PHYLOGENETIC STUDIES OF THE LAURELS AND HEATHERS (ERICACEAE: ERICOIDEAE, PHYLLODOCEAE, AND *CASSIOPE* CLADES)

BY

EMILY L. GILLESPIE

A Dissertation Submitted to the Graduate Faculty of
WAKE FOREST UNIVERSITY GRADUATE SCHOOL OF ARTS AND SCIENCES

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Biology

May 2010

Winston-Salem, North Carolina

Approved By:

Qiu-Yun (Jenny) Xiang, Ph.D., chair

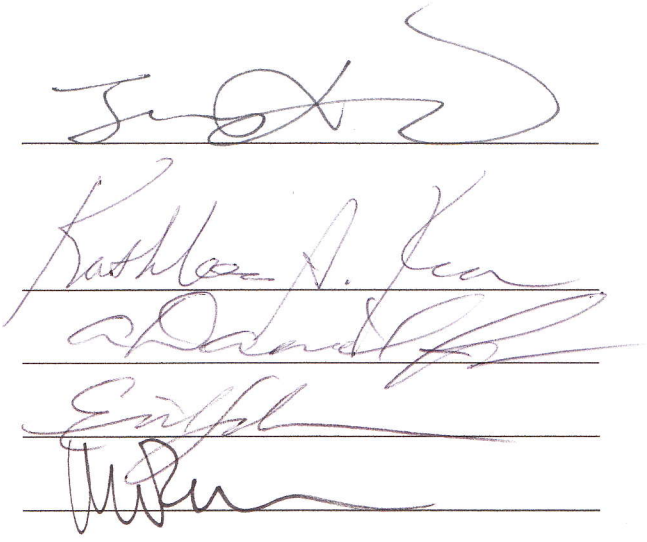
Examining Committee:

Kathleen A. Kron, Ph.D., Advisor

A. Daniel Johnson, Ph.D.

Erik C. Johnson, Ph.D.

Miles R. Silman, Ph.D.



ACKNOWLEDGEMENTS

I would like to thank my committee members, Drs. Kathleen A. Kron (WFU), A. Daniel Johnson (WFU), Erik C. Johnson (WFU), Miles R. Silman (WFU), and Qiu-Yun (Jenny) Xiang (NC State U.), for their ongoing support of my academic training. I called on them many times during the past five years, and I was always met with enthusiasm and genuine interest. Particular thanks to Kathy, for being a kind and patient advisor, for tolerating my mysterious work hours and my desire to study things like Russian, and for exchanging funny cat stories. I am most grateful for the opportunity to have learned from an expert in my field.

Funding for this study was provided by NSF grant DEB-023043 and Wake Forest University as part of a broader study of the *Rhododendron* clade. The Department of Biology provided a Vecellio Grant in order to facilitate my trips to Duke University to learn plant anatomy techniques. I received Cocke Travel Grants from the WFU Biology Department, Alumni Travel Grants from the WFU Graduate School of Arts and Sciences and Student Travel Grants from The Association of Southeastern Biologists to attend numerous meetings.

The following herbaria loaned material for this project: Harvard U. Herbaria (HUH), Swedish Museum of Natural History Herbarium (S), Museum of Nature and Human Activities, Japan (HYO), Royal Botanic Garden Edinburgh (RBGE), U. Alaska Museum of the North (ALA), Royal Botanic Gardens, Kew (RBGK), U. Washington (WTU), the Heather Society, Rhododendron Species Foundation (RSF), Florida Museum of Natural History (FLAS), California Academy of Sciences (CAS) and the Smithsonian

Institution (US). The University of British Columbia Botanical Garden allowed me to collect plants on site.

The following individuals collected or facilitated collections of plant material for these studies: Drs. Jiang Hongyuan and Du Fengguo (Forestry College of Beihua U.), Salih Terzioğlu (Karadeniz Technical U.), Robert Meinke and Aaron Liston (Oregon State U.) and Richard Abbott (Florida Museum of Natural History). Drs. Rick Ree (Field Museum), Jun Wen (Smithsonian Institution), David Boufford (Harvard U.), Peter Fritsch and Bruce Bartholomew (California Academy of Sciences) generously provided plant material from the Gaoligongshan Biodiversity Survey collections for the *Cassiope* project.

Drs. Dick White and Mel Turner (Duke U.) graciously agreed to have me visit their lab to learn various plant anatomy techniques, and Dr. Anita McCaulay (Wake Forest U.) helped me with microscopy techniques. Support staff in the WFU Department of Biology helped keep things running, including our IT guru, Mr. Jeff Muday, and our wonderful secretaries, Ms. Cindy Davis and Ms. Zella Johnson.

Numerous individuals in addition to those listed above assisted me with technical aspects of the project, made valuable suggestions and were happy to talk formal and informal 'shop' with some frequency. These include Drs. Torsten Eriksson (Swedish Museum of Natural History), Randy Small (UT-Knoxville), Zack Murrell (Appalachian State U.), Ann Powell (U. Evansville), Henry Wilbur (U. Virginia), Peter Stevens (Missouri Botanical Garden), Gary Wallace (Rancho Santa Ana Botanic Garden), Helen Murphy and Michelle DaCosta (Wake Forest U.), Jim Petranka (UNC-Asheville), Yu Yan (Sichuan U.), and Ms. A.J. Harris (Oklahoma State U.). These folks and many other

regional biologists and herbarium curators provided feedback, guidance, encouragement and friendship along the way. Many thanks to all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES AND APPENDICES

ABSTRACT

The current study explores phylogenetic studies within the laurel and heather clades within the Ericaceae. The two groups are members of two closely related clades, the Ericoideae and *Cassiope*. Phylogenetic relationships within the subfamily Ericoideae were investigated first in order to establish its main clades. Five main clades were recovered. The first four are the tribes Ericeae (African and European heathers), Empetreae (crowberries and *Diplarche*, a Rhododendron-like species), Rhodoreae (*Rhododendron* and relatives) and Bryantheae, comprised of two poorly understood genera, *Bryanthus* and *Ledothamnus*. The fifth (Phyllodoceae) includes plants such as the mountain laurels (*Kalmia*), the mountain heathers (*Phyllodoce*) and their close relatives.

The Phyllodoceae was the subject of the second part of the research. The goal here was to establish a species-level phylogeny and to examine historical biogeography of the tribe. DIVA analysis was used to infer ancestral areas of nodes, and putative vicariance events. Identification of these events was followed by relative dating of vicariance events using BEAST (Bayesian Evolutionary Analysis by Sampling Trees). This analysis revealed vicariance events in the Northern Hemisphere that likely occurred in the same general span of time. A date was hypothesized for these vicariance events, and a second BEAST analysis used these dates as calibration points to permit absolute dating of remaining nodes. These analyses revealed a likely Tertiary history of the Phyllodoceae, with distributions in North America impacted by cooling and aridification during the Eocene and Miocene, and later by glacial cycles in the Pliocene and Pleistocene.

The moss heathers (*Cassiope*) were the focus of the third project. This genus has been associated with virtually all major clades within the Ericaceae. Most recently, it was determined using molecular characters that *Cassiope* is sister to the Ericoideae. The bell heathers (*Harrimanella*) had been thought to be closely related to *Cassiope*, because its flowers and general habit are similar. Based upon anatomical characters, however, *Harrimanella* was thought to be more closely related to the blueberries (subfamily Vaccinioideae) or the Australian epacrids (subfamily Styphelioideae). In addition to estimating a species-level phylogeny for the genus *Cassiope*, the phylogenetic position of *Harrimanella* within the Ericaceae was also investigated. It was determined that *Harrimanella* is properly excluded from *Cassiope*, in agreement with anatomical data. Within *Cassiope*, morphological characters are generally homoplasious. Phylogenetic trends are evident in some anatomical characters, including vessel density and stem diameter. These results indicate that morphology in *Cassiope* is evolutionarily labile, and that anatomical characters may be tuned to local ecology.

