

MOLECULAR PHYLOGENETIC ANALYSIS OF THE NORTH-TEMPERATE
LABRADOR TEAS (ERICACEAE: RHODODENDRON SUBSECT. LEDUM)
SUGGESTS A COMPLEX GENETIC HISTORY

Andrew Hart

Department of Biological Sciences
Marshall University
One John Marshall Drive
Huntington, West Virginia 25705, U.S.A.
hart78@live.marshall.edu

Kathleen Kron

Department of Biology
Wake Forest University
1832 Wake Forest Road
Winston-Salem, North Carolina 27109, U.S.A.
kronka@wfu.edu

Emily Gillespie

Department of Biological Sciences
Marshall University
One John Marshall Drive
Huntington, West Virginia 25705, U.S.A.
gillespieE@marshall.edu

ABSTRACT

The Labrador teas are a group of nearly circumboreal shrubs or sub-shrubs inhabiting damp habitats. The 4–7 currently recognized species are classified within *Rhododendron* subg. *Rhododendron* section *Rhododendron* subsect. *Ledum*. In floral characters, these species are extremely similar. In vegetative characters, species limits in the Labrador teas have been difficult to determine because many of the traditionally used morphological characters vary continually across the geographic range. This study investigated evolutionary history and preliminary consideration of some species boundaries in the Labrador teas using DNA sequence data from five molecular markers to generate a preliminary phylogeny of *R.* subsect. *Ledum*. Data were analyzed using Maximum Parsimony, Maximum Likelihood and Bayesian methods. The nuclear data indicate a monophyletic subsect. *Ledum*, but chloroplast data indicate that the North American taxa have an evolutionary history separate from the European and Asian taxa, suggesting that one or both lineages of subsect. *Ledum* may be of hybrid origin. Additionally, our analyses suggest that taxa combined in recent treatments (i.e. *Rhododendron tomentosum*) represent separate lineages and should be recognized as distinct instead of included within more broadly defined species, however our current level of sampling cannot completely resolve this issue. This study lays the groundwork for future phylogenetic studies within subsect. *Ledum*, illustrating the need to sample more intensively across taxa in order to capture what appears to be a complex genetic and biogeographic history.

KEY WORDS: Biogeography, Ericaceae, Hybridization, Labrador Tea, *Ledum*, Molecular phylogenetics, *Rhododendron*

RESUMEN

Los tés del Labrador son un grupo de arbustos o subarbustos casi circumboreales que viven en hábitats húmedos. Las 4–7 especies reconocidas normalmente se clasifican en *Rhododendron* subg. *Rhododendron* sección *Rhododendron* subsect. *Ledum*. En los caracteres florales, estas especies son muy similares. En los caracteres vegetativos, los límites de las especies de los tés del Labrador han sido difíciles de establecer porque muchos de los caracteres morfológicos usados tradicionalmente varían continuamente a lo ancho de su rango geográfico. En este estudio se investigó la historia evolutiva y las consideraciones preliminares de las fronteras de especies en los tés del Labrador usando datos de secuencias de ADN de cinco marcadores moleculares para generar una filogenia preliminar de *R.* subsect. *Ledum*. Los datos se analizaron usando Máxima Parsimonia, Máxima Verosimilitud y métodos Bayesianos. Los datos nucleares indican una subsect. *Ledum* monofilética, pero los datos cloroplásticos indican que los taxa norteamericanos tienen una historia evolutiva separada de los taxa europeos y asiáticos, sugiriendo que uno o ambos linajes de la subsect. *Ledum* pueden ser de origen híbrido. Adicionalmente, nuestros análisis sugieren que taxa combinados en tratamientos recientes (ej. *Rhododendron tomentosum*) representan linajes separados y deben ser reconocidos como distintos en vez de ser incluidos en especies definidas más ampliamente, sin embargo, con nuestro nivel actual de muestreo no se puede resolver completamente este asunto. Este estudio pone las bases para futuros estudios filogenéticos en la subsect. *Ledum*, mostrando la necesidad de muestrear más intensivamente los taxa para capturar lo que parece ser una historia compleja genética y biogeográfica.

INTRODUCTION

The Labrador Teas (*Rhododendron* subsect. *Ledum*) are a small group of evergreen shrubs and sub-shrubs with an intermittently circumboreal distribution that are often dominant in damp ecosystems such as peat bogs,

muskegs, and acidic, sandy loams (Voss 2011). Labrador tea has been used in traditional herbal medicine (e.g., Dampc & Luczkiewicz 2013; Collin 2015; Dampc & Luczkiewicz 2015), with biological activity supported by a few pharmacological studies (reviewed in Popescu & Kopp 2013).

Historically, Labrador tea was recognized as the genus *Ledum* L. (tribe Rhodoreae) (Stevens 1971) due to the unusual presence of choripetalous flowers, multicellular trichomes and glands, and the absence of viscin threads in the pollen, all unlike *Rhododendron* s.s. (Copeland 1943). In their cladistic analysis of the tribe Rhodoreae, Kron and Judd (1990) included *Ledum* within *Rhododendron* based primarily on the shared characters of unique multicellular, glandular, peltate scales. Additionally, developmental studies (Leins 1964) demonstrated that the corolla of *Ledum* species is actually sympetalous.

Three related issues have made the species-level taxonomy of the Labrador teas complex. First, morphological cladistic analyses of the tribe Rhodoreae suggested that *Rhododendron* was paraphyletic if *Ledum* was separately recognized and these results prompted Kron and Judd (1990) to include *Ledum* within the genus *Rhododendron* as *Rhododendron* subg. *Rhododendron* sect. *Rhododendron* subsect. *Ledum*. Molecular phylogenetic studies have continued to support the inclusion of *Ledum* within *Rhododendron* (Kron 1997; Kurashige et al. 2001; Goetsch et al. 2005) but Labrador teas are still frequently listed in some floras (e.g. Flora Europaea and Flora of China) and by amateur botanists under the name "*Ledum*". Kron and Judd (1990) made new combinations for two species of Labrador tea (*R. palustre* (L.) Kron & Judd; *R. groenlandicum* (Oeder) Kron & Judd), but did not at that time include new combinations for other Labrador tea species. In response, additional new combinations were subsequently made by Harmaja (1990, 1991, 1999, 2002), but the process of transferring all *Ledum* to *Rhododendron* took a decade. Second, simple nomenclatural transfer to *Rhododendron* was not possible in some cases due to existing use of epithets within *Rhododendron*. For example, *Rhododendron macrophyllum* is an existing accepted name, and so *Ledum macrophyllum* Tolm. was renamed *Rhododendron tolmachevii* Harmaja. Third, major floras treat this small group quite differently. For example, the recent treatment of *Rhododendron* in the *Flora of North America* (FNA) (Judd & Kron 2009) combined several taxa due to their morphological similarity and the absence of reasonably clear morphological gaps with which to recognize species boundaries. Specifically, *Rhododendron neoglandulosum* Harmaja was included within *Rhododendron columbianum* (Piper) Harmaja and *Rhododendron palustre* L., *Rhododendron subarcticum* Harmaja, *Rhododendron subulatum* (Nakai) Harmaja (Harmaja 2002), *Rhododendron tolmachevii* Harmaja, and *Rhododendron diversipilosum* (Nakai) Harmaja were all included within *R. tomentosum* Harmaja. Only the circumscription of *Rhododendron groenlandicum* (Oeder) Kron & Judd was left unmodified in FNA. Because FNA focuses on North American taxa, this approach may obscure how variation on other continents is recognized if one relies on this manual when examining species that occur on continents other than North America, such as *R. tomentosum*. In contrast, the *Flora of China* (FoC) (Ruizhang & Chamberlain 2005) includes a single species, *Ledum palustre*, whose four recognized subspecies have been considered to be three separate species by Harmaja (*R. subulatum*, *R. subarcticum* and *R. tomentosum*) but all combined within *R. tomentosum* in FNA. *Flora Europaea* (Tutin et al. 1972) recognizes *Ledum groenlandicum* and several subspecies of *L. palustre*; like FoC, *Flora Europaea* does not include *Ledum* within *Rhododendron*. These numerous nomenclatural changes in a relatively short period of time and markedly different handling by commonly used floras, in combination with being found in remote locations resulting in likely incomplete distributional information, make this small group of arctic plants challenging to study phylogenetically.

