

Phylogeny and phylogeography of *Zieria* (Rutaceae)

Rosemary Ann Barrett

A thesis submitted in total fulfilment of the requirements of the degree of Doctor of
Philosophy

March 2016

School of BioSciences

Faculty of Science
The University of Melbourne

Produced on archival quality paper



Zieria smithii (DJO430), Hat Head National Park, New South Wales: photograph courtesy of Daniel Ohlsen.

Abstract

Zieria (Anthophyta: Rutaceae) is a predominately Australian genus of shrubs and small trees, consisting of 59 Australian species and one species endemic to New Caledonia. Although there has been considerable taxonomic revision of this genus in recent decades, largely based on morphology, a comprehensive molecular phylogenetic study that would include as many currently-recognised taxa as possible was clearly needed. Using five chloroplast markers (*rpl32-trnL*, *trnL-F*, *trnQ-5'rps16*, *trnS-G* and *3'trnV-ndhC*) and two nuclear DNA markers (ITS and ETS) this study undertook phylogenetic analyses in order to understand relationships, evolutionary processes and taxonomic issues in *Zieria*, and to interpret biogeographic patterns. Both Bayesian inference and maximum parsimony methods were employed. Almost all species of *Zieria* were represented in separate analyses of the cpDNA and nrDNA datasets, with the monotypic sister genus *Neobyrnnesia* used as the outgroup. The results show that relationships within *Zieria* are complex; widespread incongruence was revealed between cpDNA, nrDNA and currently recognised taxa. A combination of factors is suggested to explain this, including regional cpDNA introgression (chloroplast capture) and incomplete lineage sorting, and the need for taxonomic revision of some taxa. The nrDNA provided greater support for monophyly of species than cpDNA, and potentially a better indication of phylogeny. However, deeply divergent paralogues of nrDNA were detected in some taxa, making the assessment of phylogenetic relationships more challenging, and highlighting some possible pitfalls in using nrDNA for phylogenetic reconstructions.

The phylogenetic relationship of Australian taxa and *Z. chevalieri*, the single endemic New Caledonian species, is of particular interest in terms of the long-standing debate over the history of New Caledonia and its flora. The placement of *Z. chevalieri* differed between cpDNA and nrDNA trees. The cpDNA phylogeny placed *Z. chevalieri* as sister to all other species in the genus, suggesting that the earliest divergence was between lineages of New Caledonia and Australia, and that differentiation of species occurred through vicariance. In contrast, *Z. chevalieri* was placed higher in the nrDNA tree, which could suggest later dispersal, over water or via past exposed land in the Tasman

Sea rather than older vicariance between Australia and New Caledonia. However, the results did not favour one hypothesis over another, and the debate will continue.

A phylogeographic study of one widespread species, *Zieria arborescens*, was also undertaken, focusing on populations in the wet forests of Victoria and Tasmania. The cpDNA was not highly informative but provided some evidence of Tasmanian–eastern Victorian connections, interpreted as past geographic connectivity during periods of low sea-level now severed by Bass Strait.

Finally, a number of taxonomic issues arising from the results of this project are discussed, and recommendations are made for reappraising and possibly re-circumscribing several taxa. In some cases this might require the description of new species based on variation between populations (e.g. *Z. smithii* and *Z. furfuracea*), or re-circumscription of taxa to effect monophyly (e.g. *Z. arborescens*). Some taxa might be considered conspecific based on noteworthy similarities of morphological and genetic characters (as seen in this study) and the geographic proximity of populations (e.g. *Z. buxijugum* and *Z. parrisiae*; *Z. alata* and *Z. madida*; *Z. verrucosa* and *Z. vagans*; *Z. caducibracteata* and *Z. arborescens*; and *Z. minutiflora* and *Z. obovata*).

Declaration and statement of authorship

This is to certify that:

1. this thesis comprises only my original work towards the Ph.D. except where indicated in the Preface.
2. due acknowledgement has been made in the text to all other material used.
3. this thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Rosemary Ann Barrett

Preface

Throughout this project various people assisted with provision of fresh samples and some herbarium samples – collection details and voucher information for all accessions are provided in Appendix A. All laboratory work was done by myself except for DNA extractions of the following 16 samples which was done by Gareth Holmes prior to the commencement of this Ph.D.: *Z. adenophora* MJB2010, *Z. arborescens* MJB1868, *Z. aspalathoides* subsp. *aspalathoides* (Bean 27984), *Z. baeuerlenii* MJB2011, *Z. covenyi* MJB2013, *Z. formosa* MJB2014, *Z. fraseri* subsp. *robusta* (MBT 3739), *Z. furfuracea* subsp. *gymnocarpa* (PIF 34232), *Z. granulata* MJB2016, *Z. littoralis* MJB2017, *Z. madida* AF5320, *Z. minutiflora* subsp. *minutiflora* (PIF 34313), *Z. oreocena* (MJB1982), *Z. veronicea* subsp. *veronicea* (MJB1980), *Z. robertsiorum* (AF5349) and *Z. smithii* PIF 34139. Also the extraction of DNA and all sequencing of *Z. veronicea* subsp. *insularis* (WN14AE) was done by Will Neal during his Honours project.

Chapter 3 has been published as follows:

Barrett RA, Bayly MJ, Duretto MF, Forster PI, Ladiges PY and Cantrill DJ. (2015) A chloroplast phylogeny of *Zieria* (Rutaceae) in Australia and New Caledonia shows widespread incongruence with species-level taxonomy. *Australian Systematic Botany* **27**, 427–449.

This chapter appears in this thesis as it was published. It includes references which, for completeness, are also included in Chapter 7 (references for the whole thesis), and accession details, which are also listed in Appendix A (accessions details for the whole thesis). Michael Bayly, Marco Duretto, Paul Forster, Pauline Ladiges and David Cantrill provided assistance with comments on the manuscript for the publication. All other work, including sampling, laboratory work, analyses, and writing, was done by myself, except where stated in the previous paragraph.

Acknowledgements

This study was funded by the Australian Research Council (LP0776737), a Bush Blitz Ph.D. Research Supplement (BBC210-24), grants from the Australian Biological Resources Study (PD208-02; RFL210-17), and was supported by the Alfred Nicholas Fellowship, The University of Melbourne School of Botany Foundation, the Australasian Systematic Botany Society (Hansjörg Eichler Scientific Research Fund Grant), The Soroptimist International of the South West Pacific (Dame Margaret Blackwood Soroptimist Scholarship) and an Australian Post Graduate Award.

I would like to thank the following people for their support in the course of my PhD:

My supervisors: Dr. Michael Bayly (for his support and advice despite being completely snowed under with work at times, and his painstaking attention to detail); Prof. Pauline Ladiges (for her wisdom and fast turnaround with editorial comments, her general advice on life and pearls, and her endless generosity); and Prof. David Cantrill (for editorial assistance and providing access to the facilities at the Royal Botanic Gardens, especially in the early years of my Ph.D.).

All the people who assisted with field work or provision of specimens, especially Erin Batty, Mike Bayly, Gill Brown, Marco Duretto, Paul Forster, David Meagher, and Daniel Ohlsen.

Margaret Brookes for help sorting out the voucher specimens.

Mal Padgett who provided lunches with lessons in Photoshop.

Prof. Ed Newbigin who was my first point of contact at the university and who has offered me work from time to time which has helped keep the wolf from the door.

Prof. Gareth Nelson who is so kind and supportive and wise.

My fellow students and staff in the Systematics and Cookson Labs for friendship support and encouragement: [Erin Batty, Gill Brown, Steph Conway, Cat Clowes, Alice Crowe, Trish Downing, Rachel Fowler, Stu Gardiner, Gareth Holmes, Emma Lewis, Claire Marks, Todd McLay, David Meagher, Karen Muscat, Daniel Ohlsen, Tanja Schuster].

I express special thanks and gratitude to my children, Jen, Cassie and Will, who, for seven years, have forgone much, learnt to survive on little, but nevertheless always encouraged me to keep going. We faced considerable challenges as a family during these years. And of course huge thanks to John and Jenny Barrett, my wonderful mum and dad, whose generosity, encouragement and assistance made completion of this thesis possible.

Table of contents

Abstract.....	iv
Declaration and statement of authorship	vi
Preface	vii
Acknowledgements	viii
Table of contents	x
List of Figures.....	xv
List of Tables	xix
Key to abbreviations used in this thesis.....	xx

Chapter 1. Introduction

1.1. Introduction to the Rutaceae.....	1.1
1.2. Genus <i>Zieria</i> Sm.	1.3
History	1.3
1.3. Description of <i>Zieria</i>	1.6
Morphology	1.6
Chemical characteristics	1.9
Cytology, biology, genetics and ecology.....	1.9
1.4. Distribution and abundance of <i>Zieria</i>	1.11
1.5. Conservation.....	1.15
1.6. Phylogenetic research in <i>Zieria</i>	1.16
1.7. Biogeography.....	1.18
Connections between Australia and New Caledonia.....	1.18
Connections in south-eastern Australia (<i>Zieria arborescens</i>)	1.19
1.8. Main objectives.....	1.20

Chapter 2. General methods

2.1. Introduction	2.1
2.2. Collection of plant material	2.1
2.3. DNA extraction.....	2.2

2.4. Molecular phylogenetic research.....	2.3
Chloroplast DNA (cpDNA).....	2.3
cpDNA markers used in this study	2.6
Nuclear ribosomal DNA (nrDNA)	2.8
nrDNA markers used in this study	2.11
2.5. DNA amplification	2.12
cpDNA amplification	2.13
nrDNA amplification.....	2.13
2.6. Purification and quantification of PCR products.....	2.14
2.7. DNA sequencing.....	2.14
2.8. Sequence editing and alignment	2.15
2.9. ITS and ETS paralogues	2.16
Detection of divergent ITS and ETS paralogues	2.16
Cloning of ITS and ETS paralogues.....	2.16
Comparison of ITS and ETS paralogues	2.18
2.10. Partition homogeneity test (ILD test) of nrDNA data	2.18
2.11. Recombination analysis of nrDNA data.....	2.19
2.12. Phylogenetic analyses.....	2.20
Bayesian Inference	2.20
Parsimony Analysis	2.22
Outgroups	2.24
Inclusions and exclusions	2.25

Chapter 3. (Published) A chloroplast phylogeny of *Zieria* (Rutaceae) in Australia and New Caledonia shows widespread incongruence with species-level taxonomy

Abstract.....	3.1
Introduction	3.1
Materials and Methods	3.3
Taxon sampling	3.3
DNA isolation, polymerase chain reaction (PCR) and sequencing.....	3.3
Sequence editing and alignment.....	3.4
Phylogenetic analyses.....	3.4

Results	3.4
Overview	3.4
Monophyly of species and sharing of chloroplast haplotypes between species.....	3.5
Composition and distributions of clades C–H.....	3.7
Composition and distributions of clades I–N	3.7
Discussion.....	3.12
Incongruence between cpDNA phylogeny and species taxonomy.....	3.12
Relationship of <i>Zieria</i> in Eastern Australia and New Caledonia.....	3.13
Conclusions and prospects for further research.....	3.14
Acknowledgements	3.14
References	3.14
Appendix 1. Details of samples and sequences used in this study	3.17
Supplementary Material	3.24

Chapter 4. Phylogeny of *Zieria* based on nuclear ribosomal DNA

4.1. Introduction	4.1
4.2. Materials and methods.....	4.2
Taxon sampling	4.2
Molecular markers used in this study	4.2
ITS and ETS paralogues	4.3
4.3. Results	4.4
Amplification and sequencing of ITS and ETS: assessment of variation within genomes.....	4.4
Tests for recombination and pseudogenes in ITS and ETS sequences.....	4.6
Analyses of separate ITS and ETS datasets.....	4.7
Analyses of combined ITS + ETS datasets: overview	4.12
Monophyly of species and divergence of paralogues in combined ITS + ETS analysis	4.16
Comparison with chloroplast phylogeny of <i>Zieria</i>	4.17
Comparison with the morphological phylogeny of Armstrong (2002)	4.18
4.4. Discussion.....	4.22

Significance of divergent nrDNA paralogues	4.22
Comparison with cpDNA variation	4.25
Biogeography: relationships of the New Caledonian species, <i>Zieria</i> <i>chevalieri</i>	4.27
4.5. Concluding remarks.....	4.28

Chapter 5. Phylogeography of *Zieria arborescens*

5.1. Introduction	5.1
5.2. Quaternary climatic fluctuations	5.1
5.3. <i>Zieria arborescens</i>	5.3
5.4. Aims.....	5.7
5.5. Materials and methods.....	5.7
5.6. Results of cpDNA analysis.....	5.10
Phylogeny	5.10
Geographic paralogy (major clades C and D)	5.14
Geographic patterns.....	5.15
Similarity or divergence of haplotypes.....	5.17
Nested taxa in <i>Zieria arborescens</i>	5.19
5.7. Results of nrDNA analysis	5.20
5.8. Discussion.....	5.26
Scope of the study: the value of including multiple species.....	5.26
Overview	5.26
Incongruence between nrDNA and cpDNA and species-level taxonomy: signals of incomplete lineage sorting and introgression.....	5.27
Haplotype diversity.....	5.29
Mainland and Tasmania connections	5.30
Grampians and Otways connection	5.31
Taxonomic issues	5.32
Conclusion.....	5.33

Chapter 6. General discussion

6.1. Phylogeny of <i>Zieria</i>	6.1
6.2. Utility of chloroplast DNA markers.	6.1
6.3. Utility of nuclear ribosomal DNA markers.	6.4
6.4. Implications for species-level taxonomy	6.5
<i>Zieria smithii</i> and close relatives: species in need of taxonomic revision.....	6.6
<i>Zieria furfuracea</i> : evidence of three disparate lineages	6.10
<i>Zieria cytisoides</i> and <i>Zieria hydroscopica</i>	6.11
Relationships of the endangered coastal specialist, <i>Zieria prostrata</i>	6.11
Classification of <i>Zieria montana</i>	6.12
<i>Zieria alata</i> and <i>Zieria madida</i>	6.14
<i>Zieria arborescens</i> and related taxa.....	6.15
Taxa in clade I with glandular-verrucose leaves	6.18
Relationships of <i>Zieria verrucosa</i> , <i>Zieria vagans</i> and <i>Zieria distans</i>	6.20
<i>Zieria minutiflora</i> and related taxa	6.21
6.5. Biogeography – eastern Australia and New Caledonia	6.22
6.3. Project assessment	6.23
6.4. Conclusion and prospects for further research	6.24

Chapter 7. List of references 7.1

Appendix A. Details of samples and sequences used in this studyA.1

Appendix B. Poylmorphisms in ITS and ETS sequences.....B.1

Appendix C. Concatenation of clones sequencesC.1

Appendix D. Congruence between proposed phylogenies of *Zieria*D.1

List of Figures

Chapter 1

Figure 1.1. The illustration of <i>Zieria smithii</i> originally published in Jackson (1810).	1.4
Figure 1.2. Illustrations from Clarke and Lee (1987): (a) a typical <i>Zieria</i> flower; (b) a longitudinal section of a <i>Zieria</i> flower showing disc lobes and position of the ovaries; (c) a transverse section of the disc and ovaries showing the placentation arrangement.....	1.8
Figure 1.3. Distribution of <i>Zieria</i> and <i>Neobyrsesia</i> (sister group) in Australia and New Caledonia, adapted from Armstrong (2002).	1.11

Chapter 2

Figure 2.1. A circular gene map of the chloroplast genome of <i>Citrus sinensis</i> from Bausher <i>et al.</i> (2006), indicating the location of the chloroplast markers used in this study.....	2.4
Figure 2.2. A simplified diagram of the repeat unit structure of 18S–5.8S–28S nuclear ribosomal DNA (adapted from Shiao <i>et al.</i> 2009; not to scale).	2.9

Chapter 3 (published)

Fig. 1. Map showing the distribution of <i>Zieria</i> and <i>Neobyrsesia</i> (outgroup) in Australia and New Caledonia, adapted from Armstrong (2002). Inset shows the flower of <i>Z. chevalieri</i> , Mount Kaala, New Caledonia.....	3.2
Fig. 2. Bayesian majority-rule consensus tree of combined chloroplast data, with Bayesian posterior probabilities (PP) shown above branches, bootstrap support (BS) values for maximum parsimony analysis below branches, and nodes numbered (in grey).	3.6
Fig. 3. A section of the consensus tree from the combined chloroplast analysis, showing Clades C–H, mean branch lengths and support values.	3.8
Fig. 4. Maps showing the distribution of samples in Clades C–H.	3.9
Fig. 5. A section of the consensus tree from the combined chloroplast analysis, showing Clades I–N, mean branch lengths and support values.	3.10

Fig. 6. Maps showing the distribution of samples in Clades I–N.....	3.11
Fig. S1. Phylogeny of <i>Zieria</i> based on analysis of morphological and chemical characters by Armstrong (2002).	3.25
Fig. S2. Bayesian majority-rule consensus tree produced by analysis of combined cpDNA markers (as in Fig. 2), showing taxa colour coded according to the groups defined in the phylogeny of Armstrong (2002).	3.26

Chapter 4

Figure 4.1. Section of ITS1 showing polymorphisms typical in chromatograms of ITS sequences of <i>Zieria prostrata</i> MJB2019 when amplified using universal primers ITS5 and ITS4.	4.4
Figure 4.2. Sections of aligned ITS1 chromatograms from <i>Zieria prostrata</i> 758677, amplified by various primers (reverse primer ITS4, and forward primers ITS5, ProsMajorF and ProsMinorF).	4.5
Figure 4.3. Bayesian inference majority-rule consensus tree of ITS data.....	4.9
Figure 4.4. Bayesian inference majority-rule consensus tree of ETS data.....	4.10
Figure 4.5. Bayesian inference majority-rule consensus tree produced from combined ITS + ETS data, showing all accessions, nodes (numbered in italics), and support values.	4.13
Figure 4.6. A section of the majority-rule consensus tree from the Bayesian inference analysis of the combined ITS + ETS data.	4.14
Figure 4.7. A section of the majority-rule consensus tree (Clade A) from the Bayesian inference analysis of the combined ITS + ETS data.....	4.15
Figure 4.8. Bayesian inference majority-rule consensus tree produced from combined ITS + ETS data, with samples colour-coded according to major cpDNA clades.....	4.19
Figure 4.9. Phylogeny proposed by Armstrong (2002), based on morphological and chemical characters.....	4.20
Figure 4.10. Bayesian inference majority-rule consensus tree produced by analysis of combined ITS + ETS data (as in Figure 4.5) reproduced here with taxa colour-coded according to the groups defined in the phylogeny of Armstrong (2002), as shown in Figure 4.9.	4.21

Chapter 5

Figure 5.1 Distribution of <i>Zieria arborescens</i> in south-eastern Australia.....	5.3
Figure 5.2. <i>Zieria arborescens</i> (RAB524) in the Dandenong Ranges, east of Melbourne, Victoria.	5.4
Figure 5.3. Herbarium specimens of <i>Zieria arborescens</i> (all Victorian) showing examples of leaf variation.	5.6
Figure 5.4. Excerpt from the phylogenetic tree based on nrDNA (Chapter 4) showing relationships of <i>Zieria arborescens</i> and other closely related species.....	5.8
Figure 5.5. Distribution of the taxa in this study that are closely related to <i>Zieria arborescens</i>	5.9
Figure 5.6. Bayesian inference majority-rule consensus tree of combined cpDNA data.	5.12
Figure 5.7. Section of Bayesian inference consensus tree of combined cpDNA data showing mean branch lengths.....	5.13
Figure 5.8. Simplified Bayesian inference majority-rule consensus tree of combined cpDNA data, showing mean branch lengths of branches outside clade A.....	5.14
Figure 5.9. Distribution of cpDNA haplotypes in the two main clades, C and D, indicating geographic paralogy.	5.15
Figure 5.10. Location of the four cpDNA haplotypes in subclade L (different ones colour-coded), showing a connection between Tasmanian and the mainland in <i>Zieria arborescens</i>	5.17
Figure 5.11. Location of the five cpDNA haplotypes (colour-coded) in subclade M, showing a connection between Tasmanian and the mainland in <i>Zieria arborescens</i>	5.17
Figure 5.12. Location of the nine cpDNA haplotypes in subclade O (colour- coded). Four haplotypes in subclade P, nested within O, are shown on the map on the right, with the remaining five in O shown on the map on the left....	5.18
Figure 5.13. Bayesian inference majority-rule consensus tree of combined nrDNA data.	5.22

Figure 5.14. Bayesian inference majority-rule consensus tree of combined nrDNA data, showing mean branch lengths.....	5.23
Figure 5.15. Detailed section of the Bayesian inference tree (Clade A) based on nrDNA, showing mean branch lengths and accessions colour coded by provenance.....	5.24
Figure 5.16. Geographic location of samples with identical ITS + ETS sequences, which were part of the polytomy at the basal node of clade A.	5.25

Chapter 6

Figure 6.1. Distribution of <i>Zieria smithii</i> and related taxa (grey shapes). Coloured shapes indicate the location of samples used in this study, resolved in (colour-coded) clades by nrDNA data.....	6.8
Figure 6.2. Distribution of taxa that were previously included in a broader circumscription of <i>Zieria montana</i> (i.e., taxa with asterisks in the key) by Armstrong (2002), and geographic location of samples resolved in the nrDNA clade B (magenta shapes).	6.13
Figure 6.3. Distribution of <i>Zieria alata</i> and <i>Z. madida</i> , based on records from Australia's Virtual Herbarium.	6.14
Figure 6.4. Distribution of <i>Zieria arborescens</i> and closely related taxa (light grey shapes), and the geographic location of samples resolved in clade D (dark grey shapes) of the nrDNA tree in Chapter 4, Figure 4.5.....	6.16
Figure 6.5. Distribution of samples in clade I (<i>Zieria buxijugum</i> , <i>Z. formosa</i> , <i>Z. granulata</i> , <i>Z. parrisiae</i> and <i>Z. tuberculata</i>). This map also indicates the (restricted) distribution of each species.	6.19
Figure 6.6. Distribution of <i>Zieria verrucosa</i> , <i>Z. vagans</i> and <i>Z. distans</i> . Symbols representing the samples used in this study are shaded dark grey.....	6.21

List of tables

Chapter 1

Table 1.1. Recently described, newly recognised or reinstated taxa and their morphological affinities with previously described taxa (Duretto and Forster 2007).....	1.6
--	-----

Table 1.2. List of <i>Zieria</i> taxa indicating occurrence and distribution (common or rare).....	1.13
--	------

Chapter 2

Table 2.1. cpDNA primers used in this study.....	2.8
--	-----

Table 2.2. nrDNA primers used in this study.....	2.12
--	------

Table 2.3. Primers used for amplification of cloned PCR products.....	2.17
---	------

Chapter 3

Table 1. Details of chloroplast datasets and results from maximum parsimony analyses.....	3.5
---	-----

Appendix 1. Details of samples and sequences used in this study (including GenBank numbers for cpDNA sequences).....	3.17
--	------

Chapter 4

Table 4.1. Summary of bacterial cloning and sequencing of nrDNA PCR products....	4.6
--	-----

Table 4.2. Details of (separate) maximum parsimony analyses of ITS and ETS datasets.	4.8
---	-----

Table 4.3. ITS and ETS copies (identified by copy number) paired for the combined analysis.	4.11
--	------

Table 4.4. Details of heuristic search results in the maximum parsimony analysis of combined ITS + ETS data.	4.12
---	------

Chapter 5

Table 5.1. Details of heuristic search results for the maximum parsimony analysis of cpDNA data.....	5.11
--	------

Table 5.2. Details of heuristic search results for the maximum parsimony analysis of nrDNA data.	5.20
---	------

Key to abbreviations used in this thesis

°C	degrees Celsius
µL	microlitre
µM	micromolar
A	adenine
ABG	Australian Botanic Garden
AGRF	Australian Genome Research Facility
ANBG	Australian National Botanic Gardens
BRI	The Queensland Herbarium
BSA	Bovine Serum Albumin
C	cytosine
CANB	The Australian National Herbarium
CI	consistency index
Ck	Creek
CNS	The Australian Tropical Herbarium
cpDNA	chloroplast DNA
cult.	cultivated specimen
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleoside triphosphates
E	east
EDTA	Ethylenediaminetetraacetic acid
G	guanine
Hwy	Highway
IR	inverted repeat
km	kilometres
LS	longitudinal section
LSC	large single copy (region of cpDNA)
m	metres
M	Molar
MEL	The National Herbarium of Victoria
MELU	The University of Melbourne Herbarium
MgCl ₂	magnesium chloride
Mt	Mount
N	north
ng	nanograms
NP	National Park
nrDNA	nuclear ribosomal DNA
NSW	New South Wales
NSW	The National Herbarium of New South Wales
NT	Northern Territory

p.	page
pmol	picomole
pp.	pages
Qld	Queensland
RBG	Royal Botanic Gardens Victoria (Melbourne)
Rd	Road
rDNA	ribosomal DNA
RI	retention index
S	south
SA	South Australia
SF	State Forest
SSC	small single copy (region of cpDNA)
subsp.	subspecies
T	thymine
TS	transverse section
Tas	Tasmania
TBE	Tris/Borate/EDTA
v	version
Vic	Victoria
W	west

Chapter 1. Introduction

1.1. Introduction to the Rutaceae

The family Rutaceae falls within the Eudicotyledonae (eudicots), a major monophyletic group of the core angiosperms, consisting of just over 300 families, around 165,000 species (Judd and Olmstead 2004), and accounting for approximately 75% of extant angiosperm diversity (Drinnan *et al.* 1994). Key features of eudicots are tricolpate pollen (pollen with three apertures), a character enabling improved contact between pollen grain and stigma, and flower parts commonly in fours or fives or multiples of these numbers.

Rutaceae is placed in the order Sapindales, and together with orders Brassicales, Huerteales and Malvales form a clade known as Malvids (or Euosids II), a major subclade of the large Rosid clade (Stevens 2012). There are nine families within the Sapindales, the major ones being Anacardiaceae, Burseraceae, Meliaceae, Rutaceae, Sapindaceae and Simaroubaceae (Stevens 2012). Rutaceae is a very large family with estimates of 150–162 genera and 1500–2100 species worldwide (Kubitzki *et al.* 2011; Groppo *et al.* 2012; Hartley 2013). Evidence from *rbcL* and *atpB* DNA sequences indicates that, as presently circumscribed, Rutaceae is paraphyletic (Chase *et al.* 1999; Groppo *et al.* 2008) and re-circumscription has been recommended; monophyly is achieved by the inclusion of *Harrisonia*, *Cneorum* and *Ptaeroxylon* in the family (Chase *et al.* 1999). Rutaceae are distinguished from their morphologically similar allies (Simaroubaceae, Cneoraceae, most Meliaceae and Zygophyllaceae) by the presence of secretory cavities in the leaves, stems and fruit (Hartley 2013). These are often referred to as pellucid dots, and contain aromatic oils, which frequently give Rutaceae species distinctive odours when crushed (Duretto 1999; Harden 2002). Mostly trees or shrubs, or more rarely herbs, Rutaceae occur abundantly throughout tropical and temperate regions of the southern hemisphere and are particularly diverse in sclerophyllous vegetation in southern Africa and Australia (Duretto 1999; Harden 2002; Bayly *et al.* 2013) and in tropical America (Groppo *et al.* 2012).

Engler (1931) divided the Rutaceae into seven subfamilies, which he further divided into tribes and subtribes. Numerous phylogenetic studies have highlighted problems

with Engler's classification (see Chase *et al.* 1999; Groppo *et al.* 2008; Bayly *et al.* 2013; Appelhans *et al.* 2014; and references therein). There are currently several relatively recent subfamilial classifications of Rutaceae, but none fully reconcile morphological, phytochemical and molecular data, and monophyly of the various subfamilies that have been proposed remains uncertain. Morton and Telmer (2014) proposed four subfamilies, Aurantioideae, Cneoroideae, Rutoideae and Amyridoideae. Groppo *et al.* (2012) proposed just two, Cneoroideae and Rutoideae (by merging the subfamilies Aurantioideae, Flindersioideae, Rutoideae and Toddalioidae). Several authors (e.g. Thorne and Reveal 2007; Kubitzki *et al.* 2011; Bayly *et al.* 2013) recognise three subfamilies, Aurantioideae, Cneoroideae and Rutoideae, although Rutoideae is demonstrably paraphyletic using this classification. In most circumscriptions Rutoideae is the largest subfamily, and in Australasia it consists of around 46 genera and about 660 species (Bayly *et al.* 2013).

The name Rutaceae is derived from *Ruta graveolens*, the common rue or herb of grace, an odorous herb occurring in southern Europe and northern Africa, grown since antiquity for its medicinal properties, and sometimes used in cooking (Duretto 1999). *Ruta* is the type genus for the family. The most economically important and well-known Rutaceae are *Citrus* species that produce edible fruits. Two historically notable Australian genera, *Eremocitrus* (Desert Lime) and *Microcitrus* (Finger Lime)—now included in *Citrus* (Mabberley 1998)—were used by early settlers to make jam and drinks. *Correa alba* was also used by sealers on the islands of Bass Strait as a substitute for tea (Duretto 1999). Other species currently used domestically are: *Murraya koenigii*, the curry leaf tree; *Angostura trifoliata*, an ingredient of some commercial bitters; and Sichuan pepper, which is obtained from *Zanthoxylum* species. *Pilocarpus* is the source of pilocarpine, a drug used to treat glaucoma (Chase *et al.* 1999). Some of the strongly scented species produce essential oils for use in perfumery, e.g. *Boronia megastigma* from Western Australia (Duretto 1999). Various species are grown commercially as garden ornamentals, including species of *Boronia*, *Diosma*, *Correa*, *Choisya*, *Eriostemon*, *Murraya*, *Phebalium* and *Philotheca*, and there are others grown regionally for their timber, e.g. species of *Amyris*, *Chloroxylon*, *Flindersia*, *Halfordia* and *Zanthoxylum* (Harden 2000; Armstrong and Harden 2002; Bayly *et al.* 2013).

The Flora of Australia reports that there are 43 genera of Rutaceae in Australia, occurring in all States, and 24 are endemic (Hartley 2013). Some of the more well-known Australian genera are (with species numbers given here from the Flora of Australia, Volume 26): *Asterolasia* (16 species), *Boronia* (148 species), *Correa* (11 species), *Crowea* (three species), *Eriostemon* (two species), *Leionema* (24 species), *Microcybe* (four species), *Phebalium* (28 species), *Philotheca* (53 species), and *Zieria* (60 species).

1.2. Genus *Zieria* Sm.

Zieria is a predominantly Australian genus of shrubs and small trees, with 59 species occurring in the eastern states of Australia, and one species endemic to New Caledonia; a total of 60 species, and 18 subspecies, are currently recognised (Armstrong 2002; George *et al.* 2013). More detail about the distribution and abundance of *Zieria* is given in section 1.4. *Zieria* has been resolved as monophyletic in morphological (Armstrong 1991) and molecular (Bayly *et al.* 2013; Morton 2015) phylogenetic analyses, and, based on *rbcL* and *atpB* DNA sequences, it is most closely related to the monotypic genus, *Neobyrnesia*, which is endemic to the East Alligator River area in Arnhem land, northern Australia (Armstrong and Powell 1980; Bayly *et al.* 2013). The authority names of taxa are given in Appendix A, which lists the samples used in this study, including specimen voucher details and localities.

History

Europeans first collected specimens of *Zieria* in the late 18th century. These specimens, probably *Z. pilosa*, were collected by Banks and Solander at Kurnell Peninsula, Botany Bay, near Sydney, during Cook's voyage to Eastern Australia in 1770 (Armstrong 2002), but descriptions of the specimens were not published at that time. In 1798 James Smith named the genus¹ after his late friend, Polish botanist, Johannis [John] Zier², and described the essential characters of the genus as follows:

¹ The name *Zieria* Sm. (1798) predates *Zieria* Schimp (1856) (Bryophyta: Bryaceae), now *Plagiobryum* Lindb. Schimper, W.P. (1856), *Corollarium Bryologiae Europaeae*: 68 [tax. nov.] APNI 277456.

² 'In memoriam pie defuncti *Johannis Zier*, Soc. Linn., quondam Sodalis, botanici indefessi, nobis non obliviscendi, quanvis alio sub nomine labores qjus saepitis inclaruerint.' [In affectionate memory of the late *John Zier*, former member of the Linnean Society, a tireless botanist, not forgotten by us, although his work has often become famous under the name of another.] (Smith 1798).

‘Calyx divided into four parts. Petals four. Stamens glabrous, with glands attached. Style simple. Stigma 4-lobed. Capsules 4, joined. Seeds arillate. Shrubs with opposite leaves, 3-lobed, flowers white. Genus diagnosed by single stamens with large glands attached.’



Figure 1.1. The illustration of *Zieria smithii* originally published in Jackson (1810).

A species circumscription was not written until some years later, when Jackson (1810) described and named *Zieria smithii*, the designated lectotype for the genus (Armstrong 1991), using samples sent to him from New Holland (Australia). The description appeared with an accompanying illustration (reproduced in Figure 1.1):

...a dwarf, warty shrub with opposite branches, leaves which are composed of three lanceolate leaflets, rough with transparent dots, which probably secrete an oil, as the bruised leaves are very fragrant. A few small scattered hairs are just visible on their upper surface. The panicles of flowers arise from the axils of the leaves, which they often excel in length, and branch in the same opposite manner as the stem, with linear bracts at the divisions. The cups are very short, finely haired and dotted as the leaves. The petals are broad-lanced, pointed, and slightly tinged with purple on the outside.

Descriptions of four species of *Zieria* were published by Bonpland (1815), then five by Smith (1818), although not all were new species. Von Mueller synonymised *Zieria* as a section of *Boronia* in the 1850s, but Bentham (1863) determined it to be a distinct genus in his 'Flora Australiensis'. In this treatment, 11 new taxa were circumscribed, and a comprehensive key was provided including distribution information, descriptions and synonyms; there were ten species recognised at that time. White (1942) described five additional species, and the endemic New Caledonian species, *Zieria chevalieri*, was later described by Viot (1953). After that there was little taxonomic work done on the genus until Armstrong (1991) prepared a monograph of the genus as part of a Ph.D. thesis on pollination and systematics in the Australian Rutaceae. His taxonomic revision of *Zieria*, which was published separately (Armstrong 2002), recognised 42 species and 16 subspecies.

There was little revision of the genus until Duretto and Forster (2007) published a taxonomic revision of the Queensland *Zieria*, in which 16 new species were described and illustrated, one variety was promoted to species rank, one species previously treated as a subspecies was reinstated, and one subspecies was newly recognised. Many of these newly recognised taxa share strong morphological affinities with previously recognised species (Table 1.1). Duretto and Forster (2007) also flagged several taxa that showed morphological variation between populations as requiring more research to investigate whether further taxonomic recognition is warranted (i.e. *Z. arborescens*,

Z. aspalathoides, *Z. cytisoides*, *Z. eungellaensis*, *Z. furfuracea*, *Z. inexpectata*, *Z. laevigata*, *Z. robertsiorum* and *Z. smithii*). A species from New South Wales, *Z. odorifera*, was also revised by Duretto and Forster (2008) and four subspecies are now recognised, possibly representing speciation resulting from habitat isolation and genetic drift. A brief taxonomic history of *Zieria* taxa is provided in Armstrong (2002), covering the period 1798 to 2001, but not including the more recently described taxa (Duretto and Forster 2007, 2008).

Table 1.1 Recently described, newly recognised or reinstated taxa, and their morphological affinities with previously described taxa (Duretto and Forster 2007). Naming authorities are given here only for taxon names not provided in Appendix A (i.e. not sampled for this study).

	New or newly recognized	Shows affinities with...
1.	<i>Z. actites</i>	<i>Z. compacta</i>
2.	<i>Z. cephalophila</i>	<i>Z. compacta</i>
3.	<i>Z. exsul</i>	<i>Z. compacta</i>
4.	<i>Z. compacta</i>	<i>Z. fraseri</i> subsp. <i>compacta</i> (C.T.White) J.A.Armstr.
5.	<i>Z. fraseri</i> subsp. <i>robusta</i>	<i>Z. fraseri</i> subsp. <i>fraseri</i> , <i>Z. compacta</i> var. <i>robusta</i> C.T.White, <i>Z. compacta</i> var. <i>glabrata</i> C.T.White
6.	<i>Z. alata</i>	<i>Z. montana</i>
7.	<i>Z. boolbunda</i>	<i>Z. montana</i>
8.	<i>Z. eungellaensis</i>	<i>Z. montana</i>
9.	<i>Z. madida</i>	<i>Z. montana</i>
10.	<i>Z. scopulus</i>	<i>Z. montana</i>
11.	<i>Z. bifida</i>	<i>Z. obovata</i> , <i>Z. exsul</i> , <i>Z. minutiflora</i>
12.	<i>Z. distans</i>	<i>Z. verrucosa</i>
13.	<i>Z. vagans</i>	<i>Z. verrucosa</i>
14.	<i>Z. graniticola</i>	<i>Z. aspalathoides</i>
15.	<i>Z. inexpectata</i>	<i>Z. aspalathoides</i>
16.	<i>Z. whitei</i>	<i>Z. aspalathoides</i>
17.	<i>Z. hydroscopica</i>	<i>Z. smithii</i>
18.	<i>Z. insularis</i>	<i>Z. smithii</i>
19.	<i>Z. tenuis</i>	<i>Z. collina</i> , <i>Z. cytisoides</i>

1.3. Description of *Zieria*

Morphology

Zieria can be distinguished from other Rutaceae genera by the following combination of characters: opposite leaves, conspicuous and four-merous flowers, free petals, four stamens, free filaments, deeply four-lobed disc and dry dehiscent fruits (Smith 1798; Duretto 1999; Armstrong 2002; George *et al.* 2013). In particular the autapomorphy ‘antesepalous four-lobed disc’ distinguishes *Zieria* from other similar genera, such as *Boronia*, and defines it morphologically as a monophyletic group (Armstrong 1991; Bayly *et al.* 2013). Although disc lobes are present in *Boronia*, they are not swollen or

prominent, whilst in *Zieria* they are swollen and project prominently upwards above the base of the ovary. There are eight stamens present in *Boronia* compared to four in *Zieria*. The main differences between *Zieria* and its sister genus, *Neobyrnesia*, are: the petals and stamens are persistent in the fruit in *Neobyrnesia* but not in *Zieria*; the leaves are simple in *Neobyrnesia*, but usually trifoliate in *Zieria*; and *Neobyrnesia* lacks the disc with four distinct antesealous lobes that is typical of *Zieria* (Armstrong 1980).

Most *Zieria* species are medium to tall, multi-branched, bushy shrubs to small trees. The few taxa that attain tree proportions are: *Z. arborescens*, *Z. caducibracteata*, *Z. lasiocaulis*, *Z. montana* and *Z. southwellii*, with *Z. arborescens* up to 10 metres in height (Armstrong 2002). *Zieria littoralis*, *Z. minutiflora* and *Z. veronicea* are low spreading shrubs up to 1 m tall. *Zieria citriodora*, *Z. robertsiorum* and *Z. prostrata* are procumbent shrubs, up to about 15 cm high.

Zieria species may be either glabrous or possess trichomes, which may be simple, bifid, or stellate. Only non-glandular trichomes have been observed in *Zieria*, located on branchlets, foliage, inflorescences and fruits. The type, length, density, and location of hairs are frequently diagnostic characters for different species and subspecies (Armstrong 2002).

Zieria species are unarmed, and have opposite, usually trifoliate leaves, only rarely appearing whorled. Two species, *Z. veronicea* and *Z. chevalieri*, have unifoliate leaves, and two species, *Z. involucrata* and *Z. murphyi*, commonly have heterophyllous leaves, with both trifoliate and unifoliate leaves occurring on the same plant. Margins of leaflets are entire or, in some species, tubercles on leaf margins give a dentate appearance. Terminal and lateral leaflets are similar in appearance except that the terminal ones are usually longer. The size and shape of leaflet blades and the prominence of venation are characters that help distinguish taxa.

Zieria inflorescences are axillary and cymose (i.e. the initial stem terminates in a flower that opens before those below it or to its side). The flowers are bisexual, conspicuous and have all floral parts in fours. Variation in flower size sometimes helps to distinguish taxa (e.g. *Z. minutiflora*). The floral disc has four distinct lobes alternating with the petals. The calyx is basally fused, the petals are free, valvate or imbricate and not persistent in fruit. The four carpels contain two ovules each and are fused basally. The

styles are fused and typically inserted in a small depression between the lobes of the ovary, and the stigma has a knob-like head (capitate). The stamens are erect, free, incurved and located opposite the sepals. Characters of the stamens such as filament morphology and the degree of pubescence may be diagnostic in some species. Apiculate anthers, i.e. anthers with a small terminal appendage, may also distinguish taxa. Details of floral structures are shown in Figure 1.2.

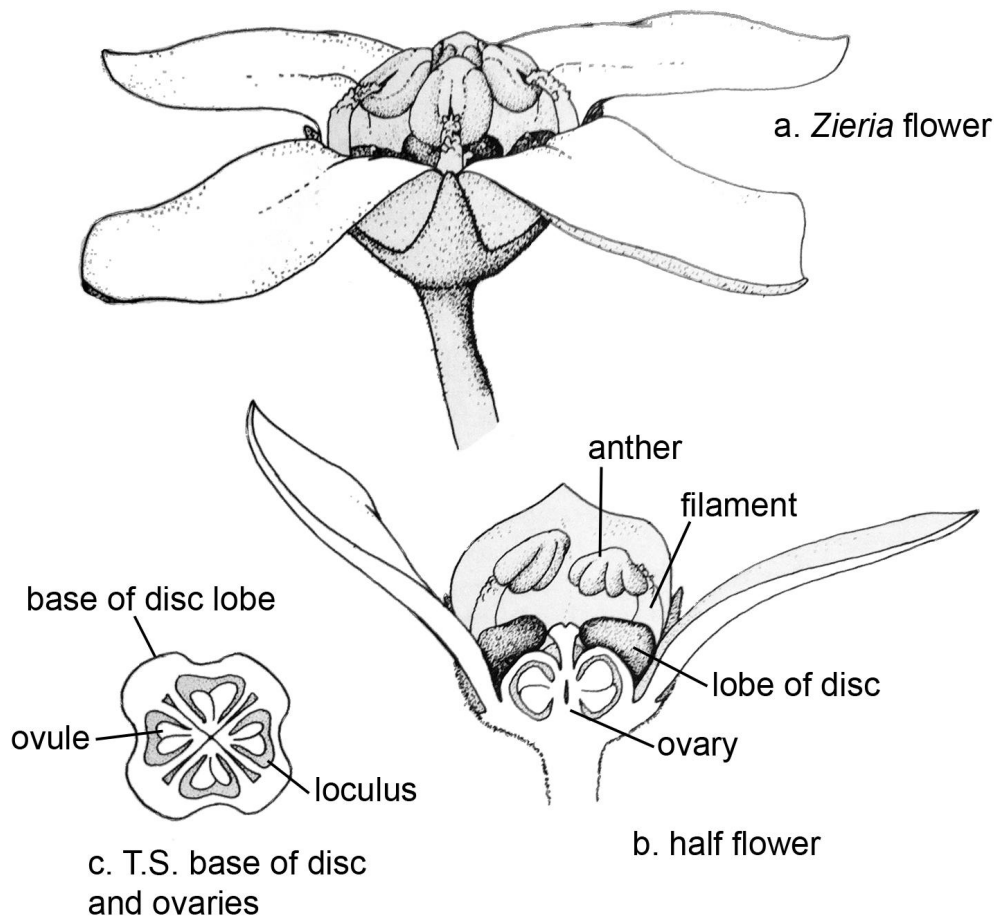


Figure 1.2. Illustrations from Clarke and Lee (1987): (a) a typical *Zieria* flower; (b) a longitudinal section of a *Zieria* flower showing disc lobes and position of the ovaries; (c) a transverse section of the disc and ovaries showing the placentation arrangement.

The dry dehiscent fruits of *Zieria* consist of one to four cocci, fused basally. Seeds are ejected forcibly from mature cocci enabling short-range initial dispersal. Seeds possess an ant-attracting elaiosome, aiding secondary dispersal. The seed surface structure of 23 species of *Zieria* was examined in detail by Powell and Armstrong (1980); most seeds are ridged, with little variation, and thus surface features are of limited use for taxonomy.

Chemical characteristics

Zieria species are commonly odoriferous, either pleasantly or unpleasantly, due to the presence of pellucid oil glands. Sometimes these glands are large and raised forming tubercles, a useful diagnostic character (Armstrong 2002). (Note: in descriptions of *Zieria* the terms tuberculate and verrucose are interchangeable, and indicate the presence of warty protuberances). The essential oils contained in the leaves have been used for chemotaxonomic investigations, and complex chemical variation within the genus has been observed (Flynn and Southwell 1987a; Southwell and Armstrong 1987). Subdivision of taxa into groups based on phytochemical characters (chemotypes) does not correlate well with groups based on morphological similarities. Nevertheless, chemical characters were used by Armstrong (2002) as potentially informative characters in a phylogenetic analysis. Essential oils and secondary metabolites of some *Zieria* species have been reported as having antimicrobial properties and these species may have potential as medicinal plants or in the development of new drugs (Islam and Ahsan 1997; Griffin *et al.* 1998). A chemical fruit fly attractant has been identified in *Z. smithii* (Fletcher *et al.* 1975), and may have potential for use in pest management. Several *Zieria* species that contain cyanogenic compounds are potentially toxic to animals that ingest them (Flynn and Southwell 1987b). *Zieria arborescens* is reported to have caused pulmonary oedema and death, following ingestion, in Tasmanian cattle (Munday 1968). Similar poisoning effects have been seen in rabbits in experimental trials (Munday *et al.* 1974). The particular toxin or toxins that cause the poisoning are not known, but are probably contained in the oils, which can be extracted from the leaves by steam-distillation (Munday *et al.* 1974).

Cytology, biology, genetics and ecology

Cytological data may be useful in indicating phylogenetic relationships and analysing phytogeographic patterns in the Australian Rutaceae (Smith-White 1954). Chromosome numbers for 19 *Zieria* species have been documented (Smith-White 1954; Stace and Armstrong 1992; Armstrong 2002) and show *Zieria* to be mostly diploid or tetraploid, with a base chromosome number $n = 18$ characteristic of the genus. *Zieria covenyi*, a putative hybrid, is known to be triploid, has sterile pollen and reproduces only

vegetatively. It is not known whether polyploidy differentiates subspecies (Stace and Armstrong 1992).

Detailed studies of the biology of *Zieria* are few. The biology of individual *Zieria* species is documented briefly in Armstrong (2002) including limited information on asexual and sexual reproduction, pollen and pollination, and seed biology in some cases. Most species have viable pollen and are insect pollinated, mainly by flies and beetles, but sometimes by bees and butterflies. Some species with very small inconspicuous flowers are self-compatible and produce viable seeds following self-pollination. Root suckering, stem layering and proliferation from rhizomes have also been observed in some taxa. *Zieria baeuerlenii* and *Z. covenyi* only reproduce vegetatively; these species have sterile pollen, do not produce seeds and are known only from extremely restricted populations (Armstrong 1991; Stace and Armstrong 1992; Sharma 2001). In one study of the rare species, *Z. prostrata*, Prakash (1994) determined it to have similar reproductive biology to that of other rutaceous plants.

Genetic studies of *Zieria* taxa are also very limited. Sharma (2001) found that the threatened species, *Z. baeuerlenii*, possesses moderate genetic variation (20 unique clones detected throughout the range of the species) despite having asexual reproduction and a distribution limited to an area of less than one square kilometre. Hogbin and Crisp (2003) studied relationships between populations of *Z. prostrata* and geographically close populations of *Z. smithii*. They showed the rare and endangered headland species, *Z. prostrata*, to be a distinct if somewhat incipient species, despite having genetic similarities with populations of *Z. smithii* occurring nearby. The headland populations of *Z. smithii* were shown to be morphologically distinct, perhaps worthy of conservation and subspecific status, but not genetically distinct enough to warrant specific recognition. The leaf proteins of *Z. covenyi* were shown to have no intra-population genetic variation (J.A. Armstrong and I.K. Sharma, unpublished data, cited by Armstrong 2002), which is to be expected in a species that only reproduces clonally. A study by Neal *et al.* (2013) analysed genetic variation in the species *Z. veronicea*, supporting taxonomic recognition of the rare Kangaroo Island endemic subspecies *insularis*, which was shown to be genetically very distinct. Populations of *Z. veronicea* in Tasmania, where the species is rare, were shown to be closely related to populations in Gippsland where conservation of the species is more secure.

The ecology of *Zieria* and the Australian Rutaceae in general has been under-researched and remains poorly understood. In particular, little is known about the seed biology and ecology of the family, and the impact of fire on species (Auld 2001). Some taxa possess lignotubers so are able to regenerate after fire, or predation, but generally this character has been poorly documented. *Zieria involucreata* is fire sensitive, but shows a weak resprouting response. This species probably forms persistent long-lived soil seedbanks, an important factor in maintaining populations in the fire-prone communities where it occurs (Auld *et al.* 2000).

1.4. Distribution and abundance of *Zieria*

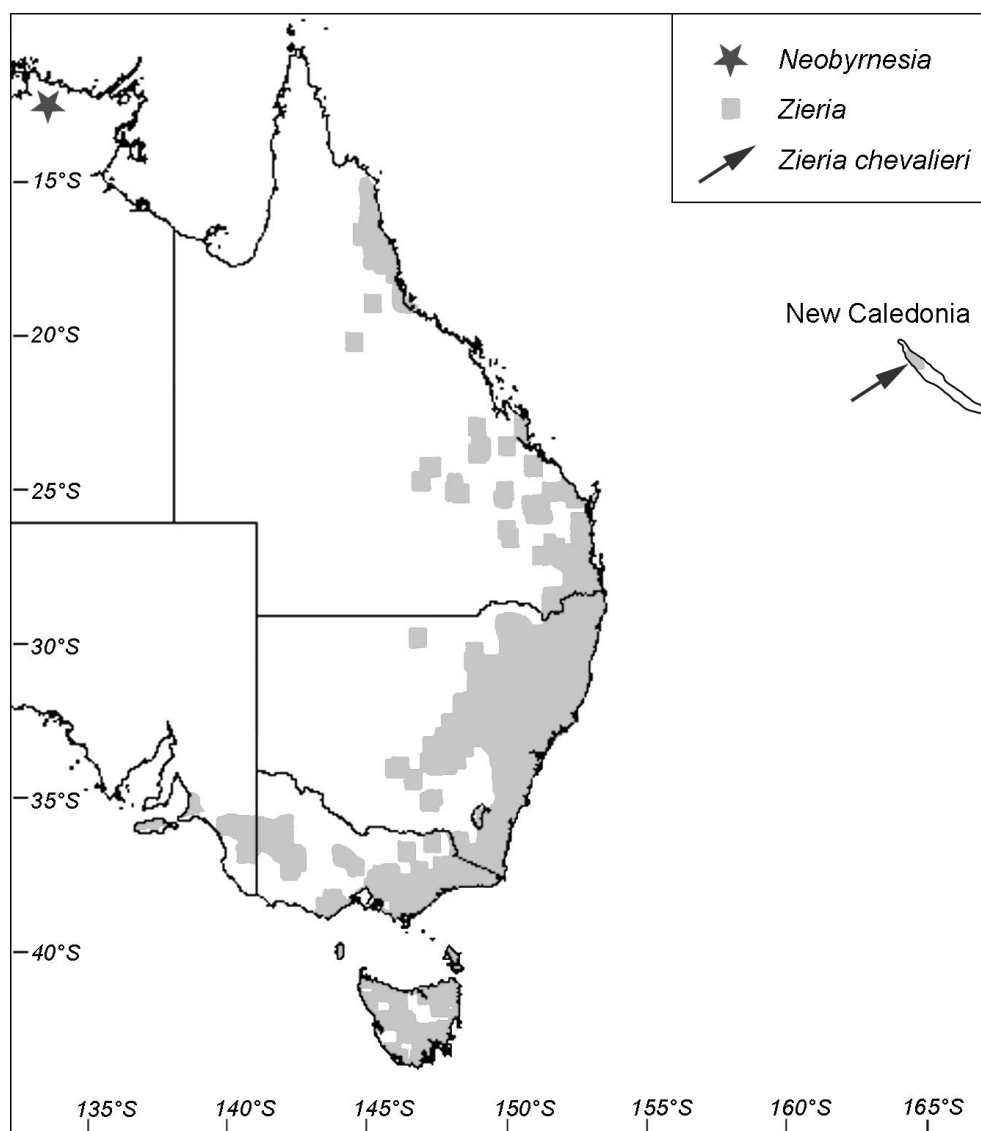


Figure 1.3. Distribution of *Zieria* and *Neobyrsesia* (sister group) in Australia and New Caledonia, adapted from Armstrong (2002).

Figure 1.3 shows the general distribution of *Zieria* although this does not reflect the varying distribution of widespread versus narrowly distributed taxa in Australia. *Zieria* occurs naturally only in the eastern states of Australia, extending as far west as Kangaroo Island in South Australia, north to the tropical regions of Queensland, and south to Tasmania (Armstrong 2002). Only one species occurs naturally outside Australia: *Z. chevalieri*, which is endemic to New Caledonia. The greatest diversity of species occurs in New South Wales (34 species) and Queensland (35 species), whilst in Victoria there are nine species, three in Tasmania and a single species in South Australia. More than half the *Zieria* taxa are restricted endemics (Table 1.2). Several are currently known only from single localities: *Z. actites*, *Z. adenophora*, *Z. arborescens* subsp. *decurrens*, *Z. baeuerlenii*, *Z. boolbunda*, *Z. buxijugum*, *Z. cephalophila*, *Z. covenyi*, *Z. floydii*, *Z. formosa*, *Z. hydroscopica* and *Z. parrisiae* (Armstrong 2002; Duretto and Forster 2007). *Zieria chevalieri* is known to occur in a single population on Mount Kaala, New Caledonia, thus is very rare and endangered (Armstrong 2002). See section 1.5 ‘Conservation’ for national listings of threatened species.

One of the most widespread species in the genus is *Z. smithii*, a species that occurs in eastern Victoria, New South Wales and Queensland, where it is common and shows considerable variation in morphology across its range (Duretto and Forster 2007). Two subspecies of *Z. smithii* (subspecies *smithii* and *tomentosa*) were described by Armstrong (2002) based on the density of the indumentum, but the distinction is not always clear and the subspecies are no longer recognised (Armstrong and Harden 2002; Duretto and Forster 2007; George *et al.* 2013). Other widespread taxa in the genus are: *Z. arborescens*, the most common species in Victoria and Tasmania but which also occurs in New South Wales and southern Queensland; *Z. aspalathoides* subsp. *aspalathoides*, widespread throughout eastern Australia, usually growing in hilly country in open eucalypt forests (subsp. *brachyphylla* is a rare taxon restricted to the Blackdown Tableland in south-eastern Queensland); *Z. compacta*, which commonly grows on rocky escarpments in a variety of habitats throughout the Great Dividing Range; *Z. cytisoides*, an understorey shrub in open eucalypt forests, which is particularly widespread in New South Wales; *Z. laevigata*, which occurs mainly in New South Wales, and is common in shrublands and open eucalypt forests on sandy soils;

and *Z. laxiflora* which occurs mainly in coastal heathlands in New South Wales and Queensland (Armstrong 2002).

Table 1.2. List of *Zieria* taxa indicating occurrence and distribution (common or rare). Asterisks denote taxa previously flagged as requiring further work to more accurately circumscribe, or to recognise new taxa (Duretto and Forster 2007).

	<i>Zieria</i> species	Subspecies	Occurrence					Distribution
			Vic	Tas	SA	NSW	Qld	
1	<i>actites</i>						✓	restricted
2	<i>adenodonta</i>					✓	✓	restricted
3	<i>adenophora</i>					✓		restricted
4	<i>alata</i>						✓	restricted
5a	<i>arborescens</i> *	<i>arborescens</i>	✓	✓		✓	✓	widespread
5b	<i>arborescens</i>	(<i>glabrifolia</i>)				✓	✓	restricted
5c	<i>arborescens</i>	(<i>decurrens</i>)				✓		restricted
6a	<i>aspalathoides</i>	<i>aspalathoides</i> *	✓			✓	✓	widespread, but rare in Vic
6b	<i>aspalathoides</i>	<i>brachyphylla</i>					✓	restricted
7	<i>baeuerlenii</i>					✓		restricted
8	<i>bifida</i>						✓	restricted
9	<i>boolbunda</i>						✓	restricted
10	<i>buxijugum</i>					✓		restricted
11	<i>caducibracteata</i>					✓		restricted
12	<i>cephalophila</i>						✓	restricted
13	<i>chevalieri</i>		(New Caledonia)					restricted
14	<i>citriflora</i>		✓			✓		rare across range
15	<i>collina</i>						✓	restricted
16	<i>compacta</i>					✓	✓	widespread
17	<i>covenyi</i>					✓		restricted
18	<i>cytisoides</i> *		✓			✓	✓	widespread, but uncommon in Vic; also in ACT
19	<i>distans</i>						✓	disjunct populations
20	<i>eungellaensis</i> *						✓	restricted
23	<i>exsul</i>						✓	restricted
22	<i>floydii</i>					✓		restricted
23	<i>formosa</i>					✓		restricted
24a	<i>fraseri</i>	<i>fraseri</i>				✓	✓	common in its range
24b	<i>fraseri</i>	<i>robusta</i>					✓	common in its range
25a	<i>furfuracea</i>	<i>furfuracea</i> *				✓		moderate
25b	<i>furfuracea</i>	<i>euthadenia</i> *					✓	common in its range
25c	<i>furfuracea</i>	<i>gymnocarpa</i> *					✓	restricted
26	<i>graniticola</i>						✓	restricted
27	<i>granulata</i>					✓		restricted
28	<i>hindii</i>					✓		restricted
29	<i>hydroscopica</i>						✓	restricted
30	<i>inexpectata</i> *						✓	restricted
31	<i>ingramii</i>					✓		restricted
32	<i>insularis</i>						✓	restricted

Chapter 1

	<i>Zieria</i> species	Subspecies	Occurrence					Distribution
			Vic	Tas	SA	NSW	Qld	
33	<i>involucrata</i>					✓		restricted
34	<i>laevigata</i> *					✓	✓	widespread
35	<i>lasiocaulis</i>					✓		restricted
36	<i>laxiflora</i>					✓	✓	widespread, mainly coastal
37	<i>littoralis</i>		✓	✓		✓		coastal areas
38	<i>madida</i>						✓	restricted
39a	<i>minutiflora</i>	<i>minutiflora</i>				✓	✓	widespread
39b	<i>minutiflora</i>	<i>trichocarpa</i>					✓	widespread
40	<i>montana</i>						✓	restricted
41	<i>murphyi</i>					✓		moderate
42	<i>obcordata</i>					✓		restricted
43	<i>obovata</i>						✓	restricted
44a	<i>odorifera</i>	<i>copelandii</i>				✓		restricted
44b	<i>odorifera</i>	<i>warrabahensis</i>				✓		restricted
44c	<i>odorifera</i>	<i>odorifera</i>				✓		
44d	<i>odorifera</i>	<i>williamsii</i>				✓		
45	<i>oreocena</i>		✓					restricted to the Grampians
46	<i>parrisiae</i>					✓		restricted
47	<i>pilosa</i>					✓		
48	<i>prostrata</i>					✓		restricted
49	<i>rimulosa</i>						✓	restricted
50	<i>robertsiorum</i> *						✓	restricted (Wet Tropics)
51	<i>robusta</i>		✓			✓		disjunct distribution
52	<i>scopulus</i>						✓	restricted
53	<i>smithii</i> *		✓			✓	✓	widespread
54	<i>southwellii</i>					✓	✓	common within its range
55	<i>tenuis</i>						✓	disjunct distribution
56	<i>tuberculata</i>					✓		restricted
57	<i>vagans</i>						✓	restricted
58a	<i>veronicea</i>	<i>veronicea</i>	✓	✓	✓			disjunct distribution
58b	<i>veronicea</i>	<i>insularis</i>			✓			restricted to Kangaroo Island
59	<i>verrucosa</i>						✓	restricted
60	<i>whitei</i>						✓	restricted
Nbr of species			9	3	1	34	35	
Total taxa			9	3	2	39	41	

Duretto and Forster (2007) recognised a speciation pattern for *Zieria* based on current distributions: a ‘core’ group of wide-ranging taxa occur in widespread, possibly more stable habitats, in conjunction with numerous narrowly distributed species (See Table 1.2). This speciation pattern is apparent more or less throughout the range of the genus. Widespread taxa such as *Z. smithii* and *Z. arborescens* are mesic forest species, requiring moderately moist habitats, whilst many of the numerous narrow endemics,

particularly those of Queensland, frequently occur in isolated pockets, in crevices or clefts on rocky substrates or outcrops (granite, rhyolite, sandstone or trachyte), in habitats surrounded by rainforest or areas more prone to fire. *Zieria* species also occur frequently as understorey shrubs in open forests and woodlands, often immediately adjacent to the isolated rocky habitats. This distribution pattern may indicate allopatric speciation, or that the isolated mountain habitats have provided refugia for species subjected to local climate change and contracting habitats. Increased speciation rates of extant taxa are known to occur on inselbergs, i.e. isolated mountains (Burke 2003). On the other hand small populations may be genetically depauperate and are particularly vulnerable to extinction through disturbance such as fire, prolonged drought, or weed infestation (Duretto and Forster 2007).

1.5. Conservation

The Australian Government's Department of the Environment website (accessed 9th December 2015) provides a summary of species that are listed as nationally threatened under the *Environment Protection and Biodiversity Conservation Act 1999* (Cwlth). Thirteen *Zieria* species are listed as endangered: *Z. adenophora*, *Z. baeuerlenii*, *Z. bifida*, *Z. buxijugum*, *Z. covenyi*, *Z. floydii*, *Z. formosa*, *Z. granulata*, *Z. ingramii*, *Z. lasiocaulis*, *Z. obcordata*, *Z. parrisiae* and *Z. prostrata*. Another eight species are listed as vulnerable: *Z. citriodora*, *Z. collina*, *Z. involucrata*, *Z. murphyi*, *Z. obovata*, *Z. rimulosa*, *Z. tuberculata* and *Z. verrucosa*.

Many of the recently described Queensland species would also probably qualify for listing under the Act if nominated, because of their restricted distribution. The Queensland Government's Nature Conservation (Wildlife) Regulation 2006 listed (as of 9th May 2014) the following *Zieria* taxa as endangered in the state: *Z. actites*, *Z. bifida* (included in the national list), *Z. exsul*, *Z. furfuracea* subsp. *gymnocarpa*, *Z. graniticola*, *Z. inexpectata*, and *Z. vagans*. This regulation also lists *Z. montana* as vulnerable and *Z. adenodonta* as near threatened, in addition to the species already listed under the national list. The New Caledonian species *Z. chevalieri* is listed as globally vulnerable on the IUCN Red List of Threatened Species (<http://www.iucnredlist.org>, accessed November 2014) under Criteria B1 and 2C, based on the assessment of Jaffre *et al.* (1998).

It is important to determine and evaluate what practical outcomes might be achieved from genetic studies of rare plants (Hogbin and Peakall 1999; Hogbin *et al.* 2000). There is also increasing evidence that information about lineage and phylogenetic diversity can be incorporated into conservation modelling and planning strategies (Faith and Baker 2006; Pollock *et al.* 2014). Given that approximately two thirds of *Zieria* species are threatened, vulnerable or endangered, this study potentially provides data to underpin conservation management decisions, and contribute to the evaluation of phylogenetic diversity in various parts of Australia and New Caledonia.

1.6. Phylogenetic research in *Zieria*

Armstrong (2002) undertook a cladistic analysis of *Zieria* based on 48 morphological and 62 phytochemical characters. It resolved six major clades, which were designated as morphological species groups based on particular combinations of characters. However the infrageneric phylogeny was proposed as ‘tentative’; relationships between and within the six major clades were poorly resolved or supported, and there was a high level of homoplasy reported (stated as ‘75%’, but a consistency index was not given). The six groups resolved in Armstrong’s (2002) phylogeny are as follows:

- Group A (14 species): *Z. adenodonta*, *Z. adenophora*, *Z. buxijugum*, *Z. collina*, *Z. floydii*, *Z. formosa*, *Z. furfuracea*, *Z. granulata*, *Z. hindii*, *Z. obcordata*, *Z. parrisiae*, *Z. robusta*, *Z. tuberculata* and *Z. verrucosa*.

These species are characterised by the presence of distinct tubercles on the younger branches, petioles, midveins, peduncles and fruits.

- Group B (five species): *Z. arborescens*, *Z. caducibracteata*, *Z. lasiocaulis*, *Z. oreocena* and *Z. southwellii*.

These species are characterised by branchlets that are terete (i.e. they do not have distinct ridges) or only slightly ridged. Their primary inflorescence bracts are navicular (boat shaped) and deciduous, leaving obvious scars on the peduncle. The abaxial surface of the calyx lobes is hirsute with mostly stellate hairs (except *Z. lasiocaulis* which has predominantly simple hairs).

- Group C (four species): *Z. chevalieri*, *Z. fraseri*, *Z. laevigata* and *Z. laxiflora*

These species are characterised by branchlets that are distinctly ridged due to the presence of prominent glabrous to subglabrous leaf decurrencies, where the base of the petiole extends down the stem producing a ridge. The lower laminar surface is velutinous, but the midveins are glabrous and dotted with pellucid oil glands. Inflorescences are equal to or longer than leaves. Calyx lobes are inflexed adaxially. Anthers are distinctly apiculate. Fruits are dotted with pellucid glands but are not tuberculate. *Zieria chevalieri*, the New Caledonian species, falls into this group and has linear leaves similar to *Z. laevigata* and *Z. laxiflora*, but differs in that it has unifoliate rather than trifoliate leaves. *Zieria fraseri* subsp. *compacta* is now recognised as a species, *Z. compacta* (Duretto and Forster 2007). It is not known which subspecies of *Z. fraseri* was used in the Armstrong (2002) analysis.

- Group D (four species): *Z. montana*, *Z. prostrata*, *Z. robertsiorum* and *Z. smithii*

These species are characterised by distinctly ridged branchlets with prominent glabrous to subglabrous leaf decurrencies, as in Group C, but differ in that the lower laminar surface is glabrous, and dotted with pellucid oil glands that turn black and become sunken on drying (a character diagnostic of this group). The petioles may be either dotted with pellucid glands or tubercles, and the midveins are glabrous to glabrescent and dotted with pellucid dots. Fruits are dotted with pellucid glands but are not tuberculate, also as in Group C.

- Group E (eight species): *Z. aspalathoides*, *Z. citriodora*, *Z. ingramii*, *Z. minutiflora*, *Z. obovata*, *Z. odorifera*, *Z. pilosa* and *Z. rimulosa*

These species are characterised by branchlets that are densely pubescent. The upper lamina leaf surfaces are covered exclusively with simple hairs (as opposed to stellate hairs). The surface of the lower lamina and the midvein are hirsute. Distinctive stamens have filaments that are warty towards the apex and prominently apiculate anthers. Ovaries are pubescent. Fruits may be glabrous or pubescent and have apiculate cocci.

- Group F (seven species): *Z. baeuerlenii*, *Z. covenyi*, *Z. cytisoides*, *Z. involucrata*, *Z. littoralis*, *Z. murphyi* and *Z. veronicea*

These species are characterised by lamina surfaces that are velutinous (i.e. with a downy appearance) and not tuberculate. Inflorescences are equal to or longer than the leaves and have navicular primary bracts. The fruits are pubescent.

Published phylogenetic studies of *Zieria* are limited. They include: the study by Armstrong (2002) described above; a study by Morton (2015) based on a combination of chloroplast DNA, nuclear DNA and Armstrong's (2002) morphological and chemical characters; and Chapter 3 of this thesis (Barrett *et al.* 2015), a study based on chloroplast DNA (mostly referred to in this thesis as 'Chapter 3'). The high level of homoplasy in Armstrong's cladogram clearly emphasised the need for a subsequent phylogenetic analysis of *Zieria* based on molecular data, in order to test relationships. Armstrong's (2002) morphological groups were also incongruent with Morton's (2015) analysis, despite Morton's inclusion of Armstrong's morphological data in combination with molecular data. Furthermore, because Armstrong's (2002) analysis predates the revision of Queensland *Zieria* taxa (Duretto and Forster 2007), there are 19 recently described, newly recognised or reinstated Queensland taxa yet to be included in a phylogenetic analysis. These new taxa were not recognised by Morton (2015) who only recognised 42 (of 60) species.

1.7. Biogeography

Connections between Australia and New Caledonia

The Rutaceae is of particular interest in the context of biogeographical connections between Australia and New Caledonia, because many genera, and some species, are shared between the two regions (Bayly *et al.* 2013). A review of biogeographic patterns and geology with respect to the Australia – New Caledonia connections was undertaken by Ladiges and Cantrill (2007), but there is continuing controversy over the history of New Caledonia and its flora (Lowry 1998; Bayly *et al.* 2013). In particular, there is debate over the importance of long-distance over-water dispersal (Grandcolas *et al.* 2008; Swenson *et al.* 2014) versus vicariance (Bauer *et al.* 2006; Ladiges and Cantrill 2007; Heads 2008) to explain biogeographic patterns such as the distribution of closely

related taxa that occur in disjunct locations in this region. Molecular phylogenies of various plant groups have been interpreted to provide support for both sides of the argument (see references in Bayly *et al.* 2013), and the occurrence of one species of *Zieria* in New Caledonia presents another opportunity for contributing to this debate. An early divergence between *Zieria* lineages in Australia and New Caledonia would potentially lend itself to a vicariant explanation for distribution of the genus. A more recent divergence of *Z. chevalieri* from an Australian sister taxon (i.e. more terminal node position on the phylogenetic tree), may be interpreted as evidence of dispersal, over-water or via past exposed land in the Tasman Sea. Any explanation, however, would have to be consistent with the very limited dispersal capacity of the ant-dispersed seeds.

Connections in south-eastern Australia (Zieria arborescens)

Hypotheses about the evolution of the Australian biota have previously been formulated on the basis of evidence from fossil records, plate tectonics, landscape development and climate change (Byrne *et al.* 2011). Molecular phylogenies, especially when time-calibrated, can augment the data and test such hypotheses (Crisp *et al.* 2004; Hugall *et al.* 2008). Some models of post-glacial recovery of mesic forest species, based on macrofossil and phylogenetic evidence, support the isolation and persistence of species in multiple refugia during periods of glaciation, followed by later expansions (Stewart and Lister 2001; McLachlan *et al.* 2005; Petit *et al.* 2008), rather than dispersal over thousands of kilometres from just a few isolated refugia (see references in Worth *et al.* 2009). Phylogenetic analyses of species that are widespread and morphologically variable have the potential to test such models of post-glacial recovery for southern hemisphere mesic forest taxa (Beheregaray 2008). *Zieria arborescens*, an eastern Australian wet forest understorey species, is one such species. It is morphologically variable and relatively widespread in south-eastern mainland Australia and Tasmania. A study of the biogeography of *Z. arborescens* will contribute to the relatively limited number of southern hemisphere phylogeographic studies that propose models of post-glacial recovery for southern hemisphere mesic forest taxa (Beheregaray 2008). Such studies potentially provide information about the resilience of mesic biota in the face of long-term habitat and climate fluctuations (global warming in particular). Importantly,

predictions about the responses of organisms could also assist in the development of climate change adaptation strategies.

