Evolution of Australasian *Plantago* (Plantaginaceae)

A thesis
submitted in part fulfillment
of the requirements for the degree of
Master of Science in Ecology and Biodiversity
at
Victoria University of Wellington
by

Mei Lin Tay

School of Biological Sciences
Victoria University of Wellington
New Zealand

Abstract

Phylogenetic analyses using molecular data were used to investigate biogeographic and evolutionary patterns of Australasian *Plantago.* The Internal Transcribed Spacers (ITS) from nuclear DNA, *ndhF-rpl32* from chloroplast DNA and *cox1* from mitochondrial DNA were selected from a primer assay of 24 primer pairs for further phylogenetic analyses. Phylogenetic reconstruction and molecular dating of a dataset concatenated from these regions comprising 20 Australasian *Plantago* species rejected a hypothesis of Gondwanan vicariance for the Australasian group. The phylogeny revealed three independent dispersal events from Australia to New Zealand that match expected direction because of West Wind Drift and ocean currents. Following this study, a dataset with 150 new ITS sequences from Australasian *Plantago*, combined with 89 *Plantago* sequences from previous studies, revealed that the New Zealand species appear to have a recent origin from Australia, not long after the formation of suitable habitats formed by the uplift of the Southern Alps (about 5 mya), followed by radiation. The ITS phylogeny also suggests that a single migration event of alpine species to lowland habitats has occurred and that recurrent polyploidy appears to be an important speciation mechanism in the genus. Species boundaries between New Zealand *Plantago* were unclear using both morphological and molecular data, which was a result of low genetic divergences and plastic morphology. The taxonomy of several New Zealand *Plantago* species need revision based on the ITS phylogeny.

Acknowledgements

I am very grateful to Professor Phil Garnock-Jones for inspiration, funding and invaluable help with fieldwork; to Dr. Heidi Meudt for help with fieldwork, her patience with endless draft revisions, guidance and support during the course of my research. Thanks also to Dr. Peter Ritchie for guidance. I would also like to acknowledge Dr. Lara Shepherd, Dr. Vincent Woo, and the Victoria University Molecular Ecology lab for much valued advice and companionship in the lab over the past two years. Peter Beveridge, Barbara Briggs, Rewi Elliot, Kerry Ford, Peter Heenan, Rodney Lewington, Peter Lockhart, Bill Malcolm, Nancy Malcolm, Colin Ogle, Ines Schönberger, Leon Perrie, Dr. Lara Shepherd and Mike Thorsen for helped with sample collection. Barbara Briggs (Herbarium of New South Wales) and Landcare Research, Christchurch supplied information about geographic distributions of *Plantago* during the course of this research. Dr. Brian Murray at Auckland University demonstrated chromosome counting techniques for the purpose of this thesis, and Lesley Mililich provided technical support essential to my research.

For financial support, I thank the Morton family (for the Alison Morton Postgraduate Scholarship for Ecology and Marine Biology); FRST OBI Defining New Zealand's Terrestrial Biota C09X0501; the Wellington Botanical Society, VUW Post Graduate Scholarship for Masters Study and the VUW small grants scheme. I would like to acknowledge the administrative staff in the School of Biological Sciences, Victoria University of Wellington: Mary Murray; Patricia Stein; and Sandra. I also thank my family and friends for their support and patience throughout this undertaking. Finally, thank you to Zhang Te and Leah Tooman for keeping me sane.

Table of Contents

List of Tables

List of Figures

Chapter One: General Introduction

1.1 Plant phylogenetics

Ever since its conception, molecular phylogenetics has been increasingly used as a means of elucidating evolutionary histories and refining taxonomy where other means fail (e.g. the study group has ambiguous morphological data or an incomplete fossil record). Methods of obtaining molecular data (such as the quantity of sample needed and the amount of time needed to process the sample through to the sequencing stage) have consequently improved to the point that molecular data are now as easy to obtain for phylogenetic purposes as any other type of data. The use of molecular data for phylogenetic studies is more common in the animal kingdom, whereas its utility has only been recently applied to plant studies. However, molecular phylogenies are extremely useful for plants, for which processes such as interbreeding species complexes, hybridisation, introgression and polyploidy are fairly common (Soltis *et al.* 2004; Hegarty and Hiscock 2005; Vriesendorp and Bakker 2005).

In New Zealand, phylogenetic analyses using molecular data have been effectively employed for many different purposes such as dating lineages (Barker *et al.* 2007; Knapp *et al.* 2007; Perrie and Brownsey 2007), clarifying taxonomy (Albach *et al.* 2005a; Heenan *et al.* 2006; de Lange *et al.* 2007; Tripp 2007), investigating biogeography (Wagstaff *et al.* 2002; Meudt and Simpson 2006; Sanmartín *et al.* 2007; Shepherd *et al.* 2007), and investigating patterns of evolution (Meudt and Bayly 2008).

Table 1.1. Recent molecular phylogenetic studies that have investigated biogeography of New Zealand plant groups.

Family or genus	**Long Distance Dispersal (LDD) vs. vicariance**	**Sequencing regions**	**References**	**Notes**
Abrotanella	LDD	ITS, *trnK-matK*	Wagstaff *et al.* 2006	Split between sister groups in Australia and South America was dated to about 3.1 million years ago (mya).
Agathis	Vicariance	*rbcL*	Stöckler *et al.* 2002	Closest relative of *Agathis australis* (the New Zealand species) is not one of the three extant Australian species of *Agathis*. I.e. presence of this species is not LDD from Australia therefore vicariance was inferred.
Asplenium	LDD	*trnL-F*	Perrie and Brownsey 2005	Multiple dispersals to New Zealand. Estimated divergence time was 43mya for the oldest pair of New Zealand and non-New Zealand group. One frequently hybridizing austral group was found.
Atherospermataceae	LDD	*rbcL*, *rpl16*, *trnL-trnF*, *trnT-trnL*, *psbA-trnH*, *atpB-rbcL*	Renner *et al.* 2000	Short fossil record and small genetic divergence indicate that the New Zealand species probably arrived by LDD from Australia or Antarctica during the last 30-50 million years (my).
Blechnaceae	LDD	*trnL-trnF*	Shepherd *et al.* 2007	Identical sequences between species from: New Zealand and Australia, New Zealand and Chile, and Australia and Hawaii.
Caltha	Vicariance	ITS, *atpB-rbcL*, *trnL-F*	Schuettpelz and Hoot 2004	Northern Hemisphere to South America and then to Australia/New Zealand.

Table 1.1. Recent molecular phylogenetic studies that have investigated biogeography of New Zealand plant groups.

Family or genus	**Long Distance Dispersal (LDD) vs. vicariance**	**Sequencing regions**	**References**	**Notes**
Chloranthaceae	LDD	*rbcL*, *rpl20-rps12*, *trnL*, *trnL-F*	Zhang and Renner 2003	Estimated time of divergence among extant *Ascarina* was 18-19mya.
Coriaria	LDD	*rbcL*, *matK*	Yokoyama *et al.* 2000	LDD was inferred from North America to the Pacific Islands, and from New Zealand to South America.
Craspedia	LDD	ITS, ETS, *psbA-trnH*	Ford *et al.* 2007	Low sequence divergence between Australia and New Zealand (1.5, 1.7, 3.0%; respectively). One trans-Tasman dispersal from Australia to New Zealand, followed by species radiation.
Elaeocarpaceae	LDD	ITS, *trnL-trnF*	Crayn *et al.* 2006	LDD for *Aristotelia* (New Zealand and Australian species form sister clades, with South American species sister to the Australasian clade). Divergence of South American and New Zealand lineages were dated to 24-27 and 3 mya, respectively.
Gentianella	LDD	ITS, *matK*, morphology	von Hagen and Kadereit 2001	Australian and New Zealand species probably dispersed only once and probably from South America (suggested by morphology; the relationships were not resolved with molecular data). New Zealand and Australian clade diverged about 2.7 mya.
Gentianella	LDD	ITS 1, ITS 2	Glenny 2004	*Gentianella* appears to have arrived in New Zealand from South America either once or twice, and probably dispersed once to Australia. First arrived in the South Island with subsequent northward expansion.

Table 1.1. Recent molecular phylogenetic studies that have investigated biogeography of New Zealand plant groups.

Family or genus	**Long Distance Dispersal (LDD) vs. vicariance**	**Sequencing regions**	**References**	**Notes**
Gleicheniaceae	LDD	*rbcL*, *trnL-F*	Perrie *et al.* 2007	All three genera have undergone LDD to/from New Zealand less than 55mya. Tasmanian and New Zealand *Gleichenia* share identical *rbc*L sequences.
Gnaphalieae	LDD	ITS	Breitwieser *et al.* 1999	Most Gnaphalieae groups have relatively recent arrival in New Zealand, followed by rapid radiation in the group and have close affinities with Australian and New Guinean taxa.
Gunnera	Vicariance	ITS, *rbcL*, *rps16*	Wanntorp and Wanntorp 2003	Vicariance of the group, with recent LDD of one species from New Zealand to Tasmania.
Herbertus	Short distance dispersal	ITS, *trnL-F*	Feldberg *et al.* 2007	Distribution cannot be explained by Gondwanan vicariance. Could be a combination of short distance dispersal and several rare LDDs.
Korthalsella	LDD	ITS, *trnL-F*	Molvray *et al.* 1999	Possibly from Malesia (where sister genus is found) outward. Direction of LDD within Australasia was unclear.
Lyallia and *Hectorella*	LDD	*rbcL*, *trnK-matK*	Wagstaff and Hennion 2007	Divergence of sister genera from sub-Antarctic Islands and New Zealand was estimated to be about 18.6 mya.
Microseris	LDD	AFLP	Vijverberg *et al.* 2002	Dispersals in south-east Australia, Tasmania, New Zealand, and the Australian mainland. Morphological differentiation is not evident in the nuclear DNA.

Table 1.1. Recent molecular phylogenetic studies that have investigated biogeography of New Zealand plant groups.

Family or genus	Long Distance Dispersal (LDD) vs. vicariance	Sequencing regions	References	Notes
Myosotis	LDD	ITS, *psbA-trnA, matK, ndhF*	Winkworth *et al.* 2002a	Multiple long distance dispersals events; most of these were in a direction opposite to predictions of the West Wind Drift.
Nothofagus	Vicariance	*rbcL*	Linder and Crisp 1995	Vicariance followed by extensive extinctions.
Nothofagus	LDD + vicariance	ITS , *rbcL, atpB-rbcL*	Swenson *et al.* 2001	LDD of one species to Tasmania and one to New Zealand. Several species in New Zealand and Australia have much less variation than expected from Gondwanan vicariance.
Nothofagus	LDD + vicariance	*atpB-psaI, trnL-trnF*	Knapp *et al.* 2005	Vicariance can explain some transoceanic relationships among *Nothofagus* but the relationships between the trans-Tasman species (in *Lophozonia* and *Fucospora*) can only be explained by mid- to late-Tertiary transoceanic dispersal.
Oreobolus	LDD	ITS, *trnL, trnL-F*	Chacón *et al.* 2006	Australasian (and Malesian, which is the closest relative) relationships unclear but hypothesized LDD from Australasia/Malesia to Hawaii, and from Australia to South America.
Oreomyrrhis	LDD	ITS	Chung *et al.* 2005	ITS phylogeny shows that the New Zealand *Oreomyrrhis* group is closely related to a group that is mostly distributed in the Northern Hemisphere. A molecular clock estimate is fairly recent (about 1.06mya).

Table 1.1. Recent molecular phylogenetic studies that have investigated biogeography of New Zealand plant groups.

Family or genus	Long Distance Dispersal (LDD) vs. vicariance	Sequencing regions	References	Notes
Ourisia	LDD	ITS, ETS, *matK*, *rps16*	Meudt and Simpson 2006	Tasmanian species are sister to a New Zealand clade. Ancestor dispersed to Australasia from South America.
Pachycladon	LDD	ITS	Heenan *et al.* 2002	Recent dispersal between New Zealand and Tasmania (1 species is endemic to Tasmania, while all others endemic to South Island, New Zealand). Origin of *Pachycladon* was estimated to be about 1.0 to 3.5 mya.
Pittosporum	LDD	ITS	Gemmill *et al.* 2002	Mostly about the Hawaiian lineages but LDD is suggested for the group, which includes one New Zealand species. Relationships to and from New Zealand are unclear.
Plagiochila	Vicariance cannot be refuted.	ITS, *rps4*	Heinrichs *et al.* 2006	Current distribution is a mixture of short distance dispersal, rare long distance events and extinction/recolonization as a result of climate changes.
Polystichum	LDD	*rps4-trnS*, AFLP	Perrie *et al.* 2003	Appears to be from Australia to New Zealand within the last 20 my.
Proteaceae	LDD + vicariance	*rbcL*, *atpB*, *atpB-rbcL*	Barker *et al.* 2007	Gondwanan vicariance with more recent LDD. Split between New Zealand and Australian species was dated to about 45.4 mya.
Pteridophytes	LDD	*rbcL*	Perrie and Brownsey 2007	Molecular dating shows that most ferns have undergone LDD. Many of the New Zealand and non New Zealand pteridophyte pairs had relatively small differences.

Table 1.1. Recent molecular phylogenetic studies that have investigated biogeography of New Zealand plant groups.

Family or genus	**Long Distance Dispersal (LDD) vs. vicariance**	**Sequencing regions**	**References**	**Notes**
Ranunculus	LDD	ITS, JSA (chloroplast)	Lockhart *et al.* 2001	Dispersal from New Zealand to Australia and to New Guinea. Dispersal origin of New Zealand species is unclear from but the group arose about 5mya.
Scaevola	LDD	ITS	Howarth *et al.* 2003	New Zealand species (shared by Tonga) is nested within an Australian lineage.
Scleranthus	LDD	ITS 1, ITS 2	Smissen *et al.* 2003	Lack of ITS sequence divergence between Australian and New Zealand populations of two species. Divergence between Australasian and Eurasian clade was estimated to be less than 35 mya.
Sophora	LDD	*atpB-rbcL*	Hurr *et al.* 1999	Long distance dispersal from Tuvalu to Lord Howe Island to New Zealand and across the South Pacific within the last 2-5 my.
Stylidiaceae	LDD	ITS, *rbcL*	Wagstaff and Wege 2002	Two dispersal events: one from Australia and the other from either South America or Australia.
sunflower alliance of families	Vicariance	*rbcL*	Bremer and Gustafsson 1997	Simple molecular dating estimated the age of the sunflower alliance to be about 96 mya.
Tetrachondra	LDD	*rbcL*	Wagstaff *et al.* 2000	Either long-distance dispersal of *Tetrachondra hamiltonii* to New Zealand, or short distance dispersal via stepping stones. East to west direction of LDD, opposite to the direction predicted by West Wind Drift.

Table 1.1. Recent molecular phylogenetic studies that have investigated biogeography of New Zealand plant groups.

Family or genus	Long Distance Dispersal (LDD) vs. vicariance	Sequencing regions	References	Notes
Veronica (*Hebe*)	LDD	ITS	Wagstaff and Garnock-Jones 1998	Recent origin and rapid radiation of the New Zealand *Hebe* group. This is followed by dispersal to the Chatham Islands, and South America (at least twice).
Veronica (*Chionohebe*, *Parahebe*)	LDD	ITS	Wagstaff and Garnock-Jones 2000	LDD from New Zealand to Australia and New Guinea, and also from New Zealand to South America and Rapa Island. Origin of New Zealand species is equivocal between Australia and Northern Hemisphere
Veronica (*Hebe*)	LDD	ITS, *rbcL*	Wagstaff *et al.* 2002	One ancestor followed by diversification in New Zealand. Two LDD events to South America, at least one to Australia and one to New Guinea; plus six others across the Pacific, including to Chatham Island and the subantarctic islands.

1.1.1 Genetic markers for plant phylogenetics

When working on molecular phylogenies, choosing genetic markers is an important process. Genetic markers need to be chosen based on their ability to differentiate between taxa at the level that the research requires. Currently, the most common genetic marker used for phylogenetic studies in plants is the Internal Transcribed Spacer (ITS) region, found in the nuclear genome of plants, fungi and animals. This refers to the region of 18S, 5.8S and 26S ribosomal genes, along with two internal transcribed spacers. It is widely applied in phylogenetic studies, from fungi to higher land plants (White *et al.* 1990; Álvarez and Wendel 2003) and recently to animals (Jorgensen *et al.* 2007; Kuriiwa *et al.* 2007). In fact, it was found that out of 244 papers published between the years 1998-2004, 66% utilized ITS data and 34% of phylogenies were published based on ITS data alone (Álvarez and Wendel 2003). The trend of utilizing ITS as a marker for reconstructing plant phylogenies is also found in New Zealand plant studies (Table 1.1), where 27 of 44 (61%) recent phylogenetic studies investigating biogeography used ITS as a marker.

The ITS region has various advantages as a marker, such as high genetic variation for most groups, concerted evolution of the region, and low functional constraints (Álvarez and Wendel 2003). In addition, the entire ribosomal DNA array region is easily amplified with universal primers (Hillis and Dixon 1991) and as a consequence of frequent use, there is a wide array of ITS sequences readily available on Genbank (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=nucleotide). However, markers from other regions may prove to be more useful than ITS, especially at lower taxonomic levels (Mort *et al.* 2007; Shaw *et al.* 2007). Markers from independent regions can also be used in combination with each other as long as they have gene

histories that do not conflict. Increasing sequence length in a dataset may allow phylogenetic trees to be more accurate, whereas sampling sites from independent regions increases the power of the phylogenetic inference (Cummings *et al.* 1995). Additionally, the use of more than one (independent) marker will allow visualization of processes such as hybridisation, introgression, reticulation and incomplete lineage sorting if these have occurred in the past (Vriesendorp and Bakker 2005). For events such as hybridisation, parental lineages may also be revealed based on the phylogenies of the different markers (e.g. Albach and Chase 2004; Smissen *et al.* 2004).

Markers from different genomes are useful in this respect because the nuclear DNA (nrDNA) is independent of the organellar DNA and they have different mechanisms of inheritance. Moreover, given that organellar DNA is generally maternally inherited and non-recombining in most plant groups, DNA from the chloroplast DNA (cpDNA) and mitochondrial DNA (mtDNA) may be able to show genetic relatedness and geographic structure of genetic variation better than nrDNA, which is biparentally inherited and undergoes recombination (Mort *et al.* 2007). Recently, the development of markers other than ITS have shown that many previously unexplored regions are useful for plant phylogenetics (e.g. chloroplast markers developed by Shaw *et al.* 2007). While mtDNA is commonly used in animal studies, it is usually not a suitable marker for plant phylogenetics because it is known to evolve at an extremely slow rate (Wolfe *et al.* 1987) and is therefore highly conserved even at the species level. However, an unusually high rate of mtDNA evolution has been found in a few plant groups, including *Plantago* (Cho *et al.* 2004), *Silene* (Mower *et al.* 2007) and Geraniaceae (Parkinson *et al.* 2005). Therefore, mtDNA may have some utility for

phylogenetic analyses within these genera and possibly other plant groups. At the population level, studies often use microsatellites, a string of one to six nucleotide sequences that are repeated multiple times (Zane *et al.* 2002), because they have frequent length polymorphism when compared among individuals in a population.

1.2 Biogeography of Southern Hemisphere plants

Molecular data have been increasingly used to reconstruct phylogenies for investigating plant biogeography. Molecular phylogenies allow for large-scale studies that would otherwise be too complex using other types of characters, such as morphology (e.g. Renner 2004; Sanmartín and Ronquist 2004; Sanmartín *et al.* 2007). For example, by plotting distribution information on 23 phylogenies for different Southern Hemisphere plant groups, Sanmartín *et al.* 2007 were able to infer that plants in the Southern Hemisphere are more likely to disperse in an eastward direction between Australia and South America, whereas they are more likely to disperse in a westward direction between South America and New Zealand.

Plant biogeography in the Southern Hemisphere has garnered much attention (Table 1.1, Setoguchi *et al.* 1998; Muñoz *et al.* 2004; Sanmartín and Ronquist 2004; Sytsma *et al.* 2004; Sanmartín *et al.* 2007). Since the conception of plate tectonics in the 1960's, distribution of Southern Hemisphere flora has been believed to be the result of vicariance from the supercontinent Gondwana (Pole 1994; de Queiroz 2005; McGlone 2005), which broke up to form the landmasses in the Southern Hemisphere. However, studies using molecular data suggest that recent long distance dispersal in Southern Hemisphere plant groups are more prevalent than previously thought (Winkworth *et al.* 1999; Winkworth *et al.* 2002b; de Queiroz 2005; McGlone 2005) (see Table 1.1).

The New Zealand landmass has been isolated by a distance of at least 1500km from its closest landmass, Australia, after the break-up of Gondwana 80 million years ago (mya) (Cooper and Cooper 1995). Recent phylogenetic analyses (Table 1.1) have revealed that divergence dates for many plant groups on different continents are too recent to be explained by vicariance (e.g. Wagstaff and Garnock-Jones 1998; Hurr *et al.* 1999; Lockhart *et al.* 2001; Radford *et al.* 2001; Swenson *et al.* 2001; von Hagen and Kadereit 2001; Heenan *et al.* 2002; Zhang and Renner 2003; Perrie and Brownsey 2005; Barker *et al.* 2007; Perrie and Brownsey 2007; Perrie *et al.* 2007). Evidence of long distance dispersal has also been found in plant lineages that are considered typical examples of Gondwanan relicts, such as *Nothofagus* (Swenson *et al.* 2001; Knapp *et al.* 2005). Some authors have gone so far as to say that the entire New Zealand flora has arrived by long distance dispersal (Pole 1994) but there is still evidence that some plant groups have origins that date back to a Gondwanan origin, e.g. *Agathis* (Stöckler *et al.* 2002; Knapp *et al.* 2007). Thus, the origins of the New Zealand flora may be a mixture of older vicariance and more recent long distance dispersal events.

Within Australasia, dispersal from Australia to New Zealand is commonly found in molecular studies (Wagstaff and Garnock-Jones 1998; Wagstaff *et al.* 1999; Wagstaff and Wege 2002; Ford *et al.* 2007). Dispersal in this direction is consistent with the prevailing West Wind Drift (Raven 1973; Winkworth *et al.* 2002b; Muñoz *et al.* 2004; Sanmartín *et al.* 2007). While not as common, there is also evidence of dispersal in the opposite direction (New Zealand to Australia) (Wagstaff and Garnock-Jones 2000; Lockhart *et al.* 2001; Wagstaff *et al.* 2002; Wanntorp and Wanntorp 2003). Other dispersal patterns can also be found, e.g. in *Veronica*, there has been evidence of

dispersals from Australia to New Zealand; New Zealand to Australia; New Zealand to Papua New Guinea; New Zealand to South America; and New Zealand to Rapa (Wagstaff *et al.* 2002). Additionally, New Zealand plant groups are often found to have established in New Zealand following one single long distance dispersal event (Wagstaff and Garnock-Jones 1998; Winkworth *et al.* 1999; Perrie *et al.* 2003; Albach *et al.* 2005b; Meudt and Simpson 2006). While molecular phylogenies have been able to reveal biogeographic patterns in many New Zealand plant groups, the biogeography of many more lineages remain to be tested.

1.3 Patterns of evolution in New Zealand plants

Overall evolutionary patterns within New Zealand plants are not well studied and in particular, the evolution of alpine species is not well-known (Lockhart *et al.* 2001). Evidence for rapid radiation in many New Zealand plant groups has been found in many molecular phylogenetic studies to date (Wagstaff and Garnock-Jones 1998; Winkworth *et al.* 1999; Heenan *et al.* 2002; Perrie *et al.* 2003; Murray *et al.* 2004; Albach *et al.* 2005b; Meudt and Simpson 2006). These radiations are often associated with speciation after recent long distance events (Breitwieser *et al.* 1999), adaptation to new habitats after the uplift of the Southern Alps (Wagstaff and Garnock-Jones 1998; Lockhart *et al.* 2001; Trewick and Morgan-Richards 2005), and climate fluctuations (Raven 1973; Lockhart *et al.* 2001; Winkworth *et al.* 2002a). Speciation in New Zealand plant groups may also have followed separation of populations due to uplift of Southern Alps (Haase *et al.* 2007), adaptation to different habitats (McBreen and Heenan 2006), disruption during glacial periods, or displacement along the alpine fault (Haase *et al.* 2007). Cook Strait, which separates the North and South Islands of New Zealand, could act as a barrier as well, although there may have been land

bridges connecting the North and South Islands during the glacial cycles of the Pliocene (Lewis *et al.* 1994). Older lineages may also have experienced population fragmentation as a result of almost total submersion of the New Zealand landmass during the Oligocene (Cooper and Cooper 1995). Additionally, a recurrent pattern within New Zealand is the dispersal of plants from the South Island to the North Island (e.g. Lockhart *et al.* 2001; Meudt and Simpson 2006).

Hybridisation is common in many New Zealand plant groups (Heenan *et al.* 2001; Lockhart *et al.* 2001; Albach and Chase 2004; Smissen *et al.* 2004), along with recurrent polyploidy (Groves and Hair 1971). Recurrent polyploidy is probably associated with hybridisation because it allows instant speciation where one or several polyploid hybrids arise sympatrically with both of (or one of) the parent populations (Petit *et al.* 1999).

1.4 Taxonomy in New Zealand plant groups

Species boundaries in New Zealand plants may be unclear using morphological characters due to recent or incomplete speciation associated with recent and rapid radiations. This often results in taxonomic problems. Many New Zealand groups are large and have high morphological variation both within and among species, such as *Myosotis* (Winkworth *et al.* 1999; Winkworth *et al.* 2002a) and *Gnaphalieae* (Smissen *et al.* 2004). However, these, along with many other New Zealand plant groups often have low genetic diversity (Hurr *et al.* 1999; Winkworth *et al.* 1999; Heenan *et al.* 2002; Mitchell and Heenan 2002; Vijverberg *et al.* 2002; Winkworth *et al.* 2002a), which makes species delimitations difficult.

Molecular phylogenies have been used for improving plant taxonomy within New Zealand. For example, molecular data revealed that the five or six genera of the *Hebe* complex have evolved within *Veronica*, and are therefore paraphyletic (Albach and Chase 2001). Circumscription of *Veronica* was subsequently enlarged to include all Southern Hemisphere *Hebe* species (Garnock-Jones *et al.* 2007). Molecular phylogenies can also be used to efficiently investigate species complexes (e.g. two species of *Crassula* were found to have frequent interbreeding and could not be separated based on genetic, morphological or karyological data (de Lange *et al.* 2007).

1.5 Introduction to study group: *Plantago*

In this thesis, a phylogeny of *Plantago* is reconstructed using molecular data. *Plantago* is a large genus (ca. 210 species) with a worldwide distribution in the family Plantaginaceae (Rahn 1996). The plants are mostly small rosettes but also form woody shrubs in the Hawaiian and Juan Fernández Islands (Rock 1920). Plants in this genus are wind-pollinated. The genus has three basic chromosome numbers, x = 4, 5, and 6 (Rahn 1996; Dhar *et al.* 2006); diploid chromosome numbers range from $2n$ = 8 (*P. ovata*) to $2n$ = 96 (*P. correae* and *P.* "sylvester").

Chemical properties and diverse breeding systems (e.g. self-incompatible or self-compatible; dichogamy; cosexuality; gynodioecy or dioecy; unisexual females or hermaphrodites) within the genus have captured the attention of many researchers (see Wolff and Schaal 1992; VanDijk and Bakx-Schotman 1997; Rønsted *et al.* 2000; Squirrell and Wolff 2001; Hale and Wolff 2003; Rønsted *et al.* 2003; Nilsson and Ågren 2006; Nilsson *et al.* 2006). Recently, *Plantago* has also garnered attention as a

result of the high evolutionary rate of mtDNA found in the genus (Cho *et al.* 2004; Bakker *et al.* 2006).

1.5.1 Previous phylogenetic studies

One phylogenetic study using morphological characters and two using genetic sequences have been performed on the genus so far.

Morphological phylogeny

Rahn (1996) reconstructed a phylogeny of the genus using morphological characters. Most Australasian species were included (one New Zealand and four Australian species were not included). Based on the morphological phylogeny, the ca. 210 *Plantago* species were classified into six subgenera, including *Littorella* (previously described as a genus). This morphological phylogeny suggested that the Australasian species do not form a monophyletic group. The Australasian species were grouped with species from South America, Asia, Europe, New Guinea and Tonga Islands (Rahn 1996).

Molecular phylogenies

Ronsted *et al.* (2002) obtained ITS and *trnL-F* sequences from 57 *Plantago* species (including two New Zealand species: *P. spathulata*, *P. raoulii*; and one from Australia: *P. debilis*). Phylogenetic analyses revealed that subg. *Albicans* was paraphyletic and circumscription of subg. *Psyllium* was revised to include those species. In a combined ITS and *trnL-F* phylogeny, the Australasian species did not form a monophyletic group and were separated by *P. stauntonii* from the Amsterdam & St. Paul Islands (located in the Indian Ocean).

Finally, Hoggard *et al.* (2003) used ITS data from 23 *Plantago* species (including one New Zealand species: *P. triandra*; and five Australian species: *P. tasmanica*, *P. daltoni*, *P. euryphylla*, *P. hispida*, *P. paradoxa*) to investigate whether *Littorella* should be included within *Plantago* or considered a separate genus. *Littorella* was found to form the sister clade to a clade of all the other *Plantago* species. The *Plantago* phylogeny (i.e., excluding *Littorella*) resolved in this study was similar to that resolved by Rønsted *et al.* 2002.

1.5.2 Biogeography of *Plantago*

Biogeography is interesting within *Plantago* because the genus is widely distributed. Some species are endemic to certain areas like the Australasian species (there are even species that are native to recently formed islands, such as Amsterdam & St. Paul Islands), whereas many species are cosmopolitan (see Rahn 1996). Based on a morphological phylogeny, Rahn (1996) invoked a vicariance model with subsequent mass extinctions to explain *Plantago* distribution in the Southern Hemisphere. In contrast, Rønsted *et al.* (2002) suggested a prevalence of long distance dispersal worldwide based on the molecular phylogeny. However, the molecular phylogenies have only included a maximum of six Australasian species (out of a total 32). South American *Plantago* are represented by 13 species that have published sequences. Thus, inclusion of Australasian *Plantago* species in the phylogeny is crucial for understanding biogeographic patterns of Southern Hemisphere plants. A molecular phylogeny including more Australasian species can be used to compare the biogeographic hypotheses of the Southern Hemisphere species as suggested by Rahn (1996) and Rønsted *et al.* (2002) and may also shed light on several lineages of the phylogeny that are unclear.

1.5.3 Taxonomy of *Plantago* in New Zealand

Plantago is placed in the family Plantaginaceae, which was traditionally classified with three genera: *Bougueria* Decne., *Littorella* P. Bergius and *Plantago* L. These were combined to form a monogeneric family based on morphology (Rahn 1996), although a molecular phylogeny has since shown evidence for recognizing *Littorella* at the generic level (Hoggard *et al.* 2003). Individuals within the genus have high morphological plasticity, which results in difficulties when circumscribing the plants into taxonomic groups, particularly at the species level (Sykes in Webb *et al.* 1988, Rahn 1996). While relationships of subgenera and sections within *Plantago* are reasonably well-resolved (in Rønsted *et al.* 2002), inter-specific relationships are still largely unclear.

In this thesis, I focus on the taxonomy of *Plantago* species that are native to New Zealand. So far there have been several taxonomic treatments for the New Zealand species (e.g. Moore in Allan 1961; Sykes in Webb *et al.* 1988) but no recent, comprehensive taxonomic monograph. Eight species of *Plantago* were accepted in the most recent Flora of New Zealand: *P. raoulii*, *P. spathulata*, *P. triandra*, *P. obconica*, *P. lanigera*, *P. unibracteata*, *P. aucklandica* and *P. triantha* (Sykes in Webb *et al.* 1988). In addition, there is a purported polyploid of 2*n*=96 given the tag name *P.* "sylvester" (Groves and Hair 1971), which is still undescribed. Like the rest of the genus, morphological differences between New Zealand species are not easily defined, resulting in problems with species delimitation and identification. Thus, genetic data may be more useful in elucidating species boundaries and relationships in the New Zealand species of *Plantago*. Discrepancies between the five most recent taxonomic treatments (Hooker 1864; Cheeseman 1906; Cheeseman 1936; Moore in Allan 1961;

Sykes in Webb *et al.* 1988) represent hypotheses of species boundaries that are investigated in this study. To date, molecular phylogenetic studies in *Plantago* have included only three out of the eight New Zealand species.

1.6 Summary and research aims

In summary, obtaining DNA sequences from multiple independent genetic markers has many advantages in phylogenetic studies. For *Plantago*, Australasian species should be integrated into the molecular phylogeny in order to make inferences about the biogeography of Southern Hemisphere *Plantago* species. A molecular phylogeny including Australasian species can also be used to look at species relationships within the group, many of which are still largely unclear. Genetic data may be useful for the genus because morphological characters are extremely plastic in the group and do not appear to be very useful for species delimitations.

For this thesis, DNA sequences from the three different genomes are obtained to reconstruct a molecular phylogeny of Australasian *Plantago* species. Selection of molecular markers from the three different plant genomes is based on an assay of various primer pairs. The molecular phylogeny is then used to: 1) investigate phylogeny and biogeography of Australasian species of *Plantago*; and 2) elucidate phylogeny, evolutionary patterns and species boundaries of the New Zealand species. In particular, I investigate if the current distribution of native Australasian *Plantago* species can be explained by vicariance or long distance dispersal. While this study cannot include a taxonomic monograph due to time constraints, I address several taxonomic issues at species rank that have plagued researchers in the past.

Chapter Two: Primer pair assays for amplification of 24 nuclear, mitochondrial, or chloroplast regions in four *Plantago* and one *Veronica* (Plantaginaceae) species

Abstract

Combining DNA sequences from all three plant genomes provides a powerful phylogenetic tool because the genomes have different modes of inheritance. DNA sequences from the chloroplast and nucleus are most often used for reconstructing the evolutionary history of plants, whereas mitochondrial DNA is not often used because of extremely low evolution rates within the mitochondrial genome. The aim of this study was to find markers with good phylogenetic signals for future studies on the genus *Plantago* (Plantaginaceae). A total of 23 different sequence regions in all three genomes were tested for amplification in four *Plantago* and one *Veronica* species. This survey revealed that primers for ITS, *cox1*, *ndhF-rpl32*, *trnK-psbA*, *trnE-trnTr*, *trnLc-trnLf*, *trnC-trnD*, *rps16* (using both sets of primers tested) and *trnH-psbA* produced consistently clean amplification of their targeted regions and were of good sequence length (about 700-1k base pairs in length each). These regions are recommended for future studies in *Plantago* or related groups. DNA microsatellites developed specifically for *Plantago major, P. coronopus* and *P. intermedia* were also trialed but they did not amplify or were not genotyped successfully. Of the regions that were sequenced, ITS and *ndhF-rpl32* had the highest sequence variation between New Zealand and Australian species (an average of 2.5% for each) and were chosen for further phylogenetic analyses of *Plantago*.

2.1 Introduction

Nuclear DNA (nrDNA) sequences are the most often used molecular markers in plant phylogenetic studies, whereas chloroplast DNA (cpDNA) has recently gained recognition as a useful marker (since the 1990's). The single most commonly used genetic marker in plant molecular phylogenetics is the Internal Transcribed Spacer (ITS) region, which is located in the nrDNA. Out of 244 plant phylogenetic papers published between the years 1998-2004, it was found that 66% utilized ITS data and 34% of these presented phylogenies based on ITS data alone (Álvarez and Wendel 2003).

DNA sequence regions from cpDNA and mtDNA, however, may prove to be more useful for phylogenetic studies than some well-established nrDNA markers. Sequence regions from cpDNA and mtDNA can also be used in conjunction with nrDNA markers to improve understanding of the evolutionary history of a plant or plant groups. Utilising genetic markers from the three different genomes may reveal different lineage histories because mtDNA and cpDNA are typically uniparentally inherited, whereas nrDNA is biparentally inherited. These contrasting modes of inheritance are especially useful for plant studies, in which polyploidy, hybridisation and introgression are fairly common processes (Soltis *et al.* 2004). For example, hybridisation events have been inferred and parent species have been elucidated from conflicting signals presented by nrDNA and organelle DNA in several phylogenetic studies (Lockhart *et al.* 2001; Albach and Chase 2004; Smissen *et al.* 2004). Where there is no conflict among lineages, combining sequences from all three plant genomes is able to provide a powerful tool for phylogenetic analysis. Concatenated

datasets have increased dataset size compared to a single region, which allows building methods to construct more robust trees, whereas analyzing data from multiple independent sources has been shown to improve the power of analyses (Cummings *et al.* 1995). Analyses of concatenated datasets may thus provide further insight into an organism's evolutionary history than can be achieved by markers from the one genome alone or even several markers from the same genome.

Mitochondrial DNA (mtDNA) is rarely found to be a suitable marker for plant phylogenetic studies. Unlike mtDNA in animals, plant mtDNA is generally assumed to evolve at an extremely slow rate (Wolfe *et al.* 1987) and is therefore highly conserved even at species level. However, an unusually high rate of mtDNA evolution has recently been reported in *Plantago* (Cho *et al.* 2004) and several other plant groups such as *Silene* (Houliston and Olson 2006) and Geraniaceae (Bakker *et al.* 2006). In the case of *Plantago*, this accelerated rate even exceeds the rate of animal mtDNA evolution by an order of magnitude (Cho *et al.* 2004), which is highly suprising. More phylogenetic work needs to be carried out on the mtDNA in the group to explore this high rate of evolution.

The genus *Plantago* (Plantaginaceae) comprises more than 200 species of wind-pollinated plants with a worldwide distribution that includes many cosmopolitan species. Plants within the genus have variable chromosome numbers. Rahn (1996) reports values from from $2n = 8$ (*P. ovata*) to $2n = 96$ (*P. correae* and *P.* "sylvester") within the genus) and diverse breeding systems (such as self-compatibility, dichogamy, cosexuality, gynodioecy, dioecy, unisexual females or hermaphrodites), which have captured the attention of many researchers (see Wolff and Schaal 1992;

VanDijk and BakxSchotman 1997; Squirrell and Wolff 2001; Hale and Wolff 2003; Nilsson and Ågren 2006; Nilsson *et al.* 2006). The plants in the genus also have interesting chemical properties, which have been investigated in various studies (Rønsted *et al.* 2000; Rønsted *et al.* 2003; Kozan *et al.* 2006; Barton 2007). Several species, such as *P. ovata*, are cultivated for medicinal uses (Dagar *et al.* 2006). Recent studies that investigate the phylogeny of *Plantago* using the nrDNA Internal Transcribed Spacer (ITS) and the cpDNA *trnL-F* region have created a framework for the evolutionary history of the group. However, these studies included sequences for only three or six Australasian species (Rønsted *et al.* 2002; Hoggard *et al.* 2003), respectively.

In this study, the usefulness of several primer pairs from each of the three plant genomes is tested for plants in the genus *Plantago*. The study group comprises two New Zealand species, one Australian and one cosmopolitan *Plantago* species. A species from a different genus in the same family, *Veronica stricta* from New Zealand, was included in the assays to establish the applicability of the primer pairs outside of the genus. The unexpected rate of evolution in *Plantago* mtDNA may be useful for phylogenetic studies involving the genus and warrants further investigation. For this reason, mtDNA primers have been included in this study. Several microsatellite primers were also included. Microsatellites are regions in any given genome, which have multiple repeats of a string of one to six nucleotides (Zane *et al.* 2002). Microsatellite regions have frequent length polymorphism, and are more variable than other types of sequence markers at the population level.

Table 2.1. Primer sequences and references for the primer pairs included in this study.

	Region	Genome	Primer sequences	Reference	Designed for
1	ITS	Nuclear	*ITS28CC* CGCCGTTACTAGGGGAATCCTTGTAAG *ITS5* GGAAGTAAAAGTCGTAACAAGG	Wagstaff and Garnock-Jones 1998 White *et al.* 1990	Universal
2	*CYC*	Nuclear	Primers are unpublished	Wang, UBC (pers. comm.)	N/A
3	*CAM*	Nuclear	*CAMX1F* AGCCTNTTCGACAAGGATGG *CAMX2R* AGTGANCGCATCACAGTT	Strand *et al.* 1997	angiosperms
4	Waxy	Nuclear	*Waxy7F* GYYTTSTGCATCCACAACATTGC *Waxy13R* GGAGTGGCRACGTTTTCCTT	Olmstead *et al.* unpubl.	Lamiales
5	*CHS*	Nuclear	*CHSX1F* AGGAAAAATTCAAGCGCATG *CHSX2RN* TTCAGTCAAGTGCATGTAACG	Strand *et al.* 1997	angiosperms
6	LFY	Nuclear	*LFY.F2* CGTGGSAAAAAGAAYGGYYTDGATTA *LFY.R3* CATTTTDGGYTTGTTKATGTA	Howarth and Baum 2005	angiosperms
7	*G3pdH*	Nuclear	*GPDX7F* GATAGATTTGGAATTGTTGAGG *GPDX9R* AAGCAATTCCAGCCTTGG	Strand *et al.* 1997	angiosperms
8	*cox1*	Mitochondrial	*cox1F4* GGATATCTAGGYATGGTTTATGC *cox1R3* AAGCTGGAGGACTTTGTAC	Cho *et al.* 2004 (pers. comm.)	*Plantago*
9	*nad1*	Mitochondrial	*nad1b* GCATTACGATCTGCAGCTCA *nad1c* GGAGCTCGATTAGTTTCTGC	Demesure *et al.* 1995	land plants
10	NIA3	Mitochondrial	*NIA-i3F* AARTAYTGGTGYTGGTGYTTYTGGTC *NIA-i3R* GAACCARCARTTGTTCATCATDCC	Howarth and Baum 2002	*Scaevola*, and angiosperms
11	*ndhF-rpl32*	Chloroplast	*ndhF* GAAAGGTATKATCCAYGMATATT *rpl32-R* CCAATATCCCTTYYTTTTCCAA	Shaw *et al.* 2007	angiosperms
13	*rpl32-trn*L	Chloroplast	*trnL* CTGCTTCCTAAGAGCAGCGT *rpl32-F* CAGTTCCAAAAAAACGTACTTC	Shaw *et al.* 2007	angiosperms
12	*trnK-psbA*	chloroplast	*trnK3F* CCGACTAGTTCCGGGTTCGAAT *PSBAR* CGCGTCTCTCTAAAATTGCAGTCAT	Winkworth *et al.* 2002a	*Myosotis*
14	*trnE-trnT*r	chloroplast	*trnE* GCC TCC TTG AAA GAG AGA TG *trnT-r* TAC CAC TGA GTT AAA AGG GC	Doyle *et al.* 1992	Poaceae

Table 2.1. Primer sequences and references for the primer pairs included in this study.

	Region	Genome	Primer sequences	Reference	Designed for
15	*trn*Lc-*trn*Lf	chloroplast	*c* CGAAATCGGTAGACGCTACG *f* ATTTGAACTGGTGACACGAG	Taberlet *et al.* 1991	land plants
16	*trnC-trnD*	chloroplast	*trnC* CCAGTTCAAATCTGGGTGTC *trnD* GGGATTGTAGTTCAATTGGT	Demesure *et al.* 1995	land plants
17	*rps16*	chloroplast	*rps16F* AAA CGA TGT GGT ARA AAG CAA C *rps16R* AAC ATC WAT TGC AAS GAT TCG ATA	Shaw *et al.* 2005	Angiosperms
18	*rbcL*	chloroplast	*aF* ATGTCACCACAAACAGAGACTAAAGC *cR* GCAGCAGCTAGTTCCGGGCTCCA	Hasebe *et al.* 1994	Ferns
19	*trnH-psbA*	chloroplast	*trnH2* CGCGCATGGTGGATTCACAATCC *psbAF* GTTATGCATGAACGTAATGCTC	Tate 2002 Sang *et al.* 1997	*Tarasa* and *Paeonia*
20	*rps16*	chloroplast	*rpsF* GTGGTAGAAAGCAACGTGCGACTT *rpsR2* TCGGGATCGAACATCAATTGCAAC	Oxelman *et al.* 1997	angiosperms (designed from *Sileneae*)
21	*trn*Lb-*trn*Ta	chloroplast	*a* CATTACAAATGCGATGCTCT *b* TCTACCGATTTCGCCATATC	Taberlet *et al.* 1991	land plants
22	PCM07	microsatellite	*PCM07F* GAGCGTCCGATCTAAACGAT *PCM07R* GACTAACGTGCATTGCCTAGC	Koorevaar *et al.* 2002	*Plantago coronopus*
23	PM6	microsatellite	*PM6F* ATATGAATTAGCCAACAAA *PM6R* CCAGCTCCAAGTCAAAGTA	Squirrell and Wolff 2001	*P. major*, *P. intermedia*
24	Jpm11	microsatellite	*Jpm11F* ATGGCATGAGTGGACCAGAT *Jpm11R* AAAAGCTGGGCACCTACAAA	Squirrell and Wolff 2001	*P. major*, *P. intermedia*

The primers used here have been previously published (Table 2.1) and were chosen based on their previous usefulness in other plant phylogenetic studies. Many of these were designed for universal use throughout land plants, including ITS regions (White *et al.* 1990; Wagstaff and Garnock-Jones 1998), *trnL-trnF* (Taberlet *et al.* 1991) and the *rps16* intron (Shaw *et al.* 2005). Other primers have been developed for specific group studies, such as *rbcL* for ferns (Hasebe *et al.* 1994), *trnE-trnTr* for Poaceae (Doyle *et al.* 1992) and *trnK-psbA* for *Myosotis* (Winkworth *et al.* 2002a). The microsatellite markers have been designed for specific species within *Plantago*; PCM07 for *P. coronopus* (Koorevaar *et al.* 2002), and JPM11 and PM6 for *P. major* and *P. intermedia* (Squirrell and Wolff 2001). Prior to this study, all of the primer pairs (with the exception of ITS and *trnL*c-*trnL*f) have not been amplified in any native New Zealand or Australian *Plantago* species. Several regions with successful amplification are then scrutinized in more detail by comparing amplified length and sequence variation between the *Plantago* samples. The regions which show successful amplification and good sequence variation can be recommended for use in future studies.

2.2 Materials and Methods

2.2.1 Study group

Within *Plantago*, two native New Zealand species (*P. spathulata* subsp. *spathulata* and *P. triandra* subsp. *triandra*), one native Australian species (*P. euryphylla*) and one cosmopolitan species that has established in New Zealand (*P. lanceolata*) were included in this study. *P. lanceolata* and *P. spathulata* subsp. *spathulata* represent the furthest

divergence between taxa in the genus and are placed in distant clades in the phylogeny (Rønsted *et al.* 2002), thereby covering a wide range of the molecular variation in the genus. For the microsatellites, the *Plantago* species that the primers were designed for were included as positive controls. Thus, *P. coronopus* was included as the positive control for PCM07, and *P. major* was included as the positive control for both JPM11 and PM6.

In addition, *Veronica stricta* was included to see if the primers had phylogenetic utilities beyond the genus *Plantago*, especially for primers that were designed for specific use, e.g. the microsatellite primer pairs. Sample collection details for the species included in this study are given in Appendix I (A).

