NORDIC JOURNAL OF
BOTANY



*Research*

# Allopolyploidy in the Wintergreen Group of tribe Gaultherieae (Ericaceae) inferred from low-copy nuclear genes

**Lu Lu, Peter W. Fritsch, Catherine M. Bush, Hong Wang, Kathleen A. Kron and De-Zhu Li**

*L. Lu (https://orcid.org/0000-0002-7276-5358) ✉ (lulukmu@163.com), School of Pharmaceutical Sciences and Yunnan Key Laboratory of Pharmacology for Natural Products, Kunming Medical Univ., Kunming, Yunnan, P. R. China. – P. W. Fritsch, Botanical Research Inst. of Texas, Fort Worth, TX, USA. – C. M. Bush, Dept of Biology, Elon Univ., Elon, NC, USA. – H. Wang, Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Inst. of Botany, Chinese Academy of Sciences, Kunming, Yunnan, P. R. China. – H. Wang, D.-Z. Li (dzl@mail.kib.ac.cn), Plant Germplasm and Genomics Center, Germplasm Bank of Wild Species, Kunming Inst. of Botany, Chinese Academy of Sciences, Kunming, Yunnan, P. R. China. – K. A. Kron, Dept of Biology, Wake Forest Univ., Winston-Salem, NC, USA.*

**Nordic Journal of Botany**
**2019: e02077**
doi: 10.1111/njb.02077

Subject Editor and
Editor-in-Chief: Torbjörn Tyler
Accepted 15 March 2019

DNA sequence data from the low-copy nuclear genes *waxy* (GBSSI) and *leafy* were compared with plastid and ITS sequence data from prior studies to reconstruct phylogenetic relationships in the Wintergreen Group of tribe Gaultherieae (Ericaceae). We conducted phylogenetic analysis with 102 species that includes representatives of all 15 major clades previously identified within the Wintergreen Group and that together span its circum-Pacific distribution. Results yielded two distinct homeologous copies of *waxy* for two of the clades, each in widely separated parts of the tree. It also yielded two copies of *leafy* for one of the clades; only one copy of *leafy* was found for the other clade, but it was placed in the same major clade as its *waxy* counterpart and well away from its placement in a prior plastid analysis. A combined four-locus (*waxy*, *leafy*, ITS and plastid data) phylogenetic analysis of all available relevant data placed the copies of each of the clades in two distinct positions in the phylogeny with strong overall statistical support. In combination with evidence from morphology, reproductive biology and cytology, the results suggest that these clades arose through allopolyploid hybridization between lineages deep in the phylogeny but relatively close geographically. This finding confirms previous assumptions that hybridization has played an important role in the evolution of the Gaultherieae.

Keywords: allopolyploidy, Gaultherieae, hybridization, low-copy nuclear genes, phylogeny, Wintergreen Group

## Introduction

The 'Wintergreen Group' of tribe Gaultherieae (Ericaceae) comprises three genera (*Diplycosia* Blume, *Gaultheria* Kalm ex L. and *Tepuia* Camp) and ca 280 species with a combined circum-Pacific distribution encompassing a wide range of ecoregions and elevations. Two of its species (*G. leucocarpa* Blume and *G. procumbens* L.) have served as a primary natural source of wintergreen oil (methyl salicylate), used widely in the confection and pharmaceutical industries (Foster and Johnson 2008). Phylogenetic



NORDIC SOCIETY OIKOS

www.nordicjbotany.org

studies of the Gaultherieae based on DNA sequence data have been conducted under various taxon and gene region sampling schemes (Powell and Kron 2001, Kron et al. 2002, Bush et al. 2009a, b, Lu et al. 2010, Fritsch et al. 2011). The results of these studies are consistent in supporting the monophyly of the Wintergreen Group, with the putative synapomorphies of fleshy fruits and the presence of methyl salicylate (Powell and Kron 2001, Bush et al. 2009a). They are also consistent in the nested placement of *Diplycosia* and *Tepuia* within *Gaultheria*, rendering the latter paraphyletic. The most recent phylogenetic study of the group analyzed six genic regions (nuclear ribosomal (nr) ITS, and plastid *matK*, *ndhF*, *rpl16*, *trnL-trnF* and *trnS-trnG*) with parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI), yielding 15 major clades with strong support (Fritsch et al. 2011, Table 1). The analysis challenged many aspects of the most recent global classification of *Gaultheria* (Middleton 1991a): out of 10 sections and 21 series total, monophyly was supported for only one of the five sections and four of the 11 series that either have more than one species or for which more than one sample was included.

Hybridization has been thought to have played a critical role in the evolution of the Wintergreen Group, particularly in *Gaultheria* (Sleumer 1952, Baas 1985, Franklin 1964, Middleton and Wilcock 1990, Middleton 1991b, Luteyn 1995a, b). The chromosome numbers for species in the Gaultherieae are 2n = 22, 24, 26, 36, 44, 48, 66, 88 and ~96, and the numbers are usually species-specific, indicating that polyploid speciation commonly occurs in the group (Goldblatt and Johnson 1979, Middleton and Wilcock 1990). The strongly conflicting placements recovered for some species in nuclear versus plastid trees have been inferred to have resulted from geographically localized reticulation (hybridization or introgression) at shallow phylogenetic levels (Lu et al. 2010, Fritsch et al. 2011). However, ITS sequence data yielded a poorly resolved tree in Fritsch et al. (2011), and thus it was thus not clear whether low support for some clades in the combined nuclear and plastid analysis resulted from reticulation events, weak phylogenetic signal, or both. Furthermore, whether polyploidy in the Wintergreen Group is through autopolyploidy, allopolyploidy, or both, has not previously been explored.

Further sampling of loci from the nuclear genome could help to increase the resolving power of the data for phylogenetic reconstruction and assess the frequency, pattern and significance of reticulate evolution in the Wintergreen Group. Single- and low-copy nuclear genes often evolve more rapidly than plastid and nuclear ribosomal gene regions and therefore offer a rich source of potentially informative data for phylogenetic reconstruction (Sang 2002, Zhang et al. 2012, Zimmer and Wen 2015). Because they are biparentally inherited and less susceptible to concerted evolution than nuclear ribosomal genes, low-copy nuclear genes also have the potential to reconstruct allopolyploidization (Sang 2002). Two low-copy nuclear loci, *waxy* (the granule-bound starch synthase gene, i.e. GBSSI, functioning in the synthesis of amylose; Mason-Gamer et al. 1998, Peralta and Spooner 2001) and *leafy* (one of the key regulatory genes involved in the formation of the floral meristem; Frohlich and Meyerowitz 1997), have been widely used in phylogenetic analysis. *Waxy* has provided phylogenetic resolution in both diploid (Clark et al. 1991, Wang et al. 1990, 1999) and polyploid (Smedmark et al. 2003, Mason-Gamer 2004, Winkworth and Donoghue 2004, Fortune et al. 2008, Rousseau-Gueutin et al. 2009) plant groups. *Leafy* has also proven useful for phylogenetic reconstruction at the interspecific level (Nishimoto et al. 2003, Oh and Potter 2003, Grob et al. 2004, Kim et al. 2008). Bush et al. (2009b) used *waxy* and *leafy* in combination with ITS data and several plastid gene regions to reconstruct the phylogeny of the Australian/New Zealand species of *Gaultheria*, finding that including the data from *waxy* and *leafy* increased phylogenetic resolution for this clade, thus demonstrating that these regions can likely provide phylogenetically informative sequence variation at deeper levels in the Wintergreen Group.

Here we investigate the interspecific relationships of the Wintergreen Group with DNA sequence data from *waxy* and *leafy* and compare and combine results with those from prior data from the ITS region and five plastid regions (*rpl16*, *matK*, *ndhF*, *trnL-trnF* and *trnS-trnG*; Fritsch et al. 2011).