The current study provides a preliminary species-level molecular phylogeny of *Rhododendron* subsect. *Ledum* using five markers from both the nucleus and chloroplast. This phylogeny represents a necessary starting point for future fine-scale investigations of species boundaries and historical biogeography by providing strong evidence of a reticulate evolutionary history as well as identifying non-monophyletic taxa that have been combined taxonomically (e.g., *Rhododendron tomentosum*), which should be investigated by using less inclusive taxonomic/biological entities in future efforts. This study also provides a basis for future biogeographic studies of north temperate Asia, Europe and North America.

METHODS

Taxon sampling.—Included exemplars/species, voucher information and Genbank numbers are reported in Table 1. Ingroup taxa were sampled following Harmaja (1991, 1999, 2002), who recognizes more species than are in FNA (Judd & Kron 2009). *Rhododendron columbianum* (Piper) Harmaja and *Rhododendron neoglandulosum* Harmaja were both included and these are together recognized in FNA (Judd & Kron 2009) as *Rhododendron columbianum* (Piper) Harmaja. *Rhododendron groenlandicum* (Oeder) Kron & Judd was included and is also recognized in FNA. *Rhododendron diversipilosum* (Nakai) Harmaja, *Rhododendron hypoleucum* (Kom.) Harmaja, *Rhododendron tolmachevii* Harmaja, and *Rhododendron tomentosum* Harmaja s.s. were included and all four of these taxa are together recognized in FNA (Judd & Kron 2009) as *Rhododendron tomentosum* Harmaja. Additionally, *Rhododendron palustre* L. was included from Siberia following FoC (Ruizhang & Chamberlain 2005). *Rhododendron subulatum* (Nakai) Harmaja and *Rhododendron subarcticum* Harmaja were not sampled due to inadequate tissue for DNA extraction. *Rhododendron subulatum* is considered likely morphologically indistinguishable from *Rhododendron palustre* in FoC (Ruizhang & Chamberlain 2005) and *R. subarcticum* is considered to be synonymous with *R. tomentosum* s.l. in FNA, but these two taxonomic problems each require further clarification.

Previous studies (Kron & Judd 1990; Kurashige et al. 1998) have placed subsect. *Ledum* as a clade within the lepidote lineage *Rhododendron* subg. *Rhododendron* L. (Cullen 1980). Therefore, outgroup sampling included representatives from five additional *Rhododendron* species from this subgenus, including *Rhododendron baileyi* Balf.f., *Rhododendron afghanicum* Aitch. & Hemsl., *Rhododendron sargentianum* Rehder & E.H. Wilson, *Rhododendron orbiculatum* Ridley and *Rhododendron hippophaeoides* Balf. f. & Sm. Five lepidote *Rhododendron* subgenera (following Goetsch et al. 2005) were represented by *Rhododendron quinquefolium* Bisset & Moore (subg. *Azaleastrum* Planch.), *Rhododendron albiflorum* Hook. (subg. *Candidastrum* Franch.), *Rhododendron arborescens* (Pursh) Torr. (subg. *Pentanthera* (G. Don) Pojarkova), *Rhododendron grande* Wight (subg. *Hymenanthes* (Blume) K. Koch) and *Rhododendron camtschaticum* (Pall.) Sm. (subg. *Therorhodon* (Maxim.) A. Gray). Two closely related non-*Rhododendron* tribes were represented by *Ceratiola ericoides* Michx. (tribe Empetreae Horan.) and *Calluna vulgaris* (L.) Hull (tribe Ericaceae DC. ex Duby). *Cassiope mertensiana* G. Don was included to represent the Cassiopoideae H.T. Cox ex P.F. Stevens, the most closely related subfamily to Ericoideae (Kron & Chase 1993; Kron 1997; Kron et al. 2002; Gillespie & Kron 2010).

Molecular methods.—Total DNA was extracted from fresh samples, herbarium specimens or silica-dried material using a modified cetyltrimethyl ammonium bromide (CTAB) protocol (Doyle & Doyle 1987) or the DNeasy Plant Mini Kit (Qiagen, Valencia CA, USA) with modifications following Drábková et al. (2002).

Chloroplast regions 3'*trnV*^(UAC)-*ndhC*, *trnS*^(GCU)-*trnG*^(UUC) (both from the Large Single Copy region), *rpL32-trnL*^(UAG) (from the Small Single Copy region), and nuclear regions GBSS-1/*waxy* (exons 9–11) and nuclear ribosomal Internal Transcribed Spacer (nrITS) were utilized. Chloroplast regions were selected based on their predicted utility as described in Shaw et al. (2007) and amplified using Shaw's recommended primers and PCR parameters. The low-copy nuclear gene GBSS-1/*waxy* (exons 9–11) (Evans et al. 2000) and nrITS (Baldwin 1992) have each been shown to be valuable for low-level taxonomic studies despite potential complications associated with nrITS, addressed below. Primer selection and PCR amplification of nrITS and *waxy* followed Gillespie and Kron (2010). QiagenTM QIAquick Gel Isolation Kit was used to isolate amplified fragments (Qiagen, Valencia CA, USA).