CHAPTER 1

INTRODUCTION

The current study attempts to use a combination of modern and classical evidence to investigate evolution within a group of Northern Hemisphere plants in the family Ericaceae. The study is DNA-based, in terms of the strategy used to reconstruct evolutionary relationships. Morphology, anatomy, and geographic evidence are examined within a phylogenetic context independent of the phylogeny reconstruction. The overall approach used here was to examine a variety of molecular markers from multiple genomes, and to analyze them using multiple analytical approaches, in an attempt to generate a consensus of phylogenetic relationships and evolutionary patterns within each clade studied. Below, the primary questions or problems addressed in each taxonomic group are described. This section is followed by a discussion of the molecular markers used for phylogeny reconstruction and the rationale for those choices. Last, the analytical approaches employed in this study are described in terms of their strengths and limitations, including a rationale for their use in this study.

Research Questions

Chapter 2 describes phylogenetic studies within the subfamily Ericoideae. The Ericoideae comprise approximately 1790 species in five tribes. The circumscription of some of these tribes, such as the Ericeae (African heathers) and Empetreae (crowberries) has been taxonomically stable over time, largely because they are morphologically distinct from other tribes. Membership of the tribe Phyllodoceae, however, has changed frequently over time, as different authors placed importance on different kinds of

morphological characters (e.g. embryological, anatomical, etc…). Taxonomic studies based upon these different characters often generated contradictory circumscriptions of the tribe Phyllodoceae, largely because most genera that have been included within this tribe are morphologically heterogeneous. Therefore, in order to proceed with phylogenetic studies within the Phyllodoceae, it was necessary to clarify the circumscription of the tribe using molecular characters. Because so many ericoid genera had been included within the Phyllodoceae by past authorities (see Table 2.1), any attempt to develop a modern circumscription of the tribe Phyllodoceae required a broader study of the subfamily Ericoideae. Chapter 2 details this research and concludes with a rearrangement of the generic membership of several Ericoid tribes and a strongly supported monophyletic tribe Phyllodoceae, which is the focus of Chapter 3.

Chapter 3 describes phylogenetic studies within the Phyllodoceae, whose generic membership was established in a broader study of the subfamily Ericoideae (Chapter 2). The goal of this study was to generate a species-level phylogeny of the tribe using molecular characters. Subsequently, an examination of morphological characters was conducted within this phylogenetic framework, resulting in the discovery of several synapomorphic characters for clades within the tribe. The ultimate goal of this study was to examine the historical biogeography of the tribe. Of particular interest was estimation of the ancestral areas of each ancestral node, which was conducted using dispersal-vicariance (DIVA) analysis. Estimation of the timing of divergence events, whether resulting from vicariance or long distance dispersal, is highly desirable in biogeographic studies. This is particularly problematic in Northern Hemisphere biogeography, where similar events (e.g., a vicariance event resulting in the same extant distributions) may or

may not have resulted from the same temporal event. For example, it is known that numerous taxa have apparently been impacted by exposure of the Bering Land Bridge (BLB) connecting Western North America and Asia. However, the BLB has been exposed on multiple occasions over time, and therefore taxa exhibiting the same BLB phylogenetic 'pattern' (e.g., sister taxa in Western North America and Eastern Asia) may not have acquired this pattern at the same time—they are not truly the same event. Confidently dated fossils can provide dates with which to calibrate internal nodes, which then permits estimation of dates for other nodes. However, no such fossils are available for taxa within the Phyllodoceae. To deal with this obstacle, an iterative approach to dating was employed. First, relative dates for nodes were determined in order to assess whether multiple vicariance events could have co-occurred in time. This was found to be the case, and subsequent analyses were calibrated based on this hypothesized event in order to refine dates for other nodes. It was found that this approach, when used with caution, can provide a way around this pervasive problem in Northern Hemisphere historical biogeography.

Chapter 4 describes preliminary phylogenetic studies in the genus *Cassiope* (i.e., subfamily Cassiopoideae), which is phylogenetically sister to the subfamily Ericoideae. Until the present study, this genus has never been the subject of a phylogenetic study, although taxonomists have described various aspects of wood and leaf anatomy for some representatives. One widespread species, *Cassiope tetragona*, has received attention from ecophysiologists as a representative of the general biology of alpine plants. In general, however, the genus is poorly studied and documented. The full extent of morphological variation is unclear, as is the geographical distribution of most species. Although a

comprehensive study of this genus is not currently possible because of these issues, preliminary phylogenetic analyses were undertaken primarily in order to target future studies. A three-gene analysis provided a basic phylogenetic framework within which future studies can be conducted, and the known morphological and anatomical evidence was examined within the context of this phylogeny.

Choice of Molecular Markers

The primary factor involved in choosing DNA regions for phylogenetic studies is the appropriate rate of mutation accumulation to provide an adequate numbers of characters with which to estimate the phylogeny. Inherent within this consideration is that the DNA sequence must be alignable across taxa in order to analyze homologous bases, and that the sequences should be as independently evolving as possible. With these considerations in mind, six genes were chosen that range from relatively 'slow' to relatively 'fast' in terms of their rates of evolution. The reportedly slowest DNA regions used in the current study are *rbcL* (the gene that encodes Rubisco) and *ndhF* (the gene that encodes the F subunit of NADH dehydrogenase). Another coding region, *matK*, which encodes for a maturase involved in splicing Group II introns, has a mixture of slow and fast regions that makes it variable, but easy to align. The fastest chloroplast region used in the current study is *trnS*^{GCU}-*G*^{UUC}-*G*^{UUC} intergenic spacer, which is non-coding and therefore not under as much constraint as the other three DNA regions. Of these, *trnS-G* spacer, *rbcL* and *matK* are found in the large single copy region of the chloroplast, whereas *ndhF* is found in the small single copy region. Since none of the regions are

found in either of the inverted repeat regions, multiple copies were not anticipated (and were not detected).

Two nuclear regions were included in this study. Nuclear Ribosomal Internal Transcribed Spacer region (nrITS) is a relatively quickly evolving DNA region that is found as a tandem array of 18S—ITS1—5.8S—ITS2—26S. This entire unit encodes the ribosome, and therefore some regions are highly conserved (e.g., those involved in functional sites), whereas other regions are less constrained (e.g., those involved in forming hairpin structures). The short amplified regions of 3' 18S, 5' 26S and the short 5.8S region provide highly conserved regions that aid in reliable alignment, while ITS1 and ITS2 provide the majority of variable sites. Granule-Bound Starch Synthase I (GBSS1/*waxy*) is a low-copy nuclear gene that is relatively fast and is found as a single copy in nearly all angiosperm taxa investigated to date, including the Ericaceae. Therefore, it was not necessary to characterize the evolutionary dynamics of this gene in detail prior to use in the current study.

In the Ericoideae and Phyllodoceae projects (Chapters 2 and 3), all six DNA regions were used. The deepest nodes (i.e. clades corresponding with tribes) benefitted from the inclusion of slower regions such as *rbc*L, *ndh*F and portions of *mat*K, which are sufficiently non-variable as to be alignable across relatively distant taxa. The faster genes (trnS-G spacer, nrITS, GBSS1 and portions of matK) were included in order to resolve the more terminal nodes. For the *Cassiope* project (Chapter 4), the three fastest regions were included, because of the expectation that DNA sequences within a single genus would be relatively more similar and that fast DNA regions would provide sufficient characters to resolve relationships, whereas the slower genes would lack that variability.

Choice of Analytical Approaches

Four distinct analytical strategies were employed for all three projects. These are Bayesian Inference, RAxML Maximum Likelihood, PhyML Maximum Likelihood, and Maximum Parsimony. Each approach has strengths and weaknesses, and although the performance and reliability of all four have been examined to some extent with simulation data or 'real' data, the peculiarities of a dataset are not generally known in advance and therefore it is not possible to choose a single 'best' approach. Datasets representing complex evolutionary histories, extensive extinction, or rapid radiation are unlikely to conform precisely to the properties for which each of these four approaches are theoretically best suited, and therefore the decision was made to include all four analyses at all stages of the projects in order to produce the most robust phylogeny possible.

