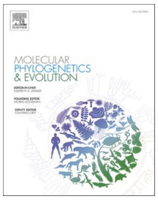
Molecular Phylogenetics and Evolution 204 (2025) 108250

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev



Forget-me-not phylogenomics: Improving the resolution and taxonomy of a rapid island and mountain radiation in Aotearoa New Zealand (*Myosotis*; Boraginaceae)

Heidi M. Meudt[a,*], Sofie M. Pearson[a,b,1], Weixuan Ning[b,2], Jessica M. Prebble[c], Jennifer A. Tate[b]

[a] *Museum of New Zealand Te Papa Tongarewa, PO Box 467, Cable St, Wellington 6140, New Zealand*
[b] *School of Agriculture and Environment, Massey University, Private Bag 11222, Palmerston North 4412, New Zealand*
[c] *Manaaki Whenua – Landcare Research, PO Box 69040, Lincoln 7640, New Zealand*

ARTICLE INFO

Keywords:
Angiosperms353
Biogeography
Genome skimming
Plastome
Southern hemisphere
Species delimitation

ABSTRACT

Island and mountain systems represent natural laboratories for studies of species radiations, but they often present several challenges for phylogenetic inference and species delimitation. The southern hemisphere forget-me-nots (*Myosotis*, Boraginaceae) comprise a geologically recent radiation centred in Aotearoa New Zealand, a mountainous archipelago, with about 50 species that are morphologically and ecologically divergent but lack genetic variation sufficient to resolve phylogenetic relationships and species boundaries using standard DNA Sanger sequencing markers, AFLPs, or microsatellites. Many of these *Myosotis* species are geographically restricted in alpine areas, uncommon or threatened, have polyploid and dysploid genomes, and are of high taxonomic and conservation priority. Here we present phylogenomic analyses using target-capture of Angiosperms353 baits, and genome skimming of whole plastomes and nrDNA, to improve resolution of the radiation, explore biogeographic and morphological patterns within it, and address specific taxonomic questions for each species. Our comprehensive sampling includes over 300 individuals representing nearly all species from New Zealand and Australia, which is ~ 2–3 × more taxon sampling and ~ 80–120 × more molecular data than previously published for *Myosotis.* Exploration of different data filtering, curation and analyses (coalescent vs. concatenation) improved the resolution of the Angiosperms353 tree, which despite short backbone branches with low support values, showed taxonomic and geographic patterns, including multiple switches between ebracteate and bracteate inflorescences and multiple expansions within New Zealand from Te Waipounamu South Island to Te Ika-a-Māui North Island, Rakiura Stewart Island, subantarctic islands, and Australia. Some of these patterns were also seen in the genome skimming datasets, and comparison of the three datasets was useful for improving our understanding of the taxonomy and resolution of this radiation. Although this phylogenomic study does not fully overcome all of the challenges regarding species delimitation of this rapid island and mountain species radiation, it nevertheless makes an important contribution to an integrative taxonomic revision of the southern hemisphere species of *Myosotis.*

1. Introduction

Islands and mountains represent natural laboratories for studies of endemic radiations of morphologically or ecologically distinct species (Meudt et al., 2021; Muellner-Riehl et al., 2024; Warren et al., 2015). Such radiations often present several challenges for phylogenetic inference and species delimitation (Lagomarsino et al., 2022; Thomas et al., 2021). This is due to a combination of recent, rapid and

---

* Corresponding author.
E-mail addresses: heidim@tepapa.govt.nz (H.M. Meudt), sofie.pearson@uq.edu.au (S.M. Pearson), weixuan@iastate.edu (W. Ning), prebblej@landcareresearch.co.nz (J.M. Prebble), j.tate@massey.ac.nz (J.A. Tate).

[1] Present address: Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Hermitage Research Facility, Warwick , QLD, 4370, Australia.

[2] Present address: Iowa State University, Ames, IA, 50011, USA.

https://doi.org/10.1016/j.ympev.2024.108250
Received 25 June 2024; Received in revised form 7 November 2024; Accepted 20 November 2024

incomplete speciation, coupled with processes such as hybridisation, polyploidisation, and incomplete lineage sorting, which can lead to complex or unresolved evolutionary histories (Kadereit and Abbott, 2021; Maurin et al., 2022; Meudt et al., 2015). Over the past three decades, molecular phylogenies based on DNA sequencing have greatly improved our knowledge and understanding of island and mountain plant species radiations – including phylogeny, biogeography, taxonomy and species limits, among many others (Barrett et al., 2016; Soltis et al., 2013; Straub et al., 2012). Yet for many plant species radiations, such advances have so far proven elusive.

The southern hemisphere forget-me-nots (*Myosotis,* Boraginaceae) are a recent species radiation whose resolution and taxonomy could improve using phylogenomic inference as part of an integrative systematic approach (Meudt, 2017; Meudt et al., 2015; Prebble et al., 2022; Winkworth et al., 2002b). About half of the c.100 species of *Myosotis* worldwide are found in the northern hemisphere (Weigend et al., 2016), and the other half comprise the monophyletic southern hemisphere lineage (Meudt et al., 2015; Winkworth et al., 2002b). Aotearoa New Zealand – a mountainous island system – is the main centre of *Myosotis* diversity in the southern hemisphere, with c. 50 species and subspecies. Only a few species are found in Australia (native *M. australis* subsp. *australis* and endemic *M. exarrhena;* Meudt et al., 2020), New Guinea (endemic *M. australis* subsp. *saruwagedica*; Meudt et al., 2020), or southern South America (endemic *M. albiflora* and native *M. antarctica* subsp. *antarctica*; Meudt and Prebble, 2018; Prebble et al., 2022).

The southern hemisphere species are morphologically and ecologically diverse and can be grouped into two main informal morphological groups (Robertson, 1989). Ebracteate-erect species have bracts at the base of the erect inflorescence only, whereas bracteate-prostrate species also have bracts subtending or near the flowers of the mostly prostrate inflorescences or branches (Meudt et al., 2015; Robertson, 1989). In the family, bracts on inflorescences are historically important but also evolutionary labile (Frohlich, 1978; Cohen, 2014; Frohlich et al., 2022). In *Myosotis,* this morphological character is the basis of a useful informal classification (Robertson, 1989), is prominent in dichotomous keys (i.e., Moore, 1961) and has guided the progression of a series of recent taxonomic studies (e.g., Meudt, 2016, 2021; Meudt and Prebble, 2018), but it is unknown whether these two morphological groups are monophyletic. Nearly all native New Zealand *Myosotis* are endemic and have restricted, habitat-specific geographic ranges, with a centre of diversity in Te Waipounamu South Island, where 82 % of the species are endemic (Table 1). Although previous phylogenies showed multiple dispersals from New Zealand to other southern hemisphere areas (Winkworth et al., 2002a, Meudt et al. 2015), finer scale phylogeographic patterns within the New Zealand archipelago require further study. The majority of New Zealand species are classified as Threatened, At Risk–Naturally Uncommon, or Data Deficient according to the New Zealand Threatened Classification System (NZTCS; De Lange et al., 2018). Therefore, taxonomic revision of *Myosotis* is a high priority, as the last comprehensive taxonomic treatment of New Zealand *Myosotis* was completed in 1961 (Moore, 1961). Since then, new species have been described (e.g., Lehnebach, 2012; Meudt et al., 2013; Meudt and Prebble, 2022, 2018; Moore and Simpson, 1973) and several (c. 20) morphologically distinct putative species that were given informal 'tag names' have been studied (e.g., Meudt and Prebble, 2022, 2018; Prebble et al., 2022). A recent focus on specific species groups over the past decade led to c. 70 % of the southern hemisphere *Myosotis* species, subspecies and tag names being revised taxonomically (Meudt, 2021; Meudt et al., 2020, 2013; Meudt and Prebble, 2022, 2018; Prebble et al., 2022, 2019). The remaining 30 % of native New Zealand species are under morphological and palynological study now for a final taxonomic revisionary paper in this series (Meudt unpubl. data).

To date, these studies have mostly relied on the comparison of morphological data, although DNA (Meudt et al., 2015, 2013; Prebble et al., 2019, 2015), pollen (Meudt, 2016; Meudt and Prebble, 2018), and ecological niche modelling (Prebble et al., 2022) have also been co-

Table 1

Species and subspecies of Southern Hemisphere *Myosotis,* including their main morphological group (based on inflorescence type and habit, either ebracteate erect or bracteate prostrate), the number of individuals sampled in this paper (*n* = 326 total), and their geographic distribution. Note that numbers of sampled individuals may be slightly lower than those listed in Table 1, depending on which samples were successfully sequenced for each of the individual datasets.

Southern hemisphere *Myosotis* species	Main morphological group (inflorescence type + habit)	No. sampled in this paper (n = 326)	Geographic distribution
M. albiflora **Banks & Sol. ex Hook.f.**	bracteate prostrate	0	Chile: Aisén, Magallanes; Argentina: Tierra del Fuego
M. albosericea **Hook.f.**	ebracteate erect	3	New Zealand: South Island
M. amabilis **Cheeseman**	ebracteate erect	0	New Zealand: North Island
M. angustata **Cheeseman**	ebracteate erect	4	New Zealand: South Island
M. antarctica **Hook. f. subsp. *antarctica***	bracteate prostrate	21	New Zealand: North, South, StewartIslands, subantarctic islands (Campbell Island); Chile: Magallanes
M. antarctica **subsp. *traillii*** **Kirk**	bracteate prostrate	12	New Zealand: North, South, Stewart Islands
M. arnoldii **L.B. Moore**	ebracteate erect	4	New Zealand: South Island
M. australis **R.Br. subsp. *australis***	ebracteate erect	22	New Zealand: North and South Islands; Australia: Western Australia, South Australia, New South Wales, Australian Capital Territory, Victoria, Tasmania
M. australis **subsp. *saruwagedica*** **(Schltr. ex Brand) Meudt et al.**	ebracteate erect	0	New Guinea
M. brevis **de Lange & Barkla**	bracteate prostrate	9	New Zealand: North and South Islands
M. brockiei **L.B. Moore & M.J.A. Simpson subsp. *brockiei***	ebracteate erect	5	New Zealand: South Island
M. brockiei **subsp. *dysis*** **Courtney & Meudt**	ebracteate erect	6	New Zealand: South Island
M. bryonoma **Meudt, Prebble & Thorsen**	bracteate prostrate	7	New Zealand: South Island
M. capitata **Hook.f.**	ebracteate erect	4	New Zealand: subantarctic islands (Auckland Islands and Campbell Island)
M. chaffeyorum **Lehnebach**	bracteate prostrate	3	New Zealand: South Island
M. cheesemanii **Petrie**	bracteate prostrate	4	New Zealand: South Island
M. **×** ***cinerascens*** **Petrie**	ebracteate erect, hybrid	0	New Zealand: South Island (Extinct)
M. colensoi **(Kirk) J. F.Macbr.**	bracteate prostrate	8	New Zealand: South Island
M. concinna **Cheeseman**	ebracteate erect	3	New Zealand: South Island
M. exarrhena **F. Muell.**	ebracteate erect	4	Australia: New South Wales, Australian Capital

(*continued on next page*)

Table 1 (*continued*)

Southern hemisphere *Myosotis* species	Main morphological group (inflorescence type + habit)	No. sampled in this paper ($n = 326$)	Geographic distribution
			Territory, Victoria, Tasmania
M. eximia Petrie	ebracteate erect	5	New Zealand: North Island
M. explanata Cheeseman	ebracteate erect	3	New Zealand: South Island
M. forsteri Lehm.	ebracteate erect	10	New Zealand: North and South Islands
M. glabrescens L.B. Moore	bracteate prostrate	3	New Zealand: South Island
M. glauca (G. Simpson & J.S. Thomson) de Lange & Barkla	bracteate prostrate	4	New Zealand: South Island
M. goyenii Petrie subsp. _goyenii_	ebracteate erect	4	New Zealand: South Island
M. goyenii subsp. _infima_ Meudt & Heenan	ebracteate erect	2	New Zealand: South Island
M. hikuwai Meudt, Prebble & G.M. Rogers	bracteate prostrate	3	New Zealand: South Island
M. laeta Cheeseman	ebracteate erect	3	New Zealand: South Island
M. laingii Cheeseman	ebracteate erect	0	New Zealand: South Island (Extinct)
M. lyallii subsp. _elderi_ (L.B.Moore) Meudt & Prebble	bracteate prostrate	14	New Zealand: South Island
M. lyallii Hook.f. subsp. _lyallii_	bracteate prostrate	8	New Zealand: South Island
M. macrantha (Hook.f.) Cheeseman	ebracteate erect	13	New Zealand: South Island
M. matthewsii L.B. Moore	bracteate prostrate	3	New Zealand: North Island
M. monroi Cheeseman	ebracteate erect	3	New Zealand: South Island
M. oreophila Petrie	ebracteate erect	4	New Zealand: South Island
M. pansa (L.B. Moore) Meudt, Prebble, R.J. Stanley & Thorsen subsp. _pansa_	ebracteate erect	6	New Zealand: North Island
M. pansa subsp. _praeceps_ (L.B. Moore) Meudt, Prebble, R.J. Stanley & Thorsen	ebracteate erect	3	New Zealand: North Island
M. petiolata Hook.f.	ebracteate erect	2	New Zealand: North Island
M. pottsiana (L.B. Moore) Meudt, Prebble, R.J. Stanley & Thorsen	ebracteate erect	4	New Zealand: North Island
M. pulvinaris Hook. f.	bracteate prostrate	7	New Zealand: South Island
M. rakiura L.B. Moore	ebracteate erect	6	New Zealand: South and Stewart Islands
M. retrorsa Meudt, Prebble & Hindm.-Walls	bracteate prostrate	6	New Zealand: South Island
M. saxatilis Petrie	ebracteate erect	10	New Zealand: South Island
M. saxosa Hook.f.	ebracteate erect	8	New Zealand: North Island
M. spatulata G. Forst.	bracteate prostrate	7	New Zealand: North, South and Chatham Islands
M. suavis Petrie	ebracteate erect	3	New Zealand: South Island
M. tenericaulis Petrie	bracteate prostrate	6	New Zealand: North, South, and Stewart Islands
M. traversii subsp. _cantabrica_ (L.B. Moore) Meudt	ebracteate erect	4	New Zealand: South Island
M. traversii Hook.f. subsp. _traversii_	ebracteate erect	6	New Zealand: South Island
M. ultramafica Meudt, Prebble & Rance	bracteate prostrate	7	New Zealand: South Island
M. umbrosa Meudt, Prebble & Thorsen	bracteate prostrate	3	New Zealand: South Island
M. uniflora Hook.f.	bracteate prostrate	3	New Zealand: South Island
M. venosa Colenso	ebracteate erect	3	New Zealand: North and South Islands
M. venticola Meudt & Prebble	bracteate prostrate	5	New Zealand: South Island
_M._ "Asbestos"	ebracteate erect	3	New Zealand: South Island
_M._ "Fiordland" [_Myosotis_ (c) (CHR 198630; Fiordland)]	ebracteate erect	3	New Zealand: South Island
_M._ "Livingstone" [_Myosotis_ (f) (CHR 405203; Livingstone)]	ebracteate erect	10	New Zealand: South Island
_M._ "Takitimu"	ebracteate erect	5	New Zealand: South Island
M. cheesemanii × _pulvinaris_	bracteate prostrate, hybrid	3	New Zealand: South Island
M. colensoi × _saxatilis_	hybrid	2	New Zealand: South Island

analysed. Previous genetic studies on southern hemisphere *Myosotis* have applied standard DNA sequencing markers (Meudt et al., 2015; Winkworth et al., 2002b); amplified fragment length polymorphisms (AFLPs, Meudt et al., 2013) or microsatellites (Prebble et al., 2019, 2015) to study species limits, evolutionary relationships, and population genetics of certain species groups of New Zealand *Myosotis*. However, as none of these studies were able to fully resolve species relationships, a new approach is therefore needed to generate a phylogeny of southern hemisphere *Myosotis*.

We chose targeted sequencing using the universal Angiosperms353 baits for the present study. Target capture selects short sequences with highly specific baits, which can work well when using degraded DNA from herbarium specimens (Brewer et al., 2019; Hart et al., 2016; McKain et al., 2018; Shee et al., 2020). The Angiosperms353 baits kit is a set of universal baits that can capture up to 353 single (or low) copy nuclear genes in any flowering plant group (Johnson et al., 2019). Although these baits were designed to capture exon sequences, introns (or non-coding regions) and flanking regions can also be recovered (Johnson et al., 2019). The combination of these regions may be useful for reconstructing shallow phylogenies, including those from rapid species radiations (Johnson et al., 2019; Thomas et al. 2021). To our knowledge, this is the first genus-level study in Boraginaceae to include Angiosperms353 data.

The Angiosperms353 baits have been used on a few New Zealand island and mountain plant species radiations to date: *Veronica* (Thomas et al. 2021; Thomas et al. 2023), *Senecio* (Schmidt-Lebuhn et al., 2022), Celmisiinae (Asteraceae; Nicol et al., 2024), and *Azorella* (Ning 2023; Ning et al. unpubl. data). In these studies, silica-dried and herbarium material of varying ages were successfully sequenced for one or more individuals per species, and the resulting phylogenies showed improved support and resolution compared to those from standard sequencing markers. *Myosotis* is similar to these genera in being a Quaternary species radiation (Kadereit and Abbott, 2021; Maurin et al., 2022; Meudt et al., 2015), with low genetic variation (Kadereit and Abbott, 2021; Maurin et al., 2022; Meudt et al., 2015), and having a history of polyploidy and dysploidy (Grau, 1964; Meudt et al., 2015; Murray and de Lange, 2013). For such genera, "the availability of hundreds of nuclear genes may not necessarily resolve a bifurcating species tree with certainty" (Johnson et al. 2019:603), but the data should nevertheless increase our understanding of their evolutionary history, and improve the taxonomy and delimitation of species, especially when used concertedly with independent (non-molecular) biological and morphological data and knowledge of the study organisms (Frost et al., 2024). Although Angiosperms353 loci are assumed to be low copy nuclear genes (Johnson et al., 2019), this may not always be the case, particularly for polyploids, and there may be a range of loci from truly single copy to multi copy (including loci with multiple alleles or paralogs). A number of strategies are available for dealing with single- and multi-copy loci in downstream analyses of such data (Frost et al., 2024; Gardner et al., 2020; Kandziora et al., 2021; Morales-Briones et al., 2022; Ufimov et al., 2022).

We also complemented the low-copy bi-parentally inherited nuclear dataset with two high-copy DNA genome skimming datasets, i.e., a uniparental plastome dataset and a nuclear ribosomal DNA dataset (Dodsworth, 2015; Straub et al., 2012; Weitemier et al., 2014). From genome skimming, near-complete plastomes can be recovered (Bakker, 2017; Jin et al., 2020; Twyford and Ness, 2017) as well as the full nuclear ribosomal DNA (nrDNA) cistron, comprising 18S, 5.8S and 26S rRNA genes, the internal transcribed spacers ITS1 and ITS2, and the intergenic spacer, IGS, which comprises the external transcribed spacer, ETS, and the non-transcribed spacer, NTS. Such near-complete sequencing of these high-copy regions can improve phylogenetic resolution relative to studies that use only a handful of standard chloroplast markers (Frankel et al., 2022; Fu et al., 2022; Harrison and Kidner, 2011; van Kleinwee et al., 2022) or that include only the internal transcribed spacer (ITS) region, which is a small part of the entire nrDNA tandem repeat. Phylogenetic inference based on plastome DNA may reflect maternal genealogy or phylogeography rather than the organisms' evolutionary history (Dodsworth et al., 2019). Incongruence between plastome and nuclear phylogenies, and chloroplast haplotype sharing between species is common in plants (e.g., Meudt and Bayly, 2008; Rieseberg and Soltis, 1991; Smissen et al., 2004; Wendel and Doyle, 1998) and may be due to convergent mutations, incomplete lineage sorting, or introgressive hybridisation (e.g., French et al., 2016; McKinnon et al., 2001; Meudt and Bayly, 2008). Comparing phylogenetic data from the independent datasets representing the nuclear and chloroplast genomes can provide additional insights into the origins, speciation processes, taxonomy, biogeography, and evolutionary history of a particular group of plants (e.g., Harrison and Kidner, 2011; Stubbs et al., 2020). At least one previously published *Myosotis* phylogenetic study found some evidence of incongruence between nuclear and chloroplast datasets (Meudt et al., 2015).