2.2.2 Molecular methods

Total DNA was extracted from silica-dried leaves or whole plants with either the QIAGEN DNeasy Plant Mini Extraction Kit (Hilden, Germany) or a cetyltrimethylammonium bromide (CTAB) extraction method modified from Doyle and Doyle (1990). PCR amplification was performed using Eppendorf Mastercycler ep gradient S (Hamburg, Germany) in a final volume of 25 µl of the following: 16.35 µl water, 10X ThermoPol reaction buffer (10 mM KCl, 10 mM $(NH_4)2SO_4$, 20 mM Tris-HCl (pH8.8), 2 mM $MgSO_4$, 0.1% Triton X-100)(New England BioLabs), 0.4 mg/mL BSA, 250 µmol dNTPs, 10 ρmol each primer, 0.75 U of *Taq* DNA polymerase (New England BioLabs) and 0.4 µl DNA template. Products were amplified with a thermocycling profile of an initial 3 minutes at 95°C; then 35 cycles of 30 seconds at

Table 2.2. PCR amplification and sequencing results for the four *Plantago* and one *Veronica* species included in this study.

	Region	*Veronica stricta*	*Plantago lanceolata*	*Plantago euryphylla*	*Plantago spathulata*	*Plantago triandra*	Annealing temperature	No. of base pairs	Sequencing
1	ITS	**S**	**S**	**S**	**S**	**S**	50℃	~800	+
2	*CYC*	M	—	M	M	M	40℃	~500-600	+
3	*CAM*	M	—	—	—	—	40℃	~200-1300*	N/A
4	Waxy	—	—	—	—	—	N/A	?	N/A
5	*CHS*	—	—	—	—	—	N/A	~150-1200*	N/A
6	LFY	—	—	—	—	—	N/A	~500*	N/A
7	*G3pdH*	—	—	—	—	—	N/A	~150-1600*	N/A
8	*cox1*	**S**	**S**	**S**	**S**	**S**	55℃	~800	+
9	*nad1*	**S**	**S**	—	—	—	50℃	~200	N/A
10	NIA3	—	—	—	—	—	N/A	~85-1646*	N/A
11	*ndhF-rpl32*	**S**	**S**	**S**	**S**	**S**	50℃	~900-1000	+
13	*rpl32-trn*L	**S**	**S**	S/M	**S**	S/M	50℃	~800-900	+
12	*trnK-psbA*	**S**	**S**	**S**	**S**	**S**	50℃	~300	+
14	*trnE-trn*Tr	**S**	**S**	**S**	**S**	**S**	50℃	~500-700	+
15	*trn*Lc-*trn*Lf	**S**	**S**	**S**	**S**	**S**	50℃	~1000	+
16	*trnC-trnD*	**S**	**S**	**S**	**S**	**S**	50℃	>2000	+
17	*rps16*	**S**	**S**	**S**	**S**	**S**	50℃	~1000	N/A
18	*rbcL*	S/M	S/M	S/M	**S**	S/M	50℃	~1400	N/A
19	*trnH-psbA*	**S**	**S**	**S**	**S**	**S**	50℃	~400	+
20	*rps16*	**S**	**S**	**S**	**S**	**S**	50℃	~1000	N/A
21	*trn*Lb-*trn*Ta	**S**	—	—	—	—	50℃	~700*	N/A
22	PCM07	M	M	M	M	M	55℃	~190*	+
23	PM6	—	—	—	—	—	N/A	~100*	N/A
24	JPM11	M	M	M	M	M	55℃	~200	0

S = single banded amplified product, M = multiple banded amplified product, S/M = multiple banded product but band of amplified region much brighter, — = no amplified product, * estimated length from previous studies, + = successful sequencing or genotyping, 0 = sequencing or genotyping failed, N/A = data not available (i.e. not sequenced).

95°C, 30 seconds at 50°C, and 2 minutes at 72°C; and a final extension period of 10 minutes at 72°C on an Eppendorf Mastercycler ep gradient S (Hamburg, Germany). The same PCR protocol was used for all primer pairs but some of the primer pairs needed optimization of annealing temperatures for successful single band amplification of the targeted region (Table 2.2).

The microsatellite regions were amplified using the same PCR cocktail as above, with the following modifications: 2 ρmol of the forward primer with an M13 tail was added, and 8 ρmol of a reverse complement of the tail labeled with a fluorescent 6FAM dye was added to the cocktail. The method used here follows the PCR protocol as described in Schuelke (2000). These regions were amplified using a two-step thermocycling profile of: an initial 5 minutes at 94°C; then 30 cycles of 30 seconds at 94°C, 30 seconds at 55-57°C, and 2 minutes at 72°C; followed by 8 cycles of 30 seconds at 94°C, 30 seconds at 53°C, and 2 minutes at 72°C; and a final extension period of 10 minutes at 72°C.

Amplified lengths for all samples were checked using a 100 base pair (bp) DNA ladder (Roche, Penzberg, Germany) on a 1.5% agarose gel. The amplified products of *Plantago* species of ITS, *CYC*, *cox1*, *ndhF-rpl32*, *rpl32-trn*L, *trnK-psbA*, *trnE-trnTr*, *trnLc-LF*, *trnC-trnD*, *trnH-psbA*, PCM07 and JPM11 were cleaned using ROCHE High Pure PCR Product Purification Kit (Manheim, Germany) and these were sequenced/genotyped using an ABI3730 Genetic Analyzer by the Allan Wilson Centre Genome Service (Massey University, Palmerston North, New Zealand). Genetic sequences were aligned and sequence variation was calculated using

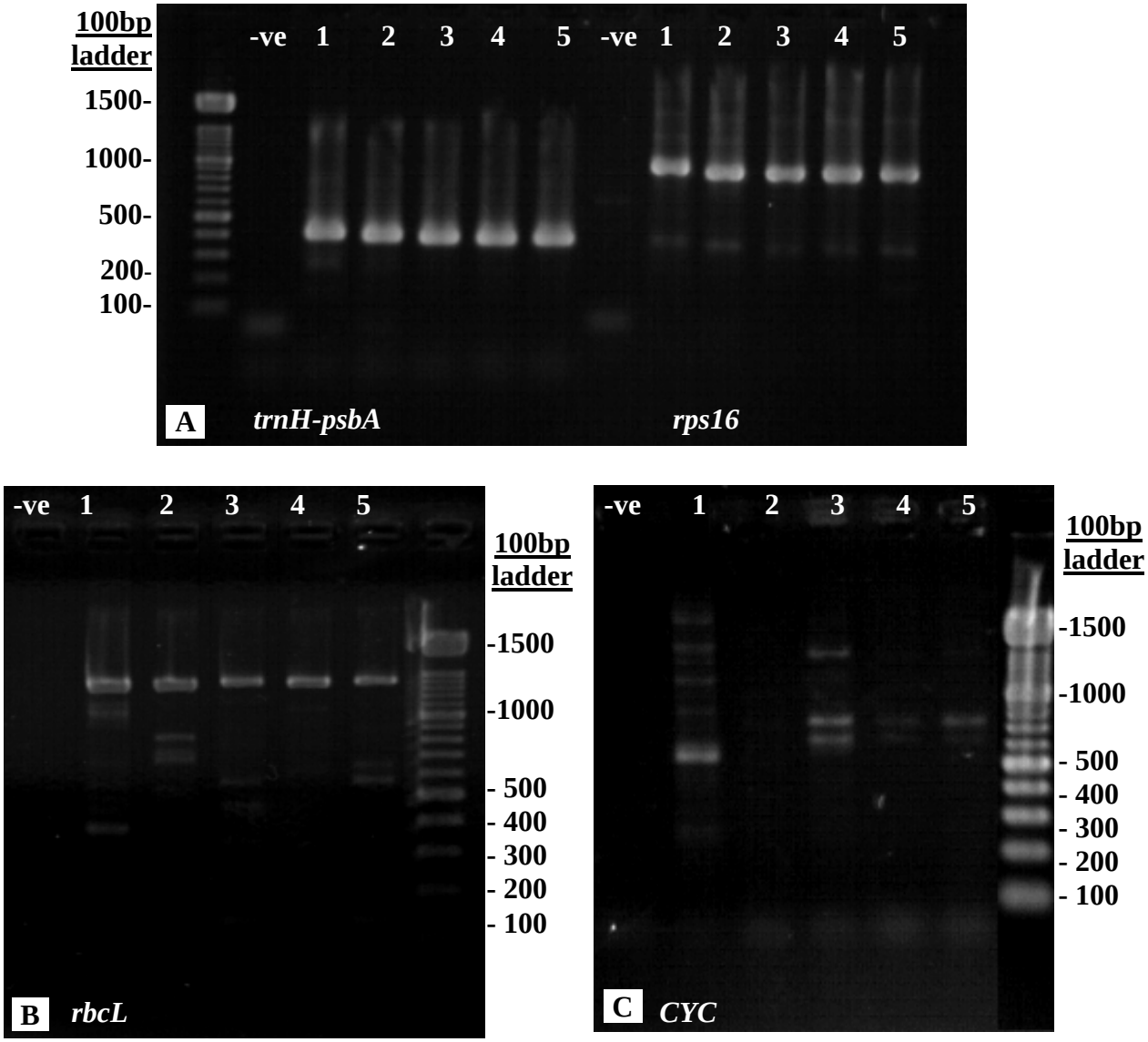


Figure 2.1. Gels illustrating various PCR products. Lanes: 1-5 = *Veronica stricta*, *Plantago lanceolata*, *P. euryphylla*, *P. spathulata* subsp. *spathulata* and *P. triandra* subsp. *triandra* for each region; A: *trnH-psbA* and *rps16* (amplified using primers from Oxelman *et al.* 1997), B: *rbcL* and C: *CYC*. Products in A show single banded amplification of targeted area (referred to as S in Table 2.2) whereas B and C are a mix of multiple bands (S/M and M in Table 2.2).

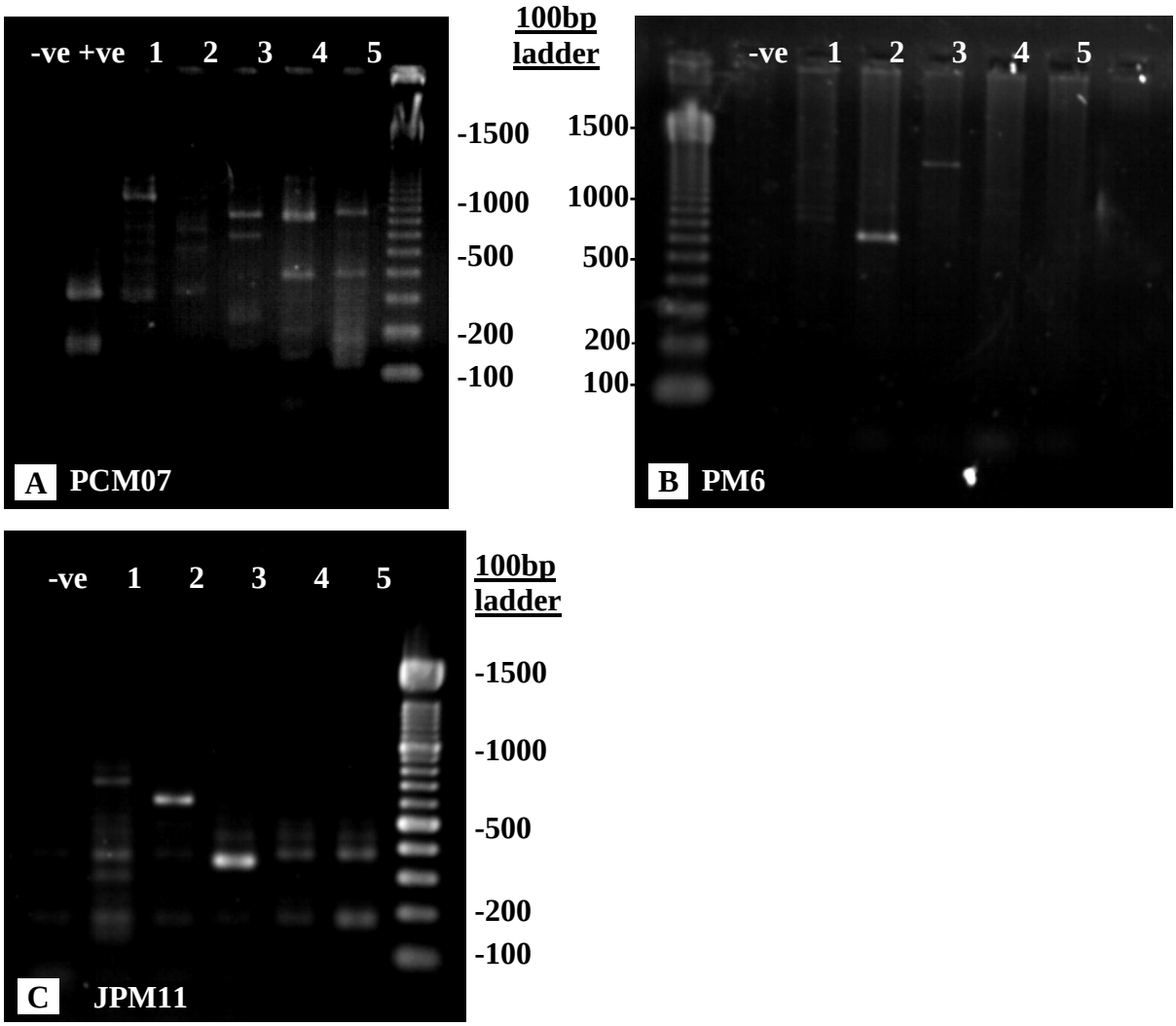


Figure 2.2. Gels illustrating the amplification of the microsatellite regions. Lanes: 1-5 = *Veronica stricta*, *Plantago lanceolata*, *P. euryphylla*, *P. spathulata* subsp. *spathulata* and *P. triandra* subsp. *triandra* for each region; A: PCM07, B: PM6, C: JPM11. Positive controls for B and C were run on a separate gel (not shown). PCM07 and JPM11 yielded product with multiple bands, whereas PM6 failed to amplify the targeted area.

MEGA3.1 (Kumar *et al.* 2004). Microsatellites were analysed using GeneMapper v3.7 (Applied Biosystems). DNA sequences were deposited into Genbank (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=nucleotide).

2.3 Results

Different primer pairs amplified with various results (see Figs. 2.1 and 2.2) and the optimal temperatures for each region varied from 40ºC to 55ºC (Table 2.2). Primer pairs that resulted in single band amplifications for all samples trialed were: ITS, *cox1*, *ndhF-rpl32*, *trnK-psbA*, *trnE-trnTr*, *trnLc-trnLf*, *trnC-trnD*, *rps16* (using both sets of primers) and *trnH-psbA*. *CAM* yielded amplified products only for *Veronica stricta*, whereas *nad1* amplified successfully in the *Veronica stricta* and the cosmopolitan *Plantago* species (*P. lanceolata*) but not in the Australasian *Plantago* species. *CYC* amplified multiple banded products in all samples but did not amplify in *P. lanceolata*. Primers that consistently failed to amplify any product despite several attempts were Waxy, Leafy, *CHS* and *G3pdH* and NIA3.

Table 2.3. Comparison of sequence variation (%) between the three sequenced regions, which represent the three plant genomes. Each of the values presented here is the sequence variation as compared with *Plantago spathulata* subsp. *spathulata*.

	Australasian		**Cosmopolitan**
	Plantago triandra* subsp. *triandra	***P. euryphylla***	***P. lanceolata***
ITS	2.84	2.09	9.12
ndhF-rpl32	2.13	3.32	11.49
Cox1	0.17	0.17	17.98

Several of the amplified products were sequenced but these are not directly comparable because products from different species were sequenced for each region. However, sequences were obtained from all four *Plantago* species for ITS, *ndhF-rpl32* and *cox1*, representing one region from each genome. Percentages of sequence variation among the *Plantago* species for these regions (Table 2.3) showed that the interspecific genetic divergence found in ITS and *ndhF-rpl32* are quite similar but there was relatively lower interspecific genetic divergence found in *cox1* among Australasian species. However, *cox1* displayed higher sequence divergence between the New Zealand *P. spathulata* subsp. *spathulata* and the cosmopolitan *P. lanceolata* than ITS and *ndhF-rpl32*.

The microsatellite primers PCM07 and JPM11 yielded multiple-banded product for all *Plantago* and *Veronica* samples, whereas PM6 did not amplify the targeted region (Fig. 2.2). PCM07 and JPM11 amplified the targeted product in the positive control but PM6 failed to amplify in its positive control. For PCM07 and JPM11, which had successful amplification, we analyzed the product for length and heterogeneity in the samples. Genotyping was unsuccessful for PCM07, whereas the JPM11 microsatellites failed to show heterogeneity in the *Plantago* samples.

2.4 Discussion

Primer pairs for amplification of regions from all three plant genomes were included in this study to test their usefulness for future phylogenetic studies focusing on *Plantago* or related genera. Regions from all three plant genomes were tested, including several regions that are not usually utilised in plant phylogenetic studies, e.g.

regions in the short single copy of the cpDNA and the mtDNA regions. With the exception of ITS and *trnL*c-*trnL*f, the primers included in this study have never been tested in *Plantago*.

2.4.1 PCR amplification results

Most of the primers tested here amplified products for all *Plantago* and *Veronica* species, including a few that were designed specifically for other plant groups (such as *aF* and *cR* from Hasebe *et al.* (1994), which were designed for Leptosporangiate ferns). Surprisingly, other primers such as those for Leafy and *CAM*, which were designed for universal use among angiosperms, consistently failed to amplify PCR products (Leafy) or only produced PCR products for *Veronica stricta* (*CAM*). The primers for the *nad1* region, shown to have worked in previous studies involving *Plantago* species (Bakker *et al.* 2006), amplified DNA from the cosmopolitan *P. lanceolata* and from the outgroup (*V. stricta*) but failed to amplify DNA from native New Zealand and Australian species. Poor primer specificity or differences in the quality of the DNA used in PCR could be the cause for non-amplification of the targeted region. However, the latter seems less probable as the method of DNA extraction was the same for all of the samples and the quality of DNA for all samples were similar when checked on a gel.

Nine regions amplified consistently with good PCR product across all samples: ITS, *cox1*, *ndhF-rpl32*, *trnK-psbA*, *trnE-trnTr*, *trnLc-trnLf*, *trnC-trnD*, *rps16* (using both sets of primers) and *trnH-psbA* (Table 2.2). Among these, one is from nrDNA, one from mtDNA and seven are from cpDNA. This result may simply be proportional to the number of cpDNA markers included in this study but appears to suggest that

regions from the cpDNA have easier amplification compared to nrDNA or mtDNA. Microsatellites failed to amplify using PCR (Fig. 2.2), or failed genotyping (Table 2.2), suggesting that primers for these regions cannot be extended to other species in the genus.

While the *nad1* region has been previously reported as >1 kbp in sequence length, the amplified *nad1* product in *Plantago lanceolata* was about 200 bp. This finding is consistent with previous studies reporting a loss of this mtDNA intron in *Plantago* (Bakker *et al.* 2006). This deletion also occurs in *Pelargonium, Geranium,* and *Sarcocaulon* (Bakker *et al.* 2000).

2.4.2 Sequence divergence compared in ITS, *ndhF-rpl32* and *cox1*

Usefulness of ITS, *ndhF-rpl32* and *cox1* are compared in this study using sequence divergence. Among these, the nuclear ITS region is the most commonly used marker in plants whereas the mtDNA *cox1* is the most commonly used marker used in animal phylogenetics. The cpDNA *ndhF-rpl32* intergenic spacer is a fairly new marker that has not yet been utilised in plant phylogenetics but was found to have the most variation among other primers designed by Shaw *et al.* (2007) for universal angiosperm use. Among the three regions, *cox1* had much higher sequence divergence between the native New Zealand species *P. spathulata* subsp. *spathulata* and the introduced *P. lanceolata* (Table 2.3). This may be support for the fast evolutionary rate of mtDNA in *Plantago* as reported by Cho *et al.* 2004. However, the same rate does not seem to extend to within the New Zealand species, where *cox1* has the least amount of variation (only one nucleotide substitution between the two New Zealand species) compared with ITS or *ndhF-rpl32*. Further studies are needed to determine

the extent of the elevated mtDNA rate in *Plantago*. The *ndhF-rpl32* region appears to have the fastest rate between Australasian species (Australian *P. euryphylla* and New Zealand *P. spathulata* subsp. *spathulata*), whereas *cox1* and ITS did not have as much variation per sequence length for these species. Thus, different primers may be better suited to different research objectives. For example, the fast evolving cpDNA marker *ndhF-rpl32* may be appropriate when working with lower taxonomic levels, while the mtDNA *cox1* may be better suited when analyzing taxa at higher taxonomic levels.

2.4.3 Conclusions

In this study, amplification of regions from nuclear, mitochondrial and chloroplast genomes in four species of *Plantago* and one *Veronica* (more than 10,000 bp in total) have shown that several regions are more easily amplified than others. Cross-amplification of microsatellites designed for other species in *Plantago* failed to amplify or sequence for these four *Plantago* and one *Veronica* species. The thermocycling profile and optimal annealing temperatures reported here can also be further optimized for each individual primer to achieve cleaner amplification product.

ITS had the most sequence variation between the New Zealand species in this study. This region is the most commonly used marker in most plant phylogenetic studies and has proven to be reliable in studies despite several drawbacks (see Feliner and Rosselló 2007). However, researchers have now been looking for alternative markers that may be more useful for phylogenetic analyses (e.g. Shaw *et al.* 2007). It is shown in this study that there may be more information in markers in the nrDNA (such as ITS) or in the other two genomes, depending on the needs of the researcher. Researchers intending to work on *Plantago* and any closely related group such as

Veronica should be able to use the primer pairs that consistently amplified the targeted products (Tables 2.1 and 2.2). Based on results from this study, it is not recommended to use the microsatellite primers tested here for species other than those they were developed for.

For further phylogenetic analyses of Australasian *Plantago* species (Chapters three and four), I chose to use ITS and *ndhF-rpl32* because these markers appear to have substantial variation and are a reasonable length (about 800-1000 bp long). *Cox1* had the best result out of all the mitochondrial regions tested and is therefore also selected for further phylogenetic analyses. The sequence divergence observed among Australasian species here indicates that the elevated rate of evolution is inconsistent across the genus and sequencing of *cox1* for more Australasian species will allow for testing of this hypothesis (Cho *et al.* 2004 only included a few *Plantago* species and no Australasian species).

Chapter Three: A three genome phylogeny for Australasian *Plantago* (Plantaginaceae) species reveals multiple trans-Tasman dispersal patterns[1]

Abstract

Recent phylogenetic studies have shown that the occurrence of recent long distance dispersal in Southern Hemisphere plants is far more prevalent than the expected distribution pattern of ancient Gondwanan vicariance. The phylogeny of New Zealand and Australian *Plantago* (Plantaginaceae) was reconstructed using DNA sequences from the Internal Transcribed Spacers, *ndhF-rpl32* and *cox1* from the nrDNA, cpDNA and mtDNA, respectively. The resolved phylogeny shows evidence for at least three long distance directional dispersal events, which does not support a Gondwanan vicariance origin for Southern Hemisphere *Plantago* species. Molecular dating also showed support for the conclusion that the distribution of Australasian *Plantago* was influenced by relatively recent long distance dispersal and not vicariance events. A concatenated ITS, *ndhF-rpl32* and *cox1* phylogeny indicated that the three trans-Tasman dispersals have occurred from Australia to New Zealand, consistent with the direction of the West Wind Drift. The finding of three dispersal events to New Zealand differs from the common pattern of a single origin followed by rapid radiation in many New Zealand plant groups. The apparent success of this group in New Zealand may be attributed to the fact that the plants were pre-adapted to the environments they established in.

[1]**Note:** This chapter has been prepared as a manuscript for submission in a journal and has been retained in this format. Co-authors of this paper are Phil Garnock-Jones, Heidi Meudt and Peter Ritchie, whose roles were supervision and advice. Mei Lin Tay collected sequence data and wrote the manuscript. Mei Lin Tay, Phil Garnock-Jones, and Heidi Meudt collaborated for the collection of *Plantago* samples

3.1 Introduction

For the last ca. 45 years, it has been the common view that Southern Hemisphere flora is a relic of the supercontinent Gondwana and that their current distributions are the result of vicariance (see Pole 1994; de Queiroz 2005; McGlone 2005). However, recent studies using molecular data revealed that most plant groups in the Southern Hemisphere have instead achieved their current distributions through more recent long distance dispersal events (Winkworth *et al.* 1999; Winkworth *et al.* 2002b; de Queiroz 2005; McGlone 2005). This paradigm shift is an element that is reflected in New Zealand plant biogeography. New Zealand has always been featured in studies about Southern Hemisphere biogeography because of its unique geologic history. New Zealand has been separated by at least 1500km from its closest landmass (Australia) since the breakup of the supercontinent Gondwana about 80 million years ago (mya) (Cooper and Cooper 1995). In spite of this, there is evidence that even plant groups regarded as classic examples of "ancient New Zealand lineages" (Cooper and Millener 1993) have previously undergone long distance dispersal. For example, it has been found that several lineages of southern beeches (*Nothofagus*) have arrived relatively recently through long distance dispersal (Swenson *et al.* 2001; Knapp *et al.* 2005). Molecular dating has shown that many plant lineages have diversified recently and are probably the result of long distance dispersal, not relicts of Gondwana (Hurr *et al.* 1999; Lockhart *et al.* 2001; Heenan *et al.* 2002; Zhang and Renner 2003; Perrie and Brownsey 2005; Barker *et al.* 2007; Perrie and Brownsey 2007). Some authors have gone as far as to say that the entire New Zealand flora arrived by long distance dispersal (Pole 1994), but there are still examples appearing in the literature of plant groups that date back to a Gondwanan origin, e.g. *Agathis* (Stöckler *et al.* 2002;

Knapp *et al.* 2007). Thus, the origins of the New Zealand flora may be the result of a mixture of older vicariance and more recent long distance dispersal events.

A common observation among New Zealand plant groups with a history of long distance dispersal is a lack of DNA sequence divergence between species from Australia and species from New Zealand (Hurr *et al.* 1999; Renner *et al.* 2000; Swenson *et al.* 2001; Wagstaff *et al.* 2002; Smissen *et al.* 2003; Ford *et al.* 2007; Sanmartín *et al.* 2007) and a recurrent pattern of New Zealand and Australian species being sister groups (Linder and Crisp 1995; Sanmartín and Ronquist 2004; Perrie and Brownsey 2005; Crayn *et al.* 2006; Meudt and Simpson 2006). A brief survey of recent phylogenetic studies involving Australasian plants revealed that many New Zealand plant groups have arrived from Australia (e.g. Wagstaff *et al.* 1999; Wagstaff and Wege 2002; Ford *et al.* 2007). This asymmetric West to East direction of dispersal is consistent with expectations of the West Wind Drift, which predicts more dispersal events from west to east because of westerly winds and ocean currents (Raven 1973; Winkworth *et al.* 2002b; Muñoz *et al.* 2004; Sanmartín *et al.* 2007). Another pattern that is frequently found is a single dispersal to New Zealand followed by rapid species radiation (Wagstaff and Garnock-Jones 1998; Winkworth *et al.* 1999; Perrie *et al.* 2003; Albach *et al.* 2005b; Meudt and Simpson 2006). Other dispersal biogeographic patterns also exist (Winkworth *et al.* 2002b). However, these patterns can be difficult to elucidate because species relationships are often unresolved polytomies in large scale studies (i.e. at the genus level or higher) (Swenson *et al.* 2001; von Hagen and Kadereit 2001). Species relationships have previously been found to be unclear in studies focusing on Australasian groups, e.g. Radford *et al.*

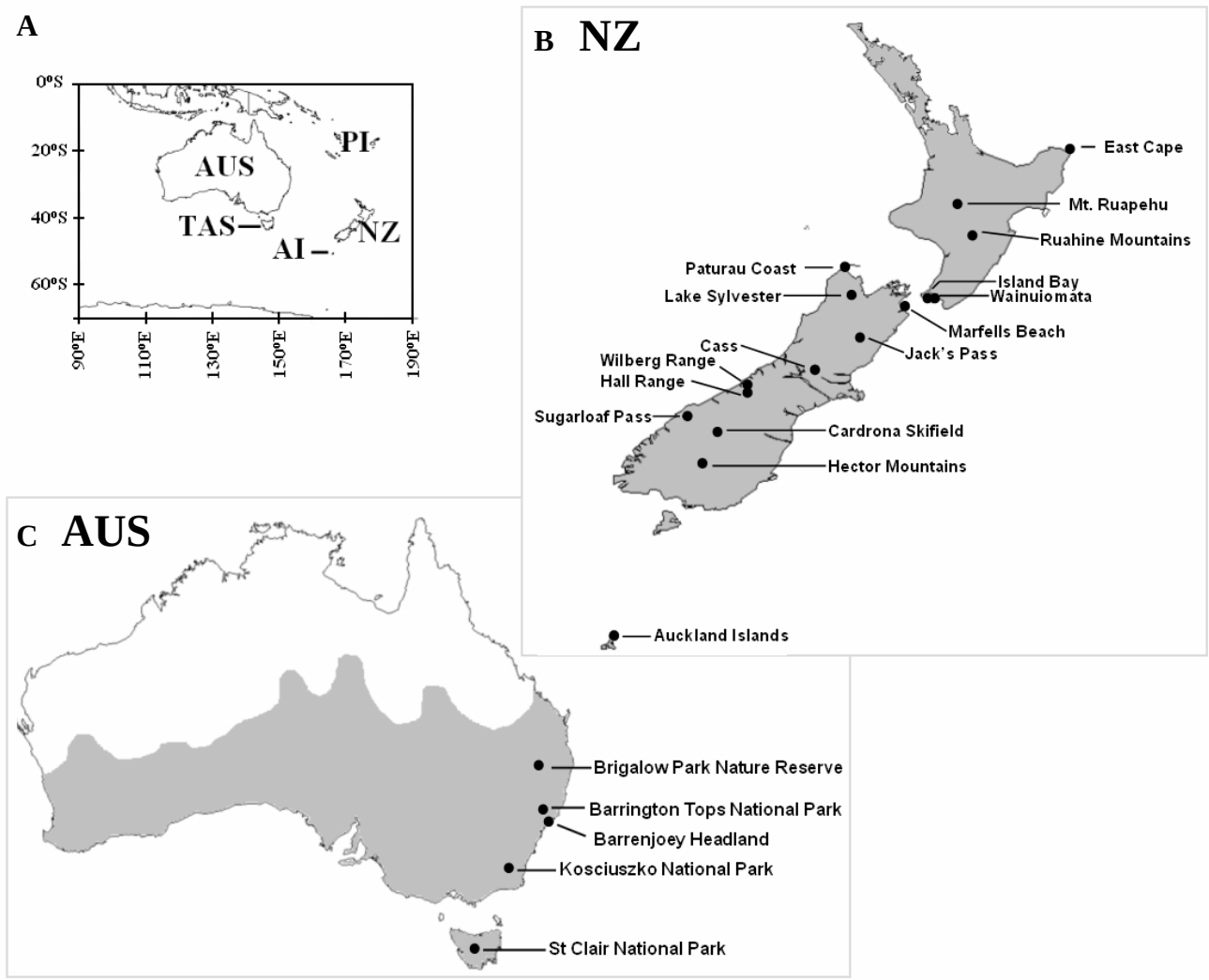


Figure 3.1. A: The Australasian region AI = Auckland Islands, AUS = Australia, NZ = New Zealand, PI = Pacific Islands and TAS = Tasmania. B: Distribution of native species of New Zealand *Plantago* from the databases of the Allan Herbarium (Landcare Research, Christchurch, NZ) and Victoria University of Wellington (Wellington, NZ). C: Distribution of native Australian *Plantago* species from the Australian Virtual Herbarium (http://www.flora.sa.gov.au/avh/). For B and C, shaded areas illustrate the distribution of native Australasian species and black circles represent the localities where native Australasian species were sampled for this molecular study

.

(2001); Perrie *et al.* (2007). Thus, more studies focused on lower taxonomic levels are needed to fully understand the historical biogeography of Australasian plants.

The molecular phylogeny of Australasian *Plantago* species was reconstructed in order to resolve species relationships and to investigate regional biogeographic patterns, e.g. to test the occurrence of long distance dispersal versus vicariance in the evolutionary history of the group. *Plantago* is a large worldwide genus in the family Plantaginaceae with six subgenera and more than 200 species (Rahn 1996). There are eight native *Plantago* species in New Zealand (Sykes in Webb *et al.* 1988), and an undescribed polyploid species (*P.* "sylvester") (Groves and Hair 1971). There are 24 species that are native to Australia (Briggs 1992), all of which are placed in subg. *Plantago* (Rahn 1996). New Zealand species of *Plantago* are distributed widely over both the North and South Islands (Fig. 3.1). In Australia, they are found on the mainland below 25^{o}S and in Tasmania. Australasian *Plantago* also occur on offshore islands; *P. aucklandica* is endemic to the Auckland Islands, and *P. triantha* is endemic to the Auckland Islands and Tasmania. The plants are mostly small rosette herbs and are usually found in damp areas (such as seepage areas and near bogs, tarns, and river edges) from alpine regions to lowland herbfields and coastal areas. Morphological differences between taxa are not easily defined, which has resulted in problems with species delimitation and identification in the past. Classification of subgenera and sections in this paper follows Rahn (1996), whereas classification of Australasian species follows Sykes (in Webb *et al.* 1988) and Briggs (1992).

When Rahn reconstructed the phylogeny of *Plantago* using morphological characters, the New Zealand species were paraphyletic to multiple South American, Australian

and other Pacific species. It was suggested that *Plantago* had a distribution consistent with vicariance with subsequent extinctions (Rahn 1996). Molecular phylogenetics can often form a robust and well-resolved evolutionary history where morphological characters and the fossil record cannot. In a worldwide phylogenetic study of *Plantago*, Rønsted *et al.* (2002) inferred that the distribution of the species was largely influenced by long distance dispersal. However, only three Australasian species were included in the study. Six Australasian species were included in another molecular phylogenetic study by Hoggard *et al.* (2003) but these studies were not focused on species relationships of the Australasian species and thus provided no further resolution among them.

Including all the Southern Hemisphere *Plantago* species is key to testing dispersal vs. Gondwanan vicariance. In a Gondwanan vicariance scenario, three clades are expected: a South American clade that is sister to an Australian clade, and a New Zealand clade that is sister to (and with a node that is basal to) the South American-Australian group. In addition, the node of the New Zealand clade should be at least 80 mya according to the break-up sequence of Gondwana. Since this study is focused on the Australasian *Plantago* species, we can conclude that long distance dispersal has occured if: (1) the New Zealand and Australian groups are not reciprocally monophyletic, (2) the New Zealand species are in a clade that is derived from the Australian species (concordant with the direction of wind flow but does not match the pattern expected from Gondwanan vicariance) and/or (3) the split between New Zealand and Australian groups is found to be less than 80mya.

The Internal Transcribed Spacers (ITS) from nuclear DNA (nrDNA), *ndhF-rpl32* spacer from chloroplast DNA (cpDNA) and *cox1* from mitochondrial DNA (mtDNA) were chosen to elucidate the evolutionary history of the Australasian *Plantago*. The *cox1* marker is particularly interesting for *Plantago* because unprecendented elevated rates of mtDNA evolution, even several times higher than human mtDNA evolution rates, have been reported in the genus (Cho *et al.* 2004). This is surprising because mtDNA in plants is normally characterized by slow evolving rates (Wolfe *et al.* 1987). A mtDNA marker would also provide good phylogenetic resolution if the elevated rate was present in the Australasian species. Using markers from the three genomes will allow a comparison of evolutionary patterns in the three plant genomes, and also an investigation of whether the mtDNA of Australasian *Plantago* has a faster rate of evolution than nrDNA and cpDNA. The aims of this study are thus to test for evidence of either Gondwanan vicariance or long distance dispersal in the evolutionary history of this plant group, to investigate Australasian biogeographic patterns, and to examine if fast evolving mtDNA is present in Australasian *Plantago*.

3.2 Materials and Methods

3.2.1 Study group

Samples for this study were collected from wild populations across the North and South Islands of New Zealand from previously known locations in order to cover the

Table 3.1: Collection details and Genbank accession numbers for species included in this study.

Species	Indigenous distribution	Collected from	Collection number	Genbank accession number			Herbarium voucher
				ITS	*ndhF-rpl32*	*cox1*	
Plantago alpestris	AUS	Kosciuszko National Park, NSW, AUS	*BGB9748*				NSW742962
P. aucklandica	AI	AI, NZ (cult. Otari-Wilton Bush)	N/A				WELTU20185
P. australis	America	Hunua Ranges, Auckland, NZ	*MLT019 & PGJ*				WELTU20181
P. cladarophylla	AUS	Barrington Tops National Park, NSW, AUS	*JRHosking 2682*				WELTU20189
P. coronopus	Mediterranean, Europe	Island Bay, Wellington, NZ	*PGJ2549*				WELTU20183
P. cunninghamii	AUS	Brigalow Park Nature Reserve, NSW, AUS	*JRHosking 2752*				WELTU20186
P. daltonii	AUS	St. Clair National Park, AUS, Tasmania	*BGB9782*				NSW743874
P. debilis	AUS	Barrenjoey Headland, NSW, AUS	*BGB9738*				NSW 742894
P. euryphylla	AUS	Kosciuszko National Park, NSW, AUS	*BGB9743*				NSW742956
P. glacialis	AUS	Kosciuszko National Park, NSW, AUS	*BGB9753*				NSW743813
P. lanceolata	cosmopolitan	Karori, Wellington, NZ	*PGJ2551*				WELTU20184
P. lanigera (1)	NZ	Sugar Loaf Pass, Aspiring National Park, NZ	*Mike Thorsen s.n.*				WELTU20133
P. lanigera (2)	NZ	Sugar Loaf Pass, Aspiring National Park, NZ	*Mike Thorsen s. n.*				WELTU20133
P. lanigera (3)	NZ	Hall Range, Canterbury, NZ	*PBH s.n.*				WELTU20143
P. lanigera (4)	NZ	Wilberg Range, Westland, NZ	*MLT027 et al.*				WELTU20147
P. lanigera (5)	NZ	Wilberg Range, Westland, NZ	*MLT027 et al.*				WELTU20147
P. major	cosmopolitan	Ben Burn Park, Karori, NZ	*PGJ2550*				WELTU20180
P. muelleri	AUS	Kosciuszko National Park, NSW, AUS	*BGB9752*				NSW743812
P. obconica (1)	NZ	Hector Mountains, Otago, NZ	N/A				CHR573261
P. obconica (2)	NZ	Cardrona Ski Field, Wanaka, NZ	*PGJ2600 et al.*				WELTU20121
P. paradoxa	AUS	St. Clair National Park, AUS, Tasmania	*BGB9781*				WELTU20187
P. raoulii (1)	NZ	Lake Sarah, Cass, NZ	*PGJ2559*				WELTU20153
P. raoulii (2)	NZ	Wainuiomata Valley, Wellington, NZ	*PB & RL s.n.*				WELTU20155
P. sp.*	?	Pukerua Bay, Wellington, NZ	*PGJ2566 & MLT*				WELTU20178

Table 3.1: Collection details and Genbank accession numbers for species included in this study.

Species	Indigenous distribution	Collected from	Collection number	Genbank accession number			Indigenous distribution
				ITS	*ndhF-rpl32*	*cox1*	
P. spathulata subsp. *spathulata*	NZ	Marfells Beach, Marlborough, NZ	*PGJ2629 & MLT*				WELTU20117
P. "sylvester"	NZ	Lake Sylvester, Nelson, NZ	*MLT022 & PGJ*				WELTU20150
P. tasmanica var. *tasmanica*	AUS	St. Clair National Park, AUS, Tasmania	*BGB9780*				WELTU20188
P. triandra subsp. *masoniae*	NZ	Paturau Coast, Nelson, NZ	*PJL s. n.*				WELTU20168
P. triandra subsp. *triandra*	NZ	Lake Sylvester, Nelson, NZ	*MLT021 & PGJ*				WELTU20163
P. triantha	TAS / AI	Enderby Island, AI, NZ	*VT55*				WELTU20177
P. unibracteata (1)	NZ	Mt Ruapehu, NZ	*MLT053 et al.*				WELTU20173
P. unibracteata (2)	NZ	Lake Sylvester, Nelson, NZ	*MLT024 & PGJ*				WELTU20175
P. varia	AUS	Kosciuszko National Park, NSW, AUS	*BGB9767*				NSW743869
Veronica hookeriana	NZ	Whanahuia Range, Ruahine Mountains, NZ	*PGJ2458*				WELTU
Veronica salicornioides	NZ	Jacks Pass, Hanmer, Canterbury, NZ	N/A				CHR512475

Locations: AI = Auckland Islands, AUS = Australia, NSW = New South Wales, NZ = New Zealand and TAS = Tasmania. Collectors: *BGB* = Barbara Briggs, *MLT* = Mei Lin Tay, *PBH* = Peter Heenan, *PGJ* = Phil Garnock-Jones, *PB* = Peter Beveridge, *PJL* = Peter Lockhart, *RL* = Rodney Lewington and *VT* = Vanessa Thorne. Herbarium vouchers: CHR = Allan Herbarium, Landcare Research, Christchurch, New Zealand; NSW = National Herbarium of New South Wales, Australia; WELT = Herbarium of Museum of New Zealand, Te Papa Tongarewa, Wellington, New Zealand; and WELTU = H. D. Gordon Herbarium in Victoria University of Wellington, New Zealand.

*The taxon referred to as *P.* sp. in this study is an introduced species that has never been reported in New Zealand. The ITS sequence for this species was identical to a *P. asiatica* sequence downloaded from Genbank (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=nucleotide) but further morphological work is needed to clarify the identity of this sample.

Note: Accession numbers will be inserted into table before publication of the paper.

range of species distributions, and also from Australia (Fig. 3.1). Location of samples, along with voucher information and Genbank accession numbers are presented in Table 3.1. Multiple individuals were obtained for several of the Australasian species. DNA sequences were obtained from all eight known native New Zealand *Plantago* species, including the recently described New Zealand species, *P. obconica* (Sykes 1988) and the undescribed New Zealand polyploid specimen, tag-named *P.* "sylvester" (Groves and Hair 1971). The chromosome number of $2n = 96$ reported by Groves and Hair (1971) was confirmed using one sample from the populations referred to as *P.* "sylvester" in this study. DNA sequences were also obtained for 12 of the 24 native Australian species including *P. triantha*, which is also found in New Zealand. Introduced *Plantago* species that have established in New Zealand (*P. coronopus, P. lanceolata, P. major* and *P. australis*) were included in this study. *P. major* and *P. australis* are in the same subgenus as the Australasian species (subg. *Plantago*). *P. coronopus* and *P. lanceolata* are in subg. *Coronopus* and *Albicans*, respectively (following Rahn 1996).

P. coronopus and *P. lanceolata* were used as close outgroups as they represent the sister groups of subg. *Plantago* (Rønsted *et al.* 2002), whereas *Veronica hookeriana* and *V. salicornioides* are used as distant outgroup species for *Plantago* because *Veronica* is a close relative and has been used as an outgroup in the past (Rønsted *et al.* 2002; Cho *et al.* 2004). The dataset of one region (*ndhF-rpl32*) contained only *V. hookeriana* as a distant outgroup due to sequencing difficulties.

3.2.2 Genetic markers

In order to reconstruct the molecular phylogeny, the following genetic markers were chosen. ITS, found in the nuclear genome of plants and animals, has various advantages as a marker (Álvarez and Wendel 2003) and is widely applied; 66% of published plant phylogenetic papers at the genus level or below between 1998 and 2002 utilised this marker (Feliner and Rosselló 2007). The second molecular marker used is the *ndhF-rpl32* intergenic spacer. This region is located in the small single copy region of the chloroplast and universal primers for this region have been published recently (Shaw *et al.* 2007), where this marker was found to have the highest variation among sequences out of many other amplified chloroplast regions. Thirdly, *cox1*, a gene in mtDNA that codes for the cyclooxygenase enzyme was used. The three regions described above each come from a different genome. By combining sequences from different independent sources (at least between the biparentally inherited ITS and the other two maternally inherited regions), incongruences (if any) in the evolutionary history of the different genomes may be visualised. Additionally, it has been shown that the increased length of variable sequences and sampling from different sites both increase accuracy and support for phylogenetic analyses (Cummings *et al.* 1995) and thus, concatenation of these three regions will be useful if they do not present conflicting signals.

3.2.3 Molecular techniques

Tissue samples were either preserved in silica gel from field collections, or were obtained from existing herbarium specimens. DNA extractions were performed using the DNEasy Plant Mini kit (Qiagen, Hilden, Germany) or using a cetyltrimethylammonium bromide (CTAB) protocol modified from Doyle and Doyle

(1990), after manual disruption of dried tissue using a pestle and mortar. The primers used to amplify the various DNA regions are: *ITS28CC*-CGCCGTTACTAGGGGAATCCTTGTAAG (Wagstaff and Garnock-Jones 1998) and *ITS5*-GGAAGTAAAAGTCGTAACAAGG (White *et al.* 1990); *ndhF*-AAAGGTATKATCCAYGMATATT and *rpl32R*-AATATCCCTTYYTTTTCCAA from (Shaw *et al.* 2007); and *cox1F4*-GGATATCTAGGYATGGTTTATGC and *cox1R3*-AAGCTGGAGGACTTTGTAC (Cho *et al.* 2004 pers. comm.) These primers amplify the ITS region (ITS1, 5.8s nuclear rDNA and ITS2), *ndhF-rpl32* spacer region, and *cox1* mitochondrial gene, respectively.

PCR amplification was performed using Eppendorf Mastercycler ep gradient S (Hamburg, Germany) in a final volume of 25 μl of the following: 16.35 μl water, 10X ThermoPol reaction buffer (10 mM KCl, 10 mM $(NH_4)2SO_4$, 20 mM Tris-HCl (pH8.8), 2 mM $MgSO_4$, 0.1% Triton X-100)(New England BioLabs), 0.4 mg/mL BSA, 250 μmol dNTPs, 10 ρmol each primer, 0.75 U of *Taq* DNA polymerase (New England BioLabs), and 0.4 μl DNA template. The amplification was carried out with a thermocycling profile of an initial 2 minutes at 94°C, followed by 30 cycles of 1 minute at 94°C, 1.5 minutes at 50°C, 1 minute at 72°C, and ending with a final extension time of 5 minutes at 72°C. This PCR protocol was used to successfully amplify regions for all primers pairs. PCR products were visualised on 1.5% agarose gels before being purified with the High Pure PCR Purification Kit (Roche Diagnostics, Manheim, Germany). The purified PCR products were sequenced in both directions using an ABI3730 Genetic Analyzer by the Allan Wilson Centre Genome Service (Massey University, Palmerston North, New Zealand).

Table 3.2: Summary of statistics for each dataset used in this study.

