New Zealand Journal of Botany

ISSN: 0028-825X (Print) 1175-8643 (Online) Journal homepage: www.tandfonline.com/journals/tnzb20

Sequence data support greater taxonomic diversity and a complicated history of hybridisation in New Zealand's sun orchids - *Thelymitra* (Orchidaceae)

Hayden R. Jones, Carlos A. Lehnebach & Jennifer A. Tate

To cite this article: Hayden R. Jones, Carlos A. Lehnebach & Jennifer A. Tate (26 Jun 2025): Sequence data support greater taxonomic diversity and a complicated history of hybridisation in New Zealand's sun orchids - *Thelymitra* (Orchidaceae), New Zealand Journal of Botany, DOI: 10.1080/0028825X.2025.2523007

To link to this article: https://doi.org/10.1080/0028825X.2025.2523007

NEW ZEALAND JOURNAL OF BOTANY
https://doi.org/10.1080/0028825X.2025.2523007

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Sequence data support greater taxonomic diversity and a complicated history of hybridisation in New Zealand's sun orchids - *Thelymitra* (Orchidaceae)

Hayden R. Jones [a], Carlos A. Lehnebach [b] and Jennifer A. Tate [a]

[a]School of Agriculture and Environment, Massey University, Palmerston North, New Zealand; [b]Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand

ABSTRACT

Thelymitra, commonly known as sun orchids, is a genus of orchids present in New Zealand (NZ), known for their radially symmetrical and brightly coloured flowers. A long history of nomenclatural changes, and a recent and rapid circumscription of Australian (AU) taxa has left our understanding of this genus in NZ lagging. In this study we report two phylogenies of NZ *Thelymitra* species, based on nuclear (*LFY*) and plastid (*ycf1*) sequence data. The phylogenies identify support for several taxa currently only known by tag-names as being distinct from those that are currently described both in NZ and AU, as well as evidence supporting an extensive history of hybrid speciation in this genus in NZ, which had been hypothesised previously for six species. This work has also found evidence for species complexes within NZ *Thelymitra* that need further molecular and taxonomic work to fully resolve species boundaries. In this study we also demonstrate a novel application of nanopore sequencing of amplicons, that enables the recovery of multiple sequence copies in allopolyploid taxa without the prior need for cloning.

ARTICLE HISTORY
Received 17 December 2024
Accepted 16 June 2025

HANDLING EDITOR
Todd McLay

KEYWORDS
Allopolyploidy; hybridisation; *LFY*; molecular phylogeny; nanopore sequencing; New Zealand; orchids; taxonomy; *Thelymitra*; *Ycf1*

Introduction

Comprising approximately 28,000 species (Christenhusz and Byng 2016) and present on every continent except Antarctica, Orchidaceae is one of the most diverse families of flowering plants. In New Zealand (NZ), this family is represented by 27 genera and 117 species, of which 82 are endemic to NZ (Allan Herbarium 2025). One particularly diverse genus present in NZ is *Thelymitra* J.R.Forst. and G.Forst. (*Diurideae*; *Orchidoideae*), commonly known as sun orchids. This genus is distributed throughout Australasia and is comprised of approximately 120 species (POWO 2024; WFO 2024), although the centre of diversity for *Thelymitra* is in Australia (Nauheimer et al. 2018). In New Zealand, on the other hand, 15 species and one sterile hybrid are included in the most current checklist of NZ vascular plants (Allan Herbarium 2025). Several entities of uncertain

CONTACT Hayden R. Jones hjones3@massey.ac.nz; haydenrjones@hotmail.co.nz
Supplemental data for this article can be accessed online at https://doi.org/10.1080/0028825X.2025.2523007.

taxonomic status have also been reported in NZ over the years (e.g. Rolfe and de Lange 2010; de Lange et al. 2024).

Thelymitra is unique amongst orchids as it produces flowers that typically lack a modified labellum (e.g. its shape and size is similar to that of the other petals and sepals), and as a result their flowers are nearly radially symmetrical (Figure 1A). Another distinctive feature of *Thelymitra* flowers is the mitra (Figure 1A–C); a structure found within the column formed by the fusion of the staminodes, style, and anthers (Burns-Balogh and Bernhardt 1988). This structure is also present in *Calochilus* R.Br. and *Epiblema* R.Br., two sister genera also in the subtribe *Thelymitrineae*. The mitra is one of the most taxonomically informative characters in this genus (Moore 1968; Rolfe and de Lange 2010; Jones and Lehnebach 2022) and has been the basis for many species circumscriptions. The diversity of shapes of this structure in NZ *Thelymitra* is shown in Figure 2.

The earliest attempts at reconstructing the phylogeny of *Thelymitra* were based on morphological characters (Burns-Balogh and Funk 1986; Dressler 1986; Burns-Balogh and Bernhardt 1988; Freudenstein and Rasmussen 1999). The most comprehensive work was by Burns-Balogh and Bernhardt (1988), who assessed 32 of the 40 *Thelymitra* species described at the time. Their study concluded that based on the presence or absence of certain mitra characters eight discrete species groups could be recognised. These authors also identified several possible hybrids between these groups. Molecular phylogenies that include *Thelymitra*, on the other hand, have often been very broad in scope, typically aiming to reconstruct the relationships across the major subfamilies within Orchidaceae (Kores et al. 1997; Cameron et al. 1999; Zhang et al. 2023; Pérez-Escobar et al. 2024), or within the tribe Diurideae (Kores et al. 2000; Kores et al. 2001; Clements et al. 2002; Weston et al. 2014; Wong and Peakall 2022; Zhang et al. 2023; Pérez-Escobar et al. 2024). While these studies have served to confirm the placement of *Thelymitra* in the tribe *Diurideae*, they have included only a limited number of *Thelymitra* species from AU, and thus offer very little insight into the diversity and infrageneric relationships within the genus, particularly for NZ taxa.

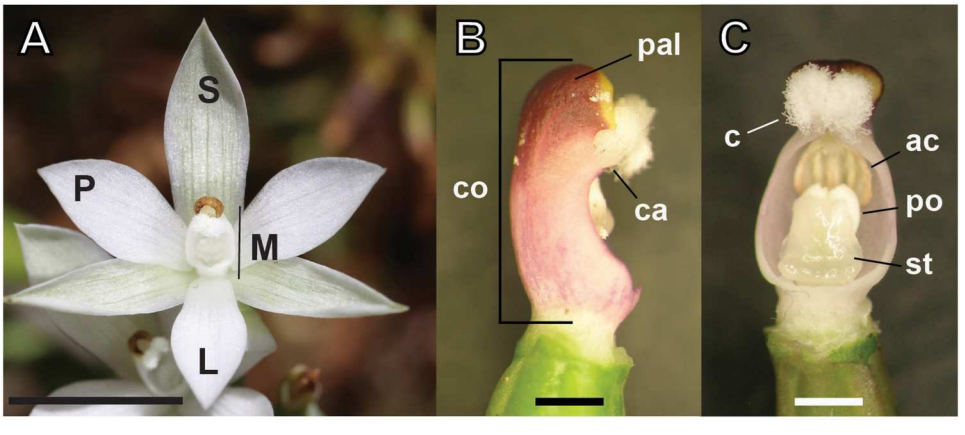

Figure 1. A flower of *Thelymitra longifolia* J.R.Forst. & G.Forst **A**, and macro photos of the column **B,C**, with the structures labelled. Structures labelled are S: Dorsal Sepal; P: Lateral Petal; L: Labellum; M: Mitra complex; co: Column; pal: Post anther lobe; ca: Column Arms; c: Cilia; ac: Anther; po: Polinia; st: Stigma. Scale bar ~1 cm.

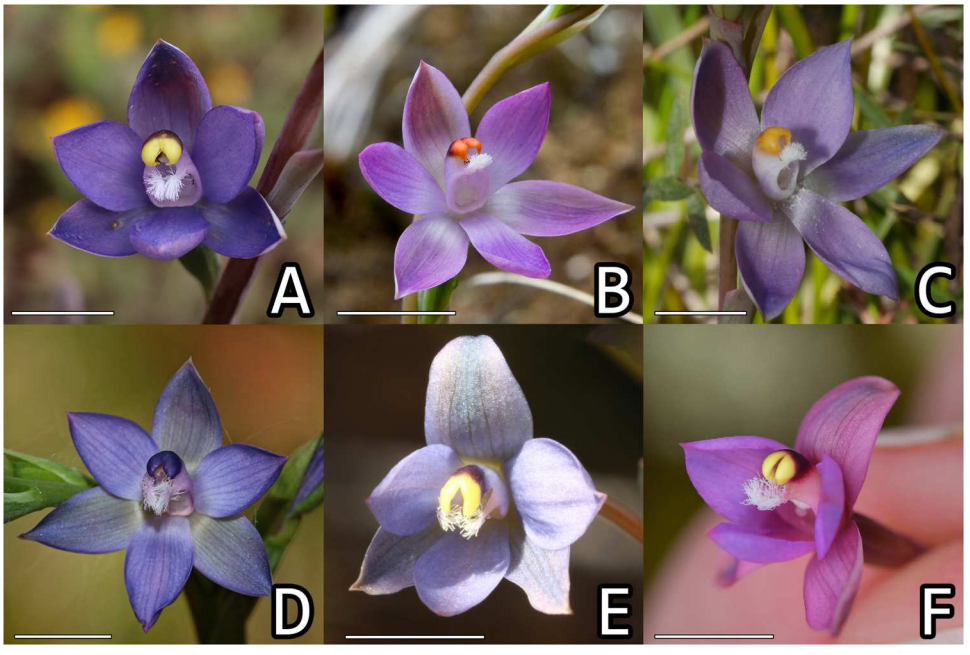

Figure 2. Examples of species in the *Thelymitra pauciflora* complex present in New Zealand. **A,** *T. pauciflora* R.Br.; **B,** *T.* aff. *pauciflora* "Orange Top"; **C,** *T. "*Rough Leaf" photo by Matt Ward; **D,** *T. cyanapicata* Jeanes photo by Kevin Matthews; **E,** *T. "*Ahipara" photo by Bill Campbell; **F,** *T.* "Darkie" photo by Christopher Stephens. Scale bar ~1 cm.