1.8. Main objectives

The overarching aim of this research was to gain insight into the evolution and classification of *Zieria* by producing an infrageneric phylogeny of the genus using chloroplast and nuclear ribosomal DNA markers and including as many species as possible. The major project objectives were to:

1. Produce a comprehensive molecular phylogeny of *Zieria* in Australia and New Caledonia, based on chloroplast and nuclear DNA data.
2. Assess the monophyly of taxa and consider issues of taxonomy suggested by the results of this study.
3. Explore patterns of speciation, e.g. relationships of widespread and narrowly endemic taxa.
4. Examine patterns of variation in widespread species.
5. Consider the relationships and distinctiveness of highly restricted (and rare) species that will allow the validity of such taxa, and their conservation value, to be assessed.
6. Infer biogeographic connections between Australia and New Caledonia, with reference to the molecular phylogenetic relationships between eastern Australian *Zieria* species and the single New Caledonia species.
7. Investigate genetic variation and phylogeographic patterns in the widespread species, *Zieria arborescens*, focussing on Victorian and Tasmanian populations.

Chapter 2. General methods

2.1. Introduction

Many of the same or similar materials and methods were used for sequencing and phylogenetic analysis of *Zieria* using chloroplast and nuclear DNA markers (Chapters 3 and 4), and in the phylogeographic study of *Z. arborescens* (Chapter 5). Therefore a summary of the general methods used is presented here. Some specific details are given in each chapter where appropriate. The authority names of taxa sampled for this study are given in Appendix A, with details of voucher specimens and geographic localities of all samples. The Materials and Methods section of Chapter 3 (published) is condensed to some extent for publication and more details may be found in this chapter, but there is unavoidably some repetition.

2.2. Collection of plant material

Ingroup taxa were selected to include as many species of *Zieria* as possible. DNA was extracted from 169 accessions representing 59 of the 60 currently recognized species and 16 (of 18) subspecies (although subspecies were not the focus of this study). *Zieria floydii* J.A.Armstr., *Z. aspalathoides* subsp. *brachyphylla* J.A.Armstr. and *Z. odorifera* subsp. *copelandii* Durretto & P.I.Forst. were the only taxa not sampled. Multiple accessions of some species were included when material was available, or when increased sampling was thought to be important, for example, for the widely distributed, morphologically variable taxon, *Z. smithii*. *Zieria arborescens* was sampled widely for a detailed phylogeographic study (Chapter 5), but a more limited number of accessions of that species was included in the large phylogenetic analyses of the genus (Chapters 3 and 4). Also included in this study was the monotypic genus, *Neobyrnesia*, strongly supported as the sister genus to *Zieria* based on *rbcL* and *atpB* DNA sequences (Bayly *et al.* 2013).

Plant material was obtained from fresh field collections, living cultivated collections of known provenance (Australian National Botanic Gardens, Canberra, and Royal Botanic Gardens Victoria), and from herbarium specimens. Plant collecting permits were provided by: the Victorian Department of Sustainability and Environment; New South Wales National Parks and Wildlife Service; Queensland Environmental Protection

Agency; Royal Botanic Gardens; Australian National Botanic Gardens; the Australian Department of the Environment; and conservation authorities of the North Province of New Caledonia (DDEE). Fresh samples for DNA extraction were dried in silica gel and surplus material has been stored in silica gel in the Systematics Laboratory (Botany), School of BioSciences, at the University of Melbourne. Herbarium specimens for examination were borrowed from BRI, CBG, and NSW. Permission for destructive sampling of some herbarium specimens was granted by the directors of BRI, MEL, MELU and NSW.

The analyses presented here used samples and DNA sequences obtained specifically for this project; i.e., the small chloroplast and ITS datasets of Morton (2015) were not included. This was primarily so that cpDNA and nrDNA phylogenies based on the same accessions could be directly compared. It was not practical to combine the data obtained specifically for this study with the datasets of Morton (2015) which contained significantly fewer taxa, did not include ETS sequences in the nrDNA dataset, and whose cpDNA dataset was based on fewer markers; missing data unnecessarily lower resolution in phylogenetic analyses. All the species included in Morton (2015) are also represented in this project (in six cases, using the same accessions based on plants cultivated at the Australian National Botanic Gardens).

Specimens were identified using current taxonomic concepts (George *et al.* 2013), except for *Z. arborescens*, a variable, widespread species in which separation of the subspecies is not as clear as identification keys suggest (Duretto and Forster 2007; George *et al.* 2013; R. A. Barrett, pers. obs.). Accessions of *Z. arborescens* were not identified to subspecies, but in figures and tables those accessions collected near type localities for the two segregate subspecies are noted (subsp. *glabrifolia* and subsp. *decurrens*).

2.3. DNA extraction

DNA was extracted using a DNeasy Plant Mini Kit (Qiagen, Melbourne, Victoria, Australia) or ISOLATE Plant DNA mini Kit (Bioline, Sydney, New South Wales, Australia). For each sample, 20–25 mg of dried leaf material was ground manually in a mortar and pestle with the aid of acid washed sand (Ajax, Finechem, Australia), and then the manufacturer's instructions followed from Step 7 (page 25 of the DNeasy Plant

Handbook, July 2006) or from step 2 (page 11 of the Bioline product manual). DNA was eluted in elution buffer in two steps of 50 μL each, giving a final elution volume of 100 μL . For some herbarium specimens a modified kit protocol was used to increase DNA yield (T.M. Schuster pers. comm.); this included the addition of extra reagents during the lysis stage: 8 μL of proteinase K, 4 μL RNase A and 30 μL of beta mercaptoethanol, and incubation at 60° C for one hour. Extracted DNA was stored in the elution buffer at –20° C.

2.4. Use of molecular phylogenetic markers

Chloroplast DNA (cpDNA)

Molecular studies are frequently undertaken to augment morphological studies and test species boundaries, with different markers appropriate for studies at different taxonomic levels. In recent years chloroplast markers have been widely used for resolving phylogenetic relationships in plants (Davis *et al.* 2014), including Rutaceae (e.g., Chase *et al.* 1999; Morton *et al.* 2003; Poon *et al.* 2007; Groppo *et al.* 2008, 2012; Harbaugh *et al.* 2008; Pfeil and Crisp 2008; Salvo *et al.* 2008, 2011; Bayer *et al.* 2009; Othman *et al.* 2010; Appelhans *et al.* 2013, 2014; Bayly *et al.* 2013, 2015; Morton *et al.* 2014; French *et al.* 2016).

The chloroplast genome has been studied extensively and is thought to be a relictual molecule, derived from an endosymbiotic relationship between once free-living photosynthetic cyanobacteria-like (prokaryotic) organisms and eukaryotic organisms (reviewed in Palmer 1987; Palmer and Delwiche 1996, 1998). It encodes about 100 genetic functions (see Palmer 1991; Clegg *et al.* 1994; Steane 2005). In land plants the chloroplast genome is circular and, with few exceptions, has a characteristic structure (Figure 2.1), which consists of a large and a small single copy region separated by two inverted repeats (Palmer 1987; Soltis and Soltis 1998). Many complete chloroplast genomes have now been sequenced, representing a wide variety of plant types, including *Citrus sinensis* (sweet orange) in the Rutaceae (Bausher *et al.* 2006).

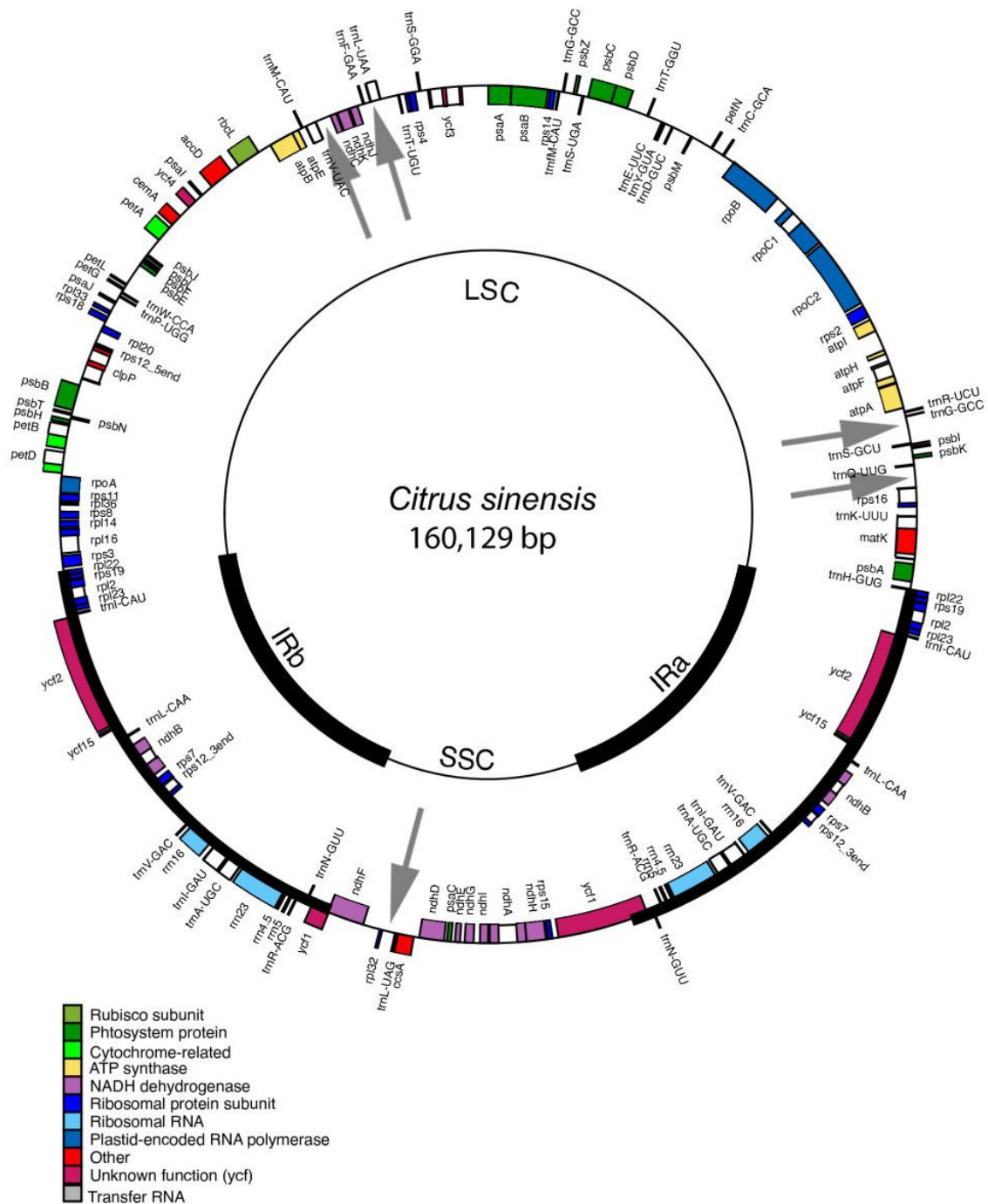


Figure 2.1. A circular gene map of the chloroplast genome of *Citrus sinensis* from Bausher *et al.* (2006), indicating the location of the chloroplast markers used in this study (grey arrows). The thick lines represent the inverted repeats (IRa and IRb, 26,996 base pairs), which separate the genome into small and large single copy regions (SSC, 18,393 base pairs and LSC, 87,744 base pairs respectively). Transcription occurs in a clockwise direction in genes on the outside of the map, and anti-clockwise in genes on the inside of the map.

The chloroplast genome is relatively stable, and well conserved across land plants in terms of size, structure, gene content and the order of the genes (Soltis and Soltis 1998). Consequently, changes in structure, arrangement and DNA content are likely to have phylogenetic significance. Typically, cpDNA is inherited uniparentally, usually maternally in flowering plants, and is passed on only by seeds (Palmer 1987; Petit *et al.* 1993), enabling comparisons to be made between seed and pollen gene flow. Comparing uniparentally inherited cpDNA to biparentally inherited nuclear DNA also assists the identification of hybridization events (Bonatelli *et al.* 2012). Other characteristics that make cpDNA particularly useful for evolutionary studies in higher plants (see Palmer 1987) are that: it occurs in high numbers in green plant cells, a few thousand per leaf cell; it is relatively small in size compared to nuclear DNA, typically consisting of about 120 to 200 thousand base pairs long; it consists of mostly single copy genes and a large number of conserved regions (whilst most nuclear genes are members of multi-gene families, which can diminish their phylogenetic usefulness); and although different regions of cpDNA display differing rates of evolution (Soltis and Soltis 1998; Shaw *et al.* 2007), its nucleotide substitution rate is two to three times slower than that of nuclear genes, but faster than that of mitochondrial genes (Clegg *et al.* 1994; Avise 2009). Also, the stable structure and conserved gene order of the chloroplast genome has allowed a number of universal primers to be designed (Taberlet *et al.* 1991; Small *et al.* 1998; Shaw *et al.* 2005, 2007).

Intraspecific variation is frequently reported in cpDNA, despite its evolutionary conservation, and there are potentially large numbers of informative characters, which can be used in phylogenetics. Non-coding regions such as intergenic spacers and introns are particularly useful for resolving relationships in closely related taxa (Kelchner 2000; Shaw *et al.* 2007), with intergenic spacer regions having been found to be more variable than introns (Shaw *et al.* 2007). The most frequent types of mutation in these non-coding regions are point mutations (single nucleotide polymorphisms; SNPs) and insertion-deletion (indel) mutations, but variation associated with indels in non-coding regions is prone to homoplasy and should be treated carefully in phylogenetic analyses (Bonatelli *et al.* 2012). Inversions and indels in coding regions are rarer due to functional constraints (Gielly and Taberlet 1994; Sang *et al.* 1997).

cpDNA markers used in this study

The choice of cpDNA markers for use in this project was based largely on the findings of Shaw *et al.* (2007), which ranked 34 chloroplast regions on the basis of their variability, according to their number of potentially informative characters (PICs). Preliminary screening of several cpDNA markers was undertaken to assess which markers could be amplified and give sufficient variation for a phylogenetic analysis of *Zieria*, and for a more focussed phylogeographic study of a selected species, *Z. arborescens*. Six chloroplast markers were trialled and four were subsequently selected for use in the phylogenetic analysis of the genus:

- the intergenic spacer *rpl32–trnL*^(UAG)
- the *trnL* intron and *trnL–trnF* intergenic spacer (*trnL–F*)
- the intergenic spacer *trnQ*^(UUG)–5'*rps16*
- the *trnS–trnG* intergenic spacer and *trnG* intron (*trnS–G*).

The *ndhF–rpl32* region proved difficult to amplify successfully in *Zieria* and was not pursued. One additional cpDNA marker, the 3'*trnV*^(UAC)–*ndhC* intergenic spacer, was used in the phylogeographic study of *Z. arborescens* (Chapter 5) to increase the number of variable characters for this finer scale investigation. The putative positions of the chloroplast markers used in this study are indicated (grey arrows) on the circular map of the *Citrus sinensis* chloroplast genome (Bausher *et al.* 2006; Figure 2.1); the genes in *Zieria* are expected to be in similar positions given the conservation of gene order in cpDNA.

The cpDNA markers used are briefly outlined below, and the primers and their sequences are listed in Table 2.1.

1. The *rpl32–trnL*^(UAG) intergenic spacer (*rpl32–trnL*) is located on the SSC region of the chloroplast genome, flanked by the ribosomal protein (*rp*) *l32* gene and the *trnL* gene which encodes the transfer RNA (tRNA) for the amino acid leucine. This region is highly variable (Timme *et al.* 2007), but was not used in sequence-based studies prior to 2007 (Shaw *et al.* 2007). It has since been found to be one of the chloroplast regions providing the highest number of PICs (Shaw *et al.* 2007). In this study the *rpl32–trnL* region was amplified using the primers *trnL*^(UAG) and *rpl32–F* (Shaw *et al.* 2007).

2. The *trnL* intron and *trnL*–*trnF* intergenic spacer (*trnL*–*F*) was one of the first chloroplast regions to be employed in phylogenetic studies (Gielly and Taberlet 1994; van Ham *et al.* 1994), and has been widely used in studies at various taxonomic levels, partly due to the availability of universal primers that were designed for this region by Taberlet *et al.* (1991). It includes two non-coding regions, the *trnL* intron and *trnL*–*F* intergenic spacer. The *trnL* and *trnF* genes encode tRNA for the amino acids leucine and phenylalanine, respectively. In this study the *trnL*–*F* region was amplified using the primers c and f of Taberlet *et al.* (1991), and in some cases where amplification was unsuccessful initially, the region was amplified in two overlapping fragments using the following combinations of internal primers (Taberlet *et al.* 1991): e paired with f; and d paired with c.

3. The *trnQ*^(UUG)–5'*rps16* (*trnQ*–5'*rps16*) intergenic spacer has been found to be highly variable (Daniell *et al.* 2006; Shaw *et al.* 2007; Timme *et al.* 2007). The spacer is located in the large single copy (LSC) region of the genome, flanked by the *trnQ*^(UUG) gene which codes for the tRNA for glutamine, and the 5' exon of *rps16* which codes for the 30S ribosomal protein CS16. In this study the *trnQ*–5'*rps16* region was amplified using the primers *trnQ*^(UUG) and *rps16x1* (Shaw *et al.* 2007).

4. The *trnS*–*trnG*–*trnG* (*trnS*–*G*) region is located in the LSC region of the genome and includes the *trnS*–*trnG* intergenic spacer and the *trnG* intron. The *trnS* and *trnG* genes encode tRNA for the amino acids serine and glycine respectively. In this study the *trnS*–*G* region was most commonly amplified using the primers of Shaw *et al.* (2005), *trnG*^(UUC) and *trnS*^(GCU), but in a few cases it was amplified in two overlapping fragments using the internal primers 5'*trnG2G* paired with *trnG*^(UUC) and 5'*trnG2S* paired with *trnS*^(GCU).

5. The 3'*trnV*^(UAC)–*ndhC* (3'*trnV*–*ndhC*) intergenic spacer is located in the LSC region of the chloroplast genome, flanked by the *trnV* gene which encodes tRNA for the amino acid valine, and the *ndhC* gene which encodes NADH dehydrogenase subunit A. It has been found to be highly variable (Shaw *et al.* 2007; Timme *et al.* 2007) but there are limited phylogenetic studies that have confirmed this. In this study the 3'*trnV*–*ndhC* regions was amplified using the primers *rpl32*–*R* and *ndhC* (Shaw *et al.* 2007).

Chapter 2

Table 2.1. cpDNA primers used in this study.

Region	Primer	Sequence (5' – 3')	Reference
<i>rpl32–trnL</i>	<i>trnL</i> ^(UAG)	CTG CTT CCT AAG AGC AGC GT	Shaw <i>et al.</i> 2007
	<i>rpl32–F</i>	CAG TTC CAA AAA AAC GTA CTT C	Shaw <i>et al.</i> 2007
<i>trnL–F</i>	Tab c	GCA AAT CGG TAG ACG CTA CG	Taberlet <i>et al.</i> 1991
	Tab f	ATT TGA ACT GGT GAC ACG AG	Taberlet <i>et al.</i> 1991
	Tab e	GGT TCA AGT CCC TCT ATC CC	Taberlet <i>et al.</i> 1991
	Tab d	GGG GAT AGA GGG ACT TGA AC	Taberlet <i>et al.</i> 1991
<i>trnQ–5'rps16</i>	<i>trnQ</i> ^(UUG)	GCG TGG CCA AGY GGT AAG GC	Shaw <i>et al.</i> 2007
	rpS16x1	GTT GCT TTY TAC CAC ATC GTT T	Shaw <i>et al.</i> 2007
<i>trnS–G</i>	<i>trnG</i> ^(UUC)	GAA TCG AAC CCG CAT CGT TAG	Shaw <i>et al.</i> 2007
	<i>trnS</i> ^(GCU)	AAC TCG TAC AAC GGA TTA GCA ATC	Shaw <i>et al.</i> 2007
	5' <i>trnG2G</i>	GCG GGT ATA GTT TAG TGG TAA AA	Shaw <i>et al.</i> 2005
	5' <i>trnG2S</i>	TTT TAC CAC TAA ACT ATA CCC GC	Shaw <i>et al.</i> 2005
<i>3'trnV–ndhC</i>	<i>trnV</i> ^(UAC) x2	GTC TAC GGT TCG ART CCG TA	Shaw <i>et al.</i> 2007
	<i>ndhC</i>	TAT TAT TAG AAA TGY CCA RAA AAT ATC ATA TTC	Shaw <i>et al.</i> 2007
<i>ndhF–rpl32</i>	<i>rpl32–R</i>	CCA ATA TCC CTT YYT TTT CCA A	Shaw <i>et al.</i> 2007
	<i>ndhF</i>	GAA AGG TAT KAT CCA YGM ATA TT	Shaw <i>et al.</i> 2007

Nuclear ribosomal DNA (nrDNA)

Compared to the wide variety of cpDNA markers that have been used in phylogenetic analyses in recent years, fewer genes in the nuclear genome of plants have been investigated, despite its larger size and the great variety of genes it contains (Soltis and Soltis 1998). However, since about 1992 nuclear ribosomal DNA (nrDNA) has been extensively used for phylogenetic reconstruction in plants (Soltis and Soltis 1998; Sang 2002; Alvarez and Wendel 2003).

The multi-copy gene families of nrDNA are composed of arrays of genes and spacers, including the 18S–5.8S–28S cistron which codes for the 18S, 5.8S and 28S ribosomal subunit genes and contains internal transcribed spacers (ITS) and external transcribed spacer (ETS) regions (Feliner and Rosselló 2007). The ITS regions consist of two spacer regions, referred to as ITS1 and ITS2, which are adjacent to the conserved 5.8S ribosomal RNA (rRNA) subunit gene, and flanked by the conserved 18S and 28S rRNA genes (Figure 2.2). The 5' end of the 18S gene is flanked by the ETS region, which is then flanked by a non-transcribed spacer (NTS). The NTS and ETS regions collectively form an intergenic spacer. The whole unit is repeated in tandem thousands of times,

with each transcribed region being separated by a non-transcribed region (Feliner and Rosselló 2007). The cistron may be present on several different chromosomes (Thompson and Flavell 1988).



Figure 2.2. A simplified diagram of the repeat unit structure of 18S–5.8S–28S nuclear ribosomal DNA (adapted from Shiao *et al.* 2009; not to scale). NTS = non-transcribed spacer; ETS = external transcribed spacer; ITS1 and ITS2 = internal transcribed spacers 1 and 2.

Although the ITS1 and ITS2 spacer regions are not incorporated into completed ribosomes, they are involved in the production, processing and maturation of ribosomes (van der Sande 1992; Baldwin *et al.* 1995). The whole transcriptional unit (consisting of the coding and ITS regions, and ETS) is processed to produce mature RNAs, subunits of cytoplasmic ribosomes (Feliner and Rosselló 2007). It is thought that ETS might have a similar but independent role to that of ITS in the maturation of ribosomes (Baldwin and Markos 1998), and given that both regions belong to the same transcriptional unit, they possibly evolve at similar rates, under similar functional constraints. Although the spacer regions are not entirely ‘non-functional’, they are less functionally constrained than the flanking gene sequences, such as the 5.8S subunit (Yokota *et al.* 1989). Thus the spacer regions potentially provide a degree of intergenomic variability that can be used for inferring relationships in phylogenetic studies.

The ITS and ETS regions have been widely used for resolving phylogenetic relationships in plants, including Rutaceae (e.g., Scott *et al.* 2000; Mole *et al.* 2004; Poon *et al.* 2007; Othman *et al.* 2010; Appelhans *et al.* 2013, 2014; Amar *et al.* 2014; Bayly *et al.* 2015; French *et al.* 2016). ITS in particular has been used extensively for phylogenetic studies, especially for analyses of relationships within genera or between closely related genera, and for barcoding and taxonomic studies in wide range of plant groups (Baldwin *et al.* 1995; Alvarez and Wendel 2003; Hughes *et al.* 2006; Song *et al.* 2012). The phylogenetic utility of ETS has been explored in fewer plant groups, but it

often augments ITS data, yielding additional characters for phylogenetic inferences in angiosperms (Baldwin and Markos 1998; Bena *et al.* 1998).

ITS and ETS have several characteristics that make them useful phylogenetic markers (reviewed in Feliner and Rosselló 2007). For example, in contrast to the maternally-inherited chloroplast markers, nuclear markers are biparentally inherited, and can reveal reticulate evolution, hybridization resulting in speciation, and the parentage of polyploids (Rieseberg and Soltis 1991; Baldwin 1992; Baldwin *et al.* 1995). In most plants the 18S–5.8S–28S cistron occurs at one or more chromosomal loci in tandem arrays composed of hundreds to thousands of copies (reviewed in Hamby and Zimmer 1992; Small *et al.* 2004), which, compared to low-copy nuclear cistrons, makes them relatively easy to detect, amplify, clone and sequence, even from herbarium specimens (Baldwin *et al.* 1995; Alvarez and Wendel 2003; Feliner and Rosselló 2007). The ‘universal’ ITS primers described by White *et al.* (1990) have been successfully employed across a range of plant and fungal taxa, eliminating the need to design specific primers, which requires prior knowledge of the sequence. In flowering plants ITS sequences are also relatively short, typically 500 to 700 base pairs in length, further facilitating PCR amplification and Sanger sequencing, often without the need for internal primers (Baldwin *et al.* 1995; Feliner and Rosselló 2007). In some groups the structurally complex ETS region is longer and more phylogenetically informative than ITS1 and ITS2 combined (Linder *et al.* 2000), but because the 5' end of the spacer lacks a highly conserved region, universal ETS primers for use in a wide number of plant families have not been developed (Linder *et al.* 2000). ETS is flanked by the highly variable non-transcribed spacer (NTS) which evolves too rapidly in sequence and length to provide a universal primer site, although the highly conserved 18S gene provides primer sites options downstream from the 3' end of the spacer (Baldwin and Markos 1998).

Multigene families, including the 18S–5.8S–28S cistron, commonly undergo a process known as concerted evolution, in which all the member genes evolve as a unit (‘in concert’). This tends to homogenize the reiterated sequences within and between loci, and is assumed to provide a degree of intergenomic uniformity in sequences that would otherwise be accumulating different mutations (Zimmer *et al.* 1980; Arnheim 1983; Baldwin *et al.* 1995), and eliminating problems with coalescence (Linder *et al.* 2000).

The process happens by repeated unequal crossing over between sister chromosomes, or gene conversion (Alvarez and Wendel 2003; Nei and Rooney 2005), and occurs frequently because the long regions of homologous DNA sequence are good substrates for the general genetic recombination process. Concerted evolution ideally results in a single predominant sequence across all copies and arrays, allowing PCR products to be directly amplified and sequenced, presumably to yield the predominant copy (Small *et al.* 2004).

However if concerted evolution is slow or does not act on all copies, divergent copies of nrDNA can persist (Childs *et al.* 1981). If undetected, the presence of divergent paralogues within a single genome can confound attempts at phylogenetic reconstruction (Sanderson and Doyle 1992). Divergent paralogues of ITS in particular have been reported in numerous plant and fungal groups in recent years (see references in Bayly *et al.* 2008).

The potential pitfalls of using nrDNA in phylogenetic reconstructions have been well documented (Balwin 1992; Mayol and Rosselló 2001; Sang 2002; Álvarez and Wendel 2003; Small *et al.* 2004; Feliner and Rosselló. 2007). They include: the presence of heterogeneous repeat types (divergent paralogues) or pseudogenes (genes that have degenerated and lost functionality); recombination of divergent copies; incomplete concerted evolution (described above); problems in alignment due to the accumulation of indels; incorrect assumptions of orthology; inherent secondary structure which can lead to compensatory mutations; contamination, due to universality of sequences and the use of universal primers. These issues can all potentially lead to increased homoplasy in ITS data sets, and high levels of homoplasy has indeed been reported in nrDNA sequence data (Álvarez and Wendel 2003).

nrDNA markers used in this study

Two nrDNA regions were chosen for this phylogenetic study, ITS and ETS, as described above. The ITS region was commonly amplified using ITS4 and ITS5 (White *et al.* 1990), although for some problematic templates different primer combinations were used including alternative forward primers, ITS18 or S3 (Käss and Wink 1997), and the reverse primer ITS26 (Käss and Wink 1997). In a few cases ITS was amplified in two overlapping fragments using the internal primers of Käss and Wink (1997), S5

and S6. Although the primers S3, S5, and S6 were designed as sequencing primers (Käss and Wink 1997) they can be useful in overcoming the problems of secondary structure in ITS and were used successfully in this study to amplify various fragments of ITS. For some problematic templates, amplification was affected or enhanced by substituting 10 µL of water with 10 µL of a trehalose-based additive, TBT-PAR (Samarakoon *et al.* 2013), in the PCR reactions. The ETS region was amplified using the primers ETS-18S (Wright *et al.* 2001) and ETS-MyrtFb (designed by Eve Lucas, Royal Botanic Gardens, Kew, and published by Bayly *et al.* 2015). The nuclear primers and their sequences are presented in Table 2.2.

Table 2.2. nrDNA primers used in this study.

Region	Primer	Sequence (5' – 3')	Reference
ITS	ITS4	TCC TCC GCT TAT TGA TAT GC	White <i>et al.</i> 1990
	ITS5	GGA AGT AAA AGT CGT AAC AAG G	White <i>et al.</i> 1990
	S3	AAC CTG CGG AAG GAT CAT TG	Käss and Wink 1997
	S5	TTC GGG CGC AAC TTG CGT TC	Käss and Wink 1997
	S6	ATA TCT CGG CTC TTG CAT CG	Käss and Wink 1997
	ITS18	GTC CAC TGA ACC TTA TCA TTT AGA CC	Käss and Wink 1997
	ITS26	GCC GTT ACT AAG GGA ATC CTT GTT AG	Käss and Wink 1997
	prosMajorF	GCG CTC CCC TCT ACC TCA	This study, designed for <i>Z. prostrata</i>
	prosMinorF	GCG CTC CCC TCT GCT ACA	This study, designed for <i>Z. prostrata</i>
ETS	ETS-18S	GAG CCA TTC GCA GTT TCA CAG	Wright <i>et al.</i> 2001
	ETS-MyrtFb	GCA GGC TCC GTG CTG GTG C	See note above, in the text

2.5. DNA amplification

All polymerase chain reactions (PCRs) were performed in a MyCycler Thermal Cycler System (Bio-Rad Laboratories, Gladesville, New South Wales, Australia) in total reaction volumes of 25 µL. Thermal cycler lids were preheated to 105° C for PCR amplifications. All sets of PCR reactions included a negative control to test for contamination, and in most cases a positive control was included to test for reagent or protocol errors.

cpDNA amplification

Reactions for amplification of cpDNA typically included 10 pmol of each primer, 5 mM of each dNTP, 1–2 µL of extracted DNA, 1.25 units HotStar *Taq* DNA polymerase and its accompanying 10 × PCR buffer (QIAGEN), 0.5 µL Fermentas non-acetylated BSA (20 mg/mL), and 2 mM MgCl₂. These were made up to the 25 µL final volume with ultrapure water.

Thermo-cycling conditions varied according to the target region. Cycling conditions for *rpl32-trnL* and *trnS-G* were: template denaturing at 95° C for 15 minutes, followed by 30 cycles of 95° C for 1 minute (denaturing), primer annealing at 50° C for 1 minute, 65° C extension for 4 minutes, then a final 5 minute extension at 65° C. Cycling conditions for *trnQ-5'rps16* were 95° C for 15 minutes, 30 cycles of 95° C for 1 minute, 50° C for 1 minute with a 0.5° C temperature increase at each cycle, 65° C for 4 minutes, then a final 5 minute extension at 65° C. A 'touchdown' protocol was used to amplify *trnL-F*: 95° C for 15 minutes, then six cycles of 95° C for 30 seconds, primer annealing at 64° C for 30 seconds with a 2° C temperature decrease at each cycle, extension for 1 minute at 72° C, then a further 30 cycles of 94° C for 30 seconds, 54° C for 30 seconds, 72° C for 1 minute, and a final extension for 10 minutes at 72° C.

nrDNA amplification

Reactions for ITS and ETS amplification typically included 10 pmol of each primer, 5 mM of each dNTP, 1–2 µL of extracted DNA, 1.25 U HotStarTaq DNA polymerase, 2.5 µL of the accompanying 10 × PCR buffer (QIAGEN), 1.25 µL dimethyl sulfoxide (DMSO), and were made up to the final volume of 25 µL with ultrapure water.

Typical thermo-cycling conditions for ITS were as follows: initial template denaturing at 95° C for 15 minutes followed by 35 cycles of 94° C for 1 minute (denaturing), primer annealing at 50° C for 1 minute, 72° C extension for 1 minute, and then a final 10 minute extension at 72° C. In some cases PCR conditions were adjusted; for example, the annealing temperature was raised to increase specificity of the primers, or, when nothing was amplified using the usual protocol, the annealing temperature was lowered to decrease the specificity of the primers. Thermo-cycling conditions for ETS were as follows: 95° C for 15 minutes, 35 cycles of 94° C for 1 minute, 50° C for 1 minute, 72° C for 30 seconds and then a final 5 minute extension at 72° C.

2.6. Purification and quantification of PCR products

The length and quantity of PCR products were assessed by electrophoresis, typically using 4 μL aliquots of PCR product mixed with 1 μL of loading dye containing glycerol, 1.5% w/v agarose gel containing ethidium bromide in 1 X TBE buffer, and electrophoresed at 80 volts for 25 minutes. Electrophoresis gels were then assessed under UV light. Suitable PCR products were purified to degrade primers and dNTPs that were not consumed during PCR. Purification of PCR products was carried out using either: a kit as per the manufacturer's instructions, i.e., QIAquick PCR Purification Kit (QIAGEN, Melbourne, Victoria, Australia) or PureLink Kit (Invitrogen, Melbourne, Victoria, Australia); or single step enzymatic clean-up of PCR products using ExoSAP-IT (Affymetrix, Santa Clara, California, USA) or ExoSTAR (GE Healthcare, Freiburg Germany). For enzymatic clean-ups, the calculated quantity of PCR product required for sequencing was purified using a ratio of 1:0.4 PCR product to enzyme. Enzyme treatment was carried out for 15 minutes at 37° C, followed by a 15 minute incubation at 80° C to inactivate the enzymes.

Following kit purification, DNA concentration of PCR products was measured using a NanoDrop 2000 UV-Vis spectrophotometer (Thermo-Scientific, Melbourne, Victoria, Australia) and 1 μL of the sample. However, a spectrophotometer cannot be used to measure DNA concentration following enzymatic clean-up, therefore in these instances DNA concentration was estimated following electrophoresis using DNA reference ladders, EasyLadder 1 (Bioline) or HyperLadder 1 (Bioline).

2.7. DNA sequencing

Purified PCR products were prepared for sequencing using ABI Prism BIG Dye v3.1 Terminator Cycle Sequencing Ready Reaction Kits (Applied Biosystems, Foster City, California, USA). Sequencing reactions contained 3.5 μL of sequencing buffer, 1 μL of Big Dye Terminator mix, 0.16 μM of each primer, 4–25 ng of PCR product depending on its length, and ultrapure water to make up the total volume of 20 μL . Amplification parameters were 96° C for 2 minutes, then 30 cycles of 96° C for 10 seconds, 50° C for 5 seconds, and extension of 4 minutes at 60° C.

Extension products were precipitated in 1.5 µL microcentrifuge tubes using the ethanol/sodium acetate precipitation method (Applied Biosystems, Foster City, CA, USA) as follows: 2 µL of 3 M sodium acetate (pH 4.6) were mixed with 50 µL of 95% ethanol, then added to the 20 µL sequencing reaction and briefly vortexed. Tubes were left at room temperature for 15 minutes for precipitation of products, then centrifuged at maximum speed (14,000 rpm) for 20 minutes. The supernatant was discarded, and the tubes were wick dried. The pellets were washed with 250 µL of 70% ethanol, then centrifuged at maximum speed for a further 5 minutes. The supernatant was discarded and the tubes were dried again thoroughly on a heat block at 90° C for 1 minute before being sealed. Samples prepared this way were sent to the Australian Genome Research Facility (AGRF), Melbourne or Brisbane, in dried down pellet form for resuspending and loading directly onto an ABI 3730xl 96-capillary automated DNA sequencer (Applied Biosystems). Alternatively, purified DNA was delivered to AGRF in plates containing the recommended amount of DNA template (depending on the length of the PCR product), 9.6 pmol of primer, and ultrapure water to make up a total volume of 12 µL per reaction. Direct sequencing of PCR products was performed using the same forward and reverse primers that were used for amplification, and also using internal sequencing primers when required, i.e., for all *trnS*–G sequences due to the length of the region (about 1600 base pairs), for some *trnL*–F sequences, and for some ITS sequences. Raw sequence data were received electronically as chromatograms, and fasta files.

2.8. Sequence editing and alignment

Contiguous sequences for each region were assembled and edited using Sequencher v.4.8 (Gene Codes Corporation, Ann Arbor, Michigan, USA), and ambiguous sites were edited manually. Visual inspection of chromatograms enabled detection of polymorphic loci within individuals. Edited contiguous sequences were exported and aligned manually using Se-Al Sequence Alignment Editor v. 2.0a11 (Rambaut 2002). Published sequences are available in GenBank (see the appendix in Chapter 3).

2.9. ITS and ETS paralogues

Detection of divergent ITS and ETS paralogues

Many of the direct sequences of ITS and ETS PCR products showed some within-genome polymorphisms, i.e., double peaks in sequencing chromatograms (see Appendix B). Visual inspection of chromatograms revealed distinct ‘major’ and ‘minor’ peaks at polymorphic loci in some samples, indicating potentially divergent sequences. A trial separation of potentially divergent copies of ITS and ETS was undertaken ‘by eye’ for a few selected samples, using base calls from major peaks to represent one sequence and those from minor peaks for a paralogous copy. Preliminary phylogenetic analyses in PAUP* 4.0 beta 10 (Phylogenetic Analysis Using Parsimony; Swofford 1991) suggested that these sequences (if correctly called) were deeply divergent and there was strong bootstrap support for their placement in different parts of the nrDNA tree (trees not presented). To confirm the results, copy-specific ITS primers were designed (prosMajorF and prosMinorF, see Table 2.2) for one sample that showed particularly high numbers of polymorphisms in both ITS and ETS sequences (*Z. prostrata* MJB2019); this was to trial separating divergent ITS copies, the presence of which had been predicted by visual assessment of the chromatograms as described above. The copy-specific primers were located in ITS1 where there was a cluster of polymorphic loci, approximately 65 base pairs from the end of the 18S gene. These forward primers were used in conjunction with the reverse primer ITS4 (White *et al.* 1990). The presence of deeply divergent paralogues of ITS was subsequently confirmed in *Z. prostrata* MJB2019; the sequences matched those determined ‘by eye’ exactly, although there was one polymorphism in one of the copies, indicating the presence of a third copy differing by a single base.

Cloning of ITS and ETS paralogues

Following this trial, four samples with numerous polymorphisms were selected for cloning, with the main aim being to amplify notably divergent copies of both ITS and ETS, but not necessarily to detect *all* copies; it was assumed that copies differing by only a base or two would be unlikely to give a different phylogenetic signal. The four samples selected for cloning were: *Z. exsul* MTM535, *Z. insularis* PIF34666, *Z. odorifera* subsp. *williamsii* MJB2018, and *Z. prostrata* MJB2019.

Amplification of sequences for cloning was carried out as previously described for ITS and ETS PCRs, except MyTaq DNA polymerase (Bioline) was used, with its accompanying 5× reaction buffer, which also contains the nucleosides. PCR products were purified using Invitrogen PureLink purification kits (Invitrogen, Melbourne, Victoria, Australia) according to the manufacturer's instructions and eluted into 30 µL elution buffer. DNA was quantified using a NanoDrop 2000 UV-Vis spectrophotometer (Thermo-Scientific, Melbourne, Victoria, Australia).

Cloning of ITS and ETS of the selected taxa was performed using a TA Cloning Kit (Invitrogen) according to the manufacturer's instructions, except where specified. PCR products were ligated into the vector pCR®2.1. Quantities of 9 ng of ITS PCR product (about 700 base pairs) and 6.4 ng of ETS PCR product (about 500 base pairs) were calculated to be optimal for ligation efficiency, using the formula given in the product manual. Ligation reactions were incubated at room temperature for 30 minutes, then on ice for 30 minutes. Immediately following shaking, 200 µL of culture was spread onto pre-warmed plates of LB medium containing 100 µg/mL of ampicillin, and 40 µL of X-gal (40 mg/µL) applied to the surface of the medium. Transformed cells were incubated for 16 hours at 37° C. For each region five to twenty colonies per sample were picked from the plates and placed directly into 30 µL pH8 Tris buffer. Colonies were incubated at 94° C for 10 minutes to ensure that all bacteria were killed. Inserts were amplified directly from up to ten positive colonies using the M13F and M13R primers (Table 2.3) and other reagents as previously described. Cycling parameters were 95° C for 5 minutes, 33 cycles of 95° C for 1 minute, 55° C for 1 minute, 65° C for 4 minutes, then a final 10 minute extension at 65° C. PCR products were assessed and sequenced as previously described.

Table 2.3. Primers used for amplification of cloned PCR products.

Primer	Sequence (5' – 3')	Reference
M13-F	GTA AAA CGA CGG CCA G	Supplied with Invitrogen TA Cloning Kit
M13-R	CAG GAA ACA GCT ATG AC	Supplied with Invitrogen TA Cloning Kit

Sequences were screened and unique copies were included in preliminary (separate) ITS and ETS phylogenetic analyses to evaluate their usefulness and placement in the trees. Cloned ITS and ETS sequences from the same accession were concatenated for the

combined ITS+ETS analysis if they were shown to be giving a congruent phylogenetic signal (see Appendix C for details).

Comparison of ITS and ETS paralogues

Paralogues within and among individuals were compared in order to discern potentially functional copies of nrDNA from pseudogenes. Comparisons included G/C content, the number of CpG and CpNpG methylation sites and the presence of conserved motifs, which have all been shown to vary between functional and pseudogenic nrDNA (Buckler *et al.* 1997; Bayly and Ladiges 2007). Sequences were assessed for the presence of four motifs in the 5.8S gene previously identified as highly conserved in angiosperms, and one motif in the ITS1 spacer. In the 5.8S gene the motifs of interest were: (1) 5'-GAA TT CA AA CC (Jobes and Thien 1997); (2) an EcoRV restriction site (GA ATC) near the 5' end of the gene (Liston *et al.* 1996); (3) and (4) the motifs M1 and M3 identified by Harpke and Peterson (2008). Also of interest was the conserved motif 5'-CAA GGA A, identified in many flowering plants by Liu and Schardl (1994), located near the centre of the ITS1 region and usually associated with a hairpin structure. The distribution of variable sites across the 5.8S gene and the spacers, ITS1 and ITS2, was compared among paralogues. Boundaries between these regions were identified by comparisons with previously published sequences (Bayly and Ladiges 2007).

2.10. Partition homogeneity test (ILD test) of nrDNA data

The question of whether or not to combine data sets for phylogenetic analyses is complex and there are many arguments for and against (reviewed in Nixon and Carpenter 1996). The incongruence length difference (ILD) test (Farris *et al.* 1995a), or partition homogeneity test as implemented in PAUP* 4.0 beta 10 (Phylogenetic Analysis Using Parsimony; Swofford 1991), is a statistical test of incongruence between partitions within data sets. It is sometimes used to decide whether to combine data sets. When little or no incongruence is detected it is considered reasonable to combine datasets into a simultaneous analysis. It is standard to assume incongruence between data sets if the ILD test result is significant at $p \leq 0.05$ less than 5%, although some

authors regard a more stringent value of $p \leq 0.01$ as indicative of significant incongruence between data sets (Groppo *et al.* 2008; Hong-Wa 2013).

Several authors report the failure or inappropriateness of the ILD test (for reviews see Mason-Gamer and Kellogg 1996; Cunningham 1997; Dolphin *et al.* 2000; Yoder *et al.* 2001; Dowton and Austin 2002; Barker and Lutzoni 2002; Darlu and Lecointre 2002; Ramírez 2006). Significant results in the test are not always due to conflicting phylogenetic signal, but can be generated by noise in the data (Dolphin *et al.* 2000), differences in size of the partitions (Dowton and Austin 2002), insufficient informative characters (Darlu and Lecointre 2002), or differences in evolutionary rates of the characters (Barker and Lutzoni, 2002; Darlu and Lecointre, 2002). Others suggest using the test with caution or as a first step to determine how much incongruence might be present (Mason-Gamer and Kellogg 1996; Hipp *et al.* 2004). Nixon and Carpenter (1996) recommend combining data sets even if the amount of incongruence is deemed significant by the ILD test.

Despite these conflicting views, for completeness, the partition homogeneity test was performed on the nrDNA dataset using the Partition Homogeneity (HomPart) option in PAUP, with data partitioned into ITS and ETS sets, and using only informative characters. The test used 1000 random repartitions and a heuristic search. Search options were: start by stepwise addition, closest addition sequence, hold one tree at each step, TBR branch swapping, and Maxtrees set to 2000 with no increase when that limit was reached. The ILD test was not performed on the cpDNA datasets because cpDNA data is readily combinable, being inherited as a unit from one parent in flowering plants (usually maternally).

2.11. Recombination analysis of nrDNA data

Recombination explains a substantial amount of genetic diversity in natural populations (Posada and Crandall 2001), and the results of phylogenetic analyses can be compromised by the presence of recombinant sequences (McDade 1992; Posada 2002). Therefore, the understanding and interpretation of phylogenetic relationships are greatly improved if it is known whether or not the sequences have undergone recombination (McDade 1992).

A recombination analysis of the aligned nrDNA dataset was performed using the computer program RDP3 (Heath *et al.* 2006) to detect evidence of recombination between different paralogues and orthologues of nrDNA in *Zieria*. In groups of sequences that were identical to each another, all but one sequence were discarded. The analysis implemented the following seven recombination detection methods: RDP (Martin and Rybicki 2000), GENECONV (Padidam *et al.* 1999), BOOTSCAN (Martin *et al.* 2005), MAXIMUM CHI SQUARE (Maynard Smith 1992), CHIMAERA (Posada and Crandall 2001), SISCAN (Gibbs *et al.* 2000) and 3SEQ (Boni *et al.* 2007). The program's default settings were used for all routines, except in General Recombination Detection Options, where sequences were recognized as 'linear', and in Data Processing Options, 'disentangle recombination signals' was selected, as recommended in the manual. The highest acceptable probability value for reporting potential recombination events was set to 0.05, with Bonferroni correction for multiple comparisons (both default settings). Also, following the recommendation of Posada (2002) that conclusions about recombination should not be based on the results of a single method, the program was set to return only recombination events recovered by more than one method.

Chloroplast genomes do not undergo recombination (Doyle 1992), so cpDNA datasets were not subjected to recombination analysis.

2.12. Phylogenetic analyses

Bayesian inference (BI) and maximum parsimony (MP) methods were used for phylogenetic analyses of cpDNA and nrDNA datasets (Chapters 3 and 4), and the phylogeographic study of *Z. arborescens* (Chapter 5).

Bayesian inference

Bayesian inference (BI) is now widely used to infer phylogenies, having become particularly popular with advances in computing capabilities. Bayesian inference is based on Bayes' theorem, first developed by Reverend Thomas Bayes, and published posthumously (Bayes 1763). The theorem uses the prior probability distribution (i.e. observations or data summarized in a hypothesis or model) and the likelihood of the data to generate a posterior probability distribution. In phylogenetics, for example, it

describes the probability of an outcome (a phylogenetic reconstruction or tree) taking into account conditions that might be related to that outcome (a model of evolution given the data). It infers the most likely phylogenetic tree for the given data, i.e., the tree with the highest posterior probability is chosen as the best estimate of phylogeny for the given dataset. One of the main differences between BI and traditional methods such as maximum parsimony or maximum likelihood is the incorporation of prior information (simple or complex models of evolution) that previously limited computational analyses using traditional methods (Huelsenbeck and Ronquist 2001).

Comparing posterior probabilities with bootstrap values as statistical measures of support for phylogenies, some authors (Suzuki *et al.* 2002, Erixon *et al.* 2003) argue that oversimplified models can give higher posterior probabilities in Bayesian analyses than bootstrap probabilities, which are generally more conservative and sometimes considered more reliable. Given that phylogenies inferred using Bayesian methods are directly linked to the model or models of evolution chosen, it is important to choose a model that fits the observed data, or inferences of phylogeny are likely to be incorrect. In this study, appropriate models of evolution were chosen using the Akaike Information Criterion (AIC) as implemented in MrModeltest 2.3 (Nylander 2004). The models used are given in each chapter.

Bayesian inference (BI) analyses were undertaken using MrBayes version 3.2.1 (Ronquist and Huelsenbeck 2003). All data sets were partitioned by DNA region, and parameters between partitions were unlinked. All BI analyses used the default settings in MrBayes and a sufficient number of generations of concurrent Markov chain Monte Carlo runs for convergence. Details of the cpDNA data analysis are given in Chapter 3. Otherwise all analyses consisted of two runs of four chains for varying numbers of generations as follows: (1) in Chapter 4, 2.1 million for ITS data, 2.4 million for ETS data, and 1.4 million for the combined ITS+ETS data; (2) in Chapter 5, 3 million for the combined cpDNA data, and 1.3 million for the combined ITS+ETS dataset. In all analyses trees were sampled every 500 generations and a majority rule consensus was computed, with trees from the first 10% of generations discarded as burn-in. Tracer v.1.5 (Rambaut 2007) was used to check that the burn-in period was adequate for each analysis and that the runs had converged on a stationary distribution. This was judged by comparing the distribution likelihood values and the standard deviation of split

frequencies, which were < 0.01 at the end of the runs. Clades with posterior probability (PP) values ≥ 0.95 were considered to be strongly supported.

Parsimony Analysis

Parsimony analysis is based on the proposal that the simplest explanation is usually the correct one, often referred to as Ockham's Razor, and attributed to William of Ockham, (c. 1280 – 1349), a 14th century English Franciscan friar, philosopher and theologian,¹. Parsimony may be used to choose between competing scientific hypotheses to explain relationships, and is widely used to infer phylogenies on the premise that the best hypothesis when building phylogenetic trees is the one requiring the fewest evolutionary changes (Farris 1983; Swofford *et al.* 1996; Doyle and Davis 1998; Kitching *et al.* 1998).

Hennig (1965) stated that phylogenetic relationships are relative, i.e., simply, the idea that two species, or species groups, are more closely related to each other than they are to a third. The link to the relationship is via a shared common ancestor. Hypotheses for discerning these sister-groups are based on the analysis of discrete characters. Hennig's cladistic methods were based on the recognition of plesiomorphies (ancestral character states) and apomorphies (derived character states), which are identified prior to construction of a cladogram. Only shared derived characters, or synapomorphies, provide evidence of common ancestry. In cladistic phylogenetic analysis using parsimony, the aim is to produce a phylogenetic tree with the least number of steps or evolutionary changes required to explain the observed result.

Hennig (1966) identified three types of groups based on ancestry and descent (discussed in Forey *et al.* 1992; Kitching *et al.* 1998). These groups are: (1) monophyletic groups, which contain the most recent common ancestor and all its descendants, and are discovered through synapomorphies; (2) paraphyletic groups, which do not include all the descendants of a common ancestor, but are what remains when part of a monophyletic group has been removed; they are based on symplesiomorphies (shared ancestral character states) and are recognised by characters they do not have, i.e., those

¹ 'Essentia non sunt multiplicanda praeter necessitate' [Entities should not be multiplied unnecessarily] or 'Frustra fit per plura, quod fieri potest per pauciora' [It is vain to do with more what can be done with less.] <http://www-groups.dcs.st-and.ac.uk/~history/Quotations/Ockham.html>

characters possessed only by more derived groups²; (3) polyphyletic groups, which are founded on convergent or homoplasious characters, i.e., characters that appear the same or similar but which were not inherited from a common ancestor (wings in birds and bats for example). Polyphyletic groups tend to be thought of as artificial by most systematists.

Many studies now use both Bayesian inference and maximum parsimony methods to infer phylogenies. The use of parsimony analysis for molecular data is sometimes criticized on statistical grounds (Felsenstein 1978; Ronquist 2004), because artefacts of the analysis such as ‘long branch attraction’ can be misleading (see Bergsten 2005 for a review), and because parsimony analysis is not based on any particular model of evolution (see Swofford *et al.* 1996; Steel and Penny 2000).

In this project maximum parsimony (MP) analysis was undertaken using the computer software PAUP* 4.0 beta 10 (Phylogenetic Analysis Using Parsimony; Swofford 1991). The heuristic search option was applied with the following settings: all characters unordered and of equal weight, gaps treated as a fifth state, starting trees obtained via stepwise addition, closest addition sequence followed by tree bisection-reconnection (TBR) branch swapping. Multistate characters in individual sequences were treated as polymorphisms in the nuclear analysis, and as uncertainties in the chloroplast analysis (although there were no polymorphisms of this sort in the chloroplast data).

When initial phylogenetic analysis yielded many thousands of equally parsimonious trees, an expeditious search strategy was employed as follows. An initial heuristic search was commenced using the above settings, and terminated when 50,000 trees were obtained, then a strict consensus of equally parsimonious trees was calculated. This consensus tree was loaded as a constraint in a second analysis that searched only for trees that were as short as or shorter than the shortest tree of the first analysis, but were inconsistent with the constraint tree. This search used 1000 random addition sequences, each followed by TBR branch swapping. Each replicate was aborted when 2000 trees with lengths exceeding that of the first analysis were obtained. When this strategy was used, if trees of equal length to that of the first analysis were obtained, *inconsistent* with the consensus from the first analysis, a further consensus was

² Nelson (1989) believes that Hennig’s most original, important and lasting contribution to systematics was the recognition (and elimination) of paraphyletic groups

computed including the topology from the new analysis. The second step of the analysis was repeated until no more new equally short but inconsistent trees were found.

The search strategy described above was used for the combined cpDNA analysis, individual ITS and ETS analyses, and the combined ITS+ETS analysis, which all yielded many thousands of equally parsimonious trees during the initial heuristic search. The aim of this search strategy was to produce a strict consensus tree that was likely to represent the full set of equally most parsimonious trees. A strict consensus tree, which is the most conservative consensus tree (Kitching *et al.* 1998), is a summary of the information common to a set of ‘rival’ (equally most parsimonious) trees and only contains groups appearing in all rival trees (Swofford 1993). Comparisons of tree topologies employed in consensus methods aim to detect overall stability among individual cladograms, but do not maximize parsimony because they do not take into consideration the relative strength of character evidence (Miyamoto 1985). Therefore a strict consensus tree is not an optimal tree for any given data set and is usually longer than the shortest trees in the group (Kitching *et al.* 1998). Polytomous nodes in a strict consensus indicate uncertain resolution rather than the simultaneous divergence of multiple lineages.

For all analyses, bootstrap support (Felsenstein 1985) was estimated using a full heuristic search with 1000 bootstrap replicates and the search settings as described above. A maximum of 2000 trees $\geq$ score 1 was saved per replicate. Groups with frequencies of less than 50% were eliminated.

Outgroups

A phylogenetic study of the Australasian Rutoideae (Bayly *et al.* 2013) showed the monotypic genus *Neobynesia* to be strongly supported as the sister taxon of *Zieria*, and *Neobynesia suberosa* was therefore set as the outgroup for the phylogenetic analyses of both cpDNA and nrDNA datasets (Chapters 3 and 4). In the phylogeography study of *Z. arborescens* (Chapter 5) the accession *Z. smithii* MJB2029 was selected as the outgroup for both cpDNA and nrDNA datasets based on the results of analyses presented in Chapters 3 and 4.

Inclusions and exclusions

1. Missing data

Accessions with large amounts of missing sequence data (e.g. a whole region in a combined dataset) were generally excluded from analyses. This is because misleading results or a dramatic increase in the number of equally parsimonious trees can be obtained when a large proportion of missing data is included in phylogenetic analyses (Wiens 1998). This is a potential problem when combining data from several markers and a whole region is not able to be amplified for some samples. In this study, preliminary analysis of the data, first including, then excluding samples with large amounts of missing data, found that the inclusion of samples with missing data often had the effect of collapsing branches or reducing statistical support for some nodes.

2. Polymorphisms

Accessions were generally excluded from analyses if their ITS or ETS sequences contained four or more polymorphisms (double or multiple peaks in chromatograms). Exceptions were the four samples that were selected for bacterial sequence cloning (see Section 2.9 ITS and ETS paralogues). The number of polymorphisms (> 0) in the ITS and ETS sequences of the samples is given in Appendix B; samples with high numbers of polymorphisms that were subsequently excluded are also noted.

3. Mononucleotide repeats

Variable-length mononucleotide repeat regions (poly-structures or single nucleotide microsatellites) are often homoplasious and exact numbers of repeats are difficult to determine by direct sequencing of PCR products (Wenpan *et al.* 2012). Consequently, all variable-length polynucleotide repeat regions, present only in the chloroplast data, were excluded from the analyses.

4. Indel coding and missing data

Indel characters were included in all analyses, unless there was uncertainty about the alignment in that part of the sequence. Although Kellerman and Udovicic (2008) found that including large homoplasious indel characters in a phylogenetic analysis can obscure results, Bonatelli *et al.* (2013) advocate the use of indels because they are a common feature in cpDNA and important for resolving haplotype relationships.

In this study, multi-base indels were common in the cpDNA data but not in the nrDNA data. Indels were mostly coded using a method equivalent to the ‘simple coding’ of Simmons and Ochoterena (2000), except in the BI analysis of the cpDNA data in Chapter 5 where they were coded separately and partitioned. Typically, each indel, single- or multi-base, was represented by one character and all other characters in that indel were excluded, to avoid overweighting indel characters. In cases where there was variation within indels, selected characters were included to maximize informativeness but not to overweight characters within that indel. For example, all variable base positions within multi-base indels were included if possible, and, in sequences with the deletion, a single character was used to denote the deletion (usually a dash to represent a gap as a fifth state) and the other variable positions within that indel were recorded as ‘missing’ data (usually a question mark).

Chapter 3. A chloroplast phylogeny of *Zieria*

This study was published in the journal *Australian Systematic Botany* in 2015. A copy of the pdf file is presented here.

SPECIAL ISSUE

CSIRO PUBLISHING

Australian Systematic Botany, 2014, 27, 427–449
<http://dx.doi.org/10.1071/SB14033>

A chloroplast phylogeny of *Zieria* (Rutaceae) in Australia and New Caledonia shows widespread incongruence with species-level taxonomy

Rosemary A. Barrett^A, Michael J. Bayly^{A,E}, Marco F. Duretto^B, Paul I. Forster^C,
Pauline Y. Ladiges^A and David J. Cantrill^D

^ASchool of Botany, The University of Melbourne, Vic. 3010, Australia.

^BNational Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Mrs Macquaries Road, Sydney, NSW 2000, Australia.

^CQueensland Herbarium, Department of Science, Information Technology, Innovation and the Arts, Brisbane Botanic Gardens, Toowong, Qld 4066, Australia.

^DRoyal Botanic Gardens Melbourne, Birdwood Avenue, South Yarra, Vic. 3141, Australia.

^ECorresponding author. Email: mbayly@unimelb.edu.au

Abstract. This study presents a molecular phylogeny of *Zieria* Sm., a genus of shrubs and small trees, with 59 species in Australia and one endemic to New Caledonia. The phylogeny is based on four cpDNA markers and 116 samples representing all species of *Zieria* except one, and the monotypic outgroup *Neobyrsesia suberosa*. The New Caledonian species, *Z. chevalieri*, was resolved as sister to a well supported clade of all Australian taxa. There was widespread incongruence between the cpDNA tree and species-level taxonomy, with 14 species shown as polyphyletic or paraphyletic. These included widespread species (e.g. *Z. smithii* and *Z. arborescens*, each falling in at least four well supported clades) and some with narrow geographic ranges (e.g. *Z. alata* and *Z. oreocena*). No species represented by three or more samples was resolved as monophyletic. We suggest that a combination of factors explains this incongruence, including regional cpDNA introgression (chloroplast capture), incomplete lineage sorting and inappropriate taxonomic boundaries. The cpDNA phylogeny provides useful insight into the evolution of *Zieria* but, because of its complexity, does not provide a clear basis for assessing phylogenetic relationships and monophyly of taxa. Better understanding of relationships, taxon limits and evolutionary processes in *Zieria* will require comparisons with nuclear DNA markers and critical assessment of morphological and genetic variation in widespread species.

Additional keywords: biogeography, chloroplast lineage sorting, hybrids, molecular phylogeny, taxonomy.

Received 30 September 2014, accepted 18 December 2014, published online 29 June 2015

Introduction

Zieria Sm. is a predominantly Australian genus of shrubs and small trees, with 59 species and 18 subspecies in Australia and one species endemic to New Caledonia (Fig. 1). In Australia, *Zieria* occurs naturally only in the eastern states, extending as far west as Kangaroo Island in South Australia, as far south as Tasmania, and north to the tropical regions of Queensland (Armstrong 2002; George *et al.* 2013). The greatest diversity of species occurs in New South Wales (34 species) and Queensland (35 species), whereas in Victoria there are nine species, three in Tasmania and a single species in South Australia. The species occupy a range of habitats, including rainforests and rainforest margins, sometimes at high altitudes, vine thickets, wet and dry sclerophyll forest, eucalypt and *Banksia* woodlands (including mallee), heathlands, rock outcrops on dry inland slopes, and exposed coastal headlands.

The genus is traditionally placed in the Australasian tribe Boronieae (*sensu* Engler 1931), and is monophyletic on the basis of morphological (Armstrong 1991) and molecular (Bayly *et al.* 2013) phylogenetic analyses of that group. It is distinguished by a combination of morphological features including opposite leaves, conspicuous and four-merous flowers, free petals, four stamens, free filaments, deeply four-lobed disc and dry dehiscent fruits (Smith 1798; Duretto 1999; Armstrong 2002; George *et al.* 2013). It is most closely related to the monotypic genus *Neobyrsesia* J.A.Armstr. endemic to Arnhem Land, northern Australia (Armstrong and Powell 1980; Bayly *et al.* 2013); *Neobyrsesia* and *Zieria* are allopatric (Fig. 1).

Classification of *Zieria* has undergone extensive revision in recent decades chiefly on the basis of morphology. A comprehensive monograph was provided by Armstrong (2002, based on his earlier PhD thesis, Armstrong 1991), who

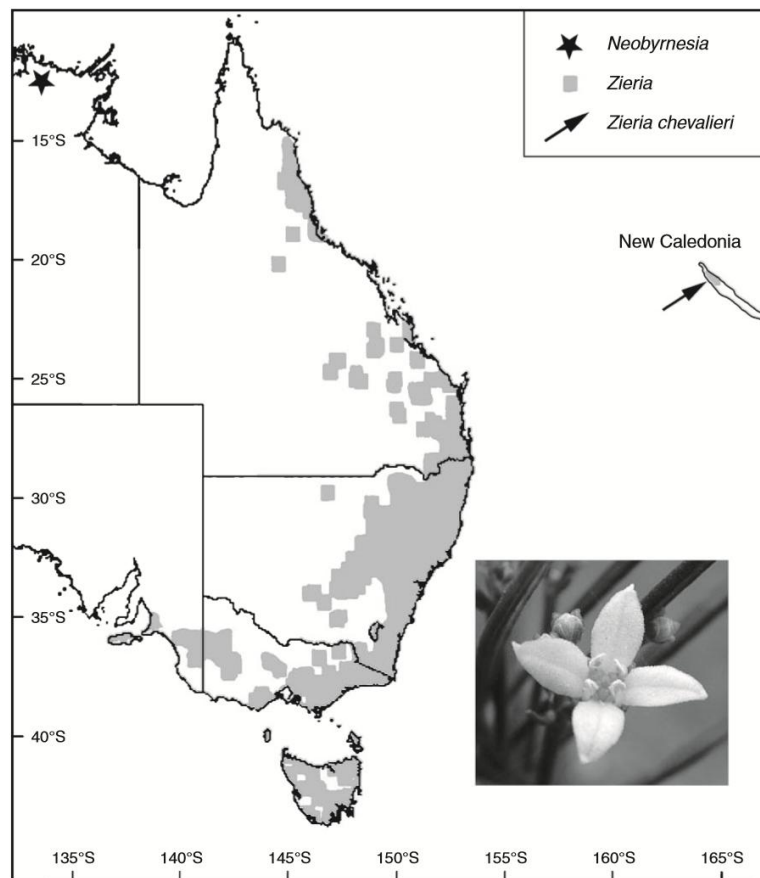


Fig. 1. Map showing the distribution of *Zieria* and *Neobyrsesia* (outgroup) in Australia and New Caledonia, adapted from Armstrong (2002). Inset shows the flower of *Z. chevalieri*, Mount Kaala, New Caledonia.

recognised 42 species and 16 subspecies, with 28 taxa being newly described. Duretto and Forster (2007) published a subsequent revision of the Queensland taxa, in which 16 new species and one new subspecies were described, one variety was promoted to species rank, and one species previously treated as a subspecies was reinstated. Duretto and Forster (2008) also reviewed the classification of *Z. odorifera*^A, a species endemic to New South Wales, recognising four subspecies, with three newly described. The most recent account of the genus is in *Flora of Australia* (George *et al.* 2013).

Some species of *Zieria* are relatively widespread (e.g. *Z. arborescens*, *Z. aspalathoides*, *Z. cytisoides*, *Z. smithii*); however, the majority are narrow-range endemics. Several taxa are currently known only from single localities, including *Z. actites*, *Z. adenophora*, *Z. arborescens* subsp. *decurrens*, *Z. baeuerlenii*, *Z. boolbunda*, *Z. buxijugum*, *Z. cephalophila*, *Z. chevalieri*, *Z. covenyi*, *Z. floydii*, *Z. formosa*, *Z. hydroscopica*

and *Z. parrisiae* (Armstrong 2002; Duretto and Forster 2007; George *et al.* 2013). Duretto and Forster (2007) discussed the speciation patterns in *Zieria*, including how such narrow endemics, often restricted to rocky areas on mountains, may be related to more widespread species, often present in forests or more mesic vegetation. Narrow endemics may represent taxa differentiated in longstanding refugial habitats, or recent derivatives of more widespread lineages that have locally colonised new habitats (e.g. from forests to rock outcrops).

Of the large number of narrowly endemic taxa in *Zieria*, some are considered threatened under state and federal conservation legislation. These include 13 listed as *Endangered* and eight listed as *Vulnerable* under the national *Environment Protection and Biodiversity Conservation (EPBC) Act* 1999, 13 under the *Queensland Nature Conservation (Wildlife) Regulation* 2006, 17 under the *New South Wales Threatened Species Conservation Act* 1995, one under Victoria's *Flora and Fauna Guarantee Act*

^AAuthorities for *Zieria* species and infraspecific taxa included in the study are given in Appendix 1.

1988, two under South Australia's *National Parks and Wildlife Act* 1972, and two under Tasmania's *Threatened Species Protection Act* 1995.

Some threatened species have been the subject of genetic research, including *Z. baeuerlenii* (Sharma 2001), *Z. granulata* (Sharma and Young 2001), *Z. covenyi* (J. A. Armstrong and I. K. Sharma, unpubl. data, cited by Armstrong 2002) and *Z. prostrata* (Hogbin and Crisp 2003). These studies assessed the levels of genetic variation to provide data to underpin conservation management, and variously reported moderate variation to no variation (e.g. in the clonally reproducing *Z. covenyi*). Most focussed on rare species in isolation and did not consider their broader relationships with other taxa, the exception being that of Hogbin and Crisp (2003), who studied relationships between populations of *Z. prostrata* and geographically close populations of *Z. smithii*.

The phylogeny of *Zieria* produced by Armstrong (2002) was based on a dataset of 48 morphological and 62 chemical characters. The phylogeny showed six major clades, defined by combinations of characters rather than clear synapomorphies, and relationships between and within the clades were poorly resolved or supported. The results were regarded as 'very tentative' by Armstrong, given the high levels of homoplasy, with the consistency index for the analysis being low at 0.25. Since publication of Armstrong's (2002) phylogeny, taxonomic revisions (Duretto and Forster 2007, 2008) have resulted in narrower or significantly altered limits for some taxa (e.g. *Z. montana* and *Z. fraseri*), and 22 newly recognised or reinstated taxa are yet to be included in a phylogenetic analysis. A new, molecular phylogenetic study that includes a greater proportion of currently recognised taxa offers potential to further test relationships and interpret morphological homoplasy demonstrated by Armstrong (2002).

The aim of our current study was to produce a comprehensive molecular phylogeny of *Zieria* in Australia and New Caledonia, based on sequencing four regions of cpDNA. This was done to provide better insight into evolution and classification of the genus, including: the monophyly of taxa, patterns of speciation, e.g. relationships of widespread and narrowly endemic taxa; the distinctiveness of restricted taxa of conservation interest; patterns of variation in widespread species; patterns of habitat diversification; biogeographic history, especially the connection between Australia and New Caledonia.

Materials and methods

Taxon sampling

Ingroup sampling included 116 accessions representing 59 (of 60) *Zieria* species and 16 (of 18) subspecies. *Zieria floydii* J.A.Armstr., *Z. aspalathoides* subsp. *brachyphylla* J.A.Armstr. and *Z. odorifera* subsp. *copelandii* Duretto & P.I.Forst. were the only taxa not sampled. *Neohyrsesia suberosa* was included as the outgroup. Plant material was obtained from field collections, cultivated collections of known provenance, and herbarium specimens. Fresh samples for DNA extraction were dried in silica gel. Details of sampling locations and voucher specimens for all samples are given in Appendix 1. Identification of taxa was straightforward using existing taxonomic concepts (George *et al.* 2013), except for *Z. arborescens*, a variable,

widespread species in which separation of the subspecies is not as clear as identification keys suggest (Duretto and Forster 2007; George *et al.* 2013; R. A. Barrett, pers. obs.). We did not identify accessions of *Z. arborescens* to subspecies, but we note in figures and tables those accessions collected near type localities for the two segregate subspecies (subsp. *glabrifolia* and subsp. *decurrens*).

DNA isolation, polymerase chain reaction (PCR) and sequencing

DNA was extracted from 20–25 mg of dried, ground leaf material using a DNeasy Plant Mini Kit (QIAGEN, Melbourne, Vic., Australia) or ISOLATE Plant DNA Mini Kit (Bioline, Sydney, NSW, Australia), following the manufacturers' instructions, with a final elution volume of 100 µL. For some herbarium specimens, a modified protocol was used, which included the addition of the following extra reagents during the lysis stage: 8 µL of proteinase K, 4 µL RNase A and 30 µL of β mercaptoethanol, and incubation at 60°C for 1 h.