The main aim of the current study is to reconstruct the relationships of the southern hemisphere *Myosotis* species radiation using phylogenomic sequencing data from three different molecular datasets to quantify congruence with the current taxonomy at multiple levels, explore biogeographic and morphological patterns, and present these phylogenies as frameworks to inform future species-specific research and conservation. Specifically, we aim to compare congruent and incongruent patterns in the Angiosperms353, nrDNA and plastome phylogenies to address: Are the two main morphological groups (ebracteate-erect vs bracteate-prostrate) monophyletic? Has species radiation primarily occurred from Te Waipounamu South Island to other islands (such as Te Ika-a-Māui North Island, Rakiura Stewart Island, and others), i.e., is the South Island both a centre of diversity and a centre of origin? Are species boundaries clearly identifiable based on the phylogenies?

2. Materials and methods

2.1. Taxon sampling

We sampled a total of 332 ingroup and outgroup individuals, including 316 native New Zealand and 10 native Australian *Myosotis* specimens representing 54 different species, subspecies or tag names (hereafter, "taxa"; Table 1; Appendix A Table S1; Fig. 1). As of 23 April 2024, vouchers for 235 (71 %) of these individuals were available on the Global Biodiversity Information Facility (GBIF, 2024). Each of six outgroup species was represented by only one individual, including representatives from two other genera in subfamily Cynoglossoideae sampled from China: *Trigonotis* sp. (which is very closely related to *Myosotis* in tribe Myosotideae; Chacón et al., 2016; Cohen, 2014; Weigend et al., 2013), and *Microula trichocarpa* (Maxim.) I.M.Johnst. in tribe Cynoglosseae, as well as four northern hemisphere *Myosotis* species: *M. arvensis, M. laxa* subsp. *caespitosa*, *M. sylvatica*, and *M. discolor* (Table S1). New Zealand *Myosotis* vouchers and leaf samples stored on silica gel were collected in the field throughout New Zealand from 2005 to 2022, mostly during field trips led by HMM and JMP, with additional collections from colleagues (Table S1; also see 9. Acknowledgments). Our sampling includes multiple individuals (range: 1–22; median: 4; mean: 5) of nearly all native *Myosotis* taxa throughout their ranges in New Zealand and Australia (Table 1; Table S1). Sampling of the two native Australian species was limited to three individuals of *M. australis* subsp. *australis* and four of *M. exarrhena* (Table S1). We were unable to include samples of *M.* × *cinerascens* and *M. laingii* (considered to be extinct), *M. australis* subsp. *saruwagedica* (New Guinea), *M. albiflora* (Chile), or *M. amabilis* (unable to be relocated during field work in New Zealand). The georeferenced localities for all 332 sampled individuals were mapped using R version 4.3.0 and the R package 'maps' (Fig. 2).

2.2. DNA extraction

Total DNA was extracted from silica-dried material using a modified CTAB protocol by Doyle and Doyle (1987) or a modified STE/CTAB protocol (Shepherd and McLay, 2011). The integrity of all DNA extracts was confirmed by electrophoresis in a 1 % Tris-borate-EDTA (TBE) agarose gel. DNA purity and quantities were checked using Nano-Drop™ One (Thermo Scientific, Madison, WI, USA) and Qubit™ dsDNA High-Sensitivity Assay Kit (Invitrogen, ThermoFisher Scientific, Oregon, USA).

2.3. Genomic library preparation and targeted sequence capture using Angiosperms353 baits

Genomic libraries were prepared using the NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina® (New England Biolabs, Ipswich, MA, USA) with all recommended volumes halved for all steps (Johnson et al., 2019). DNA fragmentation was performed using the NEBNext® Ultra II FS Enzyme Mix, with an expected insertion size around 300 bp, and we adjusted the enzyme fragmentation time from 3 to 9 min according to the sample degradation level and input amount (ranged from 16 to 500 ng). Final DNA library concentration and average fragment length were measured using Qubit™ dsDNA High-Sensitivity Assay Kit and Genomic DNA LabChip® Assay to confirm fragment sizes of 400–700 bp. The DNA libraries were subsequently pooled in equimolar amounts prior to hybridizing with the Angiosperms353 bait set (Table S1). Following the

Fig. 1. Images of all 61 southern hemisphere *Myosotis* taxa from Table 1. Attributions are listed in Appendix **E**. Images are listed by row, from left to right. Row A: *M. albiflora, M. albosericea, M. amabilis, M. angustata, M. antarctica* subsp. *antarctica, M. antarctica* subsp. *traillii.* Row B: *M. arnoldii, M. australis* subsp. *australis, M. australis* subsp. *saruwagedica, M. brevis, M. brockiei* subsp. *brockiei, M. brockiei* subsp. *dysis.* Row C: *M. bryonoma, M. capitata, M. chaffeyorum, M. cheesemanii, M.* × *cinerascens, M. colensoi, M. concinna.* Row D: *M. exarrhena, M. eximia, M. explanata, M. forsteri, M. glabrescens, M. glauca.* Row E: *M. goyenii* subsp. *goyenii, M. goyenii* subsp. *infima, M. hikuwai, M. laeta, M. laingii, M. lyallii* subsp. *elderi.* Row F: *M. lyallii* subsp. *lyallii, M. macrantha, M. matthewsii, M. monroi, M. oreophila, M. pansa* subsp. *pansa.* Row G: *M. pansa* subsp. *praeceps, M. petiolata, M. pottsiana, M. pulvinaris, M. rakiura, M. retrorsa, M. saxatilis.* Row H: *M. saxosa, M. spatulata, M. suavis, M. tenericaulis, M. traversii* subsp. *cantabrica.* Row I: *M. traversii* subsp. *traversii, M. ultramafica, M. umbrosa, M. uniflora, M. venosa, M. venticola.* Row J: *M. "Asbestos", M. "Fiordland", M. "Livingstone", M. "Takitimu", M. cheesemanii* × *pulvinaris, M. colensoi* × *saxatilis.*

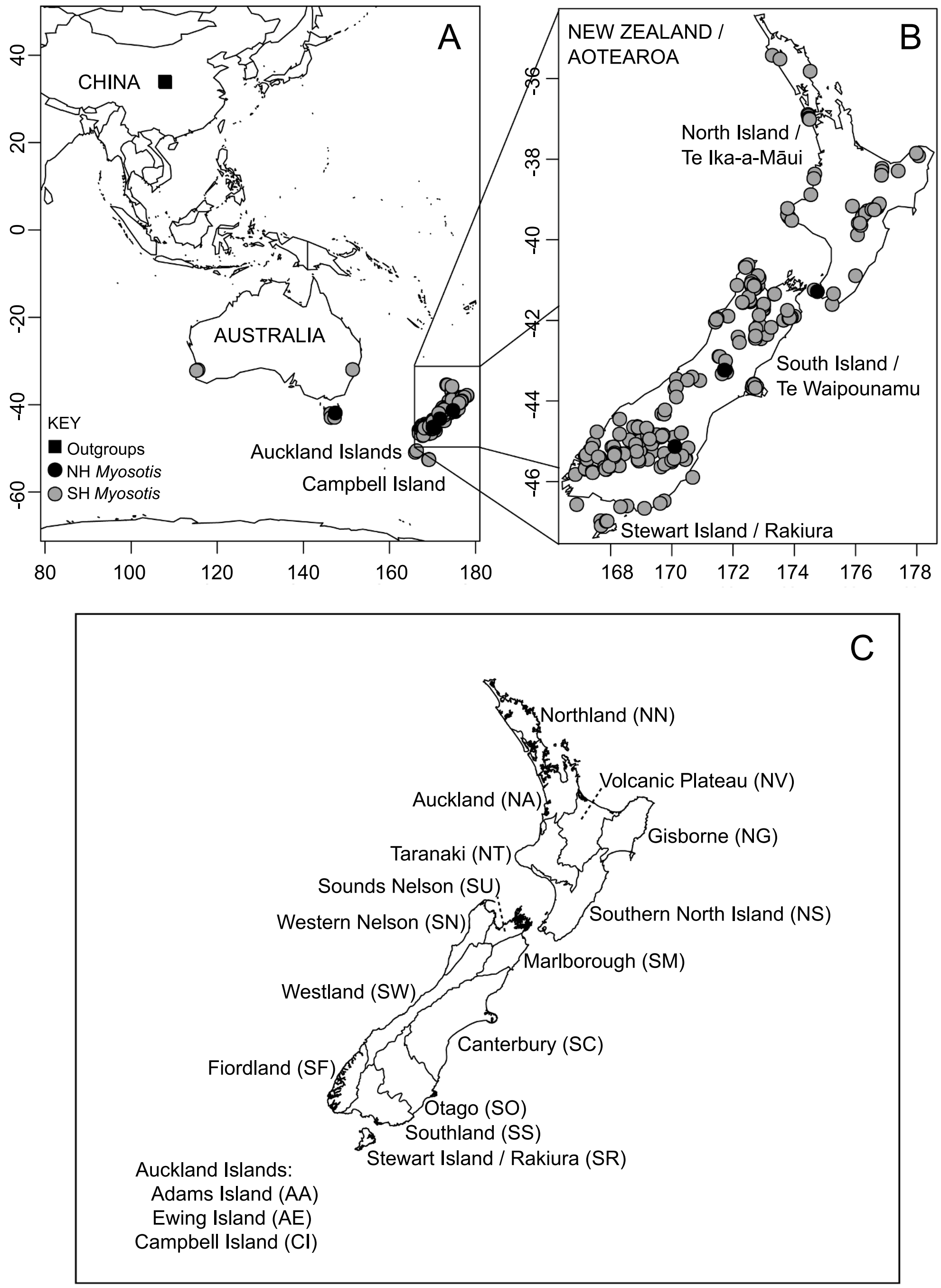


Fig. 2. A. Map showing the collection location of 332 samples included in the study. B. Map of 316 New Zealand collections included in the study. C. Ecological Provinces of Aotearoa New Zealand. For A & B, dark grey circles represent Southern Hemisphere (SH) native *Myosotis* (ingroup samples). Black indicates outgroup samples, circles for Northern Hemisphere (NH) native *Myosotis* introduced to the Southern Hemisphere, squares for other Boraginaceae outgroups (see Table S1 for complete list of specimens sampled and vouchers). The map layer in C was created by Craig Briggs with input from David Glenny (both Manaaki Whenua – Landcare Research). It is based on the classification of Wardle (1991), and was created, with modification, using the Ecological Regions layer (Department of Conservation) and an unknown coastline layer. Helen Curtis (Geospatial Programme Coordinator, Department of Conservation) gave permission to make this layer available.

myBaits™ Hybridization Capture for Targeted NGS Manual v 5.01 using the Standard Protocol, the genomic libraries and Angiosperms353 bait set were incubated 20 to 21 h at 65 °C to complete the hybridisation. After clean-up, the final post-captured library pools were checked for quality with qPCR, Qubit and LabChip at the Massey Genome Service (Massey University, Palmerston North, New Zealand).

A subset of 28 pooled libraries were initially used to determine if the target capture protocol was viable across a range of *Myosotis* species. Hence, 28 samples were sequenced on an Illumina MiSeq producing 150-bp paired-end reads (Illumina, San Diego, CA, USA) at Massey Genome Service (Massey University, Palmerston North, New Zealand). The remaining 304 pooled libraries were sequenced on an Illumina HiSeq X (NovogeneAIT, Singapore) producing 150-bp paired-end reads.

2.4. Genome skim sequencing

Genome skimming for 259 DNA libraries that had sufficient quality and quantity was undertaken. Libraries were pooled into three batches containing 83, 89 and 90 libraries with 900–1000 ng total input. The resulting three batches of pooled libraries were shipped in DNAstable

Tube kits (Biomatrica, product code: 93021–001) and sequenced on three runs of an Illumina HiSeq X (NovogeneAIT, Singapore) producing 150-bp paired-end reads.

2.5. Bioinformatics and phylogenetic analyses

2.5.1. Angiosperms353 dataset

Detailed methods are available in Appendix B, and full scripts are available at https://github.com/SofiePearson/Myosotis. In summary, gene sequences were retrieved by mapping the trimmed Angiosperms353 reads (via Trimmomatic; Bolger et al., 2014) to an Angiosperms353 target reference file (McLay et al., 2021) using BWA v 0.7.17 (Li and Durbin, 2009) in HybPiper v2 (Johnson et al., 2016) Mapped reads were then assembled into contigs with SPAdes v 3.13.1 (Bankevich et al., 2012), and exon sequences were extracted from contigs with exonerate v 2.2.0 (Slater and Birney, 2005). Exon and supercontig lengths were calculated using HybPiper and AMAS.py (Borowiec, 2016). The retrieved gene sequences of 332 individuals (see Appendix B) were aligned using MAFFT v 7.429 (Katoh and Standley, 2013), and then trimmed via trimAl v 1.4.1 (Capella-Gutiérrez et al., 2009).

Paralogs were tabulated for 329 individuals using 'paralog_retriever' in HybPiper. Individuals were filtered to reduce missing and low-quality data (e.g., Hosner et al., 2016), resulting in 291 individuals retained for further analyses. HybPhaser v 2 (Nauheimer et al., 2021) and HybPiper were then used to calculate locus heterozygosity (proportion of loci that contain SNPs) for the 291-individuals in the 322-loci dataset.

In addition, Geneious v 10.2.6 (https://www.geneious.com, accessed 12 June 2024), FigTree v 1.4.4 (https://github.com/rambaut/figtree, accessed 2 September 2024) and HybPiper were used to categorise loci as putative orthologs (single copy) or artifactual orthologs (multi-copy) following the criteria of Frost et al. (2024). Three alignments were subsequently analysed: the full dataset of all retrieved loci (hereafter 322-loci dataset), a reduced single-copy locus dataset (hereafter 56-loci single-copy dataset), and a further reduced single-copy dataset (hereafter 25-loci single-copy dataset).

Phylogenetic trees using multiple-species coalescent and concatenation models of the resulting filtered and curated data subsets (322-loci, 56-loci single-copy and 25-loci single-copy datasets; see Results) were reconstructed using ASTRAL-III (Zhang et al., 2018) with gene tree branches with < 10 % bootstrap (BS) support collapsed using 'nw_ed' (Junier and Zdobnov, 2010), and partition models by locus in IQ-TREE2 v 2.1.3 (Minh et al., 2020) with ModelFinder (Kalyaanamoorthy et al., 2017), respectively. In the first round of analyses of generating the species trees using gene trees, we did not constrain each taxon that had multiple individuals sampled to be monophyletic in ASTRAL-III, as a way to produce a tree showing the phylogenetic relationships of all sampled individuals of all taxa (hereafter, "fully sampled species tree"), many of which were not monophyletic (see Results). In addition, we reran the ASTRAL-III analysis of the 322-loci dataset twice, with a mapping file (–a option) by assigning the individuals of the same taxon to the same tip of 51 taxa (species) or 57 taxa (species and subspecies), to build the species tree (i.e., one branch per taxon) (Rabiee et al., 2019).

Finally, gene concordance factors (gcf) were calculated using PhyParts (Smith et al., 2015) to measure tree support and visualised with pie charts using the R package 'ggtree' (Yu et al., 2017).

2.6. Phylogenies of complete plastome and nrDNA data

Initially, *de novo* assembly of the complete plastome sequences and nrDNA from genome skimming reads was attempted using GetOrganelle (Jin et al., 2020) using default settings. The nrDNA cistrons of all sequenced samples were successfully assembled, whereas plastomes of only 13 out of 259 sequenced individuals were able to be completely *de novo* assembled in GetOrganelle (data not shown). Therefore, a 'mapping to reference' based method was subsequently implemented (see https://github.com/SofiePearson/Myosotis for all scripts), by mapping genome skimming reads of each of the remaining 259 individuals to the selected reference (capi_AA_Dick1, an individual of *M. capitata* that was one of the New Zealand *de novo* assembled *Myosotis* plastomes) and extracting the consensus sequences using SAMtools and BCFtools for each individual (Danecek et al., 2021).

nrDNA and plastome sequences were aligned separately using MAFFT, and the nrDNA alignment was subsequently manually trimmed because the intergenic spacer region (IGS) contained many variable sites which caused difficulties to produce a reliable alignment. Phylogenies reconstructed from the final trimmed nrDNA alignment and plastome alignment (comprising the small and large single-copy regions and one copy of the inverted repeat) were each analysed using IQ-TREE2 with default settings, automatic model detection via ModelFinder and 1000 bootstrapping iterations (−B 1000).

2.7. Comparison of multiple datasets

To determine evolutionary and biogeographic patterns among species of Australasian *Myosotis*, inflorescence type + habit (bracteate-prostrate vs. ebracteate-erect; Meudt et al., 2015; Robertson, 1989) and geography were reconstructed using the Angiosperms353 322-loci final phylogenetic tree, as well as the nrDNA and plastome phylogenies, using the package 'ggtree' in R (see https://github.com/SofiePearson/Myosotis). Tree topologies of the different datasets were also described and compared in detail to allow dissection and understanding of taxonomic and geographic patterns at multiple taxonomic levels. The major clades of each tree were also mapped onto the other trees to visualise congruence and incongruence using the package 'ggtree' in R.

3. Results

3.1. Angiosperms353 baits sequencing

The Illumina sequencing results (Table 2) and the standard HybPiper target capture summary statistics (Table S2) are given for all 329 sequenced individuals. The number of supercontigs, genes with stitched contigs and genes with paralogs are all positively correlated with the number of mapped reads for ingroup *Myosotis* and most outgroups in the 322-loci dataset (Table S2, Fig. S1). As expected, individuals sequenced on the MiSeq had fewer mapped reads compared to the majority of those sequenced on HiSeq (Fig. S1). Outgroups *Microula trichocarpa* (Maxim.) I.M.Johnst., *Trigonotis* sp. and *Myosotis laxa* contained fewer genes with paralogs (25–29) relative to the other three outgroups, *M. arvensis, M. discolor* and *M. sylvatica* (68–78), and most of the ingroup individuals (15–94, mean: 58, median: 69; Table S2, Fig. S1C).

After filtering and trimming of the Angiosperms353 dataset comprised 291 individuals and 322 loci totalling 574,445 bp (Tables 3 and 4; Fig. 3; Fig. S2). Of the individuals sequenced on HiSeq, 92 % (281/304) remained in the final dataset, whereas only 40 % (10/28) of the individuals sequenced on MiSeq individuals were retained. In the HybPhaser output, regardless of platform, most ingroup *Myosotis* individuals showed about the same average SNP percentage (range: 3.1–4.3 %), high values of locus heterozygosity (range: 93.1–99.7), and loci with > 1 % SNPs (range: 88.2–95.1 %; Table S3).

When assessed by eye, the vast majority of the 322 loci (296/322; 92 %) showed characteristics of artifactual orthologs, whereas only 25 loci (8 %) appeared to be single copy and passed the filtering steps. The 296 alignments and trees identified as likely multi-copy clearly had two or more sequence types and resulting clades, respectively, indicating the presence of different alleles or different paralogs. Many of these 296 alignments contained evidence of stitched exons (genes with sequence derived from multiple SPAdes contigs), and their paralogs ('warning by depth' in HybPiper) ranged from 0 ($n = 43$; 14 %) to 322 (median: 69; mean: 125), with 95 of 296 (32 %) having a paralogs warning below 16. By contrast, 24 of the 25 loci (92 %) that were assessed by eye as being

Table 2
HybPiper results of Angiosperms353 baits sequencing by platform and for all 329 ingroup *Myosotis* and outgroup individuals sequenced.