Recent work has been more comprehensive, particularly that by Nauheimer et al. (2018), who specifically focused on the phylogenetic relationships and biogeographic history of the entire genus. That study utilised one nuclear (ITS) and three plastid markers (*matK*, *psbJ-petA* and *ycf1*) and included 74 species of *Thelymitra* representing species from Australia (AU), New Caledonia (NC), and NZ. Despite the large number of species Nauheimer et al. (2018) included in their study, many Australian taxa described in the last two decades by Bates (2010) and Jeanes (2001, 2004, 2006, 2009, 2011, 2012, 2013), five species endemic to NZ (i.e. *Thelymitra colensoi* Hook.f., *T. formosa* Colenso, *T. hatchii* L.B.Moore, *T. nervosa* Colenso, *T. pulchella* Hook.f.) and a threatened undescribed entity (*T.* "Ahipara" (de Lange et al. 1991)) were not included. While this omission limits our understanding of NZ *Thelymitra*, Nauheimer et al. (2018) provides an excellent starting point for future work aiming to uncover the origin and affinities of these yet unstudied species. This phylogenetic study unravelled species relationships among *Thelymitra* with the NZ species included occurring across five clades (i.e. antennifera, aristata, pauciflora/nuda, variegata and venosa). The three NZ taxonomically undetermined taxa they included were recovered within the pauciflora/nuda clade. Nauheimer et al. (2018) stated that NZ *Thelymitra* is relatively young, and their phylogenies supported at least seven introductions of *Thelymitra* to NZ from AU, all during the Pleistocene (2.58 Ma to 11.7 Ka). Nauheimer et al. (2018) also detected incongruence between nuclear and plastid phylogenies and recovered multiple sequence copies for Internal Transcribed Spacer (ITS) sequence data for some taxa. The authors concluded these results supported previous hypotheses highlighting the importance of hybridisation

and polyploidy in the origin of *Thelymitra* species (e.g. Molloy and Dawson 1998; Peakall and Molloy 1998; Dawson et al. 2007).

The taxonomic study of NZ *Thelymitra* has lagged behind that of AU in recent years, as Australian orchid taxonomists have made changes, particularly in the *T. nuda* R.Br. and *T. pauciflora* R.Br. complexes, which are also present in NZ (Figure 2) but have not been studied in detail. Currently, two tag-named entities are included in the *T. pauciflora* aggregate; *Thelymitra* "Ahipara" (WELT SP79140) and *Thelymitra* "Darkie" (CHR 518036) (Rolfe and de Lange 2010). These entities have been ranked Threatened-Nationally Critical and At Risk – Declining, respectively, in the latest national conservation status assessment by de Lange et al. (2024). Another species complex in NZ is that of *T. longifolia* and allied taxa (Figure 3), which have remained largely unassessed by taxonomists and have a complicated nomenclatural history (see Cheeseman 1925; Hatch 1952; Allan Herbarium 2025) with the reinstated use of many synonymous names by local botanists and orchid enthusiasts (e.g. *T. purpureo-fusca* Colenso) or the creation of new tag-names (e.g. *Thelymitra longifolia* "Fusca" by the NZ Native Orchid Group). Rolfe and de Lange (2010) included *Thelymitra* aff. *longifolia* (CHR 537579, Whakapapa) and *Thelymitra* "Rough Leaf" (AK 229531) in the *T. longifolia* aggregate suggesting the latter has no close affinities in NZ. The conservation status of these two entities has been ranked as At Risk – Naturally Uncommon (de Lange et al. 2024). Furthermore, the specific details regarding type material and localities of NZ *Thelymitra*, particularly for those species described by William Colenso (1811–1899), is largely unknown. This is due in part to poor record keeping by Colenso, but also the relabelling of Colenso's herbarium by Cheeseman after Colenso's death (Lehnebach and Jones 2024).

One possible contributing factor to the complicated taxonomic interpretations of *Thelymitra* species is allopolyploidy, a phenomenon in which two species hybridise and the resultant offspring has a complete set of chromosomes from each parent (Grant 1981). This process has been hypothesised for the origin of at least five species present in NZ (Figure 4) (Molloy and Dawson 1998). However, these hypotheses have largely relied on investigations based on floral morphology and chromosome counts, with molecular work investigating this phenomenon in the genus limited to *T. nervosa* only (Peakall and Molloy 1998).

Our study aims to identify the phylogenetic affinities of those *Thelymitra* species not included in Nauheimer et al. (2018), as well as the taxonomically unresolved entities

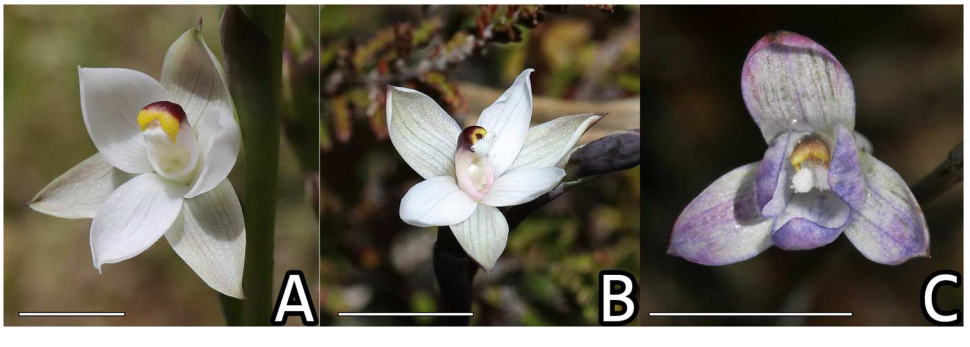

Figure 3. Examples of taxa in the *Thelymitra longifolia* complex. **A,** *T. longifolia*; **B,** *T.* aff. *longifolia* "Whakapapa"; **C,** *T.* aff. *longifolia* "Fusca" photo by Matt Ward. Scale bar ~ 1 cm.

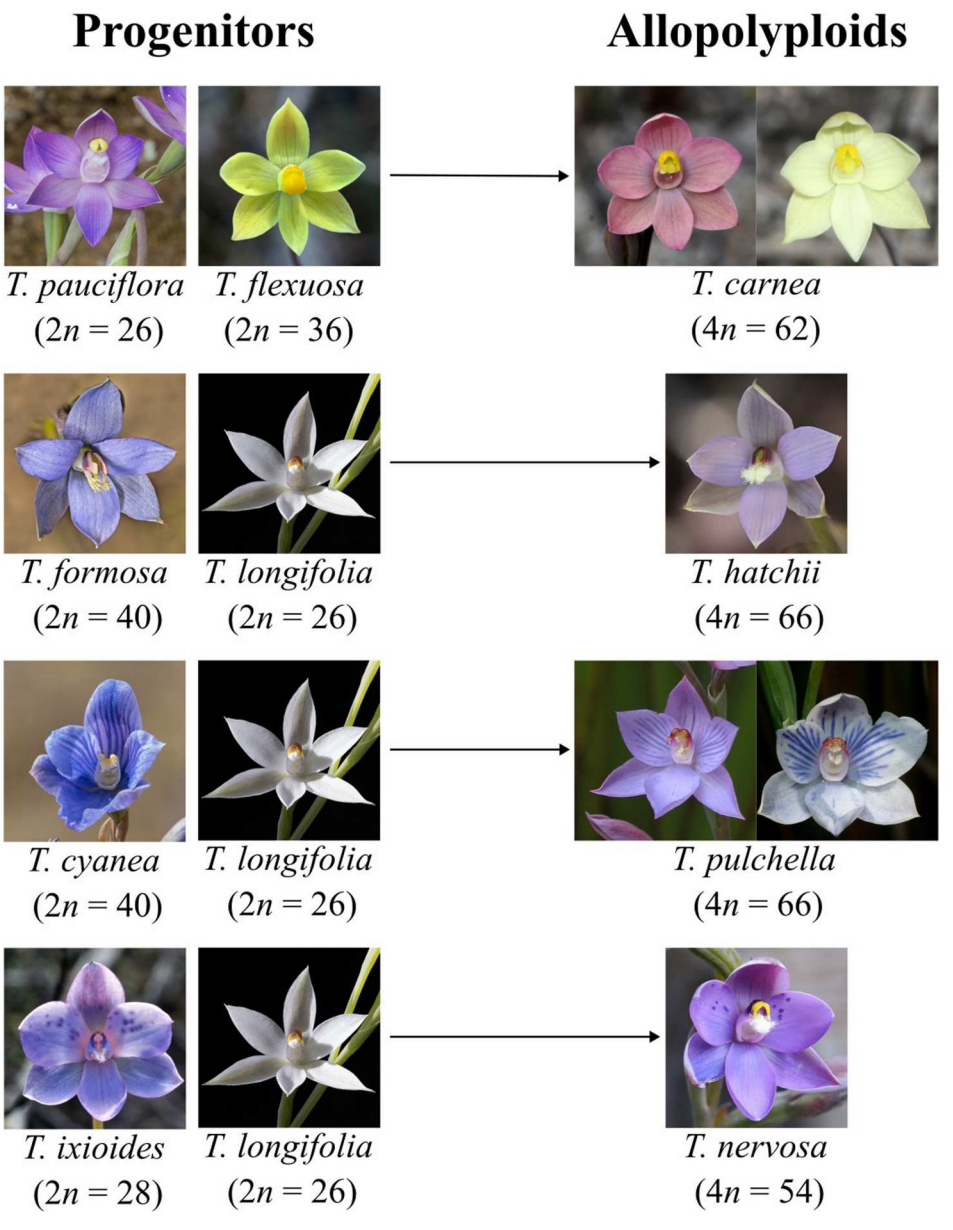

Figure 4. The putative parent and derivative allopolyploid relationships for four *Thelymitra* species proposed by Molloy and Dawson (1998) and investigated in this study. Under each image is the diploid chromosome count reported by Molloy and Dawson (1998), where the suspected polyploid species has an additive chromosome sets from both parents. (*T. carnea*, *T. ixioides* photos by Bill Campbell. *T. formosa* photo by Jeremy Rolfe. *T. flexuosa* photo by Wacrakey via iNaturalist.)

occurring in NZ. We also aim to investigate whether hybridisation and allopolyploidy have been involved in the evolution of the species present in NZ. Our results will help us to understand how tag-named entities in NZ relate to currently recognised species and their origin (e.g. local speciation, long distance dispersal, or hybridisation). This

information will aid in future work aiming to circumscribe these taxa and the implementation and prioritisation of conservation strategies by the Department of Conservation.

