost all alleles found in the southern cluster are found in northern populations and there are very few private alleles in the southern cluster. In Psecra3b there is a high proportion of alleles around 116 bp long found in the southern cluster, while in the northern populations these are uncommon. The Waimate population, at the southern edge of the northern cluster, has some of these alleles, while a high proportion of these alleles occur in the Tokomairiro population, at the northern edge of the southern cluster. In the other alleles there is no strong signal. A conclusion based on one allele is not a strong one but this fits with the general pattern. The topography of the South Island does mean that the recolonisation may have occurred on two fronts either side of the Southern Alps. In this case the Northwest Nelson population could be a candidate for a progenitor. This does not account for the apparent homogeneity of structure and limited differentiation in the southern STRUCTURE cluster

4.3 Ecological Tolerances

Pseudopanax crassifolius is a hardy plant of lowland to montane habitats. Growing throughout New Zealand it is exposed to a wide range of conditions, from frosts and snow to extended dry and warm conditions. The temperature ranges experienced throughout contemporary New Zealand are greater than the average temperature depression of 4 – 6°C during the LGM. This level of temperature depression alone would mean that much of New Zealand, especially in the North Island and the north of the South Island, would still have remained within that tolerated by many species including *P. crassifolius* today (Drost *et al*. 2007). There has been suggestions that rather than the average temperature depression, it is the frequency of extreme conditions that drove some species distribution limitations (McGlone 1988). Given the lack of association with MAT and wide tolerances to ecological conditions in many species in New Zealand, a combination of climatic elements might be what delimits species boundaries. That the genetic cluster boundary and a geographic boundary are in concordance indicates that the conditions at the Otago ranges that extend to the coast were a limiting factor in the distribution of *P. crassifolius.* This may have been temperature alone as it is an area where there is a change in average temperatures (Fig. 1.2) or a combination of a few climatic variables.

There are similarities between *Pseudopanax crassifolius* and *P. ferox* in their breeding systems and morphology and ecological tolerances. This is reflected in their similar ranges. At present as they exist together on a broad scale they are not competitively exclusive, indicating some differences in their life history traits. This suggests no reason that they would not have been able to cohabitate in the past. Shepherd & Perrie (2011) postulated that *P. ferox* was more widespread in its distribution immediately after the LGM when there was higher substrate fertility but still dry conditions. These are factors in which *P. crassifolius* and *P. ferox* show differing responses. *Pseudopanax ferox* is found in areas with higher fertility and is more drought tolerant than *P. crassifolius*. If conditions were sufficiently more suited to *P. ferox,* then *P. crassifolius* might have been competitively excluded. If a species is excluded to a great enough extent its genetic legacy may be hidden by migrants in subsequent expansions (Waters 2011). This is a possibility for *P. crassifolius* if it was kept rare by exclusion from *P. ferox*. This allows for the possibility of survival despite lack of genetic evidence and may help explain the

reversal of diversity distribution between the two species.

4.4 Chloroplast Haplotype

The distribution of the chloroplast haplotypes in *Pseudopanax crassifolius* was dominated by the A haplotype which was recorded in all populations tested. The B haplotype present in populations from the east coast of the South Island. This shows some correlation to the pattern found in *P. ferox (Shepherd & Perrie 2011)* but not as strongly demarcated across the Cook Straight. In *P. ferox* the North Island populations tested had only the A haplotype and only B was found in the South Island. The populations of *P. crassifolius* with the B haplotype are in areas with a high density of *P. ferox*. The distribution of the haplotypes could be explained by the ancestral species to both *P. crassifolius* and *P. ferox* having both haplotypes and stochastic processes have resulted in the pattern shown today. However the pattern found in the chloroplast data, while on its own is not particularly informative about historic distributions, can be interpreted to support the possibility that *P. crassifolius* was restricted when *P. ferox* was at its maximum extent shortly after the LGM. This might have meant that low densities of *P. crassifolius* were cohabiting with *P. ferox*. *Pseudopanax* is recognised as being a genus that hybridises easily even between non-sister species. This is most commonly seen in the hybridisation between *P. crassifolius* and *P. lessonii* and occurs in areas where they are both present either naturally or have been introduced (Morgan-Richards *et al*. 2009; Perrie & Shepherd 2009). There have been examples of suspected hybridisation between *P. crassifolius* and *P. ferox* (LR Perrie pers comm.). This particular hybrid is not easily recognisable and is thought to be very uncommon with only a handful of specimens ever collected. The presence of both haplotypes in *P. crassifolius* raises the possibility of introgression through hybridisation instigated by minority cytotype disadvantage (Levin 1975) for *P. crassifolius*. Sharing of non nuclear DNA between species that are closely related is fairly common and has been reported in New Zealand trees (Soltis *et al*. 1991; Petit *et al*. 1993; Gardner *et al*. 2004). The mode of such sharing is thought to be through hybridisation and subsequent introgression reinstating or retaining specific delimitation (Tsumura & Suyama 1998).

It is plausible that the presence of the B haplotype in *P. crassifolius* is not through introgression. The presence in both species may be a shared ancestral trait predating speciation or a recent

homoplasy. If either of these possibilities was the case a wider distribution of the two haplotypes through both species might be expected.

Given the lack of A haplotypes found in *Pseudopanax ferox* in the South Island, it seems likely that the haplotype introgression between the two species is asymmetric. Asymmetric introgression can give information about species composition in an area over time. Recently developed models suggest that asymmetric introgression generally goes in the direction from resident local species into recently introduced species (Currat *et al*. 2008) and this is supported by genetic studies (Du *et al*. 2011). This model has some implications for the interpretation the results presented here. It supports the hypothesis that *P. crassifolius* remained common in the north of the South Island, as there is no reported B haplotype here. The areas with the highest rates of shared haplotypes are on the east coast of the South Island from Banks Peninsula to Tokomairiro. This haplotype distribution would suggest that *P. ferox* was common in these areas as was inferred by (Shepherd & Perrie 2011) and that *P. crassifolius* was absent and has re-colonised these areas later. This is contrary to the microsatellite data that suggest that *P. crassifolius* remained extant as far south as Waimate.

There is no obvious sign of introgression nor information which would exclude it in the microsatellite data. Allele frequency distribution charts show no obvious correlation between the *Pseudopanax crassifolius* populations that show high levels of rps4 haplotype B and *P. ferox* above what is found with any other cluster (see Appendix I).

4.5 *Pseudopanax chathamicus*

The Chatham Islands are a small group of islands lying 850 km to the east of Banks Peninsula. They have been separate from mainland New Zealand for most of New Zealand's post Gondwanan history (Heenan *et al*. 2010). Long distance dispersal has been recognised as an important driver for the species diversity found in both the flora and fauna of the Chatham Islands (Trewick 2000; Shepherd *et al*. 2009; Heenan *et al*. 2010). Many of the closest relatives of Chatham Island endemics are widespread common species on the mainland New Zealand (Heenan *et al*. 2010) and this is certainly the case for *Pseudopanax chathamicus* (Perrie &

Shepherd 2009).