The following four cpDNA markers were used: the *rpl32-trnL* intergenic spacer, the *trnL-F* region (which included the *trnL* intron and the *trnL-trnF* intergenic spacer), the *trnQ-5'rps16* intergenic spacer and the *trnS-G* region (which included the *trnG* intron and the *trnG-trnS* intergenic spacer). The *rpl32-trnL* region was amplified using the primers *trnL*^(UAG) and *rpl32-F* (Shaw *et al.* 2007). The *trnL-F* region was amplified using the primers *c* and *f* of Taberlet *et al.* (1991), and in some cases where amplification was unsuccessful, the region was amplified in two overlapping fragments using the following combinations of internal primers (Taberlet *et al.* 1991): *e* paired with *f*; and *d* paired with *c*. The *trnQ-5'rps16* region was amplified using the primers *trnQ*^(UUG) and *rps16x1* (Shaw *et al.* 2007). The *trnS-G* region was amplified using the primers of Shaw *et al.* (2005), most commonly *trnG*^(UUC) and *trnS*^(GCU), but in a few cases it was amplified in two overlapping fragments using the internal primers *5'trnG2G* paired with *trnG*^(UUC) and *5'trnG2S* paired with *trnS*^(GCU).

All polymerase chain reactions (PCRs) were performed in a MyCycler Thermal Cycler System (Bio-Rad Laboratories, Gladesville, NSW, Australia). All sets of PCRs included a negative control to test for contamination, and, in most cases, a positive control was included to test for reagent or protocol errors. Reactions typically included 10 pmol of each primer, 5 mM of each dNTP, 1–2 µL of extracted DNA, 1.25 units HotStar *Taq* DNA polymerase and its accompanying 10 × PCR buffer (QIAGEN), 0.5 µL Fermentas non-acetylated BSA (20 mg mL⁻¹), and 2 mM MgCl₂. These were made up to the 25 µL final volume with ultrapure water.

Thermo-cycling conditions varied according to the target region. Cycling conditions for *rpl32-trnL* and *trnS-G* were as follows: template denaturing at 95°C for 15 min, followed by 30 cycles of 95°C for 1 min (denaturing), primer annealing at 50°C for 1 min, 65°C extension for 4 min, and then a final 5-min extension at 65°C. Cycling conditions for *trnQ-5'rps16* were as follows: 95°C for 15 min, 30 cycles of 95°C for 1 min, 50°C for 1 min with a 0.5°C temperature increase at each cycle, 65°C for 4 min, and then a final 5-min extension at 65°C. A 'touchdown' protocol was used to amplify *trnL-F*, including 95°C for 15 min,

then six cycles of 95°C for 30 s, primer annealing at 64°C for 30 s, with a 2°C temperature decrease at each cycle, extension for 1 min at 72°C, then a further 30 cycles of 94°C for 30 s, 54°C for 30 s, 72°C for 1 min, and a final extension for 10 min at 72°C.

PCR products were purified using a QIAquick PCR Purification Kit (QIAGEN, Melbourne, Vic., Australia), PureLink Kit (Invitrogen, Melbourne, Vic., Australia), or using ExoSAP-IT (Affymetrix, Santa Clara, CA, USA) or ExoSTAR (GE Healthcare, Freiburg, Germany). PCR products were quantified on agarose gels or using a NanoDrop 2000 UV-Vis spectrophotometer (Thermo-Scientific, Melbourne, Vic., Australia) and directly sequenced using the ABI Prism BigDye Terminator v3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Foster City, CA, USA). These reactions used the original amplification primers and sometimes the internal sequencing primers d and e (Taberlet *et al.* 1991) for *trnL*-F, and 5'*trnG*2G and 5'*trnG*2S (Shaw *et al.* 2007) for *trnS*-G. Sequences were analysed on an ABI 3730x1 96-capillary automated DNA sequencer (Applied Biosystems) at the Australian Genome Research Facility, Brisbane or Melbourne.

Sequence editing and alignment

Contiguous sequences for each region were assembled and edited using Sequencher v.4.8 (Gene Codes Corporation, Ann Arbor, MI, USA), and aligned manually using Se-Al Sequence Alignment Editor v. 2.0a11 (Rambaut 2002). Insertion–deletion events (indels) were coded for parsimony analysis using a method equivalent to the ‘simple coding’ of Simmons and Ochoterena (2000), with a single character representing each indel whether single- or multi-base. The full data matrix, including indel characters, is deposited in TreeBase (<http://treebase.org/>; accession number S17027).

Phylogenetic analyses

Combined cpDNA sequences were analysed using maximum parsimony (MP) with PAUP* 4.0 β 10 (Swofford 2001) and Bayesian inference (BI) with MrBayes v.3.1.2 (Ronquist and Huelsenbeck 2003). MP analyses were performed using heuristic tree searches, a CLOSEST addition sequence and tree bisection and reconnection (TBR) branch swapping, all characters being equally weighted and gaps treated as missing data. All other settings were left on default.

When initial phylogenetic analysis yielded many thousands of equally parsimonious trees, an expeditious search strategy was employed. An initial heuristic search was commenced using the above settings, and terminated when 50 000 trees were obtained, then a strict consensus of equally parsimonious trees was calculated. This consensus tree was loaded as a constraint in a second analysis that searched only for trees that were as short as or shorter than the shortest tree of the first analysis, but were inconsistent with the constraint tree. This search used 1000 random addition sequences, each followed by TBR branch swapping. Each replicate was aborted when 2000 trees with lengths exceeding that of the first analysis were obtained. When this strategy was used, if trees of equal length to that of the first analysis were obtained (and inconsistent with the consensus from the first analysis), a further consensus was computed including the topology from the new analysis. The

second step of the analysis was repeated until no more new equally short but inconsistent trees were found. The aim of this strategy was to produce a strict consensus tree that was likely to represent the full set of equally most parsimonious trees. Bootstrap support (Felsenstein 1985) was estimated using a full heuristic search with 1000 bootstrap replicates and the search settings as described above. A maximum of 2000 trees was saved per replicate.

The models for BI analyses were GTR+R for *rpl32-trnL* and *trnL*-F, GTR+I for *trnQ-rps16* and GTR+I+G for *trnS*-G, selected using the Akaike information criterion (AIC) as implemented in MrModeltest 2.3 (Nylander 2004). Indels were not separately coded for BI analyses because preliminary analyses indicated that the level of resolution and topology in BI trees was similar to that using MP. Also indels were generally associated with branches of high posterior probability, suggesting that separate coding would add little in the way of extra support or resolution. BI analyses used the default settings of MrBayes, with parameters unlinked between partitions, and each included two runs of four chains, each run for 3 000 000 generations. Trees were sampled every 500 generations and a majority rule consensus was computed, with trees from the first 300 000 generations discarded as burn-in. Tracer v.1.5 (Rambaut 2007) was used to check that the burn-in period was adequate for each analysis and that the runs had converged on a stationary distribution. This was judged by comparing the distribution likelihood values, and the standard deviation of split frequencies, which were <0.01 at the end of the runs.

Results

Overview

Data from the combined chloroplast markers included a total of 4001 characters, of which 581 were variable and 230 were parsimony informative, including 55 parsimony-informative indel characters (statistics from the MP analysis are presented in Table 1). The *rpl32-trnL* and *trnS*-G regions together provided most of the parsimony-informative characters (74%).

The majority-rule consensus tree from the BI analysis was similar to, but slightly better resolved than, the MP strict consensus tree. The two measures of support, posterior probability (PP) for BI and bootstrap support (BS) for MP, were not always well correlated, with some high Bayesian PPs occurring when the corresponding BS values were low. However, there were no conflicting nodes between analyses, and only the Bayesian consensus tree together with both PP and BS values is presented here (Fig. 2).

Letters A–N (Fig. 2) label clades of three or more taxa that were well supported and are discussed in the text (as distinct from pairs of taxa or poorly supported nodes). Australian *Zieria* formed a strongly supported clade (Node 1, PP 1, BS 100%) with the New Caledonian species, *Z. chevalieri*, shown as its sister.

Within Australian *Zieria*, two large Clades A (Node 3) and B (Node 4), a smaller Clade C (Node 6), and a clade of the two subspecies of *Z. veronicea* (Node 5) were each strongly supported in the BI analysis, with PP values of 1. However, the trichotomous Node 2 was poorly supported (PP 0.87, BS 53%), and, therefore, it is not clear whether Clade C and *Z. veronicea* are more closely related to taxa in Clade A or Clade B, or to each other.

Table 1. Details of chloroplast datasets and results from maximum parsimony analyses
Tree statistics are based on equally most parsimonious trees produced by analyses of the combined dataset

Parameter	<i>rpl32-trnL</i>	<i>trnL-F</i>	<i>trnQ-rps16</i>	<i>trnS-G</i>	Combined data
Included characters	975	871	666	1489	4001
Variable characters	186	106	92	197	581
Parsimony-informative characters, total	88	27	32	83	230
Parsimony-informative indel characters	17	8	7	23	55
Parsimony-informative characters as a percentage of variable characters (%)	47	26	35	42	40
Number of steps on shortest trees	253	142	109	296	800
Consistency index (CI)	0.85	0.82	0.89	0.78	0.83
Retention index (RI)	0.93	0.87	0.96	0.90	0.92
Rescaled consistency index (RC)	0.79	0.71	0.86	0.702	0.76
Homoplasy index (HI)	0.15	0.18	0.11	0.22	0.18

Although the large Clade A was well supported in the BI analysis (PP 1) it was poorly supported in the MP analysis (BS 54%). It comprises five subclades, D, E, F, G and H, all strongly supported by BI (PP of 1 for all), but variously well to poorly supported by MP BS. *Zieria tenuis*, *Z. madida*, *Z. littoralis*, *Z. cephalophila* and *Z. boolbunda* formed part of a basal polytomy in Clade A. *Zieria bauerlenii* and *Z. collina* (MJB2012) were strongly supported as sister (although a second accession of *Z. collina* was resolved in a different well supported clade).

The other large Clade, B, which was strongly supported in both analyses (PP 1, BS 100%), was not well resolved, consisting of a polytomy of seven lineages (Node 4, Fig. 2). Node 41 within Clade B had strong support from the BI analysis (PP 1) but not in the MP analysis where there was no BS support. The relationship of *Z. caducibracteata* (IRT9585) and *Z. smithii* (DJO430) as sister taxa was only poorly supported. *Zieria alata* (KM639) and *Z. arborescens* (MFD2035) were resolved as sister. Samples whose relationships remained unresolved within Clade B are *Z. robusta* (MFD2049), *Z. obcordata* (608324), *Z. citriodora* (the only accession) and *Z. pilosa* (2 accessions). Clade B was further resolved into six subclades, I–N, all being well supported by PP (0.98–1.0), but with variable BS support (53–97%). Other relationships within these clades are discussed in more detail below.

In summary, there was a poor correlation between the cpDNA tree and clades identified in the morphological phylogeny of Armstrong (2002). None of Armstrong's morphological clades was identified as monophyletic and most cpDNA clades contained members of multiple morphological groups (Figs S1 and S2, available as Supplementary material for this paper). The clades C–N resolved here in the cpDNA analysis broadly overlap geographically in eastern Australia, although they include some marked taxon disjunctions. Clade K was the most geographically distinct, being restricted to Victoria and Tasmania.

Monophyly of species and sharing of chloroplast haplotypes between species

Of the species represented by at least two accessions in this analysis, the following were resolved as monophyletic: *Z. veronicea* (Node 5, two subspecies), *Z. laevigata* (Node 8), *Z. ingramii* (Node 9), *Z. littoralis* (Node 20), *Z. bauerlenii* (Node 28), *Z. pilosa* (Node 39) and *Z. involucreata* (Node 50).

Monophyly of the following species was neither supported nor rejected, with accessions being part of unresolved polytomies that included other species: *Z. lasiocaulis* (Node 32), *Z. prostrata* (Node 53) and *Z. odorifera* (Node 15; including the three sampled subspecies). All other species with multiple accessions were resolved as polyphyletic or paraphyletic. These were *Z. fraseri* (two samples, one from each subspecies, nested in Clade D), *Z. cytisoides* (three samples, in different subclades of Clade A), *Z. obcordata* (two samples from disjunct populations, occurring in different parts of the tree, one in Clade A, one in Clade B), *Z. smithii* (22 samples, some in Clade A, some in Clade B), *Z. collina* (two samples, in different subclades of Clade A), *Z. furfuracea* (three samples, representing three subspecies, one in Clade L, one in Clade M and one in Clade A), *Z. arborescens* (nine samples, in different subclades of Clade B), *Z. oreocena* (four samples, in two subclades within Clade B), *Z. alata* (two samples, in different subclades within Clade B), *Z. caducibracteata* (two samples, one in a well supported clade sister to a sample of *Z. arborescens*, the other weakly supported as sister to a sample of *Z. smithii*), *Z. minutiflora* (a sample of subsp. *trichocarpa* strongly supported as sister to a polytomy including *Z. vagans*, *Z. bifida* and two samples of subsp. *minutiflora*; Clade G), *Z. distans* (two samples in Clade H), *Z. robusta* (one sample in Clade K, one in the basal polytomy of Clade B), and *Z. murphyi* (two samples in Clade M, separated with moderate support).

On the basis of cpDNA markers, widespread species, or species for which samples were collected from highly separated locations, were commonly not found to be monophyletic. However, even some species with narrow distributions or with samples collected from relatively small geographic areas had haplotypes that placed them in different clades. For example, the two specimens of *Z. alata*, a species endemic to the Wet Tropics of Queensland, were collected within 15 km of each other (Clade B), the four specimens of *Z. oreocena* (Clade B) were collected within 50 km of each other in the Grampians in western Victoria, where the species is endemic, and the two samples of *Z. caducibracteata* were collected ~4 km apart in the Budawang Range of New South Wales.

Most chloroplast haplotypes were unique to individual specimens or species. However, the following three haplotypes were shared between specimens of different species: one between *Z. prostrata* and *Z. arborescens* (from samples collected nearly 200 km apart; Clade M, Fig. 5); one between *Z. lasiocaulis* and *Z. southwellii* (across samples collected nearly 200 km apart;

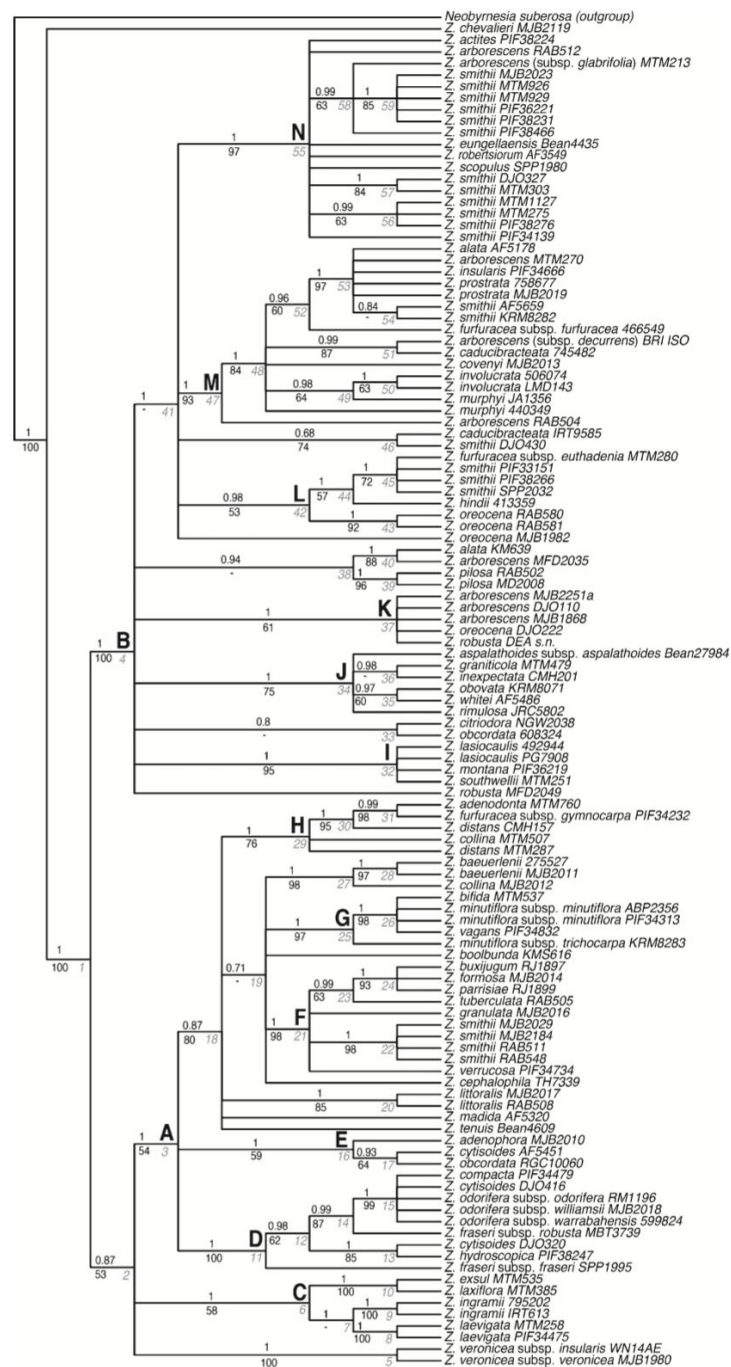


Fig. 2. Bayesian majority-rule consensus tree of combined chloroplast data, with Bayesian posterior probabilities (PP) shown above branches, bootstrap support (BS) values for maximum parsimony analysis below branches, and nodes numbered (in grey). Further details of Clades C–H and I–N are given in Figs 3 and 5 respectively.

Clade I, Fig. 5); and one between samples of *Z. odorifera*, *Z. compacta* and *Z. cytisoides* (across samples collected ~650 km apart; Clade D, Fig. 3).

Composition and distributions of Clades C–H

A portion of the Bayesian tree (Fig. 3) shows branch lengths and support values of Clade A and its subclades, and Clade C. Maps show the location of taxa within those clades (Fig. 4). Together, branch lengths and maps allow visualisation of the degree to which the chloroplast haplotypes are shared, relative divergences, and the provenance of the samples in the clades.

Clade C (PP 1, BS 58%) included *Z. exsul* and *Z. laxiflora* (both from coastal southern Queensland), *Z. ingramii* (inland New South Wales), and *Z. laevigata* (near the New South Wales–Queensland border). Both accessions of *Z. ingramii*, which were from Goonoo State Forest, ~15 km apart, had identical haplotypes. In contrast, in the more common species *Z. laevigata*, haplotypes were resolved as related (Fig. 3) but divergent (differing by 13 steps), despite accessions being collected from populations only ~30 km apart in southern Queensland.

In general, the haplotypes in Clade D were from specimens of taxa that typically occur in more inland locations than do most other *Zieria* taxa, on the western slopes or west of the Great Dividing Range. The accession of *Z. fraseri* subsp. *fraseri* was collected closest to the coast, and was well supported as sister to the other (more inland) specimens in the clade. Both subspecies of *Z. fraseri* were in this clade, but were not resolved as sister. Accessions of the three subspecies sampled of *Z. odorifera* (of four) were in a well supported subclade, with highly similar but not identical haplotypes. *Zieria odorifera* subsp. *williamsii*, *Z. compacta* and one sample of *Z. cytisoides* (DJO416) had identical haplotypes that were highly similar to the haplotypes of the two other *Z. odorifera* subspecies, which differed only by one change each. The samples sharing the identical haplotype were from localities up to 680 km apart. Another sample of *Z. cytisoides* (DJO320) was strongly supported as the sister to *Z. hydroscopica*, although their haplotypes differed by seven changes in the MP analysis. The two divergent samples of *Z. cytisoides* in this clade (differing by 16 steps in the MP analysis) were collected from localities more than 270 km apart in southern Queensland.

Clade E included three samples, one each from *Z. cytisoides*, *Z. obcordata* and *Z. adenophora*. Although it was highly supported by BI, Clade E had poor support in the MP analysis. Branch lengths were long, indicating deep divergences between the three samples, which were remotely separated geographically, from north Queensland to southern New South Wales. Of the three species in Clade E, *Z. cytisoides* is widespread and morphologically very different from the other two species; its taxonomic relationship may be better reflected by the position of two other accessions of *Z. cytisoides* that were resolved in Clade D. In contrast, *Z. obcordata* and *Z. adenophora* of Clade E are rare species that share some morphological affinities. *Zieria adenophora* is known only from a single population, whereas *Z. obcordata* has a disjunct distribution. The accession of *Z. obcordata* (RGC10060) in Clade E was from the more northern population, whereas an accession from a population

~100 km to the south (608324) was placed in major Clade B (Figs 2, 5), with strong support.

Clade F included a group of morphologically very similar species, namely, *Z. verrucosa*, *Z. granulata*, *Z. tuberculata*, *Z. parrisiae*, *Z. formosa* and *Z. buxijugum*, plus four samples of *Z. smithii* from Victoria that had identical or highly similar haplotypes. All of the other accessions of *Z. smithii* in this analysis were resolved in Clade B and were from Queensland or northern New South Wales. The morphologically similar species in this clade occur in southern New South Wales, except *Z. verrucosa*, which occurs only in Queensland, ~1000 km away. The three most closely related haplotypes of that group, representing *Z. buxijugum*, *Z. formosa* and *Z. parrisiae*, were from samples located within ~6 km of each other, but they were not identical.

All accessions in Clade G were from Queensland. *Zieria minutiflora* subsp. *trichocarpa* from northern Queensland was strongly supported as the sister to samples of *Z. bifida*, *Z. minutiflora* subsp. *minutiflora* and *Z. vagans*, which were all from southern Queensland. The haplotypes of the three taxa from southern Queensland were highly similar but not identical; *Z. bifida* and *Z. vagans* are very similar in morphology but different from *Z. minutiflora*.

All accessions in Clade H were also from Queensland. This clade was well supported by BI (PP 1), but less so by MP analysis (BS 76%). It included a sample of *Z. furfuracea* subsp. *gymnocarpa*, which was separated from other subspecies in *Z. furfuracea* in clades L and M. The relationship of *Z. furfuracea* subsp. *gymnocarpa* to *Z. distans* (CMH157) and *Z. adenodonta* was strongly supported (PP 1, BS 95%). However, *Z. distans* included two different haplotypes that did not form a monophyletic group within Clade H. Similarly, *Z. collina* appeared to be polyphyletic with one accession (MTM507) in Clade H and the other haplotype (MJB2012) outside that clade, related to *Z. baeuerlenii* (Node 27, Fig. 2), despite the two accessions of *Z. collina* being collected from populations less than 6 km apart.

Composition and distributions of Clades I–N

A portion of the Bayesian tree (Fig. 5) shows branch lengths and node-support values for Clade B and its subclades, I–N. Maps show the location of taxa within these clades (Fig. 6).

Clade I consisted of three morphologically similar species. *Zieria southwellii* and *Z. montana* are restricted to the area around the New South Wales–Queensland border, and the endangered *Z. lasiocaulis* is restricted to the headwaters of the Wilson River in the North Coast District of New South Wales. Two accessions of *Z. lasiocaulis*, which were collected from populations less than 3 km apart, had slightly different sequences; however, the sequence of *Z. lasiocaulis* (492944) was identical to that of *Z. southwellii*, from more than 300 km away.

Five species in Clade J are narrow endemics, three occurring in northern Queensland (*Z. whitei*, *Z. rimulosa* and *Z. obovata*) and two in southern Queensland (*Z. graniticola* and *Z. inexpectata*). The sixth sample in the clade is *Z. aspalathoides* subsp. *aspalathoides*, a widespread and variable taxon; the accession here was from southern Queensland. BI analysis resolved *Z. graniticola* and *Z. inexpectata* samples as sisters, and *Z. obovata* and *Z. whitei* samples as sisters, both with

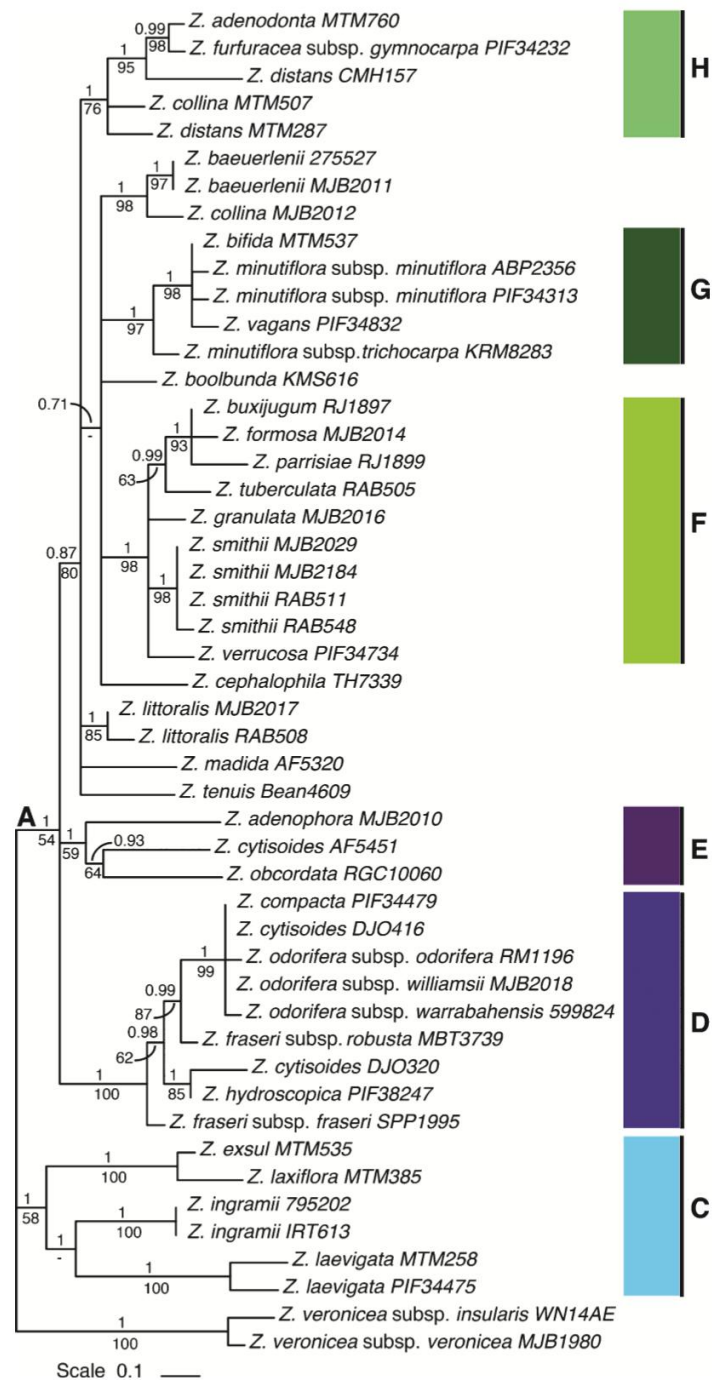


Fig. 3. A section of the Bayesian consensus tree from the combined chloroplast analysis, showing Clades C–H, mean branch lengths and support values. Bayesian inference posterior probability (PP) values are given above branches, maximum parsimony bootstrap (BS) percentages are below branches.

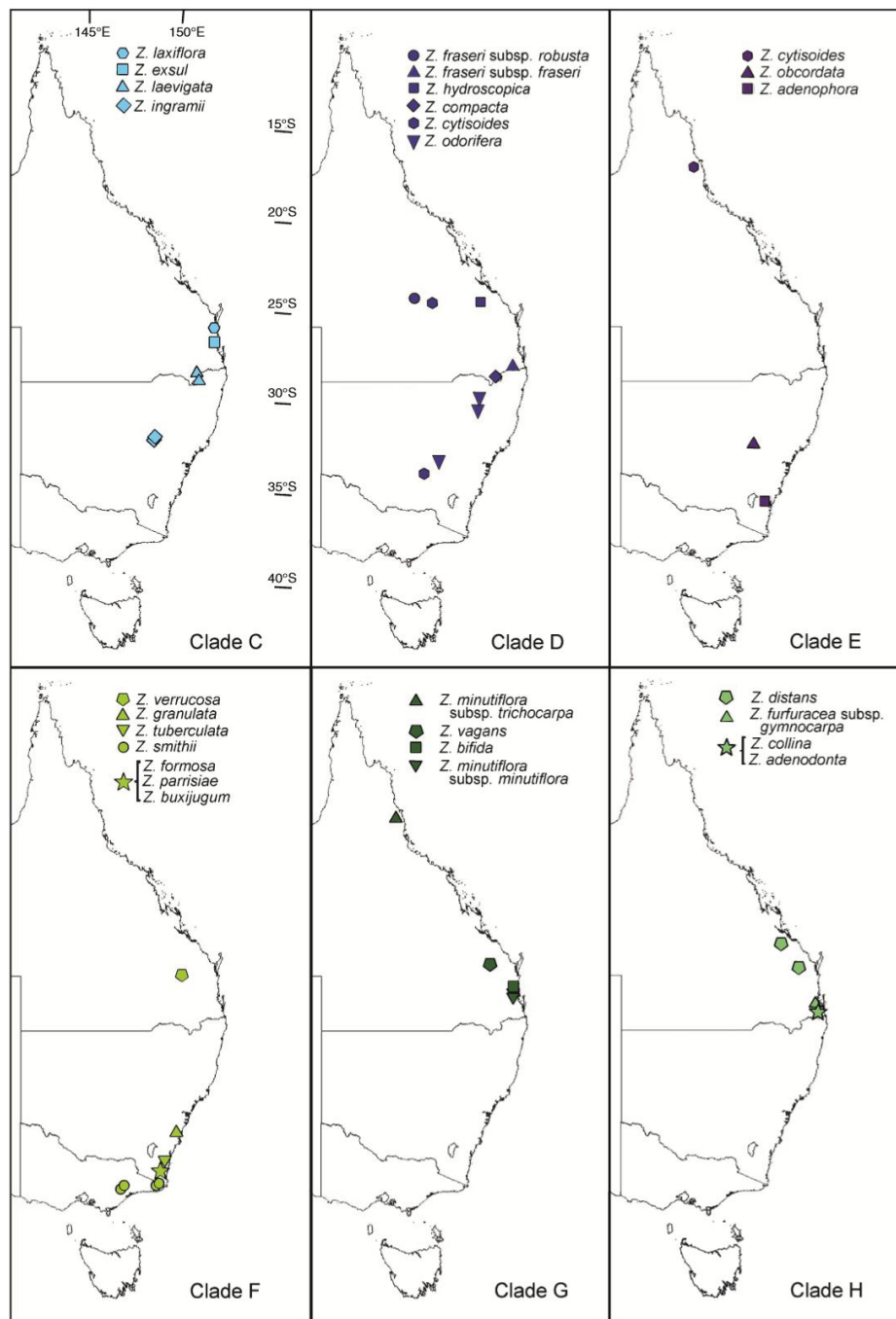


Fig. 4. Maps showing the distribution of samples in Clades C–H.

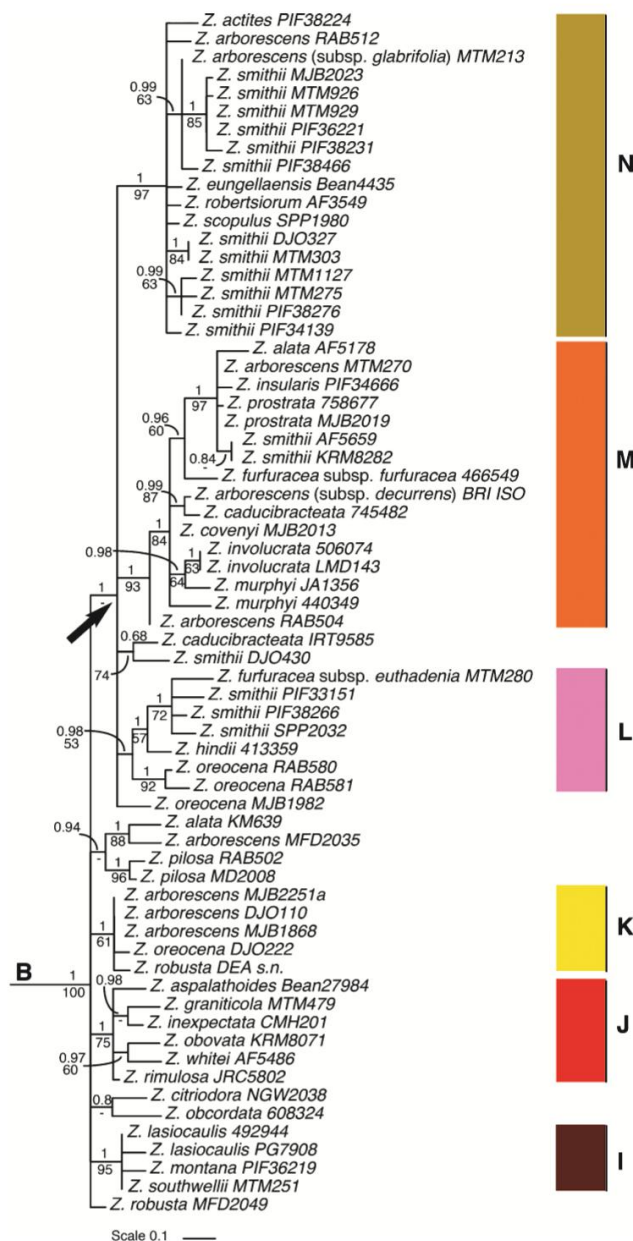


Fig. 5. A section of the Bayesian consensus tree from the combined chloroplast analysis, showing Clades I–N, mean branch lengths and support values. Posterior probability values are given above branches, maximum parsimony bootstrap percentages are below branches.

moderate to strong Bayesian support (respectively PP 0.98 and 0.97), but these relationships were poorly or not at all supported by MP analysis. *Zieria inexpectata*, *Z. graniticola* and *Z. whitei*

(but not *Z. obovata*) have strong morphological affinities with *Z. aspalathoides*. *Zieria whitei*, for instance, was initially described as a variety of *Z. aspalathoides* (var. *intermedia*;

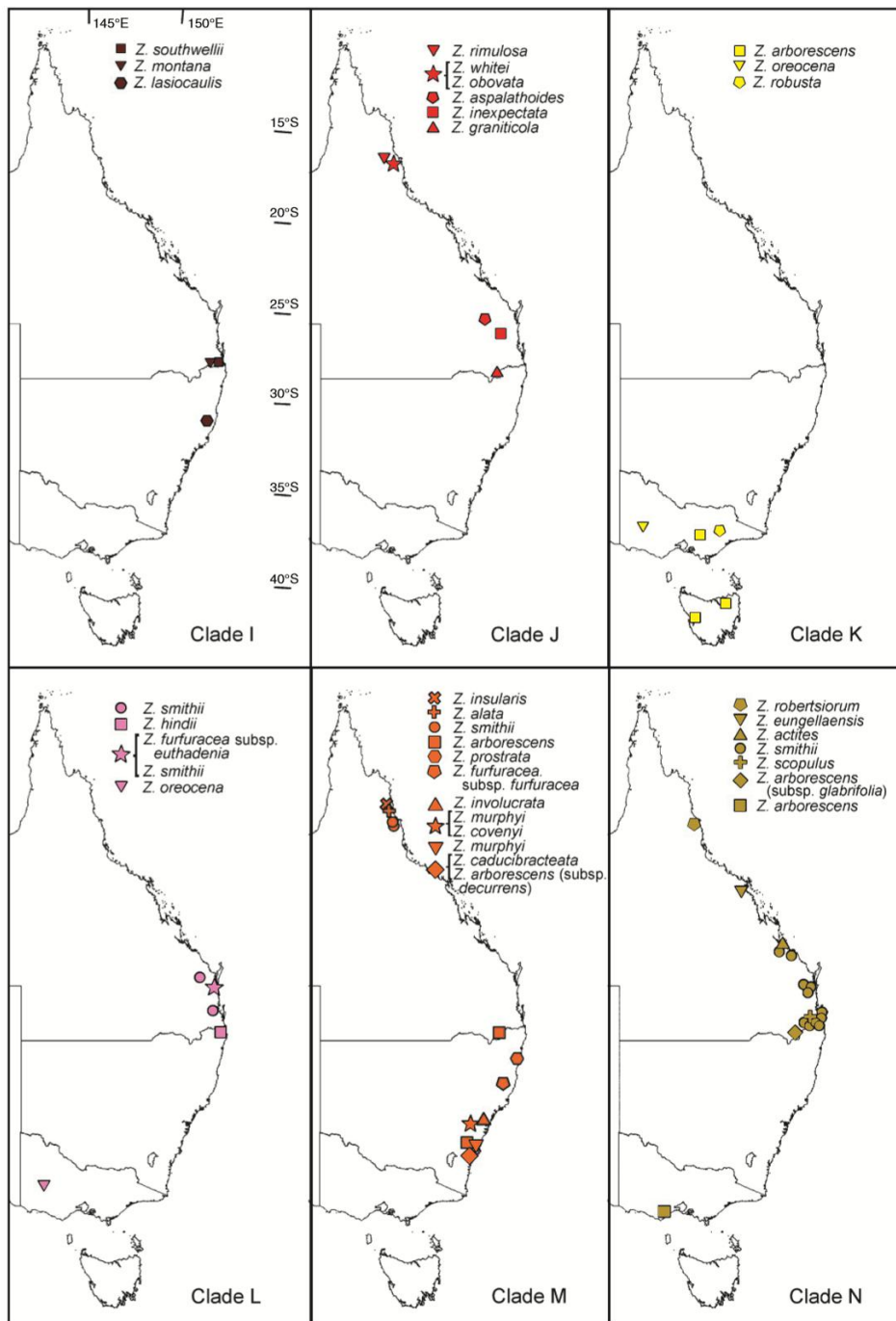


Fig. 6. Maps showing the distribution of samples in Clades I–N.

White 1942). *Zieria obovata* is morphologically very similar to *Z. minutiflora* (not in this clade), and the two species are easily confused. *Zieria rimulosa*, from Mount Mulligan, northern Queensland, is morphologically distinct, although it does have some characters in common with *Z. aspalathoides* and *Z. obovata* (northern Queensland), such as densely pubescent branchlets, and upper lamina leaf surfaces covered exclusively by simple (vs stellate) hairs.

Clade K (PP 1 but BS low at 61%) comprised Victorian and Tasmanian accessions of *Z. arborescens* (identical haplotypes), one accession of *Z. oreocena*, and one accession of *Z. robusta*, a morphologically distinctive species in the genus, very different from *Z. arborescens* and *Z. oreocena*. *Zieria oreocena* is morphologically very similar to *Z. arborescens* but geographically separated, being restricted to the Grampians in western Victoria. *Zieria robusta* has a disjunct distribution, and two accessions were sequenced. The accession in Clade K was from eastern Victoria; the other accession, from nearly 500 km to the north in New South Wales, also fell in Clade B, but at the basal node outside Clade K. All accessions of *Z. arborescens* from outside Victoria and Tasmania were placed elsewhere in Clade B. Surprisingly, *Z. arborescens* (RAB512) from the Otways, Victoria, was also not in this clade. Of the four accessions of *Z. oreocena*, only one occurred in Clade K.

Clade L contained two accessions of *Z. oreocena*, collected from close but distinct habitats at different altitudes in the northern Grampians, Victoria (Fig. 6). One population was in a sheltered valley beside a creek and the other was in a sheltered rocky outcrop near the summit of Hollow Mountain. The accessions were strongly supported as sister to a subclade of *Z. smithii* (three accessions) and *Z. furfuracea* subsp. *euthadenia*, all from southern Queensland, and *Z. hindii*, from northern New South Wales, some 1500 km further north of the Grampians. A third accession of *Z. oreocena* (MJB1982), collected ~40 km to the south, was on the basal node of Clade B, and a fourth accession (DJO222) collected less than 4 km from MJB1982, was resolved in Clade K (PP 1, but low BS of 61%). Thus, phylogenetic relationships here supported by the BI analysis were not necessarily correlated with geographic distance.

Clade M consisted of 11 taxa, including species from a wide geographical range, from the wet tropics in Queensland to the south-east of New South Wales, but with a clear disjunction in the mid- and southern parts of Queensland. Some, but not all, taxa within the clade share morphological affinities. Two accessions of the widespread *Z. arborescens* occurred in this clade (one from Queensland and one from New South Wales), as well as an accession from near the type locality of the highly restricted *Z. arborescens* subsp. *decurrens*, and one accession of *Z. caducibracteata*, a species that is morphologically similar to *Z. arborescens*. *Zieria arborescens* (MTM270) and *Z. prostrata* (MJB2019) had identical haplotypes despite being from locations nearly 200 km apart. Clade M also included the morphologically distinctive species, *Z. involucrata* and *Z. murphyi*, as well as *Z. covenyi*, a putative triploid hybrid (Armstrong 2002). *Zieria insularis* and *Z. smithii* have morphological affinities, although *Z. smithii* is highly variable. The seven samples (representing five species) that were resolved in a well supported subclade of Clade M (PP 1 BS 97%, Node 53, Fig. 2) were all collected in northern New South Wales or Queensland, with *Z. furfuracea*

subsp. *furfuracea* (466549), from further south, moderately supported as the sister. All other accessions of Clade M were from locations even further south, in southern New South Wales.

Clade N consisted of six taxa, including several accessions of *Z. smithii*, all collected from Queensland (but not all accessions of *Z. smithii* from Queensland were in this clade). The clade also included the sample of *Z. arborescens* (subsp. *glabrifolia*), from southern Queensland. Interestingly, one accession of *Z. arborescens* (RAB512), from the Otways, Victoria, was nested in this clade, indicating that this haplotype is more closely related to those of other taxa occurring in Queensland, than to haplotypes of its own species occurring elsewhere in Victoria; this was the only sample from outside Queensland that occurred in this clade.

Discussion

This study presents the first molecular phylogenetic study of *Zieria*, including almost all recognised species and subspecies, based on cpDNA. Interpretation of relationships is complicated because of widespread discordance between the cpDNA gene tree, existing species-level classification and the morphological groups proposed by Armstrong (2002). It is now widely recognised that species trees are not always congruent with gene trees because evolutionary processes often produce discrepancies between gene trees and the phylogenies of the organisms within which the genes evolved (Doyle 1992; Maddison 1997; Mallo *et al.* 2014). We conclude that the cpDNA phylogeny provides useful insight into evolutionary processes in *Zieria* (discussed below) but does not provide a clear basis for assessing phylogenetic relationships of taxa, speciation or biogeographic patterns within Australia, which were among our original aims.

Incongruence between cpDNA phylogeny and species taxonomy

The cpDNA phylogeny showed widespread incongruence with current species taxonomy in *Zieria*. There is a range of possible explanations for such incongruence, including the following: (1) homoplasy in the cpDNA sequence data, i.e. parallel evolution of similar sequences in different taxa; (2) inappropriate taxonomy, confounded by homoplasy of morphological characters; (3) introgression between species (chloroplast capture); and (4) variable retention of ancestral chloroplast lineages across different taxa (incomplete lineage sorting). Homoplasy of molecular characters becomes a less likely explanation as the number of characters supporting different clades of sequences increases (McKinnon *et al.* 1999; Holmes *et al.* 2014), and it is a poor explanation for many of the well supported clades in our dataset. The other three explanations are each more plausible and it is likely that a combination of these factors is at play in *Zieria*.

Introgression and incomplete chloroplast lineage sorting can be expected to leave different genetic and geographic signatures. With introgression, there is a chance that more derived haplotypes (near the tips of trees or networks) will be shared between species, and that interspecific sharing of chloroplast haplotypes or groups will be geographically concentrated, reflecting the area in which introgression has occurred (McKinnon *et al.* 1999; Muir and Schlöterer 2005; Nevill *et al.* 2014). With chloroplast lineage

sorting, there is less chance that tip haplotypes will be shared between species (i.e. there is bias towards sharing of ancestral haplotypes), and interspecific sharing of chloroplast haplotypes or groups need not be geographically concentrated (Schaal *et al.* 1998; Muir and Schlötterer 2005). In the current study, there are few haplotypes shared among different species that might provide evidence for recent introgression between species (discussed below). However, this could be an artefact of low and uneven sampling within species (despite including multiple representatives of most). It could also reflect dilution over time of signals of chloroplast lineage sorting or introgression, e.g. patterns of haplotype distribution could be affected by past range expansions or contractions that might have been considerable (e.g. Mellick *et al.* 2012; Rossetto *et al.* 2012) across multiple glacial or interglacial cycles.

The greatest insight into the processes responsible for the taxonomic incongruence of cpDNA sequences in *Zieria* potentially comes from the two species in our dataset that were the most highly sampled, among the most geographically widespread, and had clearly polyphyletic cpDNA sequences. These two species are *Z. smithii*, which occurs from northern Queensland to eastern Victoria, and *Z. arborescens*, which occurs from southern Queensland to Tasmania and includes three accepted subspecies, although the distinction between the subspecies is not clear cut (see note in Materials and methods).

The cpDNA sequences of *Zieria smithii* show clear geographic patterns. Both samples from northern Queensland group with other taxa from that area in Clade M, all samples from eastern Victoria and southern New South Wales group with other taxa from southern New South Wales in Clade F, and samples from south-eastern Queensland fall in two clades (N and L) with other taxa from that area. This geographic pattern of chloroplast variation and similarity to other taxa would be consistent with regional introgression of chloroplasts. However, it could also be consistent with poor taxonomy, i.e. that *Z. smithii* encompasses multiple, cryptic taxa, yet to be recognised, that are each related to different species. The RAPD data of Hogbin and Crisp (2003), for instance, suggested that the coastal species *Z. prostrata* was incompletely differentiated from nearby populations of *Z. smithii* (although this was not recovered in our cpDNA sampling where one sample of *Z. prostrata* had an identical haplotype to *Z. arborescens* from Mount Norman, almost 200 km away). The considerable morphological variation seen in *Z. smithii*, including the northern Queensland populations previously recognised as subsp. *tomentosa*, was discussed by Dureto and Forster (2007).

Zieria arborescens displays some cpDNA variation that is geographically correlated, but also variation that suggests an influence of chloroplast lineage sorting. Of particular note is the placement of the sample of *Z. arborescens* from the Otway Ranges in south-western Victoria (RAB512) in Clade N, in which all other samples are from Queensland, separated by more than 1200 km. This geographic separation, although it could be reduced with further geographic sampling, is well beyond the realms of plausibility for recent chloroplast introgression, especially in an entomophilous genus that has seeds of limited dispersal (Armstrong 1979, 2002). Placement of the Otway sample in Clade N is well supported (BS 97%, PP 1, and with six characters supporting the node in parsimony reconstructions),

also making homoplasy unlikely. As such, this example provides the best evidence in our dataset that morphologically defined species of *Zieria* could harbour highly divergent chloroplast lineages that pre-date speciation events. In other cases where similar haplotypes in different species are widely geographically separated (e.g. those in clade B of subsp. *arborescens* MFD2035 from New South Wales and *Z. alata* KM639 from north Queensland), the inferred relationships have lower bootstrap or character support.

The patterns of variation in *Z. smithii* and *Z. arborescens* provide good evidence that introgression, lineage sorting and inappropriate taxonomy all potentially account for the observed cpDNA patterns. In other non-monophyletic taxa that are represented by fewer samples, or where divergence of chloroplast haplotypes is not so strong, it is difficult to make firm conclusions. For instance, in the case of *Z. robusta*, the two samples represent disjunct populations in Victoria and New South Wales, and it is not clear whether divergence between them could reflect introgression (e.g. similarity of the Victorian sample of *Z. robusta* to Victorian samples of *Z. arborescens* in Clade K), lineage sorting, or geographic variation associated with morphological differentiation of disjunct populations, e.g. Victorian populations have leaflets that are more deeply crenate, and petioles and undersurfaces of primary veins on leaflets that are more prominently tuberculate. Likewise for *Z. furfuracea*, it is not clear whether the three highly divergent haplotypes in Clades L, M and H, each from a different subspecies, reflect introgression with other species (e.g. they each group with different species from the regions in which they were sampled), variable retention of deeply divergent chloroplast lineages, or that the species is not monophyletic and each of the three subspecies are separate evolutionary lineages. Similarly, contrasting explanations could be plausible for each of the other species in our analysis that are not resolved as monophyletic.

Relationship of Zieria in Eastern Australia and New Caledonia

The chloroplast phylogeny shows the deepest divergence in *Zieria*, with strong support (BS 100%, PP 1.0), to be that between the single New Caledonian species, *Z. chevalieri*, and Australian taxa. As with other taxa, the position of *Z. chevalieri* is inconsistent with the morphological assessment and phylogeny of Armstrong (2002) that grouped *Z. chevalieri* with *Z. laevigata*, *Z. laxiflora* and *Z. fraseri*, which are variously placed in Clades C and D in the cpDNA phylogeny.

An early divergence between plants of these two areas (as opposed to a more nested position of *Z. chevalieri*) potentially lends itself to a vicariant explanation for distribution of the genus. Such an explanation would be consistent with the very limited dispersal capacity of the ant-dispersed seeds and is interesting, given controversy over the history of New Caledonia and its flora (Lowry 1998), and debate over the importance of long-distance dispersal (Grandcolas *et al.* 2008; Swenson *et al.* 2014) versus vicariance (Bauer *et al.* 2006; Ladiges and Cantrill 2007; Heads 2008). Controversy stems from molecular dating of clades where it is concluded that clades are too young to support a vicariance explanation. Molecular dating within Rutaceae by Bayly *et al.*

(2013) estimated that the divergence between *Z. chevalieri* and Australian taxa occurred (6.5) 11.9 (17.9) million years ago. This is younger than the initial rifting of Zealandia from eastern Australia, which commenced ~85 million years ago (McLoughlin 2001), the re-emergence ~37 million years ago of most of present-day New Caledonia after being below sea level from the Late Cretaceous to the Eocene (Aitchison *et al.* 1995; Pelletier 2006), and the likely presence of substantial areas of exposed land between Australia and New Caledonia into the Eocene and Oligocene (De Beuque *et al.* 1998; Exon *et al.* 2006; Ladiges and Cantrill 2007; Sutherland *et al.* 2010; Bache *et al.* 2012). However, it would be unwise to emphasise the molecular dating as a basis for confirming or rejecting biogeographic hypotheses in *Zieria* because, as previously suggested (Bayly *et al.* 2013), large parts of the Rutaceae tree are not supported by good fossil calibrations, including in *Zieria*, and could be prone to substantial underestimates of clade ages. Furthermore, as mentioned above, over water, long-distance seed dispersal is considered highly improbable for *Zieria*.

Because the cpDNA of *Z. chevalieri* in New Caledonia is substantially divergent from that of Australian taxa, the species is significant in terms of the phylogenetic diversity of the genus, making it of conservation concern. It is known only from Mount Kaala in northern Grande Terre (New Caledonia). We have seen only a small number of plants in the field (fewer than 100 in May 2009), and Armstrong (2002) reported a total population of ~200 plants. *Zieria chevalieri* is currently listed as *Vulnerable* under the IUCN Red List of Threatened Species (<http://www.iucnredlist.org>, accessed November 2014) under Criteria B1 and 2C, based on the assessment of Jaffre *et al.* (1998). A population survey and revision of the conservation status of *Z. chevalieri* is warranted, given the high level of mining activity around Mount Kaala and our field observations in 2009 that the population had been affected by road building and recent fire.

Conclusions and prospects for further research

As discussed above, we propose that the cpDNA tree presented here does not overall represent a hypothesis for a species tree in *Zieria*. It largely reflects evolutionary processes other than those of phylogenetic descent, such as introgression between species and chloroplast lineage sorting, as well as inappropriate species-level taxonomy. Gaining further insight into the history and systematics of *Zieria* will require additional data, and priorities should be to obtain phylogenetic data from the nuclear genome, more carefully scrutinise the morphology of some species and more thoroughly sample some species.

Assessing the degree of congruence between phylogenies produced by nuclear and chloroplast markers could help ascertain the underlying causes of observed phylogenetic patterns. For instance, it would be valuable to know how taxa resolved here as polyphyletic might be resolved by nuclear markers. If they were shown to be monophyletic using nuclear markers, current taxonomy would be supported and chloroplast capture or lineage sorting implicated as explanations. A finding of non-monophyly would highlight a need for taxonomic revision.

Additional data on morphology would be useful to compare with genetic variation. This is especially true for widespread and morphologically variable species such as *Z. arborescens*,

Z. aspalathoides, *Z. cytisoides* and *Z. smithii*. *Zieria* contains a striking mixture of a few very widespread species, and a large number of narrowly endemic species. It is possible that widespread taxa could have complicated relationships with a range of endemics that have differentiated through peripheral isolation or that widespread species harbour geographic variants worthy of taxonomic recognition. Correlation of geographic patterns of morphological and genetic variation would allow a more critical assessment of species limits.

A strength of the present study was the inclusion of multiple representatives of many taxa (41% of included species, i.e. 24 of 59). This allowed for the detection of widespread cpDNA polyphyly; more limited sampling would have provided a poorer understanding of the complexity of cpDNA variation. Nonetheless, many taxa in *Zieria* were not well sampled relative to their geographic distributions and morphological variation (e.g. *Z. aspalathoides*, *Z. caducibracteata*, *Z. cytisoides* and *Z. odorifera*). No species represented by more than two samples in our dataset was shown to be monophyletic, although 51 species were represented here by only one or two samples. Further sampling would give a clearer picture of the complexities of cpDNA variation.

Acknowledgments

For assistance with field work or provision of specimens we thank Erin Batty, Gill Brown, Paul Carmen, Mike Mathieson, David Meagher, Tundra Morschke, Jérôme Munzinger, Daniel Ohlsen, Megan Thomas and Zoe Smith. Plant collecting permits were provided by the Victorian Department of Sustainability and Environment, NSW National Parks and Wildlife Service, Queensland Environmental Protection Agency, Royal Botanic Gardens Melbourne, Australian National Botanic Gardens, the Australian Department of the Environment, and conservation authorities of the North Province of New Caledonia (DDEE). We thank the directors of BRI, MEL and NSW for specimen loans and for permission to destructively sample material for molecular analysis. This study was funded by the Australian Research Council (LP0776737), a Bush Blitz PhD Research Supplement to R. Barrett (BBC210-24), grants from the Australian Biological Resources Study (PD208-02; RFL210-17), and was supported by the Alfred Nicholas Fellowship, The University of Melbourne School of Botany Foundation, the Australasian Systematic Botany Society (Hansjörg Eichler Scientific Research Fund Grant), The Soroptimist International of the South West Pacific (Dame Margaret Blackwood Soroptimist Scholarship) and an Australian Post Graduate Award.

References

- Aitchison JC, Clarke GL, Meffre S, Cluzel D (1995) Eocene arc-continent collision in New Caledonia and implications for regional southwest Pacific tectonic evolution. *Geology* **23**, 161–164. doi:10.1130/0091-7613(1995)023<0161:EACCIN>2.3.CO;2
- Armstrong JA (1979) Biotic pollination mechanisms in the Australian flora: a review. *New Zealand Journal of Botany* **17**, 467–508. doi:10.1080/0028825X.1979.10432565
- Armstrong JA (1991) Studies on pollination and systematics in the Australian Rutaceae. PhD Thesis, University of New South Wales, Sydney.
- Armstrong JA (2002) *Zieria* (Rutaceae): a systematic and evolutionary study. *Australian Systematic Botany* **15**, 277–463. doi:10.1071/SB00040
- Armstrong JA, Powell JM (1980) *Neobyrsesia* (Rutaceae): a new genus endemic to northern Australia. *Telopea* **1**, 399–408.
- Bache F, Sutherland R, Stagpoole V, Herzer R, Collot J, Rouillard P (2012) Stratigraphy of the southern Norfolk Ridge and the Reinga Basin: a record of initiation of Tonga–Kermadec–Northland subduction in the southwest

- Pacific. *Earth and Planetary Science Letters* **321–322**, 41–53. doi:10.1016/j.epsl.2011.12.041
- Bauer AM, Jackman T, Sadler RA, Whitaker AH (2006) A revision of the *Bavaya validiclavus* group (Squamata: Gekkota: Diplodactylidae), a clade of New Caledonian geckos exhibiting microendemism. *Proceedings of the California Academy of Science, Fourth Series* **57**, 503–547.
- Bayly MJ, Holmes GD, Forster PI, Cantrill DJ, Ladiges PY (2013) Major clades of Australasian Rutoideae (Rutaceae) based on *rbcL* and *atpB* Sequences. *PLoS ONE* **8**, e72493. doi:10.1371/journal.pone.0072493
- De Beuque SV, Auzende J-M, Lafoy Y, Missegue F (1998) Tectonique et volcanisme tertiaire sur la ride de Lord Howe (Sud-Ouest Pacifique). *Comptes Rendus de l'Académie des Sciences. Series IIA. Earth and Planetary Science* **326**, 663–669.
- Doyle JJ (1992) Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany* **17**, 144–163. doi:10.2307/2419070
- Duretto MF (1999) Rutaceae. In 'Flora of Victoria. Vol. 4'. (Eds N Walsh, T Entwistle) pp. 153–197. (Inkata Press: Melbourne)
- Duretto MF, Forster PI (2007) A taxonomic revision of the genus *Zieria* Sm. (Rutaceae) in Queensland. *Austrobaileya* **7**, 473–544.
- Duretto MF, Forster PI (2008) New subspecies for *Zieria odorifera* J.A. Armstr. (Rutaceae) from northern New South Wales. *Austrobaileya* **7**, 681–690.
- Engler A (1931) Rutaceae. In 'Die natürlichen Pflanzenfamilien. Teil 19a', 2nd edn. (Eds HGA Engler, K Prantl) pp. 187–359. (Wilhelm Engelmann: Leipzig, Germany)
- Exon NF, Hill PJ, Lafoy Y, Heine C, Bernardel G (2006) Kenn Plateau off northeast Australia: a continental fragment in the southwest Pacific jigsaw. *Australian Journal of Earth Sciences* **53**, 541–564. doi:10.1080/08120090600632300
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791. doi:10.2307/2408678
- George AS, Duretto MF, Forster PI (2013) *Zieria*. *Flora of Australia* **26**, 282–336.
- Grandcolas P, Murienne J, Robillard T, Desutter-Grandcolas L, Jourdan H, Guilbert E, Deharveng L (2008) New Caledonia: a very old Darwinian Island? *Philosophical Transactions of the Royal Society of London. B. Biological Sciences* **363**, 3309–3317. doi:10.1098/rstb.2008.0122
- Heads M (2008) Biological disjunction along the West Caledonian Fault, New Caledonia: a synthesis of molecular phylogenetics and pan biogeography. *Botanical Journal of the Linnean Society* **158**, 470–488. doi:10.1111/j.1095-8339.2008.00866.x
- Hogbin P, Crisp M (2003) Evolution of the coastal neospecies *Zieria prostrata* (Rutaceae) and its relationship to the *Zieria smithii* species complex. *Australian Systematic Botany* **16**, 515–525. doi:10.1071/SB02031
- Holmes GD, Downing TL, James EA, Blacket MJ, Hoffman AA, Bayly MJ (2014) Phylogeny of the holly grevilleas (Proteaceae) based on nuclear ribosomal and chloroplast DNA. *Australian Systematic Botany* **27**, 56–77.
- Jaffre T, Bouchet P, Veillon J-M (1998) Threatened plants of New Caledonia: is the system of protected areas adequate? *Biodiversity and Conservation* **7**, 109–135. doi:10.1023/A:1008815930865
- Ladiges PY, Cantrill D (2007) New Caledonia–Australian connections: biogeographic patterns and geology. *Australian Systematic Botany* **20**, 383–389. doi:10.1071/SB07018
- Lowry PP II (1998) Diversity, endemism, and extinction in the flora of New Caledonia: a review. In 'Rare, threatened, and endangered flora of Asia and the Pacific rim: proceedings of the international symposium on rare, threatened, and endangered floras of Asia and the Pacific rim. Academia Sinica Monograph Series 16'. (Eds C-I Peng, PP Lowry II) pp 181–206. (Institute of Botany, Academia Sinica: Taipei, Taiwan)
- Maddison WP (1997) Gene trees in species trees. *Systematic Biology* **46**, 523–536. doi:10.1093/sysbio/46.3.523
- Mallo D, Martins LDO, Posada D (2014) Unsorted homology within locus and species trees. *Systematic Biology* **63**, 988–992. doi:10.1093/sysbio/syu050
- McKinnon GE, Steane DA, Potts BM, Vaillancourt RE (1999) Incongruence between chloroplast and species phylogenies in *Eucalyptus* subgenus *Monocalyptus* (Myrtaceae). *American Journal of Botany* **86**, 1038–1046. doi:10.2307/2656621
- McLoughlin S (2001) The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany* **49**, 271–300. doi:10.1071/BT00023
- Mellick R, Lowe A, Allen C, Hill RS, Rossetto M (2012) Palaeodistribution modelling and genetic evidence highlight differential post-glacial range shifts of a rain forest conifer distributed across a latitudinal gradient. *Journal of Biogeography* **39**, 2292–2302. doi:10.1111/j.1365-2699.2012.02747.x
- Muir G, Schlötterer C (2005) Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Molecular Ecology* **14**, 549–561. doi:10.1111/j.1365-294X.2004.02418.x
- Nevill PG, Despres T, Bayly MJ, Bossinger G, Ades PK (2014) Shared phylogeographic patterns and widespread chloroplast haplotype sharing in *Eucalyptus* species with different ecological tolerances. *Tree Genetics & Genomes* **10**, 1079–1092. doi:10.1007/s11295-014-0744-y
- Nylander J (2004) 'MrModeltest v2.' (Program distributed by the author. Evolutionary Biology Centre: Uppsala University, Sweden)
- Pelletier B (2006) Geology of the New Caledonia region and its implications for the study of the New Caledonian biodiversity. Compendium of marine species from New Caledonia, Nouvelle-Calédonie. Documents Scientifiques et Techniques IRD, II 7. Institut de Recherche pour le Développement, Nouméa, New Caledonia. pp. 17–30.
- Rambaut A (2002) Se-AL: Sequence Alignment Editor version 2.0a11. Available at <http://tree.bio.ed.ac.uk/software/> [Verified April 2014]
- Rambaut A (2007) Tracer v1.5. Available at <http://tree.bio.ed.ac.uk/software/tracer/> [Verified April 2014]
- Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574. doi:10.1093/bioinformatics/btg180
- Rossetto R, Allen CB, Thurlby KAG, Weston PH, Milner ML (2012) Genetic structure and bio-climatic modeling support allopatric over parapatric speciation along a latitudinal gradient. *BMC Evolutionary Biology* **12**, 149. doi:10.1186/1471-2148-12-149
- Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology* **7**, 465–474. doi:10.1046/j.1365-294x.1998.00318.x
- Sharma IK (2001) Understanding clonal diversity patterns through allozyme polymorphism in an endangered and geographically restricted Australian shrub, *Zieria bauerlenii* and its implications for conservation. *Biochemical Systematics and Ecology* **29**, 681–695. doi:10.1016/S0305-1978(00)00113-7
- Sharma IK, Young AG (2001) Genetic diversity analysis among fragmented populations of *Zieria granulata* within the Tabbagong Forest and RSA land. Report prepared for the NSW Roads and Traffic Authority Southern Region, Wollongong, NSW.
- Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **92**, 142–166. doi:10.3732/ajb.92.1.142
- Shaw J, Lickey E, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* **94**, 275–288. doi:10.3732/ajb.94.3.275
- Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49**, 369–381. doi:10.1093/sysbio/49.2.369
- Smith JE (1978) The characters of twenty new genera of plants. *Transactions of the Linnean Society of London* **4**, 213–223. doi:10.1111/j.1096-3642.1798.tb00530.x

- Sutherland R, Collot J, Lafoy Y, Logan GA, Hackney R, Stagpoole V, Uruski C, Hashimoto T, Higgins K, Herzer RH, Wood R, Mortimer N, Rollet N (2010) Lithosphere delamination with foundering of lower crust and mantle caused permanent subsidence of New Caledonia Trough and transient uplift of Lord Howe Rise during Eocene and Oligocene initiation of Tonga–Kermadec subduction, western Pacific. *Tectonics* **29**, TC2004. doi:[10.1029/2009TC002476](https://doi.org/10.1029/2009TC002476)
- Swenson U, Nylander S, Munzinger J (2014) Sapotaceae biogeography supports New Caledonia being an old Darwinian Island. *Journal of Biogeography* **41**, 797–809. doi:[10.1111/jbi.12246](https://doi.org/10.1111/jbi.12246)
- Swofford DL (2001) 'PAUP*. Phylogenetic analysis using parsimony. Version 4.0b10.' (Illinois Natural History Survey, Smithsonian Institution: Champaign, IL)
- Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA *Plant Molecular Biology* **17**, 1105–1109. doi:[10.1007/BF00037152](https://doi.org/10.1007/BF00037152)
- White CT (1942) Contributions to the Queensland flora, number 7. *Proceedings of the Royal Society of Queensland* **53**, 201–228.

Note added in proof

In the time between acceptance and publication of this paper another phylogenetic study of *Zieria* has been published: Morton CM (2015) Phylogenetic relationships of *Zieria* (Rutaceae) inferred from chloroplast, nuclear, and morphological data. *PhytoKeys* **44**, 15–38.

The chloroplast DNA markers used in that paper were a subset of those used herein. Taxon sampling included 32 species, all of which are also represented here. Because species were each represented only by single accessions, that study did not detect the incongruence between cpDNA variation and species-level taxonomy reported here.