Maximum Parsimony (MP) searches for trees that minimize the number of character state changes required to explain the data. This approach is simple and intuitive. Because every possible tree topology is virtually impossible to search within the limitations of computer memory, a heuristic search strategy is usually employed that quickly disregards some tree topologies. This greatly shortens the time required to complete an analysis, but any heuristic search may not return the most parsimonious tree, and this is a primary disadvantage of this method. However, because the starting tree topology is randomized, conducting multiple independent analyses can essentially circumvent this problem. This strategy was employed in the current study, although individual analyses are not shown. Other potential pitfalls with MP include artifacts such as Long Branch Attraction (LBA), which occurs when two sequences are similar not

because they have a recent ancestral sequence in common, but instead by chance. Experimenting with removal of suspected ‘Long Branch’ taxa can help substantiate the presence of LBA, but strategies to limit the impact of these taxa on MP analyses are limited.

Maximum Likelihood is also a model-based method that generates a single tree with the best likelihood score. The single tree recovered is the primary disadvantage of Likelihood methods. One ML program, PhyML, does not allow data partitioning and separate models, and it is computationally expensive to perform bootstrap analysis, which is performed after the Likelihood tree is recovered. The other ML program used in this study, RAxML, does allow partitioning and separate models and determines internally the optimal number of bootstrap replicates to perform. Thus, bootstrap analysis is performed simultaneously with the search for the best-scoring tree, significantly reducing computational time. Since the two ML approaches are fundamentally different in their implementation of the Likelihood calculation, they are both included here.

Bayesian Inference uses a Markov Chain to generate large samples of trees, and then produces a posterior probability for each node. In phylogenetics, a posterior probability of 0.90 or above is considered strong Bayesian support, whereas anything lower is considered a lack of statistical support. This is in contrast to bootstrap methods, where values below 50% are considered a lack of statistical support, and the higher the boostrap value, the higher the support. The primary advantage of Bayesian analysis is the large tree sample that is generated. In addition, programs such as MrBayes and BEAST allow data to be partitioned by gene or gene region, and separate models can be applied to each partition.

The taxonomic groups studied here present a variety of methodological challenges, including deep nodes resulting in some very high levels of sequence divergence, highly divergent morphologies, an unreliable or depauperate fossil record, and largely undocumented geographical distributions. The methodologies used herein reflect the best practices available in modern plant molecular phylogenetics. DNA regions were chosen that have been well-characterized by past research, represent a variety of reported mutation rates, and represent as independent a picture of evolution as possible. The methods of phylogenetic analysis represent fundamentally different approaches with different strengths and limitations, and the limitations have been minimized to the extent possible. The absence of ideal calibration points for historical biogeography analyses is acknowledged while examining a novel approach to making progress in the area of Northern Hemisphere biogeography.

CHAPTER 2

MOLECULAR PHYLOGENETIC RELATIONSHIPS AND A REVISED CLASSIFICATION OF THE SUBFAMILY ERICOIDEAE (ERICACEAE)

E. GILLESPIE, K. KRON

The following manuscript is published in Molecular Phylogenetics and Evolution, doi:10.1016/j.ympev.2010.02.028, and is reprinted with permission. Stylistic variations are due to the requirements of the journal. E. Gillespie performed the experiments and prepared the manuscript. Dr. Kathleen Kron acted in an advisory and editorial capacity.

ABSTRACT

Subfamily Ericoideae (Ericaceae) includes 18 genera in five recognized tribes. Relationships involving the deepest nodes have been difficult to resolve, limiting the potential for further cladistic studies within the Ericoideae. The current study analyses six molecular markers using Bayesian, Maximum Likelihood and Maximum Parsimony methods to improve phylogenetic resolution within the Ericoideae. Two large clades were discovered. One clade includes the Phyllodoceae and *Bejaria*. The sister clade includes the Empetreae + *Diplarche*, Ericeae, Rhodoreae, and a clade comprised of *Bryanthus* and *Ledothamnus*. The current study improves upon the resolution of the phylogeny of the Ericoideae, particularly demonstrating support for the deepest nodes. Based on these results, we propose to retain the Ericeae, expand the Phyllodoceae to include *Bejaria*, expand the Empetreae to include *Diplarche*, retain the Rhodoreae (without *Diplarche*), dismantle the Bejarieae, and construct a new tribe, Bryantheae (*Bryanthus* and *Ledothamnus*).

INTRODUCTION

The subfamily Ericoideae (Ericaceae) is a morphologically diverse and geographically widespread group of plants comprised of approximately 1790 species in 19 genera (Mabberley, 1997; Luteyn, 1995; Stevens, 2001). These include two large genera: *Rhododendron* L. (~850 spp.) and *Erica* L. (~765 spp.), and several genera of small to moderate size: *Bejaria* Mutis ex L. (15 spp.), *Elliottia* Muhl. ex Elliott (4 spp.), *Kalmia* L. (8-10 spp.), *Ledothamnus* Meisn. (7 spp.), and *Phyllodoce* Salisb. (7-9 spp.). Most genera in the Ericoideae have 1-3 species and include *Bryanthus* Gmel. (1 sp.), *Calluna* Salisb. (1 sp.), *Ceratiola* Michx. (1 sp.), *Corema* D. Don (1 sp.), *Daboecia* D. Don (1 sp.), *Diplarche* Hook. & Thompson (2 spp.), *Empetrum* L. (3-18 spp.), *Epigaea* L. (3 spp.), *Kalmiopsis* Rehder (1 sp.), *Rhodothamnus* Rchb. (2 spp.) and *Therorhodion* Small (3 spp.).

While many species in the Ericoideae are north temperate in distribution (e.g., *Kalmia*, *Phyllodoce*, many *Rhododendron*), genera such as *Bejaria* (mostly South American), *Ledothamnus* (Guyanan tepuis), *Erica* (mostly South African), *Empetrum* (amphi-polar) and a number of southeast Asian *Rhododendron* extend the distribution of Ericoideae well into the Southern hemisphere (Mabberley, 1997; Stevens, 2001).

The Ericoideae are morphologically diverse in many aspects. Leaves may be plane or revolute; the inrolled form being the ericoid leaf seen in *Erica* and several other members of Ericoideae. Indumentum is highly variable within Ericoideae. The more complex types of indumentum include hairs with intricate three-dimensional branching and multicellular lepidote scales (e.g., some *Rhododendron*). Floral morphology in

Ericoideae ranges from polypetalous to gamopetalous. Some gamopetalous flowers appear to be secondarily polypetalous; e.g., *Loiseleuria* and *Leiophyllum* are now known to be derived from within the gamopetalous genus *Kalmia* (Kron and King, 1996). Although usually actinomorphic, many species possess flowers that are slightly bilaterally symmetrical (e.g. *Rhododendron*). Urceolate, funnelform and rotate floral shapes occur in this group and there is also variation in style articulation and orientation (Kron et al., 2002).

The taxonomic history of the Ericoideae is complex (Table 2.1), but a brief introduction helps to illustrate the degree to which the divergent morphologies of this group have been troublesome to many experts interested in reconstructing its evolutionary history. Repeated taxonomic changes associated with some genera (e.g. *Elliottia*, whose four extant species have belonged to as many as three genera and two tribes) have created a fragmented taxonomic history (Copeland, 1943; Cox, 1948; Stevens, 1971), which reflects the degree of morphological divergence seen in the Ericoideae.

Copeland (1943) recognized 20 genera and four tribes (Table 2.1) within the subfamily Rhododendroideae (Kron et al., 2002). All 20 of these genera are currently recognized in the subfamily Ericoideae (Kron 1997). Copeland based his circumscription of the subfamily on a study of anatomical and embryological characters. Copeland defined the Ericoideae by the presence of deciduous corollas, anthers without awns, and septicidally dehiscent capsular fruits. Tribes were circumscribed based primarily on the pattern of anther dehiscence. Two tribes, Bejarieae (*Bejaria*) and Cladothamneae (*Cladothamnus* and *Elliottia* [including *Tripetaleia*]), have anthers that dehisce via

resorption tissue as well as viscin threads associated with pollen tetrads. Copeland's tribe Phyllodoceae (*Bryanthus*, *Daboecia, Diplarche*, *Kalmia, Ledothamnus*, *Leiophyllum* [then called *Dendrium*], *Loiseleuria*, *Phyllodoce* and *Rhodothamnus* [including *Kalmiopsis*]) have resorption tissue, but lack viscin threads. The Rhodoreae (*Azalea*, *Azaleastrum*, *Hymenanthes*, *Ledum*, *Menziesia*, *Rhododendron*, *Therorhodion* and *Tsusiophyllum*) have collapse tissue rather than resorption tissue, but also possess viscin threads. This example demonstrates the difficulty of determining relationships using two 'prominent' characters to the exclusion of others, as in Copeland's (1943) treatment.

