Conservation genetics of the endemic root holoparasite, *Dactylanthus taylorii*

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

Master of Science

in

Plant Biology

At Massey University, Palmerston North,

New Zealand

Todd Graham Bruce McLay

2011

Abstract

Dactylanthus taylorii is a New Zealand endemic root holoparasite in the family Balanophoraceae. The vegetative body is extremely reduced and the majority of the plant grows underground attached to its host root. Flowers are dioecious and are pollinated by the short tail bat; a native seed disperser is unknown. Pollen records indicate *D. taylorii* was formerly more widespread over the North Island of New Zealand but habitat fragmentation and browsing by introduced mammals is thought to have severely reduced population sizes. The species is classified as nationally vulnerable and conservation management is overseen by the New Zealand Department of Conservation. Here, a conservation genetics approach was taken in order to understand the genetic variation and structure of *D. taylorii* populations.

Previously, the genetic structure of 17 populations had been studied using randomly amplified polymorphic DNAs (RAPDs) and a broad geographical pattern was identified, but was not suitable for conservation management. Alternative molecular markers are microsatellites, which have many favourable attributes for use in conservation genetic studies. Microsatellites allow for estimates of allele frequencies and therefore can be used to determine heterozygosity and their high mutation rate can be used to detect more recent changes in the genetic structure of populations. Next-generation sequencing was used in order to develop microsatellite markers as this method does not require cloning. From the 62,000 sequences obtained, 4,000 microsatellites were identified and primers were able to be designed for 750 repeats. From this primer pool, 72 were chosen to be screened and ten microsatellite loci were found to be polymorphic and consistently amplifiable. These ten were used to genotype 241 *D. taylorii* individuals from 31 populations.

Across all populations a high number of alleles were identified, although a high percentage of these were private alleles. Within-population assessment of genetic variation indicated that many populations have low levels of genetic diversity and a high proportion of homozygotes. A high degree of genetic differentiation was detected and was found to be strongly correlated to geographic distance between populations. Also, populations grouped into two, three or eight clusters that were reflective of geography. Possible explanations for the geographic pattern observed include

volcanism, mountains as physical barriers to gene flow, habitat availability and gene flow mediated by the short tail bat.

This information suggests that although populations are secure from environmental risks such as habitat loss or herbivory, there is a genetic threat to extinction. In order to increase genetic variation within-populations, translocation of genetic variation (i.e. pollen or seed) is suggested between geographically adjacent populations.

Acknowledgements

Firstly, I would like to thank my supervisors Jennifer Tate and Vaughan Symonds for giving me the opportunity to do exactly what I wanted for my thesis. Your advice and support is hugely appreciated.

I am also grateful for the support given to me by Avi Holzapfel and Paul Cashmore from the Department of Conservation. Avi, thank you for your enthusiasm towards the project from its development, and for sharing your wealth of knowledge about *Dactylanthus*. Paul, without your support and the backing of the Recovery Group I never would have got close to collecting from as many populations as I have. I also have to thank countless DOC staff from all around the North Island for collecting samples for me and being welcoming when I tagged along during monitoring. I hope the results from this study aids conservation of *Dactylanthus* and the time you all put in worthwhile.

Without trying a new method of DNA extraction from the advice of Lara Shepherd, it's entirely possible I would still be trying to get something other than a brown sludge from my samples. Likewise, without the computer expertise of Patrick Biggs I would still be sitting at a computer trying to extract my 454 sequences from the Otago University server. Also, Chrissen Gemmill supplied the DNA for several populations collected from a previous study. Thank you all for your brief but extremely important help.

Thanks to all the members in LoST for undoubtedly at one stage or another hearing me express frustration at any one of DNA extraction, genotyping or plotting points on a map, and responding with a smile (or smirk). You guys are alright.

Thank you Mum, Dad, Jodie and Scott for letting me wax lyrical *ad nauseum* about *Dactylanthus* and pretending to know what I was talking about. Most importantly, thanks for being proud of me.

I would also like to thank Massey University for the Massey Masterate, J.P Skipworth and Coombs Memorial scholarships.

Abbreviations

A	-	number of alleles
AFLP	-	Amplified Fragment Length Polymorphism
BLAST	-	basic local alignment search tool
bp	-	base pairs
cpDNA	-	chloroplast DNA
CASS	-	cheaply amplified size standard
CTAB	-	hexa-decetylammonium bromide
DATA	-	*Dactylanthus taylorii* primer for microsatellite loci
DNA	-	deoxyribonucleic acid
dNTP	-	deoxyribonucleotide triphosphate
DOC	-	Department of Conservation (New Zealand)
EDTA	-	ethylene diamine tetra-acetic acid
ERMA	-	Environmental Risk Management Authority
EST	-	expressed sequence tag
F_{IS}	-	component of Wright's (1921) fixation index, used to define within population structure by calculating the average observed heterozygosity of an individual relative to the expected heterozygosity of individuals in the population it belongs to
F_{ST}	-	component of Wright's (1921) fixation index, used to define between population structure by comparing the expected heterozygosity of individuals within a subpopulation to the total expected heterozygosity of individuals across all populations
H_E	-	expected heterozygosity
H_O	-	observed heterozygosity
H_T	-	species-wide expected heterozygosity
HWE	-	Hardy-Weinberg equilbrium
IUCN	-	International Union for Conservation of Nature
N_A	-	number of alleles per locus within a population
N_E	-	effective number of alleles per locus within a population
Mb	-	mega base pairs (=1,000,000 bp)
mtDNA	-	mitochondrial DNA
PDL	-	primer designed locus

PCR	-	polymerase chain reaction
QTL	-	quantitative trait loci
RAPD	-	Randomly Amplified Polymorphic DNA
R_{ST}	-	derivation of F_{ST} that accounts for the high-rate stepwise mutation model of microsatellites (Slatkin 1995)
SNP	-	Single Nucleotide Polymorphism
SSR	-	simple sequence repeat (microsatellite)
STE	-	sucrose, TRIS, EDTA
Tris	-	tris(hydroxymethyl)aminomethane
TVZ	-	Taupo Volcanic Zone
UPGMA	-	Unweighted Pair Group Method with Arithmetic Mean
%P	-	percentage of polymorphic loci

Contents

Figures

1 Introduction

2 Microsatellite marker development for the endangered root parasite *Dactylanthus taylorii* using high-throughput sequencing

3 Conservation genetics of the endemic root holoparasite, *Dactylanthus taylorii*

Tables

1 Introduction

2 Microsatellite marker development for the endangered root parasite *Dactylanthus taylorii* using high-throughput sequencing

3 Conservation genetics of the endemic root holoparasite *Dactylanthus taylorii*

1 Introduction

1.1 *Dactylanthus taylorii*

Dactylanthus taylorii (pua o te reinga or dactylanthus) is a fully parasitic angiosperm endemic to New Zealand (Holzapfel 2001). *Dactylanthus* is a monotypic genus and represents the most southern member of an otherwise tropical and subtropical family of obligate root parasites, the Balanophoraceae (Nickrent *et al.* 2005). Balanophoraceae is comprised of 17 genera and 44 species that are all characterised by greatly reduced vegetative and floral organs and an absence of chlorophyll. The loss of roots, stems and leaves and reduction in flower parts has made placement of the Balanophoraceae within the angiosperms based on morphological characters difficult. Owing to a similar parasitic life form, the family is believed to be within the parasitic order Santalales, but this placement is currently not well supported and similarities could be due to convergent evolution (Kuijt 1969). Recent molecular work reinforces the placement of Balanophoraceae near Santalales (Nickrent *et al.* 2005) but further study is required.

Dactylanthus taylorii obtains nutrients and water from a host tree root by attaching via a subterranean haustorial connection, which causes hypertrophic growth of the root creating a fluted disc-like woodrose (Holzapfel 2001). The underground growth and parasitic attachment of *D. taylorii* means it is difficult to identify the host tree, but there are an estimated 30 species identified as putative hosts (La Cock *et al.* 2005). The body of the plant has been referred to as a rhizome or tuber and groups of individuals as a clump (Figure 1.1). The homology of this structure is uncertain (Holzapfel 2001). Estimates of lifespan have been made dating rings of large woodrose specimens with the maximum age found of 30 years (Ecroyd 1995), and some sites have been monitored for 20 years. Seeding trials have shown that plants can flower as early as four years after germination. The number of flowers per individual is variable, and clumps have been observed with greater than twenty flowers in a single season (pers. observation). Plants do not flower every year, and it can be difficult to tell if a tuber is still alive if it is not flowering (A. Holzapfel, pers. comms.).

Figure 1.1: A typical *Dactylanthus taylorii* clump with male inflorescences emerging. The arrow indicates the body of the plant. Tree roots can be seen growing over the top of the clump and may be a potential host through vegetative reproduction. Photo credit: Dave Mudge, Nga Manu Trust

Dactylanthus taylorii is mostly dioecious, but rarely hermaphrodites can occur (<1%), with thousands of minute flowers (<1cm) borne on an inflorescence (Holzapfel 2001) (Figure 1.2). An inflorescence develops within the tuber body below the soil and only breaches the surface when it is in the final stages of development and is preparing to open (Moore 1940). The male flower is composed of a two anthers with fused filaments while the female flower consists of a single ovary with a single, elongated style and papillate stigma (Holzapfel, pers. com). Flowers of either sex have two reduced perianth members (Holzapfel 2001). Both sexes produce copious amounts of nectar which has been shown to contain a mammalian hormone, squaline (Ecroyd *et al.* 1995). The short-tailed bat (*Mystacina tuberculata*) is the only known natural pollinator (Ecroyd 1996a). It is generally assumed that gravity is the main mode of dispersal as a native seed disperser is unknown (Ecroyd 1996b), but annelidochory (worm-dispersal) has been suggested (Thorsen *et al.* 2009). Plants are also able to reproduce vegetatively via infectious root-like structures that form at the base of an inflorescence (Figure 1.2)

(Moore 1940). When a host root comes into contact with these organs they are able to recognise and infect the root to create a parasitic connection that will eventually form a new plant (Holzapfel 2001).

Figure 1.2: Close-up of a male *Dactylanthus taylorii* inflorescence. Photo credit: Dave Mudge, Nga Manu Trust.

Pollen records indicate that *Dactylanthus taylorii* was once widespread on the North Island and the top of the South Island (Macphail & Mildenhall 1980; McGlone & Topping 1983; Mildenhall & Alloway 2008). The present distribution is greatly reduced compared to the possible past distribution based on suitable habitat/host environments (Figure 1.2). The northern limit is Puketi Forest and the southernmost plants are found at Pukaha/Mount Bruce (both populations possess few individuals) and most populations are now found around Lake Taupo and the Central Plateau towards Lake Waikaremoana. Population sizes range from fewer than ten plants to greater than 1000, though exact counts are difficult. Plants are generally found in forest margins or areas of regrowth as this is the typical habitat for the most common of supposed host tree species such as lancewood (*Pseudopanax crassifolius*), lemonwood (*Pittosporum eugenioides*), fivefinger (*Pseudopanax arboreus*), kohuhu (*Pittosporum tenuifolium*), mahoe

(*Melicytus ramiflorus*), and putaputaweta (*Carpodetus serratus*) (Ecroyd 1996b; Holzapfel 2001).

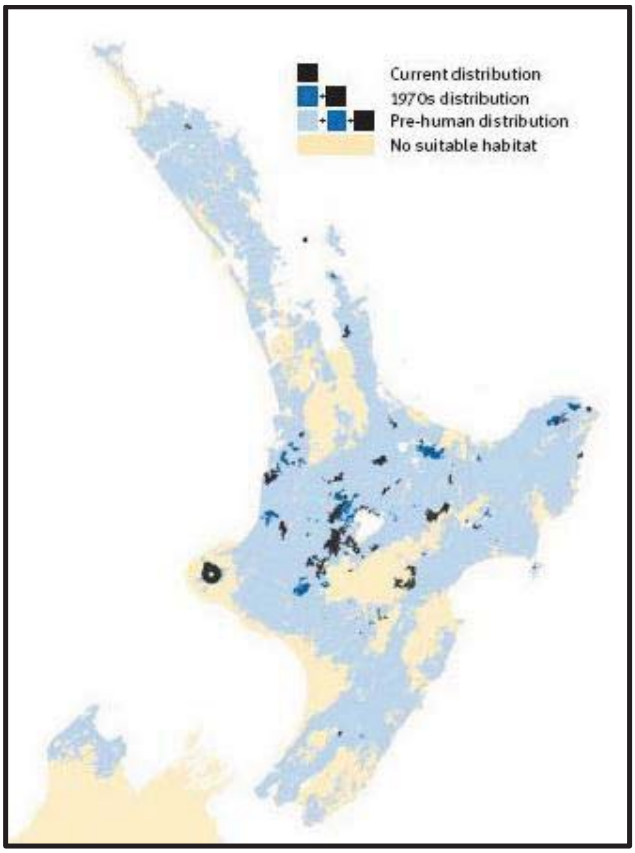

Figure 1.3: Past and present (2007) distribution of *Dactylanthus taylorii* in New Zealand (van Bunnik *et al.* 2007)

1.2 Conservation of *Dactylanthus taylorii*

The Department of Conservation (DOC) formed the Dactylanthus Recovery Group in 1993 because of concerns relating to the persistence of the species in the wild and a decline in its distribution over the past 50 years (Holzapfel 2001), which was first observed in the 1980s (Given 1981). Habitat destruction would have been the initial agent of decline. Before human settlement, New Zealand is believed to have been 80-90% forest vegetation; lowland forest sites such as those favoured by *Dactylanthus taylorii* host trees would have been among the first targets for removal by early

colonisers using fire (McGlone 1989). Seral or scrubby habitat are also more likely candidates for conversion to exotic forest in the present day (Ewers *et al.* 2006). Collectors dig up the *D. taylorii*-root complex and boil to remove the tuber and sell the woodrose (Holzapfel 2001), which is formed at the parasite-host root connection and is a feature found in several other plant parasites (Dzerefos *et al.* 1999). While this practice has been discouraged by DOC and collecting has been reduced (La Cock *et al.* 2005), plants have been dug up at sites used in this study and specimens occasionally appear on TradeMe for sale (pers. comms. Paul Cashmore). The biggest threat to *D. taylorii* persistence in the wild is browsing of inflorescences by introduced mammals, especially the brush-tailed possum, *Trichosurus vulpecula*, but also rats, mice and pigs (Ecroyd 1996b; Ferreira 2005). Attracted by the nectar, a possum will remove an inflorescence from the tuber and consume the entire upper portion including all the flowers, potentially removing an entire seasons worth of flowering (Figure 1.4). Rats and mice usually do considerably less damage, and potentially play a role in pollination and seed dispersal in populations that lack bats (Ecroyd 1996b). Wasps have been observed robbing nectar by chewing through bracts and thus never coming into contact with the anther or stigma to facilitate pollination (Holzapfel 2001). To combat browsing of inflorescences, most populations are managed closely by DOC and individuals are caged to keep possums out, but allow bats in to pollinate (La Cock *et al.* 2005).

Figure 1.4: A possum browsing an inflorescence. The *Dactylanthus taylorii* is seen moss covered in the foreground, further illustrating the cryptic nature of the plant. Photo credit: Dave Mudge, Nga Manu Images.

Many populations do not overlap with short-tailed bat ranges and therefore many populations only set seed via pollination by rats and mice or by hand pollination (Ferreira 2005). Short-tail bats have also undergone a decrease in population size since human arrival and significant effort has gone into conserving this species (Lloyd 2003; O'Donnell *et al.* 2010). *D. taylorii* is listed as "Nationally Vulnerable" in the most recent list of New Zealand threatened plants (de Lange *et al.* 2009), a positive change from the 2004 classification of "Serious Decline"(de Lange *et al.* 2004). It is also ranked as a species of highest conservation importance by DOC (Dopson *et al.* 1999).

An important objective of the *Dactylanthus* Recovery Group is to assess the amount of genetic variation across the entire range of *D. taylorii* (La Cock *et al.* 2005). Holzapfel *et al.* (2002) studied levels of genetic variation within and between 16 populations using randomly amplified polymorphic DNAs (RAPDs). They found that the populations studied were unique and that geographic distance was not a good indicator of genetic similarity, but there was a distinct eastern-western division shown by UPGMA cluster analysis (Holzapfel *et al.* 2002). The main cause of this division was hypothesised to be

volcanism, specifically the continual volcanic activity from the Taupo caldera creating the Taupo Volcanic Zone. The TVZ is a 300km long area of intense vulcanism situated north-to-south of the North Island (Manville *et al.* 2009) which reflects the division of genetic variation found in *D.taylorii* as well as several other plants and animals (Shepherd & Perrie 2011).

The discovery of new geographic outlier populations and the limited resolution provided by the RAPD study means that current understanding is not sufficient to guide management (La Cock *et al.* 2005). Most known populations are not at risk from the loss of habitat as they are legally protected and individuals within populations are protected from browsing by cages, so immediate protection of the species is secured. However, it is unknown what the effects of reduction in population sizes, population isolation and loss of pollinator has had on the genetic variation of the species which may have long term implications for the survival of individual populations and *D. taylorii* as a species.

The genetic diversity identified within populations could reveal the historical impact of small population sizes, such as drift and inbreeding, and rare natural pollination. Comparisons of genetic structure between populations could add to interpretation of the east-west demarcation and the effects of habitat fragmentation and geographic isolation on gene flow. An objective of the Recovery Group is to establish new populations via seed translocation, so seed sources can be selected based on how the genetic structure is reflected by geography. If direct action is required to increase genetic variation in isolated populations then potential sources of seed or pollen could be identified. Information obtained using a conservation genetics approach could be paramount to the immediate and on-going management of *D. taylorii* and the persistence of the species.

1.3 The Field of Conservation Biology

Conservation biology has been described as a "crisis discipline", an area of science that desperately needs its patterns and processes described and understood (DeSalle & Amato 2004). The current extinction rate of Earth's taxa is higher than any of the five great geological extinctions and is almost solely attributed to anthropogenic causes. The most influential human induced factors causing extinction were described by Diamond

(1989) as the "evil quartet": habitat degradation and fragmentation, overharvesting, exotic species, and chains of extinction.

Conservation biology was formed in reaction to the need for a discipline with an ultimate goal of protecting and perpetuating biological diversity (Meine *et al.* 2006), and is influenced by many different disciplines including ecology, genetics, wildlife biology and resource management (Frankham *et al.* 2002). The International Union for Conservation of Nature (IUCN) defines three components of biodiversity – community/ecosystem, population/species and genetic, which are studied in markedly different ways but ultimately contribute towards understanding and protecting biological diversity (Redford & Richter 1999).

1.3.1 Conservation genetics

In their landmark book *Conservation and Evolution*, Frankel and Soule (1981) presented a scientific framework illustrating the importance of understanding the evolutionary history and genetic variation of endangered species for conservation. The central dogma of conservation genetics is that all genetic diversity is significant and therefore important to conserve (Pertoldi *et al.* 2007). The underpinning theory is that evolution acts on genetic diversity, so organisms or populations with low genetic diversity are more likely to be unable to adapt to new environmental conditions presented to them and may subsequently go extinct (Falk & Holsinger 1991). This is especially important when considering the anthropogenic pressure many organisms must overcome to survive (Stockwell *et al.* 2003). Conservationists concerned with endangered plant species are confronted with an important yet intimidating assignment, as many rare species are reduced to a few populations or individuals and often require urgent study to understand the genetics and biology to guide management (Holsinger & Gottlieb 1991).

1.3.2 The conservation genetics paradigm for plants

Many rare plant species exist only in small or isolated populations in fragmented habitats (Young *et al.* 1996) and so are at risk from random processes including: demographic stochasticity, environmental stochasticity, unpredictable catastrophic events and genetic stochasticity (Ouborg *et al.* 2006). Endangered plants generally

suffer from the consequences of small population size, such as genetic drift, inbreeding and reduced gene flow due to isolation. These factors will decrease genetic variation within a population and increase variation among populations, reducing the evolutionary potential of the population or species (Ellstrand & Elam 1993). Conservation genetic studies in plants aim to understand or predict the genetic threats to population extinction in order to guide conservation management (Ouborg *et al.* 2006).

1.3.3 Why is the organism rare?

Plant rarity is a function of species abundance as well as the range it occupies, which is the result of either natural or man-made causes, or a combination of both (Gaston 1994a). This can make establishing conservation priority ranking difficult, as taxa that are naturally rare are often overlooked despite potentially suffering from the same risks to extinction as the formerly abundant species. Alternatively, naturally rare taxa might be optimally adapted to the environmental conditions they have evolved in and can buffer many of the deleterious effects of small population size, whereas species that are recently rare are not as resistant (Huenekke 1991). Therefore, while rarity is often evaluated using geographical range, population size and habitat occurrence, the evolutionary history of the organism is perhaps the most critical determinant of rarity (Qiao *et al.* 2010).

Geographical studies of rarity have indicated the high proportion of endangered species on islands (Rieseberg & Swensen 1996). Islands typically have a high diversity of habitats over a small geographical range, so plant populations are small and are often adapted and restricted to specific habitats they have diversified into (Sakai *et al.* 2002). Specialisation for an uncommon habitat is frequently a cause for natural rarity in plant species (Gaston 1994b). The isolation of islands leads to higher levels of inbreeding and endemic species which are evolutionarily naïve to human-introduced biota such as pests, herbivores and pathogens (Kelly & Sullivan 2010). These factors are believed to influence lower levels of genetic variation in island taxa when compared to mainland species and lead to increased extinction frequencies (Frankham 1997). However, island species are not rare by default and more study is required in order to corroborate this theory (Jamieson 2007) with a focus on understanding differences in genetic structure of island compared to mainland rare species (Franks 2010).