Concerns have been raised in recent years (e.g., Álvarez & Wendel 2003; Poczai & Hyvönen 2010) about the phylogenetic utility of nrITS in plant phylogenetic studies, primarily because of the possibility of amplifying paralogous sequences. Although careful use of this region is clearly warranted, nrITS is well-understood and represents a useful source of nuclear DNA data in *Rhododendron*. Here, a combined strategy of spot-cloning (five randomly-chosen clones per taxon, for at least one taxon per subgenus) and vigilant examination of chromatogram data was used to minimize the likelihood of employing paralogous sequences. Specifically, nrITS sequences were examined 'blind' (i.e. individually, outside the context of a data matrix) as well as within the context of the complete data matrix, so that chromatograms at particularly variable regions could be

TABLE 1. Taxon sampling, voucher information and Genbank accessions. ARS = American Rhododendron Society Seed Exchange Program; BOON = Appalachian State University Herbarium; MUHW = Marshall University Herbarium; RBGE or E = Royal Botanical Garden, Edinburgh; RBGK or K = Royal Botanic Garden, Kew; Rhodogarden = Private garden maintained by K. Theqvist in Turku, Finland; RSF = Rhododendron Species Foundation; S = Swedish Museum of Natural History Herbarium; WFU = Wake Forest University Herbarium; WTU = University of Washington Herbarium/Burke Museum of Natural History and Culture. For the ingroup (*Rhododendron* subsect. *Ledum*) the taxonomic concept followed is indicated. Bold numbers indicate sequences newly generated for this study.

Species	Specimen Authority	Genbank/Accession Concept	Source or Collector	Collection #	Deposition	Geography	nITS	GBSS-1/wxy	rpl32-tml	trnS-trnG	trnN-ndhC
<i>Rhododendron</i> subg.											
<i>Rhododendron</i>		Chamberlain 1996									
<i>R. columbianum</i>	L. (Piper) Harmaja	Harmaja 1991	Arnot	73	WTU	Washington, USA	KY356327	KY356312	KY356265	KY356280	KY356300
<i>R. neoglandulosum</i>	Harmaja	Harmaja 1991	Kron	s.n.	WFU	California, USA	KY356326	-	KY356264	KY356279	KY356299
<i>R. diversipilosum</i>	(Nakai) Harmaja	Harmaja 1999	Theqvist	2004-0770 garden	Rhodo-	Japan	KY356322	KY356308	KY356259	KY356274	KY356294
<i>R. groenlandicum</i>	(Oeder) Kron & Judd	Harmaja 1991	Habegger	399	WTU	Washington, USA	KY356328	KY356313	KY356266	KY356281	KY356301
<i>R. hypoleucum</i>	(Kom.) Harmaja	Harmaja 1991	Larsen	87/04	ARS	NE Asia	KY356321	KY356307	KY356260	KY356275	KY356295
<i>R. tolmachevii</i>	(Tolm.) Harmaja	Harmaja 1991	Theqvist	2004-0806 garden	Rhodo-	Amur, Russia	KY356323	KY356309	KY356261	KY356276	KY356296
<i>R. palustre</i>	(L.) Kron & Judd	Kron & Judd 1990	Chase	MWC869	K	Kurai, Siberia	KY356325	KY356310	KY356262	KY356277	KY356298
<i>R. tomentosum</i> ssp. <i>tomentosum</i>	(Stokes) Harmaja Balfour	Harmaja 1991	Sakata	1978-03135A	E	Japan	KY356324	KY356311	KY356263	KY356278	KY356297
<i>R. baileyi</i>			RSF	64/146	E	Xizang, China	KY356318	KY356306	KY356256	KY356271	-
<i>R. sargentianum</i>	Rehder & E.H. Wilson		RSF	74/104	E	Sichuan, China	KY356320	KY356305	KY356253	KY356268	KY356290
<i>R. hippophaeoides</i>	Balf. F. & Sm.		RBGE	1932-1022	E	Yunnan, China	GU176634	GU176667	KY356258	GU176707	KY356291
<i>R. orbiculatum</i>	Ridl.		RSF	80/004	E	Malaysia	KY356319	-	KY356257	KY356272	KY356292
<i>R. afghanicum</i>	Aitsch. & Hemsli.		RSF	80/083	E	Afghanistan	KY356314	-	KY356255	KY356273	KY356293
Subg. Azaleastrum	Planch.	Chamberlain 1996									
<i>R. quinquefolium</i>	Bisset & Moore		RSF	80/028	E	Japan	KY356316	KY356303	KY356251	KY356267	KY356289
Subg. Candidastrum	Franch.	Chamberlain 1996									
<i>R. albiflorum</i>	Hook.		Gillespie	08-113	WFU	Washington, USA	KY356315	KY356302	KY356250	KY356269	KY356288

TABLE 1. continued

Species	Specimen Authority	Genbank/Accession Concept	Source or Collector	Collection #	Deposition	Geography	nrITS	GBSS-1/waxy	rpl32-tml	trnS-trnG	trnV-ndhC
Subg. Pentanthera		Chamberlain 1996									
<i>R. arborescens</i>	(G. Don) Pojarkova (Pursch) Torr.	Chamberlain 1996	Brant	MUHW018651 BOON23743	MUHW BOON	N. Carolina, USA	- KY356317	KY356304 - -	KY356254 - -	- KY356270	KY356287- -
Subg. Hymenanthes		Chamberlain 1996									
<i>R. grande</i>	(Blume) K. Koch Wight	Chamberlain 1996	RBGE	1969-8606	E	cultivated	GU176633	EU669886	KY356252	GU176706	KY356286
Subg. Therorhodion	(Maxim.) A. Gray	Chamberlain 1996									
<i>R. camtschaticum</i>	(Pall.) Sm. Tribe Empetreae	Kron et al. 2002	RSF	73/054	E	Japan	GU176637	DQ000608	KY356249	GU176710	KY356285
<i>Ceratiola ericoides</i>	Michx. Tribe Ericaceae	Kron et al. 2002	Kron	2069	WFU	Florida, USA	AF519552	DQ000599	KY356248	GU176684	KY356284
<i>Calluna vulgaris</i>	(L.) Hull	Kron et al. 2002	RBGE	1972-1443	E	cultivated	GU176627	GU176655	KY356247	GU176687	KY356283
Subfam. Cassiopoideae	H.T. Cox ex P.F. Stevens	Kron et al. 2002									
<i>Cassiope mertensiana</i>	G. Don	Anderberg		75-83	S	cultivated	AF419798	DQ000598	KY356246	GU176710	KY356282

specifically examined for evidence of multiple amplicons. No evidence of paralogs for nrITS or waxy/GBSS-I was discovered for this study. Gel-isolated amplicons or cloned plasmids were sequenced at the Nevada Genomics Center (Reno, NV, USA) or at Eton Biosciences (Raleigh, NC, USA).

Geneious versions 8.1 through 9.1.4 (Biomatters, Ltd.) (<http://www.geneious.com>; Kearse et al. 2012) were used in the storage and curation, manual alignment, and minimal editing of DNA sequences. Any ambiguous base calls were checked against the chromatogram and conservatively edited. Bases at each end of the raw sequence with more than a 5% probability of base-call error or peak quality lower than Q20 were trimmed and excluded.