	ITS (nuclear)	***ndhF-rpl32* (chloroplast)**	***cox1* (mitochondrial)**	**concatenated**
Base frequencies of all sites (%)	T- 23.0 C- 26.4 A- 22.2 G- 28.4	T- 37.0 C- 12.5 A- 37.3 G- 13.2	T- 32.8 C- 23.2 A- 23.3 G- 20.6	T- 31.2 C- 20.4 A- 28.0 G- 20.4
Base frequencies of variable sites (%)	T- 31.6 C- 26.2 A- 22.5 G- 19.6	T- 33.0 C- 19.0 A- 25.4 G- 22.5	T- 26.1 C- 34.1 A- 25.5 G- 14.3	T- 30.5 C- 25.1 A- 24.7 G- 19.7
Range of sequence length (aligned length) in bp including outgroups	577-622 (639)	548-737 (942)	613 (613)	1774-1972 (2194)
No. of recoded gaps appended and aligned length (with recoded gaps)	26 (665)	36 (978)	0 (613)	61 (2255)
No. of variable sites and parsimony informative sites including outgroups (%)	211 (20%)	286 (10%)	150 (19%)	638 (13%)
No. of variable sites and parsimony informative sitesfor ingroup sequences (%)	78 (6%)	95 (6%)	39 (2%)	212 (5%)
Modeltest model (AIC)	GTR + G	TVM + G	TVM + G	TIM + I + G
Gamma shape estimate	0.4645	2.0817	0.2834	0.9143
Pinvar	0	0	0	0.2309
No. of transitions (all sites included)	17	14	10	38
No. of transversions (all sites included)	10	17	11	38
Transition/transversion ratio (all sites included)	1.6	0.8	0.9	1.0
% missing data	0.43%	0.64%	0.44%	0.52%

3.2.4 Dataset alignment and analyses

The program MEGA v3.1 (Kumar *et al.* 2004) was used to assemble and align sequences for each accession. Before alignment, the sequences were submitted to BLAST (http://www.ncbi.nlm.nih.gov/blast/) to check that they were the correct region. The aim of this study is to look at the overall biogeographic patterns of the different austral species and not to analyse samples at the population level. Therefore, population replicates were removed according to the availability of sequences in all three regions. This was done to improve the efficiency of analyses and to visually simplify the analyses, as well as to allow for concatenation of the datasets where a different number of taxa was sequenced for each region. There were 35 individuals included in the *ndhF-rpl32* dataset, whereas the ITS and *cox1* datasets included sequences from 36 individuals (*Veronica salicornioides* was not included in the *ndhF-rpl32* dataset). For the concatenated dataset, we used the 35 taxa included in the *ndhF-rpl32* dataset. Exploration of sequence datasets was performed using functions provided in MEGA v3.1 and DAMBE v4.5.9 (Xia and Xie 2001).

Analyses were first performed individually for each of the datasets and then on a concatenated dataset. Gaps were coded using modified complex indel coding (MCIC) method as implemented in SeqState v1.32 (Müller 2005) and these were included in the dataset. Preliminary analyses of our datasets suggested that gaps are informative and should be included, and previous analyses have found that MCIC coded gaps often outperform analyses in which gaps are treated as missing data or excluded (Simmons *et al.* 2007). Datasets with gaps coded resulted in 26 indel characters added onto the end of the ITS and 36 onto the chloroplast sequence matrix. There were no indels in the *cox1* dataset, as expected, because it is a coding region. An indel was

present in the aligned ITS sequence of *Veronica salicornioides*, one of the distant outgroups, but the sequence was not included in the concatenated dataset because of sequencing problems. Therefore, the concatenated dataset had 61 indel characters added onto the end of the alignment (Table 3.2). Ambiguous characters present in the datasets were excluded from phylogenetic analyses.

Maximum parsimony (MP) and maximum likelihood (ML) analyses were conducted using PAUP* v4.0b10 (Swofford 2002). A heuristic search was conducted under a MP criterion using 10,000 replicates of random sequence addition and tree-bisection-reconnection (TBR) branch swapping. A maximum of 10,000 trees was set for the MP analysis. A heuristic search was also conducted under a ML criterion, with 100 replicates of random sequence addition, and TBR branch swapping. The substitution model was selected by Modeltest v3.7 (Posada and Crandall 1998), which tests the fit among 56 different models of different complexity using a hierarchical likelihood ratio test. Models were selected using Akaike Information Weights criterion (AIC) because it has advantages over the hierarchical likelihood ratio test (Posada and Buckley 2004). Chosen models and parameter estimates are presented in Table 3.2. Nonparametric bootstrap support was assessed using 200 replicates for both MP and ML analyses, with random sequence addition and TBR branch swapping.

MrBayes v3.1.1 (Huelsenbeck and Ronquist 2001) was used to conduct heuristic searches, also implementing the AIC model that was selected by Modeltest. Each dataset was initially analysed with four chains and 500,000 generations. Additional generations were added if necessary in order to reach a standard deviation of split frequencies of less than 0.01, which ensures that the runs have converged on a

stationary distribution. For the concatenated dataset, two analyses were run. First one model was used across the whole dataset, and then a partitioned dataset was used such that separate models were used for each dataset as recommended by Modeltest for each individual dataset. For each run, 25% of the trees were excluded as burn-in.

Topological congruence was examined firstly by visually comparing trees generated from each of the datasets for each analysis. A partition-homogeneity test was also run in PAUP* (1000 replicates, TBR branch swapping, 100 random replicates of random taxa addition, and MaxTrees = 10,000) for all combinations of datasets. Additionally, supertree networks were constructed from ML trees from the three datasets using Splitstree v4.8 (Huson and Bryant 2006). The ITS, cpDNA and mtDNA datasets were then concatenated and characters (e.g. geographic distributions and habitats) were mapped onto a concatenated tree in order to infer evolutionary patterns.

3.2.5 Molecular dating

An ITS dataset consisting of representatives from each Australasian *Plantago* species, several *Veronica* sequences, and sequences of *Aragoa* and *Littorella* obtained from Genbank (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=nucleotide) were used to date nodes for testing the hypothesis of a Gondwanan origin for the Australasian (see Appendix I (B) for details of the dataset). ITS sequences were used because only ITS sequences for *Litorella* and *Aragoa* are readily available. Rate heterogeneity of the dataset was checked by estimating the likelihood score of a tree with a molecular clock constraint and comparing this to the likelihood score of a tree obtained without this constraint using a likelihood ratio test (Felsenstein 1981). It was found that the ITS sequences did not evolve in a clock-like manner. Therefore, a relaxed clock

model was implemented in the following molecular dating analyses. It has also been shown through simulations of different datasets that a relaxed–clock approach is more accurate and more precise than other models at reconstructing phylogenies (Drummond *et al.* 2006). BEAST v1.4.6 (Drummond *et al.* 2006) was used to implement a relaxed-clock model within a Bayesian framework to estimate divergence times for the lineages in our dataset. The program BEAUti v1.4.6 (Drummond *et al.* 2006) was used to set up groups that were used for dating nodes. Only monophyletic groups with high support in the ML tree were set up. These were: 1) a clade containing *Plantago* and *Aragoa* sequences, 2) a clade of the *Plantago* species, 3) a clade with all *Plantago* but excluding *P. coronopus* and *P. lanceolata*, and 4) Clade I, excluding *P. cunninghamii*. Analyses in BEAST were set up with 10,000,000 runs initially, with additional 10,000,000 runs until estimated sample sizes (ESS) for each parameter were large enough, resulting in 20,000,000 to 40,000,000 runs for each analysis. A GTR + I + G model was implemented for each analysis as recommended by Modeltest. A gamma distribution of rates is preferable in molecular dating (Sanderson and Doyle 2001). TRACER v1.4 (Rambaut and Drummond 2007) was used to analyse results. Exploration of the trees reconstructed by BEAST showed identical topology of the ITS tree reconstructed using ML analysis in all cases.

Testing Gondwanan origin of Australasian species

The hypothesis of Gondwanan origin for the Australasian species was tested in two separate analyses. First, a divergence date for a group was estimated to obtain a substitution rate for the sequences. The divergence date for the node of the Australasian *Plantago, P. major, P.* sp. and *P. australis* clade is expected to be at least 80 mya if the New Zealand species were present before New Zealand split from

Gondwana. Thus, the node of this clade was constrained at 80mya. Secondly, substitution rates were applied to estimate divergence of the nodes. The slowest rate reported across all herbaceous angiosperms lineages in a study by Kay *et al.* (2006) of 1.72 $x10^{-9}$ substitutions/site/year was chosen in order to be conservative.

The resulting substitution rates and node dates were then subjected to a t-test to compare it to the expected values. The test involves calculating a 99% confidence interval for the expected value and determining if the obtained value from BEAST is within that 99% interval. If values obtained from BEAST are found outside the interval, then the values are significantly different at the 1% level. Estimated substitution rates were then compared with the lowest substitution rate found in herbaceous angiosperms and estimated divergence dates were compared to 80mya.

Estimating age of the Australasian species

We used the dates 5.47 MYA for the radiation of the *Plantago-Littorella* group and 7.1 MY since the divergence of this group and *Aragoa* (see Rønsted *et al.* 2002) to calibrate the tree to obtain an estimate for the divergence dates of Australasian species.

3.3 Results

3.3.1 Dataset statistics

Exploration of the three datasets revealed that the *ndhF-rpl32* dataset had more variable sites relative to the ITS and *cox1* datasets, but a lower percentage of parsimony-informative sites than the ITS dataset (Table 3.2). When gaps were coded

as separate indel characters, the ITS and *ndhF-rpl32* datasets had slightly more parsimony informative sites than when gaps were ignored. In comparison to these two regions, the *cox1* dataset had fewer variable sites but a higher percentage of parsimony informative sites than the *ndhF-rpl32* dataset, and a similar percentage of parsimony informative sites as the ITS dataset. The variable sites in *cox1* provided little resolution within the Australasian *Plantago* clade, but were useful in delimiting the clade and separating *Plantago* from *Veronica*.

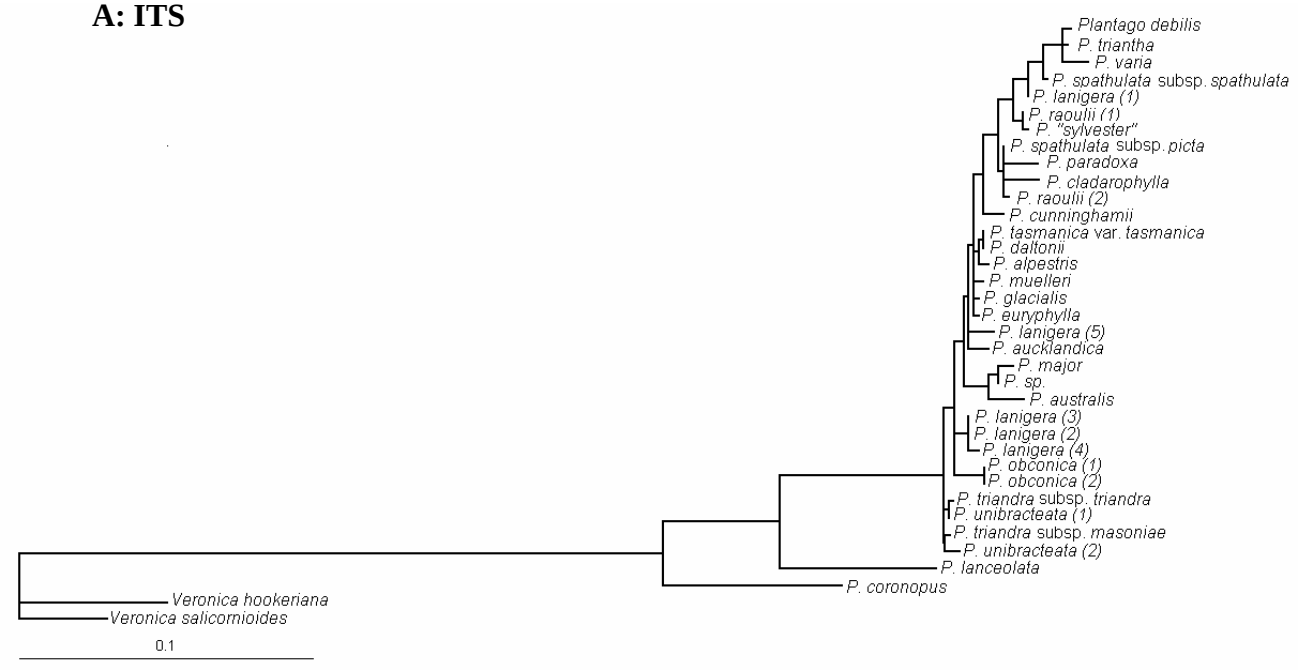
All ITS and *ndhF-rpl32* sequences were unique, whereas the *cox1* dataset only had 15 unique sequences. This result illustrates the lack of *cox1* sequence divergence between members of Australasian *Plantago*. In fact, some Australian species had identical *cox1* sequence to some New Zealand species (*P. muelleri* was identical to *P. triandra* and *P. unibracteata*; *P. debilis* and *P. triantha* were identical to *P. raoulii* and *P. spathulata* subsp. *picta*; and *P. euryphylla* and *P. glacialis* were identical to *P. lanigera*, *P. novae-zelandiae*, *P. obconica* and *P. aucklandica*).

Chi-squared tests of all base frequencies only showed significant base heterogeneity for the *ndhF-rpl32* dataset and among variable sites for the *cox1* dataset. In all three datasets, average G-C (guanine and cytosine) content was slightly less than A-T (adenine and thymine). This could be explained by transversions from A/T to C/G events, which are common in chloroplast genomes and is caused by oxidative damage (GuhaMajumdar and Sears 2005). In the *cox1* dataset, the third codon position had less G-C content than the first and second positions. The G-C content in plant genomes has been found to vary substantially although the implications of this are still unknown (Carels and Bernardi 2000).

Within the *cox1* dataset, the substitution rate was higher in the third codon position than the other two codon positions, as expected because the third codon position is functionally redundant for many amino acids. Transition and transversion rates differed among the three regions, with transitions being more frequent than transversions in the ITS dataset but the other way around for the *ndhF-rpl32* and the *cox1* datasets (Table 3.2). Transitions may have a bias over transversions because of the different molecular structures of purines (A, G) and pyrimidines (C, T) (Decker-Walters *et al.* 2004) but Keller *et al.* (2007) report that the bias is not universal and the results from the organellar DNA datasets support the latter situation. The transition/ transversion ratio was higher in the third codon position for the *cox1* gene.

Homogeneity partition (HP) tests indicated that the datasets were not congruent (p-value=0.001, 0.002, 0.001, 0.016 for ITS + *ndhF-rpl32* + *cox1*, ITS + *cox1*, ITS + *ndhF-rpl32*, and *ndhF-rpl32* + *cox1*, respectively). While the organellar DNA datasets (*ndhF-rpl32* and *cox1*) showed the least conflict, they were still significantly different. Visual analysis of the topology and Splitstree analyses (appendix II (N)) indicated that only four species (*Plantago muelleri*, *P. paradoxa*, *P. triandra* and *P. unibracteata*) provided conflicting signal and may be the cause of the low p-values of the HP test. However, the three datasets were still significantly incongruent when HP tests were run with these four species removed. Although traditionally used to check congruence of phylogenies from different regions, the HP test has been shown to be highly inaccurate even when the topologies of trees are congruent (Reeves *et al.* 2001; Yoder *et al.* 2001) and may continue to indicate conflict even after the incongruent taxa are

A: ITS



B: *ndhF-rpl32*

Figure 3.2: Maximum Likelihood phylogenies with branch lengths for each of the three datasets (A, ITS; B, *ndhF-rpl32*; C, *cox1*). The *cox1* tree displays a short branch length from the outgroup *Plantago* species (*P. coronopus* and *P. lanceolata*) to the *Veronica* species, and a long branch between the Australasian and outgroup *Plantago* species. In contrast, the ITS and *ndhF-rpl32* phylogenies have long branches from the outgroup *Plantago* to the *Veronica* species and a short branch between Australasian and outgroup *Plantago* species. Likelihood scores for these trees are -2513.31, -3045.43, and -8393.42, respectively.

C: *cox1*

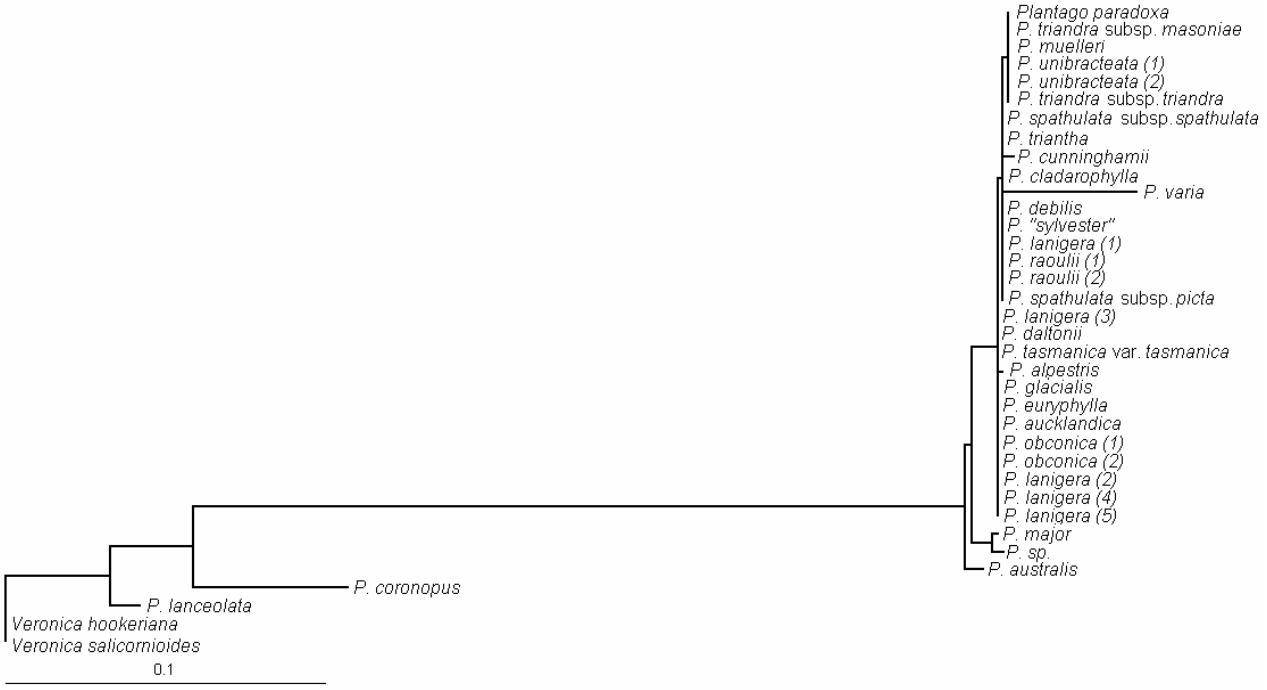


Figure 3.2 (continued): Maximum Likelihood phylogenies with branch lengths for each of the three datasets (A, ITS; B, *ndhF-rpl32*; C, *cox1*). The *cox1* tree displays a short branch length from the outgroup *Plantago* species (*P. coronopus* and *P. lanceolata*) to the *Veronica* species, and a long branch between the Australasian and outgroup *Plantago* species. In contrast, the ITS and *ndhF-rpl32* phylogenies have long branches from the outgroup *Plantago* to the *Veronica* species and a short branch between Australasian and outgroup *Plantago* species. Likelihood scores for these trees are -2513.31, -3045.43, and -8393.42, respectively.

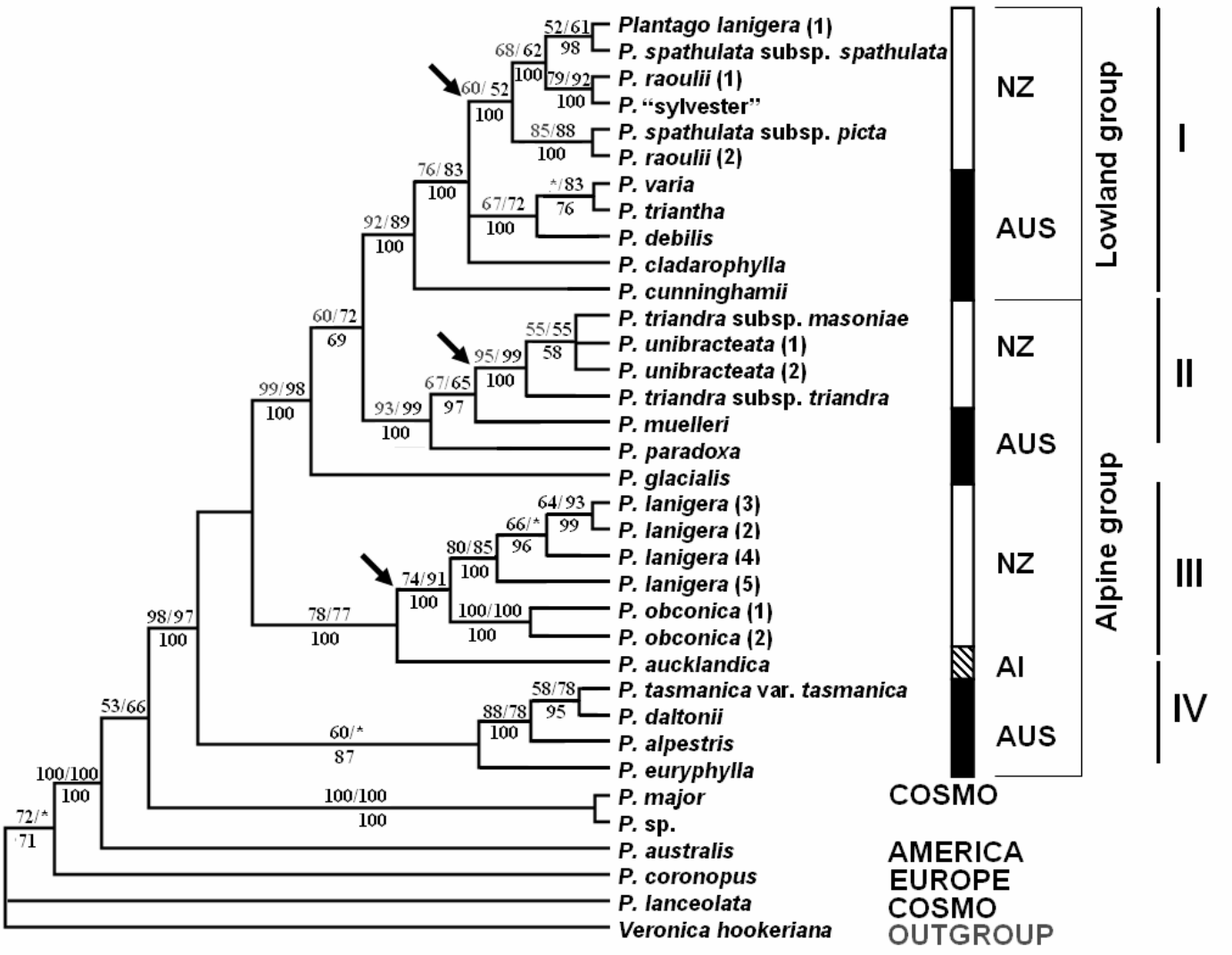


Figure 3.3. Maximum Likelihood (ML) tree reconstructed using concatenated data (ITS, *ndhF-rpl32* and *cox1*) showing indigenous distributions and habitats of native and introduced Australasian *Plantago* species (likelihood score = -7738.89). Values displayed above branches are from Maximum Parsimony/ML bootstrap analyses, whereas values presented below branches are posterior probabilities from MrBayes analyses (* indicates less than 50% support for the node). Arrows indicate a trans-Tasman dispersal event to New Zealand from an ancestral Australian population. AUS = Australia, AI = Auckland Islands, COSMO = cosmopolitan and NZ = New Zealand.

removed (Manos *et al.* 1999). Thus, analyses of a concatenated dataset were carried out because the tree topologies were not vastly incongruent in this case and there was no support for the placement of the conflicting species in the ITS tree. In the ITS dataset, some sequences had a few ambiguous sites that may result from multiple copies of the rDNA array in the genome. All ambiguous sites in the dataset were ignored during tree reconstruction analyses.

3.3.2 Phylogenetic analyses

The topologies of trees reconstructed using different tree construction methods were very similar for the ITS, *ndhF-rpl32* and *cox1* datasets (see appendix II (A-M)). Using one model versus individual models for each marker in MrBayes resulted in trees with the same topology. Posterior probabilities for trees made under Bayesian inference of phylogeny were higher compared to MP bootstrap values, and were similar to bootstrap support values obtained using ML. Therefore, only trees made using ML are presented here, with support values from all three analyses displayed on the concatenated tree (Fig. 3.3). ML trees for the individual datasets are displayed as phylograms in Fig. 3.2.

Concatenated dataset

In the tree reconstructed from the concatenated dataset (Fig. 3.3), the distant outgroup species *Veronica hookeriana* is in a polytomy with *Plantago lanceolata*. The nodes of these species are basal to the node of *P. coronopus*, which is in turn basal to the node of *P. australis*, followed by a clade containing *P. major* and *P.* sp. The Australasian species group into the clade following *P. australis*, within which there are four highly supported monophyletic groups. Clade IV is Australian and has a node that is basal to

the rest of the clades. Clade III comprises species from New Zealand, and *P. aucklandica* from the Auckland Islands. This clade is sister to *P. glacialis* (Australian), and Clades I and II, which are a mix of New Zealand and Australian species. Clade I and Clade II both contain a grade of Australian species at the base of a New Zealand clade. Thus, the pattern here is that there are multiple clades containing both New Zealand and Australian species, where the node of the New Zealand species appears in more derived positions in the tree than the node of the Australian species. In addition, Clade I is a group of Australasian *Plantago* found in lowland areas whereas Clades II, III, IV, along with *P. glacialis* are found in alpine and sub-alpine areas.

Individuals of *Plantago raoulii*, *P. lanigera*, and subspecies of *P. spathulata* and *P. triandra* do not form monophyletic groups, respectively, in the tree resolved using the concatenated dataset. This is discussed further in a separate paper focused on taxonomic issues and relationships of individuals at the population level (Chapter four). Supernetwork analysis using Splitstree shows support for the topology of the concatenated tree (appendix II (N)). The same four distinct groups are visible, with some uncertainties regarding the placement of *P. muelleri*, *P. paradoxa*, *P. triandra* subs. *triandra* and *P. unibracteata*, which was expected due to conflict between datasets.

A peculiar pattern found in the ML trees (Fig. 3.2) is a long branch separating the ingroup *Plantago* species from the outgroup *Plantago* species and an unexpectedly short branch length between *Plantago* as a whole and the *Veronica* outgroup in *cox1* (Fig. 3.2). In contrast, the ITS and *ndhF-rpl32* datasets have short branch lengths

between the ingroup species but a long branch between *Plantago* as a whole and the *Veronica* species.

ITS

In the ITS ML tree (Fig. 3.2A), the close outgroups differed in placement from the concatenated tree (*Plantago lanceolata* and *P. coronopus* have switched positions). *P. major*, *P.* sp. and *P. australis* form a clade that is nested within the Australasian species as opposed to having a node at the base of the tree in the concatenated phylogeny. The ingroup clades were similar but *P. muelleri* and *P. paradoxa* do not form a clade with *P. unibracteata* and *P. triandra*. Instead, *P. unibracteata* and *P. triandra* are placed at the base of the tree, whereas *P. muelleri* and *P. paradoxa* are placed as sister species to other Australian species further up in the tree. Furthermore, *P. debilis*, *P. varia* and *P. triantha* form a clade with *P. spathulata* subsp. *spathulata*. These results are incongruent with the concatenated tree and both organellar DNA trees.

ndhF-rpl32

The close outgroups (*P. lanceolata* and *P. coronopus*) form a clade, in contrast to the concatenated tree, where these species do not form a clade (Fig. 3.2B). When the dataset is analysed after concatenation, the topology of the reconstructed trees most resembled that of the *ndhF-rpl32* trees. This outcome may be because there is more signal in the *ndhF-rpl32* dataset.

Cox1

The *cox1* dataset resulted in very little resolution in the ML analysis (Fig. 3.2C), with two Australasian nodes resolved and only one with good bootstrap support (81% BP). Interestingly, the well supported clade is also present in the *ndhF-rpl32* dataset, and in the concatenated dataset, and is the clade that causes conflict between the ITS tree and the organelle trees. Poorly resolved nodes within the Australasian group may suggest a recent separation of Australian and New Zealand taxa, or unsuitability of this marker for reconstructing the phylogeny of this group of plants.

3.3.3 Molecular Dating

Testing Gondwanan origin of Australasian species

Molecular dating using ITS sequences rejected the hypothesis of Gondwanan vicariance for Australasian *Plantago*. When the divergence of the node including Australasian species, *P. major*, *P.* sp. and *P. australis* was set to 80mya, the final substitution rate estimated by BEAST was 5.806 (±0.240) $x10^{-10}$ substitutions/site/year. This rate is extremely low; more than ten times slower than the lowest substitution rate found in the literature for herbaceous angiosperms (1.72 $x10^{-9}$ (Kay *et al.* 2006)). When compared, the slowest rate and the observed rate were significantly different at the 1% level. Secondly, when the substitution rate was estimated at the lowest substitution rate found in herbaceous angiosperms, BEAST gave a divergence date of 24.804 (±0.663) mya for a clade of the Australasian species, *P. major, P.* sp. and *P. australis*. The node for the genus *Plantago* was estimated to be 57.717 (±0.700) mya. This is not consistent with the expectation that the Australasian *Plantago* group arose before New Zealand broke off from Gondwana (at least 80mya)

and the expected and observed values are significantly different. Thus, the Gondwanan hypothesis is rejected for Southern Hemisphere *Plantago*.

Estimating age of Australasian species

Molecular dating of the ITS dataset using BEAST provided a date of 2.291 (±0.0039) mya for the divergence a clade comprising the Australasian *Plantago* species, *P. major*, and *P.* sp. Unfortunately, there was not enough support in the tree to separate the Australasian clades from *P. major* and *P.* sp. in the ITS tree. However, the node of Clade I (excluding *P. cunninghamii*) (Fig. 3.3) was dated to 1.520 (±0.030) mya, which provides a range of possible dates of origin for the Australasian species.

3.4 Discussion

The aim of this study was primarily to investigate whether the distribution of the Australasian *Plantago* was the result of Gondwanan vicariance or long distance dispersal (or a mixture of both), and to elucidate biogeographic patterns for the Australasian species of *Plantago.* Even though Rahn (1996) attributed geographic disjunctions in the genus to vicariance and subsequent extinctions, the topology of the morphological phylogeny could also be attributed to multiple more recent dispersal events, and it was inferred from a recent molecular phylogeny that long distance dispersal is extremely common within the group (Rønsted *et al.* 2002). Subsequently, if dispersal was found to be the process behind the biogeography of Australasian *Plantago*, the goal of this study was to determine the number of and direction of dispersal events between Australia and New Zealand. This is important because many

large scale or austral (Swenson *et al.* 2001; von Hagen and Kadereit 2001) and even Australasian-focused plant phylogenetic studies (Radford *et al.* 2001; Perrie *et al.* 2007) lack signal in molecular markers between closely related Australasian taxa. This lack of signal often results in ambiguous relationships between Australasian taxa. The dataset used in this study was 2194 bp in length (2255 when indel characters are included as coded characters) and is able to resolve most of the relationships among Australasian *Plantago*.

3.4.1 Biogeography of Australasian *Plantago*

The hypothesis of Gondwanan vicariance can be rejected for Australasian *Plantago*. The concatenated tree (Fig. 3.3) reveals that multiple New Zealand clades are nested within the Australian species, which is consistent with a dispersal scenario; but not vicariance, as the cause of current distribution. There is also a lack of sequence divergence among the Australasian species, which we would not expect from a long isolation of the New Zealand species. Additionally, the nodes of the New Zealand clades are placed in a more derived position than nodes of the Australian species in the tree. Rønsted *et al.* (2002) estimated an Australasian clade of *P. spathulata* and *P. debilis* to be 0.5-0.7 mya (this clade was not found in the phylogeny presented here) and estimated 5.47 mya for the genus *Plantago* as a whole. All of this suggests that the genus is too young to be a remnant of Gondwana. This view is supported by our own molecular dating. Statistical tests rejected the hypothesis of Gondwanan vicariance for the Australasian species for both tests of: 1) when the node of the Australasian clade was set to 80mya (as would be expected for a Gondwanan vicariance hypothesis), or 2) when rates used for dating was set to the lowest known rate across herbaceous angiosperm lineages. Thus, both tree topology and molecular

dating strongly suggest that the Australasian *Plantago* group achieved their current distribution through long distance dispersal and not Gondwanan vicariance.

The clades formed by the Australasian species in the molecular phylogeny presented in this study can be interpreted as multiple directional, long distance trans-Tasman events. This is consistent with other plant phylogenetic studies that report Australian and New Zealand species as sister groups (Linder and Crisp 1995; Sanmartín and Ronquist 2004; Perrie and Brownsey 2005; Crayn *et al.* 2006; Meudt and Simpson 2006). The lack of sequence divergence between Australian and New Zealand species is especially evident in analysis of the *cox1* dataset, where several Australian and New Zealand species share identical sequences, suggesting recent diversification/speciation of the trans-Tasman species. The timing of arrival and diversification of the New Zealand species matches the geological dates of the uplift of the New Zealand mountains (5-2mya) (Winkworth *et al.* 2002b).

At least three independent long distance dispersal events to New Zealand, followed by speciation are evident in the reconstructed phylogeny (Fig. 3.3). All three dispersals appear to be directional from Australia to New Zealand. There also appears to be two dispersal events from Australia to the Auckland Islands. Dispersal in this direction is congruent with expectations because prevailing winds from the West would result in frequent west to east direction of dispersals (refered to as West Wind Drift) (Raven 1973; Winkworth *et al.* 2002b; Sanmartín *et al.* 2007). The finding of multiple dispersals across the Tasman differs from the common pattern of one dispersal followed by rapid radiation found in many other New Zealand plant groups (Wagstaff and Garnock-Jones 1998; Winkworth *et al.* 1999; Perrie *et al.* 2003; Albach *et al.*

2005b; Meudt and Simpson 2006). However, the fact that there are many cosmopolitan species in *Plantago* and even small oceanic islands have a mixture of native and introduced species from the genus indicates that the plants are able to disperse over long distances fairly frequently. Within the Australasian taxa, species are commonly distributed across geographic barriers that often separate other plants, like the Cook Strait in New Zealand, which separates the North and South Islands, and Bass Strait, which separates Tasmania from Australia.

The three dispersal events from Australia reflect two independent dispersals to the alpine/sub-alpine region in New Zealand (and the subantarctic islands) and one to the lowlands, where the closest Australian relative is also from similar habitats (Fig. 3.3). One explanation for this pattern is that the New Zealand species may have evolved from a common ancestor that had already established in that bioclimatic zone (i.e., alpine to alpine, lowland to lowland), which may have increased the success of establishment in New Zealand. This has been suggested previously (e.g. Winkworth *et al.*, 2002), albeit in the other direction. For example, it was suggested that *Craspedia* radiated from only one successful dispersal event because of establishment difficulties (Ford *et al.* 2007). Further, many New Zealand genera that contain mostly alpine species also contain lowland species, although patterns of habitat colonization are not clear (Buckley and Simon 2007). The lowland New Zealand species form a clade that is nested within an alpine clade, which suggests that alpine groups have subsequently dispersed to coastal regions (Fig. 3.3). In this study, we show that in the Australasian species, the alpine species appear to have one migration to lowland habitats, followed by an additional dispersal across the Tasman. More studies regarding the New Zealand alpine species are needed to determine patterns in habitat preferences.

Multiple dispersals across the Tasman appear fitting for *Plantago*, which is widely distributed. However, dispersal mechanisms of the group are largely unknown. Many *Plantago* seeds produce mucilage when wet and it has been hypothesized that a method of dispersal is by sticking to oceanic birds (Rønsted *et al.* 2002). Another New Zealand plant group that has also undergone multiple dispersals is *Lepidium* (Mummenhoff *et al.* 2001), which also has mucilaginous seeds. Other *Plantago* species, such as *P. lagopus*, have hairs on their capsules, which have been shown to facilitate overland dispersal by sticking to the coat of migrating ungulates (Manzano and Malo 2006). Seeds may also be dispersed by wind or water but there is no evidence for this.

We have shown that Australasian *Plantago* have undergone long distance dispersal but we are unable to make inferences about the relationships between South American and Australasian species of *Plantago* due to a lack of South American material. However, ITS sequences for 13 South American *Plantago* species were obtained from GenBank. Inclusion of these sequences in phylogenetic analyses (data not shown here; see Chapter 4) still shows support for long distance dispersal throughout the Southern Hemisphere, i.e. South American, Australian and New Zealand species do not form monophyletic groups in the tree, and the Australian species are more closely related to the New Zealand species than they are to South American species. Thus, inferences made from this study remain valid.

3.4.2 Incongruence of phylogenies from nuclear DNA and organellar DNA

In this study, trees reconstructed from the different genomes yielded similar topologies but there were a few discrepancies. The Australian species *P. muelleri* and

P. paradoxa and the New Zealand species of *P. triandra* and *P. unibracteata* appear to have conflicting evolutionary histories of nuclear and organellar genomes. While the four species form a clade in the *ndhF-rpl32* tree and the tree of the concatenated dataset, they are interspersed between other Australasian species in the ITS tree. One explanation for this conflict may be hybridisation, which is not surprising because the group is relatively young. Although trans-Tasman hybridisation appears unlikely, it has been reported to occur in plants (Perrie and Brownsey 2005). The conflict may also be a result of very little sequence divergence in the ITS sequences, which does not resolve the placement of *P. triandra* and *P. unibracteata* in the tree. Further analysis of the ITS dataset suggests that lack of resolution may be the most likely explanation for the conflict (see Chapter four).

3.4.3 Reduced rates of mtDNA in Australasian *Plantago*

Finally, the elevated rates of mtDNA substitutions in *Plantago* that were reported by Cho *et al.* (2004) were not evident in the Australasian species. There was a long branch between close *Plantago* outgroups and ingroup *Plantago* but a short branch between *Plantago* and *Veronica* in the *cox1* marker. The opposite was found in the other two genomic markers. Any long branch is again reduced in the Australasian species, which may be a result of recent rapid radiations, a process not uncommon in endemic New Zealand groups (Wagstaff and Garnock-Jones 1998; Heenan *et al.* 2002; Murray *et al.* 2004) or a decrease in substitution rate, which has been found in other lineages of *Plantago* (Cho *et al.* 2004). The latter is more plausible, as there is no evidence of reduced substitution rates in ITS or *ndhF-rpl32* sequences in the Australasian species, which would be expected if the reduction in the *cox1* substitution rates was a result of recent rapid radiations of the Australasian species.

3.4.4 Summary and future directions

In summary, this study has shown evidence for long distance dispersal between Australia and New Zealand for the native Australasian *Plantago*. This study revealed an uncommon pattern of three independent directional dispersals across the Tasman Sea from Australia to New Zealand. In all cases, the ancestors of the dispersed species appear to have been pre-adapted to the habitats they established in. It was shown through molecular dating that the node of the Australian clade is about 2.291 to 1.520 mya. Further studies are needed to clarify if there are any more trans-Tasman dispersals and if trans-Tasman hybridisation is possible, if the high rate of mtDNA evolution has slowed down in *Plantago* from other locations, and to determine the dispersal mechanism of this group. Lastly, the dataset could be expanded to include all known Southern Hemisphere species in order to reveal other Southern Hemisphere patterns and relationships.

Chapter Four: Evolutionary patterns and species boundaries in New Zealand *Plantago* (Plantaginaceae) species using ITS

Abstract

Major geological and climatic changes in New Zealand have affected the evolution of plants and are linked to events such as recent speciation. Recent speciation in many New Zealand plant lineages may have given rise to large groups with high morphological variation, for which taxonomy is difficult. In this study, phylogenetic analyses were carried out for the Australasian species of *Plantago* (Plantaginaceae) using Internal Transcribed Spacer (ITS) regions from nuclear DNA to investigate evolutionary patterns and species boundaries. An ITS phylogeny suggests that the *Plantago* species are separated allopatrically corresponding to habitat type and altitude (e.g. coastal, lowland or alpine/subalpine). Sympatric separation may also have occurred in the past and is influenced by various factors including polyploidy and different flowering times. The phylogeny also revealed that the species boundaries of several New Zealand species, including *P. spathulata, P. triandra*, and classifications of *P. obconica, P. triantha* and *P. aucklandica* may need revision. Low sequence divergence but high morphological variation was found in this study and could be caused by various factors, including recent speciation, presence of species complexes, hybridisation, and/or interbreeding.

4.1 Introduction

4.1.1 Evolution of New Zealand plant groups

Since its separation from Gondwana, New Zealand has undergone major geological and climatic changes, which have affected the evolutionary history and distribution of its biota (McDowall 2008). The main large-scale climatic and geological events that may have had an impact on speciation patterns of groups with recent origins are: 1) Pleistocene glaciation cycles (Trewick and Morgan-Richards 2005; McDowall 2008) and 2) final uplift of the Southern Alps ca. 5mya (Cox and Findlay 1995; Batt *et al.* 2000; Haase *et al.* 2007). These events may have had a dramatic effect on speciation patterns by causing physical separation of populations or through the creation of new habitats, which allow for divergence of populations as adaptations to these new habitats take place. Over time, the separation of populations could have led to speciation. Cook Strait, which separates the North and South Islands, may also have been a barrier to gene exchange in the past, although there may have been land bridges connecting the North and South Islands during the Pliocene glacial cycles (Lewis *et al.* 1994). Other phylogeographic hypotheses leading to speciation include glacial refugia (Wright *et al.* 2000; Greaves *et al.* 2007), displacement along the alpine fault (Buckley and Simon 2007), or separation by distance. Sympatric speciation is possible as well; a common pattern in New Zealand plant groups is recurrent polyploidy (Connor 1985), which may lead to rapid speciation of polyploids (i.e. when they are sympatric with parental lineages).

Molecular data have been used successfully to investigate patterns of evolution of New Zealand biota. For example, phylogenetic analyses using molecular data showed evidence that the *Hebe sensu lato* (*Veronica*) complex had two radiations: one coinciding with uplift of the Southern Alps and the other with Pleistocene glaciation cycles (Wagstaff and Garnock-Jones 1998). In addition, it has been found that many groups have dispersed from the South Island (S. I.) to the North Island (N. I.), and that lowland species often have a different origin from the alpine species (e.g. Lockhart *et al.* 2001).

Recent molecular studies found that many New Zealand plant groups show evidence of recent radiations. These events occurred either following long distance dispersal (Breitwieser *et al.* 1999; Winkworth *et al.* 1999; McGlone *et al.* 2001; Winkworth *et al.* 2002b; Sanmartín and Ronquist 2004), or are associated with physical changes such as uplift of mountain ranges and climate fluctuations (Raven 1973; Lockhart *et al.* 2001; Winkworth *et al.* 2002a). Due to recent or incomplete speciation, species boundaries in New Zealand plants may be unclear using morphological characters and this often results in taxonomic problems. Many New Zealand genera are large and have high morphological variation both within and among species, such as *Myosotis* (Winkworth *et al.* 1999) and multiple genera within Gnaphalieae (Smissen *et al.* 2004). However, most genera studied to date are also found to have little genetic variation in contrast to morphology (Winkworth *et al.* 1999; Heenan *et al.* 2002; Vijverberg *et al.* 2002; Winkworth *et al.* 2002a).

4.1.2 Taxonomy of New Zealand plant groups

The increasing use of molecular data has also helped with the systematics of many New Zealand plant groups (e.g. Heenan *et al.* 2002; Glenny 2004; Perrie and Brownsey 2005; Ford *et al.* 2007). Molecular data led to the conclusion that *Veronica* (previously circumscribed to exclude the Southern Hemisphere *Hebe* complex) was paraphyletic (Albach and Chase 2001) and the genus was subsequently enlarged by addition of *Hebe* and its relatives (Garnock-Jones *et al.* 2007). Molecular data (along with karyological information) also revealed that two species of *Crassula* (*C. hunua* and *C. ruamahanga*) were actually a species complex with interspecific hybridisation and polyploidy and these were reduced to one species (de Lange *et al.* 2007). Additionally, molecular data have been shown in several studies to provide better resolution when morphology and molecular data are congruent (Winkworth *et al.* 2002a; Meudt and Simpson 2007).

4.1.3 The genus *Plantago*

The aim of this study was to reconstruct a molecular phylogeny for the genus *Plantago* (Plantaginaceae), focusing on species relationships between and delimitations within the Australasian species. *Plantago* is a large genus of wind-pollinated plants, with over 200 species distributed worldwide. The genus has three basic chromosome numbers (x = 4, 5, and 6) (Rahn 1996; Dhar *et al.* 2006), and diploid chromosome numbers range from $2n$ = 8 (*P. ovata*) to $2n$ = 96 (*P. correae* and *P.* "sylvester"). While extensive morphological variation can occur within each species because of the plasticity of characters in some species, at the same time there may be few morphological differences between some species (Rock 1920; Sykes in Webb *et al.* 1998).

Phylogenetic relationships, species boundaries and evolutionary patterns among Australasian *Plantago* species in particular are not well understood. To date, one morphological and two molecular studies have reconstructed *Plantago* phylogeny with the inclusion of some Australasian species. The morphological phylogeny of all *Plantago* species was the first phylogenetic study of the genus and included most of the Australasian species (one New Zealand and four Australian species were not included) (Rahn 1996).

In the first study using genetic sequences, Rønsted *et al.* (2002) combined ITS (Internal Transcribed Spacer) and *trnL-F* sequences to reconstruct the phylogeny of 57 *Plantago* species (including *P. spathulata* and *P. raoulii* from New Zealand, and *P. debilis* from Australia/Tasmania). In the second molecular phylogenetic study, Hoggard *et al.* (2003) obtained ITS sequences for 23 *Plantago* species (including *P. tasmanica*, *P. daltonii*, *P. euryphylla*, *P. hispida* and *P. paradoxa*, which are endemic to Australia/Tasmania, and *P. triandra* from New Zealand). The latter focused on the relationships of *Littorella*, which was included as a subgenus within *Plantago* by Rahn (1996).

These molecular phylogenies conflict with the morphological phylogeny mainly in the circumscription of several subgenera, e.g. it was found that subg. *Albicans* was paraphyletic (Rønsted *et al.* 2002). Species of *Littorella* were found to be in a sister clade to a clade of all other *Plantago* species and it was subsequently suggested that *Littorella* be recognised at genus rank (Hoggard *et al.* 2003). The recent molecular

Table 4.1. A comparison of previous taxonomic treatments of New Zealand *Plantago*. Shaded areas indicate species and subspecies accepted in the most recent Flora of New Zealand (Webb *et al.* 1988) and are also the classifications that are used in this study.

Name	Type	Hooker 1864	Cheeseman 1906	Cheeseman 1925	Moore (in Allan 1961)	Sykes (in Webb 1988)
Plantago aucklandica Hook.f. 1844	Auckland Islands (K)	*aucklandica*	*aucklandica*	*aucklandica*	*aucklandica*	*aucklandica*
P. triantha Spreng.1825, ≡ *P. carnosa* R.Br., non Lam., ≡ *P. subantarctica* Cockayne 1928 nom. illeg. (superfl.) = (≡?) *P. brownii* Rapin (1827)	(K/BM?) the name should be based on R.Br. specimens	As *brownii* Rapin	As *brownii* Rapin	As *brownii* Rapin	*triantha*	*triantha*
P. triandra Bergg. 1877	Kelly Range (Lund?)	–	*triandra*	*triandra*	*triandra*	*triandra* subsp. *triandra*
P. hamiltonii Kirk 1879	Greymouth (WELT)	–	*triandra*	*triandra*	*triandra*	*triandra* subsp. *triandra*
P. masoniae Cheesem. 1921 (as *masonae*)	Manaia (AK)	–	–	*masoniae* (as *masonae*)	*triandra*	*triandra* subsp. *masoniae*
P. spathulata Hook.f. 1854	Pahawa (Pahaoa) (K; iso WELT?)	*spathulata*	*spathulata*	*spathulata*	*spathulata*	*spathulata* subsp. *spathulata*

Table 4.1. A comparison of previous taxonomic treatments of New Zealand *Plantago*. Shaded areas indicate species and subspecies accepted in the most recent Flora of New Zealand (Webb *et al.* 1988) and are also the classifications that are used in this study.