Naturally rare species tend to have life history traits that allow persistence at low population numbers, such as adaptation to inbreeding or to low pollinator availability (Huenneke 1991). Pollination systems are often linked to causes for rarity as co-adaptation has led to pollinator reliance for seed production and dispersal in many taxa and reduction in pollinator densities has a direct negative impact on plant reproduction (Dutra *et al.* 2009; Huenneke 1991). A result of pollinator loss and population isolation can be inbreeding, which increases homozygosity in a population and can lead to the accumulation of deleterious alleles (Lopez *et al.* 2009). However, in locally adapted populations it may be more beneficial to inbreed than outcross (Lopez *et al.* 2009; Hereford 2010) if there is a periodic purging of the genetic load via outcrossing (Porcher *et al.* 2009). Studies of genetic variation in plants believed to be naturally rare have shown that the small population size effects that are expected to cause low levels of genetic diversity were not significant and in fact genetic variation was high, suggesting that rarity is not always correlated with risk of extinction. If the habitat of the rare plant is stable, it may not require active management for conservation if it is adapted to a small population size (Qiao *et al.* 2010).

1.4 Marker selection in conservation genetics

Selecting the most suitable genetic marker system is vital to the success of any conservation genetics study, but the increasingly large number of potential tools available makes the decision of which marker to use difficult. Discussing the technical details of each marker used in conservation genetic studies is beyond the scope of this review, but is covered well in Semagn *et al*. (2006). Instead, the focus will be on the favourable attributes of markers commonly used in conservation genetic studies, presented in Figure 1.5, Table 1.1 and Table 1.2.

Figure 1.5: Favourable marker attributes for conservation genetics studies

Modified from Sunnucks (2000). See Table 1.1 for a comparison of commonly used markers

Single or multi-locus*:* The decision between using a single locus marker and a multi-locus marker involves a trade-off between accuracy and technical difficulty. Multi-locus markers are generally simpler to design. However this has significant limitations when considering the variation detected, because single locus markers provide more accurate and detailed data.

Assayable by PCR*:* PCR-based methods use DNA, which is preferable to work with over proteins required for allozyme analysis. DNA can be extracted from old material, such as herbarium specimens, PCR can work on low quantities of degraded DNA, and sample collection for DNA extraction is simpler and less invasive than for protein analysis.

Codominant*:* Dominant markers can only be scored as present or absent, whereas for codominant markers each of the two alleles at a given locus in an individual can be scored and analysed. With codominant markers, heterozygotes and homozygotes can be distinguished, and estimates of allele frequencies can be made and compared between populations.

Number of loci available: Population genetic studies require the use of enough variable loci to detect a signal, so markers that have many potential loci in the genome should be utilised. Many separate loci should be used in order to decrease the chances of using linked loci with similar biological histories and marker coverage is representative of the entire genome.

Low technical requirement: Markers that do not require any prior knowledge of DNA sequence are often less technically demanding to create, but often yield less information relative to cost. Transferability of some markers between taxa is good (cpDNA, nuclear DNA) whereas other markers can only rarely be used in related taxa (microsatellites) or not at all (RAPDs, AFLPs). The time and effort required to screen markers is also highly variable.

Comparability of data among studies: Direct comparability refers to the amplification of homologous DNA loci over a taxonomic range; comparable data can derive conclusions from a wider range of data. Indirect comparability can be made using models of molecular evolution.

Overall variability of marker: Marker variability depends on the amount of variation per locus and number of loci obtained readily. This assessment is based on the approximate outcome of an average marker system development.

Table 1.1: Molecular marker attributes important for conservation genetics – modified from Sunnucks (2000). These attributes are considered to be important in order to answer significant questions in conservation genetics (Table 1.2).

Marker attributes	Single locus	PCR assay	Codominant	Number of loci available	Technical requirement	Comparability of data among studies	Overall variability
Allozymes	Yes	No	Yes	Moderate	Moderate	Direct	Low-moderate
Microsatellites	Yes	Yes	Yes	Many	High	Indirect	High
SNPs	Yes	Yes	Yes	Many	High	Indirect	Low-high
RAPD	No	Yes	No	Many	Moderate	Limited	High
AFLP	No	Yes	No	Many	Moderate	Limited	High
Nuclear gene sequence	Yes	Yes	Yes	Moderate	Low	Direct	Moderate
cpDNA/mtDNA	Yes	Yes	No	Moderate	Low	Direct	Low-high

Table 1.2: Marker application to different conservation genetic questions – modified from Wayne and Morin (2004) and Ouborg *et al*. (1999). This represents a subjective analysis from reviewing the application of molecular markers to common conservation genetic questions. - not suitable; +/- should be avoided; + suitable; ++ well suited; +++ very well suited

Marker uses	Assessing genetic variation	Resolving population structure	Determining gene flow	Estimating effective popn. size	Detecting hybridisation	Resolving taxonomic uncertainty	Identifying targets for conservation	Informing translocation regimes
Allozymes	++	++	+	-/+	++	+	+	-/+
Microsatellites	+++	+++	+++	++	++	-/+	++	++
SNPs	++	++	+++	++	++	+	++	++
RAPD	+/-	++	+	-/+	-/+	-	+	+
AFLP	++	++	+	-/+	++	-	+	+
Nuclear gene sequence	+	+	++	+	+	+++	+++	-/+
cpDNA/mtDNA	+	+	++	++	++	+++	++	+

1.5 The scope of conservation genetics in plants

The field of conservation genetics covers many aspects important to conservation management of rare plant species including: detecting hybrids (Allendorf *et al.* 2001), resolving taxonomic uncertainty (Callmander *et al.* 2005) and estimating the effective population size (Wang 2005). Here the focus will be on areas of conservation genetics relevant to *Dactylanthus taylorii*.

1.5.1 Assessing genetic variation

The level of genetic variation identified within and among populations over the range of the taxa concerned provides information about the species chance of persistence, with low levels of genetic variation suggesting a greater chance of extinction due to reduced evolutionary potential (Wayne & Morin 2004). Small, isolated populations often have low levels of genetic diversity within populations but high levels of genetic diversity between populations (Ellstrand & Elam 1993). An assessment of the level of genetic diversity provides an estimation of the evolutionary potential of populations and the species (Ouborg *et al.* 2006), which can be used to guide conservation management (DeSalle & Amato 2004). Identifying patterns of genetic diversity is often the first objective of a conservation genetics study in order to inform conservation activities, maximise diversity, and minimise genetic erosion (Kramer & Havens 2009). This is typically done by using highly variable and selectively neutral markers under the assumption that neutral variation reflects the genome-wide genetic variation, including adaptive and neutral variation that may be selectively important for the organism's persistence (Hedrick 2001).

Various markers have been used to assess rare species' genetic variation, e.g., RAPDs (Holzapfel *et al.* 2002), amplified fragment length polymorphisms(AFLPs) (Honnay *et al.* 2007) and allozymes (Fleishman *et al.* 2001), but microsatellites are being utilised more frequently as they are codominant, so estimates of heterozygosity can be made, and are also polymerase chain reaction (PCR) based (Selkoe & Toonen 2006). For codominant loci, genetic variation is typically estimated by allelic richness (i.e., the average number of alleles for all loci in a population, percentage of loci that are polymorphic and the private alleles) or allele frequencies (i.e., observed and expected heterozygosity) (Hughes *et al.* 2008). Allelic richness is expected to determine a species

evolutionary response over generations whereas allelic frequencies are more likely to influence response to selection in the short term, therefore both measures of genetic diversity are important to conserve for an endangered species (Neel & Cummings 2003).

Once genetic variation has been measured, the next step is to compare the values obtained to those found in other studies. Where possible, comparisons should be made between studies that use the same genetic marker type as they not only measure diversity using different parameters, but due to differing marker biology they can yield vastly different estimates of genetic variation. For example, despite both being codominant markers, microsatellites generally have a far larger number of alleles than allozymes and so show significantly different levels of genetic variation even within the same species (Gao *et al.* 2002; Takahashi *et al.* 2011). In conservation genetic studies, comparisons are typically focused on one or more of the following: species from the same geographical region, as they are assumed to have a similar biogeographical history which could affect genetic variation (Riley *et al.* 2010); between other rare species, as isolation, reduction in population size and other characteristics of endangered taxa would affect genetic variation (Furches *et al.* 2009); and species with similar life histories (Aegisdottir *et al.* 2009). Life history traits such as growth form, seed dispersal mechanism and breeding system are regarded as the major drivers of evolution in flowering plants (Smith & Donoghue 2008) and the main factors affecting genetic diversity and structure within and between population (Loveless & Hamrick 1984). Meta-analyses of studies comparing between the effect of life histories on levels of genetic variation find that outcrossing perennials exhibit higher levels of diversity and lower levels of population structure than short-lived inbreeders (Hamrick *et al.* 1991; Hamrick & Godt 1996; Nybom 2004).

1.5.2 Population structure

A population is the basic unit of evolution and also serves as a unit for conservation. Therefore, understanding the population structure of an endangered species is an important focus for a conservation genetics study (Frankham *et al.* 2002). Studies using codominant microsatellites use either F_{ST} (Wright 1921, 1951) or R_{ST} (Slatkin 1995), to estimate and interpret the genetic structuring of populations. F_{ST} and R_{ST} take into

account different mutational models of microsatellites, producing different values depending on the microsatellite data.

F_{ST} values of 0 represent a group of populations with no genetic variation (perfectly mixed), while a value of 1 indicates high levels of genetic variation. Actual values are rarely 0 or 1 so require some interpretation to be biologically meaningful. Using allozymes, it has been suggested that values of 0-0.05 indicate little variation, 0.05-0.15 moderate variation, 0.15-0.25 great variation and levels above that very great variation (Wright 1951), but should ultimately be assessed case-by-case (Balloux & Lugon-Moulin 2002). Understanding the level of genetic structure can be used in setting conservation guidelines. For example, populations of Minorcan cork oak were found to have high levels of genetic differentiation ($F_{ST} = 0.104$). Therefore, each population was genetically unique and conservation management was established based on this information (Lorenzo *et al.* 2009). The F-statistic, F_{IS} (fixation index or inbreeding coefficient) can use microsatellite data to estimate levels of inbreeding within a population by comparing levels of allele fixation (Aegisdottir *et al.* 2009). Allele fixation can be used to indicate the occurrence or likelihood of inbreeding (Lopez *et al.* 2009).

Gene flow in plant populations is determined by seed dispersal and pollen movement. Gene flow is important in maintaining population connectivity and transferring genetic variation, but high levels impede local adaptation and speciation (Balloux & Lugon-Moulin 2002). Levels of gene flow between populations can be estimated indirectly using F_{ST} values. Because F_{ST} can describe the distribution of allele frequencies among populations, comparative values can be used to infer levels of gene flow. However, this is not a direct measure as there are many underlying evolutionary forces which can produce similar patterns (Ouborg *et al.* 1999). An extension of this indirect method utilises chloroplast DNA, as it is maternally inherited and therefore can differentiate between pollen and seed mediated gene flow. Direct measures of gene flow can be estimated in some cases using highly variable microsatellite loci (Aegisdottir *et al.* 2009).

1.5.3 Informing translocation regimes

The ideal solution for protection of endangered species is to preserve the habitats they exist in and conserve the plants *in situ*. This strategy protects both the rare taxa and the entire ecosystem, but in many cases this is not an option (Templeton 1991). Often, *ex situ* conservation is required to prevent extinction, through collection of a sample of the remaining population (typically seeds). Ultimately, the goal of collection of endangered taxa will be reestablishment of historic populations, enriching wild populations or establishment of new populations in risk free environments that the taxon was previously not known from (Brown & Briggs 1991).

Selecting an appropriate sample size for the translocated population is important to ensure a significant amount of the genetic variation is reflected in the new population, but is best determined in a case-by-case situation (Sinclair & Hobbs 2009). Biological traits of the study species can affect the genetic variation in the newly translocated population and ultimately, genetic diversity of the new population will depend on genetic variation and sample size of the founders, mating system of the species, and population growth rate (Hufford & Mazer 2003). For example, to conserve the critically endangered Corrigin grevillia (*Grevillea scapigera*), of which only 47 plants remain, a sample of 10 plants were chosen for micropropogation and translocation that represented 87% of the species total genetic variation detected by RAPDs. Studies of the translocated populations 6-7 years later showed that while many plants had established, AFLP studies found that despite almost complete outcrossing, due to differential reproductive success the overall heterozygosity had decreased 20% and inbreeding had increased 22% (Krauss *et al.* 2002). A review of plant translocations by Godefroid *et al.* (2011) identified factors that seem to be important for successful translocations and suggested improvements to methodology for improved survival of translocated populations.

Translocations to areas known to be ecologically suitable for a species but which it is currently absent from is potentially beneficial in decreasing isolation between populations Additionally, translocation of genetic material (i.e. seed or pollen) can be made into populations that are genetically depauperate in order to increase diversity (Godefroid *et al.* 2011) However, genetic swamping is a possible outcome due to genotypes with differential fitness coming together or a numerical advantage of one

genotype (Hufford & Mazer 2003). Using cpDNA this was detected in North American *Phragmites australis*, where an introduced European variant has started to dominate via a cryptic invasion, leading to a loss in genetic diversity of the native genotype (Saltonstall 2002). Secondary contact between two genetically diverged populations can result in reduced fitness of the hybrids due to outbreeding depression. Hybrid breakdown can be caused by the mating of two locally adapted genotypes that produces offspring comparatively unfit for either environment. Molecular information should be used to determine the best individuals for relocation and test for genetic swamping after translocation (Hufford & Mazer 2003).

1.6 Future directions of conservation genetics

1.6.1 Traits relevant to adaptation and survival – neutral variation evaluation

Conservation genetics relies on the use of neutral variation as an indicator of the amount and pattern of potentially adaptive or detrimental variation. If this association is considered to be good, then it allows for a prediction on the amount of detrimental variation that could lower fitness and the amount of adaptive variation that is present to deal with future evolutionary challenges (Hedrick 2001). Neutral genetic variation in populations is considered to be controlled by interactions of mutation, drift and migration, however the ability of small-scale neutral marker surveys to identify variation of conservation relevance is controversial (Kohn *et al.* 2006). Empirical studies have found that there is a significant positive correlation between population size, fitness and genetic variation (Leimu *et al.* 2006) but there is also evidence that neutral variation does not correlate with detrimental variation or adaptive quantitative traits (Merila & Crnokrak 2001). While neutral markers are useful for describing population structure and tracking the effects of gene flow, inbreeding and genetic drift, there is uncertainty between the correlation of neutral variation in understanding adaptive variation and the fitness of populations in their environment. This area needs resolution if conservation genetic practices continue to ascertain adaptive variation based on neutral sequence variation (Holderegger *et al.* 2006).

Several new techniques have been developed in order to directly study traits that are of adaptive significance. For example, the development of the Variance component

quantitative trait loci (QTL) technique enables creation of QTL maps without the need for backcrosses and inbred lines, which are difficult in wild endangered populations. This technique has the potential to study traits of adaptive significance in less controlled situations commonly encountered by a conservation geneticist (Besnier *et al.* 2010). Alternatively, Expressed Sequence Tag – Simple Sequence Repeats (EST-SSRs) are microsatellites that occur within expressed gene regions. Because they are within genic regions, EST-SSRs are subject to selection and thus adaptive variation may be identified (Bouck & Vision 2007; Ellis & Burke 2007). Selecting a candidate gene believed to be of evolutionary significance could be used in order to determine adaptive genetic diversity within and between populations. This approach requires the selection of a gene that is predicted to influence fitness, and identifying variation in the species that may account for future selective advantages (Piertney & Webster 2010). The use of adaptive markers in tandem with neutral markers could guide conservation efforts by providing a more detailed picture of a species evolutionary history and possible future (Kramer & Havens 2009).

1.6.2 Genomics

The ability of next-generation sequence technology to produce large amounts of sequence will affect research in all areas of ecology and evolutionary biology (Hudson 2008). Current application of genomic tools to the study of organisms or populations in their natural environments relies on the transfer of techniques, approaches and data developed in model organisms. A genomics approach to studying conservation questions requires a transition from model to non-model organisms, from the lab to natural environment and development of resources and tools necessary for detailed study at the individual or population level (Ouborg *et al.* 2010b).

Conservation genomics, or more broadly ecogenomics, will enable a detailed insight into the levels of functionally important variation within and between populations, as well as the association between adaptive and neutral variation (Kohn *et al.* 2006). Conservation genetics will continue to reconstruct important population and demographic processes, but genomics has the potential to expand research into understanding the mechanisms behind the reduced fitness of endangered and small populations (Ouborg *et al.* 2010a) including the effects of selection and the environment on sequence variation and gene expression. Figure 2 (Ouborg *et al.* 2010b) illustrates

the additional layers of understanding genomics could provide to conservation. Current conservation genetic studies focus on the patterns of genetic diversity in endangered populations, the magnitude of information from an organisms genome that can now be accessed may allow a deeper understanding on the evolutionary processes in rare populations that can lead to extinction (Avise 2010).

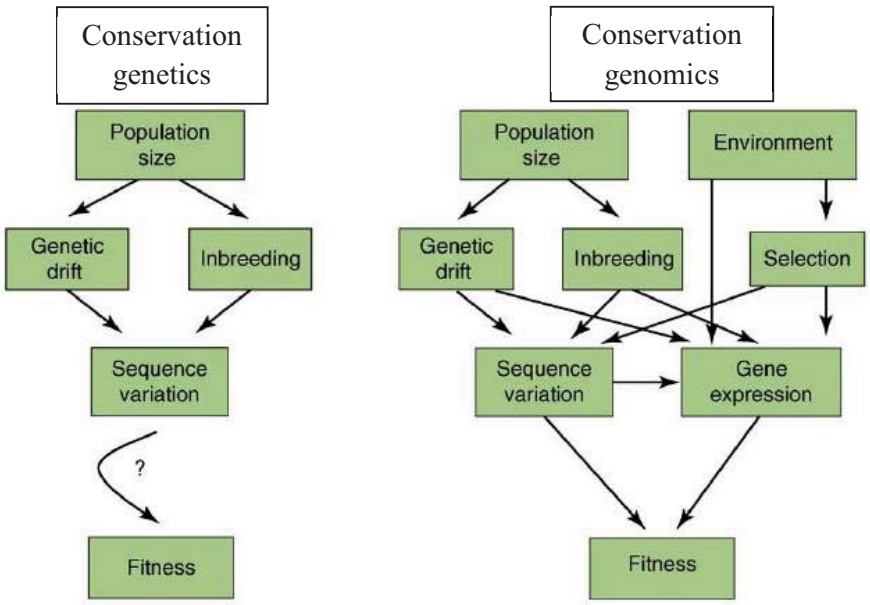

Figure 1.6: Schematic representation of the conservation genetics approach compared to the conservation genomics approach from Ouborg et al. (2010b). While conservation genetics measures the effect of population size on inbreeding and genetic drift using neutral sequence variation to estimate fitness, conservation genomics has the potential to measure the effect of the environment and selection as well as differential gene expression in order to give an overall understanding of a population's fitness.

1.7 Focus of this research

Dactylanthus taylorii is currently listed as Nationally Vulnerable on the New Zealand threatened plant list. Populations are found throughout the North Island , but many are geographically isolated in fragments of remnant indigenous forest. Habitat fragmentation, population isolation, browsing of inflorescences by introduced mammals and a reduction in population size and range of the short-tail bat would have contributed to the vulnerable status as well as affecting the genetic variation and structure of the species. This goal of this study is to use a conservation genetics approach in order to assess the genetic diversity and structure of populations to aid conservation management.

Objective 1: Design novel *D. taylorii* microsatellite markers using genomic sequence.

Objective 2: Utilise the microsatellite markers in order to assess genetic variation and structure of populations throughout the geographic range of *D. taylorii.*

Objective 3: Use the information from the genetic study in order to determine conservation implications to aid future management plans for the species.

1.8 References

Aegisdottir HH, Kuss P, Stocklin J (2009) Isolated populations of a rare alpine plant show high genetic diversity and considerable population differentiation. *Annals of Botany* **104**, 1313-1322.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution* **16**, 613-622.

Avise JC (2010) Perspective: conservation genetics enters the genomics era. *Conservation Genetics* **11**, 665-669.

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

Besnier F, Le Rouzic A, Alvarez-Castro JM (2010) Applying QTL analysis to conservation genetics. *Conservation Genetics* **11**, 399-408.

Bouck A, Vision T (2007) The molecular ecologist's guide to expressed sequence tags. *Molecular Ecology* **16**, 907-924.

Brown AHD, Briggs JD (1991) Sampling strategies for Genetic Variation in *ex situ* Collections of Endangered Plant Species. In: *Genetics and Conservation of Rare Plants* (eds. DA F, Holsinger K). Oxford University Press, New York.

Callmander KW, Schatz GE, Lowry PP (2005) IUCN Red List assessment and the Global Strategy for Plant Conservation: taxonomists must act now. *Taxon* **54**, 1047-1050.

de Lange P, Norton D, Heenan P*, et al.* (2004) Threatened and uncommon plants of New Zealand. *New Zealand Journal of Botany* **42**, 45-76.

de Lange PJ, Norton DA, Courtney SP*, et al.* (2009) Threatened and uncommon plants of New Zealand (2008 revision). *New Zealand Journal of Botany* **47**, 61-96.

DeSalle R, Amato G (2004) The expansion of conservation genetics. *Nat Rev Genet* **5**, 702-712.