Phylogenetic analyses.—Maximum Parsimony (MP) analyses were conducted on separate matrices of combined nuclear data and combined chloroplast data using PAUP*4.0a147, Jan. 2016 release (Swofford 2003) with the following parameters: Phylogenetically uninformative (i.e., non-variable or autapomorphic) characters were excluded, informative characters were equally weighted and unordered, gaps were treated as missing data, and multiple independent heuristic searches were conducted with TBR branch swapping and 1000 random stepwise addition replicates. Clade support was evaluated using bootstrap analysis (Felsenstein 1985; Felsenstein 1988) with 10,000 replicates. *Cassiope mertensiana* was identified as the outgroup taxon.

Maximum Likelihood (ML) analyses were carried out using the online RAxML tool (Stamatakis et al. 2014) through the Cipres Web Portal (http://www.phylo.org/sub_sections/portal/). Matrices of combined nuclear and combined chloroplast data were analyzed, with datasets partitioned as separate DNA regions under the GAMMA model recommended for smaller sample sizes, with *Cassiope mertensiana* identified as the outgroup taxon.

Bayesian analyses were run based on models tested for individual markers within jModelTest 2.1.3 (Guindon & Gascuel 2003; Durrin et al. 2012) using the corrected Aikake Information Criterion (AICc) which corrects for small sample size (fewer than approximately 40 terminals). Bayesian MCMC analysis (Yang & Rannala 1997) was conducted on combined, partitioned nuclear and chloroplast data as implemented in Mr. Bayes version 3.2.6 × 86 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Analyses consisted of 50,000,000 generations with a burn-in of 25% of the tree set and a split deviation of <0.01.

The Incongruence Length Difference (ILD) test (Farris et al. 1995), implemented as the Partition Homogeneity Test in PAUP*4.0a147 (Swofford 2003), was conducted to determine whether nuclear and chloroplast data recovered sufficiently congruent phylogenies to be combined into a total molecular evidence analysis. A heuristic search with 1000 replicates was performed using TBR branch swapping and simple, stepwise addition of taxa.

RESULTS

The aligned length, best fitting evolutionary model according to jModeltest, and amount of missing data (number of taxa and percent of taxa) for each partition are reported in Table 2. For chloroplast and nuclear data, branch support values are presented in the following format: Bayesian posterior probability / ML bootstrap / MP bootstrap. Bootstrap values (MP or ML) below 50% and Bayesian posteriors below 0.90 are shown as dashes (-) and are considered here to be essentially non-existent branch support. Data matrices for individual markers, combined nuclear and chloroplast matrices and both nuclear and chloroplast phylogenies are deposited in TreeBASE (TreeBASE.org) under number S19409.

Nuclear data (Fig. 1).—Analyses of combined nuclear data (nrITS and waxy exons 9–11 strongly support (0.99 / 87 / 77) a monophyletic subsect. *Ledum* (hereafter ‘*Ledum* clade’). This clade plus the remaining lepidote taxa sampled (*R. afghanicum*, *R. baileyi*, *R. hippophaeoides*, *R. orbiculatum*, and *R. sargentianum*) form a well-supported clade (0.97 / 90 / 72) and these two clades are sister to each other (1.00 / 98 / 96).

Within the *Ledum* clade, *R. groenlandicum* is sister to *R. hypoleucum* (1.00 / 95 / 79) and these two taxa are moderately to weakly supported as sister to *R. tomentosum* (0.91 / 65 / 51). *Rhododendron columbianum* + *R. tolmachevii* (- / 73 / 63) are together resolved as sister to *R. groenlandicum* + *R. hypoleucum* + *R. tomentosum*, but with fairly weak support in ML and MP only (- / 65 / 57). These five taxa are sister the *R. diversipilosum* + *R.*

TABLE 2. Marker, aligned length (in base pairs), best-fitting evolutionary model, and number of missing taxa (and %). * indicates that a single taxon has an incomplete sequence.

Marker	Aligned Length (bp)	Model (AICc)	# Missing taxa (%)
nrITS	617	TPM1uf+G	0
GBSS-1/waxy	570	HKY+G	3 (14.2%)*
<i>trnS</i> ^(GCU) - <i>trnG</i> ^(UUC)	1374	TVM+G	0
<i>rpl32-trnL</i> ^(UAG)	848	TVM+G	0
3' <i>trnV</i> ^(UAC) - <i>ndhC</i>	490	HKY+G	1 (4.7%)

palustre clade (0.92 / 48 / 57) but again with weak support from only ML and MP analyses (- / 69 / 64). The inset phylogram showing branch lengths was generated from the same Bayesian topology as the cladogram and the terminals are in the same order.

Chloroplast data (Fig. 2).—In contrast to the nuclear results, analyses of the combined chloroplast data (*trnS*^(GCU)-*trnG*^(UUC), *rpl32-trnL*^(UAG), and 3'*trnV*^(UAC)-*ndhC*) indicate that *R. subsect. Ledum* is not monophyletic. *Rhododendron diversipilosum*, *R. hypoleucum*, *R. palustre*, *R. tolmachevii* and *R. tomentosum*, are resolved as a clade with very strong support (1.00 / 100 / 99). This clade of *R. subsect. Ledum* is sister to *R. sargentianum* + *R. baileyi* (0.94 / 78 / 58) with Bayesian posterior support (0.90 / - / 50). The other three *R. subsect. Ledum* representatives (*R. groenlandicum*, *R. neoglandulosum*, *R. columbianum*) are resolved as a clade with very strong support (1.00 / 100 / 100), and they are nested within a clade of other sampled lepidote *Rhododendron* including *R. orbiculatum*, *R. afghanicum* and *R. hippophaeoides* with moderate to strong support (0.91 / 86 / 66). This clade of six lepidote *Rhododendron* are sister to *R. albiflorum* (0.94 / 78 / 76) and then *R. quinquefolium* (1.00 / 99 / 87). All of these species form a clade along with *R. arborescens* (subg. *Pentanthera*) and *R. grande* (subg. *Hymenantes*) with strong support (1.00 / 100 / 99). Finally, *R. camtschaticum* (subg. *Therorhodion*) is strongly supported (1.00 / 98 / 98) as sister to all other *Rhododendron*.

In addition to numerous single-nucleotide indels, two longer sections of DNA in the chloroplast data support the close relationship of the North American subsect. *Ledum* species (*R. columbianum*, *R. groenlandicum* and *R. neoglandulosum*) to *Rhododendron* species outside subsect. *Ledum* rather than to the Asian subsect. *Ledum* species (*R. diversipilosum*, *R. tomentosum*, *R. hypoleucum*, *R. tolmachevii* and *R. palustre*). In the *trnS*^(GCU)-*trnG*^(UUC) intergenic spacer, bases 506–514 of (5'-AATTCCTAA-3') is a likely insertion in the Asian subsect. *Ledum* species that is absent in all other taxa sampled, including the three North America subsect. *Ledum* species. In the *rpl32-trnL*^(UAG) intergenic spacer, bases 274–299 (5'-GAAAAAAGAATAAACAGAAGATACAA-3') is found in the North American subsect. *Ledum* taxa and in *R. orbiculatum*, *R. hippophaeoides*, *R. albiflorum* and *R. quinquefolium*, all of which form a strongly supported clade (1.00 / 99 / 87). Most other *Rhododendron* taxa, including the Asian subsect. *Ledum* taxa, have the sequence 5'-TGATCTTCTGTTTATTCTTTTTTC-3' in the same position.