Despite *Pseudopanax chathamicus* being recognised as a separate species with some distinct morphological differences to *Pseudopanax crassifolius,* there is no strong genetic evidence to support the delimitation of *P. chathamicus*. The placing of *P. chathamicus* as inferred by AFLP data from (Perrie & Shepherd 2009) indicates that *P. chathamicus* is a monophyletic group but is nested within *P. crassifolius*. *Pseudopanax chathamicus is* not strongly differentiated from *P. crassifolius* when included in the present analyses. Pairwise F_{ST} values are not significant. *Pseudopanax chathamicus* grouped closest to populations from the southern cluster at low K in STRUCTURE. *Pseudopanax chathamicus* does become resolved as a separate group when K is increased (Fig. 3.5) but this delineation is weak. The easiest explanation for the lack of genetic differentiation found between *P. chathamicus* and *P. crassifolius* is time. If the colonisation of the Chatham Islands by *Pseudopanax* was recent genetic differentiation might not have had time to accumulate. Interconnectivity though gene flow between mainland *P. crassifolius* and *P. chathamicus* would also break down differentiation. Given the isolation of the Chatham Islands, on going gene flow seems unlikely to be sufficiently common to maintain homogeneity between the two populations.

The east coast of the South Island is the most likely source for founders, as it is the most geographically proximate area and the genetic evidence is consistent with this, but where exactly from within this area cannot be resolved to any finer detail. No single mainland population tested contains all the alleles present in *Pseudopanax chathamicus*. The three neighbouring populations of Kaikoura, Banks Peninsular and Waimate do however contain most of the alleles. In these cases there is only one or two alleles that differ by one repeat from those found in *P. chathamicus* populations. While it is unwise to rely on a single locus for information it can give some insight. In the *P. chathamicus* population the locus Psecra2 has only three alleles, dominated by allele 156, which accounts for 80% of the recorded alleles, (13 out of the 16). This allele is rare throughout the *P. crassifolius* populations occurring in only seven individuals. Of these the populations from the east coast of the South Island account for five, with the Taranaki and Northwest Nelson populations being the other each having one individual with the allele. Other alleles common in *P. chathamicus* from other loci are not so rare throughout *P. crassifolius* but are found in populations from the east coast of the South

Island. The STRUCTURE results suggest that, at least at low K, that *P. chathamicus* is more closely aligned to the southern cluster, contrasting with the suggestion that potential source populations are those within the northern cluster. At higher K, *P. chathamicus* does become separate indicating that the affinity to the southern group is not a strong one. This does not amount to enough evidence to suggest that founder populations were located elsewhere.

Assuming complete heterozygosity, any seed can carry two alleles per locus. *Pseudopanax crassifolius* has a possible five seeds per berry (Allan 1961) making the maximum alleles carried in one fruit 10 per locus. In seed analysis, a maximum of six alleles was found at one locus from 14 seeds taken from 4 fruit from two trees. This indicates that two berries would be enough to account for all alleles found in *P. chathamicus*. The most polymorphic allele in *P. chathamicus* was Psecra 3b with 10 alleles. All other alleles could be accounted for with the variation found in one berry. This is strong support for an origin by long-distance dispersal, with very few founding events being required to establish this species. Two berries being blown across to the Chatham Islands in westerly storms, either with a bird or a raft, is a reasonable assumption. This would provide enough founders and does not require continued gene flow to maintain the levels of diversity found.

Given the different morphologic characteristics and genetic data it seems the most likely explanation is that *Pseudopanax chathamicus* is an example of incipient speciation, and probably the result of a recent dispersal event. Given that *P. chathamicus* is very closely related to *P. crassifolius*, steps should be taken to preserve its genetic integrity, as it belongs to a group that hybridises easily providing an avenue for loss of genetic identity (Morgan-Richards *et al*. 2009; Perrie & Shepherd 2009). For example *P. crassifolius* and *P. lessonii* should not be planted on the Chatham Islands

Trying to delimit the species boundaries brings into questions of species concepts. Differing species concepts provide different answers depending on what characteristics are being considered. It is when and where speciation is occurring that most confusion arises between different concepts occurs. *Pseudopanax chathamicus* and *P. crassifolius* are examples of this. A morphologic concept would separate the two species whereas genetic concepts would suggest otherwise, or at least be less certain. Using the concept whereby a separately evolving meta

population is regarded as a species (Queiroz 2007) *P. chathamicus* and *P. crassifolius* would be regarded different species. They are separated geographically, and because of this genetically, making them at least meta populations. The genetic data also indicate they are separately evolving. The monophyly of *P. chathamicus* clade is supportive of this. Given enough time it is likely that the monophyly of both species would become apparent.

4.6 Future Directions

The scope of this study has enabled many conclusions to be drawn about the phylogeography of *Pseudopanax crassifolius and P. chathamicus*. As with any endeavour such as this, more samples and more genetic markers would be welcome and would likely enable stronger conclusions to be drawn, at perhaps finer detail. In this study there are some areas where a different approach would be of some value as well. There are many well-established methods that could add important information. Isozyme markers and AFLPs have all been used with some success to infer the phylogeography of New Zealand species. The use of these different markers as well as giving a different point of view might facilitate easier comparisons between other studies. Microsatellites being highly variable nuclear markers could likely be complemented with data from organelle DNA markers or more stable genes. This might be able to clarify some questions of the historic distributions, both in the northern cluster where microsatellites show no discernable pattern and in the east coast of the South Island where there is some disparity between interpretations from rps4 and microsatellite data. However in some cases such as in *Pseudopanax* there is little chloroplast variation found between closely related species (Perrie & Shepherd 2009). Slower evolving markers may provide insight into deeper history, and even into pre Pleistocene patterns. Investigations using molecular clocks could help resolve issues especially around species delimitation. We see this with the rps4 haplotype that gives a similar result to that found with microsatellites but with some contrasts in the interpretation. Methods such as 'next generation' sequencing and SNPS, which through producing large amounts of data, can give more detail. A genomic approach would be of similar potential. All these methods would be of use both within species and to help refine species delimitation.

References

Allan HH (1961) Flora of New Zealand. Volume I. *Flora of New Zealand. Volume I.*

Alloway BV, Lowe DJ, Barrell DJA 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**, 9–35.

Anderson LL, Hu FS, Nelson DM, Petit RJ, Paige KN (2006) Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 12447–12450.

Angers B, Bernatchez L (1998) Combined use of SMM and Non-SMM methods to infer fine structure and evolutionary history of closely related Brook Charr (Salvelinus Fontinalis, Salmonidea) populations from Microsatellites. *Molecular Biology and Evolution*, **15**, 143.