Table 1. Informal clade designations of clades in the Wintergreen Group used in this study, as compared to those used in Fritsch et al. (2011: Fig. 3) based on the classification of Middleton (1991a), with clade support (maximum likelihood bootstraps/Bayesian inference posterior probabilities) from *waxy* and *leafy* sampled in the present study and plastid and ITS from Fritsch et al. (2011), plus support from the combined four-locus analysis in the present study. Clades represented by a single sample are indicated as N/A. –, not sampled.

| This study | Fritsch et al. (2011) | *Waxy* | *Leafy* | Plastid | ITS | Combined four-locus |
|---|---|---|---|---|---|---|
| Amblyandra | *G.* sect. *Amblyandra* | 61/0.84 | N/A | 100/1.00 | 75/0.98 | 100/1.00 |
| Buxifolia | Buxifolia Group | < 60/< 0.80 | N/A | 100/1.00 | < 60/< 0.80 | 100/1.00 |
| Chiogenopsis | *G.* ser. *Hispidulae* | N/A | N/A | 100/1.00 | < 60/< 0.80 | 97/1.00 |
| Diplycosia | *Diplycosia* | 81/0.99 | 100/1.00 | 100/1.00 | 98/1.00 | 100/1.00 |
| Leucothoides copy A | Core East Asian Clade | < 60/< 0.80 | 100/1.00 | 100/1.00 | 63/0.99 | 100/1.00 |
| Monoanthemona | Australia/New Zealand Clade | < 60/< 0.80 | 90/1.00 | 98/1.00 | < 60/< 0.80 | 100/1.00 |
| Myrtilloideae | *G.* ser. *Myrtilloideae* | < 60/< 0.80 | 99/1.00 | 100/1.00 | 60/0.87 | 100/1.00 |
| Pernettya | *G.* sect. *Pernettya* p.p. | 89/1.00 | < 60/< 0.80 | 100/1.00 | < 60/< 0.80 | 100/1.00 |
| Shallonium | *G.* subsect. *Dasyphyta* p.p. | < 60/< 0.80 | 100/1.00 | 100/1.00 | < 60/< 0.80 | 99/1.00 |
| Sympodiobotrys | Sympodial clade | 92/< 0.80 | 100/1.00 | 100/1.00 | 80/1.00 | 100/1.00 |
| Tepuia | *Tepuia* | 100/1.00 | – | 100/1.00 | 99/1.00 | 100/1.00 |

We assess the potential of *waxy* and *leafy* for yielding a more resolved phylogeny and, on the discovery of duplicated copies of these genes in one species and one entire clade, use the results to examine the allopolyploid origin of these entities.

## Material and methods

We used 117 accessions of 113 species in the analyses (newly obtained data from waxy and leafy referring to 114 accessions of 110 species, Supplementary material Appendix 1). The sampling covers all 15 major clades recognized in the most recent phylogenetic framework of the Wintergreen Group in Fritsch et al. (2011). For clarity in discussing clade relationships within Gaultheria, we use newly created informal clade designations (Table 1) as substitute for those in Fritsch et al. (2011), who employed names based on the most recent global classification of Gaultheria (Middleton 1991a) as modified to accomodate monophyly. All clades corresponding to other genera were left as is. We sampled 3/3 species of Amblyandra, 5/9 species of Buxifolia, 1/1 species of Chamaephyta, 2/3 species of Chiogenopsis, 18/~121 species of Diplycosia, 1/1 species of Gaultheria (sensu stricto; s.s. hereafter), 1/6 species of Gymnobotrys, 1/1 species of Gymnocaulos, 23/~62 species of Leucothoides, 14/15 species of Monoanthemona, 5/6 species of Myrtilloideae, 8/15 species of Pernettya, 14/~24 species of Shallonium, 4/4 species of Sympodiobotrys and 2/7 species of Tepuia. The sample composition is similar to that in Fritsch et al. (2011) and covers the circum-Pacific distribution of the genus, including East Asia, Southeast Asia, the Americas and Oceania; an elevation range of 0–4500 m a.s.l.; and most of the key morphological characters in the classification of Middleton (1991a). Nine samples from the other four genera of the Gaultherieae were included in the outgroup: 1/1 species of *Chamaedaphne* Moench, 1/1 species of *Eubotryoides* (T. Nakai) H. Hara, 2/2 species of *Eubotrys* Nutt. and 5/5 species of *Leucothoe* D. Don. One accession each of two species from tribe Andromedeae, i.e. *Andromeda polifolia* L. and *Zenobia pulverulenta* (W. Bartram ex Willd.) Pollard, were used to root the trees (Kron et al. 2002) except in the *leafy* analyses, for which sequence data from these species were not available.

We extracted total DNA from silica gel-dried leaves with the CTAB method (Doyle and Doyle 1987) and amplified, cloned in vivo, and sequenced DNA with the primers in Bush et al. (2009b). We performed PCR amplification, clean-up and DNA sequencing as in Fritsch et al. (2011). We used the *waxy* region between exons 9 and 11 out of 13 translated exons (Mason-Gamer et al. 1998, Peralta and Spooner 2001) and the *leafy* intron region between exons 2 and 3 (2i3), as in Bush et al. (2009b). The PCR primers are from Mason-Gamer et al. (1998) and Peralta and Spooner (2001) for *waxy*, and Frohlich and Meyerowitz (1997) for *leafy*. PCR amplification consisted of a hot start at 95°C for 10 min; 35–40 cycles of denaturation at 95°C for between 30 s and 1 min, primer annealing at 48–52°C for 1 min, primer extension at 72°C for 5 min, and a final extension at 72°C for 7 min. Sequencing was performed at the Bowman Grey Technical Center, Wake Forest Univ., NC, USA, on an ABI 377 automated sequencer.

Although *waxy* and *leafy* have been recognized as single-copy genes in many dicots (Mérida et al. 1990, Wang et al. 1999, Oh and Potter 2003), multiple copies have been detected in some species of the Ericaceae (K. Kron, unpubl.) and perhaps for *waxy* in *Gaultheria tasmanica* (Hook. f.) D.J. Middleton of the Wintergreen Group from Australia and New Zealand (Bush et al. 2009b). We therefore conducted in vivo cloning procedures for both *waxy* and *leafy*. We followed the procedure of Bush et al. (2009b), using the Topo Cloning Kit for Sequencing (Invitrogen), with up to six clones sequenced for each taxon. We edited *waxy* and *leafy* sequences with Sequencher 4.8 (Gene Codes Corp., Inc.).

Clone sequences were selected for further phylogenetic analyses with the method of Triplett et al. (2014), in which distinct clone types that may be putative homeologs duplicated through allopolyploidy could be visually identified in both *waxy* and *leafy*. PCR recombinants can occur by homologous recombination or PCR strand swapping due to the unavailability of a nonproofreading DNA polymerase, and could potentially increase homoplasy and distort the resulting phylogeny. We therefore first identified and eliminated PCR recombinants by inspecting the *waxy*/*leafy* gene alignments as a splits network in SplitsTree4 ver. 4.14.6 (Huson 1998, Huson and Bryant 2006) with the NJ distance-based method. To minimize the inclusion of sequencing errors, a consensus sequence (a single sequence for each sequence type) per sample was generated in Geneious ver. 8.0.2 (<www.geneious.com>). Original sequences were ultimately submitted to GenBank. We aligned the final consensus sequences with MAFFT ver. 6.602b (Katoh et al. 2002), manually adjusted the alignment in Se-Alv.2.0a11 (Rambaut 2002), and calculated the number of variable sites in Mega 6.0 (Tamura et al. 2013). The sequences were considered to represent different homeologs if they clustered in two different major clades.

We conducted separate phylogenetic analyses on the *waxy* and *leafy* datasets. The *waxy* dataset comprised 114 clones from 104 species, and the leafy dataset comprised 79 clones from 76 species. Samples of the Andromedeae were not available for *leafy* and so we rooted the *leafy* tree on the two species of *Leucothoe* sampled, based on prior results that placed this genus outside of the Wintergreen Group (Powell and Kron 2001, Bush et al. 2009a, Fritsch et al. 2011). We performed ML analyses with RAxML 7.0.4 (Stamatakis 2006) using a bootstrap analysis (2000 replicates) simultaneously with the ML analysis (option '-f a') under the GTRGAMMA model. We performed BI analyses with MrBayes ver. 3.2.1 (Ronquist et al. 2012) on the Cipres Science Gateway (Miller et al. 2011; <www.phylo.org>) with 10 million generations, four Monte Carlo Markov chains (MCMCs), and a sampling frequency of 1000. We selected substitution models for each genic region by applying the Akaike information criterion as implemented in jModelTest 2 (Darriba et al. 2012). The best-fitting model for phylogenetic reconstruction was found to be GTR + I + G for *leafy*, and GTR + G for *waxy*.