Sequencing platform	Total number of samples	Mean number of sequenced reads	Mean number of mapped reads	Mean mapped percentage	Mean number of genes with contigs	Mean number of extracted supercontigs (=GenesAt150pct)	average exon recovery for each sample across 353 loci, percentage	average length of gene assembled, bp (=HybPiper_length_output_heatmap)
MiSeq	28	689,179	267,931	38.8	312	224	–	–
HiSeq	301	2,281,075	780,229	33.9	333	272	–	–
All	329	2,145,594	736,629	34.3	331	268	90.0	1,792

Table 3
Taxon sampling of the three main southern hemisphere *Myosotis* phylogenomic datasets discussed in this paper: Angiosperms353, nuclear ribosomal DNA (nrDNA) and plastome.

dataset	no. individuals	no. ingroup taxa	no. hybrids	no. outgroups	no. individuals per taxon
Angiosperms353	322	54	2	5	1–19
nrDNA	259	54	2	5	1–17
plastome	259	54	2	5	1–17

Table 4
Comparison of statistics for the three main southern hemisphere *Myosotis* phylogenomic datasets discussed in this paper: Angiosperms353, nuclear ribosomal DNA (nrDNA) and plastome.

dataset	no. individuals	no. loci	total length (bp)	no. parsimony informative sites (bp)	% parsimony informative sites	no. constant sites (bp)	no. singletons (bp)
Angiosperms353	291	322	574,445	122,689	21.0 %	–	–
nrDNA	259	1	6,829	566	8.2 %	59	398
Plastome (second copy of inverted repeat removed)	259	1	138,124	3,268	2.4 %	129,634	5,222

single-copy had paralogs less than 16, and 14 of these (54 %) had a paralogs warning of 0 and thus overlapped with the second single-copy dataset of all 56 loci that had 0 paralogs.

3.2. Angiosperms353 – ASTRAL-III species tree reconstruction and analyses

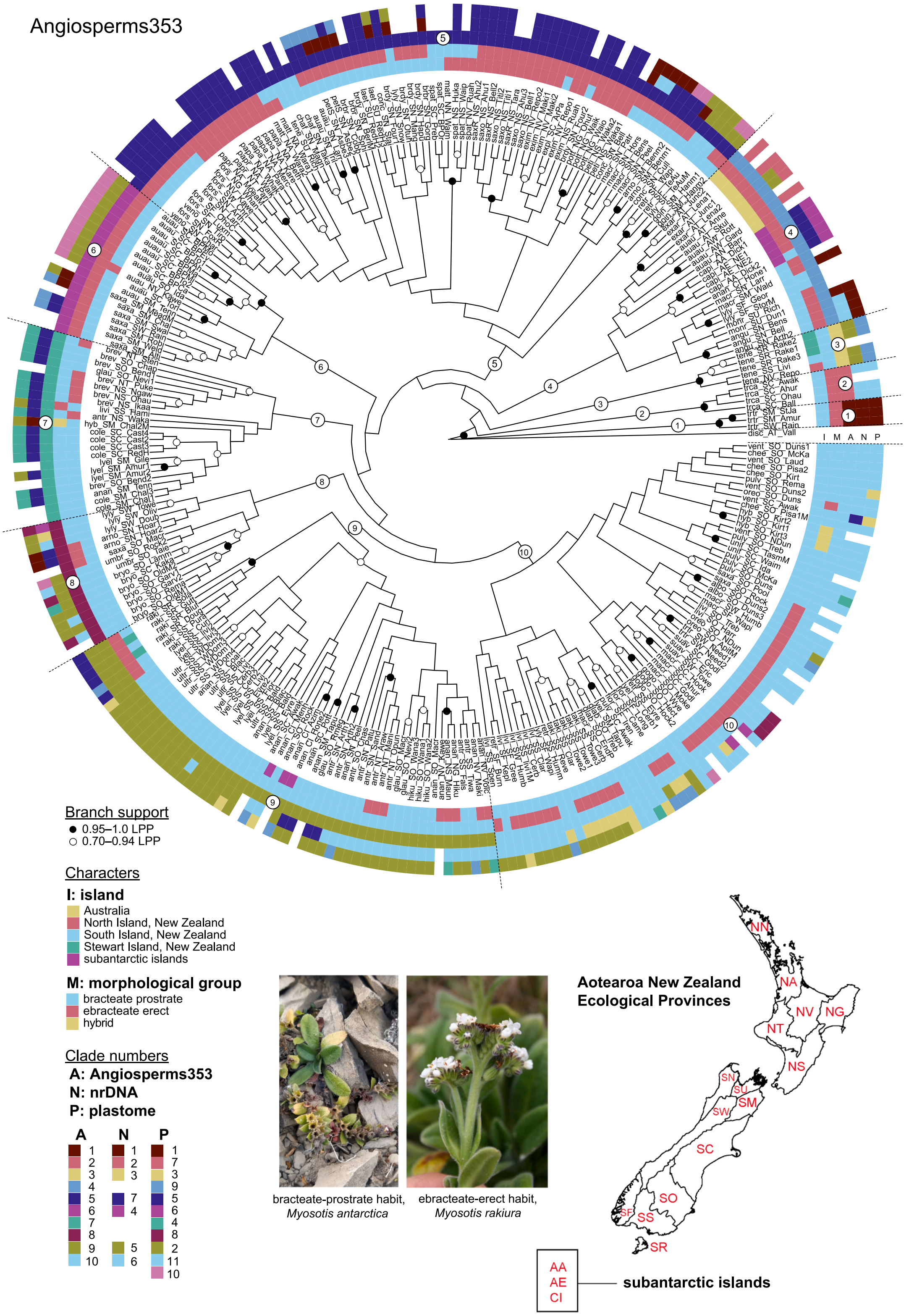
When comparing the LPP (local posterior probabilities) for specific clades in the 322-loci, 56-loci single-copy and 25-loci single-copy fully sampled species trees (for gene tree branch BS collapse values of 0 %, <10 %, <20 %, <30 %, <50 %,< 70 % and < 90 %), the 322-loci trees had generally higher LPP branch support than the other two datasets, and were more likely to find the same relationships irrespective of BS collapse value (data not shown). All trees showed some clustering of species based on ebracteate-erect and bracteate-prostrate inflorescence type and habit, although this was clearest in the 322-loci and 56-loci single-copy datasets compared to the 25-loci single-copy dataset (data not shown). For all three datasets, trees with the < 20 % BS collapse value had the highest LPP for the specific clades selected for comparison (data not shown). In general, the branches along the backbone of all ASTRAL-III trees were short and had low support (<0.5 LPP; e.g. Fig. 3, S2–S5).

Of all the ASTRAL-III fully sampled species trees from the three data subsets (322-loci, 56-loci single-copy and 25-loci single-copy; Fig. 3, S2–S5), the one based on the 322-loci dataset (with < 20 % BS collapse value) was the most resolved and highly supported overall, and made the most sense biologically (Fig. 3; Fig. S2). The branches for all four species or groups of individuals (see Appendix B, section 2.5.1.4) that were checked in all trees had higher LPP support in the 322-loci tree than the other two trees (Table 5; Fig. 3; Fig. S2). The two ASTRAL-III species trees from the 322-loci dataset analysed using mapping files (−a option) for 51 species or 57 species + subspecies showed similar species relationships to each other and low LPP values (ranging from 0.29 to 0.91 LPP; Fig. S3). In addition, branch lengths, which are given for all internal branches and terminal branches, were very short, ranging from 0.0 to 0.38 coalescence units (CU), but only 18 terminal branches were $\geq$ 0.1 CU (Fig. S3). Since branch lengths in these species trees are a direct measure of gene tree discordance, short branches indicate higher discordance among gene trees (Mirarab, 2019) and thus branch lengths are prone to underestimation (https://github.com/smirarab/ASTRAL/blob/master/astral-tutorial.md#branch-length-and-support).

Gene concordance factors (gcf) ran to completion in the 56- and 25-loci single-copy datasets, but not in the large 322-loci dataset with 291 individuals (despite being run for 30 days). In all ingroup nodes, in both of these single-copy trees, the proportion of underlying gene trees supporting each node was very low, with most pie charts largely pink (i.e., having a majority of uninformative loci), however the 25-loci single-copy dataset showed more concordance than the 56-loci single-copy dataset (Figs. S4, S5).

3.3. Angiosperms353 – Concatenated analyses

For all three Angiosperms353 datasets, the four specific clades selected for comparison (see Appendix B, section 2.5.1.4) were less often monophyletic in the concatenated trees relative to their corresponding ASTRAL-III trees (Table 5). Of the three trees, the 25-loci single-copy tree (the one with the fewest total loci and the fewest multi-copy loci) showed a modest amount of taxonomic structuring compared to the other two concatenated trees (Fig. S6). Nine taxa were monophyletic in this tree (compared to 22 in the 322-loci ASTRAL-III 353 tree), eight of which were also monophyletic in the ASTRAL-III 353 tree. Most branches including the backbone of the 322-loci IQ-TREE concatenated topology were highly supported with BS values 70–100 % (86.2 %), and the majority of these are 95 % or higher (67 %), whereas only 2.4 % of nodes have < 50 % BS (Fig. S6). Although branch support was generally much higher in all concatenated Angiosperms353 trees vs the ASTRAL-III trees, the outcomes for most taxa were less resolved and more



bracteate-prostrate habit, *Myosotis antarctica*

ebracteate-erect habit, *Myosotis rakiura*

Fig. 3. Fully sampled species tree of *Myosotis* Angiosperms353 322-loci and 291-individual dataset with 20 % gene tree bootstrap cutoff using ASTRAL-III. The ten main clades are numbered inside white circles. Local posterior probabilities (LPP) are shown above branches as filled circles (0.95–1.0 LPP), open circles (0.70–0.94 LPP), or no symbol (<0.70 PP). Coloured bars on the outside of the tree show island where that individual was collected (I), inflorescence and habit morphology (M), and clade number in this Angsioperms353 tree (A; Figs. 3, S2), the nrDNA tree (N; Figs. 4, S8) or the plastome tree (P; Figs. 5, S9). Left photo shows the bracteate-prostrate habit: *Myosotis antarctica* subsp. *antarctica*, WELT SP111328 (https://collections.tepapa.govt.nz/object/1955255) and the right photo shows the ebracteate-erect habit: *Myosotis rakiura,* WELT SP105593 (https://collections.tepapa.govt.nz/object/1591954), both photos: Museum of New Zealand Te Papa Tongarewa, CC BY 4.0 (https://creativecommons.org/licenses/by/4.0). A map of the Aotearoa New Zealand Ecological Provinces and their two-letter codes used in the sample names is also shown. See Table S1 for sample names, Fig. S2 for the ladderised Angiosperms353 tree with LPP values for all branches, and Appendix C for descriptions of each clade.

Table 5

Support values for the four clades compared among the three main southern hemisphere *Myosotis* phylogenomic datasets discussed in this paper: Angiosperms353, nuclear ribosomal DNA (nrDNA) and plastome. Support values for clades in the ASTRAL-III analyses are given as local posterior probabilities (LPP) whereas those for the IQ-Tree analyses are given as bootstrap (BS). Fractions given in brackets after a support value means the clade comprised a subset of the total number of individuals only. An asterisk in the *M. brevis* column means one individual that was identified a priori as *M. glauca* (glau_SO_Nevi) was also contained in the *M. brevis* clade. See section 4.3 and Appendix D for more details.

dataset	no. loci	no. individuals	analysis	*M. brevis*	*M. capitata*	*M. australis* (Banks Peninsula)	*M. exarrhena*
				$n = 8$	$n = 4$	$n = 8$	$n = 4$ ($n = 2$ in plastome & nrDNA)
Angiosperms353	322	291	ASTRAL-III	0.92 (7/8*)	0.97	1.0	1.0
Angiosperms353	56	291	ASTRAL-III	0.70 (7/8*)	0.63	0.85	0.49
Angiosperms353	25	291	ASTRAL-III	0.84 (4/8*)	–	1.0	0.96
Angiosperms353	322	291	IQ-Tree (concatenated)	100 % (5/8*)	100 % (3/4)	100 %	100 %
Angiosperms353	56	291	IQ-Tree (concatenated)	–	100 % (3/4)	–	100 % (3/4)
Angiosperms353	25	291	IQ-Tree (concatenated)	100 % (5/8*)	100 %	100 %	100 %
nrDNA	1	259	IQ-Tree	100 % (7/8*)	87 %	100 %	100 %
plastome	1	259	IQ-Tree	–	100 %	100 %	100 %

polyphyletic (i.e., individuals from one taxon found in a higher number of clades) than the ASTRAL-III Angiosperms353 tree, and nrDNA and plastome trees. For example, in the 25-loci single copy concatenated tree, 24/54 taxa (44 %) were less resolved and more polyphyletic than the two other trees, and 59 % were less resolved and more polyphyletic than the ASTRAL-III 353 tree. For only two taxa (4 %) was this tree slightly better resolved than the ASTRAL-III 353 tree – *M. petiolata* (both individuals monophyletic instead of just in the same main clade) and *M. antarctica* subsp. *traillii* (individuals polyphyletic in three clades instead of four). The majority of the underlying relationships that were strongly supported by the concatenated datasets were often very difficult to interpret due to the lower resolution, added polyphyly, and very few logical taxonomic, geographic or morphological patterns. Therefore, the 322-loci ASTRAL-III fully sampled species tree with $<$ 20 % BS collapse value was chosen as the optimal Angiosperms353 tree and will be discussed below.

3.4. Angiosperms353 phylogeny – Summary of the selected ASTRAL-III 322-loci species trees

Overall, the selected ASTRAL-III Angiosperms353 topology (291-individual and 322-loci ASTRAL-III fully sampled species tree with $<$ 20 % BS collapse value) is not well supported by LPP values, as nearly half (49.3 %) of all 288 nodes in the tree have $<$ 0.5 LPP, whereas 29.2 % have 0.7–1.0 LPP (and only 13.5 % have 0.95 LPP or higher) (Fig. 3; Fig. S2). The backbone of the tree in particular is poorly supported. The tree can be roughly divided into 10 main clades. Approximately 68 % of the sampled individuals (199/291) and about 65 % of the sampled taxa (35/54) are found in the three largest clades (Clades 5, 9 and 10). A summary of the 10 main clades of this tree, the taxonomic and geographic patterns within them, and patterns of monophyly and non-monophyly of the taxa can be found in Appendix C.

In general, there is a mixture of taxonomic and geographic structure in this tree, and there were some patterns regarding the main informal morphological groups, bracteate-prostrate vs. ebracteate-erect (Fig. 3; Fig. S2). Five of the clades comprise largely ebracteate-erect individuals and species (clades 1, 2, 4–6), four are bracteate-prostrate (clades 3, 7–9), and one (clade 10) is mixed. Clades 1–6 comprise 132 individuals, 82 % of which are ebracteate-erect, and the remaining bracteate-prostrate individuals in this clade are mostly of the creeping species *M. chaffeyorum, M. matthewsii, M. spatulata* and *M. tenericaulis*. Clades 7–9 include 91 individuals, of which 89 % are bracteate-prostrate, whereas clade 10 comprises 67 individuals, 52 % of which are bracteate-prostrate.

Regarding geographic distribution, most sampled ingroup *Myosotis* are from the South Island (217/290 individuals; ~75 %), which is the main centre of diversity of the genus in the southern hemisphere, and these are found throughout the tree. All nine Australian individuals of *M. exarrhena* and *M. australis* form a subclade (0.74 LPP, within Clade 4), as do all four subantarctic island individuals of *M. capitata* (0.97 LPP, Clade 4) plus one of *M. antarctica* (0.9 LPP); the remaining subantarctic island individuals of *M. antarctica* group together with other *M. antarctica* (clade 9) in two different clades. The majority of the 50 sampled North Island *Myosotis* individuals (and their corresponding North Island taxa) are in Clade 5, in three mostly North Island + northern South Island subclades: 1) All *M. pottsiana, M. petiolata, M. eximia* and *M. saxosa* (plus one *M. explanata*, expl_SC_ArPa, and one *M. brockiei* subsp. *dysis,* brdy_SN_Four2; 0.43 LPP); 2) all *M. spatulata* (North and South Islands; 1.0 LPP); 3) all *M. pansa* subsp. *praeceps, M. pansa* subsp. *pansa, M. matthewsii,* plus all *M. venosa* (South Island) and all *M. forsteri* (North and South Islands; 0.62 LPP).

Other North Island individuals of more widespread taxa are placed in other clades containing other individuals of their respective taxa (*M. antarctica, M. australis, M. tenericaulis* and *M. brevis*); the same is true of Stewart Island individuals of *M. rakiura* (Clade 9) and *M. tenericaulis* (Clade 3). The clustering of some individuals from certain regions within the North and South Islands is evident in many clades, and there are also several examples of individuals of different species but from the same or nearby locality clustering together (Appendix C). For example, some southern South Island individuals and taxa cluster together, including 30 individuals mostly from Otago, which include all *M. albosericea, M. oreophila, M. cheesemanii, M. venticola,* and *M. uniflora* (Subclade 10C, 0.39 LPP), which is in turn part of a larger southern South Island clade of 67 total individuals (Clade 10, 0.36 LPP); and Southland individuals of *M. ultramafica, M. lyallii* subsp. *elderi* (8 of 14)*,* and some *M. antarctica* subsp. *antarctica* (0.48 LPP) (Appendix C). Finally, taxa with fewer individuals sampled tended to be monophyletic, whereas taxa with higher sampling tended to be polyphyletic (Fig. S7), and species that are more widespread were generally found in a higher number of clades compared to localised species (data not shown).

In general, the two ASTRAL-III Angiosperms353 species trees (291-individual and 322-loci; Fig. S3A, Fig. S3B) are not well supported by LPP values and have very short branch lengths for both terminal and internal branches. The few relationships that have $>$ 0.5 LPP support are: *M. eximia* + *M. explanata* (0.91); *M. forsteri* + *M. venosa* (0.78–0.81) + *M. pansa* (0.59–0.62); *M. petiolata* + *M. pottsiana* (0.66); *M. matthewsii* + *M. spatulata* (0.65); *M. chaffeyorum* + *M.* "Asbestos" (0.62); and *M. colensoi* + *M. brevis* (0.58) + *M. colensoi* × *saxatilis* (0.54–0.56 LPP).

Some of the clustering of species in these species trees mirrors the clades in the fully sampled species tree, especially clades 4, 5, and 7–10 (Fig. S3). The two hybrids cluster with one of their respective parental species, and for species with subspecies (Fig. S3B), three have closely related subspecies (*M. antarctica, M. brockiei* and *M. pansa*), whereas the other three do not (*M. goyenii, M. lyallii* and *M. traversii*). The branch lengths largely correspond to whether a species is monophyletic in the fully sampled tree, i.e., species that are monophyletic have longer branch lengths (0.01 – 0.38 CU in the species + subspecies tree, Fig. S3B; mean: 0.14; median: 0.12), not monophyletic but are in the same clade (0.0–0.11; mean: 0.04; median: 0.03), or polyphyletic (0.0–0.09; mean: 0.03; median: 0.02) (branch lengths in the other species tree were similar, Fig. S3A).

3.5. Nuclear ribosomal DNA (nrDNA) dataset – Summary of the IQ-TREE phylogeny

The nrDNA dataset comprised 259 individuals totalling 6,829 bp (Tables 3 and 4; Fig. 4; Fig. S8). Overall, the resulting IQ-TREE nrDNA topology is not well resolved nor supported by BS values, as 19.5 % of all 256 nodes in the tree have zero branch lengths (and no corresponding BS value), and another 23 % have < 50 % BS, whereas 48.8 % have 70–100 % BS (with 34 % having 95 % BS or higher) (Fig. 4; Fig. S8). The backbone of the tree in particular is poorly supported. The nrDNA IQ-TREE topology can be roughly divided into seven main clades. Approximately 63 % of the sampled individuals (160/254) and 69 % of the sampled taxa (37/54) are found in the two largest clades (Clades 6 and 7). A summary of the seven main clades of the nrDNA tree, in order from smallest to largest (and also from bottom of the tree to top), the taxonomic and geographic patterns within them, and patterns of monophyly and non-monophyly of the taxa can be found in Appendix C.