Methods

Taxon sampling

Fieldwork was conducted during the flowering seasons of 2020–2023, during which above ground structures (leaf, flowering stem) of *Thelymitra* plants were collected. Because species identification in *Thelymitra* is almost impossible if flowers are not present, only specimens with fertile material were included in our study. From each specimen a portion of the leaf was excised and dried using silica beads, and the remaining material lodged in herbaria (MPN or WELT, Table 1). Additional leaf material was sourced from herbarium collections at CANB, CHR, MPN, WELT and existing silica gel collections of *Thelymitra* by Carlos Lehnebach (CL) stored at WELT. Herbarium acronyms follow Thiers (2023). Additional DNA samples were sourced from the Kew DNA Bank. In total, 52 samples of *Thelymitra* were included, comprising seventeen species and six tag-named entities included in de Lange et al. (2024), and comprising all species currently recognised in NZ (Allan Herbarium 2025). Two samples of the sister genus *Calochilus* were also obtained and used as an outgroup (Nauheimer et al. 2018).

DNA extraction

DNA was extracted using the Qiagen DNeasy plant minikit (QIAGEN Inc.) following manufacturer's recommendations or an STE-CTAB method, including an overnight lysis step (Shepherd and McLay 2011). DNA was visualised on a 1% agarose gel stained with ethidium bromide using a transilluminator (Thermo Fisher Scientific).

PCR amplification and sequencing

Plastid *ycf1* and nuclear *LFY* (partial exons 1 and 2, full intron 1) regions were PCR-amplified using published primers and thermal cycling conditions (Schlüter et al. 2007; Neubig et al. 2009). The *ycf1* region was selected as previous studies have shown it to be highly informative in this genus (Nauheimer et al. 2018), and *LFY* was chosen as nuclear region not susceptible to concerted evolution (Small et al. 2004). Preliminary testing by the first author found *LFY* to be potentially informative in allopolyploid taxa. Additional primers were designed for *LFY* (Table 2) to improve the consistency of amplification as the original primers were found to form heterodimers. All PCR reactions were achieved with TaKaRa Extaq DNA Polymerase (TaKaRa, Japan) in 20 µL reactions using 1 unit of TaKaRa Taq, 10× TaKaRa PCR buffer, 0.5 mM of each primer, 0.5 mM dNTPs, 5–200 ng DNA template, and 0.05 mM DMSO. Successful PCR amplification was verified by agarose gel electrophoresis.

PCR products were purified by digestion with 0.5 U shrimp alkaline phosphatase (rSAP; New England Biolabs) and 2.5 U exonuclease I (ExoI; New England Biolabs) at 37°C for 15 min, followed by inactivation of the enzymes for 15 min at 80°C. Cleaned products were submitted to Massey Genome Service (MGS) for cycle sequencing using

Table 1. Samples of *Thelymitra* and *Calochilus* included in this study and their associated GenBank accession numbers for DNA sequence data.

Sample	Locality/source	Voucher details	*LFY*	*ycf1*
C. campestris	from Nauheimer et al. (2018)	CANB *Clements 9933*	–	MG695276
C. paludosus	from Nauheimer et al. (2018)	CANB *Clements 11442*	–	MG695278
C. robertsonii 1	New Zealand, Waikato, Iwitahi	MPN 53220	PV815893	PV815831
C. robertsonii 2	Australia, Victoria	KEW DNA 0-570	PV815856	PV815788
T. "Ahipara" 1	New Zealand, Northland, Kaimaumau	WELT SP110816	PV815901	PV815836
T. "Ahipara" 2	New Zealand, Northland, Kaimaumau	WELT SP111004	PV815902	PV815837
T. "Darkie" 1	New Zealand, Northland, Kaingaroa	MPN 53215	PV815878 / PV815879	PV815821
T. "Darkie" 2	New Zealand, Northland, Whatuwhuwhu	MPN 53216	PV815880 / PV815881	PV815822
T. "Darkie" 3	New Zealand, Northland, Whatuwhuwhu	WELT SP119843	–	PV815802
T. "Rough Leaf" 1	New Zealand, Northland, Ahipara	MPN 53217	PV815886	PV815826
T. "Rough Leaf" 2	New Zealand, Northland, Kaimaumau	WELT SP119844	–	PV815804
T. aemula 1	New Zealand, Northland, Whatuwhuwhu	MPN 53199	PV815877	PV815820
T. aemula 2	New Zealand, Northland, Ahipara	MPN 53200	PV815883	PV815824
T. aemula 3	from Nauheimer et al. (2018)	CHR *Molloy 126/99*	–	MG695295
T. aff. *brevifolia*	New Zealand, Northland, Otangaroa Forest	WELT SP110984	PV815900	PV815834
T. aff. *holmesii*	Australia, Western Australia	KEW DNA S-5547	PV815903	PV815838
T. aff. *longifolia* "Fusca" 1	New Zealand, Otago, Queenstown	WELT SP123375	PV815896	PV815815
T. aff. *longifolia* "Fusca" 2	New Zealand, Canterbury, Wharfdale	MPN	–	PV815828
T. aff. *longifolia* "Whakapapa" 1	New Zealand, Manawatu-Whanganui, Ohakune	MPN 53218	PV815864	PV815817
T. aff. *longifolia* "Whakapapa" 2	New Zealand, Manawatu-Whanganui, Ohakune	MPN 53219	–	PV815799
***T.* aff. *longifolia* "Whakapapa" 3**	from Nauheimer et al. (2018)	CHR *Molloy 245/00*	–	MG695381
T. aff. *pauciflora*	Australia, Western Australia	KEW DNA S-5549	PV815904	PV815839
T. aff. *pauciflora* "Orange Top"	New Zealand, Bay of Plenty, Katikati	MPN 52487	–	PV815803
T. albiflora	from Nauheimer et al. (2018)	CANB *Jones 18012*	–	MG695297
T. alpicola	from Nauheimer et al. (2018)	MEL *Jeanes 938*	–	MG695301
T. antennifera 1	Australia, Western Australia, Windinie Hill	CANB 605270	PV815907 / PV815908	PV815841
T. antennifera 2	from Nauheimer et al. (2018)	CANB *Clements 10408*	–	MG695303
T. arenaria	from Nauheimer et al. (2018)	MEL *Alcock 889*	–	MG695304
T. carnea 1	New Zealand, Northland, Whatuwhuwhu	MPN 53201	PV815875 / PV815876	PV815819
T. carnea 2	New Zealand, Northland, Ahipara	MPN 53202	PV815884 / PV815885	PV815825
T. carnea 3	from Nauheimer et al. (2018)	MEL *Rouse 756*	–	MG695313
T. circumsepta	from Nauheimer et al. (2018)	CHR *Molloy 048/98*	–	MG695314
T. colensoi 1	New Zealand, Wellington, Days Bay	WELT SP119833	PV815866	PV815807
T. colensoi 2	New Zealand, Wellington, Kaitoke	MPN 52484	PV815859 / PV815860	PV815794
T. colensoi 3	New Zealand, Manawatu-Whanganui, Palmerston North	MPN 53203	–	PV815806
T. cyanapicata	New Zealand, Northland, Pukenui	MPN 53221	PV815899	PV815833
T. cyanea 1	New Zealand, Waikato, Whangamarino	WELT SP119892	PV815861	PV815795
T. flexuosa	from Nauheimer et al. (2018)	CANB *Clements 11971B*	–	MG695389
T. formosa	New Zealand, Waikato, Whangamarino	WELT SP119890	PV815871	PV815814
T. hatchii 1	New Zealand, Bay of Plenty, Tahawai	No voucher	PV815869 / PV815870	PV815812

(*Continued*)

Table 1. Continued.

Sample	Locality/source	Voucher details	*LFY*	*ycf1*
T. hatchii 2	New Zealand, Wellington, Mt. Holdsworth	WELT *Brooks s.n.*	PV815873 / PV815874	PV815835
T. ixioides 1	New Zealand, Northland, Ahipara	MPN 53205	PV815882	PV815823
T. ixioides 2	from Nauheimer et al. (2018)	MEL *Jeanes 893*	–	MG695340
T. longifolia 1	New Zealand, Nelson, Nelson lakes	WELT SP123379	PV815897	PV815816
T. longifolia 2	New Zealand, Northland, Kaimaumau	WELT SP119834	PV815863	PV815797
T. longifolia 3	New Zealand, Canterbury, Port Hills	MPN 53208	PV815888	PV815829
T. longifolia 4	New Zealand, Manawatu-Whanganui, Palmerston North	MPN 53206	PV815865	PV815805
T. longifolia 5	New Zealand, Wellington, Days Bay	MPN 52473	–	PV815790
T. longifolia 6	New Zealand, Wellington, Days Bay	MPN 52469	PV815857	PV815789
T. longifolia 7	New Zealand, Wellington, Days Bay	MPN 52476	–	PV815791
T. longifolia 8	New Zealand, Tasman, Takaka	WELT SP123419	–	PV815801
T. longifolia 9	New Zealand, Northland, Kaimaumau	WELT SP119834	–	PV815792
T. longifolia 10	New Zealand, Tasman, Takaka	WELT SP123377	–	PV815800
T. longifolia 11	New Zealand, Waikato, Karangahake	MPN	–	PV815798
T. longifolia 12	from Nauheimer et al. (2018)	CHR *Molloy 098/99*	–	MG695346
T. longiloba	from Nauheimer et al. (2018)	MEL *Branwhite1026*	–	MG695347
T. macrophylla	from Nauheimer et al. (2018)	CANB *Clements2387*	–	MG695348
T. malvina 1	from Nauheimer et al. (2018)	MEL *Jeanes 1192*	–	MG695352
T. malvina 2	from Nauheimer et al. (2018)	CANB *Clements 11694*	–	MG695350
T. matthewsii	from Nauheimer et al. (2018)	CHR *Molloy 189/00*	–	MG695353
T. media	from Nauheimer et al. (2018)	MEL *Rouse 890*	–	MG695354
T. mucida	from Nauheimer et al. (2018)	CANB *Clements 10670*	–	MG695357
T. nervosa 1	New Zealand, Wellington, Days Bay	MPN 52472	PV815894 / PV815895	PV815813
T. nervosa 2	New Zealand, Waikato, Iwitahi	MPN 53209	PV815889 / PV815890	PV815830
T. nervosa 3	New Zealand, Waikato, Iwitahi	MPN 53210	PV815891 / PV815892	PV815810
T. nuda	from Nauheimer et al. (2018)	CANB *Clements 11883*	–	MG695361
T. pauciflora 1	New Zealand, Bay of Plenty, Katikati	MPN 53211	PV815867	PV815808
T. pauciflora 2	New Zealand, Northland, Cooper's Beach	WELT SP119861	PV815858	PV815793
T. pauciflora 4	New Zealand, Bay of Plenty, Golden Cross	MPN 53213	PV815898	PV815832
T. pauciflora 5	New Zealand, Manawatu-Whanganui, Palmerston North	MPN 52486	PV815862	PV815796
T. pauciflora 6	New Zealand, Bay of Plenty, Katikati	MPN 53212	–	PV815809
T. pauciflora 7	from Nauheimer et al. (2018)	CANB *Clements 12052*	–	MG695364
T. pulchella 1	New Zealand, Auckland, Waikumete	WELT SP 123374	PV815868	PV815811
T. pulchella 2	New Zealand, Northland	MPN 53214	PV815872	PV815818
T. pulchella 3	New Zealand, Wellington, Mt. Holdsworth	WELT *Brooks s.n.*	PV815905 / PV815906	PV815840
T. queenslandica	from Nauheimer et al. (2018)	CNS *Micheneau 004*	–	MG695370
T. rubra	from Nauheimer et al. (2018)	CANB *Clements 11679*	–	MG695371
T. sanscilia 1	New Zealand, Northland	No voucher	PV815887	PV815827
T. sanscilia 2	from Nauheimer et al. (2018)	CHR *Molloy 212/00*	–	MG695372
T. sarasiniana	from Nauheimer et al. (2018)	CANB *Clements 11233*	–	MG695374
T. tholiformis	from Nauheimer et al. (2018)	CHR *Molloy 127/99*	–	MG695390
T. venosa	from Nauheimer et al. (2018)	CHR *Molloy 058/98*	–	MG695397