We examined the results with Tracer ver. 1.7 (Rambaut et al. 2018) to ensure that the analyses reached convergence and that the effective sample size of each parameter was > 200. We generated a posterior probability (PP) consensus tree with the 'sumt' option in MrBayes and a burn-in of 10%.

We constructed a four-locus dataset by combining our data from *waxy* and *leafy* with the six-gene (two-locus, i.e. ITS and plastid) data from Fritsch et al. (2011; ITS and plastid *rpl16*, *matK*, *ndhF*, *trnL-trnF* and *trnS-trnG*; TreeBASE repository No. S11331). Combining these data into a concatenated dataset allowed the duplicated genes to be placed within a more robust phylogenetic framework than by separate *waxy* and *leafy* analyses, with higher resolution and clade support. The dataset consisted of 92 accessions from 90 species from all genera of the Gaultherieae and the same two species of the Andromedeae as above. The dataset included 85 *waxy* sequences from 81 samples of 79 species in the Gaultherieae plus two from the Andromedeae, and 61 *leafy* sequences from 59 samples and species within the Gaultherieae.

To maximize bootstrap support within major clades and to prevent conflicting data from potentially interfering with the deeper-level placements of clades and the duplicated loci (Wagner 1980, Judd et al. 2015), we excluded 22 accessions from the combined analysis that exhibited topological conflicts within or among major clades between any two combinations of *waxy*, *leafy*, ITS or plastid analyses, including those found in Fritsch et al. (2011) between ITS and plastid trees (no strong conflicts were exhibited between *waxy* and *leafy*). The conflicts involving these accessions may result from hybridization among closely related taxa within major clades, e.g. in New Zealand *Gaultheria* where hybridization is known to be rampant (Franklin 1964), or among the major clades Shallonium and Buxifolia (*G. amoena* A.C. Sm., *G. insipida* Benth., *G. megalodonta* A.C. Sm. and *G. sclerophylla* Cuatrec.), although coalescent stochasticity (lineage sorting) could also be involved in some cases. Accessions were excluded if either ML BP was <0.80 or PP < 0.95.

We retained *Gaultheria procumbens* in the combined analysis despite evident data partition conflicts (Fritsch et al. 2011) because it is the only species in its section. With this dataset, we conducted ML and BI analyses as above. The best-fitting models were found to be GTR + I + G for the plastid *rpl*16 and *trnL-trnF* regions, and GTR + G for the other regions.

We used ASTRAL (Mirarab et al. 2014) to compare the concatenated results with a coalescent-based method of combining data partitions. ASTRAL improves on other coalescent-based methods in statistical consistency and is often more accurate than concatenation with maximum likelihood (Mirarab et al. 2014). We used the tree with the highest likelihood from each of the four separate ML analyses of individual loci as input.

Chromosome number evolution was inferred by tracing chromosome number with parsimony onto a modified version of the four-locus ML tree in which many of the major clades were reduced to single terminals because of extensive missing data in species chromosome counts, and in which the two allotetraploid clades/species were excluded. Chromosome numbers are from Middleton and Wilcock (1990). Parsimony analysis was conducted with the program Mesquite ver. 3.51 (Maddison and Maddison 2018).

### Data deposition

DNA sequences newly obtained are deposited in the GenBank Nucleotide Sequence Databases, and aligned molecular data matrices are available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.2gh0ht2> (Lu et al. 2019).

## Results

The aligned *waxy* dataset was 560 bp long with 205 variable sites (37%), and the *leafy* dataset was 3950 bp long with 1162 variable sites (29%). In the analyses of both *waxy* and *leafy* (Supplementary material Appendix 1 Fig. A1, A2, respectively), some species have clones that exhibit two phylogenetically distinct sequences, each of which falls into two separate parts of the tree corresponding to *Gaultheria myrsinoides* Kunth and Leucothoides. The sequences of these two clades represent distinct copies of each of these genes, and the clades in which they occur are here arbitrarily designated 'copy A' and 'copy B,' with copy A concordant with the placement of these clades in prior plastid analyses, and copy B placed in another region of the trees (Fritsch et al. 2011).

For *Gaultheria myrsinoides*, we recovered two copies of *waxy* from the two accessions sampled of this species (Supplementary material Appendix 1 Fig. A1). The two accessions with copy A group with each other (BP = 76, PP = 0.99) within Pernettya (BP = 89, PP = 1.00), whereas the two accessions of copy B group as unresolved within Shallonium (BP = 70, PP = 0.84; Supplementary material Appendix 1 Fig. A1). Despite several sequencing attempts, we were only able to recover one clone type of *leafy* for *G. myrsinoides*, the placement of which is sister to *G. schultesii* Camp (BP < 60, PP < 0.80) within Shallonium (BP = 99, PP = 1.00; Supplementary material Appendix 1 Fig. A2) and we thus considered it to be comparable to copy B of the *waxy* gene.

The species of Leucothoides are represented by two distinct copies in both *waxy* and *leafy*. In the *waxy* tree, the sequences of copy A form a clade (BP < 60, PP = 0.93; Amblyandra and three species of Shallonium also group here but with low support) within a large clade comprising Buxifolia, Gymnocaulos, Pernettya, Chamaephyta, Monoanthemona, Myrtilloideae, Shallonium, Sympodiobotrys and Tepuia (BP < 60, PP = 0.91; Supplementary material Appendix 1 Fig. A1). The copy B *waxy* sequences form a clade (BP = 63, PP = 0.80; including Chiogenopsis) that groups with a clade comprising Gaultheria s.s., Gymnobotrys and Diplycosia (BP = 96, PP = 1.00). Four species (*Gaultheria heteromera* R.C. Fang, *G. pseudonotabilis* H. Li ex R.C. Fang, *G. semiinfera* (C.B. Clarke) Airy Shaw and *G. straminea* R.C. Fang) have sequences placed in both the copy A and copy B clades.

One species of Leucothoides, *G. hookeri* C.B. Clarke, has a sequence that groups outside of the copy A clade but with only weak support (BP < 60, PP = 0.91). Because another clone of this species nests within the copy B clade, this sequence is likely copy A (BP < 76, PP = 0.80).

Similarly, *leafy* sequences of copy A from Leucothoides form a clade (BP = 100; PP = 1.00) that is sister to Amblyandra (BP = 84; PP = 1.00), and the sequences from copy B form a clade (BP = 97, PP = 1.00) that is sister to a clade comprising Gaultheria s.s. and Chiogenopsis (BP = 80, PP = 0.99; Supplementary material Appendix 1 Fig. A2). No *leafy* sequences of Leucothoides are placed outside of these two clades. Two species (*Gaultheria fragrantissima* Wall. and *G. hypochlora* Airy Shaw) have sequences placed within both the copy A and B clades.

The results of the concatenated four-locus analysis largely reflect those of the plastid results in Fritsch et al. (2011), likely because the tree from the plastid locus exhibits the highest resolution and clade support of the four loci used (Fig. 1). Bootstrap values of many major clades are at or near 100 (ML) or 1.00 (PP; Table 1). Some other strongly supported clades are also consistent with those recovered in prior plastid/ITS analyses, i.e. the two clades comprising the first divergence within the Wintergreen Group clade (BP = 95, PP = 1.00; BP = 98, PP = 1.00); the 'large-leaved' clade in Leucothoides copy A (*Gaultheria pseudonotabilis* through *G. hookeri* in Fig. 1; BP = 100, PP = 1.00) and the 'small-leaved' clade of Leucothoides copy A (i.e. *G.* ser. *Trichophyllae* Airy Shaw, *G. dolichopoda* Airy Shaw through *G. ciliisepala* Airy Shaw ex P.W. Fritsch & Lu Lu in Fig. 1; BP = 100, PP = 1.00); Gymnobotrys + Diplycosia (BP = 100, PP = 1.00). The species of Australia (*G. appressa* A.W. Hill, *G. hispida* R. Br., *G. lanceolata* Hook. f. and *G. viridicarpa* J.B. Williams) within Monoanthemona also form a strongly supported clade (BP = 92, PP = 1.00), as in Bush et al. (2009b).