The ILD test indicated that the nuclear and chloroplast data partitions were in conflict ($p = 0.01$) and should not be combined for total evidence analyses. Valid questions about the relevance of the ILD test in phylogenetics have been raised (e.g., Cunningham 1997; Yoder et al. 2001; Barker & Lutzoni 2002), due in part to the likelihood of making a Type I error (i.e., failing to combine the datasets when they are not truly incongruent—i.e. a 'soft' polytomy), or disagreement about the most appropriate p -value for this test (Cunningham 1997). In the current study, however, the conflicting nodes in the nuclear versus chloroplast analyses are very strongly supported. Analyses of combined total data (i.e., nuclear + chloroplast) resulted in a nearly complete loss of resolution and support in the region of the conflicts and are therefore not informative (and not shown).

DISCUSSION

This study is the first to use molecular data to address the evolutionary history of *Rhododendron* subsect. *Ledum*, a small group of geographically widespread, yet morphologically similar species. The number of taxa in

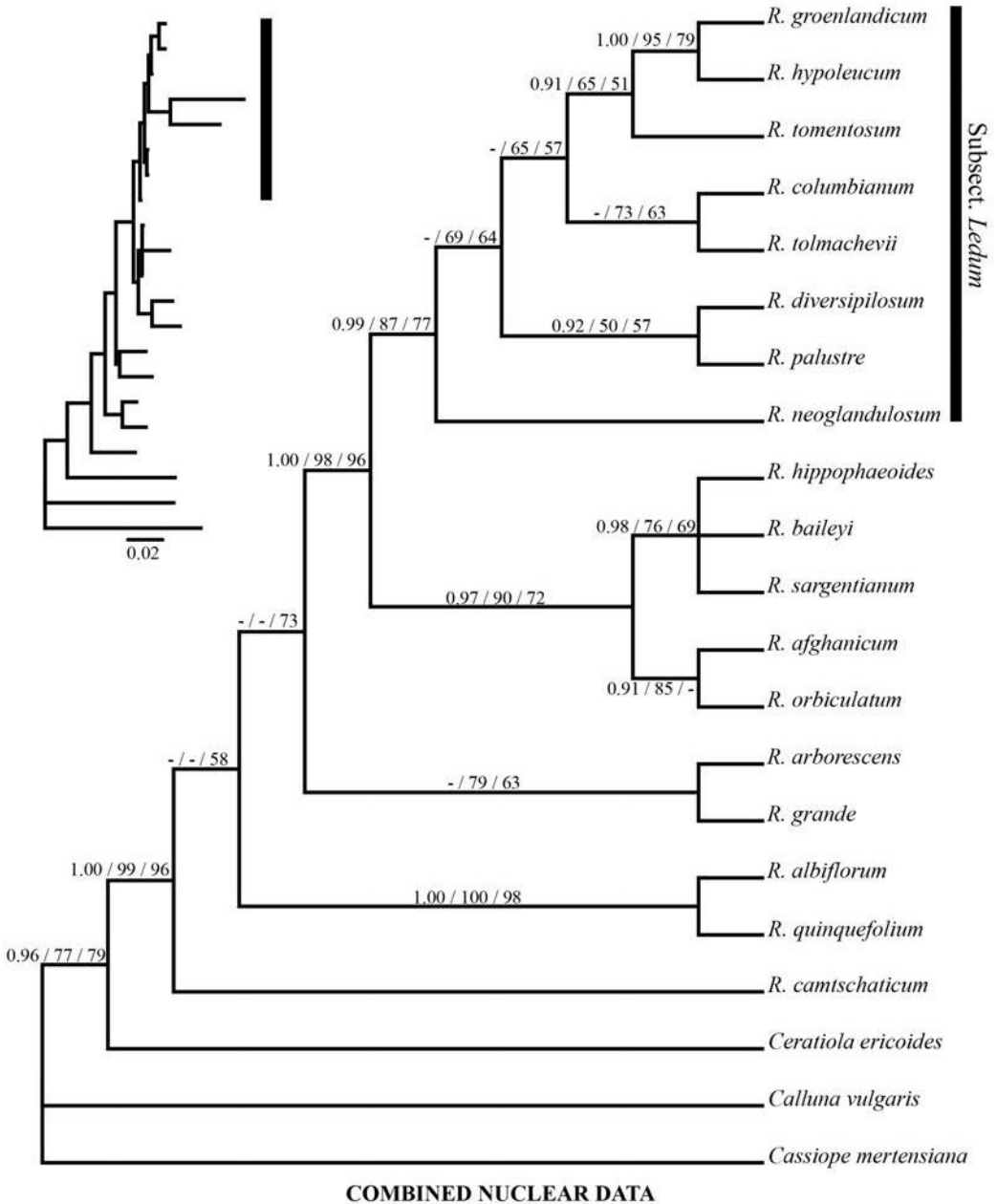


Fig. 1. Total combined nuclear data (nrITS and exons 9-11 of GBSS-1/waxy). Support values are to the left of nodes in the format (Bayesian posterior probability/ML bootstrap/MP bootstrap). Support derived from 50 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Parsimony tree length = 225, CI = 0.6533, RI = 0.6977. RAxML Likelihood score = -4603.601284. Black bar indicates *Rhododendron* subsect. *Ledum*. The Bayesian topology is shown for both the cladogram and the inset phylogram.