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

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T ,. Neigel JE, Reeb CA and Saunders NC. (1987) Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

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

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

Beatty GE, Provan J (2010) Refugial persistence and postglacial recolonization of North America by the cold-tolerant herbaceous plant Orthilia secunda. *Molecular Ecology*, **19**, 5009–5021.

Beatty GE, Provan J (2013) Post-glacial dispersal, rather than in situ glacial survival, best explains the disjunct distribution of the Lusitanian plant species Daboecia cantabrica (Ericaceae). *Journal of Biogeography*, **40**, 335–344.

Beaumont MA, Nichols RA (1996) Evaluating Loci for Use in the Genetic Analysis of Population Structure. *Proceedings: Biological Sciences*, **263**, 1619–1626.

Beerli P, Felsenstein J (1999) Maximum-Likelihood Estimation of Migration Rates and Effective Population Numbers in Two Populations Using a Coalescent Approach. *Genetics*, **152**, 763–773.

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**, 3–12.

Beheregaray LB (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, **17**, 3754–3774.

Billington HL (1991) Effect of Population Size on Genetic Variation in a Dioecious Conifer. *Conservation Biology*, **5**, 115–119.

Birks HJB, Willis KJ (2008) Alpines, trees, and refugia in Europe. *Plant Ecology & Diversity*, **1**, 147–160.

Booy G, Hendriks RJJ, Smulders MJM, Van Groenendael JM, Vosman B (2000) Genetic Diversity and the Survival of Populations. *Plant Biology*, **2**, 379–395.

Broadhurst LM, Young AG, Murray BG (2008) AFLPs reveal an absence of geographical genetic structure among remnant populations of pohutukawa (Metrosideros excelsa, Myrtaceae). *New Zealand Journal of Botany*, **46**, 13–21.

Buchanan J, Zuccarello GC (2012) Decoupling of Short- and Long-Distance Dispersal Pathways in the Endemic New Zealand Seaweed Carpophyllum Maschalocarpum (phaeophyceae, Fucales). *Journal of Phycology*, **48**, 518–529.

Buckley TR, Marske KA, 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**, 4650–4663.

Buckley TR, 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**, 682–695.

Buckley TR, Simon C, Chambers GK (2001) Phylogeography of the New Zealand Cicada Maoricicada campbelli Based on Mitochondrial DNA Sequences: Ancient Clades Associated with Cenozoic Environmental Change. *Evolution*, **55**, 1395–1407.

Bunce M, Worthy TH, Phillips MJ *et al.* (2009) The evolutionary history of the extinct ratite moa and New Zealand Neogene paleogeography. *Proceedings of the National Academy of Sciences*, **106**, 20646–20651.

Carlsson J (2008) Effects of Microsatellite Null Alleles on Assignment Testing. *Journal of Heredity*, **99**, 616 –623.

Chapin FS, Shaver GR (1985) Individualistic Growth Response of Tundra Plant Species to

Environmental Manipulations in the Field. *Ecology*, **66**, 564–576.

Chapple DG, Daugherty CH, Ritchie PA (2008) Comparative phylogeography reveals pre-decline population structure of New Zealand Cyclodina (Reptilia: Scincidae) species. *Biological Journal of the Linnean Society*, **95**, 388–408.

Chapple DG, Ritchie PA, Daugherty CH (2009) Origin, diversification, and systematics of the New Zealand skink fauna (Reptilia: Scincidae). *Molecular Phylogenetics and Evolution*, **52**, 470–487.

Chapuis M-P, Estoup A (2007) Microsatellite Null Alleles and Estimation of Population Differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Cooper A, Cooper RA (1995) The Oligocene Bottleneck and New Zealand Biota: Genetic Record of a past Environmental Crisis. *Proceedings: Biological Sciences*, **261**, 293–302.

Cornuet JM, Luikart G (1996) Description and Power Analysis of Two Tests for Detecting Recent Population Bottlenecks From Allele Frequency Data. *Genetics*, **144**, 2001–2014.

Cox SC, Findlay RH (1995) The Main Divide Fault Zone and its role in formation of the Southern Alps, New Zealand. *New Zealand Journal of Geology and Geophysics*, **38**, 489–499.

Crawford NG (2010) smogd: software for the measurement of genetic diversity. *Molecular Ecology Resources*, **10**, 556–557.

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The Hidden Side of Invasions: Massive Introgression by Local Genes. *Evolution*, **62**, 1908–1920.

Daugherty CH, Gibbs GW, Hitchmough RA (1993) Mega-island or micro-continent? New Zealand and its fauna. *Trends in Ecology & Evolution*, **8**, 437–442.

DeYoung RW, Honeycutt RL (2005) The Molecular Toolbox: Genetic Techniques in Wildlife Ecology and Management. *The Journal of Wildlife Management*, **69**, 1362–1384.

Dobrowski SZ (2011) A climatic basis for microrefugia: the influence of terrain on climate. *Global Change Biology*, **17**, 1022–1035.

Doyle J, Doyle J (1990) Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.

Drost F, Renwick J, Bhaskaran B, Oliver H, McGregor J (2007) A simulation of New Zealand's climate during the Last Glacial Maximum. *Quaternary Science Reviews*, **26**, 2505–2525.

Drummond RSM, Keeling DJ, Richardson TE, Gardner RC, Wright SD (2000) Genetic analysis and conservation of 31 surviving individuals of a rare New Zealand tree, Metrosideros bartlettii (Myrtaceae). *Molecular Ecology*, **9**, 1149–1157.

Du FK, Peng XL, Liu JQ *et al.* (2011) Direction and extent of organelle DNA introgression between two spruce species in the Qinghai-Tibetan Plateau. *New Phytologist*, **192**, 1024–1033.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571–2581.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, **5**, 435–445.

Ellstrand NC, Elam DR (1993) Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. *Annual Review of Ecology and Systematics*, **24**, 217–242.

Emerson BC, Wallis GP, Patrick BH (1997) Biogeographic area relationships in southern New Zealand: a cladistic analysis of Lepidoptera distributions. *Journal of Biogeography*, **24**,

89–99.

Estoup A, Cornuet JM (1999) Microsatellites: evolution and applications (DB Goldstein, C Schlotterer, Eds,).

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

Ewers RM, Kliskey AD, Walker S *et al.* (2006) Past and future trajectories of forest loss in New Zealand. *Biological Conservation*, **133**, 312–325.

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

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Frankham R (1996) Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology*, **10**, 1500–1508.

Franklin IR (1980) *Conservation biology: an evolutionary-ecological perspective* (ME Soulé, Ed,). Sinauer Associates.

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

Gardner RC, Fls DL, J P *et al.* (2004) A late Quaternary phylogeography for Metrosideros (Myrtaceae) in New Zealand inferred from chloroplast DNA haplotypes. *Biological Journal of the Linnean Society*, **83**, 399–412.

Gibbs G (2006) *Ghosts of Gondwana*. Craig Potton Publishing, Nelson, New Zealand.

Goudet J (1995) FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *Journal of Heredity*, **86**, 485–486.