Appendix 1. Details of samples and sequences used in this study

Herbarium abbreviations follow Index Herbariorum. Other abbreviations are as follows: ABG, Australian Botanic Garden; ABP, A.B. Pollock; AF, A. Ford; ANBG, Australian National Botanic Gardens; Bean, A.R. Bean; CMH, C.M. Haskard; DEA, D.E. Albrecht; DJO, D.J. Ohlsen; IRT, I.R. Telford; JA, J. Armstrong; JRC, J.R. Clarkson; KM, K. Muscat; KMS, K.M. Sparshott; KRM, K.R. McDonald; LMD, L. McDougal; MBT, M.B. Thomas; MFD, M.F. Duretto; MJB, M.J. Bayly; MTM, M.T. Mathieson; NGW, N.G. Walsh; NP, National Park; NSW, New South Wales; NT, Northern Territory; PG, P. Gilmore; PIF, P.I. Forster; Qld, Queensland; RAB, R.A. Barrett; RGC, R.G. Coveny; RJ, R. Johnston; RM, R. Makinson; SA, South Australia; SF, State Forest; SPP, S.P. Phillips; Tas, Tasmania; Vic., Victoria; WN, W. Neal. Collections of *Zieria arborescens* are not identified to subspecies, but those from near the type locations of subsp. *decurrens* and subsp. *glabrifolia* are indicated (see 'Materials and methods' for details)

Taxon	Collection ID	Location	Voucher	<i>rp32-trnL</i>	<i>trnL-F</i>	<i>trnQ-rps16</i>	<i>trnS-G</i>
<i>Neobynnesia suberosa</i> J.A. Armstr.	MJB1904	NT, Kakadu NP, East Alligator area, Barbedjilidji Sandstone Walk, ~300 m SW of carpark, near base of sandstone cliff	MEL	KP177988	KP188925	KP188987	KP189143
<i>Z. actites</i> Duretto & P.I. Forst.	PIF38224	Qld, Mount Larcom summit area, 2.5 km WSW of Targinie	BRI AQ818047	KP177977	KP188845	KP188995	KP189076
<i>Z. adenodonta</i> (F. Muell.) J.A. Armstr.	MTM760	Qld, Lower Bellbird Circuit, Binna Burra, Lamington NP	BRI AQ756722	KP177997	KP188858	KP188992	KP189110
<i>Z. adenophora</i> (F. Muell.) J.A. Armstr.	MJB2010	NSW, cult. ANBG, ex Araluen ~500 m ESE of Bells Creek Falls on rocky escarpment (same stock as CBG 8805884)	MELU, CBG 8805884	KP178010	KP188886	KP188982	KP189114
<i>Z. alata</i> Duretto & P.I. Forst.	AF5178	Qld, Daintree NP, 'landing rock' headwaters of Roots Creek, SE of topo '1296', site 171	BRI AQ739620	KP178042	KP188874	KP189035	KP189180
<i>Z. alata</i> & P.I. Forst.	KM639	Qld, Mount Lewis NP	MELU	KP178064	KP188885	KP189029	KP189157
<i>Z. arborescens</i> Sims	DJO110	Tas., Saint Patricks River, on the S side of the Mount Barrow road, 10 km from its junction with the Tasman Highway	MELU D 105582	KP177976	KP188883	KP189048	KP189153
<i>Z. arborescens</i>	MFD2035 (AE892158)	NSW, cult. ABG, Mount Arman. Wild source: Blue Mountains NP, Walking track to Wollongambi River, near Mount Wilson	NSW	KP178001	KP188939	KP188986	KP189175
<i>Z. arborescens</i>	MJB1868	Vic., Noojee SF, Starling Gap picnic area	MELU	KP178079	KP188949	KP188969	KP189161
<i>Z. arborescens</i>	MJB2251a	Tas., Strahan	MELU	KP178039	KP188901	KP188965	KP189120
<i>Z. arborescens</i>	MTM270	Qld, Mount Norman, S side, Girraween NP	BRI AQ745993	KP178057	KP188854	KP188958	KP189078
<i>Z. arborescens</i>	RAB504	NSW, Barren Grounds Nature Reserve, next to the entrance gate	MELU	KP177979	KP188941	KP189054	KP189136
<i>Z. arborescens</i>	RAB512	Vic., Otway Ranges, Mount Sabine Benwerrin Road	MELU	KP177972	KP188944	KP189063	KP189112
<i>Z. arborescens</i> (subsp. <i>decurrens</i>) J.A. Armstr.	BRI ISO	NSW, 1.5 miles [2.4 km] along the Caves Beach Road, near Jervis Bay	BRI AQ695444, NSW 458804	KP178014	KP188894	KP189010	KP189144
<i>Z. arborescens</i> (subsp. <i>glabrifolia</i>) J.A. Armstr.	MTM213	Qld, Donnellys Castle, ~6.5 km W of Thulimbah off Donnellys Castle Road	BRI AQ745477	KP178052	KP188897	KP189021	KP189147
<i>Z. aspalathoides</i> A. Cunn. ex Benth. subsp. <i>aspalathoides</i>	Bean27984	Qld, near junction of Delubra and Cadarga creeks, 35 km SW of Mundubbera	BRI AQ820153	KP178045	KP188903	KP189026	KP189150

(continued next page)

Appendix 1. (continued)

Taxon	Collection ID	Location	Voucher	<i>rpl32-trnL</i>	<i>trnL-F</i>	<i>trnQ-rps16</i>	<i>trnS-G</i>
<i>Z. baeuerlenii</i> J.A. Armstr.	275527	NSW, Bomaderry Creek Walking Track, cliff top N of number '4'	NSW 275527	KP178069	KP188898	KP188994	KP189104
<i>Z. baeuerlenii</i>	MJB2011	NSW, cult. ANBG, ex Bomaderry Creek Reserve, 300 m S of Wier on W side, 10 m behind bench seat (same stock as CBG 9104885)	MELU	KP178048	KP188908	KP188984	KP189085
<i>Z. bifida</i> Duretto & P.I.Forst.	MTM537	Qld, Triunia NP, adjacent to Dulong Road W of Woombye	BRI AQ792099	KP177974	KP188923	KP188951	KP189158
<i>Z. boobunda</i> Duretto & P.I.Forst.	KMS616	Qld, Boolbunda Rock, ~9 km NE of Mount Perry township	BRI AQ640789	KP178024	KP188857	KP188966	KP189171
<i>Z. buxijugum</i> J.D.Briggs & J.A.Armstr.	RJ1897	NSW, Box Range Farm ~10 km W of Pambula	MEL 2312322	KP178055	KP188930	KP189065	KP189072
<i>Z. caducibracteata</i> J.A.Armstr.	745482	NSW, Budawangs Track to Curog [Corang] River from Wog camping ground	NSW 745482	KP177984	KP188876	KP188988	KP189159
<i>Z. caducibracteata</i>	IRT9585	NSW, northern Budawang Range, Quiliys Mountain, 1.5 km SE of Endrick trig	MEL 654142	KP177993	KP188850	KP189024	KP189092
<i>Z. cephalophila</i> Duretto & P.I.Forst.	TH7339	Qld, SE edge of Sydney Heads, Britton Range, 6.3 km NE-ENE of Homevale Homestead	BRI AQ579129	KP178035	KP188913	KP188979	KP189121
<i>Z. chevalieri</i> Viro	MJB2119	New Caledonia, Province Nord (N Province), Mount Kaala near saddle N of summit on S end of range	MEL, BRI, NOU	KP178058	KP188893	KP188960	KP189105
<i>Z. citriodora</i> J.A.Armstr.	NGW2038	Vic., cult. RBG Melbourne, grown from NGW 2038, on S bank of Limestone Creek, 1.5 km direct NE from ford on Limestone Creek track	MEL 222166	KP178038	KP188890	KP189034	KP189148
<i>Z. collina</i> C.T.White	MJB2012	Qld, cult. ANBG. Wild source: Cedar Creek NP, main circuit walking track N side (same stock as CBG 8604923)	MELU	KP178019	KP188934	KP188954	KP189139
<i>Z. collina</i>	MTM507	Qld, Tamborine NP, Panorama Point section, Wongawallan Road, Mount Tamborine	BRI AQ871795	KP178034	KP188914	KP188985	KP189173
<i>Z. compacta</i> C.T.White	PIF34479	Qld, Roberts Range, ascent to Devil's Elbow, Wyberba	BRI AQ745295	KP178046	KP188871	KP189056	KP189127
<i>Z. covenyi</i> J.A.Armstr.	MJB2013	NSW, cult. ANBG ex Narrow Neck Plateau ~5.5 km from Cliff Drive along road to Bushwalkers Hill (same stock as CBG 8411672)	MELU	KP177995	KP188838	KP189053	KP189081
<i>Z. cyrtoides</i> Sm.	AF5451	Qld, Great Dividing Range, NW of Mount Emerald near topo '1089', W of Tolga	BRI AQ849111	KP178003	KP188860	KP188952	KP189130
<i>Z. cyrtoides</i>	DJO320	Qld, Camarvon Gorge NP, 50 m upstream of the first creek encountered on the Moss Gardens walking track, 15 m above the creek near the base of the sandstone cliffs to the E of the creek	MELU	KP178066	KP188888	KP189030	KP189097

<i>Z. cytisoides</i>	DJO416	NSW, Waddin Mountains NP, Bertha's Gully, on the N side of the creek next to the walking track, ~500 m E of the junction with the track, W of the gully at the base of a sandstone rock wall	MELU	KP178005	KP188896	KP189052	KP189178
<i>Z. distans</i> Duretto & P.I.Forst.	CMH157	Qld, Coalston Lakes, Walla Range	BRI AQ756300	KP178047	KP188915	KP189037	KP189155
<i>Z. distans</i>	MTM287	Qld, Kroombit Tops NP, Razorback break, ~10 km SW of barracks	BRI AQ813602	KP178009	KP188935	KP189028	KP189115
<i>Z. eungellaensis</i> Duretto & P.I.Forst.	Bean4435	Qld, WSW of Mount David, Eungella NP	MEL 2324537, BRI AQ544202	KP178027	KP188892	KP188976	KP189118
<i>Z. exim</i> Duretto & P.I.Forst.	MTM535	Qld, narrow strip of vegetation adjacent to Sunshine Motorway, Buderim, adjacent to Buderim Waste Landfill Centre	BRI AQ792100	KP177985	KP188911	KP188981	KP189134
<i>Z. formosa</i> J.D.Briggs & J.A.Armstr.	MJB2014	NSW, cult. ANBG, Wild source: Lochiel, 250 m from junction of Pambula-Wyndham road with Back Creek Road, towards Pambula, just S of road (same stock as CBG 8604998)	MELU	KP178015	KP188946	KP188957	KP189172
<i>Z. fraseri</i> Hook. subsp. <i>fraseri</i>	SPP1995	Qld, near the summit of the south peak of Mount Maroon, W of Rathdowney	BRI AQ748157	KP178044	KP188937	KP189017	KP189095
<i>Z. fraseri</i> subsp. <i>robusta</i> (C.T.White) Duretto & P.I.Forst.	MBT3739	Qld, Salvatore Rosa NP, 160 km by road SW of Springsure, at the base of The Sentinel, ~1.5 km W of camping ground	BRI AQ830008	KP178054	KP188932	KP188964	KP189165
<i>Z. furfuracea</i> subsp. <i>eulhadenia</i> J.A.Armstr.	MTM280	Qld, Wide Bay, corner of Cooloola Way and Harry's Hut Road, N of Kin Kin	BRI AQ745998	KP177968	KP188900	KP189045	KP189077
<i>Z. furfuracea</i> R.Br. ex Benth. subsp. <i>furfuracea</i>	466549	NSW, ~5 km direct WNW of Yarras, ~700 m E along the Oxley Highway from the Mount Seaview Resort turnoff, towards Wauchope	NSW 466549	KP177990	KP188865	KP188970	KP189146
<i>Z. furfuracea</i> subsp. <i>gymnocarpa</i> J.A.Armstr.	PIF34232	Qld, Belmont Gateway Motorway, thin strip of land between Motorway and Mount Petrie Road, 1.5 km S of Old Cleveland Road	BRI AQ744205, MEL 2324803	KP177998	KP188945	KP189016	KP189131
<i>Z. graniticola</i> J.A.Armstr. ex Duretto & P.I.Forst.	MTM479	Qld, Passchendaele SF, via Carniel Lane Bapaume	BRI AQ792107	KP178070	KP188943	KP188993	KP189109
<i>Z. granulata</i> (F.Muell.) C.Moore ex Benth.	MJB2016	NSW, cult. ANBG, Wild source: Long Brush Road, Kiama (same stock as CBG 8501509)	MELU	KP178063	KP188849	KP188972	KP189083
<i>Z. hindii</i> J.A.Armstr.	413395	NSW, Mount Jerusalem NP, Blackbutt Plateau area	NSW 413359	KP178051	KP188912	KP188977	KP189084
<i>Z. hydrosopica</i> Duretto & P.I.Forst.	PIF38247	Qld, Rocky Gorge Coomanglah SF, Rocky Gorge Waterhole	BRI AQ818058	KP178021	KP188870	KP189066	KP189133
<i>Z. inexpectata</i> Duretto & P.I.Forst.	CMH201	Qld, Burnett District, Pryde's Road, on the southern road reserve, ~700 m from the intersection with the Burnett Highway	BRI	KP177987	KP188853	KP189027	KP189119
<i>Z. ingramii</i> J.A.Armstr.	759202	NSW, Goonoo State Conservation Area, Freeman's Road, ~400 m E of Mount Carl Road	NSW 759202	KP178037	KP188861	KP189007	KP189125

Appendix 1. (continued)

Taxon	Collection ID	Location	Voucher	<i>rp132-trnL</i>	<i>trnL-F</i>	<i>trnQ-rps16</i>	<i>trnS-G</i>
<i>Z. ingramii</i>	IRT613	NSW, Goonoo SF, ~53 km NE of Dubbo, ~600 m N of Demire Road from a point 1.75 km E of Mount Carl Road	MEL 2218233	KP178074	KP188843	KP188971	KP189100
<i>Z. insularis</i> Duretto & P.I.Forst.	PIF34666	Qld, Mount Windsor NP, Spencer Creek downstream of forestry barracks	BRI AQ745463	KP178012	KP188848	KP188973	KP189145
<i>Z. involucreata</i> R.Br. ex Benth.	506074	NSW, Maroota, 3.2 km along Laughtondale Gully Road from Old Northern Road	NSW 506074	KP178008	KP188846	KP189013	KP189179
<i>Z. involucreata</i>	LMD143	NSW, Laughtondale Gully Road, ~3 km from Old Northern Road	MEL 2040124	KP178062	KP188940	KP189049	KP189137
<i>Z. laevigata</i> Bonpl.	MTM258	Qld, Passchendaele SF, Forestry Lane	BRI AQ813610	KP177969	KP188950	KP189031	KP189088
<i>Z. laevigata</i>	PIF34475	Qld, Girraween NP, junction of Ramsey and Bald Rock creeks	BRI AQ813611	KP178031	KP188882	KP189006	KP189073
<i>Z. lasiocaulis</i> J.A.Armstr.	492944	NSW, N side of Banda Banda Flora Reserve, 3.4 km by road, N of intersection of Banda Banda Road and Hastings Forest Way	NSW 492944	KP178056	KP188879	KP189060	KP189182
<i>Z. lasiocaulis</i>	PG7908	NSW, ~1 km W of Mount Banda Banda at the headwaters of the Wilson River, Banda Banda Flora Reserve	MEL 2199590	KP178080	KP188942	KP189046	KP189132
<i>Z. laxiflora</i> (Benth.) Domin	MTM385	Qld, Great Sandy NP, adjacent to Cooloola Coast cemetery, Rainbow Beach Road	BRI AQ755776	KP178061	KP188880	KP189064	KP189069
<i>Z. littoralis</i> J.A.Armstr.	MJB2017	Cult. ANBG. Wild source unknown (same stock as CANB 643996)	MELU	KP177982	KP188869	KP188967	KP189079
<i>Z. littoralis</i>	RAB508	NSW, Tathra Head, 50 m from the car park at Tathra Head on the walking track leading to the lookout over the ocean	MELU	KP177980	KP188884	KP188959	KP189138
<i>Z. madida</i> Duretto & P.I.Forst.	AF5320	Qld, Daintree NP, Hilda Creek area near camping area, ~800 m WNW of Thornton Peak Site 190	BRI AQ863277	KP178078	KP188834	KP189023	KP189122
<i>Z. minutiflora</i> Domin subsp. <i>minutiflora</i>	ABP2356	Qld, ~2 km due NE of Landsborough, on margin of private property and Ewen Maddock Council Reserve	BRI AQ815342	KP177983	KP188895	KP189002	KP189093
<i>Z. minutiflora</i> subsp. <i>minutiflora</i>	PIF34313	Qld, Glass House Mountains NP, Mount Coonowrin (Crookneck)	BRI AQ862319	KP177991	KP188919	KP189015	KP189163
<i>Z. minutiflora</i> subsp. <i>trichocarpa</i> J.A.Armstr.	KRM8283	Qld, Lot 5, Webster Road, near Wondecla	BRI AQ747078	KP178000	KP188931	KP189014	KP189123
<i>Z. montana</i> J.A.Armstr.	PIF36219	Qld, Mount Barney NP, southern slopes (Peasant's Ridge) of Mount Barney, steep S-facing ridge	BRI AQ813884	KP178041	KP188928	KP189001	KP189116
<i>Z. murphyi</i> Blakely	JAI1356	NSW, Mount Tomah, on ledge between Gaping Gill and Camels Hump	MEL 2102796	KP178016	KP188878	KP189058	KP189099
<i>Z. murphyi</i>	440349	NSW, Cones Walk, Bundanoon	NSW 440349	KP178068	KP188899	KP189062	KP189166
<i>Z. obcordata</i> A.Cumm.	608324	NSW, 'The Rocks' property, W of Bathurst	NSW 608324	KP178067	KP188933	KP188975	KP189098
<i>Z. obcordata</i>	RGC10060	NSW, 'Bulbudgerie' property, near Wuuluman, 15 km ENE of Wellington	MEL 2067510	KP178026	KP188841	KP188997	KP189126

<i>Z. obovata</i> (C.T. White) J.A. Armstr.	KRM8071	Qld, Grant Creek, Anniversary Falls on Rush Track, W of Herberton	BRI AQ861639	KP178023	KP188924	KP188956	KP189074
<i>Z. odorifera</i> J.A. Armstr. subsp. <i>odorifera</i>	RM1196	NSW, ~9 km (direct) SW of Molong Bocabie Gap rocky bluff immediately S of road and railway line in gap near base of bluff	MEL 717377, BRI AQ591376	KP178004	KP188887	KP189018	KP189142
<i>Z. odorifera</i> subsp. <i>warrahensis</i>	599824	NSW, Warrahah, W of Kingstown, on private property	NSW 599824	KP177970	KP188851	KP189047	KP189096
Duretto & P.I. Forst. <i>Z. odorifera</i> subsp. <i>williamsii</i> Duretto & P.I. Forst.	MJB2018	NSW, cult. ANBG. Wild source: 2.4 km from Howell on track to Copeton Lake Recreation area (same stock as CBG 8505944)	MELU	KP178017	KP188864	KP188998	KP189128
<i>Z. oreocena</i> J.A. Armstr.	DJO222	Vic., Grampians NP, Cultivation Creek Gorge, 100 m upstream from where Buandik Falls walking track ends, 1 m W of the creek	MELU	KP177975	KP188936	KP189011	KP189149
<i>Z. oreocena</i>	MJB1982	Vic., Grampians NP, Deep Creek, ~150 m upstream of crossing on Harrop Track	MEL	KP178071	KP188904	KP189003	KP189089
<i>Z. oreocena</i>	RAB580	Vic., Hollow Mountain, Northern Grampians, in sheltered rock outcrop, S side of upper cave, just below summit	MELU	KP178006	KP188889	KP189036	KP189091
<i>Z. oreocena</i>	RAB581	Vic., Northern Grampians, Hollow Mountain, walking track to Summer Day Valley climbing area	MELU	KP178002	KP188852	KP189055	KP189164
<i>Z. parrisiae</i> J.D. Briggs & J.A. Armstr.	RJ1899	NSW, Box Range Farm, ~10 km W of Pambula	MEL 2312304	KP178073	KP188862	KP188980	KP189152
<i>Z. pilosa</i> Rudge <i>Z. pilosa</i>	MFD2008 RAB502	NSW, Berowra NP (Sydney Sandstone North) NSW, Georges River NP, Picnic Pointon ridge track behind the second block of toilets from the road entrance	NSW MELU	KP177981 KP177973	KP188917 KP188921	KP189019 KP189032	KP189174 KP189106
<i>Z. prostrata</i> J.A. Armstr. <i>Z. prostrata</i>	758677 MJB2019	NSW, Look-at-me-now Headland, ~150 m SE of car park NSW, cult. ANBG. Wild source: Dammerels Headland (same stock as CBG 8803922)	NSW 758677 MELU	KP178075 KP178032	KP188916 KP188873	KP188999 KP189025	KP189101 KP189070
<i>Z. rimulosa</i> C.T. White	JRC5802	Qld, Mount Mulligan, ~40 km NW of Dimbulah, top of mountain, ~0.5 km SE of dam, rock outcrop	MEL 682024, BRI AQ398829	KP178050	KP188847	KP189059	KP189177
<i>Z. robertsonium</i> J.A. Armstr. <i>Z. robusta</i> Maiden & Bethe	AF5349 DEA s.n.	Qld, Malbon Thompson Forest Reserve, NE of Bell Peak, North site 194 Vic., 3.7 km S of Little Mount Kent, along walking track to Moroka Gorge from Horseyard Flat above N facing sandstone cliff	BRI AQ848776 MEL 2019411 A	KP178011 KP177966	KP188868 KP188855	KP189022 KP189051	KP189162 KP189135
<i>Z. robusta</i>	MFD2049 (AD 950185)	NSW, cult. ABG, Mount Annan. Wild source: Kanangra-Boyd NP, Kanangra Brook Falls, Kanangra Walls, below waterfall	NSW	KP178077	KP188842	KP188978	KP189151
<i>Z. scopulius</i> Duretto & P.I. Forst.	SPP1980	Qld, Flinders Peak, ~22 km S of Ipswich	BRI AQ747759	KP178007	KP188926	KP189043	KP189168

Appendix 1. (continued)

Taxon	Collection ID	Location	Voucher	<i>rpl32-trnL</i>	<i>trnL-F</i>	<i>trnQ-rps16</i>	<i>trnS-G</i>
<i>Z. smithii</i> Jacks.	AF5659	Qld, Herberton Range SF, steep upper slope of Mount Misch, W of Tolga	CNS 134680	KP178060	KP188938	KP189050	KP189075
<i>Z. smithii</i>	DJO327	Qld, Kroombit Tops NP, W of Dry Creek, 60 m downstream from the road crossing of Dry Creek which is 300 m E of the Barracks NSW, Hat Head NP, Korogoro Point, next to the Korogoro Point Track circuit, 100 m from the start when walking the circuit in an anticlockwise direction	MELU	KP177986	KP188872	KP189039	KP189154
<i>Z. smithii</i>	DJO430	Qld, Lot 5, Webster Road, near Wondecia Qld, Currigee, South Stradbroke Island Vic., Briagolong–Stockdale road, ~300 m W of Stratford–Stockdale Road, Gippsland, Vic.	MELU	KP178029	KP188947	KP189020	KP189170
<i>Z. smithii</i>	KRM8282	Qld, Clewley's Gap area, Kroombit Tops NP	BRI AQ747081	KP178053	KP188835	KP189057	KP189129
<i>Z. smithii</i>	MJB2023	Qld, Kureelpa Falls area, Mapleton SF	MELU	KP178025	KP188948	KP189044	KP189082
<i>Z. smithii</i>	MJB2029	Qld, Kureelpa Falls, South Maroochy River, ~4 km NE of Mapleton via Kureelpa	MELU	KP177999	KP188836	KP188990	KP189080
<i>Z. smithii</i>	MJB2184	Qld, North Stradbroke Island, E of Amity Point		KP178028	KP188881	KP188968	KP189068
<i>Z. smithii</i>	MTM303	Qld, Cainbale Creek catchment, just W of Cainbale Falls, Lamington NP		KP178065	KP188910	KP188961	KP189167
<i>Z. smithii</i>	MTM926	Qld, Mount Walsh NP, Palm Valley area, 15 km S of Biggenden Coast Range	BRI AQ757195	KP178020	KP188840	KP189004	KP189181
<i>Z. smithii</i>	PIF33151	Qld, Main Range NP, start of Mount Cordeaux walking track	BRI AQ729276	KP178036	KP188859	KP189012	KP189090
<i>Z. smithii</i>	PIF34139	Qld, Mount Barney NP, southern base of Mount Barney	BRI AQ743453	KP177971	KP188907	KP188991	KP189087
<i>Z. smithii</i>	PIF36221	Qld, Bulburin NP, Daves Range Road	BRI AQ813882	KP178030	KP188920	KP189042	KP189160
<i>Z. smithii</i>	PIF38231	Qld, Toolara SF (SF 1004), Harrys Hut Road, ~200 m from Cooloola Way	BRI AQ818053	KP178033	KP188867	KP188953	KP189094
<i>Z. smithii</i>	PIF38266	Qld, Conondale NP (formerly Imbil SF 2, Ryan LA)	BRI AQ823442	KP178081	KP188918	KP188963	KP189086
<i>Z. smithii</i>	PIF38276	Qld, Lyrebird Ridge Road, Springbrook	BRI AQ818086	KP178018	KP188839	KP188996	KP189103
<i>Z. smithii</i>	PIF38466	Vic., Wingan Inlet Croaringalong NP, Rams Head Track	BRI AQ818295	KP177989	KP188906	KP189038	KP189071
<i>Z. smithii</i>	RAB511	Vic., Moilun Creek Bridge, Stockdale to Glenaladale Road (Beverleys Road), just upstream from crossing	MELU	KP178072	KP188844	KP188983	KP189111
<i>Z. smithii</i>	RAB548	Qld, Mount Nebo, near Brisbane, Scrub Road ~4 km directly SE of Jolllys Lookout, Mount Nebo D'Aguiar Range	MELU		KP188877	KP188989	KP189184
<i>Z. smithii</i>	SPP2032	Qld, Boy-ull Creek, ~100 m upstream from Twin Falls, Springbrook NP	BRI AQ748343	KP178076	KP188922	KP189005	KP189107
<i>Z. southwellii</i> J.A. Armstr.	MTM251		BRI AQ745450	KP178049	KP188856	KP189000	KP189113

Chloroplast phylogeny of *Zieria* (Rutaceae)

<i>Z. tenuis</i> Duretto & P.I.Forst.	Bean4609	Qld, Bertya Creek, W of Warang, White Mountains NP	MEL 719859, BRI AQ544935	KP178082	KP188866	KP189061	KP189176
<i>Z. tuberculata</i> J.A.Armstr.	RAB505	NSW, Tilba Central, on the hill with the water tower	MELU	KP177992	KP188902	KP189041	KP189156
<i>Z. vagans</i> Duretto & P.I.Forst.	PIF34832	Qld, Gurgeena Plateau, SF 172, near Meredith's Road	BRI AQ745514	KP177996	KP188927	KP188955	KP189108
<i>Z. veronicea</i> subsp. <i>insularis</i> J.A.Armstr.	WN14AE	SA, Kangaroo Island, Finders Chase NP, Yacca Flat Road	MELU D 108693	KP177978	KP188863	KP189008	KP189124
<i>Z. veronicea</i> (F.Muell.) Benth. subsp. <i>veronicea</i> J.A.Armstr.	MJB1980	SA, Cox Scrub Conservation Reserve	MELU	KP178040	KP188929	KP189040	KP189169
<i>Z. verrucosa</i> J.A.Armstr.	PIF34734	Qld, Blue Hills, 3 km NNE of Monogorilby	BRI AQ745491	KP177994	KP188875	KP188962	KP189102
<i>Z. whitei</i> J.A.Armstr. ex Duretto & P.I.Forst.	AF5486	Qld, Oaky Creek Falls, ~1 km SW of Mount Emerald	BRI AQ755971	KP177967	KP188909	KP189033	KP189183

Supplementary material

A chloroplast phylogeny of *Zieria* (Rutaceae) in Australia and New Caledonia shows widespread incongruence with species-level taxonomy

*Rosemary A. Barrett^A, Michael J. Bayly^{A,E}, Marco F. Duretto^B, Paul I. Forster^C, Pauline Y. Ladiges^A
and David J. Cantrill^D*

^ASchool of Botany, The University of Melbourne, Vic. 3010, Australia.

^BNational Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Mrs Macquaries Road, Sydney, NSW 2000, Australia.

^CQueensland Herbarium, Department of Science, Information Technology, Innovation and the Arts, Brisbane Botanic Gardens, Toowong, Qld 4066, Australia.

^DRoyal Botanic Gardens Melbourne, Birdwood Avenue, South Yarra, Vic. 3141, Australia.

^ECorresponding author. Email: mbayly@unimelb.edu.au

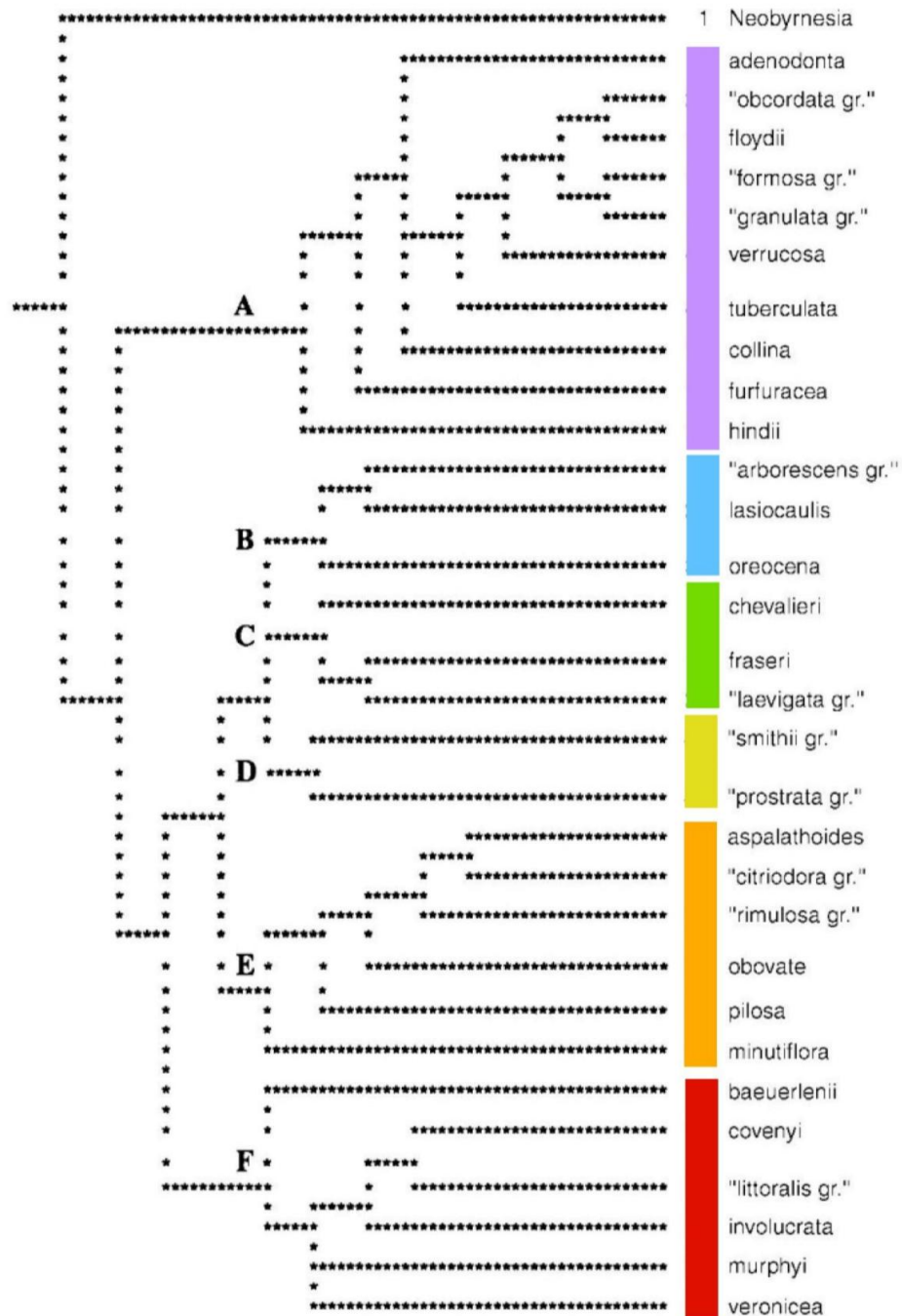


Fig. S1. Phylogeny of *Zieria* based on analysis of morphological and chemical characters by Armstrong (2002). Species within groups labelled 'littoralis gr.', for example, were not defined by Armstrong.

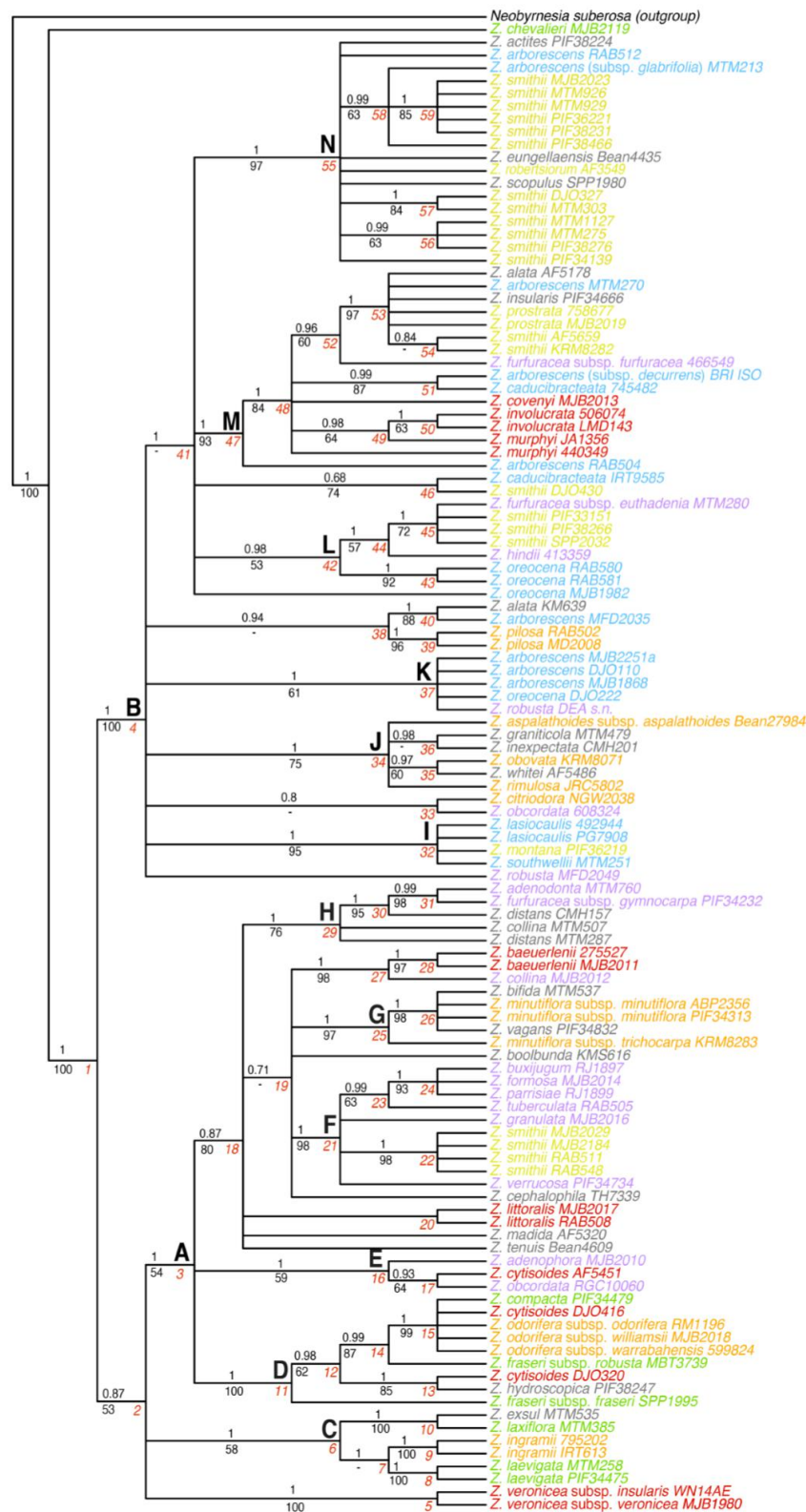


Fig. S2. Bayesian majority-rule consensus tree produced by analysis of combined cpDNA markers (as in Fig. 2), showing taxa colour coded according to the groups defined in the phylogeny of Armstrong (2002). Colours of groups follow those indicated on Fig. S1. Taxa shown in grey were not included in Armstrong's (2002) analysis. Bayesian inference (BI) posterior probabilities (PP) are shown above the branches and maximum parsimony bootstrap support (BS) values are shown below the branches.

References

- Armstrong JA (2002) *Zieria* (Rutaceae): a systematic and evolutionary study. *Australian Systematic Botany* **15**, 277–463.
doi:10.1071/SB00040

Chapter 4. Phylogeny of *Zieria* based on nuclear ribosomal DNA

4.1. Introduction

There have been no phylogenetic studies of the genus *Zieria* based solely on nuclear ribosomal DNA (nrDNA) to date. Previous phylogenetic studies of *Zieria* have been based on morphological and chemical characters (Armstrong 2002), chloroplast DNA (Chapter 3 of this thesis, published, see Barrett *et al.* 2015), and a combination of cpDNA, nrDNA and Armstrong's (2002) morphological and chemical characters (Morton 2015). The phylogeny of Armstrong (2002) was based on 48 morphological and 62 chemical characters, and was noted as being 'very tentative' by Armstrong himself. It showed high levels of homoplasy, a low consistency index (0.25), and relationships between and within the six major clades were poorly resolved or poorly supported. Armstrong's (2002) morphological groups were also poorly correlated with those based on molecular data in Chapter 3 of this thesis and the study of Morton (2015). The cpDNA study (Chapter 3) included 59 (of 60) species, most with multiple accessions (109 in total), and revealed widespread incongruence between the cpDNA tree and species-level taxonomy based on morphology. This incongruence was not detected in the more limited study of Morton (2015) which included only single accessions of 32 species and was based on a smaller cpDNA dataset, including only *rpl32-trnL* and *trnL-F*, whereas the cpDNA study (Chapter 3) also included data for *trnQ-rps16* and *trnS-G*. A combination of factors was suggested to explain the incongruence between the cpDNA data and species taxonomy, including regional cpDNA introgression (chloroplast capture), incomplete lineage sorting (deep coalescence) and inappropriate taxonomic boundaries. Also highlighted was the need for an analysis based on nuclear data, which should include multiple accessions of species, to help interpret the complex relationships within *Zieria*, better understand the underlying evolutionary processes, and possibly review taxon circumscription.

Within this context, the main aims in this chapter were to: (1) produce a comprehensive phylogeny of *Zieria* based on sequencing two nuclear markers, ITS and ETS; (2) compare the results with relationships resolved by previous analyses of chloroplast sequence data (Chapter 3) and combined molecular and morphological data (Morton

2015); (3) assess the congruence of the nrDNA phylogeny with current taxonomy within the genus.

4.2. Materials and methods

Materials and methods used for this study (including DNA extraction, amplification, cloning and sequencing, sequence alignment and phylogenetic analysis) are described in detail in General Methods (Chapter 2), but some information of particular relevance to this chapter is included below.

Taxon sampling

Where possible, to enable direct comparisons between studies, the same accessions were used in this nrDNA analysis as in the cpDNA analysis of *Zieria* (Chapter 3), except for a few samples that proved difficult to sequence. Ingroup sampling included 109 accessions representing 58 (of 60) *Zieria* species and 13 (of 18) subspecies. The only taxa not represented in this nrDNA study were: *Z. actites* Duretto & P.I.Forst., *Z. aspalathoides* subsp. *brachyphylla* J.A.Armstr., *Z. floydii* J.A.Armstr., *Z. odorifera* subsp. *copelandii* Duretto & P.I.Forst., *Z. odorifera* J.A.Armstr. subsp. *odorifera*, and *Z. odorifera* subsp. *warrabahensis* Duretto & P.I.Forst. This was mostly due to suitable material being unavailable. Of three subspecies of *Z. odorifera* originally sampled for the study, two were excluded (subsp. *odorifera* and subsp. *warrabahensis*) following preliminary analysis because of the high number of polymorphisms detected in their ITS and ETS regions using direct sequencing. This species was represented in this analysis by a single accession of *Z. odorifera* subsp. *williamsii*. For reasons given in General Methods (Chapter 2) accessions of *Z. arborescens* were not identified to subspecies, but accessions that were collected near type localities for the two segregate subspecies are noted in figures and tables (subsp. *glabrifolia* and subsp. *decurrens*). Details of all samples are presented in Appendix A, including authority names, sampling locations and voucher specimen identification numbers.

Molecular markers used in this study

Two nrDNA markers were sequenced for this study, the internal (ITS) and external (ETS) transcribed spacer regions of the 18S–5.8S–26S cistron (see Chapter 2 for general information about these markers plus details of primers and PCR protocols that

were used). The models used for Bayesian inference (BI) analyses were selected using the Akaike information criterion (AIC) as implemented in MrModeltest 2.3 (Nylander 2004). They were as follows: (1) for the analysis of the individual datasets, SYM+I+G for ITS and GTR+G for ETS; (2) for the analysis of the combined datasets, SYM+G for ITS and GTR+G for ETS. (Note that the analyses of the individual datasets included *all* the cloned sequences, whereas the combined dataset contained selected cloned sequences only, which explains why different models were applied to ITS in the different analyses). Indels were not separately coded for analyses because there were no multi-base indels in the nrDNA sequences.

Preliminary analyses of the individual ITS and ETS datasets were undertaken in order to: (1) evaluate the validity of combining the ITS and ETS datasets (assess congruency); (2) discover the placement within the trees of sequences obtained by cloning of four selected samples; (3) assess which ITS and ETS copies were likely to have originated from the same nrDNA array within an individual, in order to justify concatenation of copies for the combined analysis. An incongruence length difference (ILD) test and a recombination analysis were performed on the datasets as described in General Methods (Chapter 2). Following the preliminary analyses of the ITS and ETS datasets, a combined analysis of the datasets was undertaken.

ITS and ETS paralogues

Polymorphic sequences in some samples indicated the possibility of heterogeneous ITS and ETS repeat types. Most sequences with more than four polymorphisms in either ITS or ETS were excluded from further consideration in this analysis. Sequences containing only one or a few polymorphisms in an otherwise ‘clean’ trace were included in the dataset with polymorphisms coded using IUPAC nucleotide ambiguity codes. Deeply divergent paralogues of ITS and ETS were detected in some samples and four were subsequently selected for cloning (see Chapter 2 for details). The main aim was to amplify and sequence notably divergent copies of ITS and ETS, not necessarily *all* copies; it was assumed that copies differing by only a base or two would be unlikely to give a different phylogenetic signal.

Concatenated cloned sequences in the combined ITS + ETS analyses are referred to using the taxon name and identification number for this study, followed by two numbers

in brackets representing the respective cloned ITS and ETS sequences, e.g., *Z. exsul* MTM535 (2661 2698), *Z. exsul* MTM535 (2660 2691).

4.3. Results

Amplification and sequencing of ITS and ETS: assessment of variation within genomes

Most amplification reactions resulted in a single ITS or ETS PCR product per sample, and direct sequencing mostly yielded ‘clean’ sequences; 86% of ITS and 88% of ETS sequences had three or fewer polymorphisms in chromatograms from direct sequences. Other samples contained many polymorphisms (e.g. up to 33 in one marker), indicating that distinctly different copies of nrDNA might be present in those individuals. The number of polymorphisms (greater than zero) within ITS and ETS sequences is given in Appendix B.

To assess nrDNA variation in one sample with especially polymorphic nrDNA sequences (*Zieria prostrata* MJB2019) copy-specific ITS primers were designed for that sample as described in General Methods (Chapter 2). Using these primers, two highly divergent copies of ITS were successfully amplified and sequenced, differing in 17 base pairs across 548 base pairs of sequence. Differences between these two sequences reflected the polymorphisms detected in the initial direct sequencing of ITS for this sample i.e., using universal primers (Figure 4.1).

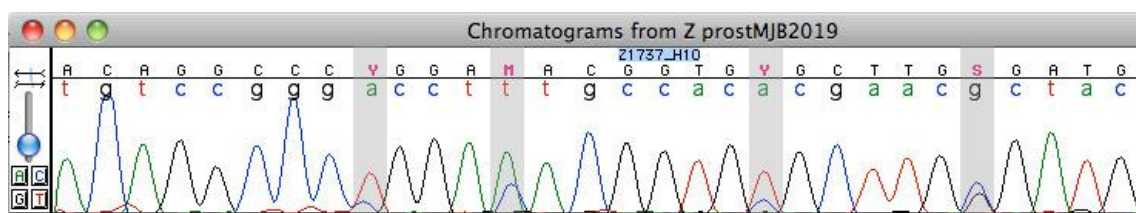


Figure 4.1 Section of ITS1 showing polymorphisms typical in chromatograms of ITS sequences of *Zieria prostrata* MJB2019 when amplified using universal primers ITS5 and ITS4. Grey columns highlight polymorphic loci.

The copy-specific primers were also trialled on another sample of *Z. prostrata* (758677), which had shown no obvious polymorphisms in direct sequencing of ITS (Figure 4.2, sequences A and B), but which subsequently yielded two divergent ITS sequences (C and D) using the copy-specific primers. These divergent sequences also differed in 17 base pairs across 548 base pairs of sequence.

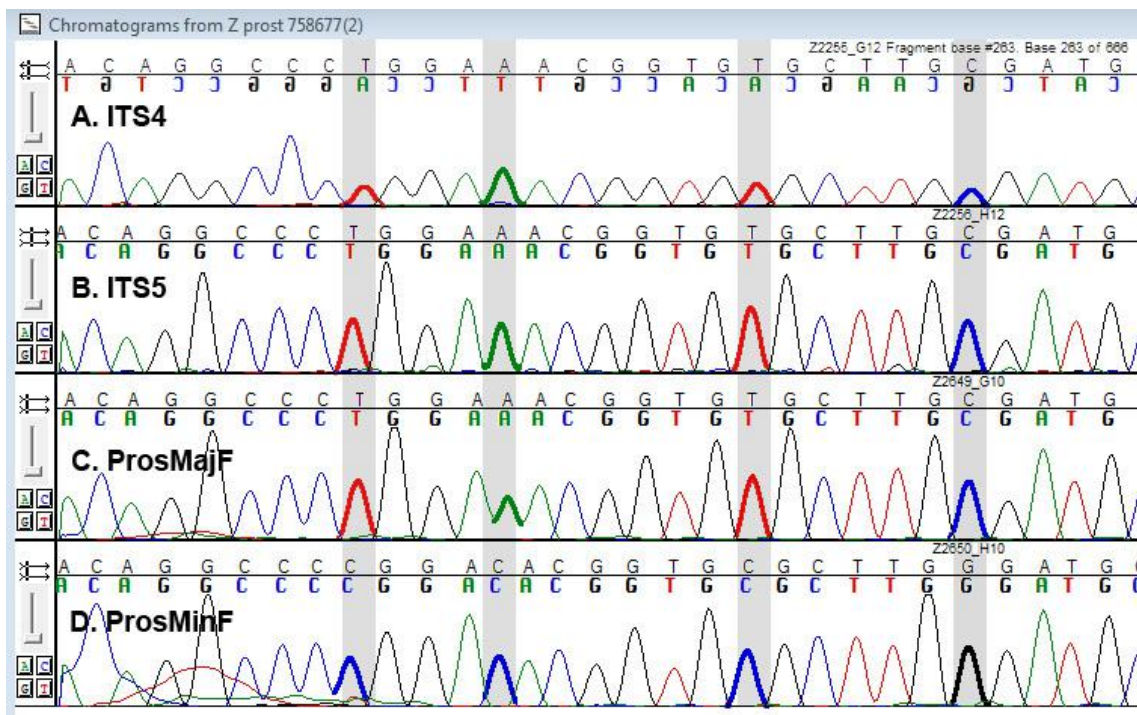


Figure 4.2 Sections of aligned ITS1 chromatograms from *Zieria prostrata* 758677, amplified by various primers (reverse primer ITS4, and forward primers ITS5, ProsMajorF and ProsMinorF). Chromatograms are from the same ITS1 region as shown in Figure 4.1, for *Z. prostrata* MJB2019. Only primer ProsMinorF amplified sequence D in this sample. Grey areas highlight polymorphic loci, which are identical to those in sample MJB2019.

Variation in nrDNA within individuals was also assessed by cloning the PCR products of both ITS and ETS for four samples with highly polymorphic nrDNA. Up to six variants of ITS and ETS were identified in individual samples (Table 4.1.). Divergence between copies was greatest in ITS sequences, with up to 28 differences between copies within sample *Zieria prostrata* MJB2019. Most divergence was in the spacer regions (showing up to 6.3% divergence within samples) with less in the 5.8S gene (showing up to 1.9% divergence within samples). For ETS, every cloned sequence from all samples was unique (up to 6 variants per sample) but cloning of ETS was less successful than ITS for two of the four samples. In the ETS dataset there were no variable characters in the 18S gene (both within and between samples) but only about 56 base pairs of this gene was included.

Table 4.1. Summary of bacterial cloning and sequencing of nrDNA PCR products.

Sample	nrDNA region	Number of colonies picked	Number sequenced successfully	Number of variants	Maximum number of pairwise differences between copies	Pairwise divergence in 5.8S gene	Pairwise divergence in ITS1, ITS2 and ETS
<i>Z. exsul</i> MTM535	ITS	20	11	6	27	0 – 1.9%	0 – 4.7%
<i>Z. insularis</i> PIF34666	ITS	30	6	5	24	0 – 1.9%	0.2 – 5.4%
<i>Z. odorifera</i> subsp. <i>williamsii</i> MJB2018	ITS	10	6	5	26	0	0.7 – 5.8%
<i>Z. prostrata</i> MJB2019	ITS	5	5	2	28	0	6.3%
<i>Z. exsul</i> MTM535	ETS	10	5	5	10	–	0 – 2.4%
<i>Z. insularis</i> PIF34666	ETS	10	6	6	14	–	0.2 – 3.4%
<i>Z. odorifera</i> subsp. <i>williamsii</i> MJB2018	ETS	5	3	3	7	–	0 – 1.7%
<i>Z. prostrata</i> MJB2019	ETS	10	3	3	11	–	0.2 – 2.7%

Tests for recombination and pseudogenes in ITS and ETS sequences

When divergent nrDNA genes are present in individual genomes there is a possibility that there has been recombination between gene copies, or that some copies might represent (non-functional) pseudogenes, or both. The results of a recombination analysis showed no evidence of recombination between divergent copies of either ITS or ETS, using the range of recombination detection protocols implemented in the RDP3 package (Heath *et al.* 2006).

Sequence comparisons aimed at detecting possible pseudogenes gave mixed signals; they provided little evidence for the presence of pseudogenes, except possibly some sequences from one sample, *Z. insularis* (ITS copies 2654, 2657 and 2718), discussed below. Typical signatures of pseudogenes were generally not apparent in the dataset. There was no significant decrease in guanine and cytosine (G + C) content: overall the ITS sequences had 55.7–58.1% G + C content, and the ITS paralogues had 56.0–57.2%; the ETS sequences had 48.5–50.9% G + C content overall and the ETS paralogues had 48.8–50.6%. Also, there was no clearly reduced number of methylation sites: there were 80–90 methylation sites in the ITS paralogues and 78–91 in the ITS sequences overall; and 32–38 methylation sites in the ETS paralogues and 30–41 in the ETS sequences overall. Both measures (G + C content and number of methylation sites) fell within a tight range, varying minimally but continuously across samples in the dataset, and not

falling into discrete classes, as would be expected among highly divergent pseudogenes (e.g. Holmes *et al.* 2014).

In angiosperms, evidence for potential pseudogenes is strongest when based on patterns of variation in the 5.8S gene, including in motifs that are generally conserved. In *Zieria*, variation in the 5.8S gene occurred at low levels across all sequences in the ITS dataset (including all paralogues), with a total of 25 changes (tree steps) occurring in the 5.8S gene (12 variable sites in 162 base pairs), compared to 233 changes in ITS1 (116 variable sites in 225 base pairs) and 202 changes in ITS2 (99 variable sites in 230 base pairs). This equates to percentages of variable sites within regions as follows: 7% in the 5.8s gene; 52% in ITS1; and 43% in ITS2. Nineteen of the 25 changes in the 5.8S gene were from cytosine to thymine and, of these, 15 were at methylation sites. There was a small cluster of changes in the 5.8S gene in a highly supported clade containing three ITS copies from *Z. insularis* PIF34666 (ITS copies 2654, 2657 and 2718; see below for details of the ITS tree). Other than this, changes in the 5.8S gene were scattered throughout the tree.

The evidence for ITS pseudogenes based on conserved motifs was also not conclusive. The motifs investigated (see page 2.18 in Chapter 2, General Methods) were conserved across most sequences, but there was some variation in divergent (cloned) sequences of *Z. insularis*, where three ITS copies (2656, 2657 and 2718) did not contain the conserved motif (5'-GAATTGCAGAATCC) identified by Jobes and Thien (1997). The motif in ITS1 identified by (Liu and Schardl 1994) was conserved in 90 sequences, but there was one base pair difference in this motif in 35 sequences, including in three of the cloned sequences (*Z. prostrata* MJB2019 ITS copy 2714, *Z. odorifera* subsp. *williamsii* MJB2018 ITS copies 2704 and 2705).

For ETS sequences, comparison of variation in genes versus spacers provided no evidence of pseudogenes. However, this was not unexpected because the ETS dataset included only 56 base pairs of the 18S gene. There were no variable characters in this part of the gene; all changes occurred in the spacer.

Analyses of separate ITS and ETS datasets

Each dataset represented 58 (of 60) *Zieria* species and 13 subspecies, and the outgroup; four *Zieria* accessions were represented by varying numbers of divergent ITS and ETS

copies obtained by cloning of PCR products. The ITS analysis was based on 125 sequences from 109 accessions and included 617 characters, 126 of which were parsimony informative. The ETS analysis was based on 122 sequences from those same accessions and included 476 characters, 99 of which were parsimony informative. Maximum parsimony (MP) analyses resulted in many thousands of equally parsimonious trees for each dataset (see Chapter 2, section 2.12, for the heuristic search strategy used). Statistics from the MP analyses are presented in Table 4.2.

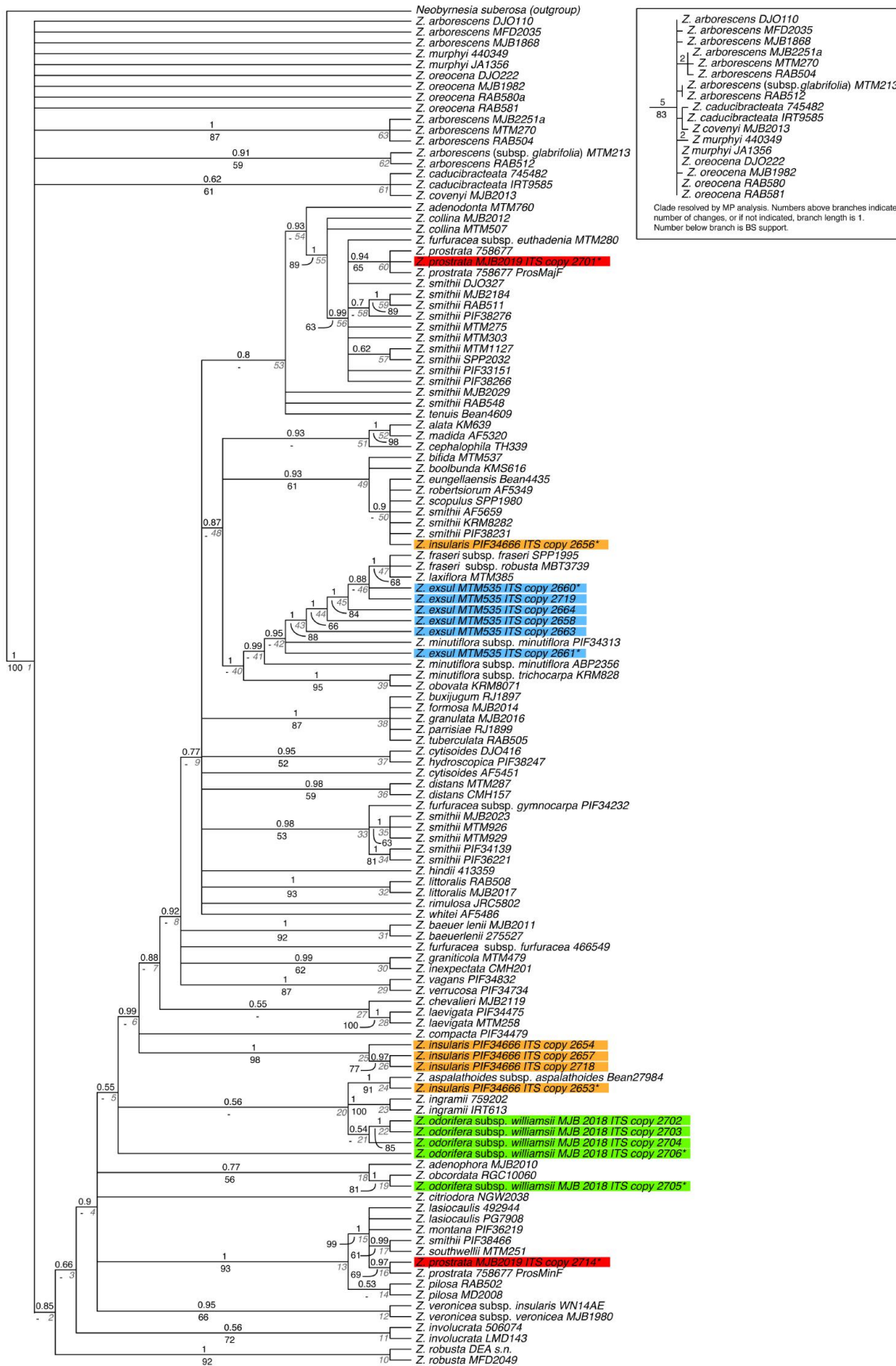
Table 4.2. Details of (separate) maximum parsimony analyses of ITS and ETS datasets.

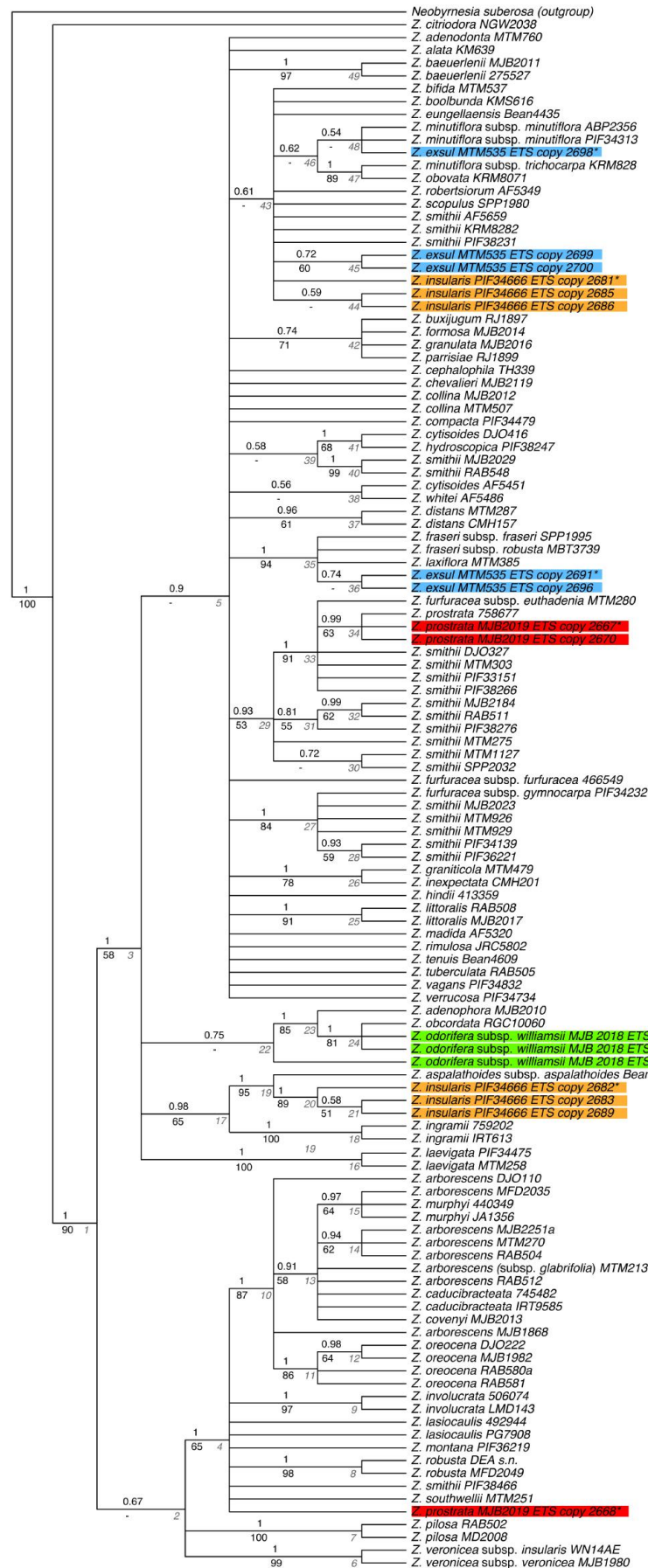
	ITS	ETS
Number of included characters	617	476
Number of parsimony informative characters	126	99
Tree Length (shortest trees)	460	355
Consistency Index (CI)	0.69	0.74
Retention Index (RI)	0.89	0.87
Rescaled Consistency Index (RC)	0.61	0.65

Tree topology from Bayesian inference (BI) and MP analyses were similar, but BI trees were generally more resolved than the MP strict consensus (with the exception of one node as shown on Figure 4.3). The BI 50% majority rule consensus trees were poorly resolved for both ITS (Figure 4.3) and ETS (Figure 4.4) datasets, ETS in particular; each tree resolved small terminal clades (variously supported) with some larger unsupported polytomies at deeper nodes. *Zieria citriodora* was strongly resolved by ETS as the sister to all other *Zieria* accessions (PP 1, BS 90%), and although not resolved by ITS, this relationship was not conflicting in ITS. The backbone of the ITS tree (nodes 2 to 9; Figure 4.3) had mostly poor support in the BI analysis (with the exception of node 6, PP 0.99), and most of these nodes were either unresolved in the MP strict consensus or had less than 50% BS support. The BI tree based on ETS included two major clades (nodes 2 and 3), although node 2 had low PP and was not resolved in the MP strict consensus; node 3 was highly supported by BI (PP 1) but only poorly supported by MP (BS 58%).

Following page. Figure 4.3. Bayesian inference majority-rule consensus tree of ITS data. Placement of the divergent ITS copies are highlighted (a different colour for each sample), and asterisks indicate those selected for pairing with ETS sequences for inclusion in a combined analysis. PP values are shown above branches, BS values for MP analysis are below branches, and nodes are numbered in grey.

Nuclear phylogeny of *Zieria*





The nodes that were resolved in the separate analyses of the ITS and ETS datasets were broadly congruent. Although the incongruence length difference (ILD) test revealed some conflict between the two datasets ($p = 0.024$), comparisons of trees showed that conflict was restricted to four pairs of nodes. Of these, three conflicts were poorly supported, i.e., with PP support of 0.62 or less for at least one of the conflicting nodes, and all BS values less than 50%. These three conflicts were between ITS node 27 and ETS node 5, ITS node 40 and ETS node 46, and ITS node 53 and ETS node 39. The other conflicting nodes were ITS node 13 and ETS node 4, which both had PP = 1 but lower BS (65% for ETS node 4) and resolved different relationships for just one species, *Z. pilosa*.

Given this level of congruence between ITS and ETS datasets it was considered reasonable to combine the ITS and ETS data for further analyses. This is done routinely in phylogenetic analysis, especially given that ITS and ETS are from the same nrDNA cistron, and that combined analyses are likely to provide more resolution than those performed on each marker separately (Baldwin and Markos 1998; Bena *et al.* 1998; Clevinger and Panero 2000). A complication with combining these datasets comes from the presence in some samples of divergent copies of ITS and ETS. Therefore, prior to undertaking the combined analysis (discussed in the next section), it was necessary to ascertain which paralogues (cloned copies) could reasonably be concatenated. The ITS and ETS paralogues joined for subsequent analyses are listed in Table 4.3 and the rationale for these combinations is given in Appendix C.

Table 4.3. ITS and ETS copies (identified by copy number) paired for the combined analysis.

Sample	ITS copy		ETS copy
<i>Z. exsul</i> MTM535	2660	paired with	2691
	2661	paired with	2698
<i>Z. insularis</i> PIF34666	2653	paired with	2682
	2656	paired with	2681
<i>Z. odorifera</i> subsp. <i>williamsii</i> MJB2019	2705	paired with	2679
	2706	paired with	2678
<i>Z. prostrata</i> MJB2019	2714	paired with	2668
	2701	paired with	2667

Preceding page, Figure 4.4. Bayesian inference majority-rule consensus tree of ETS data. Placement of the divergent ETS copies are highlighted (a different colour for each sample), and asterisks indicate those selected for pairing with ITS sequences for inclusion in a combined analysis. PP values are shown above branches, BS values are below branches, and nodes are numbered in grey.

Analyses of combined ITS + ETS datasets: overview

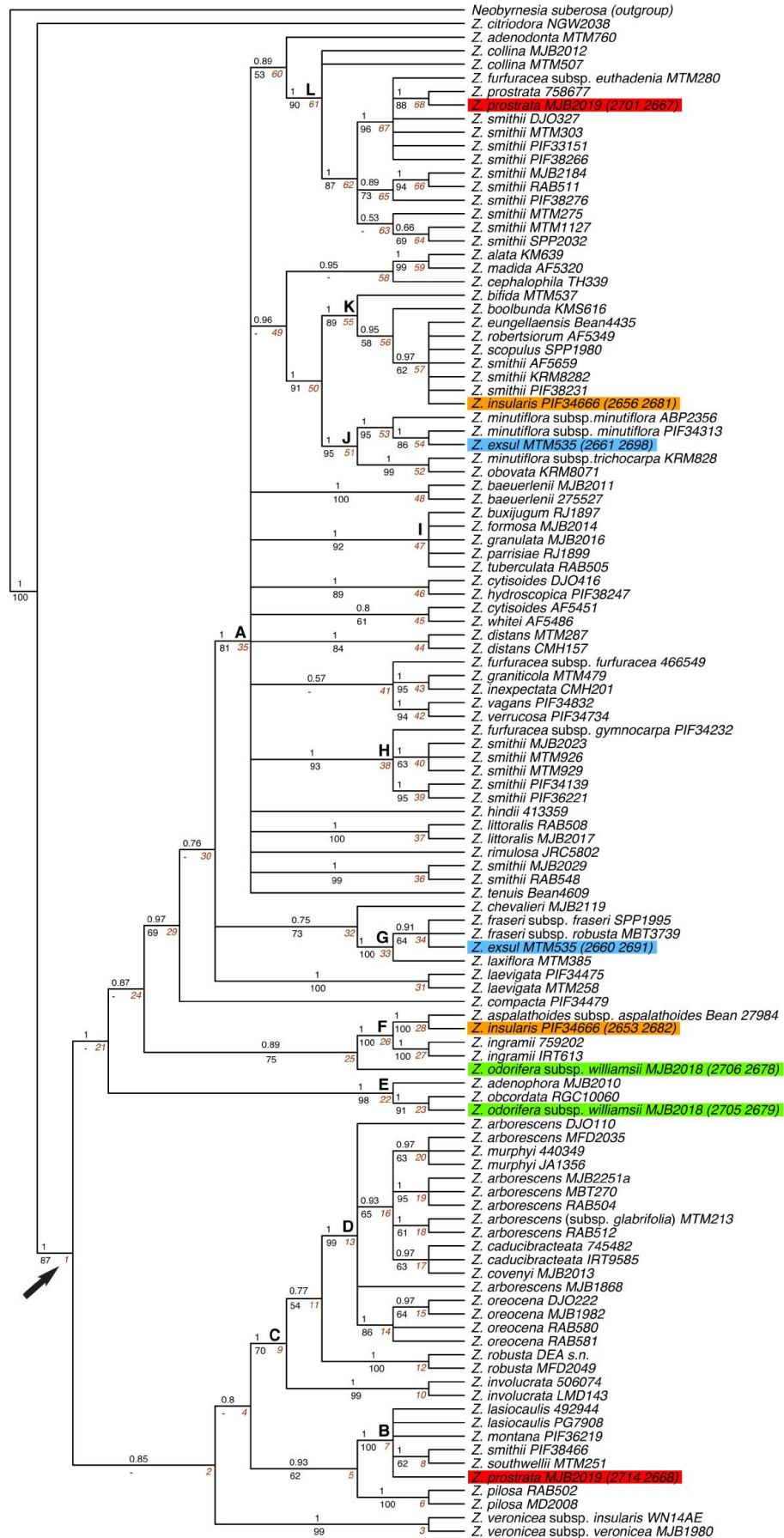
The combined ITS + ETS datasets included 109 accessions representing 58 (of 60) *Zieria* species and 13 subspecies, plus *Neobyrnesia suberosa* as the outgroup. Four samples were represented by multiple, cloned sequences. BI and MP analyses each included 1093 characters of which 413 were variable and 214 were parsimony informative. The MP analysis (Table 4.4) resulted in many thousands of trees each of length 772. The topology of the BI tree was similar to, but slightly better resolved than, the MP strict consensus tree. The two measures of support (PP and BS) were not always well correlated, with some high PP values occurring when the corresponding BS support values were low or below 50%. However, as there were no conflicting nodes, the BI majority-rule consensus tree only is presented (Figure 4.5), with BS values mapped onto it. Detailed sections of the BI tree are also presented (clades B to G in Figure 4.6, and the large polytomous clade A in Figure 4.7) with support values and branch lengths from MP analysis mapped onto them where relevant, and some well-supported nodes labelled A to L.

Table 4.4 Details of heuristic search results in the maximum parsimony analysis of combined ITS + ETS data.

	ITS	ETS	Combined
Number of included characters	617	476	1093
Number of variable characters	186	227	413
Number of parsimony informative characters	118	96	214
Parsimony informative characters as a percentage of variable characters	63%	42%	52%
Tree Length (shortest trees)	417	343	772
Consistency Index (CI)	0.76	0.77	0.75
Retention Index (RI)	0.91	0.88	0.89
Rescaled Consistency Index (RC)	0.69	0.67	0.67

Following page. Figure 4.5. Bayesian inference majority-rule consensus tree produced from combined ITS + ETS data, showing all accessions, nodes (numbered in *italics*), and support values. PP values are shown above the branches and BS values below. Divergent paralogues of four taxa are highlighted (a different colour for each sample).

Nuclear phylogeny of *Zieria*



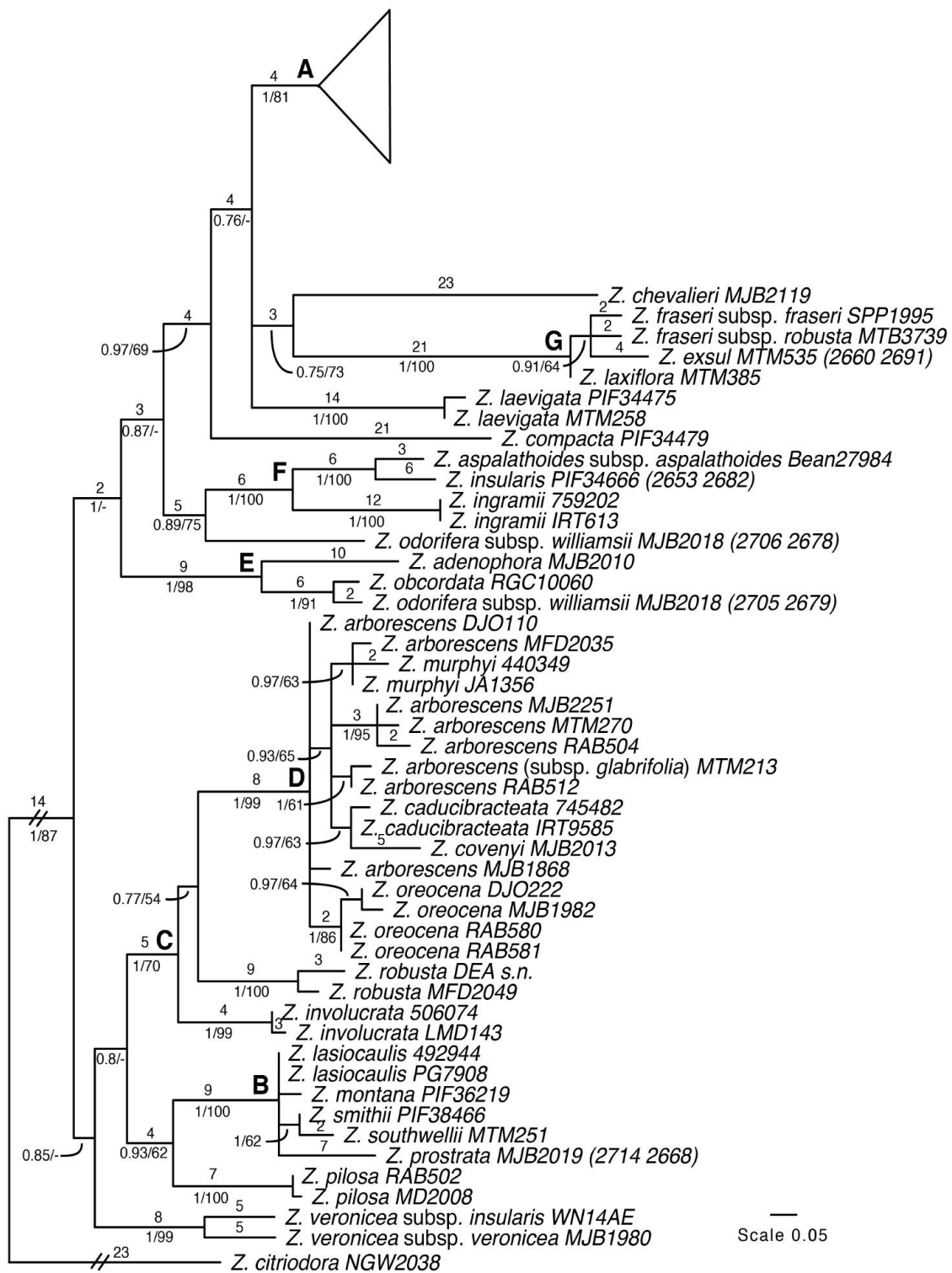


Figure 4.6. A section of the majority-rule consensus tree from the Bayesian inference analysis of the combined of ITS + ETS data. PP and BS values are below branches. Branch lengths from the maximum parsimony analysis have been mapped onto the tree above branches where relevant. Short branches without a label are one step in length. Parallel diagonal lines indicate shortening of branch lengths for illustrative purposes only. The scale bar represents the number of expected substitutions per site (BI analysis).

