

# **The Evolution of *Wahlenbergia* (Campanulaceae) in Australasia**

By

Jessica Mary Prebble

A thesis submitted to the Victoria University of  
Wellington in fulfilment of the requirement for the  
degree of Masters in Ecology and Biodiversity

**Te Whare Wananga o te Upoko o te Ika a Maui**

April 2010





# Abstract

*Wahlenbergia* is a large genus of flowering plants within the family Campanulaceae. In this thesis the first molecular phylogeny of *Wahlenbergia* was reconstructed from approximately 20% of the genus, based on the nuclear ribosomal ITS (nrITS) DNA marker and the chloroplast *trnL-F* DNA marker, with samples from South Africa, Europe, Australia and New Zealand. Additionally a large phylogeny with increased within-species sampling focusing on addressing taxonomic questions among the 45 Australasian species of *Wahlenbergia* was also reconstructed based on nrITS and *trnL-F*, plus an additional chloroplast DNA marker, *trnK*. Relationships and species limits of the New Zealand species of *Wahlenbergia* were further analysed using amplified fragment length polymorphisms (AFLPs).

*Wahlenbergia* was found to be polyphyletic, though most of the species form a clade. Tree topologies and molecular dating analysis showed that the genus originated in South Africa about 16.2 million years ago (mya), then dispersed to Australasia before radiating there about 3.7 mya, thus refuting the hypothesis of Gondwanan vicariance for the Australasian species. Two dispersals from Australia to New Zealand are hypothesised, one leading to a radiation of species with the rhizomatous growth from about 1.0 mya and the other leading to a radiation of species with the radicate growth from 0.49 mya, although the radicate species might not form a clade. Low levels of genetic variation among individuals from Australia and New Zealand was revealed with all markers, and the phylogenies were poorly resolved as a result. The low genetic diversity is probably due to rapid and recent evolution during a period of geological and climatic change, coupled with incomplete lineage sorting and hybridisation. Phylogenies reconstructed using AFLPs were also poorly resolved, although AFLPs were found to be useful for species delimitation, as has been shown in studies of other plant groups.

Despite the poor resolution, several morphological species and subspecies were recovered as monophyletic with DNA sequence data, notably the morphologically distinctive New Zealand *W. cartilaginea*, *W. matthewsii* and *W. congesta* subsp.

*congesta*. Further research into species boundaries within the *W. albomarginata*/*W. pygmaea* complex is needed. Members of the New Zealand lowland radicate *W. gracilis* complex may all belong to the same morphologically variable species, although further research is needed to justify such a taxonomic change. The other New Zealand radicate species, *W. vernicosa*, is probably a separately evolving lineage, and is not conspecific with the Australian *W. littoricola*.

# Acknowledgements

I would like to express my sincere thanks to the many people who have helped me prepare this thesis. I would especially like to thank my supervisors, Heidi Meudt and Phil Garnock-Jones, for your generous support, guidance and endless re-reading of drafts. Chris Cupido kindly shared his DNA sequences of South African *Wahlenbergia*, as well as his knowledge of the genus.

Thank you to Peter de Lange, Ewen Cameron, Mike Bayly, David Lyttle, Graeme Atkins, Hugh Wilson, Jeremy Bruhl, Benjamin Magaña-Rodríguez, Hamish Carson, Tony Silbery, Phil Garnock-Jones and Heidi Meudt for helping with specimen collection, and to all my field assistants: Gesine Pufal, Teresa Herleth, Mark Prebble, Lesley Bagnall, Fiona Hodge and Meg Prebble. Graham Petterson generously shared Judith Petterson's notes and photographs of *Wahlenbergia*, as well as some plants. Thanks also to helpful DoC staff all around the country. Barbara Briggs and others at the Sydney Botanical Gardens thoughtfully organised a NSW collection permit for me, and Mike Bayly and Alison Kellow provided accommodation in Melbourne.

I am thankful to the Institute of Natural Resources and the Allan Wilson Centre at Massey University for use of the Ancient DNA Facility, and especially Dr Steve Trewick for advice. Barbara Holland gave invaluable help running python scripts for the AFLP chapter. A big thanks to my proof-readers and formatters, Travis Lealand-Mapelsden, Lesley Bagnall and Aaron Packard.

I am grateful to the Auckland Botanical Society for granting me their Lucy Cranwell field grant, to Te Papa for their Te Papa MSc scholarship in molecular systematics and to Victoria University for the VUW post-grad scholarship. Everybody at the Te Papa herbarium has also been especially supportive.

Finally, thank you to my lovely family and friends.

# Table of Contents

<b>ABSTRACT</b>	<b>III</b>
<b>ACKNOWLEDGEMENTS</b>	<b>V</b>
<b>CONTENTS</b>	<b>VI</b>
<b>CHAPTER ONE: GENERAL INTRODUCTION</b>	<b>1</b>
<b>The New Zealand flora: origins, biogeography and evolution</b>	<b>1</b>
<b>Molecular phylogenetics</b>	<b>3</b>
Choosing DNA markers for molecular phylogenetic study	5
Organellar DNA (chloroplast and mitochondrial)	5
Nuclear ribosomal DNA	6
Low-copy nuclear markers	7
AFLPs	8
Molecular clocks	9
<b>Introduction to study group</b>	<b>10</b>
Brief taxonomic history of <i>Wahlenbergia</i> in New Zealand	13
New Zealand radicate species	15
New Zealand rhizomatous species	17
<b>A note on species concepts</b>	<b>17</b>
<b>Study aims</b>	<b>18</b>
<b>CHAPTER 2: BIOGEOGRAPHY AND A FIRST PHYLOGENY OF WAHLENBERGIA (CAMPANULACEAE)</b>	<b>19</b>
<b>Abstract</b>	<b>19</b>
<b>Introduction</b>	<b>19</b>
<b>Materials and Methods</b>	<b>22</b>
Genetic markers	22
Study group	23
DNA extraction, amplification and sequencing	24
Dataset alignment and analysis	25
Phylogenetic and dating analysis	25
Biogeography and character mapping	27
<b>Results</b>	<b>28</b>
Nuclear (ITS) phylogeny	28
Chloroplast ( <i>trnL-F</i> ) phylogeny	28
Analysis of concatenated dataset	30
Combined dataset	32
Molecular dating analysis	32

Biogeography and character mapping	34
<b>Discussion</b>	<b>35</b>
<b>Summary and Conclusions</b>	<b>39</b>

## **CHAPTER THREE: PHYLOGENETIC ANALYSIS OF THREE DNA SEQUENCING MARKERS TO EXPLORE RELATIONSHIPS WITHIN AUSTRALASIAN *WAHLENBERGIA***

<b>Abstract</b>	<b>41</b>
<b>Introduction</b>	<b>42</b>
Research aims	43
Taxonomic introduction	44
New Zealand radicate species	44
New Zealand rhizomatous species	45
Australian species	48
<b>Methods</b>	<b>49</b>
Taxon sampling	49
DNA isolation, amplification and sequencing	49
Sequence alignment and phylogenetic analysis	51
Biogeographic and character analysis	53
<b>Results</b>	<b>53</b>
DNA extractions and primer selection	53
aDNA extractions	54
ITS analysis	59
Chloroplast markers analysis	62
Analysis of concatenated dataset	66
Combined <i>ITS</i> , <i>trnK</i> and <i>trnL-F</i> analysis	67
<b>Discussion</b>	<b>69</b>
Evolution of Australasian <i>Wahlenbergia</i>	69
Testing the New Zealand <i>Wahlenbergia</i> taxonomy	71
New Zealand radicate species	71
The position of <i>W. littoricola</i> subsp. <i>vernicosa</i>	73
New Zealand rhizomatous species	74
Notes on the Australian <i>Wahlenbergia</i> taxonomy	76
Australian radicate species	76
Australian rhizomatous species	78
Newly described species	78
<b>Summary and conclusions</b>	<b>79</b>

## **CHAPTER FOUR: USING AFLPS TO RESOLVE PHYLOGENETIC RELATIONSHIPS IN A PLANT SPECIES COMPLEX WHEN NUCLEAR AND CHLOROPLAST GENES FAIL TO REVEAL SUFFICIENT VARIABILITY<sup>81</sup>**

<b>Abstract</b>	<b>81</b>
<b>Introduction</b>	<b>81</b>

<b>Methods</b>	<b>84</b>
Sampling and DNA extraction	84
Generation of AFLP data	85
Parameter optimisation and automated scoring of AFLP data	86
Phylogenetic analysis	88
Principal coordinate analysis (PCoA) of dominant data	89
<b>Results</b>	<b>89</b>
Selection of optimal AFLP automated scoring parameters	89
Phylogeny of New Zealand <i>Wahlenbergia</i> based on AFLP data	94
Cluster analysis using principal coordinates analysis	101
<b>Discussion</b>	<b>98</b>
Phylogenetic patterns and species delimitation in New Zealand <i>Wahlenbergia</i> based on AFLP	101
Radicale species	101
Rhizomatous species	103
Trans-Tasman relationships	104
<b>Summary and conclusions</b>	<b>104</b>
 <b>CHAPTER FIVE: GENERAL DISCUSSION</b>	 <b>106</b>
<b>Large scale phylogeny and biogeography</b>	<b>106</b>
<b>Medium and small scale: Relationships between species and taxonomy</b>	<b>107</b>
<b>Assessment of molecular techniques used in this thesis</b>	<b>113</b>
<b>Future directions</b>	<b>113</b>
<b>Conclusions</b>	<b>115</b>
 <b>REFERENCES</b>	 <b>116</b>
 <b>APPENDIX</b>	 <b>136</b>
<b>Appendix 1</b> Table showing details of individuals used in each analysis	<b>136</b>
<b>Appendix 2.1</b> ITS neighbour net network	<b>144</b>
<b>Appendix 2.2</b> <i>trnL-F+trnK</i> neighbour net network	<b>145</b>
<b>Appendix 2.3</b> Combined ITS+ <i>trnL-F+trnK</i> neighbour net network	<b>146</b>

## LIST OF TABLES

<b>Table 1.1</b> Taxonomic treatments of New Zealand <i>Wahlenbergia</i>	<b>14</b>
<b>Table 3.1</b> Primer sequences and references for 22 primer pairs trialled in this study	<b>57</b>
<b>Table 3.2</b> Uncorrected p distances calculated from the ITS and <i>trnL-F</i> + <i>trnK</i> Australasian <i>Wahlenbergia</i> datasets	<b>66</b>
<b>Table 4.1</b> Dataset statistics for the 36 AFLP datasets generated in GeneMarker including and excluding 6FAM characters	<b>91</b>
<b>Table 5.1</b> Synthesis table regarding species delimitation	<b>109</b>

## LIST OF FIGURES

<b>Figure 1.1</b> Growth forms of New Zealand <i>Wahlenbergia</i>	<b>12</b>
<b>Figure 1.2</b> Photos of New Zealand <i>Wahlenbergia</i>	<b>16</b>
<b>Figure 2.1</b> Bayesian 50% majority rule consensus phylogeny of <i>Wahlenbergia</i> samples from South Africa, Europe, Australia and New Zealand reconstructed using ITS	<b>29</b>
<b>Figure 2.2</b> Bayesian 50% majority rule consensus phylogeny of <i>Wahlenbergia</i> samples from South Africa, Europe, Australia and New Zealand reconstructed using <i>trnL-F</i>	<b>31</b>
<b>Figure 2.3</b> Bayesian maximum clade credibility tree of <i>Wahlenbergia</i> samples from South Africa, Europe, Australia and New Zealand reconstructed using a concatenated ITS + <i>trnL-F</i> dataset, showing divergence time estimates and with growth forms and locations mapped	<b>33</b>
<b>Figure 3.1</b> Photographs of gels showing chloroplast markers trialled	<b>55</b>
<b>Figure 3.2</b> Photographs of gels showing nuclear and mitochondrial markers trialled	<b>56</b>
<b>Figure 3.3</b> Bayesian 50% majority rule consensus phylogeny of <i>Wahlenbergia</i> samples from Australasia reconstructed using ITS, with MP and ML bootstrap values added	<b>60</b>
<b>Figure 3.4</b> Bayesian 50% majority rule consensus phylogeny of <i>Wahlenbergia</i> samples from Australasia reconstructed using <i>trnL-F</i> , with MP and ML bootstrap values added	<b>63</b>
<b>Figure 3.5</b> Bayesian 50% majority rule consensus phylogeny of <i>Wahlenbergia</i> samples from Australasia reconstructed using a combined ITS + <i>trnL-F</i> + <i>trnK</i> dataset, with MP and ML bootstrap values added	<b>68</b>
<b>Figure 4.1</b> Consensus network of the 36 AFLP datasets generated in GeneMarker excluding 6FAM characters	<b>93</b>
<b>Figure 4.2</b> One of the five most parsimonious trees reconstructed from Australasian <i>Wahlenbergia</i> AFLP data with bootstrap values greater than 50% added	<b>95</b>
<b>Figure 4.3</b> Bayesian 50% majority rule consensus phylogeny of Australasian <i>Wahlenbergia</i> AFLP data showing posterior probability values	<b>96</b>

**Figure 4.4** Principle coordinates analysis of Australasian *Wahlenbergia* AFLP data showing three dimensions

99

**Figure 4.5** Principle coordinates analysis of a reduced Australasian *Wahlenbergia* AFLP dataset showing the *W. albomarginata*/*W. pygmaea* complex in three dimensions

100



*Wahlenbergia albomarginata* subsp. *albomarginata* growing on the Waimakariri riverbed, near Arthur's Pass



# Chapter One: General Introduction

---

## The New Zealand flora: origins, biogeography and evolution

The New Zealand flora has its roots on the ancient supercontinent Gondwanaland, yet New Zealand has been increasingly isolated over the last 80 million years (my), and has been separated from its nearest large neighbour, Australia, by 1500km for at least 60 my (Kamp, 1986; Wallis and Trewick, 2009; Winkworth et al., 2005). This long isolation has fostered debate regarding the origins of the flora, a debate that became unnecessarily polarised for a time (Wallis and Trewick, 2009) between supporters of vicariance (Wardle, 1963), and long distance dispersal (Pole, 1994). It is generally accepted that a mixture of the two modes has affected the formation of the New Zealand flora as we know it, though Pole (1994) and Macphail (1997) suggested that all of the flora must have arrived in New Zealand by long distance dispersal, as they contend that all of Zealandia (New Zealand plus its large continental shelf) would have been under water during the Oligocene period (the “Oligocene drowning”). There is indeed molecular phylogenetic evidence that a large proportion of the flora has arrived by long distance dispersal after the separation of Zealandia from Gondwanaland (Winkworth et al., 1999), but nevertheless this hypothesis disregards fossil and molecular evidence that suggests some plant lineages e.g. kauri (*Agathis*) do have a Gondwanan history (Knapp et al., 2007).

Many of the New Zealand plant groups that have a long-distance dispersal event in their history have arrived from Australia (e.g. (Ford et al., 2007; Wagstaff et al., 1999)). This west to east direction of dispersal is consistent with the forces of the west wind drift, which predicts more dispersal events from west to east because of westerly winds and ocean currents (Raven, 1973; Sanmartín et al., 2007; Winkworth et al., 2002b). These New Zealand plant groups are often found to have been

established by a single dispersal event to New Zealand (implicated if the New Zealand representatives form a monophyletic group) followed by rapid species radiation e.g. (Ford et al., 2007). There is often little DNA sequence divergence between species from Australia and species from New Zealand e.g. *Sophora* (Hurr et al., 1999), *Veronica* (Wagstaff et al., 2002) and others (Ford et al., 2007; Sanmartín et al., 2007). Other locations have also been implicated as being the origins of long distance dispersal events to New Zealand, e.g. *Fuchsia* (Berry et al., 2004), *Gaultheria* (Bush et al., 2009) and the tribe Coronanthereae (Gesneriaceae) (Woo, 2007) from South America. However, these events are rarer than introductions from Australia and likely involve Antarctica as a stepping stone while it was still partly ice-free and connected to Australia. There are also examples of multiple introductions from Australia e.g. *Drosera* (Rivadavia et al., 2003) and *Plantago* (Tay et al., 2010).

There is molecular phylogenetic evidence from multiple studies that the ancestors of many of these plant groups radiated rapidly once they established in New Zealand (Winkworth et al., 1999) e.g. *Veronica* (Wagstaff and Garnock-Jones, 1998, 2000), *Pachycladon* (Heenan et al., 2002), *Polystichum* (Perrie et al., 2003), *Pratia* (Murray et al., 2004) and *Ourisia* (Meudt et al., 2009). These radiations are often associated with speciation after geologically recent long distance dispersal events; for example molecular clock analyses of a number of genera suggest that their contemporary diversity arose within the last 5 million years e.g. *Plantago* (Tay et al., 2010), *Ourisia* (Meudt et al., 2009), *Myosotis* (Winkworth et al., 2002a) and *Ranunculus* (Lockhart et al., 2001). The past 5 million years was a period of rapid geological uplift and numerous climatic oscillations in New Zealand, leading to many different situations that could influence speciation such as adaptation to new habitats after the uplift of the Southern Alps (Haase et al., 2007; Lockhart et al., 2001; Trewick and Morgan-Richards, 2005; Wagstaff and Garnock-Jones, 1998), and climate fluctuations (Raven, 1973; Winkworth et al., 2002b). Mountain building has obvious implications for species formation, in terms of both creating new habitats and creating barriers to gene flow, consequently promoting the evolution of local variants. Frequent climatic fluctuations are also hypothesised to affect the radiations of plant lineages in that climate change produces new habitats that allow range expansion and provide opportunities for the differentiation of local morphological or

ecological variants. Ongoing climate change can also lead to range contraction, which may result in the formation of additional local variants either by local differentiation or perhaps hybridisation (Winkworth et al., 2005).

Several cycles of this pattern (such as repeated glacial cycles) might be expected to produce many variants, especially if migration paths result in hybridisation or introgression between previously isolated forms (Winkworth et al., 2005). This model of species diversification would be expected to result in considerable mixing of genotypes, and therefore complex patterns of phylogenetic relationships. Recent molecular phylogenetic studies on New Zealand plants support this suggestion (Winkworth et al., 2005). Several studies have found complex patterns of relationships, using chloroplast markers for example (e.g. *Veronica* (as *Chionohebe*) (Meudt and Bayly, 2008) and *Leucogenes* (Smitsen and Breitwieser, 2008). Additionally, species with conspicuous morphological and ecological differences exhibit little or no genetic differentiation with both nuclear and chloroplast loci, suggesting that morphological differentiation has occurred over a relatively short period of time. Winkworth *et al.* (1999) report that in the New Zealand flora a small amount of genetic change in neutral and cpDNA markers can underlie dramatic morphological differences in recently evolved plant species. The southern hemisphere members of *Myosotis* have very little genetic diversity (measured with the *matK* and ITS loci) compared to northern hemisphere representatives, yet they displayed greater morphological variation. Speciation in New Zealand plant groups may also have followed displacement along the alpine fault (Haase et al., 2007). Cook Strait, which separates the North and South Islands of New Zealand, currently acts 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).

## **Molecular phylogenetics**

Molecular phylogenetics has been increasingly used as a means of clarifying evolutionary histories and refining taxonomy when traditional methods, such as morphological data and fossil records, are ambiguous. The true evolutionary history of a given set of sequences is rarely known, and so phylogenetic trees represent

estimates of ‘true trees’ (Nei and Kumar, 2000). Phylogenies treat speciation as a dichotomously branching event, an assumption that is frequently violated especially in recently evolving lineages where hybridisation and incomplete lineage sorting can confuse the picture (Woolley et al., 2008).

Hybridisation, as either homoploid hybrid speciation or allopolyploidy, has contributed extensively to angiosperm diversity. Whereas allopatric speciation via vicariance or long distance dispersal and subsequent establishment may be responsible for most speciation events, the evidence for hybridisation is nevertheless widespread (Soltis and Soltis, 2009). Allopolyploidy (hybridisation followed by genome duplication) allows instant speciation when one or several polyploid hybrids arise sympatrically with one (or both) of their parent populations (Petit et al., 1999). Allopolyploidy also leads to rapid diversification due to ‘genomic shock’ that can lead to massive genome restructuring. Because of genetic and genomic changes, individuals may arise with a modified phenotype and ecological preferences and hence are able to exploit new niches or to out compete progenitor species (Soltis and Soltis, 2009). Hybridisation and polyploidy are common in many New Zealand plant groups (Lockhart et al., 2001; Morgan-Richards et al., 2009; Wagstaff et al., 2010). Hybridisation can be a confounding factor in phylogenetic reconstruction, leading to reticulation, which cannot be displayed on a dichotomously branching tree (Woolley et al., 2008). Alternative methods such as networks can prove useful in these instances. Reconstructing phylogenetic trees from independent sources such as chloroplast and nuclear DNA is also a useful method for exploring whether or not hybridisation has occurred because conflict between loci from the two different genomes can implicate hybridisation (although it can also implicate incomplete lineage sorting, or that the markers in question are inappropriate).

As well as the assumption of dichotomous speciation, molecular phylogenetic inference relies on the selection of useful genetic markers, assumption of correct alignment of sequences, selection of a model of evolution and the choice of tree building method. Tree building methods used in this thesis include maximum parsimony, maximum likelihood and Bayesian approaches. Under parsimony, the preferred phylogenetic tree is the tree that requires the least evolutionary change to explain some observed data. The maximum-likelihood method takes a model of

sequence evolution (essentially a set of parameters that describe the pattern of substitutions) and searches for the combination of parameter values that gives the greatest probability of obtaining the observed sequences. The Bayesian approach selects the tree that has the greatest posterior probability (the probability that the tree is correct), under a specific model of substitution (Bromham and Penny, 2003). In plants, genetic markers commonly used for phylogenetic analysis include chloroplast DNA and nuclear DNA (and to a lesser extent mitochondrial DNA), though DNA fingerprinting methods such as amplified fragment length polymorphisms (AFLPs) have been gaining in popularity, especially when DNA sequence markers are not variable enough for such purposes.

### **Choosing DNA markers for molecular phylogenetic study**

Choosing which genetic markers to use is an important process when attempting to reconstruct molecular phylogenies. Ideally, markers are chosen based on their ability to differentiate between taxa at the level that the research requires. Different molecular tools are required for different questions, because of varying rates of sequence evolution among genomes, genes and gene regions (Small et al., 2004) and the probability that historical signal may be swamped by more recent noise in rapidly-evolving sequences. There is a vast array of molecular tools available e.g. DNA or amino acid sequencing, micro-satellites and AFLPs. The plant systematics community is currently using only a small fraction of tools on hand, with the vast majority of studies considering either chloroplast DNA (cpDNA) or nuclear ribosomal DNA (rDNA).

#### ***Organellar DNA (chloroplast and mitochondrial)***

One advantage of cpDNA as a tool in molecular systematics is its relatively simple genome. Chloroplast genomes vary little in size, structure and gene content over all angiosperms. The genome typically ranges in size from 135kb to 160kb and is characterized by a large (ca. 25kb) inverted repeat which divides the remainder of the genome into one large and one small single copy region (Wolfe and Randle, 2004). This conservatism, and that the chloroplast genome has been sequenced in its entirety for a number of genomes, means it is possible to develop universal primers (Wolfe and Randle, 2004). The large number of copies in each cell (there may be as

many as 900 cpDNA molecules per chloroplast and 50 000 per cell (Kelchner, 2000), allows for easy extraction and amplification (Cowan et al., 2006). Additionally, organelles are usually inherited uni-parentally, which potentially results in more straightforward interpretation (Wolfe and Randle, 2004).

Ironically the properties that make cpDNA an attractive tool for molecular systematics also hinder its usefulness in phylogenetic analysis: phylogenies based on chloroplast DNA may not reveal the true evolutionary history of a taxon for a number of reasons. Because cpDNA is uni-parentally inherited and haploid, it reveals only half of the parentage in plants of hybrid or polyploid origin (Small et al., 2004). Chloroplast DNA alone is unable to identify phylogenetic conflicts arising from hybrid ancestry.

An additional limitation of cpDNA in molecular phylogenetic studies is its relatively slow rate of evolution, which means that even non-coding cpDNA regions often fail to resolve phylogenies at low taxonomic levels. It is for this reason that mitochondrial DNA is rarely used in plant systematics studies (Ennos et al., 1999). 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). Twelve chloroplast and two mitochondrial markers are trialed in this study (see Table 3.1), and two chloroplast markers are selected for use (*trnL-F* and *trnK*).

### ***Nuclear ribosomal DNA***

Nuclear ribosomal DNA has been widely used in order to complement organellar DNA by obtaining additional, independent, bi-parentally inherited phylogenetic estimates that evolve at a much faster rate. The internal transcribed spacer (ITS) region, found in the nuclear genome of plants, fungi and animals is now the most commonly used marker in plant systematics (Alvarez and Wendel, 2003). This refers to the region of 18S, 5.8S and 26S ribosomal genes, along with two internal transcribed spacers. Ribosomal genes exist in tandem arrays of genes composed of hundreds to thousands of copies per array (Small et al., 2004). The ITS region is easily amplified with universal primers due to its conserved priming sites (Hillis and

Dixon, 1991), which flank non-coding regions with high genetic variation in most groups (Alvarez and Wendel, 2003). As a consequence of repeated use, there is a wide assortment of ITS sequences readily available on Genbank (<http://www.ncbi.nlm.nih.gov>).

There are a number of concerns with using ITS as a marker, especially if it is the only marker used in a study. One concern is the high level of sequence variation, which can lead to difficulties in aligning, potentially introducing homoplasy, especially when including outgroups (Alvarez and Wendel, 2003). The biggest concern however is that the copies of ribosomal genes may evolve separately, leading to sequence variants within and between arrays. Individuals included in phylogenetic studies may therefore possess a multiplicity of sequence types. This is a problem if pseudogenes are preferentially amplified over functional rDNA, which can result in incorrect inference of relationships between samples (Alvarez and Wendel, 2003). Cloning all sequence variants may overcome this problem, but there is still a chance that not all variants will have been amplified. Despite these concerns ITS is used in the present study, partly due to the relative ease of sequencing it and the high level of variation it contains, but mainly for the purpose of linking in with previous research (thus continuing the bandwagon effect described in Alvarez and Wendel (2003)). The use of low-copy nuclear genes, still relatively new to the field, may provide a useful alternative (Hughes et al., 2006; Small et al., 2004; Steele et al., 2008)

### ***Low-copy nuclear markers***

The nuclear genome of angiosperms contains a large number of potential genes for phylogenetic analysis—predicted to be from 26,500 nuclear genes in *Arabidopsis thaliana* to 41,000 in *Oryza sativa* (Sterck et al., 2007). As with rDNA, there may be problems with distinguishing between orthologous and paralogous gene copies, but this can usually be overcome by sequencing an adequate number of clones (Steele et al., 2008). The main advantages of low-copy nuclear genes include the higher rate of evolution and the opportunity to accumulate data sets from multiple, unlinked loci. The drawbacks are the lack of universal primers, which means a greater investment in extensive screening is required to identify the most promising loci (Hughes et al.,

2006). As there is no *a priori* reason to expect that any particular gene or gene family will be useful at any given phylogenetic depth for any given group of plants it can be difficult to know where to start looking for a useful gene to sequence (Small et al., 2004). There are several solutions to this. Some researchers prefer to use genes that have proven phylogenetically useful in the past, others conduct screens of unknown genes. Padolina (2006) took a computational approach to design universal angiosperm primers by querying the MoBIoS database (Miranker et al., 2003) to compare the genomes of the monocot *O. sativa* and the eudicot *A. thaliana* to search for primer combinations that occurred only once in each of the two genomes. This search (with additional criteria described in Padolina (2006)) resulted in 141 primer combinations available for any researcher to trial on their own plant group of interest (available at <http://www.sbs.utexas.edu/roxisteele/>). Steele et al. (2008) further trialed these primers and selected 32 primer combinations that amplified regions in representatives of at least two out of five distantly related angiosperm families. It was outside the scope of the present study to trial all 32 suggested primers, but nevertheless I selected four of these primers to trial based on their success in amplifying across the highest number of families in Steele et al.'s (2008) study. I also trialed an additional three low-copy nuclear DNA loci selected from other published studies (Levin et al., 2009; Li et al., 2008; Peralta and Spooner, 2001), however none of these were chosen for further analysis in the present study.

Corroboration of phylogenetic hypotheses by independent datasets increases confidence in a given tree. Sequences can be used in combination with each other if their gene histories do not conflict (Cummings et al., 1995). Incongruence between datasets can be of use to 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. within the Gnaphalieae (Smitsen et al., 2004).

### **AFLPs**

When both cpDNA and ITS sequencing fail to resolve phylogenies, the amplified fragment length polymorphism (AFLP) approach has the potential to overcome such



difficulties. AFLPs are generated by complete restriction endonuclease digestion of total genomic DNA, followed by selective PCR amplification and electrophoresis of a subset of the fragments, resulting in a unique, reproducible fingerprint (or profile) for each individual (Meudt and Clarke, 2007). AFLP markers are sampled throughout the genome and they can therefore reveal rare genetic differences in groups with low sequence variations e.g. among closely related species, crop species, or at the intraspecific level (Meudt and Clarke, 2007). This simultaneous analysis of many loci representing the whole genome is also more likely to generate a true species tree, rather than generating a particular gene tree as sequencing an individual locus is more prone to do (Després et al., 2003). AFLPs have not been widely used in phylogenetic studies though, and instead have been more frequently used to delimit species using ordination methods. Some consider AFLP data inappropriate for any kind of phylogenetic study e.g. (Kosman and Leonard, 2005) but others have found useful phylogenetic signal in AFLP datasets (Koopman, 2005; Meudt and Clarke, 2007). For AFLPs to be suitable for phylogenetic analysis there are two requirements: fragments must have evolved independently, and fragments of equal length must be homologous (Koopman, 2005). There is very little study into the independent evolution of the fragments, but several people have explored the homology or otherwise of same-sized fragments, and have found that in general the more closely related two samples are the more similar the sequences (and therefore homology) of same sized fragments. It is difficult to universally apply these results as taxonomic rank is not equivalent across groups (Meudt and Clarke, 2007), but Koopman et al. (2008) estimated that AFLPs are able to recover phylogenetic relationships among plant samples if the nrITS divergence is between 10-30 nucleotides between sequences. AFLPs have proven useful for exploring variation in the New Zealand flora in a number of groups including *Ourisia* (Meudt et al., 2009), *Pseudopanax*, (Perrie and Shepherd, 2009), and *Leucogenes* (Smitsen and Breitwieser, 2008) and I apply them here to New Zealand *Wahlenbergia*.

### **Molecular clocks**

The molecular clock arose from the observation that the amount of difference between the DNA of two species is a function of the time since their evolutionary separation. This provides a universal tool for placing past evolutionary events in time

(Bromham and Penny, 2003). Since their first conception, molecular clock dating methods have rapidly increased in sophistication and reliability. When first proposed it was necessary to assume a constant rate of evolution throughout the tree, in order to infer divergence dates (the strict molecular clock). This practice has long been challenged by results from datasets showing considerable departures from clocklike evolution (Britten, 1986) and when rate variation does exist among lineages it can seriously mislead not only divergence date estimation but also phylogenetic inference (Drummond et al., 2006). Recent developments allowing for a relaxed molecular clock algorithm, which can concurrently estimate the phylogeny and the divergence rate with different rates of evolution in certain lineages, has vastly improved the accuracy of molecular clock estimates (Drummond et al., 2006). To use genetic distance to estimate divergence time a calibration rate is needed that states the amount of genetic change expected per unit of evolutionary time. The calibration rate is usually calculated for each data set using a known date of divergence to estimate the rate for the whole phylogeny (Bromham and Penny, 2003). The known date is most commonly taken from the fossil record, but other sources include biogeography (for example, the formation of an island). The choice of calibration date is crucial to the accuracy of molecular dates, as shown by the difference in date estimates from studies using the same data and methods but different calibration dates e.g. dating the Cambrian explosion (Smith and Peterson, 2002). In this thesis a relaxed molecular clock using a Bayesian algorithm as implemented in BEAST (Drummond and Rambaut, 2007) is used. The Bayesian algorithm has its strengths and weaknesses; but the explicit nature of the Bayesian method makes its flaws obvious and makes it relatively easy to determine which assumptions should be examined in more detail (Sanderson et al., 2004).

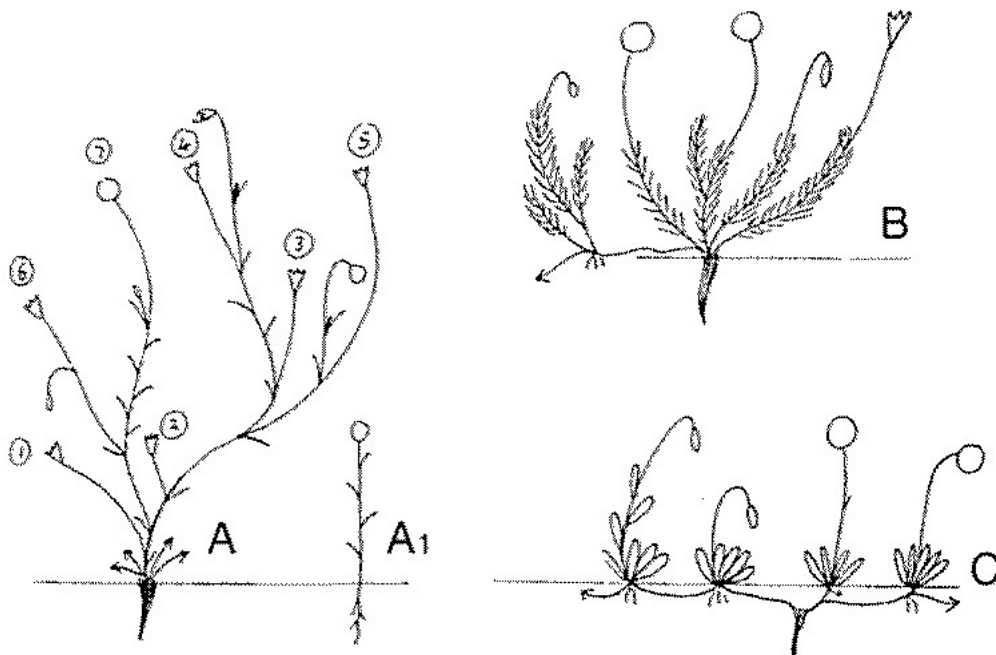
## **Introduction to study group**

In this thesis the phylogeny of the New Zealand species of *Wahlenbergia* Schrader ex Roth (Campanulaceae) and their relationships with overseas *Wahlenbergia* are explored using DNA sequencing and DNA fingerprinting data. *Wahlenbergia* is a large genus (ca. 260 spp.) with a mainly southern hemisphere distribution. The centre of diversity is South Africa (ca. 150-200 spp.), with a secondary centre in Australasia

(ca. 36 spp.), although the genus is also indigenous to New Caledonia, New Guinea, Malaysia, Japan, Taiwan, China, India, Europe, South America and some of the small volcanic Pacific islands e.g. Juan Fernández and Tonga (Pettersson, 1997b). The name *Wahlenbergia* was first introduced by Schrader (1814) but was invalidly published (van Steenis, 1960), which means the first validly published description was by Roth (1821). *Cervicinia* (Delile, 1813) is an earlier validly published name for the genus, but the later *Wahlenbergia* has been conserved against it (Smith, 1992). The type is a South African individual of *W. capensis* (van Steenis, 1960), a South African species that has subsequently been introduced into Australia (Smith, 1992).

*Wahlenbergia* belongs to the large, almost cosmopolitan angiosperm family, Campanulaceae. When considered in the strict sense (excluding Cyphiaceae, Cyphocarpaceae, Lobeliaceae, and Nemacladaceae, which are sometimes treated as part of the Campanulaceae (Haberle et al., 2009)) several recent studies have confirmed the family's monophyly, using morphological (Gustafsson and Bremer, 1995) and molecular data (Lundberg and Bremer, 2003). In their summary of the taxonomic history of the family, Haberle et al. (2009) noted that the first comprehensive monograph of the family was published by De Candolle in 1830 with 21 genera and 234 species. De Candolle divided the family into two groups based primarily on capsule dehiscence: Campanuleae, which dehisce by lateral pores or valves, and Wahlenbergieae, which dehisce by apical valves or pores. The largest genus in the family is the mainly northern hemisphere *Campanula* (ca. 421 species), which forms a broad geographic and taxonomic complement to the second largest (and mostly southern hemisphere) genus, *Wahlenbergia* (Smith, 1992). Haberle et al. (2009) further note that later authors have added to and altered De Candolle's subfamilial divisions, for example Schönland (1889) recognized three main family subdivisions based on capsule dehiscence and the position of ovary locules in relation to the calyx lobes. These were: Campanulinae (having carpels that are generally opposite the calyx lobes, with laterally dehiscent capsules), Wahlenberginae (sepals generally opposite the carpels, but chiefly apically dehiscent capsules) and the Platycodoninae (carpels alternate with the lobes, and mostly apically dehiscent capsules).

More recent phylogenetic studies have again (though informally) revised the subfamilial names within the Campanulaceae, for example Eddie (2003) and Haberle et al. (2009) use the term “campanuloid” to refer to the mostly northern hemisphere taxa allied with *Campanula*, “wahlenbergioid” to refer to the predominantly southern hemisphere taxa traditionally associated with *Wahlenbergia*, and “platycodonoid” for the chiefly Asian colporate/colpate taxa. This thesis will follow the informal subfamilial nomenclature of Eddie (2003) and Haberle et al. (2009). At lower taxonomic levels, the delineation of genera and species in the family has varied widely among authors, who have subdivided the family into 35 to 55 genera and up to 1,046 species (Haberle et al., 2009).



**Figure 1.1** Growth forms of New Zealand *Wahlenbergia*. A = radicate growth form, numbers indicate order of flowering, A1 = false annuals, B = suffructose rhizomatous and C = creeping rhizomatous. Drawing from Petterson (1997b).

New Zealand *Wahlenbergia* plants are all herbaceous perennials (Petterson, 1997b), though annuals are common overseas (e.g. *W. gracilentia* in Australia) and shrubby *Wahlenbergia* occur in South Africa and on islands such as Juan Fernández and Ascension (Smith, 1992). Plants in this genus are generally insect pollinated (Petterson, 1997b) and are conspicuously protandrous – pollen is shed onto the pollen-presenting stylar hairs while the flower is in bud, and these hairs then collapse

soon after the flower is open (Lloyd and Yates, 1982). The stigma of the same flower then becomes receptive (Smith, 1992).

A number of chromosome counts have been undertaken on several *Wahlenbergia* species. Throughout Australasia and southern Asia, the genus has the base chromosome number of  $x=9$ , though many ploidy levels have been found, e.g.,  $2n=2x=18$ ,  $2n=4x=36$ ,  $2n=6x=54$ ,  $2n=8x=72$ , and  $2n=10x=90$  (Petterson et al., 1995; Smith, 1992). In New Zealand three ploidy levels have been reported  $2n=36$ , 54 and 72 (Petterson et al., 1995), which correspond to different growth forms (Figure 1.1). In contrast, chromosome numbers in Africa have been reported with base chromosome numbers of  $x=7$ , 8 and 9, with only one example of a polyploid (Smith 1992). Recent research into *Wahlenbergia* has focused on identifying new species (Plunkett et al., 2009) assessing their use as indicators of climate change (Gallagher et al., 2009), exploring their palatability (Odhav et al., 2007), their seed longevity (Kochanek et al., 2009) and their part in floral deception (Peter and Johnson, 2008).

No phylogenetic studies of the Australasian *Wahlenbergia* have been undertaken so far. Cupido (2009) has recently conducted a molecular systematic study of the South African Campanulaceae using ITS and *trnL-F* in which he included 53 *Wahlenbergia* samples. One recent molecular phylogenetic study of the Campanulaceae family as a whole has found *Wahlenbergia* to be polyphyletic (Haberle et al., 2009) though that study included few *Wahlenbergia* samples. Haberle et al. (2009) used three chloroplast genes – *atpB*, *rbcL* and *matK* to reconstruct a phylogeny of the Campanulaceae that included 102 taxa from 41 genera. They included five *Wahlenbergia* species, two from St Helena Island, one from Juan Fernández Island, one from Australia and one from Europe. These showed *Wahlenbergia* to be polyphyletic as they were recovered in three separate clades, the European sample (*W. hederacea*) not even grouping with the wahlenbergioids.

### **Brief taxonomic history of *Wahlenbergia* in New Zealand**

In this thesis, I focus on the evolution of *Wahlenbergia* species that are native to New Zealand. The current *Wahlenbergia* taxonomy in New Zealand is mostly based on a recent morphological revision by Judith Petterson (nee Hay) (1997b). In that

**Table 1.1** Comparison of the treatments of the New Zealand *Wahlenbergia* in Hooker's handbook of the New Zealand Flora (1864), Allan's (1961) Flora of New Zealand (assisted by Judith Petterson, then Judith Hay) and Judith Petterson's (1997b) revision.