Greaves SNJ, Chapple DG, Gleeson DM, Daugherty CH, Ritchie PA (2007) Phylogeography of the spotted skink (Oligosoma lineoocellatum) and green skink (O. chloronoton) species complex (Lacertilia: Scincidae) in New Zealand reveals pre-Pleistocene divergence. *Molecular phylogenetics and evolution*, **45**, 729–739.

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

Haase M, Marshall B, Hogg I (2007) Disentangling causes of disjunction on the South Island of New Zealand: the Alpine fault hypothesis of vicariance revisited. *Biological Journal of the Linnean Society*, **91**, 361–374.

Hall GMJ, McGlone MS (2006) Potential forest cover of New Zealand as determined by an ecosystem process model. *New Zealand Journal of Botany*, **44**, 211–232.

Hamrick JL, Godt MJW (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**, 1291–1298.

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. *New Forests*, **6**, 95–124.

Hardy OJ, Vekemans X (2002) spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Harris LN, Taylor EB (2010) Pleistocene glaciations and contemporary genetic diversity in a Beringian fish, the broad whitefish, Coregonus nasus (Pallas): inferences from microsatellite DNA variation. *Journal of Evolutionary Biology*, **23**, 72–86.

Hawkins BJ, Sweet GB (1989) Genetic variation in rimu — an investigation using isozyme analysis. *New Zealand Journal of Botany*, **27**, 83–90.

Hedrick PW (2005) A Standardized Genetic Differentiation Measure. *Evolution*, **59**, 1633–1638.

Heenan PB, Mitchell AD (2003) Phylogeny, biogeography and adaptive radiation of Pachycladon (Brassicaceae) in the mountains of South Island, New Zealand. *Journal of Biogeography*, **30**, 1737–1749.

Heenan P, Mitchell A, de Lange P, Keeling J, Paterson A (2010) Late-Cenozoic origin and diversification of Chatham Islands endemic plant species revealed by analyses of DNA sequence data. *New Zealand Journal of Botany*, **48**, 83–136.

Hewitt G (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

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

Hewitt GM (2004) Genetic Consequences of Climatic Oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **359**, 183–195.

Hicks S (2006) When no pollen does not mean no trees. *Vegetation History and Archaeobotany*, **15**, 253–261.

Hill KBR, Simon C, Marshall DC, Chambers GK (2009) Surviving glacial ages within the Biotic Gap: phylogeography of the New Zealand cicada Maoricicada campbelli. *Journal of Biogeography*, **36**, 675–692.

Holland MM, Parson W (2011) GeneMarker® HID: A Reliable Software Tool for the Analysis of Forensic STR Data. *Journal of Forensic Sciences*, **56**, 29–35.

Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. *BioTechniques*, **46**, 511–517.

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

Huntley B (1990) European vegetation history: Palaeovegetation maps from pollen data - 13 000 yr BP to present. *Journal of Quaternary Science*, **5**, 103–122.

Jamieson IG, Allendorf FW (2012) How does the 50/500 rule apply to MVPs? *Trends in Ecology & Evolution*, **27**, 578–584.

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

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

Kaplan MR, Schaefer JM, Denton GH, Schaefer JM, Denton JH, Barrell DJA, Chinn TJA, Putnam AE, Andersen BG Finkel RC, Schwartz R, & Doughty AM. Glacier retreat in New Zealand during the Younger Dryas stadial. Nature, 467, 194–197.

Kelly JK, Willis JH (2002) A Manipulative Experiment to Estimate Biparental Inbreeding in Monkeyflowers. *International Journal of Plant Sciences*, **163**, 575–579.

Kimura M, Crow JF (1964) The Number of Alleles That Can Be Maintained in a Finite Population. *Genetics*, **49**, 725–738.

Kimura M, Ohta T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences of the United States of America*, **75**, 2868–2872.

Knapp M, Mudaliar R, Havell D, Wagstaff SJ, Lockhart PJ (2007) The Drowning of New Zealand and the Problem of Agathis. *Systematic Biology*, **56**, 862–870.

Lee DE, Lee WG, Mortimer N (2001) Where and why have all the flowers gone? Depletion and turnover in the New Zealand Cenozoic angiosperm flora in relation to palaeogeography and climate. *Australian Journal of Botany*, **49**, 341–356.

Leschen L, Richard A B., Buckley TR, Harman HM, Shulmeister J (2008) Determining the origin and age of the Westland beech (Nothofagus) gap, New Zealand, using fungus beetle genetics. *Molecular Ecology*, **17**, 1256–1276.

Levin DA (1975) Minority Cytotype Exclusion in Local Plant Populations. *Taxon*, **24**, 35–43.

Li Y-C, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, **11**, 2453–2465.

Liggins L, Chapple DG, Daugherty CH, Ritchie PA (2008) A SINE of restricted gene flow across the Alpine Fault: phylogeography of the New Zealand common skink (Oligosoma nigriplantare polychroma) RID B-9073-2008. *Molecular Ecology*, **17**, 3668–3683.

Lloyd BD (2003) The demographic history of the New Zealand short-tailed bat *Mystacina tuberculata* inferred from modified control region sequences. *Molecular Ecology*, **12**, 1895–1911.

Loveless MD, Hamrick JL (1984) Ecological Determinants of Genetic Structure in Plant Populations. *Annual Review of Ecology and Systematics*, **15**, 65–95.

Lutz JA, van Wagtendonk JW, Franklin JF (2010) Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. *Journal of Biogeography*, **37**, 936–950.

Marra MJ, Thackray GD (2010) Glacial forest refugium in Howard Valley, South Island, New Zealand. *Journal of Quaternary Science*, **25**, 309–319.

Marshall DC, Hill KBR, Fontaine KM, Buckley TR, Simon C (2009) Glacial refugia in a maritime temperate climate: Cicada (Kikihia subalpina) mtDNA phylogeography in New Zealand. *Molecular Ecology*, **18**, 1995–2009.

Marske KA, Leschen RAB, Barker GM, Buckley TR (2009) Phylogeography and ecological niche modelling implicate coastal refugia and trans-alpine dispersal of a New Zealand fungus beetle. *Molecular Ecology*, **18**, 5126–5142.

Marske KA, Leschen RAB, Buckley TR (2011) Reconciling phylogeography and ecological niche models for New Zealand beetles: Looking beyond glacial refugia. *Molecular Phylogenetics and Evolution*, **59**, 89–102.

Marske KA, Leschen RAB, Buckley TR (2012) Concerted Versus Independent Evolution and the Search for Multiple Refugia: Comparative Phylogeography of Four Forest Beetles. *Evolution*, **66**, 1862–1877.

McGlone MS (1985) Plant Biogeography And The Late Cenozoic History Of New-Zealand. *New Zealand Journal of Botany*, **23**, 723–749.

McGlone MS (1988) *Vegetation History* (T Webb, B Huntley, Eds,). Dordrecht ; Boston : Kluwer Academic Publishers.