Name	Type	Hooker 1864	Cheeseman 1906	Cheeseman 1925	Moore (in Allan 1961)	Sykes (in Webb 1988)
P. picta Colenso 1890	Tolaga Bay (WELT & K)	–	In syn of *P. brownie*	In syn of *P. brownie*	*picta*	*spathulata* subsp. *picta*
P. raoulii Decne 1852	Banks Pen. (P)	*raoulii*	*raoulii*	*raoulii*	*raoulii*	*raoulii*
P. dasyphylla Colenso 1892	Dannevirke (WELT)	–	*raoulii*	*raoulii*	*raoulii*	*raoulii*
P. novae-zelandiae Moore 1961	Mt Hikurangi (WELT)	–	*P. brownii* auctt. NZ non Rapin	*P. brownii* auctt. NZ non Rapin	*novae-zelandiae*	*lanigera*
P. lanigera Hook.f. 1864	Otago Lake Distr., (K &WELT)	*lanigera*	*lanigera*	*lanigera*	*lanigera*	*lanigera*
P. lanigera var. ***petriei*** Cheesem. 1906	Mt Kyeburn (AK)	–	*lanigera* var. *petriei*	*lanigera* var. *petriei*	*lanigera*	*lanigera*
P. unibracteata Rahn 1996, ≡ *P. uniflora* Hook.f. non L.≡ *P. triandra* var. *uniflora* (Hook.f.) Pilger	Ruahine Mts (K)	*uniflora* Hook.f.	*uniflora* Hook.f.	*uniflora* Hook.f.	*uniflora* Hook.f.	*uniflora* Hook.f.
P. obconica Sykes 1988	(CHR)	–	–	–	–	*obconica*

phylogenies indicate that Australasian *Plantago* species cluster together and that *P. stauntoni* from Amsterdam & St. Paul Islands is nested within the Australasian group (Rønsted *et al.* 2002; Hoggard *et al.* 2003). However, relationships among the Australasian species and the placement of *P. stauntoni* were either unsupported or unresolved. Little is known regarding the evolutionary history of Australasian species because few Australasian species were included in these molecular phylogenies.

4.1.4 New Zealand *Plantago*

In New Zealand, eight native species of *Plantago* were accepted in the most recent Flora (Sykes in Webb *et al.* 1998; Table 4.1) and one (*P.* "sylvester") is an undescribed polyploid (Groves and Hair 1971). The chromosome number of $2n = 96$ reported by Groves and Hair (1971) was confirmed using one sample from *P.* "sylvester" populations obtained in this study. Twenty four species are native to Australia (Briggs 1992). All of the Australasian species are placed in subg. *Plantago* and mainly in sect. *Oliganthos* and sect. *Mesembrynia*; one species from the Auckland Islands (*P. aucklandica*) is placed in sect. *Plantago* (Rahn 1996).

Out of the eight native New Zealand species, *Plantago lanigera*, *P. obconica*, *P. unibracteata*, *P. aucklandica* and *P triandra* subsp. *triandra* are found in the alpine/subalpine region, whereas *P. triantha*, *P. triandra* subsp. *masoniae* and *P. spathulata* subsp. *picta* are found in lowland coastal areas (although sea level is considered sub-alpine in the Auckland Islands where *P. triantha* is found). *P. spathulata* subsp. *spathulata* and *P. raoulii* are found in both lowland and subalpine areas from coastal regions to altitudes of ca. 1000 and 1500m, respectively (Sykes in Webb *et al.* 1988).

A. *Plantago spathulata* subsp. *spathulata* (Marfells Beach)

B. *P. raoulii* (Owhiro Bay)

C. *P.* "sylvester" (Lake Sylvester)

D. *P. lanigera* (Shotover Saddle)

Figure 4.1. Morphological habits of cultivated samples of the New Zealand *Plantago* species, illustrating morphological plasticity of the plants. Figures on the left were photographed in February 2007, whereas the figures on the right were taken in cultivation in November 2007.

E. *P. triandra* subsp. *masoniae* (Manaia)

F. *P. triandra* subsp. *triandra* (Lake Sylvester)

G. *P. unibracteata* (Mt. Wilberg)

H. *P. obconica* (Cardrona Skifield)

Figure 4.1. Morphological habits of cultivated samples of the New Zealand *Plantago* species, illustrating morphological plasticity of the plants. Figures on the left were photographed in February 2007, whereas the figures on the right were taken in cultivation in November 2007.

Like all *Plantago*, morphological differences in the New Zealand species are not easily defined, resulting in problems with species delimitation and identification. Although Rahn (1996) reported that hair and seed morphology appeared to be the most useful characters for the genus, characters such as leaf size, shape, hairiness and teeth, scape length, and the number of flowers may vary within a population according to environmental factors (Sykes in Webb *et al.* 1988; see also Fig. 4.1). Taxonomic discrepancies are therefore common between previous Flora treatments (Table 4.1) and are used to represent hypotheses of species boundaries in this study using a molecular phylogenetic approach. Below, I discuss the taxonomy of the New Zealand *Plantago* species according to the order that they later appear in the ITS phylogeny and the specific hypotheses that will be tested in each case, including testing species and subspecies delimitations (*P. spathulata*, *P. raoulii*, *P. lanigera* and *P. triandra*) and closest relatives (*P.* "sylvester" and *P. obconica*).

Species and subspecies delimitations

Plantago spathulata

Sykes (in Webb *et al.* 1988) included two subspecies under *P. spathulata*, subsp. *spathulata* (found on southern N. I. coasts, and both coasts and inland in the S. I.) and subsp. *picta* (endemic to the coastal Poverty Bay region and East Cape in Gisborne). These were previously treated as two species, *P. spathulata* and *P. picta* (Table 4.1). Separation of the two was based on morphological characters, such as a persistent taproot with no adventitious roots and scapes up to 7 cm long (*P. picta*), whereas *P. spathulata* only had adventitious roots and scapes of 3-12-(22) cm long. Sykes (in Webb *et al.* 1988) recognized these entities at the lower rank of subspecies because populations of the two were found close to one another geographically, and the

characters used to distinguish the two species by Moore (in Allan 1961) were inconsistent due to plastic morphology within populations. Molecular data can be used to test the hypothesis that these two are conspecific by reconstructing a population level phylogeny that includes representatives of both groups.

Plantago raoulii

P. raoulii has been noted to have two different forms, with one form widespread on both the North and South Islands, and the other common in Taranaki and around the Cook Strait (Sykes in Webb *et al.* 1988). The latter is described as having thicker leaves that are also wider and shorter. Both forms can be found in coastal regions and in altitudes of up to 1500 m. The hypothesis of whether the narrow-leaved and broad-leaved individuals can be separated based on DNA sequence data is tested in this group using a population level phylogeny for both forms.

Plantago lanigera

P. lanigera as circumscribed by Sykes (in Webb *et al.* 1988) was previously considered to be two distinct species (e.g. Moore in Allan 1961 treated the group as as *P. lanigera* and *P. novae-zelandiae*, see Table 4.1). Various morphological characters were used to differentiate between the two forms, such as crowded long hairs, rhomboid shaped lamina, and bracts <2 mm long in *P. lanigera* vs. sparse long hairs, elliptic lamina, and bracts 2.5-3 mm long in *P. novae-zelandiae*. Sykes (in Webb *et al.* 1988) recognized only one species (*P. lanigera*) because these morphological characters were inconsistent and the ranges of populations of the two forms overlap. It was found in a separate study that plants identified as *P. lanigera* had a chromosome number of $2n = 12$, whereas plants identified as *P. novae-zelandiae* exhibit both $2n =$

12 and 2*n* = 24 (Spence and Sykes 1989). While the means of morphological characters such as seed size, shape and weight, ovule numbers, number of leaves per plant, number of capsules per inflorescence, number of seeds per capsule, and inflorescence height were significantly different between the two forms, the measurement ranges of most of these characters overlap (and therefore cannot conclusively be used to separate one species from the other); only seed weight and size were found to be useful in this respect (Spence and Sykes 1989). The hypothesis that these two forms (both found in alpine regions) are conspecific is tested using a population level phylogeny that includes multiple representatives of both forms.

Plantago triandra

Two subspecies of *P. triandra* were recognized by Sykes (in Webb *et al.* 1988). Subsp. *masoniae* (found in coastal areas) is described as having fleshier leaves with smaller sessile flowers compared to subsp. *triandra* (which are mostly found on inland mountains). Only one species (*P. triandra*) was recognised by Moore (in Allan 1961) because morphological characters used to distinguish between the coastal and inland forms were regarded as inconsistent and also very plastic. *P. triandra* has also been previously treated as three species (*P. triandra*, *P. masoniae* and *P. hamiltonii*) based on morphology, but *P. hamiltonii sensu lato* was regarded as synonymous with *P. triandra* subsp. *triandra* in the most recent Flora (Sykes in Webb *et al.* 1988). Molecular data will be used to test the hypothesis that two different forms exist in *P. triandra* using a population level molecular phylogeny to evaluate if sequence divergence is present between subsp. *triandra* and subsp. *masoniae*.

Species relationships

Plantago "sylvester"

Moore (in Allan 1961) and Sykes (in Webb *et al.* 1988) both mention several plants in northwest Nelson with several different morphological characters in the description for *P. raoulii*. Based on morphology, Moore included the plants in *P. spathulata*, whereas Sykes grouped them with *P. raoulii*. Chromosome counts made for the plants (given the tag name *P.* "sylvester") suggest that they are 16-ploids with a chromosome number of $2n = 96$, whereas *P. raoulii* and *P. spathulata* have a chromosome number of $2n = 48$ (Groves and Hair 1971). Thus, using a molecular phylogeny, the relationship of *P.* "sylvester" is tested (i.e. is it more closely related to *P. raoulii* or *P. spathulata*?)

Plantago obconica

P. obconica is a recently described species found in alpine regions. It is suggested to be closely related to *P. triandra* subsp. *triandra* because of morphological similarities such as fine subulate leaves and a calyx that is shorter than the capsule (Sykes 1988). The species has not previously been included in any phylogenetic studies. The hypothesis that *P. obconica* is closely related to *P. triandra* is tested by examining relationships in a species level phylogeny.

4.1.5 Study aims

In this chapter, an ITS phylogeny of Australasian *Plantago* species, integrated with other *Plantago* species for which ITS sequences are readily available, is presented.

Table 4.2. Details of *Plantago* samples that were included in the ITS phylogeny. Samples that were collected and sequenced (DNA) for this study are indicated in bold, with details of collection location.

Ref. in Rahn (1996)	Species	Indigenous distribution	Collection location	Accession number	Reference or voucher
155	*Plantago afra*	S Africa		AY101892	Rønsted *et al.* 2002
184	*P. albicans*	Mediterranean		AY101905	Rønsted *et al.* 2002
75	***P. alpestris***	**Australia**	**Kosciuszko National Park, NSW, AUS**		**NSW742962 (*BGB9748*) / NSW742963 (*BGB9749*)**
132	*P. alpine*	Europe		AY101877	Jensen *et al.* 1996
175	*P. amplexicaulis*	Mediterranean		AY101900	Rønsted *et al.* 2002
146	*P. arborescens*	Macaronesia		AY101886 / AJ548954	Rønsted *et al.* 2000 / Hoggard *et al.* 2003
154	*P. arenaria*	Mediterranean		AY101891 / AY692082	Rønsted *et al.* 2002 / Dhar *et al.* 2006
210	*P. aristata*	E USA		AY101911 / AJ548983	Rønsted *et al.* 2002 / Hoggard *et al.* 2003
29	*P. asiatica*	S & E Asia		AY101862 / AJ548977	Rønsted *et al.* 2002 / Hoggard *et al.* 2003
166	*P. atrata*	Europe, W Asia		AY101895	Rønsted *et al.* 2002
15	***P. aucklandica***	**Auckland Islands**	**Auckland Islands (cult. in Otari-Wilton Bush, Wellington, NZ)**		**Otari cult. (*2003.2037*)**
108	***P. australis***	**America**	**Hunua Ranges, Auckland, NZ**		**WELTU20181 (*MLT019 & PGJ)***
			Waverly, NZ (cult. population)		**WELTU20182 (*Colin Ogle s. n.*)**
				AY101874 / AF313038 /	Rønsted *et al.* 2000 / Albach & Chase 2001
178	*P. bellardii*	Mediterranean		AY101902	Rønsted *et al.* 2000
52	*P. camtschatica*	E Asia		AJ548971	Hoggard *et al.* 2003
190	*P. ciliata*	Mediterranean		AY101906	Rønsted *et al.* 2002
67	***P. cladarophylla***	**Australia**	**Barrington Tops National Park, NSW, AUS**		**NSW744803 (*J. R. Hosking 2682*)**
23	*P. cornuti*	S Europe		AY101859	Rønsted *et al.* 2002
140	***P. coronopus***	**Mediterranean, Europe**	**Island Bay, Wellington, NZ**		**WELTU20183 (*PGJ2549*)**

Table 4.2. Details of *Plantago* samples that were included in the ITS phylogeny. Samples that were collected and sequenced (DNA) for this study are indicated in bold, with details of collection location.

Ref. in Rahn (1996)	Species	Indigenous distribution	Collection location	Accession number	Reference or voucher
140	***P. coronopus***	**Mediterranean, Europe**	**Island Bay, Wellington, NZ**	AY101882 / AJ548987	Rønsted *et al.* 2002 / Hoggard *et al.* 2003
137	*P. crassifolia*	Mediterranean, S Africa		AY101881	Rønsted *et al.* 2002
176	*P. cretica*	E Mediterranean		AY101901	Rønsted *et al.* 2002
57	***P. cunninghamii***	**Australia**	**Brigalow Park Nature Reserve, NSW, AUS**		**NSW744804 (*J. R. Hosking 2752*)**
74	***P. daltonii***	**Tasmania**	**St. Clair National Park, AUS, Tasmania**		**NSW743874 (*BGB9782*)**
				AJ548968	Hoggard *et al.* 2003
60	***P. debilis***	**Australia**	**Barrenjoey Headland, NSW, AUS**		**NSW 742894 (*BGB9738*)**
				AY101868	Rønsted *et al.* 2002
46	*P. elongata*	W USA		AJ548974	Hoggard *et al.* 2003
207	*P. erecta*	W USA		AY101909 / AJ548982	Rønsted *et al.* 2002 / Hoggard *et al.* 2003
72	***P. euryphylla***	**Australia**	**Kosciuszko National Park, NSW, AUS**		**NSW743824 (*BGB9760*) / NSW743822 (*BGB9758*) / NSW742956 (*BGB9743*) / NSW (*BGB9741*)**
				AJ548966	Hoggard *et al.* 2003
148	*P. famarae*	Macaronesia		AY101888	Rønsted *et al.* 2002
129	***P. glacialis***	**Australia**	**Kosciuszko National Park, NSW, AUS**		**NSW742960 (*BGB9746*) / NSW743813 (*BGB9753*)**
45	*P. heterophylla*	SE USA		AJ548975	Hoggard *et al.* 2003
77	*P. hispida*	Tasmania, E Australia		AJ548967	Hoggard *et al.* 2003
212	*P. hookeriana*	S USA		AY101913	Rønsted *et al.* 2002
169	*P. lagopus*	Mediterranean		AY101897 / AY692078	Rønsted *et al.* 2002 / Dhar *et al.* 2006
170	***P. lanceolata***	**cosmopolitan**	**Karori, Wellington, NZ**		**WELTU20184 (*PGJ2551*)**
				AY101898 / AF313036 / AJ548984 / AY692077	Albach & Chase 2001 / Rønsted *et al.* 2002 / Hoggard *et al.* 2003 / Dhar *et al.* 2006

Table 4.2. Details of *Plantago* samples that were included in the ITS phylogeny. Samples that were collected and sequenced (DNA) for this study are indicated in bold, with details of collection location.

Ref. in Rahn (1996)	Species	Indigenous distribution	Collection location	Accession number	Reference or voucher
120, 122	***P. lanigera***	**New Zealand**	**Hall Range, Canterbury, NZ**		**WELTU20131 (*PBH s. n.*) / WELTU20143 (*PBH s. n.*)**
			Sealy Tarns, Mt. Cook, NZ		**WELTU20125 (*PBH s. n.*) / WELTU20124 (*PBH s. n.*)**
			Sugarloaf Pass, Otago, NZ		**WELTU20133 (*Mike Thorsen s. n.*)**
			The Remarkables, Queenstown, NZ		**WELTU20142 (*PGJ2611 et al.*) / WELTU20141 (*PGJ2607 et al.*) / WELTU20128 (*MLT040 et al.*) / WELTU20140 (*MLT036 et al.*) / WELTU20136 (*MLT039 & PGJ*)**
			Mt. Hikurangi, East Cape, NZ		**WELTU20134 (*Mike Thorsen s. n.*)**
			Wilberg Range, Westland, NZ		**WELTU20147 (*MLT027 et al.*)**
			Shotover Saddle, Otago, NZ		**WELTU20132 (*MLT034 & PGJ*)**
			Cardrona Skifield, Wanaka, NZ		**WELTU20138 (*MLT031 et al.*)**
			Rock and Pillar Range, Otago, NZ		**WELTU20130 (*HMM273/2 & BS*)**
			Ruahine Ranges, North Island, NZ		**WELTU20145 (*MLT051 et al.*) / WELTU20139 (*MLT052 et al.*)**
			Tararua Ranges, Wellington, NZ		**WELTU20123 [Otari cult. (*2005.0122*)]**
			Thomson Mountains, Otago, NZ		**WELTU20148 [Otari cult. (*2005.0112*)]**
			Eyre Mountains, Otago, NZ		**CHR580877 (*Kerry Ford s. n.*)**
171	*P. leiopetala*	Madeira		AY101899 / AJ548985	Rønsted *et al.* 2002 / Hoggard *et al.* 2003
194	*P. lundborgii*	San Ambrosio Is.		AY101907	Rønsted *et al.* 2002
141	*P. macrorhiza*	Mediterranean		AY101883	Rønsted *et al.* 2002
26	***P. major***	**cosmopolitan**	**Kingsland, Auckland, NZ**		**WELTU20179 (*MLT018 & LT*)**
			Karori, Wellington, NZ		**WELTU20180 (*PGJ2550*)**
				AY101861 / AY692079	Rønsted *et al.* 2003 / Dhar *et al.* 2006

Table 4.2. Details of *Plantago* samples that were included in the ITS phylogeny. Samples that were collected and sequenced (DNA) for this study are indicated in bold, with details of collection location.

Ref. in Rahn (1996)	Species	Indigenous distribution	Collection location	Accession number	Reference or voucher
135	*P. maritima*	cosmopolitan		AY101879 / AJ548986	Rønsted *et al.* 2002 / Hoggard *et al.* 2003
151	*P. mauritanica*	NW Africa		AY101890	Rønsted *et al.* 2002
40	*P. maxima*	E Europe, C Asia		AY101864	Rønsted *et al.* 2002
41	*P. media*	Europe, C Asia		AY101865 / AJ548964	Rønsted *et al.* 2002 / Hoggard *et al.* 2003
131	***P. muelleri***	**Australia**	**Kosciuszko National Park, NSW, AUS**		**NSW743812 (*BGB9752*) / NSW742951 (*BGB9742*)**
91	*P. myosuros*	S America		AY101873	Rønsted *et al.* 2002
168	*P. nivalis*	S Spain		AY101896	Rønsted *et al.* 2000
162	*P. nubicola*	Peru, Bolivia, NW Argentina		AJ548972	Hoggard *et al.* 2003
–	***P. obconica***	**New Zealand**	**Hector Mountains, Otago, NZ**		**CHR573261**
			Cardrona Skifield, Wanaka, NZ		**WELTU20122 (*MLT030 et al.*) / WELTU20121 (*PGJ2600 et al.*)**
179	*P. ovata*	Mediterranean,W USA		AY101903 / AJ548973 / AY692076	Rønsted *et al.* 2002 / Hoggard *et al.* 2003 / Dhar *et al.* 2006
24	*P. palmata*	trop. Africa		AY101860	Rønsted *et al.* 2002
	P. paradoxa	**Tasmania**	**St. Clair National Park, NSW, AUS**		**WELTU20187 (*BGB9781*)**
121				AJ548969	Hoggard *et al.* 2003
211	*P. patagonica*	W USA, Argentina		AY101912	Rønsted *et al.* 2000
68	***P. raoulii***	**New Zealand**	**Manaia, Taranaki, NZ**		**WELTU20156 (*MLT054 et al.*)**
			George Creek, Wainuiomata, Wellington, NZ		**WELTU20152 (*PB & RH s. n.*)**
			Owhiro Bay, Wellington, NZ		**WELTU20151 (*MLT016 et al.*) / WELTU20157 (*MLT017 et al.*)**
			Maungatawharau, Waiouru, NZ		**Cult. in Victoria University (*Colin Ogle s. n.*)**
			Maungaharuru, Hawkes Bay, NZ		**WELT (*LP4448 & LS*)**
			Tararua Ranges, Wellington, NZ		**WELT (*LP4539 & LS*)**

Table 4.2. Details of *Plantago* samples that were included in the ITS phylogeny. Samples that were collected and sequenced (DNA) for this study are indicated in bold, with details of collection location.

Ref. in Rahn (1996)	Species	Indigenous distribution	Collection location	Accession number	Reference or voucher
68	***P. raoulii***	**New Zealand**	**Cape Palliser, Wairarapa, Wellington, NZ**		**WELT (*LP4996 & LS*)**
			Lake Sarah, Cass, NZ		**WELTU20153 (*PGJ2559*)**
22	*P. reniformis*	SE Europe		AY101858 / AJ548978	Rønsted *et al.* 2002 / Hoggard *et al.* 2003
94	*P. rhodosperma*	S USA, N Mexico		AJ548976 / AY692081	Hoggard *et al.* 2003 / Dhar *et al.* 2006
118	*P. rigida*	Andes		AY101876 / AF313037 / DQ006026	Rønsted *et al.* 2002 / Albach & Chase 2001 / Kress *et al.* 2005
35	*P. rugelii*	E N America		AY101863 / AY692080	Rønsted *et al.* 2002 / Dhar *et al.* 2006
160	*P. sarcophylla*	E Mediterranean		AY101893	Rønsted *et al.* 2002
149	*P. sempervirens*	SW Europe		AY101889	Rønsted *et al.* 2002
200	*P. sericea*	Andes		AY101910	Rønsted *et al.* 2003
136	*P. serraria*	Mediterranean		AY101880	Rønsted *et al.* 2002
	***P.* sp.**	**?**	**Pukerua Bay, Wellington, NZ**		**WELTU20178 (*PGJ2566 & MLT*)**
20	*P. sparsiflora*	SE USA		AJ548979	Hoggard *et al.* 2003
56	***P. spathulata subsp. picta***	**New Zealand**	**East Cape, Gisborne, NZ**		**CHR439486**
76	***P. spathulata* subsp. *spathulata***	**New Zealand**	**Cass, NZ**		**WELTU20118 (*PGJ2557*)**
			Cape Palliser, Wairarapa, Wellington, NZ		**WELTU20120 (*PGJ2567 & MLT*) / WELTU20119 (*PGJ2568 & MLT*)**
			Marfells Beach, Marlborough, NZ		**WELTU20117 (*PGJ2629 & MLT*)**
161	*P. squarrosa*	E Mediterranean		AY101894	Rønsted *et al.* 2002
78	*P. stauntoni*	Amsterdam & St. Paul Is.		AY101870	Rønsted *et al.* 2000
182	*P. stocksii*	W Asia		AY101904	Rønsted *et al.* 2002
142	*P. subspathulata*	Madeira		AY101884	Rønsted *et al.* 2002
133	*P. subulata*	Mediterranean		AY101878	Rønsted *et al.* 2000

Table 4.2. Details of *Plantago* samples that were included in the ITS phylogeny. Samples that were collected and sequenced (DNA) for this study are indicated in bold, with details of collection location.

Ref. in Rahn (1996)	Species	Indigenous distribution	Collection location	Accession number	Reference or voucher
–	***P.* "sylvester"**	**New Zealand**	**Lake Sylvester, Nelson, NZ**		**WELTU20150 (*MLT022*)**
					WELTU20149 (*MLT023*)
205	*P. tandilensis*	E Argentina		AY101908	Rønsted *et al.* 2002
	P. tasmanica	**Tasmania**	**St. Clair National Park, NSW, AUS**		**NSW743928 (*BGB9780*)**
73				AJ548970	Hoggard *et al.* 2003
43	*P. tenuiflora*	E Europe, C Asia		AY101866	Rønsted *et al.* 2002
84	*P. tomentosa*	S America		AY101872	Rønsted *et al.* 2002
124	***P. triandra* subsp. *triandra***	**New Zealand**	**Kettlehole Tarn, Cass, NZ**		**WELTU20158 (*PGJ2558*)**
			St. Arnaud, Nelson, NZ		**WELTU20162 (*MLT025 & PGJ*)**
			Lake Sylvester, Nelson, NZ		**WELTU20163 (*MLT021 et al.*)**
			Harihari, Westland, NZ		**WELTU20164 (*MLT029 et al.*)**
			Waipapa River, Northland, NZ (cult. in Otari-Wilton Bush, Wellington, NZ)		**WELTU20165 [Otari cult. (*1005.0115*)]**
				AJ548965	Hoggard *et al.* 2003
–	***P. triandra* subsp. *masoniae***	**New Zealand**	**Paturau Coast, Nelson, NZ**		**WELTU20160 (*PJL s. n.*)**
			Manaia, Taranaki, NZ		**WELTU20167 (*MLT055 et al.*)**
109	***P. triantha***	**Tasmania, Subantarctic Islands**	**Enderby Island, Auckland Islands**		**WELTU20177 (*VT55*)**
81	*P. trinitatis*	Ilha Trinidade		AY101871	Rønsted *et al.* 2002
119	***P. unibracteata***	**New Zealand**	**Cardrona Skifield, Wanaka, NZ**		**WELTU20171 (*PGJ2603 & MLT*)**
			Wilberg Range, Westland, NZ		**WELTU20160 (*PGJ2599 et al.*) / WELTU20172 (*MLT028 et al.*)**
			Lake Sylvester, Nelson, NZ		**WELTU20175 (*MLT024 & PGJ*)**
			Rainbow Skifield, Nelson, NZ		**WELTU20174 (*MLT026 & PGJ*)**
			The Remarkables, Queenstown, NZ		**WELTU20176 (*MLT038 & PGJ*)**
			Ruapehu Skifield, Mt. Ruapehu, NZ		**WELTU20173 (*MLT053 & PGJ*)**
114	*P. uniglumis*	S America		AY101875	Rønsted *et al.* 2002

Table 4.2. Details of *Plantago* samples that were included in the ITS phylogeny. Samples that were collected and sequenced (DNA) for this study are indicated in bold, with details of collection location.

Ref. in Rahn (1996)	Species	Indigenous distribution	Collection location	Accession number	Reference or voucher
65	***P. varia***	**Australia**	**Kosciuszko National Park, Australia**		**NSW 743869 (*BGB9766*) / NSW743869 (*BGB9767*)**
147	*P. webbii*	Macaronesia		AY101887	Rønsted *et al.* 2002
143	*Litorella uniflora*	Europe		AJ548962 / AJ548960 / AJ548963 / AJ548961 / AF515218	Hoggard *et al.* 2003 / Hoggard *et al.* 2003 / Hoggard *et al.* 2003 / Hoggard *et al.* 2003 / Albach *et al.* 2004 /
144	*Litorella americana*	North America		AJ548956 / AJ548958 / AJ548957 / AJ548955	Hoggard *et al.* 2003 / Hoggard *et al.* 2003 / Hoggard *et al.* 2003 / Hoggard *et al.* 2003
145	*Litorella australis*	South America		AJ548959	Hoggard *et al.* 2003
–	*Aragoa corrugatifolia*	Colombia		AJ548980	Hoggard *et al.* 2004
–	*Aragoa cupressina*	Northern Andes		AJ459402	Bello *et al.* 2002
–	***Veronica hookeriana***	**New Zealand**	**Whanahuia Range, Ruahine Mts, NZ**		**WELTU (*PGJ2458*)**
–	***Veronica salicornioides***	**New Zealand**	**Jacks Pass, Hanmer, Canterbury, NZ**		**CHR512475**

Collectors: *BGB* = Barbara G. Briggs, *HMM* = Heidi Meudt, *LP* = Leon Perrie, *LS* = Lara Shepherd, *LT* = Leah Tooman, *MLT* = Mei-Lin Tay, *PBH* = Peter Heenan, *PGJ* = Phil Garnock-Jones, *PB* = Peter Beveridge, *PJL* = Peter J. Lockhart, *RL* = Rodney Lewington and *VT* = Vanessa Thorn. Herbarium vouchers: CHR = Allan Herbarium, Landcare Research, Christchurch, New Zealand; NSW = National Herbarium of New South Wales, Australia; WELT = Herbarium of Museum of New Zealand, Te Papa Tongarewa, Wellington, New Zealand; and WELTU = H. D. Gordon Herbarium in Victoria University of Wellington, New Zealand. Samples in bold are new samples collected and sequenced for this study.

Note: Genbank accession numbers will be inserted into the table at before publication of papers.

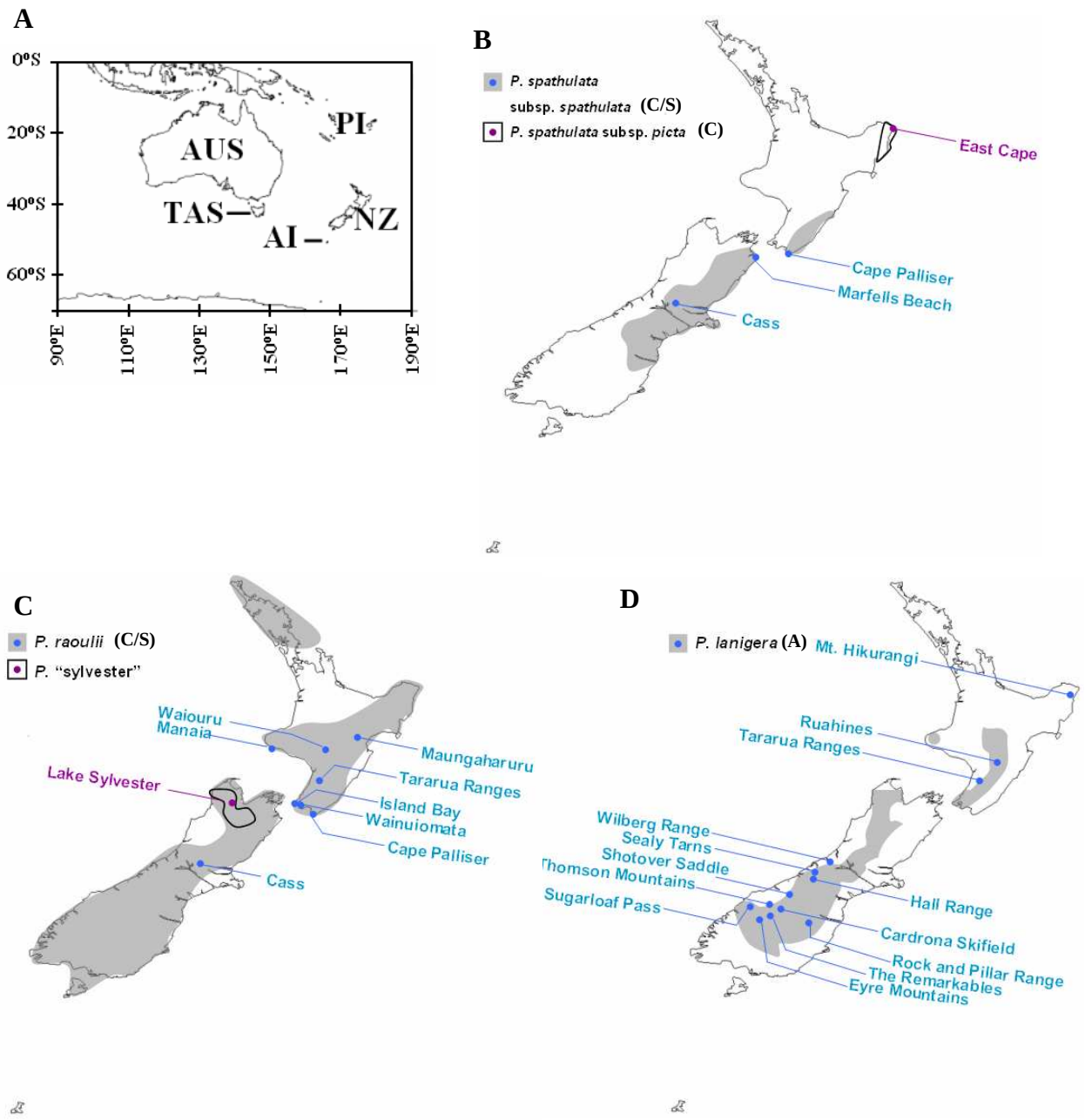


Figure 4.2: Distributions of New Zealand *Plantago* species based on collection of previous herbarium specimens (from Allan Herbarium, Landcare Research, Christchurch, New Zealand) and locations of samples collected for this study. A: The Australasian region (AI = Auckland Islands, AUS = Australia, NZ = New Zealand, PI = Pacific Islands, and TAS = Tasmania); B: *P. spathulata*; C: *P. raoulii* and *P.* "sylvester"; D: *P. lanigera* (A = alpine/subalpine habitats; C = coastal; C/S = coastal to altitudes of 1000-1500m).

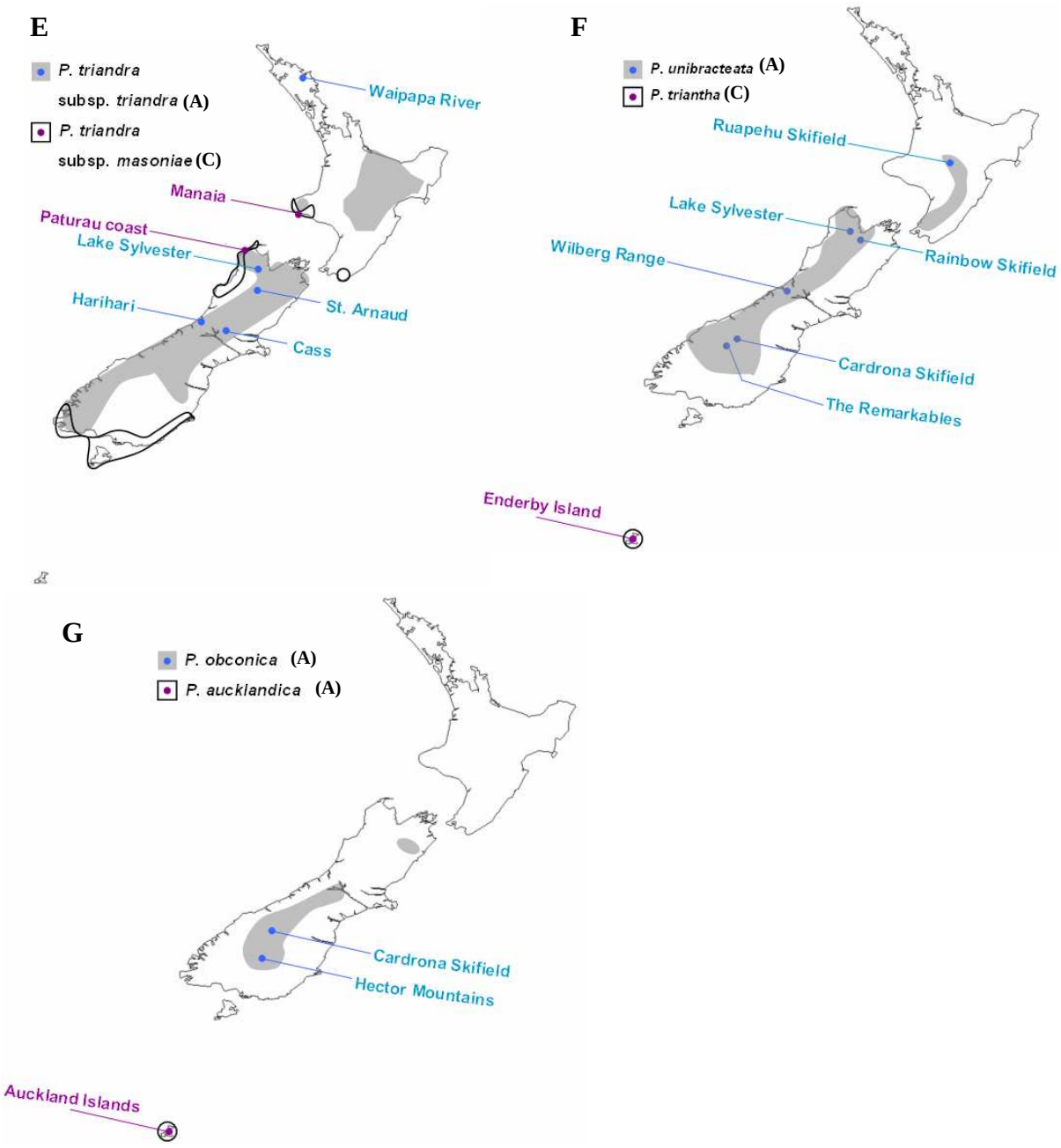


Figure 4.2 (continued): Distributions of New Zealand *Plantago* species based on collection of previous herbarium specimens (from Allan Herbarium, Landcare Research, Christchurch, New Zealand) and locations of samples collected for this study. E: *P. triandra*; F: *P. unibracteata* and *P. triantha*; and G: *P. obconica* and *P. aucklandica* (A = alpine/subalpine habitats; C = coastal; C/S = coastal to altitudes of 1000-1500m).

The inclusion of a comprehensive set of New Zealand *Plantago* population samples and representatives of Australasian species is essential for determining overall species relationships and biogeographic patterns of the genus in the Southern Hemisphere. In the previous chapter, I discussed the implications of new molecular data of Australasian *Plantago* for Southern Hemisphere biogeography. The aims of this chapter are to use ITS sequences: 1) to place the Australasian species within a larger phylogenetic framework, and 2) to examine delimitation and relationships, taxonomy and evolutionary patterns of the native New Zealand species using a molecular phylogeny (i.e. testing presence of different species or subspecies in *P. spathulata*, *P. raoulii*, *P. lanigera* and *P. triandra*; and elucidating species relationships of two undescribed or recently described species (*P.* "sylvester" and *P. obconica*).

4.2 Materials and Methods

4.2.1 Study group

Collection locations of samples, along with voucher information and Genbank accession numbers are presented in Table 4.2. Samples were collected from the wild across the North and South Islands of New Zealand (Fig. 4.2) from previously known locations, and from Australia.

Multiple individuals (ranging from one to ten) from several populations (ranging from one to fourteen) were obtained for each species. All eight native New Zealand *Plantago* species were included in the phylogeny, including the recently described *Plantago obconica* (Sykes 1988) and the undescribed New Zealand polyploid *P.* "sylvester" (Groves and Hair 1971). Twelve out of the 24 native Australian species

were obtained (including *P. triantha*, which is also found in New Zealand). In addition, introduced *Plantago* species, which have established in New Zealand (*P. coronopus, P. lanceolata, P. major, P. australis* and an unidentified specimen referred to as *P.* sp. in this study) were also inlcuded. *P. major* and *P. australis* are in the same subgenus as the Australasian species (subg. *Plantago*), whereas *P. coronopus* and *P. lanceolata* are in subg. *Coronopus* and subg. *Albicans*, respectively (following Rahn 1996). Additional sequences from previous studies were obtained from GenBank.

Plantago lanigera and *P. novae-zelandiae* are both referred to as *P. lanigera* here following Sykes (in Webb *et al.* 1988) but the ITS phylogeny is scrutinised to see if two groups are evident. Both forms were collected, including specimens with larger leaves and sparser hairs thought to be *P. novae-zelandiae* and specimens with smaller leaves that were densely covered with hair thought to represent *P. lanigera* as defined by Moore (in Allan 1961).

Several *Veronica* and *Aragoa* species were used as outgroups in the phylogeny. *Veronica* is used as a distant outgroup to *Plantago* because it is a close relative and has been used as an outgroup in the past (Rønsted *et al.* 2002; Hoggard *et al.* 2003; Cho *et al.* 2004), whereas *Aragoa* is the sister group to the *Plantago-Littorella* clade (Bello *et al.* 2002). For the purpose of this paper, the *Littorella* clade is treated as an outgroup as it has been shown to form the sister clade to a clade of all other *Plantago* species (Rønsted *et al.* 2002; Hoggard *et al.* 2003).

4.2.2 Molecular techniques

Tissue samples were mostly preserved in silica gel from field collections, but a few were obtained from existing herbarium specimens. DNA extractions were performed using the DNEasy Plant Mini kit (Qiagen, Hilden, Germany) or using a cetyltrimethylammonium bromide (CTAB) protocol modified from Doyle and Doyle (1990), after manual disruption of dried tissue using a pestle and mortar. The primers used to amplify the ITS DNA region are: *ITS28CC:* CGCCGTTACTAGGGGAATCCTTGTAAG (Wagstaff and Garnock-Jones 1998) and *ITS5:* GGAAGTAAAAGTCGTAACAAGG (White *et al.* 1990). Only ITS is used in this study because ITS sequences for *Plantago* are available on Genbank (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=nucleotide). Additionally, apart from a few conflicts, the ITS phylogeny is identical to the phylogeny reconstructed using chloroplast and mitochondrial DNA. The ITS region has also been used for everything from fungi to higher plants (White *et al.* 1990, Álvarez and Wendel 2003) to animals (Jorgensen *et al.* 2007; Kuriiwa *et al.* 2007), and often provides good resolution even at shallow phylogenetic levels (Álvarez and Wendel 2003). Thus, the ITS phylogeny will be used to explore taxonomic issues and evolutionary patterns of the native New Zealand *Plantago* species in this study.

PCR amplification was performed using Eppendorf Mastercycler ep gradient S (Hamburg, Germany) in a final volume of 25 µl of the following: 16.35 µl water, 10X ThermoPol reaction buffer (10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl (pH8.8), 2 mM $MgSO_4$, 0.1% Triton X-100)(New England BioLabs), 0.4 mg/mL BSA, 250 µmol dNTPs, 10 ρmol each primer, and 0.75 U of *Taq* DNA polymerase (New England BioLabs) and 0.4 µl DNA template. The amplification was carried out

with a thermocycling profile of: an initial 2 minutes at 94°C, followed by 30 cycles of 1 minute at 94°C, 1.5 minutes at 50°C, 1 minute at 72°C, and ending with a final extension time of 5 minutes at 72°C. This PCR protocol was used to successfully amplify regions for all the primers. PCR products were visualised on 1.5% agarose gels before being purified with the High Pure PCR Purification Kit (Roche Diagnostics, Manheim, Germany). The purified PCR products were sequenced in both directions using an ABI3730 Genetic Analyzer by the Allan Wilson Centre Genome Service (Massey University, Palmerston North, New Zealand).

4.2.3 Dataset alignment and phylogenetic analyses

The program MEGA v3.1 (Kumar *et al.* 2004) was used to assemble and align sequences for each sample. The final ITS dataset consisted of 252 sequences from 239 species of *Plantago*, 10 sequences from *Littorella*, one sequence from *Aragoa* and two sequences from *Veronica*. Of these, 152 sequences (25 *Plantago* species and 2 *Veronica* species) were obtained for this study, and 100 sequences were obtained from GenBank. Identical sequences were combined in order to improve efficiency of analyses. Thus, the final dataset had 177 sequences of length 656 bp including inferred gaps, with 321 (48.9%) variable sites and 269 (41%) parsimony-informative sites. Gaps were coded using modified complex indel coding (MCIC) method as implemented in SeqState v1.32 (Müller 2005) and were included in the dataset because the gaps may contain informative characters (shown in Chapter three). There were 733 characters in the dataset when indels were coded as characters concatenated onto the end of the sequences. The gapcoded dataset was then explored using MEGA v3.1 and DAMBE v4.5.9 (Xia and Xie 2001) and used for subsequent analyses.

The substitution model was selected using Modeltest v3.7 (Posada and Crandall 1998), which tests the fit among 56 different models of different complexity using a hierarchical likelihood ratio test. Models were chosen using Akaike Information Weights criterion (AIC) because it has advantages over the hierarchical likelihood ratio test (Posada and Buckley 2004). Thus, the model implemented for the ITS dataset was a GTR + I + G model, with base frequencies of variable sites: T = 26.3%, C = 29.4%, A = 20.1%, G = 24.3% ($p(\chi^2) = 0.5560$), scaled transition and transversion substitutions in the dataset: AC = 0.7695, AG = 1.6289, AT = 1.2942, CG = 0.2850, CT = 4.0783, GT = 1.000, a gamma shape parameter = 0.8876 and a proportion of invariable sites (pinvar) = 0.3241.

MrBayes v3.1.1 (Huelsenbeck and Ronquist 2001) was used to conduct heuristic searches under a Maximum Likelihood (ML) criterion. The analysis was conducted using 1,000,000 generations, resulting in 10,000 trees, and run until the standard deviation was less than 0.01 to ensure that the runs had converged on a stationary distribution (25% of the runs were discarded as burn-in). Branches with posterior probabilities less than 50% were collapsed in the phylogeny. Current taxonomical classifications were then plotted on the branches of the resolved phylogeny to check if the molecular phylogeny agreed with the phylogeny based on morphological characters (Rahn 1996). One of the *Plantago spathulata* subsp. *picta* sequences was removed from further analyses, because DNA was extracted from some leaves that were not in very good condition and was from a herbarium specimen more than 20 years old. The resulting sequence had multiple ambiguous sites and was placed in a different position from a more recently dried sample of *P. spathulata* subsp. *picta* with a clearer sequence (Appendix II (O)).

The Bayesian phylogeny resulted in several well-supported clades but these had poor internal resolution. In order to clarify these internal relationships, each group was analysed separately in three groups to resolve parts of the full phylogeny that did not resolve because these species had a relatively lower rate of evolution than species in other parts of the tree. These phylogenies were used to investigate relationships between the New Zealand species in more detail. A model for each dataset was estimated using Modeltest v3.7 and these are presented in Appendix III (A), along with other data statistics. MrBayes was used to conduct heuristic searches for each of the three New Zealand datasets to check initial topology. The analyses were run with 1,000,000 generations, yielding 10,000 trees, and each analysis had a final standard deviation that was less than 0.01 (25% of the trees were discarded as burn-in). Replicates of the same population that clustered in the same clade in the resulting MrBayes phylogeny were removed and ML analyses were run on the reduced datasets to improve efficiency. ML analyses were conducted using PAUP* version 4.0b10 (Swofford 2002) with 100 replicates of random sequence addition and TBR branch swapping. Nonparametric bootstrap support was assessed using 200 replicates, with random sequence addition and TBR branch swapping.

Complex evolutionary patterns, such as those arising from hybridisation, introgression and species radiations, cannot be properly displayed on traditional bifurcating trees (Lockhart *et al.* 2001; Huson and Bryant 2006). Therefore, Splitstree v4.8 (Huson and Bryant 2006) was used to conduct neighbour net analyses on each of the datasets in order to better visualise relationships between the New Zealand species. Splitstree uses genetic distances to compute split networks, which allows conflicting signals of species similarities to be visualised if they are present in a dataset.

4.3 Results

During alignment of the dataset, several pairs of identical sequences were found (presented here with Genbank accession numbers): *Plantago debilis* (NSW 742894) and *P. hispida* (AJ548967); *P.* sp. (WELTU (*PGJ2566 & MLT*)) and *P. asiatica* (AY101862); and *P. daltonii* (AJ548968) and *P. tasmanica* (AJ548970).

In general, there was very little intraspecific genetic variation within Australasian *Plantago* species sampled more than once; most of the species have less than 1% sequence divergence (Table 4.3). The highest intraspecific sequence divergence, found in *P. raoulii* and in *P. lanigera*, is still low (about 2.1-2.5%). Uncorrected p-distances between all pairwise comparisons of Australasian species range from 0% to only 4.3% (results not shown).

Table 4.3. Population statistics for the native New Zealand *Plantago* species (with more than one DNA sequence) included in the ITS phylogeny.