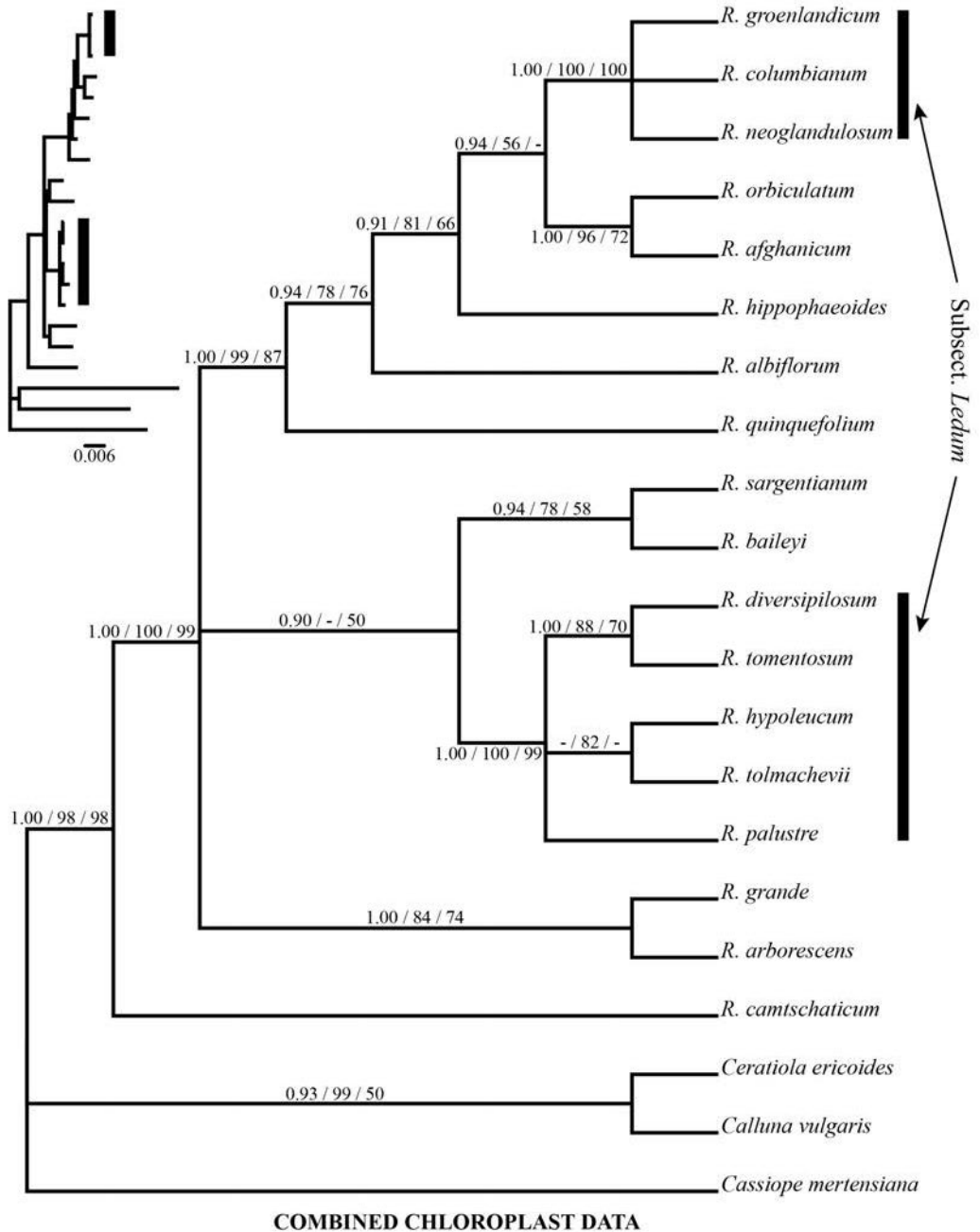


FIG. 2. Total combined chloroplast data (*trnS*^(GCU)-*trnG*^(UUC), *rpl32-trnL*^(UAG) and 3'*trnV*^(UAC)-*ndhC*). Support values are to the left of nodes in the format (Bayesian posterior probability/ML bootstrap/MP bootstrap). Support derived from 50 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Parsimony tree length = 223, CI = 0.6278, RI = 0.7629. RAxML Likelihood score = -6863.805129. Black bar indicates *Rhododendron* subsect. *Ledum*. The Bayesian topology is shown for both the cladogram and the inset phylogram.

this lineage has been the subject of debate and the phylogeny we present here is a necessary first step to addressing finer-scale questions in the future, especially questions related to species boundaries and genome evolution.

Overall, analyses of combined nuclear and combined chloroplast data yield very good resolution and support despite limited sampling outside subsection *Ledum* relative to the size and diversity of the *Rhododendron* lineage (~1,000 species). The most prominent finding here is that information from the nuclear and chloroplast genomes of *R. subsect. Ledum* do not indicate the same evolutionary history. The nuclear data strongly indicate a monophyletic *R. subsect. Ledum* that is sister to a clade containing lepidote species from other subgroups within the genus, but the combined chloroplast data strongly indicate that the primarily North American taxa, *R. groenlandicum*, *R. columbianum* and *R. neoglandulosum*, are more closely related to species of *Rhododendron* from other subsections than to the other members of *R. subsect. Ledum*. A total of 35 potentially parsimony-informative (PPI) characters in the chloroplast genome (26.1% of the total number of PPI characters) support the conclusion that the North American *R. subsect. Ledum* taxa are more closely related to other subsections than to other taxa of *R. subsect. Ledum*. This result may suggest that these taxa are of hybrid origin with egg parents from other subsections within the *R. subg. Rhododendron*, specifically subsection *Afghanica* (represented here by *R. afghanicum*), subsection *Lapponicum* (represented by *R. hippophaeoides*) or from subg. *Vireya* (represented by *R. orbiculatum*). The Asian subsection *Ledum* taxa may be more closely related to section *Anthopogon* (represented here by *R. sargentianum*) or to subsection *Baileya* (represented here by *R. baileyi*). These results permit an alternate hypothesis, that it is the Asian subsection *Ledum* taxa which are of hybrid origin, their ancestor having acquired its chloroplast genome from a species of subsection *Baileya* or sect. *Anthopogon*. Still a third scenario is possible, where both subsection *Ledum* clades are of hybrid origin with maternal parents from different lepidote lineages. The current study included representatives of six *Rhododendron* subgenera and three subsections within subg. *Rhododendron* in addition to subsection *Ledum*. However, in such a large genus as *Rhododendron*, this sampling is not sufficient to further address the question of parentage and was beyond the scope of this initial attempt to reconstruct the evolutionary history of subsection *Ledum*. In order to clarify which subsection *Ledum* clade is of hybrid origin (or potentially, both clades), future sampling must include representatives of most if not all of the 29 currently recognized subsections of *Rhododendron* subg. *Rhododendron*. Many lepidote *Rhododendron* species are geographically near by or even overlapping with subsection *Ledum*.