McGlone MS, Bathgate JL (1983) Vegetation and climate history of the Longwood Range, South Island, New Zealand, 12 000 B. P. to the present. *New Zealand Journal of Botany*, **21**, 293–315.

McGlone MS, Duncan RP, Heenan PB (2001) Endemism, Species Selection and the Origin and Distribution of the Vascular Plant Flora of New Zealand. *Journal of Biogeography*, **28**, 199–216.

McLachlan J, Clark J (2004) Reconstructing historical ranges with fossil data at continental scales. *Forest Ecology and Management*, **197**, 139–147.

McLachlan J, Clark J, Manos P (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology*, **86**, 2088–2098.

McWethy DB, Whitlock C, Wilmshurst JM *et al.* (2010) Rapid landscape transformation in South Island, New Zealand, following initial Polynesian settlement. *Proceedings of the National Academy of Sciences*, **107**, 21343–21348.

Meudt HM, Lockhart PJ, Bryant D (2009) Species delimitation and phylogeny of a New Zealand plant species radiation. *BMC Evolutionary Biology*, **9**, 111.

Mills LS, Allendorf FW (1996) The One-Migrant-per-Generation Rule in Conservation and Management. *Conservation Biology*, **10**, 1509–1518.

Morgan-Richards M, Smissen RD, Shepherd LD *et al.* (2009) A review of genetic analyses of hybridisation in New Zealand. *Journal of the Royal Society of New Zealand*, **39**, 15–34.

Morgan-Richards M, Trewick SA, Wallis GP (2001) Chromosome races with Pliocene origins: evidence from mtDNA. *Heredity*, **86**, 303–312.

Muller F, Voccia M, Bâ A, Bouvet JM (2009) Genetic diversity and gene flow in a Caribbean tree Pterocarpus officinalis Jacq.: A study based on chloroplast and nuclear microsatellites. *Genetica, 135 (2) pp. 185-198, 2009*.

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

Neiman M, Lively CM (2004) Pleistocene glaciation is implicated in the phylogeographical structure of Potamopyrgus antipodarum, a New Zealand snail. *Molecular Ecology*, **13**, 3085–3098.

Newnham RM (2012) The vegetation cover of New Zealand during the Last Glacial Maximum. *Quaternary International*, **279–280**, 350–351.

Norton DA, McGlone MS, Wigley TML (1986) Quantitative analyses of modern pollen-climate relationships in New Zealand indigenous forests. *New Zealand Journal of Botany*, **24**, 331–342.

O'Neill SB, Chapple DG, Daugherty CH, Ritchie PA (2008) Phylogeography of two New Zealand lizards: McCann's skink (Oligosoma maccanni) and the brown skink (O. zelandicum). *Molecular Phylogenetics and Evolution*, **48**, 1168–1177.

Ogden J, Basher L, McGlone M (1998) Fire, Forest Regeneration and Links with Early Human Habitation: Evidence from New Zealand. *Annals of Botany*, **81**, 687–696.

Ogden J, Wilson A, Hendy C, Newnham RM, Hogg AG (1992) The Late Quaternary History of Kauri (Agathis australis) in New Zealand and Its Climatic Significance. *Journal of Biogeography*, **19**, 611–622.

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

Opell BD (2006) Molecular Phylogenetic Evidence For The Parallel Evolution Of Rock Ecomorphs In The New Zealand Orb-Weaving Spider Waitkera Waitakerensis (Family Uloboridae). *Journal of Arachnology*, **34**, 467–475.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, **13**, 55–65.

Parducci L, Jørgensen T, Tollefsrud MM *et al.* (2012) Glacial Survival of Boreal Trees in Northern Scandinavia. *Science*, **335**, 1083–1086.

Peakall R, Smouse PE (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Perrie LR, Shepherd LD (2009) Reconstructing the species phylogeny of Pseudopanax (Araliaceae), a genus of hybridising trees. *Molecular Phylogenetics and Evolution*, **52**, 774–783.

Petit RJ, Hampe A (2006) Some Evolutionary Consequences of Being a Tree. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 187–214.

Petit JR, Jouzel J, Raynaud D *et al.* (1999) Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature*, **399**, 429–436.

Petit RJ, Kremer A, Wagner DB (1993) Geographic structure of chloroplast DNA polymorphisms in European oaks. *Theoretical and Applied Genetics*, **87**, 122–128.

Piry S, Alapetite A, Cornuet J-M *et al.* (2004) GENECLASS2: A Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity*, **95**, 536–539.

Piry S, Luikart G, Cornuet J-M (1999) http://www.ensam.inra.fr/URLB/bottleneck/bottleneck.html

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

Putnam AE, Denton GH, Schaefer JM *et al.* (2010) Glacier advance in southern middle-latitudes during the Antarctic Cold Reversal. *Nature Geoscience*, **3**, 700–704.

Queiroz KD (2007) Species Concepts and Species Delimitation. *Systematic Biology*, **56**, 879–886.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences*, **94**, 9197–9201.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rice WR (1989) Analyzing Tables of Statistical Tests. *Evolution*, **43**, 223–225.

Schmidt-Adam G, Young AG, Murray BG (2000) Low outcrossing rates and shift in pollinators in New Zealand pohutukawa (Metrosideros excelsa; Myrtaceae). *American Journal of Botany*, **87**, 1265–1271.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, **18**, 233–234.

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

Shepherd LD, De Lange PJ, Perrie LR (2009) Multiple colonizations of a remote oceanic archipelago by one species: how common is long-distance dispersal? *Journal of Biogeography*, **36**, 1972–1977.

Shepherd LD, Perrie LR (2011) Microsatellite DNA analyses of a highly disjunct New Zealand tree reveal strong differentiation and imply a formerly more continuous distribution. *Molecular Ecology*, **20**, 1389–1400.

Shepherd LD, Perrie LR, Brownsey PJ (2007) Fire and ice: volcanic and glacial impacts on the phylogeography of the New Zealand forest fern Asplenium hookerianum. *Molecular Ecology*, **16**, 4536–4549.

Shepherd LD, Zuccarello GC, Perrie LR (2008) Microsatellite markers for Pseudopanax trees and their cross-species amplification in the Araliaceae. *Conservation Genetics*, **9**, 1687–1689.

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

Soltis DE, Mayer MS, Soltis PS, Edgerton M (1991) Chloroplast-DNA Variation in *Tellima grandiflora* (Saxifragaceae). *American Journal of Botany*, **78**, 1379–1390.

Stephenson NL (1990) Climatic Control of Vegetation Distribution: The Role of the Water Balance. *The American Naturalist*, **135**, 649–670.

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia Revisited: Individualistic Responses of Species in Space and Time. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 661–671.

Suggate RP, Almond PC (2005) The Last Glacial Maximum (LGM) in western South Island, New Zealand: implications for the global LGM and MIS 2. *Quaternary Science Reviews*, **24**, 1923–1940.