Species	No. of sequences	No. of localities	No. of haplotypes	No. of variable sites	Average % sequence difference
Plantago spathulata	15	4	4	5	0.76
P. spathulata subsp. *spathulata*	14	3	3	1	0.15
P. raoulii	18	9	14	14	2.13
P. "sylvester"	12	1	6	1	0.15
P. lanigera	28	14	20	19	2.90
P. obconica	6	2	2	0	0.00
P. triandra	19	7	7	5	0.76
P. triandra subsp. *triandra*	13	5	4	3	0.46
P. triandra subsp.*masoniae*	6	2	4	4	0.61
P. unibracteata	10	7	8	5	0.76

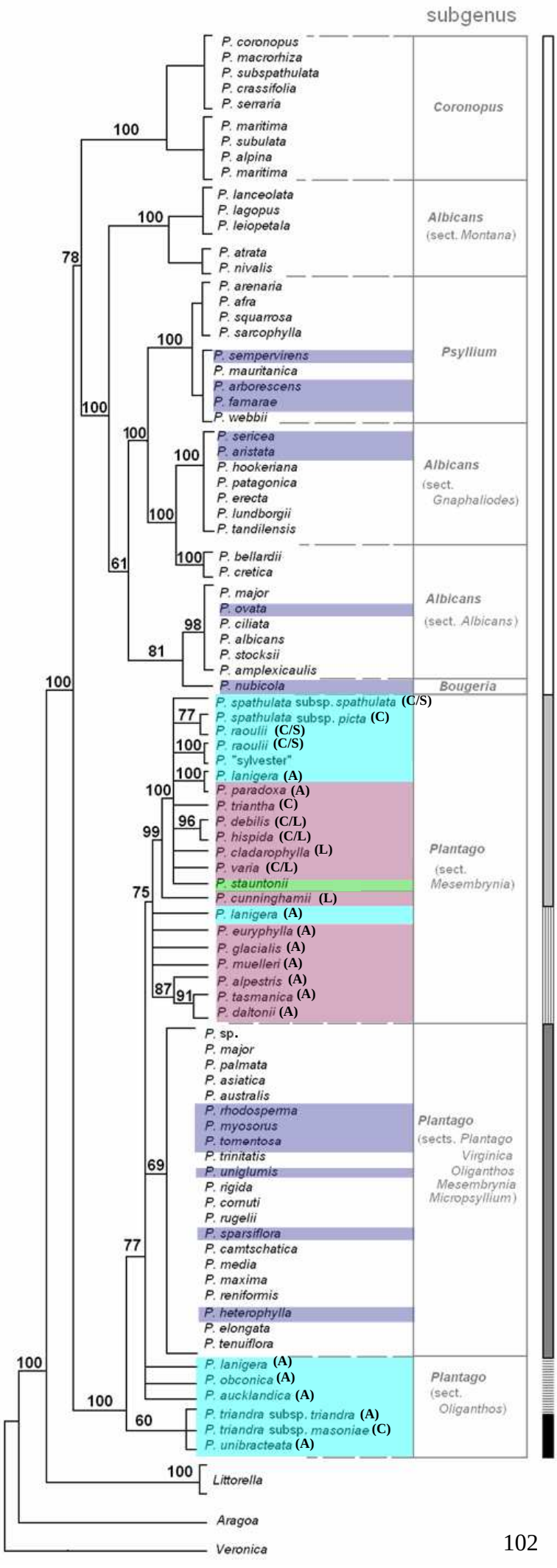


Figure 4.3: Simplified 50% consensus tree using Bayesian analysis (average tree likelihood score = -7531.24), with classifications according to Rahn (1996). Geographic locations of Southern Hemisphere species are indicated. Each Australasian species in the tree represents 1-25 individual sequences (A = alpine/subalpine habitats; C = coastal; C/S = coastal to altitudes of 1000-1500m; C/L = coastal/lowlands; L = lowland habitats).

Phylogeny using ITS

The 50% MrBayes consensus phylogeny reconstructed using the complete dataset was divided into four clades and two grades (Fig. 4.3). The focus of this paper is on the placement of Australasian species within a larger phylogeny, and species relationships within native New Zealand species. New Zealand species relationships are elucidated separately in the reduced ML analyses, thus, only the main structures of the clades are displayed in this figure for easier visual interpretation. Further resolved structure in of the clades can be found in the original tree, which is appended (Appendix II (O)).

Two main clades are resolved within *Plantago*. The first (Clade I) contains subg. *Coronopus*, subg. *Albicans*, subg. *Psyllium*, and subg. *Bougeria* (Fig. 4.3). In this clade, *Plantago nubicola* (subg. *Bougeria*) is nested within a paraphyletic *Albicans*. Most of the *P. major* replicates cluster together in *Plantago* Clade II, one sequence is placed within *Plantago* Clade I. This suggests a probable misidentification of the latter specimen, with Genbank number AY692079. Clade I contains no native New Zealand species and is not discussed further.

The second resolved clade represents subg. *Plantago* (sect. *Mesembrynia*, sect. *Oliganthos* and sect. *Virginica*) and includes all the Australasian species (1.0 PP). Subg. *Plantago* can be further divided to form three clades (Clades II, III and IV) and two grades (Grades A and B). Of the grades, Grade A contains only Australian species and is not further explored in this study. Previous analyses with a larger sequence dataset of three different genetic regions provided strong support that Grade B is a clade, even though in analyses of ITS alone the group appears to be a grade. Grade B is further explored in this study. The Australasian species do not form a clade

in the ITS phylogeny, with the widely distributed Clade III and *P. stauntoni* (from the Amsterdam & St. Paul Islands) nested within the Australasian species. However, there is low support for this finding (≤ 0.7BP for all resolved groups). Clade III contains species that are native to the Mediterranean, Europe, Africa, America and Asia, and is a mix of sect. *Oliganthos*, sect. *Mesembrynia* and sect. *Virginica*. South American species are present in Clade I and Clade III.

There is not much resolution at species level for the Australasian species (Clade II, IV and Grade A, B). Sect. *Mesembrynia* and sect. *Oliganthos* are polyphyletic in the phylogeny. Clade II and Grade A are classified as sect. *Mesembrynia*, and Clade IV and Grade B are sect. *Oliganthos*. Individuals identified as *P. lanigera* did not form a clade and were placed in Clade II, Grade A and Grade B. *P. paradoxa* forms a clade with samples from New Zealand that were identified to *P. lanigera* from Eyre Mountains. *Plantago paradoxa*, *P. glacialis*, *P. muelleri* and *P. triantha*, which were classified in sect. *Oliganthos* by Rahn (1996) are clustered among species from sect. *Mesembrynia*. *P. aucklandica* is classified in sect. *Plantago* but is not placed with other species from that section (in Clade III). *P. aucklandica*, placed by Rahn (1996) in sect. *Plantago*, is only loosely associated with Clade III, which contains all the other species of sect. *Plantago*. *Plantago triantha* is classified in sect. *Oliganthos* (Rahn 1996) but is placed with species from sect. *Mesembrynia* in the resolved ITS phylogeny (Fig. 4.3). It thus seems likely that reduction to solitary flowers, one of the defining characteristics of sect. *Oliganthos*, has occurred several times in the clade (e.g. *P. obconica*, *P. triantha* and *Littorella uniflora*).

A : Clade II

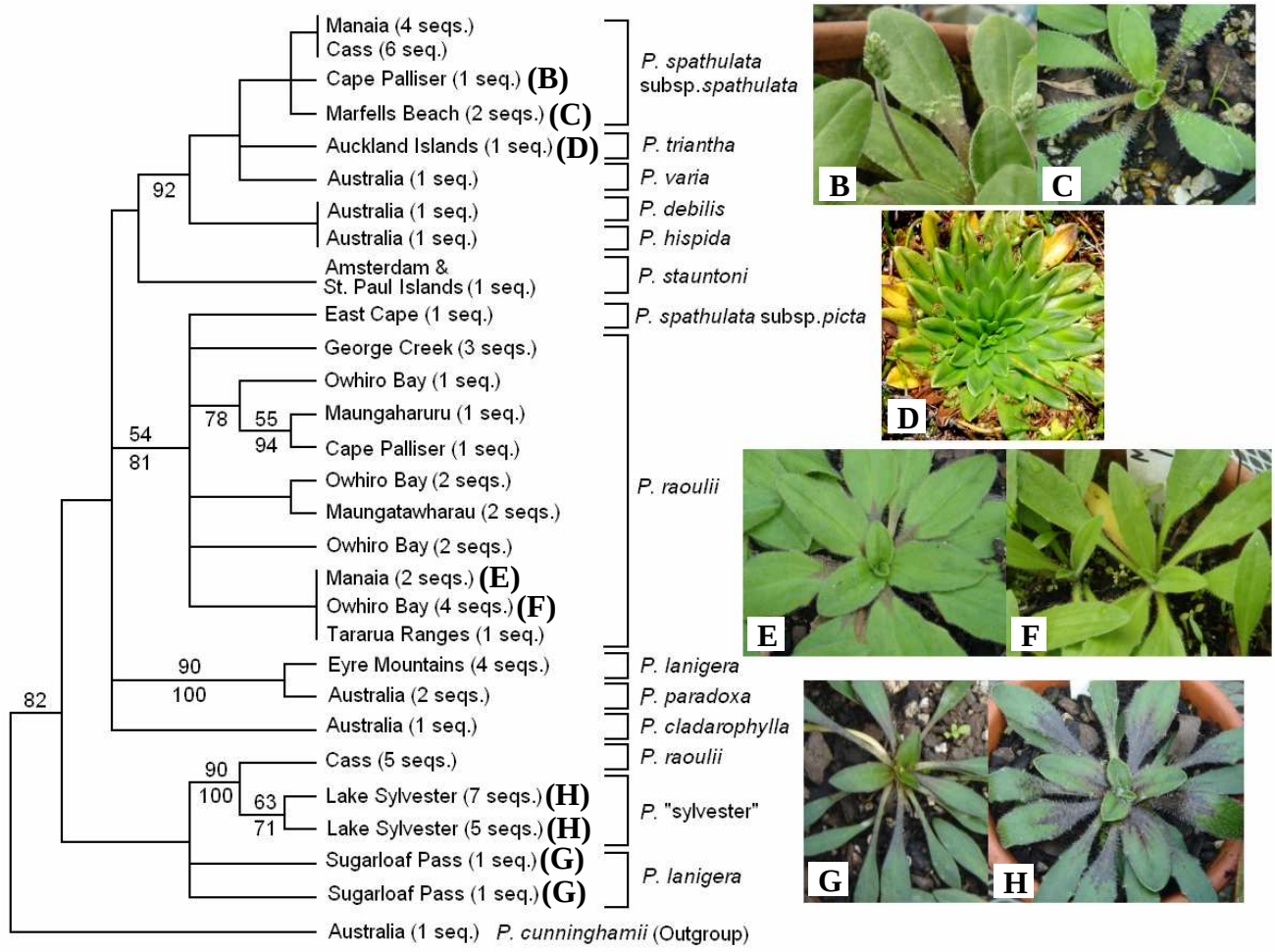


Figure 4.4. A: One of the reconstructed Maximum Likelihood (ML) trees of Clade II, which comprises lowland (including coastal) and alpine/subalpine species (tree score = 1087.60). Each sequence in A represents an individual plant. B and C: habit of *P. spathulata* subsp. *spathulata*; D: *P. triantha* from the Auckland Islands; E and F: two different forms of *P. raoulii*; G: specimen identified as *P. lanigera* from Sugarloaf Pass; and F: specimen from Lake Sylvester. ML bootstrap values are displayed above branches, whereas MrBayes posterior probabilities are displayed below branches (only support values >50% are shown). Plant pictures presented here are from plants that have been cultivated in the greenhouse for about a year.

A: Grade B

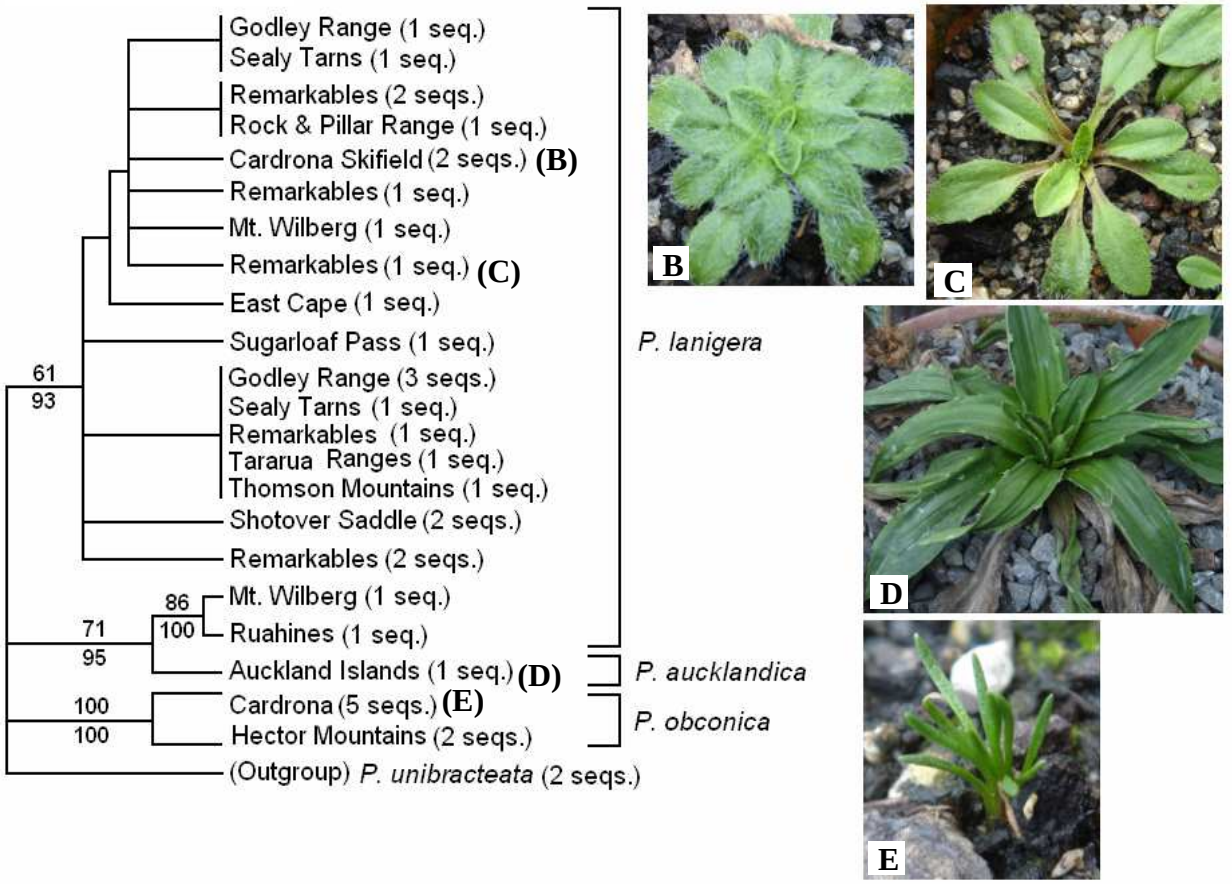


Figure 4.5. A: One of the reconstructed Maximum Likelihood (ML) trees of Grade B, which is a clade of alpine/subalpine species (tree score = 1065.00). Each sequence in A represents an individual plant. B and C: two different forms currently classified under *P. lanigera* (previously *P. lanigera* and *P. novae-zelandiae* respectively); D: *P. aucklandica*, endemic to the Auckland Islands; E: habit of *P. obconica*. ML bootstrap values are displayed above branches, whereas MrBayes posterior probabilites are displayed below branches (only support values >50% are shown). Plant pictures presented here are from plants that have been cultivated in the greenhouse for about a year.

A: Clade IV

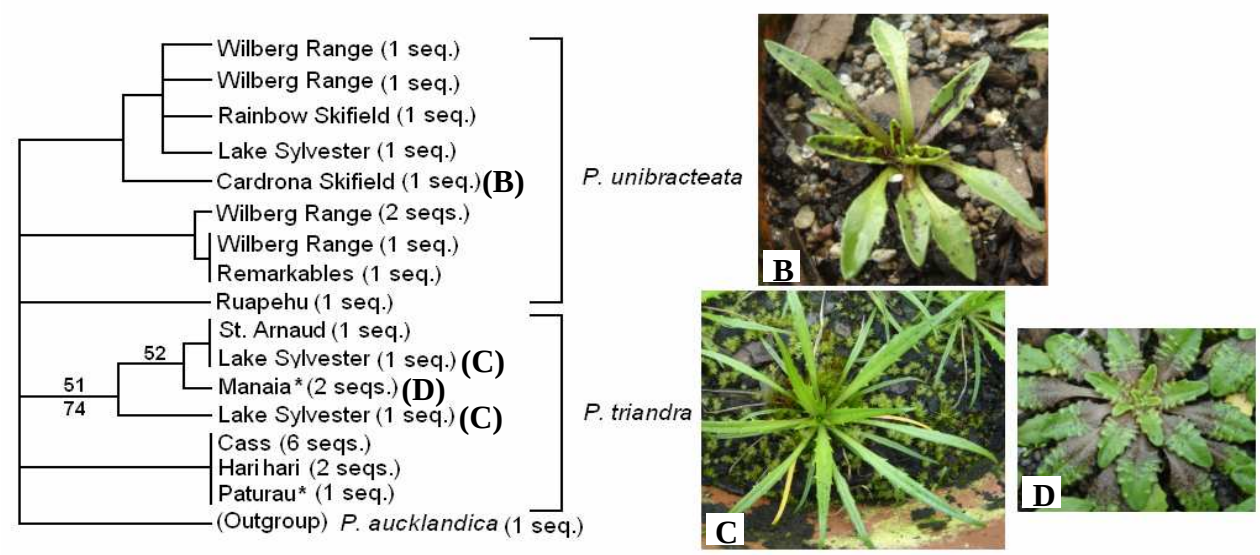


Figure 4.6. A: One of the reconstructed Maximum Likelihood (ML) trees of Clade IV, a clade of alpine/subalpine and coastal (*P. triandra* subsp. *masoniae*) species (tree score = 976.76). Each sequence in A represents an individual plant. B: *P. unibracteata*; C: *P. triandra* subsp. *triandra*; D: *P. triandra* subsp. *masoniae*. Samples of *P. triandra* marked with a * indicate populations of *P. triandra* subsp. *masoniae*, whereas the others are subsp. *triandra*. ML bootstrap values are displayed above branches, whereas MrBayes posterior probabilites are displayed below branches (only support values >50% are shown). Plant pictures presented here are from plants that have been cultivated in the greenhouse for about a year.

When the dataset was partitioned into the three Australasian groups of interest (Clade II, Grade B and Clade IV) and analysed separately, the reconstructed phylogenies were able to provide slightly better resolution of species. Proportions of nucleotide bases were equivalent among all sites in the three datasets (Appendix III (A)), but bases were heterogenous among variable sites for Clade IV and Grade B. Bayesian analyses have greater time efficiency over ML for large datasets like the large ITS dataset, but ML trees are preferred for smaller datasets because MrBayes generates numerous trees (10,000 in these analyses). The Bayesian and ML analyses of the datasets had congruent topologies and equivalent support values. Therefore, results of the reduced analyses are discussed using the phylogenies recovered using ML (Figs. 4.3A, 4.4A, 4.5A).

Clade II (sect. *Oliganthos*)

Clade II includes the two subspecies of *Plantago spathulata* (four localities, 15 individuals), *P. raoulii* (four localities, 18 individuals), and *P.* "sylvester" (one locality, 12 individuals) from New Zealand; five Australian species (*P. varia*, *P. debilis*, *P. hispida*, *P. paradoxa* and *P. cladarophylla*); one species (*P. stauntoni*) from the Amsterdam & St. Paul Islands in the South Indian Ocean; and one species, *P. triantha* (one locality, one individual), native to both New Zealand and Australia. The clade also contains several individuals that were identified as *P. lanigera* but do not form a clade with the other samples of *P. lanigera*, which are placed in Grade B.

ML analysis suggests that the two subspecies of *P. spathulata* as currently circumscribed are not sister taxa (Fig. 4.4A). *P. spathulata* subsp. *spathulata* forms a clade with *P. triantha*, *P. varia*, *P. debilis* and *P. hispida* (<50BP, 0.92PP).

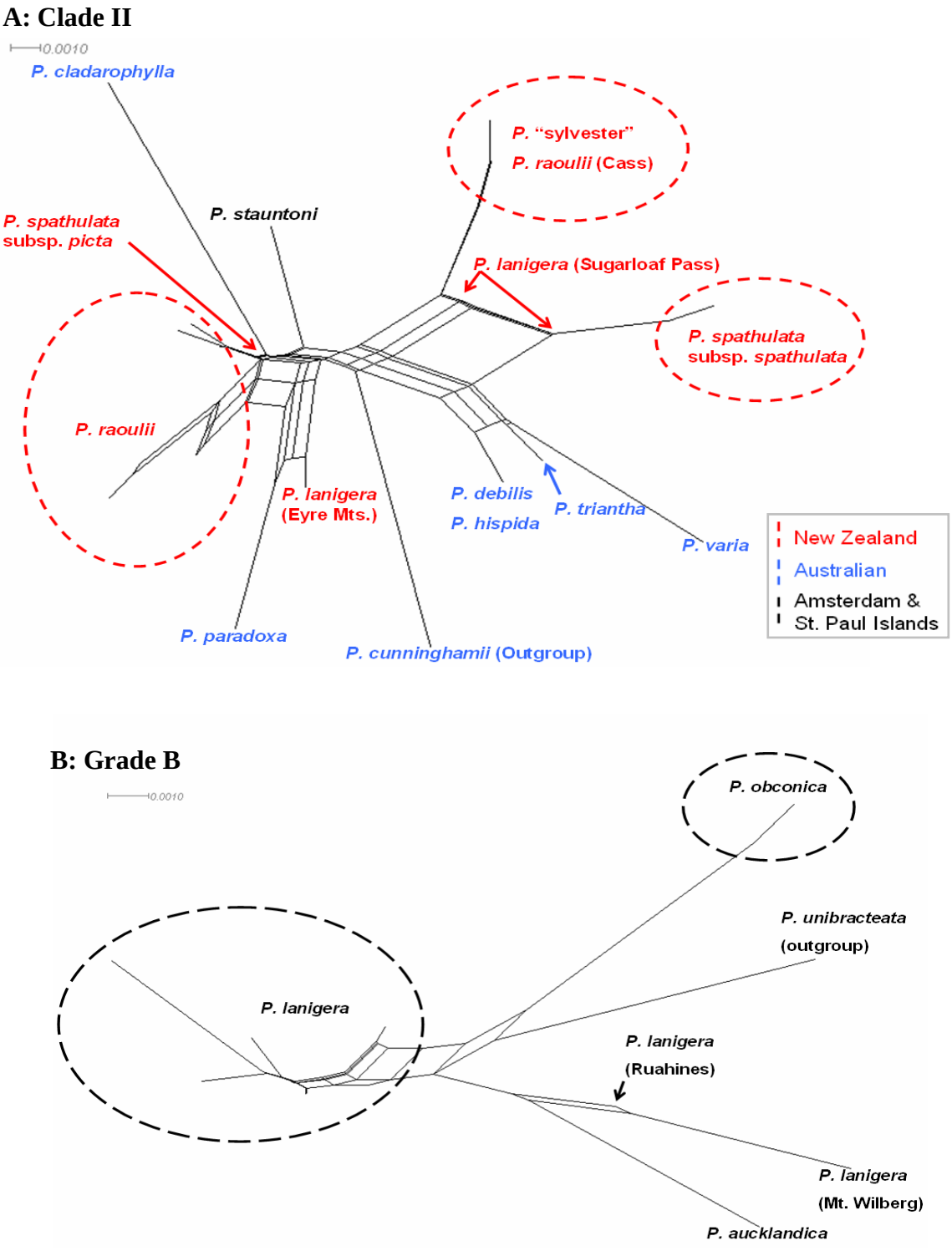


Figure 4.7. Neighbour-net analyses using Splitstree. A: Clade II; B: Grade B; and C: Clade IV. Dotted circles indicate the different species. In A colours indicate geographic location (red = New Zealand; blue = Australian; black = Amsterdam & St. Paul Islands).

C: Clade IV

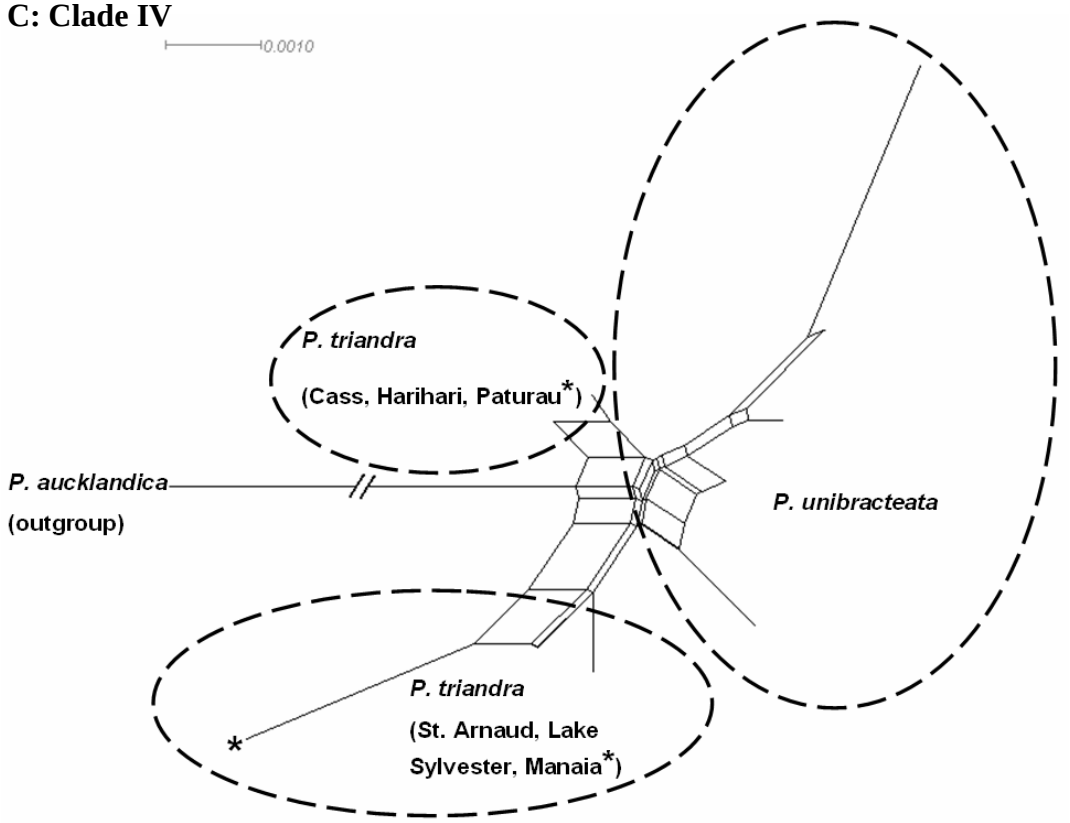


Figure 4.7 (continued). Neighbour-net analyses using Splitstree. A: Clade II; B: Grade B; and C: Clade IV. Dotted circles indicate the different species. * in C indicate populations of *P. triandra* subsp. *masoniae*.

The sister to this clade is *P. stauntoni* but there is low support for this relationship. *P. raoulii* is separated into two groups based on geographic locations. The N. I. *P. raoulii* are grouped together in a clade with *P. spathulata* subsp. *picta* (54BP, 0.81PP), whereas the sole sample of the S. I. *P. raoulii* forms a clade with *P.* "sylvester" (90BP, 1.0PP). One of the populations identified as *P. lanigera* from Sugarloaf Pass also clusters with the sample of S. I. *P. raoulii* and *P.* "sylvester" but there is low support for this group (<50BP and PP). The other *P. lanigera* population in this group is from the Eyre Mountains and forms a clade with the Australian *P. paradoxa* (90BP, 1.0PP). ML analyses yielded six trees, which differed only by rearrangements in the *P. raoulii-P. spathulata* subsp. *picta* clade. In one of the trees, the node of the branch leading to *P. spathulata* subsp. *spathulata* and the node of the *P. raoulii* clade were placed in a polytomy. *P. raoulii* had the most sequence divergence within Clade I (2.13%), *P.* "sylvester" only had 0.15% sequence divergence, and there was 0.76% sequence divergence between the two subspecies of *P. spathulata*.

Splitstree analysis of Clade II (Fig. 4.7A) revealed two groups. Each group contained both New Zealand and Australian species. *Plantago spathulata* subsp. *picta* is connected by a box to both *P. cladarophylla* and the N.I. *P. raoulii*, whereas the sample of *P. triantha* (from Auckland Islands) is attached to boxes connecting Australian species *P. debilis*, *P. hispida* and *P. varia*; but species relationships among these are not well-defined. Further, branches of one of the specimens of *P. lanigera* from Sugarloaf Pass is connected to *P. spathulata* subsp. *spathulata*, while the other specimen is connected to the branch leading to *P. raoulii* from Cass and *P.* "sylvester". Placement of *P. stauntoni* is uncertain; it is placed between the two Australasian groups, but closest to *P. cladarophylla*.

Grade B

Grade B contains *Plantago lanigera* (14 localities, 28 individuals), *P. obconica* (two localities, six individuals) and *P. aucklandica* (one locality, one individual). This grade was not resolved in the ITS phylogeny (Fig. 4.3) but there was high support for a clade comprising these species in previous combined analyses (see Fig. 3.2). Three main clades were resolved within the group (Fig. 4.5A), although relationships among them were not well-resolved. The first contains most of the *P. lanigera* samples (61BP, 93PP) with poorly supported (<50BP and PP) resolution of relationships within the clade. Secondly, there is a clade of two populations of *P. lanigera* (86BP, 100PP), which together are sister to *P. aucklandica* (71BP, 95PP). The third clade contains the two populations of *P. obconica* (100BP, 1.0PP). ML analyses resulted in 24 trees, where the only differences were rearrangements within the larger *P. lanigera* clade. Samples thought to represent *P. novae-zelandiae* are inseparable from the other *P. lanigera* samples based on ITS sequences, i.e. clustering within the large *P. lanigera* clade does not match either morphological differences or geographic location. There also does not appear to be geographical difference in the samples of *P. lanigera*. Among all the Australasian species, *P. lanigera* had the highest intraspecies sequence divergence: 2.90 % (Table 4.3) but this value is still low.

The distinction between species in Grade B (Fig. 4.7B) is clear aside from the two populations of *P. lanigera* that do not cluster with the other *P. lanigera*. The neighbour-net of *P. lanigera* indicates possible reticulation in the past, which may be a result of complex history perhaps involving hybridisation or incomplete lineage sorting.

Clade IV (sect. *Oliganthos*)

Clade IV consists of *P. unibracteata* (seven localities, ten individuals) and the two subspecies of *P. triandra* (seven populations, 19 individuals). ML analyses provided very little resolution for this clade (Fig. 4.6A). There were two resolved but poorly supported clades of *P. unibracteata* (both with <50BP), and one resolved clade for *P. triandra* (51BP). These, along with other samples from both species, form a polytomy at the base of Clade IV. The only difference among the three resulting ML trees was the placement of an individual *P. unibracteata* sequence from Mt. Wilberg. This sequence was placed alternatively as (1) sister to the other populations of the *P. unibracteata* clade, (2) at the most derived position in the *P. unibracteata* clade, and (3) in a polytomy with the *P. unibracteata* clade.

The two subspecies of *P. triandra* do not group according to the classification of Sykes (in Webb *et al.* 1988), which was based on morphology. The two clades within *P. triandra* do not appear to group according to geographic location. In fact, several populations from *P. triandra* subsp. *triandra* share an identical ITS sequence with a population of *P. triandra* subsp. *masoniae*. *P. unibracteata* populations also did not group according to geographic locations (e.g. replicates from a Mt. Wilberg population were placed in separate clades). Intraspecific and interspecific sequence divergence was low for both species. Sequences within each species differed by only 5 nucleotide substitutions out of 656 bp of ITS sequence (uncorrected p-distances range from 0.002-0.008). Sequence divergence between the two species was represented by 9 nucleotide substitutions and an extremely low range of uncorrected p-distances of 0.002-0.008.

In the Splitstree analysis of Clade IV (Fig. 4.7C), *P. unibracteata* and the two subspecies of *P. triandra* are inseparable genetically. The length of the branches connecting *P. triandra* and *P. unibracteata* show that *P. triandra* forms two groups that are more divergent genetically than either is from *P. unibracteata*. Further, representatives from both subspecies are present in each of the two groups of *P. triandra*.

4.4 Discussion

The aim of this study was firstly to integrate Australasian species of *Plantago* into a framework of an ITS phylogeny that included species representatives from around the world. This phylogeny was then used to evaluate the current taxonomy of the genus and overall relationships of the Australasian species. Secondly, phylogenetic analyses of multiple population samples per species were used to investigate evolutionary patterns and to address taxonomic issues in the New Zealand species of *Plantago*.

4.4.1 Evolution of New Zealand *Plantago*

Separation of groups according to elevation is evident in the ITS phylogeny presented here (Fig. 4.3), and also in the concatenated (ITS, *ndhF-rpl32* and *cox1*) phylogeny in Chapter three (Fig. 3.2), where the Australasian lowland species are separated from the alpine/subalpine species. The lowland species appear to have one more recent migration event into lowland habitats from alpine/subalpine groups. A similar pattern is found in alpine *Ranunculus* species (Lockart *et al.* 2001) and New Zealand *Veronica* sect. *Hebe* (Wagstaff and Garnock-Jones 1998; Wagstaff *et al.* 2002). The distribution of the New Zealand *Plantago* species also suggests that physical

separation because of adaptations to different habitats may be an important speciation mechanism for some taxa in the group. For example., within *Plantago triandra*, subsp. *triandra* is distributed on inland mountains and sometimes low altitudes in the S. I., whereas subsp. *masoniae* grows in coastal areas that are separate from populations of *P. triandra* subsp. *triandra* (Sykes in Webb *et al.* 1988).

Although the processes and mechanisms involved in the evolution of alpine species in New Zealand are not well known (Winkworth *et al.* 2005), the origin of the New Zealand species of *Plantago* was dated to be in the range of 2.3 to 1.5 mya in the previous chapter, which is within the period of uplift of the Southern Alps (starting about 5 mya). Speciation of the alpine and subalpine *Plantago* species following the uplift of these mountain ranges agrees with previous observations in other plant groups (Lockhart *et al.* 2001; Winkworth *et al.* 2002a). The Southern Alps may also present a physical barrier that limits the distribution of some species, such as *P. spathulata* subsp. *spathulata*, which is reportedly found only east of the main divide (Sykes in Webb *et al.* 1988).

The ITS phylogeny of Clade II suggests that Cook Strait may be a physical barrier for *P. raoulii* populations because N.I. populations of *P. raoulii* are separated from the S.I. population in the ML phylogeny (Fig. 4.4A), which is also evident in the Splitstree analysis (Fig. 4.7A). However, only one population of *P. raoulii* from S. I. was included and testing of this hypothesis would benefit from wider sampling. Cook Strait does not appear to present a barrier for the other species because closely related populations occur on both sides of the strait (Figs. 4.4A, 4.5A, 4.6A).

Table 4.4. Comparison of allopatry vs. sympatry and same vs. different chromosome numbers in New Zealand *Plantago* species.

	Plantago aucklandica	*P. lanigera*	*P. obconica*	*P. raoulii*	*P. spathulata subsp. spathulata*	*P. spathulata subsp. picta*	*P. "sylvester"*	*P. triandra subsp. triandra*	*P. triandra subsp. masoniae*	*P. triantha*	*P. unibracteata*
Plantago aucklandica	—	A	A	A	A	A	A	A	A	S	A
P. lanigera	N/A	—	S	A	S	A	S	S	A	A	S
P. obconica	N/A	X	—	A	S	A	A	S	A	A	S
P. raoulii	N/A	X	X	—	S	A	S	S	S	A	S
P. spathulata subsp. spathulata	N/A	X	X	I	—	A	A	S	A	A	S
P. spathulata subsp. picta	N/A	X	X	I	I	—	A	A	A	A	A
P. "sylvester"	N/A	X	X	X	X	X	—	S	S	A	S
P. triandra subsp. triandra	N/A	X	X	I	I	I	X	—	A	A	S
P. triandra subsp. masoniae	N/A	X	X	I	I	I	X	I	—	A	A
P. triantha	N/A	I/X	I	X	X	X	X	X	X	—	A
P. unibracteata	N/A	X	X	X	X	X	X	X	X	X	—

Upper right side is a comparision of whether the species are allopatric or sympatric (A = populations are allopatric, S = populations are sympatric). Lower left side is a comparison of chromosome numbers (X = different chromosome numbers, I = identical chromosome numbers, I/X = one of the chromosome numbers is the same (*P. lanigera* has two chromosome numbers 2*n* = 12, 24)). Colours represent different combinations; yellow boxes are sympatric species with different chromosome numbers; orange boxes are sympatric species with identical chromosome numbers; red boxes are allopatric species with same chromosome numbers; blue boxes are allopathic species with identical chromosome numbers. N/A = data not available (chromosome numbers for *P. aucklandica* unknown). Chromosome numbers from Groves and Hair (1971) and Rahn (1996).

Some of the New Zealand *Plantago* species occur in sympatric populations. The species have variable chromosome numbers (2*n* = 12, 24, 48, 60 and 96) which may act as a reproductive barrier where distributions overlap. Most species that have sympatric populations have different chromosome numbers (Table 4.4). For example, *P. unibracteata* and *P. lanigera* can be found growing at the same location but they have different chromosome numbers (2*n* = 60 and 2*n* = 12/24, respectively). Polyploidy may have allowed rapid speciation within the group, e.g. *P.* "sylvester" (2*n* = 96), which is suggested to be a hybrid of *P. spathulata* subsp. *spathulata* and *P. raoulii* (both 2*n* = 48) (see below). Different flowering times, along with the limited dispersal of pollen and seeds may also present a biological barrier to genetic exchange, and therefore may promote speciation.

Possible hybrids

Within the ITS dataset, several populations with unexpected relationships were found in a basal polytomy in Clade II. The samples (a few per population) were identified as *Plantago lanigera* but they do not group with the other *P. lanigera* populations, which are placed in Grade B. The morphology of these individuals is certainly peculiar; the habits of both *P. lanigera* from Eyre Mountains and *P. lanigera* from Sugarloaf Pass resemble the specimens collected from Lake Sylvester (*P.* "sylvester") with respect to coloration, leaf shape and plant size (see Fig. 4.4E & F). Their positions are also odd in the Splitstree analyses (Fig. 4.7). *P. lanigera* from the Eyre Mountains is closest to *P. paradoxa*, whereas *P. lanigera* from Sugarloaf Pass are attached between *P. spathulata* subsp. *spathulata* and a clade of *P.* "sylvester" and *P. raoulii* (from Cass). These populations could represent misidentifications but the unexpected placements

in the ML and Splitstree analyses could also suggest hybridisation. Further sampling and a study of morphology characters may help clarify relationships of these species.

4.4.2 Taxonomy of New Zealand *Plantago*

Circumscription of subgenera and sections

In the ITS phylogeny, the Australasian species fall into a clade of subg. *Plantago*, which is sister to a clade containing the other four subgenera (subg. *Albicans*, subg. *Bougeria*, subg. *Psyllium* and subg. *Coronopus*). The ITS phylogeny also shows that most of the subgenera identified by Rahn (1996) form monophyletic groups (Fig. 4.3) and the topology of the tree does not contradict the topology recovered from previous phylogenetic studies (Rønsted *et al.* 2002; Hoggard *et al.* 2003).

While there is some resolution in Clade I, there is not much resolution among Australasian species (Clades II and IV, and Grades A and B). This is probably because of low sequence divergence within and between the Australasian species (Table 4.3, Appendix III (B)). This finding is consistent with the long branch lengths found in *Plantago coronopus* and *P. lanceolata*, but short branch lengths in Australasian species were formed when phylogeny was reconstructed in the previous chapter (Fig. 3.3). Slightly better resolution of species relationships within Australasian groups was obtained when species with relatively high substitution rates were removed to form three reduced Australasian datasets. The ML phylogenies of these smaller datasets are used to address taxonomic issues and to look at the evolution of *Plantago* species in New Zealand below.

Circumscription of New Zealand species and subspecies

Plantago spathulata

In the ML tree (Fig. 4.4A), *Plantago spathulata* subsp. *picta* appears to be more divergent from *P. spathulata* subsp. *spathulata* and closer to *P. raoulii* than expected based on morphology and previous classifications. Sykes (1988) noted that the main difference is a persistent taproot in subsp. *picta*, although they may also be differentiated by hairs either on the keels of bracts and sepals (subsp. *spathulata*) or only on the margins of bracts and sepals (subsp. *picta*) (Moore in Allan 1961). This is mirrored in Rahn's (1996) morphological phylogeny, where subsp. *picta* is separated from subp. *spathulata* by only 5 steps (3 characters) which are: absence/presence of adventitious roots, whether or not the sepals are glabrous abaxially, and 1-4 ovules per ovary with a rudiment of an upper compartment on the adaxial side of the placenta. This suggests that although the two taxa are genetically divergent, they may share many plesiomorphic character states. The divergence of the two subspecies and the placement of subsp. *picta* as closely related to the population of *P. raoulii* from Cass is also evident in the Splitstree analysis of Clade II (Fig. 4.7A). They are allopatric: subsp. *picta* is endemic to East Cape and the Poverty Bay Coast in the N. I. (Sykes in Webb *et al.* 1988), whereas subsp. *spathulata* is found in the Wairarapa coast and throughout the S. I. Phylogenetic analyses suggest that the two taxa are not each other's closest relatives. However, additional samples of subsp. *picta* should be sequenced and although these results suggest that they are distinct species, it would be impractical to recognise them taxonomically without reliable and consistent morphological differences.

Plantago raoulii

There are reportedly two sympatric forms of *P. raoulii* (Sykes in Webb *et al.* 1988, C. Ogle pers. comm.) distinguished by broad vs. narrow leaves. These two distinct forms were collected for this study (Figs. 4.1B, 4.4E & F). However, both ML and Splitstree analyses indicated that there was no genetic differentiation between populations of N. I. *P. raoulii* found here that are consistent with morphology, i.e., there was no difference in the morphology of the plants that formed the separate poorly supported clades, and there were individuals from the same population that occurred in separate clades (but again no difference in morphology). Thus, it appears that the two forms of *P. raoulii* may simply be ecotypes or represent simple polymorphisms.

The phylogenetic analyses also show that *P. raoulii* is separated by geographic location (i.e. N. I. and Cass populations were in different groups). It is possible that the *P. raoulii* population from Cass may be related to *P.* "sylvester" because they group together with high support. The habitats of the Cass and Lake Sylvester samples were similar (flush among *Schoenus*) and their morphology and flowering times in cultivation are very similar (Garnock-Jones pers. comm.), but *P. raoulii* from a location near to Cass has been found to have $2n = 48$ (Groves and Hair 1971). Further studies and additional samples of *P. raoulii* from S. I. are needed to clarify this observation.

Plantago lanigera

In agreement with Sykes (in Webb *et al.* 1988), there was not enough molecular evidence to distinguish between the two forms previously known as *P. lanigera* and *P. novae-zelandiae* (e.g. Moore in Allan 1961) in the dataset. There was weak support

for a *P. lanigera* clade that included the two different forms (excluding two samples of *P. lanigera*) but no support for any difference between populations within the group (Fig. 4.5A). Thus, there is no evidence to refute the conspecificity of *P. lanigera* and *P. novae-zelandiae* (Sykes in Webb *et al.* 1988). There appears to be no barrier to genetic interchange among populations, although some reticulation in the history of *P. lanigera* is suggested from Splitstree analysis (Fig. 4.7B). An odd finding in the ML phylogeny is that some but not all individuals from two populations (Mt. Wilberg and Ruahines) were placed in a clade with *P. aucklandica*, which is sister to the larger *P. lanigera* clade. Samples from Mt. Wilberg were present in both clades, which indicate that they are not grouped according to geographic location. Two scenarios may explain this finding: that there were two different species collected in both locations, or that more variation exists within populations than between populations. The latter appears to be more plausible as samples collected from either population were morphologically indistinguishable.

Plantago triandra

Sampling in this study included populations referable to both *P. triandra* subsp. *triandra* and subsp. *masoniae* based on morphology. Although two clades of *P. triandra* were resolved, each contained both samples from both subsp. *triandra* and subsp. *masoniae* (Fig. 4.6A). Populations of the two were also not separated in the Splitstree analysis (Fig. 4.7C). Sykes (1988) distinguished subsp. *masoniae* from *P. triandra* subsp. *triandra* by fleshier leaves that differ in size. Morphological differences between coastal and inland plants appear to be cosistent (see Fig. 4.1E & F), but unfortunately, some components relating to the thickness of leaves, size, shape and even whether the leaves are toothed or entire appear to be plastic depending on

environment and seasonal changes (personal observation from plants in the greenhouse). Thus, although two morphological groups can be distinguished, there is no genetic evidence for separation of these two into separate subspecies. Several populations of subsp. *triandra* even share identical ITS sequences with individuals from a population of subsp. *masoniae*. These two morphological forms appear to be ecotypes and recognition at a lower rank (i.e. variety) might be more appropriate.

Species relationships

Plantago "sylvester"

The 16-ploid *P.* "sylvester" may be an allopolyploid between *P. raoulii* and *P. spathulata* subsp. *spathulata*. There are morphological characteristics that link it to these two species: *P.* "sylvester" plants resemble *P. spathulata* by having two long seeds in each locule, and broad glabrous corolla lobes (Moore in Allan 1961); whereas they resemble *P. raoulii* in only having four vertical seeds, and the habit of the plants appear to resemble *P. raoulii* morphologically (Sykes in Webb *et al.* 1988). *P.* "sylvester" has a chromosome number of $2n = 96$, while both *P. spathulata* and *P. raoulii* have chromosome numbers of $2n = 48$. *P.* "sylvester" is sister to a sample of *P. raoulii* from Cass (S. I.) in the ML analysis (Figs. 3A) and these are at almost equal distances from the N. I. *P. raoulii* and *P. spathulata* subsp. *spathulata* in the Splitstree analysis (Fig. 4.7A). Until the morphological recognition and geographic range of *P.* "sylvester" are better known, it would be imprudent to recognise it as a new species.

Plantago obconica

The ITS phylogeny reconstructed in this study (Fig. 4.3) and a three genome phylogeny discussed previously (Fig. 3.2) both show that *P. obconica* is closely

related to *P. lanigera* and not *P. triandra* as previously suggested based on morphology (Sykes in Webb *et al.* 1988). Thus, morphological similarities between *P. obconica* and *P. triandra* appear to have arisen convergently. *P. obconica* should be classified within sect. *Oliganthos*.

Plantago triantha and Plantago aucklandica

Plantago triantha is currently placed within sect. *Oliganthos* (Rahn 1996). However, it is placed with species from sect. *Mesembrynia* in the ITS phylogeny (Fig. 4.3) because it is very similar to *P. spathulata* subsp. *spathulata* genetically. *P. aucklandica* is placed in sect. *Plantago* (Rahn 1996) but groups with other species from sect. *Oliganthos* (Fig. 4.3 and Fig. 3.2). *P. triantha* and *P. aucklandica* should be reclassified in those groups.