Some insights into the biogeographic history of subsection *Ledum* are possible from the current results. Not all nodes are strongly supported, but several are, and these allow formation of biogeographical hypotheses that can be addressed in future studies with additional sampling. Nuclear data strongly suggest that *R. groenlandicum* (widespread in northern North America and present in Europe) is most closely related to *R. tomentosum* and *R. hypoleucum* (from Japan and Norway, respectively). Other nodes within the nuclear data are not well-supported, including nodes that place the other North American taxa (*R. columbianum* and *R. neoglandulosum*). Chloroplast data unambiguously place *R. groenlandicum* in a clade with *R. columbianum* and *R. neoglandulosum* (both from the Pacific Northwest of North America). Since the three North American taxa are most closely related to *R. orbiculatum* and *R. afghanicum*, this may suggest that the ancestor of these three taxa acquired its chloroplast genome from a lepidote ancestor related to these two species of *Rhododendron*. This further suggests that this hybridization event was coincident with migration into North America from Europe or Asia. The direction of migration cannot currently be resolved with confidence, but a picture is beginning to emerge. Matthews (1979) proposed that a belt of arctic biome existed in the northern hemisphere by approximately three million years ago, providing new niches for ancestors inhabiting mountain chains farther south (e.g. Hultén, 1937) to migrate into, during glacial minima. Certainly the center of diversity of *Rhododendron* supports this scenario, with many species occurring in and around the Himalayas. Taken together, this information may suggest that during the last three million years, an ancestor of the modern-day North American subsection *Ledum* species acquired a chloroplast genome from a mountain-inhabiting lepidote *Rhododendron*, possibly in the Himalayas, and migrated across Beringia into North America where speciation occurred, generating a widespread taxon (*R. groenlandicum*) and two more narrowly distributed taxa (*R. columbianum* and *R. neoglandulo-*

sum). In Asia, differentiation into extremely similar forms (*R. tomentosum* s.l.) with or without hybridization has occurred during the same period of time. These species are all very closely related, as illustrated by their branch lengths and the ability to form artificial hybrids (Kihlman 2004), but the Eurasian group and the North American group may be on fairly separate evolutionary trajectories at this time.

The ploidy level of all subsect. *Ledum* taxa have not yet been determined, but Lantai & Kihlman (1995) examined two formerly recognized *Ledum* taxa, *L. palustre* L. and *L. palustre* L. ssp. *decumbens* (Ait.) Hultén, both currently included within *Rhododendron tomentosum* Harmaja in FNA (Judd & Kron 2009). Lantai and Kihlman found that *L. palustre* s.s. from Sweden was a polyploid ($2n=52$) but that *L. palustre* ssp. *decumbens* was diploid ($2n=26$). The current study included an accession of *R. tomentosum* (= *L. palustre*) from Siberia and found no evidence of polyploidy in the chromatograms. Kristian Theqvist (unpubl. data) has found using flow cytometry that *R. diversipilosum*, *R. groenlandicum* and *R. hypoleucum* are likely polyploids. Our study determined that *R. groenlandicum* is possibly of hybrid origin and therefore possibly an allopolyploid, however investigation of chromatograms revealed no evidence of paralogs in any DNA marker used in this study. In the current study, *R. diversipilosum* and *R. hypoleucum* were placed unambiguously within the 'tomentosum' subclade using chloroplast data and with all subsect. *Ledum* taxa in the nuclear data, exhibiting no phylogenetic behavior generally consistent with a hybrid origin. Additionally, no evidence of hybridization was discovered in the chromatograms of these two species. These results may illustrate that the phylogenetic behavior of polyploids is unpredictable in phylogenetic analyses, or perhaps a hybridization event occurred recently enough that more intensive sampling would be required to reveal such behavior. An additional or alternative interpretation is that both in the current study and in past attempts to understand ploidy levels within subsect. *Ledum*, sampling density has not captured the complexity of the genetic landscape. In other words, the taxonomic complexity of this group may in fact reflect the biological reality that small pockets of populations may be morphologically virtually indistinguishable but genetically divergent. Future studies should take into account the possibility that terminals representing entire species or even subspecies/varieties may not be sufficient for capturing genetic diversity, and ideally multiple terminals per biological entity should be included.

Clearly, future investigations of this group must include dense sampling across the range of most if not all currently or formally recognized species, in order to fully describe its complex genetic and biogeographic history, particularly within the context of post-glacial migration and northern hemisphere vicariance and dispersal. Specific goals should include, 1) Sampling of as many of the 29 current *Rhododendron* subsections as possible. This will permit clarification regarding whether one or both subsect. *Ledum* lineages are of hybrid origin and to specify the nearest related subsection and therefore the likely source(s) of their chloroplast genome, 2) A herbarium-based accounting of each subsect. *Ledum* species' distribution. Now that many herbaria are becoming digitized, it is likely that distributions of these remote taxa may be expanded beyond 'broad brush' dot maps and this may inform sampling and/or interpretation of biogeography, and 3) Extensive sampling of extant taxa across their ranges, including where overlap with other subsect. *Ledum* species occur. It will be important in future studies to treat this small but complex group of plants in a more fine-scale way, given the intra-species genetic structure we have presented here, in order to fully illuminate their evolutionary history.

ACKNOWLEDGMENTS

The authors wish to thank Marshall University for startup funding for Gillespie in addition to the Marshall University Yeager Scholar Program for providing funding for Hart to attending conferences and present this work. Kristian Theqvist (Finnish *Rhododendron* Society) provided important plant tissue. Staff at the University of Washington Herbarium (WTU) Appalachian State University herbarium (BOON) and Wake Forest University Herbarium (WFU) provided some specimens for sampling. We thank Walter Judd and one anonymous reviewer for their helpful comments and suggestions on the manuscript.