Taylor EB, Tamkee P, Keeley ER, Parkinson EA (2010) Conservation prioritization in widespread species: the use of genetic and morphological data to assess population distinctiveness in rainbow trout (Oncorhynchus mykiss) from British Columbia, Canada. *Evolutionary Applications*, **4**, 100–115.

Trewick SA (2000) Molecular evidence for dispersal rather than vicariance as the origin of flightless insect species on the Chatham Islands, New Zealand. *Journal of Biogeography*, **27**, 1189–1200.

Trewick SA, Wallis GP (2001) Bridging the "beech-gap": New Zealand invertebrate phylogeography implicates Pleistocene glaciation and Pliocene isolation. *Evolution; International Journal of Organic Evolution*, **55**, 2170–2180.

Trewick SA, Wallis GP, Morgan-Richards M (2000) Phylogeographical pattern correlates with Pliocene mountain building in the alpine scree weta (Orthoptera, Anostostomatidae). *Molecular Ecology*, **9**, 657–666.

Trewick SA, Wallis GP, Morgan-Richards M (2011) The Invertebrate Life of New Zealand: A Phylogeographic Approach. *Insects*, **2**, 297–325.

Tsumura Y, Suyama Y (1998) Differentiation of Mitochondrial DNA Polymophisms in Populations of Five Japanese Abies Species. *Evolution*, **52**, 1031–1042.

Wallis GP, Judge KF, Bland J, Waters JM, Berra TM (2001) Genetic diversity in New Zealand Galaxias vulgaris sensu lato (Teleostei: Osmeriformes: Galaxiidae): a test of a biogeographic hypothesis. *Journal of Biogeography*, **28**, 59–67.

Wallis GP, Trewick SA (2001) Finding Fault with Vicariance: A Critique of Heads (1998). *Systematic Biology*, **50**, 602–609.

Waples RS, Gaggiotti O (2006) Invited review: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419–1439.

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

Waters JM (2011) Competitive exclusion: phylogeography's "elephant in the room"? *Molecular Ecology*, **20**, 4388–4394.

Waters JM, Craw D (2006) Goodbye Gondwana? New Zealand Biogeography, Geology, and the Problem of Circularity. *Systematic Biology*, **55**, 351–356.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, **38**, 1358–1370.

Widmer A, Lexer C (2001) Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology & Evolution*, **16**, 267–269.

Wilmshurst JM, McGlone MS, Leathwick JR, Newnham RM (2007) A pre-deforestation pollen-climate calibration model for New Zealand and quantitative temperature reconstructions for the past 18 000 years BP. *Journal of Quaternary Science*, **22**, 535–547.

Wilton AD, Breitwieser I (2000) Composition of the New Zealand seed plant flora. *New Zealand Journal of Botany*, **38**, 537–549.

Winkworth RC, Wagstaff SJ, Glenny D, Lockhart PJ (2002) Plant dispersal N.E.W.S from New Zealand. *Trends in Ecology and Evolution*, **17**, 514–520.

Winkworth RC, Wagstaff SJ, Glenny D, Lockhart PJ (2005) Evolution of the New Zealand mountain flora: Origins, diversification and dispersal. *Organisms Diversity and Evolution*, **5**, 237–247.

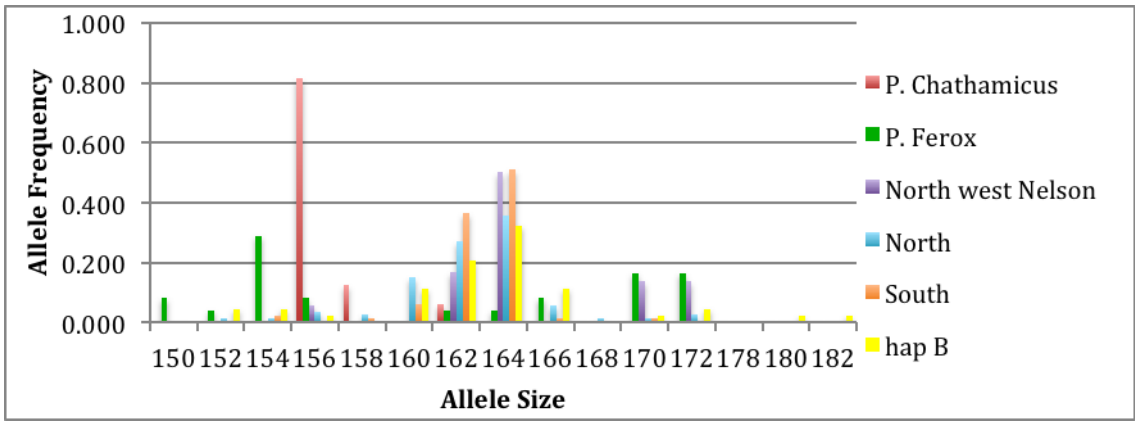
Wright S (1943) Isolation by Distance. *Genetics*, **28**, 114–138.

Young AG, Schmidt-Adam G, Murray BG (2001) Genetic variation and structure of remnant stands of pohutukawa (Metrosideros excelsa, Myrtaceae). *New Zealand Journal of Botany*, **39**, 133.

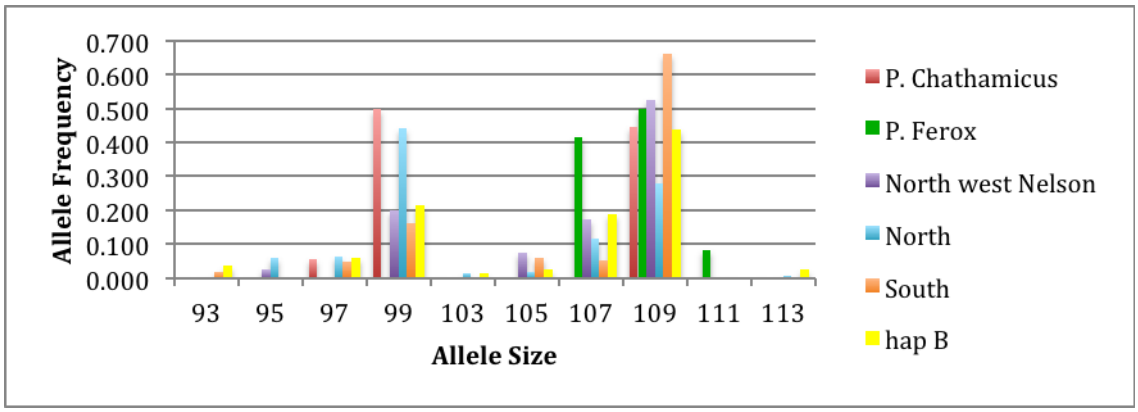
Appendix I

Allele frequencies for population clusters. Clusters based on Structure analysis with Northwest Nelson and Murchison populations being kept together as a separate cluster. A cluster of populations with high rps4 haplotype B was also created.