In general, there is a mixture of taxonomic and geographic structure in the nrDNA tree, but there are no obvious patterns regarding morphological characters (Fig. 4; Fig. S8). Individuals from the main informal morphological groups, bracteate-prostrate vs. ebracteate-erect, sometimes group together but there are many exceptions. The bracteate-prostrate creeping species *M. chaffeyorum, M. matthewsii* and *M. spatulata* group with ebracteate-erect species – the latter two with other North Island species and *M. chaffeyorum* with other northern South Island species – whereas the remaining creeping species, *M. tenericaulis,* is in a clade of other mostly southern South Island species representing both habit types.

Regarding geographic distribution, most sampled ingroup *Myosotis* are from the South Island (201/254 individuals; 79 %), and these are found throughout the tree. Only three of the five Australian individuals of *M. exarrhena* and *M. australis* form a clade (Clade 2) as do all four subantarctic island individuals of *M. capitata* (87 % BS, Clade 7) plus one of *M. antarctica* ((97 % BS); of the other two subantarctic *M. antarctica,* one groups with other *M. antarctica* and the other is with *M. tenericaulis* with low support. The majority of the 40 sampled North Island *Myosotis* individuals (and their corresponding North Island taxa) are in three highly supported, mostly North Island subclades: 1) All *M. pansa* subsp. *pansa, M. saxosa, M. petiolata* and 3 of 4 *M. spatulata* (100 % BS); 2) All *M. matthewsii* and *M. pansa* subsp. *praeceps* (plus one *M. spatulata,* spat_NN_Mana; 100 % BS); 3) All *M. pottsiana, M. eximia* and (northern South Island) *M. brockiei* subsp. *dysis,* (plus one *M. explanata,* expl_SC_ArPa; 100 % BS).

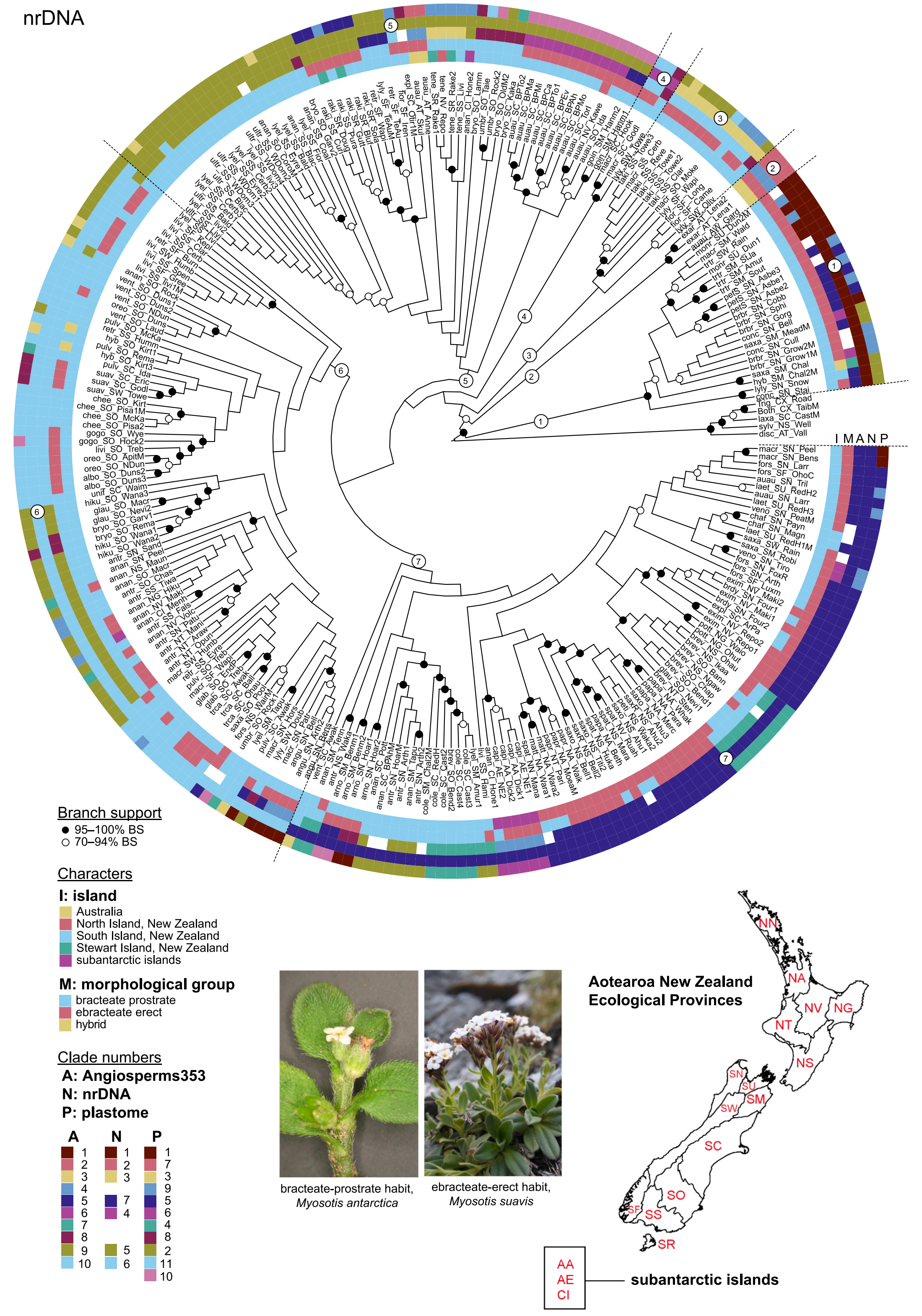
Other North Island individuals of more widespread taxa are placed in other clades containing other individuals of their respective taxa (*M. antarctica, M. australis, M. tenericaulis* and *M. brevis*); the same is true of Stewart Island individuals of *M. rakiura* (Clade 5, 100 % BS) and *M. tenericaulis* (Clade 5, 96 % BS). The clustering of some individuals from certain regions within the North and South Islands is evident in many clades, and there are also several examples of individuals of different species but from the same or nearby locality sharing very similar nrDNA sequences (Appendix C). For example, some South Island individuals and taxa cluster together with high support, including most *M. albosericea* and *M. oreophila* (except oreo_SO_Duns) from the Dunstan Mountains (100 % BS); a clade of 22 mostly ebracteate-erect taxa and individuals all from the northern South Island (Clade 1, 99 % BS); and Southland individuals of *M. ultramafica, M. lyallii* subsp. *elderi* (7 of 9), and some *M. antarctica* subsp. *antarctica* (90 % BS) (Appendix C). Finally, in general, higher intraspecific sampling led to increased polyphyly (Fig. S7).

3.6. Plastome dataset – Summary of the IQ-TREE phylogeny

The plastome dataset comprised 259 individuals totalling 149,043 bp (138,124 bp when the second copy of the inverted repeat was removed prior to the final analysis; Tables 3 and 4; Fig. 5; Figs. S9, S10). The vast majority of branches – from the larger subclades or lineages to the tips of the tree – are highly supported with BS values 80–100 % (91.5 %), and the majority of these are 95 % or higher (65.1 %), whereas only 3 % of nodes have < 50 % BS (Fig. 5; Fig. S9). The plastome IQ-TREE topology can be roughly divided into 11 main clades. Approximately 60 % of the sampled individuals (150/254) and 60 % of the sampled taxa (33/54) are found in the three largest clades (Clades 2, 5 and 11). A summary of the 11 main clades of the plastome tree, in order from largest to smallest, the taxonomic and geographic patterns within them, and patterns of monophyly and non-monophyly, can be seen in Appendix C.

Overall, there is a mixture of taxonomic and geographic structure in the plastome tree, but there are no obvious, strong overarching patterns regarding inflorescence type + habit (Fig. 5; Fig. S9). Specifically, regarding the main informal morphological groups of bracteate-prostrate vs. ebracteate-erect, although there are some clades dominated by individuals from one group or the other (Appendix C), there are multiple clades with individuals and taxa from both groups. The bracteate-prostrate creeping species *M. chaffeyorum, M. matthewsii* and *M. spatulata* group with ebracteate-erect species in Clade 5, whereas the remaining creeping species, *M. tenericaulis,* is in Clade 9 with both inflorescence + habit types.

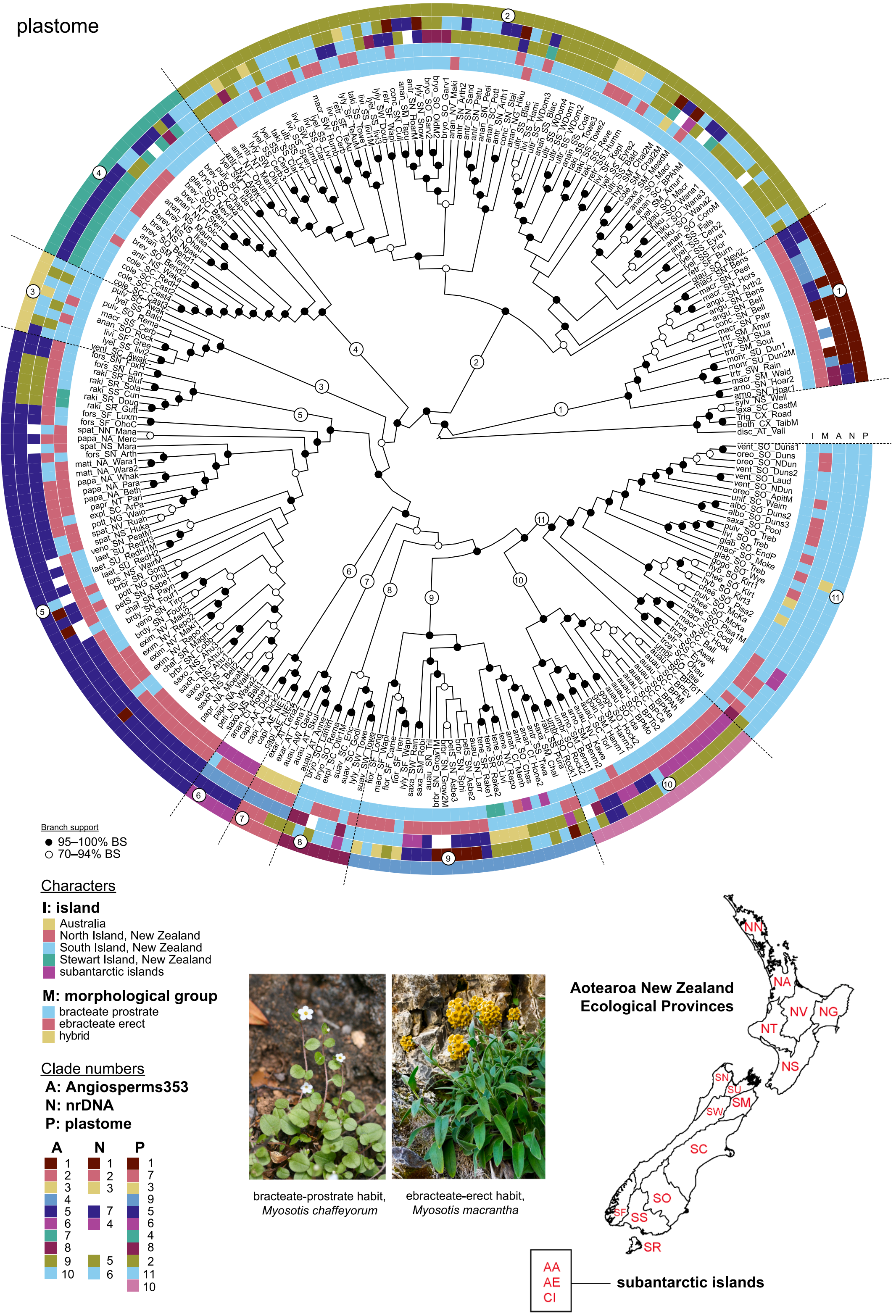
Regarding geographic distribution, most sampled ingroup *Myosotis* are from the South Island (201/254 individuals; about 80 %), and these are found throughout the tree. All five Australian individuals of *M. exarrhena* (100 % BS; Table 5) and *M. australis* (100 % BS) form a clade (Clade 7, 100 % BS) as do all four subantarctic island individuals of *M. capitata* (100 % BS, Clade 6; Table 5) plus one of *M. antarctica* (100 % BS); the other two subantarctic individuals of *M. antarctica* group with high support with other *M. antarctica* (Clade 9, 100 % BS). The majority of the 40 sampled North Island *Myosotis* individuals (and their corresponding North Island taxa) are in Clade 5, in two main subclades: 1) all *M. saxosa, M. petiolata, M. eximia* and 2/3 *M. pansa* subsp. *praeceps* (plus northern South Island *M. venosa, M. brockiei, M. chaffeyorum, M. laeta, M.* "Asbestos", and one *M. forsteri* (100 %); 2) all *M. pottsiana* and *M. matthewsii,* 3/4 *M. pansa* subsp. *pansa* (and one *M. pansa* subsp. *praeceps*), and 3/4 *M. spatulata* (plus one *M. explanata*, expl_SC_ArPa, and one *M. forsteri*; 100 %). Other North Island individuals of more widespread taxa are placed in other clades containing other individuals of their respective taxa (*M. antarctica, M. australis, M. tenericaulis* and *M. brevis*); the same is true of Stewart Island individuals of *M. rakiura* (Clade 5) and *M. tenericaulis* (Clade 9). The clustering of some individuals from certain regions within the North and South Islands is evident in many clades, and there are also several examples of individuals of different species but from the same or nearby locality sharing very similar chloroplast haplotypes (Appendix C). For example, some southern South Island individuals and taxa cluster together, including 24 individuals from Otago including all *M. albosericea, M. oreophila, M. cheesemanii, M. venticola,* and *M. uniflora* (Subclade 11B, 100 BS); and a clade of 17 ebracteate-erect taxa and individuals all from the northern South Island (Clade 1, 100 % BS) (Appendix C). Finally, as with the Angiosperms353 and nrDNA datasets, polyphyly was more



bracteate-prostrate habit, *Myosotis antarctica*

ebracteate-erect habit, *Myosotis suavis*

Fig. 4. Phylogenetic tree of the *Myosotis* nrDNA 259-individual dataset using IQ-TREE. The seven main clades are numbered inside white circles. Bootstrap support (BS) values are shown above branches as filled circles (95–100 % BS), open circles (70–94 % BS), or no symbol (<70 % BS). Coloured bars on the outside of the tree show island where that individual was collected (I), inflorescence and habit morphology (M), and clade number in the Angsioperms353 tree (A; Figs. 3, S2), this nrDNA tree (N; Figs. 4, S8) or the plastome tree (P; Figs. 5, S9). Left photo shows the bracteate-prostrate habit: *Myosotis antarctica* subsp. *traillii*, WELT SP090544 (https://collections.tepapa.govt.nz/object/1254156) and the right photo shows the ebracteate-erect habit: *Myosotis suavis,* WELT SP107452/A (https://collections.tepapa.govt.nz/object/1768787); both photos: Museum of New Zealand Te Papa Tongarewa, CC BY 4.0 (https://creativecommons.org/licenses/by/4.0). A map of the Aotearoa New Zealand Ecological Provinces and their two-letter codes used in the sample names is also shown. See Table S1 for sample names, Fig. S8 for the ladderised nrDNA tree with LPP values for all branches, and Appendix C for descriptions of each clade.



bracteate-prostrate habit, *Myosotis chaffeyorum*

ebracteate-erect habit, *Myosotis macrantha*

Fig. 5. Phylogenetic tree of the *Myosotis* plastome 259-individual dataset using IQ-TREE. The eleven main clades are numbered inside white circles. Bootstrap support (BS) values are shown above branches as filled circles (95–100 % BS), open circles (70–94 % BS), or no symbol (<70 % BS). Coloured bars on the outside of the tree show island where that individual was collected (I), inflorescence and habit morphology (M), and clade number in the Angsioperms353 tree (A; Figs. 3, S2), the nrDNA tree (N; Figs. 4, S8) or this plastome tree (P; Figs. 5, S9). Left lower photo shows the bracteate-prostrate habit: *Myosotis chaffeyorum*, WELT SP105614 and the right photo shows the ebracteate-erect habit: *Myosotis macrantha,* WELT SP108821 (https://collections.tepapa.govt.nz/object/1863566); both photos: Museum of New Zealand Te Papa Tongarewa, CC BY 4.0 (https://creativecommons.org/licenses/by/4.0). A map of the Aotearoa New Zealand Ecological Provinces and their two-letter codes used in the sample names is also shown. See Table S1 for sample names, Fig. S9 for the ladderised plastome tree with BS values for all branches, and Appendix C for descriptions of each clade.

prevalent in taxa with greater infraspecific sampling (Fig. S7).

3.7. All three phylogenies – Areas of topological congruence and patterns of monophyly and non-monophyly of southern hemisphere *Myosotis* taxa

There are several areas of broad topological congruence in the three trees (353, nrDNA, plastome) that can be observed by comparing the coloured bars on the outside of the trees (Figs. 3–5, S2, S8 and S9), the trees themselves (Figs. 3–5, S2, S8 and S9), and the detailed clade descriptions (Appendix C). 1) Clade 5 (353) comprises mostly North Island and northern South Island ebracteate-erect species and is similar in composition to Clade 7 (nrDNA) and Clade 5 (plastome). 2) Clade 10 (353) comprises southern South Island ebracteate-erect and bracteate-prostrate species which is similar to Clade 6 (nrDNA) and Clade 11 (plastome). 3) Clade 7 (353) contains the bracteate-prostrate species *M. brevis* and *M. colensoi* which is similar to Clade 4 (plastome). 4) Clade 9 (353) comprises mostly southern South Island bracteate-prostrate species which is similar to Clade 2 (plastome). Similarly, there are other examples of smaller groups of taxa that consistently group together in the three trees, such as *M. antarctica, M. lyallii* subsp. *elderi* and *M. ultramafica* in Subclades 9B and 9C (353), which are similar in composition to Subclade 5B1 (nrDNA) and Subclade 2A (plastome). Another example is the Australian *M. exarrhena* and the Australian individuals of *M. australis* grouping together in Clade 4 (353), Clade 2 (nrDNA) and Clade 7 (plastome).

The 54 taxa were classified as either being monophyletic (28–41 % of taxa), having individuals in the same main clade or lineage (30–35 %), or polyphyletic (20–40 %), in each of the three trees (353, nrDNA, plastome). Thus, about half of the taxa (28/54, 52 %) were monophyletic in at least one of the three trees (353, nrDNA, plastome). Similarly, 26 taxa (48 %) were either monophyletic or found in the same main clade (i.e., never polyphyletic) in all of the trees. Of these 26 taxa, five of them (9 %) were always monophyletic: *M. albosericea, M. capitata, M. exarrhena, M. goyenii* subsp. *infima,* and *M. uniflora.* Another 16 taxa (30 %) were monophyletic in 1 or 2 trees and in the same main clade in the remaining tree(s) (*M. angustata, M. chaffeyorum, M. cheesemanii, M. eximia, M. glabrescens, M. hikuwai, M. laeta, M. matthewsii, M. monroi, M. pansa* subsp. *pansa, M. pansa* subsp. *praeceps, M. pottsiana, M. saxosa, M. suavis, M.* sp. “Takitimu”, and *M. traversii* subsp. *cantabrica*). Finally, the remaining five taxa (9 %) were always in the same main clade (i.e., never monophyletic or polyphyletic; *M. brockiei* subsp. *dysis, M. petiolata, M. spatulata, M. ultramafica* and *M. venosa*). By contrast, 16 taxa (30 %) showed a mixture of monophyly, being in the same clade, and polyphyly (*M. arnoldii, M.* sp. “Asbestos”, *M. brevis, M. brockiei* subsp. *brockiei, M. bryonoma, M. colensoi, M. concinna, M.* sp. “Fiordland”, *M. forsteri, M. goyenii* subsp. *goyenii, M. oreophila, M. rakiura, M. tenericaulis, M. traversii* subsp. *traversii, M. umbrosa, M. venticola*). For these 16 taxa, their polyphyly involved individuals being placed in only two (rarely three) different clades. Finally, 12 taxa (22 %) were always polyphyletic in all three trees (*M. antarctica* subsp. *antarctica, M. antarctica* subsp. *traillii, M. australis, M. explanata, M. glauca, M. lyallii* subps. *elderi, M. lyallii* subsp. *lyallii, M. macrantha, M. pulvinaris, M.* sp. “Livingstone”, *M. saxatilis* and *M. retrorsa*), and the polyphyly of these 12 taxa generally involved individuals being placed in 3 to 5 different clades (rarely only 2 or as many as 6; Fig. S7).