Supplementary sequences generated by Nauheimer et al. (2018) are indicated in bold along with their voucher's details (herbarium and collector's last name and number). CANB: Australian National Herbarium, CHR: Allan Herbarium, CNS: Australian Tropical Herbarium, KEW DNA: Kew DNA bank. MEL: National Herbarium of Victoria. MPN: Massey University Herbarium. WELT: Te Papa Tongarewa Herbarium.

Table 2. Chloroplast and nuclear markers used in this study.

Marker	Primer name	Sequence (5′–3′)	Source
ycf1	3720F	TACGTATGTAATGAACGAATGG	Neubig et al. (2009)
	IntR	TTTGATTGGGATGATCCAAGG	Neubig et al. (2009)
LFY	E1Cf	ATGGTGCTGGCCACATCGCAGCAACA	Schlüter et al. (2007)
	E2Gr	GAAGAGGTAATCGAGCCCGTTCTTCTTAGCYC	Schlüter et al. (2007)
	LFY-29F	CTCCGTCACGATGAACGGGT	This study
	LFY-1437R	ATGGCGCGGATTGGGGATT	This study

Primer name, sequences and annealing temperature are indicated for each. Sources of previously published primers are also indicated.

the amplification primers and ABI BigDye v. 3.1 terminator chemistry and run on an ABI 3730 sequencer (Thermo Fisher Scientific).

Because some species were known or suspected to be allopolyploids (Molloy and Dawson 1998) they would be unsuitable for Sanger di-deoxy sequencing due to the presence of multiple gene copies which would be collapsed into a single chimeric sequence, and unusable for analysis. PCR products of *LFY* were instead sequenced through Oxford Nanopore Sequencing (ONT). In this case, *LFY* PCR products were cleaned using a ZYMO DNA Clean & Concentrator kit (Zymo Research) and libraries prepared with the Native Barcoding kit (SQK-NBD114.24; Oxford Nanopore Technologies) following the Ligation sequencing amplicons procedure by Oxford Nanopore (https://nanoporetech.com/document/ligation-sequencing-amplicons-sqk-lsk114). Prepared libraries were sequenced using a MinION Mk1C and flongle flow cells (Oxford Nanopore Technology). For both *ycf1* and *LFY*, newly generated sequences were deposited to GenBank (Table 1).

Data processing

Forward and reverse sequences from Sanger sequencing data for *ycf1* amplicons were assembled de novo in Geneious Prime 2023.3.1. (http://www.geneious.com/) using the Geneious assembler (Medium Sensitivity). Once all sequences were assembled, additional sequences of *ycf1* used by Nauheimer et al. (2018) were obtained from NCBI GenBank (Table 1), where available, to increase coverage for endemic NZ and AU species. These sequences were then aligned in Geneious using the Geneious aligner with a 65% cost matrix and then trimmed to ensure all sequences were the same length.

Nanopore reads of *LFY* were base-called using Dorado v0.7.2 with the sup v5 quality model (https://github.com/nanoporetech/dorado), and mapped against a reference sequence of *LFY* obtained from previous Sanger sequencing of *LFY* from *T. longifolia* using Minimap2 v2.28 (Li 2018) in Map-ont mode. Mapped reads were selected using samtools view and clustered using MeShClust3 (Girgis 2022). MeShClust takes in reads and outputs a list of similar sequences based on pairwise distances, the output of which was parsed using a custom script (stripSeqs.py: https://github.com/HaydenRJones/sequence-tools) to produce a fasta file of reads for each cluster identified. In the cases of diploid taxa only one cluster was identified, but for suspected allopolyploids two clusters were produced. Each cluster of reads were aligned using MAFFT v7.526 and the FFT-NS-2 algorithm (Katoh and Standley 2013), from which a consensus was then made using EMBOSS (Rice et al. 2000). Finally all consensus sequences were merged into a single fasta file and aligned again using MAFFT using the L-INS-I

algorithm. This alignment of consensus sequences was then trimmed in Geneious to ensure all sequences were the same length.

Data analysis

Model testing for phylogenetic analysis was achieved with ModelTest-NG (Flouri et al. 2015; Darriba et al. 2020). Models were filtered based on their availability in software packages used downstream (e.g. RAxML-NG, MrBayes) and selected on a per dataset basis using the highest Akaike information criterion score (AIC) following Nauheimer et al. (2018). RAxML-NG (Kozlov et al. 2019) was used for maximum likelihood analysis using the models identified previously. Support was calculated using 1000 bootstrap replicates. Bayesian phylogenetic inference was conducted using MrBayes 3.2.7a (Ronquist et al. 2012) with the following settings: two runs, three heated chains and two million generations with a burn-in of 25%. Convergence of runs was verified using Tracer (Rambaut et al. 2018) based on the average standard deviation of split frequencies being less than 0.01 (*LFY:* avg. std. deviation of splits = 0.004069; *ycf1:* avg. std. deviation of splits = 0.004868).

Results

Sequencing

The *ycf1* alignment included sequences from 82 accessions with a sequence length of 648 bp. Of these 82 sequences, 54 were newly generated and 28 were sourced from GenBank (Table 1). Eleven accessions were found to have multiple distinct copies of *LFY* (Table 3), thus the alignment used in this study comprised 52 sequences from 40 accessions, and had a sequence length of 1364 bp. The best fit substitution model for both datasets was GTR + G as identified by ModelTest-NG (Darriba et al. 2020) based on the highest AIC scores.

Phylogenetic analysis

For each marker, both methods of phylogenetic inference (maximum likelihood and Bayesian) produced trees with the same topology. The topologies differed however

Table 3. *Thelymitra* accessions with multiple distinct copies of *LFY* and the closest clade when included in a phylogeny.

Sample	Maternal copy ♀	Paternal copy ♂
T. carnea 1	Pauciflora	Antennifera
T. carnea 2	Pauciflora	Antennifera
T. colensoi 2	aff. Brevifolia	Unknown
T. hatchii 1	Formosa	Longifolia
T. hatchii 2	Formosa	Pauciflora
T. nervosa 1	Ixioides	Longifolia
T. nervosa 2	Ixioides	Longifolia
T. nervosa 3	Ixioides	Longifolia
T. "Darkie" 1	Unknown	Pauciflora
T. "Darkie" 2	Unknown	Pauciflora
T. *pulchella* 2	Cyanea	Pauciflora

Maternal inheritance is inferred based on congruence with the plastid *ycf1* phylogeny.

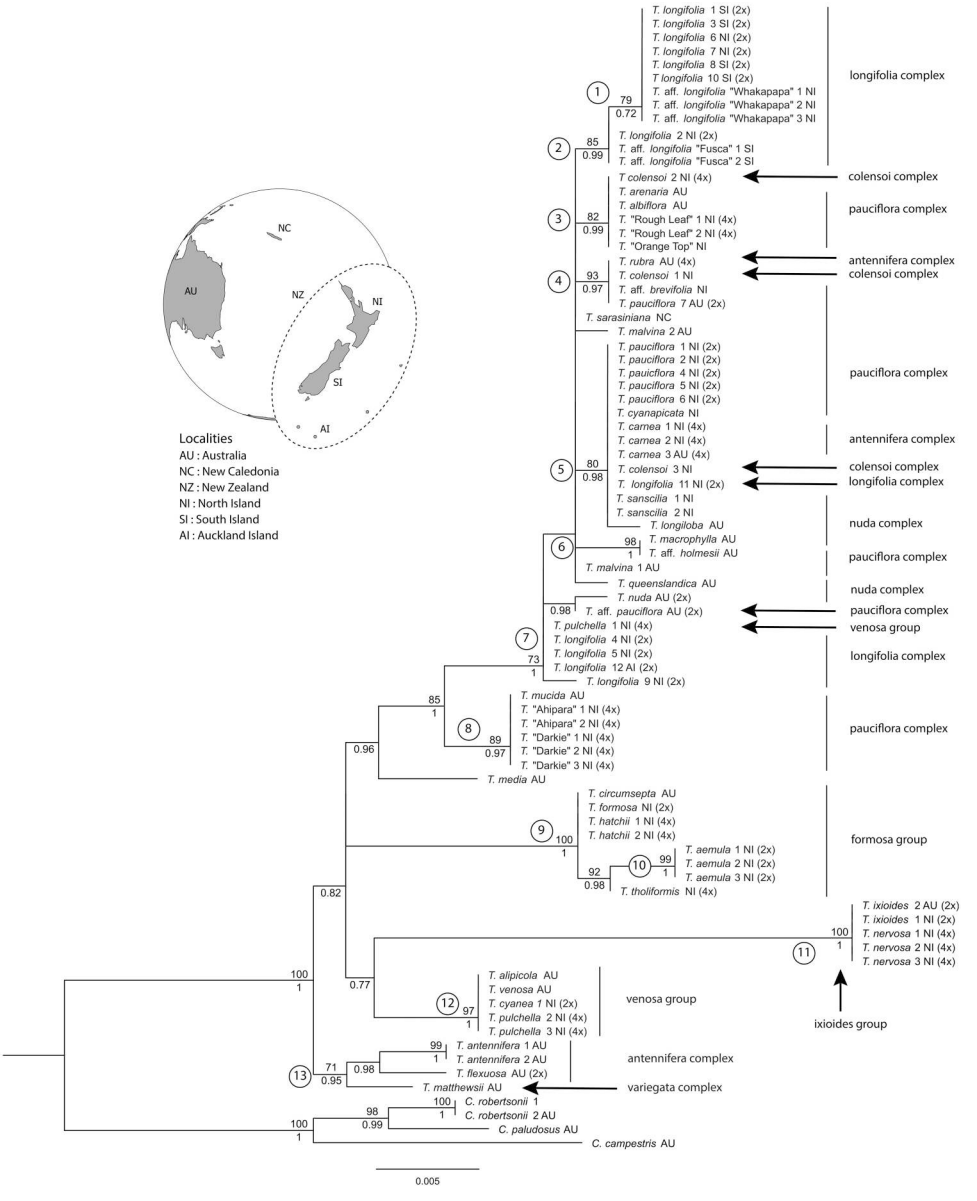

Figure 5. Maximum likelihood phylogeny of *Thelymitra* based on *ycf1* sequences. Bootstrap support is reported above the branches and posterior probability from Bayesian inference is below the branches. Clades are labelled by species group and geographic location is reported by accession following taxon name. Known or suspected allopolyploid taxa are labelled 4*x* and diploid taxa labelled 2x.