Diamond JM (1989) Overview of recent extinctions. In: *In Conservation for the Twenty-first Century* (ed. Western DaP, M.C), pp. 37-41. Oxford University Press.

Dopson S, de Lange PJ, Ogle C (1999) *The conservation requirements of New Zealand's nationally threatened vascular plants* Biodiversity Recovery Unit, Department of Conservation.

Dutra D, Kane ME, Adams CR, Richardson L (2009) Reproductive biology of Cyrtopodium punctatum in situ: implications for conservation of an endangered Florida orchid. *Plant Species Biology* **24**, 92-103.

Dzerefos CM, Shackleton CM, Witkowski ETF (1999) Sustainable utilization of woodrose-producing mistletoes (loranthaceae) in South Africa. *Economic Botany* **53**, 439-447.

Ecroyd CE (1995) Dactylanthus taylorii recovery plan. In: *Threatened species recovery plan Series 16*. Department of Conservation (NZ), Wellington.

Ecroyd CE (1996a) Dactylanthus and bats: The link between two unique endangered New Zealand species and the role of the community in their survival. In: *Nature Conservation 4: The Role of Networks* (eds. DA. S, JL. C, EM M), pp. 78-87. Surrey Beatty & Sons, Chipping Norton NSW.

Ecroyd CE (1996b) The ecology of Dactylanthus taylorii and threats to its survival. *New Zealand Journal of Ecology* **20**, 81-100.

Ecroyd CE, Franich RA, Kroese HW, Steward D (1995) Volatile constituents of Dactylanthus taylorii flower nectar in relation to flower pollination and browsing by animals. *Phytochemistry* **40**.

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. *Heredity* **99**, 125-132.

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.

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

Falk D, Holsinger K (1991) Genetics and Conservation of Rare Plants. Oxford University Press, New York.

Ferreira SM (2005) Individual-level management trade-offs for populations of Dactylanthus taylorii (Balanophoraceae). *New Zealand Journal of Botany* **43**, 415-424.

Fleishman E, Launer AE, Switky KR*, et al.* (2001) Rules and exceptions in conservation genetics: genetic assessment of the endangered plant Cordylanthus palmatus and its implications for management planning. *Biological Conservation* **98**, 45-53.

Frankel OH, Soule ME (1981) *Conservation and Evolution.* Cambridge University Press, Cambridge

Frankham R (1997) Do island populations have less genetic variation than mainland populations? *Heredity* **78**, 311-327.

Frankham R, Ballou J, Briscoe D (2002) *Introduction to Conservation Genetics* Cambridge University Press.

Franks SJ (2010) Genetics, Evolution, and Conservation of Island Plants. *Journal of Plant Biology* **53**, 1-9.

Furches MS, Wallace LE, Helenurm K (2009) High genetic divergence characterizes populations of the endemic plant Lithophragma maximum (Saxifragaceae) on San Clemente Island. *Conservation Genetics* **10**, 115-126.

Gao LZ, Schaal BA, Zhang CH, Jia JZ, Dong YS (2002) Assessment of population genetic structure in common wild rice Oryza rufipogon Griff. using microsatellite and allozyme markers. *Theoretical and Applied Genetics* **106**, 173-180.

Gaston KJ (1994a) Biodiversity - Measurement. *Progress in Physical Geography* **18**, 565-574.

Gaston KJ (1994b) Measuring geographic range sizes. *Ecography* **17**, 198-205.

Given D (1981) *Rare and endangered plants of New Zealand* Reed, Wellington.

Godefroid S, Piazza C, Rossi G*, et al.* (2011) How successful are plant species reintroductions? *Biological Conservation* **144**, 672-682.

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, Murawski DA, Loveless MD (1991) *Correlations between species traits and allozyme diversity - implications for Conservation Biology* Oxford Univ Press, New York.

Hedrick PW (2001) Conservation genetics: where are we now? *Trends in Ecology & Evolution* **16**, 629-636.

Hereford J (2010) Does selfing or outcrossing promote local adaptation? *American Journal of Botany* **97**, 298-302.

Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology* **21**, 797-807.
Holsinger K, Gottlieb L (1991) Conservation of Rare and Endangered Plants: Principles and Prospects. In: *Genetics and Conservation of Rare Plants* (eds. DA F, Holsinger K). Oxford University Press, New York.
Holzapfel S (2001) Studies of the New Zealand Root-Parasite *Dactylanthus taylorii* (Balanophoraceae). *Englera* **22**, 3-176.
Holzapfel S, Faville MZ, Gemmill CEC (2002) Genetic variation of the endangered holoparasite *Dactylanthus taylorii* (Balanophoraceae) in New Zealand. *Journal of Biogeography* **29**, 663-676.
Honnay O, Adriaens D, Coart E, Jacquemyn H, Roldan-Ruiz I (2007) Genetic diversity within and between remnant populations of the endangered calcareous grassland plant Globularia bisnagarica L. *Conservation Genetics* **8**, 293-303.
Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular Ecology Resources* **8**, 3-17.
Huenneke L (1991) Ecological implications of genetic variation in plant populations. In: *Genetics and conservation of rare plants* (ed. Falk DA HK). Oxford, University Press, New York.
Hufford KM, Mazer SJ (2003) Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution* **18**, 147-155.
Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M (2008) Ecological consequences of genetic diversity. *Ecology Letters* **11**, 609-623.
Jamieson IG (2007) Has the debate over genetics and extinction of island endemics truly been resolved? *Animal Conservation* **10**, 139-144.
Kelly D, Sullivan JJ (2010) Life histories, dispersal, invasions, and global change: progress and prospects in New Zealand ecology, 1989-2029. *New Zealand Journal of Ecology* **34**, 207-217.
Kohn MH, Murphy WJ, Ostrander EA, Wayne RK (2006) Genomics and conservation genetics. *Trends in Ecology & Evolution* **21**, 629-637.
Kramer AT, Havens K (2009) Plant conservation genetics in a changing world. *Trends in Plant Science* **14**, 599-607.
Krauss SL, Dixon B, Dixon KW (2002) Rapid genetic decline in a translocated population of the endangered plant Grevillea scapigera. *Conservation Biology* **16**, 986-994.
Kuijt J (1969) *The biology of parasitic flowering plants. University of California Press, Berkeley*.
La Cock GD, Holzapfel S, King D, Singers N (2005) *Dactylanthus taylorii recovery plan, 2004–14* Department of Conservation (NZ), Wellington
Leimu R, Mutikainen P, Koricheva J, Fischer M (2006) How general are positive relationships between plant population size, fitness and genetic variation? *Journal of Ecology* **94**, 942-952.
Lloyd BD (2003) Intraspecific Phylogeny of the New Zealand short-tailed bat Mystacina tuberculata inferred from multiple mitochondrial gene sequences. *Systematic Biology* **52**, 460-476.
Lopez S, Rousset F, Shaw FH, Shaw RG, Ronce O (2009) Joint Effects of Inbreeding and Local Adaptation on the Evolution of Genetic Load after Fragmentation. *Conservation Biology* **23**, 1618-1627.
Lorenzo Z, Burgarella C, de Heredia UL*, et al.* (2009) Relevance of genetics for conservation policies: the case of Minorcan cork oaks. *Ann Bot* **104**, 1069-1076.

Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* **15**, 65-95.
Macphail MK, Mildenhall DC (1980) Dactylanthus taylorii - in Northwest Nelson, New Zealand. *New Zealand Journal of Botany* **18**, 149-152.
Manville V, Segschneider B, Newton E*, et al.* (2009) Environmental impact of the 1.8 ka Taupo eruption, New Zealand: Landscape responses to a large-scale explosive rhyolite eruption. *Sedimentary Geology* **220**, 318-336.
McGlone MS (1989) The Polynesian settlement of New Zealand in relation to environmental and biotic changes. *New Zealand Journal of Ecology* **12**, 115-129.
McGlone MS, Topping WW (1983) Late Quarternary vegetation, Tongariro region, Central North Island, New Zealand. *New Zealand Journal of Botany* **21**, 53-76.
Meine C, Soule M, Noss RF (2006) "A mission-driven discipline": the growth of conservation biology. *Conservation Biology* **20**, 631-651.
Merila J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology* **14**, 892-903.
Mildenhall DC, Alloway BV (2008) A widespread ca. 1.1 Ma TVZ silicic tephra preserved near Wellington, New Zealand: Implications for regional reconstruction of mid-Pleistocene vegetation. *Quaternary International* **178**, 167-182.
Moore LB (1940) The structure and life-history of the root parasite Dactylanthus taylori Hook. f. *New Zealand Journal of Science and Technology, Section B* **21**, 206-224.
Neel MC, Cummings MP (2003) Effectiveness of conservation targets in capturing genetic diversity. *Conservation Biology* **17**, 219-229.
Nickrent DL, Der JP, Anderson FE (2005) Discovery of the photosynthetic relatives of the "Maltese mushroom" Cynomorium. *Bmc Evolutionary Biology* **5**, 11.
Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**, 1143-1155.
O'Donnell CFJ, Christie JE, Hitchmough RA, Lloyd B, Parsons S (2010) The conservation status of New Zealand bats, 2009. *New Zealand Journal of Zoology* **37**, 297-311.
Ouborg NJ, Angeloni F, Vergeer P (2010a) An essay on the necessity and feasibility of conservation genomics. *Conservation Genetics* **11**, 643-653.
Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma R, Hedrick PW (2010b) Conservation genetics in transition to conservation genomics. *Trends in Genetics* **26**, 177-187.
Ouborg NJ, Piquot Y, Van Groenendael JM (1999) Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology* **87**, 551-568.
Ouborg NJ, Vergeer P, Mix C (2006) The rough edges of the conservation genetics paradigm for plants. *Journal of Ecology* **94**, 1233-1248.
Pertoldi C, Bijlsma R, Loeschcke V (2007) Conservation genetics in a globally changing environment: present problems, paradoxes and future challenges. *Biodiversity and Conservation* **16**, 4147-4163.
Piertney SB, Webster LMI (2010) Characterising functionally important and ecologically meaningful genetic diversity using a candidate gene approach. *Genetica* **138**, 419-432.
Porcher E, Kelly JK, Cheptou PO*, et al.* (2009) The genetic consequences of fluctuating inbreeding depression and the evolution of plant selfing rates. *Journal of Evolutionary Biology* **22**, 708-717.

Qiao Q, Zhang CQ, Milne RI (2010) Population genetics and breeding system of Tupistra pingbianensis (Liliaceae), a naturally rare plant endemic to SW China. *Journal of Systematics and Evolution* **48**, 47-57.

Redford KH, Richter BD (1999) Conservation of biodiversity in a world of use. *Conservation Biology* **13**, 1246-1256.

Rieseberg LH, Swensen SM (1996) Conservation Genetics of Endangered Island Plants. In: *Conservation Genetics: Case Histories from Nature* (eds. Avise JC, Hamrick JL). Chapman and Hall, New York.

Riley L, McGlaughlin ME, Helenurm K (2010) Genetic diversity following demographic recovery in the insular endemic plant Galium catalinense subspecies acrispum. *Conservation Genetics* **11**, 2015-2025.

Sakai AK, Wagner WL, Mehrhoff LA (2002) Patterns of endangerment in the Hawaiian flora. *Systematic Biology* **51**, 276-302.

Saltonstall K (2002) Cryptic invasion by a non-native genotype of the common reed, Phragmites australis, into North America. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 2445-2449.

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

Semagn K, Bjornstad A, Ndjiondjop MN (2006) An overview of molecular marker methods for plants. *African Journal of Biotechnology* **5**, 2540-2568.

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.

Sinclair EA, Hobbs RJ (2009) Sample Size Effects on Estimates of Population Genetic Structure: Implications for Ecological Restoration. *Restoration Ecology* **17**, 837-844.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 1463-1463.

Smith SA, Donoghue MJ (2008) Rates of molecular evolution are linked to life history in flowering plants. *Science* **322**, 86-89.

Stockwell CA, Hendry AP, Kinnison MT (2003) Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution* **18**, 94-101.

Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology & Evolution* **15**, 199-203.

Takahashi Y, Takahashi H, Maki M (2011) Comparison of genetic variation and differentiation using microsatellite markers among three rare threatened and one widespread toad lily species of Tricyrtis section Flavae (Convallariaceae) in Japan. *Plant Species Biology* **26**, 13-23.

Templeton R (1991) Off-site Breeding of Animals and Implications for Plant Conservation Strategies. In: *Genetics and Conservation of Rare Plants* (eds. Falk DA, KE H). Oxford University Press, New York.

Thorsen MJ, Dickinson KJM, Seddon PJ (2009) Seed dispersal systems in the New Zealand flora. *Perspectives in Plant Ecology Evolution and Systematics* **11**, 285-309.

van Bunnik A, Pollock A, Somerset E*, et al.* (2007) Environment New Zealand 2007. In: *Environment New Zealand 2007*. Landcare Research, Ministry for the Environment.

Wang JL (2005) Estimation of effective population sizes from data on genetic markers. *Philosophical Transactions of the Royal Society B-Biological Sciences* **360**, 1395-1409.

Wayne RK, Morin PA (2004) Conservation Genetics in the New Molecular Age. *Frontiers in Ecology and the Environment* **2**, 89-97.
Wright S (1921) Systems of mating. *Genetics* **6**, 111-178.
Wright S (1951) The Genetical Structure of Populations. *Annals of Eugenics* **15**, 323-354.
Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution* **11**, 413-418.

2 Microsatellite marker development for the endangered root parasite *Dactylanthus taylorii* using high-throughput sequencing

2.1 Abstract

Microsatellites are highly variable markers that are used frequently in conservation genetic studies. However, markers must be designed *de novo* and typical microsatellite isolation protocols can be time consuming, expensive and technically difficult, as well as require cloning to sequence identified repeat loci. Here, high-throughput sequencing of genomic *Dactylanthus taylorii* DNA was utilised to produce 19Mb of sequence reads that are on average 313bp long. MSATCOMMANDER (ver 0.8.1) identified 2730 sequences containing repeats, from which 753 primer pairs were designed for potentially amplifiable microsatellite loci. Using repeat number, repeat motif and product size as variable parameters to reduce the total number of primer pairs to screen, 72 primer pairs were selected for screening. From these, 10 were found to be variable and consistently amplifiable for use in determining the genetic structure and variation of *D. taylorii* populations. This method would be easily replicated in many species, and will be aided by developments in next-generation sequencing technology and bioinformatic tools.

2.2 Introduction

Microsatellite loci are tandem repeats of 1-6 nucleotides that are commonly used as genetic markers in population and conservation genetic studies (Selkoe & Toonen 2006). Features of microsatellites that make them highly informative markers for evolutionary studies are a high mutation rate, codominant Mendelian inheritance and the ability for polymerase chain reaction (PCR) amplification by primers in conserved flanking regions (Sunnucks 2000). However, these benefits are often outweighed by the difficulties associated with microsatellite isolation as compared to other anonymous marker systems, such as amplified fragment length polymorphisms (AFLPs) (Squirrell *et al.* 2003), and poor transferability of markers between systems, meaning new species require *de novo* marker design (Sunnucks 2000).

Traditional methods of isolating microsatellites can be technically limiting as well as time consuming and expensive. A typical isolation protocol involves the construction of an enriched library by digesting genomic DNA and ligation of fragments to specially designed adaptors, hybridisation of fragments to 5'- biotinylated oligo repeats followed by PCR amplification of products, and transformation of products into vectors for cloning. Finally, clones are sequenced, microsatellite repeats are identified and primers are designed in flanking regions (Techen *et al.* 2010). The materials required to perform these steps are often specialised and are not easily available, especially to those without access to an established molecular biology lab, including many ecologists and conservation biologists (Selkoe & Toonen 2006). The time required from DNA extraction to primer design could be a week to three months, depending on researcher expertise and study organism, while the overall cost can vary greatly (Squirrell *et al.* 2003).

The requirement of cloning is also restrictive in countries where cultural and biosecurity issues restrict transformation of indigenous organisms or creation of new organisms. In New Zealand, permission to clone must be obtained from the Environmental Risk Management Authority (ERMA), and a permit granted after a laboratory inspection. Permission to clone native organisms requires additional Maori consultation and approval (Satterfield & Roberts 2008). The process of obtaining cloning permission adds to the financial cost and time requirement to design microsatellites.

The advent of high-throughput genomic sequencing provides an alternative method to identify potentially useful microsatellite loci. Using this approach, genomic DNA is randomly sequenced and then screened for microsatellite repeats from which primers can be designed. This method has been used successfully to find microsatellites in blue duck (Abdelkrim *et al.* 2009), extinct moa species from ancient DNA (Allentoft *et al.* 2009), coral (Baums *et al.* 2009) and a bulrush (Csencsics *et al.* 2010). Additionally, the sequencing of an enriched library was used to isolate microsatellites in a fungus, a wasp, and a wasp-parasite (Santana *et al.* 2009). In all cases, between $1/4^{th}$ and $1/16^{th}$ of a 454 (Roche) Picotiterplate run was used to shotgun sequence DNA, which was then analysed bioinformatically to identify repeats and design primers in the flanking regions. This procedure essentially removes all the difficult lab bench work of the traditional microsatellite isolation protocol, aside from DNA extraction, and replaces it with computer-based work. In publications that have used this method, the time and cost per marker required to isolate microsatellites was significantly less than traditional methods in all cases.

In this study, genomic DNA of *Dactylanthus taylorii* (Balanophoraceae) was sequenced using 454 GS FLX in order to isolate microsatellites. *D. taylorii* is a root holoparasite endemic to New Zealand, and is of conservation significance due to loss of plants from habitat destruction and possum browsing (Ecroyd 1996b). Fragmented populations are found throughout the North Island, and populations are assumed to have low levels of gene flow as the endangered short-tail bat (*Mystacina tuberculata*) is the only known native pollinator (Ecroyd 1996a). Parasitic plants from the Santalales are known to have high rates of molecular evolution, including base pair substitutions or insertion/deletion events (Nickrent & Starr 1994). This suggests that microsatellite variability could be high, but little is known about the nature of microsatellite evolution in Balanophoraceae as there have been few molecular studies on the family. Microsatellite loci identified in this study will be used to determine the genetic variation and structure of *D. taylorii* over the entire species range.

2.3 Methods

2.3.1 DNA extraction and next-generation sequencing

Genomic DNA was extracted from inflorescence tissue using a modified hexa-decetylammonium bromide (CTAB) protocol (Doyle & Doyle 1987a) with an initial sucrose, TRIS, EDTA (STE) wash to remove polysaccharides. DNA was combined from three population samples; Mangaweka, Pureora and Mahia. The mixture had a concentration of 248ng/uL and a 260/280 of 2.05, quantified using spectrophotometry on a Nanodrop 2000 (Thermo-Scientific). Genomic DNA was then subjected to shotgun sequencing on the 454 GS-FLX (Roche) at Otago University following manufacturer's protocols and quality control steps. A detailed description of this technique can be found in Marguilies *et al*. (2005). Fragments were sequenced on $1/16^{th}$ of an LR670 plate.

2.3.2 Microsatellite discovery from 454 sequence and primer design

Total sequence data were converted to a single text file and then submitted to MSATCOMMANDER ver. 0.8.1 (Faircloth 2008), a program written in Python that has a simple to use graphical user interface. It searches an input file (text or FASTA format) for microsatellites with the user able to select the repeat class (mononucleotide to hexanucleotide) and repeat length; the results output can be viewed in Excel. Primers were designed in flanking regions for sequences with microsatellite repeats automatically using PRIMER3 (Rozen & Skaletsky 2000), the custom tag option was used to place an 18bp tag (CACGACGTTGTAAAACGA) at the 5' end of all forward primers. This is an adapter for a fluorescent primer used for genotyping PCR products. Primers were designed using default criteria; product size within 100-450 bp, optimal GC content 50%, optimal melting 60°C and a GC clamp. MSATCOMMANDER has been used to screen microsatellites and design primers in several studies using 454 sequence output (Abdelkrim *et al.* 2009; Allentoft *et al.* 2009; Baums *et al.* 2009; Csencsics *et al.* 2010). The *D. taylorii* sequence reads were searched for di-, tri-, and tetranucleotides with a minimum repeat size of six, four and four repeats, respectively.

The outputs from MSATCOMMANDER and PRIMER3 were combined and first sorted into di-, tri- and tetranucleotide classes and then by repeat motif (nucleotide composition of

repeat). Microsatellite reads with primers designed to amplify loci will be referred to here-after as "primer designed locus" (PDL), and indicate potentially useful microsatellite markers. The percentage of PDL found within the total number of repeats identified was then compared based on repeat class, and on a finer scale, repeat motif within each class. The product sizes of the PDL were also compared to determine average length of products, and to see if there was a bias towards small (less than 150bp) or large product sizes (greater than 350bp).

2.3.3 Screening for polymorphic loci

Based on criteria related to sequence quality and repeat structure, 72 primer pairs were screened using individuals from seven populations to identify consistently amplifiable and variable loci. Each reaction contained a final concentration of 1X New England Biolabs (NEB) PCR Buffer, 50 μM of each dNTP, 20 nM of forward primer, 450 nM reverse primer, 450 nM dye-labelled M13 tail primer, 1.6 μL 5M (8 uM) betaine, 0.5 units of *Taq* polymerase (NEB), and approximately 50 ng of genomic DNA. The M13 tail primers were labelled with either VIC or FAM fluorescent dyes (Applied Biosystems). PCR amplification cycles were as follows: 95°C for 3 min, then 29 cycles of 95°C for 30 s, 52°C for 40 s and 72°C for 40 s, followed by a hold at 72°C for 20 min. Genotyping was performed on an ABI DNA Analyser 3730 at the Massey Genome Service, Palmerston North, New Zealand. Alleles were sized using CASS (Symonds & Lloyd 2004) and scored manually using GENEMAPPER (version 4.0, Applied Biosystems).