	Hooker Handbook of the NZ Flora 1864	Allan Flora of NZ 1961	Judith Petterson 1997
<b>Radiccate species</b>	<i>W. gracilis</i> var. <i>gracilis</i>	<i>W. gracilis</i>	<i>W. violacea</i>
			<i>W. akaroa</i>
	<i>W. gracilis</i> var. <i>capillarius</i>		<i>W. ramosa</i>
			<i>W. rupestris</i>
			<i>W. vernicosa</i> *
<b>Suffructose rhizomatose species</b>		<i>W. colensoi</i>	JAP decided the type of <i>W. colensoi</i> was made up of starved forms of both <i>W. ramosa</i> and <i>W. rupestris</i>
		<i>W. trychogyna</i>	<i>W. stricta</i> (endemic to Australia)+
		<i>W. mathewsii</i>	<i>W. mathewsii</i>
	<i>W. saxicola</i>	<i>W. brockiei</i>	<i>W. albomarginata</i> subsp. <i>albomarginata</i>
		<i>W. albomarginata</i>	
<b>Creeping rhizomatous species</b>			<i>W. albomarginata</i> subsp. <i>laxa</i>
			<i>W. albomarginata</i> subsp. <i>olivina</i>
			<i>W. albomarginata</i> subsp. <i>decora</i>
		<i>W. simpsonii</i>	<i>W. albomarginata</i> subsp. <i>flexilis</i>
		<i>W. pygmaea</i>	<i>W. pygmaea</i> subsp. <i>pygmaea</i>
			<i>W. pygmaea</i> subsp. <i>drucei</i>
			<i>W. pygmaea</i> subsp. <i>tararua</i>
		<i>W. congesta</i>	<i>W. congesta</i> subsp. <i>congesta</i>
			<i>W. congesta</i> subsp. <i>haastii</i>
	<i>W. cartilaginea</i>	<i>W. cartilaginea</i>	<i>W. cartilaginea</i>

\* Since reclassified as *W. littoricola* subsp. *vernicaosa* (de Lange and Cameron, 1999)

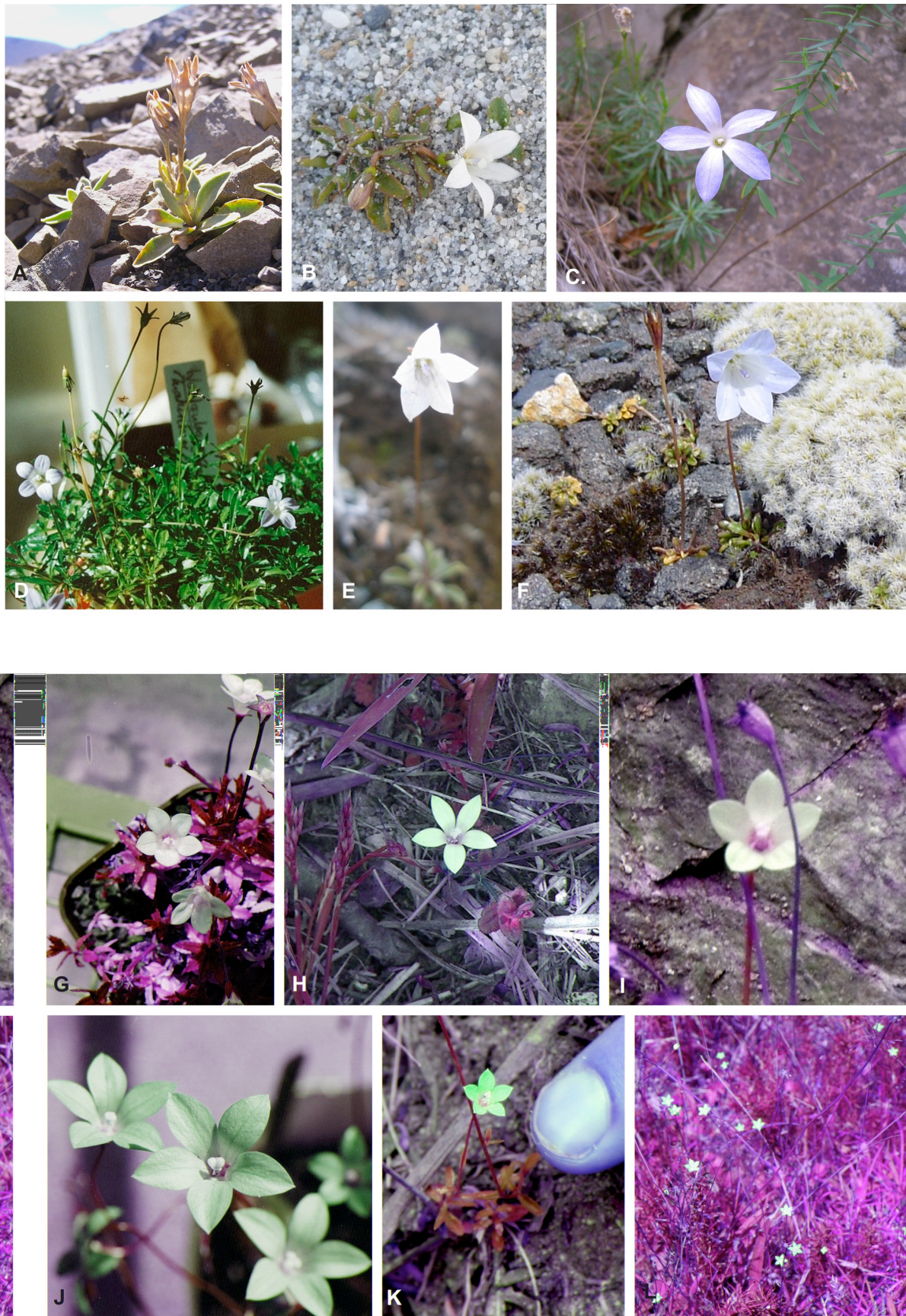
+ One population in NZ near Rarangi, Marlborough, was possibly introduced on sheep from Australia (Petterson, 1997b). No samples where found when I searched the area in March 2009.

revision she provided a taxonomic treatment for nine endemic and one indigenous species (also found on the Kermadecs and possibly Tonga), with a further seven endemic subspecies (Table 1.1). The New Zealand *Wahlenbergia* can be separated into four convenient groups based on growth form and chromosome number. These groups are temperate radicate ( $2n=72$ , 4 species), subtropical radicate ( $2n=54$ , 1 species), suffructose rhizomatous ( $2n=36$  1 species) and creeping rhizomatous ( $2n=36$ , 4 species) (Pettersen, 1997b; Pettersen et al., 1995). The only subsequent taxonomic change has been to reduce *W. vernicosa* to a subspecies of the Australian *W. littoricola* (de Lange and Cameron, 1999). Earlier taxonomic treatments for the New Zealand species (e.g. Hooker (1864) and Hay in Allan (1961)) included fewer (and different) species to Pettersen, which is indicated in Table 1.1.

### *New Zealand radicate species*

The five New Zealand radicate species all inhabit the lowland. They were originally named as one species – *Campanula gracilis* (Forster, 1786). There is a great deal of confusion over whether the type of this species is housed at Kew or Gottingen, and whether it is made up of specimens collected in New Zealand or New Calendonia (or both) (Nicolson and Fosberg, 2004). In 1913 (after the valid publication of *Wahlenbergia* and the change of genus from *Campanula* for the New Zealand species) N. E. Brown published *Wahlenbergia colensoi*, a species he based on small radicate *Wahlenbergia* specimens collected by William Colenso, and sent by him to Sir W. J. Hooker at Kew in the 1840s (Pettersen, 1997a). Descriptions of two further radicate *Wahlenbergia* species, *W. ramosa* and *W. rupestris* were also published (Simpson, 1945, 1952). In her 1997 revision Pettersen recognised 5 radicate species, none of which she called *W. gracilis* as she considered the type to have been collected in New Caledonia, and thought that none of the New Zealand *Wahlenbergia* were conspecific with the type (Pettersen, 1997a). She retained both of Simpson's species (*W. ramosa* and *W. rupestris*), but reduced *W. colensoi* to a synonym of both, having decided that it represented 'starved forms' of the white flowered *W. ramosa* and *W. rupestris* (Pettersen, 1997a). She described three new species, *W. akaroa*, *W. violacea* and *W. vernicosa*. The subtropical *W. vernicosa* has a different chromosome count ( $2n=54$ ) relative to the other New Zealand radicates. The other four temperate radicate species (*W. violacea*, *W. akaroa*, *W. ramosa* and





**Figure 1.2** New Zealand *Wahlenbergia*. A-F have the rhizomatous growth form. A = *W. cartilaginea*, B = *W. congesta* subsp. *haastii*, C = *W. matthewsii* (suffructose rhizomatous growth form), D = *W. pygmaea* subsp. *pygmaea*, E = *W. albomarginata* subsp. *olivina*, F = *W. albomarginata* subsp. *albomarginata*. G-L have the radicate growth form. G = *W. vernicosa*. H-L illustrate the *W. gracilis* complex. H = *W. rupestris*, I = *W. ramosa*, J = *W. akaroa*, K and L = *W. violacea*. Photos A, C, F and H © P. Garnock-Jones, B, E, I, K and L © J. Prebble and D, G and J © G. Petterson.



*W. rupestris*) share a chromosome number of  $2n=72$ , and have since been referred to as a single species: *Wahlenbergia gracilis* (Webb and Simpson, 2001), hereafter referred to as the '*W. gracilis* complex'. Although there is a certain amount of morphological variation within the *W. gracilis* complex (see figure 1.2) it is difficult to find characters that consistently distinguish different lineages within this group.

### *New Zealand rhizomatous species*

The three most morphologically distinct rhizomatous New Zealand *Wahlenbergia* have the simplest taxonomic histories. They were each named once, and these names have been retained until the present (*W. cartilaginea*, *W. congesta* and the suffructose *W. matthewsii* see Figure 1.2). The other rhizomatous *Wahlenbergia* in New Zealand have a more involved taxonomic history. Originally, they were referred to as *Campanula saxicola*, now known as *Wahlenbergia saxicola* and considered endemic to Tasmania. Despite the valid publication of *W. albomarginata* in 1852, alpine *Wahlenbergia* in New Zealand continued to be referred to *W. saxicola* for some time (e.g. see Table 1.1). The name *W. pygmaea* Colenso was then published in 1899 to describe small alpine rhizomatous *Wahlenbergia*. These are the two species recognized by Petterson (1997b). Four further species in this complex were described in the 20<sup>th</sup> century. Petterson (1997b) considered two to be synonymous (*W. flexilis* and *W. simpsonii*) and reduced them both to a subspecies of *W. albomarginata* (*W. albomarginata* subsp. *flexilis*); another (*W. laxa*) was also reduced to a subspecies of *W. albomarginata* and the fourth (*W. brockiei*) she considered to be synonymous with *W. albomarginata*.

## **A note on species concepts**

There has been much discussion in recent times about different species concepts, with vast numbers of different definitions leading to different conclusions concerning the boundaries and number of species (De Queiroz, 2007). Mayer (1997) listed 24 differently named species concepts, the most well-known being the biological (Mayr, 1942), ecological, evolutionary and phylogenetic concepts. While these concepts differ in their methods for delimiting species, they share the common idea that species are separately evolving metapopulation lineages (the 'unified species

concept' of De Queiroz (2007) and it is essentially a confusion between detection protocols and concepts that has lead to the differences in opinion (Hey, 2006). De Queiroz (2007) demonstrates that the different concepts (e.g. biological, ecological, phylogenetic) are better thought of as subcategories, and that evidence from different lines of investigation can be brought to identify the same speciation event. Indeed, a highly supported hypothesis of lineage divergence (i.e. speciation) requires evidence from more than one line of investigation. However, the absence of evidence for speciation from any one or more of the lines of investigation does not constitute evidence against a hypothesis of lineage separation, as the lineage simply may have not yet evolved the property in question. This is particularly pertinent in recently evolving lineages, which is where most of the controversy in delimiting species arises (De Queiroz, 2007; Hey, 2006; Shaffer and Thomson, 2007). Both Smith (1992) and Petterson (1997b) followed a morphological species concept in their revisions of *Wahlenbergia*, though neither were explicit about levels of morphological differentiation used to delimit species or subspecies. The unified species concept will be adopted in this study, with the central line of investigation involving monophyly at one or multiple DNA loci, coupled with morphological, chromosomal and/or ecological distinctness.

## Study aims

The overall aim of this Masters project is to explore the evolution of the Australasian *Wahlenbergia* at a number of different scales. At the large scale I will undertake a molecular phylogenetic analysis exploring the biogeography of *Wahlenbergia* including samples from South Africa, Europe, Australia and New Zealand. At the smaller scale I will assess the current taxonomy of the New Zealand *Wahlenbergia*, using analysis of DNA sequencing and AFLP markers.

# Chapter 2: Biogeography and a first phylogeny of *Wahlenbergia* (Campanulaceae)

---

## Abstract

The first phylogeny of *Wahlenbergia* was reconstructed using about 20% of the genus based on the nuclear ribosomal ITS marker and the chloroplast *trnL-F* marker with samples from South Africa, Europe, Australia and New Zealand. *Wahlenbergia* was confirmed to be polyphyletic, though most of the species form a clade.

*Wahlenbergia* originated in South Africa about 16.2 mya, then dispersed to Australasia and radiated into 45 species and subspecies about 3.7 mya, thus refuting the hypothesis of Gondwanan vicariance for the Australasian group. Two dispersals from Australia to New Zealand are hypothesised, one leading to a radiation of species with the rhizomatous herbaceous growth from ca. 1.0 mya and the other leading to a radiation of species with the radicate growth form 0.49 mya. Dispersals from Australia to New Zealand match the expected direction, following the west wind drift and ocean currents. An herbaceous growth form was shown to be ancestral for the genus as a whole, and chromosome levels have increased as the genus has evolved in Australasia.

## Introduction

The genus *Wahlenbergia* is one of the largest genera within the Campanulaceae, comprising upwards of 260 species (Haberle et al., 2009) mainly distributed in the southern hemisphere. The centre of diversity is South Africa, with a secondary centre in Australasia (Pettersson, 1997b; Smith, 1992) although the genus is also indigenous to New Caledonia, New Guinea, Malaysia, Japan, Taiwan, China, India, Europe, South America and some of the small volcanic Pacific islands, e.g., Juan Fernández

and Tonga (Pettersson, 1997b). There are no *Wahlenbergia* native to North America. The majority of *Wahlenbergia* species are annual or perennial herbs, though a few species have a shrubby habit (Smith, 1992). Their flowers have five petals and range in colour from white to blue-purple and are generally insect pollinated (Pettersson, 1997b). Their seeds are light and are dispersed by wind (Thorsen et al., 2009). The aim of this research is to produce the first phylogeny of *Wahlenbergia* to determine whether *Wahlenbergia* is monophyletic, whether the South African, Australian and New Zealand *Wahlenbergia* are reciprocally monophyletic, and to establish when the South African and Australasian *Wahlenbergia* are likely to have evolved.

Recent molecular studies into the phylogeny of Campanulaceae have included few samples of *Wahlenbergia* i.e. 3 samples in Roquet et al. (2009) and 5 samples in Haberle et al. (2009). Both of these studies found *Wahlenbergia* to be polyphyletic. In Haberle et al.'s (2009) study based on three chloroplast markers, the five *Wahlenbergia* samples were placed in three separate clades on their tree. Interestingly the Australian sample (*W. gloriosa*) grouped with only one of the samples from St Helena Island (*W. linifolia*). The other St Helena sample (*W. angustifolia*) was sister to the sample from Juan Fernández (*W. berteroi*) and the European sample grouped with no other *Wahlenbergia*, a relationship also found by Roquet et al. (2009). As the centre of diversity is South Africa, and the type of *Wahlenbergia* is a South African species (*W. capensis*) it is critical to understand the relationships between these peripheral species that have so far been included in phylogenetic studies and the core of the *Wahlenbergia*.

Recent studies of the New Zealand flora have used molecular phylogenetics to contrast two biogeographic scenarios: Gondwanan vicariance vs. Neogene dispersal, although these are two extremes in a range of potential hypotheses. Few studies support a vicariance scenario (e.g. Stöckler et al. (2002) for *Agathis*), whereas there is increasing evidence from molecular data that much of the biodiversity in New Zealand is the product of rapid evolution following relatively recent colonization (Winkworth et al., 2002a). This has led to much discussion over whether any of New Zealand has remained above water since splitting from Gondwanaland, or whether it has more recently (about 26 million years ago) surfaced after a submerged period – termed the Oligocene drowning (Trewick et al., 2007).

The current distribution of the southern hemisphere *Wahlenbergia* species could suggest a Gondwanan ancestry, with a few long distance dispersal events to Europe and Asia. However, a recent study into molecular dating of the Campanulaceae, which included three *Wahlenbergia* samples, dated the split between the “platycodonoid” species and the ancestor of the “campanuloid” and “wahlenbergioid” clades to between 36.1 and 41 million years ago (mya) (Roquet et al., 2009). These informal terms for groups within the Campanulaceae (“platycodonoid”, “campanuloid” and “wahlenbergioid”) were coined by Eddie et al. (2003) and are roughly equivalent to the subfamily terms used in Roquet et al. (2009). The dates from Roquet et al. (2009) are significantly later than the split up of Gondwanaland, and the absence of extant campanuloid species in Australasia suggests that long distance dispersal rather than vicariance is the most parsimonious explanation for the presence of *Wahlenbergia* there.

Roquet et al. (2009) used two chloroplast markers (*rbcL* and *trnL-F*) and set their molecular clock using four calibration points. First, they placed a minimum age of 16 mya on *Campanula* and related genera based on a *Campanula* seed fossil. Second, they used the endemism of *Azorina vidalii* (Campanulaceae) to the “island” Ponta Delgada (which is actually a city on Santa Maria Island) in the Azores, which was formed 8 mya (Abdel-Monem et al., 1975), to date the split of this species from its sister species. Third, the endemism of *Musschia aurea* (Campanulaceae) to Madeira, dated as being no more than 5.2 mya, gave a date for the split of *M. aurea* from its sister species. Finally, they set a maximum age of 41 mya for the root node based on a large-scale dating analysis of the angiosperm families (Wikström et al., 2001).

Another recent study involving molecular dating analysis of the Campanulaceae included five *Wahlenbergia* samples, but the focus was on Cretan *Campanula* species (Cellinese et al., 2009). This study used three chloroplast loci and was calibrated at two points but did not report any dates regarding the *Wahlenbergia* clades. First, they calibrated the molecular clock using the same fossil seed as (Roquet et al., 2009) but, by making more assumptions about the closest living relatives of the seed fossil they placed it at a different node. Second, they also dated the root of the tree, but they set the age of the split between Campanulaceae and

Lobeliaceae to between 45 mya and 80 mya, with this earlier date representing the age of the split between the Rousseeaceae and the lineage that leads to Campanulaceae/Lobeliaceae (Wikström et al., 2003). They do not report their estimate for the age of the split between the subfamily Platycodoneae and the ancestor of the subfamilies Campanuleae and Wahlenbergieae (equivalent to the platycodonoid, campanuloid and wahlenbergioid of Roquet et al. (2009)), but their estimations are in general older than those of Roquet et al.'s (2009) study (compare Figure 4 in Cellinese et al. (2009) with Figure 3 in Roquet et al. (2009)).

The aims of this chapter are to: 1) produce the first phylogeny of *Wahlenbergia* including about 20% of the approximately 260 species in the genus, 2) date key events for *Wahlenbergia* with a focus on the Australasian and South African clades by extending the dating efforts of previous studies and including a larger group of *Wahlenbergia* samples, 3) map characters of interest onto the phylogeny, and 4) explore the biogeography of the genus and to test for evidence of either Gondwanan vicariance or long distance dispersal in the evolutionary history of this plant group in the southern hemisphere, with a particular focus on New Zealand.

## Materials and Methods

### Genetic markers

Chloroplast DNA (cpDNA) is widely used as a genetic marker – the chloroplast genome is a double stranded circular DNA molecule and has been sequenced in its entirety in a number of organisms, allowing for the development of universal primers (Wolfe and Randle, 2004). The *trnL-F* region comprises the *trnL* (UAA) intron, and an intergenic spacer between the *trnL* (UAA) 3' exon and *trnF* (GAA) (Taberlet et al., 1991). When it was first proposed as a cpDNA marker the intergenic spacer was considered potentially useful for evolutionary studies of related species and probably of populations of the same species (Taberlet et al., 1991). *trnL-F* has been used as a marker in Campanulaceae molecular phylogenetic studies previously (Roquet et al., 2009) which makes it a practical choice in this case as outgroup samples are available on GenBank (<http://www.ncbi.nlm.nih.gov/Genbank/>).

Nuclear ribosomal DNA has been widely used in order to complement organelle DNA as an additional, independent, bi-parentally inherited marker that evolves at a much higher rate. The internal transcribed spacer (ITS) region, found in the nuclear ribosomal genome of plants, fungi and animals is the most commonly used marker in plant systematics (Alvarez and Wendel, 2003). There are a number of concerns with using ITS as a marker, especially if it is the only marker used in a study. One concern is the high level of sequence variation that can occur among genera or even species, which can lead to difficulties in alignment, especially when including outgroups, which can potentially introduce homoplasy into the data set (Alvarez and Wendel, 2003). Like *trnL-F*, ITS has also been used in previous molecular phylogenetic studies of the Campanulaceae (Eddie et al., 2003).

### **Study group**

Location of samples, collection details, voucher information, GenBank accession numbers and references are presented in Appendix 1. The study group comprised 75 samples, 58 of which were *Wahlenbergia*. Fifteen samples of New Zealand *Wahlenbergia* were included, which represented all ten native species. A selection of the Australian native *Wahlenbergia* (ten out of 27) and South African *Wahlenbergia* (31 out of ca. 150-200 spp.) were also included. The South African samples were all recently sequenced by Chris Cupido (Cupido, 2009). An additional two European *Wahlenbergia* sequences downloaded from GenBank (<http://www.ncbi.nlm.nih.gov/Genbank/>) made up the ingroup. These represent all of the *Wahlenbergia* species currently sequenced for both ITS and *trnL-F*. Outgroup species were selected following Haberle et al. (2009) and some outgroup sequences were downloaded from GenBank. Name abbreviations (tag names) for ingroup taxa used in all figures and tables have the first four letters of the species or subspecies name in all caps, followed by the general location, followed by the specific locality (i.e. VIOL\_SI\_Dun is *W. violacea* from the Dun Mountains of the South Island of New Zealand). Tag names for outgroup taxa have the first four letters of the genus, followed by the first four letters of the species epithet (i.e., LOBE\_CARD is *Lobelia cardinalis*; see Appendix 1 and figure legends for more details).

## DNA extraction, amplification and sequencing

Total DNA was extracted from leaf fragments dried in silica gel using a modification of the hot CTAB method of Doyle and Doyle (1990). Ground tissue from approximately 0.5–1.0cm<sup>2</sup> of sample leaf was transferred to a 1.5mL disposable Eppendorf tube and incubated on a heat block for 45 minutes at 60°C in extraction buffer [100mM tris-HCL pH 8.0, 1.4 M NaCl<sub>2</sub> 20 mM EDTA, 2% CTAB (hexadecyltrimethylammonium bromide), 1% PVP40 (polyvinylpyrrolidone)]. 0.6mL chloroform was added and the solutions were gently mixed before spinning at 10 000 rotations per minute (rpm) for 10 seconds. Following centrifugation the clear supernatant (upper phase) was transferred to a new Eppendorf tube and 0.6 mL of isopropanol was added before the solution was stood on ice for 5–10 minutes. If no DNA precipitation was obvious at this stage the extract solution was left overnight at 4°C to precipitate DNA. The following day the DNA precipitate was briefly spun (6000 rpm for 30 seconds) and transferred to a new Eppendorf tube containing 0.8 mL 80% ethanol using a cut pipette tip so as to avoid damaging the DNA. This wash step was repeated, followed by a final spin (6000 rpm 30-120 seconds) and the 80% ethanol was poured off. The DNA pellet was left to air dry, and then re-suspended in 50µL TE buffer (10mM Tris-HLC pH 8.0, 1mM EDTA).

Two regions were then amplified by polymerase chain reaction (PCR) using the following primer sequences: Chloroplast *trnL-Ff*: ATT TGA ACT GGT GAC ACG AG and *trnL-Fc* CGA AAT CGG TAG ACG CTA CG (Taberlet et al., 1991) and nuclear ribosomal ITS5: GGA AGT AAA AGT CGT AAC AAG AAG G and ITS28cc: CGC CGT TAC TAG GGG AAT CCT TGT AAG.

PCR amplification was performed in a Biometra T gradient machine (Whatman, Germany) in a final volume of 25µL. Each 25µL volume contained 2µL of DNA template (nanodrop concentrations ranged from 20ng/µL to 100ng/µL), 1x ThermoPol reaction buffer (10mM KCl, 10mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 20mM Tris-HCL (pH 8.8 @25°C), 2 mM MgSO<sub>4</sub>, 0.1% Triton X-100) (New England BioLabs), 1.5mM MgCl<sub>2</sub> (Bioline), 0.4 mg/mL BSA (Bovine Serum Albumin SIGMA A-2153, fraction V), 250 µmol dNTPs, 0.4 µM of each primer, 0.75 Unit of Taq DNA polymerase (New England Biolabs) and ddH<sub>2</sub>O to make 25µL.



Products were amplified with the following thermalcycler profile: there was an initial denaturation for 2 minutes at 94°C, followed by 35-40 cycles of denaturation for 30 seconds at 94°C, annealing for 30 seconds at 52°C, extension for 1 minute at 72°C, and a final extension step for 5 minutes at 72°C.

Size of the amplified products for all samples was checked using a 100 base pair (bp) DNA ladder (Roche, Penzberg, Germany) on a 1.5% agarose gel. The amplified products were cleaned using 1-5 µL of exosap (Global Science, Auckland) per 25 µL of PCR product. The purified PCR product was then sequenced on an AB13730 Genetic Analyzer by the Allan Wilson Centre Genome Service (Massey University, Palmerston North, New Zealand). Reverse compliments were sequenced to confirm the reads for a number of samples.

### **Dataset alignment and analysis**

Genetic sequences were aligned using Geneious Pro 3.8.5 (Drummond et al., 2007) with a gap-opening penalty of 12 and a gap-extension penalty of 3. The final alignment was checked by eye and adjustments made where necessary. Before alignment the sequences were submitted to a BLAST search (<http://www.ncbi.nlm.nih.gov/blast/>) to confirm that the correct DNA marker had been sequenced. Samples that had been sequenced from forward and reverse directions were combined using Geneious Pro 3.8.5 (Drummond et al., 2007).

### **Phylogenetic and dating analysis**

MrBayes v3.1.1 (Huelsenbeck, 2001) was used to conduct heuristic searches, implementing the AIC model that was selected by jModeltest (Posada, 2008). The *trnL-F* and ITS datasets were analysed separately with four chains and between 2 to 5 million generations, until the standard deviation of split frequencies had fallen below 0.01, which indicates the runs are stationary. The consensus trees were created excluding 10% of the trees as burn-in. Ten percent was selected as an appropriate amount to ensure that the runs had converged on a stationary distribution by considering the standard deviation of split frequencies and exploring the MrBayes output in Tracer (Rambaut and Drummond, 2007). Fifty-percent majority rule

consensus trees were created for both the *trnL-F* and ITS analyses. The congruence of the two datasets was assessed visually by comparing the topology of the chloroplast and ITS Bayesian trees and also by conducting a homogeneity partition test implemented in PAUP\* (Swofford, 2002) with 100 replicates, TBR branch-swapping and maximum of 1000 trees.

The datasets were combined and then analysed using BEAST v1.4.6 (Drummond and Rambaut, 2007), which was used to estimate divergence times from both gene regions simultaneously. Beast uses a relaxed phylogenetic model, where topology and branch lengths are estimated simultaneously from the data. The topology, including placement of the root node, is not specified *a priori* (Drummond et al., 2006). The BEAST .xml input file was created with BEAUTi v 1.4.6 (Drummond and Rambaut, 2007), and it was not necessary to estimate parameters independently between data partitions as both the *trnL-F* and ITS data sets were best estimated by the same model of evolution as determined by jModeltest (GTR + G). A relaxed Bayesian clock with rates for each branch drawn independently from a lognormal distribution was implemented, along with a Yule prior (constant rate of speciation per lineage) (Drummond and Rambaut, 2007). A BEAST run of 10 000 000 iterations was performed and convergence dates and estimated sample size (ESS) were assessed using Tracer v 1.5 (Drummond and Rambaut, 2007). Means and 95% higher posterior densities (HPD) of age estimate were obtained from the outputs using Tracer. The 95% HPD represents the shortest interval that contains 95% of the sampled values from the posterior (Drummond and Rambaut, 2007). After discarding the first 10% of samples as burn-in, the samples from the posterior were summarized on the maximum clade credibility tree (the tree that has the maximum sum of posterior probabilities on its internal nodes (Drummond and Rambaut, 2007)) using TreeAnnotator v 1.5.3 with posterior probability limit set to 0.5 and summarizing mean node heights on. The tree was visualized using FigTree v 1.3.1 (Drummond and Rambaut, 2007).

The relaxed molecular clock (Drummond et al., 2006) was calibrated at two points. First, an internal clade containing *Azorina vidalii* and *Campanula polyclada* was given a mean age (tmrca in BEAST) of 6.5 mya with a normal distribution and standard deviation of 0.8 my. These settings gave a prior 95% confidence interval of

4.9 mya to 8.1 mya, which correspond to the 90% highest posterior density interval found by Roquet et al. (2009) for the same node using the penalized likelihood method, and also included their initial maximum age estimate of 8 mya. Roquet et al. (2009) used the endemism of *Azorina vidalii* (Campanulaceae) on Santa Maria Island in the Azores, which was formed 8 mya (Abdel-Monem et al., 1975), to date the split of *A. vidalii* species from its sister species *Campanula dimorphantha*, *C. polyclada*, *C. dichotoma*, *C. molis* and *C. balfourii*. They set a maximum age of 8 mya for this node, and after analyses estimated this node to range from 5.1 to 8 mya (penalized likelihood dating method,) or 3.8 to 7.2 mya (Bayesian relaxed clock method).

The second calibration point was that of the root node (treemodel.RootHeight in BEAST) which corresponds to the split between the platycodonoids and the wahlenbergioids identified by a previous molecular dating study (Roquet et al. 2009). In the present study the prior was set with a normal distribution and a mean at 38.5 mya, the midpoint of Roquet et al.'s (2009) range from the penalized likelihood dating method, with a range spanning 1.3 standard deviations around this mean. This gave a 95% confidence interval of 35.95 mya to 41.05 mya, which included in it the estimation of Wikström et al. (2001). Roquet et al. (2009) estimated this node to range from 36 to 41 mya (penalized likelihood dating method), or from 20 to 27 mya (Bayesian relaxed clock method). They had initially set this node with a maximum age of 41 mya based on the dating analysis of Wikström et al. (2001).

### **Biogeography and character mapping**

Several character traits were mapped onto the nuclear chloroplast and combined Bayesian trees using MacClade v 4.08 (Maddison, 2001). Characters included geographical location (New Zealand, Australia, South Africa, Europe), growth form (herbaceous vs. dwarf-shrub for all areas and radicate vs. rhizomatous for Australasian species), and ploidy levels (Pettersson et al., 1995; Smith, 1992).

## Results

### Nuclear (ITS) phylogeny

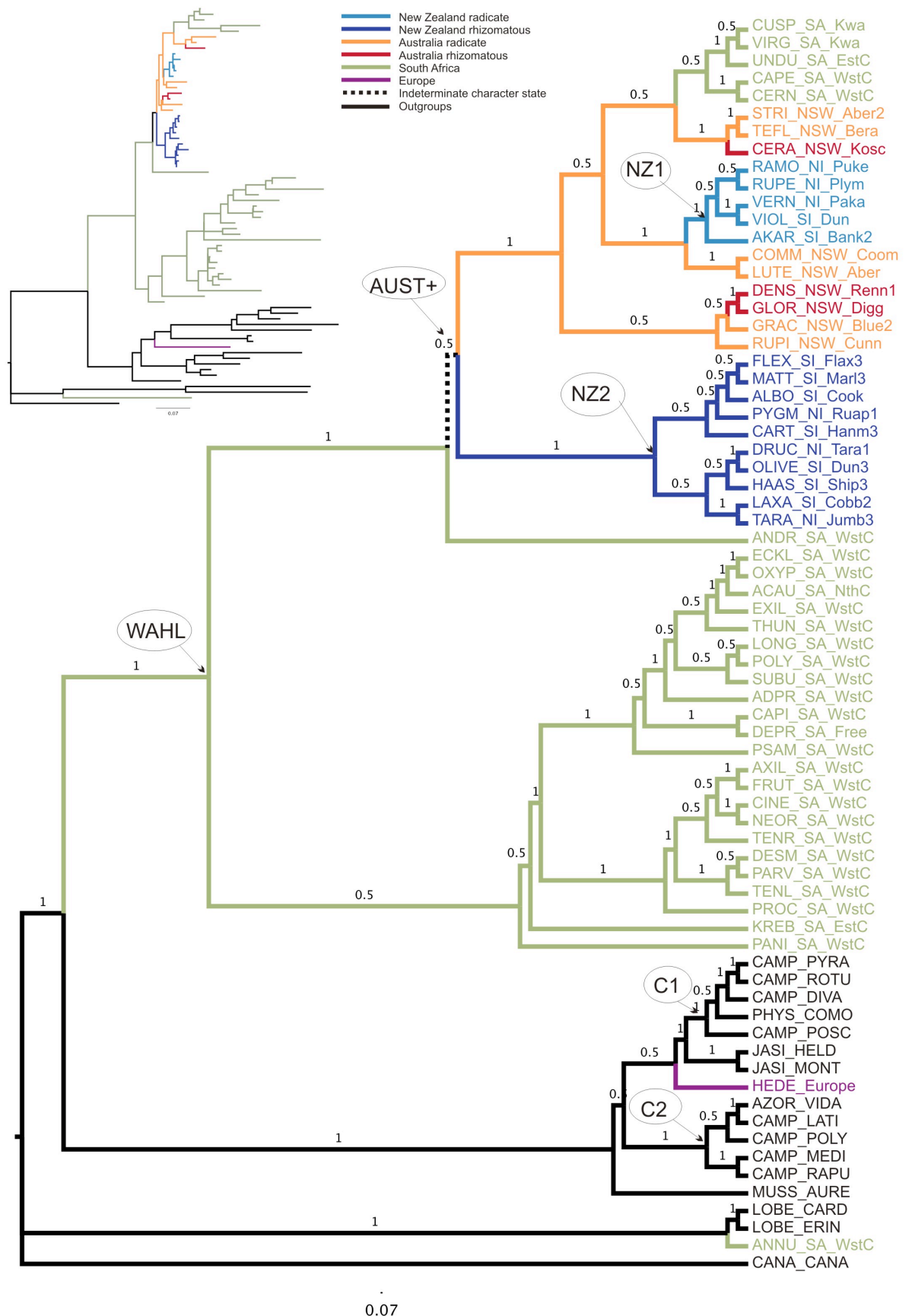
The aligned ITS dataset contained 71 samples and was originally 746 characters long. I had difficulties in aligning both the Australasian samples with the South African samples, and all of the *Wahlenbergia* samples to the outgroup due to high levels of variability in this marker. Homoplasy could thus be a problem in this dataset (Alvarez and Wendel, 2003). Two regions in the alignment, involving bases 50 - 170 and 270 - 510 were excluded, which resulted in a final alignment of 384 characters. Of these 149 were constant. The model of evolution selected by jModeltest was the GTR+G model. The Bayesian analysis ran for 5 million generations, by which time the standard deviation of split frequencies was <0.01.

The 50% majority consensus tree contains a large number of poorly supported branches (e.g. <50% posterior probability (pp) Figure 2.1). Most of the *Wahlenbergia* samples form a clade (labelled “WAHL” in Figure 2.1, 1.0 pp), with the exception of *W. hederacea* (HEDE\_Europe) and *W. annularis* (ANNU\_SA\_WstC). The *W. hederacea* sample (HEDE\_Europe) groups with

some *Campanula*, *Jasione* and *Physoplexus* samples, and *W. annularis* (ANNU\_SA\_WstC) forms a highly supported clade with the two *Lobelia* samples. Within the large *Wahlenbergia* clade there are two main lineages, one containing 23 of the South African species (0.50 pp), and the other the remaining 6 South African species and all of the New Zealand and Australian *Wahlenbergia* (AUST+ 1.0 pp). All of the New Zealand species with a radicate growth form a single clade (NZ1, 1.0 pp), as do all of the New Zealand species with a rhizomatous habit (NZ2, 1.0 pp). Interestingly a poorly supported clade of South African species (0.50 pp), containing the *Wahlenbergia* type species *W. capensis* (CAPE\_SA\_WstC), appears to be derived from within the Australasian *Wahlenbergia*.

### Chloroplast (*trnL-F*) phylogeny

The aligned *trnL-F* dataset of 75 individuals was 945 bases long, and 478 nucleotides were conserved across all samples. This dataset was relatively easy to align by eye.



**Figure 2.1** Bayesian 50% majority rule consensus tree of *Wahlenbergia* (colours) and outgroup (black) based on the ITS dataset. Inset is the phylogram of the same tree. Numbers above branches are posterior probability (pp) values. WAHL = *Wahlenbergia* clade, AUST+ = (mostly) Australasian clade, NZ1 = New Zealand species with a radicate growth form NZ2 = New Zealand species with a rhizomatous growth form. C1 and C2 refer to clades in Haberle et al. (2009). See Appendix 1 for an explanation of the tag names and voucher information.

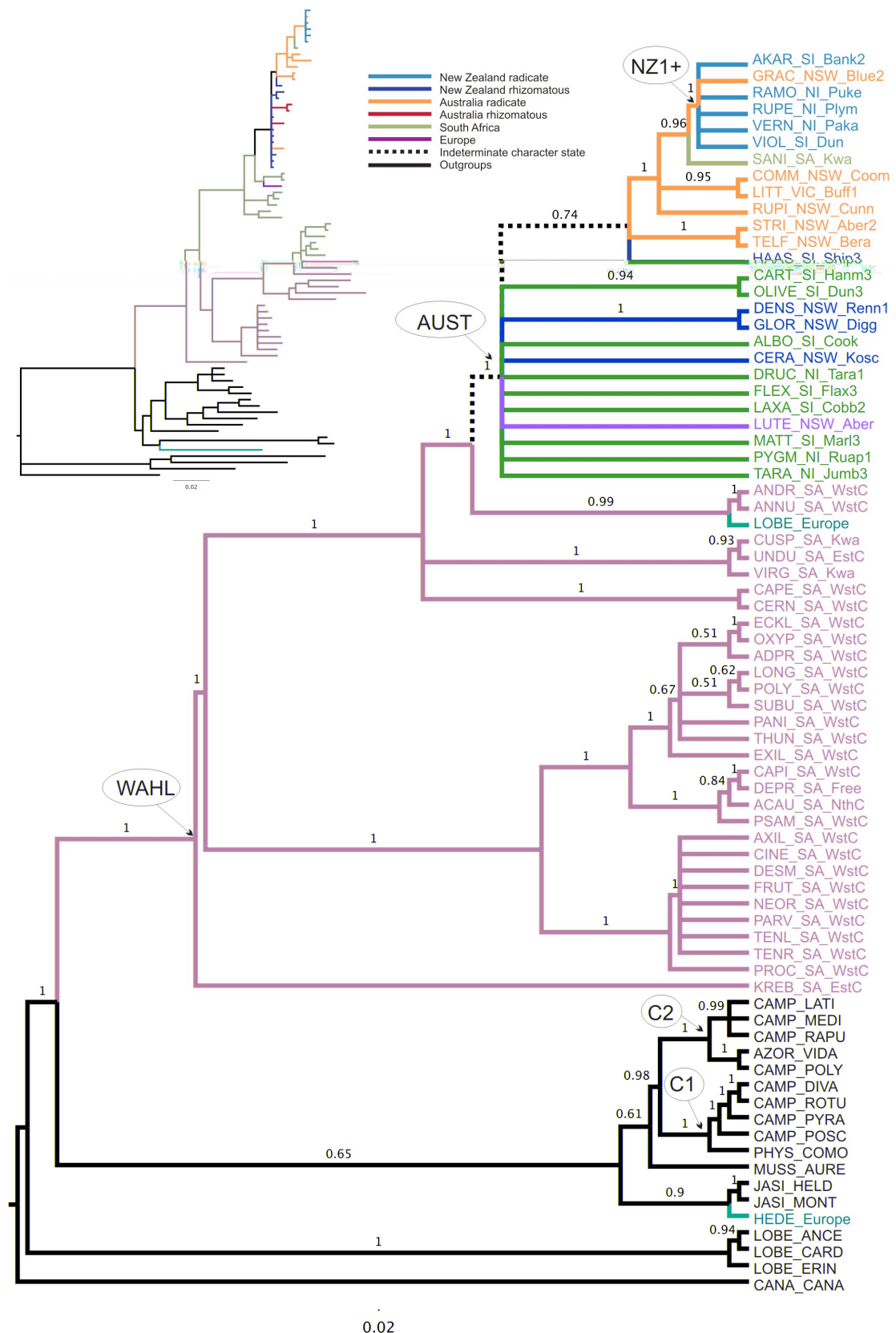
The model of evolution selected by jModeltest was the GTR+G model. The Bayesian analysis ran for 2 million generations, by which time the standard deviation of split frequencies was <0.01.

All *Wahlenbergia* samples, excluding *W. hederacea* (HEDE\_Europe), form a highly supported clade (WHAL 1.0 pp) in the 50% majority rule tree (Figure 2.2).

*Wahlenbergia hederacea* groups with two *Jasione* species. Within the *Wahlenbergia* clade there are three main lineages, the first containing only the South African *W. krebsii* (KREB\_SA\_EstC), the second comprising 22 of the South African species and the third including the remaining 8 South African species, plus all of the Australasian species sampled. In this way, the Australasian *Wahlenbergia*, one unidentified South African specimen and the European *W. lobeliodes* (LOBE\_Europe) are derived with respect to the South African species in a clade that has 1.0 posterior probability.

### **Analysis of concatenated dataset**

The homogeneity partition test showed the nuclear vs. chloroplast alignments to be significantly incongruent ( $p=0.01$ ). However, this test can be highly inaccurate even when the topologies of trees are congruent (Reeves et al., 2001; Yoder et al., 2001) although in this case visual assessment of the topologies does reveal a level of incongruence. Most of the incongruence appears to be “soft” incongruence, meaning that a particular relationship was resolved in one dataset but not in the other. An example of this is the New Zealand rhizomatous samples all of which form a clade in the ITS tree (clade NZ2 Figure 2.1) but are largely unresolved in the *trnL-F* tree (clade AUST+ in Figure 2.2). An example of a highly supported “hard” incongruence between the datasets is the placement of a group of five South African species (*W. cuspidata*, *W. virgata*, *W. undulata*, *W. capensis*, and *W. cernua*), which are highly supported (1.0 pp) as being basal to the Australasian clade in the *trnL-F* tree, but are derived within the Australasian species on the ITS tree. Another striking example of a hard incongruence is the placement of *W. annularis* (ANNU\_SA\_WstC), which formed a highly supported clade with two *Lobelia* samples in the ITS tree, yet groups with the European *W. lobeliodes* (LOBE\_Europe) and the South African *W. androsacea* (ANDR\_SA\_WstC) in the *trnL-F* tree. I



**Figure 2.2** Bayesian 50% majority rule consensus tree of *Wahlenbergia* (colours) and outgroup (black) based on the *trnL-F* dataset. Inset is the phylogram of the same tree. Numbers above branches are posterior probability (pp) values. WAHL = *Wahlenbergia* clade, AUST = Australasian clade, NZ1+ = (mostly) New Zealand species with a radicate growth. New Zealand species with a rhizomatous growth form do not form a clade. C1 and C2 refer to clades in Haberle et al. (2009). See Appendix 1 for an explanation of the tag names and voucher information.

nevertheless proceeded with analysis of the concatenated dataset, with the caveat that the results should be treated with caution due to the incongruence discussed above.