The four-locus concatenated results place the two accessions of *Gaultheria myrsinoides* as sister to each other in both copy A and copy B (BP = 100, PP = 1.00), with copy A grouping within Pernettya (BP = 100, PP = 1.00) and copy B grouping within Shallonium (BP = 99, PP = 1.00; Fig. 1). The copy A clade of Leucothoides (BP = 100, PP = 1.00) groups as sister (BP = 98, PP = 1.00) to a large clade (BP < 60, PP = 0.96) comprising Amblyandra, Buxifolia, Chamaephyta, Gymnocaulos, Monoanthemona, Myrtilloideae, Pernettya, Shallonium and Sympodiobotrys. The copy B clade of Leucothoides (BP = 84, PP = 1.00) groups as sister (BP < 60, PP < 0.80) to Chiogenopsis (BP = 97, PP = 1.00) within a larger clade that also includes Gaultheria s.s., Diplycosia, Gymnobotrys and Tepuia (BP = 95, PP = 1.00).

The results of the coalescent-based analysis from ASTRAL largely reflect those from the concatenated dataset, but with less overall resolution and slight differences in topology in poorly supported regions (Fig. 2). The results of the parsimony analysis of chromosome number yielded an ancestral number of n = 11 in the Wintergreen Group, with all other numbers derived (Fig. 3). When polyploidy occurs, it does so independently among the five major clades in which it has been found (Gaultheria s.s., Gymnobotrys, Leucothoides, Pernettya and Shallonium).

## Discussion

The results from ML and BI applied to either separate or combined datasets in the concatenated or coalescent-based analyses generally produced similar results; the few topological differences lacked strong support (i.e. BP were <0.6, and PP < 0.8). The *leafy* tree is generally more resolved and exhibits higher clade support than the *waxy* tree, likely due to the relatively low number of variable sites in the *waxy* dataset. Although there are some topological differences between the trees from *waxy* and *leafy*, none are strongly supported. The results from both of these genes are also generally consistent with those of Fritsch et al. (2011) based on ITS and plastid sequence data. The main differences involve the relative placements of Amblyandra, Leucothoides, Sympodiobotrys and Shallonium. For example, in the concatenated analysis Amblyandra is sister to Leucothoides copy A (BP = 84, PP = 1.00) in the *leafy* tree, whereas it is sister to Shallonium in the combined plastid + ITS results (BP = 86, PP = 1.00). Also, Sympodiobotrys + Shallonium is sister to the clade of Gymnocaulos + Pernettya + Chamaephyta + Myrtilloideae + Monoanthemona in the *leafy* tree (BP = 96, PP = 1.00), whereas only Sympodiobotrys is sister to this clade in the combined plastid + ITS results (BP = 79, PP = 1.00). The basis for these conflicts are unclear and may involve rapid cladogenesis and/or ancient reticulate evolution. The rest of the phylogeny is consistent with that from Fritsch et al. (2011), particularly regarding the placement of the major clades highlighted in that study. On this basis, the classification of the Wintergreen Group will be revised to reflect strictly monophyletic groups (K. A. Kron, W. S. Judd, P. W. Fritsch, L. Lu unpubl.).

Our results yielded two distinct homeologous copies of *waxy* for two of the clades in the Wintergreen Group (*Gaultheria myrsinoides* in Pernettya, and the Leucothoides clade), each in widely separated parts of the tree. The placement of one of these copies in each case (copy A) was consistent with the placement of these clades in a prior plastid analysis and plastid + ITS analysis (Fritsch et al. 2011). The results also yielded two copies of *leafy* for one of the clades (Leucothoides), and again, the placement of one of these (copy A) was consistent with that from the prior plastid and plastid + ITS analyses. Only one copy of *leafy* was found for the other clade (*G. myrsinoides*), but it was placed in the same major clade as its *waxy* counterpart and well away from its placement in a prior plastid analysis. We thus infer that we did not detect copy A of this species for *leafy*. Our combined four-locus analysis with *waxy*, *leafy*, plastid and ITS data placed the copy A and B clades in robust positions in the phylogeny, distinct from each other. All other clones from the other 13 major clades exhibit no duplicated distinct copies.

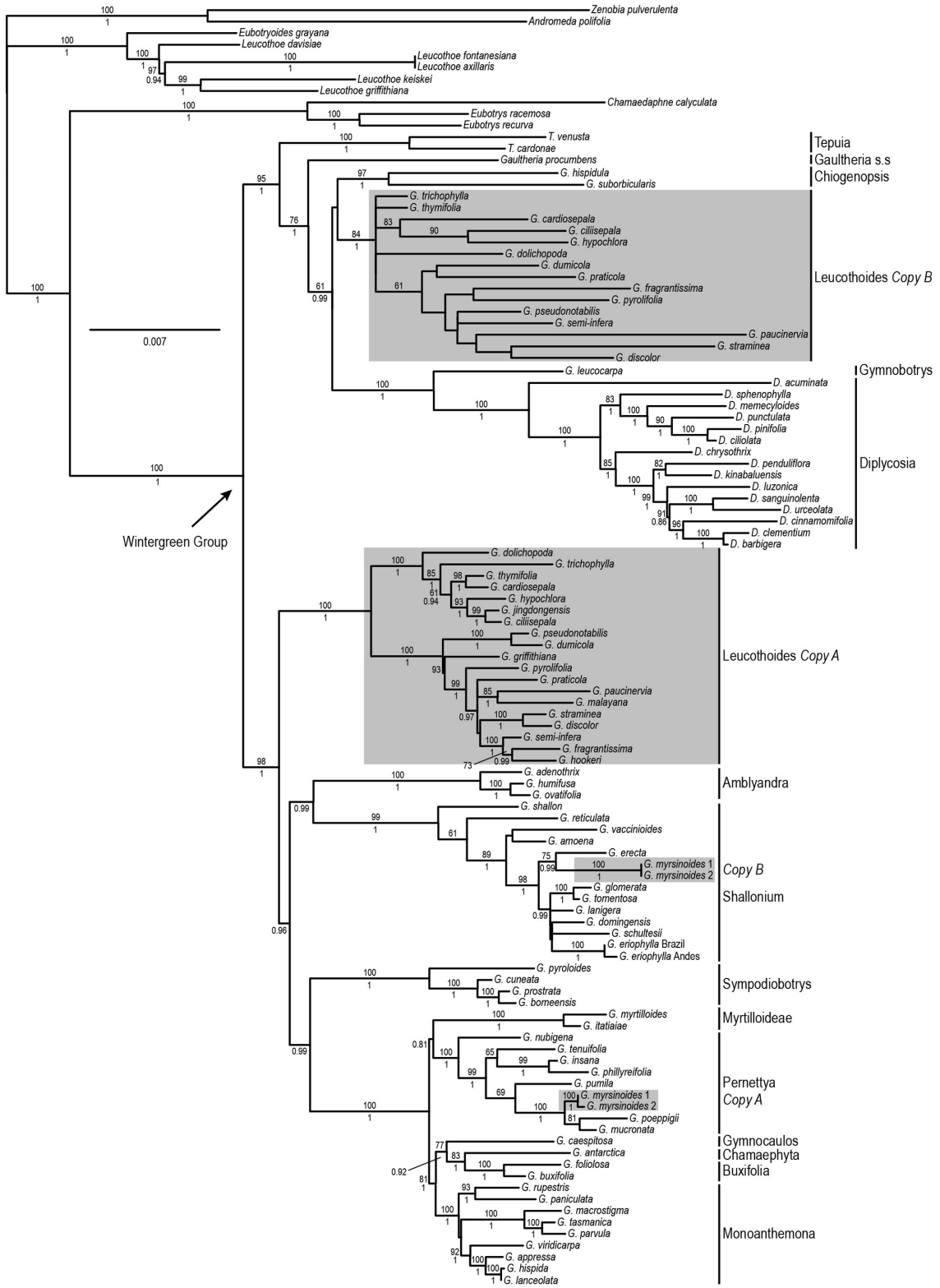


Figure 1. Best tree from a maximum likelihood (ML) analysis based on a concatenated four-locus dataset (*waxy*, *leafy*, ITS and plastid data) with 92 samples and 90 species of the Gaultherieae plus two outgroup species from the Andromedeae. Numbers above branches are ML bootstraps (BP); those below branches are posterior probabilities (PP) from a separate Bayesian inference analysis as mapped onto the ML tree. Only values ≥ 60 BP or ≥ 0.80 PP are shown. Clades involved in duplicated *waxy* genes and their two distinct copies (A and B; see text) are indicated with shading. Clade names are as in Supplementary material Appendix 1.