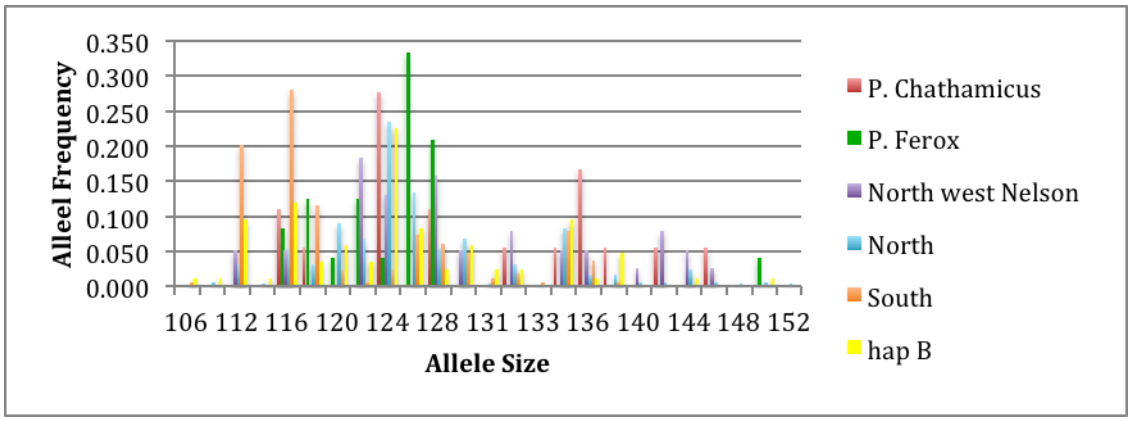
Allele frequencies for population clusters Locus Psecra 2



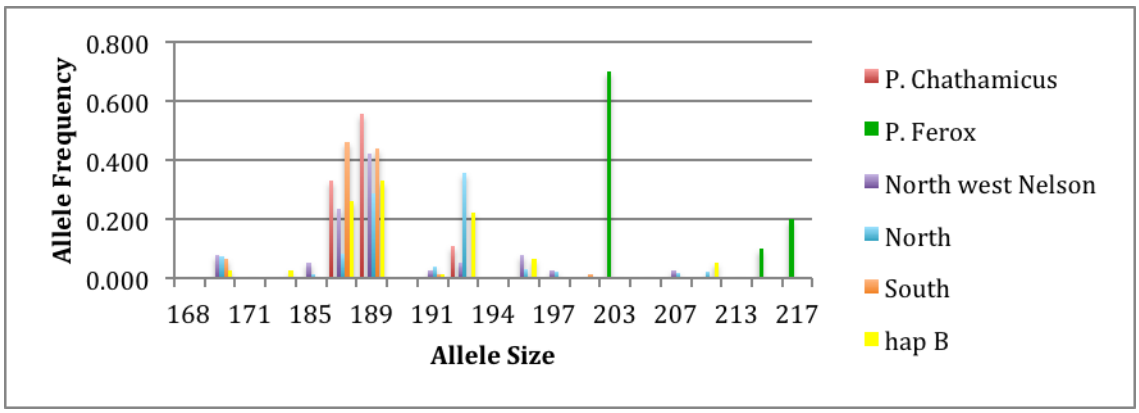
Allele frequencies for population clusters Locus Psecra 3a



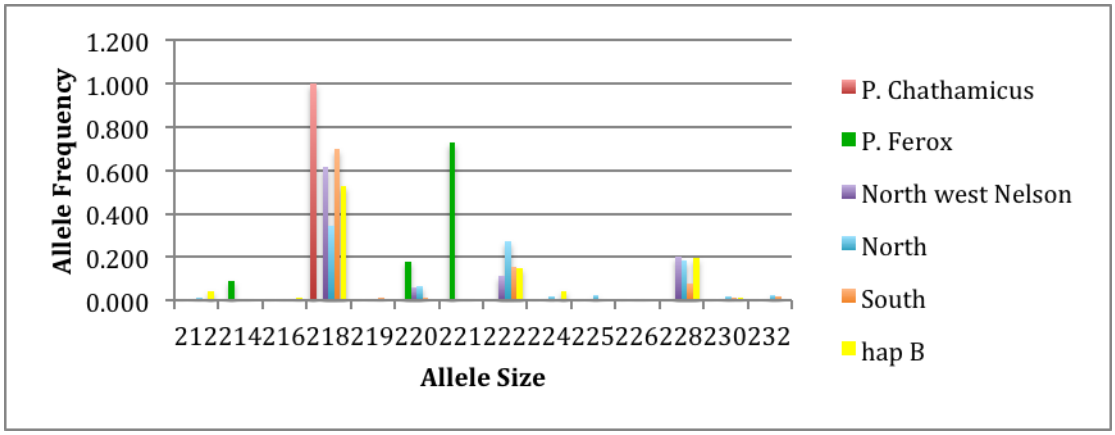
Allele frequencies for population clusters Locus Psecra 3b



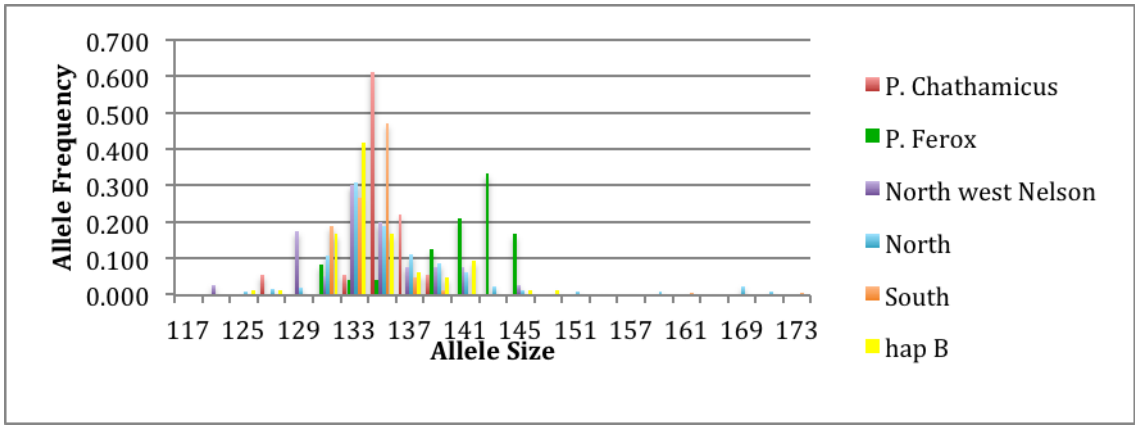
Allele frequencies for population clusters Locus Psecra 5