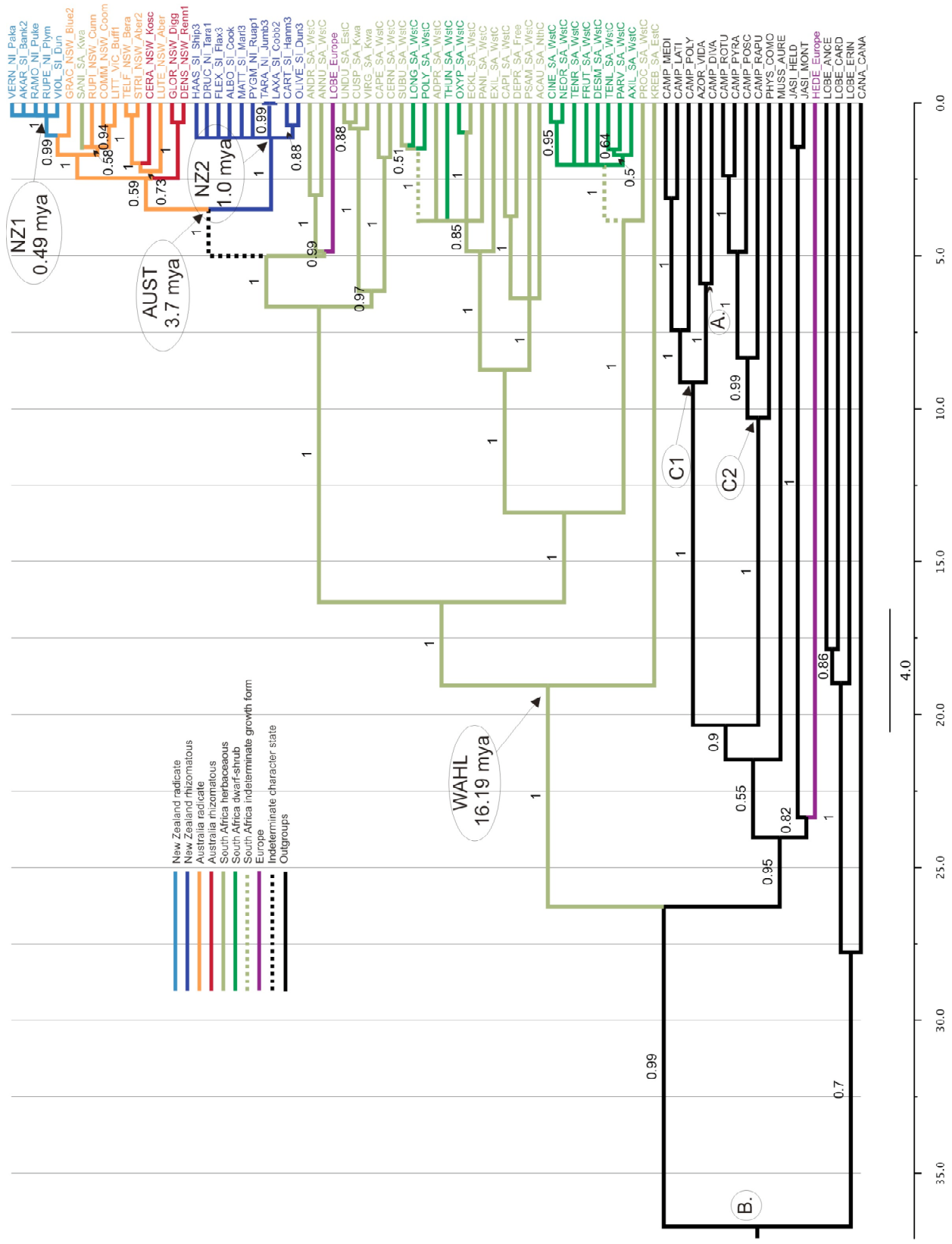
### **Combined dataset**

The final concatenated dataset was 1329 characters long. The four samples that were missing from the ITS dataset (CAMP\_RAPU, LOBE\_ANCE, LITT\_VIC\_Buffl and LOBE\_Europe) were coded as missing data for the ITS characters in the concatenated alignment. The Bayesian phylogeny (Figure 2.3) is generally well resolved at the deeper branches, but there are several polytomies towards the tips of the branches. All *Wahlenbergia* samples, excluding *W. hederacea* (HEDE\_Europe), form a highly supported clade (clade WAHL 1.0 pp in Figure 2.3). *Wahlenbergia hederacea* (HEDE\_Europe) groups with two *Jasione* species. Within the *Wahlenbergia* clade there are three main lineages, the first contains only the South African *W. krebsii* (KREB\_SA\_EstC), the second comprises 22 of the South African species and the third includes the remaining 8 South African species, plus all of the Australasian species sampled. Similar to the *trnL-F* tree the Australasian *Wahlenbergia*, one unidentified South African specimen (SANI\_SA\_Kwa) and the European *W. lobeliodes* (LOBE\_Europe) are derived with respect to the South African species in a clade that has 1.0 posterior probability. Within this clade the New Zealand rhizomatous species form a highly supported monophyletic group (NZ2 1.0 pp) as do the New Zealand radicate species (NZ1 1.0p p). With respect to the outgroups, the *Campanula* species sampled fall into two clades, one of which corresponds with the C1 clade in Haberle et al. (2009), and the other the C2 clade (labelled as such in Figures 2.1 - 2.3). All other outgroup relationships in all three separate and combined are also similar to previous studies.

### **Molecular dating analysis**

Divergence time estimates are shown on Figure 2.3. The *Wahlenbergia* clade (excluding *W. hederacea*) is estimated to have diverged 16.19 mya, with a 95% highest posterior density (HPD) of 10.65 to 21.64 mya (clade WAHL in Figure 2.3). The Australasian *Wahlenbergia* is estimated to have diverged 3.70 mya, with a 95% HPD of 2.33 to 5.25 mya (clade AUST Figure 2.3). Within the Australasian clade the dates of divergence of the two New Zealand clades were also estimated. The clade





**Figure 2.3** Bayesian maximum clade credibility tree of *Wahlenbergia* (colours) and outgroups (black) based on the combined (ITS + *trnL-F*) dataset. Numbers above branches are posterior probability (pp). A and B indicate the calibration points; the node by A was given a mean age of 6.5 mya with a normal distribution and standard deviation of 0.8 my. The node by B was given a mean of 38.5 mya, with a normal distribution and standard deviation of 1.3 my. The scale line is in millions of years before present. WAHL = *Wahlenbergia* clade, AUST = Australasian clade, NZ1 = New Zealand species with a rhizomatous growth form, NZ2 = New Zealand species with a rhizomatous growth form. C1 and C2 refer to clades in Haberle et al. (2009). See Appendix 1 for an explanation of the tag names and voucher information.

containing all of the New Zealand species with a radicate growth form diverged 0.49 mya, with a 95% HPD of 0.16 to 0.89 mya (clade NZ1 Figure 2.3). The clade containing all of the New Zealand *Wahlenbergia* with a rhizomatous growth form diverged 1.00 mya, with a 95% HPD of 0.40 to 1.75 mya (clade NZ2 Figure 2.3).

### **Biogeography and character mapping**

The results of mapping the geographical location of *Wahlenbergia* samples onto the individual nuclear, chloroplast and combined trees (displayed on Figures 2.1-2.3) showed the Australasian species to be largely derived with respect to the South African species, which supports a hypothesis of a South African origin for the genus. In the *trnL-F* and combined trees the Australasian samples form a clade, indicating a single introduction from South Africa to Australasia. Within this Australasian clade the New Zealand and Australian species are not reciprocally monophyletic and instead there are two clades of New Zealand species, (evident in Figure 2.1 and Figure 2.3, but not Figure 2.2), which indicates two separate introductions to, and radiations in, New Zealand.

Plant habit is mapped on to the tree in Figure 2.3. An herbaceous growth form is shown to be the ancestral state for *Wahlenbergia* as a whole, as well as for the Australasian and South African species, and the dwarf-shrub growth form has evolved multiple times in South Africa. Whether the herbaceous South African *Wahlenbergia* have a radicate or rhizomatous growth form has not been well recorded, and as it is difficult to determine from herbarium material, it is therefore unknown for most of the species (Chris Cupido pers .comm. 2010). As I was unable to get information relating to this character for the South African species, I only used these terms for the Australian and New Zealand species (Figures 2.1 - 2.3). The New Zealand radicate species form a clade, as do the New Zealand rhizomatous species. The Australian radicate species are paraphyletic however, as two of the Australian rhizomatous species are sister species but the third is nested within a clade of Australian radicate species.

I was unable to satisfactorily map ploidy levels on to the trees, as very few counts have been completed for South African *Wahlenbergia*, and of the Australian

*Wahlenbergia* samples most species have been found to have multiple ploidy levels within the same species. The counts that have been made of South African species have a base of  $x = 7, 8$  or  $9$  with only one example of polyploidy (Smith, 1992). The Australasian samples that have had their chromosomes counted have all had a base of  $x = 9$ , and the majority are polyploids.

## Discussion

This study is the first to produce a phylogeny of *Wahlenbergia* including 51 out of ca. 260 species (ca. 20%). Below, the relationships of the *Wahlenbergia* samples included in this study are discussed. Biogeographic relationships among *Wahlenbergia* from different regions (particularly the two main centres of diversity, South Africa and Australasia) and the origin and diversification of *Wahlenbergia* based on molecular dating techniques are also discussed.

Within the large *Wahlenbergia* clade (labelled WAHL in Figure 2.3) the South African *Wahlenbergia* are paraphyletic, with the Australasian *Wahlenbergia* nested within them in a highly supported clade (1.0 pp). According to the molecular dating analysis, *Wahlenbergia* evolved ca. 16.19 mya. This suggests the Australasian species, estimated to have arisen ca. 3.7 mya, have all diverged following one long-distance dispersal event from South Africa. This dispersal could have been facilitated by a step-wise progression through Asia, as there are extant species of Asian *Wahlenbergia*. Nevertheless, long-distance dispersal events from South Africa to Australia are not unknown: for example the Australasian Gnaphalidae (Asteraceae) taxa were founded by a single trans-Indian Ocean long-distance dispersal event (Bergh and Linder, 2009). Long-distance dispersal over ca. 8000km of open ocean seems improbable, but may be facilitated by the westerly winds, which increase in force with greater latitude in the Southern Ocean ('roaring forties', 'furious fifties' and 'shrieking sixties'). Another mode of dispersal is oceanic rafting, which would be made possible by the west wind drift, a large surface current flowing from west to east in the Southern Ocean. The west wind drift would take roughly a year to float objects between South Africa and Australia (Waters and Roy, 2004). Bergh and Linder (2009) discounted the hypothesis that the Gnaphalidae could have had a

stepping-stone dispersal through Asia as the Australasian taxa were not nested within a Eurasian clade. The European *W. lobelioides* included in this study was nested within the South African *Wahlenbergia* independently of the Australasian *Wahlenbergia*, and therefore does not implicate a stepping stone dispersal for them, but including Asian *Wahlenbergia* in future studies will be key to distinguishing between the different biogeographic scenarios.

The phylogenetic and molecular dating results of this study do not support the hypothesis of a Gondwanan history for *Wahlenbergia*. Rather they suggest the genus probably diverged in Africa during the Miocene, and then dispersed to Australasia where it diverged there relatively recently during the Pleistocene or Pliocene. Further evidence to suggest *Wahlenbergia* has only been in Australasia for a short while can be found in the fossil pollen record. *Wahlenbergia* has a distinctive pollen type which is known in New Zealand only from the mid-Pliocene (Waipipian Stage ca. 3.6 mya), and younger sediments in New Zealand (pers comm. Dallas Mildenhall May 2009 and (Muller, 1981)). In Australia there are no published accounts of *Wahlenbergia* pollen older than quaternary (e.g. (Prebble et al., 2005)), though the absence of evidence cannot be taken as evidence of the absence of *Wahlenbergia* from Australia before this. A number of recent phylogenetic studies on New Zealand plant groups have found similarly recent dates of arrival and divergence e.g. *Pachycladon* (Joly et al., 2009), *Ourisia* (Meudt and Simpson, 2006), *Ranunculus* (Lockhart et al., 2001) and *Sophora* (Hurr et al., 1999).

In this study relationships within the South African *Wahlenbergia* are often resolved with high support. Chris Cupido will discuss relationships within the South African *Wahlenbergia* further when this chapter is readied for publication as they are his data and I am unknowledgeable about the South African samples.

Relationships within the highly supported Australasian clade (AUST) are poorly resolved, but it is clear that the Australian and New Zealand species are not reciprocally monophyletic. The inclusion of one unidentified South African sample in the Australasian clade is very interesting, and could possibly represent an Australasian species recently introduced by anthropogenic means to South Africa, though attempts to identify it using keys for the Australasian species have been

unsuccessful (Chris Cupido pers comm. 2010). Alternatively, it could represent a natural long-distance dispersal event from Australia to South Africa, which has been documented in other plant groups e.g. *Patersonia* and *Geosiris* (Iridaceae) (Goldblatt et al., 2002). In a literature survey by Bergh and Linder (2008) dispersals in the easterly direction across the Indian Ocean have occurred eight times in the last 55 million years, though they ventured no possible mechanisms for dispersals in this direction. Additionally the South African species nested within the Australasian clade in the ITS tree could represent an additional dispersal from Australia to South Africa, but this requires further investigation.

New Zealand specimens are grouped into two highly supported clades within the Australasian clade (NZ1 and NZ2 in Figure 2.3). All of the samples in NZ1 have the radicate growth form, whereas those in NZ2 have the rhizomatous growth form. The position of NZ2 is unresolved, but NZ1 is nested within a clade containing four Australian species and the unidentified South African species. This topology probably suggests two long-distance dispersal events from Australia resulting in clades NZ1 and NZ2 (even though the placement of NZ2 is unresolved). Long distance dispersal from Australia to New Zealand is a common occurrence (e.g. *Pittosporum* (Chandler et al., 2007; Gemmill et al., 2002), *Elaeocarpus* (Crayn et al., 2006), *Charmichaelia* (Wagstaff et al., 1999) and *Polystichum* (Perrie et al., 2003)). The mechanisms for dispersal across the 2000km of Tasman Sea that separates New Zealand and Australia remain unclear (Winkworth et al., 2002b) although wind dispersal has been advocated for plants with small, lightweight seeds. The seeds of *Wahlenbergia* are lightweight, but without any obvious modifications for wind-dispersal (Pettersen, 1997b). In other cases biotic vectors such as migratory birds, or sea birds have been implicated (Winkworth et al., 2002b).

An herbaceous growth form was shown to be the ancestral state for *Wahlenbergia* as a whole, as well as for the Australasian and South African species. Unfortunately, whether species of herbaceous South African *Wahlenbergia* have a rhizomatous or radicate growth form is largely unrecorded (Chris Cupido pers. comm. 2010). It would be of great interest to determine what growth form the ancestor of the Australasian species that dispersed from South Africa had.

Chromosome numbers were unable to be mapped, as few counts have been made on African species, and most of the Australian species with the radicate growth form have more than one ploidy race. Nevertheless, there does appear to be a pattern of increasing ploidy levels as the genera has evolved, as South African species are mostly diploids, and with a base of  $x=7$ , 8 or 9, whereas the Australasian species all have a base of  $x=9$  and all are polyploids. It appears likely that the ancestor of the Australian species that dispersed from South Africa was of the  $x=9$  group, as Smith (1992) refers to a South African “*W. undulata* group”, which consists of 6 species including *W. undulata* and *W. virgata* (included in this study). The group is characterized by a base chromosome number of  $x=9$ , a distinct corolla-tube and usually stilar glands. Smith (1992) hypothesised that the general affinities of the native Australian species were with the *W. undulata* group in South Africa, and the placement of *W. undulata* and *W. virgata* with regards to the Australasian species in this study (Figure 2.3) offers some support for that hypothesis. However, further chromosome counts of both the South African and Australian species are necessary to explore this issue further.

*Wahlenbergia* is confirmed to be non-monophyletic by this study, with *W. hederacea* from Europe grouping instead to the two sampled species of the outgroup *Jasione* with high support (Figure 2.3). All remaining *Wahlenbergia* included in this study form a highly supported clade (1.0 pp in *trnL-F* and combined datasets). Note however, that my sampling includes only species from South Africa, Australia, New Zealand and southern Europe. Unfortunately, no *Wahlenbergia* samples from Asia, South America or any Pacific islands were able to be included in this study. To fully assess the monophyly of *Wahlenbergia* the next important step will be to include *Wahlenbergia* from these unsampled locations, as well as a selection of other closely related wahlenbergioid species (e.g. *Microcodon glomeratum* and *Theilera guthriei* included in Haberle et al. (2009)) since previous studies have shown *Wahlenbergia* to be polyphyletic even within the wahlenbergioid clade (Haberle et al. 2009, Roquet et al. 2009).

The relationship of *W. hederacea* with the *Jasione* species in the campanuloid clade has been recovered in all studies that have included these species (e.g. (Eddie et al., 2003), (Haberle et al., 2009), and (Cellinese et al., 2009)). Their relationship to the

rest of Campanulaceae was unresolved in Eddie et al.'s (2003) ITS phylogeny, but in all subsequent studies (e.g. chloroplast gene order (Cosner et al., 2004), ITS & *trnL-F* (Roquet et al., 2009) and *atpB*, *matK* and *rbcL* (Haberle et al., 2009)) they have been resolved as part of the campanuloids. Based on their apical dehiscence *W. hederaceae* and *Jasione* species were included in subtribe Wahlenbergeae by De Candolle (1830). However, as noted by Haberle et al. (2009) *Jasione* in particular also share characters with the campanuloids such as *Campanula jacquinii* with which they share characters like flowers in terminal congested heads, subtended by a bract. These results taken together strongly suggest that *Wahlenbergia hederacea* is currently in the incorrect genus and should be considered a campanuloid rather than a wahlenbergioid species. Further studies to find its sister species would be necessary before any taxonomic changes could be undertaken.

The ITS dataset also had one South Africa *Wahlenbergia* (ANNU\_SA\_WstC) grouping with two *Lobelia* samples rather than with the other South African *Wahlenbergia* as it did in the *trnL-F* and combined datasets. While this could represent the true relationship, potentially indicating hybridisation with *Lobelia*, it could also be an artefact of the difficulties in aligning the ITS dataset. Due to the high level of sequence divergence, homoplasy could well have been introduced into this dataset (Alvarez and Wendel, 2003).

Other subfamilial relationships found in this study, including two clades of largely *Campanula* species, match subfamilial relationships found in Roquet et al. (2009), Haberle et al. (2009) and Eddie et al. (2003) but with more limited sampling in this case. The clade labeled C1 in Figure 2.3 matches Haberle et al.'s (2009) C1 clade, and Roquet et al.'s (2009) *Campanula sensu strictu*. The clade labeled C2 in Figure 2.3 matches Haberle et al.'s C2 clade, and Roquet et al.'s *Rapunculus* clade.

## Summary and Conclusions

In this study, the genus *Wahlenbergia* is confirmed to be polyphyletic, although the majority of species form a clade including the type species *W. capensis*. Further sampling is required to confirm whether this monophyly extends to Asian and Pacific

*Wahlenbergia*, and whether it is maintained when other wahlenbergioid species are included in the dataset. The genus originally evolved in South Africa and began diversifying about 16.2 mya before dispersing to Australasia and radiating about 3.7 mya. It is unclear whether this dispersal event included a stepping stone path through Asia, as unfortunately no Asian *Wahlenbergia* were included in this study. Two introductions to New Zealand appear likely, leading to two radiations: one of radicate and one of rhizomatous species. The New Zealand rhizomatous species arrived and diverged about 1 mya, and the radicate species about 0.5 mya. An herbaceous habit was shown to be the ancestral state for the genus as a whole. Although chromosome numbers were unable to be mapped there does appear to be a pattern of increasing ploidy levels as the genus has evolved. The recent evolution of the *Wahlenbergia* genus as a whole and the recent introductions to New Zealand refute the Gondwanan vicariance hypothesis for the origins of the New Zealand *Wahlenbergia* and instead provide overwhelming evidence for arrival via long distance dispersal.



# Chapter Three: Phylogenetic analysis of three DNA sequencing markers to explore relationships within Australasian *Wahlenbergia*

---

## Abstract

A phylogeny of the Australasian *Wahlenbergia* of 105 individuals, representing 64% of the species present in Australasia, was reconstructed based on the nuclear ribosomal ITS marker and the chloroplast *trnL-F* and *trnK* markers following a primer trial of 22 nuclear, chloroplast and mitochondrial markers. Low levels of genetic variation among individuals from Australia and New Zealand was revealed with all markers, and the phylogenies were poorly resolved as a result, a finding that is in line with other published molecular phylogenetic studies on other New Zealand groups. Conflict between the nrITS and chloroplast markers coupled with this low genetic diversity is probably due to rapid and recent evolution during a period of geological and climatic change, perhaps coupled with incomplete lineage sorting and/or hybridisation. Despite the poor resolution, several morphological species were recovered as monophyletic, notably the morphologically distinctive New Zealand *W. cartilaginea*, *W. matthewsii* and *W. congesta* subsp. *congesta*. The lowland radicate *W. gracilis* complex may all belong to the same morphologically variable species. In contrast, the other New Zealand radicate species, *W. vernicosa*, is probably a separately-evolving lineage, and is not conspecific with the *W. gracilis* complex or the Australian *W. littoricola* as previously hypothesized. Two recently described Australian species (*W. rupicola* and *W. telfordii*) formed monophyletic groups and thus the species rank was supported.

## Introduction

The genus *Wahlenbergia* is one of the largest genera within the Campanulaceae, comprising upwards of 260 species (Haberle et al., 2009). It is distributed mainly in the southern hemisphere, although also extends north of the equator into western Europe and eastern Asia (Lammers, 1996). The centre of diversity is South Africa (ca. 81% of *Wahlenbergia* species), with a secondary centre in Australasia (ca. 13%) (Petterson, 1997b; Smith, 1992). The genus is also indigenous to New Caledonia, New Guinea, Malaysia, Japan, Taiwan, China, India and some of the small volcanic Pacific islands, e.g., Juan Fernández and Tonga (Petterson, 1997b). The Australasian centre of diversity is the focus of this chapter. The previous chapter in this thesis found the Australasian *Wahlenbergia* to be monophyletic (including one unidentified South African specimen), derived with respect to the remaining South African species and with two introductions hypothesised from Australia to New Zealand. The focus of this chapter will be on testing the current morphology-based taxonomy of New Zealand *Wahlenbergia* using modern molecular systematic techniques. The present study will be the first dedicated phylogenetic study of the Australasian *Wahlenbergia* and builds on morphological revisions of New Zealand (Petterson 1997b) and Australian (Smith 1992) species.

The current *Wahlenbergia* taxonomy in New Zealand is mostly based on a recent morphological revision by Judith Petterson (1997b). Petterson's revision represents the culmination of a lifetime of study of a very difficult genus (de Lange, 2005) beginning with her Masters thesis on *Wahlenbergia* (Petterson, 1953). In her 1997 revision, she provided taxonomic treatment for ten indigenous species, nine of them endemic and one also found on the Kermadecs and possibly Tonga. Of these, three species are further divided at subspecies rank, providing an addition seven endemic subspecies (Table 1.1). Petterson primarily used morphological and ecological characters to distinguish species, and did not attempt any phylogenetic assessment or analysis. A base chromosome number of  $x=9$  is inferred for the New Zealand species, which can be separated into four convenient groups based on growth form and chromosome number: temperate radicate ( $2n=72$ , 4 octaploid species), subtropical radicate ( $2n=54$ , 1 hexaploid species), suffructose rhizomatous ( $2n=36$ , 1 tetraploid species) and creeping rhizomatous ( $2n=36$ , 4 tetraploid species) (Petterson,

1997b; Petterson et al., 1995), growth forms are illustrated in Figure 1.1. There are no diploids ( $2n=18$ ) found in New Zealand, suggesting our species might have been derived from polyploidy ancestors.

The current Australian *Wahlenbergia* revision (Smith, 1992) included 26 species, 25 native and one introduced. This treatment is generally well accepted, though some state-specific Australian flora include slightly different treatments, e.g. the Flora of Victoria considers *W. littoricola* as a synonym of *W. graniticola* (Walsh, 1999). The Australian species are all herbaceous annuals or perennials, and either have a rhizomatous or tufted growth form (Smith, 1992). The radicate growth form of Petterson (1997b) is equivalent to the tufted growth form of Smith (1992). Several chromosome counts of Australian *Wahlenbergia* have been undertaken (summarized in Smith (1992)), and a base chromosome number of  $x=9$  is evident. Plants with the rhizomatous growth form are all diploids ( $2n=2x=18$ ). Plants with the tufted growth form range from diploids ( $2n=18$ ) to octoploids ( $2n=72$ ), with most species comprising at least two chromosome races (Smith, 1992).

## Research aims

Developments in molecular techniques such as DNA sequencing have provided taxonomists with a wealth of new characters to help make taxonomic decisions, and programs for analyzing these characters (and morphological characters) by creating phylogenies are now well established in the field. Phylogenetic trees provide the framework for the most meaningful evolutionary comparisons (Soltis and Soltis, 2000). Phylogenetic studies of New Zealand plants using DNA sequences as characters have been useful in aiding taxonomic decisions in a whole range of taxa e.g. *Veronica* (Garnock-Jones et al., 2007) and *Pachycladon* (Heenan et al., 2002).

In this chapter, I present the first molecular phylogeny of the Australasian *Wahlenbergia*, with the primary aim of testing the current New Zealand *Wahlenbergia* taxonomy, although questions regarding the Australian species will also be addressed where sampling permits. Appropriate markers for DNA sequencing have been selected from a trial of organellar DNA (chloroplast and

mitochondrial) and nuclear DNA (both low-copy and ribosomal). Specific questions include:

1. species boundaries within the New Zealand *W. gracilis* complex,
2. the status of *W. littoricola* subsp. *vernica*,
3. the status of subspecies of the New Zealand rhizomatous species e.g. those of *W. albomarginata* and *W. pygmaea* and
4. monophyly of the newly described Australian species *W. rupicola* and *W. telfordii*.

These specific questions are further introduced below, along with the species and species complexes they relate to.

## **Taxonomic introduction**

### ***New Zealand radicate species***

All of the New Zealand radicate species inhabit lowland habitats. Morphologically they appear very similar – especially when dried and pressed – and precise floral measurements are often necessary to distinguish them. The five radicate species have a lively taxonomic history (Pettersen, 1997a) and were originally considered to be one species – *Campanula gracilis* (Forster, 1786). Four of these species are temperate and one, *W. vernica*, is subtropical.

The four endemic temperate radicate species (*W. violacea*, *W. akaroa*, *W. ramosa* and *W. rupestris*) share a chromosome number of  $2n=72$ . They are hard to distinguish, and hybrids (*W. violacea* × *W. rupestris*, *W. violacea* × *W. akaroa* and *W. rupestris* × *W. akaroa*) have been reported (Pettersen, 1997b). Webb and Simpson (2001) referred to the four temperate radicate species as a single species: *Wahlenbergia gracilis* (hereafter referred to as the *W. gracilis* complex). They found no way to distinguish between the seeds of any of the five radicate species, but accepted that *W. vernica* should be considered a separate species due to its different chromosome count. The radicate species are characterised mainly based on floral characters. *W. violacea* and *W. akaroa* have blue/purple flowers, *W. akaroa* is restricted to Banks Peninsula and has flowers approximately twice as large as *W.*

*violacea*. *W. ramosa* and *W. rupestris* have white or pale blue flowers, and are distinguished by *W. rupestris* having non-overlapping elliptic shaped petals, compared to *W. ramosa* having ovate shaped petals that overlap (Pettersson, 1997b). The taxonomy is further confused by the enigmatic *W. colensoi*, which was treated as a starved form of *W. ramosa*, *W. rupestris* or *W. violacea* in Pettersson (1997b), but has been found growing in sympatry with more typical samples of these species (JMP, pers. obs. based on notes on herbarium specimens). Whether these entities should be considered varieties, subspecies or species, and how many should be recognised, are unresolved questions. I address species boundaries and ranks in the *W. gracilis* complex by using DNA sequence data to test whether clades based on DNA sequences match the taxonomy based on morphological characters.

The indigenous tropical radicate species, *W. vernicosa*, has a different chromosome count of  $2n=54$  which is unique in New Zealand. Subsequent to Pettersson's revision, de Lange and Cameron (1999) reduced *W. vernicosa* to a subspecies of the otherwise Australian *W. littoricola* on the basis that the only morphological character to distinguish *W. vernicosa* from *W. littoricola* is its glossy leaf surface. Pettersson (2005) contested this reduction to subspecies rank, but without providing any further evidence for her rejection of the change this matter deserves further investigation. Here, I use DNA sequence data to examine the origins of this hexaploid species (in the absence of a diploid New Zealand species as a putative parent) and to test its proposed conspecificity with Australian *W. littoricola*.

### *New Zealand rhizomatous species*

All of the rhizomatous species in New Zealand share the same chromosome count of  $2n=36$ . Only one species has the suffructose rhizomatous growth form and four have creeping rhizomatous habit.

The suffructose *W. matthewsii* is easy to identify as a species because of its unique growth form, along with its limestone habitat. Putative hybrids between *W. matthewsii* and *W. albomarginata* from Castle Hill and Prebble Hill are represented by herbarium specimens at the Allan Herbarium (CHR). These have the creeping rhizomatous growth form characteristic of *W. albomarginata*, yet linear leaves

characteristic of *W. matthewsii*. Some of these samples (e.g. CHR 76373) were previously identified as *W. brockiei* (Pettersen, 1997b). *Wahlenbergia brockiei* was not recognized by Judith Pettersen as she considered the type to be a dry habitat form of *W. albomarginata* (Pettersen, 1997b) based on growth observations of *W. brockiei* and *W. albomarginata* in a common garden. Unfortunately, *W. brockiei* was not able to be sampled for the present study.

There are four species with the creeping rhizomatous growth form: *Wahlenbergia albomarginata*, *W. pygmaea*, *W. cartilaginea* and *W. congesta*. *W. congesta* is a relatively distinctive species, being the only consistently coastal rhizomatous *Wahlenbergia* in New Zealand. It inhabits the west coast of the South Island. Pettersen (1997) recognised two subspecies, *W. congesta* subsp. *haastii* is characterised by smaller capsules and flowers than *W. congesta* subsp. *congesta*. The most distinctive *Wahlenbergia* in New Zealand is *W. cartilaginea*, which is restricted to scree slopes in the northwestern South Island and has very thick glaucous cartilaginous leaves and calyx lobes. A number of New Zealand genera each contain just one or a few scree-adapted species (e.g., *Stellaria roughii*, *Lobelia roughii*, *Notothlaspi rosulatum*, *Ranunculus haastii* and *R. acraeus*), suggesting the scree habitat might be a relatively recent product of greywacke mountain uplift and erosion cycles that include glaciation and interglacial aridity and wind. *W. cartilaginea* is more locally distributed than many of the others, which might indicate a very recent origin.

*Wahlenbergia albomarginata* (sensu Judith Pettersen) is restricted to the South Island, primarily in alpine habitats. It has five subspecies, which are distinguished by leaf characters combined with geographic and geological (e.g. edaphic) separation. Several of the five subspecies are difficult to identify in practice (e.g. *W. albomarginata* subsp. *albomarginata* (dry areas of eastern South Island, Central Otago, Fiordland) vs. *W. albomarginata* subsp. *laxa* (inland and wetter areas of South Island) and *W. albomarginata* subsp. *decora* (higher altitude areas)), as their leaf morphology can be intermediate. *W. albomarginata* subsp. *flexilis* is easy to distinguish as it is restricted to eastern Marlborough limestone, though it would be difficult to identify based on morphological characters alone (JMP pers. obs.). The remaining subspecies, *W. albomarginata* subsp. *olivina*, is the only subspecies

consistently able to be identified using the key in Petterson (1997b) as it differs both in habitat requirements (it grows on ultramafic serpentine rock outcrops) and in stable morphological characters (it has consistently thicker leaf margins).

*W. pygmaea* sensu Judith Petterson effectively replaces *W. albomarginata* in the North Island as the common alpine species. It is generally a much smaller plant as its name suggests, and in the past small South Island *W. albomarginata* plants have been identified as *W. pygmaea* (JMP pers. obs. of herbarium specimen labels at WELT and NSW). Other than size of plants the main character that separates *W. albomarginata* from *W. pygmaea* is the shape of the flowers. *W. albomarginata* has a ‘narrow-campanulate-rotate corolla, with tube distinctly longer than broad’ whereas in *W. pygmaea* the corolla is ‘broadly campanulate, with tube as wide as, or wider than long’ (Petterson, 1997b). There are three subspecies of *W. pygmaea* currently recognised, defined primarily by geography, but also apparently by floral and foliar differences (Petterson, 1997b). *Wahlenbergia pygmaea* subsp. *pygmaea* is found growing on the central plateau of the North Island and south into the Ruahine ranges. It has leaves that are entire or with 2-4 teeth on each side. *Wahlenbergia pygmaea* subsp. *tararua* is restricted to the Tararua ranges, and only one population is known from there now. Its leaves are serrate, with 4-7 teeth on each side. Doubts have been expressed as to whether there is sufficient evidence to consider this a separate subspecies and recently it was excluded from the naturally uncommon species list, on the basis that it was taxonomically indistinct (de Lange et al., 2009). *Wahlenbergia pygmaea* subsp. *drucei* is restricted to Mt Taranaki. It subtly differs morphologically by having dentate leaves with 4-7 teeth on each side and broader capsules than *W. pygmaea* subsp. *pygmaea*.

Despite the floral differences described above it is not possible to distinguish between *W. pygmaea* and *W. albomarginata* on the basis of their seeds. However, seeds may not be a useful character in this group, because seeds of the otherwise morphologically distinct South Island west coast endemic *W. congesta* cannot be distinguished from those of these two species either (Webb and Simpson, 2001). In this thesis I use molecular characters to examine the monophyly and relationships of these two species and their intraspecific taxa as an independent test of their status.

### *Australian species*

There were 25 native species included in Smith's 1992 revision, 19 of which have a radicate ("tuffed") growth form, and six have a rhizomatous growth form. The radicate species can be further split into three groups based on morphological and cytological evidence (Smith, 1992). The *W. scopulicola* group consists of rock-crevice growing species, *W. scopulicola*, *W. glabra* and *W. islensis*. The *W. gracilenta* group consists of all of the species with an annual life cycle, *W. gracilenta*, *W. pressii*, *W. victoriensis* and *W. caryophylloides*. The *W. communis* group is the largest group, consisting of twelve species. These share the same growth form and life cycle as the New Zealand radicate species and are the most in need of revision (Smith, 1992). The members include *W. stricta*, *W. luteola*, *W. communis*, *W. queenslandica* and three subgroups of closely related species. The *W. graniticola* subgroup consists of *W. graniticola*, *W. aridicola* and *W. littoricola*. The *W. planiflora* subgroup consists of *W. planiflora*, *W. fluminalis* and *W. multicaulis* and the *W. gracilis* subgroup consists of *W. gracilis* and *W. tumidifructa*.

The rhizomatous species (termed the *W. gloriosa* group in Smith (1992)) include *W. ceracea*, *W. gloriosa*, *W. densifolia*, *W. gymnoclada*, *W. saxicola* and *W. insulae-howei*. Smith (1992) conjectured the New Zealand rhizomatous species were related to the *W. gloriosa* group, and suggested the group were most likely to have spread from Australia to New Zealand, rather than the other way around, based on the increased ploidy levels of the New Zealand specimens.

Subsequent to Smith's (1992) revision two new Australian radicate species have been described, *W. rupicola* and *W. telfordii* (Plunkett et al., 2009). A dendrogram based on phonetic analysis of morphological characters grouped *W. telfordii* with individuals of *W. glabra*, *W. ceracea* and *W. stricta*. *Wahlenbergia rupicola* grouped most closed with *W. luteola*, though had similar seed morphology to *W. ceracea*, and a superficial morphological resemblance to *W. stricta*. Here I use molecular characters to examine the monophyly and relationships of these two new species as an independent test of their status.



## Methods

### Taxon sampling

Location of samples, collection details, voucher information and GenBank accession numbers are presented in Appendix 1. A total of 105 individuals are included in this study, including 94 Australasian *Wahlenbergia* individuals (representing 29 of the 45 species and subspecies), two European *Wahlenbergia* individuals and nine outgroup individuals. Fifteen of the 17 species and subspecies of New Zealand *Wahlenbergia* (Pettersen, 1997b) are represented by 54 individuals in this study. The only New Zealand *Wahlenbergia* taxa not included in this study are *W. congesta* subsp. *congesta* and *W. albomarginata* subsp. *decora*. Six species and two subspecies are represented by more than one sampled population e.g. *W. violacea*, *W. littoricola* subsp. *vernica* and *W. albomarginata* subsp. *albomarginata*. This was not practical for species with more restricted distributions, and in some cases all samples of a particular species/subspecies are members of the same population e.g. *W. akaroa* and *W. albomarginata* subsp. *flexilis*. Ten populations are represented by more than one plant.

The Australian *Wahlenbergia* are represented in this study by 40 samples comprising between 1 and 6 individuals of 14 of the 28 species present in Australia as classified by Smith (1992) and (Plunkett et al., 2009). Not all Australian species were able to be included in this study due to collection constraints – most species present in New South Wales were collected on one trip to this state, and additional samples were extracted from herbarium specimens or sent by other researchers. Sequences of European *Wahlenbergia* and some outgroups were downloaded from GenBank (<http://www.ncbi.nlm.nih.gov/>) and were selected to represent closely related genera following Haberle et al. (2009).

### DNA isolation, amplification and sequencing

Total DNA was extracted from silica-dried leaves after manual disruption of dried tissue with a mortar and pestle using a (CTAB) method modified from Doyle & Doyle (1990) and explained in detail in Chapter Two of this thesis.

Leaves from older herbarium specimens from the herbarium at the museum of New Zealand Te Papa Tongawera (WELT) collected between 1994-1997 were removed with permission and extracted in a dedicated ancient DNA (aDNA) lab at Massey University, Palmerston North. These samples were extracted using a Qiagen DNeasy plant mini extraction kit following the manufacturers instructions. Half of the eluted aDNA was then returned to the modern lab at Victoria University for amplification and sequencing, with half stored in the aDNA lab for future use.

Twenty chloroplast, mitochondrial and nuclear DNA sequencing markers were trialled. A list of the primers trialled can be found in Table 3.1. Selection criteria for successful primers included those that amplified cleanly (single bands) and universally across a selection of different *Wahlenbergia*. The primers were trialled on the Australian radicate *W. communis* (COMM\_NSW\_Quea), the New Zealand temperate radicate *W. violacea* (VIOL\_SI\_Dun) and two New Zealand creeping rhizomatous subspecies: *W. albomarginata* subsp. *laxa* (LAXA\_SI\_Cobb1) and *W. albomarginata* subsp. *albomarginata* (ALBO\_SI\_Garv). PCR amplification and sequencing of the trial markers followed protocols outlined in the references in Table 3.1. Protocols for the markers eventually chosen for this study (the nuclear ribosomal DNA marker ITS, and the chloroplast DNA regions *trnL-F* and *trnK*) are outlined below.

The oligonucleotide primers for the three loci (*trnL-F*, *trnK* and ITS) are listed in Table 3.1. PCR amplification was performed in a Biometra T gradient machine (Whatman, Germany) in a final volume of 25µL. Each 25µL volume contained 2µL of DNA template (nanodrop concentrations ranged from 20ng/µL to 100ng/µL), 1x ThermoPol reaction buffer (10mM KCl, 10mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 20mM Tris-HCL (pH 8.8 @25°C), 2 mM MgSO<sub>4</sub>, 0.1% Triton X-100) (New England BioLabs), 1.5mM MgCl<sub>2</sub> (Bioline), 0.4 mg/mL BSA (Bovine Serum Albumin SIGMA A-2153, fraction V), 250 µmol dNTPs, 0.4 µM of each primer, 0.75 Unit of Taq DNA polymerase (New England Biolabs) and ddH<sub>2</sub>O to make 25µL. In some instances 4% v/v of Betaine (N,N,N-trimethylglycine) and/or 4% v/v DMSO (Dimethyl sulfoxide) were added to the master mix to optimise the PCR conditions when amplifying ITS

fragments. The volume of ddH<sub>2</sub>O was altered to maintain a reaction volume of 25µl in the event of such additions.

Products were amplified with the following thermalcycler profiles. For the *trnL-F* region and the ITS region there was an initial denaturation for 2 minutes at 94°C, followed by 35-40 cycles of denaturation for 30 seconds at 94°C, annealing for 30 seconds at 52°C, extension for 1 minute at 72°C, and a final extension step for 5 minutes at 72°C. The other chloroplast region, *trnK*, was amplified using the following profile: an initial denaturation at 94°C for 2 minutes followed by 35 cycles of 94°C for 1 minute, 50°C for 1 minute, 72°C for one minute and finished with 5 minutes extension at 72°C.

Size of amplified PCR products for all samples were checked using a 100 base pair (bp) DNA ladder (Roche, Penzberg, Germany), or in some instances a 123 base pair ladder (Invitrogen) on a 1.5% agarose gel (photos of gels of all trialled primers are displayed in Figures 3.1 and 3.2). The amplified products were cleaned using exo-sap with 1-5µL of exo-sap per 25 µL of PCR product. The purified PCR product was then sequenced on an AB13730 Genetic Analyzer by the Allan Wilson Centre Genome Service (Massey University, Palmerston North, New Zealand). Reverse compliments were sequenced to confirm reads for a number of samples.

### **Sequence alignment and phylogenetic analysis**

DNA sequences were aligned using Geneious Pro 3.8.5 (Drummond et al., 2007). Before alignment the sequences were submitted to a BLAST search (<http://www.ncbi.nlm.nih.gov/blast/>) to check the expected marker had indeed been amplified. Samples that had been sequenced from forward and reverse directions were combined using Geneious Pro 3.8.5 (Drummond et al., 2007). Of the 105 sequences analysed in this study six were downloaded from GenBank, 30 were sequenced in Chapter Two and 69 were newly sequenced here. All new sequences will be deposited in GenBank upon publication.

Analyses were first performed individually for chloroplast (*trnL-F* and *trnK*) and nuclear (ITS) data sets and then on a concatenated data set (*trnL-F*, *trnK* and ITS).