Loci that were consistently amplifiable and variable were then genotyped in 43 individuals from five geographically distant populations (Little Barrier Island, Taahau, Maude Track, Mahia, Mangaweka, Figure 2.1). Basic summary statistics for each locus (number of individuals amplified, number of alleles, size range, observed and expected heterozygosity) were calculated in GENALEX 6 (Peakall & Smouse 2006). Tests for deviations from Hardy-Weinberg equilibrium (HWE) were performed using GENEPOP 4.0.10 (Rousset 2008) with significant departures recorded when $p<0.05$. ARLEQUIN 3.5.1.2 (Excoffier & Lischer 2010) was used to check for linkage disequilibrium between all pairs of markers and significant deviations when $p<0.05$.

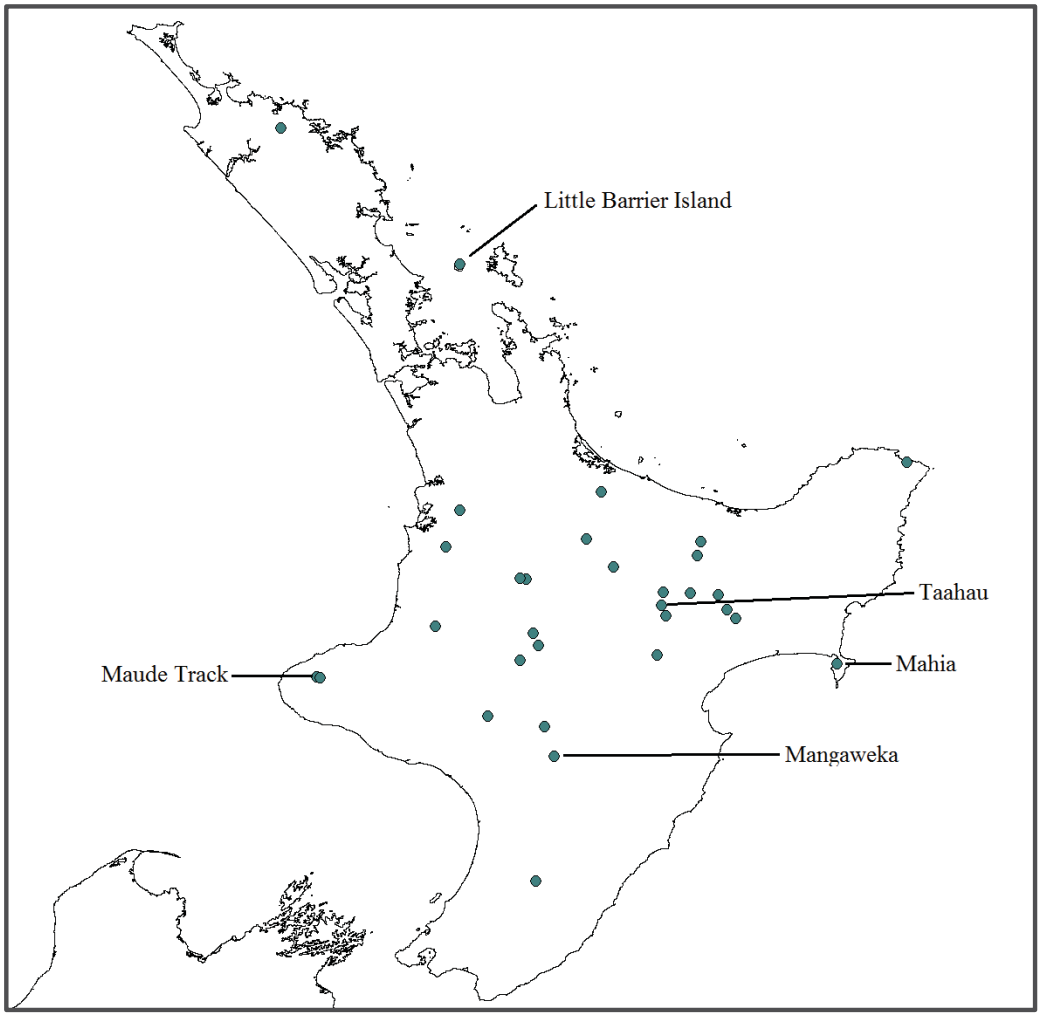

Figure 2.1: *Dactylanthus taylorii* individuals from labelled populations were used to screen microsatellite loci. Other points represent populations that will be screened with suitable loci.

2.4 Results

2.4.1 Microsatellite discovery and primer design

A total of 61,721 individual sequences were obtained, of which 61,709 passed the Newbler (Margulies *et al.* 2005) quality run. This gave a total of 19Mb of sequence, with an average read length of 313bp. MSATCOMMANDER detected a total of 2730 sequences containing repeats (4% of the quality reads) and 4044 total microsatellite repeats; the higher number of total repeats is due to the occurrence compound repeats within sequence reads. A total of 1547 dinucleotide, 2247 trinucleotide and 250 tetranucleotide repeats were identified. Reverse complement repeats were grouped together (e.g., AC and GT, ATT and AAT) to produce three dinucleotide (AT, AC, AG), 10 trinucleotide (e.g. AAG, ATT) and eight tetranucleotide (e.g. ATTT, AGAT) repeat motifs (Figure 2.2).

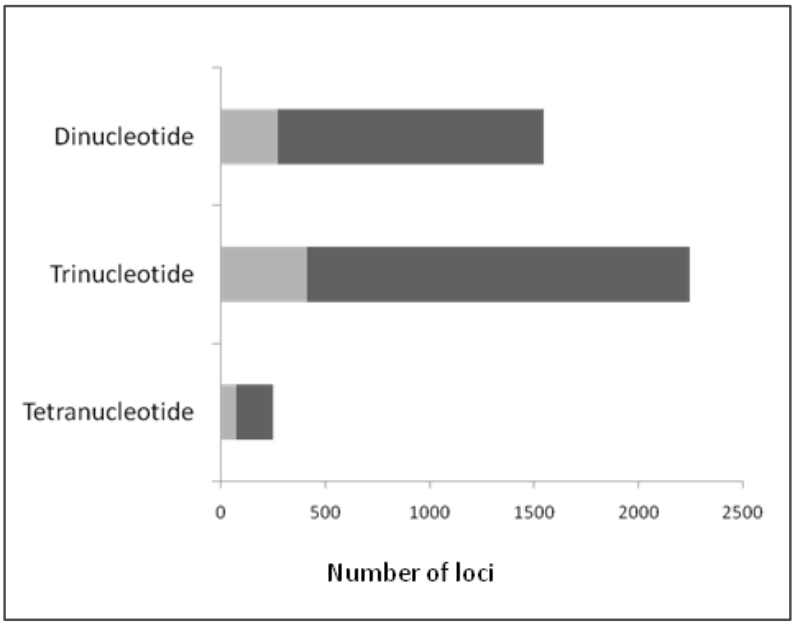

Figure 2.2: Total number of microsatellite repeat loci identified by MSATCOMMANDER (total bar), and the number of loci for which PRIMER3 was able to design primers for (PDL, light grey).

Using default criteria, PRIMER3 designed primers for 753 of the 4044 repeats (19%), representing 1.2% of the total quality sequences. Comparing between repeat classes, trinucleotide were the highest number of repeats identified with 2247 and PDL with 410 (18%). Of the 1547 dinucleotide repeats 270 (17%) were PDL, and tetra nucleotides had the fewest number of identified repeats at 250, but 73 of those repeats had PDL (29%) (Figure 2.2). The repeats also had an A/T bias, with only 23 of the 68 (33%) nucleotide monomers being a G or C.

There was an uneven distribution of repeat motifs in each of the three repeat classes (Figure 3.3). In the dinucleotide class, the repeat motif AT represented 75% of the repeats, with AC and AG accounting for 12% and 13%, respectively. However, the percentage of PDL for each motif type was reasonably even with 17% for AT, 21% for AC and 15% for AG (Figure 2.3A). Four of the six monomers (66%) of the dinucleotide repeats were an A or T. In the trinucleotide class, the AAG motif represented 65%, but only 5% of those sequences had PDL, and the AAT motif represented 24% of all repeats. Other motifs represented less than 5% of the total trinucleotide repeats. This is more than likely due to the large number of reads that consisted entirely of CTT repeats, with no appropriate regions for primer design. In contrast, many of the other trinucleotide motifs had a high percentage of PDL (Figure 2.3B). There was no A/T bias in the trinucleotide repeat monomers, with 50% GC content. The tetranucleotide class was not dominated by one motif. ATTT repeats made up 40%, AGAT made up 31% and AATT made up 14%, with 5 other repeat types accounting for 15% of the tetranucleotide repeats. Primers were designed for 25% or greater of all motifs except for AGAT for which only 3% had PDL. The tetranucleotide repeats had the greatest AT bias, with 75% of the bases being an A or T (Figure 2.3C).

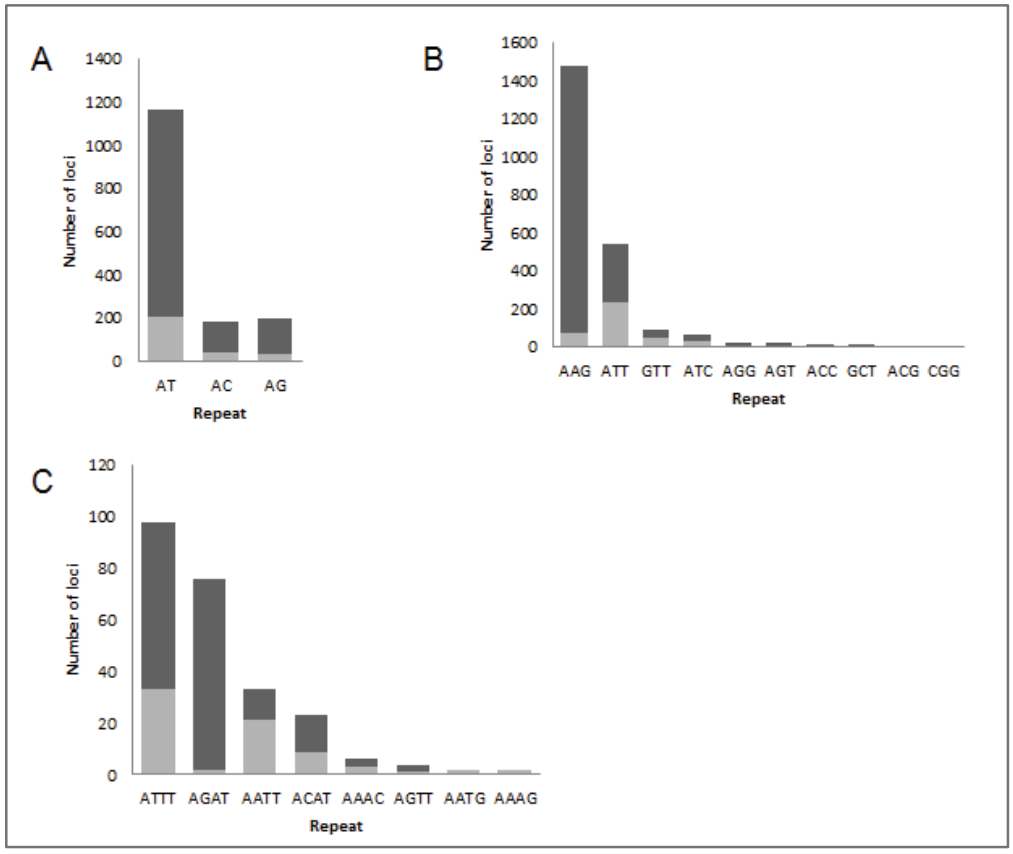

Figure 2.3: Total counts of microsatellite loci (total bar) and sub-set of loci with primers designed by PRIMER3 (light grey), compared by repeat motif type. A, dinucleotides; B, Trinucleotides and C;Tetranucleotides.

While it is useful to have so many PDL to choose from to screen and test for marker suitability, performing PCR and subsequent analysis on 753 primer pairs would be prohibitively expensive and time consuming. To reduce the total pool of markers to select and screen, the number of tandem repeats within each PDL were compared (Figure 2.4). The relative numbers of repeats were distributed fairly evenly between the three classes, with short repeats occurring more frequently than long repeats. The shortest dinucleotide repeat size was six repeats, which had twice as many PDL as the repeat size of seven (120 compared to 60). The shortest trinucleotide and tetranucleotide repeat size, four repeats, had approximately five times as many PDL as the five repeat size classes. This result is expected because long repeats will have less flanking sequence surrounding them in which suitable primers can be designed, and this

effect would be more pronounced in tri- and tetranucleotides as they have more bases per repeat.

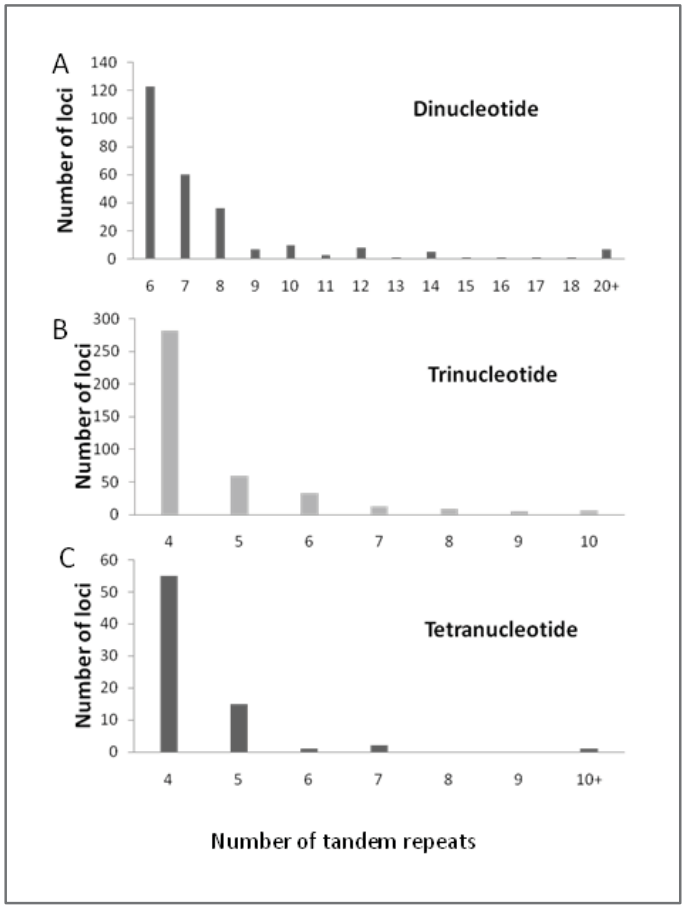

Figure 2.4: Repeat number frequency of loci with primers designed by PRIMER3 (PDL). A) dinucleotide, B) trinucleotide, C) tetranucleotide. Note that the Y-axis is not the same for all graphs.

PRIMER3 designed primers that would amplify a region between 100-450bp, which is the minimum and maximum size considered to be useful for microsatellite markers. The average product size was 197 bp, and the maximum was 449bp and most of the PDL were between 160 and 279bp in length, with only a few products greater than 280bp (Figure 2.4). Because the average sequence read size was 313bp, it is likely that there

were not many reads with microsatellites greater than 280bp that also had suitable flanking regions for primer design.

.

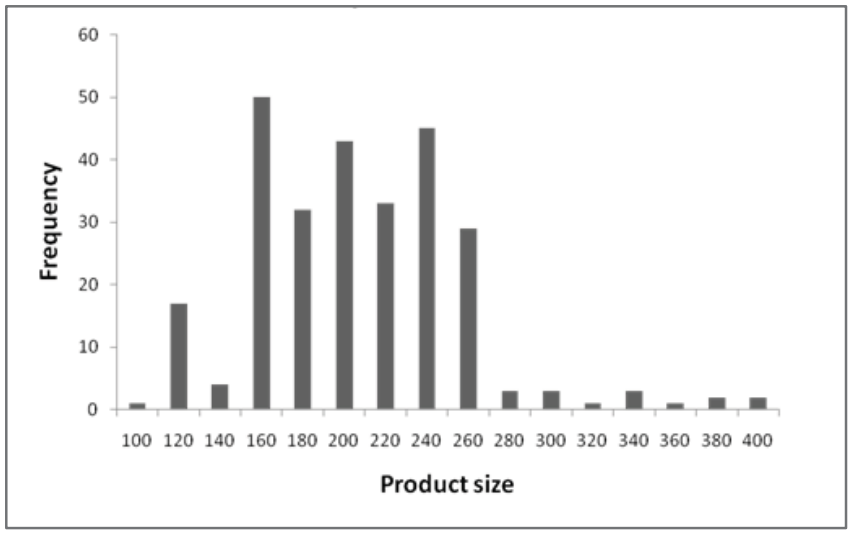

Figure 2.5: PCR product sizes of total PDL. Bin size is 20bp

2.4.2 Microsatellite screening

Of the 72 primer pairs chosen for screening (named DATA 1-72), 51 produced a product of the expected size, and 15 of these were variable. From these 15 loci, 10 were found to be variable and consistently amplified in the screen of 43 individuals. Amplification success ranged between 70-100%. Markers DATA20 and DATA65 were found to be absent from one population each after re-genotyping to check for null alleles. The number of observed alleles per locus ranged from four to sixteen, with an average of 9.7. Observed heterozygosity (H_O) ranged from 0.182-0.634 and expected heterozygosity (H_E) ranged from 0.408-0.729. DATA4, DATA20, DATA21, DATA32 and DATA62 showed significant deviations from Hardy-Weinberg equilibrium across all populations due to heterozygote deficiency. No significant evidence of linkage disequilibrium was found between any pair of loci.

Table 2.1: Characteristics of 10 microsatellite loci from 43 individuals of *Dactylanthus taylorii*, developed from 454 pyrosequencing. The forward sequence contains a 5' fluorescent tag used during PCR amplification. N, number of individuals amplified; A, total number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity.

Locus	Primer sequence (5'-3')	Repeat motif	N	A	Allele size range	H_O	H_E
DATA4	F: CACGACGTTGTAAAACGACCTGTTTAGAACCGGCCACTAC	TA(6)	39	9	306-318	0.199	0.449
	R: AAATCGGGCTCACAATGGC						
DATA19	F: CACGACGTTGTAAAACGACCAGATGGACGGCTCAGATTAG	AC(10)	39	10	150-178	0.634	0.538
	R: AACGGATATGTGTTGTGTATGAAGA						
DATA20	F: CACGACGTTGTAAAACGACTCATAAATCAGCACGCATCC	AT(10)	34	9	183-206	0.186	0.501
	R: CGTTCAACACACATCGCATA						
DATA21	F: CACGACGTTGTAAAACGACTGACTCGAAAATAAATTGGGAAA	AT(10)	43	12	240-266	0.479	0.561
	R: CCATCTTACGTTAAAATACAACAGAGA						
DATA32	F: CACGACGTTGTAAAACGACGCCATTTGTAGGTTTCCAAGA	AC(8)	41	15	209-265	0.422	0.581
	R: GGGTATAATGGGAAGGGTTGA						
DATA40	F: CACGACGTTGTAAAACGACCCAACTTCATTGTCCACCTTC	CT(8)	43	8	258-268	0.434	0.466
	R: TTCACCCCCTTTAGTGAACTC						
DATA41	F: CACGACGTTGTAAAACGACTCGATAGGTAGCATCAAGGTG	AC(8)	43	8	227-241	0.467	0.458
	R: CCATCCCTAAAATAACTTCAAACC						
DATA46	F: CACGACGTTGTAAAACGACGGTACCAAGCCCGGATAAAT	GTT(8)	43	4	250-260	0.285	0.346
	R: ACCCCATGGTGGCTTACATA						
DATA62	F: CACGACGTTGTAAAACGACACATGGCGCTTATTGCAGG	GT(15)	42	16	117-153	0.182	0.729
	R: GGTTCCATAACACACGCACG						
DATA65	F: CACGACGTTGTAAAACGACGGCCCGTGTATAGGACGC	AAT(8)	30	6	254-269	0.479	0.408
	R: TGGTACTGTGGCTGGTACG						

2.5 Discussion

High-throughput genomic sequencing has revolutionized our ability to generate genomic tools for non-model organisms. For example, from the endangered endemic root holoparasite, *Dactylanthus taylorii*, 61,721 Roche 454 Genome Sequencer FLX shotgun sequence reads were generated and from these, a total of 4044 microsatellite repeat loci were identified. From the 2730 reads in which microsatellite loci were detected, 753 primer pairs were designed, 72 were screened, and 10 were found to amplify and be suitably variable; these can be used as genetic markers for evolutionary and population genetic studies (Balloux & Lugon-Moulin 2002).

The time taken from DNA extraction to primer design was eight weeks. This time period included six weeks for sequencing, which was delayed by a shortage of necessary sequencing reagents and the relative position in the queue. Therefore, the overall time spent isolating microsatellites was approximately two weeks, including DNA extraction, sequence acquisition and bioinformatics using MSATCOMMANDER. This method was able to greatly reduce the amount of time spent obtaining microsatellites, compared to more traditional methods. During the six week wait for results, I was able to extract DNA from all 31 *Dactylanthus taylorii* population samples to be screened with the microsatellite loci.

A shotgun sequencing method is also less technically demanding than a typical isolation protocol. The bench work for this method is limited to DNA extraction and eventual screening of markers, rather than a series of steps requiring specific consumables (e.g., adapters, biotinylated beads) and researcher expertise (Techen *et al.* 2010). Using NEWBLER (Roche) to convert the sequences into FASTA format requires some understanding of using command line, but is detailed in the 454 pipeline information (Roche). MSATCOMMANDER has an easily interpreted interface and can be used with all operating systems, and was run on a desktop PC (Intel Core 2 Duo, 2.67GHz, 2.96GB RAM) with little difficulty.