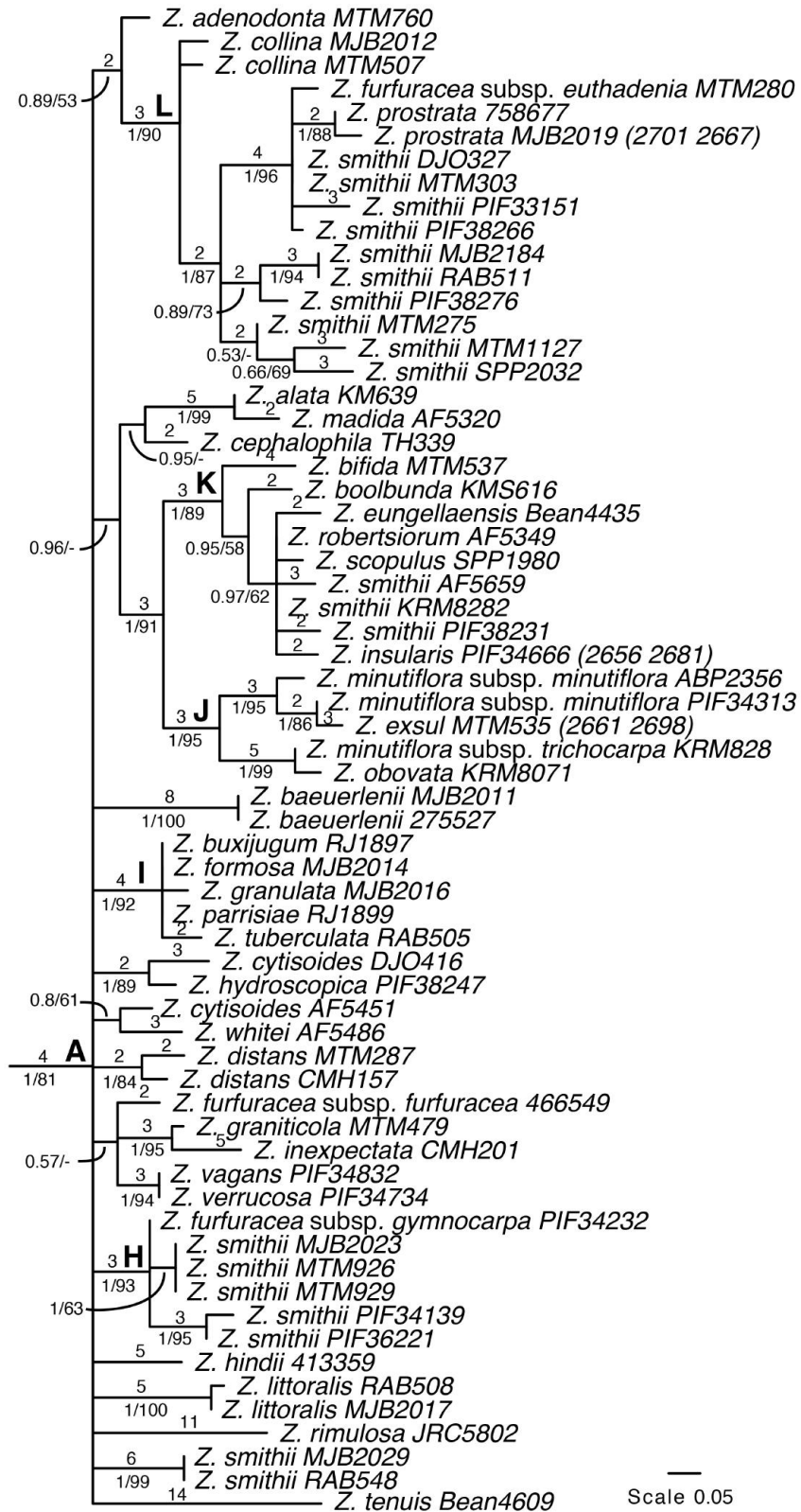


Figure 4.7. A section of the majority-rule consensus tree (Clade A) from the Bayesian inference analysis of the combined ITS + ETS data. PP and BS values are below branches. Branch lengths from the maximum parsimony analysis have been mapped onto the tree above branches where relevant. Short branches without a label are one step in length. The scale bar represents the number of expected substitutions per site.

The combined nrDNA markers provided a moderate level of phylogenetic resolution suitable for identifying some major lineages in the genus, and for comparisons with the phylogeny produced using chloroplast data (Chapter 3) and current taxonomy. In total, the BI tree resolved 68 nodes. *Zieria citriodora* was resolved as sister to a well-supported clade containing all other *Zieria* samples (Figure 4.5, node 1, indicated by a black arrow, PP 1, BS 87%). Above this node, support along the backbone of the tree was variable.

Monophyly of species and divergence of paralogues in combined ITS + ETS analysis

Of the species represented by two accessions in the combined ITS + ETS analysis (Figure 4.5), the following nine were resolved as monophyletic: *Z. veronicea* (node 3, two subspecies); *Z. pilosa* (node 6); *Z. involucrata* (node 10); *Z. robusta* (node 12); *Z. ingramii* (node 27); *Z. laevigata* (node 31); *Z. littoralis* (node 37); *Z. distans* (node 44); and *Z. baeuerlenii* (node 48). The only species represented by more than two accessions to be resolved as monophyletic was *Z. oreocena* (four accessions, node 14, PP 1, BS 86%). Monophyly of the following species was neither supported nor rejected, with accessions being part of unresolved polytomies that included other species: *Z. lasiocaulis* (node 7); *Z. caducibracteata* (node 17); *Z. murphyi* (node 20); *Z. fraseri* (two subspecies, node 34); and *Z. collina* (node 61). All other species with two or more accessions, from geographically separated populations, were polyphyletic or paraphyletic. These were: *Z. smithii* (21 samples, falling into at least six well supported clades throughout the tree); *Z. cytisoides* (two samples, one in a well-supported clade with *Z. hydroscopica*, node 46, and the other in a poorly supported clade with *Z. whitei*, node 45); *Z. furfuracea* (three samples, representing three subspecies, two in separate well supported clades, nodes 38 and 67, and one in an unsupported clade, node 41); *Z. minutiflora* (three samples, representing two subspecies, nested in a strongly supported clade, node 51, with *Z. obovata*, and one lineage of cloned sequences of *Z. exsul*); *Z. arborescens* (eight samples, nested in a strongly supported clade, node 13, with accessions of *Z. caducibracteata*, *Z. covenyi*, *Z. murphyi*, and the monophyletic subclade of *Z. oreocena* accessions).

As in the separate analyses of ITS and ETS data, there was clear evidence in the combined analysis that paralogues present within individual samples were divergent. Placement of divergent sequences from single accessions into different clades was mostly clear-cut and well supported by nodes with PP of 1 and BS $\geq$ 89%, e.g. *Z. prostrata* MJB2019, *Z. exsul* MTM535 and *Z. insularis* PIF34666 (Figure 4.5). Two ITS + ETS paralogues from *Z. odorifera* subsp. *williamsii* MJB2019 also fell in separate clades, but there was lower support for one of those clades (node 25, PP 0.89, BS 75% Figure 4.5.).

Comparison with chloroplast phylogeny of Zieria

The two nrDNA markers were nearly three times more variable than the four markers used in the cpDNA analysis (Chapter 3); 37.79% of the included nrDNA characters were variable among 109 samples, whereas only 14.12% of the cpDNA characters were variable among the same samples.

Most species of *Zieria* that were resolved as monophyletic in the analysis of combined ITS + ETS data were also resolved as monophyletic in the analysis of combined cpDNA data (Chapter 3); Figure 4.8 colour-codes the samples in the nrDNA tree according to their placement in the cpDNA tree. Notable exceptions were *Z. distans*, *Z. robusta* and *Z. oreocena*, which were all resolved as polyphyletic in the cpDNA analysis, and *Z. caducibracteata*, *Z. collina* and *Z. murphyi*, which are potentially monophyletic based on nrDNA analysis, but are moderately to strongly supported as not monophyletic based on cpDNA. Also, based on nrDNA, *Z. arborescens* (as currently circumscribed) is paraphyletic, although all the accessions were resolved within clade D, and it was well supported. Embedded within *Z. arborescens* were accessions of *Z. murphyi*, *Z. caducibracteata* and *Z. covenyi*, although the subclades containing these species were variously supported (nodes 17 and 20, both with PP 0.97 and BS 63%). In contrast, based on cpDNA *Z. arborescens* was strongly resolved as polyphyletic (Chapter 3).

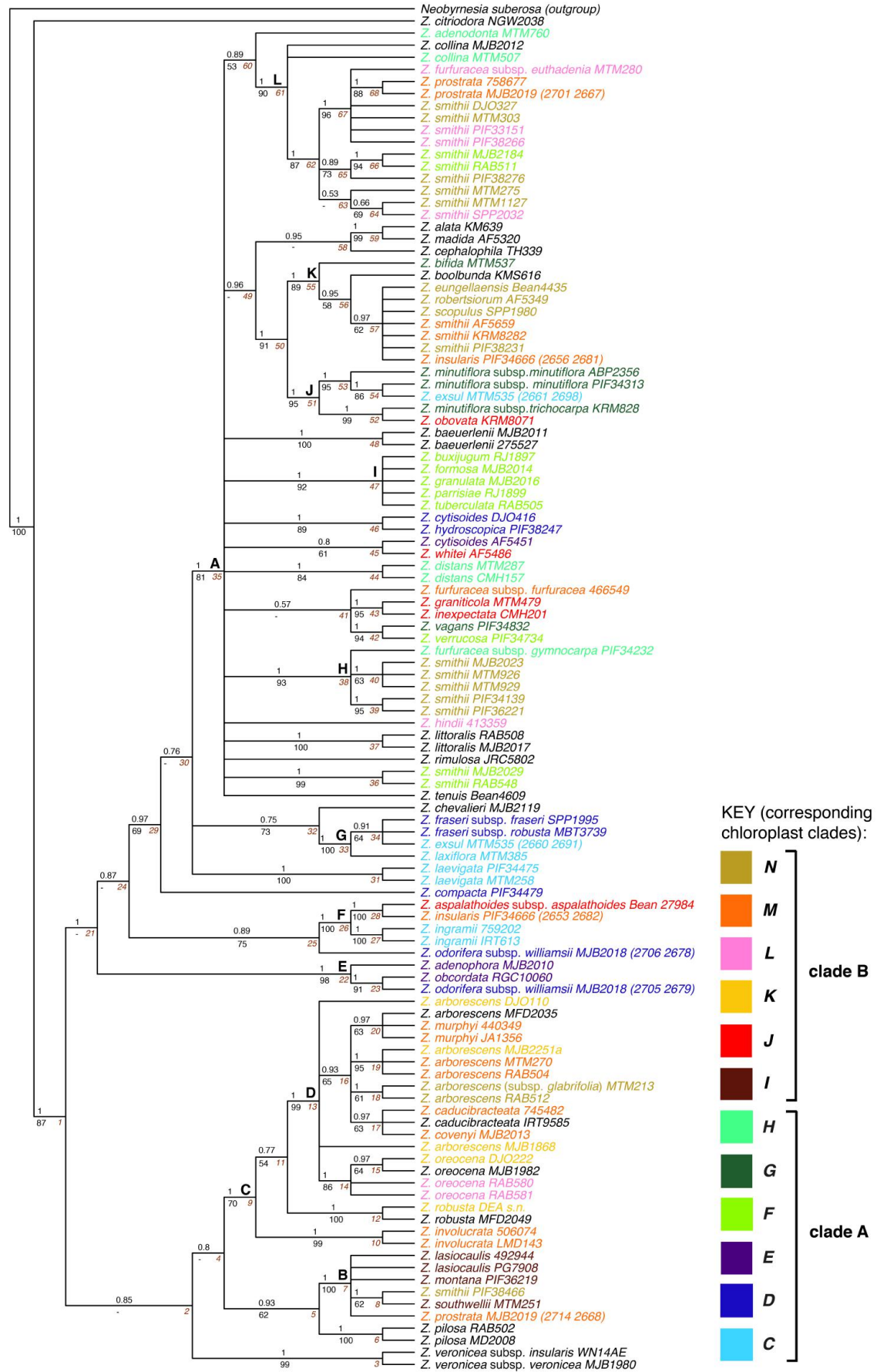
On the whole the phylogeny based on nrDNA was highly incongruent with the phylogeny based on cpDNA (Chapter 3). Of 53 highly supported nodes in the nrDNA tree (i.e. with PP $\geq$ 0.95), only 13 of them (24.5%) were congruent with highly supported nodes ($\geq$ 0.95 PP) in the cpDNA tree (see Appendix D), and 75.5% of the nodes were highly supported as in conflict. Furthermore, in 11 (of 13) instances the

congruent nodes consisted of two or three accessions of a single species. Only two non-conflicted nrDNA nodes contained multiple species: node 43, consisting of *Z. graniticola* and *Z. inexpectata* (with good PP support but less than 50% BS support for the cpDNA node); and node 47, consisting of *Z. buxijugum*, *Z. formosa*, *Z. granulata*, *Z. parrisiae* and *Z. tuberculata*.

Comparison with the morphological phylogeny of Armstrong (2002)

Results of the combined ITS+ ETS analysis shows partial congruence with the morphological and chemical analysis of Armstrong (2002), as well as some noteworthy mismatches (Figures 4.9 and 4.10). Most nrDNA clades contained members of multiple morphological groups and none of Armstrong's morphological clades was identified as monophyletic. Five nrDNA clades (F, G, I, J, and K; Figure 4.10) each contained samples from single morphological groups (based on Armstrong's phylogeny reproduced in Figure 4.9), but none contained all members of any one group; some also contained taxa that were unplaced by morphological analyses because they had not been described at the time of Armstrong's (2002) work.

Following page. Figure 4.8. Bayesian inference majority-rule consensus tree produced from combined ITS + ETS data with samples colour-coded according to major cpDNA clades (see Figures 3 and 5 in Chapter 3) which in this figure are labelled in italics in the key.



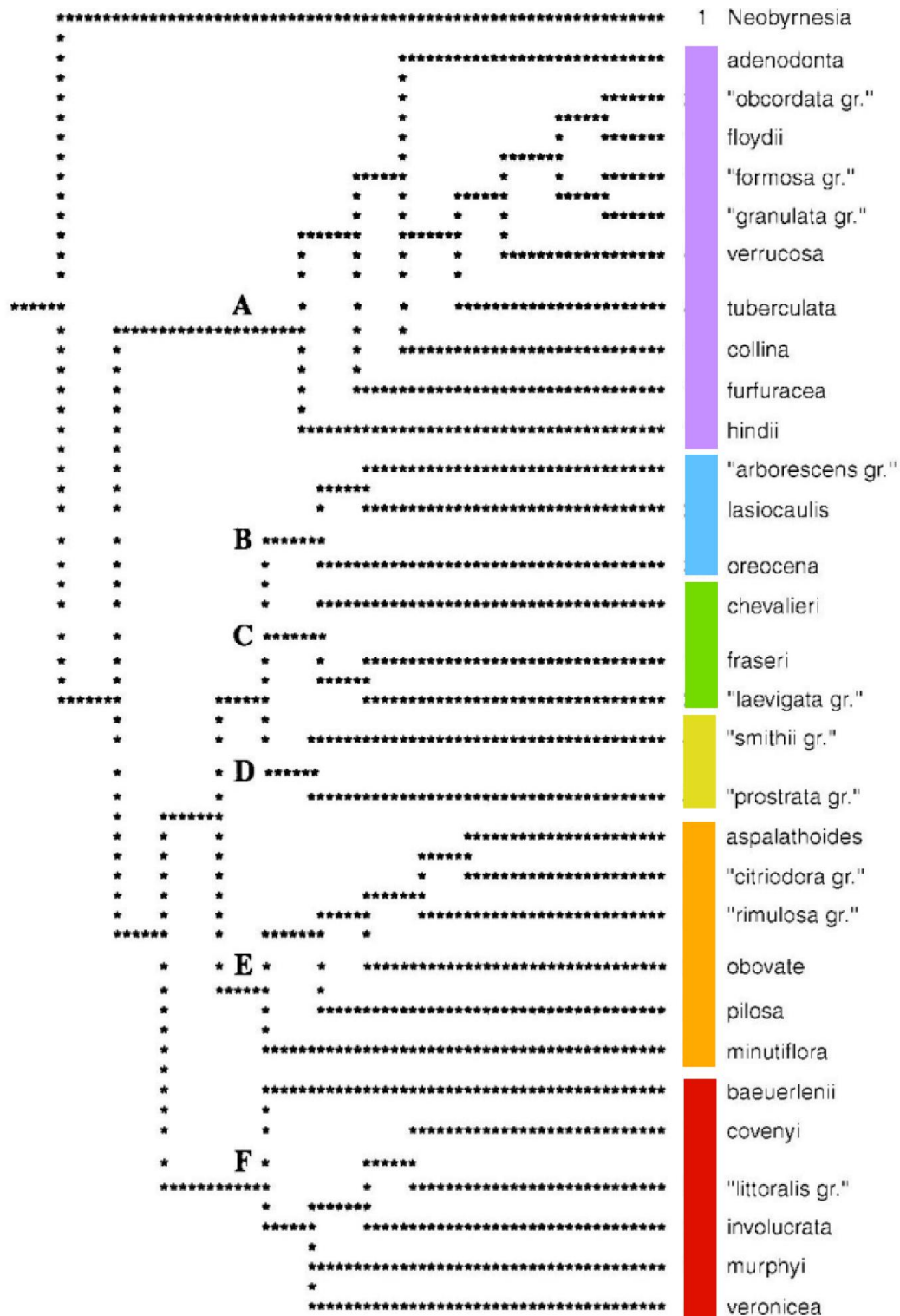
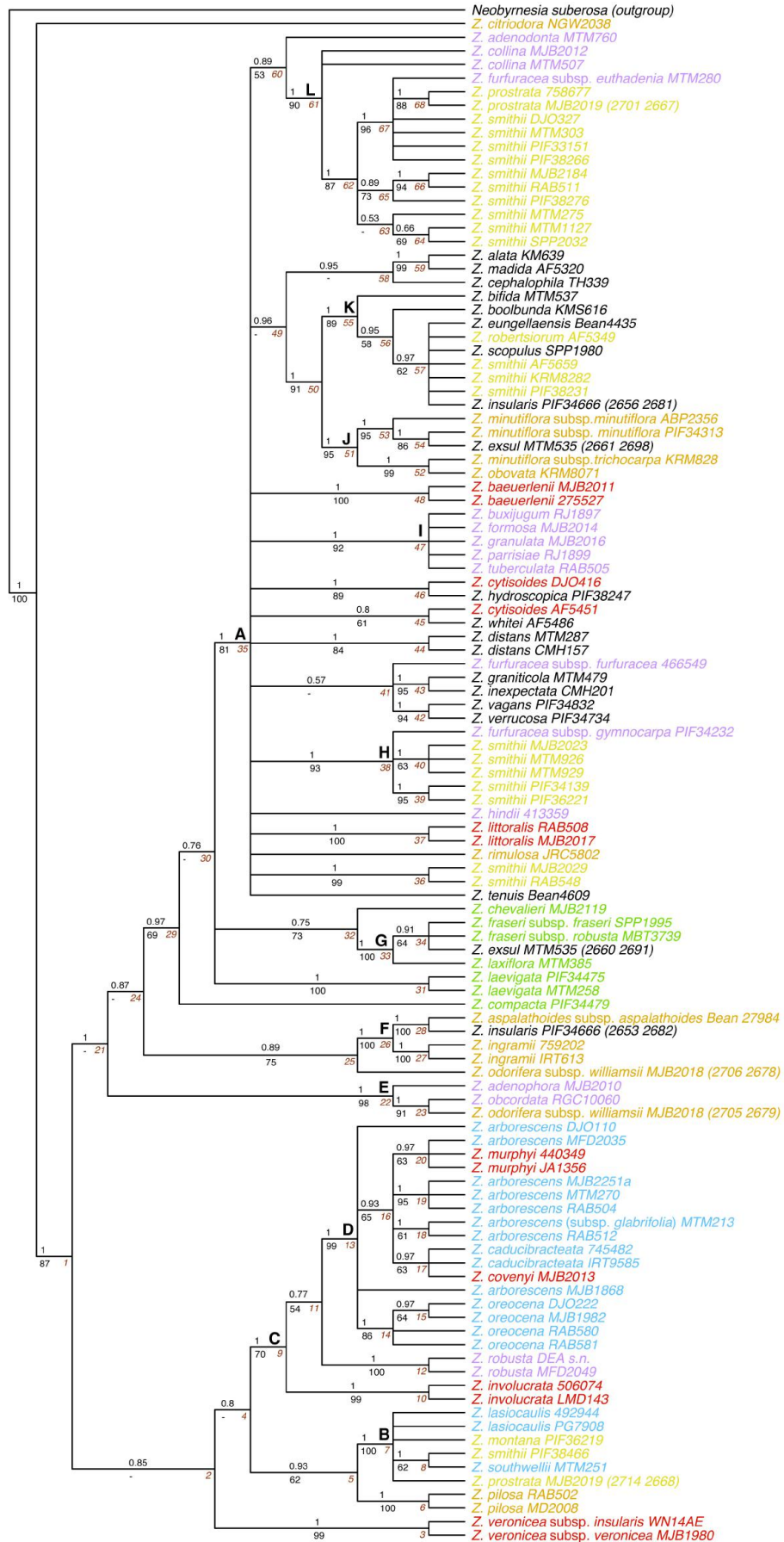


Figure 4.9. Phylogeny proposed by Armstrong (2002), based on morphological and chemical characters. Species within groups, labelled 'littoralis gr.' for example were not defined by Armstrong.

Following page. Figure 4.10. Bayesian inference majority-rule consensus tree produced by analysis of combined ITS + ETS data (as in Figure 4.5) reproduced here with taxa colour-coded according to the groups defined in the phylogeny of Armstrong (2002), as shown above in Figure 4.9. Taxa shown in grey were not included in Armstrong's (2002) analysis. PP values are shown above the branches and BS values below.

Nuclear phylogeny of *Zieria*



4.4. Discussion

Significance of divergent nrDNA paralogues

The results of this study showed strong evidence that divergent copies of nrDNA, which fell in different clades on the gene trees, are present in the genomes of some individuals of *Zieria*. This is not an uncommon phenomenon in plants (Bailey *et al.* 2003; Bayly and Ladiges 2007; Bayly *et al.* 2008; Poczai and Hyvönen 2010; Holmes *et al.* 2014), but presents challenges for phylogeny reconstruction (Buckler *et al.* 1996; Fitch 1970; Sanderson and Doyle 1992). There is an extensive literature on nrDNA paralogues in plants that has mostly focussed on ITS regions (Baldwin *et al.* 1995; Samuel *et al.* 1997; Mayol and Rosselló 2000; Feliner and Rosselló 2007). There has also been a logical inference that paralogues of ITS would be accompanied by corresponding paralogues of ETS (Poczai and Hyvönen 2010; Potts *et al.* 2013), which is part of the same nrDNA cistron, but few studies have discussed or demonstrated such paralogues (examples include King and Roalson 2008; Fehrer *et al.* 2009; Poczai and Hyvonen 2010). This study found that individuals with divergent copies of ITS clearly have corresponding divergent copies of ETS.

In most plants the 18S–5.8S–28S nrDNA cistron occurs in tandem arrays composed of hundreds to thousands of copies, and arrays can be situated at one or more chromosomal loci. Such multi-copy gene families commonly undergo ‘concerted evolution’ (reviewed in Nei and Rooney 2005; Poczai and Hyvonen 2010), in which all the member genes evolve as a unit (‘in concert’). This tends to homogenize the reiterated sequences within and between loci, conferring a degree of intergenomic uniformity to sequences that would otherwise accumulate different mutations (Zimmer *et al.* 1980; Arnheim 1983; Sanderson and Doyle 1992); it can result in a single predominant sequence across all copies and arrays. However, if concerted evolution is incomplete or slower than the rate of mutation (Childs *et al.* 1981; Baldwin *et al.* 1995) nrDNA copies can become divergent, and sometimes this leads to highly divergent, non-functional pseudogenes (Bailey *et al.* 2003; Razafimandimbison *et al.* 2004; Bayly and Ladiges 2007; Ochieng *et al.* 2007; Bayly *et al.* 2008). Studies of diploid species have shown bias in copy number towards one homeologue over another (Rauscher *et al.* 2002, 2004) sometimes as the result of PCR selection and PCR drift (Wagner *et al.* 1994), or because of an

imbalance in the relative copy number within the genome. Variation in nrDNA within genomes can also arise through hybridisation between species and introgression (Wagner 1983; Rieseberg and Wendel 1993; Soltis and Soltis 2009) producing novel combinations of nrDNA copies within genomes. These copies can variously be retained in progeny, can recombine, or be homogenised to one dominant sequence, representing either of the original parental forms, or an intermediate form (Álvarez and Wendel 2003). Divergent nrDNA copies arising through incomplete concerted evolution can be classified as ‘shallow’ or ‘deep’ paralogues (sensu Bailey *et al.* 2003) depending on whether they arise within extant species (in which case nrDNA copies within a species will form a monophyletic group on gene trees) or prior to speciation (in which case nrDNA copies within a species can be polyphyletic on gene trees).

In the case of *Zieria* the genetic history of the divergent paralogues of nrDNA is not clear, especially in the absence of more information about the genomic arrangement of the nrDNA arrays; it is not known whether highly divergent copies are located within the same or in different arrays. Also there is no strong evidence that the divergent paralogues represent pseudogenes or that they have undergone recombination. It is not clear whether deeply divergent paralogues that fall in different clades on the gene trees (e.g. those of *Z. prostrata* MJB2019, Figure 4.5) represent: (a) deep paralogues, predating speciation events, that have been variably retained in genomes of some species, or variably detected by sequencing; or (b) divergent sequences incorporated into individual genomes through hybridisation of species that are in different clades of the genus.

One example of a potential influence of introgression on nrDNA diversity is seen in *Z. exsul* (MTM535). Two paralogues of ITS and ETS were recovered in this sample. One grouped strongly in clade G with *Z. laxiflora* and *Z. fraseri*, with which *Z. exsul* has some morphological affinity¹. The other paralogue (in clade J) differed by only one or two characters from samples of *Z. minutiflora* subsp. *minutiflora*, from which it is more morphologically distinct, but with which it co-occurs (Duretto and Forster 2007). It is plausible that the paralogue in clade G is more indicative of species relationships whereas the one in clade J reflects local introgression with *Z. minutiflora*.

¹ This is based on the morphological similarity of *Z. exsul* and *Z. compacta*, as described by Duretto and Forster (2007). *Zieria compacta* is a reinstated species (Duretto and Forster 2007) but was previously within the circumscription of *Z. fraseri* and morphologically very similar.

The presence of divergent paralogues such as these within individual genomes creates difficulties for phylogenetic reconstructions because accurate phylogenies require a comparison of orthologous sequences (Fitch 1970), and failure to distinguish orthologues from paralogues frequently results in misleading results following phylogenetic analysis (Álvarez and Wendel 2003). If deep paralogues or pseudogenes of nrDNA are unknowingly compared, the phylogeny will be confounded, producing a gene tree that does not accurately reflect the species tree (Sanderson and Doyle 1992).

In *Zieria* the observed pattern of nrDNA variation makes phylogenetic inferences problematic without knowing the history of divergent sequences in individuals. It is also highly likely that the nrDNA data used in this study provide only limited insight into the extent of nrDNA variation within individuals. Different PCR conditions can influence which paralogues are amplified (Buckler *et al.* 1997; Mayol and Rosselló 2001). This was exemplified here by an accession of *Z. prostrata* (758677) in which direct sequencing using universal primers amplified what appeared to be a ‘clean’ ITS sequence (no polymorphisms), and the presence of divergent paralogues of ITS was only detected in the sample following the subsequent use of copy-specific primers. This raises the question as to how many other samples contained paralogues of ITS or ETS that were not detected using routine sequencing. One solution to this potential problem is to clone all nrDNA PCR products in the expectation that all copies will be detected, even those that occur in very small numbers. Cloning and sequencing of all PCR products, although desirable, can be prohibitively expensive and labour-intensive (Baldwin *et al.* 1995), and was beyond the scope of this study due to the large number of samples. Next-generation sequencing of genomic DNA libraries or rDNA amplicons could provide greater insight into nrDNA diversity within genomes (Straub *et al.* 2012; Weitemier *et al.* 2014) at substantially lower costs.

Another issue highlighted by this study is the need for caution in combining ITS and ETS sequences in analyses. These markers are routinely combined without careful consideration of homology (Bena *et al.* 1997; Baldwin and Markos 1998; Clevinger and Panero 2000; Wright *et al.* 2001; Appelhans *et al.* 2014), i.e., without considering whether the combined ITS and ETS sequences are orthologous. This is especially problematic in groups for which diversity has been detected within individuals in either ITS or ETS, but not in both regions (Suárez-Santiago *et al.* 2006; Garcia-Jacas *et al.*

2009; Parra-O. *et al.* 2009; Gibbs *et al.* 2009; Ladiges *et al.* 2010). These two markers are typically amplified using separate PCRs and it should not be assumed that they are necessarily orthologous. Standard tests for incongruence between markers (e.g. ILD test or partition homogeneity test; Farris *et al.* 1995a, 1995b) are especially warranted before combining ITS and ETS data. Where divergent paralogues of both markers are recovered, the method used here to identify orthologous sequences to combine for analysis is a reasonable one.

Comparison with cpDNA variation

This study provides clear evidence of conflict between phylogenetic reconstructions based on nrDNA and cpDNA data (Chapter 3), a phenomenon that is relatively widespread in plants (Soltis *et al.* 1996; Wendel and Doyle 1998; Othman *et al.* 2010; Xu *et al.* 2012; Hong-Wa and Besnard 2013; Yu *et al.* 2013; Li *et al.* 2014). A number of processes could explain the causes of this incongruence including paralogy in nrDNA, effects of hybridisation on nrDNA or cpDNA, incomplete lineage sorting (which can be more apparent in cpDNA because of lower mutation rates), allopolyploidy, and sequence homoplasy. In the phylogenetic study of *Zieria* based on cpDNA (Chapter 3) it was inferred that any or all of incomplete lineage sorting, regional cpDNA introgression (chloroplast capture) or inappropriate taxonomic boundaries could account for the incongruence between cpDNA and current taxonomy. Comparison of both nuclear and chloroplast data here potentially allows further insight into the causes of the incongruence between the datasets in *Zieria*.

Notwithstanding the issues regarding paralogy, the nrDNA potentially provided a better indication of phylogeny in *Zieria* than cpDNA. The nrDNA phylogeny, at least for some taxa, was more consistent with species-level taxonomy than the cpDNA phylogeny. Key examples are *Z. distans*, *Z. oreocena* and *Z. robusta*, which are monophyletic in the nrDNA tree but are strongly supported as polyphyletic or paraphyletic in the cpDNA tree. Likewise, *Z. caducibracteata*, *Z. collina* and *Z. murphyi* are potentially monophyletic based on the nrDNA tree but are moderately to strongly supported as not monophyletic in the cpDNA tree. Another notable example is *Z. arborescens*, which in the nrDNA tree is paraphyletic with respect to *Z. murphyi*, *Z. caducibracteata* and *Z. covenyi*, although all samples group in the same clade (D)

with strong support. This contrasts markedly with the cpDNA tree in which *Z. arborescens* is strongly supported as polyphyletic, falling in at least five different clades.

Given the level of congruence between nrDNA and current species-level taxonomy in some taxa (e.g. *Z. arborescens*, *Z. caducibracteata*, *Z. collina*, *Z. distans*, *Z. murphyi*, *Z. oreocena* and *Z. robusta*) some inferences about the processes responsible for the widespread incongruence between nrDNA and cpDNA in *Zieria* can be made. In these cases incomplete lineage sorting, rather than introgressions of cpDNA, seems to best explain the pattern. For example, given the monophyly of *Z. oreocena* in the nrDNA tree, placement of some chloroplast haplotypes of *Z. oreocena* in a clade with *Z. smithii*, *Z. furfuracea* subsp. *euthadenia* and *Z. hindii* (Chapter 3, clade L) is likely to be the result of incomplete lineage sorting because there was no evidence of introgression of nrDNA in these species and introgression is highly unlikely over the very large geographic distance separating them (more than 1500 km). A similar pattern was seen in *Z. arborescens*, in which one sample (RAB512) from the Otways, Victoria, was placed in a chloroplast clade (Chapter 3, clade N) with samples of *Z. actites*, *Z. arborescens* (subsp. *glabrifolia*), *Z. eungellaensis*, *Z. robertsiorum*, *Z. scopulus* and *Z. smithii* from Queensland. Again, hybridisation over such distances is highly unlikely and there is no evidence of nrDNA introgression on the basis of the position of *Z. arborescens* (RAB512) in the nrDNA tree. In both these cases the cpDNA divergences within species are quite deep, suggesting retention of divergent cpDNA lineages within species over extended periods of time.

The inference of incomplete lineage sorting of cpDNA in *some* taxa, including those showing deep divergences, could support the same process as a plausible explanation for much of the incongruence detected. In the majority of cases, however, the underlying causes are less clear than in the examples described above and are not discussed here in detail, but a few points of interest and inferences for some taxa are discussed in Chapter 6 (General Discussion, section 6.2, 'Implications for species-level taxonomy').

Biogeography: relationships of the New Caledonian species, Zieria chevalieri

Placement of the single endemic New Caledonian species, *Z. chevalieri*, is of biogeographic interest, and its placement differs between cpDNA and nrDNA trees. The cpDNA phylogeny (Chapter 3) placed *Z. chevalieri*, with strong support (PP 1.0, BS 100%), as sister to all other species in the genus, suggesting an early divergence of New Caledonia and Australia, as might be expected from a vicariant pattern of differentiation. In contrast, combined ITS + ETS phylogeny showed the deepest divergence in *Zieria*, with good support (PP 1, BS 87%), to be that between *Z. citriodora* (from north eastern Victoria near the New South Wales border) and all other taxa; *Z. chevalieri* was nested higher in the tree. The nested position of *Z. chevalieri* in the nrDNA tree could be seen as consistent with more recent dispersal, rather than older vicariance, between Australia and New Caledonia, but this pattern alone does not favour one hypothesis over another. Inferences hinge on interpretations of the age of divergence, the geological history of New Caledonia and surrounding areas, and the likely dispersal mechanisms, as discussed in Chapter 3.

The combined ITS + ETS tree (Figure 4.5.), showed weak support (PP 0.75, BS 73%) for *Z. chevalieri* as sister to a clade consisting of *Z. fraseri*, *Z. exsul* and *Z. laxiflora*. Together with *Z. laevigata* and *Z. compacta*, these six taxa form a grade that subtends the more species-rich Clade A. Armstrong (2002) also grouped *Z. chevalieri* with *Z. fraseri*, *Z. laxiflora* and *Z. laevigata* based on a common set of morphological characters: distinctly ridged branches, velutinous lower leaf surfaces, inflexed calyx lobes and apiculate anthers. *Zieria exsul* and *Z. compacta*, previously included under Armstrong's (2002) broad circumscription of *Z. fraseri*, have similar morphology (Duretto and Forster 2002).

However, of note is that, although similar in morphology, *Z. chevalieri* differs from any members of its sister clade by at least 44 base changes in the nrDNA MP reconstructions; Figure 4.6 shows the long branches leading to *Z. chevalieri* and its sister clade. It is therefore conceivable that the position of *Z. chevalieri* in this tree might be affected by 'long branch attraction' (Bergsten 2005). Alternatively, the long branch of *Z. chevalieri* might reflect a relatively long period of isolation from morphologically similar Australian relatives.

Given the congruence with morphology, it seems likely that the position of *Z. chevalieri* in the nrDNA tree gives a better indication of its phylogenetic relationships than the cpDNA tree (Chapter 3) – at least that it is not sister to all other species in the genus. If this is true, as suggested above, the differing cpDNA pattern could relate either to chloroplast lineage sorting, or chloroplast introgression among the Australian species. If introgression among Australian taxa has been widespread since the geographic isolation of a New Caledonian lineage, it could conceivably result in *Z. chevalieri* occupying a more isolated position in the cpDNA tree. Even so, it would be highly speculative to attempt to interpret the history of the biogeographic connection of the New Caledonian and Australian taxa on the basis of the nrDNA phylogeny presented here. How early *Z. chevalieri* and its related species diverged within *Zieria* remains equivocal, given the weak or absent support for nodes in the backbone of the tree (Figure 4.5, nodes 21, 24, 30). Although Bayly *et al.* (2013) estimated the divergence between *Z. chevalieri* and Australian taxa to have occurred (6.5)–11.9–(17.9) million years ago, they did so on the basis of a cpDNA tree. It is also unwise to emphasise these molecular divergence estimates as a basis for confirming or rejecting hypotheses of dispersal or vicariance in *Zieria* because large parts of the Rutaceae phylogenetic tree are not supported by good fossil calibrations, including in *Zieria*, and could be prone to substantial underestimates of clade ages (Bayly *et al.* 2013). Furthermore over-water, long-distance seed dispersal is considered highly improbable for *Zieria*, which has ant-dispersed seeds (Armstrong 2002).

4.5. Concluding remarks

The nrDNA tree based on combined ITS + ETS data presented in this chapter provided greater support for monophyly of species than the cpDNA data in Chapter 3, and potentially provides a hypothesis for a species tree in *Zieria*. However, relationships are complex in *Zieria*, with widespread incongruence revealed between the nuclear and chloroplast DNA gene trees and existing species-level taxonomy. Results were further confounded by the presence of deeply divergent paralogues of nrDNA in some taxa, which could be the result of hybridisation, or represent allelic variants that have existed since before speciation.

The study highlights the need for caution when combining datasets; issues of paralogy must be carefully considered. Nevertheless, ITS and ETS regions have been shown to produce more robust infrageneric phylogenies when combined, than when each region is analysed in isolation (Baldwin and Markos 1998, Bena *et al.* 1998, Clevinger and Panero 2000). In this study, there was a reasonable amount of resolution obtained by the concatenation of ITS and ETS sequence data, which ultimately produced a single tree with improved resolution and support for nodes, over the two poorly-resolved trees which resulted from separate analysis of the datasets. It is nevertheless challenging to draw definitive conclusions about many species relationships, especially for species represented by a single accession. Taxonomic issues arising from the results of this analysis are discussed in detail in Chapter 6 (General Discussion).

Future research should include multiple samples of all taxa, and bacterial cloning of PCR products is recommended where possible, given that divergent paralogues can remain undetected using routine PCR methods and primers, as shown in this study by the sample *Z. prostrata* 758677. Some next-generation sequencing approaches would also allow for identification of the divergent paralogues.

Chapter 5. Phylogeography of *Zieria arborescens*

5.1. Introduction

This chapter presents a phylogeographic study of the mesic forest understory species *Zieria arborescens*. Phylogeography uses genetic data and biogeographic patterns to study populations, and make inferences about the temporal structure of those populations (Avice 2000, 2009; Beheregaray 2008). The historical examination of genetic lineages contributes to our understanding of the biology of the species, the ecological and evolutionary processes that have shaped them, and issues concerning conservation (Beheregaray 2008; reviewed in Byrne *et al.* 2011). In this study both chloroplast and nuclear DNA sequence data were used to identify genetic lineages of *Z. arborescens* and investigate geographic distribution patterns for comparison with studies of other mesic forest taxa of south eastern Australia.

Temperature and rainfall are key factors that shape vegetation. The Australian mesic biome, thought to be the ancestral biome of the continent (Byrne *et al.* 2011), has undergone fragmentation brought about by climate change (increased aridification in particular) over the past 50 million years as Australia drifted north away from the rest of the former supercontinent, Gondwana (Barlow 1981). There has been a resulting decline of rainforests and expansion of sclerophyllous flora adapted to drier habitats (Hill *et al.* 1999; Hill 2004). Although Australia is currently dominated by arid and semi-arid habitats, flora of the mesic zone persists, including rainforest and tall open forest taxa.

5.2. Quaternary climatic fluctuations in south-eastern Australia

Australia experienced the world-wide climatic fluctuations of the Quaternary (2.5 million years ago to the present). Repeated glacial periods, most extensive in Tasmania (McKinnon *et al.* 2004), subjected the biota to cool, arid, harsh conditions. In addition, during the Quaternary (as well as during earlier times from the early Miocene; Baillie 1989) successive periods of low sea-level resulted in land bridges connecting Tasmania to the Australian mainland at various times, as recently as 7,000 years ago (Jackson 1999). As a consequence of these repeated connections, south-eastern mainland Australia and Tasmania share similar vegetation including mesic forests, which typically occur in regions of high rainfall (i.e., above 800 mm).

During the Last Glacial Maximum (LGM), which occurred in south eastern Australia about 18,000 years ago, glaciers were at their thickest, sea levels were at their lowest, temperatures were 6–6.5° C cooler than today (Colhoun *et al.* 1996), and rainfall was much reduced compared to today, especially in the summer (Hubbard 1995). Climatic conditions during the LGM were probably the least favourable for mesic forest growth than at any time over the past 2 million years (Hope 1994). A combination of greater aridity and lower temperatures than at present were key factors shaping vegetation during that time, and parts of south-eastern Australia were mostly treeless (Hope 1994; Hill, 2004).

During the LGM, ice sheet cover was not extensive in south-eastern Australia (compared to that which affected Europe and North America), being confined to the Central Plateau of Tasmania (Colhoun *et al.* 1996) and small areas in the Snowy Mountains on mainland Australia (Barrows *et al.* 2001). Consequently, there was no large-scale removal of vegetation during the LGM and many south-eastern Australian forest species that had contracted in distribution were able to recover rapidly when conditions became more favourable (Byrne *et al.* 2008; McKinnon *et al.* 2004). One model suggests that in the southern hemisphere recolonization of the landscape occurred by expansion over short distances from multiple ‘micro-refugia’ that escaped major climate changes during the glaciation period (Byrne *et al.* 2008; MacPhail and Colhoun 1985; Nevill *et al.* 2010) rather than dispersal over vast distances from only a few distinct refugia as occurred in many temperate and boreal tree species of the northern hemisphere (Jackson and Overpeck 2000). Likely micro-refugia in south-eastern Australia included some coastal areas, the exposed continental shelf, suitable microsites within larger less suitable areas, or sites at higher elevations (McKinnon *et al.* 2004). In Victoria putative refugia were likely to have existed in the Otway Ranges (‘the Otways’), the south Central Highlands, South Gippsland and East Gippsland, and in Tasmania, in the west, south-east, east and north-east (Nevill *et al.* 2014). Evidence supporting the identification of putative refugia for forest species during glacial maxima in south-eastern Australia comes from the fossil pollen record, climate modeling of species distributions, and a limited number of molecular studies (Steane *et al.* 1998; Jackson *et al.* 1999; Freeman *et al.* 2001; McKinnon *et al.* 2001, 2004; Worth *et al.* 2009, 2010, 2011; Nevill *et al.* 2010, 2014; Pollock *et al.* 2013). Genetic signatures indicating that species might have

persisted in past refugia include extant populations with high levels of genetic diversity or with low levels of diversity but unique divergent haplotypes (Nevill *et al.* 2010). Phylogenetic analyses of eastern Australian species that are widespread and morphologically variable have the potential to test models of post-glacial recovery of southern hemisphere mesic forest taxa (Beheregaray 2008). *Zieria arborescens* is one such species.

5.3. *Zieria arborescens*

Zieria arborescens is a morphologically variable small tree or shrub that typically grows in the understorey of wet forests of eastern Australia ranging from south-eastern Queensland to Tasmania (Figure 5.1).

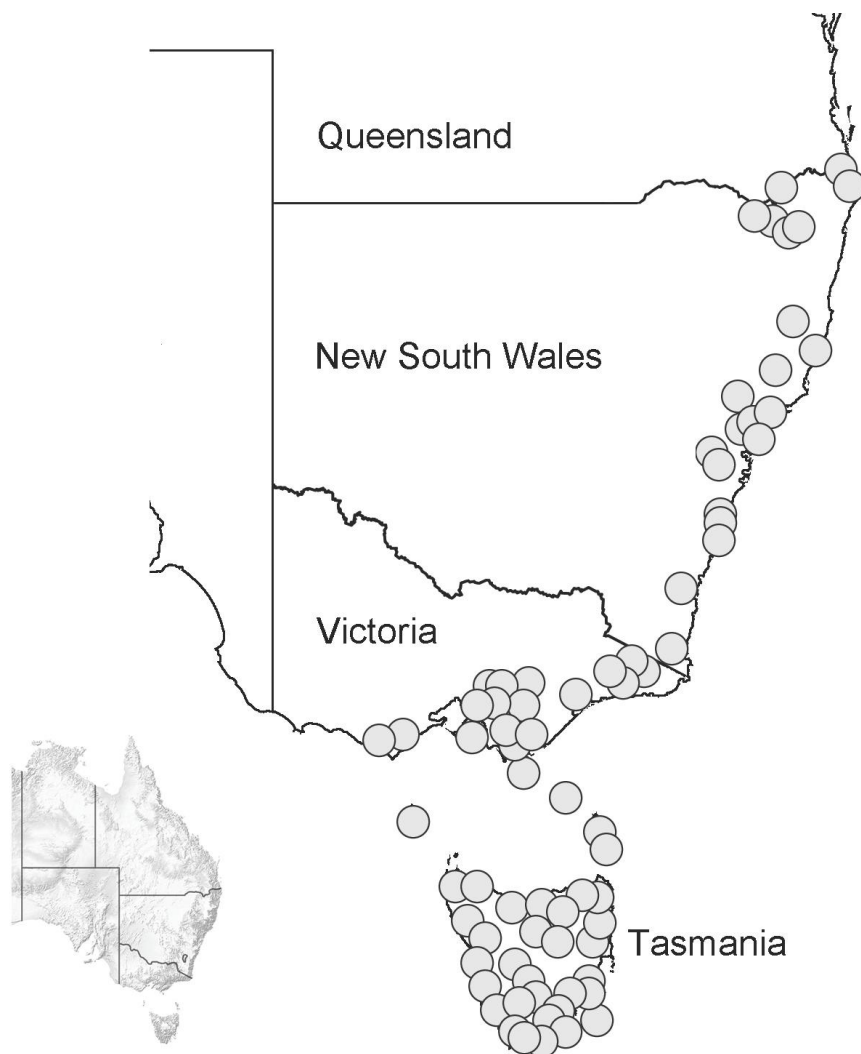


Figure 5.1 Distribution of *Zieria arborescens* in south-eastern Australia. Today Tasmania is separated from the Australian mainland by Bass Strait, a stretch of relatively shallow water (generally less than 100 m deep, and currently about 250 km wide). *Zieria arborescens* occurs on islands in Bass Strait.

It is the most common species of *Zieria* to occur in Victoria and Tasmania (including islands in Bass Strait), and is one of only a few relatively widespread species of *Zieria* that occur in New South Wales. In Queensland it occurs only in the extreme southeast corner of the state.



Figure 5.2. *Zieria arborescens* (RAB524) in the Dandenong Ranges, east of Melbourne, Victoria.

Zieria arborescens (Figure 5.2) is characterised by the following combination of characters: shrub or tree-like habit (up to 10 m); branches slightly to distinctly ridged with decurrent leaf bases, stellate pubescent (or more rarely glabrescent), and not glandular-verrucose; leaves relatively large and broad (i.e., central leaflet 4–22 mm wide and 50–90 mm long, but sometimes much larger); petioles 8–40 mm long; leaf surfaces not glandular-verrucose; leaf upper surfaces with few hairs; leaf under surfaces hirsute to densely stellate-tomentose or glabrous with a midrib not glandular-verrucose; prominent secondary venation (more obvious on the lower leaf surface of hirsute specimens); inflorescences with around 120–500 flowers, usually shorter than

the leaves; caducous floral bracts that are foliaceous, but usually much smaller than leaflets (12–20 mm long); ridged pedicels pubescent with short scattered stellate hairs (becoming densely tomentose in between ridges); ovate to lanceolate calyx lobes 1.3–2.2 mm long and 1.2–1.8 mm wide; filaments dilated towards the base; fruits glabrous (Armstrong 2002; George *et al.* 2013). Flowers are pollinated by nectar-seeking dipterans (Armstrong 2002), and as with all *Zieria* species, the seeds of *Z. arborescens* are forcibly ejected from the dehiscent coccus of the fruit, and commonly dispersed by ants with the aid of ant-attracting eliasomes (Armstrong 2002). *Zieria arborescens* is thought to regenerate after fire from seeds that persist in seedbanks in the soil (Floyd 1976).

Throughout its distributional range (Figure 5.1) *Z. arborescens* displays considerable variation in leaf size and shape and in the degree of pubescence of the indumentum on the lower surface of the leaflets. Juvenile leaves are not particularly distinctive or overly large (R.A. Barrett pers. obs.). Figure 5.3 shows leaves (at the same scale) from different populations; the leaflets of the population that regenerated following severe bushfires in early 2009 in the Cumberland Falls area, near Marysville, Victoria, are much larger than is typical, and outside the range described by Armstrong (2002) and George *et al.* (2013), whereas the leaflets of the Phillip Island population are quite small (although within the described range). Leaflets with glabrous under-surfaces have been observed in specimens throughout Tasmania and Victoria, in south-eastern New South Wales and in restricted populations in the Darling Downs District of south-eastern Queensland (there recognised as subsp. *glabrifolia*, see below); those with a dense indumentum (as in the Grand Ridge Road population shown in Figure 5.3) are found throughout the species range (Armstrong 2002, R.A. Barrett pers. obs.).



Figure 5.3. Herbarium specimens of *Zieria arborescens* (all Victorian) showing examples of leaf variation. Clockwise from top: RAB554b from Cumberland Falls walk, eastern Victoria (very large leaves); RAB569a from Grand Ridge Road, Gippsland (pronounced discolourous leaves, due to the dense indumentum on the underside of leaflets); RAB520 from Phillip Island (smaller leaves).

Armstrong (2002) recognised three subspecies of *Z. arborescens*: (1) subsp. *arborescens*, which is widespread and occurs in all the states in eastern Australia; (2) subsp. *glabrifolia*, which Armstrong considered to be endemic to Queensland, but which occurs in south east Queensland and northern New South Wales following the circumscriptions of Duretto and Forster (2007) and George *et al.* (2013); (3) subsp. *decurrens*, which is restricted to the Caves Beach area of Jervis Bay in New South Wales (Armstrong 2002; George *et al.* 2013). The key to the subspecies provided by Armstrong (2002) does not work with material from outside Queensland (Duretto and Forster 2007) and the characters that distinguish the subspecies are not clear (Duretto and Forster 2007; George *et al.* 2013; R.A. Barrett pers. obs.). In the current study, as elsewhere in this Ph.D. project, samples of *Z. arborescens* were not identified to subspecies, but accessions collected near type localities for the two segregate subspecies are noted in brackets in the text and in figures and tables, i.e., (subsp. *glabrifolia*) and (subsp. *decurrens*).

5.4. Aims

This study aimed to identify major evolutionary divergences or disjunctions across the distributional range of *Z. arborescens* with a particular focus on the biogeographic connections between Victoria and Tasmania. Reasons for focussing on the southern distributional range were twofold. Firstly, it was intended to place the phylogeography of *Z. arborescens* in the context of similar studies on wet forest taxa within a similar geographic range, such as eucalypts (McKinnon *et al.* 1999; Freeman *et al.* 2001; Nevill *et al.* 2010, 2014), *Nothofagus* (Worth *et al.* 2009), *Tasmannia* (Worth *et al.* 2010), and *Atherosperma* (Worth *et al.* 2011). Secondly, in Victoria and Tasmania *Z. arborescens* generally does not co-occur with other *Zieria* species (see Figure 5.5), therefore in these populations evolutionary processes such as hybridization and introgression are less likely to confound the results of genetic studies resulting in a clearer picture of the biogeographic connections in south-eastern Australia. Specimens of *Z. arborescens* collected from its more northern distributional range were included in the study to establish a broad phylogenetic context.

5.5. Materials and methods

This study employed Bayesian inference (BI) and maximum parsimony (MP) methods to analyse cpDNA and nrDNA data from accessions of *Z. arborescens* and

closely related taxa, including outgroup taxa. The common materials and methods used here, including taxon sampling, DNA extraction, PCRs, sequencing, alignment, and phylogenetic analysis, are outlined in General Methods (Chapter 2). Where possible the same accessions were used in both cpDNA and nrDNA analyses, to enable direct comparison of results.

Fifty seven accessions of *Z. arborescens* were sequenced, broadly covering the geographical range of the species, but focussing on Victoria and Tasmania. One accession, *Z. arborescens* 622747, was excluded from the nrDNA analysis because of the large number of polymorphisms detected in its nrDNA sequences (27 in ITS and 19 in ETS). The analysis included accessions of ten other species previously noted as closely related to *Z. arborescens* morphologically (Armstrong 2002) or genetically (Chapters 3 and 4): *Z. caducibracteata*, *Z. covenyi*, *Z. murphyi*, *Z. oreocena*, *Z. robusta*, *Z. involucrata*, *Z. lasiocaulis*, *Z. southwellii*, *Z. montana* and *Z. smithii*. An excerpt of the nrDNA tree from Chapter 4 (Figure 5.4) shows the relationships of *Z. arborescens* (in clade D) to these other taxa (in clades B, C and D).

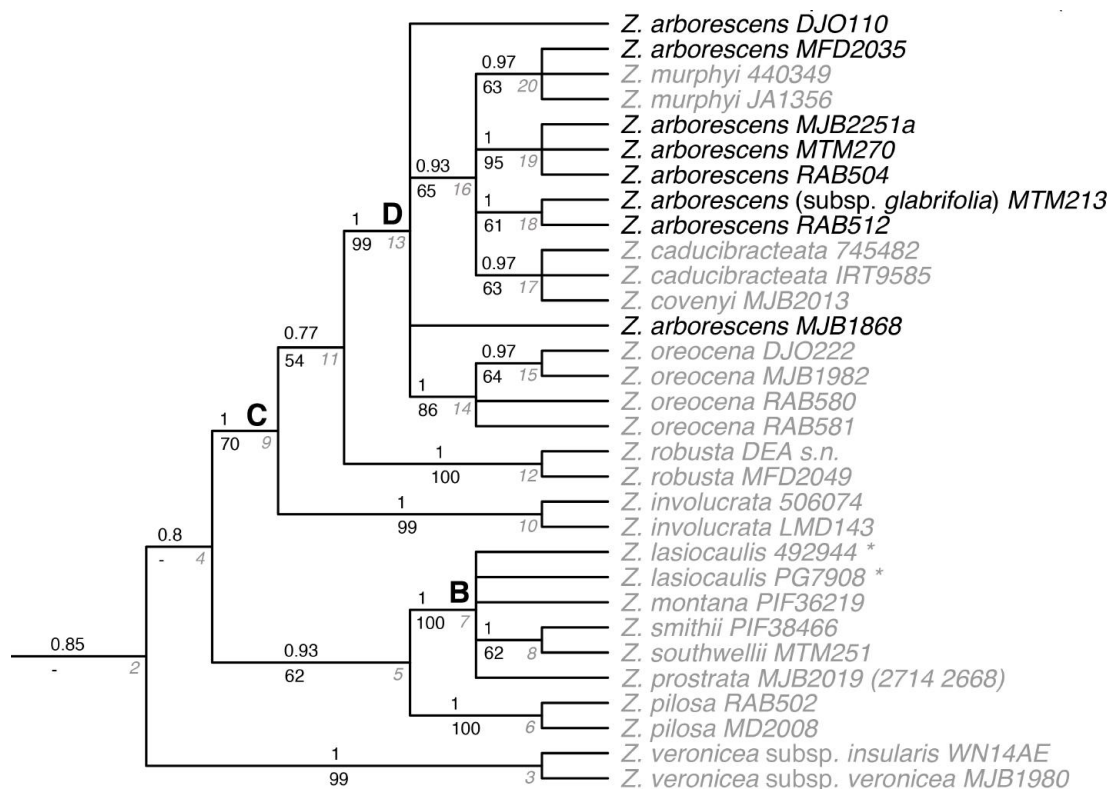


Figure 5.4. Excerpt from the phylogenetic tree based on nrDNA (Chapter 4) showing relationships of *Zieria arborescens* and other closely related species. PP and BS values are given above and below the line, respectively, and node numbers are in grey italics. Accessions of *Z. arborescens* are shown in black; all other taxa are shown in grey. Placement of the morphologically similar species *Z. lasiocaulis* is highlighted with asterisks.

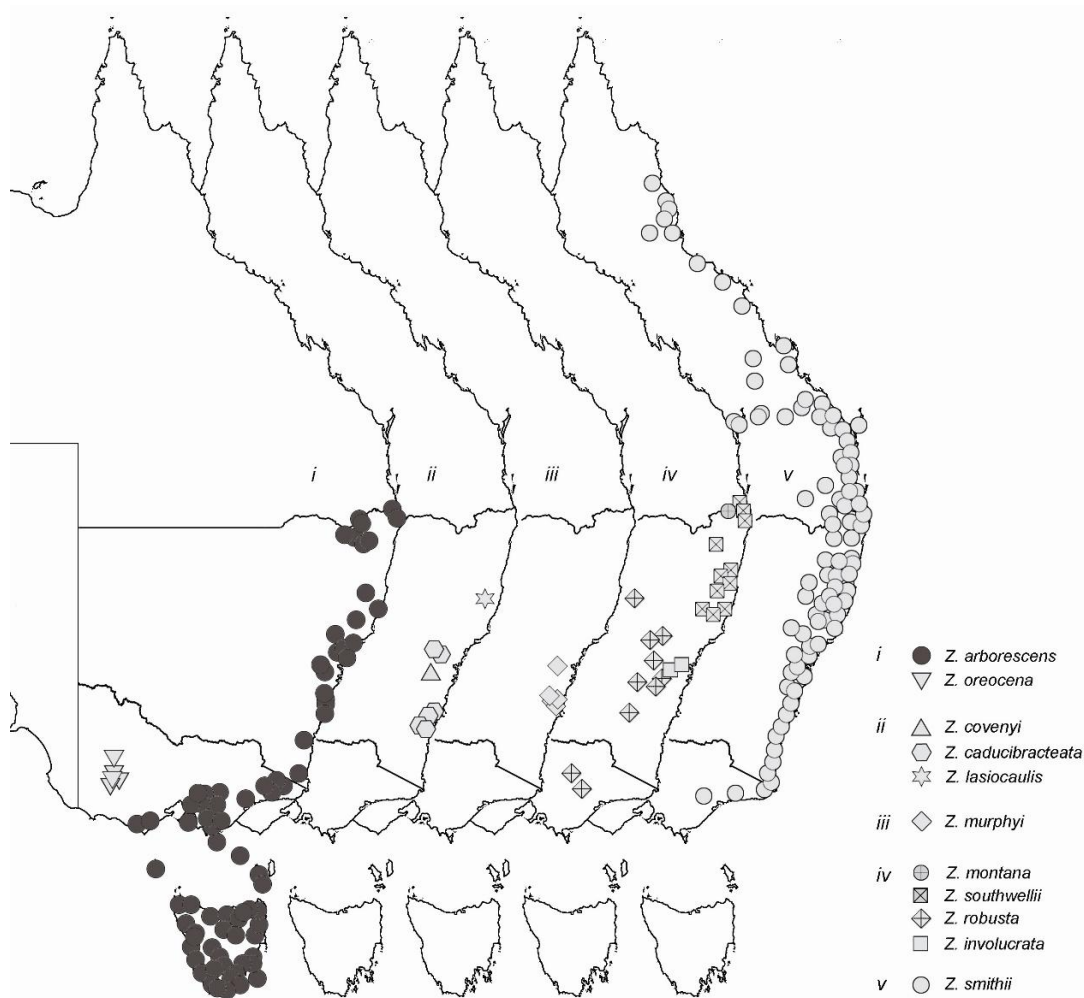


Figure 5.5. Distribution of taxa in this study that are closely related to *Zieria arborescens* (see Figure 5.4).

The distributions of the important taxa included in this study are shown in Figure 5.5. Of note is *Z. lasiocaulis* (Figure 5.4, clade B), considered to be morphologically similar to *Z. arborescens* (Armstrong 2002). Four taxa in clade B and eight other taxa (for which sequence data were available) provided phylogenetic context and potential outgroups. *Zieria prostrata* (clade B), which was shown to have divergent nrDNA paralogues (discussed in Chapter 4), was not included. A single accession, *Z. smithii* MJB2029, was selected as the outgroup for all BI analyses (which only allows one outgroup), and for consistency this sample was set as the functional outgroup in MP analyses.

The chloroplast markers used here were the same as those used for the phylogenetic study of the genus (Chapter 3, published, see Barrett *et al.* 2015) with the addition of a fifth marker, *trnV-ndhC*. Information about the markers and primers used is provided in Chapter 2 (General Methods). Sequence data from the five chloroplast

markers were combined to improve resolution and increase statistical support for nodes, and because chloroplast genomes do not undergo recombination, being generally inherited maternally as a unit (Soltis and Soltis 1998; Miller *et al.* 2009).

The nuclear markers, ITS and ETS, were combined because preliminary analysis of the markers separately produced poorly resolved trees each consisting mainly of a large polytomy. ITS and ETS are from the same nrDNA cistron and datasets are routinely combined in phylogenetic analysis to improve resolution (Baldwin and Markos 1998; Bena *et al.* 1998; Clevinger and Panero 2000; Harrington and Gadek 2004), as discussed in more detail in Chapter 4.

The models used in the BI analyses for both chloroplast and nuclear datasets were chosen using the Akaike information Criterion (AIC), implemented in MrModeltest 2.3 (Nylander 2004). They were: GTR+G for *rpl32-trnL* and ETS; GTR+I for *trnL-F* and *trnV-ndhC*; F81+I for *trnQ-5'rps16* and *trnS-G*; and SYM+G for ITS. Indel characters were coded separately. Numerous variable mononucleotide repeat regions detected in the chloroplast DNA sequence alignment were excluded from both MP and BI analyses. MP analyses that yielded many thousands of equally parsimonious trees subsequently employed the search strategy described in the General Methods (Chapter 2).

5.6. Results of cpDNA analysis

Phylogeny

Sequences from 84 samples were included in the cpDNA analysis: 57 accessions of *Z. arborescens*, 17 of species closely related to *Z. arborescens*, and ten potential outgroup samples. Analysis of data from the five combined markers included 4541 characters of which 238 were variable, 96 were parsimony informative and 43 were indels (Table 5.1). A heuristic search in the MP analysis of the combined cpDNA data resulted in many thousands of equally parsimonious trees each of length 293 steps and a CI = 0.84. Indels provided the highest percentage of parsimony informative characters (27% of the total), followed by the *rpl32-trnL* and *trnS-G* regions (24% and 21% respectively).

Table 5.1. Details of heuristic search results for the maximum parsimony analysis of cpDNA data.

	<i>rpl32– trnL</i>	<i>trnL– trnF</i>	<i>trnQ– 5'rps16</i>	<i>trnS– trnG</i>	<i>trnV– ndhC</i>	Indels	Combined data
Number of included characters	1023	897	680	1453	445	43	4541
Number of variable characters per region (excluding indels)	68	30	28	56	13	43	238
Number of indel characters per region	15	2	4	15	7		43
Total number of variable characters per region	83	32	32	71	20		238
Number of parsimony informative characters	23	9	12	20	6	26	96
Consistency Index (CI)	0.91	0.89	1.00	0.92	0.93	0.72	0.84
Retention Index (RI)	0.95	0.95	1.00	0.95	0.98	0.92	0.93
Rescaled Consistency Index (RC)	0.87	0.84	1.00	0.88	0.91	0.66	0.79

Thirty two unique haplotypes were identified in *Z. arborescens*, and haplotypes of some nested taxa differed from those of *Z. arborescens* by one or two base pairs only. Tree topology of the MP strict consensus was highly similar to that of the BI majority-rule consensus, with no conflicting nodes. The BI majority-rule consensus tree (Figure 5.6) resolved 39 nodes, four more than the MP consensus tree (not shown). The two measures of statistical support for nodes were not always well correlated, with some clades having high PP values but low BS support, e.g. clades D, G, O and P, and node 4 in the backbone, which had a BS value of 68%. The backbone of the tree was otherwise well supported: nodes 2 to 7 had PP values ranging from 0.97 to 1, and BS values between 88 and 100% (with the exception of node 4). To indicate the degree of divergence between haplotypes branch lengths are shown in Figure 5.7 (showing ingroup samples) and Figure 5.8 (mainly showing outgroups).

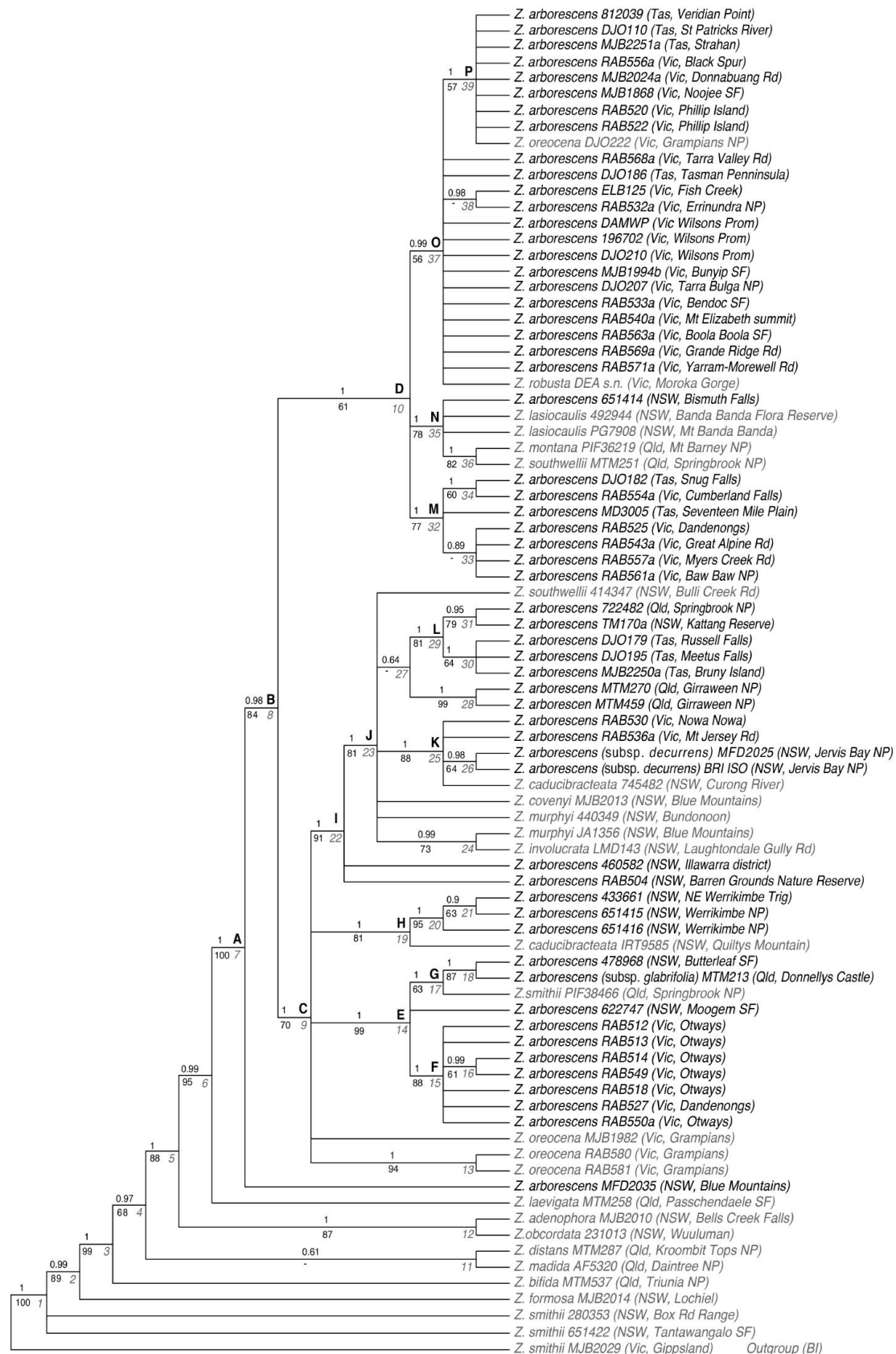


Figure 5.6. Bayesian inference majority-rule consensus tree of combined cpDNA data, with PP values above branches, BS values below branches, and nodes numbered (in grey). Major clades and subclades are labeled A to P. Taxa other than *Z. arborescens* are in grey.

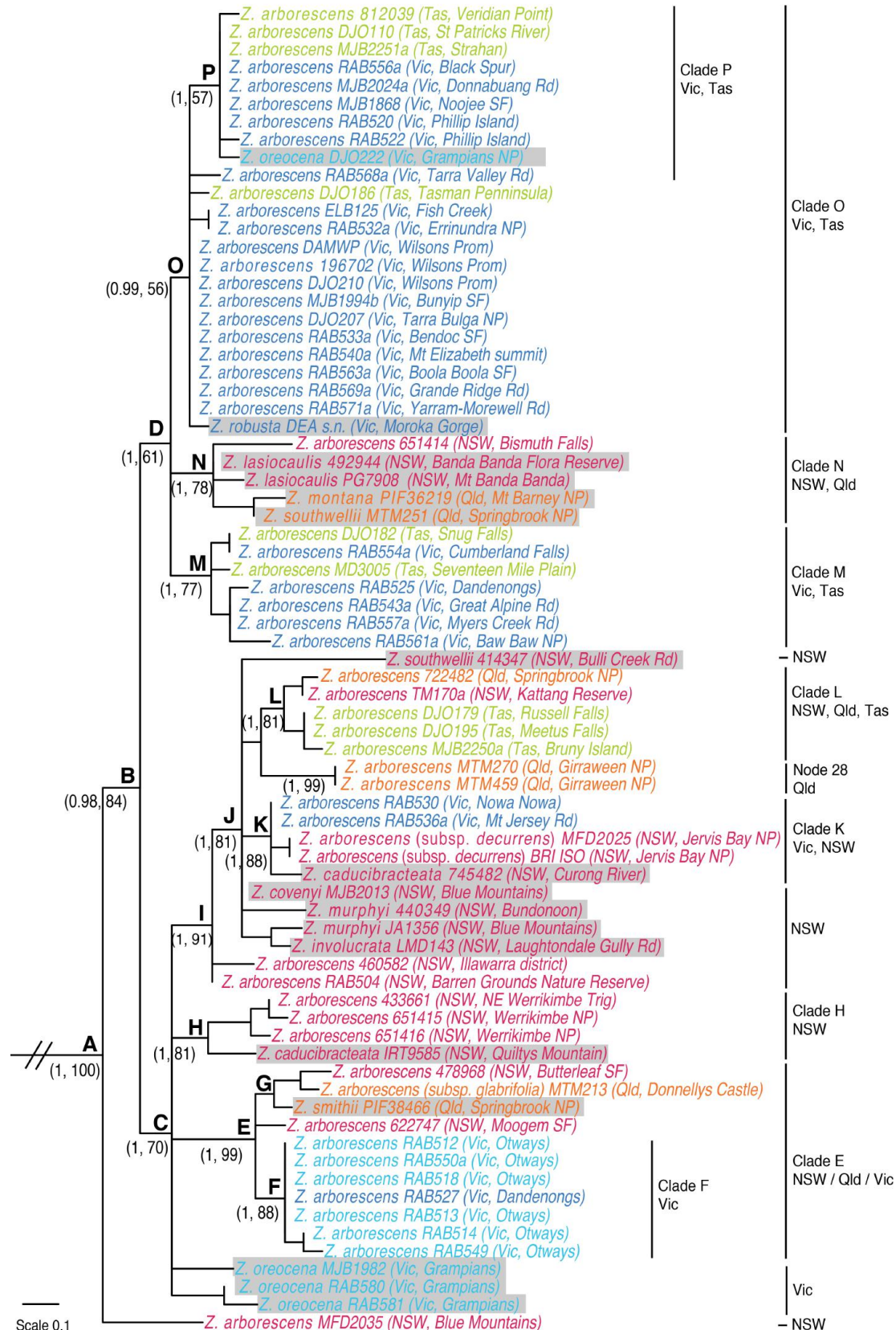


Figure 5.7. Section of Bayesian inference consensus tree of combined cpDNA data showing mean branch lengths. Double lines on branch at node A indicate interrupted branch length scale. Accessions of taxa other than *Z. arborescens* are highlighted in grey. Some PP and BS values are given in brackets (also see Figure 5.6). Samples are colour-coded by location as follows: light blue = western Victoria, dark blue = eastern Victoria, green = Tasmania, orange = Queensland, crimson = New South Wales.