Some analyses were run on the freely available Bioportal ([www.bioportal.uio.no](http://www.bioportal.uio.no)). Congruence between the ITS and chloroplast datasets was assessed by visual assessment of the topologies and by running an homogeneity partition test in PAUP\* (Swofford, 2002) with 100 replicates, TBR branch swapping and maximum of 1000 trees. Maximum Parsimony (MP) analyses were conducted using PAUP\* v4.0b10 (Swofford, 2002). Due to the large number of most parsimonious trees (MPTs), MP analysis was performed in a two-step search strategy. First, multiple islands were searched with 10 000 random addition sequence replicates, nchuck = 5, chuckscore = 1 and maxtrees = 10 000. The resulting trees were then swapped to completion with the same settings but chuckscore = no. Support for clades was assessed using 1000 bootstrap replicates, 10 random addition replicates, and MAXTREES = 20 000 in PAUP\*.

Maximum likelihood (ML) tests were conducted using Garli (Zwickl, 2006) using mostly default settings, except that 1000 bootstrap replicates were conducted and the model of evolution was altered to fit those selected by jModelTest (Posada, 2008). JModelTest was used to test the fit amongst 88 different models of different complexity. Models were selected using Akaike information weights criterion (AIC).

MrBayes v3.1.1 (Huelsenbeck, 2001) was used to conduct heuristic searches, also implementing the AIC model that was selected by jModelTest (Posada, 2008). Each dataset was analysed with four chains and either 7 or 10 million generations, in most cases until the standard deviation of split frequencies had fallen below 0.01, which indicates the runs are stationary. The consensus trees were created excluding 10% of the trees as burn-in. Ten percent was selected as an appropriate amount to ensure that the runs had converged on a stationary distribution by considering the standard deviation of split frequencies and exploring the MrBayes output in Tracer v. 1.5 (Rambaut and Drummond, 2007). Fifty-percent majority rule consensus trees were created for all Bayesian analyses.

SplitsTree v4.10 (Huson and Bryant, 2006) was used to create neighbour net networks to explore conflicting signals both with the chloroplast and nuclear ITS datasets separately, and within the combined datasets. Outgroup samples were excluded from these analyses.

## Biogeographic and character analysis

Several character traits were mapped onto the nuclear chloroplast and combined Bayesian trees using MacClade v 4.08 (Maddison, 2001). Characters included geographical location (New Zealand, Australia, Europe) and herbaceous growth form (radicate vs. rhizomatous). No additional characters were mapped as analysis in Chapter Two showed that mapping ploidy levels was unfeasible for this study, and preliminary analysis showed that additional geographical characters were not informative and appeared to explain less of the topology of the phylogenetic trees than the large scale distribution (data not shown, three different scales were explored: 1. North or South Island, 2. west or east of the alpine fault, and 3. North Island, north of the South Island, middle of the South Island and south of the South Island).

## Results

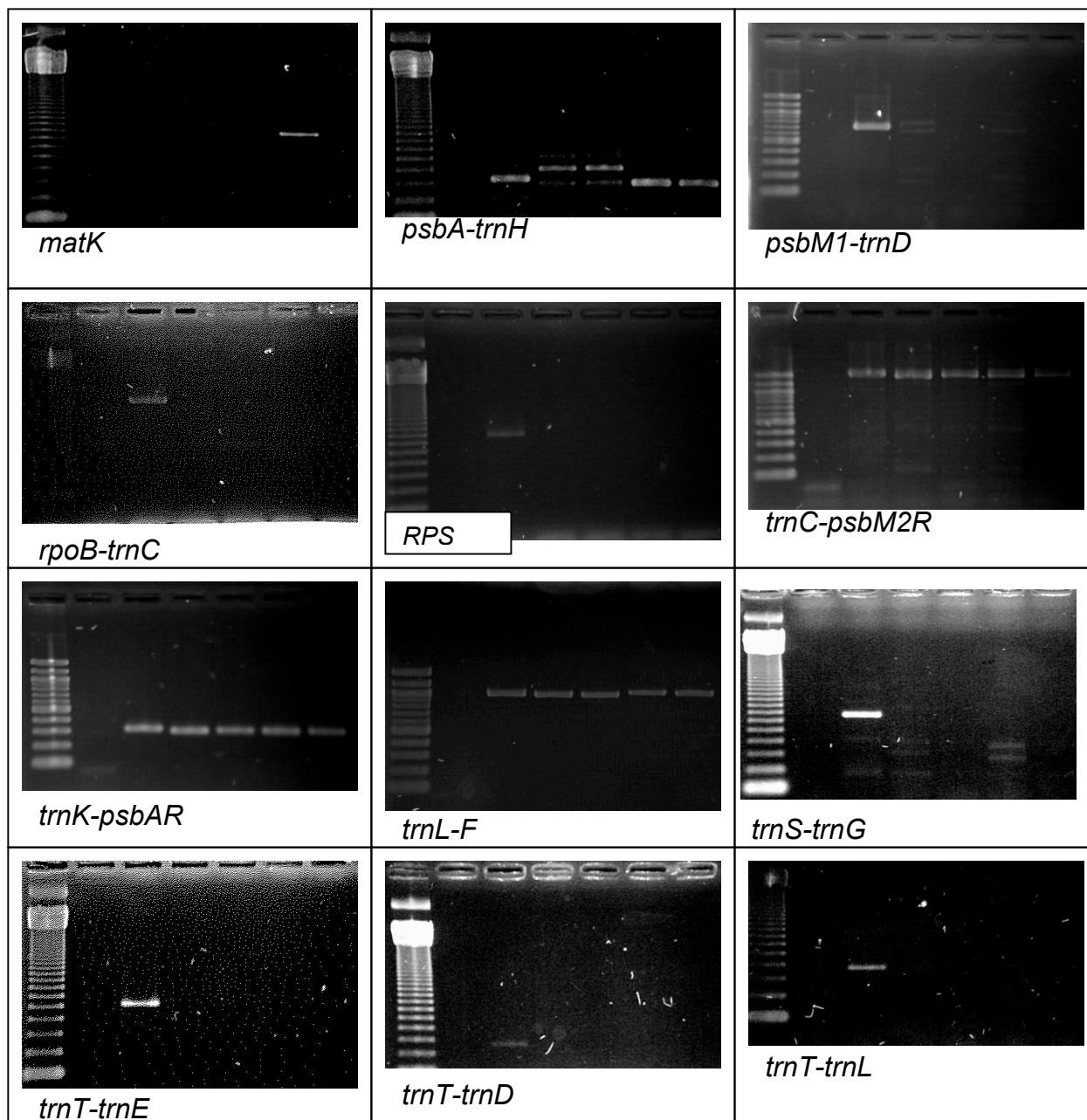
### DNA extractions and primer selection

Large amounts of high quality DNA were extracted from 99 samples of *Wahlenbergia*, *Campanula* and *Lobelia*. Nanodrop concentrations measured ranged from 72.6 ng/μl to 2005.4 ng/μl. Of the chloroplast markers trialled only four amplified across all four trial *Wahlenbergia* samples: *psbM1-trnD*, *psbA-trnH*, *trnL-F* and *trnK-psbAR* (gels in Figure 3.1, primer sequences and references Table 3.1). Of these *psbM1-trnD* and *psbA-trnH* failed to sequence cleanly. Both *trnL-F* and *trnK-psbAR* (hereafter referred to as *trnK*) amplified and sequenced readily, and contained some level of variation, so were selected for use in this study. Of the nuclear markers trialled, four amplified relatively cleanly across all four *Wahlenbergia* samples (Figure 3.2): the low-copy markers ‘Actin’, ‘atpB’ and Eif3E, and the nuclear ribosomal ITS. Note that ‘Actin’ and ‘atpB’ are tag names given to these primer pairs and do not refer the chloroplast regions commonly given these names (Steele et al., 2008). Both ‘Actin’ and ‘atpB’ sequenced relatively cleanly but contained only a limited amount of variation between sequences, mostly due to heterozygous positions, so were assessed as having too low a level of phylogenetic information to be useful. The third low-copy marker Eif3E could not be sequenced

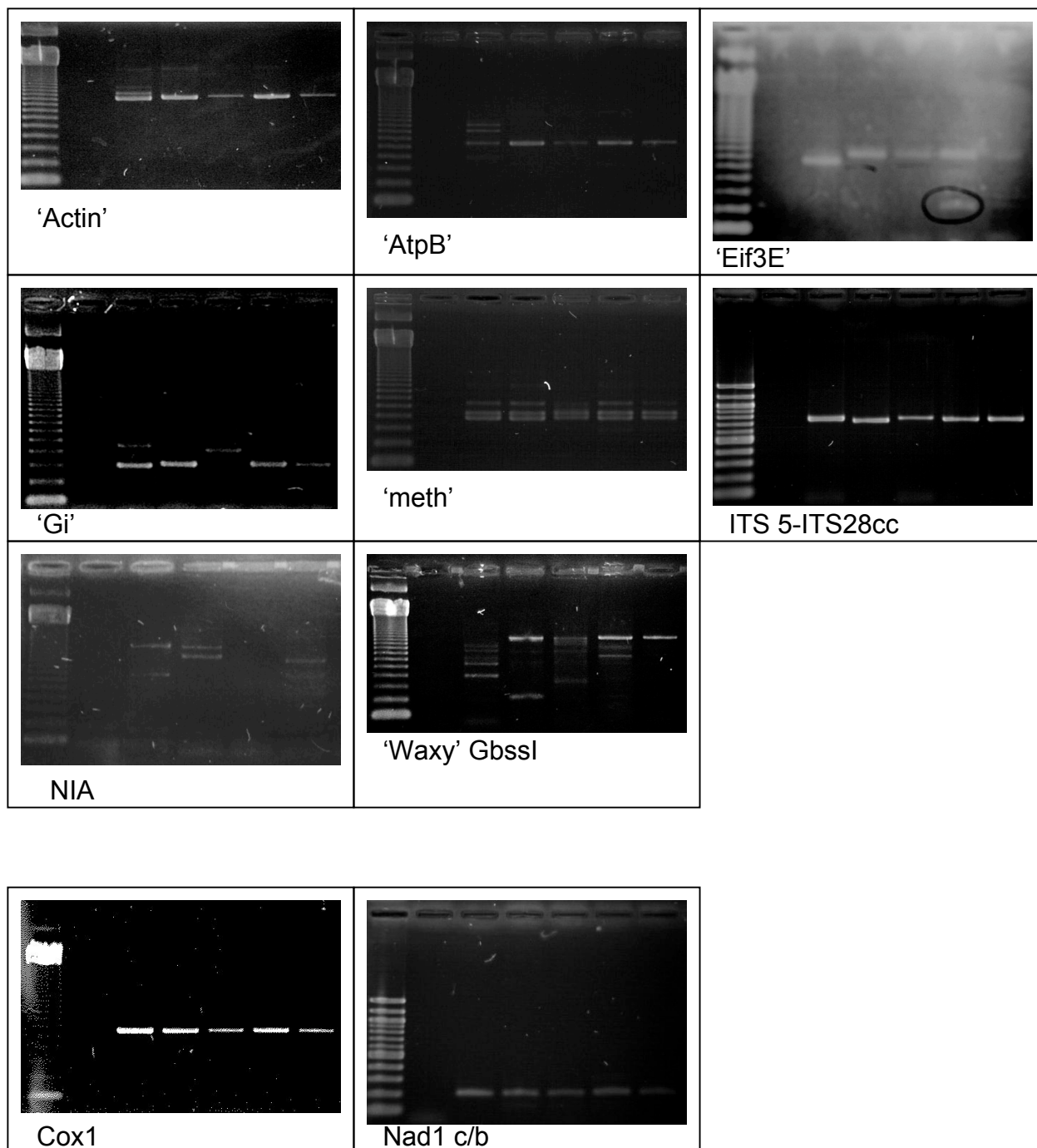
cleanly. The nuclear ribosomal ITS was a challenge to amplify and sequence and one base length indels were found to be present in a few samples, but was nevertheless used in this study. In some cases these sequences could still be used by obtaining the reverse sequence and creating contigs, but in other cases obtaining ITS sequences for certain samples had to be abandoned. Both of the mitochondrial markers trialled amplified and sequenced cleanly (Figure 3.2), but all four *Wahlenbergia* samples shared the identical haplotype for both markers, so were deemed of no use for phylogenetic analysis at the species level and were not pursued further.

### **aDNA extractions**

DNA was successfully extracted from six herbarium samples using the aDNA method outlined above (four samples of *W. littoricola* subsp. *littoricola* (LITT\_NSW\_Buff1-4), one of *W. preissii* (PREI\_WA\_Mund) and one of *W. tumidifructa* (TUME\_VIC\_Black). Because the samples were brought back to a modern lab for PCR and preparation for sequencing there is a chance they may have been contaminated with *Wahlenbergia* DNA present in the modern lab. This appears unlikely however, as three distinct chloroplast haplotypes were sequenced from the six samples, one of which (*W. tumidifructa*) had a unique haplotype in the dataset as a whole. Unfortunately no nrITS sequences were able to be generated for the aDNA samples, probably due to the small amount and lower quality of the DNA extracted compared to that extracted from more recently collected samples.



**Figure 3.1** Photographs of gels showing chloroplast primers trialled. The ladders are either 123 or 100bp ladders. The columns for each from left to right are ladder, – ve control, + control, *W. communis* (COMM\_NSW\_Quea), *W. violacea* (VIOL\_SI\_Dun), *W. albomarginata* subsp. *laxa* (LAXA\_SI\_Cobb1) and *W. albomarginata* subsp. *albomarginata* (ALBO\_SI\_Garv) except for *matK* in which the columns are reversed. Primer sequences and references are listed in Table 3.1. Explanation of tag names and voucher information can be found in Appendix 1.



**Figure 3.2** Photographs of gels showing nuclear (top) and mitochondrial (bottom) primers trialled. The ladders are either 123 or 100bp ladders. The columns for each are from left to right ladder, -ve control, + control, *W. communis* (COMM\_NSW\_Quea), *W. violacea* (VIOL\_SI\_Dun), *W. albomarginata* subsp. *laxa* (LAXA\_SI\_Cobb1) and *W. albomarginata* subsp. *albomarginata* (ALBO\_SI\_Garv). Primer sequences and references are listed in Table 3.2. Explanation of tag names and voucher information can be found in Appendix 1.



Table 3.1. Primer sequences and references for primer pairs included in this study

<b>Markers</b>	<b>Location</b>	<b>Primer name</b>	<b>Primer sequence (5'-3')</b>	<b>Reference</b>
<i>matK</i>	Chloroplast	390 F 1326R	CGA TCT ATT CAT TCA ATA TTT C TCT AGC ACA CGA AAG TCG AAG T	(Cuenoud et al., 2002)
<i>psbA-trnH</i>	Chloroplast	TrnHGUG	CGC GCA TGG TGG ATT CAC AAT CC	(Tate and Simpson, 2003)
<i>psbM1-trnD</i>	Chloroplast	PsbA psbM	GTT ATG CAT GAA CGT AAT GCT C GCG GTA GGA ACT AGA ATA AAT AG	(Sang et al., 1997) (Lee and Wen, 2004)
<i>rpoB-trnC</i>	Chloroplast	trnD RpoB trnC	GGG ATT GTA GTT CAA TTG GT CKA CAA AAY CCY TCR AAT TG CAC CCR GAT TYG AAC TGG GG	(Shaw et al., 2005)
<i>RPS f/r</i>	Chloroplast	tpsF tpsR2	GTGGTAGAAAGCAACGTGCGACTT TCGGGATCGAACATCAATTGCAAC	(Oxelman et al., 1997)
<i>trnC-psbM2R</i>	Chloroplast	trnC psbM2R	CCA GTT CAA ATC TGG GTG TC TTC TTG CAT TTA TTG CTA CTG C	(Lee and Wen, 2004)
<i>trnK-psbA</i>	Chloroplast	trnK3F psbA R	CCG ACT AGT TCC GGG TTC GAA TC CGC GTC TCT CTA AAA TTG CAG TCA T	(Winkworth et al., 2002)
<i>trnL-trnF</i>	Chloroplast	trnLf trnLc	ATT TGA ACT GGT GAC ACG AG CGA AAT CGG TAG ACG CTA CG	(Taberlet et al., 1991)
<i>trnS-trnG</i>	Chloroplast	trnSGCU trnG	AGA TAG GGA TTC GAA CCC TCG GT TTT TAC CAC TAA ACT ATA CCC GC	(Shaw et al., 2005)
<i>Trn-trnE</i>	Chloroplast	trnTr trnE	TAC CAC TGA GTT AA AGG GC GCC TCC TTG AAA GAG AGA TG	(Zimmer et al., 2002)
<i>trnT-trnD</i>	Chloroplast	trnTGGU trnDGCU-F	CTA CCA CTG AGT TAA AAG GG ACC AAT TGA ACT ACA ATC CC	(Shaw et al., 2005)
<i>trnT-trnL</i>	Chloroplast	trnTGGU	CTA CCA CTG AGT TAA AAG GG	(Shaw et al., 2005)

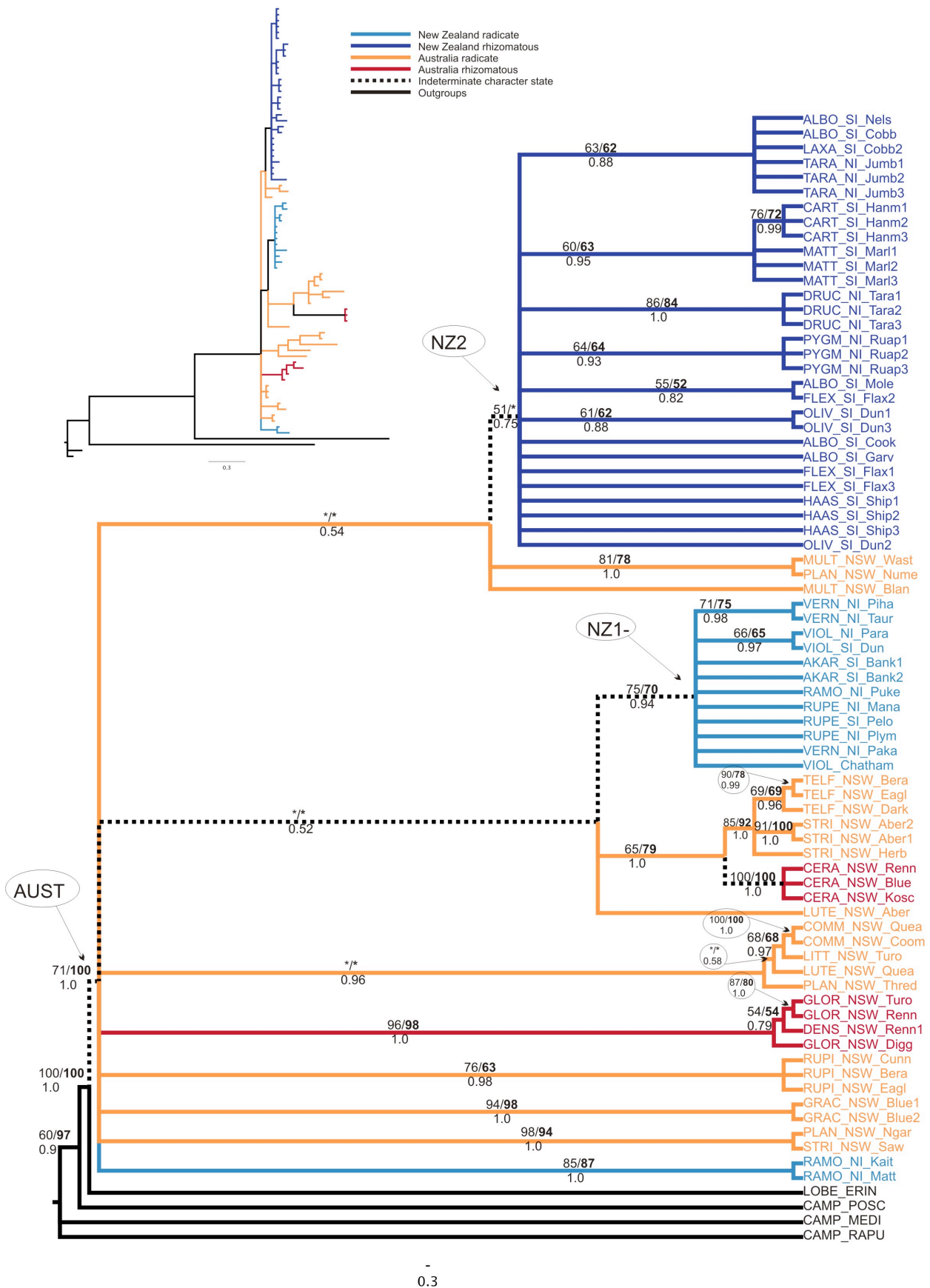
<b>Markers</b>	<b>Location</b>	<b>Primer name</b>	<b>Primer sequence (5'-3')</b>	<b>Reference</b>
Cox1	Mitochondrial	TrnLb cox1F4	TCT ACC GAT TTC GCC ATA TC	(Taberlet et al., 1991)
		cox1R3	GGATATCTAGGYATGGTTTATGC	(Cho et al., 2004)
Nad1 c/b	Mitochondrial	nad1b	AAGCTGGAGGACTTTGTAC	
		nad1c	GCATTACGATCTGCAGCTCA	(Demesure et al., 1995)
'Actin'	Nuclear	61F	GGAGCTCGATTAGTTTCTGC	
		61R	ATG GGA CAA AAA GAT GCT TA	(Steele et al., 2008)
'AtpB'	Nuclear	51F	TAG AAG CAC TTC CTG TGG AC	
		51R	CCT AGC TTG ATG ACA CCA C	(Steele et al., 2008)
'Eif3E'	Nuclear	Eif3E F	CTT GGA CGT ATC ATG AAT G	
		Eif3E R	ttt gaa tgt ggc aac tay tct rgt gct gc	(Li et al., 2008)
'Gi'	Nuclear	Gi F	acc tct tca cac tcy ytc atc tt	
		Gi R	tgg gct aca gat gca ctt ga	(Steele et al., 2008)
'Meth'	Nuclear	86 F	cac gca aga aat gca rat gca tcc a	
		86 R	ATG GTT GAG TAC TTT GGT GAG CAG	(Steele et al., 2008)
ITS	Nuclear	ITS28cc	GCC CAG TCC AAG TAG AA GC	
			CGC CGT TAC TAG GGG AAT CCT TGT AAG	(Wagstaff and Garnock-Jones, 1998)
NIA	Nuclear	ITS5	GGA AGT AAA AGT CGT AAC AAG AAG G	(White, 1990)
		NIA-i1F	TCG GAA AGC TAT TAT CAT TAC AAG G	(Levin et al., 2009)
		NIA-i2R	CCA TGT CTC TCC TCC ATC CA	
Waxy (GbssI)	Nuclear	B	CAC TGC TAT AAA CGT GGG GTT GA	(Peralta and Spooner, 2001)

## ITS analysis

The final ITS dataset contained 77 individuals (74 *Wahlenbergia* plus four outgroups) and was 816 characters long, 279 of which were variable and 133 parsimony informative. The model of evolution selected by jModelTest was GTR+G with base frequencies of variable sites: A = 17.6% C = 29.1% G = 31.4% T = 21.9%, scaled transitions and transversion substitutions in the dataset: AC = 1.3169 AG = 3.5376 AT = 2.0637 CG = 0.8530 CT = 1.1670 and a gamma shape parameter = 0.3879. The maximum parsimony (MP) analysis returned a tree with the best score of 400. The full number of most parsimonious trees was not recorded due to the sampling method, but was at least 10 000. The Bayesian analysis ran for 7 million generations, at which time the standard deviation of split frequencies was <0.01, which increases the likelihood that the runs had converged on a stationary distribution. The score of the best maximum likelihood (ML) tree was -3355.77.

The MP, Bayesian and ML methods returned 50% majority rule consensus trees with similar topologies, differing only in that some poorly supported clades evident in the Bayesian analysis were missing from the MP and ML trees. The (Australasian) *Wahlenbergia* samples formed a highly supported clade (1.0 posterior probability, 100% ML bootstrap, 100% MP bootstrap; hereafter 1.0/100/100 see Figure 3.3) separate from the outgroup samples (European *Wahlenbergia* species were not included in the ITS dataset). Within this clade relationships were poorly resolved, indicated by the large polytomy containing eight clades. Within this polytomy the clades that contained fewer samples were more highly supported than those containing more samples, which generally only had low support from the Bayesian analysis and were absent from the MP or ML trees.

The New Zealand species fell into three clades: first, a poorly supported one containing all of the New Zealand rhizomatous species (clade NZ2 0.75/<50/51); secondly, a clade with most of the New Zealand radicate species (clade NZ1-, 0.94/75/70, and thirdly a well-supported clade containing two (radicate) *W. ramosa* samples (1.0/85/87, Figure 3.3).



**Figure 3.3** Bayesian 50% majority rule consensus tree of *Wahlenbergia* (colours) and outgroup (black) based on the ITS dataset. Inset is the phylogram of the same tree. Numbers above branches are bootstrap support from the ML/MP analyses. Numbers below branches are posterior probability. AUST = Australasian *Wahlenbergia* clade, NZ1- = most samples of New Zealand species with a radicate growth form. NZ2 = New Zealand species with a rhizomatous growth form. See Appendix 1 for an explanation of the tag names and voucher information.

None of the New Zealand radicate species formed monophyletic groups within clade NZ1 (Figure 3.3), and samples from each species are either unresolved (e.g. *W. akaroa* and *W. rupestris*), paraphyletic (e.g. *W. littoricola* subsp. *vernica* and *W. violacea*) or polyphyletic (e.g. *W. ramosa*). Thus, relationships within the *W. gracilis* complex are unresolved, and even the sub-tropical *W. littoricola* subsp. *vernica* does not form its own lineage. Genetic distances among the New Zealand radicate species were low, as evidence by the p values (range 0 – 0.025, Table 3.2). Interestingly the Australian *W. littoricola* subsp. *littoricola* sample does not group with the New Zealand *W. littoricola* subsp. *vernica* samples.

The only New Zealand rhizomatous species to form a monophyletic group within clade NZ2 (Figure 3.3) was *W. cartilaginea* (0.99/76/72), although a few subspecies did too e.g. *W. albomarginata* subsp. *drucei* (1.0/86/84) and *W. pygmaea* subsp. *pygmaea* (0.93/64/64). The majority of species/subspecies relationships were either unresolved (e.g. *W. pygmaea* subsp. *tararua*, *W. matthewsii*, *W. congesta* subsp. *haastii*) or paraphyletic (e.g. *W. albomarginata* subsp. *albomarginata*, *W. albomarginata* subsp. *olivina*, *W. albomarginata* subsp. *flexilis*). Thus, while the subspecies of *W. albomarginata* and *W. pygmaea* often formed monophyletic groups the species themselves did not. Genetic distances among New Zealand rhizomatous species were also low, with a range in p distance values of 0 to 0.015 (Table 3.2).

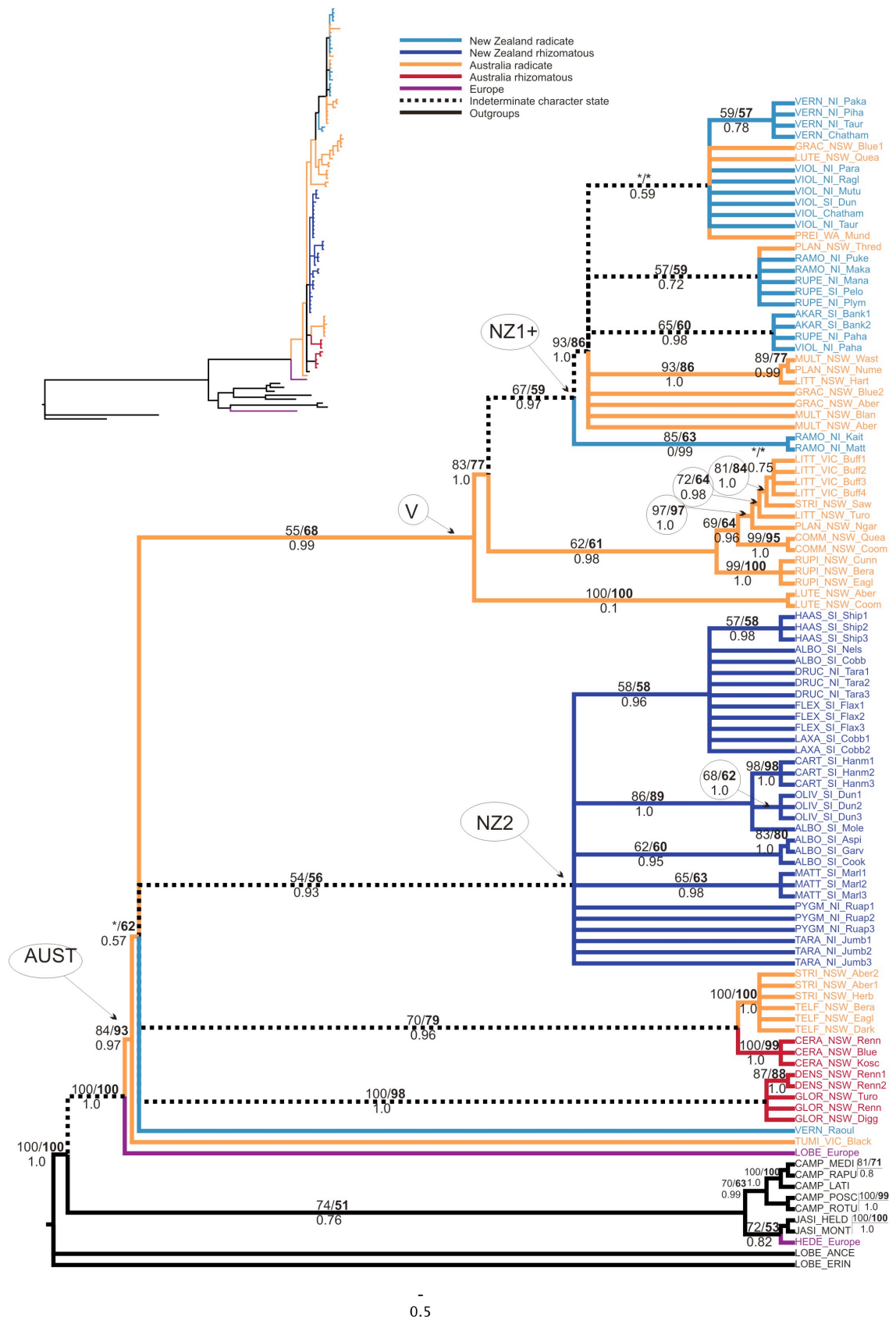
Although large-scale relationships of the Australian species were largely unresolved, in a number of cases all samples from the same species formed well-supported clades (e.g. *W. ceracea*, *W. communis*, *W. gracilis* and the two newly described Australian species *W. rupicola* and *W. telfordii*). However, the majority of samples from one species are either paraphyletic (e.g. *W. multicaulis* and *W. gloriosa*) or polyphyletic (e.g. *W. planiflora*, *W. luteola* and *W. stricta*). While the newly described *W. telfordii* formed a highly supported clade with the most of the samples of *W. stricta* and all of the samples of *W. ceracea*, the position of the newly described *W. rupicola* was unresolved. The Australian rhizomatous species fall in two separate clades, and there is no relationship between them and the New Zealand rhizomatous species. The genetic distances among Australian species was slightly higher than those among the New Zealand species, among the Australian rhizomatous species p distance values

ranged from 0 and 0.054 and among the Australian radicate species p distance values ranged from 0 and 0.065 (Table 3.2)

The neighbour net network (Appendix 2.1) revealed high levels of conflicting signal in the ITS dataset, particularly regarding the Australian radicate samples. Also, as expected due to their low genetic variation for the ITS marker, the New Zealand rhizomatous and radicate species have star-like topologies in the network.

### **Chloroplast markers analysis**

The final chloroplast dataset contained 105 individuals (96 *Wahlenbergia* plus 9 outgroups) and was 1131 characters long, 889 characters from the *trnL-F* marker and 242 characters from the *trnK* marker. There were 370 variable characters, 241 of which were parsimony informative. The 10 samples that were sequenced only for either *trnL-F* or *trnK* were coded as missing data in the concatenated alignment. The model of evolution selected by jModelTest was GTR+G with base frequencies of variable sites A = 32.6% C = 17.4% G = 18.7% T = 31.3%, scaled transitions and transversion substitutions in the dataset: AC = 1.0481 AG = 0.6020 AT = 0.1343 CG = 1.2275 CT = 0.8801 and a gamma shape parameter = 0.9499. The parsimony analysis returned a tree with the best score of 531. The full number of most parsimonious trees was not recorded due to the sampling method, but was at least 10 000. The Bayesian analysis ran for 7 million generations, by which time the standard deviation of split frequencies was below the 0.01 recommended. The score of the best ML tree was -4575.56. The MP, Bayesian and ML methods returned 50% majority rule consensus trees with very similar topologies, differing only in that the ML tree contained one clade not supported by the MP or Bayesian trees (data not shown, discussed further below) and that one poorly supported clade evident in the Bayesian analysis was missing from the MP and ML trees. The *Wahlenbergia* samples (excluding *W. hederacea* (HEDE\_Europe), the placement of which is discussed in more detail in Chapter Two of this thesis) formed a highly supported clade (1.0/100/100) separate from the outgroup samples (Figure 3.4). Within this *Wahlenbergia* clade, the European *W. lobelioides* (LOBE\_Europe) and the Victorian *W. tumidifructa* (TUME\_VIC\_Black) are successively sister to the remainder



**Figure 3.4** Bayesian 50% majority rule consensus tree of *Wahlenbergia* (colours) and outgroup (black) based on the *trnL-F* + *trnK* dataset. Inset is the phylogram of the same tree. Numbers above branches are bootstrap values (%) from the ML/MP analyses. Number below branches are pp values. AUST = Australasian *Wahlenbergia* clade, NZ1+ = (mostly) New Zealand species with a radicate growth form. NZ2 = New Zealand species with a rhizomatous growth form. V = clade VERN\_Raoul is sister to in ML analysis. See Appendix 1 for an explanation of the tag names and voucher information.

of the Australasian *Wahlenbergia*. Within this main Australasian clade (0.97/84/93), the topology and support values are similar to the ITS tree in that relationships between species are generally poorly resolved, even though some species/subspecies are resolved as monophyletic.

Similar to the ITS tree the New Zealand radicate species do not form a monophyletic group, although all but one of them fall into the same clade (labelled NZ1+ in Figure 3.4 0.97/67/59). Like the ITS tree, none of the New Zealand radicate species within this clade are themselves monophyletic. *Wahlenbergia akaroa* is unresolved, and the remainder (i.e., *W. littoricola* subsp. *vernica*, *W. violacea*, *W. ramosa* and *W. rupestris*) are all polyphyletic. The *W. littoricola* subsp. *vernica* samples mostly fall in a moderately supported (0.78/5759/) clade of their own and are polyphyletic only due to the placement of the *W. littoricola* subsp. *vernica* sample collected from Raoul Island (VERN\_Raoul), the position of which is unresolved. In the MP and Bayesian analysis this sample formed part of a large unresolved polytomy, yet in the ML tree it is grouped as sister to clade V in Figure 3.4, with 57% ML bootstrap support. It is clear the Australian *W. littoricola* samples do not form a clade with their proposed conspecific New Zealand *W. littoricola* subsp. *vernica* samples however, as these all fall in a separate clade. Genetic distances among the New Zealand radicate species were low, as evidenced by p distances (range 0 – 0.008 Table 3.2).

Like the ITS tree, the New Zealand rhizomatous species form a clade (NZ2 0.96/54/56) and three species are resolved as monophyletic (i.e., *W. cartilaginea* (1.0/98/98), *W. congesta* subsp. *haastii* (0.98/57/58) and *W. matthewsii* (0.98/65/63)). Additionally a few subspecies grouped in clades (e.g. *W. albomarginata* subsp. *olivina*, *W. pygmaea* subsp. *drucei*, *W. pygmaea* subsp. *pygmaea* and *W. pygmaea* subsp. *tararua*). The remainder were either unresolved (*W. albomarginata* subsp. *flexilis* and *W. albomarginata* subsp. *laxa*,) or polyphyletic (*W. albomarginata* subsp. *albomarginata*). Genetic distances are also low among the New Zealand rhizomatous species (p distances range from 0 to 0.009 Table 3.2).



Although large-scale relationships of the Australian species were (like the ITS tree) largely unresolved in the chloroplast tree, in a number of cases all samples from the same species formed well-supported clades on their own (e.g. *W. ceracea* (1.0/100/99), *W. densifolia* (1.0/87/88), *W. communis* (1.0/99/95) and the newly described species *W. rupicola* (1.0/99/100)) or together with another species (e.g., *W. gloriosa* and *W. densifolia* (1.0/100/98), or *W. telfordii* and *W. stricta* (1.0/100/100)). However, the majority of samples from one species are either unresolved (e.g. *W. telfordii*), paraphyletic (e.g. *W. gloriosa*, *W. gracilis* and *W. multicaulis*) or polyphyletic (e.g. *W. luteola*, *W. planiflora*, *W. littoricola* subsp. *littoricola* and *W. stricta*) within the large clade comprising most of the NZ and Australian radicate species. Interestingly the four Australian *W. littoricola* subsp. *littoricola* samples extracted from herbarium specimens (LITT\_NSW\_Buff1-4) formed a highly supported clade, which also includes a sample identified as *W. stricta* (STRI\_NSW\_Saw) and an additional recently collected *W. littoricola* (LITT\_NSW\_Turo) sample. The voucher specimen for the *W. stricta* sample is rather lacking however (it is a poor specimen with few leaves and only one shrivelled flower), and I am not confident in my identification (i.e. there is a chance it could in fact be a *W. littoricola*). A sample identified as *W. planiflora* (PLAN\_NSW\_Ngar) is sister to this clade, and the identification of this specimen at least as not a *W. littoricola* is believable, as its flowers appear rotate (lacking a corolla tube), whereas *W. littoricola* flowers are generally shortly campanulate (Smith, 1992). The other sample identified as *W. littoricola* in this dataset (LITT\_NSW\_Hart), which does not group with the other samples mentioned above, is another recent collection, and again has a rather poor voucher specimen (few leaves, solo flower), making confirmation of the identification impossible. As in the ITS analysis the Australian species exhibit greater genetic variation in the chloroplast sequences than the New Zealand *Wahlenbergia* based on p distance values (a range of 0 to 0.025 among the Australian radicates, and 0 to 0.014 among the Australian rhizomatous Table 3.2).

The neighbour net network (Appendix 2.2) revealed less conflicting signal than the ITS neighbour net network, though conflict still is present. The main split in the network separates the majority of the NZ and Australian radicate species (excluding *W. luteola*, *W. stricta*, *W. telfordii*, *W. tumidifructa* and *W. vernicosa* from Raoul

Island) from the rhizomatous species, with two samples of *W. luteola* exactly intermediate between these groups.

### Analysis of concatenated dataset

The homogeneity partition test showed the nuclear vs. chloroplast alignments to be significantly incongruent ( $p=0.01$ ). However, this test can be highly inaccurate even when the topologies of trees are congruent (Reeves et al., 2001; Yoder et al., 2001) although in this case visual assessment of the topologies does reveal a level of incongruence. Most of the incongruence appears to be “soft” incongruence, meaning that a particular relationship was resolved in one dataset but not in the other. An example of this is the *W. littoricola* subsp. *vernica* samples all of which are in the New Zealand radicate clade in the ITS

**Table 3.2** Uncorrected p distances calculated from the ITS and combined *trnL-F+trnK* datasets. Readings for the New Zealand radicate group were calculated excluding the *W. vernica* sample from Raoul Island.

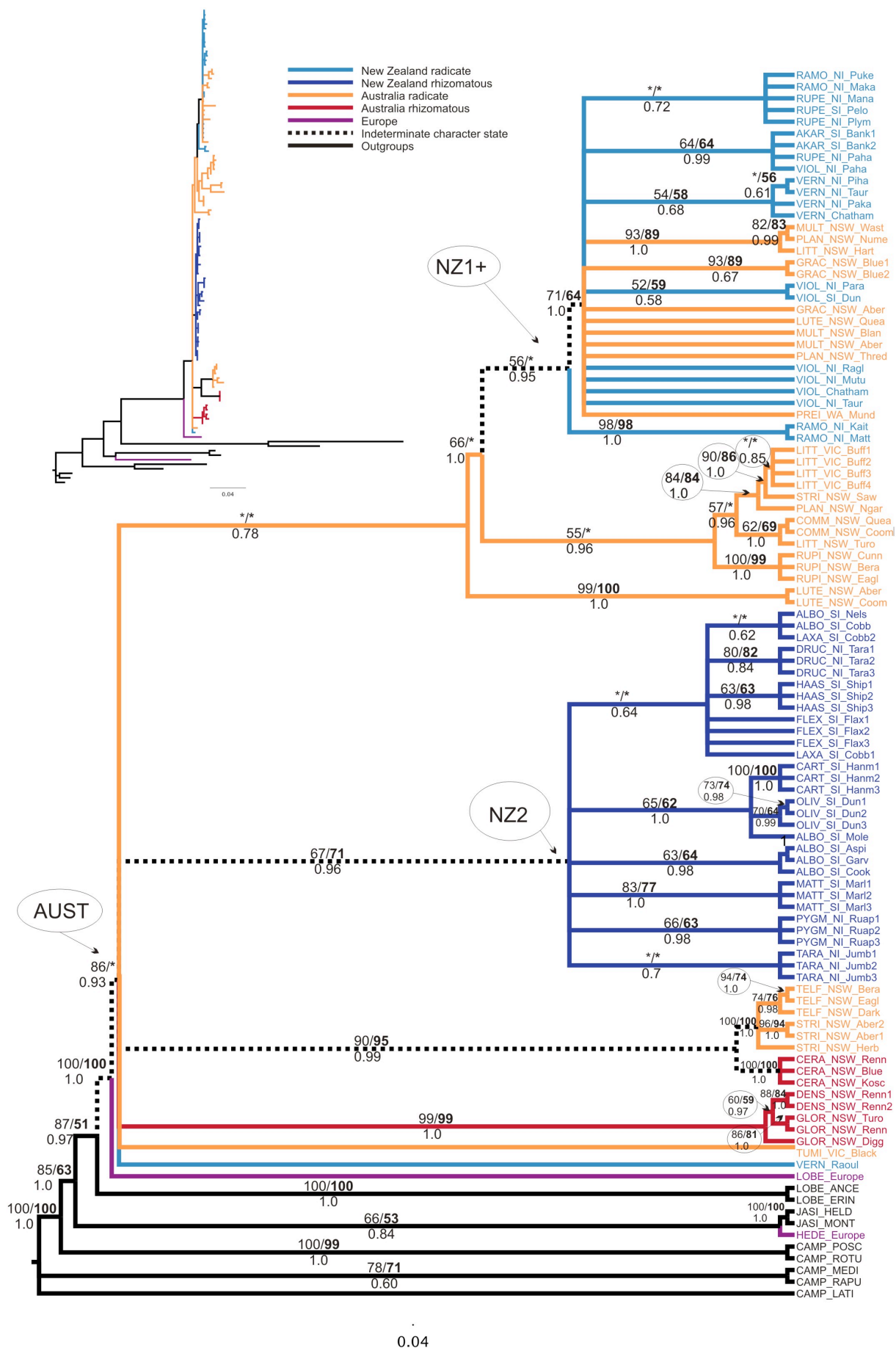
	Species group	Average uncorrected p distance	Standard deviation	Minimum uncorrected p distance	Maximum uncorrected p distance	Samples with maximum uncorrected p distance
ITS	New Zealand radicate	0.008	0.006	0	0.025	RAMO_NI_Puke RUPE_SI_Pelo
	New Zealand rhizomatous	0.004	0.003	0	0.015	CART_SI_Hamn1 OLVE_SI_Dun2
	Australian radicate	0.029	0.014	0	0.065	LUTE_NSW_Quea STRI_NSW_Aber2
	Australian rhizomatous	0.029	0.023	0	0.054	CERA_NSW_Renn GLOR_NSW_Renn
<i>trnL-F</i> + <i>trnK</i>	New Zealand radicate	0.003	0.002	0	0.008	RAMO_NI_Kai VERN_Chatham
	New Zealand rhizomatous	0.003	0.002	0	0.009	ALBO_SI_Aspi CART_SI_Hanm3
	Australian radicate	0.013	0.006	0	0.025	LUTE_NSW_Quea STRI_NSW_Herb
	Australian rhizomatous	0.008	0.006	0	0.014	CERA_NSW_Blue DENS_NSW_Renn1

tree (clade NZ1- in Figure 3.3), and two of the three samples form a moderately supported clade within this. In contrast all of the *W. littoricola* subsp. *vernica* samples form a moderately supported clade in the chloroplast tree (Figure 3.4)

(excluding the Raoul Island sample which was not included in the ITS dataset). A rare example of a highly supported “hard” incongruence between the datasets is the relationship of the *W. cartilaginea* samples to the *W. matthewsii* samples. In the ITS tree a clade containing samples from these species has moderate support (0.95/60/63), but this clade is absent from the *trnL-F* + *trnK* tree, where instead a clade containing the *W. cartilaginea* samples, the *W. albomarginata* subsp. *olivina* samples and one *W. albomarginata* subsp. *albomarginata* sample has high support (1.0/86/89). The neighbour net network (Appendix 2.3) of the concatenated dataset also revealed a high level of conflicting signals within the dataset, though this contained conflict within the individual ITS and chloroplast datasets, and it is therefore difficult to determine the source of the conflict. I nevertheless proceeded with analysis of the concatenated dataset, with the caveat that all results should be treated with caution due to the incongruence discussed above.