Enrichment libraries typically target specific repeat classes, repeat motifs and repeat sizes using probes selected *a priori*. Limiting the subset of microsatellites could limit the number of potentially useful microsatellite loci identified, as well as affect the diversity of markers and overall variability detected (Bachtrog *et al.* 2000). This

potential source of bias does not occur with shotgun sequencing, as the microsatellites detected should be a random sampling from the genome. However, in this study not all repeat motifs were isolated from the sequence reads (e.g., CG/GC) with only a few motifs represented in the tri- and tetranucleotide class, but the cause of this is unclear.

There was also an AT bias within the repeats, with an average of 67% of the nucleotide composition of the microsatellites being an A or T. This may reflect the AT bias observed in many plants species (Kawabe & Miyashita 2003). The trinucleotide class had an even mixture of AT:CG (50:50) base composition, which may reflect that repeats of three or six (or multiples thereof) are more likely to be found in genes, which typically have a higher GC content (Carels *et al.* 1998), and where they are less likely to cause a frame-shift mutation (Buschiazzo & Gemmell 2006). Microsatellites isolated using probes in other Santalales do not have a high A/T bias (Bottin *et al.* 2005; Lhuillier *et al.* 2006). However, there does appear to be an A/T bias of repeat motifs that were isolated from 454 sequences (Abdelkrim *et al.* 2009; Baums *et al.* 2009; Csencsics *et al.* 2010). This is an effect that may warrant further investigation for future users of next-generation sequencing to isolate microsatellites.

While the MSATCOMMANDER microsatellite mining parameters used were relaxed in order to find the maximum number of repeat loci, by adjusting the settings to perform a more stringent search the overall number of potential primers to screen would be reduced. However, there are no "ideal" microsatellite parameters that are universally adhered to in order to create markers that will best detect variability. Locus variability is associated with its mutation rate, and mutation rate can be affected by the repeat class, repeat size and whether the repeat is perfect or compound (Symonds & Lloyd 2003; Bhargava & Fuentes 2010). Therefore it is up to individual researchers to determine what repeat parameters are best for their study, but by having more to choose from better informed decisions can be made. As a typical conservation genetics study only uses 10-20 microsatellite markers, stricter parameter settings would reduce the overall number of potential primers to screen. The 10 found here represent 13% of the 72 primers pairs screened from a total pool of 753 PDL, however it is unlikely that 13% of the 753 will be found to be suitable markers as the loci initially chosen for screening were selected based on parameters expected to perform best as microsatellite markers. This strike rate of usable markers is lower than that found in other studies using similar methods (Abdelkrim *et al.* 2009; Csencsics *et al.* 2010).

While this approach is ultimately intended for studying the genetic structure of *D. taylorii* populations across the species distribution, it also provides millions of other base pairs that can be utilised to understand genome evolution in this parasitic plant. As *D. taylorii* currently has only two sequences deposited in Genbank (AY957447, maturase R gene, and AY957443, 18S ribosomal RNA), the data generated here will be an invaluable genetic resource for future comparative genomic studies. For example, a batch BLAST of all 67,000 sequences will be performed to determine if any chloroplast genes remain in this holoparasite (Krause 2008), and to look for genes such as flowering (MADS-box) genes that are potentially useful for resolving the Balanophoraceae phylogeny.

The ten novel polymorphic microsatellite markers obtained will be genotyped in *D. taylorii* populations across its entire geographical range. Results from the five populations used here to assess marker suitability suggest that *Dactylanthus taylorii* has a high allelic richness with an average of 9.7 alleles per locus and a low allelic diversity (H_E=0.504) compared to averages found in a meta-analysis of microsatellite variation in plants (Nybom 2004). Five markers were found to deviate from Hardy-Weinberg equilibrium due to homozygote excess which may indicate inbreeding, reduced diversity due to population bottleneck or other potential causes of reduced genetic variation (Ellstrand & Elam 1993; Ouborg *et al.* 2006). The markers developed here will be used to study the genetic structure and variation of *D. taylorii* in order to inform conservation management decisions.

2.6 Conclusion

Shotgun sequencing of *Dactylanthus taylorii* genomic DNA and subsequent bioinformatical analysis has identified 753 potential microsatellite loci and from a screen of 72 primers a total of 10 usable markers were found. This method offers a fast and simple alternative to traditional microsatellite isolation, while generating ten times as many potential markers as a typical probe-based protocol. Next generation sequencing also eliminates the requirement of cloning for designing microsatellites, an issue which is particularly relevant to researchers of New Zealand native species. This method is currently only limited by the extraction of high quality DNA from an organism, and will continue to improve as developments are made in the areas of high-throughput sequencing and bioinformatic tools for analysis.

2.7 References

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

Allentoft ME, Schuster SC, Holdaway RN*, et al.* (2009) Identification of microsatellites from an extinct moa species using high-throughput (454) sequence data. *Biotechniques* **46**, 195-+.

Bachtrog D, Agis M, Imhof M, Schlotterer C (2000) Microsatellite Variability Differs Between Dinucleotide Repeat Motifs--Evidence from Drosophila melanogaster. *Molecular Biology and Evolution* **17**, 1277.

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

Baums IB, Devlin-Durante K, Brown L, Pinzon JH (2009) Nine novel, polymorphic microsatellite markers for the study of threatened Caribbean acroporid corals. *Molecular Ecology Resources* **9**, 1155-1158.

Bhargava A, Fuentes FF (2010) Mutational Dynamics of Microsatellites. *Molecular Biotechnology* **44**, 250-266.

Bottin L, Vaillant A, Sire P, Cardi C, Bouvet JM (2005) Isolation and characterization of microsatellite loci in Santalum austrocaledonicum, Santalaceae. *Molecular Ecology Notes* **5**, 800-802.

Buschiazzo E, Gemmell NJ (2006) The rise, fall and renaissance of microsatellites in eukaryotic genomes. *BioEssays* **28**, 1040-1050.

Carels N, Hatey P, Jabbari K, Bernardi G (1998) Compositional properties of homologous coding sequences from plants. *Journal of Molecular Evolution* **46**, 45-53.

Csencsics D, Brodbeck S, Holderegger R (2010) Cost-Effective, Species-Specific Microsatellite Development for the Endangered Dwarf Bulrush (Typha minima) Using Next-Generation Sequencing Technology. *Journal of Heredity* **101**, 789-793.

Doyle JJ, Doyle JL (1987a) A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochemistry Bulletin* **19**, 11-15.

Ecroyd CE (1996a) Dactylanthus and bats: The link between two unique endangered New Zealand species and the role of the community in their survival. In: *Nature Conservation 4: The Role of Networks* (eds. DA. S, JL. C, EM M), pp. 78-87. Surrey Beatty & Sons, Chipping Norton NSW.

Ecroyd CE (1996b) The ecology of Dactylanthus taylorii and threats to its survival. *New Zealand Journal of Ecology* **20**, 81-100.

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.

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.

Faircloth BC (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* **8**, 92-94.

Kawabe A, Miyashita NT (2003) Patterns of codon usage bias in three dicot and four monocot plant species. *Genes & Genetic Systems* **78**, 343-352.

Krause K (2008) From chloroplasts to "cryptic" plastids: evolution of plastid genomes in parasitic plants. *Current Genetics* **54**, 111-121.

Lhuillier E, Vaillant A, Butaud JF, Bouvet JM (2006) Isolation and characterization of microsatellite loci in Santalum insulare, Santalaceae. *Molecular Ecology Notes* **6**, 653-655.

Margulies M, Egholm M, Altman WE*, et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380.

Nickrent DL, Starr EM (1994) High Rates of Nucleotide Substitution in Nuclear Small-Subunit (18S) rDNA from Holoparasitic Flowering Plants. *Journal of Molecular Evolution* **39**, 62-70.

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**, 1143-1155.

Ouborg NJ, Vergeer P, Mix C (2006) The rough edges of the conservation genetics paradigm for plants. *Journal of Ecology* **94**, 1233-1248.

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

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

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

Santana QC, Coetzee MPA, Steenkamp ET*, et al.* (2009) Microsatellite discovery by deep sequencing of enriched genomic libraries. *Biotechniques* **46**, 217-223.

Satterfield T, Roberts M (2008) Incommensurate risks and the regulator's dilemma: considering culture in the governance of genetically modified organisms. *New Genetics and Society* **27**, 201-216.

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

Squirrell J, Hollingsworth PM, Woodhead M*, et al.* (2003) How much effort is required to isolate nuclear microsatellites from plants? *Molecular Ecology* **12**, 1339-1348.

Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology & Evolution* **15**, 199-203.

Symonds VV, Lloyd AM (2003) An analysis of microsatellite loci in Arabidopsis thaliana: Mutational dynamics and application. *Genetics* **165**, 1475-1488.

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

Techen N, Arias RS, Glynn NC*, et al.* (2010) Optimized construction of microsatellite-enriched libraries. *Molecular Ecology Resources* **10**, 508-515.

3 Conservation genetics of the endemic root holoparasite *Dactylanthus taylorii*

3.1 Abstract

Species with isolated, small populations often show low levels of genetic diversity which can limit their ability to respond to evolutionary pressures and consequently lead to extinction of the species. *Dactylanthus taylorii* (Balanophoraceae) is a root holoparasite endemic to New Zealand which is threatened due to habitat fragmentation and browsing of inflorescences by introduced herbivores. Ten polymorphic microsatellite marker loci were used to estimate the genetic variation and structure of 31 populations to determine the genetic impact of isolation and reduction in population size. Across all populations a high number of alleles was detected (185), although many of them were private (56). On average a population had a low number of alleles per loci (N_A=3.155), a low expected heterozygosity (H_E=0.468) and a high F_{IS} (0.240). A large number of the populations were found to deviate from Hardy-Weinberg equilibrium due to homozygote excess. A high degree of genetic differentiation between populations was found (F_{ST}=0.316) and differentiation correlated significantly with geographical distance. Populations were found to fall into two, three or eight genetic clusters which largely reflected geographical distribution of populations. These results suggest that populations are genetically depauperate and may be at risk to extinction due to genetic factors. The genetic clusters found could be useful as a guide for translocation of genetic material via seed or pollen in order to increase genetic diversity in populations.

3.2 Introduction

Conservation of intraspecific genetic variation is one of the three focal areas for biodiversity conservation (Schatz 2009). Many rare plant species exist only in small or isolated populations in fragmented habitats (Young *et al.* 1996) and suffer from the consequences of small population size, such as genetic drift, inbreeding and reduced gene flow due to isolation (Ouborg *et al.* 2006). These factors will decrease genetic variation within a population and increase variation among populations, reducing the evolutionary potential of the population or species (Ellstrand & Elam 1993). Knowledge of the genetic variation that exists within and between populations of endangered species and the genetic structure of endangered populations (i.e. levels of isolation and gene flow between populations) is an important resource for conservation regimes (Wayne & Morin 2004).

The most recent assessment of New Zealand's indigenous plant taxa estimates that 898 species, or 38% of the native flora, are threatened in some way (de Lange *et al.* 2009). New Zealand's geological history, including early separation from Gondwana and its present isolation, volcanism, climate change, fault-line movements and mountain building (McGlone *et al.* 2001), has led to the evolution of a flora with a high level of endemism (~85%) (de Lange *et al.* 2006). As such, the New Zealand flora is regarded as one of the world's biodiversity hotspots (Kelly & Sullivan 2010). The arrival of humans in New Zealand (~1300AD) and subsequent habitat loss or degradation and introduction of pests (including animals and plants) have led to a decline in the distribution of many plant taxa (Craig *et al.* 2000). Conservation practices in New Zealand initially focused on protecting habitats, as ecosystems are often more feasible to protect and demographic factors are a more immediate risk to endangered species, but have more recently utilised genetic information in order to guide conservation management of endangered species, (e.g *Clianthus* sp. (Song *et al.* 2008), *Hebe speciosa* (Armstrong & De Lange 2005), *Metrosideros excelsa* (Young et al. 2001), *Olearia gardneri* (Barnaud & Houliston 2010)).

Dactylanthus taylorii Hook.f. (pua o te reinga or dactylanthus) is a fully parasitic angiosperm endemic to New Zealand (Holzapfel 2001) and is ranked as a species of highest conservation importance by the New Zealand Department of Conservation (Dopson *et al.* 1999). *Dactylanthus* is a monotypic genus and represents the most

southern member of an otherwise tropical and subtropical family of obligate root parasites, the Balanophoraceae (Nickrent *et al.* 2005). This family is currently unplaced within the angiosperm phylogeny (APG 2009) and is comprised of 17 genera and 44 species that are all characterised by greatly reduced vegetative and floral organs. *D. taylorii* obtains nutrients and water from a host tree root by attaching via a subterranean haustorial connection, which causes hypertrophic growth of the root creating a fluted disc-like "woodrose"(Holzapfel 2001). *D. taylorii* is dioecious, and an inflorescence develops within the tuber body below the soil breaching the surface when it is in the final stages of development and preparing to open (Moore 1940). The endemic short-tailed bat (*Mystacina tuberculata*) is the only known natural pollinator (Ecroyd 1996a). A native seed disperser is unknown, although it is generally assumed that gravity is the main mode of dispersal (Ecroyd 1996b). Population sizes range from fewer than ten plants to several thousand and are found throughout the North Island (La Cock *et al.* 2005). Plants are generally found in forest margins or areas of regrowth as this is the habitat for most of the 30 supposed host tree species (Holzapfel 2001).

A decline of *Dactylanthus taylorii* in its distribution was first observed in the 1980s (Given 1981). Habitat destruction, collection of woodrose specimens and browsing of inflorescences by the introduced brush-tailed possum (*Trichosurus vulpecula)* are believed to be the main agents of decline (Holzapfel 2001). A decline in the distribution of the short-tail bat also means many populations have no natural pollination and only set seed by hand pollination or via pollination by rats and mice (Ferreira 2005). *D. taylorii* is listed as "Nationally Vulnerable" in the most recent list of New Zealand threatened plants (de Lange *et al.* 2009) ,a positive change from the 2004 classification of "Serious Decline" (de Lange *et al.* 2004).

The genetic structure of *Dactylanthus taylorii* populations has previously been studied by Holzapfel *et al*. (2002) using five randomly amplified polymorphic DNAs (RAPDs) markers on 17 populations located throughout the North Island. Each population was found to be unique, with individuals sampled from the same population clustering together in UPGMA (Unweighted Pair Group Method with Arithmetic Means) dendrograms. No significant relationship was found between genetic variation and geographic distance, but a dendrogram of Nei's genetic distance displayed a major separation between eastern and western populations, though statistical support for this division was not shown. The cause of this demarcation was hypothesised to be volcanic

activity as the split falls in the same areas as the Taupo Volcanic Zone (TVZ) (Manville *et al.* 2009).

The discovery of new geographic outlier populations and the limited resolution provided by the RAPD study means that current understanding is not sufficient to guide management (La Cock *et al.* 2005). Microsatellite loci differ from RAPDs in that they are co-dominant, rather than dominant molecular markers which allows for calculations of potentially more meaningful genetic parameters for conservation, such as observed and expected heterozygosity and F_{IS} (which can be used to estimate inbreeding). The high mutation rate and selective neutrality of microsatellites also enables inferences on gene flow and population structure to be made (Selkoe & Toonen 2006; Semagn *et al.* 2006).

In this study polymorphic microsatellite loci were used to analyse populations of *Dactylanthus taylorii* over its geographical range to: (1) assess the genetic variation within the species, and (2) determine how this diversity is distributed to evaluate the genetic structure of populations. Our results indicate the presence of considerably more structure than previously recognised and a low level of genetic variation at the population level. The evolutionary and biogeographical history of *D. taylorii* is discussed, as well as how the results of this study could be used to guide future conservation management of the species.

3.3 Materials and methods

3.3.1 Sampling

A total of 241 individuals were sampled from 31 populations of *Dactylanthus taylorii* throughout its geographical range (Figure 3.1 and Table 3.1). Most of the sampling was done between March-May 2010, apart from the Oropi population which was collected in February 2011, and DNA samples from Te Araroa, Pukerimu, Mt Pirongia and Mangaweka which were obtained from a previous study (Holzapfel *et al.* 2002); samples were collected under Department of Conservation (DOC) permits BP-27170-FLO, NO-27249-FLO and TT-27186-FLO. Inflorescences were randomly sampled throughout a population to try to maximise representation across the distribution of the population and to minimise sampling from the same plant. As many populations have a high male: female plant ratio, young male buds were preferentially sampled to reduce the effect on the breeding population. Sample size ranged from one to 15 individuals (average 7.8 individuals per population). The initial aim was to sample a minimum of ten individuals from each population, but in some cases this was not possible due to small population size or poor flowering in the season collections were undertaken. Population size was estimated by DOC staff involved with *D. taylorii* management.

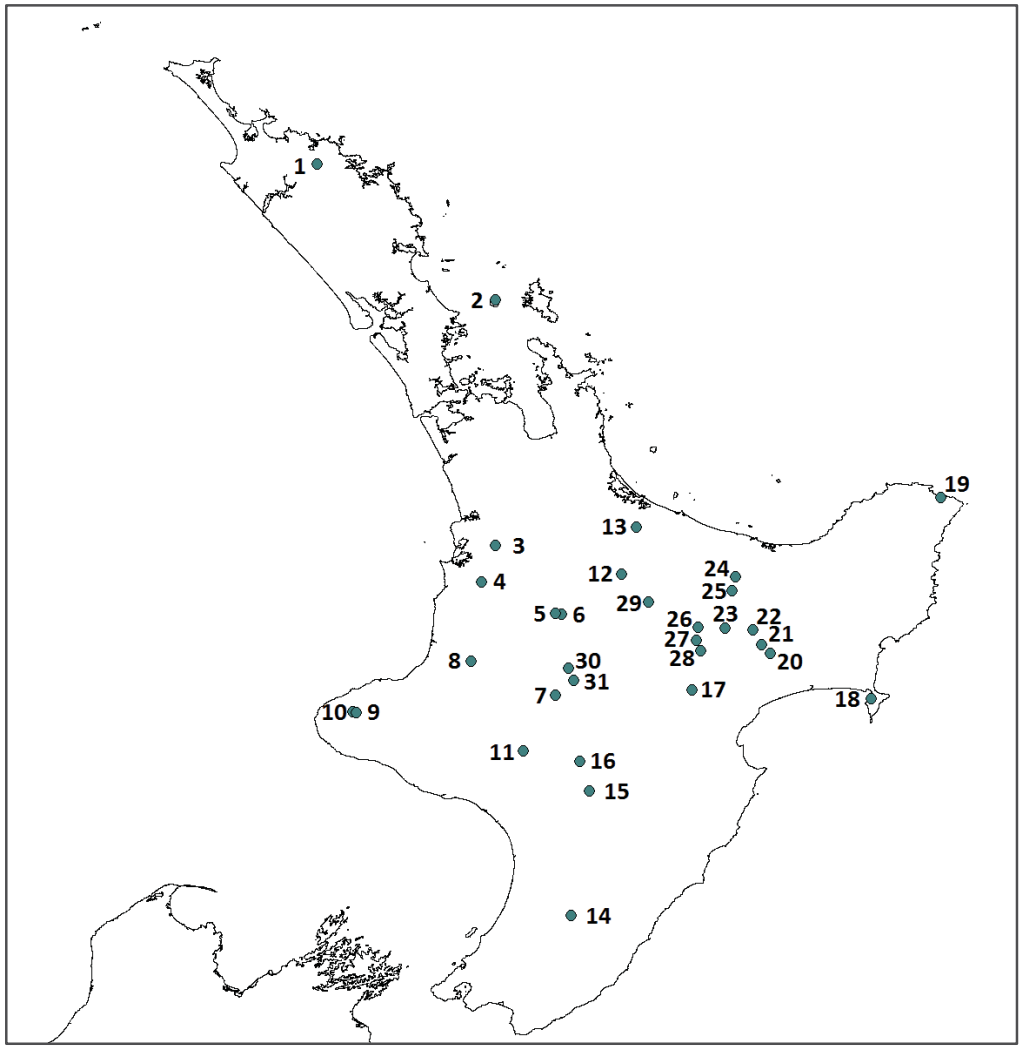

Figure 3.1: Distribution map of the populations of *Dactylanthus taylorii* sampled. Dots represent populations as indicated in Table 1. Map was created using DIVA-GIS.