Cox (1948) studied wood anatomy of the Rhododendroideae, and recognized five tribes (Table 2.1). Four of these tribes, Bejarieae, Cladothamneae, Phyllodoceae, and Rhodoreae were circumscribed similarly to those groups in Copeland's study (but see *Daboecia*, below), based on suites of anatomical characters. For example, xylem vessels of the Rhodoreae and Bejarieae have pits of three types (scalariform, elliptic-elongate and round), and tertiary wall thickening. The Phyllodoceae lack round pit types and tertiary thickening, but possess scalariform and elliptic-elongate pits. Scalariform pits and tertiary wall thickenings are present in Cox's (1948) Cladothamneae, but some members of this group possess round and elliptic-elongate pit types. In contrast to Copeland's (1943) treatment, Cox (1948) moved *Daboecia* from Phyllodoceae into its own tribe (Daboecieae). *Daboecia* has thickened, uneven vessel walls which lack scalariform perforations entirely, whereas the remainder of Cox's Phyllodoceae have thin, even vessel walls and most taxa have a large proportion (> 79%) of scalariform perforations. Some of Cox's 'advanced' character states included uneven and thick vessel walls, a shift from scalariform to porous end wall perforations, and a shift from scalariform pits to elliptic-

elongate pits followed by a shift to the further derived condition of round pits. Cox's (1948) schematic indicated a common ancestor giving rise to all five tribes, but essentially indicated a polytomy among the extant tribes.

Stevens (1969, 1971) examined anatomical and morphological characters within the Ericaceae as a whole. He recognized seven tribes (Table 2.1) within Rhododendroideae. Bejarieae, Cladothamneae, and Rhodoreae were retained in agreement with Copeland (1943) and Cox (1948). *Diplarche* was moved from Phyllodoceae into its own tribe (Diplarcheae) based on stomata presence on the abaxial calyx surface only; Phyllodoceae typically have adaxial calyx stomata only. Otherwise, Phyllodoceae membership was left unchanged relative to Copeland (1943) and Cox (1948). Stevens (1971) retained Daboecieae (*Daboecia*) as a distinct tribe, in agreement with Cox (1948), because of several embryological characters uncommon in the Ericaceae, such as a hypostase in the ovule, a very thick (7-9 layers of cells) integument and papillate seeds. He recognized, however, some similarities to the Ericoideae (see below) in leaf anatomy, petiole tissue, flower merosity, corolla shape and absence of viscin threads. In addition, Stevens recognized a new rhododendroid tribe, Epigaeae, which included *Epigaea.* This taxon was previously considered to be a relative of *Andromeda* L. in the Vaccinioideae, and therefore not included in the Rhododendroideae studies by Copeland (1943) or Cox (1948). Epigaeae was included within Rhododendroideae by Stevens (1971) due to similarities in inflorescence position, stomatal distribution, corolla shape, and stamen morphology, all of which were unlike *Andromeda*. Stevens (1971) included the tribes Calluneae (*Calluna*) and Ericeae (*Erica*) within the subfamily Ericoideae. He thought Ericoideae and Rhododendroideae might

have a close relationship, by way of *Daboecia*. *Daboecia* shares character states in common with the Ericoideae (see above), but also with the rhododendroid taxa *Bryanthus* (terminal racemose inflorescence and perforation plate structure) and *Ledothamnus* (mucilaginous epidermis).

Kron and Chase (1993) used *rbcL* data and Maximum Parsimony analyses to investigate relationships within Ericaceae. This study, followed by Kron (1997), found that the traditional Ericoideae (i.e. *Erica* and *Calluna*) as well as Empetraceae (represented by *Ceratiola*) were derived from within the Rhododendroideae, rendering Rhododendroideae paraphyletic. Naming this resulting clade required that the entire Rhododendroideae and Empetraceae be subsumed within Ericoideae (The type is *Erica* L.); the former Ericoideae was subsequently named Ericeae and the family Empetraceae was named Empetreae.

Kron et al. (2002) used *matK*, *rbcL* and 91 morphological characters to analyze relationships within Ericaceae. The total combined analysis revealed a strongly supported monophyletic Ericoideae, composed of five tribes: Bejarieae (*Bejaria*, *Bryanthus*, and *Ledothamnus*), Ericeae (*Calluna*, *Daboecia*, and *Erica*), Empetreae (*Ceratiola*, *Corema*, and *Empetrum*), Rhodoreae (*Menziesia*, *Rhododendron*, and *Therorhodion*), and Phyllodoceae (*Elliottia*, *Epigaea*, *Kalmia*, *Kalmiopsis*, *Phyllodoce*, and *Rhodothamnus*). Of these tribes, Ericeae was moderately well supported (79% bootstrap), while Empetreae, and Rhodoreae were each very strongly supported (92% and 100% respectively). However, little to no support existed for the Bejarieae and Phyllodoceae clades in these analyses and therefore the circumscription of each could not be

confidently ascertained. Additionally, the relationships among all of the tribes (the deepest nodes within Ericoideae) were not well-supported.

The objectives of the current study are to increase the resolution at the deepest nodes within the Ericoideae (i.e., tribe-level relationships) and to test the monophyly of the tribes recognized by Kron et al. (2002) using six molecular markers. A better resolved phylogeny of the Ericoideae will permit more detailed study within each of the tribes in terms of outgroup selection and taxon sampling, as well as a refined understanding of large-scale morphological patterns of evolution within the group.

MATERIALS AND METHODS

Total DNA isolation was carried out on fresh, silica-dried or herbarium material by a modified CTAB method (Doyle and Doyle, 1987), with or without cesium chloride purification, the Invisorb Spin Plant Mini Kit (Invitek GmbH, Berlin, Germany), or the Qiagen Plant Mini Kit (Qiagen, Valencia CA, USA) with modifications following Drábková et al. (2002).

Forty taxa representing all genera of Ericoideae were selected (Appendix 2.1). Taxon sampling was minimal in large clades such as Rhodoreae and Ericeae, where monophyly was not in question (Kron, 1997; Kron et al., 2002). Previous studies (e.g., Kron et al., 2002) have shown that a small number of exemplars for a large, strongly supported group is sufficient representation for that group when analyzing other related taxa. This approach was employed in this study in order to permit examination of the more problematic areas within Ericoideae such as in the small, morphologically divergent genera that have been classified within the Phyllodoceae in the past (Appendix 2.1, Table

2.1). For ingroup taxa, six species of Bejarieae were sampled (3 of 15 recognized species (3/15) of *Bejaria*), 1/1 *Bryanthus* and 2/6 *Ledothamnus*), five species of Empetreae (1/2 *Corema*, 1/1 *Ceratiola*, and 3/18 *Empetrum*), six species of Ericeae (1/1 *Calluna*, 1/1 *Daboecia*, and 4/765 *Erica*), 15 species of Phyllodoceae (4/4 *Elliottia*, 5/10 *Kalmia* [incl. *Loiseleuria* and *Leiophyllum*], 3/5 *Phyllodoce*, 1/1 *Kalmiopsis*, 1/3 *Epigaea*, and 1/2 *Rhodothamnus*), and eight species of Rhodoreae (1/2 *Diplarche*, 1/3 *Therorhodion*, 1/7 *Menziesia* and 5/865 *Rhododendron*—one from each of the five largest subgenera) were selected for analysis. Four outgroup taxa were chosen from additional subfamilies within the Ericaceae (Appendix 2.1). Based on previous research (e.g. Kron, 1997; Kron and King, 1996; Kron et al., 2002), the subfamily Cassiopoideae is strongly supported as the sister group to the subfamily Ericoideae, and this clade is represented by *Cassiope mertensiana* in the current study. More distantly related clades within the Ericaceae are represented by *Enkianthus campanulatus* (Enkianthoideae), *Arctostaphylos uva-ursi* (Arbutoideae), and *Vaccinium tenellum* (Vaccinioideae).