4.4.3 Implications of this study

The reconstructed ITS phylogeny presented here using ITS sequences did not match the concatenated ML tree presented in the previous chapter (the latter resembles and is strongly influenced by the tree of the cpDNA spacer *ndhF-rpl32*). This outcome was also encountered in the previous chapter comparing trees of the ITS and organellar regions. *Plantago triandra*, *P. unibracteata*, *P. muelleri* and *P. paradoxa* showed conflicting organellar and nuclear DNA signals. The phylogeny presented in Chapter three had many branches with high BP support compared to the ITS phylogeny presented here where support values (PP) were mostly quite low. Thus, this conflict might simply be a result of poor sequence divergence among the four species, i.e. phylogenetic analyses may not be able to place them in the tree and they are placed as a polytomy at the base of the clade instead. Trans-Tasman hybridisation

events in the past may also have caused the incongruence, but this is a less likely explanation because migration, hybridisation, and extinctions must all be postulated. All four species are currently placed in sect. *Oliganthos* but if placement of *P. paradoxa* and *P. muelleri* follows that as indicated by the ITS phylogeny, they may have to be reclassified to sect. *Mesembrynia*.

The South American species are found in two groups in the ITS phylogeny (in Clade I and Clade III). Thus, the South American species are not showing patterns consistent with those expected as a result of Gondwanan vicariance, i.e. one clade each of New Zealand, South American and Australian species, with an expected relationship of ((Australia + South America) New Zealand) according to breakup sequence or multiple paralogous repeats of this pattern. In addition, *Plantago stauntoni* from the Amsterdam & St. Paul Islands (Southern Indian Ocean) is nested within the Australasian species. The phylogeny suggests that the ancestor of *P. stauntoni* dispersed there from Australasia but it is unclear from Fig. 4.3 exactly where it dispersed from.

For the three Australasian clades, there is little resolution and low support for any resolved clades with ITS alone (Figs. 4.3 - 4.5) but it was shown in Chapter three that a concatenated dataset, including organellar DNA sequences were able to resolve species relationships (Fig. 3.2). Splitstree analyses back up the topology reconstructed by ML analyses but the neighbour net analyses revealed several additional reticulate relationships that were not illustrated in the ML trees (Fig. 4.7). However, analysis using Splitstree also showed difficulty in delineating species within the group. The lack of resolution and low support for clades may be a result of the low sequence

divergence among Australasian species (Table 4.3, Appendix III (B)). For example, the interspecific p-distance range (0.002-0.008) between *Plantago triandra* and *P. unibracteata* is extremely low and results in no resolution between the two in the ML phylogeny. The low sequence divergence coupled with variable morphology is consistent with the pattern found in most New Zealand plants that have undergone recent speciation (e.g. Winkworth *et al.* 1999; Winkworth *et al.* 2002a); changes in relatively few loci may result in huge morphological differences (Winkworth *et al.* 1999). It may also be that the ITS region is not the best marker to capture genetic variation at this level. Other molecular data such as amplified fragment length polymorphism (AFLP) and randomly amplified polymorphic DNA (RAPD) are increasingly being applied for plant groups with low genetic variability and thus, use of these resources, along with wider sampling, may be able to improve resolution of species relationships in Australasian *Plantago* (e.g. Wolff and Morgan-Richards 1999; Meudt and Bayly 2008). If the New Zealand species have had a recent origin and not much genetic divergence, concatenation of multiple independent loci may also be helpful as long as little hybridisation in the group has occurred.

It was found during phylogenetic analyses that samples of *P. triandra* and *P. unibracteata* were inseparable genetically (Figs. 4.6A, 4.7C). It is interesting to note that *P. unibracteata* has previously been treated as a variety of *P. triandra* (Table 4.1). Wider population sampling of both species will be able to clarify if this observation stands, which may lead to reclassification of *P. unibracteata*.

It is also interesting to note that *Plantago aucklandica* appears to have a recent origin from within *P. lanigera* (Fig. 4.5A). *P. aucklandica* is found in Auckland Islands and

the subantarctic islands, and could be classified as a megaherb. Thus, if this relationship is true, then *P. aucklandica* appears to have a different history from that suggested by Wagstaff *et al.* (2007), which is that subantarctic megaherb lineages are old and are sister to large New Zealand clades.

Plantago sp. might represent a recent addition to the naturalised flora of New Zealand because it does not match any of the reported species in Sykes (in Webb *et al.* 1988). While *P. asiatica* is common and widespread in Asia, it has never been found in New Zealand. The taxonomic status of *P. tasmanica* and *P. daltonii* were discussed by Brown (1981), who mentioned that morphologies of the two species are very similar. Field and glasshouse experiments revealed that differences in leaf morphology remain consistent between species, although there is some degree of plasticity in leaf morphology. There were two sequences each of *P. daltonii* and *P. tasmanica* and they were all identical (except for a few ambiguous sites in the sequences). Their close genetic similarity mirrors their close morphological similarity. This suggests that the two may not be isolated from one another but there is not enough sampling to clarify their taxonomic status in this study.

Additionally, it was found that genetic divergence in New Zealand *Plantago* is highest within populations, followed by between populations. There is the least amount of variation between species. This pattern is also not expected because the ITS region undergoes concerted evolution (Álvarez and Wendel 2003), which should reduce genetic divergence within populations. Explanations for this unexpected pattern could be incomplete speciation within many lineages, or that the plants are undergoing dispersal around the North and South Islands. Additionally, *Plantago*

plants are wind-pollinated, which suggests that genetic exchange could occur over long distances and it also appears that biotic seed dispersal is not limited, as evidenced by widespread dispersal and presence of many cosmopolitan species in the genus.

4.4.4 Conclusions and future directions

In summary, this study is important because it shows the placement of most of the Australasian species in a worldwide *Plantago* phylogeny. DNA sequences for 13 Australasian species were added to the *Plantago* phylogeny, which now includes about 80 *Plantago* species out of ca. 210 so far, excluding *Littorella*. Analysis of the ITS dataset revealed that the Australasian species have very little interspecific genetic differences, even though extensive morphological variation may exist. Taxonomy has been difficult for the genus because of plastic morphological characters and flower reduction associated with wind pollination. The New Zealand *Plantago* species are no exception; the morphology of plants in cultivation changes dramatically according to environmental and seasonal change (see Fig. 4.1) The complex evolutionary history in the New Zealand *Plantago* species as indicated by Splitstree analyses require further investigation because processes such as hybridisation or incomplete lineage sorting may be involved in this group. In addition to reticulation, recurrent polyploidy appears to be a common speciation mechanism in the group, suggested by variable chromosome numbers and demonstrated by the Australasian species and is probably associated with hybrid speciation. The groups are also separated by altitude: whether plants are found in alpine/subalpine or lowland habitats appears to relate to the separation of the New Zealand groupings.

Additionally, it is evident from the ITS phylogeny presented in this study that the current taxonomic classifications of many New Zealand species (i.e. *P. spathulata, P. triandra*, *P. obconica*, *P. aucklandica* and *P. triantha*) need revision.

It is shown here that while ITS is the most commonly used marker in plant studies (Álvarez and Wendel 2003) and also in *Plantago* phylogenetic studies (Rønsted *et al.* 2002; Hoggard *et al.* 2003), it may not be the most suitable marker for investigating genetic variation in recently diverged species. Resolution in the phylogeny of Australasian species may be increased with wider population sampling and utilising markers from faster evolving regions such as from the cpDNA or non-sequence markers such as RAPD or AFLP. The relationships between the two subspecies currently classified under *P. spathulata* require further studies, with increased sampling of *P. spathulata* subsp. *picta*.

Additional chromosome counts should be made, especially for the *P. raoulii* from the S. I., which appears to be closely related to the 16-ploid *P.* "sylvester". Also, *P. lanigera* in the ITS phylogeny included samples within two populations that were widely separated in the phylogenetic tree (Fig. 4.5A) and in the Splitstree analysis (Fig. 4.7B). Given that two chromosome numbers are known within this complex (Grove and Hair 1971; Sykes 1988), chromosome counts would also be interesting for these samples.

Chapter Five: General Discussion

The aim of this thesis was to generate molecular sequence data to: 1) elucidate the phylogeny and biogeography of the Australasian species of *Plantago* (Plantaginaceae) and 2) investigate evolutionary patterns and taxonomy of the New Zealand *Plantago* species. One region each from nuclear DNA, chloroplast DNA and mitochondrial DNA were chosen for amplification and sequencing of DNA from 20 Australasian *Plantago* species. Phylogenies for markers from each of the three genomes (ITS, *ndhF-rpl32*, and *cox1*) were reconstructed for the Australasian species to examine biogeographic patterns of the group in the Southern Hemisphere. Following this, new ITS sequences of 150 *Plantago* individuals were integrated into a larger phylogeny of *Plantago*, which included published sequences of other species distributed worldwide to investigate placement of the Australasian species. Phylogenetic analyses of the three New Zealand clades that were resolved in the ITS and concatenated phylogenies revealed several evolutionary patterns and assist species delimitations of the New Zealand species, which are outlined in this chapter.

5.1 Choosing a suitable region for phylogenetic analyses

The most commonly used region for plant phylogenetic analyses is the Internal Transcribed Spacer (ITS) from nuclear DNA. Despite its advantages, markers from other regions may sometimes prove to be more useful for reconstructing phylogenies, either by themselves or in conjunction with ITS. In this thesis, amplification and sequencing using 24 primer pairs from all three plant genomes was tested on four *Plantago* species (*P. spathulata* subsp. *spathulata* and *P. triandra* subsp. *triandra* from New Zealand; *P. euryphylla* from Australia; and the cosmopolitan *P. lanceolata*)

and a relative from the same family (*Veronica stricta* from New Zealand). Out of eleven primer pairs with consistent, clean amplification across all species (Table 2.1), the ITS from the nuclear genome and the intergenic spacer *ndhF-rpl32* from the chloroplast genome were chosen for further phylogenetic studies because of their ease of amplification and adequate sequence diversity (2.8% for ITS and 2.1% for *ndhF-rpl32*, comparing the two New Zealand species). While mitochondrial DNA (mtDNA) substitution rates are traditionally regarded as very slow in plants (Wolfe *et al.* 1987), a high rate of mtDNA evolution has been found in *Plantago* (Cho *et al.* 2004). Therefore, *cox1* from the mtDNA was also chosen to compare rates found in Australasian mtDNA with rates found in previous studies, and for phylogeny reconstruction in this study.

Comparisons between the two New Zealand species revealed that ITS and *ndhF-rpl32* had relatively higher sequence variation than *cox1*, whereas *cox1* had more variation when comparing the Australasian species with the cosmopolitan *Plantago lanceolata* (Table 2.3). In Chapter Four, phylogenetic analyses indicated that ITS is not a suitable marker for investigating New Zealand species relationships because very little genetic differences were found between species. In Chapter Three, the *ndhF-rpl32* region was found to have better variation between species, and higher resolution in the phylogeny compared to the ITS region. It was also shown that the *cox1* marker provided very little resolution in the Australasian group. Thus, the use of fast evolving chloroplast markers is recommended for further studies involving Australasian *Plantago*.

5.2 Phylogeny of *Plantago*

Previous molecular phylogenies reconstructed for *Plantago* included only three and six species (Rønsted *et al.* 2002 and Hoggard *et al.* 2003, respectively) out of a total of 32 Australasian species. Inclusion of the Australasian species is essential for elucidating biogeographic patterns of Southern Hemisphere *Plantago*, and to investigate unclear Australasian species relationships. In this thesis, inclusion of sequences from 20 Australasian species in the worldwide ITS phylogeny revealed that all the Australasian species are placed in a clade of subg. *Plantago* (Fig. 4.2). The subg. *Plantago* clade is separate from a clade comprising the four other subgenera (subg. *Psyllium*, subg. *Albicans*, subg. *Bougeria*, and subg. *Coronopus*). This clade also includes a clade of species that are native to America (including South America), Africa, Asia, Ilha Trindade, Europe, and several cosmopolitan species. *P. stauntoni* from the Amsterdam & St. Paul Islands was nested within the Australasian species.

There was some incongruence between phylogenies reconstructed using the ITS dataset and either of the organelle datasets caused by different placements of sequences of *P. triandra* and *P. unibracteata* (New Zealand species), and *P. muelleri* and *P. paradoxa* (Australian species) in the phylogenies. Explanations for this incongruence included low resolution in the ITS phylogeny within the Australasian group, or reticulation in the evolutionary history of these species (e.g. trans-Tasman hybridisation). The lack of resolution in the ITS phylogeny is most likely because there was low support for the placement of these species in the ITS phylogeny (Fig. 4.3) but high support in the concatenated phylogeny (Fig. 3.2).

The phylogeny displaying branch lengths of the *cox1* region for the Australasian species (Fig. 3.3C) revealed that while there is an elevated rate of mitochondrial evolution between Australasian species and the two introduced species (*P. lanceolata* and *P. coronopus*) that were included, the elevated rate is not present among Australasian species. This could be due to recent radiations in the group, which is not uncommon in endemic New Zealand groups (Heenan *et al.* 2002; Murray *et al.* 2004; Wagstaff and Garnock-Jones 1998). However, different *Plantago* lineages have also been found to have either substitution rate increases or rate decreases (Cho *et al.* 2004) and the branch lengths could simply indicate that there has been a decrease in substitution rate in the Australasian *Plantago* species. The latter situation is more likely to be true, as there was no decrease in substitution rate evident in the other markers (ITS or *ndhF-rpl32*), which is expected if speciation was the result of more recent rapid radiations.

5.3 Biogeography of Australasian *Plantago*

Reconstructions of molecular phylogenies for Southern Hemisphere plant groups have revealed that long distance dispersal has been an important process in most plant groups (Radford *et al.* 2001; Swenson *et al.* 2001; von Hagen and Kadereit 2001; Heenan *et al.* 2002; Hurr *et al.* 1999; Lockhart *et al.* 2001; Winkworth *et al.* 2002; Zhang and Renner 2003; Perrie and Brownsey 2005; Barker *et al.* 2007; Perrie and Brownsey 2007; Perrie *et al.* 2007; Sanmartín *et al.* 2007) but there is still evidence of Gondwanan vicariance in a few plant lineages (Stöckler *et al.* 2002, Knapp *et al.* 2007). Most New Zealand plant groups also have one dispersal event to New Zealand, followed by speciation (Wagstaff and Garnock-Jones 1998; Winkworth *et al.* 1999; Perrie *et al.* 2003; Albach *et al.* 2005; Meudt and Simpson 2006). A meta-analysis of

plant phylogenetic studies has shown that dispersal between Australia and New Zealand is asymmetrical, and most dispersal events have occurred in the direction of West to East due to West Wind Drift and oceanic currents (Sanmartín *et al.* 2007).

In this thesis, the hypothesis of Gondwanan vicariance for the Southern Hemisphere *Plantago* was rejected because there was no evidence of monophyletic groups of Australian, New Zealand and South American species, respectively, in the phylogeny. Further, New Zealand clades and species are more closely related to Australian groups than to South American clades and species. Molecular dating techniques provided evidence for long distance dispersal within the Australasian *Plantago*. It was found that there was too little genetic divergence among the Australasian species to infer a Gondwanan origin for the group. The origin for the Australasian clade was estimated to be within 1.5 to 2.2 million years ago, which is consistent with the uplift of the Southern Alps in the South Island of New Zealand.

The phylogeny of the concatenated dataset (ITS + *ndhF-rpl32* + *cox1*) suggests at least three long distance dispersal events to New Zealand from Australia (Fig. 3.2). This direction of dispersal is consistent with patterns influenced by the West-Wind Drift and ocean currents (Sanmartín *et al.* 2007) but differs from the common pattern of one origin of most New Zealand groups. These long distance events also reflect dispersal of species that were already adapted to the alpine/subalpine or lowland environment in Australian lineages before establishment in New Zealand. One habitat shift event from the alpine/subalpine zone to lowland habitats in Australasian *Plantago* can be postulated from the phylogeny. Additionally, the ancestor of *P.*

stauntoni, which is native to the Amsterdam & St. Paul Islands, appears to have dispersed from within Australasia.

5.4 Taxonomy of New Zealand *Plantago*

Taxonomy of plant groups in New Zealand has been difficult because of extreme morphological variation but a common pattern of low genetic divergence, which is associated with the recent origin or recent radiations of many plant lineages in New Zealand (Winkworth *et al.* 1999; Smissen *et al.* 2004). Taxonomy within the genus *Plantago* has also been difficult because plants in the genus have extremely plastic morphological characters (Sykes in Webb *et al.* 1988; Rahn 1996). Difficult species delimitations in the New Zealand species of *Plantago* have resulted in multiple taxonomic changes in the past (Moore in Allan 1961; Sykes in Webb *et al.* 1988). These changes were used to represent hypotheses of species questions in this study. No previous molecular studies have included more than two New Zealand species when reconstructing the phylogeny of the genus (Rønsted *et al.* 2002; Hoggard *et al.* 2003).

In this thesis, phylogenetic analyses using ITS sequences were able to show that there was no genetic difference between morphological forms recognized within each of *Plantago lanigera*, *P. triandra* and *P. raoulii*. Therefore, the rank of *P. lanigera* and *P. raoulii* should be maintained; in addition, the two subspecies described under *P. triandra* are not supported by molecular data. Genetic distinctions between *P. triandra* and *P. unibracteata* were also unclear. *P. spathulata* subsp. *picta* and *P. spathulata* subsp. *spathulata* did not group together in the phylogeny, which indicates that the two are not as closely related as previously thought based on similarity of

morphological characters. However, only one sample of *P. spathulata* subsp. *picta* was obtained for this study. With more sampling, there may be substantial evidence for *P. spathulata* subsp. *picta* to be restored to species rank (*P. picta* as in (Moore in Allan 1961). The placement of *P. obconica* in the phylogeny of a concatenated ITS, *ndhF-rpl32* and *cox1* dataset (Fig. 3.2) clearly indicated that *P. obconica* is closely related to *P. lanigera*, not *P. triandra* as previously suggested (Sykes 1988) and should therefore be classified within sect. *Oliganthos*. Additionally, *Plantago triantha* perhaps should be placed in sect. *Mesembrynia* and *P. aucklandica* in sect. *Oliganthos*.

Low resolution found in phylogenies might be a result of low genetic divergence found among Australasian species using the ITS marker (Table 4.2, appendix 4.4), and the *ndhF-rpl32* and *cox1* markers (Table 2.3). The highest genetic divergence among New Zealand *Plantago* species (2.8401%) was found in ITS, which is also low.

Splitstree analyses (Fig. 4.7) and the ML phylogeny using ITS sequences (Fig. 4.4) revealed that the placement of several specimens of *P. lanigera* (from the Eyre Mountains and Sugarloaf Pass) were not well-defined (Fig. 4.6). Further studies should include more specimens of these, along with nearby samples, to determine if these could be hybrids, which could explain their ambiguous placements in the phylogenetic analyses.

5.5 Evolutionary patterns

Phylogenetic analyses suggest that adaptation to different habitats and altitudes (i.e. alpine/subalpine, lowland or coastal) may have played an important role in speciation within New Zealand *Plantago*. There are eight native New Zealand species: *P.*

lanigera, *P. obconica*, *P. unibracteata*, *P. aucklandica* and *P triandra* subsp. *triandra* can be found in the alpine/subalpine region; *P. triantha*, *P. triandra* subsp. *masoniae* and *P. spathulata* subsp. *picta* can be found in coastal areas; and *P. spathulata* subsp. *spathulata* and *P. raoulii* can be found from coastal regions to altitudes of ca. 1000 and 1500m, respectively (Sykes in Webb *et al.* 1988). Separation of species according to altitude is backed up by the ITS phylogeny (Fig. 4.3), and also in the concatenated (ITS, *ndhF-rpl32* and *cox1*) phylogeny (Fig. 3.2), where the New Zealand lowland species are separated from the alpine/subalpine species. The molecular phylogeny also suggests that the New Zealand species were pre-adapted to the alpine/subalpine and lowland habitats, with a more recent origin for the lowland species. There appears to have been a single migration event into lowland habitats from alpine/subalpine groups.

The New Zealand species have variable chromosome numbers (2*n* = 12, 24, 48, 60 and 96) which may act as a reproductive barrier where distributions overlap. For example, *P. unibracteata* and *P. lanigera* can be found growing at the same location but have different chromosome numbers (2*n* = 60 and 2*n* = 12/24, respectively). Polyploidy may be a mechanism of rapid speciation within the group. A good example is *P.* "sylvester" (2*n* = 96), which is suggested to be a hybrid of *P. spathulata* subsp. *spathulata* and *P. raoulii* (both 2*n* = 48). Different flowering times, along with the limited dispersal of pollen and seeds may also present a biological barrier to genetic exchange, which may promote speciation.

5.6 Future directions

One of the problems encountered in this thesis is that the New Zealand *Plantago* species have very little intra- and interspecific genetic divergence, resulting in little resolution within the group in all three markers. Wider population sampling within New Zealand or the addition of further independent genetic markers may help overcome this problem. Other genetic data such as randomly amplified polymorphic DNA (RAPD) and Amplified Fragment Length Polymorphism (AFLP) are often used for lineages that have little genetic divergence (e.g. Wolff and Morgan-Richards 1999; Meudt and Bayly 2008) and may be able to provide further resolution within the New Zealand groups. Wider population sampling throughout New Zealand and addition of genetic material would also help to clarify identities of ambiguously placed populations or individuals, such as the *P. lanigera* populations from Sugarloaf Pass and Eyre Mountains.

Additional chromosome counts to those reported by Groves and Hair (1971) would be invaluable in future studies for specimens suspected of being hybrids, like the South Island specimens of *P. raoulii* (i.e. chromosome counts will be able to determine if they are 16-ploids like *P.* "sylvester"). Additional counts and sampling should be made for *P.* "sylvester" specimens, including morphologically similar plants that have been reported in the Gouland Downs (Moore in Allan 1961). In addition, the two chromosome races within the *P. lanigera* complex also require further sampling and their relationship to morphological variation, habitats, distributions, and sympatry needs further work to clarify the taxonomic status of the individuals currently placed under *P. lanigera*.

For *Plantago* phylogeny reconstruction and biogeographic analyses, the ITS and *ndhF-rpl32* datasets could be expanded to include all of the Australian species, some samples of species distributed around the Pacific (such as Hawaii, Tonga, Juan Fernández, and New Guinea), and additional South American species. This would contribute to the understanding of distribution patterns, especially in the Southern Hemisphere. Additionally, genetic data could also be used to investigate unclear species relationships for Australian *Plantago*.

5.7 Conclusions

It was shown here that the *Plantago* phylogeny reconstructed using sequences from all three plant genomes, along with some molecular dating work, rejected the hypothesis of vicariance from Gondwana. Instead, the origin of the Australasian species appears to be through three long distance dispersal events of pre-adapted lineages. In addition, a phylogram of the *cox1* region suggests that the high rate of mtDNA evolution observed in *Plantago* by Cho *et al.* (2004) has slowed down in the Australasian group. In regard to New Zealand *Plantago* taxonomy, phylogenetic analyses in the Australasian group has shown that *P. lanigera*, *P. raoulii* and *P. triandra* should remain as species, with abolition of the subspecies within *P. triandra*, and that the two subspecies of *P. spathulata* should be recognised at species rank (as in Moore in Allan 1961). *P. obconica*, *P. triantha* and *P. aucklandica* need to be reclassified. Additionally, ITS may not be a suitable marker for investigating species boundaries within this group. *NdhF-rpl32* from the fast-evolving chloroplast region may be more useful for future phylogenetic work. Chromosome counts, utilization of additional genetic markers, and expansion of the molecular datasets (both within New

Zealand and outside) will be able to provide more answers, particularly for taxonomic work on the genus.

Literature cited

Albach, D. C. and M. W. Chase (2001). Paraphyly of *Veronica* (Veroniceae; Scrophulariaceae): Evidence from the internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA. *Journal of Plant Research* **114**(1113): 9-18.

Albach, D. C. and M. W. Chase (2004). Incongruence in Veroniceae (Plantaginaceae): evidence from two plastid and a nuclear ribosomal DNA region. *Molecular Phylogenetics and Evolution* **32**(1): 183-197.

Albach, D. C., H. M. Meudt and B. Oxelman (2005a). Piecing together the "new" Plantaginaceae. *American Journal of Botany* **92**(2): 297-315.

Albach, D. C., T. Utteridge and S. J. Wagstaff (2005b). Origin of Veroniceae (Plantaginaceae, formerly scrophulariaceae) on New Guinea. *Systematic Botany* **30**(2): 412-423.

Allan, H. H. (1961). *Flora of New Zealand. Vol I. Indigenous tracheophyta.* Wellington, Government Printer.

Álvarez, I. and J. F. Wendel (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29**(3): 417-434.

Bakker, F. T., F. Breman and V. Merckx (2006). DNA sequence evolution in fast evolving mitochondrial DNA *nad*1 exons in Geraniaceae and Plantaginaceae. *Taxon* **55**(4): 887-896.

Bakker, F. T., A. Culham, C. E. Pankhurst and M. Gibb (2000). Mitochondrial and cloroplast DNA-based phylogeny of *Pelargonium* (Geraniaceae) (vol 87, pg 727, 2000). *American Journal of Botany* **87**(7): 933-933.

Barker, N. P., P. H. Weston, F. Rutschmann and H. Sauquet (2007). Molecular dating of the 'Gondwanan' plant family Proteaceae is only partially congruent with the timing of the break-up of Gondwana. **34:** 2012-2027.

Barton, K. E. (2007). Early ontogenetic patterns in chemical defense in *Plantago* (Plantaginaceae): Genetic variation and trade-offs. *American Journal of Botany* **94**(1): 56-66.

Batt, G. E., J. Braun, B. P. Kohn and I. McDougall (2000). Thermochronological analysis of the dynamics of the Southern Alps, New Zealand. *Geological Society of America Bulletin* **112**(2): 250-266.

Bello, M. A., M. W. Chase, R. G. Olmstead, N. Ronsted and D. Albach (2002). The paramo endemic *Aragoa* is the sister genus of *Plantago* (*Plantaginaceae; Lamiales*): Evidence from plastid *rbcL* and nuclear ribosomal ITS sequence data. *Kew Bulletin* **57**(3): 585-597.

Breitwieser, I., D. S. Glenny, A. Thorne and S. J. Wagstaff (1999). Phylogenetic relationships in Australasian Gnaphalieae (Compositae) inferred from ITS sequences. *New Zealand Journal of Botany* **37**(3): 399-412.

Bremer, K. and M. H. G. Gustafsson (1997). East Gondwana ancestry of the sunflower alliance of families. *Proceedings of the National Academy of Sciences of the United States of America* **94**(17): 9188-9190.

Briggs, B. G. (1992). *Plantago*. Sydney, Royal Botanic Gardens & Domain Trust.

Brown, M. J. (1981). An Experimental Taxonomic Study of *Plantago Tasmanica* Hook. F. and *P. Daltonii* Dcne. *Australian Journal of Botany* **29**(4): 441-452.

Buckley, T. R. and C. Simon (2007). Evolutionary radiation of the cicada genus *Maoricicada* Dugdale (Hemiptera : Cicadoidea) and the origins of the New Zealand alpine biota. *Biological Journal of the Linnean Society* **91**(3): 419-435.

Carels, N. and G. Bernardi (2000). Two classes of genes in plants. *Genetics* **154**(4): 1819-1825.

Chacón, J., S. Madriñan, M. W. Chase and J. J. Bruhl (2006). Molecular phylogenetics of *Oreobolus* (Cyperaceae) and the origin and diversification of the American species. *Taxon* **55**(2): 359-366.

Cheeseman, T. F. (1906). *Manual of the New Zealand flora*. Wellington, Government Printer.

Cheeseman, T. F. (1925). *Manual of the New Zealand flora*. Wellington, Government Printer.

Cho, Y., J. P. Mower, Y. L. Qiu and J. D. Palmer (2004). Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* **101**(51): 17741-17746.

Chung, K. F., C. I. Peng, S. R. Downie, K. Spalik and B. A. Schaal (2005). Molecular systematics of the trans-Pacific alpine genus *Oreomyrrhis* (Apiaceae): Phylogenetic affinities and biogeographic implications. *American Journal of Botany* **92**(12): 2054-2071.

Connor, H. E. (1985). Biosystematics of Higher Plants in New Zealand 1965-1984. *New Zealand Journal of Botany* **23**(4): 613-643.

Cooper, A. and R. A. Cooper (1995). The Oligocene Bottleneck and New-Zealand Biota: Genetic Record of a Past Environmental Crisis. *Proceedings of the Royal Society of London Series B-Biological Sciences* **261**(1362): 293-302.

Cooper, R. A. and P. R. Millener (1993). The New Zealand Biota: Historical Background and New Research. *Trends in Ecology & Evolution* **8**(12): 429-433.

Cox, S. C. and R. H. Findlay (1995). The Main Divide Fault Zone and its role in formation of the Southern Alps, New Zealand. *New Zealand Journal of Geology and Geophysics* **38**(4): 489-499.

Crayn, D. M., M. Rossetto and D. J. Maynard (2006). Molecular phylogeny and dating reveals an Oligo-Miocene radiation of dry-adapted shrubs (former Tremandraceae) from rainforest tree progenitors (Elaeocarpaceae) in Australia. *American Journal of Botany* **93**(9): 1328-1342.

Cummings, M. P., S. P. Otto and J. Wakeley (1995). Sampling Properties of DNA Sequence Data in Phylogenetic Analysis. *Molecular Biology and Evolution* **12**(5): 814-822.

Dagar, J. C., Y. Kumar and O. S. Tomar (2006). Cultivation of medicinal isabgol (*Plantago ovata*) in alkali soils in semiarid regions of Northern India. *Land Degradation & Development* **17**(3): 275-283.

de Lange, P. J., P. B. Heenan, D. J. Keeling, B. G. Murray, R. Smissen and W. R. Sykes (2007). Biosystematics and Conservation: A Case Study with Two Enigmatic and Uncommon Species of *Crassula* from New Zealand**:** mcm294.

de Queiroz, A. (2005). The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution* **20**(2): 68-73.

Decker-Walters, D. S., S. M. Chung and J. E. Staub (2004). Plastid sequence evolution: A new pattern of nucleotide substitutions in the Cucurbitaceae. *Journal of Molecular Evolution* **58**(5): 606-614.

Demesure, B., N. Sodzi and R. J. Petit (1995). A Set of Universal Primers for Amplification of Polymorphic Noncoding Regions of Mitochondrial and Chloroplast DNA in Plants. *Molecular Ecology* **4**(1): 129-131.

Dhar, M. K., B. Friebe, S. Kaul and B. S. Gill (2006). Characterization and physical mapping of ribosomal RNA gene families in *Plantago*. *Annals of Botany* **97**(4): 541-548.

Doyle, J. J., J. I. Davis, R. J. Soreng, D. Garvin and M. J. Anderson (1992). Chloroplast DNA Inversions and the Origin of the Grass Family (Poaceae). *Proceedings of the National Academy of Sciences of the United States of America* **89**(16): 7722-7726.

Doyle, J. J. and J. D. Doyle (1990). Isolation of plant DNA from fresh tissue. *Focus* **12**: 13-15.

Drummond, A. J., S. Y. W. Ho, M. J. Phillips and A. Rambaut (2006). Relaxed phylogenetics and dating with confidence. *Plos Biology* **4**(5): 699-710.

Feldberg, K., J. Hentschel, R. Wilson, D. S. Rycroft, D. Glenny and J. Heinrichs (2007). Phylogenetic biogeography of the leafy liverwort *Herbertus* (Jungermanniales, Herbertaceae) based on nuclear and chloroplast DNA sequence data: correlation between genetic variation and geographical distribution. *Journal of Biogeography* **34**(4): 688-698.

Feliner, G. N. and J. A. Rosselló (2007). Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution* **44**(2): 911-919.

Felsenstein, J. (1981). Evolutionary Trees from DNA-Sequences - a Maximum-Likelihood Approach. *Journal of Molecular Evolution* **17**(6): 368-376.

Ford, K. A., J. M. Ward, R. D. Smissen, S. J. Wagstaff and I. Breitwieser (2007). Phylogeny and biogeography of *Craspedia* (Asteraceae : Gnaphalieae) based on ITS, ETS and psbA-trnH sequence data. *Taxon* **56**(3): 783-794.

Garnock-Jones, P., D. Albach and B. G. Briggs (2007). Botanical names in Southern Hemisphere *Veronica* (Plantaginaceae): sect. *Detzneria*, sect. *Hebe*, and sect. *Labiatoides*. *Taxon* **56**(2): 571-582.

Gemmill, C. E. C., G. J. Allan, W. L. Wagner and E. A. Zimmer (2002). Evolution of insular Pacific *Pittosporum* (Pittosporaceae): Origin of the Hawaiian radiation. *Molecular Phylogenetics and Evolution* **22**(1): 31-42.

Glenny, D. (2004). A revision of the genus *Gentianella* in New Zealand. *New Zealand Journal of Botany* **42**(3): 361-530.

Greaves, S. N. J., D. G. Chapple, D. M. Gleeson, C. H. Daugherty and P. A. Ritchie (2007). Phylogeography of the spotted skink (*Oligosoma lineoocellatum*) and green skink (*O. chloronoton*) species complex (Lacertilia : Scincidae) in New Zealand reveals pre-Pleistocene divergence. *Molecular Phylogenetics and Evolution* **45**(2): 729-739.

Groves, B. E. and J. B. Hair (1971). Contributions to a chromosome atlas of the New Zealand flora - 15 Miscellaneous families. *New Zealand Journal of Botany* **9**: 569-575.

GuhaMajumdar, M. and B. B. Sears (2005). Chloroplast DNA base substitutions: an experimental assessment. *Molecular Genetics and Genomics* **273**(2): 177-183.

Haase, M., B. Marshall and I. Hogg (2007). Disentangling causes of disjunction on the South Island of New Zealand: the Alpine fault hypothesis of vicariance revisited. *Biological Journal of the Linnean Society* **91**(3): 361-374.

Hale, M. L. and K. Wolff (2003). Polymorphic microsatellite loci in *Plantago lanceolata*. *Molecular Ecology Notes* **3**(1): 134-135.

Hasebe, M., T. Omori, M. Nakazawa, T. Sano, M. Kato and K. Iwatsuki (1994). *Rbcl* Gene-Sequences Provide Evidence for the Evolutionary Lineages of Leptosporangiate Ferns. *Proceedings of the National Academy of Sciences of the United States of America* **91**(12): 5730-5734.

Heenan, P. B., P. J. de Lange and A. D. Wilton (2001). Sophora (Fabaceae) in New Zealand: taxonomy, distribution, and biogeography. *New Zealand Journal of Botany* **39**(1): 17-53.

Heenan, P. B., P. J. Lockhart, K. McBreen and D. Havell (2006). Relationships in the alpine *Ranunculus haastii* (Ranunculaceae) complex and recognition of *R. piliferus* and *R. acraeus* from southern New Zealand. *New Zealand Journal of Botany* **44**(4): 425-441.

Heenan, P. B., A. D. Mitchell and M. Koch (2002). Molecular systematics of the New Zealand *Pachycladon* (Brassicaceae) complex: generic circumscription and relationship to *Arabidopsis* sens. lat. and *Arabis* sens. lat. *New Zealand Journal of Botany* **40**(4): 543-562.

Hegarty, M. J. and S. J. Hiscock (2005). Hybrid speciation in plants: new insights from molecular studies. *New Phytologist* **165**: 411-423.

Heinrichs, J., M. Lindner, H. Groth, J. Hentschel, K. Feldberg, C. Renker, J. J. Engel, M. von Konrat, D. G. Long and H. Schneider (2006). Goodbye or welcome Gondwana? insights into the phylogenetic biogeography of the leafy liverwort *Plagiochila* with a description of *Proskauera* gen. nov (Plagiochilaceae, Jungermanniales). *Plant Systematics and Evolution* **258**(3-4): 227-250.

Hillis, D. M. and M. T. Dixon (1991). Ribosomal DNA - Molecular Evolution and Phylogenetic Inference. *Quarterly Review of Biology* **66**(4): 410-453.

Hoggard, R. K., P. J. Kores, M. Molvray, G. D. Hoggard and D. A. Broughton (2003). Molecular systematics and biogeography of the amphibious genus *Littorella* (Plantaginaceae). *American Journal of Botany* **90**(3): 429-435.

Hooker, J. D. (1864). *Handbook of the New Zealand flora: a systematic description of the native plants of New Zealand and the Chatham, Kermadecs, Lord Auckland's, Campbell's, and Macquarrie's Islands*. London, Reeve & Co.

Houliston, G. J. and M. S. Olson (2006). Nonneutral evolution of organelle genes in *Silene vulgaris*. *Genetics* **174**(4): 1983-1994.

Howarth, D. G. and D. A. Baum (2002). Phylogenetic utility of a nuclear intron from nitrate reductase for the study of closely related plant species. *Molecular Phylogenetics and Evolution* **23**(3): 525-528.

Howarth, D. G. and D. A. Baum (2005). Genealogical evidence of homoploid hybrid speciation in an adaptive radiation of *Scaevola* (goodeniaceae) in the Hawaiian Islands. *Evolution* **59**(5): 948-961.

Howarth, D. G., M. H. G. Gustafsson, D. A. Baum and T. J. Motley (2003). Phylogenetics of the genus *Scaevola* (Goodeniaceae): Implication for dispersal patterns across the Pacific Basin and colonization of the Hawaiian Islands. *American Journal of Botany* **90**(6): 915-923.

Huelsenbeck, J. P. and F. Ronquist (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**(8): 754-755.

Hurr, K. A., P. J. Lockhart, P. B. Heenan and D. Penny (1999). Evidence for the recent dispersal of *Sophora* (Leguminosae) around the Southern Oceans: molecular data. *Journal of Biogeography* **26**(3): 565-577.

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

Jorgensen, A., T. K. Kristensen and J. R. Stothard (2007). Phylogeny and biogeography of African Biomphalaria (Gastropoda : Planorbidae), with emphasis on endemic species of the great East African lakes. *Zoological Journal of the Linnean Society* **151**(2): 337-349.

Kay, K. M., J. B. Whittall and S. A. Hodges (2006). A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *Bmc Evolutionary Biology* **6:** 36.

Keller, I., D. Bensasson and R. A. Nichols (2007). Transition-transversion bias is not universal: A counter example from grasshopper pseudogenes. *Plos Genetics* **3**(2): 185-191.

Knapp, M., R. Mudaliar, D. Havell, S. J. Wagstaff and P. J. Lockhart (2007). The drowning of New Zealand and the problem of *Agathis*. *Systematic Biology* **56**(5): 862-870.

Knapp, M., K. Stöckler, D. Havell, F. Delsuc, F. Sebastiani and P. J. Lockhart (2005). Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (southern beech). *Plos Biology* **3**(1): 38-43.

Koorevaar, G. N., S. Ivanovic, J. M. M. Van Damme, H. P. Koelewijn, W. P. C. Van't Westende, M. J. M. Smulders and B. Vosman (2002). Dinucleotide repeat microsatellite markers for buck's-horn plantain (*Plantago coronopus*). *Molecular Ecology Notes* **2**(4): 524-526.

Kozan, E., E. Küpeli and E. Yesilada (2006). Evaluation of some plants used in Turkish folk medicine against parasitic infections for their in vivo anthelmintic activity. *Journal of Ethnopharmacology* **108**(2): 211-216.

Kumar, S., K. Tamura and M. Nei (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* **5**(2): 150-163.

Kuriiwa, K., N. Hanzawa, T. Yoshino, S. Kimura and M. Nishida (2007). Phylogenetic relationships and natural hybridisation in rabbitfishes (Teleostei : Siganidae) inferred from mitochondrial and nuclear DNA analyses. *Molecular Phylogenetics and Evolution* **45**(1): 69-80.

Lewis, K. B., L. Carter and F. J. Davey (1994). The Opening of Cook Strait - Interglacial Tidal Scour and Aligning Basins at a Subduction to Transform Plate Edge. *Marine Geology* **116**(3-4): 293-312.

Linder, H. P. and M. D. Crisp (1995). *Nothofagus* and Pacific biogeography. *Cladistics-the International Journal of the Willi Hennig Society* **11**(1): 5-32.

Lockhart, P. J., P. A. McLenachan, D. Harell, D. Glenny, D. Huson and U. Jensen (2001). Phylogeny, radiation, and transoceanic dispersal of New Zealand alpine buttercups: Molecular evidence under split decomposition. *Annals of the Missouri Botanical Garden* **88**(3): 458-477.

Manos, P. S., J. J. Doyle and K. C. Nixon (1999). Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). *Molecular Phylogenetics and Evolution* **12**(3): 333-349.

Manzano, P. and J. E. Malo (2006). Extreme long-distance seed dispersal via sheep. *Frontiers in Ecology and the Environment* **5**(5): 244-248.

McBreen, K. and P. B. Heenan (2006). Phylogenetic relationships of *Pachycladon* (Brassicaceae) species based on three nuclear and two chloroplast DNA markers. *New Zealand Journal of Botany* **44**(4): 377-386.

McDowall, R. M. (2008). Process and pattern in the biogeography of New Zealand - a global microcosm? *Journal of Biogeography* **35**(2): 197-212.

McGlone, M. S. (2005). Goodbye Gondwana. *Journal of Biogeography* **32**(5): 739-740.

McGlone, M. S., R. P. Duncan and P. B. Heenan (2001). Endemism, species selection and the origin and distribution of the vascular plant flora of New Zealand. *Journal of Biogeography* **28**(2): 199-216.

Meudt, H. M. and M. J. Bayly (2008). Phylogeographic patterns in the Australasian genus *Chionohebe* (*Veronica* s.l., Plantaginaceae) based on AFLP and chloroplast DNA sequences. *Molecular Phylogenetics and Evolution* **In Press, Accepted Manuscript**.

Meudt, H. M. and B. B. Simpson (2006). The biogeography of the austral, subalpine genus *Ourisia* (Plantaginaceae) based on molecular phylogenetic evidence: South American origin and dispersal to New Zealand and Tasmania. *Biological Journal of the Linnean Society* **87**(4): 479-513.

Meudt, H. M. and B. B. Simpson (2007). Phylogenetic analysis of morphological characters in *Ourisia* (Plantaginaceae): Taxonomic and evolutionary implications. *Annals of the Missouri Botanical Garden* **94**(3): 554-570.

Mitchell, A. D. and P. B. Heenan (2002). Sophora sect. Edwardsia (Fabaceae): further evidence from nrDNA sequence data of a recent and rapid radiation around the Southern Oceans. *Botanical Journal of the Linnean Society* **140**(4): 435-441.

Mitchell, A. D., C. J. Webb and S. J. Wagstaff (1998). Phylogenetic relationships of species of *Gingidia* and related genera (Apiaceae, subfamily Apioideae). *New Zealand Journal of Botany* **36**(3): 417-424.

Molvray, M., P. J. Kores and M. W. Chase (1999). Phylogenetic relationships within *Korthalsella* (Viscaceae) based on nuclear ITS and plastid *trnL-F* sequence data. *American Journal of Botany* **86**(2): 249-260.

Mort, M. E., J. K. Archibald, C. P. Randle, N. D. Levsen, T. R. O'Leary, K. Topalov, C. M. Wiegand and D. J. Crawford (2007). Inferring phylogeny at low taxonomic levels: Utility of rapidly evolving cpDNA and nuclear ITS loci. *American Journal of Botany* **94**(2): 173-183.

Mower, J. P., P. Touzet, J. S. Gummow, L. F. Delph and J. D. Palmer (2007). Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. *Bmc Evolutionary Biology* **7**(1): 135.

Müller, K. (2005). SeqState: Primer Design and Sequence Statistics for Phylogenetic DNA Datasets. *Applied Bioinformatics* **4**(1): 65-69.

Mummenhoff, K., H. Bruggemann and J. L. Bowman (2001). Chloroplast DNA phylogeny and biogeography of *Lepidium* (Brassicaceae). *American Journal of Botany* **88**(11): 2051-2063.

Muñoz, J., A. M. Felicísimo, F. Cabezas, A. R. Burgaz and I. Martínez (2004). Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* **304**(5674): 1144-1147.

Murray, B. G., P. M. Datson, E. L. Lai, K. M. Sheath and E. K. Cameron (2004). Polyploidy, hybridisation and evolution in *Pratia* (Campanulaceae). *New Zealand Journal of Botany* **42**(5): 905-920.

Nilsson, E. and J. Ågren (2006). Population size, female fecundity, and sex ratio variation in gynodioecious *Plantago maritima*. *Journal of Evolutionary Biology* **19**(3): 825-833.

Nilsson, E., N. Gyllenstrand and K. Wolff (2006). Six polymorphic microsatellite markers for *Plantago maritima*. *Molecular Ecology Notes* **6**(4): 1093-1095.

Oxelman, B., M. Liden and D. Berglund (1997). Chloroplast rps16 intron phylogeny of the tribe *Sileneae* (*Caryophyllaceae*). *Plant Systematics and Evolution* **206**(1-4): 393-410.

Parkinson, C. L., J. P. Mower, Y. L. Qiu, A. J. Shirk, K. M. Song, N. D. Young, C. W. dePamphilis and J. D. Palmer (2005). Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. *Bmc Evolutionary Biology* **5**: 12.

Perrie, L. and P. Brownsey (2007). Molecular evidence for long-distance dispersal in the New Zealand pteridophyte flora. **34:** 2028-2038.

Perrie, L. R., M. J. Bayly, C. A. Lehnebach and P. J. Brownsey (2007). Molecular phylogenetics and molecular dating of the New Zealand Gleicheniaceae. *Brittonia* **59**(2): 129-141.

Perrie, L. R. and P. J. Brownsey (2005). Insights into the biogeography and polyploid evolution of new Zealand *Asplenium* from chloroplast DNA sequence data. *American Fern Journal* **95**(1): 1-21.

Perrie, L. R., P. J. Brownsey, P. J. Lockhart, E. A. Brown and M. F. Large (2003). Biogeography of temperate Australasian *Polystichum* ferns as inferred from chloroplast sequence and AFLP. *Journal of Biogeography* **30**(11): 1729-1736.

Petit, C., F. Bretagnolle and F. Felber (1999). Evolutionary consequences of diploid-polyploid hybrid zones in wild species. *Trends in Ecology & Evolution* **14**(8): 306-311.

Pole, M. (1994). The New-Zealand Flora - Entirely Long-Distance Dispersal. *Journal of Biogeography* **21**(6): 625-635.

Posada, D. and T. R. Buckley (2004). Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* **53**(5): 793-808.

Posada, D. and K. A. Crandall (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**(9): 817-818.

Radford, E. A., M. F. Watson and J. Preston (2001). Phylogenetic relationships of species of *Aciphylla* (Apiaceae, subfamily Apioideae) and related genera using molecular, morphological, and combined data sets. *New Zealand Journal of Botany* **39**(2): 183-208.

Rahn, K. (1996). A phylogenetic study of the Plantaginaceae. *Botanical Journal of the Linnean Society* **120**(2): 145-198.

Rambaut, A. and A. J. Drummond (2007). Tracer v1.4.

Raven, P. H. (1973). Evolution of Subalpine and Alpine Plant Groups in New Zealand. *New Zealand Journal of Botany* **11**: 177-200.

Reeves, G., M. W. Chase, P. Goldblatt, P. Rudall, M. F. Fay, A. V. Cox, B. Lejeune and T. Souza-Chies (2001). Molecular systematics of Iridaceae: Evidence from four plastid DNA regions. *American Journal of Botany* **88**(11): 2074-2087.

Renner, S. (2004). Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences* **165**(4): S23-S33.

Renner, S. S., D. B. Foreman and D. Murray (2000). Timing transantarctic disjunctions in the Atherospermataceae (Laurales): Evidence from coding and noncoding chloroplast sequences. *Systematic Biology* **49**(3): 579-591.

Rock, J. F. (1920). The genus *Plantago* in Hawaii. *American Journal of Botany* **7**(5): 195-U11.