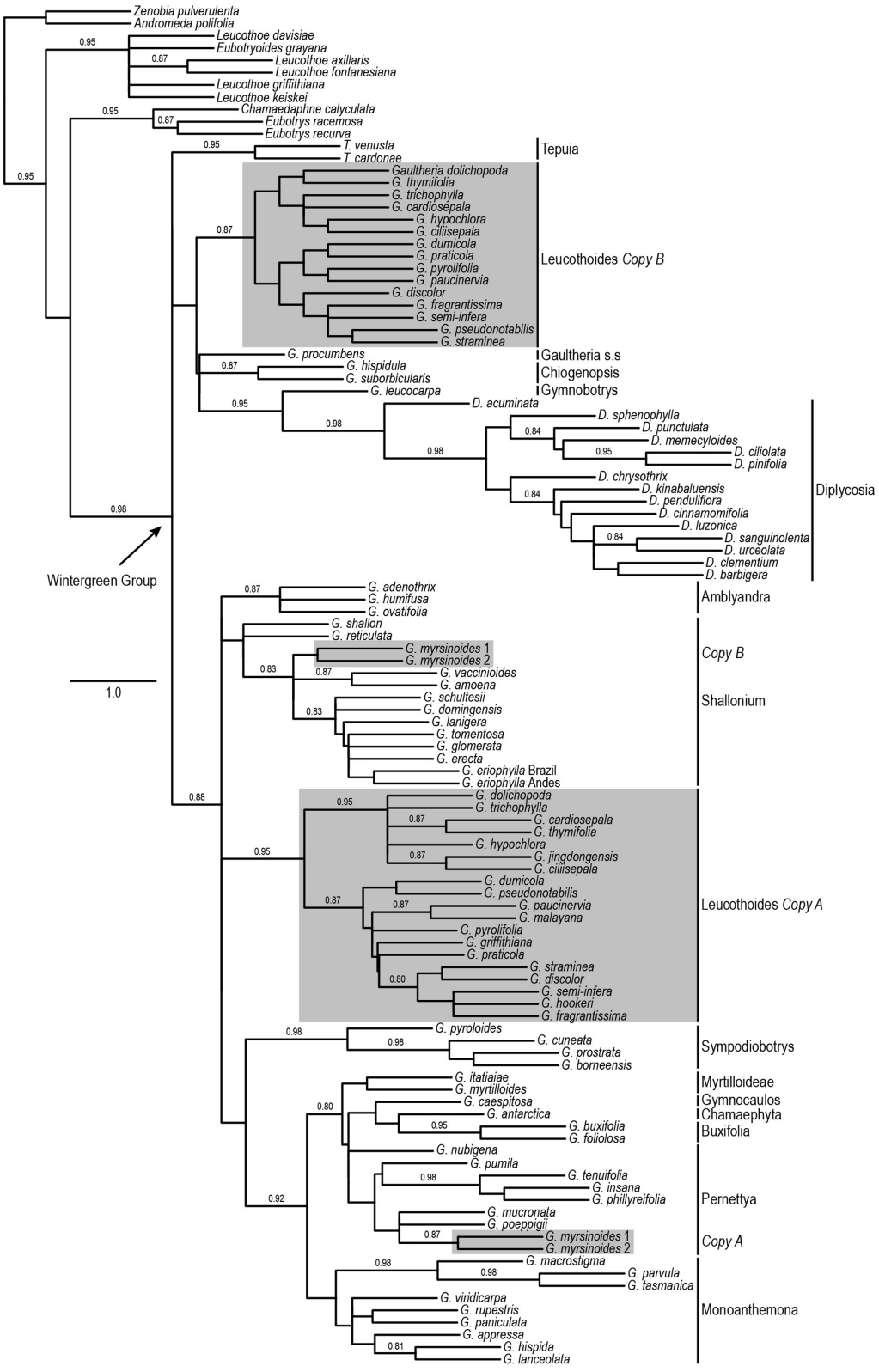


Figure 2. Tree from the coalescent-based estimate based on the best trees from each of the four separate maximum likelihood analyses on individual loci. Numbers along branches are posterior probabilities ≥ 0.80. Clades involved in duplicated *waxy* genes and their two distinct copies (A and B; see text) are indicated with shading. Clade names are as in Supplementary material Appendix 1.

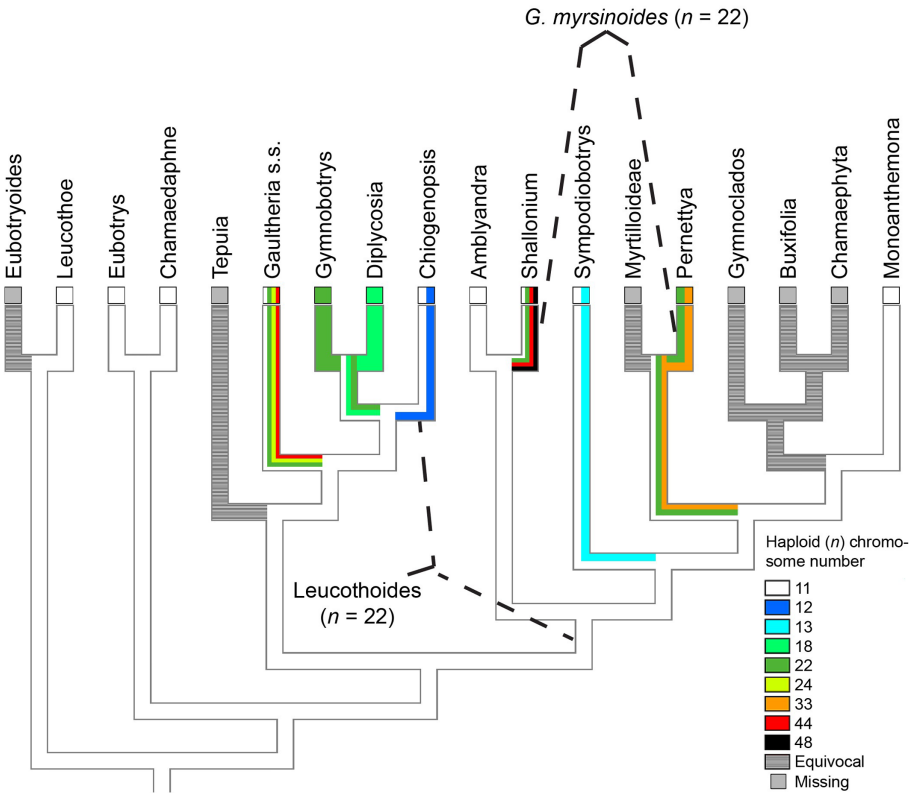


Figure 3. Chromosome numbers in the Gaultherieae traced with parsimony onto the four-locus maximum likelihood tree, as modified such that many of the major clades have been reduced to single terminals because of extensive missing data in species chromosome counts, and with the two allotetraploid clades/species excluded from the tree and added manually, with dotted lines indicating the putative parental lineages for each. Colors denote character states for the haploid number indicated. In clades with equivocal branches (hatched), chromosome counts have not been conducted. Clade names are as in Supplementary material Appendix 1.