Six DNA regions (four chloroplast and two nuclear) were included in this study (Appendix 2.2). These include chloroplast regions *rbcL, ndhF, matK,* and *trnS*^{GCU}-*trnG*^{UCC}-*trnG*^{UCC} intergenic spacer (*trnS-G-G*), and nuclear regions *waxy*/GBSS-1 (*waxy* exons 9-11) and nuclear ribosomal Internal Transcribed Spacer (nrITS). Regions were amplified using standard PCR primers and protocols (Appendix 2.2 – Baldwin, 1992; Baldwin et al., 1995; Evans et al., 2000; Johnson and Soltis, 1994; Kron and King, 1996; Olmstead et al., 1992; Olmstead and Sweere, 1994; Powell and Kron 2001; Shaw et al., 2005; Steele and Vilgalys, 1994; White et al., 1990). Primer sequences are reported in Appendix 2.2. Single gene, combined nuclear, combined chloroplast and total data

matrices are deposited in TreeBase (www.treeBASE.org), under accession number SN4740. All newly generated sequences are deposited in Genbank (www.ncbi.nlm.nih.gov) under numbers GU176622-GU176746 and are listed in Appendix 2.1.

Amplified fragments were cleaned using Qiagen™ QIAquick Gel Isolation Kit (Qiagen, Valencia CA, USA). DNA was sequenced on an ABI 377 automated sequencer at the DNA Sequencing and Gene Analysis Laboratory at the Wake Forest University School of Medicine (Winston-Salem, NC) or at Nevada Genomics Center (Reno, NV). Sequences were minimally edited and manually aligned in Geneious 4.7.4 (Drummond et al., 2006).

Nuclear regions were cloned using the TOPO TA kit (Invitrogen, Carlsbad CA, USA) for representative taxa (five clones per taxon) within the Ericaceae in order to detect multiple copies if they existed. No evidence for multiple copies of either nrITS or *waxy* was found within the Ericoideae, and therefore additional cloning was not carried out.

The Incongruence Length Difference test (Farris et al., 1995), implemented as the partition homogeneity test in PAUP*4.0b10 (Swofford, 2002), was conducted to determine if the nuclear and chloroplast partitions were sufficiently congruent to be combined into a total molecular evidence analysis. A heuristic search with 100 replicates was performed using TBR swapping and simple, stepwise addition of taxa.

Maximum Parsimony (MP) analyses were carried out using PAUP*4.0b10, (Swofford, 2002) with the following options: Parsimony-informative characters were unordered and equally weighted, gaps were treated as missing data, searches were

heuristic with TBR branch swapping and 1000 random stepwise addition replicates. Relative clade support was assessed using bootstrap (Felsenstein, 1985; Felsenstein, 1988) with the full bootstrap option in PAUP* (10,000 replicates).

Maximum Likelihood analyses were carried out using the PhyML online server (Guindon and Gascuel, 2003; Guindon et al. 2005) on the ATCG Montpellier bioinformatics platform, as well as using the PhyML online server (Guindon and Gascuel, 2003; Guindon et al., 2005) on the ATCG Montpellier bioinformatics platform.. For PhyML, total molecular evidence, as well as chloroplast and nuclear partitions were each run under a GTR model, as determined by the AICc criterion (Aikake, 1974) in MrAIC.pl 1.4.3 (Nylander, 2004). AICc was chosen as the criterion for model selection because it is not hierarchical in nature and also corrects for small sample sizes (approximately 40 and below). Bootstrap analysis (100 replicates) was conducted to determine node support. For RAxML, total molecular evidence, chloroplast, and nuclear datasets were run with genes as separate partitions, under optimal models indicated by MrAIC (Nylander, 2004). Both Maximum Likelihood strategies were used because they use fundamentally different bootstrapping strategies and are therefore essentially different approaches.

For Bayesian analyses, the data were partitioned by DNA region and evolutionary models were chosen using the AICc criterion in MrAIC.pl 1.4.3 (Nylander, 2004) (Appendix 2.2). Bayesian MCMC analyses (Yang and Rannala, 1997) as implemented in MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) consisted of 20 million generations with a burn-in of 25%. Clade support is reported as posterior probabilities.

Bayesian, Maximum Likelihood, and Maximum Parsimony analyses of individual DNA regions, as well as data partitioned by genome, were conducted to assess topology and clade support using individual data partitions. Single gene analyses revealed many short branches and a general lack of clade support at the basal nodes of the Ericoideae (i.e., the nodes of particular interest in the current study), and therefore single-gene analyses are not shown here.

RESULTS

Results for Bayesian, Maximum Likelihood and Maximum Parsimony analyses of the chloroplast partition, nuclear partition, and total molecular evidence are presented in Figures 2.1, 2.2, and 2.3, respectively. Node labels and descriptions of support within the text include Bayesian posterior probabilities (PP), ML bootstrap and MP bootstrap values, in the following format: (PP / ML / MP).

Aligned length, number of potentially parsimony informative characters (and percent), and percent missing data (not including alignment gaps) for each DNA region are given in Table 2.1. After preliminary parsimony analyses revealed a very long branch leading to *Ledothamnus guyanensis* Meissn., *Ledothamnus sessiliflorus* N. E. Br. was included in an attempt to shorten the branch leading to *Ledothamnus.* Therefore all sequence data for this taxon was newly generated for this study. Sequence data for one or two genes were commonly available for other taxa, and these were compiled with thirteen newly generated nrITS sequences (no taxa missing), 16 *waxy* (ten missing taxa), five *matK* (one missing taxa), six *rbcL* (three missing taxa), 31 *ndhF* (five missing taxa), and

39 *trnS-G-G* spacer (seven missing taxa). All *trnS-G-G* spacer sequences were newly generated for this study.

Chloroplast Data (Fig. 1)

The Ericoideae is well-supported (1.00 / - / 100 / 100) by analyses of chloroplast data. Four major clades are recovered. These are Empetreae + Rhodoreae (1.00 / 100 / 100 / 97), Phyllodoceae + *Bejaria* (1.00 / 100 / 100 / -), *Ledothamnus* + *Bryanthus* (0.99 / 79 / 60 / 69), and Ericeae (1.00 / 83 / 98 / 93). These correspond to the tribes recognized in Kron et al. (2002) (Table 2.1) except for the position of *Bejaria*. In Kron et al. (2002), Bejarieae included *Ledothamnus* and *Bryanthus* in addition to *Bejaria*. However in the chloroplast analysis of the current study, *Bejaria* is strongly supported as sister to the Phyllodoceae clade rather than in a clade with *Ledothamnus* and *Bryanthus*. *Diplarche* is placed as sister to the Empetreae with strong support (1.00 / 94 / 94 / 100). The position of *Diplarche* as sister to Empetreae was not found in the Kron et al. (2002) study that included the chloroplast regions *rbcL* and *matK*, but in this study is placed as sister to Empetreae. Within the Phyllodoceae, *Elliottia* is well-supported as sister to the remaining members by Bayesian PP and ML (1.00 / 87 / 81 /-), and *Kalmia* is moderately supported as sister (0.92 / 78 / 60 / -) to a consistently well-supported (1.00 / 100 / 100 / 100) clade comprised of *Epigaea*, *Kalmiopsis*, *Phyllodoce* and *Rhodothamnus*—hereafter referred to as the '*Phyllodoce* clade' (refer to Fig. 1). Within this clade, *Epigaea* and *Rhodothamnus* are strongly supported as sister taxa (1.00 / 100 / 100 / 99), as are *Phyllodoce* and *Kalmiopsis* (1.00 / 100 / 100 / 100).

Nuclear Data (Fig. 2)

Although some shallow nodes are strongly supported based on the nrITS and *waxy* data, overall there is less resolution and relatively low support, compared to chloroplast data. Nuclear data indicate the Ericoideae is a clade, but with weak support (0.77 / - / - / 77). *Bejaria* is strongly supported (1.00 / 100 / 100 / 100) as a clade, as are the Empetreae including *Diplarche*, which is sister to *Corema* (0.97 / - / 63 / -), although MP bootstrap support is lacking for the Empetreae (1.00 / 100 / 100 / -). Several genera are well supported by nuclear data, including *Ledothamnus* (1.00 / 99 / 99 / 100), *Erica* (1.00 / 100 / 100 / 100), *Empetrum* (1.00 / 99 / 95 / 98), and *Kalmia* (1.00 / 100 / 90 / 100). Some genera are not supported as clades, including *Phyllodoce* (which is rendered paraphyletic by placement of *Kalmiopsis* within *Phyllodoce*) and *Elliottia* (which is polyphyletic, or perhaps paraphyletic if the lack of support for the *Kalmia* + *Phyllodoce* clade is taken into consideration).