To address several detailed taxonomic questions for each species, descriptions of taxonomic, geographic and morphological patterns from the same trees (Figs. 3–5, S2, S8 and S9) are summarised in Appendix D.

4. Discussion

The current study builds on previous phylogenetic studies (Meudt et al., 2015, 2013; Prebble et al., 2019, 2015; Winkworth et al., 2002b) and represents the first phylogenomic analysis of the southern hemisphere *Myosotis* (Boraginaceae) lineage, and the first published whole chloroplast genomes for this genus. It also represents the first genus-level study in Boraginaceae to analyse Angiosperms353 data. Using a combination of data from targeted sequence capture (574,445 bp), whole chloroplast genomes (138,124 bp), and the nuclear ribosomal DNA tandem repeat cluster (6,829 bp), a total of 719,389 bp were analysed as three separate datasets with different concatenated and coalescent methods for 291 or 322 individuals of 54 *Myosotis* ingroup taxa from New Zealand and Australia. The resulting data-rich phylogenies represent ~ 120 × more base pairs (two orders of magnitude), ~80 × more loci, and ~ 2–3 × more taxon sampling than the most recent phylogenetic analysis based on Sanger sequencing of standard DNA markers (Meudt et al., 2015). This study is also one of only a handful of phylogenetic studies of native New Zealand lineages published to date to use either Angiosperms353 baits for target-capture phylogenomics (Nicol et al., 2024; Ning, 2023; Schmidt-Lebuhn et al., 2022; Thomas et al., 2023, 2021) or genome skimming for sequencing whole chloroplast genomes (Maurin et al., 2022), and the first to our knowledge to use both. Another novel aspect to this phylogenomic study is the extensive inter- and intraspecific taxon sampling, which includes near-comprehensive species sampling from southern hemisphere *Myosotis* (54/59 taxa; 92 %; Table 1), and up to 19 individuals per taxon (median: 4; mean: 5), most of which were field-collected throughout their geographical ranges by HMM and JMP over 17 years.

From field work to final figures, there are significant challenges to collect, assemble, curate, filter and analyse molecular datasets with hundreds of loci and individuals for island and mountain species radiations (Chan et al., 2000; Lagomarsino et al., 2022; Thomas et al., 2021). Overall, the phylogenomic data has not provided a fully resolved, highly supported species tree, which is especially true of the deeper phylogenetic nodes, but it does highlight both geographic and taxonomic patterns throughout the southern hemisphere *Myosotis* lineage at multiple levels, and sheds light on some evolutionary processes and species relationships. Here, we interpret the utility of different target capture and genome skimming phylogenomic datasets and analyses to resolve rapid island and mountain radiations, the phylogenetic relationships that are inferred (or unable to be inferred), and the taxonomic implications for southern hemisphere *Myosotis* at multiple levels.

4.1. Getting the most out of target capture Angiosperms353 and genome skimming datasets for species radiations

Angiosperms353 data. In general, the Angiosperms353 data were of similarly high quality and quantity for most sequenced *Myosotis* individuals (Fig. S1). However, the final Angiosperms353 ASTRAL-III fully sampled species tree was not well supported by LPP values (only 29.2 % of the branches had 0.7–1.0 LPP; Fig. 3; Fig. S2), and the backbone of the tree was unresolved; the ASTRAL-III species trees had similar issues but with much lower LPP values and short branch lengths (Fig. S3). Thus, this dataset did not fully resolve species relationships in the southern hemisphere *Myosotis* lineage, but taken together, the sheer number of loci, characters and individuals analysed, the high percentage of parsimony informative characters (21 %), and the use of data filtering, curation, and coalescent methods, all contributed to increasing the effectiveness of the Angiosperms353 data for phylogenetic inference of a group with a combination of characteristics that can be challenging, i.e., a very shallow, recent, polyploid, island and mountain species radiation.

The data here suggest the entire southern hemisphere *Myosotis* lineage has a recent allopolyploid origin in which all the species in the radiation have the same two (or more) differing ancestral subgenomes. Supporting this hypothesis are the high (and similar) average values for several HybPhaser statistics (locus heterozygosity, SNP percentage, and percentage of loci with > 1 % SNP; Table S3, Fig. S1D), which would be expected of hybrids or allopolyploids (Nauheimer et al., 2021). Indeed, molecular dating has shown that New Zealand *Myosotis* is a very recent Pleistocene radiation, with a stem age (colonisation age) of 2.8–1.0 mya and a crown age (radiation age) of 1.8–0.7 mya (Meudt et al., 2015). Many *Myosotis* species worldwide are polyploid, with chromosome numbers ranging from $2n = 14$ to $2n = 88$, and there is also dysploidy, with base numbers ranging from $x = 6$ to $x = 12$ (Grau, 1968, 1964; Luque, 1992). Consistent with an allopolyploid origin hypothesis, the few New Zealand species that have had their chromosomes counted so far are all polyploids, and they also show a dysploid series of $2n = 36$, 40, 44, 46 or 48 chromosomes (de Lange and Murray, 2002; Murray and de Lange, 2013). The yellow-flowered Iberian annual *Myosotis persoonii* Rouy & E.G.Camus ($2n = 48$; (Grau, 1968; Luque, 1992) is the closest northern hemisphere relative to the southern hemisphere lineage according to a published ITS phylogeny (Winkworth et al., 2002b), whereas the chloroplast phylogeny in the same paper was unresolved in this regard. In both of those published phylogenies (Winkworth et al., 2002b), the nine other mainly European species also in the same clade as *M. persoonii* and the southern hemisphere species – including *M. discolor*, which was sampled here – have chromosome numbers (where known) of $2n = 24$, 48 or 72 (Grau, 1968; Luque, 1992). A complementary phylogenomic study using similar markers to those used here for northern hemisphere *Myosotis* – as well as additional chromosome counts and estimation of genome size, particularly for New Zealand species (Meudt et al., 2021) – would be highly desirable to clarify the closest extant relatives and polyploid origins of the southern hemisphere lineage and fully understand and compare patterns of evolution in the genus as a whole.

The statistics from HybPiper (especially ‘GenesWithStitchedContigs’ and ‘ParalogWarningsDepth’; Table S2), as well as the very high percentage of Angiosperms353 loci that were identified by eye as being artifactual orthologs (92 %), are indicative of a high number of paralogs in this dataset, which would be expected for polyploid lineages. Compared to the 90 of 313 loci (29 %) that were identified by eye as artifactual orthologs in a published Angiosperms353 dataset for the neotropical genus *Freziera* (Pentaphylacaceae), the 92 % of loci identified here is much higher, even though *Freziera* is similar to *Myosotis* in that it is also a likely allopolyploid genus that has undergone a rapid radiation (but in the South American Andes) (Frost et al., 2024). Identifying and removing such artifactual orthologs or paralogs is one of three strategies currently employed to analyse such datasets. Although in the case of *Myosotis* this strategy did not produce more resolved, better supported or more informative trees, taking the time to visually inspect every Angiosperms353 locus alignment and gene tree, and comparing the results of the various downstream phylogenetic analyses, greatly improved our understanding of the underlying data and also shaped our final filtering and analysis strategies.

Additional strategies for dealing with paralogs not explored here include those that separate and include paralogs in downstream analyses in different ways (Gardner et al., 2020; Kandziora et al., 2021; Morales-Briones et al., 2022; Ufimov et al., 2022). So far, these pipelines have focused on exons only (not supercontigs), to the exclusion of introns and non-target “splash zones”, such that many of the variable sites crucial for a study focused on rapid radiations at the level of species would be lost (Johnson et al., 2019). Of the three strategies we used here: (1) coalescent species tree reconstruction from all loci, (2) coalescent species tree reconstruction from single-copy loci, and (3) phylogenetic analysis of concatenated loci (all loci or single-copy loci only), the first one produced the ‘best’ species tree for *Myosotis,* but there was nevertheless congruent signal to varying degrees in the 25-loci single copy coalescent and concatenated analyses. Although concatenation can be useful for shallow phylogenies where combining all information from all available loci may increase phylogenetic signal (Chen et al., 2023), as a method that is not “ILS-aware”, it may also lead to erroneously reconstructed relationships for rapid radiations (Thureborn et al., 2022) especially those that are also mainly polyploid (e.g., Nicol et al., 2024; Overson et al., 2023). This may have occurred in the current study, in which many of the branches of the concatenated trees were highly supported but were highly inconsistent with other morphological or geographic data, and with previous taxonomic or other biological knowledge, and were likely erroneous. By contrast, the coalescent method can take into account incomplete lineage sorting, but it can also be sensitive to gene tree error (Thureborn et al., 2022; Zhang et al., 2018; Howard et al., 2022), which could be high in our dataset given the potentially large number of artifactual orthologs found, and the high levels of gene tree incongruence seen in the gcf analyses of data subsets that were able to run to completion (Figs. S4, S5). The analyses here show that, in spite of incongruence and paralogy, there is merit to using as many of the Angiosperms353 loci and as much of the underlying data from each locus as possible (Maddison and Knowles, 2006; Thomas et al., 2021; Yan et al., 2022). A future strategy that can simultaneously identify, separate and incorporate paralogous sequences prior to coalescent analysis could improve the utility of such Angiosperms353 datasets for resolving the relationships, species limits and taxonomy, particularly of recent radiations with a history of polyploidy.

Testing different levels of gene tree bootstrap cutoffs was another useful tool to further curate and filter the Angiosperms353 data. The finding that the 20 % BS collapse value gave trees with the highest LPP for the chosen clades in all three 353 datasets examined is in line with Zhang et al. (2018). That study showed that removing branches with < 10 % BS (and sometimes up to 20 % BS) improves species tree accuracy, whereas more aggressive filtering reduces the ability to reconstruct the true species tree. As a final note regarding sequencing platforms, those such as MiSeq − which produce fewer mapped reads and genes with assembled contigs per individual compared to HiSeq − may not be an appropriate platform for sequencing Angiosperms353 baits in studies similar to this one. The majority of MiSeq-individuals (60 %) did not pass the average recovered supercontig length filter used here (1,500 bp). Relative to the genome skimming datasets, there was a lower proportion of MiSeq-sequenced individuals that were included in the final Angiosperms353 dataset based on the filtering scheme we used. This suggests that the DNA for these samples was of high quality and quantity (because these samples were able to be included in the genome skimming datasets). Finally, comparing the ASTRAL-III fully sampled species tree (Fig. 3, Fig. S2) with the species trees using two different mapping files (Fig. S3), and interpreting support values and branch lengths was another useful tool for disentangling the complexity of species relationships and species and subspecies limits (Rabiee et al. 2019). Although the low LPP values and very short branches corresponding to high gene tree discordance and non-monophyly in the species trees were perhaps expected for a species radiation (e.g., Chan et al., 2020), the distribution of branch lengths and the topology of the species trees added another layer of taxonomic interpretation to the Angiosperms353 dataset (Fig. S3; Appendix C, D).

Genome skimming datasets. In general, the two genome skimming datasets – whole plastome and nuclear ribosomal DNA tandem repeat cluster – comprised high quality sequences for the subset of 259 taxa (80 %) included in the Angiosperms353 target capture dataset whose genomic libraries contained sufficient quality and quantity DNA for genome skimming. Using the same genomic libraries for genome skimming was a cost-effective way of getting complete plastome and nrDNA datasets, as opposed to attempting extraction of partial plastome DNA and nrDNA as off-target reads from the Angiosperms353 target capture (where only ∼ 20 % of the plastome is usually recovered; Baker et al., 2021). Once the plastome and nrDNA sequences were assembled and aligned for each *Myosotis* individual, the downstream analyses were standard and straightforward to implement, although we used high performance computers due to the large size of the datasets. The well-resolved plastome tree had only 2 % parsimony informative characters overall, but most branches were highly supported, and this tree was thus a crucial additional, independent and uniparental marker for southern hemisphere *Myosotis,* contributing taxonomic and geographic patterns for comparison to the Angiosperms353 dataset. Although the nrDNA tree was, as expected based on previous ITS and ETS phylogenies (Meudt et al., 2015; Winkworth et al., 2002b), not well supported despite having 8.2 % parsimony informative characters (including outgroups), with concerted evolution well documented especially in polyploid species

(Small et al., 2004), it was nevertheless assembled and included as a valuable third dataset for comparison and corroboration.

4.2. Phylogenetic relationships in southern hemisphere Myosotis

The phylogenomic trees presented here contribute variously toward reconstructing the relationships of New Zealand and Australian *Myosotis.* In addition to the logistical and bioinformatic challenges encountered, there are also challenges around interpretation of the resulting large, intricate and multiple trees. These can be difficult and complex to understand, summarise and compare, particularly for a recent, rapid island (and mountain) species radiation (Kandziora et al., 2021; Lagomarsino et al., 2022; Larridon et al., 2019; Murillo-A et al., 2022; Ottenlips et al., 2021; Thomas et al., 2021) combined with a history of polyploidy and dysploidy (Murray and de Lange, 2013) with high morphological and ecological variation yet low genetic variation (Meudt et al., 2015). Extensive biological knowledge of the taxa and individuals under study is an essential part of this process (Frost et al., 2024), without which any evolutionary or taxonomic interpretations would be largely impossible. The phylogenomic trees presented here are not fully resolved, supported, congruent, or completely consistent with current taxonomy, which is not unexpected compared with other recent phylogenomic studies of other similar groups (Lagomarsino et al., 2022; Ottenlips et al., 2021; Thomas et al., 2021; Yan et al., 2024). Nevertheless, taxonomic and geographic patterns are evident especially when comparing trees from different datasets, and combining the phylogenomic data with additional, independent (non-molecular) biological and morphological data and knowledge of species. Below we compare the trees from the three main datasets (353, plastome and nrDNA) to outline the main species relationships and groups, evolution of habit and inflorescence type, and biogeographical patterns for southern hemisphere *Myosotis.*

Most species of genera that are closely related to *Myosotis* have ebracteate inflorescences, as do northern hemisphere species of *Myosotis* (Cohen, 2014). The presence or absence of bracts on the inflorescence is evolutionarily labile in this lineage of Boraginaceae, traditionally called tribe Cynoglosseae, which is also the only tribe to have some species with bracts present on the base of the inflorescence only (i.e., not with the flowers; Cohen, 2014). (Names and circumscription of subfamilies of tribes within the family has changed since then, see Chacón et al., 2016, Luebert et al. 2016). All southern hemisphere *Myosotis* have bracts on the inflorescences, including the ebracteate-erect species which have bracts at the base of the erect inflorescence only, whereas bracteate-prostrate species also have bracts subtending or near the flowers of the mostly prostrate inflorescences or branches (Meudt et al., 2015; Robertson, 1989). In the Angiosperms353 tree (Fig. 3; Fig. S2), there are five clades of mostly ebracteate-erect species (Clades 1, 2, 4–6), three clades of mostly bracteate-prostrate species (Clades 3, 7–9), and one clade that is a mixture of both inflorescence and habit types (Clade 10). Shifts from ebracteate to bracteate inflorescences (and back again) in *Myosotis* parallels the same trends seen in the tribe as a whole, where ebracteate inflorescences are ancestral and several switches to bracteate inflorescences have occurred (Cohen, 2014). One example of this in the current study is the three bracteate-prostrate New Zealand species, namely the few-flowered, creeping species *M. spatulata, M. chaffeyorum* and *M. matthewsii,* which consistently group with ebracteate-erect species in all three phylogenies (especially with *M. forsteri, M. venosa, M. laeta,* and both subspecies of both *M. pansa* and *M. brockiei*; Figs. 3–5, S2, S8 and S9). Thus, bracteate inflorescences have evolved independently in this group of closely related species compared to the other bracteate-prostrate species. Even though the inflorescence + habit character cannot be used to classify southern hemisphere *Myosotis* into two monophyletic groups*,* it nevertheless remains an extremely useful character for field identification at the species level, and for understanding the evolution of inflorescences in the genus. This is also the case for at least one other genus in the family, *Euploca* Nutt. (formerly *Heliotropium* L. section *Orthostachys* R.Br.; Boraginaceae *sensu* APG IV or Heliotropiaceae), where detailed morphological and developmental studies aided understanding of inflorescence evolution and refinement of infrageneric taxonomy since subsections based on bract presence are not monophyletic (Frohlich, 1978; Frohlich et al., 2022). Similar developmental studies could prove fruitful regarding the evolution of southern hemisphere *Myosotis* inflorescences, especially regarding the bracteate-prostrate creeping species mentioned above.

With respect to biogeographic patterns in the trees, the South Island of New Zealand is the centre of origin and diversity for southern hemisphere *Myosotis*, which is also the case for numerous other New Zealand genera such as *Ourisia* (Meudt and Simpson, 2006), *Ranunculus* (Lockhart et al., 2001), and *Veronica* (Thomas et al., 2023)*,* among others. For *Myosotis,* 80 % of the southern hemisphere species are native to the South Island (mostly in mountainous areas of the Southern Alps) and 82 % of these are endemic there (Table 1); similarly, 75 % of the samples in our study were collected on the South Island (Table S1). Consistent with previous findings (Winkworth et al., 2002a), all trees show *Myosotis* has dispersed multiple times from New Zealand, and they go a step further to show South Island as the origin of these dispersals to the following places: Australia, the New Zealand subantarctic islands, North Island, and Stewart Island.

Because the Australian endemic *M. exarrhena* always groups with all or most Australian individuals of *M. australis* in all trees*,* these two species are likely derived from a single dispersal event to Australia by a shared ancestor. *Myosotis exarrhena* may be derived from within Australian *M. australis* (353; Fig. 3; Fig. S2), which also have similar but separate plastome and nrDNA relationships (i.e., a sister relationship; Figs. 4, 5), but these patterns could also be the result of interspecific hybridisation, for which there is some morphological evidence (Meudt et al., 2020) (Appendix D). A detailed molecular study encompassing a more extensive sampling of these two species from throughout their ranges in Australia, and from *M. australis* subsp. *saruwagedica* from New Guinea, which was not included here, is required to fully understand the relationship of these two species.

By contrast, all trees show there were at least two separate dispersals to the New Zealand subantarctic islands, with separate arrivals by ebracteate-erect subantarctic endemic species *M. capitata* (once to Auckland Islands + Campbell Island, including the individual anan_-CI_Hone1; see section 4.3) and bracteate-prostrate *M. antarctica* subsp. *antarctica* (at least once to Campbell Island). The latter species is widespread in the North and South Islands and has even reached southern Chile (although no Chilean material was sampled here; (Prebble et al., 2022), so multiple dispersals of *M. antarctica* subsp. *antarctica* to Campbell Island, as potentially suggested by the nrDNA tree, might be plausible. Similarly, three species have dispersed from the South Island to nearby Stewart Island (*M. tenericaulis, M. antarctica* subsp. *traillii,* and *M. rakiura*), and Stewart Island samples for the two species sampled here (*M. tenericaulis* and *M. rakiura*) always cluster with their respective South Island conspecifics in all three trees (Figs. 3–5, S2, S8 and S9).

Even though it is not the centre of diversity for the genus, the North Island of New Zealand boasts eight endemic *Myosotis* taxa, seven of which were sampled here, as well as eight additional species (seven sampled from the North Island) that are native there (Table 1). There is evidence that the endemic North Island species are closely related to one another and to other species native to the North Island and are derived from southern North Island ancestors. Although they do not form a monophyletic group, the North Island endemic species are all found within the same clade in all three trees (*M. eximia, M. matthewsii, M. pansa* subsp. *pansa, M. pansa* subsp. *praeceps, M. petiolata, M. pottsiana* and *M. saxosa*; Figs. 3–5, S2, S3, S8 and S9). These endemic species are in turn often closely related to other species native to the North Island, such as *M. spatulata, M. forsteri* and *M. venosa,* as well as other species endemic to the northern South Island, including (in at least two of the three trees; Figs. 3–5, S2, S3, S8 and S9). *M. laeta, M. brockiei, M.*

"Asbestos", *M. chaffeyorum,* and the two individuals of the more widespread *M. australis* sampled from that area.