between the nuclear and plastid regions (a tanglegram comparing topologies is included as supplementary Figure 1). New Zealand *Thelymitra* were resolved into six and five species groups in the *ycf1* and *LFY* phylogenies, respectively (Figures 5 and 6). For some taxa, incongruence was identified between the phylogenies produced for each marker while other taxa had multiple sequence copies for *LFY* (Table 3). Parentage for these copies was inferred based on congruence with the *ycf1* phylogeny, representing

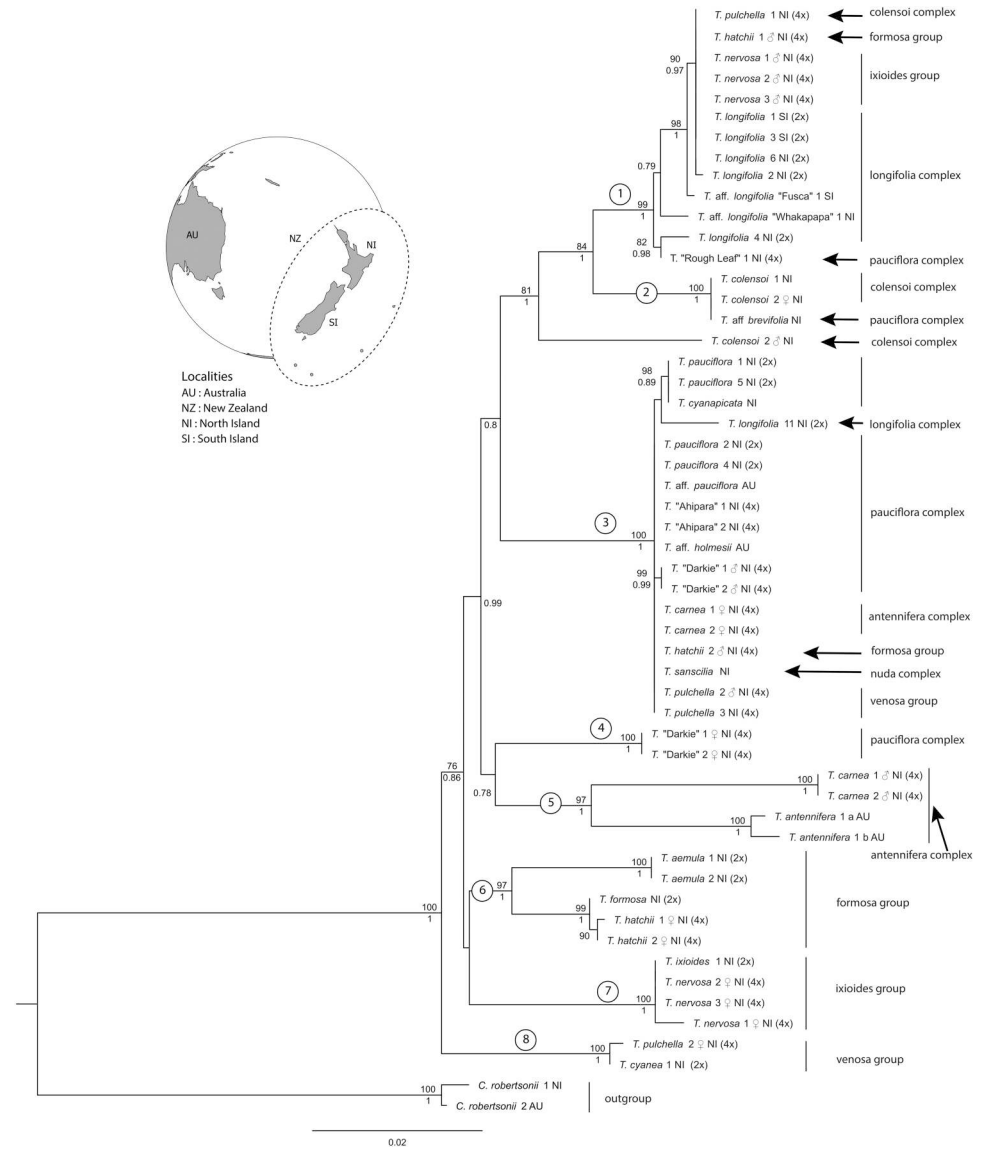

Figure 6. Maximum likelihood phylogeny of *Thelymitra* based on *LFY* sequences. Bootstrap support is reported above the branches and posterior probability from Bayesian inference is below the branches. Clades are labelled by species group and geographic location is reported by accession following taxon name. Maternal ♀/paternal ♂ inheritance for allopolyploid taxa is inferred from plastid relationships when multiple distinct *LFY* sequences were obtained. Known or suspected allopolyploid taxa are labelled 4*x* and diploid taxa labelled 2*x*.

the maternal lineages as the chloroplast is maternally inherited in orchids (Goldschmidt-Clermont 2001).

Ycf1 phylogeny

In the *ycf1* phylogeny, NZ *Thelymitra* was resolved into six species groups and thirteen distinct clades (Figure 5). These groups and clades were consistent with those reported by

Nauheimer et al. (2018). Thirteen NZ species were contained within a large clade mostly comprised of taxa from the *nuda* and *pauciflora* complexes (Figure 5; node 7). This clade also included *T. carnea* and a single accession of *T. pulchella*, which have not been associated with this complex before. The remaining nine NZ species were contained in five clades. Sequence variation within clades was usually low with most accessions contained within them having identical sequences.

***Affinities of newly sequenced taxa for* ycf1.** The newly sequences species had affinities with three species groups and one complex. For instance, two accessions of *T. pulchella* were included within a clade comprised entirely of *venosa* group species (i.e. *T. cyanea*, *T. venosa*) (Figure 5, node 12). This clade was well supported (Bootstrap support (BS): 97, Posterior probability (PP): 1) and all accessions within it had identical sequences. A third accession of *T. pulchella* (*T. pulchella* 1) was within the *pauciflora* and *nuda* complexes clade and was identical in sequence to three accessions of *T. longifolia* (Figure 5, node 7). All accessions of *T. nervosa* resolved into the *ixioides* group (Figure 5, node 11); a well-supported clade (BS: 100, PP: 1) containing two accessions of *T. ixioides* from AU and NZ. All accessions in this clade had identical sequences. Both accessions of *T. hatchii*, as well as one of *T. formosa* were in a well-supported clade (BS: 100, PP: 1) that included members of the *formosa* group (i.e. *T. aemula*, *T. circumsepta*, and *T. tholiformis*) (Figure 5, node 9). Sequences of *T. hatchii* and *T. formosa* were identical to that of *T. circumcepta*, an Australian endemic.

Thelymitra colensoi was polyphyletic with sequences from all three accessions being different from one another (Figure 5, node 3, 4 & 5), but all within the *pauciflora* and *nuda* complexes clade. The sequence of *T. colensoi* 1 was identical to an Australian accession of *T. pauciflora* and a NZ accession of *T.* aff. *brevifolia* (BS: 93, PP: 0.97) (Figure 5, node 4). The sample *T. colensoi* 2 was the same as the Australian species *T. albiflora* Jeanes and *T. arenaria* Lindl., as well as the undescribed NZ entity *T.* "Rough Leaf" (Figure 5, node 3; BS: 93, PP: 0.97). The last sample of this species, *T. colensoi* 3 fell within a clade that included NZ accessions of *T. pauciflora* (Figure 5, node 5). This clade had a BS: 80 and PP: 0.98.

***Resolution of species complexes based on* ycf1.** The *nuda* and *pauciflora* complexes were found to be paraphyletic. New Zealand accessions from these complexes resolve into six well supported clades. One subgroup was comprised entirely of accessions of *T.* "Darkie" (CHR 518036) and *T.* "Ahipara" (WELT SP79140) as well as a single accession of *T. mucida* Fitzg. from Australia and was sister to the rest of the complex (Figure 5, node 8; BS: 87, PP: 0.97). A second subgroup contained all NZ accessions of *T. pauciflora* (Figure 5, node 5; BS: 80, PP: 0.98), however these differed from AU accessions of the same species (Figure 5, node 4). This subgroup also contained *T. sanscilia* and *T. cyanapicata*. A third subgroup (Figure 5, node 3) contained two undescribed taxa *T.* "Rough leaf" (de Lange et al. 2024) and *T.* aff. *pauciflora* "Orange top" (as *T.* aff. *brevifolia* in de Lange et al. 2024), as well as two AU species *T. albiflora* and *T. arenaria*. Support for this group are BS: 82 and PP: 0.99. *Thelymitra longifolia* was found to be paraphyletic. Accessions of this species resolved into two of the subgroups (Figure 5, node 7), one similar in sequence to AU *T.* aff. *pauciflora* and was comprised of plants collected from the North Island. The second group contained plants from

the North and South Islands of NZ (Figure 5, nodes 1 and 2; BS: 85, PP: 0.99), as well as the tag named forms: *T.* aff. *longifolia* "Whakapapa" (de Lange et al. 2024) and *T.* aff. *longifolia* "Fusca". The *pauciflora* and *nuda* complexes clade also contained species from the *antennifera* complex: both AU and NZ accessions of *T. carnea* which were in the same subgroup as NZ accessions of *T. pauciflora*, and *T. rubra* which was with the AU accession of *T. pauciflora* (Figure 5, node 4; BS: 93, PP: 0.97).