When combined with evidence from chromosome numbers, the formation of distinct clades of gene copies from the same species samples has been considered evidence for allopolyploidy. In a phylogenetic analysis of *Viburnum* L., Winkworth and Donoghue (2004) found two polyploid clades with an additional copy in each of two paralogous *waxy* loci. They proposed that these duplicated clade patterns arose through allopolyploidy, and enumerated several angiosperms with similar allopolyploid patterns based on *waxy* data, e.g. *Elymus* L. and *Eragrostis* Wolf in the Poaceae (Mason-Gamer 2001, Ingram and Doyle 2003), the Maloideae of Rosaceae (Evans et al. 2000, Evans and Campbell 2002), Geinae of Rosaceae (Smedmark et al. 2003), *Fragraria* L. in the Rosaceae (Rousseau Gueutin et al. 2009) and the Rhamnaceae (Evans et al. 2000) Likewise, as based on distinct clone types of a non-single copy *leafy* gene, allopolyploid speciation or a hybrid origin was inferred in *Persicaria* (L.) Mill. in the Polygonaceae (Kim et al. 2008), Neillieae in the Rosaceae (Oh and Potter 2005), and *Pseudotsuga* Carr. in the Pinaceae (Wei et al. 2010). Two distinct copies of another nuclear low-copy gene (i.e. AAT) was explained by allopolyploidy in the Sideroxyleae of Sapotaceae (Smedmark and Anderberg 2007).

Similarly, chromosome numbers combined with phylogenetic results have provided evidence for allopolyploidy in the Wintergreen Group. Chromosome numbers in the Gaultherieae are reported to be 2n = 22 for *Chamaedaphne*, *Eubotrys* and *Leucothoe* (base number x = 11), 2n = 36 for *Diplycosia* (x = 9 or 18), and 2n = 22, 24, 26, 44, 48, 66, 88 and ~96 for *Gaultheria* (x = 11, 12 or 13), with *Eubotryoides* and *Tepuia* still unknown (Middleton and Wilcock 1990, Goldblatt and Johnson 1979 onwards). The character mapping analysis recovered a chromosome number of n = 11 as the ancestral number in the Wintergreen Group, suggesting x = 11 as the base number for the clade. Among the 35 species with chromosome numbers reported in the Gaultherieae, those from only five of the 15 major clades of the Wintergreen Group are reported as polyploids (Fig. 3). Species in Pernettya, including *G. myrsinoides* and all species sampled from Leucothoides, are reported as tetraploids (2n = 44). Tetraploids and hexaploids have been reported in other species of Pernettya; diploids, tetraploids and hexaploids in Gaultheria s.s. (*G. procumbens*); tetraploids in Gymnobotrys (only *G. leucocarpa* sampled); and diploids and tetraploids in Shallonium (with also one octoploid, *G. shallon* Pursh).

Our study provides strong evidence that tetraploidy in Leucothoides arose through a single allopolyploidization in the most recent common ancestor of the clade. *Waxy* and/or *leafy* are both duplicated among the many species of Leucothoides, suggesting whole-genome duplication for the entire clade, and the phylogenetic placement of the duplications agrees across genes. The phylogenetic placements of the duplications suggest that the clade arose between members of the diploid Chiogenopsis clade and a diploid ancestor of Leucothoides. Our separate and combined results grouped the two species of Chiogenopsis (each with only one copy of *waxy* and/or *leafy*) with copy B of *waxy* and *leafy* of Leucothoides. The placements of copy A clones from both *waxy* and *leafy* are congruent with those in the plastid + ITS tree, whereas those of copy B clones are incongruent. Our inference of this event is also supported by morphological evidence. *Gaultheria* ser. *Trichophyllae* was traditionally placed in the same morphologically based section as Chiogenopsis in the classification of Middleton (1991a), i.e. *G.* sect. *Chiogenopsis* D.J. Middleton. The species of the two series share several characters considered important in the classification of the genus, i.e. leaves usually < 1 cm wide, the presence of two bracteoles borne apically on the pedicel immediately beneath the calyx, solitary flowers and a glabrous corolla, although these characters are not widely shared among the rest of the species of Leucothoides. Furthermore, the geographic distribution of one of the species of Chiogenopsis, i.e. *G. suborbicularis* W.W. Sm., overlaps that of many species of Leucothoides in the Himalaya-Hengduan Mountains, indicating a potential sympatric area where hybridization could have occurred between ancestral elements of these two clades. The four-locus phylogeny combined with evidence from morphology and geographic distribution suggests that an ancestor within the *G. suborbicularis* lineage is the paternal diploid parent of Leucothoides. The putative maternal ancestor remains uncertain or else is extinct, and could be a deep-level ancestor in the Wintergreen Group because copy A of Leucothoides groups as sister to a large clade comprising most of the other major clades of *Gaultheria* (Fig. 1). Two other clades in this large clade, Sympodiobotrys and Amblyandra, have Asian species, and thus an ancestor common to all the Asian clades could have been involved in the hybridization event.

The data also suggest that tetraploidy in *Gaultheria myrsinoides* arose through allopolyploidization. Two copies of *leafy* were detected in different parts of the phylogeny, suggesting whole-genome duplication, although in this case only copy B of *waxy* was detected. The allopolyploid origin of *G. myrsinoides* appears to have involved a diploid ancestor in Shallonium, possibly in the lineage of *G. erecta* Vent. or *G. schultesii* (species that are near or sister to *G. myrsinoides* copy B in the *leafy* and combined analysis) with the other parent an ancestral lineage in Pernettya (Fig. 1). The latter parent might be the ancestor within the *G. mucronata* (L. f.) Hook. & Arn./*G. poeppigii* DC. lineage, which forms the sister clade of *G. myrsinoides* copy A. However, both of these species are documented hexaploids, and so the immediate ancestor of these species is unlikely to have formed the allotetraploid *G. myrsinoides* with a diploid parent. Diploidy is unknown in Pernettya but might occur in one or more of the unsampled species of this clade. Alternatively, the clade may have originally been diploid, the polyploidy in these species having arisen at least in part though autopolyploidy. This is supported by the detection of only a single copy of *waxy* and/or *leafy* in *G. insana* (Molina) D.J. Middleton, *G. mucronata*, *G. poeppigii* and *G. pumila* (L. f.) D.J. Middleton, all from Pernettya. The allopolyploid event giving rise to this species would likely have occurred in the region of overlap of Pernettya and Shallonium, from central Mexico to the central Andes of South America.

*Gaultheria myrsinoides* exhibits high morphological variation and a wide geographic range from Mexico to Argentina, with many species segregates named but synonymized by Luteyn (1995a). It has been found to hybridize with both *G. erecta* in Mexico (with which it was placed as sister in the combined analysis) and *G. reticulata* Kunth in South America (Camp 1939, Middleton 1991b). It is unclear from our data as to whether one of these species is one of the ancestors of *G. myrsinoides* in its tetraploid evolution. Data with higher resolving power could shed more light on the ancestry of the species.

The lack of some of the copies of genes otherwise expected in Leucothoides and *Gaultheria myrsinoides* could be due either to clone sampling artifact or the loss of the copy, the latter of which commonly occurs for one of the homeologs of allopolyploids (Doyle et al. 2008). Although it is possible that an early whole-genome duplication or polyploid event occurred early in the evolution of the Wintergreen Group clade followed by the loss of one of the homeologs, this would require independent losses throughout the clade to explain the recovery of single-copy genes in most of its species, versus only the two allopolyploid gains in the most recent common ancestor of the two *Gaultheria* lineages. Moreover, it is unlikely that our sampling would not have recovered one of the copies for one of the two genes in other large clades. In Leucothoides, for example, we recovered either copy A or copy B or both in 15 species of the clade sampled, and for other large clades the same should have been expected if they were of allopolyploid origin. A fossil-calibrated dating analysis of the major clades of the Gaultherieae based on a combined plastid and ITS phylogeny yield a time of divergence for Leucothoides crown and stem nodes of ca 8 and 13 Ma (Lu et al. 2019). This represents a substantial amount of time in which gene loss among various species of the clade could have occurred. The divergence time of *G. myrsinoides* from its sister species is much more recent at ca 1.3 Ma, but still would allow ample time for possible gene loss.