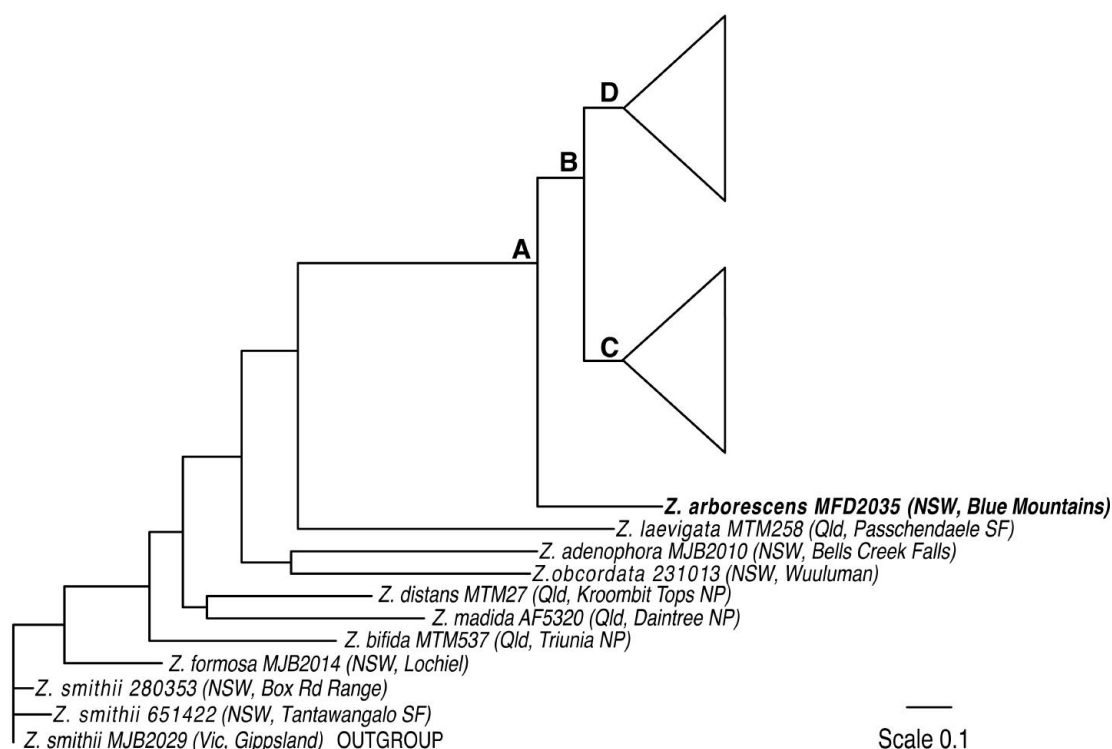


Figure 5.8. Simplified Bayesian inference majority-rule consensus tree of combined cpDNA data, showing mean branch lengths of branches outside clade A. The accession in bold (*Z. arborescens* MFD2035) is the sister to all other samples in clade A, which contains all other accessions of *Z. arborescens*.

Geographic paralogy (major clades C and D)

Figure 5.9 shows the distribution of the samples in the two main clades, C (PP 1, BS 70%) and D (PP1, BS 61%). Both clades contain cpDNA haplotypes of *Z. arborescens* from southeast Queensland, northern New South Wales, eastern Victoria and Tasmania, although most of the haplotypes from eastern Victorian were in Clade D. The geographic overlap of these clades provides evidence of geographic paralogy, and therefore the clades can be evaluated separately (Nelson and Ladiges 1996). In addition, both clades contained haplotypes of *Z. arborescens* as well as other species, indicating deep divergences within *Z. arborescens* that possibly predate speciation.

Clade C consisted of 29 haplotypes in total (18 for *Z. arborescens*) from samples occurring from Tasmania to southeast Queensland. All three haplotypes from the Otways, Victoria, are in this clade, together with most (but not all) of the haplotypes of *Z. arborescens* from New South Wales.

Clade D consisted of 19 haplotypes in total (13 in *Z. arborescens*) from samples that were mostly from Victoria (particularly eastern Victoria) and Tasmania, with the exception of those in subclade N which were from New South Wales and Queensland (although only one of these was a haplotype of *Z. arborescens*).

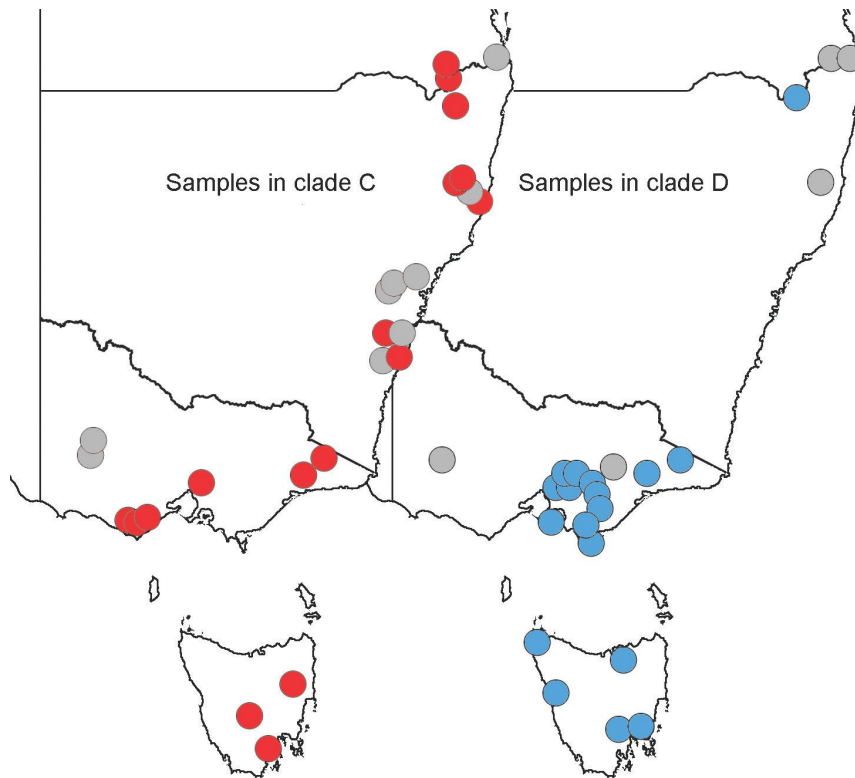


Figure 5.9. Distribution of cpDNA haplotypes in the two main clades, C and D, indicating geographic paralogy. Samples of *Zieria arborescens* are in colour and other taxa are in grey.

Geographic patterns

On a finer scale the results revealed several different geographic patterns. The first showed, in a few instances, grouping of cpDNA haplotypes from the same geographic state or region, although not necessarily in close spatial proximity. This pattern is evident in subclades F, H and G (Figure 5.7). The haplotypes in subclade F were all of *Z. arborescens* from Victoria, i.e., six from the Otways, and one from the Dandenong Ranges, which is about 150 km northeast. Subclade H consisted of three haplotypes of *Z. arborescens* from the Werrikimbe National Park, New South Wales; sister to these and differing markedly by five base pairs was one haplotype of *Z. caducibracteata* (IRT9585) from Quilts Mountain in Budawang National Park, also in New South Wales, but about 500 km to the south. Subclade G contained three haplotypes from within a radius of about 100 km in south-eastern Queensland and north eastern New South Wales (two of *Z. arborescens* and one of *Z. smithii*).

A second pattern was the finding of cpDNA haplotypes in very close geographical proximity but with deep divergences (see Figure 5.7 for branch lengths). For example, divergent haplotypes of *Z. arborescens* were recorded from the Dandenong Ranges, Victoria, with one grouping with haplotypes from the Otways in clade C (subclade F) and another grouping with haplotypes from eastern Victoria and Tasmania in clade D (subclade M). Samples of *Z. oreocena* (from the Grampians in western Victoria) also had notably divergent haplotypes. One haplotype was clustered in clade D, subclade P, differing by a single base pair from haplotypes of *Z. arborescens* that were not from the Grampians but which occurred from Tasmania to central Victoria. All of these samples in subclade P were nested within subclade O, which included a further 13 samples of *Z. arborescens* from eastern Victoria, one from Tasmania and one of *Z. robusta*. In contrast, the other three haplotypes of *Z. oreocena* were resolved at the basal node of clade C (the other main clade). Two samples of *Z. caducibracteata*, collected less than 5 km apart in New South Wales, were also divergent and fell into subclades H and K, nested with samples of *Z. arborescens*.

A third pattern showed the grouping together of haplotypes from disjunct areas across a broad geographic range; for example, subclade E (PP 1, BS 0.99%) included haplotypes from southern Victoria (subclade F), northern New South Wales, and southeast Queensland (subclade G). Subclade L contained haplotypes from Queensland, New South Wales and Tasmania (Figure 5.10). Although it clearly shows a large geographic disjunction, subclade L nevertheless provides an example of connections in *Z. arborescens* between Tasmania and the mainland.

Subclades M and O (clade D) also show connections between Tasmania and the mainland (Figures 5.11 and 5.12). Identical haplotypes of *Z. arborescens* occur in both states (subclades M and P).

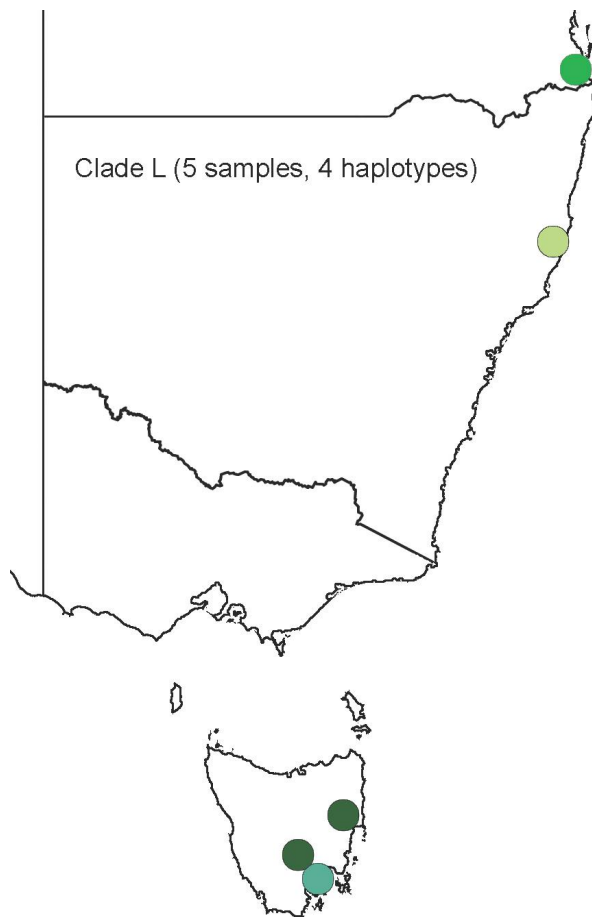


Figure 5.10. Location of the four cpDNA haplotypes in subclade L (different ones colour-coded), showing a connection between Tasmanian and the mainland in *Zieria arborescens*.

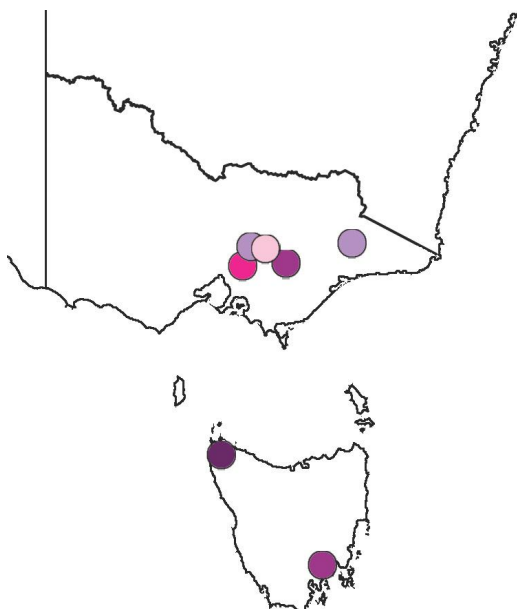


Figure 5.11. Location of the five cpDNA haplotypes (colour-coded) in subclade M, showing a connection between Tasmanian and the mainland in *Zieria arborescens*.

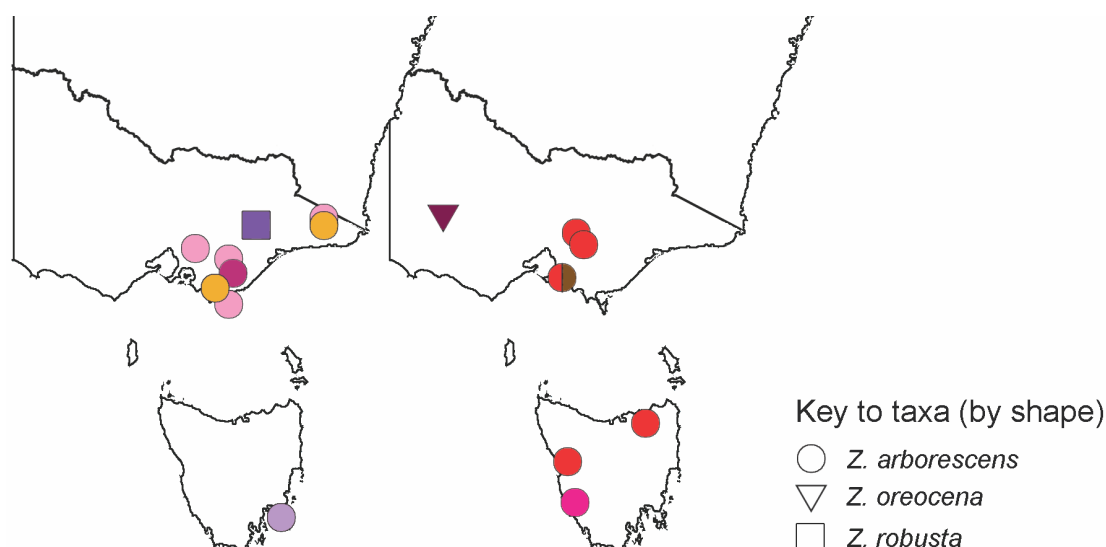


Figure 5.12. Location of the nine cpDNA haplotypes in subclade O (colour-coded). Four haplotypes in subclade P, nested within O, are shown on the map on the right, with the remaining five in O shown on the map on the left.

Similarity or divergence of haplotypes

Several subclades (F, K, L, M, O and P) contained samples with identical haplotypes (Figure 5.7). As mentioned above, in subclade F the haplotype from a sample in the Dandenong Ranges was identical to those of four samples from the Otways, about 150 km away. In subclade M one sample from Victoria (RAB554a) had an identical haplotype to another from Tasmania (DJO182). Likewise, in subclade O haplotypes of three samples from Wilsons Promontory were identical to those of seven samples from elsewhere in eastern Victoria, and, in subclade P, haplotypes of two samples from Tasmania were identical to those of four eastern Victorian samples (also see maps in Figures 5.11 and 5.12).

In contrast to the identical or highly similar haplotypes were the observed deep divergences between other haplotypes. The haplotype of *Z. southwellii* (414347) had the longest terminal branch of ingroup samples, with twelve autapomorphies detected by MP analysis (MP tree not shown), followed by *Z. arborescens* (MFD2035) with eight (sister to all others in the ingroup), *Z. arborescens* (651414) with seven, and *Z. murphyi* (404349) and *Z. oreocena* (MJB1982) each with five autapomorphies. The branch leading to subclade E (containing the Otways samples) was the longest internal branch, with the haplotypes of this group differing from others by at least seven steps in the MP analysis (Figure 5.7). Other relatively long internal branches, indicating noteworthy divergence of haplotypes within groups, tended to be among

samples from New South Wales and Queensland (subclades G, H, J and N), whereas closely related haplotypes from Victoria and Tasmania mostly differed by only a few base pairs, i.e., within subclades F, K, M and O (see Figure 5.7 for branch lengths and Figures 5.10, 5.11 and 5.12 for locality maps of samples in subclades L, M and O).

Taxa nested in Zieria arborescens

The results confirm that, based on cpDNA data, *Z. arborescens* is not monophyletic. Clade A (PP 1, BS 100%) contained all included accessions of *Z. arborescens*, plus accessions of the following ten species: *Z. oreocena* (four accessions); *Z. smithii* (one accession); *Z. caducibracteata* (two accessions); *Z. involucrata* (one accession); *Z. murphyi* (two accessions); *Z. covenyi* (one accession); *Z. southwellii* (two accessions); *Z. montana* (one accession); *Z. lasiocaulis* (two accessions); *Z. robusta* (one accession). In a few cases the haplotypes of *Z. arborescens* differed from those of other species by a single base pair (Figure 5.7) in a total of 4541 (e.g. *Z. robusta* in subclade O, *Z. oreocena* in subclade P, *Z. caducibracteata* in subclade K, *Z. smithii* in subclade E). One accession of *Z. arborescens* (MFD2035; highlighted in bold in Figure 5.8) from the Blue Mountains in New South Wales, was resolved as sister to all the other samples in clade A and was characterised by eight autapomorphies as detected by MP analysis (MP tree not shown). Only subclades F, L, and M contained samples of *Z. arborescens* and no other species.

Subclade N showed four species to be related, but included only one *Z. arborescens* haplotype (651414), which was unique, differing from others in the group by five or more base pairs. Subclade G showed a haplotype of *Z. smithii* from Springbrook NP, Queensland, to be related to two haplotypes of *Z. arborescens*, one from Queensland and one from New South Wales. Subclade H related one New South Wales haplotype of *Z. caducibracteata* to three haplotypes of *Z. arborescens* also from New South Wales. In subclade I (PP 1, BS 91%) haplotypes of five narrowly distributed New South Wales species (*Z. caducibracteata*, *Z. covenyi*, *Z. involucrata*, *Z. murphyi*, and *Z. southwellii*) were all nested with samples of *Z. arborescens* ranging from Tasmania to east Gippsland, New South Wales and Queensland. Single accessions of two Victorian species, *Z. oreocena* (Grampian Ranges) and *Z. robusta* (Moroka Gorge) were nested in subclade O with Victorian and Tasmania haplotypes of *Z. arborescens*,

while haplotypes of another three samples of *Z. oreocena* were divergent and fell on the basal node of major clade C.

5.7. Results of nrDNA analysis

Eighty three samples were included in the nrDNA analyses including 56 accessions of *Z. arborescens*, 17 accessions representing ten taxa closely related to *Z. arborescens*, and 10 potential outgroup samples. Most amplification reactions yielded a single ITS or ETS PCR product, and produced ‘clean’ sequences using direct sequencing methods. Most samples in the study had zero, one or two polymorphisms in their ITS or ETS sequences; the exceptions were one sample with three polymorphisms in ITS (*Z. robusta* DEA s.n.), two samples with four polymorphisms in ITS (*Z. arborescens* DJO207 and RAB536a), and one sample with three polymorphisms in ETS (*Z. covenyi* MJB2013). Data from the two combined markers included 1163 characters of which 179 were variable and 75 were parsimony informative (i.e., 6.4% of the total characters, and 42% of the variable characters). The ITS region provided the majority of the parsimony-informative characters (48 in ITS compared with 27 in ETS; Table 5.2).

Table 5.2. Details of heuristic search results for the maximum parsimony analysis of nrDNA data.

	ITS	ETS	Combined data
Number of included characters	674	489	1163
Number of variable characters	104	75	179
Number of parsimony informative characters	48	27	75
Parsimony informative characters as a percentage of variable characters	46%	36%	42%
Tree Length (shortest trees)	152	111	266
Consistency Index (CI)	0.88	0.92	0.89
Retention Index (RI)	0.95	0.94	0.94
Rescaled Consistency Index (RC)	0.83	0.87	0.83

The results from MP (Table 5.2) and BI analyses of nrDNA data were similar, and neither tree was well resolved. The BI majority-rule consensus tree was slightly better resolved than the MP strict consensus tree, and the BI tree only is presented here (Figure 5.13) with the main clades and subclades labelled A to K. Branch lengths are shown in Figures 5.14 and 5.15 (clade A in detail). The BI tree resolved 23 variously supported nodes. The major clade (A, node 7; Figure 5.13) was well supported (PP 1,

BS 100%) and contained all the accessions of *Z. arborescens* as well as *Z. oreocena* (four samples), *Z. murphyi* (two samples), *Z. covenyi* (one sample) and *Z. caducibracteata* (two samples). Node 7 is a large polytomy of 35 individual accessions of *Z. arborescens* and four main subclades (B to E, PP 0.88 to 1, BS 62% to 87%). Subclade E was resolved by BI (PP 0.88) but not by MP using the 'strict consensus' option, although it had moderate BS support (81%). Subclade E (node 21) included two further resolved small subclades G (PP 0.97, BS 64%) and F (PP 1, BS 61%).

Zieria robusta (DEA s.n. from Moroka Gorge, Victoria) was resolved as sister to all the accessions in clade A but this node (6) was not highly supported (PP 0.81, BS 76%). In the backbone of the tree node 4 was well supported in both analyses (PP 1, BS 89%), and nodes 2, 3 and 5 were moderately to well supported by BI but poorly or not supported by MP analysis. Also, resolved outside the main clade (A) was clade H (node 10; PP 1, BS 92%), which contained *Z. southwellii*, *Z. montana* and three accession of *Z. smithii*, and two accessions of *Z. lasiocaulis*, a species previously shown to have strong morphological affinities with *Z. arborescens* (Armstrong 2002). Of the outgroups, *Z. adenophora* and *Z. obcordata* were strongly supported as sister taxa (node 9, PP 1, BS 100%), but this was based on only one representative accession of each species.

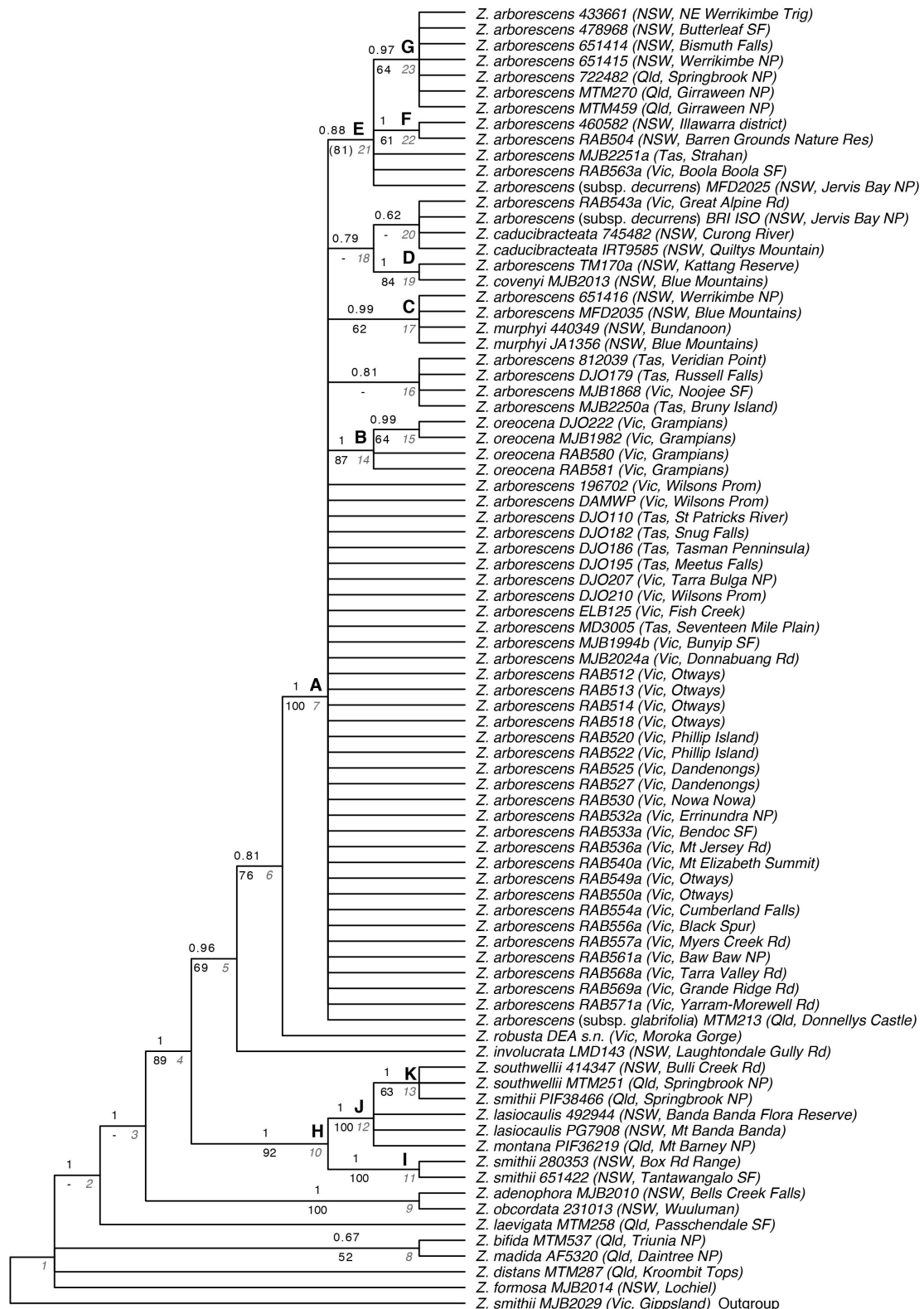


Figure 5.13. Bayesian inference majority-rule consensus tree of combined nrDNA data, with PP values above branches and BS values below, and nodes numbered (in grey). Clades and subclades are labelled A to K. Subclade E had moderate support in the bootstrap analysis (81%, shown in brackets) but was not resolved in the maximum parsimony strict consensus tree.

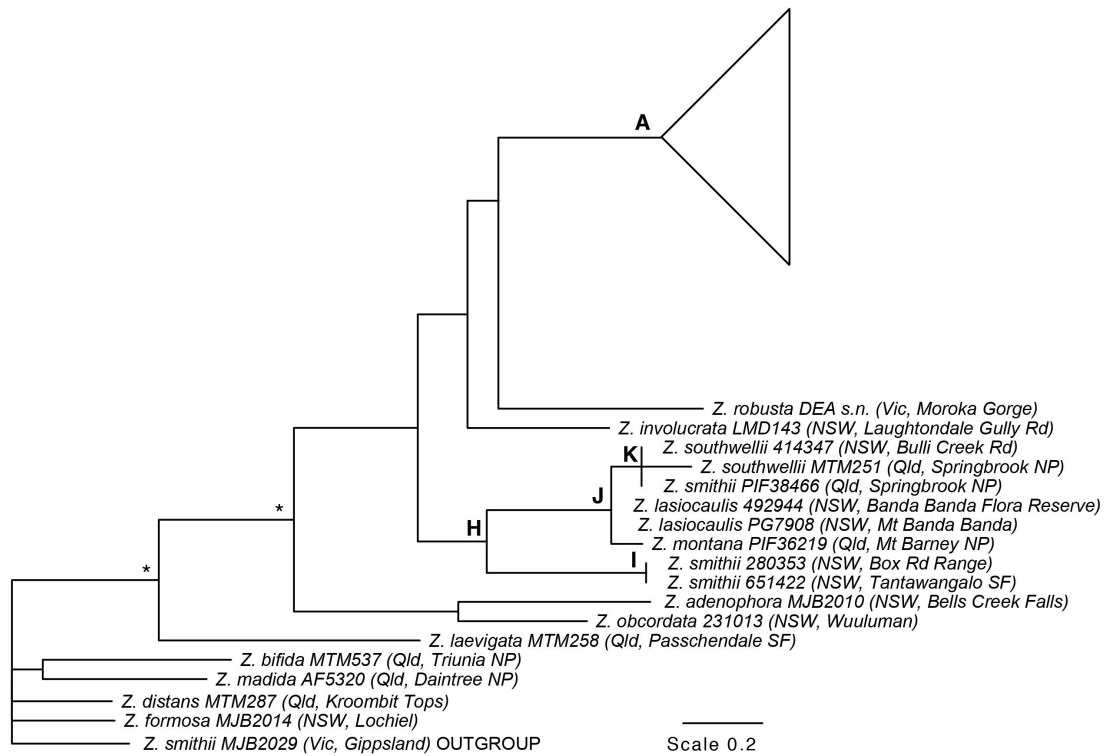


Figure 5.14. Bayesian inference majority-rule consensus tree of combined nrDNA data, showing mean branch lengths (clade A is shown in detail in Figure 5.15). Nodes with an asterisk were not resolved by maximum parsimony using the strict consensus option.

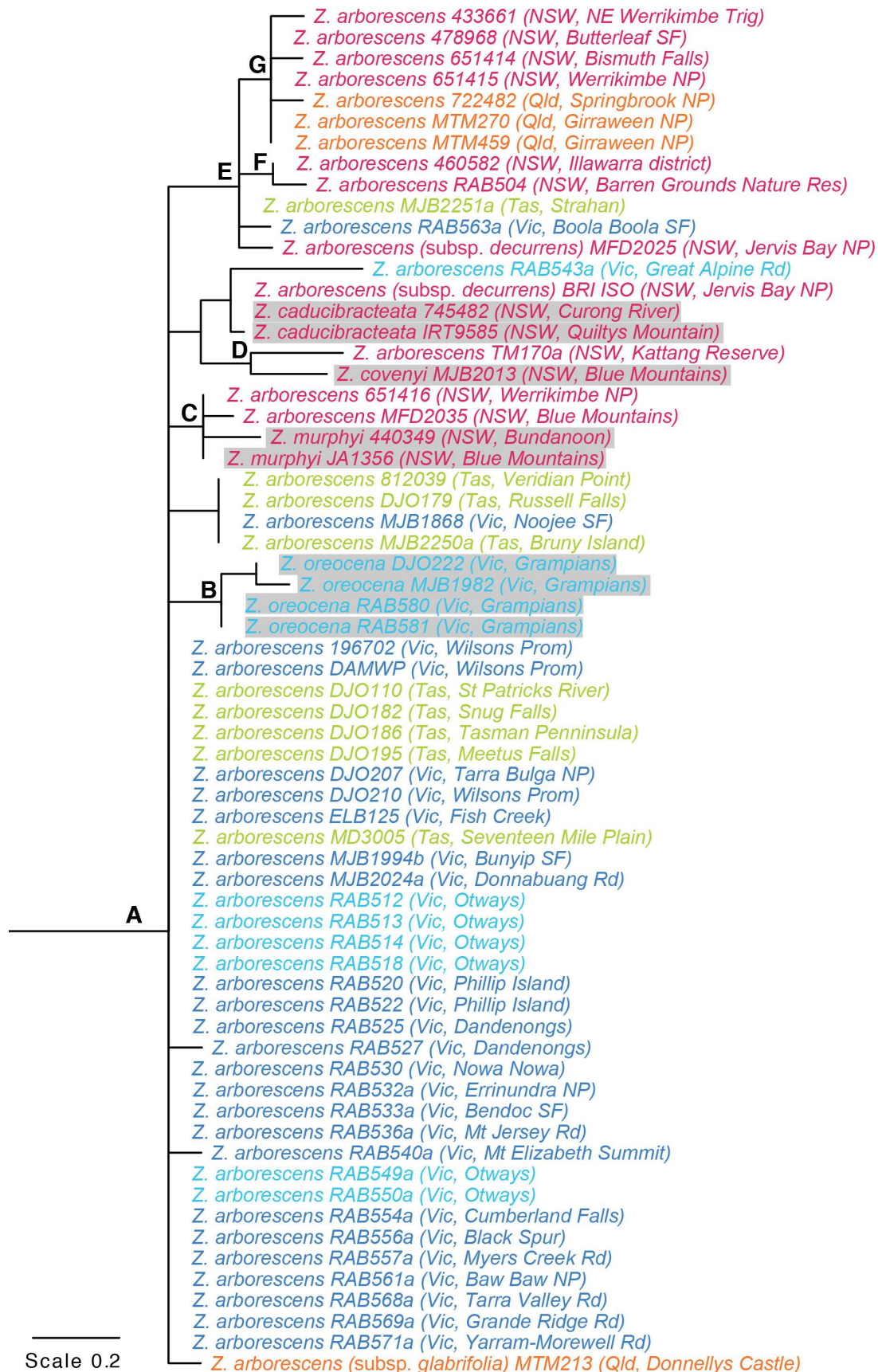


Figure 5.15. Detailed section of the Bayesian inference tree (Clade A) based on nrDNA, showing mean branch lengths and accessions colour coded by provenance as follows: light blue = western Victoria, dark blue = eastern Victoria, green = Tasmania, orange = Queensland, red = New South Wales. Accessions of taxa other than *Z. arborescens* are highlighted in grey.

The results (Figures 5.13, 5.14 and 5.15) show ITS + ETS sequence data to be largely uninformative of phylogeographic pattern in *Z. arborescens*. Thirty-two samples of *Z. arborescens* had identical ITS + ETS sequences (Figure 5.15); these were from Tasmania, eastern Victoria and the Otways in western Victoria. Three samples from different parts of Tasmania had identical sequences to one from Noojee State Forest in eastern Victoria, and although resolved as a group, their relationship is not supported, with node 16 (see Figure 5.13) having only 0.81 PP and less than 50% BS support.

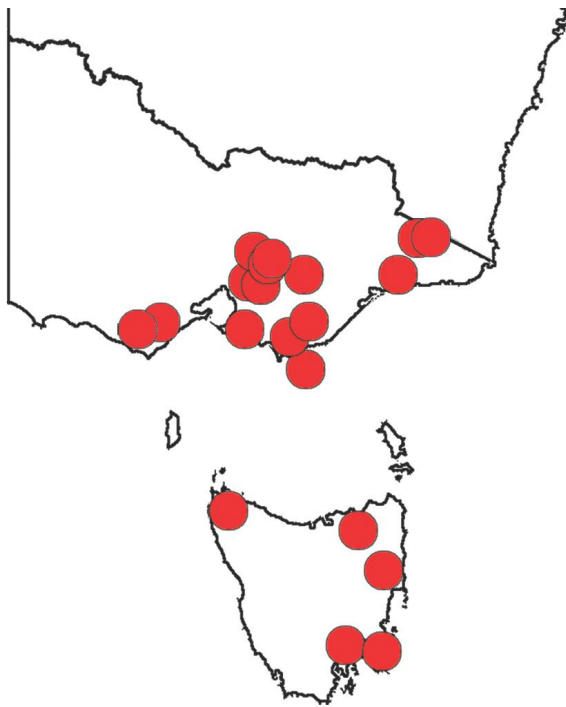


Figure 5.16. Geographic location of samples with identical ITS + ETS sequences, which were part of the polytomy at the basal node of clade A.

Clade B included all samples of *Z. oreocena* (including the one that had a divergent chloroplast haplotype; Figure 5.7), supporting the uniqueness and monophyly of that species, which is endemic to the Grampians, Victoria. However its relationship to *Z. arborescens* is equivocal being part of the polytomy at node A. Subclade C relates the two samples of the New South Wales endemic, *Z. murphyi*, to two New South Wales samples of *Z. arborescens*. One sample of *Z. murphyi* (JA1356) had identical ITS + ETS sequences to that of *Z. arborescens* (MFD2035) collected only 10 km away. The other samples, *Z. murphyi* (440349, from Bundanoon) and *Z. arborescens* (651426, from Werrikimbe NP) differed by two base pairs (in ITS) and were geographically separated by about 120 km. Subclade D (PP 1, BS 84%) relates the

Blue Mountains narrow endemic *Z. covenyi* to an accession of *Z. arborescens* from the Kattang Reserve (TM170a) in New South Wales more than 300 km away, although these two sequences differed markedly, by seven base pairs.

There was no support for suggesting that *Z. caducibracteata* is more closely related to two samples of *Z. arborescens* from Victoria or to subclade D than to other samples, since neither nodes 20 nor 18 had bootstrap support and only PP of 0.62 and 0.79 respectively (Figure 5.13). The two samples of *Z. caducibracteata*, collected from populations less than 5 km apart, had sequences that differed by a single nucleotide (in ITS), and the ITS and ETS sequences of one of them (*Z. caducibracteata* 745482) were identical to those of *Z. arborescens* subsp. *decurrens* from Jervis Bay (BRI ISO) about 45 km away.

5.8. Discussion

Scope of the study: the value of including multiple species

This chapter presents a phylogeographic analysis of *Z. arborescens*, a relatively widespread species of *Zieria*, focussing on the southern part of its distribution (Victoria and Tasmania). The study is based on molecular markers and also expands on previous phylogenetic studies of the genus (Chapters 3 and 4) that included fewer samples of *Z. arborescens*. An original aim of this study was to complement previous similar studies of wet forest taxa from south-eastern Australia and models of postglacial recolonisation. Accessions of taxa closely related to *Z. arborescens* were included in this study, and more distantly related taxa were included to provide a broader phylogenetic context to the results. Inclusion of these additional taxa is in contrast to most phylogeographic studies which typically focus on a single species without considering gene flow between species or the role of deeper historical processes that can generate patterns of diversity within species (e.g. lineage divergence that predates speciation). This can be a major limitation when interpreting results in studies with such a narrow perspective. This study attempted to avoid such a limitation by incorporating a number of potentially related or hybridising species.

Overview

Analysis of the cpDNA and nrDNA datasets here produced results that support the conclusions of Chapters 3 and 4, i.e., that, in general, phylogenetic relationships

within *Zieria* are complex (especially those based on cpDNA data), and that molecular gene trees frequently conflict with current taxonomic boundaries. The cpDNA data showed ten taxa nested within *Z. arborescens* and complex geographic relationships between haplotypes. The results appear to be somewhat complicated by underlying evolutionary processes as discussed below, and provide a poor model for inferring overarching phylogeographic patterns in *Z. arborescens*. The phylogeny from the nrDNA data was less problematic for taxonomy but unfortunately largely uninformative for phylogeography given the low sequence variation in the ITS and ETS markers.

Nevertheless, prior to this study it was not known which or how many markers would be needed to provide sufficient resolution for phylogeographic deductions, in particular, how much variation existed in ITS and ETS markers in *Z. arborescens*, and what sampling regime would be appropriate for phylogeographic studies in this species. One outcome of this study was to gain evidence that variation in ITS and ETS is low in *Z. arborescens* and closely related species. Subsequently, future research will require additional, more informative nuclear markers, and next-generation sequencing approaches could be particularly useful.

Incongruence between nrDNA and cpDNA and species-level taxonomy: signals of incomplete lineage sorting and introgression

Variation in nrDNA and cpDNA and current species-level taxonomy were found to be incongruent for *Z. arborescens* and closely related taxa sampled in this study. As discussed in Chapters 3 and 4, incongruence may reflect: (1) incomplete lineage sorting, which refers to variable retention of different ancestral chloroplast lineages across different taxa and which still exist within *Z. arborescens*; (2) hybridisation and introgression between species that co-occur or co-occurred in the past, resulting in ‘chloroplast capture’; (3) imperfect species-level taxonomy which might also be confounded by homoplasy of morphological characters; and (4) parallel evolution of similar cpDNA sequences in different taxa (homoplasy) and in different populations of *Z. arborescens*. In the cpDNA tree many of the clades were well supported statistically with up to six characters supporting some clades (e.g. subclade E), therefore homoplasy of molecular characters is an unlikely explanation of the patterns seen in the cpDNA (McKinnon *et al.* 1999; Holmes *et al.* 2014). Resolution in the nrDNA tree was poor, subclades were supported by fewer characters and often poorly

supported by PP and BS, therefore in some instances homoplasy is a possible explanation for the results of the nrDNA analysis, especially where there was geographic disjunction between samples.

Both incomplete lineage sorting and introgression are inferred to affect patterns of genetic variation in *Z. arborescens* and related species. Particular examples, starting with likely cases of incomplete lineage sorting, are discussed in the following paragraphs.

Zieria oreocena is endemic to the Grampians, western Victoria, and morphologically very similar to *Z. arborescens*. The two species are sympatric. Although the nrDNA analysis resolved *Z. oreocena* as distinct and monophyletic, one of the cpDNA haplotypes (*Z. oreocena* DJO222) was resolved in subclade P (in the main clade D, Figure 5.7), while the others were in the main clade C. This relationship of accession DJO222 to a putative ancestral haplotype in eastern Victoria (see above) is challenging to explain. Ancestral haplotypes that are shared between species, especially where there is no geographic concentration of haplotypes, are likely to reflect incomplete lineage sorting (Schaal *et al.* 1998; Muir and Schlötterer 2005). Thus given the deep divergence of chloroplasts within *Z. oreocena* the nested position of the DJO222 haplotype within *Z. arborescens* is most likely explained by incomplete lineage sorting (deep coalescence) of chloroplasts prior to speciation of *Z. oreocena*. Similarly the highly supported relationships of cpDNA in the Otways, Victoria, with New South Wales and Queensland (clade E) suggest a strong influence of incomplete lineage sorting.

Subclade N (Figure 5.7), part of the basal polytomy of clade D, included one highly divergent haplotype of *Z. arborescens* from New South Wales, related to three other species of *Zieria* from New South Wales and Queensland. The divergent haplotype of *Z. arborescens* may indicate hybridisation with other species, while the nesting of haplotypes of three other Queensland and New South Wales species in clade D with haplotypes of *Z. arborescens* from Tasmania and Victoria suggests the persistence of ancestral chloroplast lineages.

In contrast, haplotypes that are more derived (near the tips of trees) and shared between species that exist in close geographic proximity suggest that introgression of chloroplasts might have occurred in that area relatively recently (McKinnon *et al.*

1999; Muir and Schlötterer 2005; Nevill *et al.* 2014). There are few haplotypes in this study that provide strong evidence for recent introgression between species, although this could be an artefact of low or patchy sampling. A possible exception is *Z. caducibracteata* (745482) which has a highly similar haplotype (one nucleotide different) to samples of *Z. arborescens* from Jervis Bay less than 50 km to the east. On the other hand, *Z. caducibracteata* is highly similar to *Z. arborescens* morphologically and, as discussed in Chapter 6, the classification and possibly the taxon boundaries of these two species should be reassessed.

Zieria robusta provides another example of possible introgression. This species has a disjunct distribution, with numerous populations in New South Wales and fewer in Victoria (Figure 5.5, map iv), and was shown to be polyphyletic based on cpDNA (Chapter 3). The sample in this study (DEA s.n., from Victoria) had a haplotype that differed by one base pair from haplotypes of *Z. arborescens* from eastern Victoria and Wilsons Promontory (at the basal polytomy of subclade O; Figure 5.6). Introgression of chloroplasts is possible in this instance given the close geographic proximity of *Z. robusta* (DEA s.n.) and populations of *Z. arborescens* in Victoria (see the map in Figure 5.12). However alternative processes are also possible in *Z. robusta* (as previously discussed in Chapter 3), including incomplete lineage sorting, and variation associated with morphological differentiation of the disjunct populations.

In the nrDNA phylogeny, subclades C and D, which have PP and BS support, may be evidence of hybridisation and introgression of *Z. arborescens* with co-occurring species *Z. murphyi* and *Z. covenyi*. Any interaction with *Z. caducibracteata* is unclear since its position in the nrDNA phylogeny lacks support.

Haplotype diversity

The sampling in this study was not designed to enable comparisons of haplotype diversity between areas, so only broad comparisons are feasible. Some areas appeared to be dominated by single haplotypes or a small number of closely related haplotypes. Such areas included: the Otways, Victoria (subclade F); much of the central highlands of Victoria (subclade M); south and east Gippsland including Wilson Promontory and Phillip Island (subclade O). This pattern shows some similarity to that in *Eucalyptus regnans*, *E. delegatensis* and *E. obliqua* (Nevill *et al.* 2014), which are also dominated by a single haplotype in parts of eastern Victoria and the central

highlands including areas postulated to have been treeless during the Last Glacial Maximum (LGM) and later re-colonised e.g. the central highlands of Victoria. Adjacent areas, i.e., putative glacial refugia, showed higher haplotype diversity in those species. Thus areas with a mixture of divergent lineages might be evidence of refugia, where there has been stable, long-term persistence of species that has allowed maintenance of genetic diversity over time, or of other processes, such as suture zones, where distinct maternal lineages from populations expanding from different directions have converged, but chloroplast lineages have persisted (Ferris *et al.* 1998; Nevill *et al.* 2014). In this study some areas also showed a mixture of chloroplast lineages. Representatives from both the main cpDNA clades (C and D) occurred in reasonably close proximity, i.e., haplotypes in the Grampians (*Z. oreocena*), the Dandenongs, Nowa Nowa and Mt Elizabeth, Mt Jersey Rd and Errinundra (*Z. arborescens*). Also, haplotypes of *Z. caducibracteata* in Budawang National Park fell in different subclades of clade C.

Mainland and Tasmania connections

There was evidence of connectivity between Tasmania and the mainland in *Z. arborescens* through central and eastern Victoria. For example, the relationships of cpDNA haplotypes in subclade P (clade D) show connections between the central Dividing Range of Victoria, south and east Gippsland (i.e., eastern Victoria, including Philip Island and Wilsons Promontory) and western and north-east Tasmania. The more inclusive subclade O included the same or similar haplotypes from Wilsons Promontory, south and east Gippsland, and one from the Tasman Peninsula, south-east Tasmania, which may represent isolation and fragmentation of the species distribution that was once across an easterly corridor (including Wilsons Promontory and eastern Victoria). An identical haplotype in subclade M (clade D) links south-eastern Tasmania (Snug Falls) and eastern Victoria (Cumberland Falls). By comparison, the Tasmanian haplotypes in clade C are not as closely related to those of the mainland, differing by a greater number of base pairs and being most closely related to mainland haplotypes that are in New South Wales and Queensland (rather than Victoria). This suggests that older, relictual chloroplast lineages are variably retained in clade C, and the geographic history of those relationships is not clear.

This eastern connection between Tasmania and the mainland in *Z. arborescens* contrasts with patterns seen in some eucalypts such as *Eucalyptus globulus* (Freeman *et al.* 2001), other related species in the subgenus *Eucalyptus* (McKinnon *et al.* 1999), *E. ovata* (McKinnon *et al.* 2003), and *E. regnans* (Nevill *et al.* 2010, 2014), where the most recent connections (based on cpDNA data) seem to be via a western link, shown in particular by shared haplotypes in the Otways, Victoria, and the central west coast of Tasmania. The pattern in *Nothofagus* (Worth *et al.* 2009) remains unclear however, as the common haplotype of *N. cunninghamii* in the Otways is very widespread, occurring across the entire species ranges, including Tasmania, the Otways, Wilsons Promontory and south-eastern Victoria.

This study showed evidence of a distinctive population of cpDNA haplotypes of *Z. arborescens* in the Otways, which today is a bioregion more or less isolated by climate and geology. The link between the haplotypes in the Otways with some from Queensland and New South Wales harbours a significant geographic disjunction, and may reflect a different (more ancient) historical time period than that reflected in the patterns discussed for Clade D.

On the other hand, it is possible that there is a cpDNA connection between the Otways and Tasmania in *Z. arborescens* that was undetected in the current study because of low sampling in Tasmania. One way to begin to test the likelihood of such a connection would be to sample *Z. arborescens* on King Island, to determine whether the haplotypes of the samples there are related to Tasmanian or Otways populations; if haplotypes on King Island are more related to those in the Otways then more extensive sampling in Tasmania might be worthwhile.

Grampians and Otways connection

There was no evidence of cpDNA interaction between the Grampians and the Otways, even though both ranges are in western Victoria. This is congruent with the results of the nrDNA. A similar pattern was observed in *Tasmania lanceolata* (Worth *et al.* 2010) which showed unexpectedly strong phylogeographic structuring of cpDNA haplotypes in four major clades, the distribution of which mostly did not overlap, and in particular that haplotypes in the Grampians populations of *T. lanceolata* were endemic. Distinctive and geographically localised haplotypes such as these, and those of *Z. arborescens* in the Otways and of *Z. oreocena* in the Grampians, might be

evidence of either historical differentiation or isolation preventing further interaction between lineages. *Zieria oreocena* in particular is likely to be an example of allopatric speciation, especially given its monophyly based on nrDNA data, and that the Grampians are relatively isolated geographically with a high level of endemism among plants (Crisp *et al.* 2001). Haplotypes of populations of *Z. arborescens* in the Otways are quite distinctive (apart from one haplotype shared with a sample from the Dandenong Ranges, east of Melbourne). This pattern for the Otways was not observed in some other wet forest groups such as *Nothofagus* (Worth *et al.* 2009), *E. globulus* (Freeman *et al.* 2001) and *Tasmannia* (Worth *et al.* 2010) in which populations in the Otways had haplotypes that were common and widespread. Notably, ancestral haplotypes in these species were shared between the Otways and Tasmania, which could reflect a collective movement of forest species across Bass Strait during past favourable climates (Worth *et al.* 2010). On the other hand, Nevill *et al.* (2014) found a pattern similar to that in *Z. arborescens*, i.e., haplotypes were relatively restricted in the Otways in populations of *E. regnans*, and to a lesser extent, *E. obliqua*, where although there was a relationship to Tasmanian lineages the haplotypes were not widespread.

Taxonomic issues

Both cpDNA and nrDNA data showed *Z. arborescens* to be non-monophyletic as currently circumscribed. For *Z. arborescens* to be considered monophyletic the phylogenetic position of four taxa nested in the nrDNA tree need to be explained: *Z. caducibracteata*, *Z. covenyi*, *Z. murphyi* and *Z. oreocena*. Taxonomic issues of these four species are discussed in more detail in Chapter 6 (section 6.2.7). In brief, it is recommended that *Z. caducibracteata* be sunk as a species or recognised at subspecific rank within *Z. arborescens*. The endangered *Z. covenyi*, known from one restricted locality in the Blue Mountains, is a sterile triploid, putatively a hybrid of *Z. caducibracteata* and *Z. cytisoides* (Armstrong 2002), which could explain its placement in the nrDNA tree; it could be argued that because of its distinctive morphology, chromosome number and conservation status (Endangered under the EPBC Act 1999), this population warrants ongoing recognition taxonomically. Despite the highly similar morphology of *Z. arborescens* and *Z. oreocena*, a species endemic to the Grampians, Victoria, it is recommended that *Z. oreocena* be retained as a species given its monophyly in the nrDNA tree and the uncertainty about its

placement in the nrDNA tree. The nesting of *Z. murphyi* within *Z. arborescens* is taxonomically more challenging because *Z. murphyi* is morphologically distinct and therefore would not sit comfortably under a broader circumscription of *Z. arborescens*. Greater sampling of both taxa in the Blue Mountains region, testing for introgression, and a detailed taxonomic treatment of *Z. arborescens* could help to resolve this issue. This would require the use of more variable and phylogenetically informative molecular markers that are better able to unambiguously resolve within-species relationships, coupled with a detailed study of morphological variation between populations across the species range.

Conclusion

This study provides a solid foundation for further phylogeography or population studies in *Z. arborescens* and closely related taxa. It highlights a common problem with phylogeographic studies of plants, which is the limited availability of genetic markers known to resolve relationships. This is especially so compared to the range of variable mitochondrial DNA markers that are used in animal studies (Avice 1986; Schaal *et al.* 1998; Newton *et al.* 1999).

The broader perspective taken in this study contrasts to that of most phylogenetic studies, which focus on single species, and the results illustrate the importance of considering interactions between species, especially when interpreting patterns in cpDNA data. *Zieria arborescens* was found to harbour deep cpDNA divergence which crossed species boundaries. The results are partially consistent with other studies on south-eastern Australian species that show chloroplasts to be variable within species (McKinnon *et al.* 2004; Worth *et al.* 2009, 2010, 2011; Nevill *et al.* 2010, 2014), but inconsistent in that cpDNA variation in *Z. arborescens* is not strongly geographically structured. The patchiness of the distribution of *Z. arborescens* and the patterns of its haplotype distribution could be the result of potentially large range expansions or contractions throughout multiple glacial and interglacial cycles in the past (e.g. Mellick *et al.* 2012; Rossetto *et al.* 2012), but do not provide evidence that current populations are derived from nearby glacial refugia via short-distance dispersal following the Last Glacial Maximum.

In summary, results of the molecular analyses, especially that of cpDNA, provide useful information about evolutionary processes within *Z. arborescens* and some

closely related taxa (discussed below and in previous chapters), but some of these processes confound the interpretation of phylogeographic connections and speciation patterns in this species in Victoria and Tasmania, which was one of the main aims of this study. Future phylogeographic research using additional chloroplast data might provide further information about phylogeographic patterns in *Z. arborescens* but must consider underlying processes such as introgression and incomplete lineage sorting which have been shown to complicate phylogenetic reconstructions in *Zieria*. Also recommended in future studies of *Z. arborescens* is the use of low-copy nuclear markers (given that resolution in the ITS + ETS data was poor); a sampling regime that includes more populations and more samples per population; and a critical assessment of morphological and genetic variation in *Z. arborescens* and closely related taxa such as *Z. murphyi*, *Z. caducibracteata*, *Z. covenyi* and *Z. oreocena*.

Chapter 6. General discussion

6.1. Phylogeny of *Zieria*

This study aimed to construct a phylogeny of *Zieria* (Rutaceae), a predominately Australian genus of shrubs and small trees, consisting of 59 Australian species and one species endemic to New Caledonia. Phylogenetic analyses were based on chloroplast and nuclear ribosomal DNA data, and biogeographic patterns were assessed. The results showed that relationships within this genus are complex, and revealed widespread incongruence between cpDNA, nrDNA and currently recognised taxa. It is likely that a combination of factors explains the incongruence of these results, including regional cpDNA introgression (chloroplast capture), incomplete lineage sorting, and the need for recircumscription of some taxa. Because of its complexity, the cpDNA analysis did not provide a clear basis for assessing monophyly of taxa and inferring phylogenetic relationships, but nevertheless provided interesting and useful insights into the evolution of *Zieria*. Following the results of the cpDNA analysis it also became clear that a comparison of cpDNA data with nrDNA data was needed in order to better understand the relationships, evolutionary processes and taxonomic issues in *Zieria*. Analysis of nrDNA data subsequently provided greater support for monophyly of species than cpDNA, potentially providing a better indication of phylogeny than cpDNA. However, in some taxa deeply divergent paralogues of nrDNA were detected that could predate speciation or be the result of more recent hybridisation between species, highlighting issues of paralogy that must be considered when using nrDNA for phylogenetic reconstructions.

6.2. Utility of chloroplast DNA markers

The cpDNA phylogeny of *Zieria* (Chapter 3) was based on combining four cpDNA markers (*rpl32-trnL*, *trnL-F*, *trnQ-5'rps16* and *trnS-G*) and 116 samples (including multiple accessions for some taxa) representing all species of *Zieria* except one, and the monotypic outgroup *Neobyrsesia suberosa*. There was widespread incongruence between the cpDNA tree and species-level taxonomy, with 14 species shown as polyphyletic or paraphyletic, including some widespread species (e.g. *Z. smithii* and *Z. arborescens*, each of which were resolved in at least four well supported clades) and

some species with narrow geographic ranges (e.g. *Z. alata* and *Z. oreocena*). No species represented by three or more samples was resolved as monophyletic by cpDNA. In contrast, Morton (2015) did not detect incongruence in her phylogenetic analysis of *Zieria*. Her study was more limited than the analyses presented here. She only recognised 42 species (and sampled only 32), in all likelihood because she was following the treatment of Armstrong (2002) without considering the additional species recognised in the subsequent works of Duretto and Forster (2007) and George *et al.* (2013), which were not cited in her paper. Morton (2015) included only a single accession for each of the 32 included species, and only two cpDNA markers, *rpl32* and *trnL-trnF*, whereas the cpDNA analysis in this study also included *trnQ-rps16* and *trnS-trnG* (four markers in total), plus additional taxon sampling. Different findings between the two studies highlight the need for extensive sampling and the use of a number of genetic markers to uncover patterns of relationship.

One challenge facing molecular systematists is to find suitable molecular markers for use in phylogenetic studies at various levels; different markers will be suitable for use in high-level phylogenetic studies to those suitable for low-level genetics studies. In low-level studies, such as within genera, a useful marker typically needs to be relatively rapidly evolving, easily amplifiable, able to be aligned unambiguously and sufficiently variable to provide a meaningful phylogenetic signal (Baldwin *et al.* 1995). At the time this study was undertaken there was no published literature on the phylogenetic utility of genetic markers in *Zieria*.

For the combined dataset of four cpDNA markers reported here about 40% of the variable characters were parsimony-informative, but the level of variation and informativeness among the individual markers used was variable. The *rpl32-trnL* intergenic spacer was the most useful of the markers, being relatively straightforward to amplify and sequence, and having 88 parsimony informative characters (compared to 83 in *trnS-G*, 32 in *trnQ-5'rps16* and 27 in *trnL-F*). The *trnS-trnG* spacer and the *trnG* intron (*trnS-G* region) were co-amplified using two primers (one forward and one reverse), but due to the length of the co-amplified region (about 1600 base pairs), sequencing required the use of internal primers in all samples, more or less doubling the cost of sequencing for this region. *trnS-G* was the longest marker used, and despite being nearly 500 base pairs longer than *rpl32-trnL*, it only provided 11 additional

variable characters, and fewer of its variable characters were parsimony-informative. It also had the highest homoplasy index (HI). The *trnQ*–5′*rps16* intergenic spacer was also a useful marker, being relatively trouble free to amplify and sequence, and having the highest consistency index (CI) and retention index (RI), therefore the lowest HI of the markers. In addition, although it was the shortest of the chloroplast markers and had the lowest number of variable characters, more of its variable characters were parsimony-informative than those of the slightly longer and more troublesome *trnL*–F region. *trnL*–F was often problematic to amplify, frequently requiring the use of internal primers for sequencing, amplification or both, and often resulting in poor quality PCR products and yield. It had the lowest RI because it had the highest number of parsimony uninformative variable characters. The results of this study agree with the findings of Shaw *et al.* (2005, 2007) that evolutionary rate heterogeneity exists among non-coding cpDNA regions and that, although *trnL*–F is one of the most widely used non-coding cpDNA regions in infrageneric systematic investigations, in *Zieria* it provides fewer parsimony-informative characters than some other commonly used regions.

An additional cpDNA marker, the *trnV*–*ndhC* intergenic spacer, was used for the phylogeographic study of *Z. arborescens* (Chapter 5), which also included the markers used for the cpDNA phylogeny of the genus (Chapter 3), i.e., *rpl32*–*trnL*, *trnL*–F, *trnQ*–5′*rps16* and *trnS*–G. The *trnV*–*ndhC* intergenic spacer was the shortest of the cpDNA markers (445 included characters), had the lowest number of parsimony-informative characters (six) and the highest homoplasy index (0.28). Alignment was also problematic in some areas of the sequence, which included one inverted repeat, and its overall usefulness was limited other than to add a few extra characters to the dataset to try to increase resolution at this fine level of analysis.

Amplification of the *ndhF*–*rpl32* intergenic spacer was trialled, having been noted by Timme *et al.* (2007) as a highly variable region, and by Shaw *et al.* (2007) as one of the best choices for low-level molecular studies. Despite a concentrated effort using the primers *ndhF* and *rpl32*-R (Shaw *et al.* 2007), several samples of *Zieria*, and a considerable number different PCR protocols, there were continual and varied problems with the PCR products obtained (i.e., either more than one PCR product was amplified or nothing was amplified) and eventually this marker was abandoned as a candidate for these molecular phylogenetic studies of *Zieria*.

6.3. Utility of nuclear ribosomal DNA markers

The phylogenetic analysis of *Zieria* in Chapter 4 was based on two markers, the ITS and ETS regions of nrDNA. The phylogeny included 109 samples, representing 58 of the 60 currently recognised species of *Zieria*, with multiple accessions of most, and the monotypic *Neobyrnesia suberosa* was used as the outgroup. Ten species were resolved as monophyletic based on two, or in one case four, samples. Monophyly of four species was neither supported nor rejected, and all other species with more than one accession were resolved as polyphyletic or paraphyletic.

Both the ITS region and the 3' end of the ETS region proved to be useful markers for resolving phylogenetic relationships between species of *Zieria*, but were less useful for detecting variation within a species. The availability of universal primers enabled easy amplification in most taxa. Alignment of ingroup sequences was mostly straightforward, with only the outgroup having some regions of difficult alignment.

The ETS region provided more variable characters in the combined dataset (227) than the ITS region (186) but, of those, more ITS characters were parsimony-informative (118 in ITS compared to 96 in ETS). In total the combined nrDNA data (two markers) had 214 parsimony-informative characters, compared with 230 parsimony-informative characters in the combined cpDNA analysis (four markers) in Chapter 3. Employment of ETS data clearly augmented the ITS dataset, almost doubling the number of parsimony-informative characters. The phylogenetic signals in the ITS and ETS sequences were broadly congruent, but with some 'soft' incongruence, mainly due to lack of resolution in parts of one or the other tree; there was no strong conflict between the results of the separate analyses and those of the combined analysis.

Subsequently a reasonable amount of resolution was obtained by the combination of ITS and ETS sequence data. Nevertheless, interpretation of relationships was challenging, especially for species represented by a single accession, because of widespread incongruence between the nuclear and chloroplast DNA gene trees, existing species-level classification and the morphological groups proposed by Armstrong (2002). When there is conflict between nuclear and chloroplast gene trees, the nuclear gene tree is often thought to be more likely to represent the species tree, although evolutionary processes commonly produce discrepancies between gene trees and the

phylogenies of the organisms within which the genes evolved (Doyle 1992; Maddison 1997). It is reasonable to conclude that the nrDNA phylogeny provided a better indication of phylogeny in *Zieria* than the cpDNA analysis given the greater level of congruence, in some species at least, between nrDNA and species-level taxonomy based on morphology.

The underlying evolutionary processes that gave rise to the divergent nrDNA paralogues observed in some individuals are uncertain. Divergent paralogues within the genomes of some samples could predate speciation (Muir *et al.* 2001; Bailey *et al.* 2003), and be variably retained or variably detected, i.e., not always detected by routine PCR methods (Buckler *et al.* 1997); alternatively, they could represent novel nrDNA combinations formed through hybridisation after speciation (Rauscher *et al.* 2002; Álvarez and Wendel 2003; French *et al.* 2016). There was no strong evidence for recombination between paralogues or that the paralogues represent pseudogenes (see Bayly *et al.* 2008).

The results of this nrDNA study highlight both the usefulness and the pitfalls of ITS and ETS sequences in phylogenetic analyses, which have been widely documented (Baldwin 1992; Baldwin *et al.* 1995; Baldwin and Markos, 1998; Wright *et al.* 2001; Álvarez and Wendel 2003; Feliner and Rosselló JA 2007). Recommendations for the methodology of future research, particularly of *Zieria*, include: use of sequencing methods that would separate different rDNA paralogues, e.g., cloning of all PCR products when using nrDNA markers or using next-generation DNA sequencing methods (whether targeted amplicon sequencing or skimming of high copy rDNA sequences from genomic sequencing); the use of low-copy nuclear markers; more extensive sampling (i.e., multiple accessions of taxa, ideally from multiple populations); close scrutiny of the morphology of some taxa.

6.4. Implications for species-level taxonomy

The nrDNA phylogeny showed greater support for monophyly of species than cpDNA and was more congruent with species-level taxonomy. Key examples were *Z. distans*, *Z. oreocena* and *Z. robusta*, which were monophyletic in the nrDNA tree but were strongly supported as polyphyletic or paraphyletic in the cpDNA tree. Likewise, *Z. caducibracteata*, *Z. collina* and *Z. murphyi* were potentially monophyletic based on

the nrDNA tree but were moderately to strongly supported as not monophyletic in the cpDNA tree. *Zieria arborescens*, (discussed in more detail below), was resolved as paraphyletic in the nrDNA tree with respect to *Z. murphyi*, *Z. caducibracteata* and *Z. covenyi*, although all samples of *Z. arborescens* grouped in the same clade, with strong support. This contrasted markedly with the cpDNA tree in which *Z. arborescens* was strongly supported as polyphyletic, falling in at least five different clades. Implications for the taxonomy of some taxa are discussed below, predominantly in the context of the results of the nrDNA phylogeny.

Zieria smithii and close relatives: species in need of taxonomic revision

Zieria smithii, the type species for the genus (designated lectotype by Armstrong 2002), is one of the most widespread and probably the most abundant species in the genus. It occurs from north Queensland through eastern New South Wales to Victoria (Figure 6.1 a) in a wide variety of habitats including rainforest, rainforest-sclerophyll ecotones, sclerophyll forest, open woodland, heathland and swamp (Duretto and Forster 2007). It also shows considerable morphological character variation across its range, including: the presence or absence of ridges and leaf decurrencies on branchlets; density of indumentum on various organs; degree of verrucosity of stems, petioles, leaflets and sepals; type of hairs; size and shape of leaflets, petals and sepals; length of petioles and peduncles; and number of flowers per inflorescence (Duretto and Forster 2007). Some of this variation is at least partly correlated with geography (Armstrong 2002, Duretto and Forster 2007, George *et al.* 2013), but patterns are not clear. Armstrong (2002) divided *Z. smithii* into two subspecies (subsp. *smithii* and subsp. *tomentosa*) based on the density of the indumentum on younger branches, petioles and abaxial leaf surfaces, but the distinction between the subspecies is problematic and they are not generally recognised (Armstrong and Harden 2002, Duretto and Forster 2007, George *et al.* 2013). Because of the substantial variation in *Z. smithii* the species is difficult to clearly delimit on the basis of consistent morphological characters and it commonly appears multiple times in identification keys to the genus, e.g., twice in Armstrong's (2002) key, three times in George *et al.* (2013), and four times in the key to *Zieria* species found in Queensland (Duretto and Forster 2007).

In this study *Zieria smithii* was represented by 21 samples from populations across its range. Analysis of the nrDNA data resolved *Z. smithii* as polyphyletic, with samples falling into at least six well-supported clades throughout the nrDNA tree (see Figure 4.5 for the nrDNA tree and Figure 6.1 for distribution maps of taxa and samples). Also, embedded, with strong support, within nrDNA clades containing multiple samples of *Z. smithii* were samples of several other taxa: *Z. collina*, *Z. furfuracea* subsp. *euthadenia* and *Z. prostrata* in Clade L; *Z. bifida*, *Z. boolbunda*, *Z. eungellaensis*, *Z. robertsiorum*, *Z. scopulus* and *Z. insularis* in Clade K; *Z. furfuracea* subsp. *gymnocarpa* in Clade H. Some of these nrDNA clades showed partial or strong geographic patterns. For example: all six samples of *Z. smithii* in Clade H were collected in south-east Queensland; 13 of the 15 samples in Clade L were from southern Queensland (south of Rockhampton) or northern New South Wales, and two were from far east Gippsland in Victoria; all nine samples in Clade K were from Queensland, although the most northern ones were from Mt Windsor in the Wet Tropics, north of Cairns, and the most southern ones were from south-east Queensland over 1500 km away; two samples resolved as a pair (Figure 4.5, node 36) were from Gippsland in Victoria.

Most of the taxa embedded in *Z. smithii* in the nrDNA phylogeny were narrow endemics, restricted to small geographic areas, mostly within the distributional range of *Z. smithii*, or peripheral, as in the case of *Z. robertsiorum*, a population of which occurs on the summit of Mt Finnegan, Queensland, about 60 km further north than the most northern populations of *Z. smithii*. Some of the embedded taxa, especially in clade L (*Z. boolbunda*, *Z. eungellaensis*, *Z. robertsiorum*, *Z. scopulus* and *Z. insularis*) are specialists of rocky outcrops. This pattern of relationships is consistent with these narrowly-distributed taxa having arisen through local differentiation from a more widespread collection of lineages currently included in *Z. smithii*. Speciation through local isolation and adaptation, in some cases associated with habitat differentiation, is a plausible explanation for the contrasting distribution patterns of widespread versus narrowly endemic taxa that are prevalent in *Zieria* as previously discussed by Duretto and Forster (2007).

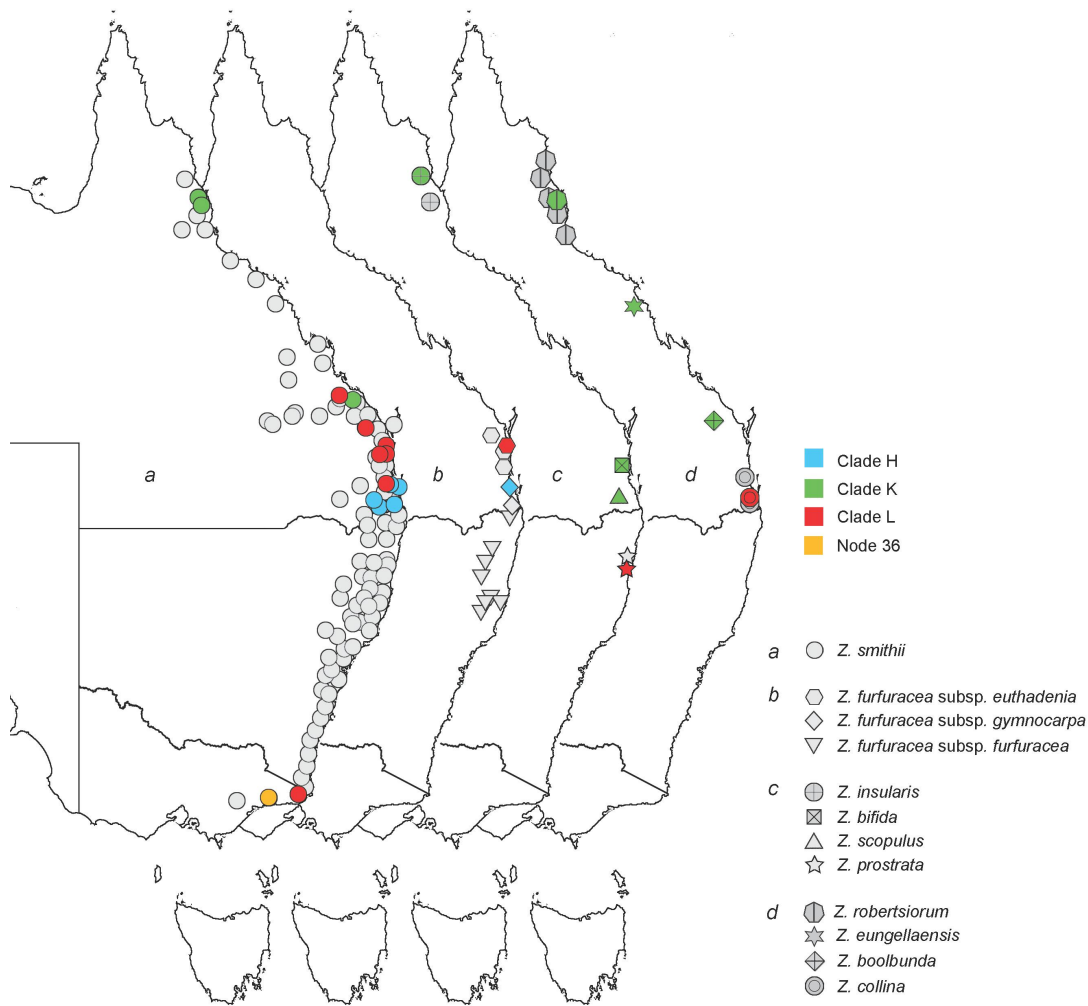


Figure 6.1. Distribution of *Zieria smithii* and related taxa (grey shapes). Coloured shapes indicate the location of samples used in this study, resolved in (colour-coded) clades by nrDNA data (Figure 4.5).