LFY phylogeny

For *LFY,* both phylogenetic methods produced trees with the same topology, however, these differed from the plastid phylogenies. NZ species resolved into seven clades (Figure 6), with six of these similar to the species groups recovered by Nauheimer et al. (2018) in their nuclear phylogeny. These groups formed clades with high support both for Maximum likelihood and Bayesian analysis (generally >97/0.99), although support at splits deeper in the tree were usually low, such as that between the *longifolia* and *pauciflora* groups (Figure 6, nodes 1 v/s 3; BS: 33, PP: 0.8) and between the *formosa* and *ixioides* groups (Figure 6, nodes 6 v/s 7; BS: 30, PP: 0.61).

Recovery of multiple sequence copies and evidence for hybridisation. Two diverged copies of *LFY* were recovered for 11 samples covering five species and one tag-named entity (Table 3). These were from accessions of *T. carnea*, *T. colensoi*, *T. hatchii*, *T. nervosa*, *T.* "Darkie", and *T. pulchella*. For all accessions of *T. carnea* the inferred paternal copy of *LFY* was within the *pauciflora* clade (Figure 6, node 5; BS: 100, PP: 1). The other copy did not correspond to anything sequenced for this project and was in its own clade (Figure 6, node 5; BS: 100, PP: 1); however, it was most similar to the sequences obtained from *T. antennifera*. For all accessions of *T. nervosa*, the inferred paternal copy of the sequence was resolved into the *longifolia* clade (Figure 6, node 1; BS: 90, PP: 0.97) and the inferred maternal copy with the *ixioides* group (Figure 6, node 7; BS: 100, PP: 1). For both accessions of *T. hatchii*, the inferred maternal copy of *LFY* was in the *formosa* group (Figure 6, node 6; BS: 90, PP: 0.67) although it differed from *T. formosa*, and the inferred paternal copy differed between accessions. In accession 1, the paternal copy was the same as those recovered from *T. longifolia* (Figure 6, node 1), and for accession 2 it was with *T. pauciflora* (Figure 6, node 3). For *T.* "Darkie" the paternal copy of *LFY* was most similar to *T. pauciflora* (Figure 6, node 3), however was different in sequence by two base pairs¸ and the maternal copy resolved into its own clade on the tree (Figure 6, node 4; BS: 100, PP: 1). This clade was sister to the clades containing *T. antennifera* and the inferred paternal copies of *T. carnea* (Figure 6, node 5). All three accessions of *T. pulchella* differed. For accessions 1 and 3, only a single copy of *LFY* was recovered with these resolving into the *longifolia* (Figure 6, node 1) and *pauciflora* (Figure 6, node 3) clades, respectively. Two copies of *LFY* were recovered for *T. pulchella* 2 with the inferred maternal copy in the *venosa* group (Figure 6, node 8) and the inferred paternal copy with *T. pauciflora*. (Figure 6, node 3). Multiple copies of *LFY* were recovered for *T. colensoi 2*. One of these was with another accession of *T. colensoi* (Figure 6, node 2) and *T.* aff. *brevifolia*, and the other copy was unique and was outside this clade.

***Resolution of species complexes based on* LFY.** Most *pauciflora* and *nuda* complexes associated taxa were placed into a polytomy with high support (Figure 6, node 3; BS: 100, PP: 1), including most NZ accessions of *T. pauciflora*, *T. sanscilia*, *T.* "Ahipara", and AU accessions of *T. pauciflora*, and *T.* aff. *holmesii*. Two NZ accessions of *T. pauciflora* and *T. cyanapicata* differed by two base pairs in their sequences and resolved into their own clade within this larger clade (BS: 98, PP: 0.89). Multiple sequence copies were recovered for accessions of *T.* "Darkie", one copy differed by three base pairs from *T. pauciflora* and is nested within taxa in node 3 (Figure 6, BS: 99, PP: 0.99), and the other is resolved into its own clade (Figure 6, node 4). Samples of *T.* aff. *brevifolia* and *T. colensoi* differed significantly from the rest of this complex (with a 45 bp indel) and resolved into a separate clade sister to the *longifolia* complex (Figure 6, node 2; BS: 84, PP: 1). *Thelymitra longifolia* was polyphyletic. Most samples resolved into a single clade (Figure 6, node 1) with variation for accessions of *T.* aff. *longifolia* "Whakapapa" (BS: 57, PP: 0.79) and *T.* aff. *longifolia* "Fusca" (BS: 98, PP: 1). One accession of *T. longifolia* from the Coromandel – northern North Island, resolved into the *pauciflora* group clade (Figure 6, node 3).

Discussion

Our study aimed to reconstruct the phylogenetic relationships of NZ species of *Thelymitra* and several taxonomically unassessed entities occurring in the country. We also investigated the contribution of processes such as hybridisation and polyploidy in their evolution. We were able to produce the most complete phylogeny of NZ *Thelymitra* to date. This study included eight newly sequenced species and several tag named entities. When supplementing our study with data from the thirteen other NZ species, this renders our *ycf1* phylogeny inclusive of all currently recognised species and tag named entities in NZ (Allan Herbarium 2025). The *LFY* phylogeny had lower coverage, including only nineteen taxa but these comprised all NZ endemic species of *Thelymitra*. For *LFY* the two missing species were *T. matthewsii* Cheeseman and *T. malvina* M.A.Clem, D.L.Jones & Molloy, however accessions of these taxa were included in the *ycf1* phylogeny, and nuclear amplicons from these species have been assessed previously by Nauheimer et al. (2018) so their phylogenetic affinities are well-known.

Our findings suggest that at least four taxa currently known by tag names only, and reportedly morphologically distinct, are also genetically distinct, enabling future work to focus on their circumscription and conservation where necessary. Multiple copies of *LFY* were recovered for six taxa (*T. colensoi*, *T. carnea*, *T. hatchii*, *T. pulchella*, *T. nervosa*, *T.* "Darkie"), indicating a potential allopolyploid origin. This origin has been previously hypothesised for some species in this genus, however this is the first time it has been demonstrated for all of these species but *T. nervosa*.

LFY was found to be a suitable molecular marker for reconstructing the phylogenetic relationships in this genus, however it has limited resolution within the *pauciflora* and *nuda* complexes clade. The amplicon used in this study only covered the first and second exons and the intervening first intron, and so it is possible that sequences covering more of this gene would be more informative. Nanopore sequencing proved to be a suitable approach for sequencing this marker due to its long length of >1400 bp, and the single molecule sequencing enabled by this platform facilitated phasing of different amplicon copies in allopolyploid taxa helping to improve its phylogenetic utility as a marker (Rothfels 2021).

Taxonomic implications of newly sequenced species

Several NZ species of *Thelymitra* sequenced for this study were identical to Australian species, such as *T. cyanea* and *T. venosa*, or *T. formosa* and *T. circumcepta*. In the case of *T. cyanea* and *T. venosa*, the current taxonomic treatment used for these species is that by Jeanes (2012). A thorough investigation of the nomenclature for these species is recommended, as it appears that the application of *T. cyanea* by Jeanes (2012) differed from that described by Lindley in 1840. The current concept of *T. cyanea* is very similar in morphology to *T. venosa*, thus it seems plausible that the current application of the name *T. cyanea* in NZ may be erroneous, and *T. venosa* may be a better fit for the plants included here as was the case historically, see Cheeseman (1925); Hatch (1952); Moore (1968). Synonymy between *T. formosa* and *T. circumcepta* also requires further investigation, however these plants are highly similar in both morphology and habitat and are good candidates for inclusion in future molecular and taxonomic work to clarify if these are different species or not.

Thelymitra colensoi was found to be polyphyletic for both markers investigated, with evidence for a potential allopolyploid origin identified for a single accession based on the presence of multiple copies of *LFY*. The nomenclature and specific application of this name is complicated, with historical treatments by Cheeseman (1925), Hatch (1952), and Moore (1968) opting not to recognise it. Today this name is applied loosely to small purple flowered species with brown arched columns, with pale yellow and deeply cleft post anther lobes (Rolfe and de Lange 2010). To complicate things further, the name *T. intermedia* is also sometimes used interchangeably with *T. colensoi* (Rolfe and de Lange 2010), despite this name being placed in synonymy with *T. tholiformis* by Molloy and Hatch (1990), and the historical concept of this name differing greatly from that of *T. colensoi* (Moore 1968). Given the nomenclatural confusion surrounding this name, it is maybe no surprise that it may cover multiple taxa. However, similar to *T. venosa* and *T. cyanea*, further genetic assessments and study of historical collections are needed to better understand both the taxonomy and clarify the nomenclature associated with this species concept.

Many of the undescribed entities known by tag names included in this work resolved differently from the currently described species they have normally been associated with (e.g. *T.* "Ahipara", *T.* "Darkie", *T.* "Rough Leaf") – particularly in the phylogenies based on *ycf1*. This result supports that these are potentially new species in need of circumscription. While these tag-named entities are currently considered morphologically different from any currently described species, work is currently underway to properly characterise these entities, and better understand their distributions. Furthermore, additional molecular work is still required to better understand their origins as it appears hybridisation or allopolyploidy may be a factor for all three taxa (see below).

Allopolyploid origins are confirmed, and new hybrids are identified

Multiple nuclear gene copies of *LFY* were detected in nanopore sequence data for six taxa of *Thelymitra*. These data were generated for four of the six species Molloy and Dawson (1998) studied (*T. carnea*, *T. hatchii*, *T. nervosa*, *T. pulchella*) and an additional two entities not previously suspected of being of hybrid or allopolyploid origin (*T. colensoi*, *T.*

"Darkie"). Phylogenetic analysis of the *LFY* gene copies indicated distinct parental origins in the allopolyploid taxa, and these inferred progenitors were largely the same as those proposed by Molloy and Dawson (1998). However, our work further indicates inferred maternal and paternal progenitor species for these allopolyploid taxa. This inference is based on maternal chloroplast inheritance, which was not something Molloy and Dawson (1998) were able to propose based solely on morphological analysis. One difference from Molloy and Dawson (1998) was the evidence here to suggest a different paternal progenitor for some samples of *T. hatchii* and *T. pulchella* with *T. pauciflora* identified as a potential progenitor for some taxa rather than *T. longifolia* as hypothesised.