Of the North Island endemics, there are some general trends (with exceptions) that can be seen in all three trees, such as the close relationship of *M. matthewsii, M. pansa* subsp. *pansa* and *M. pansa* subsp. *praeceps* (together with *M. spatulata*), and the close relationship of *M. eximia, M. petiolata* and *M. saxosa,* whereas the relationships of *M. pottsiana* varied between groups, depending on the tree (Figs. 3–5, S2, S8 and S9). The first group of species are allopatric on the northern and eastern parts of the North Island (Meudt et al., 2013; Meudt and Prebble, 2018), whereas the second group (plus *M. pottsiana*) are allopatric in the northeastern part of the island, except for the range of *M. petiolata* which overlaps with that of *M. saxosa.* This is additional evidence that the species in the informal "*M. petiolata* group" (*M. petiolata, M. pansa, M. pottsiana,* and South Island M. "Asbestos") are distinct species rather than subspecies, and that group itself may not be monophyletic (Meudt et al., 2013) but is instead part of a broader North Island endemic lineage. Finally, as with South Island taxa that have dispersed to Stewart Island, those that have dispersed to the North Island also tend to cluster with their respective conspecifics in most trees (e.g., *M. australis, M. tenericaulis, M. brevis, M. antarctica* subsp. *antarctica, M. antarctica* subsp. *traillii*) (Figs. 3–5, S2, S8 and S9).

The mountains of the southern South Island, which in addition to those in northwestern part of the South Island, are centres of diversity for New Zealand *Myosotis* (Meudt and Prebble, 2018), and, more generally, are known hot spots of taxonomic and morphological diversity and endemism in the New Zealand flora (Cockayne, 1928; Wardle, 1963). Multiple species from the southern South Island, especially high-elevation habitats in Otago, are to varying degrees intermixed with one another in the same clade in the trees presented here, and these phylogenomic patterns mirror the morphological and taxonomic complexity and diversity in the genus in this area (Meudt and Prebble, 2018). A good example of this is the bracteate-prostrate species *M. antarctica* subsp. *antarctica, M. antarctica* subsp. *traillii, M. glauca, M. hikuwai, M. lyallii* subsp. *elderi* and *M. ultramafica*, which are closely related to one another in several trees (including the 353 ASTRAL-III species tree, Fig. S3). Some of these species belong to the so-called pygmy species group (*M. antarctica* subsp. *antarctica, M. antarctica* subsp. *traillii, M. brevis, M. glauca* and *M. hikuwai*; Prebble et al., 2022, 2019). A previous study of the pygmy group using microsatellites focused on population genetics and species delimitation (Prebble et al., 2019, 2015), but species relationships were not fully resolved and monophyly of the group was not addressed. In the phylogenomic results reported here, as was the case with the *M. petiolata* group above, not all members of the pygmy group are each other's closest relatives, suggesting again that this informal species group may not be monophyletic and is instead part of a broader bracteate-prostrate lineage.

4.3. Taxonomic implications at multiple levels for New Zealand and Australian Myosotis

The phylogenomic trees contribute to the main aim of the current study by providing numerous insights that can be applied to improve the taxonomy of southern hemisphere *Myosotis,* which is currently undergoing taxonomic revision (Meudt, 2021; Meudt et al., 2020, 2013; Meudt and Prebble, 2022, 2018; Prebble et al., 2022, 2019). There were a range of outcomes at the different scales of species, subspecies and even individual samples and hybrids, for their taxonomy, monophyly, species limits, geography, and correct identification. As has been shown previously (Beck et al., 2021; Hassemer et al., 2019; Ottenlips et al., 2021; Slimp et al., 2021), the phylogenomic data presented here, including the Angiosperms353 loci and genome skimming data, were useful for improving morphology-based taxonomy, species and subspecies delimitation for some taxa. However, for other taxa, the data were insufficient to address such questions (e.g., Wenzell et al., 2021) or were complex and difficult to interpret in a taxonomic framework. Below we compare the trees from the three main datasets (353, plastome and nrDNA) to outline the main taxonomic implications for southern hemisphere *Myosotis.*

In our taxonomic revisionary papers on southern hemisphere *Myosotis* and here, we have used the general-lineage species concept (de Queiroz, 2007; de Queiroz, 1998) to delimit species as separately evolving lineages, where possible using multiple lines of evidence in an integrative framework (e.g., Meudt, 2021, 2017; Meudt et al., 2013; Prebble et al., 2022). In addition, we have used subspecies rank for allopatric taxa that have few minor distinguishing characters that may have limited overlap (Stuessy 2009; Meudt 2021). Although microsatellite data and AFLPs have provided evidence for independently evolving metapopulations for some *Myosotis* taxa in the *M. petiolata* group and the pygmy group, most molecular datasets used to date have not been informative for species or subspecies delimitation, and morphology has been used as the main operating criterion. The near exhaustive interspecific and comprehensive infraspecific sampling in this phylogenomic study is novel and has important implications, including the ability to assess not only species relationships but also monophyly, species limits, phylogeography, even underlying evolutionary processes, all of which may be much more complex than depicted in trees with much more limited sampling, especially given the evolutionary history of this group.

Speciation in the southern hemisphere *Myosotis* radiation has occurred geologically recently, over a very short period of time, and indeed may be incipient and ongoing (Meudt et al., 2015; Prebble et al., 2022, 2019). This is a common scenario for many New Zealand radiations, whose species limits are difficult to interpret in light of complex patterns seen in published phylogenies that may involve polyploidy, hybridisation and introgression, or retention of ancestral diversity (incomplete lineage sorting) or a combination of these (Heenan et al., 2022; Meudt and Bayly, 2008; Shepherd et al., 2023; Shepherd and Heenan, 2022; Winkworth et al., 2005). Many New Zealand *Myosotis* species are naturally uncommon, rare or threatened; they are often restricted to certain substrates, habitats or areas; and they often have highly reduced, allopatric or sometimes overlapping geographic ranges (Meudt et al., 2015). In *Myosotis,* species are largely clearly defined based on morphology, and there is very limited evidence for extant hybridisation and gene flow (see Appendix D), yet the results in all three phylogenomic datasets show geographic structuring, which would not be expected from incomplete lineage sorting alone (Meudt and Bayly, 2008; Shepherd et al., 2023). Rather, a previously proposed model of species diversification seems to fit the phylogenomic patterns seen here for *Myosotis,* in which climatic and environmental instability in New Zealand during the Pliocene and Pleistocene, together with glaciation cycles and vulcanism, provided the opportunity for the evolution of local, geographic lineages (Winkworth et al., 2005). Such local lineages would be geographically proximate, in different stages of speciation, and would have undergone multiple cycles of geographic expansion and contraction (and thus hybridisation and introgression with other local lineages). Thus, the process of speciation would be closely linked to geography via local adaptation to different habitats or substrates from nearby ancestors with whom cycles of hybridisation and introgression have occurred, likely involving parapatric and peripatric modes of speciation. Such a scenario could be responsible for the geographic admixture and structure found here for New Zealand *Myosotis*, on both a local scale as well as inter-island patterns discussed above (section 4.2). A similar model has been proposed for Mediterranean plants including another Boraginaceae genus, *Iberodes* M.Serrano et al., where Pliocene-Pleistocene ecological divergence without physical barriers to gene flow, a progenitor–derivative (budding) pattern of a restricted species evolving from a more widely distributed one (paraphyly), and niche differentiation were all evident (Otero et al., 2022).

Given this combination of factors and model of species diversification, a number of outcomes for species are all plausible, including monophyly vs polyphyly (or paraphyly; Otero et al., 2022), presence vs lack of geographic structure, blurry vs clear species limits, congruent vs incongruent placement in different trees, and lack of resolution and support vs highly resolved and well supported. The outcome for a given species may or may not be in agreement with the underlying morphology-based taxonomy (or it may be uninformative), and therefore the species limits and taxonomy of each taxon should be assessed in the general-lineage concept framework after comparing the phylogenomic results to all other available evidence and biological knowledge (e.g., Appendix D).

Overall, the phylogenomic results are congruent with the current taxonomy in many cases (Appendix D). About half of the *Myosotis* taxa were either monophyletic, or all of the sampled individuals were at least contained in the same larger clade, including five taxa that were always monophyletic (*M. albosericea, M. capitata, M. exarrhena, M. goyenii* subsp. *infima,* and *M. uniflora;* see section 3.10). Even for some taxa that were not consistently monophyletic, there was very high support for most of the individuals grouping together (including *M. brevis, M. rakiura, M. spatulata, M. pansa* subsp. *pansa*, among others). Thus, despite the challenges and complexities discussed above regarding the evolutionary histories of New Zealand species radiations, the phylogenomic results presented here were perhaps surprisingly informative and largely congruent with the current species limits of about half of the *Myosotis* taxa (Appendix D).

However, for the other half of the sampled taxa, the phylogenomic data were partially to largely incongruent with current species limits. For the 16 taxa (30 %; including 13 species and 3 subspecies, see section 3.10) that showed some or minor polyphyly (i.e., in two clades only, and only in one or two of the three trees), there were clear examples where the taxonomy could be improved at both the species and subspecies levels. For example, the Western Nelson (northwestern part of South Island) individuals of *M. traversii* subsp. *traversii,* whose morphology is unusual, will be re-evaluated in another study (Meudt unpubl. data) as belonging to another species, possibly *M. suavis* (as evidenced in the ASTRAL-III species tree, Fig. S3). In other cases, disjunct populations did not cluster together, such as in *M. arnoldii* and *M. colensoi,* which suggest that recognition of allopatric subspecies could be appropriate (to be addressed in another study for *M. arnoldii*; Meudt unpubl. data) or could simply be a lack of phylogenetic resolution (likely for *M. colensoi,* which is morphologically constant in both populations). The status of some taxa improved after confirmation of *a priori* misidentifications that was responsible for some of the observed "polyphyly" (e.g., *M. venticola, M. oreophila* and *M. brevis*; see further below in section 4.3). Other cases were not simple to resolve and may require additional infraspecific sampling or are tag names currently under morphological study (Meudt unpubl. data) whose species limits have yet to be defined (e.g., *M.* sp. "Fiordland" and *M.* sp. "Asbestos"; Appendix D).

For the 12 taxa (22 %) that were always polyphyletic in all three trees, several of these represent the most widespread taxa in the lineage, for which we had high intraspecific sampling, i.e., *M. antarctica* subsp. *antarctica, M. antarctica* subsp. *traillii, M. australis, M. lyallii* subps. *elderi, M. lyallii* subsp. *lyallii,* and *M. macrantha.* For example, regarding *M. australis,* there were three main lineages found: 1) most *M. australis*, 2) Western Nelson individuals, and 3) Australian individuals (Appendix C). The main *M. australis* lineage comprises all North Island, and Marlborough, Canterbury and Otago (South Island) individuals (including all white and yellow-flowered *M. australis,* and those individuals previously identified as *M. lytteltonensis*), which is consistent with the recognition of one morphologically variable and widespread species in Australasia, *M. australis* (Meudt et al., 2020). However, the exclusion of the two Western Nelson samples (one of each previously identified as *M. mooreana* and *M. australis* "small white") from this clade was unexpected, and these individuals instead cluster with other taxa from the same geographic area, as do the Australian *M. australis* (see section 4.2 and Appendix D). There is no morphological data to support their recognition as a separate taxon (Meudt et al., 2020), which could suggest the possibility of a cryptic species (Struck et al., 2018), but as the genetic data are inconclusive, further study is required. Understanding the relationship between the main lineages in *M. australis* would benefit from both increased sampling from Western Nelson (such as that done in a similar study on *Craspedia* (Asteraceae) from this same area (Breitwieser et al., 2015); and perhaps also use of different markers, for example, SNP-based markers (e.g., Shepherd et al., 2023; Shepherd and Heenan, 2022).

For the widespread New Zealand taxa *M. lyallii* and *M. antarctica,* both of which also have subspecies, the phylogenomic results were complex and difficult to interpret regarding species and subspecies limits, showing a mixture of geographic and non-geographic patterns, and perhaps mirroring to some degree the challenges of species and subspecies delimitation in these two species based on morphology (Meudt and Prebble, 2018; Prebble et al., 2022). There are apparent examples of blurry species boundaries in certain geographic areas from both morphology and phylogenomics, especially between *M. ultramafica, M. retrorsa, M. colensoi* and *M. antarctica* (for *M. lyallii*), and *M. hikuwai*, *M. glauca* and others (for *M. antarctica*) (Figs. 3, 4), which is nevertheless to be expected based on the species diversification model described above, and which may have been variable and irregular throughout the range of widespread species such as these. To give one example, *M. ultramafica* occurs in the same mountain ranges as its close relative *M. lyallii* but on ultramafic substrates (Meudt & Prebble 2022), and the data here are consistent with incipient parapatric speciation of *M. ultramafica* from the more widespread *M. lyallii* (Figs. 3, 4; Appendix D). Such a progenitor-derivative (budding) parapatry (such as that found in *Iberodes,* the Mediterranean Boraginaceae genus mentioned above; Otero et al., 2022), often involves edaphic specialisation of the more derived species (Crawford, 2010). Indeed, several additional species in multiple Boraginaceae genera have evolved on ultramafic substrates in Mediterranean Europe (e.g., species in *Halacsya* Dörfl., *Onosma* L., *Paramolktkia* Greuter, and, notably, *Myosotis alpestris* F.W. Schmidt subsp. *suaveolens* (Waldst. & Kit.) Strid; Cecchi and Selvi, 2009; Cecchi et al., 2011; Jakovljević et al., 2019), the United States (e.g., eight species of *Cryptantha,* Simpson and York, 2024), and the Pacific (including Australia, New Caledonia, and New Zealand; van der Ent et al., 2015; Lee 1992). These are often derived independently from repeated colonisation of neighbouring non-ultramafic species that may be pre-adapted to drought tolerance and basic substrates (Cecchi and Selvi, 2009; Cecchi et al., 2011). Ultramafic specialisation in New Zealand *Myosotis* has also originated independently from nearby ancestors. *Myosotis ultramafica* is one of only three New Zealand *Myosotis* species considered to be endemic to ultramafic substrates, together with *M. monroi* and *M. laeta* from northern South Island (van der Ent et al., 2015; Lee 1992), and each of these are more closely related to other geographically proximate, non-ultramafic specialist species, rather than to each other (Figs. 3, 4; Appendix D).

In other cases, the phylogenomic results were surprising as the morphological data are clear and relatively uncontroversial regarding species boundaries, such as *M. glauca, M. pulvinaris,* and *M. retrorsa.* Overall, the extensive intraspecific sampling in our study highlights the potential detrimental implications of sampling only one or a few individuals per species, especially for widespread species, and species with subspecies. Low sampling can lead to ascertainment bias and spurious reduced complexity in trees, as well as the inability to fully examine species limits. To fully understand taxonomic and geographic patterns of species radiation, multiple individuals per taxon is highly recommended (e.g., Shaw and Small, 2005).

With respect to some species pairs and intraspecific patterns, the limitations of the phylogenomic datasets were apparent, as has been noted in previous studies (e.g., Wenzell et al., 2021). For example, the phylogenomic analyses were unable to draw clear species boundaries between *M. matthewsii* and *M. spatulata*, and *M. chaffeyorum* and *M. tenericaulis,* respectively. In both of these pairs, the first species listed (*M. matthewsii, M. chaffeyorum*) has a reduced and parapatric or slightly overlapping geographic distribution compared to the latter species

(*M. spatulata, M. tenericaulis*). The phylogenomic analyses also indicated other nearby species that might also be closely related (see Appendix D), which again highlights the importance of geography to speciation in New Zealand *Myosotis.* Detailed studies with additional focused infraspecific sampling, and perhaps SNP-based markers are required, such as ddRADseq or genotype-by-sequencing, which might be better suited to these two questions of species boundaries (Appendix D). In addition, none of the six species with subspecies (*M. antarctica, M. brockiei, M. goyenii, M. lyallii, M. pansa* and *M. traversii*) were ever monophyletic in any of the phylogenomic trees, although in some cases, the underlying allopatric subspecies were monophyletic (or nearly so). Nevertheless, in the ASTRAL-III Angiosperms353 species + subspecies mapping for the species tree (Fig. S3B), the subspecies are closely related to one another in *M. antarctica, M. brockiei* and *M. pansa*, but this was not the case for *M. goyenii, M. lyallii* or *M. traversii*. For nine species, including five of the six species with subspecies, high infraspecific sampling (>10 individuals; *M. antarctica, M. australis, M. brockiei, M. forsteri, M. lyallii, M. macrantha, M. saxatilis, M. traversii* and *M.* sp. "Livingstone") generally meant higher levels of non-monophyly (Fig. S7). Every one of these taxa showed clustering of individuals by geography, which in turn clustered with other taxa from the same geographic area, which have different taxonomic interpretations (Appendix D).

With respect to hybrids, there is little evidence that natural hybridisation is frequent between southern hemisphere *Myosotis* species, and we have only seen rare examples of what appear to be hybrids with intermediate morphologies during many years of field work. At least three putative hybrids have been documented (Meudt et al., 2020; Meudt and Prebble, 2018; this paper), two of which were sampled here together with their putative parents (Table 1; Table S1). In both cases, the phylogenomic tree topologies of all three datasets are congruent with gene flow occurring between the species pairs and their putative hybrid at these locations, however it is not possible to determine which are the maternal or paternal progenitors (Appendix D, see under *Myosotis colensoi* × *saxatilis* hybrid and *Myosotis cheesemanii* × *pulvinaris* hybrid).

Interestingly, at the individual level, the phylogenomic datasets proved to be very useful to recognise and correctly identify certain samples whose placement in one or more of the trees lead us to question their *a priori* identification. Even though several of the phylogenies were not fully resolved, we were still able to use them to correctly re-identify some individuals, sometimes with high support and in multiple trees. Some of these individuals had already been flagged with puzzling morphological characters or uncertain *a priori* identification. In cases such as these, having morphological and other biological knowledge can help understand whether the unusual phylogenomic placement is the result of an erroneous taxonomy, misidentification, or incorrect or unresolved phylogeny due to another process affecting the genetic signal such as incipient speciation, incomplete lineage sorting, isolation by distance, or hybridisation. We reviewed the specimens, morphological data, previous morphological and genetic analyses for a handful of such specimens, many of which are discussed in Appendix D.

In at least five cases, the individual was found to be misidentified, and this was highly supported by both the phylogenomic data, morphology, and sometimes other evidence. For example, one subantarctic individual identified as *M. antarctica* subsp. *antarctica* (anan_CI_Hone1) was always placed in all three trees with high support with samples of the other subantarctic species, *M. capitata* (Figs. 3–5, S2, S8 and S9). This result confirmed a suspicion that both species were present and collected as a mixed collection at this location (Mt Honey, Campbell Island; JMP pers. obs.), and is congruent with previous microsatellite analysis of the same individuals (Prebble et al., 2019) (Appendix D, see under *M. capitata*). A similar highly supported outcome was found for brev_SO_Bend2 (which is not *M. brevis,* but *M. antarctica*), matt_NN_Waim (which is not *M. matthewsii,* but *M. spatulata*), oreo_SO_Duns (which is not *M. oreophila,* but *M. venticola*), and livi_SS_Hami (which is not *M.* "Livingstone" but *M. ultramafica*; Appendix D). However, there were also other individuals whose unusual placement could not be explained, including from *M. antarctica, M. brockiei, M. explanata, M. glauca, M. lyallii* subsp. *elderi,* and *M. tenericaulis.* For example, no taxonomic, morphological or geographical explanation could be found for the highly supported placement of glau_SO_Nevi within *M. brevis*. These individuals are likely causing some of the large amounts of discordance in the ASTRAL-III species trees (Fig. S3). For most of these taxa, additional intraspecific sampling for the molecular phylogenies, and in some cases additional morphological data (e.g., for *M. explanata,* which has not yet been included in the recent taxonomic revision), are required (Appendix D).