### **Combined *ITS*, *trnK* and *trnL-F* analysis**

The final concatenated dataset contained 105 samples and was 1947 characters long, 1131 (58%) from the chloroplast and 816 (42%) from the nuclear genome. There were 649 variable characters and 374 parsimony informative ones. The majority of the variable characters were from the chloroplast dataset (370, 57%), though this was in proportion with the greater number of characters contributed by the chloroplast dataset. The majority of the parsimony informative characters were also from the chloroplast data (64%), which was a slightly higher proportion than its contribution of characters. The 28 samples that were sequenced only for chloroplast markers were coded as missing data for the ITS characters in the concatenated alignment. It was not necessary to estimate parameters independently between data partitions as both the nuclear and chloroplast datasets were best estimated by the same model of evolution (GTR+G). The parsimony analysis returned a tree with the best score of 1005. The full number of most parsimonious trees was not recorded due to the sampling method, but was at least 10 000. The Bayesian analysis ran for 10 million generations, by which time the standard deviation of split frequencies was at 0.02, which is higher than the point recommended to stop the analysis, but it seemed unlikely to drop below the threshold of 0.01 within an acceptable timeframe. The score of the best ML tree was -7947.01.



**Figure 3.5** Bayesian 50% majority rule consensus tree of *Wahlenbergia* (colours) and outgroup (black) based on the combined ITS + *trnL-F* + *trnK* dataset. Inset is the phylogram of the same tree. Numbers above branches are bootstrap values (%) from the ML/MP analyses. Number below branches are pp values. AUST = Australasian *Wahlenbergia* clade, NZ1+ = (mostly) New Zealand species with a radicate growth form, NZ2 = New Zealand species with a rhizomatous growth form. See Appendix 1 for an explanation of the tag names and voucher information.

The MP, Bayesian and ML methods returned 50% majority rule consensus trees with relatively similar topologies, differing in that a number of moderately and poorly supported clades evident in the Bayesian analysis were missing from the MP and ML trees, and one clade was present in the MP tree but absent from the Bayesian and ML analyses (*W. tumidifructa* was placed as sister to all other Australasian *Wahlenbergia* with 53% bootstrap support, data not shown). The *Wahlenbergia* samples (again excluding *Wahlenbergia hederacea*) formed a highly supported clade (1.0/99/99) separate from the outgroup samples (Figure 3.5). As in the chloroplast tree, within this *Wahlenbergia* clade the European *W. lobelioides* (LOBE\_Europe) and the Victorian *W. tumidifructa* (TUME\_VIC\_Black) were sister to the remainder of the Australasian *Wahlenbergia* (only supported by 54% bootstrap support in the MP analysis, data not shown). In general the areas of incongruence between the ITS and chloroplast trees have been resolved in favour of the relationships present in the chloroplast tree, probably because the chloroplast tree was reconstructed using a greater number of characters (and higher percentage of the parsimony informative characters) and generally had slightly higher support values than the ITS tree. Unsurprisingly the pattern of low resolution at the among-species level, present in both the nuclear and chloroplast trees, is repeated in the concatenated tree. More species/subspecies are recovered as monophyletic in the concatenated tree compared to either of the nuclear or chloroplast trees though, as information resolving different species/subspecies was present in the different trees.

## Discussion

The central aim of this chapter was to generate the first phylogeny of the Australasian *Wahlenbergia*, and to use this to test the current taxonomy. The phylogenies presented here are poorly resolved, which limits their usefulness in the role of exploring relationships within *Wahlenbergia*. Nevertheless some parts of the trees, generally more towards the tips of the branches, are resolved giving some information at the species level.

## Evolution of Australasian *Wahlenbergia*

In the previous chapter in this thesis the Australasian *Wahlenbergia* were shown to

be monophyletic, and two introductions from Australia to New Zealand of *Wahlenbergia* were indicated, one leading to the rhizomatous radiation, and the other to the radicate radiation. With greater sampling of both Australian and New Zealand individuals in this chapter that clear pattern is somewhat obscured. It still appears most likely that the New Zealand rhizomatous species are the result of a single introduction from Australia, but there is no support for Smith's (1992) hypothesis that the New Zealand rhizomatous species are derived from the Australian rhizomatous species. In fact, it is unclear which Australian species they may be most closely related to as there is a general lack of support and resolution in the backbone of the Australasian *Wahlenbergia* phylogenies presented here (Figures 3.3-3.5). Unfortunately this study did not include any samples of the Tasmanian rhizomatous *W. saxicola*, which was even considered conspecific with the New Zealand rhizomatous species at one time, and this may represent the sister species to the New Zealand rhizomatous clade.

Additionally, and in contrast to the findings of Chapter Two of this thesis, there is also no clear New Zealand radicate clade. All the New Zealand radicate species appear to be very closely related to each other and to certain Australian radicate species, particularly *W. gracilis*, *W. multicaulis* and *W. planiflora*. In fact, some individuals from different species from Australia and New Zealand have identical chloroplast haplotypes (e.g. the Australian *W. gracilis* (GRAC\_NSW\_Blue1) and a number of the New Zealand *W. violacea* samples (e.g. VIOL\_NI\_Para). This low level of sequence variation suggests recent evolution, which was also indicated in the last chapter where I found the New Zealand radicate *Wahlenbergia* had radiated about 0.5 mya. Additionally there is very low sequence divergence within the New Zealand radicate species as indicated by low average uncorrected p distances (e.g. 0.003 for the chloroplast dataset and 0.008 of the nrITS).

Low genetic diversity exhibited by a range of chloroplast and nuclear markers has been found in a number of other New Zealand plant genera, which have also probably radiated within the last 5 my (e.g. *Plantago* (Tay et al., 2010), *Myosotis* (Winkworth et al., 2002a) and *Ranunculus* (Lockhart et al., 2001)). Recent evolution, combined with the geological and climatic conditions during the last 5 my, has been implicated in causing this low level of diversity. Species with conspicuous

morphological and ecological differences exhibit little or no genetic differentiation with both nuclear and chloroplast loci, suggesting that differentiation has occurred over a relatively short period of time. Examples of recent diversification in New Zealand include the Gnaphalieae (Smitsen et al., 2004), *Ourisia* (Meudt and Simpson, 2006), *Ranunculus* (Lockhart et al., 2001) and the hebe complex (Wagstaff and Garnock-Jones, 1998, 2000).

It can be very challenging to identify species in recently evolved/ evolving lineages, as insufficient time may have passed for reproductive isolation and gene-tree monophyly to have become fixed (Shaffer and Thomson, 2007). Furthermore, Hudson and Coyne (2002) demonstrated that reciprocal monophyly for a pair of species takes from 4 to 7  $N_e$  generations for 50% of nuclear genes sampled (where  $N_e$  is the historically effective population size of each descendant taxon) to 9 to 12  $N_e$  generations for 95% of sampled genes. Historical and current population sizes are usually unknown, but it seems likely that reciprocal monophyly for most genes will require from tens of thousands to millions of generations, implying a long waiting time for monophyly to evolve even in completely isolated lineages (Shaffer and Thompson, 2007). As a recently diverged lineage in which hybridisation is known (Smith 1992; Pettersson, 1997b) species of Australasian *Wahlenbergia* may not have evolved reciprocal monophyly. Thus, despite the lack of monophyly for a number of currently accepted species of Australasian *Wahlenbergia* in this chloroplast and nuclear DNA sequence analysis, other lines of evidence (e.g. morphological, cytological etc) should be considered before the species are rejected.

## **Testing the New Zealand *Wahlenbergia* taxonomy**

### *New Zealand radicate species*

The New Zealand radicate species are not recovered as a monophyletic group in any of the analyses (MP, ML and Bayesian) for either of the datasets (chloroplast or nuclear) and none of the species are recovered as monophyletic. The relationships among samples from the *W. gracilis* complex are unresolved in all analyses. The ITS tree is extremely poorly resolved with regards the *W. gracilis* complex and the chloroplast tree is not much better. The only clade that receives high support in both the ITS and chloroplast trees is that of two *W. ramosa* individuals (RAMO\_NI\_Kait

and RAMO\_NI\_Matt) that is separate from the remaining New Zealand radicate species. These two individuals were collected from hills in the Wellington region; there are no other obvious morphological or geographical characters that distinguish them.

None of the temperate radicate species (the *W. gracilis* complex) therefore have evolved monophyly according to three DNA sequence markers, but does this mean they shouldn't be considered species? Recall that the morphological differences used to separate them are largely based on floral differences regarding colour, size, and/or degree of petal overlap (Pettersen, 1997b). In addition, all have the same chromosome number  $2n=72$ , and all except *W. akaroa* have largely overlapping distributions on both the North and South Islands. Thus, the morphological differences are not great and are often difficult to score, and there is no additional genetic, cytological, or ecological data to suggest there are four separate lineages within the *W. gracilis* complex. The ability of several of the *W. gracilis* complex to hybridise further complicates the picture. Plants that are hybridising, even if they are in independent lineages, will take even longer to evolve monophyly than those that have developed complete reproductive isolation (Shaffer and Thompson 2007). Information about hybridisation among the *W. gracilis* complex has so far been restricted to observational studies, recordings of purported hybrid swarms based on field observations of morphologically intermediate individuals growing within proximity of 'pure' populations of the parent species (Pettersen, 1997b). If extensive hybridisation and gene flow are occurring, it could provide additional evidence that this complex is effectively a single lineage, and recognition of species within this lineage is therefore not appropriate. It would be useful to undertake controlled cross-pollinations in a glass house setting to determine whether hybrids do in fact set viable seed, and whether there is any difference in the viability of seeds formed from parents of the same species vs. different species. This is key to whether there are any barriers to cross-pollination between these four radicate species, which would help to determine their taxonomic status.



### *The position of W. littoricola subsp. vernicosa*

*W. littoricola* subsp. *vernicosa*, the one New Zealand radicate species to have a different chromosome count from the other New Zealand radicate species ( $2n=54$  vs.  $2n=72$  (Pettersen et al., 1995)), is found in subtropical areas of northern New Zealand, the Chatham Islands, plus several islands in the Pacific (Pettersen, 1997b). The phylogenetic analysis conducted in this study does not support the reduction of *W. vernicosa* to a subsp. of *W. littoricola* (de Lange and Cameron, 1999). Evidence to support the subspecies rank could have included samples from the two subspecies forming a monophyletic group, or given the lack of resolution at least samples from both subspecies belonging to the same unresolved grade. However, in both the chloroplast and ITS trees the *W. vernicosa* samples are in a quite separate clade from the *W. littoricola* samples. Both the chloroplast and ITS trees resolve other Australian and New Zealand species as being more closely related to the New Zealand *W. vernicosa* individuals than the Australian *W. littoricola*, though different species are implicated in each case.

In this analysis, *W. vernicosa* is polyphyletic because the sample collected from Raoul Island does not group with the remainder of the samples, which do form a monophyletic group in the chloroplast tree. Due to the uni-parental inheritance of chloroplast DNA the chloroplast tree may not reflect the species tree. Furthermore chloroplast trees have been shown to reflect geographic locations rather than species relationships in a number of cases in the New Zealand flora (e.g. *Metrosideros* (Gardner et al., 2004)). This does not appear to be the case for *Wahlenbergia* as the samples that form a monophyletic group are themselves from a range of locations (Chatham Island, Tauranga, Piha and Pakahi Island in the Hauraki Gulf see Appendix 1), which suggests the Raoul Island individual may be a different species altogether. Unfortunately no ITS sequence was generated for the Raoul Island sample identified as *W. vernicosa*, which could have given an independent confirmation of the placement on the chloroplast tree.

Another point to consider is that the two *W. vernicosa* samples that have had their chromosomes counted were from Northland: Surville Cliffs and Three Kings Islands (Pettersen et al., 1995). Detailed cytological studies of *W. vernicosa* from across its

range, including Raoul Island and other Pacific islands, would help to determine whether different chromosome races are present within this species, and even if similar ploidy levels are found, whether these polyploids have originated from different putative parental species (e.g. in *Asplenium* (Shepherd et al., 2008)). Finally, inclusion of more individuals of this species from throughout its range in the molecular phylogeny would also help determine whether Raoul Island and other Pacific island individuals are conspecific with New Zealand *W. vernicosa*.

In summary, the molecular phylogenetic data reported here provide further evidence to support the separation of *W. vernicosa* from the other radicate species of *Wahlenbergia* in New Zealand, and its recognition as a species in its own right rather than a subspecies of the Australian *W. littoricola*. However, there is no evidence to support the current taxonomy of the other four radicate species. This could indicate that the *W. gracilis* complex should be considered one species, perhaps with varieties to accommodate the morphological variation recorded by Petterson (1997b). Further research using methods able to detect variation at a smaller scale may be able to shed some light on their relationships.

### *New Zealand rhizomatous species*

The New Zealand species with a rhizomatous growth form (both creeping and suffructose) are recovered as a monophyletic group in all of the datasets (chloroplast, ITS and concatenated) though not for all types of analyses (e.g. the ITS MP analysis does not support a NZ rhizomatous clade). Within this clade the most morphologically distinctive New Zealand species, *W. cartilaginea* (Figure 1.2), is recovered as monophyletic in both the ITS and chloroplast datasets. The specimens of *W. matthewsii* and *W. congesta* subsp. *haastii* are recovered as monophyletic in the chloroplast data, but not in the nuclear data. These two species are also morphologically distinctive; *W. matthewsii* even has a different growth form (suffructose rhizomatous) that distinguishes it from all other New Zealand *Wahlenbergia* (Figure 1.1). Unfortunately I was unable to sample any *W. congesta* subsp. *congesta* individuals to test the status of subspecies in this species. Morphologically the subspecies are very similar, except that *W. congesta* subsp. *haastii* has smaller capsules and flowers (Petterson, 1997b). There is evidence to

suggest that individuals of these morphologically diverse species (*W. cartilaginea*, *W. matthewsii* and *W. congesta*) fall into genetically distinct lineages, which coupled with their morphological differentiation suggests they are appropriately circumscribed at the species level.

Conversely, neither of the other two rhizomatous species, *W. albomarginata* and *W. pygmaea* is monophyletic in either dataset. At the species level there is no DNA sequence evidence to separate *W. albomarginata* from *W. pygmaea*, and as there are few morphological characters that separate them (*W. albomarginata* has a narrow-campanulate-rotate corolla, whereas in *W. pygmaea* the corolla is broadly campanulate (Pettersen, 1997b)), recognition of these two species may be an artificial split between North and South Island individuals of what is essentially the same species. The lack of any genetic differences separating individuals from the North vs. South Islands could suggest that Cook Strait is not a barrier to gene flow between these individuals, or at least that a periodically emerging land bridge across Cook Strait during the Pleistocene may have facilitated gene-flow (Lewis et al., 1994). If further study indicates that only one species should be recognised, then *W. albomarginata* has priority.

At subspecies rank, in some cases individuals from the same subspecies are grouped together, though this is not evident in subspecies from which individuals were collected across a wider geographic range (e.g. *W. albomarginata* subsp. *albomarginata*) and therefore may be an artefact of the sampling, in that individuals from the same subspecies are often also from the same population due to sampling constraints. Of the four *W. albomarginata* subspecies sampled for this study, subsp. *olivina* is the only one to form a monophyletic group, and this only occurs in the chloroplast data (two individuals are grouped in the ITS data, the position of the third individual is unresolved). *Wahlenbergia albomarginata* subsp. *olivina* is the most morphologically distinctive subspecies, and is able to consistently be identified based on morphological characters (thick white leaf margins, leaves dark green and hairy on top, purplish and glabrous below), and edaphic characters (it only grows on ultramafic rock) (Pettersen, 1997b). The evidence of its monophyly, although limited from these analyses, suggests that it is an independent lineage, and that perhaps a subspecies rank is justified in this case. Further exploration may even reveal that an

elevation to species rank is justified. There is no evidence that subspecies rank is justified for the other subspecies of *W. albomarginata* based on this DNA sequence data, though given the recent evolution of this group it may be unrealistic to expect genetic differentiation among them.

The individual subspecies of *W. pygmaea* are unresolved by the chloroplast data, though the samples from Mt Taranaki (subsp. *drucei*) are in a separate clade from the individuals from the other two subspecies. In the ITS data the Mt Taranaki samples form a well-supported clade of their own (1.0/86/84) and to a lesser extent so do the samples from the central plateau (*W. pygmaea* subsp. *pygmaea* 0.93/64/64). The samples from the Tararua ranges however are part of a clade also containing *W. albomarginata* subsp. from the north west of the South Island. Although this gives some support for the monophyly of the different subspecies of *W. pygmaea* the conflicting signal from the nuclear and chloroplast datasets makes it difficult to interpret.

## **Notes on the Australian *Wahlenbergia* taxonomy**

### ***Australian radicate species***

Although not so extensively sampled as the New Zealand species, some interesting results relating to the Australian *Wahlenbergia* are presented here. The Australian radicate species do not form a monophyletic clade, though unlike the New Zealand radicates several do form monophyletic species. Species that form monophyletic groups include the two newly described species (discussed further below), *W. communis* and *W. gracilis* (nuclear data only). However, the majority of Australian radicate species sampled for this study do not form monophyletic groups, which could be due to several reasons such as incorrectly identified samples, incorrectly circumscribed species, hybridisation and inappropriate choice of molecular markers (Shaffer and Thomson, 2007). Incorrect identification is a real risk when working with this genus, as the characters distinguishing most of the radicate species are quite precise floral characters, which can be a challenge to apply in the field e.g. ratio of tube to lobe in flowers that are less than one centimetre long. Another factor to consider with the Australian radicate species is that they have mostly been shown to

contain at least two races of ploidy levels, e.g. the polyphyletic *W. luteola*, which actually morphologically is rather distinctive (opposite linear leaves, bi-colour flowers) has had the chromosomes counted for 10 individuals, three of which had  $2n=18$ , and 7 had  $2n=36$  (Smith 1992). It is therefore not unimaginable that the *W. luteola* sample which appears genetically quite different from the other two perhaps has a different ploidy level. Furthermore Smith (1992) reports that hexaploids ( $2n=54$ ) hybridise more readily than the diploids ( $2n=18$ ) or tetraploids ( $2n=36$ ) do. This could potentially explain the peculiar grouping of samples of *W. multicaulis* - (MULT\_NSW\_Wast), *W. planiflora* (PLAN\_NSW\_Nume) and *W. littoricola* (LITT\_NSW\_Hart) found in both the nuclear and chloroplast trees (note the ITS dataset is missing the *W. littoricola* (LITT\_NSW\_Hart) sample), as all of these species have hexaploid counts of  $2n=54$  reported (Smith, 1992).

The relationship between the New Zealand *W. gracilis* complex and the Australian *W. gracilis* species deserves a mention. As touched on in Chapter One the taxonomic history of *W. gracilis* is a confusing one, as there is uncertainty over whether the type of this species (originally *Campanula gracilis*) is housed at Kew or Gottingen, and whether it is made up of specimens collected in New Zealand or New Caledonia (or both) (Nicolson and Fosberg, 2004). All of the New Zealand radicate species were originally named as *C. gracilis* (Forster, 1786), hence the use of the term *W. gracilis* complex (Webb and Simpson, 2001), despite the fact that no species in New Zealand have that name now. Petterson (1997a) considered the type to have been collected in New Caledonia, and thought that none of the New Zealand *Wahlenbergia* were conspecific with the type. However, Smith (1992) considered the Australian entity called *W. gracilis* to belong to this widespread species described by Forster (1786) that also occurred in New Guinea, New Caledonia and New Zealand. In this study the ITS tree contains a moderately supported clade containing all of the New Zealand radicate *W. gracilis* complex individuals (labelled NZ1- in Figure 3.3). This clade has moderate support and does not include *W. gracilis* from Australia. However, the cpDNA tree shows that some cpDNA types are shared among New Zealand *W. gracilis* complex individuals and Australian *W. gracilis* (clade NZ1+ in Fig 3.4). Further research is required to determine which entities are conspecific, and which belong to the type of *W. gracilis*. Extensive sampling from New Caledonia, New

Guinea, New Zealand, Australia and Pacific islands will be necessary, analysed using sufficiently variable DNA sequence markers, or AFLPs.

### *Australian rhizomatous species*

Unlike in New Zealand the three Australian rhizomatous species are not monophyletic, although two of them, *W. densifolia* and *W. gloriosa*, do form a well-supported clade. The two species form a morphological continuum, with small samples of *W. gloriosa* sharing characters with *W. densifolia* (see descriptions in Smith (1992)). There is evidence in the chloroplast data that *W. densifolia* is differentiated from *W. gloriosa*, but not in the nuclear data. The samples here identified as *W. densifolia* were collected from the same location as one of the *W. gloriosa* samples (start of the Rennix walking track, Mt Kosciuszko see Appendix 1). So the fact that sympatric samples of *W. gloriosa* and *W. densifolia* are genetically differentiated (even only in the chloroplast data) is fairly strong evidence that they do form separate lineages. However, as the *W. densifolia* samples are nested within the *W. gloriosa* samples there may be a case for reducing their level of taxonomic recognition (e.g. to a subspecies of *W. gloriosa*). Further research with greater sampling across both species ranges would be necessary before such a decision could be made. The other rhizomatous species, *W. ceracea*, forms a clade with most of the *W. stricta* samples and the *W. telfordii* samples (discussed below). There is insufficient resolution to speculate on the evolution of the rhizomatous growth form in Australia.

### *Newly described species*

Both of the species newly described by Plunkett et al. (2009) represent distinct genetic lineages, a result that is congruent with their recent recognition at the species level. Analysis of DNA sequence data shows that *Wahlenbergia telfordii* is affiliated with *W. stricta* and the rhizomatous *W. ceracea*, and is in fact indistinguishable from *W. stricta* on chloroplast data alone, but is differentiated by the ITS data. On the other hand using morphological data Plunkett et al. (2009) noted morphological floral and leaf similarities between *W. rupicola* and *W. stricta*, and seed morphological similarities between *W. rupicola* and *W. ceracea*. It is therefore

surprising that here the *W. telfordii* individuals grouped with *W. stricta* and *W. ceracea* samples, though this had been implicated by the morphological dendrogram in Plunkett et al. (2009). The position of the *W. rupicola* individuals was less well resolved, as they grouped in a clade containing *W. littoricola* and *W. communis* samples among other species, but were entirely unresolved in the ITS tree. The lack of resolution is not surprising as Plunkett et al. (2009) hypothesised this entity was most closely related to either *W. glabra* (based on an initial morphological assessment), a species not included in this dataset or *W. luteola* (based on the phonetic dendrogram) which in the chloroplast data it does in fact group nearer to than any of the other species included in their morphological assessment.

## Summary and conclusions

The Australasian *Wahlenbergia* exhibit very low genetic diversity as measured by the chloroplast markers *trnL-F* and *trnK* and the nuclear ribosomal ITS. The low genetic diversity (especially in the New Zealand species) is probably due to rapid evolution during a period of geological and climatic change, potentially coupled with incomplete lineage sorting and hybridisation. While this low level of variation has resulted in poorly resolved phylogenetic trees a certain amount of information regarding relationships between individuals and species has resulted in an increased understanding of the evolution of the Australasian *Wahlenbergia* species.

The New Zealand radicate species as a group are probably not monophyletic, though further study is necessary. The only New Zealand radicate species clearly monophyletic in this study is the subtropical radicate *W. vernicosa*, which should be recognised at the species rank and not as a subspecies of *W. littoricola* as it is not closely related to that species. The other radicate species, the *W. gracilis* complex, should be perhaps considered as varieties of a single species, but further detailed studies are needed. The New Zealand rhizomatous species are probably monophyletic, which indicates radiation following a single introduction from Australia. Three morphologically distinctive rhizomatous species have some level of genetic distinction also (*W. cartilaginea*, *W. matthewsii* and *W. congesta*), which is consistent with their species ranking. The other two species (*W. albomarginata* and

*W. pygmaea*) are indistinguishable on the basis of this study and should perhaps be recognised as one species, though one subspecies of *W. albomarginata* might be a distinct genetic lineage (*W. albomarginata* subsp. *olivina*) and requires further study. The Australian species are very poorly resolved, though the two newly described species do form distinct lineages (*W. rupicola* and *W. telfordii*).

The poorly-resolved and complex phylogenetic relationships exhibited in this study make species delimitation on the basis of sequence data a difficult task. In this study, some species hypotheses have been supported, but a larger number of species hypotheses have not been adequately addressed by this dataset as there are too few characters to distinguish between individuals leading to very poorly resolved trees. In order to address these issues it will be necessary to explore this genus using another method, such as analysis of low-copy nuclear markers or perhaps amplified fragment length polymorphisms (AFLPs). AFLP analysis can be very useful when genetic variability is low, when studying polyploids and when hybridisation is occurring (Meudt and Clarke, 2007), three situations that all apply to the Australasian *Wahlenbergia*.



# Chapter Four: Using AFLPs to resolve phylogenetic relationships in a plant species complex when nuclear and chloroplast genes fail to reveal sufficient variability

---

## Abstract

AFLPs were used to analyse 38 individuals of Australasian *Wahlenbergia*, with a focus on the New Zealand representatives. Overall support values were slightly higher in the AFLP analyses than in the DNA sequence analysis of Chapter Three, indicating greater resolution for this technique. However, the higher support was generally found towards the tips of the branches, whereas deeper relationships were equally poorly supported using both AFLPs and sequence data. The morphologically distinctive *W. matthewsii* and *W. congesta* subsp. *congesta* were recovered as being distinctively different using AFLP analysis. Members of the *W. albomarginata*/*W. pygmaea* complex may all belong to the same species, as may the lowland radicate *W. gracilis* complex.

## Introduction

An interesting characteristic of the New Zealand flora is its relatively low rate of molecular evolution of commonly used markers such as nuclear ribosomal internal transcribed spacer (nrITS) and chloroplast loci, coupled with high rates of morphological and ecological variation (Winkworth et al., 2005). This pattern is found across several New Zealand plant genera both alpine and lowland (e.g. *Myosotis* (Winkworth et al., 1999; Winkworth et al., 2002a), *Veronica* as the hebe complex (Wagstaff et al., 2002; Wagstaff and Garnock-Jones, 1998), *Ourisia* (Meudt

and Simpson, 2006) and *Sophora* (Mitchell and Heenan, 2002)) and contributes to the difficulty in understanding species boundaries and evolutionary relationships in these groups. It also suggests that these species radiations have occurred relatively recently (Winkworth et al., 2005). Indeed molecular clock analyses of nrITS sequence data from several large genera in New Zealand (e.g. *Myosotis* (Winkworth et al., 2002a), *Ourisia* (Meudt et al., 2009) and *Ranunculus* (Lockhart et al., 2001)) suggest that these lineages radiated within the last 5 my. The last 5 million years was a period of rapid geological uplift and a changeable climate that created many new and fragmented habitats with fluctuating connectivity (Winkworth et al., 2005), which could well explain the pattern of low molecular vs. high morphological variation. Recent molecular studies of the New Zealand flora indicate this period was also a time of numerous long distance dispersal events between Australia and New Zealand and was therefore an important time for establishing trans-Tasman relationships in the flora (Winkworth et al., 1999).

Nuclear ribosomal DNA (nrDNA) and chloroplast DNA (cpDNA) are commonly used in plant phylogenetic analysis (Alvarez and Wendel, 2003). Chloroplast DNA is favoured because it is easy to extract and amplify due to the large number of copies in each cell (Kelchner, 2000) and universal primers can easily be developed for it (Wolfe and Randle, 2004). Additionally, organelles are usually inherited uniparentally, which potentially results in more straightforward interpretation. Uniparental inheritance restricts the usefulness of cpDNA though, as it reveals only half of the parentage in plants of hybrid or polyploid origin (Small et al., 2004). An additional limitation of cpDNA in molecular phylogenetic studies is its relatively slow rate of evolution, which means that even non-coding cpDNA regions often fail to resolve phylogenies at low taxonomic levels. Nuclear ribosomal DNA (e.g. the internal transcribed spacer (ITS)) has been widely used to complement organellar DNA by obtaining additional, independent, bi-parentally inherited phylogenetic estimates that evolve at a much higher rate. The high level of sequence variation can make alignment very difficult though, and there are a number of other concerns with using ITS as marker (see Alvarez and Wendel (2003) for a good summary). When both cpDNA and ITS sequences fail to resolve phylogenies, the amplified fragment length polymorphism (AFLP) approach has the potential to solve such difficulties (Meudt and Clarke, 2007).

The low level of variation in DNA sequence data frequently found in New Zealand studies means it can be difficult to resolve phylogenetic trees (e.g. *Ourisia* (Meudt and Simpson, 2006)). Molecular systematists in New Zealand have started turning to alternative molecular techniques, such as AFLPs, to explore molecular variation in the flora (e.g. *Veronica* (Meudt and Bayly, 2008), *Leocogenes* (Smitsen and Breitwieser, 2008), *Pseudopanax* (Perrie and Shepherd, 2009), *Metrosideros* (Broadhurst et al., 2008), *Ourisia* (Meudt et al., 2009) and *Polystichum* (Perrie et al., 2003)). AFLP markers are sampled throughout the (largely nuclear) genome and therefore they can display rare genetic differences in groups with low sequence variation e.g. among closely related species, crop species, or at the intraspecific level (Meudt and Clarke, 2007). This simultaneous analysis of many loci representing the whole genome is also more likely to generate a true species tree, rather than generating a particular gene tree as sequencing does (Després et al., 2003).

The New Zealand species of *Wahlenbergia* (Campanulaceae) appear typical of New Zealand alpine plant radiations – though half of the species inhabit the lowland - as they are mostly white flowered, polyploid species that occur in a diversity of habitats. Chapter two of this thesis has shown the ancestors of New Zealand *Wahlenbergia* probably arrived in New Zealand from Australia, then radiated, in the last 5 million years. Chapter three showed the low sequence variation of nrITS and two chloroplast loci and the accompanying difficulties in resolving the relationships of species of Australasian *Wahlenbergia* because of this. There are two clear lineages present within the New Zealand *Wahlenbergia*, corresponding to two different growth forms, radicate and rhizomatous. Within the New Zealand radicate group *W. vernicosa* has a unique chromosome count, and there was evidence from chloroplast DNA sequence data that it forms a distinct lineage. However, there was no evidence to falsify the hypothesis that the other four radicate species (the *W. gracilis* complex) are conspecific. Based on cpDNA and nDNA sequence data the New Zealand rhizomatous species are probably monophyletic, which indicates radiation following a single introduction from Australia. Three morphologically distinctive rhizomatous species had some level of genetic distinction based on DNA sequence data (*W. cartilaginea*, *W. matthewsii* and *W. congesta*), which is consistent with their species ranking. The other two species (*W. albomarginata* and *W. pygmaea*) were

indistinguishable on the basis of nuclear and chloroplast sequence data, though one subspecies of *W. albomarginata* formed a distinct group (*W. albomarginata* subsp. *olivina*). The New Zealand species of *Wahlenbergia* are therefore a prime candidate for analysing with AFLP data to attempt to resolve relationships and reconstruct their evolutionary history.

The aim of the present chapter is to analyse AFLP data for Australasian *Wahlenbergia* to assess species relationships and species limits, with a particular focus on New Zealand species. The current taxonomy and specific research questions of this thesis regarding Australasian *Wahlenbergia* have been described in detail in Chapter Three. Thus, the monophyly or otherwise of all of the species put forward in Petterson's (1997b) revision will be assessed, with specific interest in whether there is any evidence to falsify the hypothesis that all members of the *W. gracilis* complex are conspecific. The conspecificity of the New Zealand radicate *W. vernicosa* (treated as a subspecies of *W. littoricola* by de Lange and Cameron (1999)) with the Australian *W. littoricola* is of interest too. Evidence for species monophyly will also be assessed in species with the rhizomatous growth form, paying particular attention to the *W. albomarginata*/*W. pygmaea* complex and their subspecies. In this research, I follow de Queiroz (2007) in defining species as separately evolving metapopulation lineages. That is, a species is a demographically and genetically interconnected set of populations (metapopulation) that has continuity through time (Mayden, 1997). Although this definition does not provide an unambiguous cutoff for when speciation has occurred, it emphasizes that the primary goal of species delimitation research is lineage discovery and delimitation by using evidence from numerous sources of data.

## Methods

### Sampling and DNA extraction

Location of samples, collection details and voucher information are presented in Appendix 1. A total of 38 individuals were chosen for AFLP analysis, 31 New Zealand samples and 7 Australian samples. These individuals were selected in order that two or more representatives of each species and subspecies present in New Zealand were included in the dataset. This sampling strategy is effectively that of a

pilot study i.e. the intention was to assess the level at which the AFLPs are informative for New Zealand *Wahlenbergia*. Total DNA was extracted from silica-dried leaves after manual disruption of dried tissue with a mortar and pestle using a (CTAB) method modified from Doyle & Doyle (1990) and described fully in Chapter Two. Only DNAs of high quality and high concentration were used. Quality was determined by running out the extracted DNA on a 1.5% agarose gel and quantity was assessed by use of a nano-drop machine. Only samples with concentrations of DNA greater than 250ng/μL were used.

### **Generation of AFLP data**

AFLPs were generated based on the protocol of Vos et al. (1995) using an updated protocol for capillary detection of fluorescently labelled markers (see: [http://awcmee.massey.ac.nz/aflp/AFLP\\_Protocol.pdf](http://awcmee.massey.ac.nz/aflp/AFLP_Protocol.pdf)). Total DNAs were digested using EcoRI (Roche) and MseI (NEB) restriction enzymes at 37°C for 2 h, followed by 15 mins at 70°C to denature the enzymes. If nanodrop readings were higher than 1000ng/μL, 3μL of DNA was digested, if less than 1000ng/μL, 5μL were digested, thus ensuring a similar amount of DNA in each sample. Complete digestion was checked by running the digests out on a 1.5% agarose gel. Eco- and Mse-linkers were then ligated to the resulting DNA fragments by incubating with T4 DNA Ligase (Roche) at 37°C for 3 hours. Pre-selective PCR amplification was performed using primers Eco+A and Mse+C in a Biometra T gradient machine (Whatman, Germany) in a final volume of 20μL. Each 20μL volume contained 1μL of ligated DNA, 4μL of 5M betaine, 2.5μL of 2nM dNTPs, 2μL of 10x PCR buffer (Roche), 1μL of each of the primers (10pmol/μL), 0.2μL (corresponding to 1 unit) of Taq DNA polymerase (Roche) and ddH<sub>2</sub>O to make 20μL. The PCR protocol consisted of 20 cycles of denaturation at 94°C for 30 seconds, annealing at 56°C for 1 minute followed by extension at 72°C for 1 minute.

Selective PCR amplification was performed using fluorescently-labelled Eco+ANN primer and Mse+CNN primer in a final volume of 20μL. Each 20μL volume contained 1μL of pre-selective PCR product, 2.5μL of 25mM MgCl<sub>2</sub>, 2.5μL of 2nM dNTPs, 2μL of 10x PCR buffer (Roche), 1μL of each of the primers (10pmol/μL), 0.2μL (corresponding to 1 unit) of Taq DNA polymerase (Roche) and ddH<sub>2</sub>O to make

20µL. A step-down PCR program was used. The protocol involved an initial two-minute denaturation at 94°C followed by ten cycles of 94°C for 30 sec, 30 sec at 65°C – 56°C (dropping a degree every cycle) and 1 minute at 72°C followed by a further 30 cycles with the annealing temperature at 56°C. Four different fluorescently labelled Eco primers were used and trialled with a selection of potential combinations of eight different Mse+CNN primers. The following primer combinations were chosen based on a screen involving 4 individuals: 6FAM-Eco+ACT/Mse+CAA, VIC-Eco+ATA/Mse+CGT, NED-Eco+ACC/Mse+CAC and PET-Eco+AGG/Mse-CAA (hereafter, 6FAM, VIC, NED, PET). All primers were from Sigma except VIC-, NED- and PET- labelled primers (Applied Biosystems). For each individual, selective amplifications of each of the four dyes were mixed together in the ratio of 1:1:1:2 (6FAM:VIC:NED:PET), along with a GS-500 LIZ size standard, and 1µL of each sample was run on an Applied Biosystems Genetic Analyzer (ABI3730) at the Allan Wilson Centre Genome Service (Massey University, Palmerston North, New Zealand). To ascertain reproducibility replicate AFLP profiles were generated from independent restriction digests of the same DNA extraction for six individuals (ca. 10% of the dataset). Replicates were included in different runs on the 3730 Genetic Analyzer to ensure samples in different runs were comparable.

### **Parameter optimisation and automated scoring of AFLP data**

Automated scoring was performed on the resulting electronic AFLP profiles using GeneMarker v 1.80 (SoftGenetics). Scoring parameters were optimised following the procedure developed by Holland et al. (2008). After preliminary testing, they determined that peak height threshold (PHT), minimum fragment length (MFL), stutter peak filter (SPF) and local and global detection percentages (LGDP) were the most important parameters to be optimised (Holland et al. 2008). Thirty-six different *Wahlenbergia* AFLP character matrices were generated that varied by setting PHT to 50, 100 or 150 relative fluorescence units (rfu), MFL to 50, 75 or 100 base pairs (bp), SPF to either its default of 5%, or turned off, and LGDP to its default of 1% (both local and global) or turned off in GeneMarker. When scoring all matrices the ‘smoothing’ option was turned on and all other parameters were left at their default

values in GeneMarker. Data matrices containing characters from all four dyes were exported from GeneMarker and converted into NEXUS format files.

Comparisons of the 36 datasets to determine optimal scoring parameter settings were based on the following measures of accuracy: resolution (and normalised resolution), the number of parsimony informative characters, the number of replicate pairs that were correctly assigned as sister to one another, and the error rates (both Euclidean and Jaccard (Holland et al., 2008)). Holland et al. (2008) also developed a set of Python scripts to streamline the process of analysing the PAUP\* output and producing resolution scores, normalised resolution scores and both types of error rates. These scripts were implemented with Barbara Holland's help for this study. Resolution scores were thus calculated by performing 100 repetitions of 100 bootstrap replicates of neighbour joining (NJ) trees created using PAUP\* version 4.0b10 (Swofford, 2002). For each replicate the bootstrap scores over 50% were summed and this number was divided by the number of internal edges in each tree, to give a number between 0% and 100%. Both the number and quality of characters influence resolution scores. To get a measure of character quality independent of sequence length Holland et al. (2008) defined a normalised resolution score, which was calculated by creating datasets the length of the shortest dataset for each combination for parameter settings. The resolution score was then calculated as above. The number of parsimony informative characters and the number of replicate pairs that were correctly assigned as sister to one another were assessed from the PAUP\* NJ trees.

Both Euclidean and Jaccard error rates are measures of the replicate error. The Euclidean error rate compares the number of correct calls where a replicate pair both have no peak or both have a peak, to the number of incorrect calls where one half of the replicate pair has a peak and the other half does not. Holland et al. (2008) pointed out that this error rate may be influenced by the size of the data set, as more individuals (especially closely related ones) would increase the number of double absent calls in replicate pairs, thus apparently lowering the error rate. To counteract this Holland et al. (2008) developed an alternative error rate, which doesn't include double absent calls in the denominator, termed the Jaccard error rate.

To assess the robustness of the phylogenies to changes in parameter settings, majority rule consensus trees were created for each of the 36 datasets. The sets of 36 trees were then analysed using consensus networks as implemented in SplitsTree4 v. 10 (Huson and Bryant, 2006).

The parameter optimisation exercise was performed twice, once on the 36 datasets made up of characters from all four dyes, and once on another 36 datasets with the characters from 6FAM excluded. The 6FAM characters were excluded as further exploration of the samples in GeneMarker revealed that the 6FAM primer combination had failed in some samples, including one replicate. Optimal parameter settings were determined based upon the above analyses by assessing trends in each of the accuracy measures. Following scoring optimization and replicate checking, the six replicate profiles were removed; one additional individual (CART\_SI\_Hanm1) was also removed as NED failed in that sample.