Rønsted, N., M. W. Chase, D. C. Albach and M. A. Bello (2002). Phylogenetic relationships within *Plantago* (Plantaginaceae): evidence from nuclear ribosomal ITS and plastid trnL-F sequence data. *Botanical Journal of the Linnean Society* **139**(4): 323-338.

Rønsted, N., H. Franzyk, P. Mølgaard, J. W. Jaroszewski and S. R. Jensen (2003). Chemotaxonomy and evolution of *Plantago* L. *Plant Systematics and Evolution* **242**(1-4): 63-82.

Rønsted, N., E. Göbel, H. Franzyk, S. R. Jensen and C. E. Olsen (2000). Chemotaxonomy of *Plantago*. Iridoid glucosides and caffeoyl phenylethanoid glycosides. *Phytochemistry* **55**(4): 337-348.

Sanderson, M. J. and J. A. Doyle (2001). Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. *American Journal of Botany* **88**(8): 1499-1516.

Sang, T., D. J. Crawford and T. F. Stuessy (1997). Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84**(8): 1120-1136.

Sanmartín, I. and F. Ronquist (2004). Southern Hemisphere biogeography inferred by event-based models: Plant versus animal patterns. *Systematic Biology* **53**(2): 216-243.

Sanmartín, I., L. Wanntorp and R. C. Winkworth (2007). West Wind Drift revisited: testing for directional dispersal in the Southern Hemisphere using event-based tree fitting. *Journal of Biogeography* **34**(3): 398-416.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* **18**(2): 233-234.

Schuettpelz, E. and S. B. Hoot (2004). Phylogeny and biogeography of *Caltha* (Ranunculaceae) based on chloroplast and nuclear DNA sequences. *American Journal of Botany* **91**(2): 247-253.

Setoguchi, H., T. A. Osawa, J. C. Pintaud, T. Jaffre and J. M. Veillon (1998). Phylogenetic relationships within Araucariaceae based on *rbcL* gene sequences. *American Journal of Botany* **85**(11): 1507-1516.

Shaw, J., E. B. Lickey, J. T. Beck, S. B. Farmer, W. S. Liu, J. Miller, K. C. Siripun, C. T. Winder, E. E. Schilling and R. L. Small (2005). The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **92**(1): 142-166.

Shaw, J., E. B. Lickey, E. E. Schilling and R. L. Small (2007). Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* **94**(3): 275-288.

Shepherd, L. D., L. R. Perrie, B. S. Parris and P. J. Brownsey (2007). A molecular phylogeny for the New Zealand Blechnaceae ferns from analyses of chloroplast trnL-trnF DNA sequences. *New Zealand Journal of Botany* **45**(1): 67-80.

Simmons, M. P., K. Müller and A. P. Norton (2007). The relative performance of indel-coding methods in simulations. *Molecular Phylogenetics and Evolution* **44**(2): 724-740.

Smissen, R. D., I. Breitwieser and J. M. Ward (2004). Phylogenetic implications of trans-specific chloroplast DNA sequence polymorphism in New Zealand Gnaphalieae (Asteraceae). *Plant Systematics and Evolution* **249**(1-2): 37-53.

Smissen, R. D., P. J. Garnock-Jones and G. K. Chambers (2003). Phylogenetic analysis of ITS sequences suggests a Pliocene origin for the bipolar distribution of *Scleranthus* (Caryophyllaceae). *Australian Systematic Botany* **16**(3): 301-315.

Soltis, D. E., P. S. Soltis and J. A. Tate (2004). Advances in the study of polyploidy since Plant speciation. *New Phytologist* **161**(1): 173-191.

Spence, J. R. and W. R. Sykes (1989). Are *Plantago novae-zelandiae* L. Moore and *P. lanigera* Hook. f. (Plantaginaceae) different? *New Zealand Journal of Botany* **27**(4): 499-502.

Squirrell, J. and K. Wolff (2001). Isolation of polymorphic microsatellite loci in *Plantago major* and *P. intermedia*. *Molecular Ecology Notes* **1**(3): 179-181.

Stöckler, K., I. L. Daniel and P. J. Lockhart (2002). New Zealand Kauri (*Agathis australis* (D. don) Lindl., Araucariaceae) survives oligocene drowning. *Systematic Biology* **51**(5): 827-832.

Strand, A. E., J. LeebensMack and B. G. Milligan (1997). Nuclear DNA-based markers for plant evolutionary biology. *Molecular Ecology* **6**(2): 113-118.

Swenson, U., A. Backlund, S. McLoughlin and R. S. Hill (2001). *Nothofagus* biogeography revisited with special emphasis on the enigmatic distribution of subgenus *Brassospora* in New Caledonia. *Cladistics-the International Journal of the Willi Hennig Society* **17**(1): 28-47.

Swofford, D. L. (2002). PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sunderland, Massachusetts, Sinauer Associates.

Sykes, W. R. (1988). Notes on New Zealand *Plantago* Species. *New Zealand Journal of Botany* **26**(2): 321-323.

Sytsma, K. J., A. Litt, M. L. Zjhra, J. C. Pires, M. Nepokroeff, E. Conti, J. Walker and P. G. Wilson (2004). Clades, clocks, and continents: Historical and biogeographical analysis of Myrtaceae, Vochysiaceae, and relatives in the Southern Hemisphere. *International Journal of Plant Sciences* **165**(4): S85-S105.

Taberlet, P., L. Gielly, G. Pautou and J. Bouvet (1991). Universal Primers for Amplification of 3 Noncoding Regions of Chloroplast DNA. *Plant Molecular Biology* **17**(5): 1105-1109.

Tate, J. A. (2002). Systematics and evolution of *Tarasa* (Malvaceae): an enigmatic Andean polyploid genus. Austin, The University of Texas. **PhD**.

Trewick, S. A. and M. Morgan-Richards (2005). After the deluge: mitochondrial DNA indicates Miocene radiation and Pliocene adaptation of tree and giant weta (Orthoptera : Anostostomatidae). *Journal of Biogeography* **32**(2): 295-309.

Tripp, E. A. (2007). Evolutionary relationships within the species-rich genus *Ruellia* (Acanthaceae). *Systematic Botany* **32**(3): 628-649.

VanDijk, P. and T. Bakx-Schotman (1997). Chloroplast DNA phylogeography and cytotype geography in autopolyploid *Plantago media*. *Molecular Ecology* **6**(4): 345-352.

Vijverberg, K., L. Lie and K. Bachmann (2002). Morphological, evolutionary and taxonomic aspects of Australian and New Zealand *Microseris* (Asteraceae). *Australian Journal of Botany* **50**(1): 127-143.

von Hagen, K. B. and J. W. Kadereit (2001). The phylogeny of *Gentianella* (Gentianaceae) and its colonization of the southern hemisphere as revealed by nuclear and chloroplast DNA sequence variation. *Organisms Diversity & Evolution* **1**(1): 61-79.

Vriesendorp, B. and F. T. Bakker (2005). Reconstructing patterns of reticulate evolution in angiosperms: what can we do? *Taxon* **54**(3): 593-604.

Wagstaff, S. J., I. Breitwieser, C. Quinn and M. Ito (2007). Age of origin of enigmatic megaherbs from the subantarctic islands. *Nature Precedings* (hdl:10101/npre.2007.1272.1).

Wagstaff, S. J., M. J. Bayly, P. J. Garnock-Jones and D. C. Albach (2002). Classification, origin, and diversification of the New Zealand Hebes (Scrophulariaceae). *Annals of the Missouri Botanical Garden* **89**(1): 38-63.

Wagstaff, S. J., I. Breitwieser and U. Swenson (2006). Origin and relationships of the austral genus *Abrotanella* (Asteraceae) inferred from DNA sequences. *Taxon* **55**(1): 95-106.

Wagstaff, S. J. and P. J. Garnock-Jones (1998). Evolution and biogeography of the *Hebe* complex (Scrophulariaceae) inferred from ITS sequences. *New Zealand Journal of Botany* **36**(3): 425-437.

Wagstaff, S. J. and P. J. Garnock-Jones (2000). Patterns of diversification in *Chionohebe* and *Parahebe* (Scrophulariaceae) inferred from ITS sequences. *New Zealand Journal of Botany* **38**(3): 389-407.

Wagstaff, S. J., P. B. Heenan and M. J. Sanderson (1999). Classification, origins, and patterns of diversification in New Zealand Carmichaelinae (Fabaceae). *American Journal of Botany* **86**(9): 1346-1356.

Wagstaff, S. J. and F. Hennion (2007). Evolution and biogeography of *Lyallia* and *Hectorella* (Portulacaceae), geographically isolated sisters from the Southern Hemisphere. *Antarctic Science* **19**(4): 417-426.

Wagstaff, S. J., K. Martinsson and U. Swenson (2000). Divergence estimates of *Tetrachondra hamiltonii* and *T. patagonica* (Tetrachondraceae) and their implications for austral biogeography. *New Zealand Journal of Botany* **38**(4): 587-596.

Wagstaff, S. J. and J. Wege (2002). Patterns of diversification in New Zealand Stylidiaceae. *American Journal of Botany* **89**(5): 865-874.

Wanntorp, L. and H. E. Wanntorp (2003). The biogeography of *Gunnera* L.: vicariance and dispersal. *Journal of Biogeography* **30**(7): 979-987.

Webb, C. J., W. R. Sykes and P. J. Garnock-Jones (1988). *Flora of New Zealand. Volume IV. Naturalised Pteridophytes, Gymnosperms, Dicotyledons*. Christchurch, Botany Division DSIR.

White, T. J., T. Bruns, S. Lee and J. Taylor (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols : a guide to methods and applications* M. Innis, C. Gelfand, J. Sninsky and T. White. San Diego, Academic Press**:** 315-322.

Winkworth, R. C., J. Grau, A. W. Robertson and P. J. Lockhart (2002a). The origins and evolution of the genus *Myosotis* L. (Boraginaceae). *Molecular Phylogenetics and Evolution* **24**(2): 180-193.

Winkworth, R. C., A. W. Robertson, F. Ehrendorfer and P. J. Lockhart (1999). The importance of dispersal and recent speciation in the flora of New Zealand. *Journal of Biogeography* **26**(6): 1323-1325.

Winkworth, R. C., S. J. Wagstaff, D. Glenny and P. J. Lockhart (2002b). Plant dispersal NEWS from New Zealand. *Trends in Ecology & Evolution* **17**(11): 514-520.

Winkworth, R. C., S. J. Wagstaff, D. Glenny and P. J. Lockhart (2005). Evolution of the New Zealand mountain flora: Origins, diversification and dispersal. *Organisms Diversity & Evolution* **5(**3): 237-247.

Wolfe, K. H., W. H. Li and P. M. Sharp (1987). Rates of Nucleotide Substitution Vary Greatly among Plant Mitochondrial, Chloroplast, and Nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America* **84**(24): 9054-9058.

Wolff, K. and M. Morgan-Richards (1999). The use of RAPD data in the analysis of population genetic structure: case studies of *Alkanna* (Boraginaceae) and *Plantago* (Plantaginaceae). *Molecular systematics and plant evolution*. P. M. Hollingsworth, R. M. Bateman and R. J. Gornall. New York, Taylor & Francis.

Wolff, K. and B. Schaal (1992). Chloroplast DNA Variation within and among 5 *Plantago* Species. *Journal of Evolutionary Biology* **5**(2): 325-344.

Wright, S. D., C. G. Yong, J. W. Dawson, D. J. Whittaker and R. C. Gardner (2000). Riding the ice age El Nino? Pacific biogeography and evolution of Metrosideros subg. Metrosideros (Myrtaceae) inferred from nuclear ribosomal DNA. *Proceedings of the National Academy of Sciences of the United States of America* **97**(8): 4118-4123.

Xia, X. and Z. Xie (2001). DAMBE: Software package for data analysis in molecular biology and evolution. *Journal of Heredity* **92**(4): 371-373.

Yoder, A. D., J. A. Irwin and B. A. Payseur (2001). Failure of the ILD to determine data combinability for slow loris phylogeny. *Systematic Biology* **50**(3): 408-424.

Yokoyama, J., M. Suzuki, K. Iwatsuki and M. Hasebe (2000). Molecular phylogeny of *Coriaria*, with special emphasis on the disjunct distribution. *Molecular Phylogenetics and Evolution* **14**(1): 11-19.

Zane, L., L. Bargelloni and T. Patarnello (2002). Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**(1): 1-16.

Zhang, L. B. and S. Renner (2003). The deepest splits in Chloranthaceae as resolved by chloroplast sequences. *International Journal of Plant Sciences* **164**(5): S383-S392.

Appendix I: Details of sample collections

A. Samples used in primer pair assays (Chapter two).

Taxon	Collection location	Indigenous distribution	Voucher specimen
P. triandra subsp. *triandra*	Lake Sylvester, NW Nelson, New Zealand	NZ	WELTU20163 (*MLT021 et al.*)
P. spathulata subsp. *spathulata*	Between Kettlehole Bog and Kettlehole Tarn, Cass, New Zealand	NZ	WELTU20118 (*PGJ2257*)
P. euryphylla	Kosciuszko National Park, Southern Tablelands, NSW, Australia	AUS	NSW742956 (*BGB9743*)
P. lanceolata	Ponsonby Road, Karori, Wellington, New Zealand	COSMO	WELTU20184 (*PGJ2551*)
P. coronopus	End of Red Rocks, Owhiro Bay, Wellington, New Zealand	MED, EUR	WELTU20183 (*PGJ2549*)
P. major	Ben Burn Park, Wellington, New Zealand	COSMO	WELTU20180 (*PGJ2550*)
Veronica stricta	VUW, Mount Street, Wellington, New Zealand	NZ	WELTU (*MLT001*)

Distibution: NZ = New Zealand, AUS = Australia, COSMO = cosmopolitan, MED = Mediterranean, EUR = Europe. Samples collected from New Zealand are deposited in the herbarium at the herbarium of Victoria University of Wellington, Wellington, New Zealand (WELTU). Australian voucher specimens can be found at the National Herbarium of New South Wales (NSW).

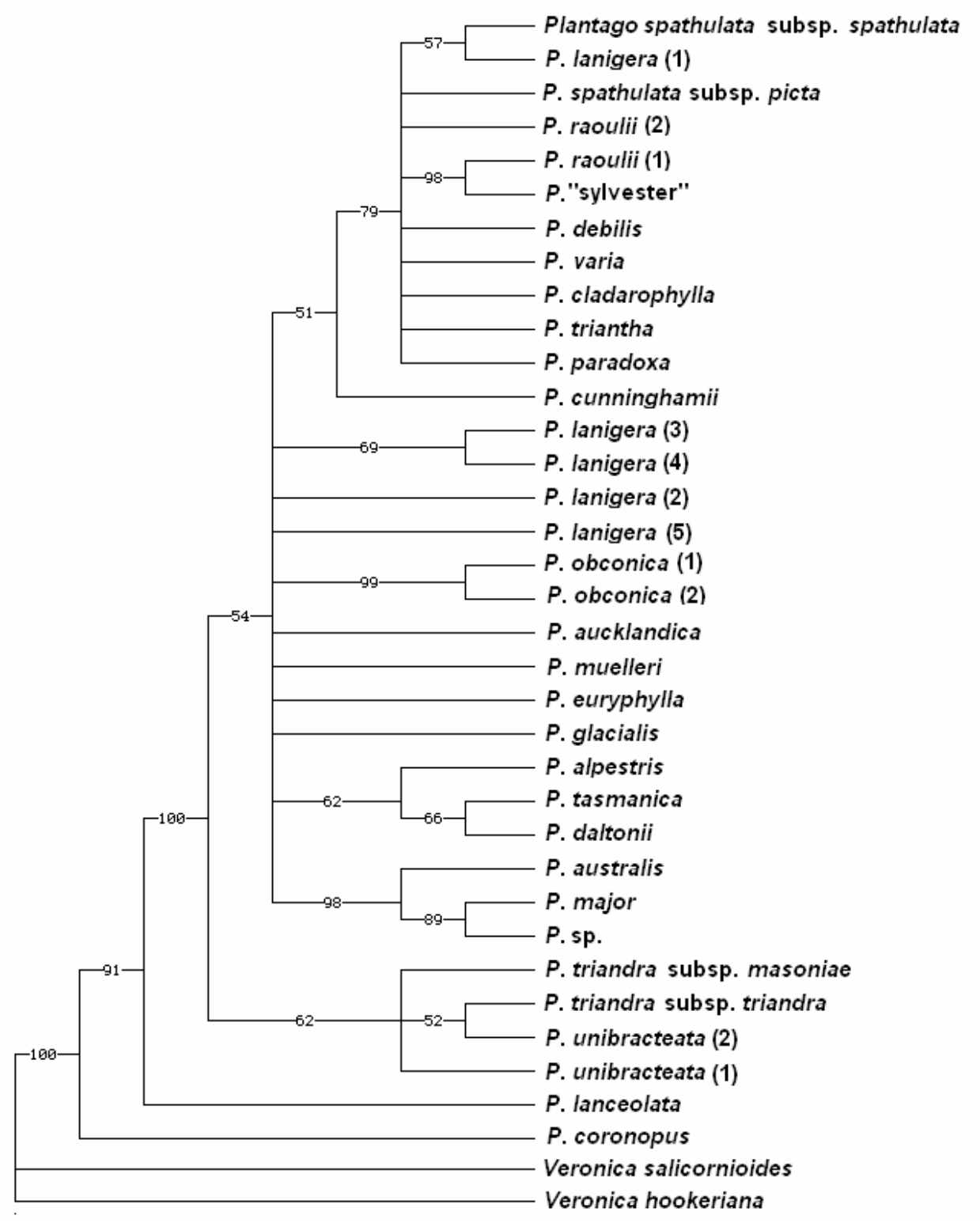
B. Species and accession numbers of ITS sequences used in molecular dating analyses (Chapter three).

Species	Collected from	Reference/Herbarium Voucher
Plantago "sylvester"	Lake Sylvester, Nelson, NZ	WELTU20150
P. alpestris	Kosciuszko National Park, NSW, AUS	NSW742962
P. aucklandica	Auckland Islands, NZ	In cult. Otari (*2003.2037*)
P. australis	Hunua Ranges, Auckland, NZ	WELTU20181
P. cladarophylla	Barrington Tops National Park, NSW, AUS	NSW744803
P. coronopus	Island Bay, Wellington, NZ	WELTU20183
P. cunninhamii	Brigalow Park Nature Reserve, NSW, AUS	NSW744804
P. daltonii	St. Clair National Park, AUS, Tasmania	NSW743874
P. debilis	Barrenjoey Headland, NSW, AUS	NSW 742894
P. euryphylla	Kosciuszko National Park, NSW, AUS	NSW742956
P. glacialis	Kosciuszko National Park, NSW, AUS	NSW743813
P. lanceolata	Karori, Wellington, NZ	WELTU20184
P. lanigera (3)	Hall Range, Canterbury, NZ	WELTU20143
P. major	Ben Burn Park, Karori, NZ	WELTU20180
P. muelleri	Kosciuszko National Park, NSW, AUS	NSW743812
P. obconica	Cardrona Skifield, Wanaka, NZ	WELTU20121.
P. paradoxa	St. Clair National Park, AUS, Tasmania	WELTU20187
P. raoulii (1)	Lake Sarah, Cass, NZ	WELTU20153
P. sp.	Pukerua Bay, Wellington, NZ	WELTU20178
P. spathulata subsp. *picta*	East Cape, Gisborne, NZ	CHR439486
P. spathulata subsp. *spathulata*	Marfells Beach, Marlborough, NZ	WELTU20117
P. tasmanica	St. Clair National Park, AUS, Tasmania	WELTU20188
P. triandra subsp. *masoniae*	Paturau Coast, Nelson, NZ	WELTU20168
P. triandra subsp. *triandra*	Lake Sylvester, Nelson, NZ	WELTU20163
P. triantha	Enderby Island, AI, NZ	WELTU20177
P. unibracteata	Lake Sylvester, Nelson, NZ	WELTU20175
P. varia	Kosciuszko National Park, NSW, AUS	NSW743869
Aragoa. corrugatifolia (AJ548980)	Colombia	Hoggard *et al.* 2003
A. cupressina (AJ459402)	northern Andes	Bello *et al.* 2002
Littorella americana (AJ548956)	Anderson Lake	Hoggard *et al.* 2007
L. americana (AJ548958)	New Brunswick	Hoggard *et al.* 2007
L. americana (AJ548957)	Trout Lake	Hoggard *et al.* 2007
L. australis (AJ548959)	Falkland Islands	Hoggard *et al.* 2007
L. uniflora (AJ548960)	Denmark	Hoggard *et al.* 2007
L. uniflora (AJ548963)	Iceland	Hoggard *et al.* 2007
L. uniflora (AJ548961)	Sweden	Hoggard *et al.* 2007
Veronica hookeriana	Whanahuia Range, Ruahine Mts, NZ	WELTU (*PGJ2458*)
Veronica salicornioides	Jacks Pass, Hanmer, Canterbury, NZ	CHR512475

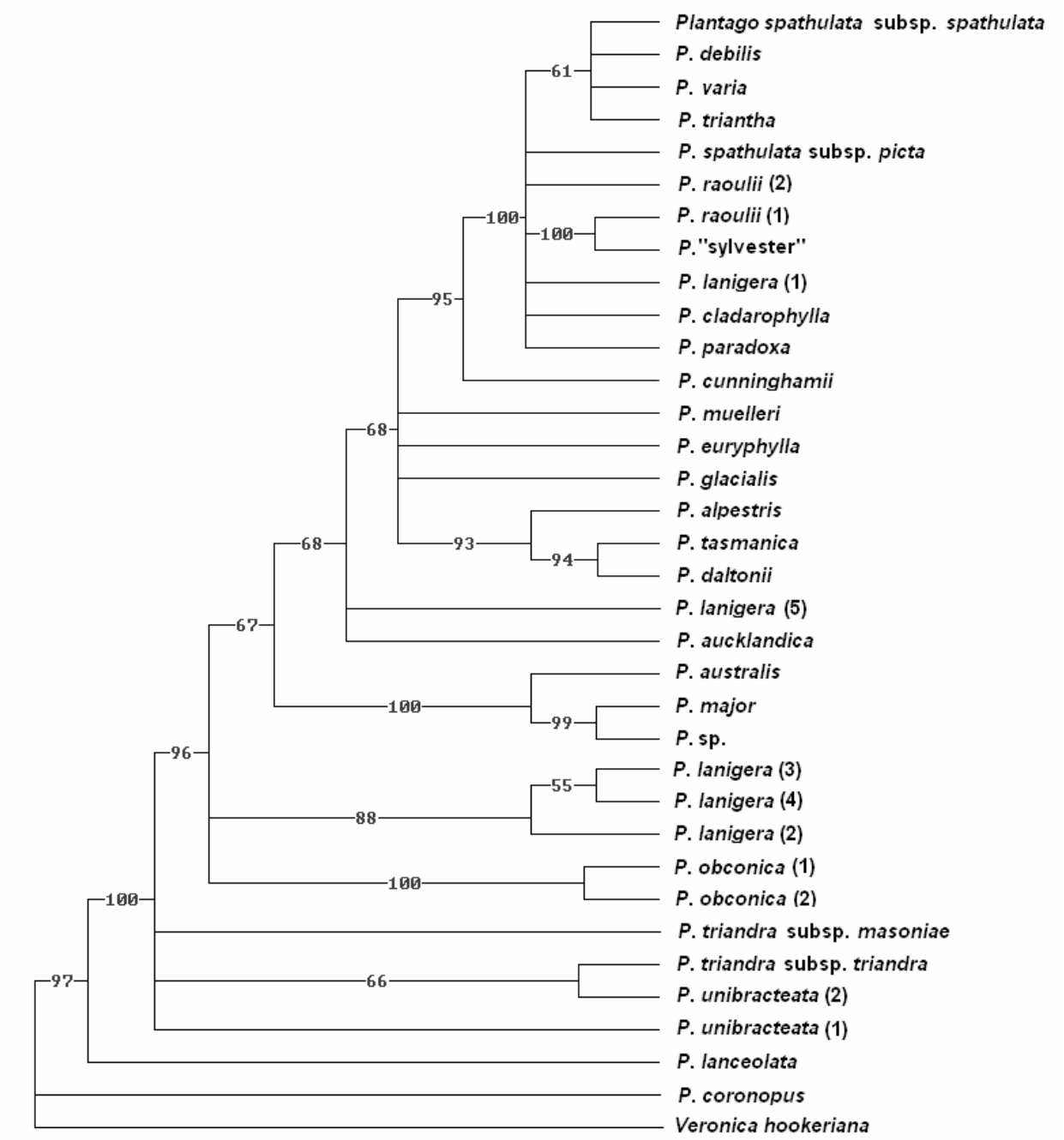
Locations: AI = Auckland Islands, AUS = Australia, NSW = New South Wales, NZ = New Zealand, RBG Kew = Royal Botanic Gardens Kew, UK = United Kingdom. Collectors: *BGB* = Barbara G. Briggs, *MLT* = Mei-Lin Tay, *PBH* = Peter Heenan, *PGJ* = Phil Garnock-Jones, *PJL* = Peter J. Lockhart and *VT* = Vanessa Thorn. Herbarium vouchers: WELTU = H. D. Gordon Herbarium in Victoria University of Wellington, New Zealand; NSW = National Herbarium of New South Wales, Australia; and CHR = Allan Herbarium, Landcare Research, Christchurch, New Zealand. The *Littorella* species included here had the same sequences for replicate samples of the same species.

Appendix II: Phylogenetic analyses

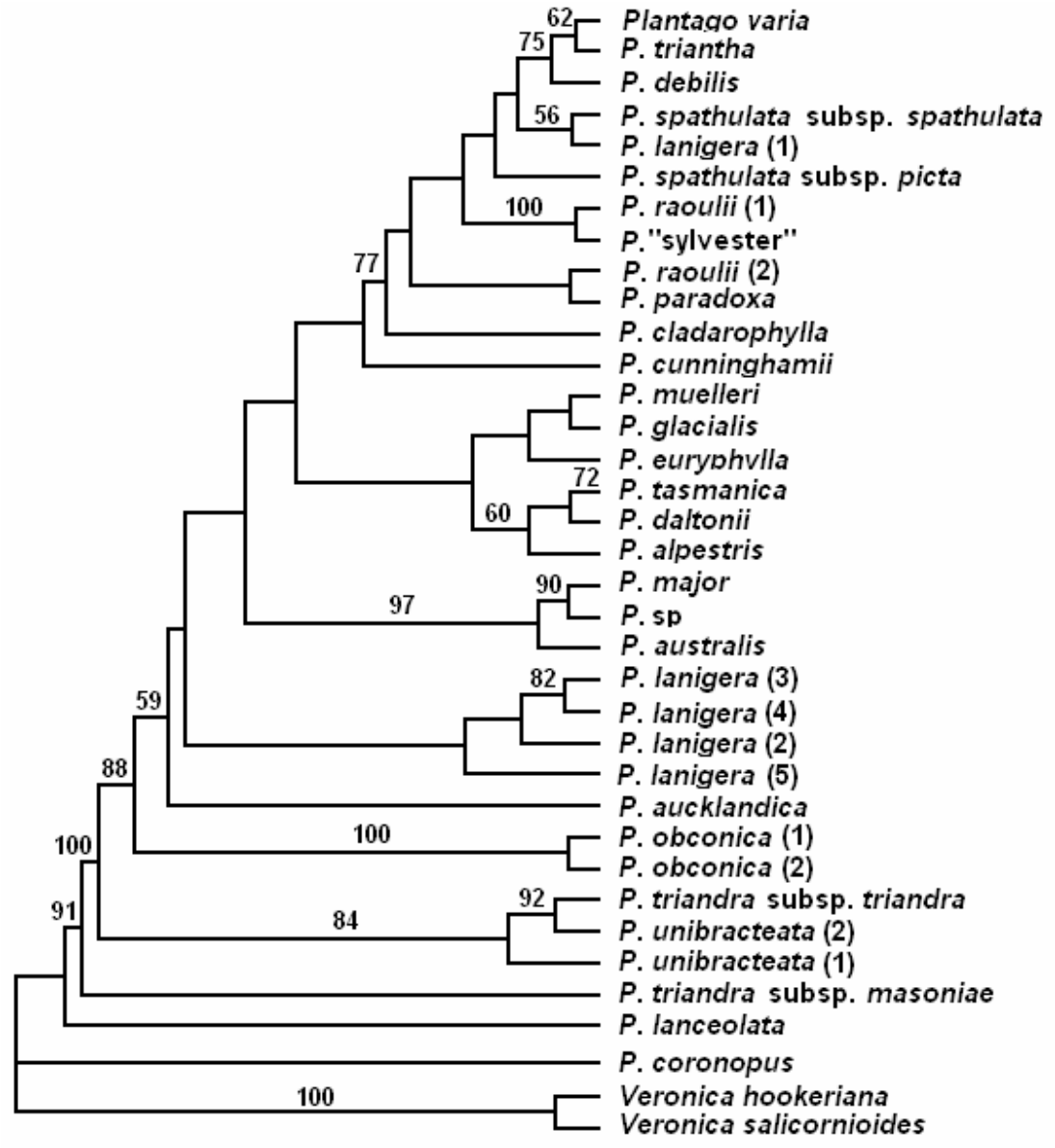
A. 50% bootstrap consensus phylogeny of the ITS dataset recovered using Maximum Parsimony (using samples listed in Table 3.1). Tree scores range from 350-356.



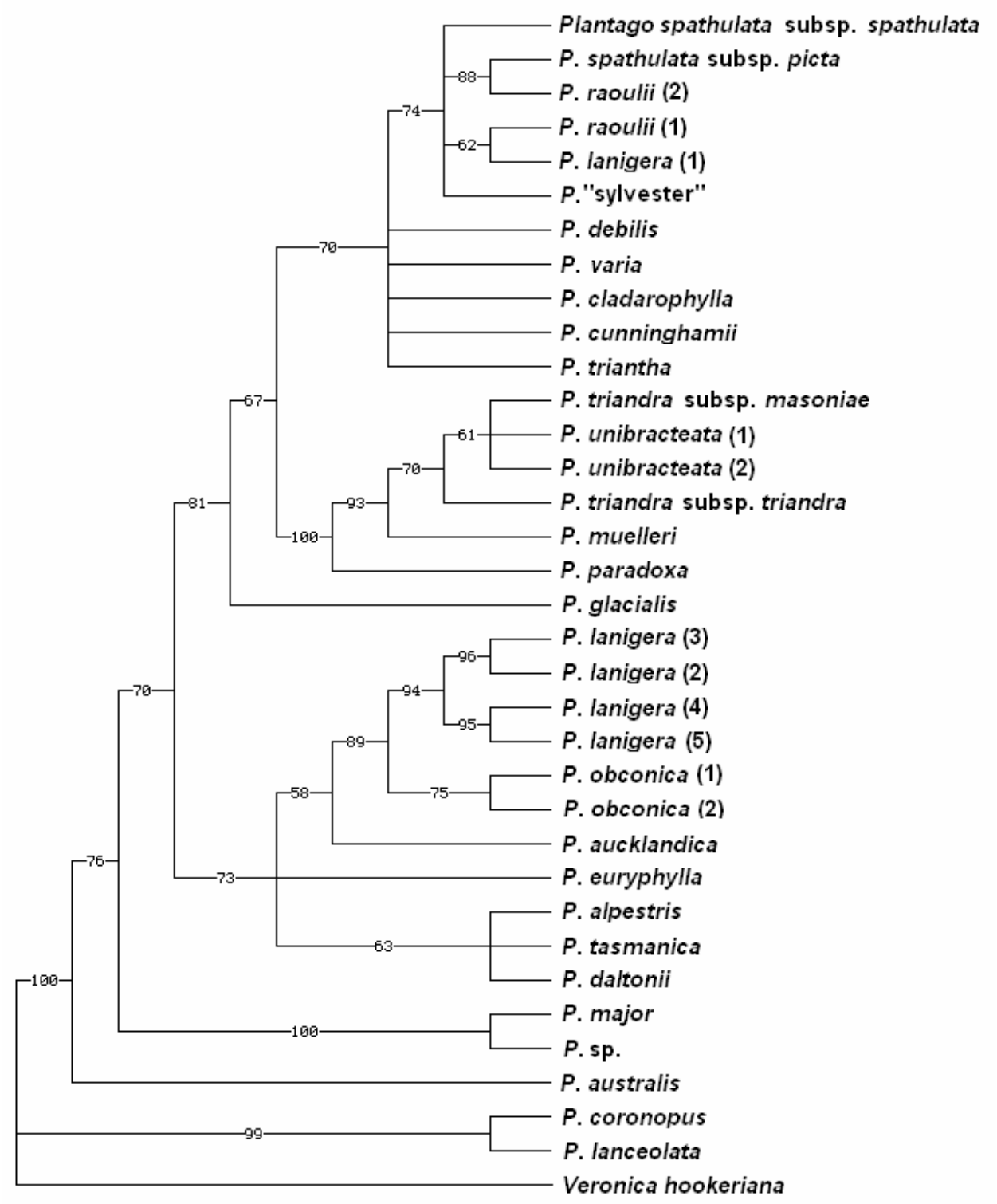
B. 50% posterior probability consensus phylogeny of the ITS dataset recovered using Bayesian analysis (using samples listed in Table 3.1). Estimated tree likelihood score = -2417.00.



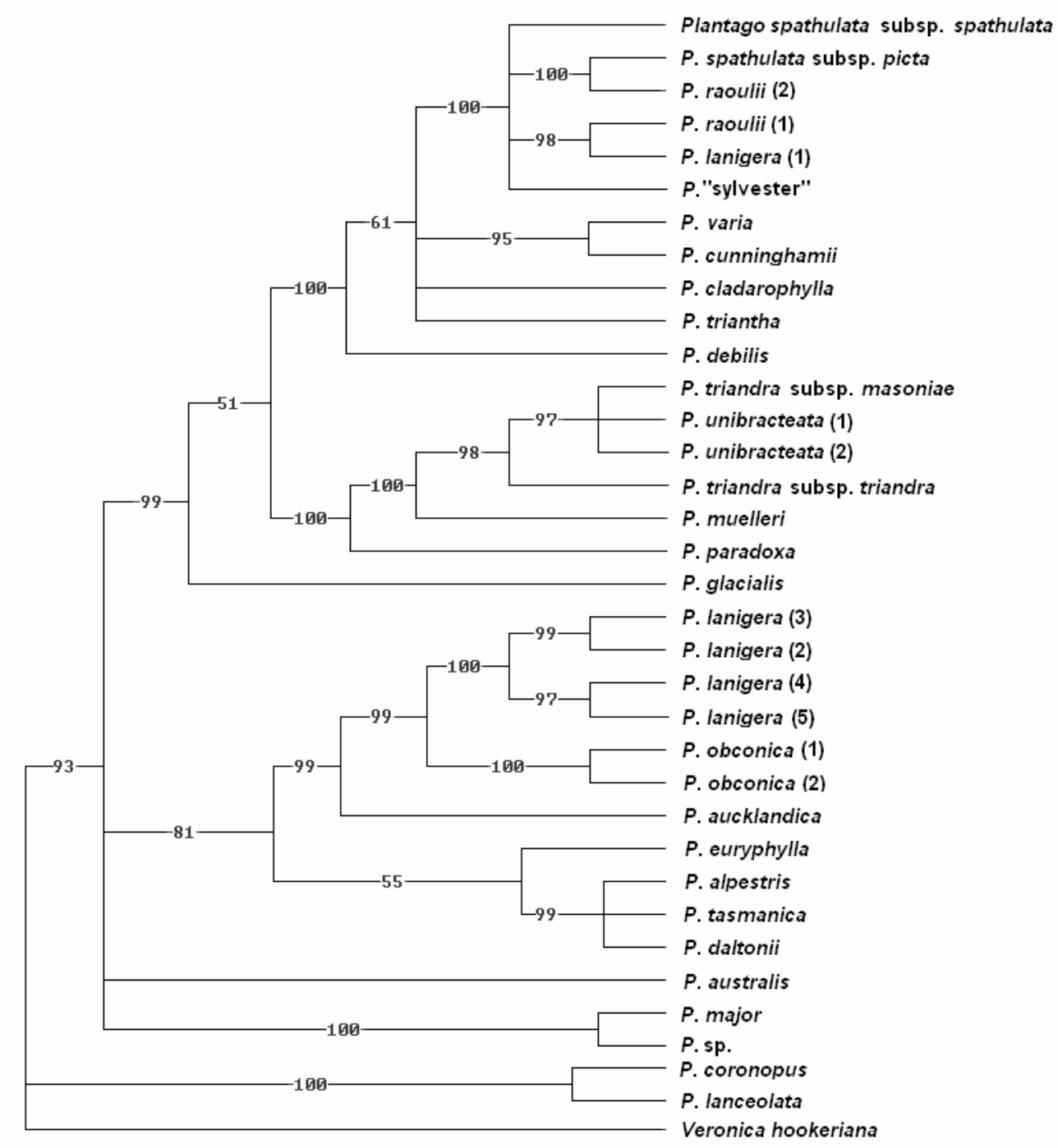
C. 50% bootstrap consensus phylogeny of the ITS dataset recovered using neighbour-joining methods (using samples listed in Table 3.1).



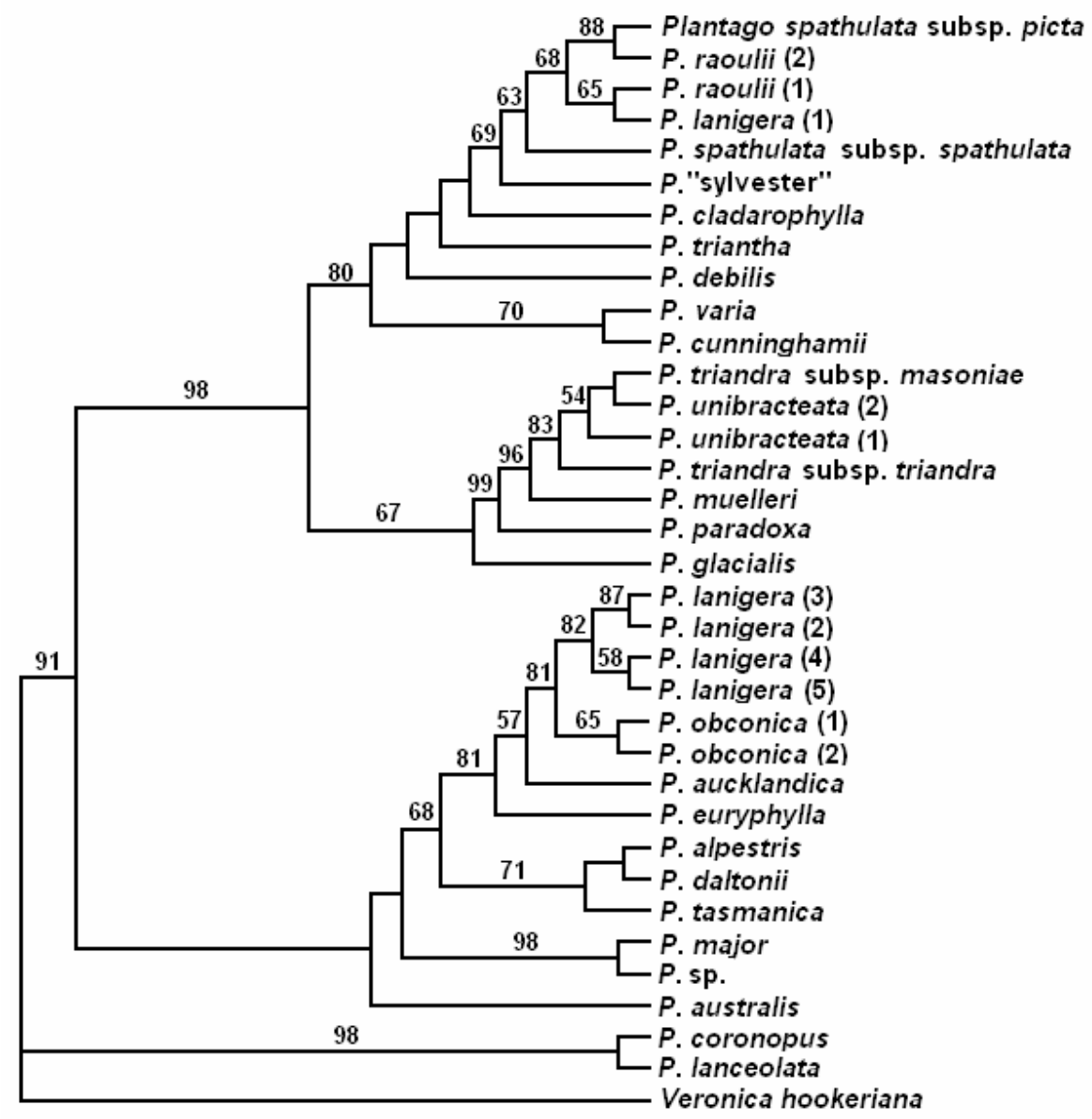
D. 50% bootstrap consensus phylogeny of the *ndhF-rpl32* dataset recovered using Maximum Parsimony (using samples listed in Table 3.1). Tree scores range from 496-500.



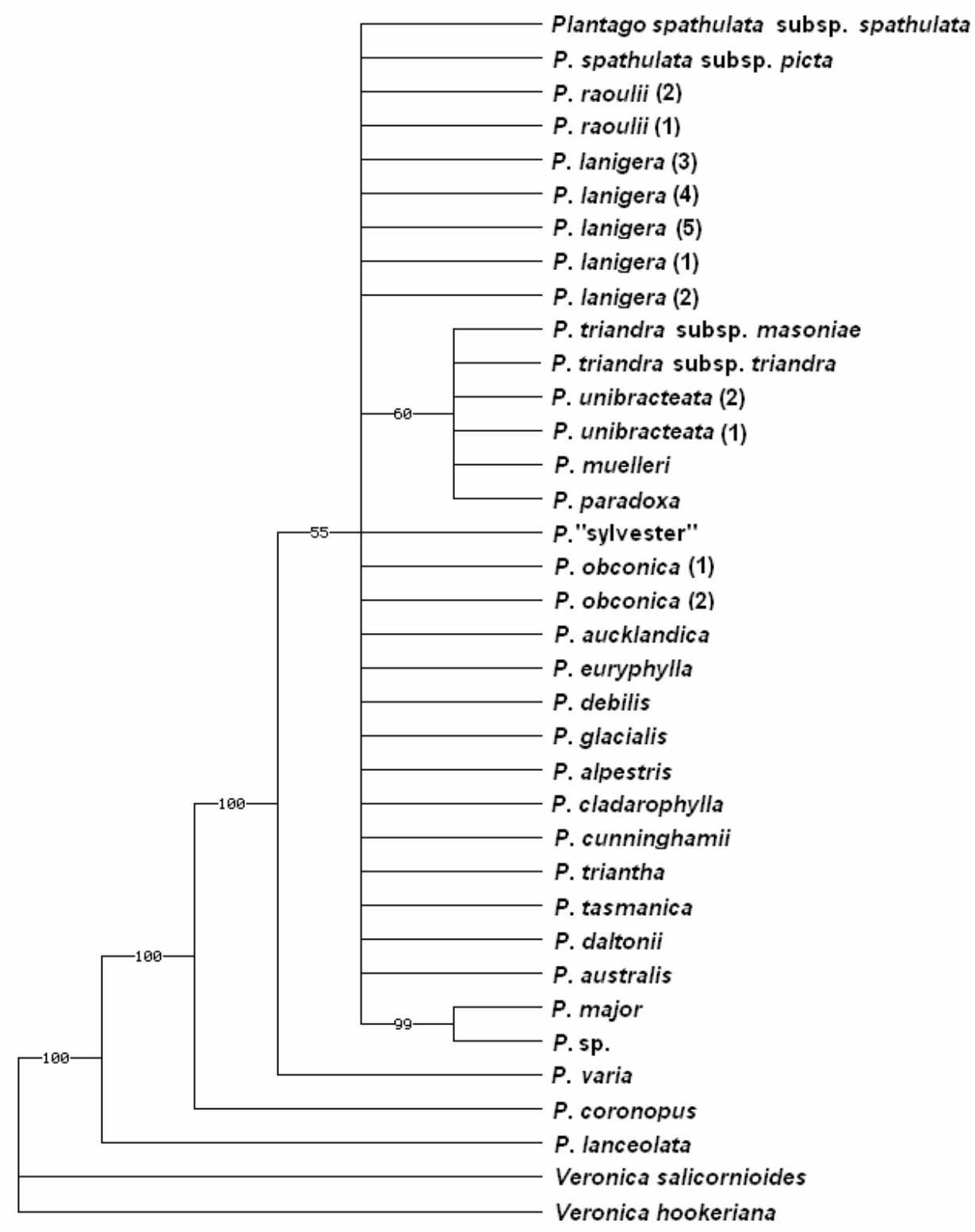
E. 50% posterior probability consensus phylogeny of the *ndhF-rpl32* dataset recovered using Bayesian analysis (using samples listed in Table 3.1). Estimated likelihood score = -3079.09.



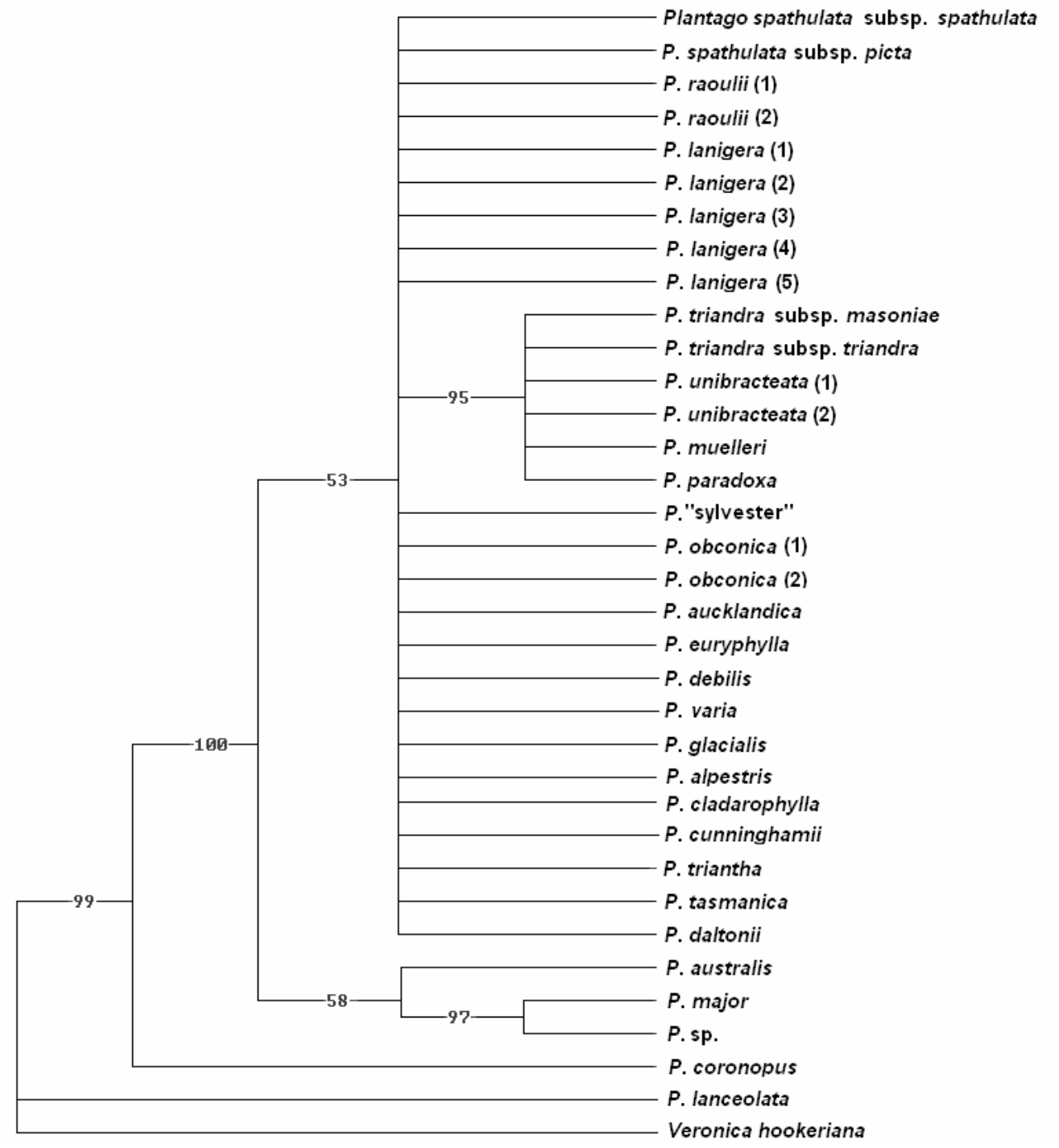
F. 50% bootstrap consensus phylogeny of the *ndhF-rpl32* dataset recovered using neighbour-joining methods (using samples listed in Table 3.1).