Table 3.1: Population information and genetic diversity estimates for 31 populations of *D. taylorii*

	Population name	# indv	Pop size	%P	P_A	N_A	N_E	H_O	H_E	F_{IS}
Pop 1	Puketi	2	25	50	1	1.500	1.460	0.050	0.238	0.889*
Pop 2	Little Barrier Island	6	150	90	2	2.800	2.094	0.307	0.437	0.401*
Pop 3	Mt Pirongia	11	300	90	9	3.100	2.135	0.291	0.428	0.369*
Pop 4	Geanges	5	16	100	2	3.400	2.811	0.430	0.611	0.427*
Pop 5	Waipapa Lodge	2	10	30	0	1.300	1.233	0.300	0.175	-0.25
Pop 6	Pikiariki	13	500	100	7	6.400	4.006	0.404	0.703	0.461*
Pop 7	Tongariro Forest	6	150	90	2	3.300	2.552	0.428	0.533	0.41*
Pop 8	Waitaanga Conservancy	7	20	90	3	2.700	1.944	0.288	0.460	0.449*
Pop 9	Maude	10	50	100	2	4.900	3.477	0.478	0.644	0.319*
Pop 10	Mangorei	9	100	80	2	4.300	3.193	0.368	0.565	0.407*
Pop 11	Parapara	4	20	70	1	2.000	1.683	0.342	0.338	0.137
Pop 12	Pukerimu	9	100	80	1	3.100	2.108	0.302	0.422	0.349*
Pop 13	Oropi	15	150	50	1	1.700	1.295	0.114	0.207	0.479*
Pop 14	Mt Bruce	1	3	10	2	0.900	0.900	0.100	0.050	NA
Pop 15	Mangaweka	7	20	100	2	2.800	2.091	0.381	0.441	0.221*
Pop 16	Ngarakehu	6	10	90	1	2.900	2.182	0.485	0.449	0.022
Pop 17	Wharangi	4	30	90	2	2.500	2.084	0.567	0.466	-0.071
Pop 18	Mahia	9	30	100	2	3.700	2.471	0.305	0.532	0.514*
Pop 19	Te Araroa	7	1000	100	3	3.900	3.039	0.460	0.589	0.294*
Pop 20	Onepoto	6	94	100	0	2.400	1.917	0.277	0.446	0.481*
Pop 21	Te Wera	8	40	100	0	4.000	2.464	0.446	0.560	0.272*
Pop 22	Waihirere	8	100	100	1	3.500	2.086	0.471	0.476	0.078
Pop 23	Te Whaiti Exclosure	3	100	70	0	2.300	2.008	0.333	0.368	0.298*
Pop 24	Waikokopu	10	60	100	2	4.500	3.373	0.536	0.666	0.248*
Pop 25	Kopuriki	3	35	90	0	2.400	2.102	0.767	0.468	-0.484
Pop 26	Pukeroa	11	60	100	0	3.400	2.577	0.551	0.569	0.091
Pop 27	Taahau	12	190	100	1	3.200	2.145	0.413	0.463	0.155*
Pop 28	Waione	12	180	90	0	3.400	2.254	0.384	0.468	0.224*
Pop 29	Te Kopia	12	170	100	6	4.800	3.185	0.557	0.650	0.188*
Pop 30	Kakaramea	11	300	100	0	3.300	2.285	0.499	0.525	0.112
Pop 31	100 Acre Bush	12	50	100	1	3.400	2.338	0.639	0.548	-0.119
	Mean	7.77		85.8	1.81	3.155	2.306	0.396	0.468	0.240

 # individuals, total samples from a population; Pop size, approximate population size; PA, number of private alleles; %P, percentage of polymorphic loci; NA, number of alleles per locus; NE, number of effective alleles per locus; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, fixation index score, * denotes significant deviations from HWE due to homozygote excess.

3.3.2 DNA extraction and genotyping

Tissue was collected fresh and kept cold until processing in the lab. The inflorescence head and bracts were removed from the peduncle, washed in 70% ethanol and the tissue was cut into approximately 100mg discs, and frozen in liquid nitrogen. Samples were then stored at -20ºC until DNA extraction.

Approximately 100mg of homogenised tissue was used in each extraction. In order to reduce polysaccharides which could potentially interfere with PCR, a modified hexadecyltrimethyl ammonium bromide (CTAB) extraction (Doyle & Doyle 1987b) was used that included an initial sucrose, Tris and EDTA (STE) wash (Shepherd & McLay 2011) and room temperature isopropanol to precipitate the DNA. DNA quantity and purity was assessed using spectrophotometry (Nanodrop, Thermo Scientific). Samples with a DNA concentration less than 50ng/µl or 260/280 nm ratio less than 1.70 were re-extracted.

All 241 individuals were screened at 10 microsatellite loci (Table 2) developed for *Dactylanthus taylorii* using Roche 454 sequencing as in Chapter 2 (McLay *et al. in prep*). Microsatellite loci were amplified individually for all DNA samples in 10 µL PCR reactions using a modified tailing method (Schuelke 2000; Boutin-Ganache *et al.* 2001). Each reaction contained a final concentration of 1X New England Biolabs (NEB) PCR Buffer, 50 µM of each dNTP, 20 nM of forward primer, 450 nM reverse primer, 450 nM dye-labelled M13 tail primer, 1.6 µL 5M (8 µM) betaine, 0.5 units of *Taq* polymerase (NEB), and approximately 80 ng of genomic DNA. The M13 tailed primers were labelled with either VIC or FAM fluorescent dyes (Applied Biosystems). PCR amplification consisted of: 95°C for 3 minutes, then 29 cycles of 95°C for 30 seconds, 52°C for 40 seconds and 72°C for 40 seconds, followed by a final extension at 72°C for 20 minutes.

In most cases, a marker labelled with FAM was pooled with a VIC labelled marker for the same individual in equal volumes. One µl of the pooled PCR was added to 8 µL Hi-Dye (Applied Biosystems) and 1 µL CASS ladder (Symonds & Lloyd 2004). Genotyping was performed on an ABI 3730 DNA Analyser (Massey Genome Service, Palmerston North). Alleles were sized using GENEMAPPER (version 4.0, Applied

Biosystems) and recorded manually. The same positive controls were included on all genotyping plates and were checked for consistency.

3.3.3 Assessing genetic variation

Each microsatellite locus was assessed for the total number of alleles (A) and the observed and expected heterozygosity (H_O and H_E). Genetic diversity for each population was assessed across all loci using the observed number of alleles (N_A), the effective number of alleles (N_E), expected heterozygosity (H_E), observed heterozygosity (H_O), F_{IS} and the percentage of polymorphic loci (%P). Weir and Cockerham's (1984) F-statistics (F_{IS} and F_{ST}) were calculated using FSTAT 2.9.3.2 (Goudet 1995) and GENALEX 6 (Peakall & Smouse 2006) was used to calculate N_A, N_E, A, %P, H_E and H_O. ARLEQUIN 3.5 (Excoffier & Lischer 2010) was used to test all markers for pairwise linkage disequilibrium. Tests for deviations from Hardy-Weinberg equilibrium were performed in FSTAT (significant values when $p<0.05$). The effect of population size and sample size on molecular diversity (H_E, N_A, F_{IS}) was assessed using regression analysis in Excel.

3.3.4 Determining genetic structure and differentiation

Isolation by distance was assessed using a Mantel test between pairwise F_{ST} values obtained from FSTAT, specifically $F_{ST}/(1-F_{ST})$ following Rousset (1997), and the natural log of the geographic distance between two populations in GENEPOP (Raymond & Rousset 1995; Rousset 2008) . AMOVA (analysis of molecular variance) was calculated in ARLEQUIN in order to determine how genetic variation was partitioned among and between populations for all 31 populations, and between major geographical regions identified using the most supported K-value identified by STRUCTURE results.

To characterise the distribution of genetic variation, STRUCTURE (version 2.2) (Pritchard *et al.* 2000) was used to identify the most likely number of genetic clusters within the *Dactylanthus taylorii* dataset. STRUCTURE uses a Bayesian method based on multi-locus genotype data to construct "ancestral" populations and assigns individuals to those populations in order to maximise HWE within-populations based on allele frequencies; no prior information of the geographical origin of the individual was included. The number of ancestral populations (K) tested for was 1-10, with 12 simulations run for

each. Burn-in was set at 100,000 and the number of iterations was 1,000,000 under the admixture model with independent allele frequencies. Mean values of posterior probability were calculated from likelihood values for each K, and the optimum K was determined using the ΔK method outlined in Evanno *et al*. (2005).

3.4 Results

3.4.1 Genetic variation

Ten polymorphic loci were genotyped successfully in 241 *D. taylorii* individuals. Across all populations a total of 185 alleles were observed with a mean of 18.5 alleles per locus and the number of alleles per locus ranged from five to 28 (Table 3.2).

Table 3.2: Characteristics of ten microsatellite loci for 241 samples of *Dactylanthus taylorii*

	A	**Size range (bp)**	**F_{ST}**	**F_{IS}**	**H_O**	**H_T**
DATA4	17	304-343	0.26	0.444	0.310	0.481
DATA19	26	148-196	0.303	0.256	0.418	0.480
DATA20	20	162-207	0.369	0.386	0.315	0.449
DATA21	24	240-266	0.235	0.358	0.366	0.517
DATA32	26	207-285	0.292	0.287	0.424	0.532
DATA40	12	257-270	0.393	0.093	0.511	0.479
DATA41	16	227-265	0.41	0.05	0.454	0.449
DATA46	5	247-260	0.26	0.361	0.247	0.339
DATA62	28	112-162	0.252	0.352	0.473	0.582
DATA65	11	254-285	0.378	-0.143	0.441	0.367
Mean	18.5		0.316	0.258	0.464	0.464

A, number of alleles per locus; size range (bp), range of allele sizes at microsatellite loci; F_{ST}/F_{IS}, estimates of Wright's fixation index for each microsatellite locus; H_O, observed heterozygosity; H_T, species-wide expected heterozygosity

A total of 56 private alleles were observed, representing 30% of the total alleles, with an average of 1.81 private alleles per population. Three populations possess 39% of the total private alleles: Mt Pirongia(9), Pikiariki (7) and Te Kopia (6). Eight populations were found to have no private alleles. Within-populations mean number of alleles per locus was 3.16, (range 0.9 to 6.4), but the number of effective alleles per locus was significantly different ($p<0.05$) with an average of 2.31 alleles per locus (range 0.9-4). For both N_A and N_E, Mt Bruce represented the lowest value and Pikiariki the highest (Table 3.1). No significant evidence for linkage disequilibrium was found between any pair of markers.

Polymorphic loci per population ranged from one to nine (10-90%), with an average of 8.58. The number of polymorphic loci was below 50% in populations that typically had a low sample size: Little Barrier Island (n=6), Waipapa Lodge (n=2), Mt Bruce (n=1), although Oropi (n=15) was an exception (Table 3.1).

Observed heterozygosity (H_O) had a large range, with 0.050 at Puketi and 0.767 at Kopuriki, with an average value of 0.396. The range of expected heterozygosity (H_E) was 0.050 at Mt Bruce and 0.703 at Pikiariki, with an average of 0.468 (Table 3.1). In most cases, H_E was greater than H_O, which is commonly observed in microsatellite data sets (Nybom 2004).

Average F_{IS} for all populations was 0.240 but F_{IS} values varied widely between populations, with the lowest being -0.484 at Kopuriki and the highest 0.889 at Puketi. Negative F_{IS} values were found for Waipapa Lodge, Kopuriki, 100 Acre Bush and Takarere populations. This indicates there are a high number of heterozygotes in these populations which is explained by outcrossing. Twenty-two populations were found to deviate significantly from Hardy-Weinberg equilibrium due to heterozygote deficiency (Table 3.1). Results from a linear regression showed that population size had no effect on F_{IS} (R^2=0.0281, P=0.36), but had a marginally significant effect on H_E ($R^{2=}$0.09, P=0.09) and was strongly correlated with N_A (R^2=0.1489, P=0.03 Figure 3.2A). To check for the effect of sample size on diversity, a regression of population size and sample size was performed (Figure 3.2B). With all populations included, an R^2 of 0.0854 was found to be marginally significant (P=0.11). Removing the outlier (Te Araroa, population size of 1000) increased the R^2 value to 0.3409 that was strongly correlated with sample size (P=<0.001). A regression of population size without the

outlier to N_A gave an R^2 of 0.2661 that was strongly correlated (P=0.003). However, the strongest correlation found was between sample size and N_A (R^2=0.3805, $p<0.001$) (Figure 3.2C).

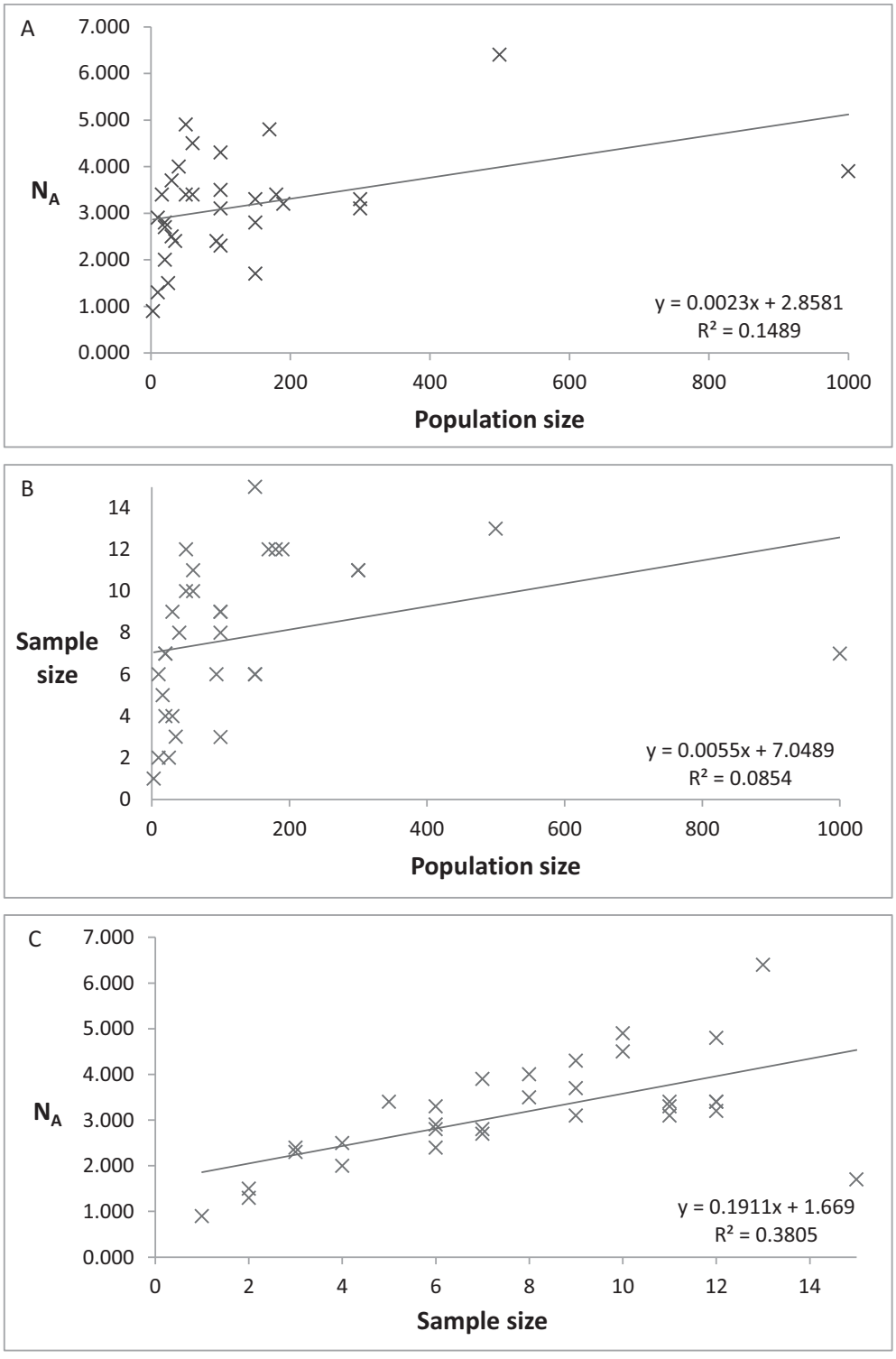

Figure 3.2: Regression analysis between population size and N_A (A), population size and sample size (B), sample size and N_A (C)

3.4.2 Genetic structure and differentiation

Average genetic diversity among all populations was found using global F_{ST} to be 0.316 (Table 2.2). A Mantel test showed that genetic differentiation between the populations correlated significantly with geographical distances (R^2=0.0671, p<0.0001, Figure 3.3). Using AMOVA to examine among population differentiation for all populations found that the majority of the variation was partitioned within populations (67.39%), with the rest of the variation (32.61%) partitioned among populations (Table 3.3).

Using the ΔK method (Evanno *et al.* 2005) for evaluating the most likely value of K from STRUCTURE likelihood results revealed that K=2 provides the best *ad hoc* statistical fit (Figure 3.4). Two clusters reveal a geographic pattern similar to that found using RAPDs (Holzapfel *et al.* 2002). Twelve populations (20-31) represent a geographically central cluster, and all nineteen (1-19) other populations are grouped together. The Little Barrier Island (pop. 2) population clusters predominantly with the central cluster, but is also admixed with the other cluster (Figure 3.5A). AMOVA comparing variation between the two STRUCTURE-derived K=2 groups identified the among region variation as 8.58%, the among population within region variation as 21.43%, and the within population variation as 70% (Table 3.3).

ΔK also showed support for K=3 and K=8 (Figure 3.4). When K=3, the central cluster persists, but the remaining populations are split into two clusters. Populations 1-11 form a cluster representative of the western populations and populations 12-19 are a mostly eastern cluster, but includes Pukerimu and Oropi (pops. 12 & 13) which are geographically separated by the central cluster (Figure 3.5B). Under K=8, the central cluster still remains, but other populations are grouped together in clusters that generally reflect geography. There is a western cluster composed of four populations in the Taranaki region (pops. 10-13), and an eastern cluster with six populations (pops.14-19) in Hawkes Bay-Manawatu region. A western-central cluster has four populations (pops. 6-9) and also shows a high level of admixture with other clusters which is likely due to its adjacency with other clusters. Several clusters are represented predominantly by one or two populations, including a Puketi cluster, a Little Barrier Island cluster (also found in several individuals from Te Araroa), a Mount Pirongia cluster and a cluster with Oropi and Te Kopia (Figure 3.5C).

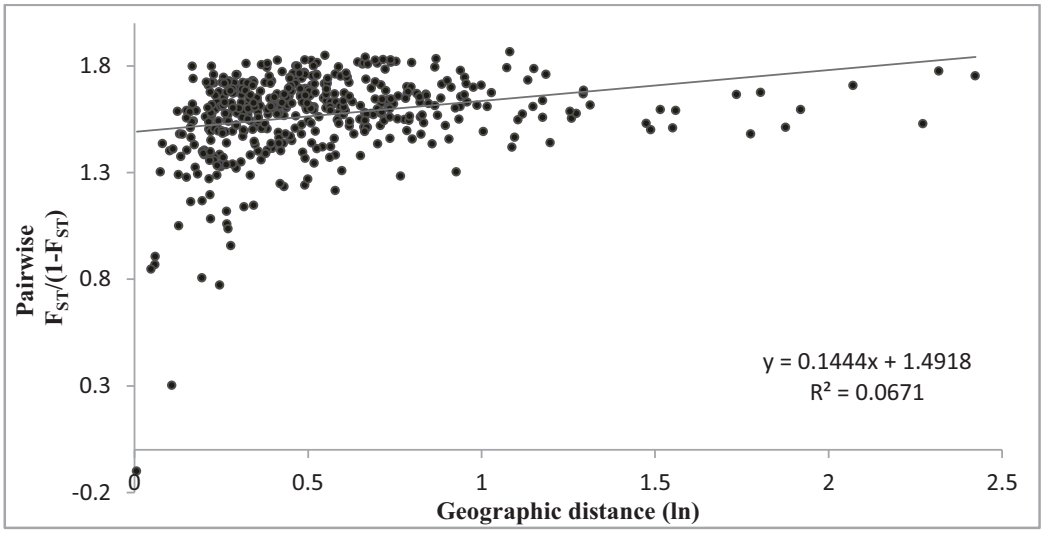

Figure 3.3: Pairwise genetic distance ($F_{ST}/(1-F_{ST})$ and the natural log of geographic distance (km) between populations of *Dactylanthus taylorii*

	Source of variation	d.f	Sum of squares	Variance components	Percentage variation
A) Total populations	Among pop.	30	382.225	0.79306	32.61
	Within pop.	414	669.402	1.63893	67.39
B) Between regions	Among regions	1	68.67	0.25609	8.57571
	Among pop. within groups	28	319.172	0.63989	21.4279
	Within pop.	414	863.338	2.09025	69.99639

Table 3.3: Values of AMOVA partitioning of microsatellite variation in: A) all populations, B) comparing between regions identified in STRUCTURE

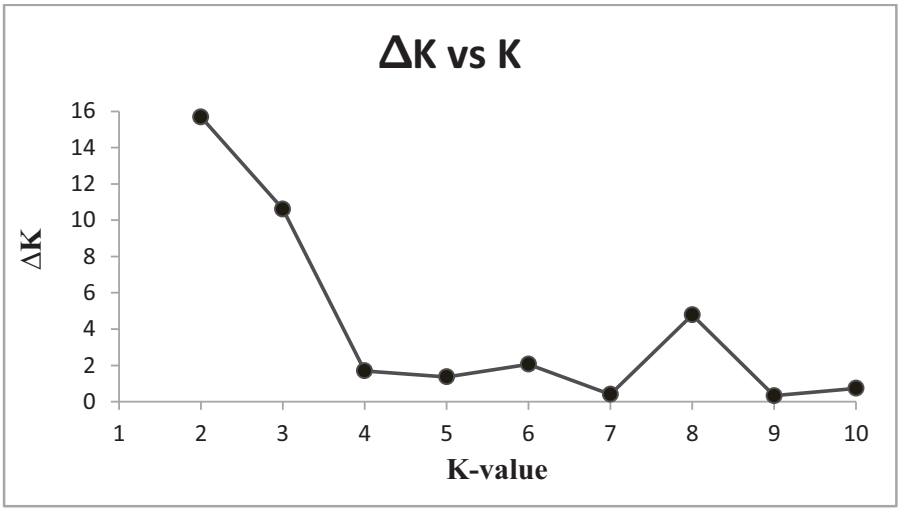

Figure 3.4: Plot of ΔK vs. K from STRUCTURE following Evanno (2005) based on 12 replicates for each value of K.