REFERENCES

- ÁLVAREZ, I. & J.F. WENDEL. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molec. Phylogen. Evol.* 29:417–434.
- BALDWIN, B. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. *Molec. Phylogen. Evol.* 1:3–16.
- BARKER, F. & F. LUTZONI. 2002. The utility of the incongruence length difference test. *Syst. Biol.* 51:625–637.
- COLLIN, G. 2015. Aromas from Quebec. IV. Chemical composition of the essential oil of *Ledum groenlandicum*: A review. *Amer. J. Essential Oils Nat. Prod.* 2(3):6–11.
- COPELAND, H. 1943. A study, anatomical and taxonomic of the genera of Rhododendroideae. *Amer. Midl. Naturalist* 30:533–625.
- CULLEN, J. 1980. A revision of *Rhododendron* 1. Subgenus *Rhododendron* Sections *Rhododendron* and *Pogonanthum*. *Roy. Bot. Gard. Edinburgh* 39:1–207.
- CUNNINGHAM, C.W. 1997. Can three incongruence tests predict when data should be combined? *Molec. Biol. Evol.* 14:733–740.
- DAMP, A. & M. LUCZKIEWICZ. 2013. *Rhododendron tomentosum* (*Ledum palustre*). A review of traditional use based on current research. *Fitoterapia* 85:130–143.
- DAMP, A. & M. LUCZKIEWICZ. 2015. Labrador tea—the aromatic beverage and spice: A review of origin, processing and safety. *J. Sci. Food Agric.* 95:1577–1583. doi:10.1002/jsfa.6889
- DARRIBA, D., G.L. TABOADA, R. DOALLO, & D. POSADA. 2012. jModelTest 2: More models, new heuristics and parallel computing. *Nature Meth.* 9:772.
- DOYLE, J. & J. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull. Bot. Soc. Amer.* 19:11–15.
- DRÁBKOVÁ, L., J. KIRSCHNER, & Č. VLČEK. 2002. Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens in Juncaceae. *Pl. Molec. Biol. Reporter* 20:161–175.
- EVANS, R.C., L.A. ALICE, C.S. CAMPBELL, E.A. KELLOGG, & T.A. DICKINSON. 2000. The granule-bound starch synthase (GBSSI) gene in the Rosaceae: Multiple loci and phylogenetic utility. *Molec. Phylogen. Evol.* 17(3):388–400.
- FARRIS, J., S.M. KALLERSJO, A.G. KLUGE, & C. BULT. 1995. Constructing a significance test for incongruence. *Syst. Biol.* 44:570–572.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.
- FELSENSTEIN, J. 1988. Phylogenies from molecular sequences: Inference and reliability. *Ann. Rev. Genet.* 22:521–565.
- GILLESPIE, E. & K. KRON. 2010. Molecular phylogenetic relationships and a revised classification of the subfamily Ericoideae (Ericaceae). *Molec. Phylogen. Evol.* 56:343–354.
- GOETSCH, L.A., A.J. ECKERT, & B. HALL. 2005. The molecular systematics of *Rhododendron* (Ericaceae): A phylogeny based upon *RPB2* gene sequences. *Syst. Bot.* 30:616–626.
- GUINDON, S. & O. GASCUEL. 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Syst. Biol.* 52:696–704.
- HARMAJA, H. 1990. New names and nomenclatural combinations in *Rhododendron* (Ericaceae). *Ann. Bot. Fenn.* 27:203–204.
- HARMAJA, H. 1991. Taxonomic notes on *Rhododendron* subsection *Ledum* (*Ledum*, Ericaceae), with a key to its species. *Ann. Bot. Fenn.* 28:171–173.
- HARMAJA, H. 1999. *Rhododendron diversipilosum*, comb. nov. (Ericaceae). *Ann. Bot. Fenn.* 35:263–264.
- HARMAJA, H. 2002. *Rhododendron subulatum*, comb. nov. (Ericaceae). *Ann. Bot. Fenn.* 39:183–184.
- HUELSENBECK, J. & F. RONQUIST. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
- HULTÉN, E. 1937. Outline of the history of arctic and boreal biota during the Quarternary period. *Lehre J Cramer, New York, U.S.A.*
- JUDD, W.S. & K.A. KRON. 2009. *Rhododendron*. In: *Flora of North America* Editorial Committee, eds. *Flora of North America north of Mexico*. Oxford University Press, New York, U.S.A., and Oxford, U.K. 8:371–372, 374–375, 451, 453, 455.
- KEARSE, M. (& 13 others). 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12):1647–1649.
- KIHLMAN, B. 2004. Hybrids between *Ledums* and lepidote *Rhododendrons*. *J. Amer. Rhododendron Soc.* 58:74–81.
- KRON, K. 1997. Phylogenetic relationships of Rhododendroideae (Ericaceae). *Amer. J. Bot.* 84:973–980.

- KRON, K. & M. CHASE. 1993. Systematics of Ericaceae, Empetraceae, Epacridaceae, and related taxa based upon *rbcl* sequence data. *Ann. Missouri Bot. Gard.* 80:735–741.
- KRON, K. & W. JUDD. 1990. Phylogenetic relationships within the Rhodoreae (Ericaceae) with specific comments on the placement of *Ledum*. *Syst. Bot.* 15:57–68.
- KRON, K.A., W.S. JUDD, P.F. STEVENS, D.M. CRAYN, A.A. ANDERBERG, P.A. GADEK, C.J. QUINN, & J.L. LUTEYN. 2002. Phylogenetic classification of Ericaceae: Molecular and morphological evidence. *Bot. Rev.* 68(3):335–423.
- KURASHIGE, Y., J. ETOH, T. HANDA, K. TAKAYANAGI, & T. YUKAWA. 2001. Sectional relationships in the genus *Rhododendron* (Ericaceae): Evidence from *matk* and *trnK* intron sequences. *Pl. Syst. Evol.* 228:1–14.
- KURASHIGE, Y., M. MINE, N. KOBAYASHI, T. HANDA, K. TAKANAYAGI, & T. YUKUWA. 1998. Investigation of sectional relationships in the genus *Rhododendron* (Ericaceae) based on *matK* sequences. *J. Jap. Bot.* 73:143–154.
- LANTAI, K. & B. KIHLMAN. 1995. The chromosome number of *Ledum palustre* ssp. *decumbens* and of some related taxa. *Hereditas (Lund.)* 122:181–184.
- LEINS, P. 1964. Entwicklungsgeschichtliche Studien an Ericales Blüten. *Bot. Jahrb. Syst.* 83:57–88.
- MATTHEWS J.V. 1979. Tertiary and Quaternary environments: Historical background for an analysis of the Canadian insect fauna. In: *Canada and its insect fauna* (ed. Danks HV), Entomological Society of Canada, Ottawa. Pp. 31–86.
- POCZAI, P. & J. HYVÖNEN. 2010. Nuclear ribosomal spacer regions in plant phylogenetics: Problems and prospects. *Molec. Biol. Reporter.* 37:1897–1912.
- POPESCU, R. & B. KÖPP. 2013. The genus *Rhododendron*: An ethnopharmacological and toxicological review. *J. Ethnopharmacol.* 147:42–62.
- RONQUIST, F. & J. HUELSENBECK. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
- RUIZHANG, F. & D.F. CHAMBERLAIN. 2005. *Ledum*. In: Wu Z.Y., P.H. Raven, & D.Y. Hong, eds. *Flora of China*. (Apiaceae through Ericaceae). Science Press and Missouri Botanical Garden Press, Beijing, China, and St. Louis, Missouri, U.S.A. 14:214.
- SHAW, J., E. LICKY, E. SCHILLING, & R. SMALL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *Amer. J. Bot.* 94:275–288.
- STAMATAKIS, A. 2014. RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. doi:10.1093/bioinformatics/btu033
- STEVENS, P.F. 1971. A classification of the Ericaceae: Subfamilies and tribes. *Bot. J. Linn. Soc.* 64:1–53.
- SWOFFORD, D.L. 2003. PAUP*: Phylogenetic analysis using parsimony (* and other methods). Version 4.0a147, January 2016 release. Sinauer Associates, Massachusetts, U.S.A.
- TUTIN, T.G., V.H. HEYWOOD, N.A. BURGESS, D.M. MOORE, D.H. VALENTINE, S.M. WALTERS, & D.A. WEBB. 1972. *Flora Europaea*, Volume 3 (Diapensiaceae to Myoporaceae). Cambridge Univ. Press, New York, New York, U.S.A.
- VOSS, D.H. 2011. A key for *Ledum* in *Rhododendron*. *J. Amer. Rhododendron Soc.* 65(2):99–101.
- YANG, Z. & B. RANNALA. 1997. Bayesian phylogenetic inference using DNA sequences: A Markov Chain Monte Carlo method. *Molec. Biol. Evol.* 14:714–724.
- YODER, A. D., J.A. IRWIN, & B.A. PAYSEUR. 2001. Failure of the ILD to determine data combinability for slow loris phylogeny. *Syst. Biol.* 50:408–424.