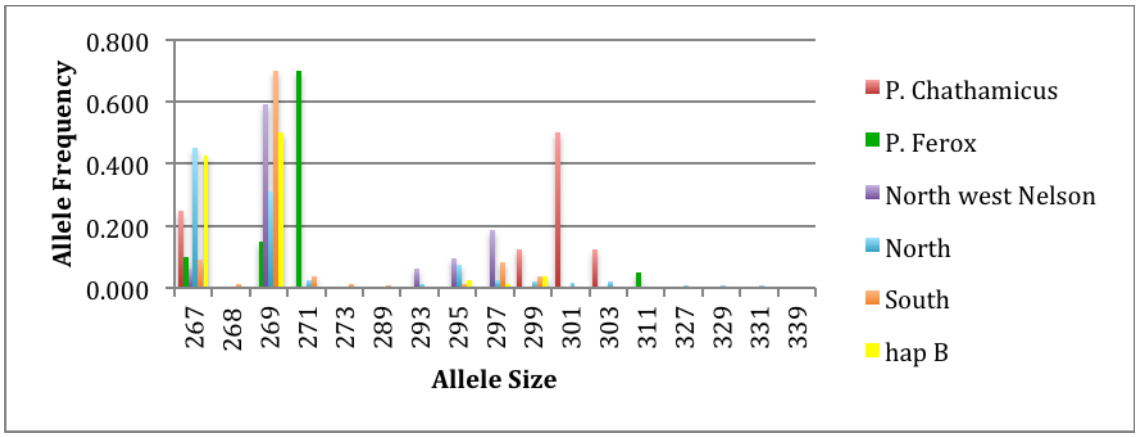
Allele frequencies for population clusters Locus Pseles 5



Allele frequencies for population clusters Locus Psearb 9

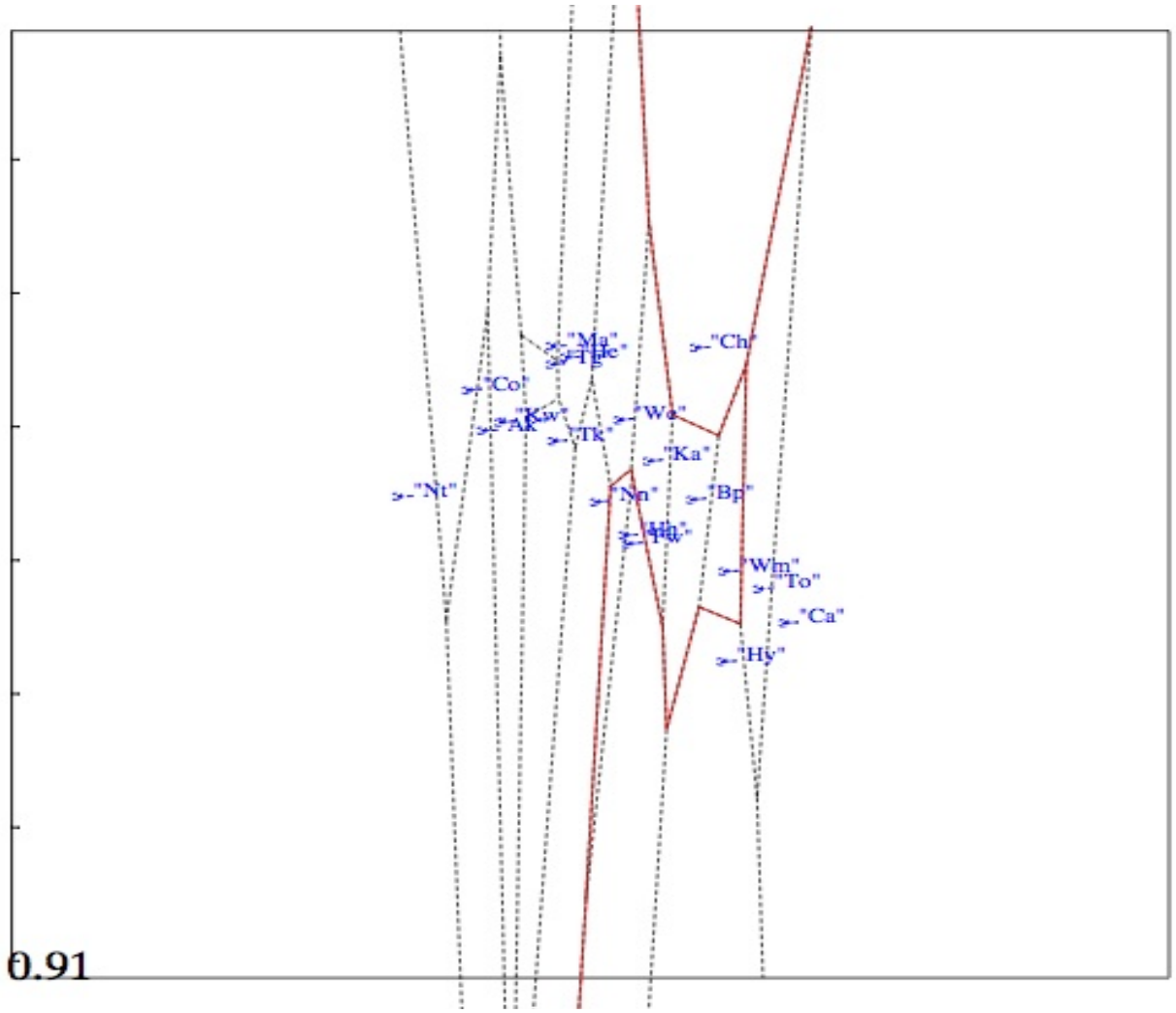


Allele frequencies for population clusters Locus M13J



Appendix II

SAMOVA output Voronoi diagram at K=3. The three clusters at K=3 are *Pseudopanax chathamicus*, North analogous to the Northern STRUCTURE cluster and south, again analogous to the southern STRUCTURE cluster.



Appendix III

Microsatellite data from parent and seed analysis.

	Psecra 2			Psecra 3a			Psecra 5			Pseles 5			Psearb9	
mgoo3 parent	162	164		93	97		187	193		218	218		135	137
mg003a	162	170		99	109		187	193		212	224		135	135
mg003b	164	164		99	109		193	193		218	218		135	137
mg003c	*	*		*	*		193	193		218	218		135	137
mg003d	162	164		95	99	109	193	193		218	218		137	137
mg003e	162	164		97	99	109	187	193		218	218		135	135
mg003f	162	164	166	*	*		193	193		218	218		-	-
mg003g	162	164	166	*	*		193	193		224	226		135	135
mg005 parent	*	*		99	109		193	193		224	224		135	135
m005h	166	158		99	109		193	193		*	*		135	137
mg005i	*	*		97	99	109	193	193		224	224		135	135
mg005j	158	166		99	109		193	193					133	135
mgoo5k	164	166		97	99	109	193	193		*	*		135	135
mg005l	164	164		99	109		193	193		224	224		135	135
mg005m	160	166		99	107	109	193	193		218	224		133	135
mg005n	164	166		99	107	109	193	193		224	226		135	135