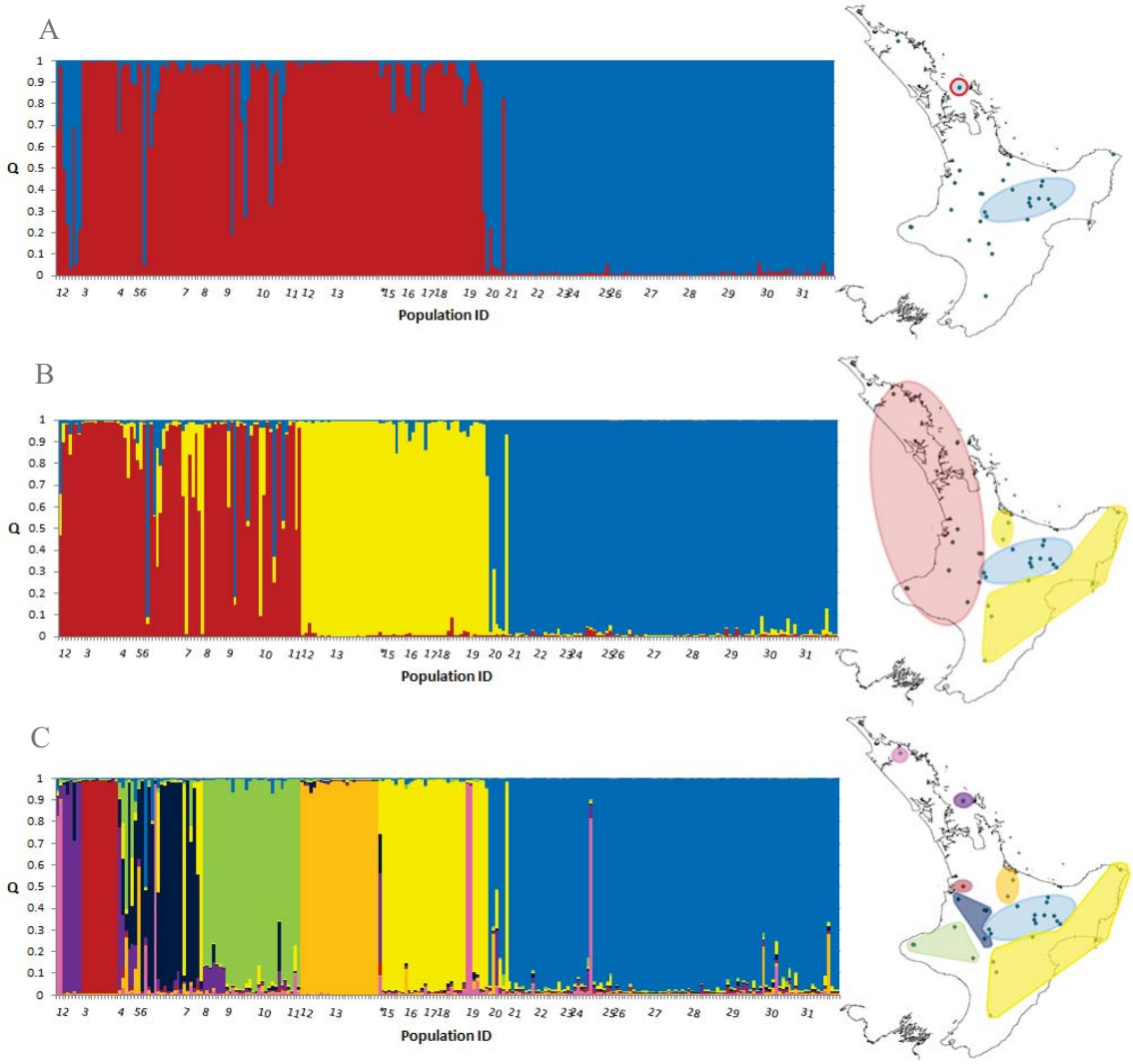

Figure 3.5: STRUCTURE cluster assignment of *Dactylanthus taylorii* individuals and geographical representation of cluster results. A, K=2, the blue shaded circle with the red outline represents the Little Barrier Island population (2) which has an even mixture of both clusters. B, K=3. C, K=8, the Mt Bruce population (14) is shown to be grouped with cluster 7 based on geography despite it showing evidence for belonging to other clusters. Population numbers as in Figure 3.1 and Table 3.1 except for Mt Bruce (14) is represented by an asterisk. Colours in the STRUCTURE graphs reflect those shown on the maps.

Table 3.4: Species used to compare genetic variation and structure with *Dactylanthus taylorii*. All taxa were outcrossing and rare in some way. Alleles/locus, average number of microsatellite alleles detected over all loci; N_A, number of alleles per locus in a population; H_O and H_T, species wide observed and expected heterozygosities; F_{ST}, global estimates of species F_{ST} (note: F_{ST} was not calculated the same way in all studies. Values from this study were not used for calculating means.

Reference	Plant	Family	Growth form	Sex system	A	N_A	H_O	H_T	F_{ST} (species)	Rare
Aegisdottir *et al.* 2009	*Campanula thyrsoides*	Campanulaceae	perennial monocarp	outcrossing	20	8	0.763	0.762	0.1	Y
Furches *et al.* 2009	*Lithophragma maximum*	Saxifragaceae	perennial herb	most species in genera self-incompatible	6.4	2.05	0.214	0.504	0.524	Y
Holmes *et al.* 2009	*Grevillea repens*	Proteaceace	shrub	preferential outcrosser	12.17	5.7	0.553	0.577	0.272	Y
O'Brien *et al.* 2010	*Banksia spinulosa* var. *spinulosa*	Proteaceae	shrub	Selfing possible	13.75	6.49	0.69	0.77	0.0614	Y
Riley *et al.* 2010	*Galium catalinense*	Rubiaceae	shrub	outcrossing	8.13	2.58*	0.494	0.55	0.104	Y
Segarra-Moragues *et al.* 2005	*Borderea chouardii*	Dioscoreaceae	geophyte	outcrossing	3.35	2.05	0.14	0.17	0.35	Y
Sosa *et al.* 2010	*Sambucus palmensis*	Adoxaceae	shrub	outcrossing	6.8	6.8	0.55	0.5	0.229	Y
Straub & Doyle 2009	*Amorpha georgiana*	Fabaceae	shrub	mixed	17.7	4.3	0.65	0.75	0.26	Y
This study	***Dactylanthus taylorii***	**Balanophoraceae**	**parasite**	**outcrossing**	**18.5**	**3.16**	**0.396**	**0.468**	**0.316**	**Y**
Yao *et al.* 2007	*Changiostyrax dolichocarpa*	Styracaceae	tree	outcrossing	7.375	5	0.594	0.643	0.064	Y
Yan *et al.* 2009	*Medicago ruthenica*	Fabaceae	perennial herb	outcrossing	16.4	7.4	0.574	0.677	0.13	Y
				Mean	11.21	5.04	0.522	0.590	0.209	

*actually N_E

3.5 Discussion

Many rare plant species exist only in small or isolated populations in fragmented habitats and so have an increased risk of extinction due to genetic factors as well as demographic and environmental stochasticity (Young *et al.* 1996; Ouborg *et al.* 2006). Genetic drift, inbreeding and reduced gene flow due to isolation are consequences of small population sizes which function to reduce genetic variation within a population and increase genetic differentiation between populations (Ellstrand & Elam 1993). These factors reduce the evolutionary potential of the species and decrease a populations ability to respond to environmental challenges. Conservation genetics studies attempt to assess the genetic variation and structure of endangered species in order to determine the genetic threats to the species survival with the goal of contributing to conservation management (Hedrick 2001).

Due to the isolated nature of the populations and small population sizes, populations of *Dactylanthus taylorii* were expected to harbour little genetic variation (Ellstrand & Elam 1993; Ouborg *et al.* 2006). Little genetic work has been performed on other members of Balanophoraceae and there have been few studies of the New Zealand flora using microsatellites, so strict comparisons between *D. taylorii* genetic diversity and species with a similar parasitic nature and biogeographical histories are not feasible. Reproductive systems and other life history traits of a species have been described as significant factors affecting genetic diversity and structure of plant populations (Hamrick *et al.* 1991; Hamrick & Godt 1996). In this case, genetic diversity found in populations of *D. taylorii* was compared to other long-lived, outcrossing species that are rare (Table 3.3).

3.5.1 Genetic variation

Microsatellite variation in *D. taylorii* measured by the number of alleles per microsatellite locus (A=18.5) was considerably higher than the average (A=9.9) found in a review by Nybom (2004) and in a compilation of studies on rare plants (Table 3.3, average A=11.21, range=3.35-20). This result is inflated by the large number of private alleles (which represent 30% of the total alleles observed) resulting in the allelic diversity within populations being low compared to other rare species (N_A=3.155 vs. average N_A=5.04). Species expected heterozygosity (H_E=0.468) was lower than the

average for other rare species (Table 3.3, H_E=0.590, range=0.17-0.77) and the average found by Nybom (2004) (H_E=0.61). Comparing between species with similar life histories (Nybom 2004), *D. taylorii* displays lower expected heterozygosity than long lived perennials (H_E=0.68) and outcrossing species (H_E=0.65), but similar heterozygosity to species with gravity dispersed seed (H_E=0.47) and early successional plants (H_E=0.46).

Perennial species with overlapping generations and outcrossing breeding systems are expected to have high genetic variation as both of these factors promote diversity within a population (Hamrick & Godt 1996). The lifespan of *Dactylanthus taylorii* has been estimated, using tree ring counts of woodroses, to be at least 30 years but a maximum age is unknown (Ecroyd 1996b). Long-lived perennial species have a high reproductive output which should function to increase or maintain genetic variability (Petit & Hampe 2006). High F_{IS} values in 22 of the 31 populations indicate deviations from Hardy Weinberg equilibrium due to homozygote excess. Despite the fact that *D. taylorii* is dioecious and must outcross to set seed, a potential cause of this is inbreeding between closely related individuals. While such high levels of inbreeding are unexpected in a plant that should be obligately outcrossing, the effect of small breeding population sizes, pollinator decline/absence and a limited mode of seed dispersal could all contribute to mating between closely related individuals. Many populations also have a much higher ratio of males to females (A. Holzapfel, pers. comm) reducing the effective size of the populations with only a few females contributing to the gene pool and lowering the chances of successful seed set. The ability of *D. taylorii* to vegetatively reproduce via specialised infectious roots at the base of the inflorescence (Holzapfel 2001) could also contribute to reduced genetic variation (de Witte & Stocklin 2010). Meta-analysis studies have suggested that self-incompatible species have an increased risk of loss of genetic variation after habitat fragmentation due to a reduction in effective population size (Leimu *et al.* 2006). Outcrossing and long lived plant life history traits are typically factors found to increase or maintain genetic variation, but a reduced ability to cross caused by pollinator loss, decreased population size or reduction in females within a population could explain why they are not functioning to maintain diversity in *D. taylorii*.

Two other *Dactylanthus taylorii* life history traits are often associated with lower levels of genetic variation, gravity dispersal of seeds and growth in early succession habitats.

The expected heterozygosity found in *D. taylorii* is similar to the average expected heterozygosities for plants that have gravity dispersed seeds and grow in developing habitats (Nybom 2004). Gravity dispersal is assumed as the fruits produced by *D. taylorii* are small and not fleshy, and no native disperser has been identified. However, something other than gravity must function to move seeds around to explain the presence of *D. taylorii* on the tops of ridges and its general distribution throughout the North Island (Holzapfel 2001). Gravity dispersal limits seed movement greatly and while an inflorescence laden with seed detaches from a plant to aid dispersal, it may not travel far from the clump they are assumed to originate from and therefore genetic information would not be moved far from the maternal source. Because of the seral nature of the most common species *D. taylorii* parasitises, it has been suggested that it is an early coloniser of disturbed habitats and provides nectar for bats and insects in the early stage of forest development (A. Holzapfel, pers. comms.). The theory of parasitic plants as habitat facilitators has been previously explored and seems to be a characteristic of many parasitic species (Watson 2009). There are many characteristics of *D. taylorii* that suggest it could play this role in ecosystem development including: primarily parasitising trees found in early successional forests (La Cock *et al.* 2005), an abundance of carbohydrate rich nectar (Ecroyd *et al.* 1995) and a short time of four to ten years from seed germination to flower production compared to other colonising plant taxa (Holzapfel & Dodgson 2010). Populations recolonising habitats would have reduced genetic diversity due to population bottlenecks and founder effects (Tian *et al.* 2009).

Due to different sample sizes, evaluating the effect of population size on genetic diversity is difficult (Figure 3.2). The theoretical implication of small population size is the increased effects of factors that reduce genetic diversity; such as drift, bottlenecks and founder effects (Ellstrand & Elam 1993; Young *et al.* 1996). Most of the populations studied here contain fewer than 200 individuals, which can be spread over a large geographic area, although estimates are difficult to make due to the cryptic nature of the plant. Additionally, the high male: female sex ratio found in most populations suggests that the effective population size is much lower. It is possible that plants are naturally uncommon in populations, but there is evidence that even small areas are able to support a large number of plants with the only limiting factor being healthy host trees (Hill 1926; Holzapfel 2001). Some evidence was found for small populations having

lower genetic variation than large populations and therefore may be at a greater risk of extinction due to genetic reasons, but these results may be strongly affected by variation in the sample sizes. There is a strong correlation between sample size and population size when the Te Araroa population is removed (as it an outlier with 1,000 plants) which suggests the sample size was dependant on population size in many cases. A potential cause for the lack of a clear relationship between population size and variation is due to the longevity of *Dactylanthus taylorii*, if populations only recently began to decrease in size there may have been insufficient generations for a reduction in diversity (Hoebee & Young 2001; Myking *et al.* 2009). The link between population or sample size and genetic variation requires further investigation before any management decisions are made based on smaller populations being at greater risk than large populations.

Low levels of genetic variation also have been observed in several other studies of New Zealand native plant species, including taxa that are common and endangered (Haase 1992a, b; Young *et al.* 2001; Barnaud & Houliston 2010). The hypothesised cause of the reduced genetic diversity in various taxa is population bottlenecks instigated by repeated glaciations and corresponding expansion and contraction of plant habitats during the Pleistocene (McGlone *et al.* 2001). Glaciation cycles during the Pleistocene are also believed to have affected speciation and genetic diversity of plant species in other parts of the world (Hewitt 1996). It is not known how cold tolerant *Dactylanthus taylorii* is (the Maude and Mangorei populations are at an elevation of 1000 metres) but the tree species it is believed to parasitise most commonly are components of lowland forests which are typically adapted to warmer climates and therefore would not have been abundant during glaciations (McGlone *et al.* 2001). Reduction in suitable habitat (i.e. hosts), during the Pleistocene could contribute to the low genetic diversity in *D. taylorii* observed in this study. Alternatively, because these studies used markers that typically show lower levels of genetic variation than microsatellites (e.g. allozymes) it is possible that the low level of genetic variation is due to the marker type rather than species biology (Shepherd & Perrie 2011).

3.5.2 Genetic structure and differentiation

The species wide F_{ST} value of 0.316 (Table 3.1) indicates a high degree of genetic structure. Compared to Nybom's (2004) meta-analysis, the F_{ST} found in this study is

higher than the average for long lived perennials (0.19) and outcrossing taxa (0.22), but similar to the values for species with gravity dispersed seed (0.34) and early successional taxa (0.37). It is also higher than the average found from a comparison of rare species (Table 3.3, F_{ST}=0.209). Many populations are almost completely isolated from each other; therefore gene flow between populations is expected to be low. A strong correlation between pairwise F_{ST} values and geographic distance suggest that if gene flow does occur it is mostly between neighbouring populations (Slatkin 1993; Wright 1945).

Populations that are greatly isolated by distance and are assumed to have been for a long period of time are expected to have accumulated a greater number of private alleles due to mutation and restricted gene flow (Segarra-Moragues *et al.* 2005; Sosa *et al.* 2010). This does not appear to be the case for *Dactylanthus taylorii*, as under this assumption the most geographically isolated populations (Puketi, Little Barrier Island, Te Araroa and Mt Bruce) would be expected to have a substantially higher number of private alleles than the rest of the populations. Instead, they have a number of private alleles close to the species average. In fact, the populations with private alleles much greater than the average (Mt Pirongia, Pikiariki, and Te Kopia) are all found within 110km of each other, and the longest distance between one of them and another population of significant size is 60km. The cause of this pattern is unclear. It's possible these populations were isolated for some other reason other than distance, or have had a higher number of generations than others, both of which could have contributed to a high number of private alleles.

The structure of genetic variation under the K-values of 2 and 3 is notable as it is similar to the geographic pattern found with RAPD markers using Nei's genetic distance, which separated eastern populations from western populations (Holzapfel *et al.* 2002). Under K=2 a cluster of populations throughout the centre of the North Island is found with the rest of the populations in the other cluster (Figure 3.5A). AMOVA indicated that 8.6% of genetic variation was found among these two regions (with 70% found within populations and 21.4% among populations) (Table 3.3). When increased to K=3 (Figure 3.5B), the central cluster persists but the rest of the populations are split into eastern and western groups. While the clusters showed broad geographic patterns, in some cases adjacent populations separated by less than 20km (e.g., Tongariro Forest and Kakaramea) were grouped into different clusters.

Volcanism was inferred to be the cause of the geographic separation of the species by Holzapfel *et al* (2002), as the Taupo Volcanic Zone (TVZ) is a 300km long area through the central plateau of the North Island (Figure 3.6) and roughly coincides with the geographical split found using RAPDs. The TVZ is characterised by having a geology derived from volcanic rock compared to the sedimentary nature of most of the North Island. There are two major eruptions that would have shaped the landscape within the TVZ; the Oruanui (26.5 thousand years ago) and the Taupo/Hatepe eruption (1.8 thousand years ago) (Manville & Wilson 2004). The more recent Taupo eruption is believed to have caused ash fall around the caldera and the associated heat would have destroyed and carbonised all vegetation within the 20,000km^2 blast zone (Manville *et al.* 2009). Supporting evidence for the effect of the TVZ on the biogeography of *D. taylorii* has been found in other plant and animal species that also display an east-west differentiation. These examples include the fern *Asplenium hookerianum* (Shepherd *et al.* 2007), two moa species (Baker *et al.* 2005), the North Island brown kiwi (Shepherd & Lambert 2008), and the short tail bat (Lloyd 2003). The K=2 result suggests that the populations found nearest the area of volcanism represent genetic remnants of an ancestral population different to the rest of the populations.

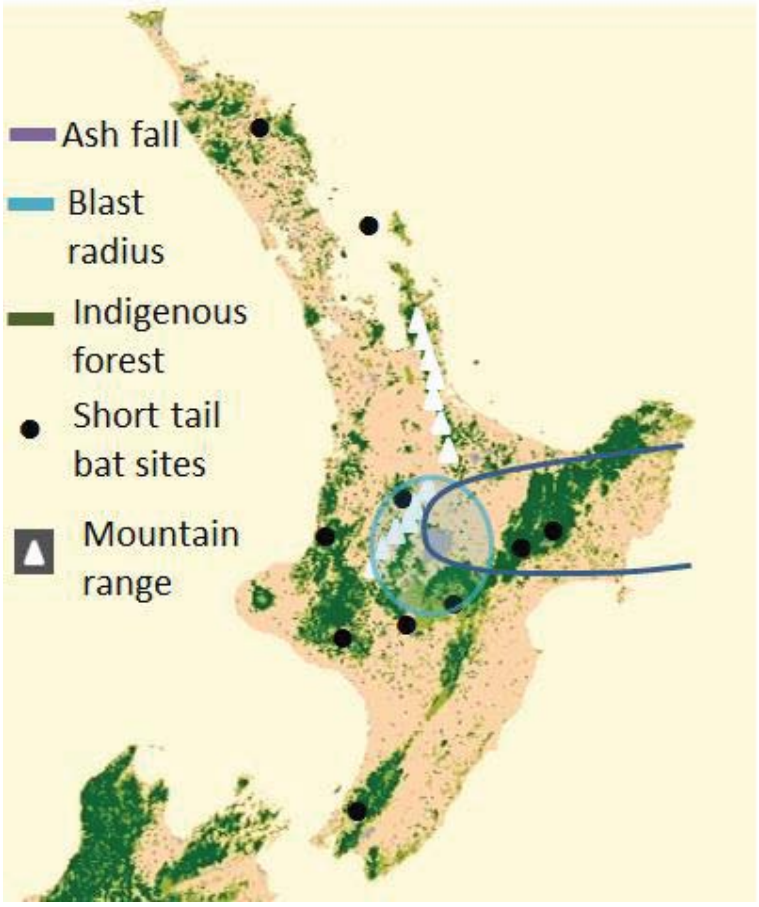

Figure 3.6: North Island map representing possible causes of past and present structuring of *Dactylanthus taylorii* populations. The remaining indigenous forest and mountain range is taken from the Environment of New Zealand 2007 report (van Bunnik *et al.* 2007), short tail bat sites are from Lloyd (2003) and the area affected by the Taupo volcanic eruption from Manville *et al.* (2009).