### **Phylogenetic analysis**

Phylogenetic trees were reconstructed using maximum parsimony (MP) and Bayesian inference models. MP trees were built using PAUP\* v4b10 (Swofford, 2002). MP analysis was performed in a two-step search strategy. First, multiple islands were searched with 10,000 random addition sequence replicates, nchuck = 5, chuckscore = 1 and maxtrees = 10,000. The resulting trees were then swapped to completion with the same settings but chuckscore = no. Support for clades was assessed using 1000 bootstrap replicates, 10 random addition replicates, and MAXTREES = 20,000 in PAUP\*. For Bayesian inference a restriction site (binary) model was used. Although the evolution of AFLP markers is far more complex than the model assumes (Luo et al., 2007), this model has been shown to be useful for analyzing AFLP data nonetheless (Koopman et al., 2008). Four independent chains with 10 million iterations were run using MrBayes v3.1 (Huelsenbeck, 2001). Convergence was assumed to have occurred when the standard deviation of split frequencies was below 0.01, but was also assessed using Tracer v1.4.1 (Rambaut and Drummond, 2007). A burn in of 5% was used.



## **Principal coordinate analysis (PCoA) of dominant data**

The program NTSYS-pc (Rohlf, 1990) was used to conduct a principal coordinate analysis (PCoA) to investigate clusters in the dataset. PCoA is an ordination method, which graphically explores the relationships within the data by mapping a similarity matrix. A PCoA allows the assessment of the dimensionality of the data and a description of the major patterns of variation within and between populations (Ishida et al., 2003). The PCoA was performed using Jaccard distances, DCENTER, EIGEN and MOD3D modules with 3 dimensions analysed and visualised. Initially the full dataset was analysed, followed by subsets of the dataset to further investigate clusters of interest.

## **Results**

### **Selection of optimal AFLP automated scoring parameters**

The resolution, normalised resolution, number of parsimony informative characters, the number of replicate pairs that were correctly assigned as sister to one another, and the error rates (both Euclidean and Jaccard) for all 36 datasets are presented in Table 4.1. The number of characters in each of the 36 datasets generated in GeneMarker using different scoring parameter settings varied between 945 and 1473 with 6FAM characters included, and between 714 and 1039 with them excluded. The datasets with the greatest number of characters also had the largest number of parsimony informative characters and the highest resolution scores. Unfortunately they also had the highest error rates and the lowest normalised resolution scores. Nevertheless, overall none of the accuracy measures varied all that much. When 6FAM characters were included the number of parsimony informative characters ranged from 731 to 1204, replicates correctly paired from 3-5, resolution scores from 56%- 65%, normalised resolution from 66%-71%, Euclidean error rates from 7%-9% and Jaccard error rates from 42%-46%. When the 6FAM characters were excluded the number of parsimony informative characters ranged from 555 to 842, replicates correctly paired from 4-5, resolution scores from 54%- 60%, normalised resolution from 80%-87%, Euclidean error rates from 7%-9% and Jaccard error rates from 42%-46%.

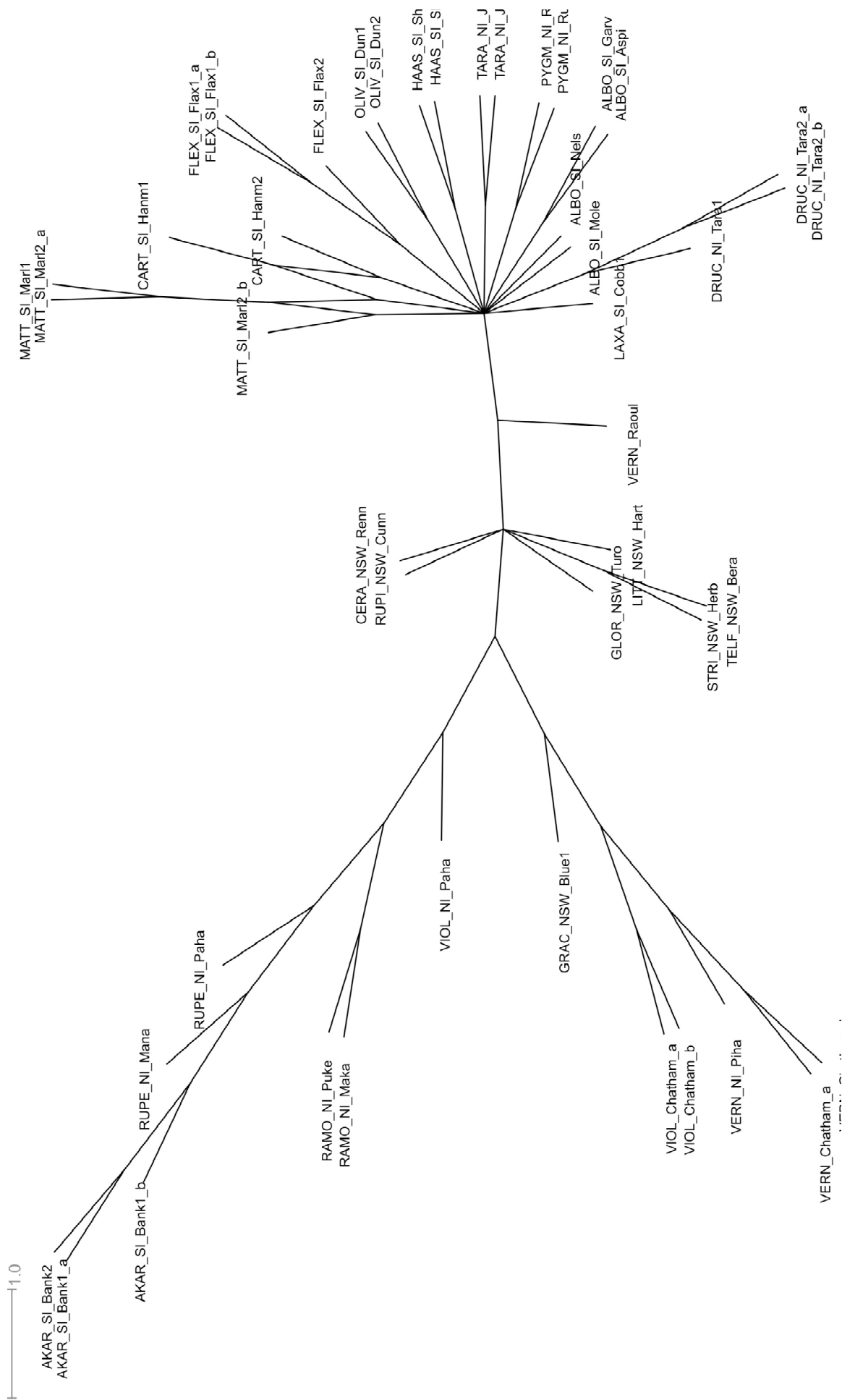
When other published AFLP studies have reported Euclidean error rates, they are usually between 2-5% (see (Bonin et al., 2004)). The higher error rate reported here could be due to the low number of correctly paired replicates (see Table 4.1). When the 6FAM characters were excluded the error rates decreased only slightly for some datasets, despite the higher replicate pairing rate. The higher pairing rate was due to the consistent pairing of the *W. vernicosa* (VERN\_NI\_Piha) pair, as 6FAM had failed for one of the replicates. The two replicates that mostly did not pair even when 6FAM characters were excluded were *W. akaroa* (AKAR\_SI\_Bank1) and *W. matthewsii* (MATT\_SI\_Marl2). In all 36 datasets the *W. akaroa* replicates grouped in a clade containing only the two replicates and the other *W. akaroa* (AKAR\_SI\_Bank2) sample (data not shown), which indicates merely that the AFLPs are unable to distinguish between individuals from the same population. In contrast there was no pattern to the relationships between the *W. matthewsii* replicates in the first 24 datasets, they were often in different clades (though the clades always had low bootstrap support). For the remaining datasets they followed the same pattern as the *W. akaroa* replicates.

As well as the low replicate pairing, the higher Euclidean error rates reported in this study could be due to the completely automated scoring process implemented here. Most published AFLP studies have used a combination of manual or semi-automated scoring. Holland et al. (2008) reported similarly high Euclidean error rates in their fully automatedly scored datasets, and after exploring their results with ReplicateError (available from <http://baseplugins.thep.lu.se/wiki/se.lu.onk.ReplicateError>) suggested the majority of errors were scoring errors (rather than PCR errors). They therefore suggested that the incorporation of improved scoring algorithms into GeneMarker would further increase its power and usefulness (Holland et al., 2008). They also nevertheless advocated use of automated scoring of AFLPs, as manual scoring is subjective, time-consuming and not repeatable in the same way that automated scoring is (Holland et al., 2008).

**Table 4.1** Dataset statistics for the 36 datasets generated in GeneMarker including and excluding the 6FAM characters

Dataset File name Number	No of Characters		No of Parsimony informative characters		No of replicates correctly paired (/6)		Resolution scores		Normalised resolution scores		Euclidean error rates		Jaccard error rates	
	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM
1 WahlAFLP1_R50_PD50_GL1_SPON.nex	1473	1039	1204	842	4	4	65	54	66	81	9	9	46	45
2 WahlAFLP2_R50_PD50_GL1_SPOFF.nex	1473	1039	1204	842	3	4	65	55	66	81	9	9	46	45
3 WahlAFLP3_R50_PD50_GL0_SPON.nex	1473	1039	1205	843	3	4	65	57	67	80	9	9	46	45
4 WahlAFLP4_R50_PD50_GL0_SPOFF.nex	1473	1039	1205	843	3	4	65	57	66	80	9	9	46	45
5 WahlAFLP5_R50_PD100_GL1_SPON.nex	1341	982	1087	792	3	4	63	59	67	81	9	8	45	44
6 WahlAFLP6_R50_PD100_GL1_SPOFF.nex	1341	982	1087	792	3	4	64	60	67	83	9	8	45	44
7 WahlAFLP7_R50_PD100_GL0_SPON.nex	1342	983	1088	793	3	4	64	59	68	82	9	9	45	44
8 WahlAFLP8_R50_PD100_GL0_SPOFF.nex	1342	983	1088	793	3	4	62	59	67	82	9	9	45	44
9 WahlAFLP9_R50_PD150_GL1_SPON.nex	1182	900	943	721	3	4	62	57	67	83	8	8	43	43
10 WahlAFLP10_R50_PD150_GL1_SPOFF.nex	1182	901	943	722	3	4	61	57	67	83	8	8	43	43
11 WahlAFLP11_R50_PD150_GL0_SPON.nex	1183	901	944	722	3	4	62	57	66	83	8	8	43	43
12 WahlAFLP12_R50_PD150_GL0_SPOFF.nex	1183	901	944	722	3	4	63	57	68	83	8	8	43	43
13 WahlAFLP13_R75_PD50_GL1_SPON.nex	1363	952	1114	767	3	4	59	57	68	84	9	8	46	45
14 WahlAFLP14_R75_PD50_GL1_SPOFF.nex	1363	952	1114	767	4	4	58	58	68	84	9	8	46	45
15 WahlAFLP15_R75_PD50_GL0_SPON.nex	1363	952	1114	767	3	4	59	57	68	83	9	8	46	45
16 WahlAFLP16_R75_PD50_GL0_SPOFF.nex	1363	952	1114	767	3	4	60	57	68	83	9	8	46	45
17 WahlAFLP17_R75_PD100_GL1_SPON.nex	1236	897	1001	720	3	4	59	58	69	84	9	8	45	44
18 WahlAFLP18_R75_PD100_GL1_SPOFF.nex	1238	897	1001	720	4	4	59	58	69	84	9	8	45	44
19 WahlAFLP19_R75_PD100_GL0_SPON.nex	1238	897	1001	720	3	5	59	58	69	84	9	8	45	44
20 WahlAFLP20_R75_PD100_GL0_SPOFF.nex	1238	897	1001	720	4	4	58	58	68	85	9	8	45	44
21 WahlAFLP21_R75_PD150_GL1_SPON.nex	1081	816	857	649	4	4	60	55	69	85	8	8	43	43

Dataset File name Number	No of Characters		No of Parsimony informative characters		No of replicates correctly paired scores (/6)		Resolution scores		Normalised resolution scores		Euclidean error rates		Jaccard error rates	
	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM	6FAM	No 6FAM
22 WahlAFLP22_R75_PD150_GL1_SPOFF.nex	1081	816	857	649	3	4	61	56	70	85	8	43	43	43
23 WahlAFLP23_R75_PD150_GL0_SPON.nex	1081	816	857	649	3	4	61	56	70	85	8	43	43	43
24 WahlAFLP24_R75_PD150_GL0_SPOFF.nex	1081	816	857	649	3	5	61	56	70	85	8	43	43	43
25 WahlAFLP25_R100_PD50_GL1_SPON.nex	1225	848	991	677	3	4	59	55	68	82	8	46	46	45
26 WahlAFLP26_R100_PD50_GL1_SPOFF.nex	1225	848	991	677	3	4	60	54	68	82	8	46	46	45
27 WahlAFLP27_R100_PD50_GL0_SPON.nex	1226	849	992	678	3	4	60	55	68	82	8	46	46	46
28 WahlAFLP28_R100_PD50_GL0_SPOFF.nex	1226	849	992	678	3	4	60	54	69	82	8	46	46	46
29 WahlAFLP29_R100_PD100_GL1_SPON.nex	1101	794	875	626	3	5	57	54	68	82	8	45	45	44
30 WahlAFLP30_R100_PD100_GL1_SPOFF.nex	1101	794	875	626	4	5	58	55	68	83	8	45	45	44
31 WahlAFLP31_R100_PD100_GL0_SPON.nex	1101	794	875	626	3	5	59	54	68	82	8	45	45	44
32 WahlAFLP32_R100_PD100_GL0_SPOFF.nex	1101	794	875	626	4	5	58	54	68	82	8	45	45	44
33 WahlAFLP33_R100_PD150_GL1_SPON.nex	945	714	730	554	3	4	55	57	71	87	7	42	42	42
34 WahlAFLP34_R100_PD150_GL1_SPOFF.nex	945	714	730	554	3	4	55	56	70	87	7	42	42	42
35 WahlAFLP35_R100_PD150_GL0_SPON.nex	945	714	731	555	3	4	55	55	71	87	7	42	42	43
36 WahlAFLP36_R100_PD150_GL0_SPOFF.nex	945	714	731	555	3	4	56	55	71	87	7	42	42	43

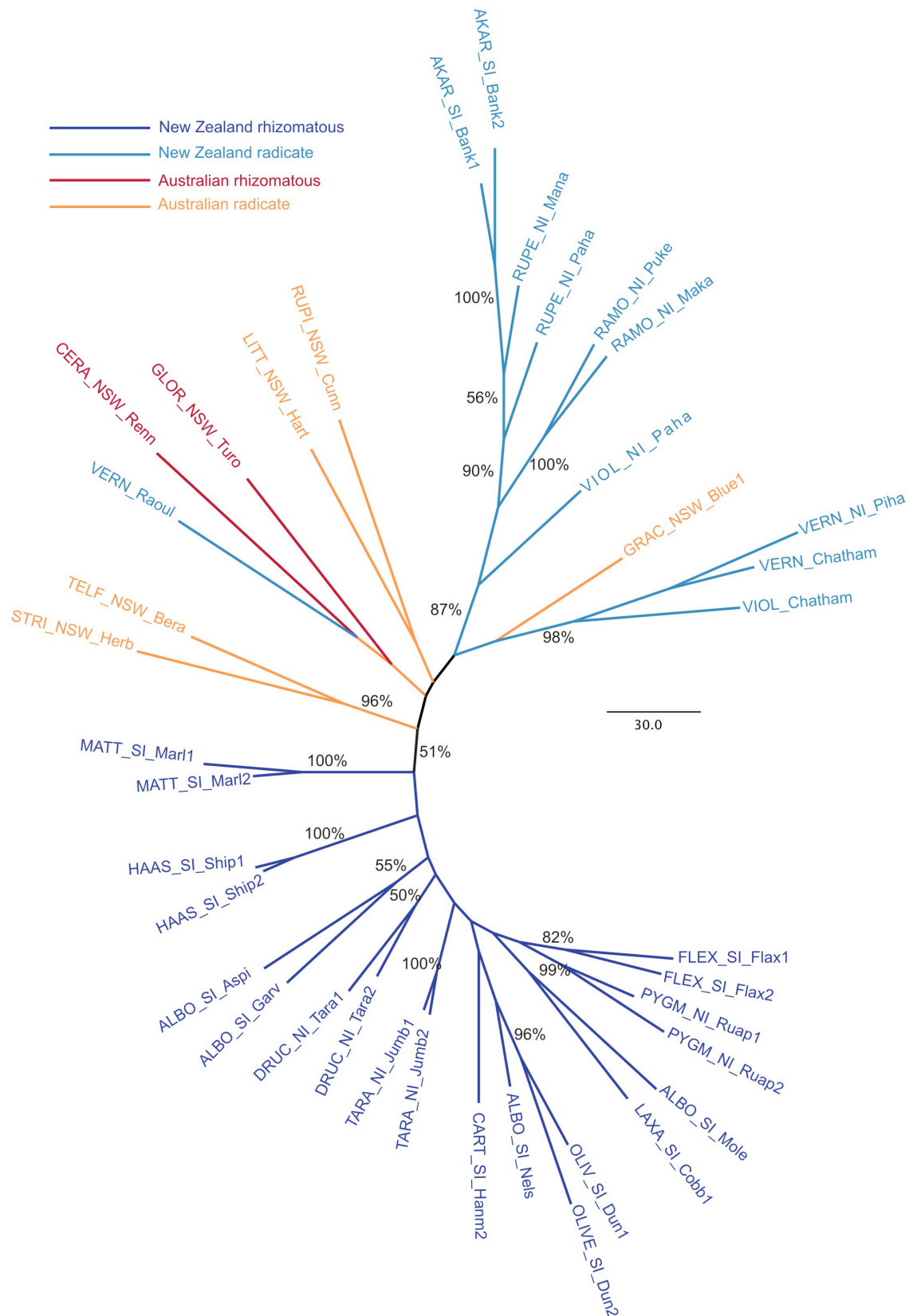


**Figure 4.1.** consensus network of the 36 Australasian *Wahlenbergia* AFLP datasets generated in GenMarker , excluding the 6FAM characters. Explanation of tag names and voucher information can be found in appendix 1.

The small amount of variation in all measures suggests this particular dataset is relatively robust to changes in the scoring parameters, which is also indicated by the lack of “boxiness” in the majority rule consensus network (Figure 4.1). The main conflict is in the placement of the *W. matthewsii* replicates. The low variation suggests it is of low importance which dataset is selected for use in this instance. Therefore, I decided to treat the correct replicate pairing as the most important indicator of which datasets to use, which meant excluding 6FAM characters. Of the datasets with 6FAM characters excluded, datasets 25-36 were better at placing the *W. matthewsii* replicates. Within these datasets I selected the dataset with the lowest error rates and highest normalised resolution score. The optimal scoring parameters were selected to be PHT of 150, MFL of 100, SPF on and LGDP off (dataset number 35, highlighted in blue in Table 4.1). With 6FAM characters included this set of scoring parameters yielded the highest normalised resolution score (71%) and the lowest error rates (Euclidean = 7, Jaccard = 42). With 6FAM excluded this set of scoring parameters again had the highest normalised resolution score (87%), and the lowest Euclidean error rate (7), and the Jaccard error rate was second lowest (43, rather than 42).

### **Phylogeny of New Zealand *Wahlenbergia* based on AFLP data**

The parsimony analysis returned five most parsimonious trees, one of which is shown here with bootstrap values added (Figure 4.2). The Bayesian analysis ran for 10 million generations, at which point the standard deviation of split frequencies was 0.003. A 50% majority rule consensus tree was created (Figure 4.3). The Bayesian and MP trees presented here are not rooted, as it is unclear which species should be considered the root of the tree based on the DNA sequence phylogenies reconstructed in chapter three. The 50% majority rule trees reconstructed with AFLP data using Bayesian and MP estimation had similar topologies, though with some important differences (note the 50% majority rule MP tree is not shown, but instead bootstrap values of greater than 50% are recorded on one of the most parsimonious trees Figure 4.2). Overall, the Bayesian tree had higher posterior probability (pp) values than the parsimony tree bootstrap (bs) values, which is in line with previous



**Figure 4.2.** One of the five most parsimonious trees reconstructed from Australasian *Wahlenbergia* AFLP data. Numbers next to branches are bootstrap values (%). Explanation of tag names and voucher information can be found in Appendix 1.







studies (Erixon et al., 2003; Wilcox et al., 2002), and can be explained by the fact that posterior probabilities and parsimony bootstrap support values measure different characteristics of the data (Alfaro et al., 2003). Posterior probabilities measure the probability of a node being correct, conditional on the evolutionary model employed and the data set. Therefore their reliability heavily depends on the accuracy of both the data sample and the evolutionary model, which has been acknowledged as being lacking in the case of analysis of AFLP data (Luo et al., 2007). Parsimony bootstrap values are different in that they measure the sensitivity of the observed results to the sampling error associated with collecting characters from a hypothesized underlying character distribution, without the use of an explicitly specified evolutionary model (Alfaro et al., 2003). Given the fact that posterior probabilities and parsimony bootstrapping measure different features of the data, the most reliable nodes are those that have high values of both.

The Bayesian and MP analysis methods both grouped the New Zealand rhizomatous species into a poorly supported clade (0.77 posterior probability (pp), 51% bootstrap support (bs)). The Bayesian 50% majority rule tree (Figure 4.3) recovered the New Zealand radicate species (excluding the *W. vernicosa* sample from Raoul Island (VERN\_Raoul)) in a clade with high (0.97) pp, but this relationship was unsupported using parsimony methods (Figure 4.2). In both trees the position of the Australian samples is unresolved between the New Zealand rhizomatous and New Zealand radicate groups. Higher support values were found usually on branches connecting terminal taxa, which corroborates the general notion that AFLP markers are most reliable at the lower taxonomic levels (Mueller and Wolfenbarger, 1999).

Within the New Zealand radicates, *W. akaroa* (1.0 pp, 100% bs) and *W. ramosa* (0.99 pp, 100% bs) formed clades in both analyses while *W. violacea* and *W. rupestris* did not. Interestingly the *W. vernicosa* samples (excluding VERN\_Raoul) formed a monophyletic group in the parsimony analysis (<50% bs), but in the Bayesian analysis the *W. vernicosa* sample from the Chatham Islands (VERN\_Chatham) formed a clade with the *W. violacea* (VIOL\_Chatham) sample also from the Chatham Islands (0.89 pp). The position of the Australian *W. littoricola* (LITT\_NSW\_Hart) sample was unresolved, but it was not within the clade containing the New Zealand *W. vernicosa* samples.

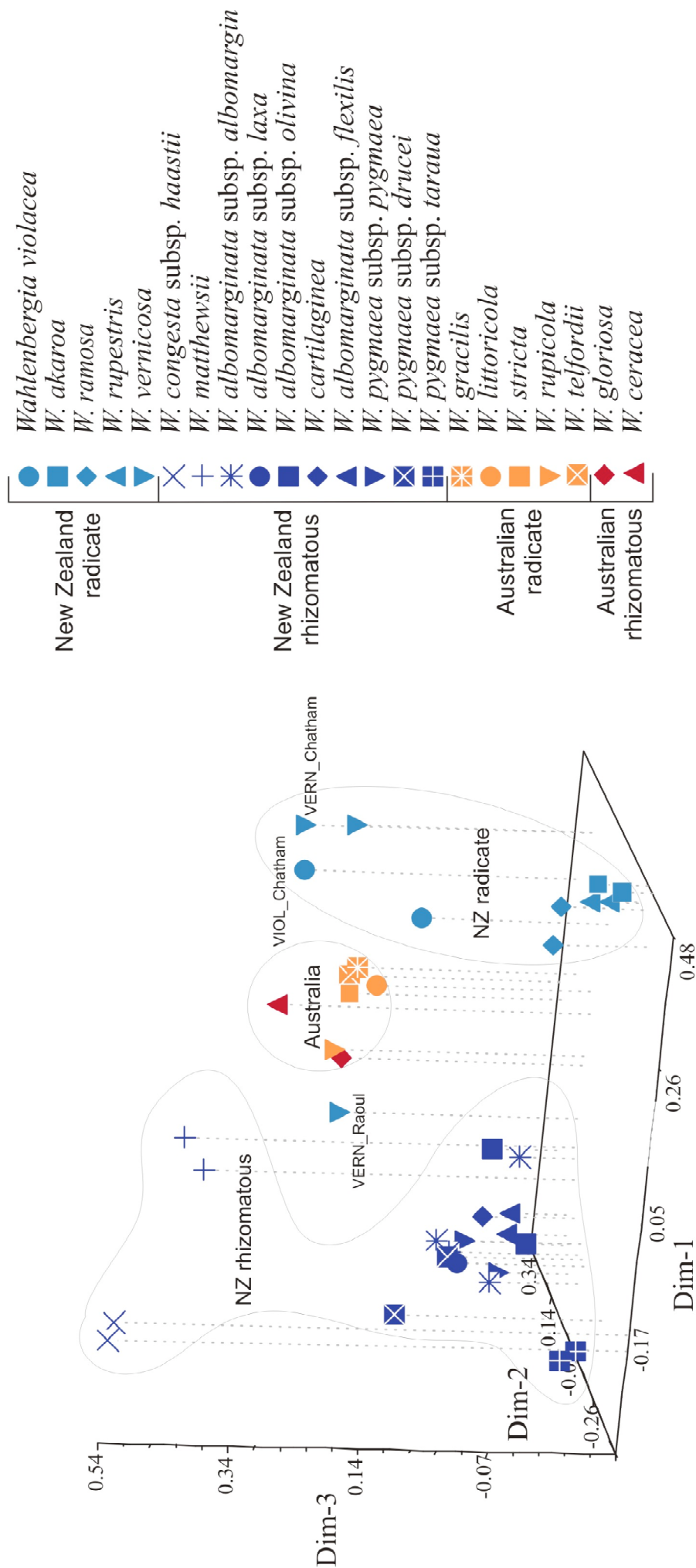
Within the New Zealand rhizomatous clade samples of *W. matthewsii* (1.0 pp, 100%bs) and *W. congesta* (1.0 pp, 100% bs) formed species clades, but the *W. albomarginata* and *W. pygmaea* samples did not. However, subspecies of these two alpine species mostly formed clades, apart from the *W. albomarginata* subsp. *albomarginata* individuals that were spread throughout the rhizomatous clade in both analyses, and the *W. pygmaea* subsp. *drucei* samples that grouped in separate clades in the Bayesian tree.

### Cluster analysis using principal coordinates analysis

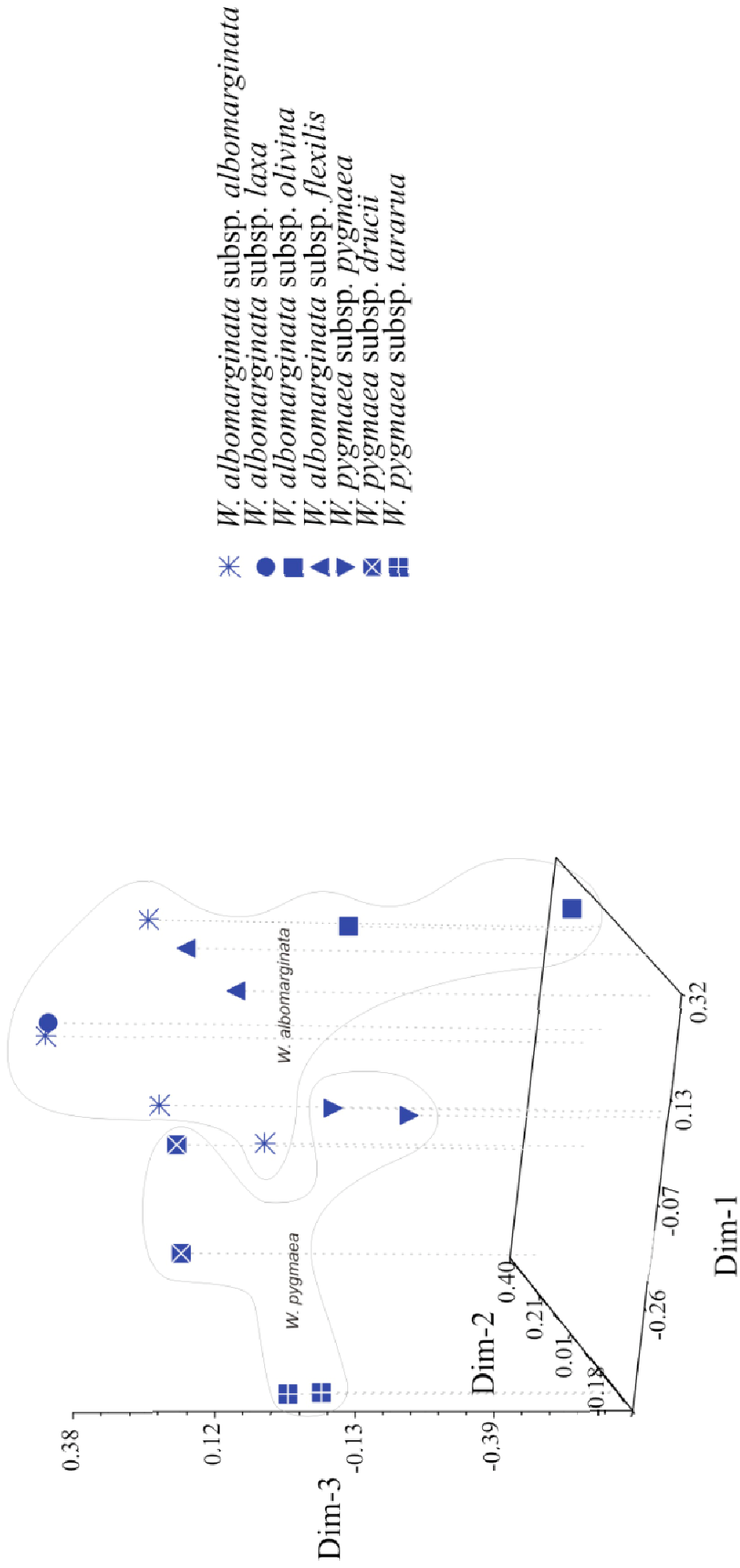
Principal coordinates analysis revealed similar patterns to the tree building methods (Figure 4.4); the NZ rhizomatous species were clearly separated from the remaining species. The New Zealand radicate species also formed a separate cluster with the exception of the *W. vernicosa* sample from Raoul Island (VERN\_Raoul), which was closer to the Australian species. The Australian samples cluster tightly together and are separate from the New Zealand samples. Dimension one explained 9.4% of the variation, dimension two 5.3% and dimension three 4.8%. The New Zealand radicate samples are fairly well separated on dimensions 1 & 2 (Figure 4.4) and the NZ rhizomatous species are mainly separated on dimensions 1 & 3, particularly the individuals of *W. matthewsii* and *W. congesta* subsp. *haastii*. Individuals of *Wahlenbergia albomarginata*, *W. pygmaea* and *W. cartilaginea* are not well separated from each other or as species. A reduced dataset containing only members of *W. albomarginata* and *W. pygmaea* was analysed (Figure 4.5). With the other samples removed, more structure is visible, and members of the same subspecies generally group together. However, there is still no clear separation between *W. albomarginata* and *W. pygmaea*. Of this reduced dataset dimension one explained 10.5% of the variation, dimension two 9.4% and dimension three 8.6%.

## Discussion

AFLPs are a useful technique to explore genetic variation within New Zealand *Wahlenbergia* based on the tree-like phylogenies recovered using MP and Bayesian techniques, and the distinct clusters formed using PCoA. Given the low posterior



**Figure 4.4** Principle coordinates analysis (PCoA) of Australasian *Wahlenbergia* AFLP dataset showing three dimensions. Dimension one explains 9.4%, dimension two an additional 5.3% and dimension four a further 4.8%. Explanation of tag names and information on voucher specimens can be found in Appendix 1.



**Figure 4.5** Principle coordinates analysis (PCoA) of a reduced Australasian *Wahlenbergia* AFLP dataset showing the *W. albomarginata*/*W. pygmaea* complex in three dimensions. Dimension one explains 10.5%, dimension two an additional 9.4% and dimension four a further 8.6%. Explanation of tag names and information on voucher specimens can be found in Appendix 1.

probabilities and bootstrap support for branches deeper in the trees it seems likely that AFLPs will be more useful at species delimitation rather than assessing phylogenetic relationships among species. Other studies on New Zealand plants using AFLP data have come to a similar conclusion (e.g. *Ourisia* (Meudt et al., 2009)) and this is consistent with the accepted view that AFLPs are more useful at lower taxonomic levels (Meudt and Clarke, 2007; Mueller and Wolfenbarger, 1999). Unfortunately, this means we have not yet managed to discover a molecular technique that is able to resolve deeper relationships within the Australasian *Wahlenbergia*. Nevertheless, AFLPs will be a useful technique for exploring species boundaries within ambiguous species complexes in this genus.

## **Phylogenetic patterns and species delimitation in New Zealand**

### ***Wahlenbergia* based on AFLP**

#### ***Radicata species***

The New Zealand radicate species (excluding *W. vernicosa* from Raoul Island) form a cluster in the PCoA analysis and a well-supported group in the Bayesian tree analysis (0.97 pp). They do not form a clade in the MP tree, where *W. gracilis* (an Australian radicate species) is nested within them in an unsupported clade which is unable to be clearly separated from the remaining Australian samples (Figure 4.2). This repeats the pattern found in Chapter Three of this thesis using nuclear and chloroplast DNA sequences, where there was also no clear separation of Australian and New Zealand radicate samples. From the dating analyses conducted in Chapter Two the radiation of all of the Australasian *Wahlenbergia* was estimated to have occurred about 3.7 mya, and it appeared the New Zealand radicate species probably radiated more recently still about 0.5 mya. This may have been insufficient time for reciprocal species monophyly to have evolved (Shaffer and Thomson, 2007) and therefore the lack of clear structure in the radicate group may simply reflect recent evolution.

Despite this recent evolution, there is some structure within the New Zealand radicates; the *W. akaroa* samples form a well-supported clade (1.0 pp, 100% bs), as do the *W. ramosa* samples (0.99 pp, 100% bs). These also appear to be closely

related based on the PCoA analysis. Neither of these species formed monophyletic groups in the analysis of DNA sequence data. That the relationships are recovered using AFLP data could indicate a greater resolution found using this method. Alternatively it could reflect the smaller sampling in this AFLP study, as two *W. ramosa* samples that grouped separately to the other New Zealand radicate sample in the nuclear and chloroplast sequence trees were not included in the AFLP dataset. However, because the two *W. akaroa* individuals that were included in all studies received high support in the AFLP analysis this suggests more information is available using AFLP data relative to DNA sequence data.

The other NZ radicate species do not form monophyletic groups in the AFLP trees. This is not surprising for *W. rupestris* and *W. violacea* as they do not form monophyletic groups in the DNA sequence data either, but the *W. vernicosa* samples (excluding the one collected from Raoul Island) were monophyletic at least in the chloroplast tree. In the parsimony AFLP analysis the clade containing the two *W. vernicosa* samples (one of which, VERN\_Chatham, is from the Chatham Islands) also includes a *W. violacea* sample that was collected from Chatham Islands (VIOL\_Chatham). Samples of *W. vernicosa* that have had their chromosomes counted are tetraploid ( $2n=54$ ), whereas *W. violacea* (and all other New Zealand temperate radicate species) are octoploid ( $2n=72$ ) (Pettersen et al., 1995). Albach (2007) found that AFLPs are sensitive to the analytical method employed (i.e. different relationships are recovered), especially if hybridisation and allopolyploidy are involved. The origins of the New Zealand polyploid *Wahlenbergia* species are unknown, but analysis of samples with different ploidy levels could explain why the relationship between the *W. violacea* and *W. vernicosa* samples changes subtly with different analysis methods (parsimony vs. Bayesian). Detailed genomic and hybridisation studies would be needed to test this hypothesis. Despite the non-monophyly of *W. vernicosa* in both the AFLP and nuclear DNA sequence data the consistent morphological differences (e.g. glossy leaves) and different chromosome number appear to set it apart from the other New Zealand radicate species.

The position of the Raoul Island sample identified as *W. vernicosa* (VERN\_Raoul) is again unresolved, though the PCoA analysis appears to group it more closely with the Australian samples than the other New Zealand samples. The single Australian

*W. littoricola* (LITT\_NSW\_Hart) sample included in the AFLP dataset is more closely related to other Australian species than it is to the *W. vernicosa* samples, a pattern also found in chloroplast and nuclear sequence data. The patterns observed in the nrDNA, cpDNA and AFLP are not those expected from conspecific samples, thus they falsify the hypothesis (de Lange and Cameron, 1999) that these entities are conspecific.

Overall this dataset yields little evidence for speciation within the *W. gracilis* complex (*W. violacea*, *W. akaroa*, *W. ramosa* and *W. rupestris*) despite the monophyly of *W. akaroa* and *W. ramosa*. The two *W. akaroa* samples were from the same population, and both of the *W. ramosa* samples were from the Wellington west coast, suggesting the AFLP could be finding population level geographic structure within a single species, rather than species level variation. Additionally the morphological characters differentiating *W. ramosa* and *W. akaroa* are no more than those differentiating the non-monophyletic *W. violacea* and *W. rupestris*. It appears that *W. akaroa* is derived from within the *W. rupestris* samples with high support (Figures 4.2 and 4.3). This result ties in with the original species diagnosis in Petterson (1997b), which described *W. akaroa* as being “similar to *W. rupestris*, differing in the pale-violet corolla...” Although these findings do not refute a hypothesis that the *W. gracilis* complex is a single variable species, the flower size differences (Petterson, 1997b) indicate that more sampling and further independent tests should be conducted.

### ***Rhizomatous species***

The New Zealand rhizomatous species form a poorly supported clade in both the parsimony and Bayesian tree building methods (0.77 pp, 51% bs). This clade was also recovered using DNA sequence data (Chapter Three), with slightly lower support values (Figures 3.3-3.5). Within this clade *W. matthewsii* and *W. congesta* subsp. *haastii* form distinct lineages in the genus based on the AFLP data. This is recovered by both the parsimony and Bayesian tree building methods, as well as the PCoA. There was also some evidence from the DNA sequence data that these species were distinct from each other and the remaining species, and morphological characters also suggest these represent distinct genetic lineages that should continue

to be treated at the species rank. Only one specimen in the AFLP analyses represented the other morphologically distinct rhizomatous *Wahlenbergia*, *W. cartilaginea* (CART\_SI\_Hanm2). Despite its very different morphology this one specimen clusters tightly with the *W. albomarginata*/*W. pygmaea* complex in the PCoA analysis and its placement is unresolved in both tree building methods for the AFLP data. However, its distinct morphology and habitat (it is a scree plant), and monophyly in phylogenies reconstructed from both nrDNA and cpDNA sequence data, is sufficient evidence to falsify a hypothesis of conspecificity with the other New Zealand rhizomatous species.

As found in the DNA sequence data the two species *W. albomarginata* and *W. pygmaea* do not form monophyletic groups. The samples from these two species were analysed using PCoA without the rest of the samples to try to tease out their relationships. This analysis, along with the tree building methods, showed that despite the non-monophyly of the two species several of their subspecies do form groups or clades. As discussed in Chapter Three this could partly be an artefact of the reduced range of some of the subspecies in that the duplicate samples are from the same population whereas samples of the more widespread (non-monophyletic) subspecies e.g. *W. albomarginata* subsp. *albomarginata* are from a range of populations. This AFLP study could not refute the hypothesis that *W. albomarginata* and *W. pygmaea* are conspecific. This question deserves further study and AFLPs are potentially useful to test this with increased sampling to represent different populations of all subspecies from throughout their ranges.

### **Trans-Tasman relationships**

AFLP data are not able to distinguish among the Australian samples, which is probably due to inadequate sampling as only one individual of each of seven species was included.

## **Summary and conclusions**

Similar patterns of relationships were found with this analysis of AFLP data relative to those found using analyses of chloroplast and nuclear DNA. Overall, support



values - which can be used as a proxy for level of resolution within the data (Holland et al., 2008) - were slightly higher in the AFLP analyses. The support values were generally higher towards the tips of the branches, whereas deeper relationships were equally poorly supported using both AFLP and sequence data, suggesting neither method has the appropriate level of variation to elucidate these relationships. It may be possible to find a more variable DNA marker (such as low copy nuclear genes), which may be able to provide information at this level. Although several low-copy nuclear markers trialled in chapter three could not be successfully sequenced, others should be attempted. It is also possible that there are no molecular techniques able to resolve the relationships in question, as there may not be the appropriate signal in the genomes. These kinds of molecular techniques rely on mutations that arise by chance, and the probability that mutations will have occurred that are able to correctly assign each individual to a species group must be relatively low, as their radiation (in Australasia) has probably occurred only in the last 3.7 my. Thus, some level of fuzziness and uncertainty in delimiting young species at the beginning of differentiation as shown here is therefore perhaps not unexpected (Shaffer and Thompson, 2007). Hence, rapid evolution and species radiation, such as that occurring within Australasian *Wahlenbergia*, continues to present a challenge to systematists and taxonomists.

# Chapter Five: General Discussion

---

The overall aim of this thesis was to explore the evolution of the Australasian *Wahlenbergia* at three different scales. At the large scale I undertook a molecular phylogenetic analysis exploring the biogeography of *Wahlenbergia* including samples from South Africa, Europe, Australia and New Zealand. At the medium scale I assessed the current taxonomy of the New Zealand *Wahlenbergia* (and Australian species where sampling permitted) using DNA sequencing, and at the small scale I further explored relationships among New Zealand *Wahlenbergia* species using AFLP techniques.

## Large scale phylogeny and biogeography

At the large scale, the genus *Wahlenbergia*, which comprises ~260 species worldwide, was confirmed to be polyphyletic in a phylogeny reconstructed using *trnL-F* and ITS sequence data from ~20% of the species in the genus, although the majority of *Wahlenbergia* species formed a clade. Further sampling is required to confirm whether this monophyly extends to include Asian and Pacific *Wahlenbergia*, and whether it is maintained when other closely related wahlenbergioid species are included in the dataset. The genus originally evolved in South Africa (with a herbaceous growth form) and began diversifying about 16.2 mya before dispersing to Australasia about 3.7 mya and radiating into 45 species and subspecies there. It is unclear whether this dispersal event included a step-wise progression through Asia, as unfortunately no Asian *Wahlenbergia* were included in this study. Two introductions from Australia to New Zealand appear likely, leading to two radiations: one of species with the radicate growth form and one of species with the rhizomatous (both creeping and suffructose) growth form. The ancestor of the New Zealand rhizomatous species arrived and began diverging about 1 mya, and the ancestor of the radicate species about 0.5 mya. The recent evolution of the *Wahlenbergia* genus as a whole, and the relatively recent introductions to New Zealand, refute the

Gondwanan vicariance hypothesis for the origins of the New Zealand *Wahlenbergia* and instead provide overwhelming evidence for arrival via long distance dispersal. Recent studies have shown long-distance dispersal to be a common component in forming the New Zealand flora e.g. *Plantago* (Tay et al., 2010), *Veronica* (as the hebe complex) (Wagstaff and Garnock-Jones, 1998) and *Pittosporum* (Gemmill et al., 2002).