The ILD test revealed that the nuclear and chloroplast datasets are in conflict (P=0.01). Most instances of conflict at deeper nodes involve poor support in the nuclear data and these are considered 'soft' conflict, following the convention of other authors (e.g. Bull et al., 1993; Mason-Gamer and Kellogg, 1996; Oh and Potter, 2005). These instances of soft conflict include the relationship of Empetreae sister to Ericeae in nuclear data (0.66 / - / - / -) and the resolution of *Bryanthus* with the Rhodoreae/Empetreae/Ericeae clade (0.84 / - / - / -). More strongly supported conflict occurs with regard to the genus *Elliottia*, which is not monophyletic in the nuclear data. This is the most serious instance of conflict between the nuclear and chloroplast data, yet

support for the positions of the two *Elliottia* clades is not strong in most analyses of nuclear data.

Questions about the applicability of the ILD test have been raised in the literature (e. g. Yoder et al., 2001; Barker and Lutzoni, 2002; Cunningham, 1997). These concerns stem in part from the likelihood of making a Type I error (i. e. failing to combine the datasets when they are not incongruent). Cunningham (1997) suggested that a P-value of between 0.01 and 0.001 would be a more reasonable critical value for the ILD test for use in phylogenetic studies. Barker and Lutzoni (2002) went so far to suggest that even these values were not appropriate to aid in determining whether to combine datasets. Therefore, because the P-value determined with the ILD test is near the critical value considered appropriate by Cunningham (1997), because the conflict in the nuclear and chloroplast estimates are limited to a small number of taxa or nodes with poor support, and because nodes not in conflict suffer lower support when data partitions are analyzed separately, the data were combined into a total evidence analysis.

Total Molecular Evidence (Fig. 3)

The Ericoideae are well-supported in all analyses (1.00 / 84 / 86 / 100) of six DNA regions. Within the Ericoideae, two large clades were resolved. One clade includes *Bejaria* and the Phyllodoceae, and is well supported by Bayesian PP and ML, but not well-supported by MP (1.00 / 100 / 100 / 50). Within the Phyllodoceae, *Elliottia* is strongly supported as a clade (1.00 / 100 / 100 / 98) and is sister to the remaining taxa (1.00 / 89 / 90 / 71). *Kalmia* is sister to the '*Phyllodoce* clade' (0.95 / 76 / 69 / 56).

Within this clade, *Epigaea* is sister to *Rhodothamnus* (1.00 / 97 / 98 / 96) and *Phyllodoce* is sister to *Kalmiopsis* (1.00 / 100 / 100 / 100).

The other large clade within Ericoideae has good support by Bayesian PP but no support by bootstrap in either the ML or MP analyses (0.92 / - / - / -). The *Ledothamnus* + *Bryanthus* clade has strong to moderate support (0.98 / 75 / 63 / 80) and the Ericeae is strongly supported (1.00 / 86 / 98 / 98). The clade comprised of the Empetreae and Rhodoreae is also well supported (1.00 / 100 / 97 / 93). The relationship between Ericeae and Empetreae + Rhodoreae is weakly supported (0.70 / - / 53 / -). *Diplarche* is placed as sister to the Empetreae, and this is the same position as supported in the chloroplast analysis. This genus has most recently been considered a member of the Rhodoreae (Kron et al., 2002), but here it is strongly supported (1.00 / 100 / 100 / 100) as sister to the Empetreae. A phylogram resulting from Bayesian analysis of total combined molecular data (Fig. 2.4) indictes very short branches at the basal-most nodes within the Ericoideae.

DISCUSSION

The current study has resulted in considerable improvement in tribe-level relationships and tribe composition within the Ericoideae, relative to the combined parsimony analysis of *matK* and *rbcL* of Kron et al. (2002). In both Kron et al. (2002) and combined analyses of the current study, the Ericoideae are strongly supported as a clade, and the sister relationship of the Ericoideae to the Cassiopoideae (*Cassiope*) is also supported in this study. Therefore the circumscription of the Ericoideae and position of the immediate outgroup *Cassiope* is confirmed.

The results here support recognition of five tribes within Ericoideae (Table 2.1). 1) Ericeae (*Calluna*, *Daboecia* and *Erica*) is retained in agreement with Kron et al. (2002). 2) Strong support exists for the inclusion of *Diplarche* within Empetreae, along with *Empetrum*, *Corema* and *Ceratiola*. We assert that recognizing the relationships among *Diplarche* and the remaining Empetreae will help encourage the inclusion of *Diplarche* in future investigations of character evolution within Empetreae, despite noteworthy morphological differences that led Stevens (1969, 1971) to recognize *Diplarche* in its own tribe, Diplarcheae. Retaining *Diplarche* in Diplarcheae would certainly facilitate simpler morphological description of both Diplarcheae and Empetreae, given the morphological distinctiveness of each. However, the nuclear phylogeny opens up the possibility that *Diplarche* is not sister to the remaining Empetreae, but rather, nested within. Inclusion of *Diplarche* within Empetreae accommodates both a scenario where *Diplarche* branches at the basal-most node, as well as a scenario where *Diplarche* is more deeply nested within the clade, should future phylogenetic studies find stronger support for it. 3) The Rhodoreae should be retained, except for the transfer of *Diplarche* to Empetreae. Therefore the Rhodoreae includes *Rhododendron* s.l., *Menziesia*, and *Therorhodion*. 4) The Phyllodoceae is expanded to include *Bejaria*. 5) *Bryanthus* and *Ledothamnus* are transferred from Bejarieae into a new tribe, Bryantheae (type = *Bryanthus gmelinii* D. Don). Therefore the Bejarieae, recognized by Copeland (1943), Cox (1948), Stevens (1971) and Kron et al., (2002) is dismantled.

The Phyllodoceae, comprised of *Elliottia*, *Kalmia* and the '*Phyllodoce* clade', is strongly supported here by Bayesian and ML analyses of chloroplast and total molecular data, and moderately supported by MP, in contrast to the Kron et al. (2002) study that

showed MP bootstrap support below 50%. Within the Phyllodoceae, relationships at the level of genera are unchanged relative to Kron et al. (2002). *Elliottia*, then *Kalmia*, are successive sister taxa to the '*Phyllodoce* clade'. However, in the current study, support for these relationships is much higher.

Analyses of combined chloroplast data strongly support most tribes and is not in conflict with the total evidence analyses. Nuclear data, however, generated a very different topology with respect to particular taxa, when compared with either the chloroplast or the total molecular trees. Missing data are always of concern in phylogenetic studies, whether directly stated or not. Although many taxa are complete for all six DNA regions, some of the taxa whose placement differs among the nuclear and chloroplast or total evidence trees are missing some data in the current study (i.e. *Elliottia* and Ericeae). However, a growing body of evidence suggests that missing data are not inherently problematic. Recent empirical and simulation studies have indicated that missing data up to 95% have no negative impact on the accuracy of phylogenetic studies, provided that the dataset is large (>2000 bp), and that the missing data are not distributed such that any taxa have completely non-overlapping data (Driskell et al., 2004; Phillipe et al., 2004; Wiens, 2003; Wiens, 2005; Wiens and Moen, 2008). All of those conditions are met in the current study. The dataset is nearly 8000 bases total, the missing data are distributed in a way that does not result in 'non-overlapping' data partitions in any taxa, and incomplete data do not approach 95% in any partition of the data for any taxa at any taxonomic level.

The most incomplete data occur in *Erica*, where 75% (3 of 4 taxa) have missing data for *waxy*, *ndhF* and *trnS-G-G* spacer. The placement of Ericeae in both the

chloroplast and total evidence analyses is sister to Rhodoreae + Empetreae, with strong support. The nuclear data place it sister to the Empetreae, but with weak support. The two early-branching taxa, *Calluna vulgaris* and *Daboecia cantabrica*, are each complete for all six DNA regions, and within the Ericeae clade, the genus *Erica* is strongly supported as monophyletic despite missing data, in agreement with other studies (e.g. Kron, 1997; Kron et al., 2002). Therefore it is unlikely that even this amount of missing data in *Erica* had a significant negative impact on the support or placement of the Ericeae in the nuclear analyses.