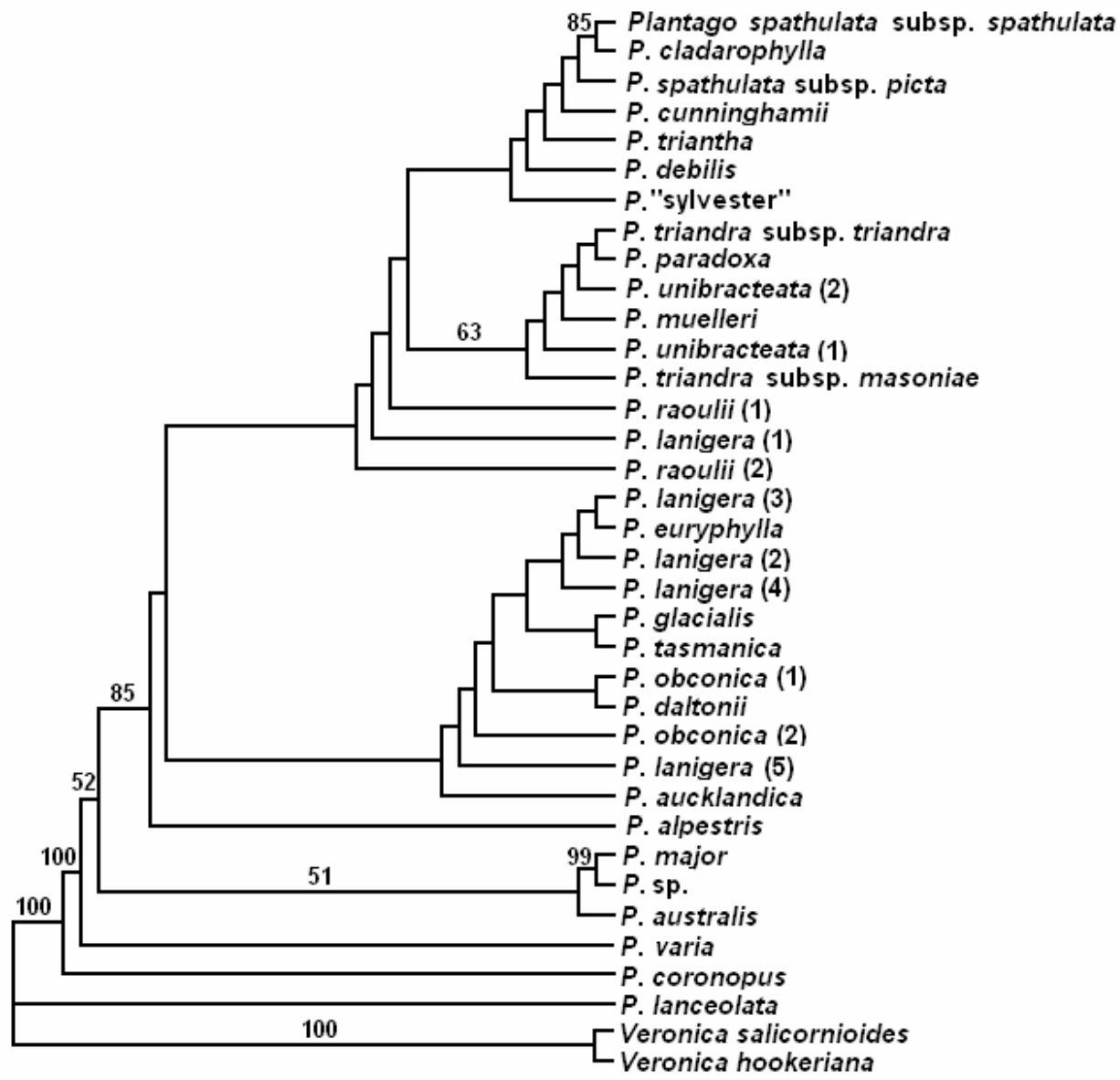
G. 50% bootstrap consensus phylogeny of the *cox1* dataset recovered using Maximum Parsimony (using samples listed in Table 3.1). Tree scores range from 198-199.



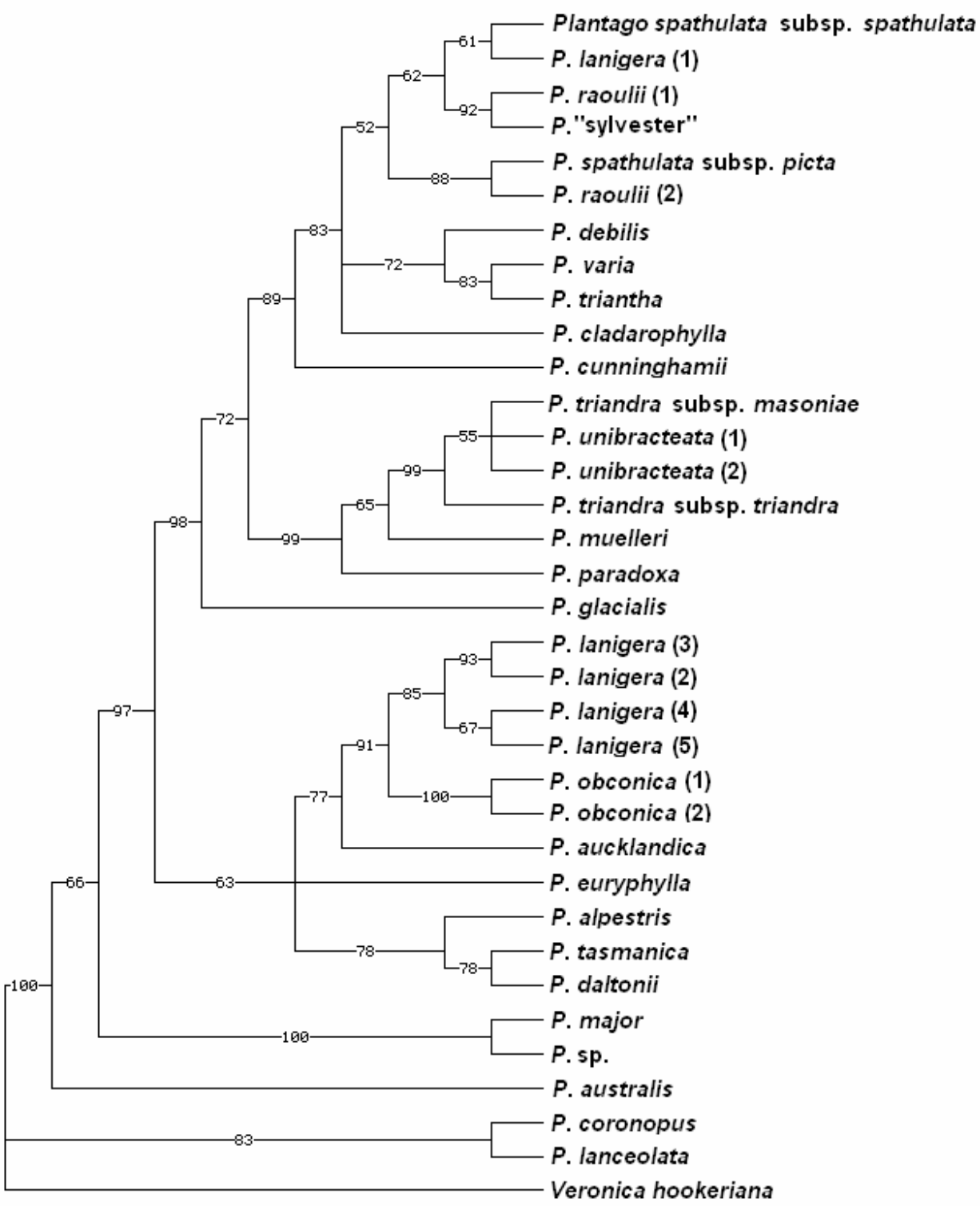
H. 50% posterior probability consensus phylogeny of the *cox1* dataset recovered using Bayesian analysis (using samples listed in Table 3.1). Estimated likelihood score = -1838.13.



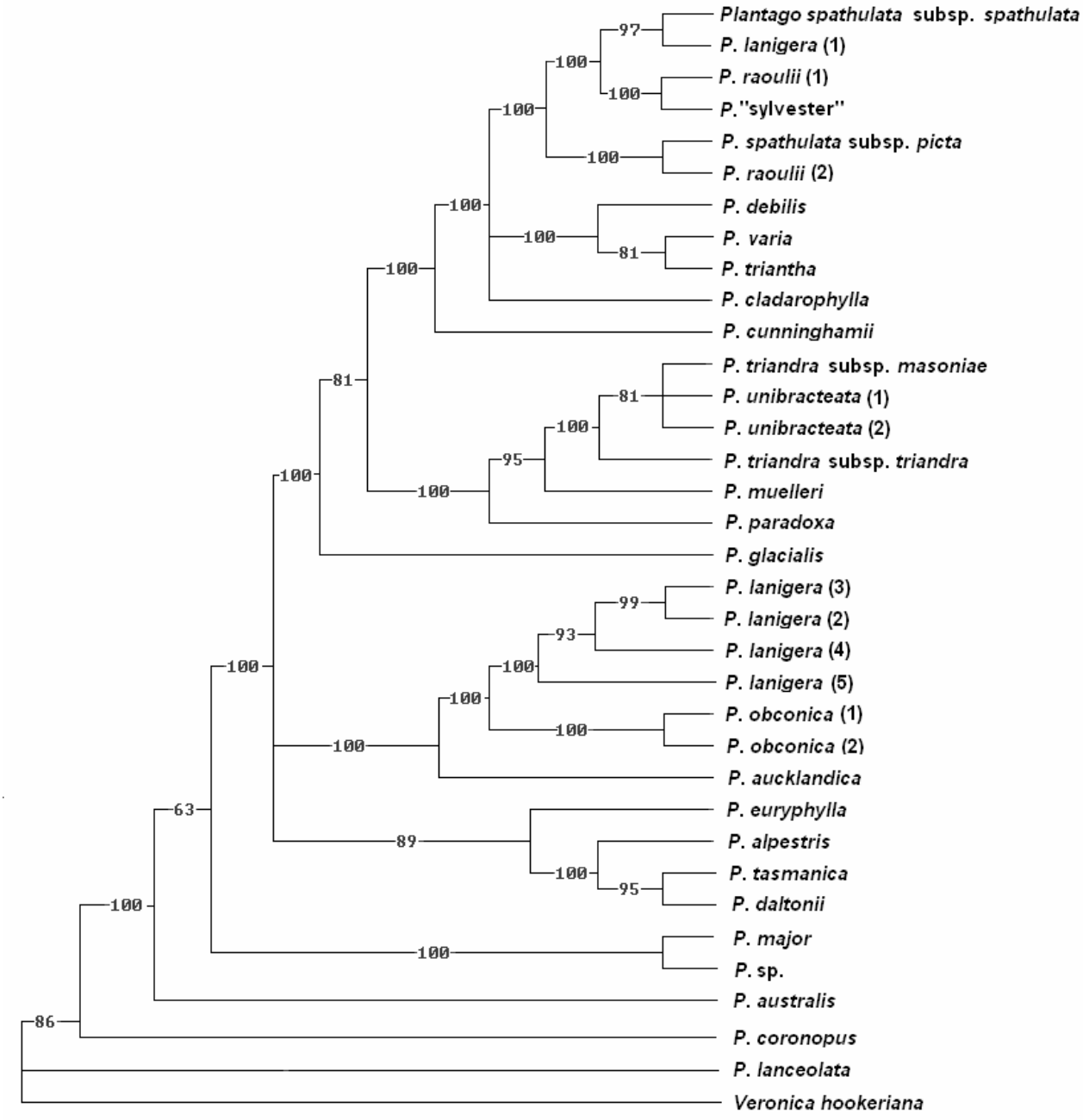
I. 50% bootstrap consensus phylogeny of the *cox1* dataset recovered using neighbour-joining methods (using samples listed in Table 3.1).



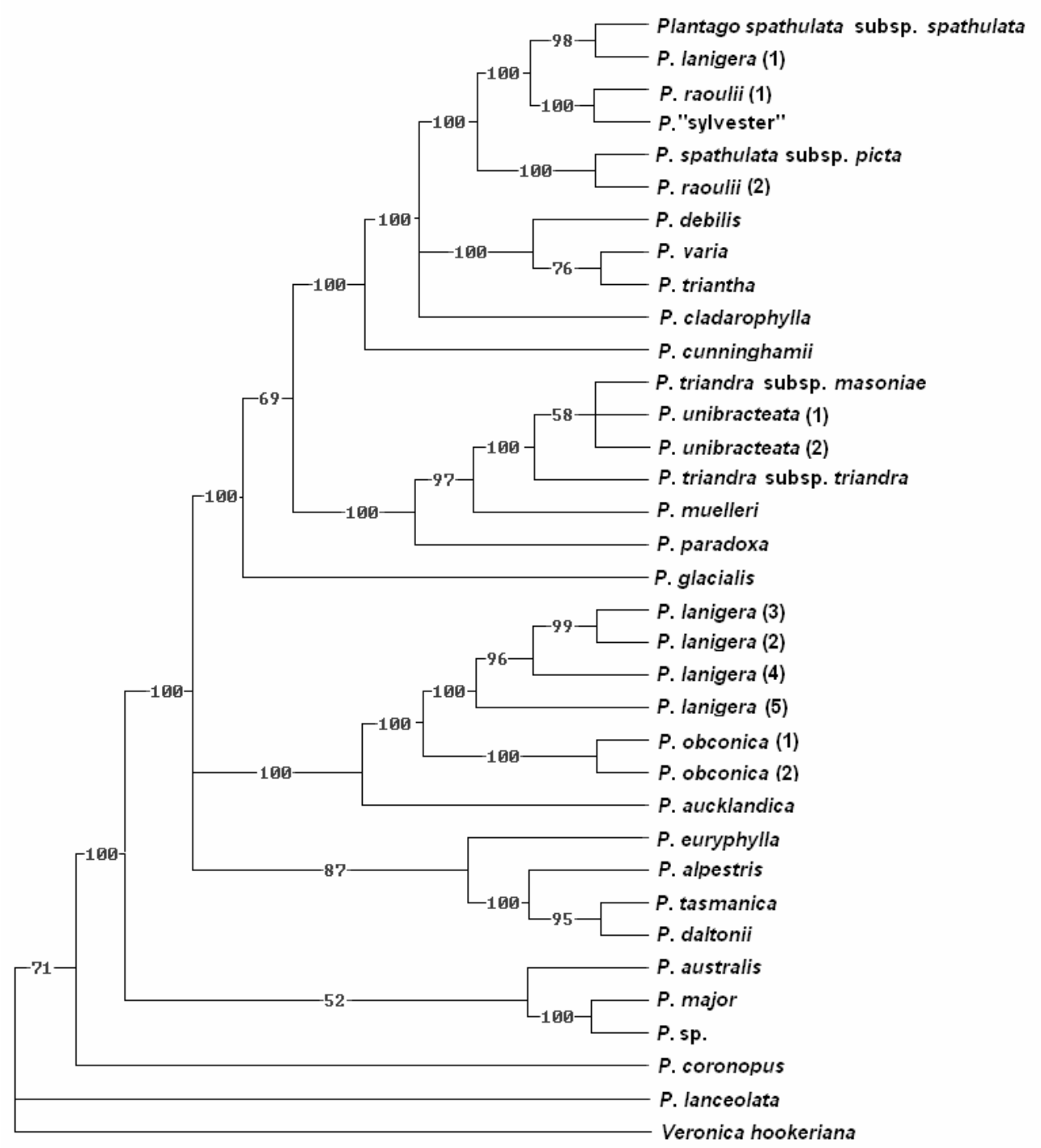
J. 50% bootstrap consensus phylogeny of the concatenated (ITS + *ndhF-rpl32* + *cox1*) dataset recovered using Maximum Parsimony (using samples listed in Table 3.1). Tree scores range from 1056-1058.



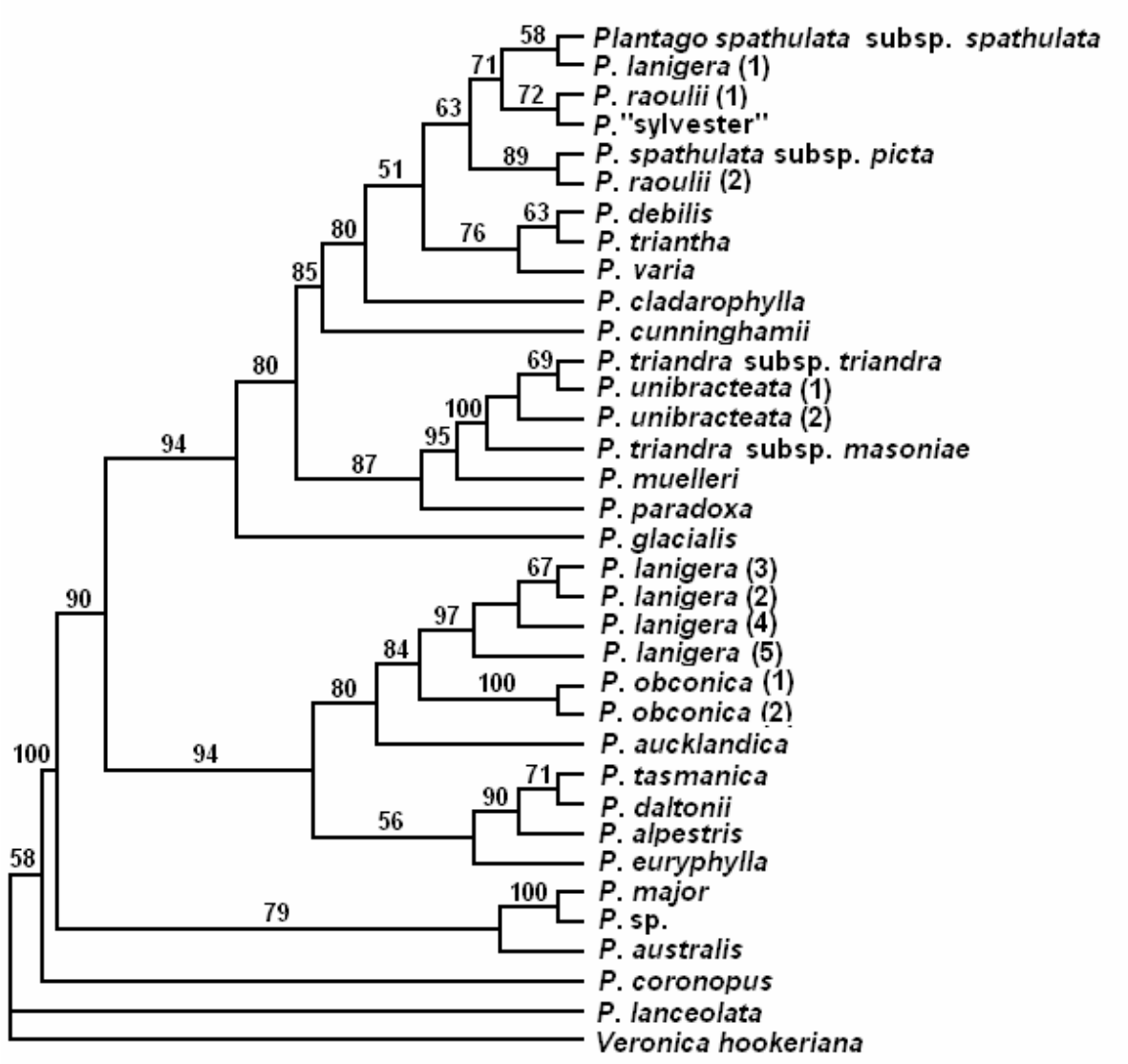
K. 50% posterior probability consensus phylogeny of the concatenated (ITS + *ndhF-rpl32* + *cox1*) dataset recovered using Bayesian analysis with one model for the dataset (using samples listed in Table 3.1). Estimated likelihood score = -7775.85.



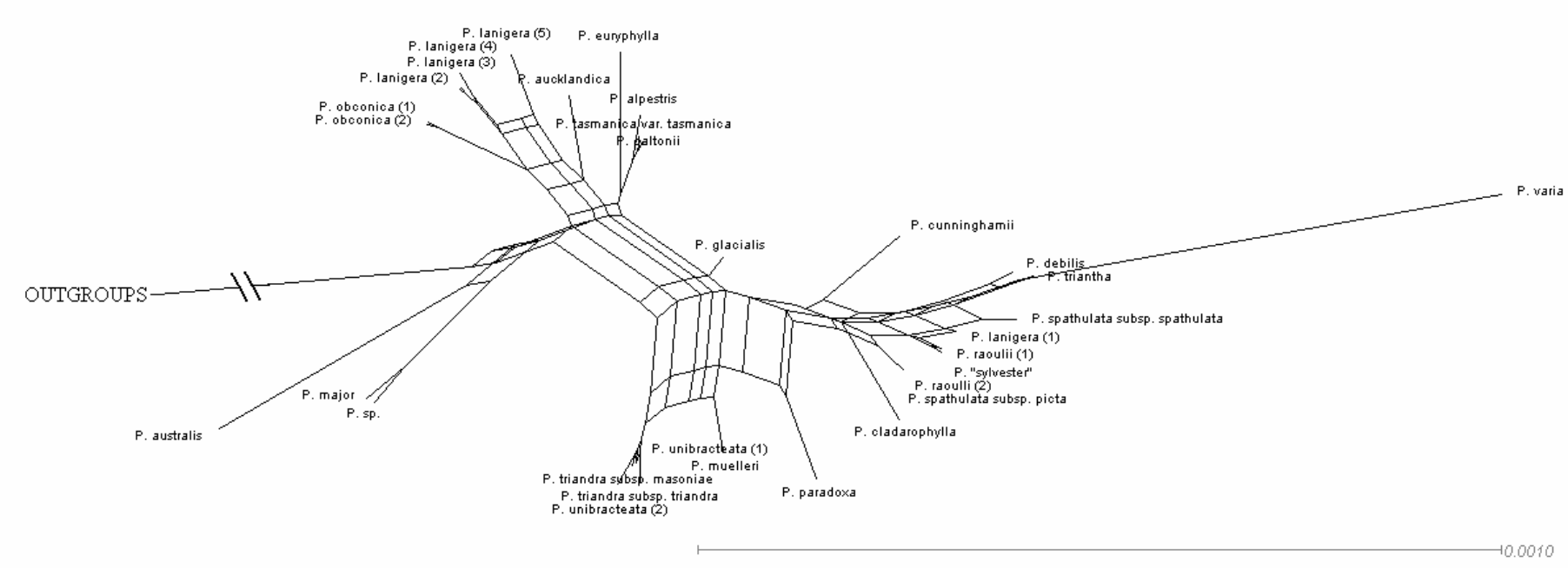
L. 50% posterior probability consensus phylogeny of the concatenated (ITS + *ndhF-rpl32* + *cox1*) dataset recovered using Bayesian analysis when the dataset was partitioned and different models were implemented for each region (using samples listed in Table 3.1). Estimated likelihood score = -7874.62.



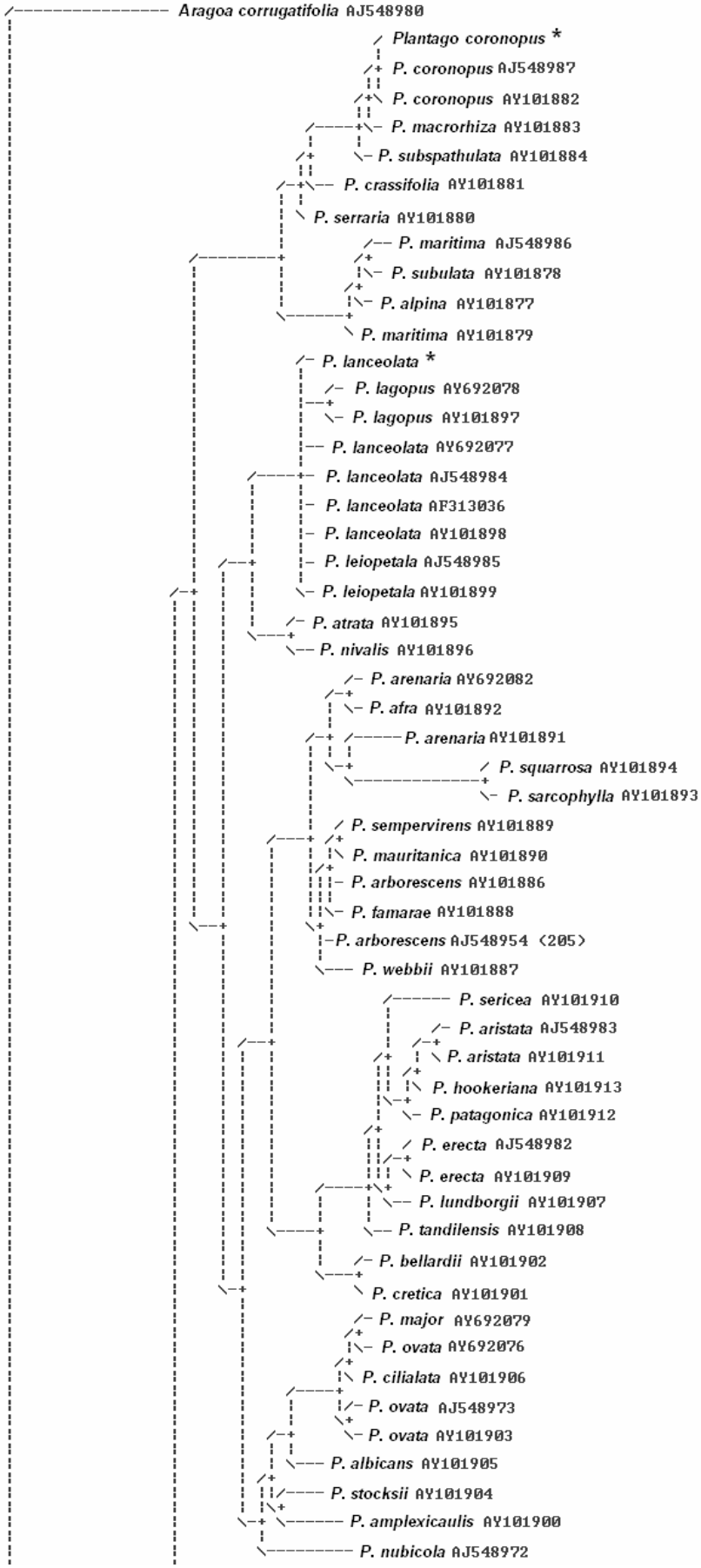
M. 50% bootstrap consensus phylogeny of the concatenated (ITS + *ndhF-rpl32* + *cox1*) dataset recovered using neighbour-joining methods (using samples listed in Table 3.1).



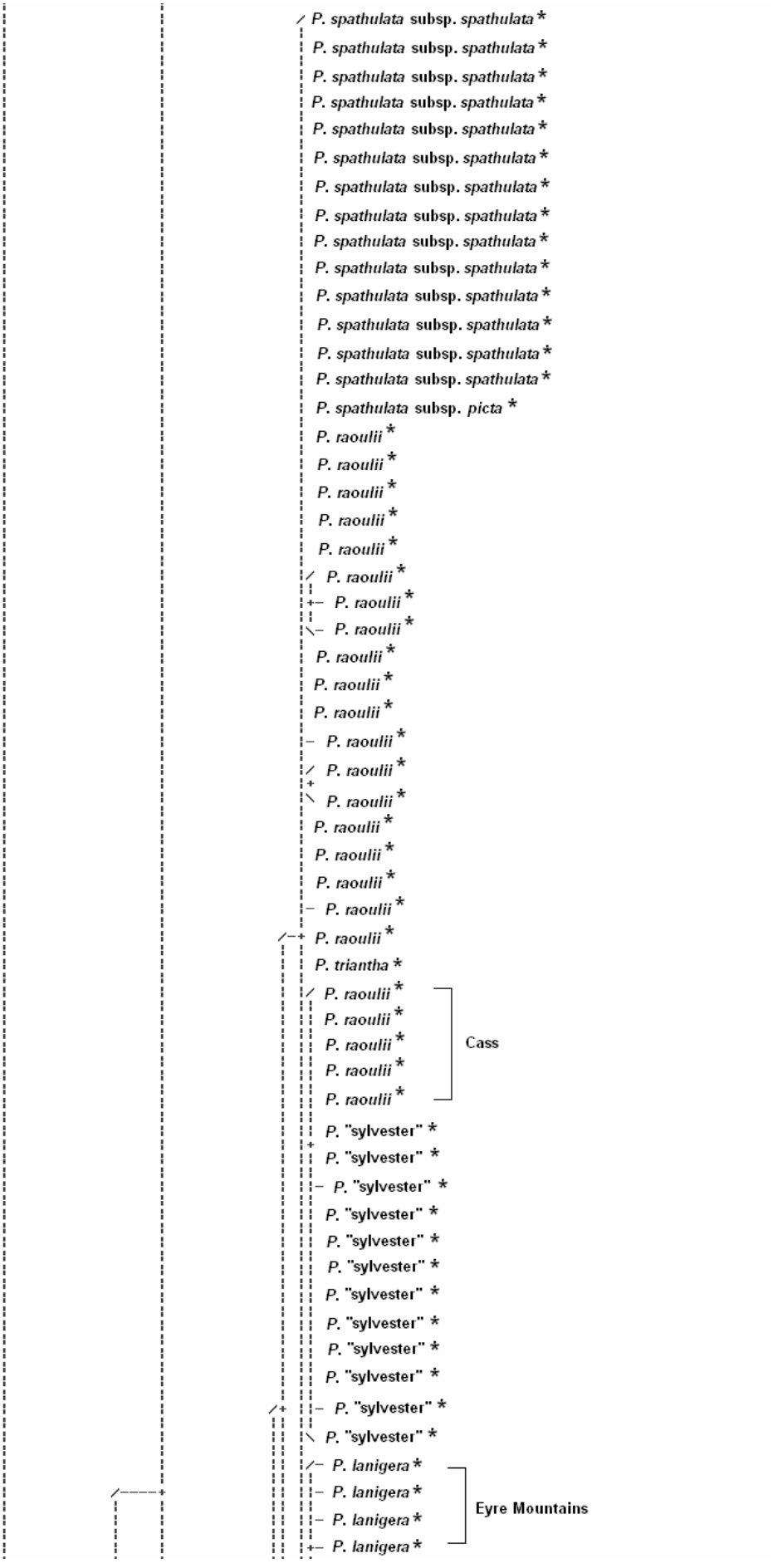
N. Neighbour-net analyses of the concatenated regions ITS, *ndhF-rpl32* and *cox1* of the Australasian species using Splitstree v4.8 (using samples listed in Table 3.1).



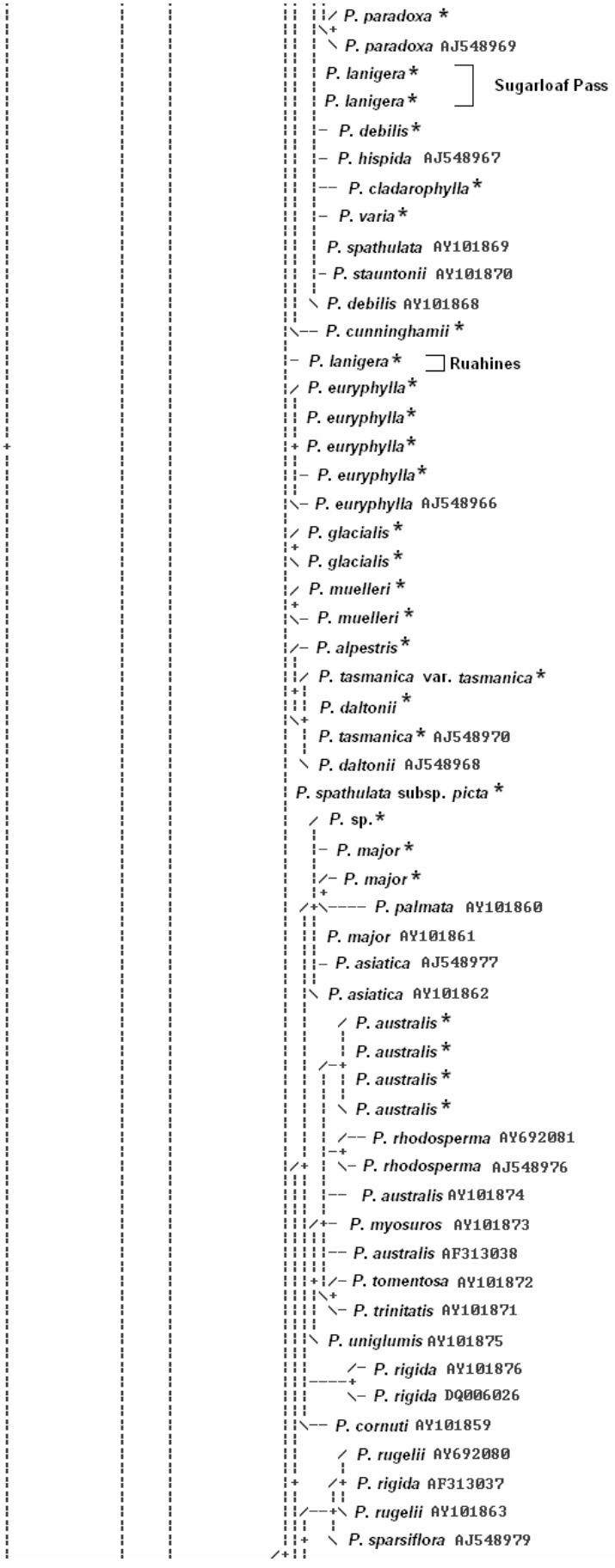
O. 50% posterior probability consensus ITS phylogeny of a worldwide *Plantago* dataset reconstructed using Bayesian analysis (samples listed in Table 4.1). Estimated likelihood score = -7531.24.



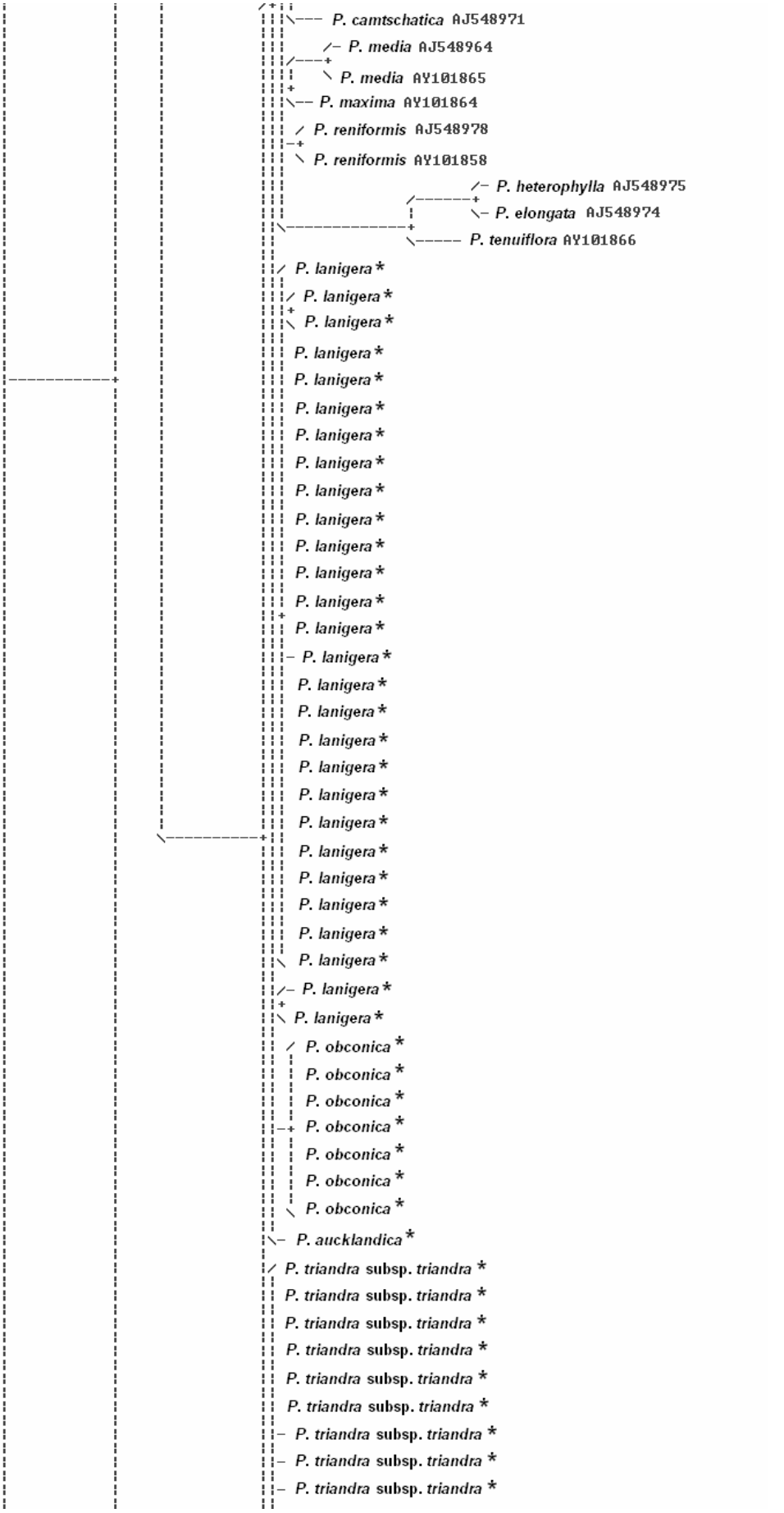
O. **(continued)** 50% posterior probability consensus ITS phylogeny of a worldwide *Plantago* dataset reconstructed using Bayesian analysis.



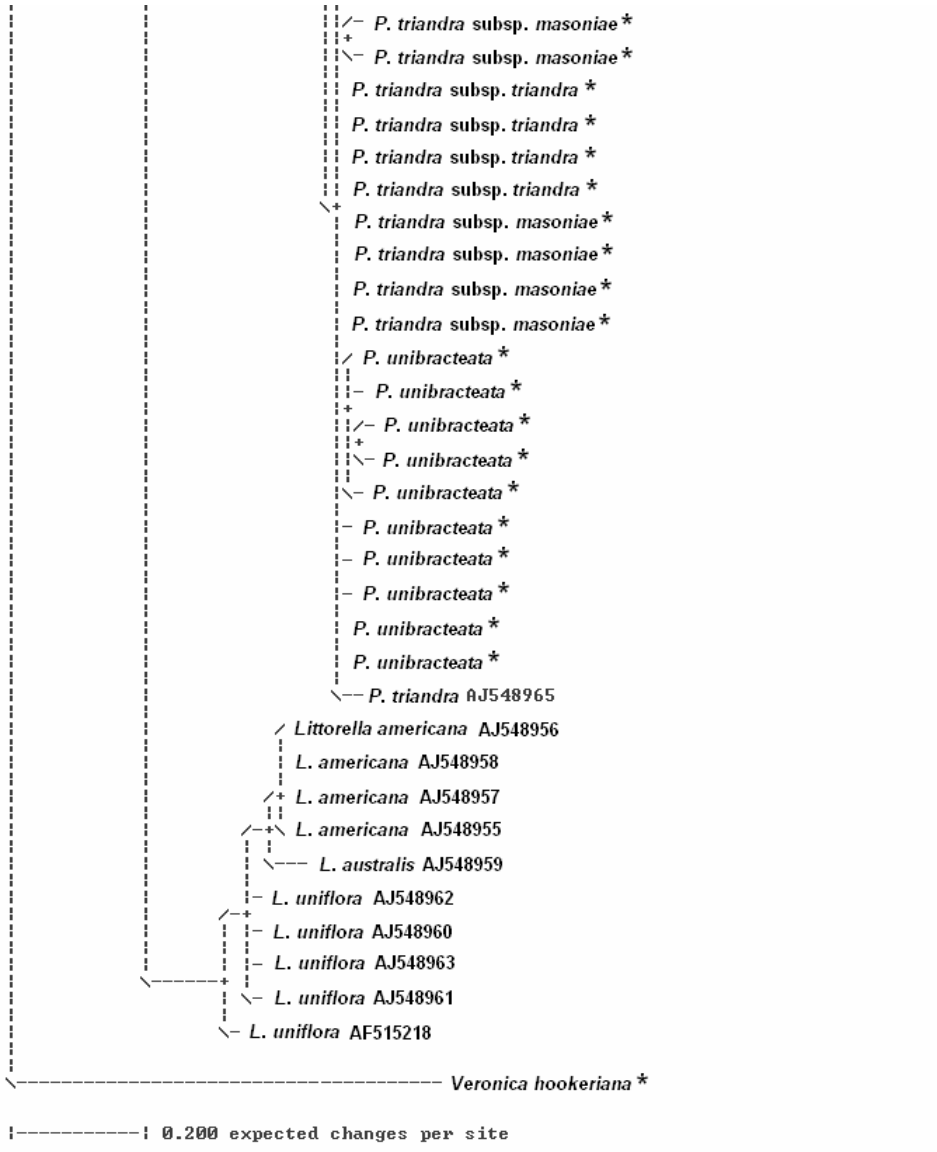
O. **(continued)** 50% posterior probability consensus ITS phylogeny of a worldwide *Plantago* dataset reconstructed using Bayesian analysis.



O. **(continued)** 50% posterior probability consensus ITS phylogeny of a worldwide *Plantago* dataset reconstructed using Bayesian analysis.



O. **(continued)** 50% posterior probability consensus ITS phylogeny of a worldwide *Plantago* dataset reconstructed using Bayesian analysis.



Appendix III. Data statistics

A. Data statistics for the three Australasian groups resolved in a worldwide ITS phylogeny (Fig. 4.2). Sample details can be found in Table 4.1.

	Clade II	Grade B	Clade IV
Base frequencies (all sites)	T- 23.1 C- 26.3 A- 22.1 G- 28.5	T- 23.0 C- 26.0 A- 22.3 G- 28.6	T- 23.2 C- 26.0 A- 22.4 G- 28.5
Base frequencies (variable sites)	T- 29.4 C- 27.9 A- 22.9 G- 19.8	T- 22.4 C- 48.0 A- 11.0 G- 18.6	T- 22.3 C- 41.3 A- 9.3 G- 27.1
Variable sites and parsimony-informative sites (%)	37(3.35%)	31(1.98%)	16(0.61%)
Modeltest model	GTR + I	TIMef	K81 + I
Pinvar	0.8224	0	0.9468
Number of transitions (all sites)	3	3	1
Number of transversions (all sites)	3	2	2
Transition/transversion ratio (all sites)	1	1.5	0.5
p-value of $\chi 2$ of base frequencies overall (variable sites)	0.7530 (0.5399)	0.7530 (0.0001)	0.7530 (0.0002)

B. Uncorrected pairwise distances of ITS sequences among Australasian species (sample details can be found in Table 4.1).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
2	0.7	—																				
3	0.8	0. 5	—																			
4	2.9	1.5	2.7	—																		
5	1.3	0. 8	1.8	2.5	—																	
6	1.8	0. 2	1.3	1.8	1.3	—																
7	1.8	0. 5	1.5	2.3	1.3	0.8	—															
8	0.5	0.3	1.0	2.0	0.5	0.8	0.8	—														
9	3.0	1.7	2.9	1.2	2.7	2.0	2.3	2.2	—													
10	2.9	1.7	2.7	1.7	2.9	2.2	2.7	2.2	1.5	—												
11	3.2	1.9	3.0	0. 7	2.9	2.2	2.7	2.4	1.3	1.8	—											
12	3.0	1.7	2.8	1.5	2.7	2.0	2.5	2.2	1.3	1.8	1.7	—										
13	2.3	1.0	2.2	1.5	2.0	1.3	1.8	1.5	1.3	1.8	1.7	0.8	—									
14	2.9	1.2	2.3	2.7	2.3	1.3	1.8	1.8	2.9	3.0	2.7	2.9	2.2	—								
15	2.2	1.0	2.0	2.3	2.2	1.5	2.0	1.5	2.2	2.3	2.5	1.8	1.2	2.3	—							
16	1.5	1.2	1.0	3.4	2.3	2.0	2.2	1.7	3.6	3.4	3.7	3.5	2.9	2.9	2.7	—						
17	2.3	1.0	2.2	1.5	2.0	1.3	1.8	1.5	1.3	1.8	1.7	1.0	0.3	2.2	1.2	0.029	—					
18	2.5	1.2	2.3	1.7	2.2	1.5	2.0	1.7	1.5	2.0	1.8	1.2	0.5	2.3	1.3	3.0	0.5	—				
19	2.7	1.3	2.5	1.8	2.3	1.7	2.2	1.8	1.7	2.2	2.0	1.3	0.7	2.2	1.5	3.2	0.7	0.8	—			
20	2.5	1.2	2.3	1.7	2.2	1.5	2.0	1.7	1.5	2.0	1.8	1.2	0.5	2.2	1.3	3.0	0.5	0.7	0.5	—		
21	2.5	1.2	2.0	2.2	2.0	1.3	0.8	1.5	2..5	3.2	2.9	2.7	2.0	2.4	2.5	2.7	2.0	2.2	2.3	2.0	—	
22	2.5	1.2	2.3	1.7	2.2	1.5	2.0	1.7	1.5	2.0	1.8	1.2	0.5	2.2	1.3	3.0	0.5	0.7	0.5	0.0	2.0	—
23	1.0	0.3	0. 5	2.5	2.0	1.2	1.7	1.2	2.7	2.5	2.9	2.7	2.0	2.2	1.8	1.2	2.0	2.2	2.3	2.2	2.2	2.2

1 = *P. spathulata* subsp. *spathulata*, 2 = *P. spathulata* subsp. *picta*, 3 = *P. triantha*, 4 = *P. triandra subspp*, 5 = *P. "sylvester"*, 6 = *P. raoulii*, 7 = *P. lanigera* (Eyre Mountains), 8 = *P. lanigera* (Sugarloaf Pass), 9 = *P. lanigera*, 10 = *P. obconica*, 11 = *P. unibracteata*, 12 = *P. aucklandica*, 13 = *P. euryphylla*, 14 = *P. cladarophylla*, 15 = *P. cunninghamii*, 16 = *P. varia*, 17 = *P. glacialis*, 18 = *P. muelleri*, 19 = *P. alpestris*, 20 = *P. tasmanica*, 21 = *P. paradoxa*, 22 = *P. daltonii*, 23 = *P. debilis & P. hispida*. Highlighted boxes indicate relatively high p-distances found in the dataset (≥3.0%). Species with multiple shaded values are *P. varia* and *P. spathulata* subsp. *picta* (representatives of each species are used).