4.4. Conclusions

In summary, this paper represents the first phylogenomic analysis of the southern hemisphere *Myosotis* lineage and is novel compared to other similar studies of island or mountain radiations in its comprehensive interspecific sampling and extensive intraspecific sampling for all three datasets (whole plastomes, nrDNA, and Angiosperms353). Relative to a previous *Myosotis* phylogeny, it has ~ 80–120 × more molecular data and ~ 2–3 × more taxon sampling (Meudt et al., 2015), and to our knowledge is the first study to publish complete plastomes for *Myosotis.* Despite short branches on the backbone of most trees, phylogenomic analyses of the three datasets showed some congruent patterns, including taxonomic and geographic patterns, multiple switches between ebracteate-erect and bracteate-prostrate habit, and multiple expansions within Aotearoa New Zealand from Te Waipounamu South Island to Te Ika-a-Māui North Island, Rakiura Stewart Island, subantarctic islands, and Australia, respectively. Although this phylogenomic study does not fully overcome all of the challenges regarding species delimitation of this rapid island and mountain species radiation, it nevertheless makes an important contribution to our understanding of species limits, evolution, and geographic patterns for the morphologically and ecologically diverse species in the southern hemisphere *Myosotis* radiation.

CRediT authorship contribution statement

Heidi M. Meudt: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing - original draft, Visualisation, Writing - review & editing. **Sofie Pearson:** Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Validation, Writing – review & editing. **Weixuan Ning:** Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Validation, Writing – review & editing. **Jessica M. Prebble:** Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing – review & editing. **Jennifer A. Tate:** Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing – review & editing.

Funding

Our *Myosotis* research was supported by the Museum of New Zealand Te Papa Tongarewa (Te Papa) [via an internal contestable fund]; the New Zealand Ministry of Business, Innovation and Employment via Manaaki Whenua – Landcare Research [via Strategic Science Investment Fund for Crown Research Institutes]; the Hutton Fund [2018; no grant number]; and the Department of Conservation [2015 & 2019; no grant numbers].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the Museum of New Zealand Te Papa (Te Papa) for funding the wet lab genetics research, technician time and bioinformatics, and Massey University for hosting one of us (SMP), for the current study. We thank the following institutions for funding and supporting the overarching *Myosotis* research project, including field work and taxonomic research; the Ministry of Business, Innovation and Employment's Science and Innovation Group, Manaaki Whenua – Landcare Research, Te Papa, the Hutton Fund, and the Department of Conservation. We gratefully acknowledge several landowners, Māori iwi (tribes) and the Department of Conservation who allowed access to field sites to collect specimens for this study over the last ~ 12 years, and the Department of Conservation who authorised field work under Manaaki Whenua – Landcare Research concession CA-31615-OTH. A special thank you to all those who assisted HMM and JMP with field work, particularly Antony Kusabs (former Te Papa Collection Manager Botany), as well as Dirk Albach, Graeme Atkins, John Barkla, Marilyn Barkla, Charlotte Boyt, Ilse Breitwieser, Hamish Brown, Mary Bruce, Zuri Burns, Di Carter, Paul Cashmore, Jan Clayton-Greene, Shannel Courtney, Michael Douglas, Chris Ecroyd, Alex Fergus, Kerry Ford, Adele Gilchrist, Jane Gosden, Phil Garnock-Jones, Peter Heenan, John Hobbs, Rowan Hindmarsh-Walls, Melissa Hutchison, Carlos Lehnebach, Mike Lusk, Melissa Jager, Justin Liu, Alan Lee, David Lyttle, Cathy Jones, Daniel Kimber, Emily King, Kristina MacDonald, Alan Mark, Santiago Martín-Bravo, Alicia McCarthy, Todd McLay, Atamira Nuku, Emma Lewis, Colin Miskelly, Chris Morse, Colin Ogle, Dhana Pillai, Kay Pilkington, Tim Poupart, Brian Rance, Mathew Rees, Kalinka Rexer-Huber, Geoff Rogers, Sam Rowland, Cara-Lisa Schloots, Alice Shanks, Tony Silbery, Barbara Simpson, Neill Simpson, Nick Singers, Barry Sneddon, Rebecca Stanley, Rebecca Teele, Jon Terry, Mike Thorsen, David Toole, Simon Walls, Tui Warmenhoven, Wiremu Wharepapa, Rainer Vogt, Hugh Wood and Andreas Zeller. Thanks to herbarium HO for allowing destructive sampling from their specimens, Jonathan Frericks and Cliff Keilty at Percy Scenic Reserve (Hutt City Council, Wellington, New Zealand) for allowing and supporting sampling of their living collection. Thanks to Max Weigend for identifying *Microula trichocarpa*. We are especially grateful to Xiaoxiao Lin at the Massey Genome Service. We wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI; https://www.nesi.org.nz) high performance computing facilities, consulting support, and training services as part of this research. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme. We also thank Neil Dixon and James Hall of Te Papa's IT team for support using NeSI at Te Papa. Finally, we appreciate the constructive feedback received from the referees who reviewed our paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2024.108250.

Data availability

Data and scripts are available at NCBI (BioProject PRJNA1125924), Dryad (https://doi.org/10.5061/dryad.v41ns1s53) and Github (https://github.com/SofiePearson/Myosotis).

References

Baker, W.J., Dodsworth, S., Forest, F., Graham, S.W., Johnson, M.G., McDonnell, A., Pokorny, L., Tate, J.A., Wicke, S., Wickett, N.J., 2021. Exploring Angiosperms353: An open, community toolkit for collaborative phylogenomic research on flowering plants. Am. J. Bot. 108, 1059–1065. https://doi.org/10.1002/ajb2.1703.

Bakker, F.T., 2017. Herbarium genomics: skimming and plastomics from archival specimens. Webbia 72, 35–45. https://doi.org/10.1080/00837792.2017.1313383.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455. https://doi.org/10.1089/cmb.2012.0021.

Barrett, C.F., Bacon, C.D., Antonelli, A., Cano, Á., Hofmann, T., 2016. An introduction to plant phylogenomics with a focus on palms. Bot. J. Linn. Soc. 182, 234–255. https://doi.org/10.1111/boj.12399.

Beck, J.B., Markley, M.L., Zielke, M.G., Thomas, J.R., Hale, H.J., Williams, L.D., Johnson, M.G., 2021. Are Palmer's elm-leaf goldenrod and the smooth elm-leaf goldenrod real? The Angiosperms353 kit provides within-species signal in *Solidago ulmifolia s.l.* Syst. Bot. 46, 1107–1113. https://doi.org/10.1600/036364421X16370109698740.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Borowiec, M.L., 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4, e1660.

Breitwieser, I., Ford, K.A., Smissen, R.D., 2015. Characterisation of SSR markers for New Zealand *Craspedia* and their application in Kahurangi National Park. N. Z. J. Bot. 53, 60–73. https://doi.org/10.1080/0028825X.2014.1001410.

Brewer, G.E., Clarkson, J.J., Maurin, O., Zuntini, A.R., Barber, V., Bellot, S., Biggs, N., Cowan, R.S., Davies, N.M.J., Dodsworth, S., Edwards, S.L., Eiserhardt, W.L., Epitawalage, N., Frisby, S., Grall, A., Kersey, P.J., Pokorny, L., Leitch, I.J., Forest, F., Baker, W.J., 2019. Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. Front. Plant Sci. 10, 1102. https://doi.org/10.3389/fpls.2019.01102.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Cecchi, L., Coppi, A., Selvi, F., 2011. Evolutionary dynamics of serpentine adaptation in *Onosma* (Boraginaceae) as revealed by ITS sequence data. Plant Syst. Evol. 297, 185–199.

Cecchi, L., Selvi, F., 2009. Phylogenetic relationships of the monotypic genera *Halacsya* and *Paramoltkia* and the origins of serpentine adaptation in circum-mediterranean Lithospermeae (Boraginaceae): insights from ITS and matK DNA sequences. Taxon 58, 700–714.

Chacón, J., Luebert, F., Hilger, H.H., Ovchinnikova, S., Selvi, F., Cecchi, L., Guilliams, C.M., Hasenstab-Lehman, K., Sutorý, K., Simpson, M.G., Weigend, M., 2016. The borage family (Boraginaceae s.str.): A revised infrafamilial classification based on new phylogenetic evidence, with emphasis on the placement of some enigmatic genera. Taxon 65, 523–546. https://doi.org/10.12705/653.6.

Chan, K.O., Hutter, C.R., Wood Jr, P.L., Grismer, L.L., Brown, R.M., 2020. Target-capture phylogenomics provide insights on gene and species tree discordances in Old World treefrogs (Anura: Rhacophoridae). Proc. R. Soc. B 287.

Chen, L., Kadereit, G., Veranso-Libalah, M.C., 2023. Combining Angiosperms353 and Sanger data provides support for the reinstatement of the genus *Myrianthemum* (Melastomataceae). Bot. J. Linn. Soc. 203, 123–133. https://doi.org/10.1093/botlinnean/boad024.

Cockayne, L., 1928. The Vegetation of New Zealand. Wilhelm Engelmann, Leipzig.

Cohen, J.I., 2014. A phylogenetic analysis of morphological and molecular characters of Boraginaceae: evolutionary relationships, taxonomy, and patterns of character evolution. Cladistics 30, 139–169. https://doi.org/10.1111/cla.12036.

Crawford, D.J., 2010. Progenitor-derivative species pairs and plant speciation. Taxon 59, 1413–1423.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. GigaScience 10. https://doi.org/10.1093/gigascience/giab008.

de Lange, P.J., Rolfe, J.R., Barkla, J.W., Courtney, S., Champion, P.D., Perrie, L.R., Beadel, S.M., Ford, K., Breitwieser, I., Schönberger, I., Hindmarsh-Walls, R., 2018. Conservation Status of New Zealand Indigenous Vascular Plants, 2017 (No. 22). Department of Conservation, Wellington, New Zealand. https://www.doc.govt.nz/Documents/science-and-technical/nztcs22entire.pdf.

de Lange, P.J., Murray, B.G., 2002. Contributions to a chromosome atlas of the New Zealand flora—37. Miscellaneous families. N. Z. J. Bot. 40, 1–23. https://doi.org/10.1080/0028825X.2002.9512767.

de Queiroz, K., 1998. The general lineage concept of species, species criteria, and the process of speciation. In: Howard, D.J., Berlocher, S.H. (Eds.), Endless Forms: Species and Speciation. Oxford University Press, pp. 57–75.

de Queiroz, K., 2007. Species concepts and species delimitation. Syst. Biol. 56, 879–886. https://doi.org/10.1080/10635150701701083.

Dodsworth, S., 2015. Genome skimming for next-generation biodiversity analysis. Trends Plant Sci. 20, 525–527. https://doi.org/10.1016/j.tplants.2015.06.012.

Dodsworth, S., Pokorny, L., Johnson, M.G., Kim, J.T., Maurin, O., Wickett, N.J., Forest, F., Baker, W.J., 2019. Hyb-Seq for flowering plant systematics. Trends Plant Sci. 24, 887–891. https://doi.org/10.1016/j.tplants.2019.07.011.

Frankel, L., Murúa, M., Espíndola, A., 2022. Biogeography and ecological drivers of evolution in the Andes: resolving the phylogenetic backbone for *Calceolaria* (Calceolariaceae). Bot. J. Linn. Soc. https://doi.org/10.1093/botlinnean/boab079.

French, P.A., Brown, G.K., Bayly, M.J., 2016. Incongruent patterns of nuclear and chloroplast variation in *Correa* (Rutaceae): introgression and biogeography in south-eastern Australia. Plant Syst. Evol. 302, 447–468. https://doi.org/10.1007/s00606-016-1277-7.

Frohlich, M.W., 1978. Systematics of *Heliotropium* section *Orthostachys* in Mexico. Harvard University. Available from ResearchGate. Unpublished PhD thesis.

Frohlich, M.W., Sage, R.F., Craven, L.A., Schuster, S., Gigot, G., Hilger, H.H., Akhani, H., Mahdavi, P., Luebert, F., Weigend, M., Thulin, M., Doyle, J.J., Doyle, J.L., Vogan, P., Forrest, A., Fulcher, T.K., Devey, D.S., Chase, M.W., 2022. Molecular phylogenetics of *Euploca* (Boraginaceae): homoplasy in many characters, including the C4 photosynthetic pathway. Bot. J. Linn. Soc. 199, 497–537. https://doi.org/10.1093/botlinnean/boab082.

Frost, L., Bedoya, A.M., Lagomarsino, L., 2024. Artifactual Orthologs and the Need for Diligent Data Exploration in Complex Phylogenomic Datasets: A museomic case study from the Andean flora. Systematic Biology. In press. DOI: 10.1093/sysbio/syad076.

Fu, C.-N., Mo, Z.-Q., Yang, J.-B., Cai, J., Ye, L.-J., Zou, J.-Y., Qin, H.-T., Zheng, W., Hollingsworth, P.M., Li, D.-Z., Gao, L.-M., 2022. Testing genome skimming for species discrimination in the large and taxonomically difficult genus *Rhododendron*. Mol. Ecol. Resour. 22, 404–414. https://doi.org/10.1111/1755-0998.13479.

Gardner, E.M., Johnson, M.G., Pereira, J.T., Puad, A.S.A., Arifiani, D., Wickett, N.J., Zerega, N.J.C., 2020. Paralogs and off-target sequences improve phylogenetic resolution in a densely-sampled study of the breadfruit genus (*Artocarpus*, Moraceae). Syst. Biol. 70, 558–575. https://doi.org/10.1093/sysbio/syaa073.

GBIF, 2024, GBIF.org Occurrence Download, 23 Apr 2024. DOI: 10.15468/dl.w47eqg.

Grau, J., 1964. Die Zytotaxonomie der *Myosotis-alpestris*- und der *Myosotis-silvatica*-Gruppe in Europa. Österreichische Botanische Zeitschrift 111, 561–617. https://www.jstor.org/stable/43337391.

Grau, J., 1968. Cytotaxonomische Bearbeitung der Gattung *Myosotis* L. III. Die Annuellen Sippen. Mitteilungen Der Botanischen Staatssammlung 7, 17–100.

Harrison, N., Kidner, C.A., 2011. Next-generation sequencing and systematics: What can a billion base pairs of DNA sequence data do for you? Taxon 60, 1552–1566. https://doi.org/10.1002/tax.606002.

Hart, M.L., Forrest, L.L., Nicholls, J.A., Kidner, C.A., 2016. Retrieval of hundreds of nuclear loci from herbarium specimens. Taxon 65, 1081–1092. https://doi.org/10.12705/655.9.

Hassemer, G., Bruun-Lund, S., Shipunov, A.B., Briggs, B.G., Meudt, H.M., Rønsted, N., 2019. The application of high-throughput sequencing for taxonomy: The case of *Plantago* subg. *Plantago* (Plantaginaceae). Mol. Phylogenet. Evol. 138, 156–173. https://doi.org/10.1016/j.ympev.2019.05.013.

Heenan, P.B., McGlone, M.S., Mitchell, C.M., Cheeseman, D.F., Houliston, G.J., 2022. Genetic variation reveals broad-scale biogeographic patterns and challenges species' classification in the *Kunzea ericoides* (kānuka; Myrtaceae) complex from New Zealand. N. Z. J. Bot. 60, 2–26. https://doi.org/10.1080/0028825X.2021.1903946.

Hosner, P.A., Faircloth, B.C., Glenn, T.C., Braun, E.L., Kimball, R.T., 2016. Avoiding Missing Data Biases in Phylogenomic Inference: An Empirical Study in the Landfowl (Aves: Galliformes). Mol. Biol. Evol. 33, 1110–1125. https://doi.org/10.1093/molbev/msv347.

Howard, C.C., Crowl, A.A., Harvey, T.S., Cellinese, N., 2022. Peeling back the layers: First phylogenomic insights into the Ledebouriinae (Scilloideae, Asparagaceae). Mol. Phylogenet. Evol. 169, 107430. https://doi.org/10.1016/j.ympev.2022.107430.

Jakovljević, K., Đurović, S., Antušević, M., Mihailović, N., Buzurović, U., Tomović, G., 2019. Heavy metal tolerance of *Pontechium maculatum* (Boraginaceae) from several ultramafic localities in Serbia. Botanica Serbica 43 (1), 73–83. https://doi.org/10.2298/BOTSERB1901073J.

Jin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., dePamphilis, C.W., Yi, T.-S., Li, D.-Z., 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21, 241. https://doi.org/10.1186/s13059-020-02154-5.

Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J.C., Wickett, N.J., 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. Appl. Plant Sci. 4. https://doi.org/10.3732/apps.1600016.

Johnson, M.G., Pokorny, L., Dodsworth, S., Botigué, L.R., Cowan, R.S., Devault, A., Eiserhardt, W.L., Epitawalage, N., Forest, F., Kim, J.T., Leebens-Mack, J.H., Leitch, I.J., Maurin, O., Soltis, D.E., Soltis, P.S., Wong, G.-K.-S., Baker, W.J., Wickett, N.J., 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. Syst. Biol. 68, 594–606. https://doi.org/10.1093/sysbio/syy086.

Junier, T., Zdobnov, E.M., 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26, 1669–1670. https://doi.org/10.1093/bioinformatics/btq243.

Kadereit, J.W., Abbott, R.J., 2021. Plant speciation in the Quaternary. Plant Ecolog. Divers. 14, 105–142. https://doi.org/10.1080/17550874.2021.2012849.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.

Kandziora, M., Sklenář, P., Kolář, F., Schmickl, R., 2021. How to Tackle Phylogenetic Discordance in Recent and Rapidly Radiating Groups? Developing a Workflow Using *Loricaria* (Asteraceae) as an Example. Front. Plant Sci. 12, 765719. https://doi.org/10.3389/fpls.2021.765719.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.

Lagomarsino, L.P., Frankel, L., Uribe-Convers, S., Antonelli, A., Muchhala, N., 2022. Increased resolution in the face of conflict: phylogenomics of the Neotropical bellflowers (Campanulaceae: Lobelioideae), a rapid plant radiation. Ann. Bot. 129, 723–736. https://doi.org/10.1093/aob/mcac046.

Larridon, I., Villaverde, T., Zuntini, A.R., Pokorny, L., Brewer, G.E., Epitawalage, N., Fairlie, I., Hahn, M., Kim, J., Maguilla, E., Maurin, O., Xanthos, M., Hipp, A.L., Forest, F., Baker, W.J., 2019. Tackling rapid radiations with targeted sequencing. Front. Plant Sci. 10, 1655. https://doi.org/10.3389/fpls.2019.01655.

Lee, W.G., 1992. In: New Zealand Ultramafics. in 'the Ecology of Areas with Serpentinized Rocks. A World View'. Kluwer Academic Publishers:, Dordrecht, The Netherlands, pp. 375–417.

Lehnebach, C.A., 2012. Two new species of forget-me-nots (*Myosotis*, Boraginaceae) from New Zealand. PhytoKeys 53–64.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Lockhart, P.J., McLenachan, P.A., Havell, D., Glenny, D., Huson, D., Jensen, U., 2001. Phylogeny, Radiation, and Transoceanic Dispersal of New Zealand Alpine Buttercups: Molecular Evidence under Split Decomposition. Ann. Mo. Bot. Gard. 88, 458–477. https://www.jstor.org/stable/3298586.

Luebert, F., Cecchi, L., Frohlich, M.W., Gottschling, M., Guilliams, C.M., Hasenstab-Lehman, K.E., Hilger, H.H., Miller, J.S., Mittelbach, M., Nazaire, M., Nepi, M., 2016. Familial classification of the Boraginales. Taxon 65, 502–522.

Luque, T., 1992. Karyological studies on Spanish Boraginaceae. VI. Contribution to the tribe Eritrichieae. Bot. J. Linn. Soc. 110, 77–94. https://doi.org/10.1111/j.1095-8339.1992.tb00417.x.

Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55, 21–30. https://doi.org/10.1080/10635150500354928.

Maurin, K.J.L., Smissen, R.D., Lusk, C.H., 2022. A dated phylogeny shows Plio-Pleistocene climates spurred evolution of antibrowsing defences in the New Zealand flora. New Phytol. 233, 546–554. https://doi.org/10.1111/nph.17766.