Alternative explanations for the polyphyly of *Z. smithii* in nrDNA phylogenies are less plausible, but could include that: (1) the nrDNA relationships of *Z. smithii* recovered here are artefacts of selective and incomplete sampling of nrDNA paralogues within genomes, and (2) close nrDNA relationships between *Z. smithii* and narrowly endemic species do not directly reflect the speciation patterns, but are a product of introgression between historically distinct genetic lineages. It is unlikely that all nrDNA paralogues were recovered for all samples in the dataset and therefore some artefacts of nrDNA paralogy are conceivable (explanation 1). However, random sampling of nrDNA paralogues from different accessions would not be expected to give the geographic patterns observed among the clades of *Z. smithii* and its relatives, making it unlikely that such artefacts have substantially affected these patterns. In relation to explanation 2, it is possible that *Z. smithii* locally hybridises with some species, but there is little

inference of such hybridisation in published monographs or floras (Duretto 1999; Armstrong 2002; Armstrong and Harden 2002; Duretto and Forster 2007; George *et al.* 2013) and it is unlikely (i.e., highly unparsimonious) that hybridisation has occurred to such an extent that it has obliterated all signal of *Z. smithii* as a monophyletic lineage (if it historically were one).

In this study, at least one possible example of hybridisation between *Z. smithii* and another species was detected: a sample of *Z. smithii* (PIF3466) from the Springbrook Plateau in SE Queensland was resolved by nrDNA as sister to a sample of *Z. southwellii* (Figure 4.5, node 8, PP 1, BS 62%) collected from less than 3 km away, with identical ITS sequences and differing by just two base pairs in ETS. All other samples of *Z. smithii* in this study were resolved in the large Clade A. Populations of *Z. smithii* in the Springbrook area occur mostly in ecotones between different plant communities and are not morphologically typical of *Z. smithii*, being a larger leaved variant and sharing various morphological characters with several species, including *Z. southwellii* and *Z. arborescens* (Forster pers. obs.). There is a possibility that the atypical Springbrook populations, thought to be a variant of *Z. smithii*, have arisen through hybridisation between *Z. smithii* and *Z. southwellii* or *Z. arborescens*, both of which occur in the Springbrook plateau area. More extensive sampling of all taxa occurring in this area would help to resolve these relationships.

Also of note is that in this study ten samples of *Z. smithii* (about a third of the total number sampled) from New South Wales and Queensland, including some from the Springbrook area, were ultimately excluded from analyses because their nrDNA sequences contained numerous polymorphic sites, indicating the presence of divergent nrDNA copies. As a result, the extent of variation in nrDNA within genomes is potentially under-represented, as is the potential extent of nrDNA variation associated with introgression.

The pattern of nrDNA relationships shown in Chapter 4 implies that the classification of *Z. smithii* and its close relatives requires revision. The current circumscription of *Z. smithii* is very broad, in terms of both morphological and genetic variation, and some populations are most closely related to other species. Because *Z. smithii* is so markedly polyphyletic, and because basal relationships in the main clade of the nrDNA tree (clade

A, Figure 4.5) were not well resolved, a clear taxonomic solution, recognising monophyletic groups is not immediately apparent. No small amount of ‘lumping’ of segregate taxa would define a monophyletic group and it is likely that the solution will involve the further ‘splitting’ of *Z. smithii*. This requires a much clearer understanding of both molecular and morphological variation in *Z. smithii* and allied taxa. This study provides a framework for focussing efforts in this direction.

Zieria furfuracea: evidence of three disparate lineages

Zieria furfuracea was first described by Bentham (1863). Armstrong (2002) subsequently established three allopatric subspecies, which were accepted by Duretto and Forster (2007) and George (2013). Two of these (subsp. *euthadenia* and subsp. *gymnocarpa*) are from south-east Queensland and the other (subsp. *furfuracea*) from north-east New South Wales (Figure 6.1 b). Duretto and Forster (2007) suggested that the three subspecies appear to form a monophyletic group on the basis of morphological resemblance, but that recognition of the subspecies at specific rank might be warranted given the range of character states defining them (especially bud aestivation), and because they are disjunct in occurrence. The results of the nrDNA study (Chapter 4) indicate that *Z. furfuracea* is not monophyletic, and that subsp. *euthadenia* and subsp. *gymnocarpa* are distinct from each other and from subsp. *furfuracea*. *Zieria furfuracea* (three samples representing the three subspecies) was also resolved as polyphyletic by cpDNA data analysis (Chapter 3).

Relationships of subsp. *furfuracea* were not well resolved by nrDNA, but subsp. *euthadenia* and subsp. *gymnocarpa* were each resolved, with strong support, in different clades containing numerous accessions of *Z. smithii* (Figure 6.1). These two subspecies have some morphological characters that are more typical of *Z. smithii* than of subsp. *furfuracea*, i.e., upper leaf surfaces that are glabrous and glandular-verrucose and, in subsp. *gymnocarpa*, petals that are imbricate in bud, and fruits that are glabrous, although the illustrations in George *et al.* (2013), contrary to descriptions, show that fruits of subsp. *furfuracea* can also be glabrous. The appropriate taxonomic status of taxa currently placed in *Z. furfuracea* should be determined in conjunction with that of related populations of *Z. smithii*.

Zieria cytisoides and *Zieria hydroscopica*

Zieria hydroscopica is morphologically very similar to *Z. smithii* (Duretto and Forster 2007), differing in that *Z. hydroscopica* has adaxial leaflet surfaces that are pilose (vs. glabrous or with sparse stellate or simple hairs), and branchlets with stellate hairs that have longer rays (George *et al.* 2013). The single accession of *Z. hydroscopica* in the nrDNA study (Chapter 4) was not resolved with *Z. smithii* but as sister to a sample of *Z. cytisoides* (DJO416) from the Central Western Slopes district of New South Wales over 1000 km away (Figure 4.5, node 46). This is an unexpected result because, as well as the vast geographic distance separating the samples, the species do not share many morphological characters. The highly restricted *Z. hydroscopica* only occurs in the Coomingleh SF near Monto in the Brigalow Belt Bioregion in southern Qld; the widespread and highly variable *Z. smithii* also occurs in the Coomingleh SF, and the widespread *Z. cytisoides* occurs from Victoria to north Queensland including in the Brigalow Belt Bioregion. If *Z. hydroscopica* resulted from hybridisation, it is possible that the parents were *Z. smithii* (because of the close morphological similarity) and *Z. cytisoides* (because of highly similar nrDNA). In addition, the sample of *Z. hydroscopica* used in this study was resolved with a sample of *Z. cytisoides* (DJO 320) in the chloroplast analysis (Chapter 3), but unfortunately amplification of both ITS and ETS was unsuccessful for that particular sample, so it could not be included in the nrDNA analysis. Sampling *Z. smithii* and *Z. cytisoides* from nearby localities is recommended for future investigations of *Z. hydroscopica*.

Relationships of the endangered coastal specialist, Zieria prostrata

Zieria prostrata is a narrowly endemic, endangered species that occurs on coastal headlands near Coffs Harbour in northern New South Wales (Figure 6.1 c). Based on morphological resemblance, Hogbin and Crisp (2003) assumed that *Z. prostrata* was closely related to the widespread species *Z. smithii*, populations of which also occur on headlands in the same vicinity and on headlands up to 300 km further south. They conducted a morphometric study of the two species and, using RAPD markers, made an assessment of genetic variation between the lineages. The results showed clear morphological separation of the species, but incomplete genetic separation, suggesting that *Z. prostrata* has descended from *Z. smithii* through local differentiation.

Whilst the assessment of Hogbin and Crisp (2003) is plausible based on the morphological assessment, nrDNA data (Chapter 4) are more complex; both samples of *Z. prostrata* in this study contained divergent paralogues of nrDNA that fell in different clades (i.e., clades L and B; see the nrDNA tree in Figure 4.5, and the location of the samples in clade L in Figure 6.1, and clade B in Figure 6.2). Based on combined ITS + ETS, one paralogue of *Z. prostrata* (MJB2019) was resolved with strong support in a clade with one sample of *Z. furfuracea* subsp. *euthadenia* and accessions of *Z. smithii* collected in south-east Queensland. The other paralogue was resolved with very high support in a clade with samples of *Z. lasiocaulis* (two accessions), *Z. montana*, *Z. southwellii*, and a sample of *Z. smithii* (PIF38466) from the Springbrook plateau near the Queensland/New South Wales border. This suggests that *Z. prostrata* has complex relationships with a number of species, but uncertainty about the history of the divergent nrDNA paralogues clouds understanding of those relationships. Data from specimens of *Z. smithii* from populations geographically close to *Z. prostrata* were unavailable for this study. In particular, amplification of both nrDNA markers was unsuccessful for a coastal headland sample of *Z. smithii* (DJO430), which was collected from Hat Head, about 100 km south of *Z. prostrata*. Establishing the placement of nrDNA sequences from that sample would have been very useful in this context. In the cpDNA analysis (Chapter 4) that sample was not resolved near *Z. prostrata*, which had cpDNA haplotypes similar to a sample of *Z. arborescens* from Mt Norman and samples of *Z. insularis*, *Z. alata* and *Z. smithii* (all from north Queensland). These chloroplast relationships, especially among samples separated by large geographic distances, could, as discussed above, relate to incomplete lineage sorting and might not be good indicators of phylogenetic descent.

Classification of Zieria montana

Zieria montana was first described by Armstrong (2002). He proposed a broad circumscription including a cluster of populations in south-east Queensland (which includes the type locality) and highly disjunct populations in the Queensland Wet Tropics, and in Eungella National Park. Duretto and Forster (2007) employed a much narrower circumscription of *Z. montana*, which they considered to be restricted to south-east Queensland (see Figure 6.2). In their classification, some south-east

Queensland populations previously included in *Z. montana* were designated as new species, i.e., *Z. scopulus* and *Z. boolbunda*; and the more northern populations were variously assigned to *Z. robertsiorum* or designated as new species, i.e., *Z. madida*, *Z. alata* and *Z. eungellaensis*.

The nrDNA tree (Figure 4.5) resolved *Z. montana* in clade B, based on a single accession (Figure 6.2); it was clearly distinct from the segregate taxa recognised by Duretto and Forster (2007), which all fell in clade A. This result supports the separation of *Z. montana* sensu stricto from the other taxa. Some of the segregates are, however, shown as closely related to each other or to accessions of *Z. smithii* (the status of *Z. madida* and *Z. alata* are discussed in more detail below).



Figure 6.2. Distribution of taxa that were previously included in a broader circumscription of *Zieria montana* (i.e., taxa with asterisks in the key) by Armstrong (2002), and geographic location of samples resolved in the nrDNA clade B (magenta shapes).

Zieria alata and *Zieria madida*

Zieria alata and *Z. madida* are two narrowly endemic species from the Queensland Wet Tropics (Figures 6.2 b and 6.3), often occurring in clefts or crevices in exposed windswept heathland or stunted closed forest, typically at altitudes of 900m or higher. They were described by Duretto and Foster (2007), who separated them from *Z. montana* sensu lato. On the basis of their similar morphology and habitats, Duretto and Forster (2007) suggested that *Z. alata* and *Z. madida* might be closely related. *Zieria alata* was described as differing from *Z. madida* by having prominently glandular verrucose leaf bases (versus not or slightly glandular verrucose), leaflets with acute tips (versus obtuse) and glandular verrucose peduncles (versus not). These characters do, however, show some variation, with some specimens assigned to *Z. madida* by Duretto and Forster (2007) being more glandular-verrucose than others (Webb and Tracey 12158) or having leaflets that are more acute (Godwin 2471).

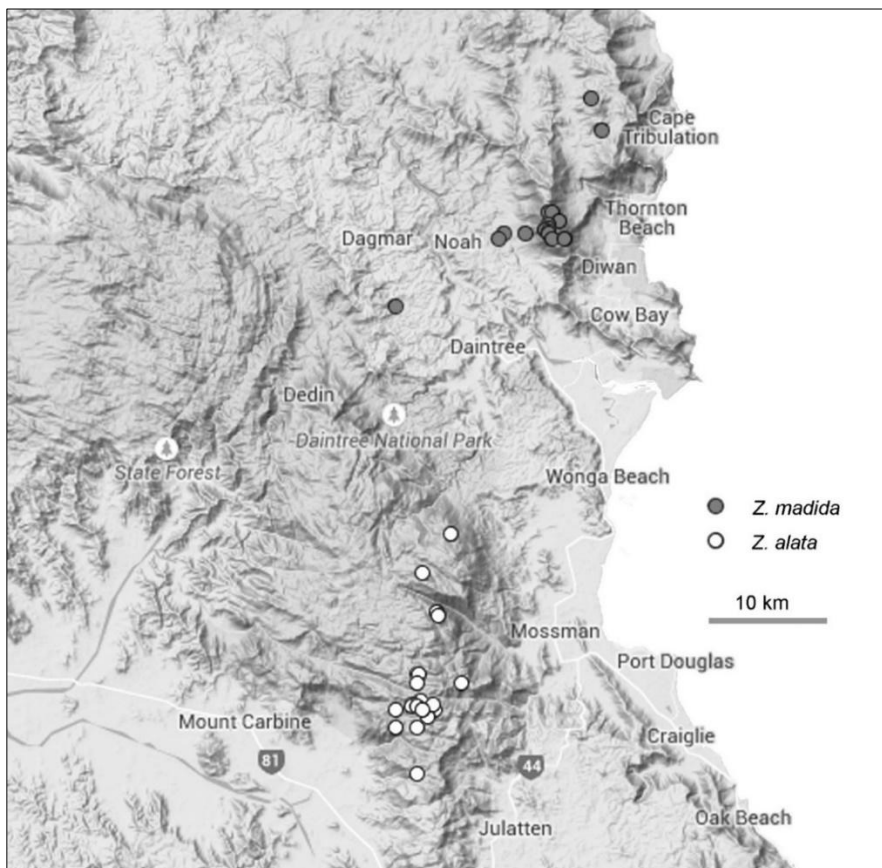


Figure 6.3. Distribution of *Zieria alata* and *Z. madida*, based on records from Australia's Virtual Herbarium.

The two species occur in an area with a radius of about 30 km (Figure 6.3) but their distributions are narrowly disjunct, with *Z. madida* occurring on Thornton Peak and Mount Pieter Botte, and *Z. alata* on the Main Coast Range to the west of the Daintree River and Mossman (including Mt Lewis and the Devil's Thumb). The nrDNA data (Chapter 4) support the sister relationship of *Z. alata* and *Z. madida* and their separation from *Z. montana*. ITS sequences of *Z. alata* and *Z. madida* differed by only two base pairs and their ETS sequences were identical. With confirmation that the two species are closely related, the morphological differences or overlap between them warrants closer inspection. They could be considered synonymous, or their sister relationship could be better indicated by subspecies rank. More sampling or use of more variable molecular markers could test whether there are consistent genetic differences between them. The two samples in this study have divergent chloroplast haplotypes (Chapter 3), but little emphasis should be placed on this, given the patterns of chloroplast DNA variation in *Zieria*.

Zieria arborescens and related taxa

Samples of the relatively widespread mesic forest species, *Zieria arborescens*, were included in the context of a broad analysis of the genus (Chapters 3 and 4) and in a more detailed investigation of populations in wet forests of Victoria and Tasmania (Chapter 5). The latter study was aimed at discovering any biogeographic pattern that may be congruent with other biogeographic studies of the mesic region of south-eastern Australia. *Zieria arborescens* was found to harbour deep chloroplast divergence, which was not strongly structured geographically, and also crossed species boundaries. These results were partially consistent with other studies on south-eastern Australian plants that also detected variable chloroplasts within species, but which, in contrast, showed strong geographic structuring of haplotypes (McKinnon 2004; Worth *et al.* 2009, 2010, 2011; Nevill *et al.* 2010, 2014). Multiple groups of closely related haplotypes in *Z. arborescens* provided some evidence of Tasmanian–eastern Victorian connections across Bass Strait (through south Gippsland), which was in contrast to a Tasmanian–western Victorian connection observed in some eucalypts (McKinnon *et al.* 1999; Freeman *et al.* 2001; McKinnon *et al.* 2004; Nevill *et al.* 2010, 2014). The complex distribution of haplotypes in *Z. arborescens* might also be partly explained by

incomplete chloroplast lineage sorting, with deep coalescence pre-dating speciation (discussed in Chapters 3 and 5). Unfortunately the results of the analysis of nrDNA in *Z. arborescens* (Chapter 5) showed poor resolution and were of limited use for determining phylogeographic patterns in that species.

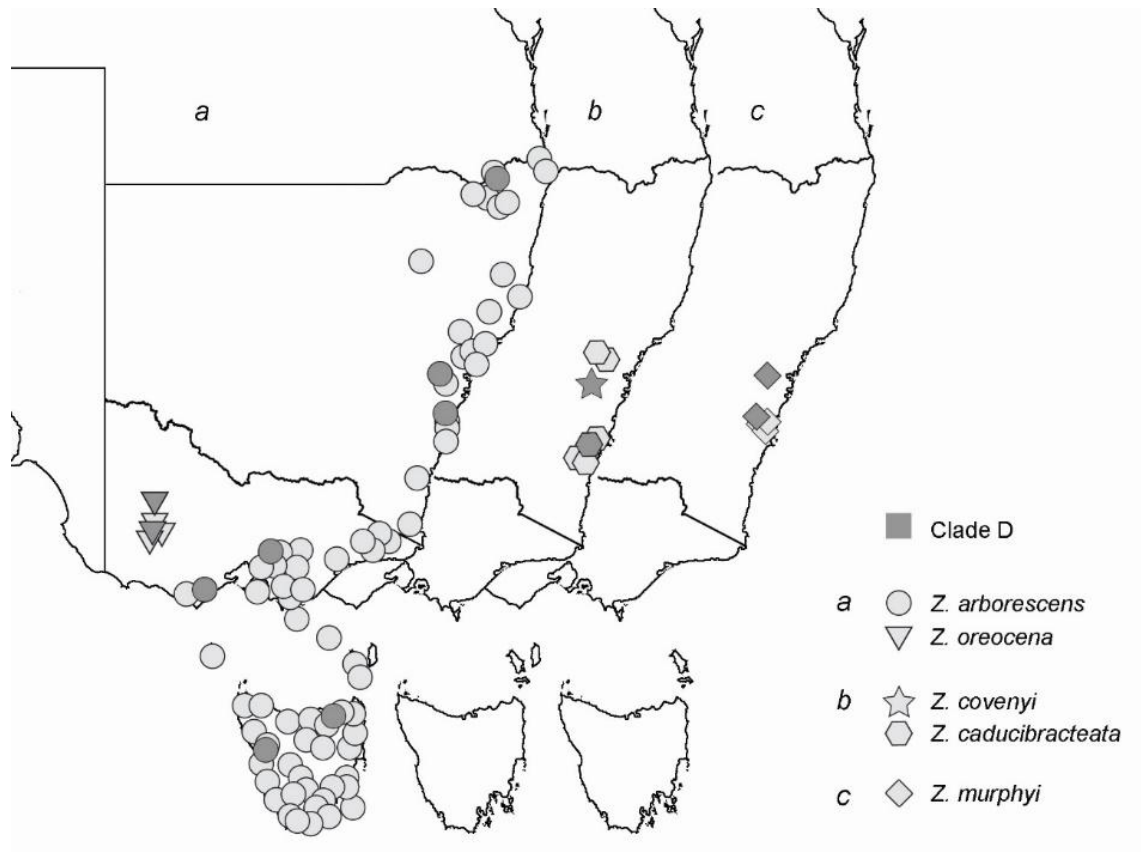


Figure 6.4. Distribution of *Zieria arborescens* and closely related taxa (light grey shapes), and the geographic location of samples resolved in clade D (dark grey shapes) of the nrDNA tree in Chapter 4, Figure 4.5.

In a broader context, based on combined ITS + ETS data (Chapter 4), eight accessions of *Z. arborescens*, were resolved in a single well-supported clade (D, Figure 6.4) with accessions of four more narrowly distributed species (*Z. caducibracteata*, *Z. covenyi*, *Z. murphyi* and *Z. oreocena*). The widespread distribution of *Z. arborescens* broadly overlaps that of *Z. caducibracteata*, with *Z. caducibracteata* effectively ‘replacing’ *Z. arborescens* in part of its range in southern New South Wales between Morton National Park and the Budawang. The two species are morphologically very similar with the main morphological differences being that *Z. caducibracteata* has young branches with a ‘soft’ indumentum (vs. ‘scurfy’ in *Z. arborescens*), triangular calyx

lobes (vs. ovate to lanceolate) and hirsute fruits (vs. glabrous). This morphological similarity and placement within the same nrDNA clade provides evidence that *Z. caducibracteata* is nested within *Z. arborescens*. In addition, in the cpDNA phylogeny (Chapter 3), one of the two samples of *Z. caducibracteata* shared a similar cpDNA haplotype to a sample of *Z. arborescens* (placement of the other sample was not well supported). It is therefore proposed that *Z. caducibracteata* should be sunk as a species or recognised at an infraspecific rank within *Z. arborescens* (which would require recognition of other infraspecific taxa based on monophyletic groups).

Zieria murphyi has a disjunct distribution in the Sydney Basin, New South Wales (Figure 6.4 c). The two samples in this study reflect this distribution, one being from the Blue Mountains, and the other from the population near Bundanoon, more than 120 km to the south. In the nrDNA phylogeny, both samples of *Z. murphyi* were nested in the clade containing samples of *Z. arborescens*, and were moderately supported in a trichotomy with one sample of *Z. arborescens* (MFD2035), also collected in the Blue Mountains, less than 10 km from the location of *Z. murphyi* (JA1356). In contrast to *Z. caducibracteata* (discussed above), *Z. murphyi* is morphologically distinct from *Z. arborescens*. *Zieria murphyi* is velutinous all over with heterophyllous leaves that are frequently unifoliate (vs. consistently trifoliate in *Z. arborescens*, the more common character state in the genus). *Zieria murphyi* shares these morphological characters and a similar chloroplast haplotype (in at least one sample) with samples of *Z. involucrata* (Chapter 3), and although these species were not resolved as sister in the combined ITS + ETS analysis, they were resolved in the same clade (C). The nesting of *Z. murphyi* within *Z. arborescens* in the nrDNA phylogeny is taxonomically challenging. Given its morphological distinctiveness it would not sit comfortably under a broad concept of *Z. arborescens*.

Also nested in the nrDNA clade containing *Z. arborescens* was a sample of the putative triploid hybrid, *Z. covenyi* (Armstrong, 2002), placed in a moderately supported trichotomy with two samples of *Z. caducibracteata*. *Zieria covenyi* only reproduces vegetatively, and is known from a single population on the Narrow Neck Peninsula in the Blue Mountains, New South Wales (Figure 6.4 b). The proposition by Armstrong (2002), based on some morphological characters, that *Z. covenyi* is a hybrid of *Z. caducibracteata* and *Z. cytisoides*, is only partly supported by these results. Both

nrDNA trees (BI and MP) and to a lesser extent cpDNA trees (Chapter 3) group *Z. covenyi* with samples of *Z. caducibracteata*, but show no close relationship to samples of *Z. cytisoides*. *Zieria covenyi* is distinct morphologically with leaves and floral bracts somewhat similar in size and shape to those of *Z. caducibracteata* (and *Z. arborescens*), but with a soft velutinous indumentum throughout, as in *Z. cytisoides* (and *Z. murphyi*). Retention of its taxonomic identity and conservation status is recommended (i.e., endangered under the EPBC Act 1999).

A clade containing four samples of *Z. oreocena* was also grouped with samples of *Z. arborescens* in clade D in the nrDNA tree (Figure 4.5). *Zieria oreocena* occurs only in the Grampians in Victoria whereas *Z. arborescens* is more widespread (Figure 6.4 a) but does not occur in the Grampians (George *et al.* 2013). The two taxa are morphologically very similar, with the main differences being that *Z. oreocena* (like *Z. arborescens* subsp. *glabrifolia*) has stems and leaves that are glabrous (vs. stellate-tomentose in the *Z. arborescens* type), and has triangular calyx lobes (vs. ovate to lanceolate). However, the indumentum of leaflets in *Z. arborescens* is highly variable, with many plants in Victoria, Tasmania and New South Wales also having glabrous abaxial leaf surfaces (George *et al.* 2013; Barrett pers. obs.). In addition, based on four chloroplast markers, one sample of *Z. oreocena* (DJO222) has a haplotype that is highly similar to three samples of *Z. arborescens*, varying by a single character (Chapter 3). The position of *Z. oreocena* on the basal polytomy of Clade D in the nrDNA tree allows the possibility that it could be sister to the rest of the clade, or that it could be nested within *Z. arborescens*. If it were clearly nested in *Z. arborescens* it should, given its strong morphological resemblance, be considered synonymous with *Z. arborescens* or potentially recognised at an infraspecific rank. However, given the uncertainty about its placement in the nrDNA tree, its monophyly in the nrDNA tree, and its allopatric distribution, it would arguably be better to retain *Z. oreocena* as a species distinct from *Z. arborescens* in the absence of stronger evidence to the contrary. It would seem expeditious to first address taxonomic issues in the more widespread and variable species, *Z. arborescens*.

Taxa in clade I with glandular-verrucose leaves

The nrDNA analysis resolved a well-supported clade (I), which consisted of five morphologically similar species with leaves that are glandular-verrucose on both surfaces. The five species, *Z. buxijugum*, *Z. formosa*, *Z. granulata*, *Z. parrisiae* and *Z. tuberculata*, are all endemic to New South Wales, and occur in the Central or South Coast regions of New South Wales (Figure 6.5). Three samples, *Z. buxijugum*, *Z. parrisiae* and *Z. formosa*, have identical ITS and ETS sequences, and each species is known from single endangered populations in relatively close proximity, between 6 and 15 km west of Pambula, New South Wales (Armstrong 2002). The southern-most population of *Z. tuberculata* occurs near Tilba Tilba, about 80 km to the north of these three species, and *Z. granulata* occurs about 200 km further north near Kiama. *Zieria buxijugum* and *Z. parrisiae* were both described in Armstrong (2002), and are particularly similar morphologically with slight differences in leaf margins, narrowness and shape of the leaves, length of inflorescences compared to leaves, floral bracts, indumentums of pedicels, calyx lobes and petals, and waxiness of seed surfaces.

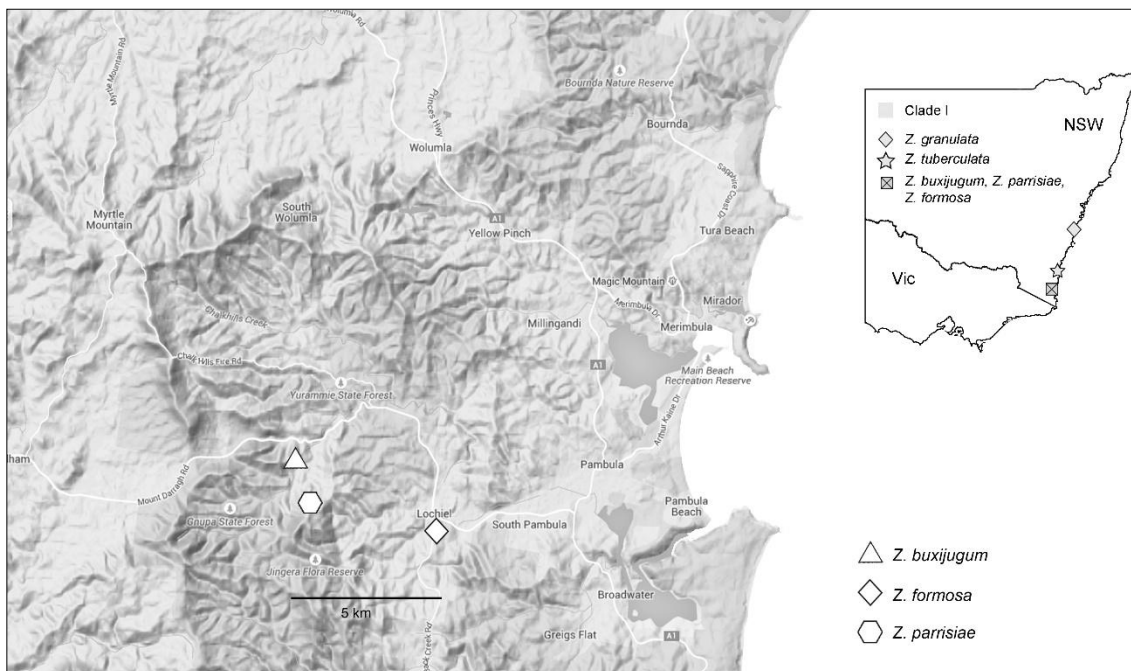


Figure 6.5. Distribution of samples in clade I (*Zieria buxijugum*, *Z. formosa*, *Z. granulata*, *Z. parrisiae* and *Z. tuberculata*). This map also indicates the (restricted) distribution of each species.

The close proximity of the populations of *Z. buxijugum* and *Z. parrisiae* (less than 2 km apart, see Figure 6.5), identical nrDNA, similar chloroplast haplotypes (differing by six

steps over 4001 base pairs) and similar morphology lead to the conclusion that they could be treated as conspecific. *Zieria formosa* is also morphologically very similar to *Z. buxijugum* and *Z. parrisiae* (and has an identical ITS and ETS sequence) but has anthers terminated by small white apiculum (versus non-apiculate in the other two species), petals that are commonly glabrous on the adaxial surface (versus moderate to sparsely pubescent), and often has broader leaves that are lighter-coloured adaxially. The status of all three species warrants more critical appraisal; detailed morphological comparisons and use of more variable molecular markers would be particularly useful given that reassessment of the taxonomic status of *Z. buxijugum*, *Z. parrisiae* and *Z. formosa* has implications for Federal conservation listings (see Chapter 1).

Relationships of Zieria verrucosa, Zieria vagans and Zieria distans

Zieria verrucosa, *Z. vagans* and *Z. distans* are narrowly distributed species that all occur in south-east Queensland within an area with a radius of about 180 km (Figure 6.6). They are not sympatric, but populations of the different species can occur within 15 km of each other, with *Z. verrucosa* and *Z. vagans* commonly in the understorey of semi-evergreen vine-thicket, and *Z. distans* always in rocky places in a range of habitats including rock pavements, tall open shrubland and fragmented semi-evergreen vine-thicket (Duretto and Forster 2007).

Zieria verrucosa was described by Armstrong (2002) and the other two species by Duretto and Forster (2007), who considered them all to be closely related based on morphology. Samples of the three species used in this study had different cpDNA haplotypes that each fell in different clades (Chapter 3), but the nrDNA analyses (Chapter 4) placed *Z. verrucosa* as sister to *Z. vagans* with strong support and identical ITS and ETS sequences (Figure 4.7) in samples collected about 70 km apart. The nrDNA analyses also resolved *Z. distans* as monophyletic, based on two samples from different populations, with different cpDNA haplotypes, and placed them on a large polytomy, separated from *Z. verrucosa* and *Z. vagans* by only one poorly supported node (see Figure 4.5, node 41; PP = 0.57, BS < 50%). Given the congruence of the nrDNA data with morphological resemblance it seems clear that *Z. verrucosa* and *Z. vagans* at least are very closely related. These nrDNA markers provide no characters for distinguishing the two species; without use of more variable molecular markers their

continued recognition rests on the morphological characters described by Duretto and Forster (2007), i.e., *Z. vagans* has narrow-elliptic leaflets that are not obviously glandular above (versus linear, strongly glandular verrucose leaflets in *Z. verrucosa*), leaves that are spreading or held erect (versus drooping) and sepals that are not obviously glandular (versus scattered glandular verrucose). The species have similar sized leaves and narrow-oblong inflorescence bracts.

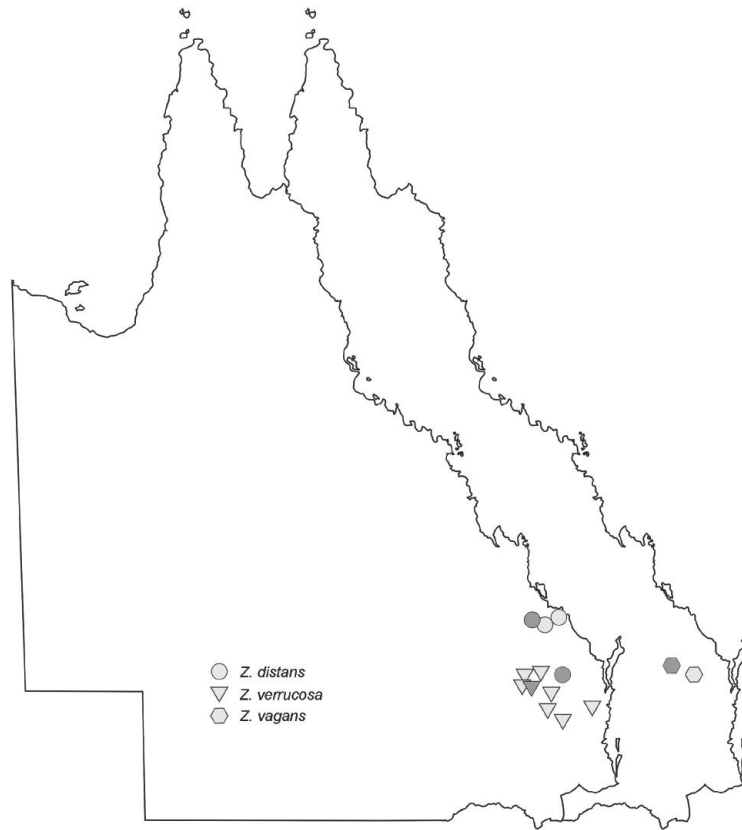


Figure 6.6. Distribution of *Zieria verrucosa*, *Z. vagans* and *Z. distans*. Symbols representing the samples used in this study are shaded dark grey.

Zieria minutiflora and related taxa

The nrDNA analysis resolved a strongly supported clade (J, Figure 4.5), which consisted of *Z. minutiflora* (two subspecies), paralogues of *Z. exsul* and one sample of *Z. obovata*. Samples of *Zieria minutiflora* subsp. *trichocarpa* and *Z. obovata* (from populations less than 12 km apart in north Queensland) differed by a single nucleotide and were resolved as sisters with five character changes and strong statistical support. *Zieria minutiflora* and *Z. obovata* were shown to have different cpDNA haplotypes (Chapter 3) but they are morphologically very similar and often confused (Armstrong

2002). Duretto and Forster (2007) considered them to be possibly conspecific, with the main distinguishing morphological character being that *Z. obovata* typically has inflorescences that are longer than the leaves whereas in *Z. minutiflora* they are usually shorter (Duretto and Forster 2007). The results of the nrDNA analysis in this study are consistent with a close relationship between the species and do not provide grounds for their taxonomic distinctiveness. It is reasonable to suggest that *Z. obovata* should be sunk as a species, or recognised at an infraspecific rank, within *Z. minutiflora*. (A possible explanation for nrDNA paralogues from *Z. exsul* also being placed in this clade is outlined in Chapter 4, section 4.4.1 *Significance of divergent nrDNA paralogues.*)

6.5. Biogeography – eastern Australia and New Caledonia

There is continuing debate about the age and history of the New Caledonian biota, how best to explain the numerous biogeographic connections between Australia and New Caledonia, and why there is a high level of endemism in the extant flora of New Caledonia (Lowry 1998; Ladiges and Cantrill 2007). In this study, phylogenetic relationships of the single endemic New Caledonian species, *Z. chevalieri*, differed between cpDNA and nrDNA trees. The cpDNA phylogeny (Chapter 3) placed *Z. chevalieri* as sister to all other species in the genus, suggesting an early divergence of New Caledonian and Australian lineages, and differentiation of species through vicariance (Bauer *et al.* 2006; Ladiges and Cantrill 2007; Heads 2008). In contrast, *Z. chevalieri* was nested higher in the nrDNA tree with *Z. fraseri*, *Z. exsul* and *Z. laxiflora* (Chapter 4). This may be interpreted as supporting more recent, over-water dispersal, as argued by some for other taxa based on molecular clock analyses (Grandcolas *et al.* 2008; Swenson *et al.* 2014). However, the phylogenetic position of *Z. chevalieri* is equivocal, the age of *Zieria* and molecular estimates of clade divergences are questionable (Bayly *et al.* 2013), the geological history of Zealandia and New Caledonia is still to be fully documented, and any long-distance dispersal hypothesis would have to account for the very limited capacity for seed dispersal of *Zieria*, which relies on ants (Armstrong 2002). The results here do not favour one hypothesis over another, as discussed in detail in Chapters 3 and 4; thus the debate is likely to continue over the history of biogeographic connections between New Caledonian and Australian taxa, and hypotheses of long distance dispersal or vicariance.

6.6. Project assessment

The project was successful overall, with most of the original aims (stated in Chapter 1 and in italics below) being addressed to a greater or lesser degree, as reviewed briefly below.

- *Produce a comprehensive molecular phylogeny of Zieria in Australia and New Caledonia, based on chloroplast and nuclear DNA data*

Phylogenies of *Zieria* were produced based on four cpDNA markers (Chapter 3) and two nrDNA markers (Chapter 4). All but one or two recognised species were included in each analysis, making them currently the most comprehensive phylogenetic studies of this genus to date.

- *Assess the monophyly of taxa and consider issues of taxonomy suggested by the results of this study*

The monophyly of species was assessed based on both datasets. The nrDNA provided greater support for monophyly of species than cpDNA, potentially providing a better indication of phylogeny than cpDNA. Implications for current species-level taxonomy were discussed in Chapter 6, section 6.2.

- *Discern patterns of speciation, e.g. relationships of widespread and narrowly endemic taxa*

Patterns of speciation were confounded in the cpDNA data by evolutionary processes such as deep coalescence and incomplete lineage sorting, resulting in incongruence between cpDNA and nrDNA datasets and current taxonomy. Possible speciation patterns for some taxa were discussed in relation to the results of the nrDNA analysis.

- *Examine patterns of variation in widespread species*

Variation was detected in both cpDNA and nrDNA in some widespread species (Chapters 3 and 4), which were discussed in Chapter 6 in the context of the molecular phylogeny of the genus and its implications for taxonomy, e.g. *Z. smithii*, *Z. arborescens*, *Z. furfuracea*.

- *Consider the relationships and distinctiveness of highly restricted (and rare) species that will allow the validity of such taxa, and their conservation value, to be assessed*

The relationships of some highly restricted taxa were considered in relation to the results of this study: e.g. *Z. alata* and *Z. madida*, *Z. buxijugum*, *Z. parrisiae* and *Z. formosa*, *Z. prostrata*, *Z. hydroscopica*. The south-eastern Queensland species, *Z. montana*, once broadly circumscribed with a disjunct distribution, but now more narrowly circumscribed, was also discussed in the context of phylogenetic resolution of its segregates.

- *Infer biogeographic connections between Australia and New Caledonia, with reference to the molecular phylogenetic relationships between eastern Australian *Zieria* species and the single New Caledonia species*

Biogeographic links between Australia and New Caledonia were discussed in Chapters 3 and 4, and some alternative hypotheses were suggested to explain the conflicting placement of the New Caledonian species, *Z. chevalieri*, in the cpDNA and nrDNA trees.

- *Investigate genetic variation and phylogeographic patterns in the widespread species, *Zieria arborescens*, focussing on Victorian and Tasmanian populations*

A phylogeographic study of *Z. arborescens* is presented in Chapter 5. Chloroplast DNA was more variable than nrDNA, which proved to be of limited value at this level of investigation.

6.7. Conclusion and prospects for further research

Through this study, insights into the evolutionary history and systematics of *Zieria* have been gained, although the widespread incongruence between the nrDNA and cpDNA data made inferring relationships in *Zieria* challenging. The nrDNA tree presented in Chapter 4 potentially represents a hypothesis for a species tree in *Zieria*, whereas the observed patterns of relationships in the cpDNA tree (Chapter 3) largely reflect evolutionary processes other than those of phylogenetic descent, such as hybridisation, introgression between species, and incomplete lineage sorting, a combination of which

potentially explains the incongruence seen between competing (nrDNA and cpDNA) phylogenies.

In instances where taxa were shown to be monophyletic using nuclear markers, current taxonomy is supported. However, some taxa were found to be in conflict with current species-level taxonomy, i.e., they were resolved as polyphyletic by nuclear or chloroplast markers. One issue may be previous over-splitting of taxa based on minor morphological characters. In this project, those taxa resolved as polyphyletic by nrDNA in particular, drew attention to the need for at least some taxonomic revision within this genus (as discussed above). The presence of divergent paralogous of ITS and ETS shed light on the genetic history of some individuals, species or the population from which they were sampled. It also reinforces the need for caution when using these markers for phylogenetic analysis of flowering plants, and the necessity for careful consideration of paralogy.

This project provides an important framework for further investigation of the genus *Zieria*. Future research should focus efforts on specific areas that have been identified throughout this project, which could include:

- Molecular and morphological variation in *Z. smithii* and allied taxa.
- A taxonomic assessment of taxa currently placed in *Z. furfuracea* (three subspecies currently recognised), which should be determined in conjunction with related populations of *Z. smithii*.
- The hypothesis that *Z. hydroscopica* is the result of hybridisation; sampling should include *Z. smithii*, *Z. cytisoides* and *Z. hydroscopica* from nearby localities.
- The complex relationships of *Z. prostrata*, which should include samples of *Z. smithii* and *Z. arborescens* from nearby populations and consider the history of divergent paralogues detected in samples of *Z. prostrata* and *Z. smithii* (and more rarely in *Z. arborescens*).
- Further investigation of *Z. alata* and *Z. madida* including more sampling, and the use of more variable molecular markers to test whether there are consistent genetic differences between them.

- A taxonomic treatment of *Z. arborescens*, including molecular and morphological data, followed by a critical reappraisal of the taxonomic boundaries of closely related taxa: *Z. caducibracteata*, *Z. covenyi*, *Z. murphyi* and *Z. oreocena*.
- A critical taxonomic appraisal of *Z. buxijugum*, *Z. parrisiae* and *Z. formosa*, including detailed morphological comparisons and the use of more variable molecular markers.
- Further investigation into the close relationship of three morphologically similar species, *Z. verrucosa*, *Z. vagans* and *Z. distans*, focussing on the phylogenetic relationships of *Z. distans*.
- The hypothesis that *Z. obovata* and *Z. minutiflora* are conspecific or at least very closely related, based on more sampling and more informative nuclear markers.
- Phylogeographic patterns and history of areas in eastern Australia, and resolution of the Australia–New Caledonia connection.

Chapter 7. List of References

- Álvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29**, 417–434.
- Amar MH, Hassan AHM, Biswas MK, Dulloo E, Xie Z-Z, Guo W-W (2014) Maximum parsimony based resolution of inter-species phylogenetic relationships in *Citrus* L. (Rutaceae) using ITS of rDNA. *Biotechnology and Biotechnological Equipment* **28**, 61–67.
- Appelhans MS, Wen J, Wagner WL (2014) A molecular phylogeny of *Acronychia*, *Euodia*, *Melicope* and relatives (Rutaceae) reveals polyphyletic genera and key innovations for species richness. *Molecular Phylogenetics and Evolution* **79**, 54–68.
- Appelhans MS, Wen J, Wood KR, Allan GJ, Zimmer EA, Wagner WL (2013) Molecular phylogenetic analysis of Hawaiian Rutaceae (*Melicope*, *Platydesma* and *Zanthoxylum*) and their different colonization patterns. *Botanical Journal of the Linnean Society*. **174**, 425–48.
- Armstrong JA (1979) Biotic pollination mechanisms in the Australian flora: a review. *New Zealand Journal of Botany* **17**, 467–508.
- Armstrong JA (1991) Studies on pollination and systematics in the Australian Rutaceae. PhD thesis. University of New South Wales, Sydney.
- Armstrong JA (2002) *Zieria* (Rutaceae): a systematic and evolutionary study. *Australian Systematic Botany* **15**, 277 – 463.
- Armstrong JA, Harden G (2002) *Zieria*. In ‘Flora of New South Wales. Volume 2’. 2nd Edition. (Ed G Harden.) pp 277–288 (UNSW Press: Kensington)
- Armstrong JA, Powell JM (1980) *Neobyrnnesia* (Rutaceae): a new genus endemic to Northern Australia. *Telopea* **1**, 399–408.
- Arnheim (1983) Concerted evolution in multigene families. In ‘Evolution of genes and proteins’. (Eds M Nei and R Koehn). pp. 38–61 (Sinauer Associates: Sunderland).
- Auld TD (2001) The ecology of the Rutaceae in the Sydney region of south-eastern Australia: Poorly known ecology of a neglected family. *Cunninghamia* **7**, 213–239.

References

- Auld TD, Keith DA, Bradstock RA (2000) Patterns in longevity of soil seedbanks in fire-prone communities of south-eastern Australia. *Australian Journal of Botany* **48**, 539–548.
- Avice JC (2000) ‘Phylogeography: the history and formation of species’. (Harvard University Press: Cambridge, Massachusetts).
- Avice JC (2009) Phylogeography: retrospect and prospect. *Journal of biogeography* **36**, 3–15.)
- Avice JC (1986) Mitochondrial DNA and the evolutionary genetics of higher animals. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **312**, 325–342.
- Bache F, Sutherland R, Stagpoole V, Herzer R, Collot J, Rouillard P (2012) Stratigraphy of the southern Norfolk Ridge and the Reinga Basin: A record of initiation of Tonga–Kermadec–Northland subduction in the southwest Pacific. *Earth and Planetary Science Letters* **321–322**, 41–53.
- Bailey CD, Carr TG, Harris SA, Hughes CE (2003) Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. *Molecular Phylogenetics and Evolution* **29**, 435–455.
- Baillie PW (1989) Jurassic–Cainozoic. In ‘Geology and mineral resources of Tasmania’. (Ed CR Burrett and EL Martin) pp. 339–409. (Geological Society of Australia: Melbourne).
- Baldwin BG (1992) Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the *Compositae*. *Molecular Phylogenetics and Evolution* **1**, 3–16.
- Baldwin BG, Markos S (1998) Phylogenetic Utility of the external transcribed spacer (ETS) of 18S–26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution* **10**, 449–463.
- Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995) The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* **82**, 247–277.

- Barlow BA (1981) The Australian flora: its origin and evolution. *Flora of Australia* **1**, 25–75
- Barker FK, Lutzoni FM (2002) The utility of the incongruence length difference test. *Systematic Biology* **51**, 625–637.
- Barrett RA, Bayly MJ, Duretto MF, Forster PI, Ladiges PY, Cantrill DJ (2015) A chloroplast phylogeny of *Zieria* (Rutaceae) in Australia and New Caledonia shows widespread incongruence with species-level taxonomy. *Australian Systematic Botany* **27**, 427–449.
- Barrows T, Stone J, Fifield L, Cresswell R (2001) Late Pleistocene glaciation of the Kosciuszko Massif, Snowy Mountains, Australia. *Quaternary Research* **55**, 179–189.
- Bauer A, Jackman T, Sadler R, Whitaker A (2006) A revision of the *Bavayia validiclavis* group (Squamata: Gekkota: Diplodactylidae), a clade of New Caledonian geckos exhibiting microendemism. *Proceedings of the California Academy of Science, Fourth Series* **57**, 503–547.
- Bausher MG, Singh ND, Lee S-B, Jansen RK, Daniell H (2006) The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var ‘Ridge Pineapple’: organization and phylogenetic relationships to other angiosperms. *BMC Plant Biology* **6**, 21–21.
- Bayer RJ, Mabberley DJ, Morton C, Miller CH, Sharma IK, Pfeil BE, Rich S, Hitchcock R, Sykes S (2009) A molecular phylogeny of the orange subfamily (Rutaceae: Aurantioideae) using nine cpDNA sequences. *American Journal of Botany* **96**, 668–685.
- Bayes T. (1763) An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F.R.S. communicated by Mr. Price, in a letter to John Canton, A.M.F.R.S. *Philosophical Transactions* **53**, 370–418.
- Bayly MJ, Duretto, MF, Holmes, GD, Forster PI, Cantrill DJ, Ladiges, PY (2015) Transfer of the New Caledonian genus *Boronella* to *Boronia* (Rutaceae) based on analyses of cpDNA and nrDNA. *Australian Systematic Botany* **28**, 111–23.

References

- Bayly MJ, Holmes GD, Forster PI, Cantrill DJ, Ladiges PY (2013) Major clades of Australasian Rutoideae (Rutaceae) based on *rbcL* and *atpB* sequences. *PLoS ONE* **8**, e72493.
- Bayly MJ, Ladiges PY (2007) Divergent paralogues of ribosomal DNA in eucalypts (Myrtaceae). *Molecular Phylogenetics and Evolution* **44**, 346–356.
- Bayly MJ, Udovicic F, Gibbs AK, Parra-O. C, Ladiges PY (2008) Ribosomal DNA pseudogenes are widespread in the eucalypt group (Myrtaceae): implications for phylogenetic analysis. *Cladistics* **24**, 131–146.
- Beheregaray L (2008) Twenty years of phylogeography: the state of the field and the challenges for the southern hemisphere. *Molecular Ecology* **7**, 3754–3774.
- Bena G, Jubier M-F, Olivieri I, Lejeune B (1998) Ribosomal external and internal transcribed spacers: combined use in the phylogenetic analysis of *Medicago* (Leguminosae). *Journal of molecular evolution* **46**, 299–306.
- Bentham G (1863) ‘Flora Australiensis: a description of the plants of the Australian territory. Volume 1’. (Lovell Reeve and Co.: London)
- Bergsten J (2005) A review of long-branch attraction. *Cladistics* **21**, 163–193.
- Bonatelli I, Zappi D, Taylor N, Moraes E (2013) Usefulness of cpDNA markers for phylogenetic and phylogeographic analyses of closely related cactus species. *Genetics and Molecular Research* **12**, 4579–4585.
- Bonpland AJA (1815) ‘Description des plantes rares cultivees “a Malmaison” et “a Navarre”, part 5.’ (Didot: Paris)
- Boni MF, Posada D, Feldman MW (2007) An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* **176**, 1035–1047.
- Buckler ES, Ippolito A, Holtsford TP (1997) The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics* **145**, 821–832.
- Burke A (2003) Inselbergs in a changing world—global trends. *Diversity and Distributions* **9**, 375–383.
- Byrne M (2008) Evidence for multiple refugia at different time scales during Pleistocene climatic oscillations in southern Australia inferred from phylogeography. *Quaternary Science Reviews* **27**, 2576–2585.

- Byrne M, Steane DA, Joseph L, David K, Yeates DK, Jordan GJ, Crayn D, Aplin K, Cantrill DJ, Cook LG, Crisp MD, Keogh JS, Melville J, Moritz C, Porch N, Kale Sniderman JM, Sunnucks P and Weston PH (2011) Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. *Journal of Biogeography*, **38**, 9, 1635–1656.
- Chase MW, Morton CM, Kallunki JA (1999) Phylogenetic relationships of Rutaceae: a cladistic analysis of the subfamilies using evidence from *rbcL* and *atpB* sequence variation. *American Journal of Botany* **86**, 1191–1199.
- Childs G, Maxson R, Cohn RH, Kedes L (1981) Orphans: dispersed genetic elements derived from tandem repetitive genes of eucaryotes. *Cell* **23**, 651–663.
- Clarke and Lee (1987) ‘Name that flower: the identification of flowering plants’ (Melbourne University Press: Melbourne)
- Clegg MT, Brandon SG, Learn GH Jr., Morton BR (1994) Rates and Patterns of Chloroplast DNA Evolution. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 6795–6801.
- Clevinger JA, Panero JL (2000) Phylogenetic Analysis of *Silphium* and subtribe Engelmanniinae (Asteraceae: Heliantheae) based on ITS and ETS sequence data. *American Journal of Botany* **87**, 565–572.
- Colhoun EA, Hannan D, Kiernan K (1996) Late Wisconsin glaciation of Tasmania. *Papers and Proceedings of the Royal Society of Tasmania* **130**, 33–45.
- Crisp M, Cook L, Steane D (2004) Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**, 1551–1571.
- Crisp MD, Laffan S, Linder HP, Monro A (2001) Endemism in the Australian flora. *Journal of Biogeography*, **28**, 183–98.
- Cunningham CW (1997) Is congruence between data partitions a reliable predictor of phylogenetic accuracy? Empirically testing an iterative procedure for choosing among phylogenetic methods. *Systematic Biology* **46**, 464–478.
- Daniell H, Lee S-B, Grevich J, Saski C, Quesada-Vargas T, Guda C, Tomkins J, Jansen RK (2006) Complete chloroplast genome sequences of *Solanum bulbocastanum*,

References

- Solanum lycopersicum* and comparative analyses with other Solanaceae genomes. *Theoretical and Applied Genetics*. **112**, 1503–1518.
- Darlu P, Lecointre G (2002) When does the incongruence length difference test fail? *Molecular Biology and Evolution* **19**, 432–437.
- Davis C, Xi Z, Mathews S (2014) Plastid phylogenomics and green plant phylogeny: almost full circle but not quite there. *BMC Biology* **12**, 11.
- De Beuque SV, Auzende J-M, Lafoy Y, Missegue F (1998) Tectonique et volcanisme tertiaire sur la ride de Lord Howe (Sud-Ouest Pacifique). *Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science* **326**, 663–669.
- Dolphin K, Belshaw R, Orme CDL, Quicke DLJ (2000) Noise and incongruence: interpreting results of the incongruence length difference test. *Molecular Phylogenetics and Evolution* **17**, 401–406.
- Dowton M, Austin AD (2002) Increased congruence does not necessarily indicate increased phylogenetic accuracy—the behavior of the incongruence length difference test in mixed-model analyses. *Systematic Biology* **51**, 19–31.
- Doyle JJ (1992) Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany* **17**, 144–163.
- Doyle JJ, Davis JI (1998) Homology in molecular phylogenetics: a parsimony perspective. In 'Molecular systematics of plants II: DNA sequencing.' (Eds DE Soltis, PS Soltis, JJ Doyle.) pp. 101–131. (Springer Science+Business Media: New York)
- Drinnan AN, Crane PR, Hoot SB (1994) 'Patterns of floral evolution in the early diversification of non-magnoliid dicotyledons (eudicots).' (Springer: Vienna)
- Duretto MF (1999) Rutaceae. In 'Flora of Victoria. Volume 4.' (Eds N Walsh, T Entwistle.) pp 153–197. (Inkata Press: Melbourne)
- Duretto MF, Forster PI (2007) A taxonomic revision of the genus *Zieria* Sm. (Rutaceae) in Queensland. *Austrobaileya* **7**, 473–544.
- Duretto MF, Forster PI (2008) New subspecies for *Zieria odorifera* J.A. Armstr. (Rutaceae) from northern New South Wales. *Austrobaileya* **7**, 681–690.

- Engler A (1931) Rutaceae. In ‘Die natürlichen Pflanzenfamilien Teil 19a’, 2nd edition (Eds HGA Engler and K Prantl). pp. 187–359. (Wilhelm Engelmann, Leipzig, Germany)
- Erixon P, Svennblad B, Britton T, Oxelman B (2003) Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Systematic Biology* **52**, 665–73.
- Exon NF, Hill PJ, Lafoy Y, Heine C, Bernardel G (2006) Kenn Plateau off northeast Australia: a continental fragment in the southwest Pacific jigsaw. *Australian Journal of Earth Sciences* **53**, 541–564.
- Farris JS (1983) The logical basis of phylogenetic analysis. *Advances in cladistics* **2**, 7–36.
- Farris JS, Källersjö M, Kluge AG, Bult C (1995a) Constructing a significance test for incongruence. *Systematic Biology* **44**, 570–2.
- Farris JS, Källersjö M, Kluge AG, Bult C (1995b) Testing significance of incongruence. *Cladistics* **10**, 315–319.
- Farris JS, Kluge AG (1985) Parsimony, synapomorphy, and explanatory power: a reply to Duncan. *Taxon* **34**, 130–135.
- Fehrer J, Krak K, Chrtek J (2009) Intra-individual polymorphism in diploid and apomictic polyploid hawkweeds (*Hieracium*, Lactuceae, Asteraceae): disentangling phylogenetic signal, reticulation, and noise. *BMC Evolutionary Biology* **9**, 239.
- Feliner GN, Rosselló JA (2007) Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution* **44**, 911–919.
- Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* **27**, 401–410.
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.
- Ferris C, King RA, Väinölä R, Hewitt GM (1998) Chloroplast DNA recognizes three refugial sources of European oaks and suggests independent eastern and western immigrations to Finland. *Heredity* (Edinburgh) **80** (Pt 5), 584–93.

References

- Fitch WM (1970) Distinguishing homologous from analogous proteins. *Systematic Zoology* **19**, 99–113.
- Fletcher BS, Bateman MA, Hart NK, Lamberton JA (1975) Identification of a fruit fly attractant in an Australian plant, *Zieria smithii*, as -O-methyl eugenol. *Journal of economic entomology* **68**, 815–816.
- ‘Flora of Australia: Meliaceae, Rutaceae, Zygophyllaceae’. (2013) Volume 26 (ABRS/CSIRO: Melbourne).
- Floyd AG (1976) Effect of burning on regeneration from seeds in wet sclerophyll forest. *Australian Forestry*, **39**, 210–220.
- Flynn TM, Southwell IA (1987a) Essential oil constituents of the genus *Zieria*. *Phytochemistry* **26**, 1673–1686
- Flynn TM, Southwell IA (1987b) Cyanogenesis in the genus *Zieria*. *Phytochemistry* **26**, 1669–1672.
- Forey PL (1992) ‘Cladistics: a practical course in systematics.’ (Clarendon Press: Oxford)
- Freeman JS, Jackson HD, Steane DA, McKinnon GE, Dutkowski GW, Potts BM, Vaillancourt RE (2001) Chloroplast DNA phylogeography of *Eucalyptus globulus*. *Australian Journal of Botany*, **49**, 585–596.
- French PA, Brown GK, Bayly MJ (2016) Incongruent patterns of nuclear and chloroplast variation in *Correa* (Rutaceae): introgression and biogeography in south-eastern Australia. *Plant Systematics and Evolution*, (published early online) 1–22.
- Garcia-Jacas N, Soltis PS, Font M, Soltis DE Vilatersana R, Susanna A (2009) The polyploid series of *Centaurea toletana*: Glacial migrations and introgression revealed by nrDNA and cpDNA sequence analyzes. *Molecular Phylogenetics and Evolution* **52**, 377–394.
- George AS, Duretto MF, Forster PI (2013) *Zieria*. *Flora of Australia: Meliaceae, Rutaceae, Zygophyllaceae* **26**, 282–336.
- Gibbs AK, Udovicic FB, Drinnan AN, Ladiges P (2009) Phylogeny and classification of *Eucalyptus* subgenus *Eudesmia* (Myrtaceae) based on nuclear ribosomal DNA, chloroplast DNA and morphology. *Australian Systematic Botany* **22**, 158–179.

- Gibbs MJ, Armstrong JS, Gibbs AJ (2000) Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* **16**, 573–582.
- Gielly L, Taberlet P (1994) The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbcL* sequences. *Molecular Biology and Evolution* **11**, 769–777.
- Grandcolas P, Murienne J, Robillard T, Desutter-Grandcolas L, Jourdan H, Guilbert E, Deharveng L (2008) New Caledonia: a very old Darwinian Island? *Philosophical Transactions of the Royal Society of London. B. Biological Sciences* **363**, 3309–3317.
- Griffin SG, Leach DN, Markham J (1998) Antimicrobial activity of essential oils from *Zieria*. *Journal of essential oil research* **10**, 165–174.
- Groppo M, Kallunki J, Pirani J, Antonelli A (2012) Chilean *Pitavia* more closely related to Oceania and Old World Rutaceae than to Neotropical groups: evidence from two cpDNA non-coding regions, with a new subfamilial classification of the family. *PhytoKeys* **19**, 9–29.
- Groppo M, Pirani JR, Salatino MLF, Blanco SR, Kallunki JA (2008) Phylogeny of Rutaceae based on two noncoding regions from cpDNA. *American Journal of Botany* **95**, 985–1005.
- Hamby and Zimmer (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In ‘Molecular systematics of plants’ (Eds P Soltis, D Soltis, and J Doyle) pp. 50–91 (Chapman and Hall: New York).
- Harbaugh DT, Wagner WL, Allan GJ, Zimmer EA (2009) The Hawaiian Archipelago is a stepping stone for dispersal in the Pacific: an example from the plant genus *Melicope* (Rutaceae). *Journal of Biogeography* **36**, 230–241.
- Harden G (Ed) (2002) ‘Flora of New South Wales.’ (UNSW University Press: Kensington)
- Harpke D, Peterson A (2008) 5.8S motifs for the identification of pseudogenic ITS regions. *Botany* **86**, 300–305.
- Harrington MG, Gadek PA (2004) Molecular systematics of the *Acmena* alliance (Myrtaceae): phylogenetic analyses and evolutionary implications with reference to Australian taxa. *Australian Systematic Botany* **17**, 63–72.

References

- Hartley TG (2013) Rutaceae. *Flora of Australia: Meliaceae, Rutaceae, Zygophyllaceae*, **26**, 43.
- Heads M (2008) Biological disjunction along the West Caledonian Fault, New Caledonia: a synthesis of molecular phylogenetics and pan biogeography. *Botanical Journal of the Linnean Society* **158**, 470–488.
- Heath L, van der Walt E, Varsani A, Martin DP (2006) Recombination patterns in aphthoviruses mirror those found in other picornaviruses. *Journal of Virology* **80**, 11827–11832.
- Hennig W (1965) Phylogenetic systematics. *Annual review of entomology* **10**, 97–116.
- Hennig W (1966) 'Phylogenetic systematics.' (University of Illinois Press: Urbana)
- Hill RS, Truswell EM, McLoughlin S, Dettmann M (1999) Evolution of the Australian flora: fossil evidence. In 'Flora of Australia, Volume 1, 2nd edition'. pp 251–320 (ABRS/CSIRO: Melbourne).
- Hill RS (2004) Origins of the south-eastern Australian vegetation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**, 1537–1549.
- Hipp AL, Hall JC, Sytsma KJ (2004) Congruence versus phylogenetic accuracy: revisiting the incongruence length difference test. *Systematic Biology* **53**, 81–89.
- Hogbin P, Crisp M (2003) Evolution of the coastal neospecies *Zieria prostrata* (Rutaceae) and its relationship to the *Zieria smithii* species complex. *Australian Systematic Botany* **16**, 515–525.
- Hogbin PM, Peakall R (1999) Evaluation of the contribution of genetic research to the management of the endangered plant *Zieria prostrata*. *Conservation Biology* **13**, 514–522.
- Hogbin PM, Peakall R, Sydes MA (2000) Achieving practical outcomes from genetic studies of rare Australian plants. *Australian Journal of Botany* **48**, 375–382.
- Holmes GD, Downing TL, James EA, Blacket MJ, Hoffman AA, Bayly MJ (2014) Phylogeny of the holly grevilleas (Proteaceae) based on nuclear ribosomal and chloroplast DNA. *Australian Systematic Botany* **27**, 56–77.
- Hong-Wa C, Besnard G (2013) Intricate patterns of phylogenetic relationships in the olive family as inferred from multi-locus plastid and nuclear DNA sequence

- analyses: a close-up on *Chionanthus* and *Noronhia* (Oleaceae). *Molecular Phylogenetics and Evolution* **67**, 367–378.
- Hope G (1994) Quaternary vegetation. In: ‘History of the Australian Vegetation: Cretaceous to Recent.’ pp. 368–388, (Ed Hill R). (Cambridge University Press: London).
- Hubbard NN (1995) In search of regional palaeoclimates – Australia, 18,000 yr BP. *Palaeogeography, Palaeoclimatology, Palaeoecology* **116**, 167–188.
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.
- Hugall AF, Foster R, Hutchinson M, Lee MS (2008) Phylogeny of Australasian agamid lizards based on nuclear and mitochondrial genes: implications for morphological evolution and biogeography. *Biological Journal of the Linnean Society* **93**, 343–58.
- Hughes, CE, Eastwood, RJ, Bailey, CD (2006) From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical Transactions: Biological Sciences* **361**, 211–225.
- Islam S, Ahsan M (1997) Biological activities of the secondary metabolites isolated from *Zieria smithii* and *Zanthoxylum elephantiasis* on microorganisms and brine shrimps. *Phytotherapy Research* **11**, 64–66.
- Jackson G (1810) *Zieria smithii*. In ‘The botanist's repository, for new and rare plants’. Vol. 9, t. 606. (Ed HC Andrews) (H. Andrews: London)
- Jackson, WD (1999) Paleohistory of vegetation change: the last 2 million years. In ‘Vegetation of Tasmania’ (Eds JB Reid, RS Hill, MJ Brown, MJ Hovenden) pp. 64–88. Flora of Australia Supplementary Series no. 8. (Australian Biological Resource Study: Canberra).
- Jackson S, Overpeck J (2000) Responses of plant populations and communities to environmental changes of the late quaternary. *Paleobiology* **26**, 194–220.
- Jaffre TA, Bouchet P, Veillon JM (1998) Threatened plants of New Caledonia: Is the system of protected areas adequate? *Biodiversity and conservation* **7**, 109–135.

References

- Jobes DV, Thien LB (1997) A conserved motif in the 5.8 S ribosomal RNA (rRNA) gene is a useful diagnostic marker for plant internal transcribed spacer (ITS) sequences. *Plant Molecular Biology Reporter* **15**, 326–334.
- Judd WS, Olmstead RG (2004) A survey of tricolpate (eudicot) phylogenetic relationships. *American Journal of Botany* **91**, 1627–44.
- Käss E, Wink M (1997) Molecular phylogeny and phylogeography of *Lupinus* (Leguminosae) inferred from nucleotide sequences of the *rbcL* gene and ITS 1 + 2 regions of rDNA. *Plant Systematics and Evolution* **208**, 139–167.
- Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden* **87**, 482–498.
- Kellermann J, Udovicic F (2008) Large indels obscure phylogeny in analysis of chloroplast DNA (*trnL*–F) sequence data: Pomaderreae (Rhamnaceae) revisited. *Telopea* **12**, 1–22.
- King MG, Roalson EH (2008) Exploring evolutionary dynamics of nrDNA in *Carex* Subgenus *Vignea* (Cyperaceae). *Systematic Botany* **33**, 514–524.
- Kitching IJ (1998) ‘Cladistics: the theory and practice of parsimony analysis.’ (Oxford University Press: New York).
- Kubitzki K, Kallunki JA, Duretto M, Wilson P (2011) Rutaceae. In ‘The Families and Genera of Vascular Plants.’ (Ed K Kubitzki.) Vol. 10 pp. 276–356. (Springer: Berlin Heidelberg).
- Ladiges P, Cantrill D (2007) New Caledonia–Australian connections: biogeographic patterns and geology. *Australian Systematic Botany* **20**, 383–389.
- Ladiges PY, Bayly MJ, Nelson GJ (2010) East-west continental vicariance in *Eucalyptus* subgenus *Eucalyptus* (Myrtaceae). In ‘Beyond cladistics: the branching of a paradigm’ (Eds DM Williams and S Knapp) pp 267–302 (University of California Press: Berkeley, California).
- Li F, Fan Q, Li Q, Chen S, Guo W, Cui D, Liao W (2014) Molecular phylogeny of *Cotoneaster* (Rosaceae) inferred from nuclear ITS and multiple chloroplast sequences. *Plant Systematics and Evolution* **300**, 1533–1546.
- Linder CR, Goertzen LR, Heuvel BV, Francisco-Ortega J, Jansen RK (2000) The complete external transcribed spacer of 18S–26S rDNA: amplification and

- phylogenetic utility at low taxonomic levels in Asteraceae and closely allied families. *Molecular Phylogenetics and Evolution* **14**, 285–303.
- Liston A, Robinson WA, Oliphant JM, Alvarez-Buylla ER (1996) Length variation in the nuclear ribosomal DNA internal transcribed spacer region of non-flowering seed plants. *Systematic Botany* 109–120.
- Liu J-S, Schardl CL (1994) A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. *Plant Molecular Biology* **26**, 775–778.
- Lowry PP II (1998) Diversity, endemism, and extinction in the flora of New Caledonia: a review. In ‘Proceedings of the international symposium on rare, threatened, and endangered floras of Asia and the Pacific rim. Academica Sinica Monograph Series 16’. (Eds C-I Peng, PP Lowry II) pp 181–206. (Institute of Botany, Academia Sinica: Taipei, Taiwan)
- Mabberley DJ (1998) Australian Citreae with notes on other Aurantioideae (Rutaceae). *Telopea* **7**, 333–344.
- Macphail MK, Colhoun EA (1985) Late last glacial vegetation, climates and fire activity in southwest Tasmania. *Search* **16**, 43–5.
- Maddison WP (1997) Gene trees in species trees. *Systematic Biology* **46**, 523–536.
- Mallo D, Martins LDO, Posada D (2014) Unsorted homology within locus and species trees. *Systematic Biology* **63**, 988–992.
- Martin D, Rybicki E (2000) RDP: detection of recombination amongst aligned sequences. *Bioinformatics* **16**, 562–563.
- Martin DP, Posada D, Crandall KA, Williamson C (2005) A modified BOOTSCAN algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Research and Human Retroviruses* **21**, 98–102.
- Mason-Gamer RJ, Kellogg EA (1996) Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). *Systematic Biology* **45**, 524–545.
- Maynard Smith J (1992) Analyzing the mosaic structure of genes. *Journal of molecular evolution* **34**, 126–129.
- Mayol M, Rosselló JA (2001) Why nuclear ribosomal DNA spacers (ITS) tell different stories in *Quercus*. *Molecular Phylogenetics and Evolution* **19**, 167–176.