## **Medium and small scale: Relationships between species and taxonomy**

The Australasian *Wahlenbergia* exhibit very low genetic diversity as measured by the chloroplast DNA markers *trnL-F* and *trnK* and the nuclear ribosomal DNA marker. The low genetic diversity (especially in the New Zealand species) is probably due to rapid evolution during a period of geological and climatic change, perhaps coupled with incomplete lineage sorting and hybridisation (Winkworth et al., 1999). While this low level of variation has resulted in poorly resolved phylogenetic trees a certain amount of information regarding relationships among individuals and species has resulted in an increased understanding of the evolution of the Australasian *Wahlenbergia* species.

With greater sampling in Chapter Three, the New Zealand radicate species no longer formed a clade as they had in Chapter Two, although further study is necessary as this could simply be due to the lack of variation in the markers used in this chapter, and does not necessarily point to ongoing gene-flow between Australia and New Zealand *Wahlenbergia*. Other studies on the Australasian flora have also found very low levels of sequence divergence between Australia and New Zealand (e.g. *Craspedia* (Ford et al., 2007), *Sophora* (Hurr et al., 1999; Mitchell and Heenan, 2002) and *Veronica* (Wagstaff and Garnock-Jones, 1998). Although the relationship was unsupported, the Australian and New Zealand radicate species did appear to fall into two separate groups when analysed with the more sensitive AFLP data for instance.

The monophyly of the New Zealand rhizomatous species was confirmed using DNA sequence and AFLP data, although as in Chapter Two, their relationship with other *Wahlenbergia* species remained unresolved. Interestingly there did not appear to be a close relationship between the Australian rhizomatous species included in this study and the New Zealand rhizomatous clade as was hypothesised by Smith (1992). Unfortunately this study did not include any samples of the Tasmanian rhizomatous *W. saxicola*, which was even considered conspecific with the New Zealand rhizomatous species at one time, as this may represent the sister species to the New Zealand rhizomatous clade. Close relationships between Tasmanian and New Zealand species has been found in other plant groups e.g. *Pachycladon* (Heenan et al., 2002).

The Australian *Wahlenbergia* do not form a monophyletic group in any of the DNA sequence analyses, but they do form a cluster in the AFLP analyses. This lack of monophyly perhaps is not surprising given the ancestors of the New Zealand *Wahlenbergia* arrived from Australia only in the last ca. 1 my, meaning reciprocal monophyly may not have had sufficient time to evolve (Shaffer and Thomson, 2007). Interestingly, neither the Australian rhizomatous nor the Australian radicate species form monophyletic groups. Smith (1992) had hypothesised that all of the Australian rhizomatous species were closely related (termed the *W. gloriosa* group), and also split the radicate species into three other groups that he hypothesised as being related to each other, termed the *W. scopulicola*, *W. gracilenta* and *W. communis* groups. Within the *W. communis* group he further identified sub-groups termed the *W. graniticola*, *W. planiflora* and *W. gracilis* groups. This study included no species belonging to the *W. scopulicola* group, and only one belonging to the *W. gracilenta* group (*W. preissii*) so monophyly of these groups was not able to be assessed. However, the *W. preissii* sample had an identical chloroplast haplotype to Australian radicates belonging to the *W. communis* and *W. gracilis* groups, as well as several New Zealand radicate samples (e.g. see placement in Figure 3.5), which indicates there is little evidence to support this hypothesised group. None of the other groups proposed by Smith (1992) were reconstructed as being monophyletic in this thesis either. However, sampling of the Australian species was not extensive in this thesis, and further phylogenetic research is warranted.

**Table 5.1** Synthesis table, showing relationships found among samples of Australian and New Zealand *Wahlenbergia* from the same morphological species in analyses conducted in Chapters Two and Three of this thesis. For the PCoA analysis YES = samples form a group, NO = samples do not form a group. Ecology and chromosome count information from (Peterson, 1997; Plunkett et al., 2009; Smith, 1992).

Morphology	cpDNA tree-building analysis	nrDNA tree-building analysis	AFLP tree-building analysis	AFLP PCoA analyses	Ecology	Chromosome Counts	Synthesis
<b>New Zealand radicates</b>							
<i>W. violacea</i>	Polyphyletic	Paraphyletic	Polyphyletic	NO	Lowland	2n=72	Single species?
<i>W. akaroa</i>	Unresolved	Unresolved	Monophyletic	NO	Lowland	2n=72	
<i>W. ramosa</i>	Polyphyletic	Polyphyletic	Monophyletic	YES	Lowland, coastal	2n=72	
<i>W. rupestris</i>	Polyphyletic	Unresolved	Paraphyletic	NO	Lowland, rivers	2n=72	
<i>W. vernicosa</i>	Monophyletic*	Paraphyletic	Monophyletic-MP*	NO	Lowland, coastal	2n=54	<i>W. vernicosa</i>
<b>New Zealand rhizomatous</b>							
<i>W. mathewsii</i>	Monophyletic	Unresolved	Monophyletic	YES	Limestone	2n=36	Single species?
<i>W. albomarginata</i>	Polyphyletic	Polyphyletic	Polyphyletic	NO	Varied	2n=36	
<i>W. albomarginata</i> subsp. <i>albomarginata</i>	Polyphyletic	Polyphyletic	Polyphyletic	NO	Alpine	2n=36	
<i>W. albomarginata</i> subsp. <i>laxa</i>	Unresolved	N/A	N/A	N/A	Alpine	2n=36	
<i>W. albomarginata</i> subsp. <i>olivina</i>	Monophyletic	Paraphyletic	Monophyletic	NO	Ultramafic	2n=36	
<i>W. albomarginata</i> subsp. <i>flexilis</i>	Unresolved	Paraphyletic	Monophyletic	NO	Limestone	2n=36	
<i>W. pygmaea</i>	Paraphyletic	Polyphyletic	Polyphyletic	NO	Alpine	2n=36	
<i>W. pygmaea</i> subsp. <i>pygmaea</i>	Unresolved	Monophyletic	Monophyletic	NO	Alpine	2n=36	
<i>W. pygmaea</i> subsp. <i>drucei</i>	Unresolved	Monophyletic	Monophyletic	NO	Alpine	2n=36	
<i>W. pygmaea</i> subsp. <i>tararua</i>	Unresolved	Unresolved	Monophyletic	YES	Alpine	2n=36	
<i>W. congesta</i> subsp. <i>haastii</i>	Monophyletic	Unresolved	Monophyletic	YES	Coastal	2n=36	<i>W. congesta</i> subsp. <i>haastii</i>

Morphology	cpDNA tree- building analysis	nrDNA tree- building analysis	AFLP tree- building analysis	AFLP PCoA analyses	Ecology	Chromosome Counts	Synthesis
<i>W. cartilaginea</i>	Monophyletic	Monophyletic	N/A	N/A	Scree	2n=36	<i>W. cartilaginea</i>
Australian radicates							
<i>W. stricta</i>	Polyphyletic	Polyphyletic	N/A	N/A	Lowland	2n=18, 36	?
<i>W. luteola</i>	Polyphyletic	Polyphyletic	N/A	N/A	Lowland	2n=18, 36	?
<i>W. communis</i>	Monophyletic	Monophyletic	N/A	N/A	Lowland	2n=18, 36	<i>W. communis</i>
<i>W. littoricola</i> subsp. <i>littoricola</i>	Paraphyletic	N/A	N/A	N/A	Lowland	2n= 54, 72	?
<i>W. planiflora</i>	Polyphyletic	Polyphyletic	N/A	N/A	Lowland	2n=18, 36, 54	?
<i>W. multicaulis</i>	Paraphyletic	Paraphyletic	N/A	N/A	Lowland	2n=36, 54, 72	?
<i>W. gracilis</i>	Paraphyletic	Monophyletic	N/A	N/A	Lowland	2n=54	?
<i>W. telfordii</i>	Unresolved	Monophyletic	N/A	N/A	Lowland	?	<i>W. telfordii</i>
<i>W. rupicola</i>	Monophyletic	Monophyletic	N/A	N/A	Lowland	?	<i>W. rupicola</i>
Australian rhizomatous							
<i>W. ceracea</i>	Monophyletic	Monophyletic	N/A	N/A	Alpine	2n=18	<i>W. ceracea</i>
<i>W. gloriosa</i>	Paraphyletic	Paraphyletic	N/A	N/A	Varied	2n=18	Single species?
<i>W. densifolia</i>	Monophyletic	Paraphyletic	N/A	N/A	Alpine	2n=18	
Australasian radicate							
<i>W. littoricola</i>	Polyphyletic	Polyphyletic	Polyphyletic	NO	Lowland	2n= 54, 72	Restricted to Australia

\* Excluding sample from Raoul Island (VERN\_Raoul)

Although phylogenetic relationships among *Wahlenbergia* species are poorly resolved and difficult to interpret in this thesis, some relationships at the species level could be clarified. Other studies on the New Zealand flora have found that molecular data (both DNA sequences and AFLP) are useful at the same taxonomic level (e.g. *Ourisia* (Meudt et al., 2009)). The complex phylogenetic relationships exhibited in this study make species delimitation on the basis of sequence data a difficult task. In this case some species hypotheses have been supported, but a larger number of species hypotheses have not been adequately addressed by this dataset as there are too few characters to distinguish between individuals leading to very poorly resolved trees. The unified species concept was adopted in this study, with the central line of investigation involving monophyly at one or multiple DNA loci and AFLP analysis, coupled with morphological, chromosomal and/or ecological distinctness. Table 5.1 displays all of the *Wahlenbergia* species represented by more than one sample in at least one analysis in this thesis, and details what each analysis, or line of evidence, suggests about each species hypothesis.

The only New Zealand radicate species clearly monophyletic based on DNA sequence data in this thesis is the subtropical radicate *W. vernicosa*, which incidentally is probably more closely related to other Australian and New Zealand species, and not *W. littoricola* as hypothesised (de Lange and Cameron, 1999) and therefore there is no evidence to support conspecificity between the two. Regarding the other four (temperate) radicate species, the so-called *W. gracilis* complex, there is very little evidence presented in this thesis to suggest that they should not be considered a single species. None of the four species form monophyletic groups in the DNA sequence data. Two do form monophyletic groups in the AFLP tree building analysis (*W. akaroa* and *W. ramosa*), but it is difficult to say whether this is representing population level structure or species boundaries, and deserves further investigation with additional sampling.

Within the New Zealand rhizomatous clade two morphologically distinctive rhizomatous species have some level of genetic distinction in both the DNA sequence and AFLP analysis (*W. matthewsii* and *W. congesta*), which is consistent with their species ranking. A third morphologically-distinctive rhizomatous species, the scree plant *W. cartilaginea*, was only represented by one sample in the AFLP

analysis, but was found to be monophyletic in the DNA sequencing data and its distinct morphology and habitat all point to recognition at the species level. The other two New Zealand rhizomatous species (*W. albomarginata* and *W. pygmaea*) are indistinguishable based on the DNA sequence and AFLP analyses conducted for this thesis. The two species are also indistinguishable based on seed morphology (Webb and Simpson, 2001) and chromosome count (Petterson, 1997b; Petterson et al., 1995). The main character that separates *W. albomarginata* from *W. pygmaea* is the shape of their flowers. *W. albomarginata* has a ‘narrow-campanulate-rotate corolla, with tube distinctly longer than broad’ whereas in *W. pygmaea* the corolla is ‘broadly campanulate, with tube as wide as, or wider than long’ (Petterson, 1997b). Further study is necessary to determine whether this is simply an artificial North/South Island split between individuals of the same species.

Interestingly although there is no monophyly at the species level several subspecies of both *W. albomarginata* and *W. pygmaea* formed monophyletic groups either in the DNA sequencing analyses or the AFLP analyses (or both). In particular one subspecies of *W. albomarginata* might be a distinct genetic lineage (*W. albomarginata* subsp. *olivina*) as the samples included in this thesis consistently formed clades, it can consistently be identified based on morphological characters, and it has a unique habitat (ultramafic rock).

The diversification of the New Zealand rhizomatous group appears to have been driven by local habitat adaptations. Of the distinct lineages, *W. cartilaginea* is found in scree habitats, *W. matthewsii* grows on limestone and *W. congesta* is a coastal plant. The variation between the subspecies of *W. albomarginata* can also be explained by habitat, for example *W. albomarginata* subsp. *olivina* grows only on ultramafic rock and *W. albomarginata* subsp. *flexilis* grows on limestone.

Multiple samples of Australian species were not included in the AFLP dataset, so there is less evidence pertaining to species delimitation for them. However, several species were shown to be monophyletic based on DNA sequence data alone, including the rhizomatous *W. ceracea* and the two newly described radicate species (*W. rupicola* and *W. telfordii*). Species relationships within the Australian radicate samples are complicated by the fact that they have mostly been shown to contain at least two races of different ploidy levels. Therefore, a dedicated research project into



the Australian (and New Zealand) radicate species is essential to resolve the taxonomy among these species.

## **Assessment of molecular techniques used in this thesis**

In this thesis, amplification and sequencing using 22 primer pairs from all three plant genomes was tested on four *Wahlenbergia* species. Two chloroplast (*trnL-F* & *trnK*) and one nuclear (ITS) marker were selected for phylogenetic reconstruction. These markers were the easiest to amplify, sequence and align and had the highest variability of markers trialed, yet still overall revealed very little genetic variability within the Australasian *Wahlenbergia* species (Table 3.2). Despite this low genetic variability within Australasian *Wahlenbergia* the level of variability between Australasian and South African *Wahlenbergia* using ITS sequences in particular was very high, which made for difficulties in aligning these sequences. This highlights the importance of selecting genetic markers for the particular study in question.

The Australasian phylogeny (Chapter Three) was poorly resolved, due to the low levels of sequence variation, so AFLP analysis was conducted with the hopes that greater variation would be recovered. In the AFLP analysis support values were generally higher towards the tips of the branches than in the DNA sequencing analysis, whereas deeper relationships were equally poorly supported using both AFLP and sequence data, suggesting neither method has the appropriate level of variation to elucidate these relationships. Unfortunately, this means we have not yet managed to discover a molecular technique that is able to resolve deeper relationships within the Australasian *Wahlenbergia*.

## **Future directions**

It may be possible to find a more variable DNA marker (such as low copy nuclear genes), which may be able to provide information at the among species level. Although several low-copy nuclear markers trialled in chapter three could not be successfully sequenced, others should be attempted. Alternatively, ‘next-generation’ sequencing methods could be employed to tackle this question. Next generation

sequencing methods are high-throughput sequencing technologies that parallelize the sequencing process, producing thousands or millions of sequences at once (Shendure and Ji, 2008). Perhaps more bases will reveal greater numbers of variable sites and hence greater resolution? Another method gaining popularity when traditional methods for finding variable DNA sequence markers has failed is one similar to that taken by Padolina (2006) to design low-copy nuclear markers. Padolina (2006) took a computational approach to design universal angiosperm primers by querying the MoBIoS (Miranker et al., 2003) database to compare the genomes of the monocot *O. sativa* and the eudicot *A. thaliana* to search for primer combinations that occurred only once in each of the two genomes. It is becoming ever more feasible to sequence whole chloroplast genomes and although this has mainly been accomplished for model or crop species such as cotton (Lee et al., 2006) to date, phylogenetic studies using fully sequenced chloroplasts are not unknown (Moore et al., 2007). If the chloroplasts of two closely related *Wahlenbergia* samples were fully sequenced, it may be possible to design primers for sequences that contained sufficient variation to create resolved phylogenetic trees, and hence resolve deeper relationships within the Australasian *Wahlenbergia*.

It is also possible that there are no molecular techniques able to resolve the relationships in question, as there may not be the appropriate signal in the genomes. These kinds of molecular techniques rely on mutations that arise by chance, and the probability that mutations will have occurred that are able to correctly assign each individual to a species group must be relatively low, as their radiation (in Australasia) has probably occurred only in the last 3.7 my. Thus, some level of fuzziness and uncertainty in delimiting young species at the beginning of differentiation as shown here is therefore perhaps not unexpected (Shaffer and Thompson 2007). A systematic morphological study of the genus, with extensive sampling and accompanying statistical analysis, may even yield more evidence regarding species boundaries than molecular techniques are able to do.

Further research into large-scale relationships of *Wahlenbergia* should include samples from Asia, the Pacific, South America and Europe. Chromosome counts would be very useful for a larger number of South African species also. At the medium scale (if useful markers can be found) it would be very interesting to assess

the relationships between the Australian and New Zealand radicate species, particularly the Australian *W. communis* group (which includes the Australian *W. gracilis*), and the New Zealand '*W. gracilis*' complex (which does not currently contain any species named *W. gracilis* (Webb and Simpson, 2001)). AFLPs have been shown here to be useful at delimiting New Zealand species of *Wahlenbergia*, and greater sampling of the *W. gracilis* complex, and the *W. albomarginata*/*W. pygmaea* complex is necessary to tease out relationships within these complexes. There are also outstanding questions relating to two uncollected New Zealand taxa, *W. congesta* subsp. *congesta* and *W. albomarginata* subsp. *decora*, which would be good to include in future datasets.

## Conclusions

In this thesis I have reconstructed the first phylogeny of *Wahlenbergia*, including ca. 20% of the genus, with samples from its two main areas of diversity, South Africa and Australasia. *Wahlenbergia* was confirmed to be polyphyletic, although most of these species formed a monophyletic group. *Wahlenbergia* originated in South Africa, dispersed to and radiated in Australia about 3.7 mya, and from there two introductions to New Zealand are hypothesised, one leading to a rhizomatous clade ca. 1 mya, the other to a radicate group ca. 0.5 mya, thus refuting the hypothesis of Gondwanan vicariance for this group. With regards the New Zealand taxonomy of the group there was no evidence to support the hypothesis that *W. vernicosa* was conspecific with the Australian *W. littoricola*, but also no evidence to refute the hypothesis that all members of the *W. gracilis* complex (*W. violacea*, *W. akaroa*, *W. ramosa* and *W. rupestris*) were conspecific. There was evidence to support species rank for three morphologically distinct species (*W. cartilaginea*, *W. matthewsii* and *W. congesta*), but no evidence to refute the conspecificity of the common alpine species *W. albomarginata* and *W. pygmaea*. The nuclear ITS and chloroplast *trnL-F* and *trnK* may not be suitable markers for exploring species relationships within this group, as phylogenies were poorly resolved at deeper levels. However, AFLPs are useful for delimiting New Zealand *Wahlenbergia* at the species level, and will be more useful for future work.

# References

- Abdel-Monem, A.A., Fernandea, L.A., & Boone, G.M. 1975. K-Ar ages from the eastern Azores group (Santa Maria, Sao Miguel and the Formigas Islands). *Lithos* 8:247-254.
- Albach, D.C. 2007. Amplified fragment length polymorphisms and sequence data in the phylogenetic analysis of polyploids: multiple origins of *Veronica cymbalaria* (Plantaginaceae). *New Phytologist* 176:481-498.
- Alfaro, M.E., Zoller, S., & Lutzoni, F. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology and Evolution* 20:255-266.
- Allan, H.H. 1961. Flora of New Zealand. Government Printer, Wellington.
- Alvarez, I., & Wendel, J.F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29:417-434.
- Antonelli, A. 2008. Higher level phylogeny and evolutionary trends in Campanulaceae subfam. Lobelioideae: Molecular signal overshadows morphology. *Molecular Phylogenetics and Evolution* 46:1-18.
- Bakker, F.T., Breman, F., & Merckx, V. 2006. DNA sequence evolution in fast evolving mitochondrial DNA nad1 exons in Geraniaceae and Plantaginaceae. *Taxon* 55:887-896.
- Bergh, N.G., & Linder, H.P. 2009. Cape diversification and repeated out-of-southern-Africa dispersal in paper daisies (Asteraceae–Gnaphalieae). *Molecular Phylogenetics and Evolution* 51:5-18.
- Berry, P.E., Hahn, W.J., Sytsma, K.J., Hall, J.C., & Mast, A. 2004. Phylogenetic relationships and biogeography of *Fuchsia* (Onagraceae) based on noncoding nuclear and chloroplast DNA data. *American Journal of Botany* 91:601-614.
- Bonin, A., Bellemain, E., Eidesen, P.B., Pompanon, F., Brochmann, C., & Taberlet, P. 2004. How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* 13:3261-3273.
- Britten, R. 1986. Rates of DNA sequence evolution differ between taxonomic groups. *Science* 231:1393–1398.

- Broadhurst, L.M., Young, A.G., & Murray, B.G. 2008. AFLPs reveal an absence of geographical genetic structure among remnant populations of pohutukawa (*Metrosideros excelsa*, Myrtaceae). *New Zealand Journal of Botany* 46:13-21.
- Bromham, L., & Penny, D. 2003. The modern molecular clock. *Nature Reviews Genetics* 4:216-224.
- Bush, C.M., Wagstaff, S.J., Fritsch, P.W., & Kron, K.A. 2009. The phylogeny, biogeography and morphological evolution of *Gaultheria* (Ericaceae) from Australia and New Zealand. *Australian Systematic Botany* 22:229-242.
- Cano-Maqueda, J., Talavera, S., Arista, M., & Catalan, P. 2008. Speciation and biogeographical history of the *Campanula lusitanica* complex (Campanulaceae) in the Western Mediterranean region. *Taxon* 57:11252-11266.
- Cellinese, N., Smith, S.A., Edwards, E.J., Kim, S.T., Haberle, R.C., Avramakis, M., & Donoghue, M.J. 2009. Historical biogeography of the endemic Campanulaceae of Crete. *Journal of Biogeography* 36:1253-1269.
- Chandler, G.T., Plunkett, G.M., Pinney, S.M., Cayzer, L.W., & Gemmill, C.E.C. 2007. Molecular and morphological agreement in Pittosporaceae: phylogenetic analysis with nuclear ITS and plastid *trnL-trnF* sequence data. *Australian Systematic Botany* 20:390-401.
- Cho, Y., Mower, J.P., Qiu, Y.L., & Palmer, J.D. 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* 101:17741-17746.
- Cosner, M.E., Raubeson, L.A., & Jansen, R.K. 2004. Chloroplast DNA rearrangements in Campanulaceae: phylogenetic utility of highly rearranged genomes. *Bmc Evolutionary Biology* 4.
- Cowan, R.S., Chase, M.W., Kress, W.J., & Savolainen, V. 2006. 300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants. *Taxon* 55:611-616.
- Crayn, D.M., Rossetto, M., & Maynard, D.J. 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* 93:1328-1342.
- Cuenoud, P., Savolainen, V., Chatrou, L.W., Powell, M., Grayer, R.J., & Chase, M.W. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. *American Journal of Botany* 89:132-144.

- Cummings, M.P., Otto, S.P., & Wakeley, J. 1995. Sampling properties of DNA-sequence data in phylogenetic analysis. *Molecular Biology and Evolution* 12:814-822.
- Cupido, C.N. 2009. Systematic studies of the South African Campanulaceae sensu stricto with an emphasis on generic delimitation. Ph.D. thesis. In. University of Cape Town.
- de Lange, P.J. 2005. A comment on *Wahlenbergia littoricola* subsp. *vernica*. *Wellington Botanical Society Newsletter*.
- de Lange, P.J., & Cameron, E.K. 1999. The vascular flora of Aorangi Island, Poor Knights Islands, northern New Zealand. *New Zealand Journal of Botany* 37:433-468.
- de Lange, P.J., Norton, D.A., Courtney, S.P., Heenan, P.B., Barkla, J.W., Cameron, E.K., Hitchmough, R., & Townsend, A.J. 2009. Threatened and uncommon plants of New Zealand (2008 revision). *New Zealand Journal of Botany* 47:61-96.
- De Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology* 56:879-886.
- Demesure, B., Sodji, N., & Petit, R.J. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4:129-131.
- Després, L., Gielly, L., Redoutet, W., & Taberlet, P. 2003. Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. *Molecular Phylogenetics and Evolution* 27:185-196.
- Doyle, J.J., & Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12:13-15.
- Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., & Wilson, A. 2007. Geneious v3.0 available from <http://www.geneious.com/>.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J., & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *Plos Biology* 4:699-710.
- Drummond, A.J., & Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *Bmc Evolutionary Biology* 7.
- Eddie, W.M.M., Shulkina, T., Gaskin, J., Haberle, R.C., & Jansen, R.K. 2003. Phylogeny of Campanulaceae s. str. inferred from its sequences of nuclear ribosomal DNA. *Annals of the Missouri Botanical Garden* 90:554-575.

- Ennos, R.A., Sinclair, W.T., Hu, X.-S., & Langdon, A. 1999. Using organelle markers to elucidate the history, ecology and evolution of plant populations. in, *Molecular systematics and plant evolution*. Taylor and Francis, London.
- Erixon, P., Svennblad, B., Britton, T., & Oxelman, B. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Systematic Biology* 52:665-673.
- Ford, K.A., Ward, J.M., Smitsen, R.D., Wagstaff, S.J., & Breitwieser, I. 2007. Phylogeny and biogeography of *Craspedia* (Asteraceae : Gnaphalieae) based on ITS, ETS and *psbA-trnH* sequence data. *Taxon* 56:783-794.
- Forster, G.J.A. 1786. *Florulae insularum australium prodromus*, Goettingen.
- Frajman, B., & Schneeweiss, G.M. 2009. A Campanulaceous Fate: the Albanian stenoendemic *Asyneuma comosiforme* in fact belongs to isophyllous *Campanula*. *Systematic Botany* 34:595-601.
- Gallagher, R.V., Hughes, L., & Leishman, M.R. 2009. Phenological trends among Australian alpine species: using herbarium records to identify climate-change indicators. *Australian Journal of Botany* 57:1-9.
- Gardner, R.C., De Lange, P.J., Keeling, D.J., Bowala, T., Brown, H.A., & Wright, S.D. 2004. A late Quaternary phylogeography for *Metrosideros* (Myrtaceae) in New Zealand inferred from chloroplast DNA haplotypes. *Biological Journal of the Linnean Society* 83:399-412.
- Garnock-Jones, P., Albach, D., & Briggs, B.G. 2007. Botanical names in Southern Hemisphere *Veronica* (Plantaginaceae): sect. *Detzneria*, sect. *Hebe*, and sect. *Labiatoidea*. *Taxon* 56:571-582.
- Gemmill, C.E.C., Allan, G.J., Wagner, W.L., & Zimmer, E.A. 2002. Evolution of insular Pacific *Pittosporum* (Pittosporaceae): Origin of the Hawaiian radiation. *Molecular Phylogenetics and Evolution* 22:31-42.
- Goldblatt, P., Savolainen, V., Porteous, O., Sostaric, I., Powell, M., Reeves, G., Manning, J.C., Barraclough, T.G., & Chase, M.W. 2002. Radiation in the Cape flora and the phylogeny of peacock irises *Moraea* (Iridaceae) based on four plastid DNA regions. *Molecular Phylogenetics and Evolution* 25:341-360.
- Gustafsson, M.H.G., & Bremer, K.r. 1995. Morphology and Phylogenetic Interrelationships of the Asteraceae, Calyceraceae, Campanulaceae, Goodeniaceae, and Related Families (Asterales). *American Journal of Botany* 82:250-265.
- Haase, M., Marshall, B., & Hogg, I. 2007. Disentangling causes of disjunction on the South Island of New Zealand: the Alpine fault hypothesis of vicariance revisited. *Biological Journal of the Linnean Society* 91:361-374.

- Haberle, R.C., Dang, A., Lee, T., Penaflor, C., Cortes-Burns, H., Oestreich, A., Raubeson, L., Cellinese, N., Edwards, E.J., Kim, S.T., Eddie, W.M.M., & Jansen, R.K. 2009. Taxonomic and biogeographic implications of a phylogenetic analysis of the Campanulaceae based on three chloroplast genes. *Taxon* 58:715-734.
- Heenan, P.B., Mitchell, A.D., & Koch, M. 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* 40:543-562.
- Hey, J. 2006. On the failure of modern species concepts. *Trends in Ecology & Evolution* 21:447-450.
- Hillis, D.M., & Dixon, M.T. 1991. Ribosomal DNA - molecular evolution and phylogenetic inference. *Quarterly Review of Biology* 66:411-453.
- Holland, B., Clarke, A., & Meudt, H. 2008. Optimizing automated AFLP scoring parameters to improve phylogenetic resolution. *Systematic Biology* 57:347-366.
- Houliston, G.J., & Olson, M.S. 2006. Nonneutral evolution of organelle genes in *Silene vulgaris*. *Genetics* 174:1983-1994.
- Hudson, R.R., & Coyne, J.A. 2002. Mathematical consequences of the genealogical species concept. *Evolution* 56:1557-1565.
- Huelsenbeck, J.P., Ronquist, F. . 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17: 754-755.
- Hughes, C.E., Eastwood, R.J., & Bailey, C.D. 2006. From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical Transactions of the Royal Society B-Biological Sciences* 361:211-225.
- Hurr, K.A., Lockhart, P.J., Heenan, P.B., & Penny, D. 1999. Evidence for the recent dispersal of *Sophora* (Leguminosae) around the Southern Oceans: molecular data. *Journal of Biogeography* 26:565-577.
- Huson, D.H., & Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23:254-267.
- Ishida, T.A., Hattori, K., Sato, H., & Kimura, M.T. 2003. Differentiation and hybridization between *Quercus crispula* and *Q. dentata* (Fagaceae): Insights from morphological traits, amplified fragment length polymorphism markers, and leafminer composition. *American Journal of Botany* 90:769-776.



- Joly, S., Heenan, P.B., & Lockhart, P.J. 2009. A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand. *Molecular Phylogenetics and Evolution* 51:365-372.
- Kamp, P.J.J. 1986. Late Cretaceous-Cenozoic tectonic development of the southwest pacific region. *Tectonophysics* 121:225-251.
- Kelchner, S.A. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden*.
- Knapp, M., Mudaliar, R., Havell, D., Wagstaff, S.J., & Lockhart, P.J. 2007. The drowning of New Zealand and the problem of *Agathis*. *Systematic Biology* 56:862-870.
- Kochanek, J., Steadman, K.J., Probert, R.J., & Adkins, S.W. 2009. Variation in seed longevity among different populations, species and genera found in collections from wild Australian plants. *Australian Journal of Botany* 57:123-131.
- Koopman, M.M., & Ayers, T.J. 2005. Nectar spur evolution in the Mexican *lobelias* (Campanulaceae : Lobelioideae). *American Journal of Botany* 92:558-562.
- Koopman, W.J.M. 2005. Phylogenetic signal in AFLP data sets. *Systematic Biology* 54:197-217.
- Koopman, W.J.M., Wissemann, V., De Cock, K., Van Huylenbroeck, J., De Riek, J., Sabatino, G.J.H., Visser, D., Vosman, B., Ritz, C.M., Maes, B., Werlemark, G., Nybom, H., Debener, T., Linde, M., & Smulders, M.J.M. 2008. AFLP markers as a tool to reconstruct complex relationships: A case study in *Rosa* (Rosaceae). *American Journal of Botany* 95:353-366.
- Kosman, E., & Leonard, K.J. 2005. Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Molecular Ecology* 14:415-424.
- Lammers, T.G. 1996. Phylogeny, biogeography, and systematics of the *Wahlenbergia fernandeziana* complex (Campanulaceae: Campanuloideae). *Systematic Botany* 21:397-415.
- Lee, C., & Wen, J. 2004. Phylogeny of *Panax* using chloroplast *trnC-trnD* intergenic region and the utility of *trnC-trnD* in interspecific studies of plants. *Molecular Phylogenetics and Evolution* 31:894-903.
- Lee, S.-B., Kaittanis, C., Jansen, R., Hostetler, J., Tallon, L., Town, C., & Daniell, H. 2006. The complete chloroplast genome sequence of *Gossypium hirsutum*: organization and phylogenetic relationships to other angiosperms. *BMC Genomics* 7:61.

- Levin, R.A., Blanton, J., & Miller, J.S. 2009. Phylogenetic utility of nuclear nitrate reductase: A multi-locus comparison of nuclear and chloroplast sequence data for inference of relationships among American *Lycieae* (Solanaceae). *Molecular Phylogenetics and Evolution* 50:608-617.
- Lewis, K.B., Carter, L., & Davey, F.J. 1994. The opening of cook strait - interglacial tidal scour and aligning basins at a subduction to transform plate edge *Marine Geology* 116:293-312.
- Li, M., Wunder, J., Bissoli, G., Scarponi, E., Gazzani, S., Barbaro, E., Saedler, H., & Varotto, C. 2008. Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. *Cladistics* 24:727-745.
- Lloyd, D.G., & Yates, J.M.A. 1982. Intra-sexual selection and the segregation of pollen and stigmas in hermaphrodite plants, exemplified by *Wahlenbergia albomarginata* (Campanulaceae). *Evolution* 36:903-913.
- Lockhart, P.J., McLenachan, P.A., Harell, D., Glenney, D., Huson, D., & Jensen, U. 2001. Phylogeny, radiation, and transoceanic dispersal of New Zealand alpine buttercups: Molecular evidence under split decomposition. *Annals of the Missouri Botanical Garden* 88:458-477.
- Lundberg, J., & Bremer, K. 2003. A phylogenetic study of the order Asterales using one morphological and three molecular data sets. *International Journal of Plant Sciences* 164:553-578.
- Luo, R.Y., Hipp, A.L., & Larget, B. 2007. A bayesian model of AFLP marker evolution and phylogenetic inference. *Statistical Applications in Genetics and Molecular Biology* 6.
- MacPhail, M.K. 1997. The New Zealand flora - Entirely long-distance dispersal? Comment. *Journal of Biogeography* 24:113-117.
- Maddison, D.R., Maddison, W.P. 2001. MacClade 4. In. Sinauer Associates, Sunderland.
- Mayden, R.L. 1997. A hierarchy of species concepts: The denouement in the saga of the species problem. Pp. 381-424 in: Claridge, M.F., Dawah, H.A., & Wilson, M.R., (eds), *Species: The units of biodiversity*. Chapman and Hall, London. p 381-424.
- Mayr, E. 1942. Systematics and origin of species. Columbia University Press, New York.
- Meudt, H., & Bayly, M. 2008. Phylogeographic patterns in the Australasian genus *Chionohebe* (*Veronica* sl, Plantaginaceae) based on AFLP and chloroplast DNA sequences. *Molecular Phylogenetics and Evolution* 47:319-338.

- Meudt, H., & Clarke, A. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science* 12:106-117.
- Meudt, H., Lockhart, P., & Bryant, D. 2009. Species delimitation and phylogeny of a New Zealand plant species radiation. *BMC Evol Biol* 9:111.
- Meudt, H.M., & Simpson, B.B. 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* 87:479-513.
- Miranker, D., Xu, W., & Mao, R. 2003. MoBioS: A Metric-Space DBMS to Support Biological Discovery. In: 15th International Conference on Scientific and Statistical Database Management. p 241.
- Mitchell, A.D., & Heenan, P.B. 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* 140:435-441.
- Moore, M.J., Bell, C.D., Soltis, P.S., & Soltis, D.E. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* 104 19363-19368.
- Morgan-Richards, M., Smissen, R.D., Shepherd, L.D., Wallis, G.P., Hayward, J.J., Chan, C.H., Chambers, G.K., & Chapman, H.M. 2009. A review of genetic analyses of hybridisation in New Zealand. *Journal of the Royal Society of New Zealand* 39:15-34.
- Mueller, U.G., & Wolfenbarger, L.L. 1999. AFLP genotyping and fingerprinting. *Trends in Ecology & Evolution* 14:389-394.
- Muller, J. 1981. Fossil pollen records of extant angiosperms. *Botanical Revue* 47.
- Murray, B.G., Datson, P.M., Lai, E.L., Sheath, K.M., & Cameron, E.K. 2004. Polyploidy, hybridization and evolution in *Pratia* (Campanulaceae). *New Zealand Journal of Botany* 42:905-920.
- Nei, M., & Kumar, S. 2000. Molecular evolution and phylogenetics. Oxford University Press, New York.
- Nicolson, D.H., & Fosberg, D.R. 2004. The Forsters and the botany of the second Cook expedition (1772-1775) 2ed. Gantner Verlag, Vienna.
- Odhav, B., Beekrum, S., Akula, U., & Baijnath, H. 2007. Preliminary assessment of nutritional value of traditional leafy vegetables in KwaZulu-Natal, South Africa. *Journal of Food Composition and Analysis* 20:430-435.

- Oxelman, B., Liden, M., & Berglund, D. 1997. Chloroplast rps16 intron phylogeny of the tribe Sileneae (Caryophyllaceae). *Plant Systematics and Evolution* 206:393-410.
- Padolina, J.M. 2006. Phylogenetic Reconstruction of *Phalaenopsis* Using Nuclear and Chloroplast DNA Sequence Data and Using *Phalaenopsis* as a Natural System for Assessing Methods to Reconstruct Hybrid Evolution in Phylogenetic Analyses. PhD Thesis. In. The University of Texas at Austin.
- Park, J.M., Kovacic, S., Liber, Z., Eddie, W.M.M., & Schneeweiss, G.M. 2006. Phylogeny and biogeography of isophyllous species of *Campanula* (Campanulaceae) in the Mediterranean area. *Systematic Botany* 31:862-880.
- Peralta, I., & Spooner, D. 2001. Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*) 1. *American Journal of Botany* 88:1888-1902.
- Perez-Espona, S., Sales, F., Hedge, I., & Moller, M. 2005. Phylogeny and species relationships in *Jasione* (Campanulaceae) with emphasis on the 'montana' complex. *Edinburgh Journal of Botany* 62:29-51.
- Perrie, L.R., Brownsey, P.J., Lockhart, P.J., Brown, E.A., & Large, M.F. 2003. Biogeography of temperate Australasian *Polystichum* ferns as inferred from chloroplast sequence and AFLP. *Journal of Biogeography* 30:1729-1736.
- Perrie, L.R., & Shepherd, L.D. 2009. Reconstructing the species phylogeny of *Pseudopanax* (Araliaceae), a genus of hybridising trees. *Molecular Phylogenetics and Evolution* 52:774-783.
- Peter, C.I., & Johnson, S.D. 2008. Mimics and magnets: The importance of color and ecological facilitation in floral deception. *Ecology* 89:1583-1595.
- Petit, C., Bretagnolle, F., & Felber, F. 1999. Evolutionary consequences of diploid-polyploid hybrid zones in wild species. *Trends in Ecology & Evolution* 14:306-311.
- Petterson, J. 1953. The genus *Wahlenbergia* in New Zealand. Unpublished MSc thesis In. Victoria University of Wellington.
- Petterson, J. 2005. The genus *Wahlenbergia*, (Campanulaceae): the harebells of New Zealand. *Wellington Botanical Society Bulletin* 49:16 - 35.
- Petterson, J.A. 1997a. Identity of the original *Wahlenbergia gracilis* (Campanulaceae) and allied species. An historical review of early collections. *New Zealand Journal of Botany* 35:55-78.
- Petterson, J.A. 1997b. Revision of the genus *Wahlenbergia* (Campanulaceae) in New Zealand. *New Zealand Journal of Botany* 35:9-54.

- Petterson, J.A., Williams, E.G., & Dawson, M.I. 1995. Contributions to a chromosome atlas of the New Zealand flora .34. *Wahlenbergia* (Campanulaceae). *New Zealand Journal of Botany* 33:489-496.
- Plunkett, G.T., Bruhl, J.J., & Telford, I.R.H. 2009. Two new, sympatric species of (Campanulaceae) from the New England Tableland escarpment, New South Wales, Australia. *Australian Systematic Botany* 22:319-331.
- Pole, M. 1994. The New Zealand Flora - entirely long-distance dispersal. *Journal of Biogeography* 21:625-635.
- Posada, D. 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution* 25:1253-1256.
- Prebble, M., Sim, R., Finn, J., & Fink, D. 2005. A Holocene pollen and diatom record from Vanderlin Island, Gulf of Carpentaria, lowland tropical Australia. *Quaternary Res* 64:357-371.
- Rambaut, A., & Drummond, A.J. 2007. Tracer v1.4 available at: <http://beast.bio.ed.ac.uk/Tracer>. In.
- Raven, P.H. 1973. Evolution of subalpine and alpine plant groups in New Zealand *New Zealand Journal of Botany* 11:177-200.
- Reeves, G., Chase, M., Goldblatt, P., Rudall, P., Fay, M., Cox, A., Lejeune, B., & Souza-Chies, T. 2001. Molecular systematics of Iridaceae: evidence from four plastid DNA regions 1. *American Journal of Botany* 88:2074-2087.
- Rivadavia, F., Kondo, K., Kato, M., & Hasebe, M. 2003. Phylogeny of the sundews, *Drosera* (Droseraceae), based on chloroplast *rbcL* and nuclear 18S ribosomal DNA sequences. *American Journal of Botany* 90:123-130.
- Rohlf, F.J. 1990. NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System. Applied Biostatistics, Setauket, New York.
- Roquet, C., Saez, L., Aldasoro, J.J., Susanna, A., Alarcon, M.L., & Garcia-Jacas, N. 2008. Ural Delineation, Molecular Phylogeny and Floral Evolution in *Campanula*. *Systematic Botany* 33:203-217.
- Roquet, C., Sanmartin, I., Garcia-Jacas, N., Saez, L., Susanna, A., Wikstrom, N., & Aldasoro, J.J. 2009. Reconstructing the history of Campanulaceae with a Bayesian approach to molecular dating and dispersal-vicariance analyses. *Molecular Phylogenetics and Evolution* 52:575-587.
- Roth, A.W. 1821. *Novae plantarum species praeserim Indiae orientalis*.