The other instance where nuclear data produced a topology different from the chloroplast or total molecular phylogenies is in the monophyly and placement of *Elliottia*. Whereas in the chloroplast and total molecular analyses *Elliottia* is monophyletic and placed as sister to *Kalmia* + the '*Phyllodoce* clade', the nuclear data place two *Elliottia* as a well-supported clade unresolved in its placement within the Ericoideae, and two *Elliottia* as sister to *Kalmia*. All evidence except analyses of nuclear data support the monophyly of *Elliottia*. The chloroplast (1.00 / 100 / 94) and total evidence data (1.00 / 100 / 98) in the current study strongly support monophyly. Moretz (2002) identified seven characters that unite *Elliottia*, including deciduous leaves, long slits in the anthers, terminal inflorescences, pollen shed prior to the opening of flowers, conduplicate leaves in bud, polypetalous corollas and flattened anther filaments. Most of these characters are homoplasious within the Ericoideae, but flattened anther filaments are unique, and therefore likely represent a truly synapomorphic character for *Elliottia*. Hebda and Chinnappa (1980) carried out meiotic chromosome counts for all four taxa and found that all four *Elliottia* are n=11. Most Ericaceae are n=12 or n=13 (aside from

Elliottia, n=11 is found in some Gaultherieae). Therefore, despite some noise in the molecular data, most molecular data, morphological data and cytological data all suggest that *Elliottia* is monophyletic.

Relationships within the *'Phyllodoce* clade' appear stable at the generic level with the current sampling. Both Kron et al. (2002) and the current study recovered a topology where *Phyllodoce* + *Kalmiopsis* is sister to *Epigaea* + *Rhodothamnus*, with all nodes strongly supported. Members of this clade are morphologically rather dissimilar, especially in leaf characters. Given the generic relationships as presently known, *Epigaea*, which has essentially flat leaves with slightly inrolled margins, is sister to *Rhodothamnus*, whose leaves have small projections along the margins. *Phyllodoce*, whose leaves are ericoid, is sister to *Kalmiopsis*, which has flat leaves with orange colored glandular hairs. Given that the molecular phylogenetic relationships within this clade appear to be very stable over several studies (Kron 1997; Kron et al., 2002; current study), future studies of leaf evolution and development within this clade would be valuable.

Members of the Phyllodoce clade also exhibit disjunct geographic distributions. Two genera (*Epigaea* and *Phyllodoce*) exhibit intercontinental disjunctions. *Epigaea* has three members, one each in North America, Asia and the Caucasus, while its sister taxon *Rhodothamnus* is limited to mountainous regions of southern Europe and the Caucasus. *Phyllodoce* has 5-7 species and is roughly circumboreal in distribution, while its sister *Kalmiopsis* is strictly limited to southwest Oregon, USA. Therefore a study of this clade including all species would be valuable, both in terms of leaf evolution and historical biogeography.

The composition of the Phyllodoceae appears to be *Elliottia, Kalmia, Phyllodoce, Kalmiopsis, Epigaea and Rhodothamnus*. *Bryanthus* and *Ledothamnus*, which have been previously included within the Phyllodoceae (Copeland 1943, Cox 1948, Stevens 1971), are not resolved as part of this clade. The clade of *Kalmia* + the '*Phyllodoce* clade' is the least well-supported node within the Phyllodoceae (0.95/69/56), and a more concentrated study of the Phyllodoceae should help clarify this node. This clade is supported by the apparently secondary loss of viscin threads associated with pollen and the presence of adaxial calyx stomata. The majority of the Rhododendroideae, including *Bejaria* and *Elliottia* possess viscin threads and lack adaxial calyx stomata.

In the current analysis, *Bejaria* is strongly supported by Bayesian and ML as sister to the Phyllodoceae sensu Kron et al. (2002), although MP support is weak. The sister relationship of the Phyllodoceae to *Bejaria* was not recovered in Kron et al. (2002). Instead *Bejaria* was sister to *Bryanthus* + *Ledothamnus* in the strict consensus tree, but without bootstrap support. The strong support for the sister relationship of *Bejaria* and the Phyllodoceae in this study is further supported by the presence of homogenous pith and articulated pedicels in both Phyllodoceae and *Bejaria* and this character could represent a morphological synapomorphy for this clade.

Kron et al. (2002) recovered both the Empetreae and the Rhodoreae as clades with very strong support in MP analyses. Both clades were recovered here, again with strong support. In addition, Empetreae was found to be sister to Rhodoreae, and the Empetreae + Rhodoreae clade is also strongly supported in all analyses of chloroplast and total molecular data. One major difference between this study and Kron et al. (2002) concerns the placement of *Diplarche*. The combined *matK* + *rbcL* analysis of Kron et al. (2002)

found *Diplarche* nested within the Rhodoreae, one node up from *Therorhodion*, with strong MP support. In the current study, *Diplarche* is placed in the Empetreae clade in all three analyses. However, the position of *Diplarche* within Empetreae differs in the nuclear data analysis from its placement as sister to remaining Empetreae in the chloroplast analysis. Combined data that use two nuclear regions and four chloroplast regions, however, place *Diplarche* as sister to the Empetreae with very strong support in Bayesian, ML and MP analyses. Copeland (1943) and Cox (1948) both placed *Diplarche* in Phyllodoceae, but Stevens (1971) considered *Diplarche* an isolated genus within Rhododendroideae because of its possession of unique characteristics such as epipetalous stamens and a septifragal capsule with valves that split into two layers. He noted other morphological similarities to some members of Phyllodoceae, and placed *Diplarche* in a separate tribe (Table 2.1).

The recognition of *Diplarche* as sister to the Empetreae provides an opportunity to reconsider the evolution of several morphological characters within the Ericoideae. For five characters mapped by Kron et al. (2002), the reconstruction of the character state of the common ancestor to the Empetreae and Rhodoreae clades is impacted in the same manner by the new placement of *Diplarche*. For example, the previous placement of *Diplarche* within Rhodoreae (both taxa have flat leaves) suggests that the common ancestor of all Rhodoreae possessed flat leaves, whereas the common ancestor of Empetreae possessed ericoid leaves. In this case, the character state for the ancestor of Empetreae + Rhodoreae would be equivocal. However, the current study's placement of *Diplarche* (flat leaves) with the Empetreae (ericoid leaves) opens up the possibility that the character state of the common ancestor of Empetreae and Rhodoreae had flat leaves, a

hypothesis requiring further study. Precisely the same pattern is found in pedicel bracts (Rhodoreae = 2, *Diplarche* = 2, Empetreae = 2+), corolla fusion (Rhodoreae = sympetalous, *Diplarche* = sympetalous, Empetreae = choripetalous), corolla persistence (Rhodoreae = deciduous, *Diplarche* = deciduous, Empetreae = persistent tepals), and fruit dehiscence (Rhodoreae = dehiscent, *Diplarche* = dehiscent, Empetreae = indehiscent).

The absence of midrib fibers (Kron et al., unpublished data) may represent a morphological synapomorphy for Empetreae + *Diplarche*. Midrib fibers are present in Rhodoreae. Previous studies (e.g. Kron et al. 2002) resolved *Diplarche* as within the Rhodoreae, which would make the absence of midrib fibers within the Rhodoreae a non-uniform character. The only other Ericoideae taxon lacking midrib fibers is *Elliottia bracteata* (Phyllodoceae), and so rather than the independent loss of bud scales in the common ancestor of Empetreae and separately in *Diplarche*, the loss of midrib fibers appears to have occurred in the common ancestor of all Empetreae and *Diplarche*, representing a potential morphological synapomorphy for a more broadly circumscribed Empetreae (including *Diplarche*). The Rhodoreae as circumscribed here are recognized by slightly zygomorphic corollas (but actinomorphic in *Menziesia*) with spots or blotches, ovoid capsules, and flowers on the shoots of the previous season.