References

- McDade LA (1992) Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* **46**, 1329–1346.
- McKinnon GE, Jordan GJ, Vaillancourt RE, Steane DA, Potts BM (2004) Glacial refugia and reticulate evolution: the case of the Tasmanian eucalypts. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **359**, 275–84.
- McKinnon GE, Steane DA, Potts BM, Vaillancourt RE (1999) Incongruence between chloroplast and species phylogenies in *Eucalyptus* Subgenus *Monocalyptus* (Myrtaceae). *American Journal of Botany* **86**, 1038–1046.
- McKinnon GE, Vaillancourt RE, Jackson HD, Potts BM (2001) Chloroplast sharing in the Tasmanian eucalypts. *Evolution* **55**, 703–711.
- McLachlan JS, Clark JS, Manos PS (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology* **86**, 2088–2098.
- McLoughlin S (2001) The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany* **49**, 271–300.
- Mellick R, Lowe A, Allen C, Hill RS, Rossetto M (2012) Palaeodistribution modelling and genetic evidence highlight differential post-glacial range shifts of a rain forest conifer distributed across a latitudinal gradient. *Journal of Biogeography* **39**, 2292–2302.
- Miller JS, Kamath A, Levin RA (2009) Do multiple tortoises equal a hare? The utility of nine noncoding plastid regions for species-level phylogenetics in tribe Lycieae (Solanaceae). *Systematic Botany* **34**, 796–804.
- Miyamoto MM (1985) Consensus cladograms and general classifications. *Cladistics* **1**, 186–189.
- Mole BJ, Udovicic F, Ladiges PY, Duretto MF (2004) Molecular phylogeny of *Phebalium* (Rutaceae: Boronieae) and related genera based on the nrDNA regions ITS 1+2. *Plant Systematics and Evolution* **249**, 197–212.
- Morton CM (2015) Phylogenetic relationships of *Zieria* (Rutaceae) inferred from chloroplast, nuclear, and morphological data. *PhytoKeys* **44**, 15–38.

- Morton CM, Grant M, Blackmore S (2003) Phylogenetic relationships of the Aurantioideae inferred from chloroplast DNA sequence data. *American Journal of Botany* **90**, 1463–1469.
- Morton CM, Telmer C (2014) New Subfamily Classification for the Rutaceae. *Annals of the Missouri Botanical Garden* **99** (4), 620–641.
- Muir G, Fleming CC, Schlötterer C (2001) Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. *Molecular Biology and Evolution* **18**, 112–9.
- Muir G, Schlötterer C (2005) Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Molecular Ecology* **14**, 549–561.
- Munday BL (1968) *Zieria arborescens* (Stinkwood) intoxication in cattle. *Australian Veterinary Journal* **44**, 501–502.
- Munday BL, Cummings R, Wilson BJ (1974) Experimental *Zieria arborescens* (Stinkwood) poisoning in rabbits. *Research in veterinary science* **17**, 270–272.
- Musters W, Boon K, Van der Sande CA, Van Heerikhuizen H, Planta RJ (1990). Functional analysis of transcribed spacers of yeast ribosomal DNA. *The EMBO journal* **9**, 3989–3996.
- Neal W (2013) Phylogeography and population genetics of *Philotheca verrucosa* and *Zieria veronicea* (Rutaceae), two plant species from south-eastern Australia. Honours thesis, University of Melbourne.
- Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics* **39**, 121–152.
- Nelson G (1989) Cladistics and evolutionary models. *Cladistics* **5**, 275–289.
- Nevill PG, Bossinger G, Ades PK (2010) Phylogeography of the world's tallest angiosperm, *Eucalyptus regnans*: evidence for multiple isolated Quaternary refugia. *Journal of Biogeography* **37**, 179–192.
- Nevill PG, Despres T, Bayly MJ, Bossinger G, Ades PK (2014) Shared phylogeographic patterns and widespread chloroplast haplotype sharing in *Eucalyptus* species with different ecological tolerances. *Tree Genetics and Genomes*, **10**, 1079–1092.

References

- Newton AC, Allnutt TR, Gillies ACM, Lowe AJ, Ennos RA (1999) Molecular phylogeography, intraspecific variation and the conservation of tree species. *Trends in Ecology and Evolution*, **14**, 140–145.
- Nixon KC, Carpenter JM (1996) On simultaneous analysis. *Cladistics* **12**, 221–241.
- Nylander J (2004) ‘MrModeltest version 2.’ (Program distributed by the author. Evolutionary Biology Centre: Uppsala University, Sweden)
- Ochieng JW, Henry RJ, Baverstock PR, Steane DA, Shepherd M (2007) Nuclear ribosomal pseudogenes resolve a corroborated monophyly of the eucalypt genus *Corymbia* despite misleading hypotheses at functional ITS paralogs. *Molecular Phylogenetics and Evolution* **44**, 752–764.
- Othman RNA, Jordan GJ, Worth JRP, Steane DA, Duretto MF (2010) Phylogeny and infrageneric classification of *Correa* Andrews (Rutaceae) on the basis of nuclear and chloroplast DNA. *Plant Systematics and Evolution* **288**, 127–138.
- Padidam M, Sawyer S, Fauquet CM (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**, 218–225.
- Palmer JD (1987) Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *The American Naturalist* **130**, S6–S29.
- Palmer JD (1991) Plastid chromosomes: structure and evolution. In ‘The Molecular Biology of Plastids’ (Eds. L Bogorad and IK Vasil) pp. 5–53. (Academic Press: San Diego)
- Palmer JD, Delwiche CF (1996) Second-hand chloroplasts and the case of the disappearing nucleus. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 7432.
- Palmer JD, Delwiche CF (1998) The origin and evolution of plastids and their genomes. In ‘Molecular systematics of plants II: DNA sequencing.’ (Eds DE Soltis, PS Soltis, JJ Doyle.) pp. 375–409. (Springer Science+Business Media: New York)
- Parra-O. C, Bayly MJ, Drinnan A, Udovicic F, Ladiges PY (2009) Phylogeny, major clades and infrageneric classification of *Corymbia* (Myrtaceae), based on nuclear ribosomal DNA and morphology. *Australian Systematic Botany* **22**, 384–399.
- Pelletier B (2006) Geology of the New Caledonia region and its implications for the study of the New Caledonian biodiversity. Compendium of marine species from

- New Caledonia, Nouvelle-Calédonie. Documents Scientifiques et Techniques IRD, II 7. Institut de Recherche pour le Développement, Nouméa, New Caledonia. pp. 17–30.
- Petit RJ, Kremer A, Wagner DB (1993) Finite island model for organelle and nuclear genes in plants. *Heredity* **71**, 630–41.
- Pfeil BE, Crisp MD (2008) The age and biogeography of *Citrus* and the orange subfamily (Rutaceae: Aurantioideae) in Australasia and New Caledonia. *American Journal of Botany* **95**, 1621–1631.
- Poczai P, Hyvönen J (2010) Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. *Molecular Biology Reports* **37**, 1897–1912.
- Pollock LJ, Bayly MJ, Nevill PG, Vesk PA (2013) Chloroplast DNA diversity associated with protected slopes and valleys for hybridizing *Eucalyptus* species on isolated ranges in south-eastern Australia. *Journal of Biogeography* **40**, 155–67.
- Pollock LJ, Tingley R, Morris WK, Golding N, O'Hara RB, Parris KM, Vesk PA, McCarthy MA (2014) Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution* **5**, 397–406.
- Poon W-S, Shaw P-C, Simmons MP, But PP-H (2007) Congruence of molecular, morphological, and biochemical profiles in Rutaceae: a cladistic analysis of the subfamilies Rutoideae and Toddalioideae. *Systematic Botany* **32**, 837–846.
- Posada D (2002) Evaluation of methods for detecting recombination from DNA sequences: empirical data. *Molecular Biology and Evolution* **19**, 708–717.
- Posada D, Crandall KA (2001) Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proceedings of the National Academy of Sciences* **98**, 13757–13762.
- Potts AJ, Hedderson TA, Grimm, GW (2014) Constructing phylogenies in the presence of intra-individual site polymorphisms (2ISPs) with a focus on the nuclear ribosomal cistron. *Systematic Biology* **63**, 1–16.
- Powell JM, Armstrong JA (1980) Seed surface structure in the genus *Zieria* Sm. (Rutaceae). *Telopea* **2**, 85–112.

References

- Prakash N (1993/94) Aspects of the Reproductive Biology of *Zieria prostrata* (Rutaceae). *Research Reports* Available at http://www.aff.org.au/AFF2_Res_repts.htm
- Rambaut A (2002) 'Se-Al: Sequence Alignment Editor version 2.0a11'. Available at <http://tree.bio.ed.ac.uk/software/> [Verified December 2015].
- Rambaut A (2007) 'Tracer version 1.5'. Available at <http://tree.bio.ed.ac.uk/software/tracer/> [Verified December 2015].
- Ramírez MJ (2006) Further problems with the incongruence length difference test: "hypercongruence" effect and multiple comparisons. *Cladistics* **22**, 289–295.
- Rauscher JT, Doyle JJ, Brown AHD (2002) Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploid complex. *Molecular Ecology* **11**, 2691–2702.
- Rauscher JT, Doyle JJ, Brown AHD (2004). Multiple origins and nrDNA internal transcribed spacer homeologue evolution in the *Glycine tomentella* (Leguminosae) allopolyploid complex. *Genetics* **166**, 987–998.
- Razafimandimbison SG, Kellogg EA, Bremer B (2004) Recent origin and phylogenetic utility of divergent ITS putative pseudogenes: a case study from *Naucleaeae* (Rubiaceae). *Systematic Biology* **53**, 177–192.
- Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* **5**, 65–84.
- Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In 'Hybrid zones and the evolutionary process.' (Ed RG Harrison.) pp. 70–109 (Oxford University Press: New York)
- Ronquist F (2004) Bayesian inference of character evolution. *Trends in Ecology and Evolution* **19**, 475–81.
- Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.
- Rossetto R, Allen CB, Thurlby KAG, Weston PH, Milner ML (2012) Genetic structure and bio-climatic modeling support allopatric over parapatric speciation along a latitudinal gradient. *BMC Evolutionary Biology* **12**, 149.

- Salvo G, Bacchetta G, Ghahremaninejad F, Conti E (2008) Phylogenetic relationships of Ruteae (Rutaceae): new evidence from the chloroplast genome and comparisons with non-molecular data. *Molecular Phylogenetics and Evolution* **49**, 736–748.
- Salvo G, Manafzadeh S, Ghahremaninejad F, Tojibaev K, Zeltner L, Conti E (2011) Phylogeny, morphology, and biogeography of *Haplophyllum* (Rutaceae), a species-rich genus of the Irano-Turanian floristic region. *Taxon* **60**, 513–527.
- Samarakoon T, Wang SY, Alford MH (2013) Enhancing PCR amplification of DNA from recalcitrant plant specimens using a trehalose-based additive. *Applications in Plant Sciences* **1**, 1200236.
- Samuel R, Bachmair A, Jobst J, Ehrendorfer F (1998) ITS sequences from nuclear rDNA suggest unexpected phylogenetic relationships between Euro-Mediterranean, East Asiatic and North American taxa of *Quercus* (Fagaceae). *Plant Systematics and Evolution* **211**, 129–139.
- Sanderson MJ, Doyle JJ (1992) Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. *Systematic Biology* **41**, 4–17.
- Sang T, Crawford D, Stuessy T (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84**, 1120–1136.
- Sang T (2002) Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecular Biology* **37**, 121–147.
- Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology* **7**, 465–474.
- Scott KD, McIntyre CL, Playford J (2000) Molecular analyses suggest a need for a significant rearrangement of Rutaceae subfamilies and a minor reassessment of species relationships within *Flindersia*. *Plant Systematics and Evolution* **223**, 15–27.
- Sharma IK (2001) Understanding clonal diversity patterns through allozyme polymorphism in an endangered and geographically restricted Australian shrub,

References

- Zieria baeuerlenii* and its implications for conservation. *Biochemical Systematics and Ecology* **29** 681–695.
- Sharma IK, Young AG (2001) Genetic diversity analysis among fragmented populations of *Zieria granulata* within the Tabbagong Forest and RSA land. Report prepared for the New South Wales Roads and Traffic Authority Southern Region, Wollongong, New South Wales.
- Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **92**, 142–166.
- Shaw J, Lickey E, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* **94**, 275–288.
- Shiao YH, Lupascu ST, Gu YD, Kasprzak W, Hwang CJ, Fields JR, Leighty RM, Quiñones O, Shapiro BA, Alvord WG, Anderson LM (2009) An intergenic non-coding rRNA correlated with expression of the rRNA and frequency of an rRNA single nucleotide polymorphism in lung cancer cells. *PLoS One* **4**, e7505.
- Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49**, 369–381.
- Small RL, Ryburn JA, Cronn RC, Seelanan T, Wendel JF (1998) The tortoise and the hare: choosing between noncoding plastome and nuclear adh sequences for phylogeny reconstruction in a recently diverged plant group. *American Journal of Botany* **85**, 1301–1315.
- Smith JE (1798) The characters of twenty new genera of plants. *Transactions of the Linnean Society of London* **4**, 213–223.
- Smith JE (1818) *Zieria*. In ‘The cyclopedia: or universal dictionary of arts, sciences and literature’. Vol. 39. (Ed A Rees) (Longman et al.: London)
- Smith-White S (1954) Chromosome numbers in the Boronieae (Rutaceae) and their bearing on the evolutionary development of the tribe in the Australian flora. *Australian Journal of Botany* **2**, 287–303.

- Soltis D, Soltis P (1998) Choosing an approach and an appropriate gene for phylogenetic analysis. In 'Molecular systematics of plants II.' (Eds D Soltis, P Soltis, J Doyle.) pp. 1–42. (Springer Science+Business Media: New York)
- Soltis DE, Johnson LA, Looney C (1996) Discordance between ITS and chloroplast topologies in the *Boykinia* group (Saxifragaceae). *Systematic Botany* 169–185.
- Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. *Annual Review of Plant Biology* 60, 561–588.
- Song J, Shi L, Li D, Sun Y, Niu Y, Chen Z, Hongmei L, Pang X, Sun Z, Liu C, Lv A, Deng Y, Larson-Rabin Z, Wilkinson M, Chen S (2012) Extensive pyrosequencing reveals frequent intra-genomic variations of internal transcribed spacer regions of nuclear ribosomal DNA. *PLoS ONE* 7, e43971.
- Southwell IA, Armstrong JA (1987) Chemical variation within the genus *Zieria*. *Phytochemistry* 26, 1687–1692.
- Stace HM, Armstrong JA (1992) New chromosome numbers for Rutaceae. *Australian Systematic Botany* 5, 501–505.
- Steane DA (2005) Complete nucleotide sequence of the chloroplast genome from the Tasmanian blue gum, *Eucalyptus globulus* (Myrtaceae). *DNA Research*. 12, 215–20.
- Steane DA, Byrne M, Vaillancourt RE, Potts BM (1998) Chloroplast DNA polymorphism signals complex interspecific interactions in *Eucalyptus* (Myrtaceae). *Australian Systematic Botany* 11, 25–40.
- Steel M, Penny D (2000) Parsimony, likelihood, and the role of models in molecular phylogenetics. *Molecular Biology and Evolution* 17, 839–850.
- Stevens PF (2012) Angiosperm Phylogeny Website. Version 12, July 2012 (and more or less continuously updated since).
- Stewart JR and Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology and Evolution*, 16, 608–613.
- Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn RC and Liston A (2012) Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *American Journal of Botany* 99:349–364.

References

- Suárez-Santiago VN, Salinas MJ, Garcia-Jacas N, Soltis PS, Soltis DE, Blanca G (2007) Reticulate evolution in the *Acrolophus* subgroup (*Centaurea* L., Compositae) from the western Mediterranean: origin and diversification of section *Willkommia* Blanca. *Molecular Phylogenetics and Evolution* **43**, 156–172.
- Sutherland R, Collot J, Lafoy Y, Logan GA, Hackney R, Stagpoole V, Uruski C, Hashimoto T, Higgins K, Herzer RH, Wood R, Mortimer N, Rollet N (2010) Lithosphere delamination with foundering of lower crust and mantle caused permanent subsidence of New Caledonia Trough and transient uplift of Lord Howe Rise during Eocene and Oligocene initiation of Tonga–Kermadec subduction, western Pacific. *Tectonics* **29**, TC2004.
- Suzuki Y, Glazko GV, Nei M (2002) Over credibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 16138–16143.
- Swenson U, Nylinder S, Munzinger J (2014) Sapotaceae biogeography supports New Caledonia being an old Darwinian island. *Journal of Biogeography* **41**, 797–809.
- Swofford D (2001) ‘PAUP*. Phylogenetic analysis using parsimony, version 4.0b10’. (Illinois Natural History Survey, Smithsonian Institution: Champaign)
- Swofford DL, GJ Olsen PJ Waddell, DM Hillis (1996) Phylogenetic inference. In ‘Molecular systematics. 2nd edition’ (Eds DM Hillis, C Moritz, and BK. Marble) pp. 407–514 (Sinauer: 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* **17**, 1105–1109.
- Thorne RF and Reveal JL (2007) An updated classification of the class Magnoliopsida (“Angiospermae”). *The Botanical Review* **73**, 67–181.
- Timme RE, Kuehl JV, Boore JL, Jansen RK (2007) A comparative analysis of the *Lactuca* and *Helianthus* (Asteraceae) plastid genomes: identification of divergent regions and categorization of shared repeats. *American Journal of Botany* **94**, 302–312.

- van der Sande CAFM, Kwa M, van Nues RW, van Heerikhuizen H, Raué HA, Planta RJ (1992) Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisiae* ribosomal DNA. *Journal of Molecular Biology* **223**, 899-910.
- van Ham RC, Hart HT, Mes TH, Sandbrink JM (1994) Molecular evolution of noncoding regions of the chloroplast genome in the Crassulaceae and related species. *Current Genetics* **25**, 558-66.
- Viot R, Guillaumin A (1953) 'Contributions a la flore de la Nouvelle Calédonie. C11—Plantes recoltees par M. R. Viot. R. Viot and A. Guillaumin.' *Memoires du Museum National D'Histoire Naturalie, Serie B., Botanique* **4**, 9–10.
- Wagner WH (1983) Reticulistics: the recognition of hybrids and their role in cladistics and classification. *Advances in cladistics* **2**, 63-79.
- Wagner A, Blackstone N, Cartwright P, Dick M, Misof B, Snow P, Wagner GP, Bartels J, Murtha M, Pendleton J (1994) Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. *Systematic Biology* **43**, 250-61.
- Weitemier K, Straub SCK, Cronn RC, Fishbein M, Schmickl R, McDonnell A, Liston A (2014) Hyb-Seq: combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences* **2**, 1400042.
- Wendel JF, Doyle JJ (1998) Phylogenetic incongruence: window into genome history and molecular evolution. In 'Molecular systematics of plants II: DNA sequencing.' (Eds DE Soltis, PS Soltis, JJ Doyle.) pp 265–296 (Springer Science+Business Media: New York)
- Wenpan D, Jing L, Jing Y, Ling W, Shiliang Z (2012) Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS ONE* **7**, 1–9.
- White CT (1942) Contributions to the Queensland flora, number 7. *Proceedings of the Royal Society of Queensland* **53**, 201–228.
- White TJ, Bruns T, Lee S, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In 'PCR protocols: a guide to methods and applications.' (Ed MA Innis, DH Gelfand, JJ Sninsky, and TJ White.) pp. 315–322. (Academic Press, Inc.: New York)

References

- Wiens JJ (1998) Does adding characters with missing data increase or decrease phylogenetic accuracy? *Systematic Biology* **47**, 625–640.
- Worth J, Jordan G, McKinnon G, Vaillancourt R (2009) The major Australian cool temperate rainforest tree *Nothofagus cunninghamii* withstood Pleistocene glacial aridity within multiple regions: evidence from the chloroplast. *New Phytologist* **182**, 519–532.
- Worth JRP, Jordan GJ, Marthick JR, McKinnon GE, Vaillancourt RE (2010) Chloroplast evidence for geographic stasis of the Australian bird-dispersed shrub *Tasmannia lanceolata* (Winteraceae). *Molecular Ecology* **19**, 2949–2963.
- Worth JRP, Marthick JR, Jordan GJ, Vaillancourt RE (2011) Low but structured chloroplast diversity in *Atherosperma moschatum* (Atherospermataceae) suggests bottlenecks in response to the Pleistocene glacials. *Annals of Botany* **108**, 1247–1256.
- Wright SD, Yong CG, Wichman SR, Dawson JW, Gardner RC (2001) Stepping stones to Hawaii: a trans-equatorial dispersal pathway for *Metrosideros* (Myrtaceae) inferred from nrDNA (ITS+ETS). *Journal of Biogeography* **28**, 769–774.
- Xu B, Wu N, Gao X-F, Zhang L-B (2012) Analysis of DNA sequences of six chloroplast and nuclear genes suggests incongruence, introgression, and incomplete lineage sorting in the evolution of *Lespedeza* (Fabaceae). *Molecular Phylogenetics and Evolution* **62**, 346–358.
- Yoder AD, Irwin JA, Payseur BA (2001) Failure of the ILD to determine data combinability for Slow Loris phylogeny. *Systematic Biology* **50**, 408–424.
- Yokota Y, Kawata T, Iida Y, Kato A, Tanifuji S (1989) Nucleotide sequences of the 5.8S rRNA gene and internal transcribed spacer regions in carrot and broad bean ribosomal DNA. *Journal of Molecular Evolution* **29**, 294–301.
- Yu W-B, Huang P-H, Li D-Z, Wang H (2013) Incongruence between nuclear and chloroplast DNA phylogenies in *Pedicularis* section *Cyathophora* (Orobanchaceae). *PLoS ONE* **8**, e74828.
- Zimmer EA, Martin SL, Beverley SM, Kan YW, Wilson AC (1980) Rapid duplication and loss of genes coding for the alpha chains of hemoglobin. *Proceedings of the National Academy of Sciences of the United States of America* **77**, 2158–2162.

Appendix A. Accession details.

This table provides details of the samples that were used in the final analyses of this study. Samples that were sequenced for this study but subsequently excluded (due to poor quality sequences, high numbers of polymorphisms, or missing data) are not documented here. Herbarium abbreviations follow Index Herbariorum. Abbreviations of collectors' names are as follows : ABP, A.B. Pollock; AF, A. Ford; Bean, A.R. Bean; CMH, C.M. Haskard; DAM, D.A. Meagher; DEA, D.E. Albrecht; DJO, D.J. Ohlsen; ELB, E.L. Batty; IRT, I.R. Telford; JA, J. Armstrong; JRC, J.R. Clarkson; KM, K. Muscat; KMS, K.M. Sparshott; KRM, K.R. McDonald; LMD, L. McDougal; MBT, M.B. Thomas; MD, M. Duretto; MFD, M.F. Duretto; MJB, M.J. Bayly; MTM, M.T. Mathieson; NGW, N.G. Walsh; PG, P. Gilmour; PIF, P.I. Forster; RAB, R.A. Barrett; RGC, R.G. Coveny; RJ, R. Johnstone; RM, R. Makinson; SPP, S.P. Phillips; TM, T. McLay; WN, W. Neal. All other abbreviations are listed at the front of the thesis, page xx. Collections of *Zieria arborescens* were not identified to subspecies, but those from near the type locations of subsp. *decurrens* and subsp. *glabrifolia* are indicated (see Chapter 2, General Methods, for details). GenBank numbers for published cpDNA sequences are provided in the appendix of Chapter 3.

* In most cases the identification (ID) number used in this study was a combination of the collector's initials and their collection number, except for some herbarium specimens where just the numerical part of the accession number was used in order to save space in figures and tables.

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Neobyrsesia suberosa</i> J.A.Armstr.	MJB1904	NT, Kakadu NP, East Alligator area, Bardedjilidji Sandstone Walk, ~300 m SW of carpark, near base of sandstone cliff	MEL 2383567	133.02	-12.55	✓	✓	
<i>Z. actites</i> Duretto & P.I.Forst.	PIF38224	Qld, Mount Larcom summit area, 2.5 km WSW of Targinie	BRI AQ818047	151.09	-23.81	✓		

Appendix A

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. adenodonta</i> (F.Muell.) J.A.Armstr.	MTM760	Qld, Lower Bellbird Circuit, Binna Burra, Lamington NP	BRI AQ756722	153.20	-28.21	✓	✓	
<i>Z. adenophora</i> (F.Muell.) J.A.Armstr.	MJB2010	NSW, cult. ANBG. Wild source: Araluen ~500 m ESE of Bells Creek Falls on rocky escarpment (voucher for stock is CBG 8805884)	MELU D 112161a	149.79	-35.60	✓	✓	✓
<i>Z. alata</i>	KM639	Qld, Mount Lewis NP	MELU D 112170a	145.27	-16.60	✓	✓	
<i>Z. alata</i> Duretto & P.I.Forst.	AF5178	Qld, Daintree NP, 'landing rock' headwaters of Roots Creek, SE of topo '1296', site 171	BRI AQ739620	145.27	-16.42	✓		
<i>Z. arborescens</i> Sims	196702	Vic, Wilsons Promontory NP, 12.8 km from carpark along maintenance road to lighthouse, SE Pt	NSW 196702	146.38	-39.12			✓
<i>Z. arborescens</i>	280353	NSW, Approximately 500 m S of the end of Box Range Rd, 11 km E of Wyndham	NSW 280353	149.77	-36.94			✓
<i>Z. arborescens</i>	414347	NSW, Bulli Creek Rd off Comboyne Rd, 1 km S of town, Byabarra	NSW 414347	152.58	-31.50			✓
<i>Z. arborescens</i>	433661	NSW, NE Werrikimbe Trig	NSW 433661	152.22	-31.25			✓
<i>Z. arborescens</i>	460582	NSW, Illawarra district, between Knights Hill and Minnamurra Falls, 'Brushstop'	NSW 460582	150.71	-34.63			✓
<i>Z. arborescens</i>	478968	NSW, Butterleaf SF, 2 km E of Mt Scott	NSW 478968	152.12	-29.52			✓
<i>Z. arborescens</i>	622747	NSW, Moogem SF, ~60 km E of Glen Innes along Gwydir Highway	NSW 622747	152.24	-29.54			cpDNA only
<i>Z. arborescens</i>	651414	NSW, Bismuth Falls, WSW of Torrington	NSW 651414	151.75	-29.28			✓
<i>Z. arborescens</i>	651415	NSW, Werrikimbe NP, 'The Racecourse', near Mt Boss SF, New England	NSW 651415	152.30	-31.12			✓
<i>Z. arborescens</i>	651416	NSW, Werrikimbe NP, along Mesa Trail via Carabeen Walking Trail	NSW 651416	152.17	-31.23			✓

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. arborescens</i>	651422	NSW, Tantawangalo SF, Tantawangalo Mountain, 14 km direct SW of Candelo	NSW 651422	149.55	-36.82			✓
<i>Z. arborescens</i>	722482	Qld, Springbrook, SW of the Gold Coast	NSW 722482	153.28	-28.23			✓
<i>Z. arborescens</i>	812039	Tas, 2 km NE of Veridian Point	NSW 812039	145.50	-42.92			✓
<i>Z. arborescens</i>	DAMWP	Vic, Wilsons Promontory. Sealers cove track ~500 m E of Femy Glade	MELU D 112160a	146.40	-39.02			✓
<i>Z. arborescens</i>	DJO110	Tas, Saint Patricks River, on the S side of the Mount Barrow road, 10 km from its junction with the Tasman Highway	MELU D 105582	147.43	-41.37	✓	✓	✓
<i>Z. arborescens</i>	DJO179	Tas, Next to the Russell Falls walking track on the E bank of Russell Falls Ck, 200 from the information centre	MELU D 112134a	146.72	-42.68			✓
<i>Z. arborescens</i>	DJO182	Tas, Snug Falls, edge of the first rainforest creek that is encountered when walking from the car park along the Snug Falls walking track	MELU D 112135a	147.21	-43.08			✓
<i>Z. arborescens</i>	DJO186	Tas, Tasman Peninsula. E side of the Fortescue Bay road 2 km from the Fortescue Camp Ground	MELU D 112136a	147.94	-43.14			✓
<i>Z. arborescens</i>	DJO195	Tas, Meetus Falls, at the walking track junction of the Meetus Falls lookout track and the Cygnet River walking track	MELU D 112137a	147.88	-41.95			✓
<i>Z. arborescens</i>	DJO207	Vic, Tarra Bulga NP, Bulga Section, 400 m from the suspension bridge next to the walking track from the main car park	MELU D 112138a	146.57	-38.43			✓
<i>Z. arborescens</i>	DJO210	Vic, Wilsons Promontory, next to Sealers Cove track, 20 m W of Sealers Ck	MELU D 112139a	146.42	-39.02			✓

Appendix A

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. arborescens</i>	ELB125	Vic, Near Fish Ck, towards Wilsons Promontory, on roadside on Meeniyan-Promontory Rd	MELU D 112132a	146.12	-38.73			✓
<i>Z. arborescens</i>	MD3005	Tas, Grunter Rd, Seventeen Mile Plain	MEL 2355532	144.90	-40.86			✓
<i>Z. arborescens</i>	MFD2035	NSW, cult. ABC, Mt Annan. Wild source: Blue Mountains NP, Walking track to Wollongambie River, near Mt Wilson	NSW (AE892158)	150.35	-33.50	✓	✓	✓
<i>Z. arborescens</i>	MJB1868	Vic, Noojee SF, Starling Gap picnic area	MELU D 120871	145.80	-37.81	✓	✓	✓
<i>Z. arborescens</i>	MJB1994b	Vic, Bunyip SP, on Link Ck, near crossing on Link Rd	MELU D 112147a	145.62	-37.96			✓
<i>Z. arborescens</i>	MJB2024a	Vic, Dona Buang Rd, 5.5 km toward Healesville from intersection with road to Donna Buang summit	MELU D 112129a	145.69	-37.73			✓
<i>Z. arborescens</i>	MJB2250a	Tas, Bruny Island, Coolangatta Rd, between intersection with Resolution Rd and Adventure Bay Rd	MELU D 112131a	147.26	-43.39			✓
<i>Z. arborescens</i>	MJB2251a	Tas, Strahan, Green St, ~40 m from the intersection of Green St and the Esplanade	MELU D 112130a	145.37	-42.13	✓	✓	✓
<i>Z. arborescens</i>	MTM270	Qld, Mount Norman, S side, Girraween NP	BRI AQ745993	151.96	-28.86	✓	✓	✓
<i>Z. arborescens</i>	MTM459	Qld, Gully between Wallangarra Ridge and Shark Fin Ridge, Girraween NP	BRI 790810	151.94	-28.88			✓
<i>Z. arborescens</i>	RAB504	NSW, Barren Grounds Nature Reserve, next to the entrance gate	MELU D 112152a	150.71	-34.67	✓	✓	✓
<i>Z. arborescens</i>	RAB512	Vic, Otway Ranges, Mount Sabine Benwerrin Road	MELU D 112099a	143.81	-38.59	✓	✓	✓
<i>Z. arborescens</i>	RAB513	Vic, Otway Ranges, Mt Sabin Benwerrin Rd	MELU D 112100a	143.80	-38.59			✓
<i>Z. arborescens</i>	RAB514	Vic, Otway Ranges, Erskine Falls carpark, near sign at steps to view falls	MELU D 112101a	143.91	-38.51			✓

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. arborescens</i>	RAB518	Vic, Otway Ranges, Erskine Falls Rd, ~1 km past Blanket Leaf Picnic Area, heading towards Eskine Falls turnoff	MELU D 112102a	143.92	-38.52			✓
<i>Z. arborescens</i>	RAB520	Vic, Phillip Island. Cowes Rhyll Rd, before tip heading E, 50 m along walking track from parking area on S side of road	MELU D 112103a	145.27	-38.47			✓
<i>Z. arborescens</i>	RAB522	Vic, Phillip Island. Cowes Rhyll Rd, just past tip heading E, S side of road where boardwalk meets road	MELU D 112104a	145.28	-38.46			✓
<i>Z. arborescens</i>	RAB525	Vic, Dandenongs, Mechanics reserve, ~20 m W of Farndons Rd on walking track	MELU D 112105a	145.36	-37.83			✓
<i>Z. arborescens</i>	RAB527	Vic, Dandenongs, The Crescent (Sassafras Creek Rd) 0.7 km before Pernins Creek Rd heading E	MELU D 112106a	145.37	-37.88			✓
<i>Z. arborescens</i>	RAB530	Vic, Nowa Nowa, 50 m downstream from bridge (Boggy Ck) on Princes Hwy, W side of bank, close to creek edge	MELU D 112108a	148.09	-37.73			✓
<i>Z. arborescens</i>	RAB532a	Vic, Errinundra NP, (via Club Terrace), W of the Ocean View Lookout on BA track (Ellergy Ck track)	MELU D 112109a	148.80	-37.37			✓
<i>Z. arborescens</i>	RAB533a	Vic, Bendoc SF, Gap Rd (east), ~2.3 km from Bonang Hwy	MELU D 112110a	148.78	-37.25			✓
<i>Z. arborescens</i>	RAB536a	Vic, Mt Jersey Rd, 6.6 km from Bonang Hwy (turnoff is slightly N of Goongerah)	MELU D 112111a	148.66	-37.33			✓
<i>Z. arborescens</i>	RAB540a	Vic, Mt Elizabeth summit	MELU D 112112a	147.93	-37.49			✓
<i>Z. arborescens</i>	RAB543a	Vic, Great Alpine Rd, ~7 km S of Tambo Crossing, beside creek bed, below where road crosses St Patricks Ck	MELU D 112113a	147.86	-37.54			✓

Appendix A

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. arborescens</i>	RAB549a	Vic, Western Otways, Devondale Rd, near Chapple Vale, ~300 m from turn-off from Cobden/Lavers Hill Rd	MELU D 112116a	143.31	-38.64			✓
<i>Z. arborescens</i>	RAB550a	Vic, Otways, Wait-a-while track, 8.2 km from turn-off from Lavers Hill/Colac Rd, near Holy Water track	MELU D 112117a	143.46	-38.72			✓
<i>Z. arborescens</i>	RAB554a	Vic, Cumberland Falls walk, east of Marysville	MELU D 112118a	145.88	-37.56			✓
<i>Z. arborescens</i>	RAB556a	Vic, Black Spur, Maroondah Hwy, ~2.5 km S of Narbethong	MELU D 112119a	145.64	-37.58			✓
<i>Z. arborescens</i>	RAB557a	Vic, Myers Creek Rd, near Monda Rd and Silvia Creek Rd	MELU D 112120a	145.51	-37.57			✓
<i>Z. arborescens</i>	RAB561a	Vic, S side of South Cascade Ck on Thomson Valley Rd, Baw Baw NP	MELU D 112121a	146.35	-37.84			✓
<i>Z. arborescens</i>	RAB563a	Vic, Boola Boola SF, W2 Track	MELU D 112122a	146.45	-38.12			✓
<i>Z. arborescens</i>	RAB568a	Vic, Tarra Valley Rd, ~26 km from Yarram by road, ~1 km before intersection with Grand Ridge Rd	MELU D 112123a	146.54	-38.44			✓
<i>Z. arborescens</i>	RAB569a	Vic, Grand Ridge Rd between Linklatters Track and Yarram Rd, S side of road (i.e. upslope)	MELU D 112124a	146.45	-38.46			✓
<i>Z. arborescens</i>	RAB571a	Vic, Yarram Morwell Rd, ~300 m from Summerfield Rd, next to logging coupe	MELU D 112125a	146.48	-38.49			✓
<i>Z. arborescens</i>	TM170a	NSW, Kattang reserve	MELU D 112128a	152.84	-31.64			✓
<i>Z. arborescens</i> (subsp. <i>decurrens</i> J.A.Armstr.)	BRI ISO	NSW, 1.5 miles [2.4 km] along the Caves Beach Road, near Jervis Bay	BRI AQ695444, NSW 458804	150.68	-35.15	✓		✓
<i>Z. arborescens</i> (subsp. <i>decurrens</i>)	MFD2025	NSW, cult. ABG, Mt Annan. Wild source: Jervis Bay NP, 200 m along Steamers Beach access track from Steamers Beach carpark	AE952505	150.72	-35.16			✓

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. arborescens</i> (subsp. <i>glabrifolia</i> J.A.Armstr.)	MTM213	Qld, Donnellys Castle, ~6.5 km W of Thulimbah off Donnellys Castle Road	BRI AQ745477	151.88	-28.56	✓	✓	✓
<i>Z. aspalathoides</i> A.Cunn. ex Benth. subsp. <i>aspalathoides</i>	Bean27984	Qld, near junction of Delubra and Cadarga creeks, 35 km SW of Mundubbera	BRI AQ820153	151.03	-25.80	✓	✓	
<i>Z. baeuerlenii</i> J.A.Armstr.	275527	NSW, Bomaderry Creek Walking Track, cliff top N of number '4'	NSW 275527	150.62	-34.85	✓	✓	
<i>Z. baeuerlenii</i>	MJB2011	NSW, cult. ANBG. Wild source: Bomaderry Creek Reserve, 300m S of Weir on W side (voucher for stock is CBG 9104885)	MELU D 112162a	150.59	-34.85	✓	✓	
<i>Z. bifida</i> Duretto & P.I.Forst.	MTM537	Qld, Triunia NP, adjacent to Dulong Road W of Woombye	BRI AQ792099	152.90	-26.65	✓	✓	✓
<i>Z. boolbunda</i> Duretto & P.I.Forst.	KMS616	Qld, Boolbunda Rock, ~9 km NE of Mount Perry township	BRI AQ640789	151.70	-25.12	✓	✓	
<i>Z. buxijugum</i> J.D.Briggs & J.A.Armstr.	RJ1897	NSW, Box Range Farm ~10 km W of Pambula	MEL 2312322	149.76	-36.93	✓	✓	
<i>Z. caducibracteata</i> J.A.Armstr.	745482	NSW, Budawang NP. Track to Curong [Corang] River from Wog Wog camping ground	NSW 745482	150.20	-35.25	✓	✓	✓
<i>Z. caducibracteata</i>	IRT9585	NSW, northern Budawang Range, Quiltys Mountain, 1.5 km SE of Endrick trig	MEL 654142	150.20	-35.22	✓	✓	✓
<i>Z. cephalophila</i> Duretto & P.I.Forst.	TH7339	Qld, SE edge of Sydney Heads, Britton Range, 6.3 km NE–ENE of Homevale Homestead	BRI AQ579129	148.58	-21.41	✓	✓	
<i>Z. chevalieri</i> Viot	MJB2119	New Caledonia, Province Nord (N Province), Mt Kaala, near saddle N of summit on S end of range	MEL 2383640	164.37	-20.61	✓	✓	

Appendix A

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. citriodora</i> J.A.Armstr.	NGW2038	Vic, cult. RBG Melbourne, grown from NGW 2038, on S bank of Limestone Creek, 1.5 km direct NE from ford on Limestone Creek track	MEL 222166	148.08	-36.80	✓	✓	
<i>Z. collina</i> C.T.White	MJB2012	Qld, cult. ANBG. Wild source: Cedar Creek NP, main circuit walking track N side (voucher for stock is CBG 8604923)	MELU D 112163a	153.16	-27.90	✓	✓	
<i>Z. collina</i>	MTM507	Qld, Tamborine NP, Panorama Point section, Wongawallan Road, Mount Tamborine	BRI AQ871795	153.22	-27.91	✓	✓	
<i>Z. compacta</i> C.T.White	PIF34479	Qld, Roberts Range, ascent to Devil's Elbow, Wyberba	BRI AQ745295	151.88	-28.88	✓	✓	
<i>Z. covenyi</i> J.A.Armstr.	MJB2013	NSW, cult. ANBG. Wild source: Narrow Neck Plateau ~5.5 km from Cliff Drive along road to Bushwalkers Hill (voucher for stock is CBG 8411672)	MELU D 112164a	150.26	-33.75	✓	✓	✓
<i>Z. cytisoides</i> Sm.	AF5451	Qld, Great Dividing Range, NW of Mount Emerald near topo '1089', W of Tolga	BRI AQ849111	145.40	-17.20	✓	✓	
<i>Z. cytisoides</i>	DJO320	Qld, Carnarvon Gorge NP, 50 m upstream of the first creek encountered on the Moss Gardens walking track	MELU D 112141a	148.22	-25.05	✓		
<i>Z. cytisoides</i>	DJO416	NSW, Weddin Mountains NP, Bertha's Gully, on the N side of the creek next to the walking track, ~500 m E of the junction with the track	MELU D 112143a	147.96	-33.92	✓	✓	
<i>Z. distans</i> Duretto & P.I.Forst.	CMH157	Qld, Coalston Lakes, Walla Range	BRI AQ756300	151.86	-25.64	✓	✓	
<i>Z. distans</i>	MTM287	Qld, Kroombit Tops NP, Razorback break, ~10 km SW of barracks	BRI AQ813602	150.89	-24.42	✓	✓	✓

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. eungellaensis</i> Duretto & P.I.Forst.	Bean4435	Qld, WSW of Mount David, Eungella NP	MEL 2324537, BRI AQ544202	148.63	-21.04	✓	✓	
<i>Z. exsul</i> Duretto & P.I.Forst.	MTM535	Qld, narrow strip of vegetation adjacent to Sunshine Motorway, Buderim, adjacent to Buderim Waste Landfill Centre	BRI AQ792100	153.10	-26.68	✓	✓	
<i>Z. formosa</i> J.D.Briggs & J.A.Armstr.	MJB2014	NSW, cult. ANBG. Wild source: Lochiel, 250 m from junction of Pambula–Wyndham road with Back Creek Road, towards Pambula, just S of road (voucher for stock is CBG 8604998)	MELU D 112165a	149.81	-36.95	✓	✓	✓
<i>Z. fraseri</i> Hook. subsp. <i>fraseri</i>	SPP1995	Qld, near the summit of the south peak of Mount Maroon, W of Rathdowney	BRI AQ748157	152.73	-28.21	✓	✓	
<i>Z. fraseri</i> subsp. <i>robusta</i> (C.T.White) Duretto & P.I.Forst.	MBT3739	Qld, Salvator Rosa NP, 160 km by road SW of Springsure, at the base of The Sentinel, ~1.5 km W of camping ground	BRI AQ830008	147.19	-24.80	✓	✓	
<i>Z. furfuracea</i> subsp. <i>euthadenia</i> J.A.Armstr.	MTM280	Qld, Wide Bay, corner of Cooloola Way and Harry's Hut Road, N of Kin Kin	BRI AQ745998	152.96	-26.19	✓	✓	
<i>Z. furfuracea</i> R.Br. ex Benth. subsp. <i>furfuracea</i>	466549	NSW, ~5 km direct WNW of Yarras, ~700 m E along the Oxley Highway from the Mount Seaview Resort turnoff, towards Wauchope	NSW 466549	152.26	-31.37	✓	✓	
<i>Z. furfuracea</i> subsp. <i>gymnocarpa</i> J.A.Armstr.	PIF34232	Qld, Belmont Gateway Motorway, thin strip of land between Motorway and Mount Petrie Road, 1.5 km S of Old Cleveland Road	BRI AQ744205, MEL 2324803	153.13	-27.52	✓	✓	
<i>Z. graniticola</i> J.A.Armstr. ex Duretto & P.I.Forst.	MTM479	Qld, Passchendaele SF, via Carniel Lane Bapaume	BRI AQ792107	151.83	-28.56	✓	✓	

Appendix A

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. granulata</i> (F.Muell.) C.Moore ex Benth.	MJB2016	NSW, cult. ANBG. Wild source: Long Brush Road, Kiama (voucher for stock is CBG 8501509)	MELU D 112166a	150.83	-34.68	✓	✓	
<i>Z. hindii</i> J.A.Armstr.	413395	NSW, Mount Jersusalum NP, Blackbutt Plateau area	NSW 413359	153.38	-28.53	✓	✓	
<i>Z. hydroscopica</i> Duretto & P.I.Forst.	PIF38247	Qld, Rocky Gorge Coomingleh SF, Rocky Gorge Waterhole	BRI AQ818058	150.93	-24.88	✓	✓	
<i>Z. inexpectata</i> Duretto & P.I.Forst.	CMH201	Qld, Burnett District, Pryde's Road, on the southern road reserve, ~700 m from the intersection with the Burnett Hwy	BRI	152.02	-26.54	✓	✓	
<i>Z. ingramii</i> J.A.Armstr.	759202	NSW, Goonoo State Conservation Area, Freeman's Road, ~400 m E of Mount Carl Road	NSW 759202	148.98	-32.07	✓	✓	
<i>Z. ingramii</i>	IRT613	NSW, Goonoo SF, ~53 km NE of Dubbo, ~600 m N of Denmire Road from a point 1.75 km E of Mount Carl Road	MEL 2218233	149.03	-31.93	✓	✓	
<i>Z. insularis</i> Duretto & P.I.Forst.	PIF34666	Qld, Mount Windsor NP, Spencer Creek downstream of forestry barracks	BRI AQ745463	145.04	-16.27	✓	✓	
<i>Z. involucrata</i> R.Br. ex Benth.	506074	NSW, Maroota, 3.2 km along Laughtondale Gully Road from Old Northern Road	NSW 506074	151.01	-33.41	✓	✓	
<i>Z. involucrata</i>	LMD143	NSW, Laughtondale Gully Road, ~3 km from Old Northern Road	MEL 2040124	151.01	-33.42	✓	✓	✓
<i>Z. laevigata</i> Bonpl.	MTM258	Qld, Passchendaele SF, Forestry Lane	BRI AQ813610	151.79	-28.61	✓	✓	✓
<i>Z. laevigata</i>	PIF34475	Qld, Girraween NP, junction of Ramsey and Bald Rock creeks	BRI AQ813611	151.92	-28.84	✓	✓	
<i>Z. lasiocaulis</i> J.A.Armstr.	492944	NSW, N side of Banda Banda Flora Reserve, 3.4 km by road, N of intersection of Banda Banda Road and Hastings Forest Way	NSW 492944	152.41	-31.16	✓	✓	✓

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. lasiocaulis</i>	PG7908	NSW, ~1 km W of Mount Banda Banda at the headwaters of the Wilson River, Banda Banda Flora Reserve	MEL 2199590	152.43	-31.16	✓	✓	✓
<i>Z. laxiflora</i> (Benth.) Domin	MTM385	Qld, Great Sandy NP, adjacent to Cooloola Coast cemetery, Rainbow Beach Road	BRI AQ755776	152.97	-25.98	✓	✓	
<i>Z. littoralis</i> J.A.Armstr.	MJB2017	Cult. ANBG. Wild source unknown (voucher for stock is CANB 643996)	MELU D 112167a	Unknown	Unknown	✓	✓	
<i>Z. littoralis</i>	RAB508	NSW, Tathra Head, 50 m from the car park on the walking track leading to the lookout over the ocean	MELU D 112155a	149.99	-36.73	✓	✓	
<i>Z. madida</i> Duretto & P.I.Forst.	AF5320	Qld, Daintree NP, Hilda Creek area near camping area, ~800 m WNW of Thornton Peak Site 190	BRI AQ863277	145.37	-16.16	✓	✓	✓
<i>Z. minutiflora</i> Domin subsp. <i>minutiflora</i>	ABP2356	Qld, ~2 km due NE of Landsborough, on margin of private property and Ewen Maddock Council Reserve	BRI AQ815342	152.98	-26.79	✓	✓	
<i>Z. minutiflora</i> subsp. <i>minutiflora</i>	PIF34313	Qld, Glass House Mountains NP, Mount Coonowrin (Crookneck)	BRI AQ862319	152.91	-26.90	✓	✓	
<i>Z. minutiflora</i> subsp. <i>trichocarpa</i> J.A.Armstr.	KRM8283	Qld, Lot 5, Webster Road, near Wondeccla	BRI AQ747078	145.46	-17.44	✓	✓	
<i>Z. montana</i> J.A.Armstr.	PIF36219	Qld, Mount Barney NP, southern slopes (Peasant's Ridge) of Mount Barney, steep S-facing ridge	BRI AQ813884	152.69	-28.29	✓	✓	✓
<i>Z. murphyi</i> Blakely	JA1356	NSW, Mt Tomah, on ledge between Gaping Gill and Camels Hump	MEL 2102796	150.42	-33.57	✓	✓	✓
<i>Z. murphyi</i>	440349	NSW, Cones Walk, Bundanoon	NSW 440349	150.30	-34.65	✓	✓	✓
<i>Z. obcordata</i> A.Cunn.	608324	NSW, 'The Rocks' property, W of Bathurst	NSW 608324	149.24	-33.42	✓		
<i>Z. obcordata</i>	RGC10060	NSW, 'Bulbudgerie' property, near Wuuluman, 15 km ENE of Wellington	MEL 2067510	149.10	-32.53	✓	✓	

Appendix A

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. obcordata</i>	231013	NSW, 'Bulbudgerie', private property, Uungala Rd, Wuuluman	MELU D 112159a	149.11	-32.52			✓
<i>Z. obovata</i> (C.T.White) J.A.Armstr.	KRM8071	Qld, Grant Creek, Anniversary Falls on Rush Track, W of Herberton	BRI AQ861639	145.37	-17.37	✓	✓	
<i>Z. odorifera</i> J.A.Armstr. subsp. <i>odorifera</i>	RM1196	NSW, ~9 km (direct) SW of Molong Bocoble Gap rocky bluff immediately S of road and railway line in gap near base of bluff	MEL 717377, BRI AQ591376, NSW 748756	148.81	-33.16	✓		
<i>Z. odorifera</i> subsp. <i>warrabahensis</i> Duretto & P.I.Forst.	599824	NSW, Warrabah, W of Kingstown, on private property	NSW 599824	151.01	-30.53	✓		
<i>Z. odorifera</i> subsp. <i>williamsii</i> Duretto & P.I.Forst.	MJB2018	NSW, cult. ANBG. Wild source: 2.4 km from Howell on track to Copeton Lake Recreation area (voucher for stock is CBG 8505944)	MELU D 112168a	151.05	-29.95	✓	✓	
<i>Z. oreocena</i> J.A.Armstr.	DJO222	Vic, Grampians NP, Cultivation Creek Gorge, 100 m upstream from where Buandik Falls walking track ends, 1 m W of the creek	MELU D 112140a	142.28	-37.25	✓	✓	✓
<i>Z. oreocena</i>	MJB1982	Vic, Grampians NP, Deep Creek, ~150 m upstream of crossing on Harrop Track	MEL D 112146a	142.27	-37.28	✓	✓	✓
<i>Z. oreocena</i>	RAB580	Vic, Hollow Mountain, Northern Grampians, in sheltered rock outcrop, S side of upper cave, just below summit	MELU D 112126a	142.38	-36.89	✓	✓	✓
<i>Z. oreocena</i>	RAB581	Vic, Northern Grampians, Hollow Mountain, walking track to Summer Day Valley climbing area	MELU D 112127a	142.39	-36.89	✓	✓	✓
<i>Z. parrisiae</i> J.D.Briggs & J.A.Armstr.	RJ1899	NSW, Box Range Farm, ~10 km W of Pambula	MEL 2312304	149.76	-36.94	✓	✓	
<i>Z. pilosa</i> Rudge	MFD2008	NSW, Berowra NP (Sydney Sandstone North)	NSW	151.15	-33.62	✓	✓	

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. pilosa</i>	RAB502	NSW, Georges River NP, Picnic Point on ridge track behind the second block of toilets from the road entrance	MELU D 112151a	151.02	-33.98	✓	✓	
<i>Z. prostrata</i> J.A.Armstr.	758677	NSW, Look-at-me-now Headland, ~150 m SE of car park	NSW 758677	153.19	-30.18	✓	✓	
<i>Z. prostrata</i>	MJB2019	NSW, cult. ANBG. Wild source: Dammerels Headland (voucher for stock is CBG 8803922)	MELU D 112169a	153.20	-30.16	✓	✓	
<i>Z. rimulosa</i> C.T.White	JRC5802	Qld, Mount Mulligan, ~40 km NW of Dimbulah, top of mountain, ~0.5 km SE of dam, rock outcrop	MEL 682024, BRI AQ398829	144.86	-16.88	✓	✓	
<i>Z. robertsiorum</i> J.A.Armstr.	AF5349	Qld, Malbon Thompson Forest Reserve, NE of Bell Peak, North site 194	BRI AQ848776	145.88	-17.09	✓	✓	
<i>Z. robusta</i> Maiden & Betche	DEA s.n.	Vic, 3.7 km S of Little Mount Kent, along walking track to Moroka Gorge from Horseyard Flat above N facing sandstone cliff	MEL 2019411	146.98	-37.46	✓	✓	✓
<i>Z. robusta</i>	MFD2049	NSW, cult. ABG, Mount Annan. Wild source: Kanangra–Boyd NP, Kanangra Brook Falls, Kanangra Walls, below waterfall	NSW, AD950185	150.11	-33.99	✓	✓	
<i>Z. scopulus</i> Duretto & P.I.Forst.	SPP1980	Qld, Flinders Peak, ~22 km S of Ipswich	BRI AQ747759	152.81	-27.81	✓	✓	
<i>Z. smithii</i> Jacks.	AF5659	Qld, Herberton Range SF, steep upper slope of Mount Misch, W of Tolga	CNS 134680	145.42	-17.23	✓	✓	
<i>Z. smithii</i>	DJO327	Qld, Kroombit Tops NP, W of Dry Creek, 60 m downstream from the road crossing of Dry Creek (300 m E of the Barracks)	MELU D 112142a	150.97	-24.37	✓	✓	
<i>Z. smithii</i>	DJO430	NSW, Hat Head NP, 100 m from the start of the Korogoro Point Track circuit, (walking anticlockwise)	MELU D 112144a	153.06	-31.06	✓		

Appendix A

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. smithii</i>	KRM8282	Qld, Lot 5, Webster Road, near Wondecla	BRI AQ747081	145.46	-17.44	✓	✓	
<i>Z. smithii</i>	MJB2023	Qld, Currigee, South Stradbroke Island	MELU D 112171a	153.42	-27.83	✓	✓	
<i>Z. smithii</i>	MJB2029	Vic, Briagolong–Stockdale road, ~300 m W of Stratford-Stockdale Road, Gippsland	MELU D 112148a	147.15	-37.85	✓	✓	✓
<i>Z. smithii</i>	MJB2184	Vic, Croagingalong NP, along walking track to Genoa Peak, ridgeline below summit	MELU D 112149a	149.63	-37.53	✓	✓	
<i>Z. smithii</i>	MTM1127	Qld, Kureelpa Falls area, Mapleton SF	BRI AQ759541	152.89	-26.60	✓	✓	
<i>Z. smithii</i>	MTM275	Qld, Kureelpa Falls, South Maroochy River, ~4 km NE of Mapleton via Kureelpa	BRI AQ746002	152.89	-26.59	✓	✓	
<i>Z. smithii</i>	MTM303	Qld, Clewley’s Gap area, Kroombit Tops NP	BRI AQ813593	150.98	-24.39	✓	✓	
<i>Z. smithii</i>	MTM926	Qld, North Stradbroke Island, E of Amity Point	BRI AQ757196	153.47	-27.41	✓	✓	
<i>Z. smithii</i>	MTM929	Qld, Cainbale Creek catchment, just W of Cainbale Falls, Lamington NP	BRI AQ757195	153.11	-28.19	✓	✓	
<i>Z. smithii</i>	PIF33151	Qld, Mount Walsh NP, Palm Valley area, 15 km S of Biggenden Coast Range	BRI AQ729276	152.04	-25.65	✓	✓	
<i>Z. smithii</i>	PIF34139	Qld, Main Range NP, start of Mount Cordeaux walking track	BRI AQ743453	152.40	-28.06	✓	✓	
<i>Z. smithii</i>	PIF36221	Qld, Mount Barney NP, southern base of Mount Barney	BRI AQ813882	152.68	-28.26	✓	✓	
<i>Z. smithii</i>	PIF38231	Qld, Bulburin NP, Dawes Range Road	BRI AQ818053	151.51	-24.59	✓	✓	
<i>Z. smithii</i>	PIF38266	Qld, Toolara SF (SF 1004), Harrys Hut Road, ~200 m from Cooloola Way	BRI AQ823442	152.96	-26.19	✓	✓	

Taxon	ID number in this study*	Location	Voucher	Longitude (decimal)	Latitude (decimal)	cpDNA study (Ch. 3)	nrDNA study (Ch. 4)	<i>Z. arborescens</i> study (Ch. 5)
<i>Z. smithii</i>	PIF38276	Qld, Conondale NP (formerly Imbil SF 2, Ryan LA)	BRI AQ818086	152.59	-26.46	✓	✓	
<i>Z. smithii</i>	PIF38466	Qld, Lyrebird Ridge Road, Springbrook	BRI AQ818295	153.26	-28.21	✓	✓	✓
<i>Z. smithii</i>	RAB511	Vic, Wingan Inlet Croagingalong NP, Rams Head Track	MELU D 112158a	149.50	-37.75	✓	✓	
<i>Z. smithii</i>	RAB548	Vic, Moilun Creek, Stockdale to Glenaladale Road (Beverleys Road), just upstream from crossing	MELU D 112115a	147.26	-37.76	✓	✓	
<i>Z. smithii</i>	SPP2032	Qld, Mount Nebo, near Brisbane, Scrub Road ~4 km directly SE of Jollys Lookout, Mount Nebo D'Aguilar Range	BRI AQ748343	152.83	-27.43	✓	✓	
<i>Z. southwellii</i> J.A.Armstr.	MTM251	Qld, Boy-ull Creek, ~100 m upstream from Twin Falls, Springbrook NP	BRI AQ745450	153.27	-28.22	✓	✓	✓
<i>Z. tenuis</i> Duretto & P.I.Forst.	Bean4609	Qld, Bertya Creek, W of Warang, White Mountains NP	MEL 719859, BRI AQ544935	144.78	-20.46	✓	✓	
<i>Z. tuberculata</i> J.A.Armstr.	RAB505	NSW, Tilba Central, on the hill with the water tower	MELU D 112153a	150.07	-36.31	✓	✓	
<i>Z. vagans</i> Duretto & P.I.Forst.	PIF34832	Qld, Gurgeena Plateau, SF 172, near Meredith's Road	BRI AQ745514	151.38	-25.45	✓	✓	
<i>Z. veronicea</i> (F.Muell.) Benth. subsp. <i>veronicea</i>	MJB1980	SA, Cox Scrub Conservation Reserve	MEL 2383600	138.73	-35.34	✓	✓	
<i>Z. veronicea</i> subsp. <i>insularis</i> J.A.Armstr.	WN14AE	SA, Kangaroo Island, Finders Chase NP, Yacca Flat Road	MELU D 108693	-36.02	136.73	✓	✓	
<i>Z. verrucosa</i> J.A.Armstr.	PIF34734	Qld, Blue Hills, 3 km NNE of Monogorilby	BRI AQ745491	151.01	-26.03	✓	✓	
<i>Z. whitei</i> J.A.Armstr. ex Duretto & P.I.Forst.	AF5486	Qld, Oaky Creek Falls, ~1 km SW of Mt Emerald	BRI AQ755971	145.40	-17.21	✓	✓	

Appendix B. Polymorphisms in ITS and ETS sequences

This table documents the number of polymorphisms in ITS and ETS sequences (Chapter 4), noting samples that were sequenced for the study but subsequently excluded from the nrDNA analyses because of the high number of polymorphisms, or missing data. Accessions included in the study but not listed here had no polymorphisms detected in either ITS or ETS sequences. The species names of samples with multiple accessions have not been repeated, but identification or collecting numbers are listed after the species name. Question marks indicate that amplification was unsuccessful for that region, so the number of polymorphisms is unknown.

Samples	ITS	ETS	Comments
<i>Z. actites</i> PIF38224	11, most in ITS 1	8	Excluded from nrDNA analyses
<i>Z. adenodonta</i> MTM760	0	1	
<i>Z. arborescens</i> (subsp. <i>glabrifolia</i>) MTM213			
<i>Z. baeuerlenii</i> 275527, MJB2011			
<i>Z. bifida</i> MTM537			
<i>Z. buxijugum</i> RJ1897			
<i>Z. distans</i> CMH157			
<i>Z. hydroscopica</i> PIF 38247			
<i>Z. laevigata</i> MTM258			
<i>Z. laxiflora</i> MTM385			
<i>Z. prostrata</i> 758677			
<i>Z. alata</i> AF 5178	?	0	Excluded from nrDNA analyses
<i>Z. arborescens</i> DJO110, MJB1868	1	0	
<i>Z. littoralis</i> RAB508			
<i>Z. parrisiae</i> RJ1899			
<i>Z. robertsiorum</i> AF5349			
<i>Z. smithii</i> MJB2029, PIF38231, RAB548			
<i>Z. verrucosa</i> PIF34734			
<i>Z. chevalieri</i> MJB2119	1	1	
<i>Z. distans</i> MTM287			
<i>Z. granulata</i> MJB2016			
<i>Z. littoralis</i> MJB2017			
<i>Z. veronicea</i> subsp. <i>insularis</i> WN14AE			
<i>Z. involucrata</i> 506074	2	2	
<i>Z. citriodora</i> NGW2038	2	0	
<i>Z. compacta</i> PIF34479			
<i>Z. murphyi</i> JA356			
<i>Z. oreocena</i> DJO222			
<i>Z. smithii</i> PIF38266			

Appendix B

Samples	ITS	ETS	Comments
<i>Z. caducibracteata</i> IRT9585	1	2	
<i>Z. collina</i> MTM507			
<i>Z. scopulus</i> SPP1980			
<i>Z. tuberculata</i> RAB505			
<i>Z. collina</i> MJB2012	2	4	
<i>Z. covenyi</i> MJB2013	1	3	
<i>Z. southwellii</i> MTM251			
<i>Z. cytisoides</i> DJO320	11 in ITS1	0	Excluded from nrDNA analyses
<i>Z. exsul</i> MTM 535	19	Many	Cloned
<i>Z. furfuracea</i> subsp. <i>furfuracea</i> 466549	0	3	
<i>Z. insularis</i> PIF34666	25	33	Cloned
<i>Z. lasiocaulis</i> 492944	0	2	
<i>Z. montana</i> PIF36219, <i>Z. smithii</i> KRM8282	3	0	
<i>Z. obcordata</i> 608324	?	0	Excluded from nrDNA analyses
<i>Z. odorifera</i> subsp. <i>warrabahensis</i> 599824	23	26	Excluded from nrDNA analyses
<i>Z. odorifera</i> subsp. <i>williamsii</i> MJB2018	33	27	Cloned
<i>Z. odorifera</i> subsp. <i>odorifera</i> RM1196	17	25	Excluded from nrDNA analyses
<i>Z. prostrata</i> MJB2019	23	11	Cloned
<i>Z. robusta</i> DEA s.n.	3	2	
<i>Z. robusta</i> MFD2049	0	6	Included in the nrDNA analyses despite the number of ETS polymorphisms; the taxon was of special interest because it was polyphyletic based on cpDNA
<i>Z. smithii</i> DJO430	?	?	Excluded from nrDNA analyses
<i>Z. smithii</i> MTM275	4	3	

Appendix C. Concatenation of cloned sequences

Given that cloning had produced divergent copies of ITS and ETS in four samples (Chapter 4), possibly representing different lineages of nrDNA in those individuals, options for handling that sequence data in a combined analysis were considered. The options were to: (1) exclude those accessions from the combined analysis because of the problem of assessing the homology of the cloned sequences, but risk losing important and interesting information about the history of particular accessions, taxa and populations. (2) include the individual cloned sequences as separate entities in the combined analysis, and have ‘missing data’ for the second region (i.e., a sequence including ITS data would have nearly 500 base pairs of missing data for the ETS part), but risk poor resolution and reduced support for nodes. (3) assess the phylogenetic signal of the individual ITS and ETS sequences, select and pair them thoughtfully for a combined data analysis (thus avoiding the introduction of large amount of missing data), with the intention of producing a single tree with improved overall resolution and increased support for nodes, rather than two poorly resolved trees.

Notwithstanding the shortcomings of option 3 (assessing and joining the cloned sequences), this was the approach adopted in this study, particularly as independent ITS and ETS analyses showed noteworthy congruence in the resolution of some cloned sequences (see below for details). Copies that showed identical or highly similar relationships in both analyses (i.e., a congruent phylogenetic signal) were assumed to be from the same nrDNA array, and arguably it was reasonable to concatenate those sequences for inclusion in the combined analysis of the datasets. ITS and ETS regions have been shown to produce more robust infrageneric phylogenies when combined, than when each region is analysed in isolation (Baldwin & Markos 1998, Bena *et al.* 1998, Clevinger & Panero 2000). In five instances in a preliminary analyses, ITS and ETS paralogues showed clear congruence, i.e., both ITS and ETS paralogues from an individual resolved identical or highly similar phylogenetic relationships. Other copies were either less well resolved or resolved in one but not in both trees (see Figures 4.3 and 4.4 showing the ITS and ETS trees, respectively). When similar copies were giving the same phylogenetic signal, a single copy was chosen to represent that lineage. In a few cases, the ITS and ETS copies were assumed to have originated from the same

nrDNA array on the basis of their placement in both trees *outside* clades containing one or more of the other copies. The justification for selecting and joining ITS and ETS copies for a combined data analysis is described below. (Sequence ‘pairs’ are summarised in Chapter 4, Table 4.3).

- *Zieria exsul* MTM535:

ITS and ETS copies from *Zieria exsul* MTM535 were joined because they were resolved in clades with *Z. fraseri* and *Z. laxiflora* in both analyses (ITS node 43, ETS node 35). *Zieria exsul* ITS copy 2661 and two accessions of *Z. minutiflora* subsp. *minutiflora* were resolved as sister to samples on ITS node 43 which contained other ITS copies from *Zieria exsul* MTM535. ETS copy 2698 was resolved with the same two accessions of *Z. minutiflora* subsp. *minutiflora* (ETS node 48) and so ITS copy 2661 and ETS copy 2698 were selected for joining.

- *Zieria insularis* PIF34666:

ITS copy 2653 and ETS copy 2682 from *Z. insularis* PIF34666 were joined because they were well supported as sister to *Z. aspalathoides* subsp. *aspalathoides* in both analyses (ITS node 24, ETS node 19). ITS copy 2656 from *Z. insularis* PIF34666 was resolved with *Z. eungellaensis*, *Z. robertsiorum*, *Z. scopulus* and three *Z. smithii* accessions (ITS node 49), and was joined with ETS copy 2681 which was resolved with the same accessions in the ETS tree (ETS node 43), but with accessions of some other taxa. (The three *Z. insularis* ITS copies on node 25 might represent a separate (third) ITS lineage within this sample, but their relationship to other ITS sequences was uncertain, and a corresponding third lineage in ETS was not detected, so none of these copies was included in the combined analysis).

- *Zieria odorifera* subsp. *williamsii* MJB2018:

ITS copy 2705 and ETS copy 2679 from *Zieria odorifera* subsp. *williamsii* MJB2018 were joined because they were resolved as sister to *Z. obcordata* RGC10060 in both analyses (ITS node 19, ETS node 24). Both trees contained other copies of ITS and ETS from *Z. odorifera* subsp. *williamsii* that were resolved outside that clade, so ITS copy 2706 (node 5) and ETS copy 2678 (node 22) were selected for joining.

- *Zieria prostrata* MJB2019:

ITS copy 2714 and ETS copy 2668 from *Zieria prostrata* MJB2019 were joined because they were resolved in both analyses in clades with *Z. southwellii*, *Z. smithii* PIF36219, *Z. montana* and *Z. lasiocaulis* (ITS node 15, ETS node 3), although the ETS clade contained accessions of other taxa as well. ITS copy 2701 and ETS copy 2667 from *Zieria prostrata* MJB2019 were joined because they were resolved in both analyses in clades with *Z. furfuracea* subsp. *furfuracea* and accessions of *Z. smithii* (ITS node 56, ETS node 33).

Appendix D. Congruence between proposed phylogenies of *Zieria*

This table indicates the high level of incongruence between proposed phylogenies of *Zieria* based on nrDNA (Figure 4.5 in Chapter 4) and cpDNA (Figure 2 in Chapter 3). In all cases there was at least one, but sometimes many, well-supported chloroplast nodes (PP ≥ 0.95) that were in conflict with well-supported nrDNA nodes (PP ≥ 0.95). Samples at congruent nodes only are listed.

ITS+ETS node	BS	PP	Conflict with cpDNA tree (≥ 0.95 PP)	Combined cpDNA node	BS	PP	Samples (at congruent nodes)
1	87	1.00	yes				
2		0.85	—				
3	99	1.00	no	5	100	1.00	<i>Z. veronicea</i> (two accessions of different subspecies)
4		0.08	—				
5	62	0.93	—				
6	100	1.00	no	39	96	1.00	<i>Z. pilosa</i> (two accessions)
7	100	1.00	yes				
8	62	1.00	yes				
9	70	1.00	yes				
10	99	1.00	no	50	63	1.00	<i>Z. involucrata</i> (two accessions)
11	54	0.77	—				
12	100	1.00	yes				
13	99	1.00	yes				
14	86	1.00	yes				
15	64	0.97	yes				
16	65	0.93	—				
17	63	0.97	yes				
18	61	1.00	yes				
19	95	1.00	yes				
20	63	0.97	yes				
21		1.00	yes				
22	98	1.00	yes				
23	91	1.00	yes				
24		0.87	—				
25	75	0.89	yes				
26	100	1.00	yes				
27	100	1.00	no	9	100	1.00	<i>Z. ingramii</i> (two accessions)

Appendix D

ITS+ETS node	BS	PP	Conflict with cpDNA tree (≥ 0.95 PP)	Combined cpDNA node	BS	PP	Samples (at congruent nodes)
28	100	1.00	yes				
29	69	0.97	yes				
30		0.076	—				
31	100	1.00	no	8	100	1.00	<i>Z. laevigata</i> (two accessions)
32	73	0.75					
33	100	1.00	yes				
34	64	0.91	—				
35	81	1.00	yes				
36	99	1.00	no	22	98	1.00	<i>Z. smithii</i> (two accessions)
37	100	1.00	no	20	85	1.00	<i>Z. littoralis</i> (two accessions)
38	93	1.00	yes				
39	95	1.00	yes				
40	63	1.00	no	59	85	1.00	<i>Z. smithii</i> (three accessions)
41		0.57	—				
42	94	1.00	yes				
43	95	1.00	no	36	0.98	—	<i>Z. graniticola</i> , <i>Z. inexpectata</i>
44	84	1.00	yes				
45	61	0.80					
46	89	1.00	yes				
47	92	1.00	no	21	98	1.00	<i>Z. buxijugum</i> , <i>Z. formosa</i> , <i>Z. granulata</i> , <i>Z. parrisiae</i> , <i>Z. tuberculata</i>
48	100	1.00	no	28	97	1.00	<i>Z. baeuerlenii</i> (two accessions)
49		0.96	yes				
50	91	1.00	yes				
51	95	1.00	yes				
52	99	1.00	yes				
53	95	1.00	yes				
54	86	1.00	yes				
55	89	1.00	yes				
56	58	0.95	yes				
57	62	0.97	yes				
58		0.95	yes				
59	99	1.00	yes				
60	53	0.89					
61	90	1.00	yes				
62	87	1.00	yes				

ITS+ETS node	BS	PP	Conflict with cpDNA tree (≥ 0.95 PP)	Combined cpDNA node	BS	PP	Samples (at congruent nodes)
63		0.53					
64	69	0.66					
65	73	0.89					
66	94	1.00	no	22	98	1.00	<i>Z. smithii</i> (two accessions)
67	96	1.00	yes				
68	88	1.00	no	53	97	1.00	<i>Z. prostrata</i> (two accessions)