- Sanderson, M.J., Thorne, J.L., Wikstrom, N., & Bremer, K. 2004. Molecular evidence on plant divergence times. *American Journal of Botany* 91:1656-1665.
- Sang, T., Crawford, D.J., & Stuessy, T.F. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). In: *American Journal of Botany*. p 1120-1136.
- Sanmartín, I., Wanntorp, L., & Winkworth, R.C. 2007. West Wind Drift revisited: testing for directional dispersal in the Southern Hemisphere using event-based tree fitting. *Journal of Biogeography* 34:398-416.
- Schrader, H.A. 1814. Plantarum semina in Horto Academico Gottingensi, Anno 1814 collecta.
- Shaffer, H., & Thomson, R. 2007. Delimiting Species in Recent Radiations. *Systematic Biology* 56:896-906.
- Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W.S., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E., & Small, R.L. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. In: *American Journal of Botany*. p 142-166.
- Shendure, J., & Ji, H. 2008. Next-generation DNA sequencing. *Nature Biotechnology* 26:1135-1145.
- Shepherd, L.D., Perrie, L.R., & Brownsey, P.J. 2008. Low-copy nuclear DNA sequences reveal a predominance of allopolyploids in a New Zealand *Asplenium* fern complex. *Molecular Phylogenetics and Evolution* 49:240-248.
- Simpson, G. 1945. Notes on some New Zealand plants and descriptions of new species (No. 4). *Transactions of the Royal Society of New Zealand* 75:187-202.
- Simpson, G. 1952. Notes on some New Zealand plants and descriptions of new species (No. 5). *Transactions and Proceedings of the Royal Society of New Zealand* 79:419-435.
- Small, R.L., Cronn, R.C., & Wendel, J.F. 2004. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17:145-170.
- Smissen, R., & Breitwieser, I. 2008. Species relationships and genetic variation in the New Zealand endemic *Leucogenes* (Asteraceae: Gnaphalieae). *New Zealand Journal of Botany* 46:65-76.
- Smissen, R.D., Breitwieser, I., & Ward, J.M. 2004. Phylogenetic implications of trans-specific chloroplast DNA sequence polymorphism in New Zealand Gnaphalieae (Asteraceae). *Plant Systematics and Evolution* 249:37-53.

- Smith, A.B., & Peterson, K.J. 2002. Dating the time of origin of major clades: molecular clocks and the fossil record. *Annual Review of Earth and Planetary Sciences* 30:65-88.
- Smith, P.J. 1992. A revision of the genus *Wahlenbergia* (Campanulaceae) in Australia. *Telopea* 5:91-175.
- Soltis, E.D., & Soltis, P.S. 2000. Contributions of plant molecular systematics to studies of molecular evolution. *Plant Molecular Biology* 42:45-75.
- Soltis, P.S., & Soltis, D.E. 2009. The role of hybridization in plant speciation. *Annual Review of Plant Biology* 60:561-588.
- Steele, P.R., Guisinger-Bellian, M., Linder, C.R., & Jansen, R.K. 2008. Phylogenetic utility of 141 low-copy nuclear regions in taxa at different taxonomic levels in two distantly related families of rosids. *Molecular Phylogenetics and Evolution* 48:1013-1026.
- Sterck, L., Rombauts, S., Vandepoele, K., Rouze, P., & Van de Peer, Y. 2007. How many genes are there in plants (. . . and why are they there)? *Current Opinion in Plant Biology* 10:199-203.
- Stöckler, K., Daniel, I.L., & Lockhart, P.J. 2002. New Zealand Kauri (*Agathis australis* (D. don) Lindl., Araucariaceae) survives oligocene drowning. *Systematic Biology* 51:827-832.
- Swofford, D.L. 2002. PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods). Version 4. In. Sinauer Associates Sunderland, Massachusetts.
- Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*.
- Tate, J.A., & Simpson, B.B. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. In: Syst Bot. p 723-737.
- Tay, M.L., Meudt, H.M., Garnock-Jones, P.J., & Ritchie, P.A. 2010. DNA sequences from three genomes reveal multiple long-distance dispersals and non-monophyly of sections in Australasian *Plantago* (Plantaginaceae). *Australian Systematic Botany* 23:47-68.
- Thorsen, M.J., Dickinson, K.J.M., & Seddon, P.J. 2009. Seed dispersal systems in the New Zealand flora. *Perspectives in Plant Ecology, Evolution and Systematics* 11:285-309.

- Trewick, S.A., & Morgan-Richards, M. 2005. After the deluge: mitochondrial DNA indicates Miocene radiation and Pliocene adaptation of tree and giant weta (Orthoptera : Anostostomatidae). *Journal of Biogeography* 32:295-309.
- Trewick, S.A., Paterson, A.M., & Campbell, H.J. 2007. Guest Editorial: Hello New Zealand. *Journal of Biogeography*.
- van Steenis, C.G.G.J. 1960. Original generic description and typification of the genus 8668 *Wahlenbergia* Schrad. ex Roth (1921) Campanulaceae. *Taxon* 9:122.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., & Zabeau, M. 1995. AFLP - a new technique for DNA-fingerprinting *Nucleic Acids Research* 23:4407-4414.
- Vriesendorp, B., & Bakker, F.T. 2005. Reconstructing patterns of reticulate evolution in angiosperms: what can we do? *Taxon* 54:593-604.
- Wagstaff, S.J., Bayly, M.J., & Garnock-Jones, P.J. 2002. Classification, origin, and diversification of the New Zealand hebes (Scrophulariaceae). *Annals of the Missouri Botanical Garden*.
- Wagstaff, S.J., & Garnock-Jones, P.J. 1998. Evolution and biogeography of the *Hebe* complex (Scrophulariaceae) inferred from ITS sequences. *New Zealand Journal of Botany* 36:425-437.
- Wagstaff, S.J., & Garnock-Jones, P.J. 2000. Patterns of diversification in *Chionohebe* and *Parahebe* (Scrophulariaceae) inferred from ITS sequences. *New Zealand Journal of Botany* 38:389-407.
- Wagstaff, S.J., Heenan, P.B., & Sanderson, M.J. 1999. Classification, origins, and patterns of diversification in New Zealand Carmichaelinae (Fabaceae). *American Journal of Botany* 86:1346-1356.
- Wagstaff, S.J., Molloy, B.P.J., & Tate, J.A. 2010. Evolutionary significance of long distance dispersal and hybridisation in the New Zealand endemic genus *Hoheria* (Malvaceae). *in Press*.
- Wallis, G.P., & Trewick, S.A. 2009. New Zealand phylogeography: evolution on a small continent. *Molecular Ecology* 18:3548-3580.
- Walsh, N.G. 1999. *Wahlenbergia*. in: Walsh, N.G., Entwistle, T.J., (ed), *Flora of Victoria Volume 4: Dicotyledons (Cornaceae to Asteraceae)*. Inkata Press, Australia.
- Wardle, P. 1963. Evolution and distribution of the New Zealand flora, as affected by quaternary climates. *New Zealand Journal of Botany* 1:3-17.



- Waters, J.M., & Roy, M.S. 2004. Out of Africa: The slow train to Australasia. *Systematic Biology* 53:18-24.
- Webb, C.J., & Simpson, M.J.A. 2001. Seeds of New Zealand gymnosperms & dicotyledons. Manuka Press, Christchurch.
- White, T. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315-322 in: M Innis, C.G., J Sninsky, T White, (ed), *PCR protocols: a guide to methods and applications*. Academic Press, San Diego. p 315-322.
- Wikström, N., Savolainen, V., & Chase, M.W. 2001. Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society B-Biological Sciences* 268:2211-2220.
- Wikström, N., Savolainen, V., & Chase, M.W. 2003. Angiosperm divergence times: congruence and incongruence between fossils and sequence divergence estimates. Pp. 142-165 in: Donoghue, P.C.J., & Smith, M.P., (eds), *Telling the Evolutionary Time: Molecular Clocks and the Fossil Record*. p 142-165.
- Wilcox, T.P., Zwickl, D.J., Heath, T.A., & Hillis, D.M. 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. *Molecular Phylogenetics and Evolution* 25:361-371.
- Winkworth, Robertson, A.W., & Ehrendorfer, F. 1999. The importance of dispersal and recent speciation in the flora of New Zealand. *Journal of Biogeography*.
- Winkworth, R., Grau, J., Robertson, A., & Lockhart, P. 2002a. The origins and evolution of the genus *Myosotis* L.(Boraginaceae). *Molecular Phylogenetics and Evolution* 24:180-193.
- Winkworth, R., Wagstaff, S.J., Glenney, D., & Lockhart, P. 2002b. Plant dispersal NEWS from New Zealand. *Trends in Ecology & Evolution* 17:514-520.
- Winkworth, R.C., Wagstaff, S.J., & Glenney, D. 2005. Evolution of the New Zealand mountain flora: Origins, diversification and dispersal. *Organisms Diversity & Evolution*.
- Wolfe, A.D., & Randle, C.P. 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: implications for plant molecular systematics. *Systematic Botany* 29:1011–1020.
- Woo, V., L. . 2007. Evolution, biogeography and taxonomy of the southwest Pacific Tribe Coronanthereae (Gesneriaceae). PhD thesis In: Botany. Victoria University of Wellington.

- Woolley, S.M., Posada, D., & Crandall, K.A. 2008. A Comparison of Phylogenetic Network Methods Using Computer Simulation. *Plos One* 3.
- Yoder, A.D., Irwin, J.A., & Payseur, B.A. 2001. Failure of the ILD to determine data combinability for slow loris phylogeny. *Systematic Biology* 50:408-424.
- Zimmer, E., Roalson, E., Skog, L., Boggan, J., & Idnurm, A. 2002. Phylogenetic relationships in the Gesnerioideae (Gesneriaceae) based on nrDNA ITS and cpDNA trn LF and trn ET spacer region sequences 1. *American Journal of Botany* 89:296-311.
- Zwickl, D.J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence data sets under the maximum likelihood criterion. Ph.D. Thesis In. The University of Texas Austin.

**Appendix 1.** Table showing details of study group for each analyses. The table is in alphabetical order by tag name. Details of collection location, collector, collection number and herbarium where the voucher specimen is located are provided. JMP = Jessie Prebble. Literature citations and GenBank numbers are given for previously published sequences.

Tag name	Species	Collection details or literature citation	Chapter two		Chapter three		Chapter four	
			ITS	trnL-F	ITS	trnL-F	trnK	AFLP
Outgroup Samples								
AZOR_VIDA	<i>Azorina vidalii</i>	ITS (Eddie et al., 2003), trnL-F (Roquet et al., 2008)	AY3222007	EF088696	No	No	No	No
CAMP_DIVA	<i>Campanula divaricata</i>	ITS (Eddie et al., 2003), trnL-F (Roquet et al., 2008)	AY3222014	EF088718	No	No	No	No
CAMP_LATI	<i>C. latifolia</i>	ITS (Eddie et al., 2003), trnL-F (Roquet et al., 2008)	AY3222024	EF088732	No	EF088732	No	No
CAMP_MEDI	<i>C. medium</i>	New Zealand, Wellington, Mahina Bay garden. JMP 09131. WELT	Yes	Yes	Yes	Yes	Yes	No
CAMP_POLY	<i>C. polyclada</i>	ITS and trnL-F (Roquet et al., 2008)	EF090539	EF088746	No	No	No	No
CAMP_POSC	<i>C. poscharskyana</i>	New Zealand, Wellington, Mahina Bay garden. JMP 09130. WELT	Yes	Yes	Yes	Yes	Yes	No
CAMP_PYRA	<i>C. pyramidalis</i>	ITS (Park et al., 2006), trnL-F (Roquet et al., 2008)	DQ304606	EF088754	No	No	No	No
CAMP_RAPU	<i>C. rapunculoides</i>	New Zealand, Wellington, Mahina Bay garden. JMP 09132. WELT	No	Yes	Yes	Yes	Yes	No
CAMP_ROTU	<i>C. rotundifolia</i>	ITS (Cano-Maqueda et al., 2008), trnL-F (Roquet et al., 2008)	FM212736	EF088759	No	EF088759	No	No
CANA_CANa	<i>Canarina canariensis</i>	ITS (Eddie et al., 2003), trnL-F (Roquet et al., 2008)	AY3222045	EF088777	No	No	No	No
JASI_HELD	<i>Jasione heldreichii</i>	ITS (Perez-Espona et al., 2005), trnL-F (Roquet et al., 2008)	DQ2222834	EF088781	No	EF088781	No	No
JASI_MONT	<i>J. montana</i>	ITS (Park et al., 2006), trnL-F (Antonelli, 2008)	DQ304566	DQ356174	No	DQ356174	No	No
LOBE_ANCE	<i>Lobelia anceps</i>	New Zealand, Kapati Coast, Queen Elizabeth Two Park H. Carson 10. WELTU	No	Yes	No	Yes	No	No
LOBE_CARD	<i>L. cardinalis</i>	ITS (Koopman and Ayers, 2005), trnL-F (Antonelli, 2008)	AY350630	DQ356231	No	No	No	No
LOBE_ERIN	<i>L. erinus</i>	New Zealand, Wellington, Mahina Bay garden. JMP 09133. WELT	Yes	Yes	Yes	Yes	No	No
MUSS_AURE	<i>Musschia aurea</i>	ITS (Eddie et al., 2003), trnL-F (Roquet et al., 2008)	AY3222067	EF088785	No	No	No	No
PHYS_COMO	<i>Physoplexis comosa</i>	ITS (Park et al., 2006), trnL-F (Frajman and Schneeweiss, 2009)	DQ304585	FJ426586	No	No	No	No

Tag name	Species	Collection details or literature citation	Chapter two		Chapter three		Chapter four
			ITS	trnL-F	ITS	trnL-F	trnK AFLP
Ingroup samples							
ACAU_SA_NthC	<i>Wahlenbergia acaulis</i>	South Africa, Northern Cape, Kamiesberg. C. Cupido 267. NGB	Yes	Yes	No	No	No
ADPR_SA_WstC	<i>W. adpressa</i>	South Africa, Western Cape, Hopefield. C. Cupido 210	Yes	Yes	No	No	No
AKAR_SI_Bank1	<i>W. akaroa</i>	New Zealand, Banks Peninsula, Palm Gully. H. Wilson 1. WELTU	No	No	Yes	Yes	Yes
AKAR_SI_Bank2	<i>W. akaroa</i>	New Zealand, Banks Peninsula, Palm Gully. H. Wilson 2. WELTU	Yes	Yes	Yes	Yes	Yes
ALBO_SI_Garv	<i>W. albomarginata</i> subsp. <i>albomarginata</i>	New Zealand, Southland, Garvie Mts. JMP 0990. WELT	No	No	Yes	Yes	Yes
ALBO_SI_Aspi	<i>W. albomarginata</i> subsp. <i>albomarginata</i>	New Zealand, Mt Aspiring NP, Shotover Saddle. JMP 0987. WELT	No	No	No	Yes	Yes
ALBO_SI_Cobb	<i>W. albomarginata</i> subsp. <i>albomarginata</i>	New Zealand, Cobb valley, Waingaro Peak. JMP 0976. WELT	No	No	Yes	Yes	No
ALBO_SI_Cook	<i>W. albomarginata</i> subsp. <i>albomarginata</i>	New Zealand, Canterbury, Mt Cook. JMP 0994. WELT	Yes	Yes	Yes	Yes	No
ALBO_SI_Mole	<i>W. albomarginata</i> subsp. <i>albomarginata</i>	New Zealand, Marlborough, Rainbow/Molesworth boundary. B. Rodriguez 209d. WELTU	No	No	Yes	Yes	Yes
ALBO_SI_Nels	<i>W. albomarginata</i> subsp. <i>albomarginata</i>	New Zealand, Nelson Lakes, Mt Robert. JMP 0983. WELT	No	No	Yes	Yes	Yes
ANDR_SA_WstC	<i>W. androsacea</i>	South Africa, Western Cape, Melkbos. C. Cupido 183. NGB	Yes	Yes	No	No	No
ANNU_SA_WstC	<i>W. annularis</i>	South Africa, Western Cape, Elandsbaai. C. Cupido 251. NGB	Yes	Yes	No	No	No
AXIL_SA_WstC	<i>W. axillaris</i>	South Africa, Western Cape, Bredasdorp. C. Cupido 107. NGB	Yes	Yes	No	No	No
CAPE_SA_WstC	<i>W. capensis</i>	South Africa, Western Cape, Malmesbury. C. Cupido 184. NGB	Yes	Yes	No	No	No
CAPI_SA_WstC	<i>W. capillacea</i>	South Africa, Western Cape, Uniondale. C. Cupido 313. NGB	Yes	Yes	No	No	No
CART_SI_Hanm1	<i>W. cartilaginea</i>	New Zealand, Hamner Springs, Island Saddle. JMP 09117a. WELT	No	No	Yes	Yes	No
CART_SI_Hanm2	<i>W. cartilaginea</i>	New Zealand, Hamner Springs, Island Saddle. JMP 09117b. WELT	No	No	Yes	Yes	Yes
CART_SI_Hanm3	<i>W. cartilaginea</i>	New Zealand, Hamner Springs, Island Saddle. JMP 09117c. WELT	Yes	Yes	Yes	Yes	No
CERA_NSW_Blue	<i>W. ceracea</i>	Australia, NSW, Kosciuszko NP, Blue Lake. JMP 0941. WELT	No	No	Yes	Yes	No

Tag name	Species	Collection details or literature citation	Chapter two		Chapter three		Chapter four	
			ITS	trnL-F	ITS	trnL-F	trnK	AFLP
CERA_NSW_Kosc	<i>W. ceracea</i>	Australia, NSW, Kosciuszko NP, summit Mt Kosciuszko JMP 0944. WELT	Yes	Yes	Yes	Yes	Yes	No
CERA_NSW_Renn	<i>W. ceracea</i>	Australia, NSW, Kosciuszko NP, Rennix walk. JMP 0937d. WELT	No	No	Yes	Yes	Yes	Yes
CERN_SA_WstC	<i>W. cernua</i>	South Africa, Western Cape, Cape Town. C. Cupido 188. NGB	Yes	Yes	No	No	No	No
CINE_SA_WstC	<i>W. cinerea</i>	South Africa, Western Cape, Genadendal. C. Cupido 222. NGB	Yes	Yes	No	No	No	No
COMM_NSW_Coom	<i>W. communis</i>	Australia, NSW, Cooma. JMP 0964a. WELT	Yes	Yes	Yes	Yes	Yes	No
COMM_NSW_Quea	<i>W. communis</i>	Australia, NSW, Queanbeyan River. JMP 0960a. WELT	No	No	Yes	Yes	Yes	No
CUSP_SA_Kwa	<i>W. cuspidata</i>	South Africa, KwaZulu-Natal, Himeville. C. Cupido 302. NGB	Yes	Yes	No	No	No	No
DENS_NSW_Renn1	<i>W. densifolia</i>	Australia, NSW, Kosciuszko NP, Rennix walk. JMP 0934. WELT	Yes	Yes	Yes	Yes	Yes	No
DENS_NSW_Renn2	<i>W. densifolia</i>	Australia, NSW, Kosciuszko NP, Rennix walk. JMP 0951. WELT	No	No	No	Yes	Yes	No
DEPR_SA_Free	<i>W. depressa</i>	South Africa, Free State, Barker's Kop. Roux 3350. NGB	Yes	Yes	No	No	No	No
DESM_SA_WstC	<i>W. desmantha</i>	South Africa, Western Cape, Albertina. C. Cupido 310. NGB	Yes	Yes	No	No	No	No
DRUC_NI_Tara1	<i>W. pygmaea subsp. drucei</i>	New Zealand, Mt Taranaki, Wilkies pools. JMP 09112. WELT	Yes	Yes	Yes	Yes	Yes	Yes
DRUC_NI_Tara2	<i>W. pygmaea subsp. drucei</i>	New Zealand, Mt Taranaki, North Egmont Rd end. JMP 09107. WELT	No	No	Yes	Yes	Yes	Yes
DRUC_NI_Tara3	<i>W. pygmaea subsp. drucei</i>	New Zealand, Mt Taranaki, North Egmont Rd end. JMP 09106. WELT	No	No	Yes	Yes	Yes	No
ECKL_SA_WstC	<i>W. ecklonii</i>	South Africa, Western Cape, Paarl. C. Cupido 206. NGB	Yes	Yes	No	No	No	No
EXIL_SA_WstC	<i>W. exilis</i>	South Africa, Western Cape, Malmesbury. C. Cupido 81. NGB	Yes	Yes	No	No	No	No
FLEX_SI_Flax1	<i>W. albomarginata subsp. flexilis</i>	New Zealand, Marlborough, Flaxbourne River mouth. JMP 09118a. WELT	No	No	Yes	Yes	Yes	Yes
FLEX_SI_Flax2	<i>W. albomarginata subsp. flexilis</i>	New Zealand, Marlborough, Flaxbourne River mouth. JMP 09118b. WELT	No	No	Yes	Yes	Yes	Yes
FLEX_SI_Flax3	<i>W. albomarginata subsp. flexilis</i>	New Zealand, Marlborough, Flaxbourne River mouth. JMP 09118c. WELT	Yes	Yes	Yes	Yes	Yes	No
FRUT_SA_WstC	<i>W. fruticosa</i>	South Africa, Western Cape, Riversdie. C. Cupido 311. NGB	Yes	Yes	No	No	No	No

Tag name	Species	Collection details or literature citation	Chapter two		Chapter three		Chapter four	
			ITS	trnL-F	ITS	trnL-F	trnK	AFLP
GLOR_NSW_Digg	<i>W. gloriosa</i>	Australia, NSW, Kosciuszko NP, Diggers Ck. JMP 0939. WELT	Yes	Yes	Yes	Yes	Yes	No
GLOR_NSW_Renn	<i>W. gloriosa</i>	Australia, NSW, Kosciuszko NP, Rennix walk. JMP 0953. WELT	No	No	Yes	Yes	Yes	No
GLOR_NSW_Turo	<i>W. gloriosa</i>	Australia, NSW, Tuross Falls Rd. JMP 0960a. WELT	No	No	Yes	Yes	Yes	Yes
GRAC_NSW_Aber	<i>W. gracilis</i>	Australia, NSW, Abercrombie Caves. JMP 0913. WELT	No	No	No	Yes	Yes	No
GRAC_NSW_Blue1	<i>W. gracilis</i>	Australia, NSW, Blue Mts. JMP 0901. WELT	No	No	No	Yes	Yes	Yes
GRAC_NSW_Blue2	<i>W. gracilis</i>	Australia, NSW, Blue Mts. JMP 0902. WELT	Yes	Yes	Yes	Yes	Yes	No
HAAS_SI_Ship1	<i>W. congesta subsp. haastii</i>	New Zealand, West Coast, Ship Ck. JMP 0986a. WELT	No	No	Yes	Yes	Yes	Yes
HAAS_SI_Ship2	<i>W. congesta subsp. haastii</i>	New Zealand, West Coast, Ship Ck. JMP 0986b. WELT	No	No	Yes	Yes	Yes	Yes
HAAS_SI_Ship3	<i>W. congesta subsp. haastii</i>	New Zealand, West Coast, Ship Ck. JMP 0986c. WELT	Yes	Yes	Yes	Yes	Yes	No
HEDE_Europe	<i>W. hederacea</i>	Europe. ITS (Eddie et al., 2003), trnL-F (Roquet et al., 2008)	AY322080	EF088792	No	EF088792	No	No
HUTT_SA_Kwa	<i>W. huttonii</i>	South Africa, KwaZulu-Natal, Himeville. C. Cupido 304.	Yes	Yes	No	No	No	No
KREB_SA_EstC	<i>W. krebsii</i>	South Africa, Eastern Cape, Hogsback. C. Cupido 294. NGB	Yes	Yes	No	No	No	No
LAXA_SI_Cobb1	<i>W. albomarginata subsp. laxa</i>	New Zealand, Cobb valley, valley floor JMP 0978. WELT	No	No	No	Yes	Yes	Yes
LAXA_SI_Cobb2	<i>W. albomarginata subsp. laxa</i>	New Zealand, Cobb valley, valley floor JMP 0979. WELT	Yes	Yes	Yes	Yes	Yes	No
LITT_NSW_Hart	<i>W. littoricola</i>	Australia, NSW, Hartley Reserve. JMP 0905. WELT	No	No	No	Yes	Yes	Yes
LITT_NSW_Turo	<i>W. littoricola</i>	Australia, NSW, Tuross Falls carpark. JMP 0961. WELT	No	No	Yes	Yes	Yes	No
LITT_VIC_Buff1	<i>W. littoricola</i>	Australia, Victoria, Mt Buffalo. J. Petterson 1. WELT	No	Yes	No	Yes	Yes	No
LITT_VIC_Buff2	<i>W. littoricola</i>	Australia, Victoria, Mt Buffalo. J. Petterson 2. WELT	No	No	No	Yes	Yes	No
LITT_VIC_Buff3	<i>W. littoricola</i>	Australia, Victoria, Mt Buffalo. J. Petterson 3. WELT	No	No	No	Yes	Yes	No
LITT_VIC_Buff4	<i>W. littoricola</i>	Australia, Victoria, Mt Buffalo. J. Petterson 4. WELT	No	No	No	Yes	Yes	No
LOBE_Europe	<i>W. lobeloides</i>	Europe. trnL-F (Roquet et al., 2008)	No	EF088793	No	EF088793	No	No
LONG_SA_WstC	<i>W. longifolia</i>	South Africa, Western Cape, Darling. C. Cupido 212. NGB	Yes	Yes	No	No	No	No
LUTE_NSW_Aber	<i>W. luteola</i>	Australia, NSW, Abercrombie Caves. JMP 0911. WELT	Yes	Yes	Yes	Yes	Yes	No
LUTE_NSW_Coom	<i>W. luteola</i>	Australia, NSW, Cooma. JMP 0966. WELT	No	No	No	Yes	Yes	No

Tag name	Species	Collection details or literature citation	Chapter two		Chapter three		Chapter four	
			ITS	trnL-F	ITS	trnL-F	trnK	AFLP
LUTE_NSW_Quea	<i>W. luteola</i>	Australia, NSW, Queanbeyan River. JMP 0928a. WELT	No	No	Yes	Yes	Yes	No
MATT_SI_Marl1	<i>W. matthewsii</i>	New Zealand, Marlborough, Woodside Ck. P.	No	No	Yes	Yes	Yes	Yes
MATT_SI_Marl2	<i>W. matthewsii</i>	Garnock-Jones 2769a. WELTU						
MATT_SI_Marl3	<i>W. matthewsii</i>	New Zealand, Marlborough, Woodside Ck. P.	No	No	Yes	Yes	Yes	Yes
MULT_NSW_Aber	<i>W. multicaulis</i>	Garnock-Jones 2769b. WELTU						
MULT_NSW_Blan	<i>W. multicaulis</i>	New Zealand, Marlborough, Woodside Ck. P.	Yes	Yes	Yes	Yes	Yes	No
MULT_NSW_Wast	<i>W. multicaulis</i>	Garnock-Jones 2769c. WELTU						
NEOR_SA_WstC	<i>W. neoridiga</i>	Australia, NSW, Abercrombie Caves. JMP 0912. WELT	No	No	No	Yes	Yes	No
OLIV_SI_Dun1	<i>W. albomarginata</i> subsp. <i>olivina</i>	Australia, NSW, Blanket Flat. JMP 0920. WELT	No	No	Yes	Yes	Yes	No
OLIV_SI_Dun2	<i>W. albomarginata</i> subsp. <i>olivina</i>	Australia, NSW, Kosciuszko NP, Waste Pt Rd. JMP 0931. WELT	No	No	Yes	Yes	Yes	No
OLIV_SI_Dun3	<i>W. albomarginata</i> subsp. <i>olivina</i>	South Africa, Western Cape, Prins Albert. C. Cupido 278. NGB	Yes	Yes	No	No	No	No
OXY_SA_WstC	<i>W. oxyphylla</i>	New Zealand, Nelson area, Dun Mts. JMP 0973a. WELT	No	No	Yes	Yes	Yes	Yes
PANI_SA_WstC	<i>W. paniculata</i>	New Zealand, Nelson area, Dun Mts. JMP 0973b. WELT	No	No	Yes	Yes	Yes	Yes
PARV_SA_WstC	<i>W. parvifolia</i>	New Zealand, Nelson area, Dun Mts. JMP 0973c. WELT	Yes	Yes	Yes	Yes	Yes	No
PLAN_NSW_Ngar	<i>W. planiflora</i>	South Africa, Western Cape, Vanrhynsdorp. C. Cupido 259	Yes	Yes	No	No	No	No
PLAN_NSW_Nume	<i>W. planiflora</i>	South Africa, West Cape, Yzerfontein. C. Cupido 181. NGB	Yes	Yes	No	No	No	No
PLAN_NSW_Thred	<i>W. planiflora</i>	South Africa, Western Cape, Cape Town. C. Cupido 119. NGB	Yes	Yes	No	No	No	No
POLY_SA_WstC	<i>W. polyantha</i>	Australia, NSW, Kosciuszko NP, Ngarigo camp ground. JMP 0949. WELT	No	No	Yes	Yes	Yes	No
PREI_WA_Mund	<i>W. preissii</i>	Australia, NSW, Numerella. JMP 0955a. WELT	No	No	Yes	Yes	Yes	No
PROC_SA_WstC	<i>W. procumbens</i>	Australia, NSW, Kosciuszko NP, Thredbo Diggings. JMP 0947. WELT	No	No	Yes	Yes	Yes	No
PSAM_SA_WstC	<i>W. psammophila</i>	South Africa, Western Cape, Albertina. C. Cupido 287. NGB	Yes	Yes	No	No	No	No
		Australia, Western Australia, S of Perth, Mundaring Weir. J. Petterson. WELT	No	No	No	Yes	Yes	No
		South Africa, Western Cape, Napier. C. Cupido 244. NGB	Yes	Yes	No	No	No	No
		South Africa, Western Cape, Vanrhynsdorp. C. Cupido 260. NGB	Yes	Yes	No	No	No	No



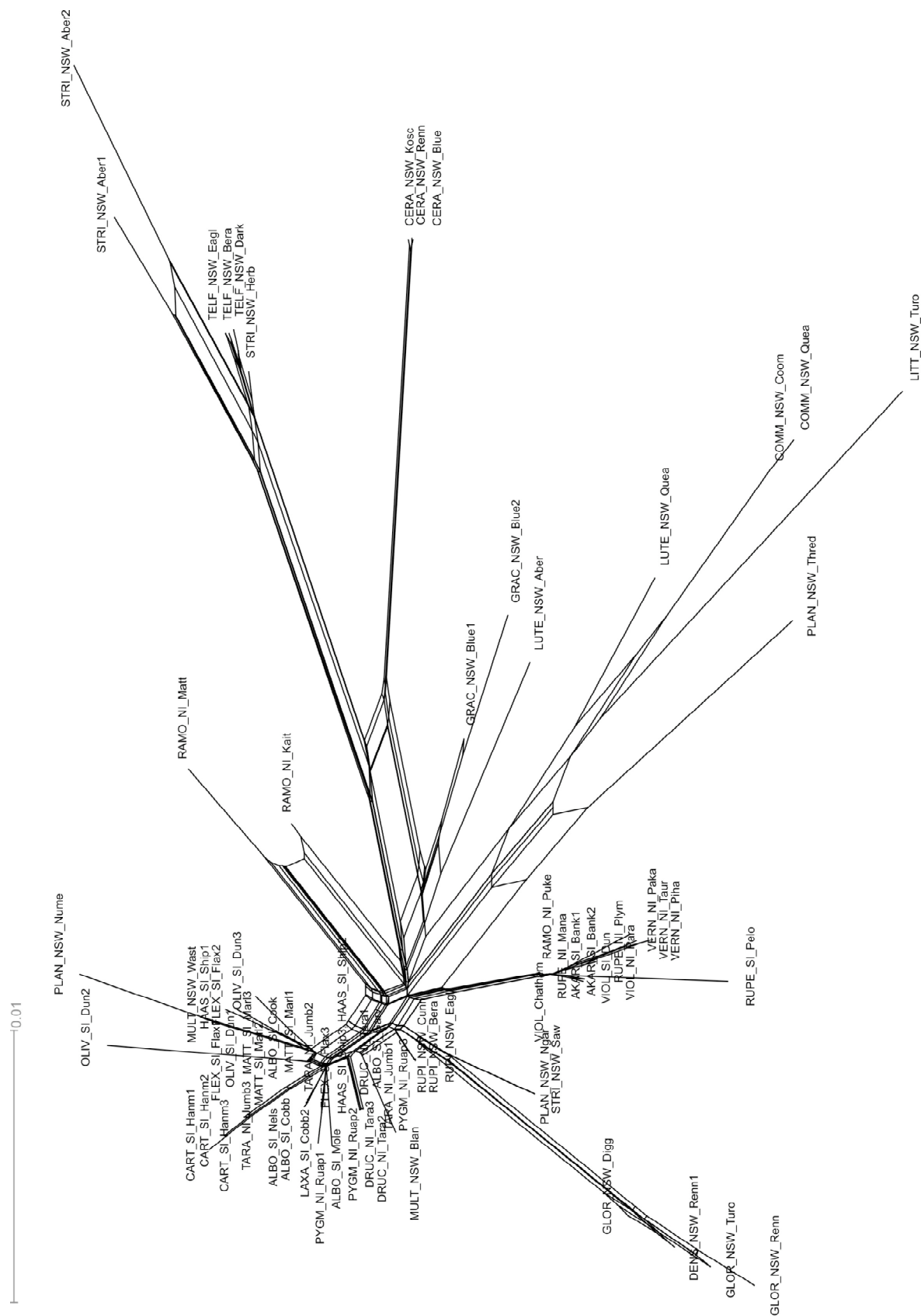
Tag name	Species	Collection details or literature citation	Chapter two		Chapter three		Chapter four
			ITS	trnL-F	ITS	trnL-F	
PYGM_NI_Ruap1	<i>W. pygmaea</i> subsp. <i>pygmaea</i>	New Zealand, Central Plateau, Mt Ruapehu, Ohakune Rd. PGJ 2770a. WELTU	Yes	Yes	Yes	Yes	Yes
PYGM_NI_Ruap2	<i>W. pygmaea</i> subsp. <i>pygmaea</i>	New Zealand, Central Plateau, Mt Ruapehu, Ohakune Rd. PGJ 2770b. WELTU	No	No	Yes	Yes	Yes
PYGM_NI_Ruap3	<i>W. pygmaea</i> subsp. <i>pygmaea</i>	New Zealand, Central Plateau, Mt Ruapehu, Ohakune Rd. PGJ 2770c. WELTU	No	No	Yes	Yes	No
RAMO_NI_Kait	<i>W. ramosa/rupestris?</i>	New Zealand, Akatarawa, Kaitoke. JMP 0857. WELT	No	No	Yes	Yes	No
RAMO_NI_Maka	<i>W. ramosa</i>	New Zealand, Wellington, Makara beach. JMP 09114. WELT	No	No	No	Yes	Yes
RAMO_NI_Matt	<i>W. ramosa/rupestris?</i>	New Zealand, Orongorongo Valley, Mt Matthews Saddle. JMP 0848. WELT	No	No	Yes	Yes	No
RAMO_NI_Puke	<i>W. ramosa</i>	New Zealand, Wellington, Pukerua Bay. JMP 09105a. WELT	Yes	Yes	Yes	Yes	Yes
RUPE_NI_Mana	<i>W. rupestris</i>	New Zealand, Wairarapa, Manawatu Gorge. JMP 0997. WELT	No	No	Yes	Yes	Yes
RUPE_NI_Paha	<i>W. rupestris</i>	New Zealand, Wairarapa, Pahaoa Gorge. JMP 09101. WELT	No	No	No	Yes	Yes
RUPE_NI_Plym	<i>W. rupestris</i>	New Zealand, New Plymouth, meeting of the Waters Scenic Reserve. JMP09109a. WELT	Yes	Yes	Yes	Yes	No
RUPE_SI_Pelo	<i>W. rupestris</i>	New Zealand, Marlborough, Pelorus Bridge. JMP 0970. WELT	No	No	Yes	Yes	No
RUPI_NSW_Bera	<i>W. rupicola</i>	Australia, NSW, New England NP, Berarngutta rest area. G. Plunkett 25. NE	No	No	Yes	Yes	No
RUPI_NSW_Cunn	<i>W. rupicola</i>	Australia, NSW, Cunnawarra NP, Beech lookout. G. Plunkett 23. NE	Yes	Yes	Yes	Yes	Yes
RUPI_NSW_Eagl	<i>W. rupicola</i>	Australia, NSW, New England NP, Eagles Nest walk. G. Plunkett 28. NE	No	No	Yes	Yes	No
SANI_SA_Kwa	<i>W. sp. 'Sani Rd'</i>	South Africa, KwaZulu-Natal, Sani Road. C. Cupido 309. NGB	No	Yes	No	No	No
STRI_NSW_Aber1	<i>W. stricta</i>	Australia, NSW, Abercrombie Caves. JMP 0915. WELT	No	No	Yes	Yes	No
STRI_NSW_Aber2	<i>W. stricta</i>	Australia, NSW, Abercrombie Caves. JMP 0916. WELT	Yes	Yes	Yes	Yes	No
STRI_NSW_Herb	<i>W. stricta</i>	Australia, NSW, Herberton. I. Telford 2562. NE	No	No	Yes	Yes	Yes
STRI_NSW_Saw	<i>W. stricta ?</i>	Australia, NSW, Kosciuszko NP, Sawpit Ck. JMP 0933b. WELT	No	No	Yes	Yes	No
SUBU_SA_WstC	<i>W. subulata</i>	South Africa, Western Cape, Somerset West. C. Cupido 207. NGB	Yes	Yes	No	No	No
TARA_NI_Jumb1	<i>W. pygmaea</i> subsp. <i>tatarua</i>	New Zealand, Tatarua ranges, Mt Jumbo. JMP 09104a. WELT	No	No	Yes	Yes	Yes



Tag name	Species	Collection details or literature citation	Chapter two		Chapter three		Chapter four	
			ITS	trnL-F	ITS	trnL-F	trnK	AFLP
TARA_NI_Jumb2	<i>W. pygmaea</i> subsp. <i>tararua</i>	New Zealand, Tararua ranges, Mt Jumbo. JMP 09104b. WELT	No	No	Yes	Yes	Yes	Yes
TARA_NI_Jumb3	<i>W. pygmaea</i> subsp. <i>tararua</i>	New Zealand, Tararua ranges, Mt Jumbo. JMP 09104c. WELT	Yes	Yes	Yes	Yes	Yes	No
TELF_NSW_Bera	<i>W. telfordii</i>	Australia, NSW, New England NP, Berarngutta rest area. G. Plunkett 24. NE	Yes	Yes	Yes	Yes	Yes	Yes
TELF_NSW_Dark	<i>W. telfordii</i>	Australia, NSW, New England NP, Darkie Pt. G. Plunkett 36. NE	No	No	Yes	Yes	Yes	No
TELF_NSW_Eagl	<i>W. telfordii</i>	Australia, NSW, New England NP, Eagles Nest lookout. G. Plunkett 30. NE	No	No	Yes	Yes	Yes	No
TENL_SA_WstC	<i>W. tenella</i>	South Africa, Western Cape, Cape Town. C. Cupido 194. NGB	Yes	Yes	No	No	No	No
TENR_SA_WstC	<i>W. tenerima</i>	South Africa, Western Cape, Prins Albert. C. Cupido 277. NGB	Yes	Yes	No	No	No	No
THUN_SA_WstC	<i>W. thunbergiana</i>	South Africa, Western Cape, Elandsbaai. C. Cupido 250. NGB	Yes	Yes	No	No	No	No
TUMI_VIC_Black	<i>W. tumidifructa</i>	Australia, Victoria, Blackburn Lake. W. May. WELT	No	No	No	Yes	Yes	No
UNDU_SA_EstC	<i>W. undulata</i>	South Africa, Eastern Cape, Hogsback. C. Cupido s.n. NGB	Yes	Yes	No	No	No	No
VERN_Chatham	<i>W. littoricola</i> subsp. <i>vernicaosa</i>	Chatham Islands. P. de Lange CH2031. AK	No	No	No	Yes	Yes	Yes
VERN_NI_Paka	<i>W. littoricola</i> subsp. <i>vernicaosa</i>	New Zealand, Hauraki Gulf, Pakahi Island. E. Cameron 15141. AK	Yes	Yes	Yes	Yes	Yes	No
VERN_NI_Piha	<i>W. littoricola</i> subsp. <i>vernicaosa</i>	New Zealand, Waitakere Ranges, Piha. P. de Lange 7617. AK	No	No	No	Yes	No	Yes
VERN_NI_Taur	<i>W. littoricola</i> subsp. <i>vernicaosa</i>	New Zealand, Bay of Plenty, Moturiki Island, Tauranga. P. de Lange 7589, AK	No	No	Yes	Yes	Yes	No
VERN_Raoul	<i>W. littoricola</i> subsp. <i>vernicaosa</i>	Raoul Island. P. de Lange ? AK	No	No	No	Yes	Yes	Yes
VIOL_Chatham	<i>W. violacea</i>	Chatham Islands. P. de Lange CH2003. AK	No	No	Yes	Yes	Yes	Yes
VIOL_NI_Mutu	<i>W. violacea</i>	New Zealand, Waikato, Muturangi Pa. P. de Lange 7677. AK	No	No	No	Yes	Yes	No
VIOL_NI_Paha	<i>W. violacea</i>	New Zealand, Wairarapa, Pahaoa Gorge. JMP 09103. WELT	No	No	No	Yes	Yes	Yes
VIOL_NI_Para	<i>W. violacea</i>	New Zealand, East Cape, Parauui Lookout. H. Meudt 282. WELT	No	No	Yes	Yes	Yes	No
IOI_NI_Ragl	<i>W. violacea</i>	New Zealand, Waikato, Te Toto Gorge, Raglan. P. de Lange 7682. AK	No	No	No	Yes	Yes	No

Tag name	Species	Collection details or literature citation	Chapter two		Chapter three		Chapter four	
			ITS	trnL-F	ITS	trnL-F	trnK	AFLP
VIOL_NI_Taur	<i>W. violacea</i>	New Zealand, Bay of Plenty, Moturiki Island, Tauranga. P. de lange 7588, AK	No	No	No	Yes	Yes	No
VIOL_SI_Dun	<i>W. violacea</i>	New Zealand, Nelson area, Dun Mts. JMP 0972. WELT	Yes	Yes	Yes	Yes	Yes	No
VIRG_SA_Kwa	<i>W. virgata</i>	South Africa, KwaZulu-Natal, Himeville. C. Cupido 299. NGB	Yes	Yes	No	No	No	No

## Appendix 2.1 Neighbour net analysis of the ITS dataset from Chapter Three created using SplitsTree v 4.10



$10^{-0.0010}$ 

**Appendix 2.3** Neighbour net analysis of the combined ITS + *trnL-F* + *trnK* dataset from chapter three created using SplitsTree v 4.10

0.0010