McKain, M.R., Johnson, M.G., Uribe-Convers, S., Eaton, D., Yang, Y., 2018. Practical considerations for plant phylogenomics. Appl. Plant Sci. 6, e1038.

McKinnon, G.E., Vaillancourt, R.E., Jackson, H.D., Potts, B.M., 2001. Chloroplast sharing in the Tasmanian eucalypts. Evolution 55, 703–711. https://doi.org/10.1111/j.0014-3820.2001.tb00806.x.

McLay, T.G.B., Birch, J.L., Gunn, B.F., Ning, W., Tate, J.A., Nauheimer, L., Joyce, E.M., Simpson, L., Schmidt-Lebuhn, A.N., Baker, W.J., Forest, F., Jackson, C.J., 2021. New targets acquired: Improving locus recovery from the Angiosperms353 probe set. Appl. Plant Sci. 9. https://doi.org/10.1002/aps3.11420.

Meudt, H.M., 2016. Pollen morphology and its taxonomic utility in the Southern Hemisphere bracteate-prostrate forget-me-nots (*Myosotis*, Boraginaceae). N. Z. J. Bot. 54, 475–497. https://doi.org/10.1080/0028825X.2016.1229343.

Meudt, H.M., 2017. Integrative, next-generation, collaborative vascular plant systematics in New Zealand. New Zealand Science Review 73, 99–106. https://doi.org/10.26686/nzsr.v73i3-4.8536.

Meudt, H.M., 2021. Taxonomic revision of five species groups of ebracteate-erect *Myosotis* (Boraginaceae) endemic to New Zealand, based on morphology, and description of new subspecies. Aust. Syst. Bot. 34, 252–304. https://doi.org/10.1071/sb20028.

Meudt, H.M., Albach, D.C., Tanentzap, A.J., Igea, J., Newmarch, S.C., Brandt, A.J., Lee, W.G., Tate, J.A., 2021. Polyploidy on Islands: Its Emergence and Importance for Diversification. Front. Plant Sci. 12, 637214. https://doi.org/10.3389/fpls.2021.637214.

Meudt, H.M., Bayly, M.J., 2008. Phylogeographic patterns in the Australasian genus *Chionohebe* (*Veronica* s.l., Plantaginaceae) based on AFLP and chloroplast DNA sequences. Mol. Phylogenet. Evol. 47, 319–338. https://doi.org/10.1016/j.ympev.2007.12.019.

Meudt, H.M., Prebble, J.M., 2018. Species limits and taxonomic revision of the bracteate-prostrate group of southern hemisphere forget-me-nots (*Myosotis*, Boraginaceae), including description of three new species endemic to New Zealand. Aust. Syst. Bot. 31, 48–105. https://doi.org/10.1071/SB17045.

Meudt, H.M., Prebble, J.M., 2022. Morphological analyses support recognition of three new threatened species of bracteate–prostrate *Myosotis* (Boraginaceae) endemic to the South Island of Aotearoa New Zealand. Aust. Syst. Bot. 35, 364–394. https://doi.org/10.1071/SB22012.

Meudt, H.M., Prebble, J.M., Stanley, R.J., Thorsen, M.J., 2013. Morphological and amplified fragment length polymorphism (AFLP) data show that New Zealand endemic *Myosotis petiolata* (Boraginaceae) comprises three rare and threatened species. Aust. Syst. Bot. 26, 210–232. https://doi.org/10.1071/SB13023.

Meudt, H.M., Prebble, J.M., Lehnebach, C.A., 2015. Native New Zealand forget-me-nots (*Myosotis*, Boraginaceae) comprise a Pleistocene species radiation with very low genetic divergence. Plant Syst. Evol. 301, 1455–1471. https://doi.org/10.1007/s00606-014-1166-x.

Meudt, H.M., Simpson, B.B., 2006. The biogeography of the austral, subalpine genus *Ourisia* (Plantaginaceae) based on molecular phylogenetic evidence: South American origin and dispersal to New Zealand and Tasmania. Biol. J. Linn. Soc. 87, 479–513. https://doi.org/10.1111/j.1095-8312.2006.00584.x.

Meudt, H.M., Thorsen, M.J., Prebble, J.M., 2020. Taxonomic revision of the *Myosotis australis* group (Boraginaceae) native to Australia, New Zealand and New Guinea. Aust. Syst. Bot. 33, 477–542. https://doi.org/10.1071/SB20014.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Mirarab, S., 2019. Species tree estimation using ASTRAL: practical considerations. *arXiv preprint arXiv:1904.03826*. DOI: 10.48550/arXiv.1904.03826.

Moore, L.B., 1961. Boraginaceae. In: Allan, H.H. (Ed.), Flora of New Zealand. Government Printer, Wellington, New Zealand, p. 323.

Moore, L.B., Simpson, M.J.A., 1973. A new *Myosotis* from north-west Nelson. N. Z. J. Bot. 11, 163–170. https://doi.org/10.1080/0028825X.1973.10430270.

Morales-Briones, D.F., Gehrke, B., Huang, C.H., 2022. Analysis of Paralogs in Target Enrichment Data Pinpoints Multiple Ancient Polyploidy Events in *Alchemilla s.l.* (Rosaceae). Syst. Biol. 71, 190–207. https://doi.org/10.1093/sysbio/syab032.

Muellner-Riehl, A.N., Anthelme, F., Ibanez, T., 2024. Past, present, and future of mountain and island systems. J. Syst. Evol. 62, 195–200. https://doi.org/10.1111/jse.13073.

Murillo-A, J., Valencia-D, J., Orozco, C.I., Parra-O, C., Neubig, K.M., 2022. Incomplete lineage sorting and reticulate evolution mask species relationships in Brunelliaceae, an Andean family with rapid, recent diversification. Am. J. Bot. 109, 1139–1156. https://doi.org/10.1002/ajb2.16025.

Murray, B.G., de Lange, P.J., 2013. Contributions to a chromosome atlas of the New Zealand flora – 40. Miscellaneous counts for 36 families. N. Z. J. Bot. 51, 31–60. https://doi.org/10.1080/0028825X.2012.747969.

Nauheimer, L., Weigner, N., Joyce, E., Crayn, D., Clarke, C., Nargar, K., 2021. HybPhaser: A workflow for the detection and phasing of hybrids in target capture data sets. Appl. Plant Sci. 9. https://doi.org/10.1002/aps3.11441.

Nicol, D.A., Saldivia, P., Summerfield, T.C., Heads, M., Lord, J.M., Khaing, E.P., Larcombe, M.J., 2024. Phylogenomics and morphology of Celmisiinae (Asteraceae: Astereae): Taxonomic and evolutionary implications. Mol. Phylogenet. Evol. 195, 108064. https://doi.org/10.1016/j.ympev.2024.108064.

Ning, W., 2023. Phylogenomics and evolution of polyploid *Azorella* (Apiaceae) in New Zealand: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Plant Biology at Massey University, Manawatū. Massey University, New Zealand https://hdl.handle.net/10179/19791.

Otero, A., Vargas, P., Fernández-Mazuecos, M., Jiménez-Mejías, P., Valcárcel, V., Villa-Machío, I., Hipp, A.L., 2022. A snapshot of progenitor–derivative speciation in *Iberodes* (Boraginaceae). Mol. Ecol. 3192–3209.

Ottenlips, M.V., Mansfield, D.H., Buerki, S., Feist, M.A.E., Downie, S.R., Dodsworth, S., Forest, F., Plunkett, G.M., Smith, J.F., 2021. Resolving species boundaries in a recent radiation with the Angiosperms353 probe set: the *Lomatium packardiae/L. anomalum* clade of the *L. triternatum* (Apiaceae) complex. Am. J. Bot. 108, 1217–1233. https://doi.org/10.1002/ajb2.1676.

Overson, R.P., Johnson, M.G., Bechen, L.L., Kinosian, S.P., Douglas, N.A., Fant, J.B., Hoch, P.C., Levin, R.A., Moore, M.J., Raguso, R.A., Wagner, W.L., Skogen, K.A., Wickett, N.J., 2023. A phylogeny of the evening primrose family (Onagraceae) using a target enrichment approach with 303 nuclear loci. BMC Ecol. Evol. 23, 66. https://doi.org/10.1186/s12862-023-02151-9.

Prebble, J.M., Tate, J.A., Meudt, H.M., Symonds, V.V., 2015. Microsatellite markers for the New Zealand endemic *Myosotis pygmaea* species group (Boraginaceae) amplify across species. Appl. Plant Sci. 3. https://doi.org/10.3732/apps.1500027.

Prebble, J.M., Meudt, H.M., Tate, J.A., Symonds, V.V., 2019. Comparing and co-analysing microsatellite and morphological data for species delimitation in the New Zealand native *Myosotis pygmaea* species group (Boraginaceae). Taxon 68, 731–750. https://doi.org/10.1002/tax.12096.

Prebble, J.M., Symonds, V.V., Tate, J.A., Meudt, H.M., 2022. Taxonomic revision of the southern hemisphere pygmy forget-me-not group (*Myosotis*; Boraginaceae) based on morphological, population genetic and climate-edaphic niche modelling data. Aust. Syst. Bot. 35, 63–94. https://doi.org/10.1071/SB21031.

Rabiee, M., Sayyari, E., Mirarab, S., 2019. Multi-allele species reconstruction using ASTRAL. Mol. Phylogenet. Evol. 130, 286–296. https://doi.org/10.1016/j.ympev.2018.10.033.

Rieseberg, L.H., Soltis, D.E., 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evolutionary Trends in Plants 5, 65–84.

Robertson, A.W., 1989. Evolution and pollination of New Zealand Myosotis (Boraginaceae): a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Botany at University of Canterbury. Christchurch, New Zealand. https://doi.org/10.26021/8460, http://hdl.handle.net/10092/5849.

Schmidt-Lebuhn, A.N., Egli, D., Grealy, A., Nicholls, J.A., Zwick, A., Dymock, J.J., Gooden, B., 2022. Genetic data confirm the presence of *Senecio madagascariensis* in New Zealand. N. Z. J. Bot. 62, 1–13. https://doi.org/10.1080/0028825X.2022.2148544.

Shaw, J., Small, R.L., 2005. Chloroplast DNA phylogeny and phylogeography of the North American plums (*Prunus* subgenus *Prunus* section *Prunocerasus*, Rosaceae). Am. J. Bot. 92, 2011–2030. https://doi.org/10.3732/ajb.92.12.2011.

Shee, Z.Q., Frodin, D.G., Cámara-Leret, R., Pokorny, L., 2020. Reconstructing the complex evolutionary history of the Papuasian *Schefflera* radiation through herbariomics. Front. Plant Sci. 11, 258. https://doi.org/10.3389/fpls.2020.00258.

Shepherd, L.D., Heenan, P.B., 2022. Phylogenomic analyses reveal a history of hybridisation and introgression between *Sophora* sect. *Edwardsia* (Fabaceae) species in New Zealand. N. Z. J. Bot. 60, 113–133. https://doi.org/10.1080/0028825X.2021.1960567.

Shepherd, L.D., McLay, T.G.B., 2011. Two micro-scale protocols for the isolation of DNA from polysaccharide-rich plant tissue. J. Plant Res. 124, 311–314. https://doi.org/10.1007/s10265-010-0379-5.

Shepherd, L.D., Townsend, A., de Lange, P.J., Perrie, L.R., 2023. Genetic analysis of a genus of plant mimics (*Alseuosmia* A.Cunn; Alseuosmiaceae) reveals incongruence between morphology and phylogeny and possible mimetic polymorphism. Evolutionary Journal of the Linnean Society 2. https://doi.org/10.1093/evolinnean/kzad005.

Simpson, M.G., York, D.A., 2024. *Cryptantha whippleae* (Boraginaceae), a new serpentine-adapted species endemic to northern California, USA. *PhytoKeys* 247, 155–172. https://doi.org/10.3897/phytokeys.247.132060.

Slater, G.S.C., Birney, E., 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinf. 6, 1–11. https://doi.org/10.1186/1471-2105-6-31.

Slimp, M., Williams, L.D., Hale, H., Johnson, M.G., 2021. On the potential of Angiosperms353 for population genomic studies. Appl. Plant Sci. 9. https://doi.org/10.1002/aps3.11419.

Small, R.L., Cronn, R.C., Wendel, J.F., 2004. Use of nuclear genes for phylogeny reconstruction in plants. Aust. Syst. Bot. 17, 145–170. https://doi.org/10.1071/SB03015.

Smissen, R.D., Breitwieser, I., Ward, J.M., 2004. Phylogenetic implications of trans-specific chloroplast DNA sequence polymorphism in New Zealand Gnaphalieae (Asteraceae). Plant Syst. Evol. 249, 37–53. https://doi.org/10.1007/s00606-004-0209-0.

Smith, S.A., Moore, M.J., Brown, J.W., Yang, Y., 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. BMC Evol. Biol. 15, 150. https://doi.org/10.1186/s12862-015-0423-0.

Soltis, D.E., Gitzendanner, M.A., Stull, G., Chester, M., Chanderbali, A., Chamala, S., Jordon-Thaden, I., Soltis, P.S., Schnable, P.S., Brad Barbazuk, W., 2013. The potential of genomics in plant systematics. Taxon 62, 886–898. https://doi.org/10.12705/625.13.

Straub, S.C.K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R.C., Liston, A., 2012. Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. Am. J. Bot. 99, 349–364. https://doi.org/10.3732/ajb.1100335.

Struck, T.H., Feder, J.L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V.I., Kistenich, S., Larsson, K.H., Liow, L.H., Nowak, M.D., Stedje, B., 2018. Finding evolutionary processes hidden in cryptic species. Trends Ecol. Evol. 33, 153–163. https://doi.org/10.1016/j.tree.2017.11.007.

Stubbs, R.L., Folk, R.A., Xiang, C.-L., Chen, S., Soltis, D.E., Cellinese, N., 2020. A phylogenomic perspective on evolution and discordance in the alpine-Arctic plant clade *Micranthes* (Saxifragaceae). Front. Plant Sci. 10, 1773. https://doi.org/10.3389/fpls.2019.01773.

Stuessy, T.F., 2009. Plant Taxonomy. The Systematic Evolution of Comparative Data, 2nd Edition. Columbia University Press, New York.

Thomas, A.E., Igea, J., Meudt, H.M., Albach, D.C., Lee, W.G., Tanentzap, A.J., 2021. Using target sequence capture to improve the phylogenetic resolution of a rapid radiation in New Zealand *Veronica*. Am. J. Bot. 108, 1289–1306. https://doi.org/10.1002/ajb2.1678.

Thomas, A., Meudt, H.M., Larcombe, M.J., Igea, J., Lee, W.G., Antonelli, A., Tanentzap, A.J., 2023. Multiple origins of mountain biodiversity in New Zealand's largest plant radiation. J. Biogeogr. 50, 947–960. https://doi.org/10.1111/jbi.14589.

Thureborn, O., Razafimandimbison, S.G., Wikström, N., Rydin, C., 2022. Target capture data resolve recalcitrant relationships in the coffee family (Rubioideae, Rubiaceae). Front. Plant Sci. 13, 967456. https://doi.org/10.3389/fpls.2022.967456.

Twyford, A.D., Ness, R.W., 2017. Strategies for complete plastid genome sequencing. Mol. Ecol. Resour. 17, 858–868. https://doi.org/10.1111/1755-0998.12626.

Ufimov, R., Gorospe, J.M., Fér, T., Kandziora, M., Salomon, L., van Loo, M., Schmickl, R., 2022. Utilizing paralogues for phylogenetic reconstruction has the potential to increase species tree support and reduce gene tree discordance in target enrichment data. Mol. Ecol. Resour. 22, 3018–3034. https://doi.org/10.1111/1755-0998.13684.

van Der Ent, A., Jaffré, T., L'Huillier, L., Gibson, N., Reeves, R.D., 2015. The flora of ultramafic soils in the Australia–Pacific Region: state of knowledge and research priorities. *Aust. J. Bot.* 63 (4), 173–190. https://doi.org/10.1071/BT15038.

van Kleinwee, I., Larridon, I., Shah, T., Bauters, K., Asselman, P., Goetghebeur, P., Leliaert, F., Veltjen, E., 2022. Plastid phylogenomics of the *Sansevieria* Clade of *Dracaena* (Asparagaceae) resolves a recent radiation. Mol. Phylogenet. Evol. 169, 107404. https://doi.org/10.1016/j.ympev.2022.107404.

Wardle, P., 1963. Evolution and distribution of the New Zealand flora, as affected by quaternary climates. N. Z. J. Bot. 1, 3–17. https://doi.org/10.1080/0028825X.1963.10429318.

Wardle, P., 1991. Vegetation of New Zealand. Cambridge, UK, Cambridge University Press.

Warren, B.H., Simberloff, D., Ricklefs, R.E., Aguilée, R., Condamine, F.L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernández-Palacios, J.M., Hengl, T., Norder, S.J., Rijsdijk, K.F., Sanmartín, I., Strasberg, D., Triantis, K.A., Valente, L.M., Whittaker, R.J., Gillespie, R.G., Emerson, B.C., Thébaud, C., 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. Ecol. Lett. 18, 200–217. https://doi.org/10.1111/ele.12398.

Weigend, M., Luebert, F., Selvi, F., Brokamp, G., Hilger, H.H., 2013. Mol. Phylogenet. Evol. 68, 604–618. https://doi.org/10.1016/j.ympev.2013.04.009.

Weigend, M., Selvi, F., Thomas, D.C., Hilger, H.H., 2016. Boraginaceae. In: Kadereit, J., Bittrich, V. (Eds.), Flowering Plants. Eudicots. The Families and Genera of Vascular Plants, Springer, Cham.

Weitemier, K., Straub, S.C.K., Cronn, R.C., Fishbein, M., Schmickl, R., McDonnell, A., Liston, A., 2014. Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. Appl. Plant Sci. 2. https://doi.org/10.3732/apps.1400042.

Wendel, J.F., Doyle, J.J., 1998. Phylogenetic incongruence: Window into genome history and molecular evolution. In: Soltis, D.E., Soltis, P.S., Doyle, J.J. (Eds.), Molecular Systematics of Plants II. Springer, Boston, MA, USA, pp. 265–296.

Wenzell, K.E., McDonnell, A.J., Wickett, N.J., Fant, J.B., Skogen, K.A., 2021. Incomplete reproductive isolation and low genetic differentiation despite floral divergence across varying geographic scales in *Castilleja*. Am. J. Bot. 108, 1270–1288. https://doi.org/10.1002/ajb2.1700.

Winkworth, R.C., Glenny, D., Lockhart, P.J., 2002a. Plant dispersal N.E.W.S from New Zealand. Trends Ecol. Evol. 17, 514–520. https://doi.org/10.1016/S0169-5347(02)02590-9.

Winkworth, R.C., Grau, J., Robertson, A.W., Lockhart, P.J., 2002b. The origins and evolution of the genus *Myosotis* L. (Boraginaceae). Mol. Phylogenet. Evol. 24, 180–193. https://doi.org/10.1016/S1055-7903(02)00210-5.

Winkworth, R.C., Wagstaff, S.J., Glenny, D., Lockhart, P.J., 2005. Evolution of the New Zealand mountain flora: Origins, diversification and dispersal. Org. Divers. Evol. 5, 237–247. https://doi.org/10.1016/j.ode.2004.12.001.

Yan, Y., da Fonseca, R.R., Rahbek, C., Borregaard, M.K., Davis, C.C., 2024. A new nuclear phylogeny of the tea family (Theaceae) unravels rapid radiations in genus *Camellia*. Mol. Phylogenet. Evol. 196, 108089. https://doi.org/10.1016/j.ympev.2024.108089.

Yan, Z., Smith, M.L., Du, P., Hahn, M.W., Nakhleh, L., 2022. Species Tree Inference Methods Intended to Deal with Incomplete Lineage Sorting Are Robust to the Presence of Paralogs. Syst. Biol. 71, 367–381. https://doi.org/10.1093/sysbio/syab056.

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.-T.-Y., 2017. Ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8, 28–36. https://doi.org/10.1111/2041-210X.12628.

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinf. 19, 153. https://doi.org/10.1186/s12859-018-2129-y.