Volcanism provides a logical explanation for the geographic patterns observed when K=2/3, but over time it is expected that *Dactylanthus taylorii* would be able to reinvade volcanically disturbed habitat as its favoured hosts are species commonly found in regenerating forests in New Zealand. Maintenance of this pattern is probably through habitat fragmentation and the reduction in numbers of the short-tail bat (O'Donnell *et al.* 2010). Figure 3.6 shows the remaining indigenous forest and populations of short-tail bat (black circles). Most of the central cluster populations are found within a contiguous

tract of forest with known bat populations, which potentially explains the genetic similarity. Populations at Te Kopia, Kakaramea and 100 Acre Bush may also be visited by bats that are known to use surrounding *Pinus radiata* plantations (Borkin & Parsons 2010). Short-tail bats are believed to fly approximately 10km on an average night, but it is known they can fly up to 40km (Lloyd 2001). Most of the populations in the central cluster are potentially within the range of a single bat's flight. *D. taylorii* populations at Tongariro Forest and Wharangi are within the range of bats that are potentially servicing nearby populations but may not be known about. The K=3 pattern suggests the eastern populations were formerly connected but separated due to the TVZ and the central cluster has formed due to persistence of populations in unfragmented habitat and exclusive bat pollination. However, it is unclear whether volcanism could explain the eastern-western pattern found under K=3 as the TVZ doesn't fall exactly where the genetic split is found. If volcanism was responsible for this pattern, you would expect that populations to the west of the TVZ would all be a single cluster, but the Oropi and Pukerimu populations cluster with eastern populations. Figure 3.6 also shows the geographic positions of ranges that run through the centre of the North Island, and could represent a barrier to gene flow between the western and two eastern clusters. Whether bats would cross these series of ranges is unknown, but the genetic structure of populations at K=3 largely reflects a pattern that could be explained by the ranges as a barrier to restrict gene flow.

Also, the genetic clusters revealed by STRUCTURE (for K=2 and 3) and the isolation by distance results suggest that historically the species may have been only a few large populations (or meta-populations). Before severe habitat fragmentation, suitable host species would have been found over much of the North Island and a continuous distribution of *Dactylanthus taylorii* could have been possible with gene flow occurring through pollination by bats. This contrasts with many New Zealand species which have disjunct distributions (McGlone *et al.* 2001; Shepherd & Perrie 2011). This idea was considered by Holzapfel (2001) but rejected on the basis of the genetic results from the RAPDs study and evidence suggesting that short-tail bats are more likely to facilitate pollination between flowers in the same population, and *D. taylorii* populations are separated by more than 40km (the assumed flight range of a bat in one night) (Holzapfel *et al.* 2002). However, in the meta-population scenario, populations may not have been

so isolated and many individuals could be accessed in a single night through continuous forest.

It has been suggested that the unique life-history strategy of parasitism means that plant parasites do not show classical patterns of genetic diversity and structure, but population genetic studies of parasitic plants are rare (Barrett *et al.* 2008). A study on the dioecious mistletoe, *Viscum album*, which has a fragmented habitat distribution, found similar results to this study. There was strong evidence for isolation by distance and genetic structuring of populations suggesting an overall lack of gene flow, and despite being obligate outcrossers there was inbreeding in populations and low genetic variation (Stanton *et al.* 2009). Other studies on parasitic plant structure have indicated high levels of gene flow (Mutikainen & Koskela 2002) and evidence of reduced gene flow between plants of the same species differentiated by host specialisation (Thorogood *et al.* 2009).

Holoparasitic plant species typically show extreme host specialisation but even generalist species have shown specialisation at the local level (Thorogood & Hiscock 2007). The number of putative hosts of *D. taylorii* is approximately 30, though by the nature of the habitat some hosts are more common than others. Generalist parasites evolve in heterogeneous environments with many potential hosts, whereas specialists are dependent on host availability (Norton & De Lange 1999). It is possible that a reduction in gene flow and adaptation to local hosts could lead to the evolution of host specific races of *D. taylorii* genotypes.

3.5.3 Conservation implications

Before the arrival of humans, *Dactylanthus taylorii* biogeography would have been affected by gene flow between populations facilitated by bats, the presence of suitable hosts in an environment and volcanism. Pollen records suggest that this species was once more widespread than it is currently (Holzapfel 2001; Mildenhall & Alloway 2008) and suitable hosts would have been present over the whole of the North Island; STRUCTURE results indicate the possibility of a few large meta-populations maintained by gene flow via pollination. Since human arrival, *D. taylorii* has been affected by habitat degradation reducing the number of individuals and creating isolation between populations, reduction in numbers of the major pollinator, and browsing of inflorescences by introduced mammals (Ecroyd 1996b). A conservation effort from the DOC and the *Dactylanthus* Recovery Group has reduced the effect of mammalian browsing by caging *D. taylorii* individuals in most populations and nearly all populations are found in areas protected by law (La Cock *et al.* 2005).

However, the impact of human arrival and subsequent effects on *Dactylanthus taylorii* ecology appear to have affected the genetic variation of populations and the species overall. Genetic diversity is lower than that found for many other rare species (Table 3.3) as well as other New Zealand taxa (Haase 1992a, b; Shepherd & Perrie 2011). While the correlation between population size and genetic variation in *D. taylorii* is unclear, there is theoretical (Ellstrand & Elam 1993) and empirical (Leimu *et al.* 2006) evidence that small populations have reduced variation compared to large populations due to population bottlenecks and genetic drift. High F_{IS} scores are unusual in dioecious species and indicate the effect of reduced pollen dispersal causing loss of variation in the species (Myking *et al.* 2009). The genetic structure of *D. taylorii* and the corresponding geographic structure of the two optimal K-values of two and three can be explained by historic geological, biological and anthropomorphic events. However, for conservation purposes it may be more useful to consider the species as eight genetic clusters rather than two or three separate clusters. The eight clusters typically reflected groupings that were geographically sensible (with exceptions between central cluster populations and non-central populations), and combined with the IBD results, represent evidence of gene flow between geographically adjacent populations.

Most populations are now in protected habitats and individuals are caged to reduce browsing. However, the lack of genetic variation within populations is cause for concern for the survival of individual populations and has long term consequences for species persistence. There are several management actions that could enhance the genetic variation of populations and the natural recovery of *Dactylanthus taylorii*. First, enhancing outcrossing within populations would promote genetic diversity and reduce inbreeding between closely related individuals and clones. Artificial pollination between plants is already undertaken in many populations but the seed production from these crosses is unknown (La Cock *et al.* 2005). Pollination between plants that are not adjacent should be encouraged and successful seed set recorded if possible. Second, in populations where seed is set, an attempt should be made to disperse it in order to promote recruitment. Natural dispersal is presumed to be predominantly by gravity; however female inflorescences with seed in cages often do not move very far (pers. observ.). The effectiveness of establishing plants by hand seeding has been proven at a trial site with great success (Holzapfel & Dodgson 2010). Third, consideration should be given to the transfer of genetic information between populations via pollen or seed in order to increase variation within populations. Pollen transfer would be easier, but seed transfer is more likely to introduce completely novel genetic information. Fourth, new populations could be established in areas with the ultimate goal of decreasing isolation between populations. For translocations of pollen or seed, results from STRUCTURE and IBD suggest movement of genetic material should be within the eight clusters and between adjacent populations to reduce the deleterious effects of outbreeding depression via coadapted gene complexes which has been observed in translocations of other rare species (Godefroid *et al.* 2011). Consideration should also be given to the hosts present at the source of genetic material and the habitat that it is being moved into due to the possibility of host-specific races. While translocation of genetic information could be used to introduce novel variation into genetically depauperate populations, there are also negative consequences of effecting local adaptation that need to be considered (Hufford & Mazer 2003). Translocations should initially be performed on a small scale and monitored carefully for signs of reduced fitness due to outbreeding depression by comparing between seed-set of females pollinated by local and introduced pollen (Heiser & Shaw 2006).

The conservation guidelines established by Holzapfel *et al.* (2002) based on their study of 16 populations using RAPDs focused on the distinctness of populations, the poor correlation they found between genetic and geographic distance (R^2=0.011, P=0.219), the east-west geographical division and that the species may be naturally isolated and therefore adapted to the consequences of small population size. The comparably high global F_{ST} found here indicates a high degree of genetic differentiation between populations and supports the distinct population clustering found using UPGMA. Similarly, geographic differentiation of populations around the TVZ was also found reinforcing the hypothesis that volcanism has influenced the biogeographical history of the species. Similar results have been found in other studies of rare plants using both RAPDs and microsatellites; broad geographical patterns are found using RAPDs but better resolution of these patterns are made using microsatellites (Aegisdottir *et al.* 2009; Segarra-Moragues *et al.* 2005). In contrast to the isolation by distance results using RAPDs, in this study a strong correlation was found between population pairwise F_{ST} values and distance. STRUCTURE results from K=8 also suggest historically there was gene flow between adjacent populations. The high level of inbreeding, low genetic variation and potential effect of population size on genetic variation is evidence contrary to adaptation to small population sizes. While RAPDs identified population structure and evidence of a biogeographic pattern, microsatellite data were able to refine that pattern as well as provide information on within-population genetic diversity and these results can be utilised to aid conservation management of *Dactylanthus taylorii.*

Despite the large number of threatened plants in New Zealand there are fewer than ten published conservation genetic studies of the flora. Conservation genetic studies of New Zealand plants have typically focused on species that have a narrow distribution and few have utilised microsatellites to estimate genetic variation or determine population structures. The finding from this study that populations of *Dactylanthus taylorii* have low levels of genetic variation has important conservation implications. Without knowing the genetic diversity of other rare native species, direct management to avoid extinction from genetic causes cannot be made. The effect of the Pleistocene glaciation on genetic variation within species is also unclear and warrants further investigation as high levels of genetic diversity have been found in the New Zealand fierce lancewood, (*Pseudopanax ferox*) using microsatellites (Shepherd & Perrie 2011), whereas low levels of diversity in other tree species has been detected using allozymes (Young *et al.*

2001). Furthermore, additional studies of widespread species would help determine the role of volcanism in shaping the biogeographical patterns observed in native plants and animals.

3.6 Conclusion

A low within-population genetic diversity has been found for the New Zealand endemic root holoparasite, *Dactylanthus taylorii*. Despite being long-lived and outbreeding, life history traits typically thought to increase or maintain variation within populations, low levels of genetic diversity were detected. Furthermore, high F_{IS} scores were observed in many populations indicating heterozygote deficiency, which could be caused by inbreeding between closely related individuals or clones. The relationship between population size and variation is unclear and warrants further research to determine the genetic consequences of population bottlenecks through habitat fragmentation and founder effects.

Considerable levels of genetic differentiation were found and the isolation by distance results suggests that if gene flow occurs it is restricted to adjacent populations. Populations grouped in clusters that reflected geography, and biogeographical structure was inferred to be due to volcanism, mountain barriers, habitat fragmentation and gene flow via pollination by bats.

The low levels of genetic variation found in *Dactylanthus taylorii* are cause for concern of the extinction of populations due to genetic consequences. Sources of genetic information for translocation of seeds or pollen to increase genetic diversity within population should be based on the between-population differentiation and population genetic structure results. The findings from this study are a significant advancement in understanding the risk of *D. taylorii* loss due to genetic reasons and the knowledge of genetic variation within-populations and genetic structure between-populations will be useful to guide future management decisions for the species.

3.7 References

Aegisdottir HH, Kuss P, Stocklin J (2009) Isolated populations of a rare alpine plant show high genetic diversity and considerable population differentiation. *Annals of Botany* **104**, 1313-1322.

APG (2009) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* **161**, 105-121.

Armstrong TTJ, De Lange PJ (2005) Conservation genetics of Hebe speciosa (Plantaginaceae) an endangered New Zealand shrub. *Botanical Journal of the Linnean Society* **149**, 229-239.

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

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.

Barrett LG, Thrall PH, Burdon JJ, Linde CC (2008) Life history determines genetic structure and evolutionary potential of host-parasite interactions. *Trends in Ecology & Evolution* **23**, 678-685.

Borkin KM, Parsons S (2010) Plantation forests are used by the lesser short-tailed bat, Mystacina tuberculata rhyacobia. *New Zealand Journal of Zoology* **37**, 13-17.

Boutin-Ganache I, Raposo M, Raymond M, Deschepper CF (2001) M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *Biotechniques* **31**, 24-+.

Craig J, Anderson S, Clout M*, et al.* (2000) Conservation issues in New Zealand. *Annual Review of Ecology and Systematics* **31**, 61-78.

de Lange P, Norton D, Heenan P*, et al.* (2004) Threatened and uncommon plants of New Zealand. *New Zealand Journal of Botany* **42**, 45-76.

de Lange PJ, Norton DA, Courtney SP*, et al.* (2009) Threatened and uncommon plants of New Zealand (2008 revision). *New Zealand Journal of Botany* **47**, 61-96.

de Lange PJ, Sawyer JWD, Rolfe J, Network NZPC (2006) *New Zealand indigenous vascular plant checklist* New Zealand Plant Conservation Network.

de Witte LC, Stocklin J (2010) Longevity of clonal plants: why it matters and how to measure it. *Annals of Botany* **106**, 859-870.

Dopson S, de Lange PJ, Ogle C (1999) *The conservation requirements of New Zealand's nationally threatened vascular plants* Biodiversity Recovery Unit, Department of Conservation.

Doyle JJ, Doyle LJ (1987b) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 11-15.

Ecroyd CE (1996a) Dactylanthus and bats: The link between two unique endangered New Zealand species and the role of the community in their survival. In: *Nature Conservation 4: The Role of Networks* (eds. DA. S, JL. C, EM M), pp. 78-87. Surrey Beatty & Sons, Chipping Norton NSW.

Ecroyd CE (1996b) The ecology of Dactylanthus taylorii and threats to its survival. *New Zealand Journal of Ecology* **20**, 81-100.

Ecroyd CE, Franich RA, Kroese HW, Steward D (1995) Volatile constituents of Dactylanthus taylorii flower nectar in relation to flower pollination and browsing by animals. *Phytochemistry* **40**.

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.

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.

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.

Ferreira SM (2005) Individual-level management trade-offs for populations of Dactylanthus taylorii (Balanophoraceae). *New Zealand Journal of Botany* **43**, 415-424.

Given D (1981) *Rare and endangered plants of New Zealand* Reed, Wellington.

Godefroid S, Piazza C, Rossi G*, et al.* (2011) How successful are plant species reintroductions? *Biological Conservation* **144**, 672-682.

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

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

Haase P (1992b) Isozyme variation and genetic-relationships in *Phyllocladus trichomanoides* and *P. alpinus* (Podocarpaceae). *New Zealand Journal of Botany* **30**, 359-363.

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, Murawski DA, Loveless MD (1991) *Correlations between species traits and allozyme diversity - implications for Conservation Biology* Oxford Univ Press, New York.

Hedrick PW (2001) Conservation genetics: where are we now? *Trends in Ecology & Evolution* **16**, 629-636.

Heiser DA, Shaw RG (2006) The fitness effects of outcrossing in Calylophus serrulatus, a permanent translocation heterozygote. *Evolution* **60**, 64-76.

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

Hill H (1926) Dactylanthus taylori. *Trans. New Zealand Inst.* **56**, 87-90.

Hoebee SE, Young AG (2001) Low neighbourhood size and high interpopulation differentiation in the endangered shrub Grevillea iaspicula McGill (Proteaceae). *Heredity* **86**, 489-496.

Holzapfel S (2001) Studies of the New Zealand Root-Parasite Dactylanthus taylorii (Balanophoraceae). *Englera* **22**, 3-176.

Holzapfel S, Dodgson J (2010) Conservation introduction of the parasitic plant dactylanthus at Waipapa, New Zealand. In: *Global Re-Introduction Perspectives: Additional case-studies from around the globe* (ed. Soorae PS). IUCN/SSC Re-introduction Specialist Group, Abu Dhabi, UAE.

Holzapfel S, Faville MZ, Gemmill CEC (2002) Genetic variation of the endangered holoparasite *Dactylanthus taylorii* (Balanophoraceae) in New Zealand. *Journal of Biogeography* **29**, 663-676.

Hufford KM, Mazer SJ (2003) Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution* **18**, 147-155.

Kelly D, Sullivan JJ (2010) Life histories, dispersal, invasions, and global change: progress and prospects in New Zealand ecology, 1989-2029. *New Zealand Journal of Ecology* **34**, 207-217.

La Cock GD, Holzapfel S, King D, Singers N (2005) *Dactylanthus taylorii recovery plan, 2004–14* Department of Conservation (NZ), Wellington

Leimu R, Mutikainen P, Koricheva J, Fischer M (2006) How general are positive relationships between plant population size, fitness and genetic variation? *Journal of Ecology* **94**, 942-952.

Lloyd BD (2001) Advances in New Zealand mammalogy 1990-2000: Short-tailed bats. *Journal of the Royal Society of New Zealand* **31**, 59-81.

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

Manville V, Segschneider B, Newton E*, et al.* (2009) Environmental impact of the 1.8 ka Taupo eruption, New Zealand: Landscape responses to a large-scale explosive rhyolite eruption. *Sedimentary Geology* **220**, 318-336.

Manville V, Wilson CJN (2004) The 26.5 ka Oruanui eruption, New Zealand: a review of the roles of volcanism and climate in the post-eruptive sedimentary response. *New Zealand Journal of Geology and Geophysics* **47**, 525-547.

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.

McLay T, Tate JA, Biggs P, Symonds VV (*in prep*) Isolation of microsatellite markers from the endangered root holoparasite Dactylanthus taylorii using 454 sequence reads. *Conservation Genetics Resources* ***In prep.***

Mildenhall DC, Alloway BV (2008) A widespread ca. 1.1 Ma TVZ silicic tephra preserved near Wellington, New Zealand: Implications for regional reconstruction of mid-Pleistocene vegetation. *Quaternary International* **178**, 167-182.

Moore LB (1940) The structure and life-history of the root parasite Dactylanthus taylori Hook. f. *New Zealand Journal of Science and Technology, Section B* **21**, 206-224.

Mutikainen P, Koskela T (2002) Population structure of a parasitic plant and its perennial host. *Heredity* **89**, 318-324.

Myking T, Vakkari P, Skroppa T (2009) Genetic variation in northern marginal Taxus baccata L. populations. Implications for conservation. *Forestry* **82**, 529-539.

Nickrent DL, Der JP, Anderson FE (2005) Discovery of the photosynthetic relatives of the "Maltese mushroom" Cynomorium. *Bmc Evolutionary Biology* **5**, 11.

Norton DA, De Lange PJ (1999) Host specificity in parasitic mistletoes (Loranthaceae) in New Zealand. *Functional Ecology* **13**, 552-559.

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**, 1143-1155.

O'Donnell CFJ, Christie JE, Hitchmough RA, Lloyd B, Parsons S (2010) The conservation status of New Zealand bats, 2009. *New Zealand Journal of Zoology* **37**, 297-311.

Ouborg NJ, Vergeer P, Mix C (2006) The rough edges of the conservation genetics paradigm for plants. *Journal of Ecology* **94**, 1233-1248.

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

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annual Review of Ecology Evolution and Systematics* **37**, 187-214.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**, 945-959.

Raymond M, Rousset F (1995) GENEPOP (VERSION-1.2) - Population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228.

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

Schatz GE (2009) Plants on the IUCN Red List: setting priorities to inform conservation. *Trends in Plant Science* **14**, 638-642.

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

Segarra-Moragues JG, Palop-Esteban M, GonzÁLez-Candelas F, CatalÁN P (2005) On the verge of extinction: genetics of the critically endangered Iberian plant species, Borderea chouardii (Dioscoreaceae) and implications for conservation management. *Molecular Ecology* **14**, 969-982.

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

Semagn K, Bjornstad A, Ndjiondjop MN (2006) An overview of molecular marker methods for plants. *African Journal of Biotechnology* **5**, 2540-2568.

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

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

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.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**, 264-279.

Song J, Murdoch J, Gardiner SE*, et al.* (2008) Molecular markers and a sequence deletion in intron 2 of the putative partial homologue of LEAFY reveal geographical structure to genetic diversity in the acutely threatened legume genus Clianthus. *Biological Conservation* **141**, 2041-2053.

Sosa PA, Gonzalez-Perez MA, Moreno C, Clarke JB (2010) Conservation genetics of the endangered endemic Sambucus palmensis Link (Sambucaceae) from the Canary Islands. *Conservation Genetics* **11**, 2357-2368.

Stanton S, Honnay O, Jacquemyn H, Roldan-Ruiz I (2009) A comparison of the population genetic structure of parasitic Viscum album from two landscapes differing in degree of fragmentation. *Plant Systematics and Evolution* **281**, 161-169.

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

Thorogood CJ, Hiscock SJ (2007) Host specificity in the parasitic plant Cytinus hypocistis. *Research Letters in Ecology* **2007**, 84234.

Thorogood CJ, Rumsey FJ, Harris SA, Hiscock SJ (2009) Gene flow between alien and native races of the holoparasitic angiosperm Orobanche minor (Orobanchaceae). *Plant Systematics and Evolution* **282**, 31-42.

Tian B, Liu RR, Wang LY*, et al.* (2009) Phylogeographic analyses suggest that a deciduous species (Ostryopsis davidiana Decne., Betulaceae) survived in northern China during the Last Glacial Maximum. *Journal of Biogeography* **36**, 2148-2155.

van Bunnik A, Pollock A, Somerset E*, et al.* (2007) Environment New Zealand 2007. In: *Environment New Zealand 2007*. Landcare Research, Ministry for the Environment.

Watson DM (2009) Parasitic plants as facilitators: more Dryad than Dracula? *Journal of Ecology* **97**, 1151-1159.

Wayne RK, Morin PA (2004) Conservation Genetics in the New Molecular Age. *Frontiers in Ecology and the Environment* **2**, 89-97.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Wright S (1945) Isolation by distance under diverse systems of mating. *Genetics* **30**, 571-572.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution* **11**, 413-418.

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**, 133-140.