Some of the known polyploid species in the Wintergreen Group have only single copies of *waxy* and *leafy*, i.e. *Gaultheria leucocarpa* of Gymnobotrys (tetraploid), *G. reticulata* and *G. shallon* of Shallonium (tetraploid and octoploid, respectively), the species of Pernettya (tetraploid and hexaploid),

and *G. procumbens* of Gaultheria s.s. (diploid, tetraploid and hexaploid). As suggested for *Viburnum* (Winkworth and Donoghue 2004), these polyploids may have arisen independently, either through autopolyploidy or after hybridization between closely related clades, and therefore the homeologs in these species/clades may not have distinct sequences. Alternatively, our sample size precluded the detection of the other copy in at least some of these species, or one of the two homeologs was lost.

The allopolyploid origins for the species within Leucothoides and Pernettya are consistent with their geographic locations in montane regions. The species of Leucothoides have their center of diversity in the Himalaya and Hengduan Mountains in East Asia, and the latter in the Andes of South America. These regions are recognized as 'species pump' areas (Terborgh 1992, Jetz et al. 2004), with highly variable environmental parameters involving area, elevation, topography and climate (Körner 2004, Tang et al. 2006, Schwery et al. 2015, Yu et al. 2018) promoting hybridization and subsequent (allo)polyploid speciation (Rauscher 2002, Nie et al. 2005, Schmidt-Lebuhn et al. 2010).

The ITS sequences in the allopolyploids (ITS sequences are all based on direct Sanger sequencing methodology) do not exhibit extensive polymorphic sites, suggesting that they have been homogenized through the process of concerted evolution (Zimmer et al. 1980). The placements of the allopolyploids in the ITS tree in Fritsch et al. (2011) generally match those based on the plastid results, suggesting that homogenization has favored the maternal parent (the ITS placement of *Gaultheria myrsinoides* in that work is with Leucothoides but merely with weak support and is effectively unresolved with respect to Pernettya). This nuclear region therefore has limited capacity to provide insight into polyploid evolution within the Wintergreen Group, although it has facilitated the discovery of some species-level reticulate evolution in the clade, e.g. the putative hybrid origin of *G. procumbens* (Lu et al. 2010, Fritsch et al. 2011).

Our study has generated the most highly resolved phylogenetic estimate of tribe Gaultherieae and the Wintergreen Group with the strongest overall clade support, and also has provided the first evidence of allopolyploid evolution in the tribe. A follow up study could involve restriction-site Associated DNA (RAD) tags or target capture with high-throughput sequencing technology, which effectively overcome problems associated with the extensive labor involved in cloning and Sanger sequencing of multiple-copy nuclear genes in polyploid organisms (Qi et al. 2015), and thus can address phylogenetic relationships and polyploid origins within the Wintergreen Group in more detail.

*Acknowledgements* – The authors thank Boni Cruz, Brian Simison and Jun-Bo Yang for laboratory advice and assistance, and Morgan Gostel for advice on data analysis.
*Funding* – This study was supported by grants from National Natural Science Foundation of China (41671052, 31100163 to LL), Applied Fundamental Research Foundation of Yunnan Province (2017FB072 to LL), the U.S. National Science Foundation (DEB-0717711 to PWF and DEB-0717744 to KAK), and the Hundred-Talent Program of Kunming Medical Univ. (60117190444, 60118090109 to LL).
*Author contributions* – LL, PWF, DZL and HW designed research; LL analyzed data; LL, PWF, DZL and HW wrote the paper; CMB and KAK provided molecular data. The authors declare no conflict of interest.

## References

Baas, P. 1985. Comparative leaf anatomy of *Pernettya* Gaud. (Ericaceae). – Bot. Jahrb. Syst. 105: 481–495.
Bush, C. M. et al. 2009a. Phylogeny of Gaultherieae (Ericaceae: Vaccinioideae) based on DNA sequence data from *matK*, *ndhF* and *nrITS*. – Int. J. Plant Sci. 170: 355–364.
Bush, C. M. et al. 2009b. The phylogeny, biogeography and morphological evolution of *Gaultheria* (Ericaceae) from Australia and New Zealand. – Aust. Syst. Bot. 22: 229–242.
Camp, W. H. 1939. Studies in the Ericales. IV. Notes on *Chimaphila*, *Gaultheria* and *Pernettya* in Mexico and adjacent regions. – Bull. Torrey Bot. Club 66: 7–28.
Clark, J. R. et al. 1991. Nucleotide sequence of a wheat (*Triticum aestivum* L.) cDNA clone encoding the *waxy* protein. – Plant Mol. Biol. 16: 1099–1101.
Darriba, D. et al. 2012. jModelTest 2: more models, new heuristics and parallel computing. – Nat. Methods 9: 772.
Doyle, J. J. and Doyle, J. L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. – Phytochem. Bull. 19: 11–15.
Doyle, J. J. et al. 2008. Evolutionary genetics of genome merger and doubling in plants. – Annu. Rev. Genet. 42: 443–461.
Evans, R. C. and Campbell, C. S. 2002. The origin of the apple sub-family (Maloideae; Rosaceae) is clarified by DNA sequence data from duplicate GBSSI genes. – Am. J. Bot. 89: 1478–1484.
Evans, R. C. et al. 2000. The granule-bound starch synthase (GBSSI) gene in the Rosaceae: multiple loci and phylogenetic utility. – Mol. Phylogenet. Evol. 17: 388–400.
Fortune, P. M. et al. 2008. Molecular phylogeny and reticulate origins of the polyploid *Bromus* species from section *Genea* (Poaceae). – Am. J. Bot. 95: 454–464.
Foster, S. and Johnson, R. L. 2008. National Geographic desk reference to nature's medicine. – National Geographic.
Franklin, D. A. 1964. *Gaultheria* hybrids on Rainbow Mountain. – New Zeal. J. Bot. 2: 34–43.
Fritsch, P. W. et al. 2011. Phylogenetic analysis of the Wintergreen Group (Ericaceae) based on six genic regions. – Syst. Bot. 36: 990–1003.
Frohlich, M. W. and Meyerowitz, E. M. 1997. The search for flower homeotic gene homologs in basal angiosperms and Gnetales: a potential new source of data on the evolutionary origin of flowers. – Int. J. Plant Sci. 158: S131–S142.
Goldblatt, P. and Johnson, D. E. 1979 onwards. Index to plant chromosome numbers. Missouri Botanical Garden. – <http://www.tropicos.org/Project/IPCN>, accessed 8 March 2019.
Grob, G. B. J. et al. 2004. Potential phylogenetic utility of the nuclear FLORICAULA/LEAFY second intron: comparison with three chloroplast DNA regions in *Amorphophallus* (Araceae). – Mol. Phylogenet. Evol. 30: 13–23.