In the case of *T. nervosa,* this species was determined to be a cross between *T. longifolia* and *T. ixioides* by Peakall and Molloy (1998), and our results were consistent with this. For the accessions we investigated it was found that *T. ixioides* was always the maternal progenitor based on the maternal inheritance of *ycf1*. For *T. hatchii* it had been proposed by Molloy and Dawson (1998) that this species was a cross between *T. formosa* and *T. longifolia*. Our results suggested that the accessions investigated had differing paternal species and were derived from crosses between *T. longifolia* and *T. formosa*, or *T. pauciflora* and *T. formosa*. Based on *ycf1* sequences, *T. formosa* was the maternal progenitor for both samples. Molloy and Dawson (1998) hypothesised that *T. carnea* was a cross between *T. flexuosa* and *T. pauciflora*. Our results found that for this species the maternal parent for *T. carnea* is close to NZ accessions of *T. pauciflora*, and not *T. pauciflora* senso stricto from AU. The paternal parent for this species is currently unknown as *T. flexuosa* was unable to be included in this study, however it is similar in sequence to *T. antennifera*. Work by Nauheimer et al. (2018) showed that *T. antennifera* and *T. flexuosa* are related to each other in phylogenies based on nuclear and plastid sequence data, and so the original hypothesis by Molloy and Dawson (1998) may still hold.

Different results were found for all accessions of *T. pulchella* investigated for this study, including different placements in the nuclear and plastid phylogenies, and multiple sequence copies for *LFY*. This may suggest that the current concept of *T. pulchella* covers different entities that have arisen from multiple hybridisation events involving different progenitor species. This may explain why the species can be so variable in floral morphology, however future work will likely need to reassess the taxonomy of this species to account for this apparent varied origin. Only one accession of *T. pulchella* was identified as having multiple copies of *LFY*, with these corresponding to *T. pauciflora* and *T. cyanea*. The other two accessions of *T. pulchella* did not show variation and instead only a single copy of *LFY* was recovered. For one accession this copy was similar to *T. longifolia* and the other accession *T. pauciflora*.

Multiple copies of *LFY* were recovered from two entities not previously suggested as being hybrids or allopolyploids: *T.* "Darkie" and *T. colensoi*. For *T.* "Darkie" it appears that the paternal parent is closely related to *T. pauciflora*, and the maternal parent is currently unknown, but based on *ycf1* sequence may be related to *T. mucida* from Australia, as is the case for *T.* "Ahipara". The chromosome number for *T.* "Darkie" and *T.* "Ahipara" is much higher than many other species of *Thelymitra* at $2n = 60$ (Dawson et al. 2007), which would be consistent with a polyploid origin. At present there are no known combinations of chromosome numbers that could give us this value (Dawson et al. 2007) however this may involve aneuploidy or species that are currently uncounted and have novel chromosome numbers (e.g. $2n = 34$).

One major difference between the plastid and nuclear phylogenies was the placement of some clades between them. An example of this is the *T. cyanea* group, which in the nuclear phylogeny was sister to the rest of the genus, and in the plastid phylogeny this was sister to the *T. ixioides* group. The *T. formosa* group also changes position in the phylogenies, being sister to the *T. ixioides* group in the nuclear phylogeny, and sister to the *T. pauciflora* and *T. nuda* clades for the plastid phylogeny. Both of the clades mentioned are notable as they have chromosome counts of $2n = 40$, which is on the higher side reported for this genus. One explanation put forth by Dawson et al. (2007) to explain these high counts is a possible hybrid or allopolyploid origin both for this group and the other higher chromosome count groups in *Thelymitra* (e.g. $2n = 36$ for *T. flexuosa,* $2n = 70$ for *T. antennifera*), although currently there are no documented counts that could be combined to form these ploidy levels.

The phylogenetic incongruence between loci we identified could be evidence for an ancient hybridisation event in some of these clades, although no evidence supporting multiple gene copies was identified for the nuclear marker sequenced in the *T. cyanea* or *T. formosa* accessions. Furthermore, other plausible explanations for these incongruent relationships exist, such as incomplete lineage sorting (Joly et al. 2009), and without more information i.e. chromosome counts, sequencing more species, or other more robust molecular techniques, no conclusion can be drawn regarding an origin for these groups. Future work may benefit from analysing more nuclear markers – such as those offered by target enrichment (Johnson et al. 2019). Such an endeavour could identify discordance between loci consistent with different gene origins, or multiple gene copies indicating potential hybridisations or an allopolyploid origin.

While no direct evidence of allopolyploidy was identified for *T.* "Ahipara" or *T.* "Rough leaf", there was evidence to potentially suggest it as plausible. This included significant incongruence between nuclear and chloroplast phylogenies, and very high chromosome numbers for these entities ($2n = 60$, $2n = 84$, respectively). Work by Nauheimer et al. (2018) also proposed that *T.* "Rough leaf" may have been a hybrid based on detected variation in ITS sequence data, however this work did not investigate the nature of this variation.

Limited resolution in species complexes

Neither *ycf1* nor *LFY* showed high resolution within the *pauciflora* and *nuda* complexes. For *ycf1* all accessions of NZ *T. pauciflora* were the same, as well as accessions of *T. sanscilia*, *T. cyanapicata* and one *T. colensoi,* despite these samples covering a range of morphological variation observed. This may indicate that this is instead just a highly variable species that has been oversplit, and some of these different names may instead be better suited as synonyms. As an example, the synonymy of *T. cyanapicata* with *T. pauciflora* has also been brought into question by Australian researchers (Daniel Duval pers. comm.), however, work is still ongoing to investigate if these species are truly distinct. Alternatively, this lack of resolution may be a result of insufficient variation in the markers sequenced, and therefore other molecular approaches (i.e. GBS, RADseq) may be better suited for resolving these differences.

NZ accessions of *T. pauciflora* all differed from Australian accessions of the same species, suggesting that NZ *T. pauciflora* may in fact be a different species. A considerable amount of revision has been done by Australian authors on this group (Jeanes 2004,

2013; Bates 2010), and it is possible that what is called *T. pauciflora* in NZ may better fit under some of these new names. Unfortunately, many of these new species have yet to be sequenced and so until this happens, we cannot be quite certain on how NZ *T. pauciflora* fits into these revisions. Jeanes (2013) suggested that NZ *T. pauciflora* may have affinities with *T. peniculata* Jeanes, which is similar in morphology, however no molecular work has been done yet for that species.

The *T. longifolia* complex was found to be variable for both markers, however the phylogeny based on *ycf1* was more consistent with morphology and resolved the tag named forms separately. This marker broadly split the species into two groups, one consisting of large (leaf longer than 30 cm) and many flowered (15+) plants from the North Island of NZ (Coromandel and Northland regions), and the other comprised of a range of morphologies (i.e. leaf under 30 cm, flowers 1–10) and localities from both the North and South Islands, as well as two tag-named forms "Whakapapa" and "fusca". One accession of *T. longifolia* clustered with *T. pauciflora* for both markers, although this result may be due to a misidentification due to this specimen being collected late in the flowering season and most diagnostic characters being in poor condition (i.e. decaying). This result may be supportive of this species being a complex as hypothesised by Hatch (1952); Rolfe and de Lange (2010).

Conclusion

Overall, this work supports the hypothesis that there are yet undescribed species of *Thelymitra* in NZ, although more work is needed to confirm exactly where some species boundaries are. It appears that hybridisation events have played a role in the formation of species in this genus in NZ. It seems likely that there are yet more examples of this to be discovered (particularly in AU), however more work is needed to uncover them (i.e. chromosome counts to identify allopolyploids and progenitors, molecular work to confirm hybridisations). One major issue with this genus is the current nomenclatural confusion that surrounds many of the species and species groups, particularly in the *pauciflora* and *nuda* complexes. A comprehensive revision of this genus is long overdue, which would not only benefit future authors working on this group, but also conservationists concerned with species boundaries for rare taxa. Additional revision would also help in closing the gap in our understanding with respect to nomenclatural differences that currently exist between Australia and New Zealand.

Acknowledgements

This work was funded in part by an Hansjörg Eichler grant to HRJ, a grant from the APSF (19047) to CAL, and the New Zealand Native Orchid Group. HRJ also thanks the Te Papa Foundation for a MSc scholarship. HRJ would also like to thank everyone who helped with collecting samples for this study, and everyone who helped with lab work, especially Trish McLenachan, and Peter Lockhart. HRJ would also like to thank curators and staff at the AK, CANB, CHR, MPN, and WELT herbaria for their help in obtaining samples and permitting destructive sampling of material.

Authors contribution

Hayden R. Jones conducted the research and wrote the initial manuscript. Carlos A. Lehnebach and Jennifer A. Tate provided guidance on research and revisions to the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Australasian Systematic Botany Society and Australia and Pacific Science Foundation [grant number 19047].

ORCID

Hayden R. Jones http://orcid.org/0009-0000-6591-8718
Carlos A. Lehnebach http://orcid.org/0000-0001-7368-013X
Jennifer A. Tate http://orcid.org/0000-0001-5138-2115

References

Allan Herbarium. 2025 Checklist of the New Zealand Flora Seed Plants 2025. A New Zealand Plant Names Database Report. doi:10.26065/py9d-ch68.

Bates RJ. 2010. The *Thelymitra pauciflora* R.Br. complex (Orchidaceae) in South Australia with the description of seven new taxa. Journal of the Adelaide Botanic Gardens. 24:17–32. https://www.jstor.org/stable/23874281

Burns-Balogh P, Bernhardt P. 1988. Floral evolution and phylogeny in the tribe Thelymitreae (Orchidaceae: Neottioideae). Plant Systematics and Evolution. 159(1-2):19–47. doi:10.1007/bf00937424.

Burns-Balogh P, Funk VA. 1986. A phylogenetic analysis of the Orchidaceae. Washington: Smithsonian Institution Press. doi:10.5962/bhl.title.123271.

Cameron KM, Chase MW, Whitten WM, Kores PJ, Jarrell DC, Albert VA, Yukawa T, Hills HG, Goldman DH. 1999. A phylogenetic analysis of the Orchidaceae: evidence from *rbc*L nucleotide sequences. American Journal of Botany. 86(2):208–224. doi:10.2307/2656938.

Cheeseman TF. 1925. Manual of the New Zealand flora (2nd revised and enlarged by the author; edited by W.R.B. Oliver). Wellington: Govt. Printer. https://www.biodiversitylibrary.org/item/321773.

Christenhusz MJM, Byng JW. 2016. The number of known plants species in the world and its annual increase. Phytotaxa. 261(3):201–217. doi:10.11646/phytotaxa.261.3.1.

Clements MA, Jones DL, Sharma IK, Nightingale ME, Marron GJ, Fitzgerald KJ, Mackenzie AM, Molloy BPJ. 2002. Phylogenetics of Diurideae (Orchidaceae) based on the internal transcribed spacer (ITS) region of nuclear ribosomal DNA. Lidleyana. 17(3):135–117.

Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. 2020. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. Molecular Biology and Evolution. 37(1):291–294. doi:10.1093/molbev/msz189.