Huson, D. H. 1998. SplitsTree: analyzing and visualizing evolutionary data. – Bioinformatics 14: 68–73.
Huson, D. H. and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. – Mol. Biol. Evol. 23: 254–267.
Ingram, A. L. and Doyle, J. J. 2003. The origin and evolution of *Eragrostis tef* (Poaceae) and related polyploids: evidence from nuclear *waxy* and plastid *rps16*. – Am. J. Bot. 90: 116–122.
Jetz, W. et al. 2004. The coincidence of rarity and richness and the potential signature of history in centres of endemism. – Ecol. Lett. 7: 1180–1191.
Judd, W. S. et al. 2015. Plant systematics: a phylogenetic approach, 4th ed. – Sinauer.
Katoh, K. et al. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. – Nucleic Acids Res. 30: 3059–3066.
Kim, S. T. et al. 2008. Allopolyploid speciation in *Persicaria* (Polygonaceae): insights from a low-copy nuclear region. – Proc. Natl Acad. Sci. USA 105: 12370–12375.
Körner, C. 2004. Mountain biodiversity, its causes and function. – Ambio 13: 11–17.
Kron, K. A. et al. 2002. Phylogenetic classification of Ericaceae: molecular and morphological evidence. – Bot. Rev. 68: 335–423.
Lu, L. et al. 2010. Reticulate evolution, cryptic species and character convergence in the core East Asian clade of *Gaultheria* (Ericaceae). – Mol. Phylogenet. Evol. 57: 364–379.
Lu, L. et al. 2019. Data from: Allopolyploidy in the Wintergreen Group of tribe Gaultherieae (Ericaceae) inferred from low-copy nuclear genes. – Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.2gh0ht2>.
Lu, L. et al. 2019. Why is fruit colour so variable? Phylogenetic analyses reveal relationships between fruit-colour evolution, biogeography and diversification. – Global Ecol. Biogeogr. doi: 10.1111/geb.12900.
Luteyn, J. L. 1995a. *Pernettya*. – In: Luteyn, J. L. (ed.), Flora Neotropica monograph, Monograph 66. Ericaceae Part II. The superior-ovaried genera. NY Bot. Gard., pp. 365–383.
Luteyn, J. L. 1995b. *Gaultheria*. – In: Luteyn, J. L. (ed.), Flora Neotropica monograph, Monograph 66. Ericaceae Part II. The superior-ovaried genera. N.Y. Bot. Gard., pp. 384–488.
Maddison, W. P. and Maddison, D. R. 2018. Mesquite: a modular system for evolutionary analysis, ver. 3.51. – <www.mesquiteproject.org>, accessed 8 March 2019.
Mason-Gamer, R. J. 2001. Origin of the North American *Elymus* (Poaceae: Triticeae) allotetraploids based on granule-bound starch synthase gene sequences. – Syst. Bot. 26: 757–768.
Mason-Gamer, R. J. 2004. Reticulate evolution, introgression and intertribal gene capture in an allohexaploid grass. – Syst. Biol. 53: 25–37
Mason-Gamer, R. J. et al. 1998. Granule-bound starch synthase: structure, function and phylogenetic utility. – Mol. Biol. Evol. 15: 1658–1673.
Mérida, A. et al. 1990. Expression of the granule bound starch synthase I (*Waxy*) gene from snapdragon is developmentally and circadian clock regulated. – Plant Physiol. 120: 401–409.
Middleton, D. J. 1991a. Infrageneric classification of the genus *Gaultheria* L. (Ericaceae). – Bot. J. Linn. Soc. 106: 229–258.
Middleton, D. J. 1991b. Ecology, reproductive biology and hybridization in *Gaultheria* L. – Edinb. J. Bot. 48: 81–89.
Middleton, D. J. and Wilcock, C. C. 1990. Chromosome counts in *Gaultheria* and related genera. – Edinb. J. Bot. 47: 303–313.
Miller, M. A. et al. 2011. The CIPRES science gateway: a community resource for phylogenetic analyses. Proceedings of the 2011 TeraGrid Conference: extreme digital discovery. – ACM. 2011: 41.
Mirarab, S. et al. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. – Bioinformatics 30: i541–i548.
Nie, Z. L. et al. 2005. Polyploidy in the flora of the Hengduan Mountains hotspot, southwestern China. – Ann. Miss. Bot. Gard. 92: 275–306.
Nishimoto, Y. et al. 2003. Topological incongruence between nuclear and chloroplast DNA trees suggesting hybridization in the urophyllum group of the genus *Fagopyrum* (Polygonaceae). – Genes Genet. Syst. 78: 139–153.
Oh, S.-H. and Potter, D. 2003. Phylogenetic utility of the second intron of LEAFY in *Neillia* and *Stephanandra* (Rosaceae) and implications for the origin of *Stephanandra*. – Mol. Phylogenet. Evol. 29: 203–215
Oh, S.-H. and Potter, D. 2005. Molecular phylogenetic systematics and biogeography of tribe Neillieae (Rosaceae) using DNA sequences of cpDNA, rDNA and LEAFY. – Am. J. Bot. 92: 179–192.
Peralta, I. and Spooner, D. M. 2001. Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*). – Am. J. Bot. 88: 1888–1902.
Powell, E. A. and Kron, K. A. 2001. An analysis of the phylogenetic relationships in the wintergreen group (*Diplycosia*, *Gaultheria*, *Pernettya*, *Tepuia*; Ericaceae). – Syst. Bot. 26: 808–817.
Qi, Z. et al. 2015. Phylogenomics of polyploid *Fothergilla* (Hamamelidaceae) by RAD-tag based GBS – insights into species origin and effects of software pipelines. – J. Syst. Evol. 53: 432–447.
Rambaut, A. 2002. Se-Al sequence alignment editor, ver. 2.0a11. – Inst. of Evolutionary Biology. <http://tree.bio.ed.ac.uk/software/>, accessed 8 March 2019.
Rambaut, A. et al. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. – Syst. Biol. syy032.
Rauscher, J. T. 2002. Molecular phylogenetics of the *Espeletia* complex (Asteraceae): evidence from nrDNA ITS sequences on the closest relatives of an Andean adaptive radiation. – Am. J. Bot. 89: 1074–1084.
Ronquist, F. et al. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. – Syst. Biol. 61: 539–542.
Rousseau-Gueutin, M. et al. 2009. Tracking the evolutionary history of polyploidy in *Fragaria* L. (strawberry): new insights from phylogenetic analyses of low-copy nuclear genes. – Mol. Phylogenet. Evol. 51: 515–530.
Sang, T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. – Crit. Rev. Biochem. Mol. Biol. 37: 121–147.
Schmidt-Lebuhn, A. N. et al. 2010. An Andean radiation: polyploidy in the tree genus *Polylepis* (Rosaceae, Sanguisorbeae). – Plant Biol. 12: 917–926.
Schwery, O. et al. 2015. As old as the mountains: the radiations of the Ericaceae. – New Phytol. 207: 355–367.
Sleumer, H. O. 1952. Die Arten der Gattung *Gaultheria* L. in Brasilien. – Bot. Jahrb. 75: 443–450.
Smedmark, J. E. E. and Anderberg, A. A. 2007. Boreotropical migration explains hybridization between geographically distant lineages in the pantropical clade Sideroxyleae (Sapotaceae). – Am. J. Bot. 94: 1491–1505.

Smedmark, J. E. E. et al. 2003. Ancient allopolyploid speciation in Geinae (Rosaceae): evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. – Syst. Biol. 52: 374–385.
Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. – Bioinformatics 22: 2688–2690.
Tamura, K. et al. 2013. MEGA6: molecular evolutionary genetics analysis ver. 6.0. – Mol. Biol. Evol. 30: 2725–2729.
Tang, Z. et al. 2006. Biodiversity in China's mountains. – Front. Ecol. Environ. 4: 347–352.
Terborgh, J. 1992. Diversity and the tropical rain forest. – Scientific American Library.
Triplett, J. K. et al. 2014. Independent allopolyploidization events preceded speciation in the temperate and tropical woody bamboos. – New Phytol. 204: 66–73.
Wagner, W. H. 1980. Origin and philosophy of the groundplan-divergence method of cladistics. – Syst. Bot. 5: 173–193.
Wang, S. J. et al. 1999. Molecular characterization and expression of granule-bound starch synthase in the sink and source tissues of sweet potato. – Physiol. Plant. 106: 253–261.
Wang, Z. et al. 1990. Nucleotide sequence of the rice waxy gene. – Nucleic Acids. Res. 18: 5898–5898.
Wei, X. X. et al. 2010. Molecular phylogeny and biogeography of *Pseudotsuga* (Pinaceae): insights into the floristic relationship between Taiwan and its adjacent areas. – Mol. Phylogenet. Evol. 55: 776–785.
Winkworth, R. C. and Donoghue, M. J. 2004. *Viburnum* phylogeny: evidence from the duplicated nuclear gene GBSSI. – Mol. Phylogenet. Evol. 33: 109–126.
Yu, W. B. et al. 2018. The hemiparasitic plant *Phtheirospermum* (Orobanchaceae) is polyphyletic and contains cryptic species in the Hengduan Mountains of southwest China. – Front. Plant Sci. 9: 142.
Zhang, N. et al. 2012. Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. – New Phytol. 195: 923–937.
Zimmer, E. A. et al. 1980. Rapid duplication and loss of genes coding for the alpha chains of hemoglobin. – Proc. Natl Acad. Sci. USA 77: 2158–2162.
Zimmer, E. A. and Wen, J. 2015. Using nuclear gene data for plant phylogenetics: progress and prospects II. Next-gen approaches. – J. Syst. Evol. 53: 371–379.

Supplementary material (available online as Appendix njb-02077 at <www.nordicjbotany/appendix/njb-02077>). Appendix 1.