Dawson MI, Molloy BPJ, Beuzenberg EJ. 2007. Contributions to a chromosome atlas of the New Zealand flora—39. Orchidaceae. New Zealand Journal of Botany. 45(4):611–684. doi:10.1080/00288250709509743.

de Lange PJ, Crowcroft GM, Forester L. 1991 *Thelymitra* "Ahipara" an endangered orchid transferred, with notes on its taxonomic status, distribution and ecology. Head Office, Dept. of Conservation.

de Lange PJ, Gosden J, Courtney SP, Fergus AJ, Barkla JW, Beadel SM, Champion PD, Hindmarsh-Walls R, Makan T, Michel P. 2024. Conservation status of vascular plants in Aotearoa New Zealand, 2023 in Aotearoa New Zealand, 2023. www.doc.govt.nz.

Dressler RL. 1986. Recent advances in orchid phylogeny. Lindleyana. 1(1):5–20.

Flouri T, Izquierdo-Carrasco F, Darriba D, Aberer AJ, Nguyen LT, Minh BQ, Von Haeseler A, Stamatakis A. 2015. The phylogenetic likelihood library. Systematic Biology. 64(2):356–362. doi:10.1093/sysbio/syu084.

Freudenstein JV, Rasmussen FN. 1999. What does morphology tell us about orchid relationships? —A cladistic analysis. American Journal of Botany. 86(2):225–248. doi:10.2307/2656939.
Girgis HZ. 2022. Meshclust v3.0: high-quality clustering of DNA sequences using the mean shift algorithm and alignment-free identity scores. BMC Genomics. 23(1):423. doi:10.1186/s12864-022-08619-0.
Goldschmidt-Clermont M. 2001. Genetics of photosynthesis. In: Brenner S, Miller JH, editors. Encyclopedia of genetics. Cambridge: Academic Press; p. 1459–1465. doi:10.1006/rwgn.2001.1680.
Grant, V. 1981. Plant speciation. New York: Columbia University Press. doi:10.7312/gran92318.
Hatch ED. 1952. The New Zealand forms of *Thelymitra* J. R. and G. Forster and appendices. Transactions and Proceedings of the Royal Society of New Zealand. 79:386–402.
Jeanes JA. 2001. Resolution of the *Thelymitra canaliculata* R.Br. (Orchidaceae) complex in southern Australia. Muelleria: An Australian Journal of Botany. 15(1):75–89. doi:10.5962/p.237633.
Jeanes JA. 2004. A revision of the *Thelymitra pauciflora* R.Br. (Orchidaceae) complex in Australia. Muelleria: An Australian Journal of Botany. 19:19–79. doi:10.5962/p.291362.
Jeanes JA. 2006. Resolution of the *Thelymitra fuscolutea* R. Br. (Orchidaceae) complex of southern Australia. Muelleria: An Australian Journal of Botany. 24:3–24. doi:10.5962/p.291581.
Jeanes JA. 2009. Resolution of the *Thelymitra variegata* (Lindl.) F.Muell. (Orchidaceae) complex of southern Australia and New Zealand. Muelleria: An Australian Journal of Botany. 27(2):149–170. doi:10.5962/p.291950.
Jeanes JA. 2011. Resolution of the *Thelymitra aristata* (Orchidaceae) complex of south-eastern Australia. Muelleria: An Australian Journal of Botany. 29(2):110–129. doi:10.5962/p.292518.
Jeanes JA. 2012. Two new rare species in the *Thelymitra venosa* complex (Orchidaceae) from south-eastern mainland Australia. Muelleria: An Australian Journal of Botany. 30(1):8–22. doi:10.5962/p.292239.
Jeanes JA. 2013. An overview of the *Thelymitra nuda* (Orchidaceae) complex in Australia including the description of six new species. Muelleria: An Australian Journal of Botany. 31(1):3–30. doi:10.5962/p.295674.
Johnson MG, Pokorny L, Dodsworth S, Botigué LR, Cowan RS, Devault A, Eiserhardt WL, Epitawalage N, Forest F, Kim JT, et al. 2019. A universal probe Set for targeted sequencing of 353 nuclear genes from Any flowering plant designed using k-medoids clustering. Systematic Biology. 68(4):594–606. doi:10.1093/sysbio/syy086.
Joly S, McLenachan PA, Lockhart PJ. 2009. A statistical approach for distinguishing hybridization and incomplete lineage sorting. The American Naturalist. 174(2):E54–E70. doi:10.1086/600082.
Jones HR, Lehnebach CA. 2022. We need your help! your observations can help improve our understanding of the New Zealand white sun orchid (*Thelymtira longifolia*). The New Zealand Native Orchid Journal. 166:10–13.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution. 30(4):772–780. doi:10.1093/molbev/mst010.
Kores PJ, Cameron KM, Molvray M, Chase MW. 1997. The phylogenetic relationships of Orchidoideae and Spiranthoideae (Orchidaceae) as inferred from *rbc*L plastid sequences. Lidleyana. 12(1):1–11.
Kores PJ, Molvray M, Weston PH, Hopper SD, Brown AP, Cameron KM, Chase MW. 2000. Phylogenetic relationships within the Diurideae (Orchidaceae); inferences from plastid matK DNA sequences. In: Wilson KL, Morrison DA, editors. Monocots: systematics and evolution. Clayton: CSIRO Publishing; p. 449–456. https://www.ebsco.com/terms-of-use.
Kores PJ, Molvray M, Weston PH, Hopper SD, Brown AP, Cameron KM, Chase MW. 2001. A phylogenetic analysis of Diurideae (Orchidaceae) based on plastid DNA sequence data. American Journal of Botany. 88(10):1903–1914. doi:10.2307/3558366.
Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics. 35(21):4453–4455. doi:10.1093/bioinformatics/btz305.

Lehnebach CA, Jones HR. 2024. Taxonomic notes on New Zealand orchids I: typification of nine orchid names described by William Colenso in 1886. New Zealand Journal of Botany. 62:1–24. doi:10.1080/0028825X.2024.2353917.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 34(18):3094–3100. doi:10.1093/bioinformatics/bty191.

Molloy BPJ, Dawson MI. 1998. Speciation in *Thelymitra* (Orchidaceae) by natural hybridism and amphidiploidy. ecosystems, entolomogy, and plants. The Royal Society of New Zealand. 48:103–113.

Molloy BPJ, Hatch ED. 1990. *Thelymitra tholiformis* (Orchidaceae) — a new species endemic to New Zealand, with notes on associated taxa. New Zealand Journal of Botany. 28(2):105–114. doi:10.1080/0028825X.1990.10412350.

Moore LB. 1968. Taxonomic notes on New Zealand monocotyledons. New Zealand Journal of Botany. 6(4):473–492. doi:10.1080/0028825x.1968.10428586.

Nauheimer L, Schley RJ, Clements MA, Micheneau C, Nargar K. 2018. Australasian orchid biogeography at continental scale: molecular phylogenetic insights from the Sun Orchids (*Thelymitra*, Orchidaceae). Molecular Phylogenetics and Evolution. 127:304–319. doi:10.1016/j.ympev.2018.05.031.

Neubig KM, Whitten WM, Carlsward BS, Blanco MA, Endara L, Williams NH, Moore M. 2009. Phylogenetic utility of ycf1 in orchids: a plastid gene more variable than matK. Plant Systematics and Evolution. 277(1-2):75–84. doi:10.1007/s00606-008-0105-0.

Peakall R, Molloy BPJ. 1998. Isozyme evidence for the amphidiploid origin of *Thelymitra decora* (Orchidaceae). *ecosystems, entolomogy, and plants*. The Royal Society of New Zealand. 48:115–119.

Pérez-Escobar OA, Bogarín D, Przelomska NAS, Ackerman JD, Balbuena JA, Bellot S, Bühlmann RP, Cabrera B, Cano JA, Charitonidou M, et al. 2024. The origin and speciation of orchids. New Phytologist. 242(2):700–716. doi:10.1111/nph.19580.

POWO. 2024. Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. [retrieved 2024 Oct 31]. https://powo.science.kew.org/.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using tracer 1.7. Systematic Biology. 67(5):901–904. doi:10.1093/sysbio/syy032.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. Trends in Genetics. 16(6):276–277. doi:10.1016/s0168-9525(00)02024-2.

Rolfe J, de Lange PJ. 2010. Illustrated guide to New Zealand sun orchids, Thelymitra (Orchidaceae).

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. Mrbayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology. 61(3):539–542. doi:10.1093/sysbio/sys029.

Rothfels CJ. 2021. Polyploid phylogenetics. New Phytologist. 230(1):66–72. doi:10.1111/nph.17105.

Schlüter PM, Kohl G, Stuessy TF, Paulus HF. 2007. A screen of low-copy nuclear genes reveals the LFY gene as phylogenetically informative in closely related species of orchids (*Ophrys*). Taxon. 56(2):493–504. doi:10.1002/tax.562019.

Shepherd LD, McLay TG. 2011. Two micro-scale protocols for the isolation of DNA from polysaccharide-rich plant tissue. Journal of Plant Research. 124(2):311–314. doi:10.1007/s10265-010-0379-5.

Small RL, Cronn RC, Wendel JF. 2004. Use of nuclear genes for phylogeny reconstruction in plants. Australian Systematic Botany. 17(2):145–170. doi:10.1071/SB03015.

Thiers B. 2023. Index herbariorum. https://sweetgum.nybg.org/science/ih/.

Weston PH, Perkins AJ, Indsto JO, Clements MA. 2014. 5: Phylogeny of Orchidaceae tribe Diurideae and its implications for the evolution of pollination systems. In: Ednes-Meier R, Bernhardt P, editors. Dawin's orchids: then and Now. Chicago: University of Chicago Press; p. 92–154. doi:10.7208/chicago/9780226173641.003.0005.

WFO. 2024. Thelymitra J.R. Forst & G. Forst. [retrieved 2024 Oct 31]. https://www.worldfloraonline.org/taxon/wfo-4000038139.

Wong DCJ, Peakall R. 2022. Orchid phylotranscriptomics: The prospects of repurposing multi-tissue transcriptomes for phylogenetic analysis and beyond [perspective]. Frontiers in Plant Science. 13:1–9. doi:10.3389/fpls.2022.910362.

Zhang G, Hu Y, Huang M-Z, Huang W-C, Liu D-K, Zhang D, Hu H, Downing JL, Liu Z-J, Ma H. 2023. Comprehensive phylogenetic analyses of Orchidaceae using nuclear genes and evolutionary insights into epiphytism. Journal of Integrative Plant Biology. 65(5):1204–1225. doi:10.1111/jipb.13462.