The clade comprised of *Bryanthus* and *Ledothamnus*, which is itself well-supported in Bayesian and MP analyses in the current study, is sister to [Ericeae + [Empetreae + Rhodoreae]], with strong PP support, but weak bootstrap support. These relationships did not appear in the parsimony analyses of Kron et al. (2002), since bootstrap support based on two genes was often weak at the deeper nodes. Characters uniting *Bryanthus* + *Ledothamnus* may include ericoid leaves, stomata on the adaxial

surface of the calyx, the style articulated with the ovary and the absence of an endothecium (although these characters states are unclear for other Ericoideae). The similarities Stevens (1971) noted among Ericeae (then called Ericoideae) including *Daboecia* and *Bryanthus* + *Ledothamnus* are particularly interesting given the results presented here. A more detailed study of the *Bryanthus* + *Ledothamnus* [Ericeae [Empetreae + Rhodoreae] clade would be valuable, using both additional molecular data and a novel examination of anatomical and morphological characters. From a biogeographical perspective, this is also a very interesting relationship because the monotypic *Bryanthus* (*B. gmelini*) is endemic to Japan and Kamchatka, and *Ledothamnus* (7 spp.) is restricted to the Guyanan shield of South America. This disjunct distribution clearly warrants further investigation.

The current study represents an improvement of our understanding of evolution within the Ericoideae. It confirms the close relationship of Empetreae to Rhodoreae, and also strongly suggests that *Diplarche* should be included within Empetreae. Combined molecular analyses in the current study show the Ericeae as sister to the Empetreae/*Diplarche* + Rhodoreae, but with weak support in all three analyses. The circumscription of the Phyllodoceae is improved by the current study; it is now clear that *Bejaria* should be included within the Phyllodoceae, but that *Bryanthus* and *Ledothamnus* should be recognized as a separate tribe. Inclusion of *Diplarche* as part of the Empetreae, recognition of a new tribe, Bryantheae (*Bryanthus* and *Ledothamnus*), and inclusion of *Bejaria* within the Phyllodoceae should encourage proper sampling of these clades for future studies and more informed choice of outgroups for phylogenetic studies.

ACKNOWLEDGEMENTS

The authors thank the following institutions and individuals for their assistance in obtaining plant material: AA, GH, NCU, NYBG, RBGE, RBGK, WTU, Royal Botanic Garden, Edinburgh, Royal Botanic Gardens, Kew, Rhododendron Species Foundation, the Heather Society, and Arne Anderberg. Technical advice was provided by Torsten Eriksson, and anonymous reviewers provided comments on the manuscript. Funding for this study was provided by NSF grant DEB-023043 and Wake Forest University.

LITERATURE CITED

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Automatic Control 19, 716–723.

Baldwin, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Mol. Phylogen. Evol. 1, 3-16.

Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S., Donoghue, M.J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Mo. Bot. Gard. 82, 247–277.

Barker, F.K., Lutzoni, F. 2002. The utility of the incongruence length difference test. Syst. Biol. 51, 625–637.

Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L., Waddell, P.J. 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42, 384–397.

Copeland, H.F. 1943. A study, anatomical and taxonomic, of the genera of Rhododendroideae. Am. Midl. Nat. 30(3), 533-625.

Cox, H.T. 1948. Studies in the comparative anatomy of the Ericales I-Subfamily Rhododendroideae. Am. Midl. Nat. 39(1), 220-245.

Cunningham, C.W. 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. 14, 733–740.

Doyle, J., Doyle, J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11-15.

Drábková, L., Kirschner, J., Vlček, Č. 2002. Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens in Juncaceae. Plant Mol. Biol. Rep. 20, 161-175.

Driskell A.C., Ané, C., Burleigh, J.G., McMahon, M.M., O'Meara, B.C., Sanderson, M.J. 2004. Prospects for building the Tree of Life from large sequence databases. Science 306, 1172–1174.

Drummond, A.J., Kearse, M., Heled, J., Moir, R., Thierer, T., Ashton, B., Wilson, A., Stones-Havas, S. (2006) Geneious v2.5.

Evans, R.C., Alice, L.A., Campbell, C.S., Kellogg, E.A., Dickinson, T.A. 2000. The granule-bound starch synthase (GBSS1) gene in the Rosaceae: Multiple loci and phylogenetic utility. Mol. Phylogen. Evol. 17(3), 388-400.

Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C. 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. of Genet. 22, 521–565.

Guindon, S., Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52(5), 696-704.

Guindon S., Lethiec, F., Duroux, P., Gascuel, O. 2005. PHYML Online—A web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. 1, 33 (Web Server issue):W557-9.

Hebda, R.J., Chinnappa, C.C. 1980. Cytological evidence for generic limits in tribe Cladothamneae (Ericaceae). Can. J. Bot. 58, 1119-1120.

Huelsenbeck, J.P., Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754-755.

Johnson, L.A., Soltis, D.E. 1994. matK DNA sequence and phylogenetic reconstruction in Saxifragaceae s.s. Syst. Bot. 19, 143–156.

Kron, K.A. 1997. Phylogenetic relationships of Rhododendroideae (Ericaceae). Am. J. Bot. 84(7), 973-980.

Kron, K.A., Chase, M.W. 1993. Systematics of the Ericaceae, Empetraceae, Epacridaceae and related taxa based on rbcL sequence. Ann. Mo. Bot. Gard. 80(3), 735-741.

Kron, K.A., King, J.M. 1996. Cladistic relationships of *Kalmia*, *Leiophyllum*, and *Loiseleuria* (Phyllodoceae, Ericaceae) based on *rbcL* and nrITS data. Syst. Bot. 21(1), 17-29.

Kron, K.A., Judd, W.S., Stevens, P.F., Crayn, D.M., Anderberg, A.A., Gadek, P.A., Quinn, C.J., Luteyn, J.L. 2002. Phylogenetic classification of Ericaceae: Molecular and morphological evidence. Bot. Rev. 68(3), 335-423.

Luteyn, J.L. 1995. Ericaceae--Part II. The Superior-Ovaried Genera (Monotropoideae, Pyroloideae, Rhododendroideae, and Vaccinioideae p.p.). Flora Neotrop. Monograph 66, 107-122.

Mabberley, D.J. 1997. The Plant-Book. 2nd ed. Cambridge, UK, The Press Syndicate of the University of Cambridge. 858 pp.

Mason-Gamer, R.J., Kellogg, E.A. 1996. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). Syst. Biol. 45, 524–545.

Moretz, C.C. 2002. A systematic revision of *Elliottia* using molecular and morphological characters. M.S. Thesis. Wake Forest University.

Nylander, J.A.A. 2004. MrAIC.pl. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Oh, S-H., 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.

Olmstead, R.G., Sweere, J.A. 1994. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. Syst. Biol. 43, 467–481.

Olmstead, R.G., Michaels, H.J., Scott, K.M., Palmer, J.D. 1992. Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbcL*. Ann. Mo. Bot. Gard. 79, 249-265.

Philippe H.E., Snell, E.A., Bapteste, E., Lopez, P., Holland, P.W.H., Casane, D. 2004. Phylogenomics of eukaryotes: impact of missing data on large alignments. Mol. Biol. Evol. 21, 1740–1752.

Powell, E.A., Kron, K.A. 2001. Phylogenetics of the wintergreen group (*Gaultheria*, *Pernettya*, *Diplycosia*; Ericaceae). Syst. Bot. 26, 808–817.

Ronquist, F., Huelsenbeck, J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E., Small, R.L. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92(1), 142–166.

Steele, K.P., Vilgalys, R. 1994. Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene *matK*. Syst. Bot. 19, 126–142.

Stevens, P.F. 1969. Taxonomic studies in the Ericaceae. Ph.D. Thesis. University of Edinburgh.

Stevens, P.F. 1971. A classification of the Ericaceae: Subfamilies and tribes. Bot. J. Linn. Soc. 64, 1-53.

Stevens, P.F. (2001 onwards). Angiosperm Phylogeny Website. Version 9, June 2008 [and more or less continuously updated since].

Swofford, D.L. 2002. Paup* 4.0b10: Phylogenetic analysis using parsimony (* and other methods). Sunderland: Sinauer Associates.

White, T.J., Bruns, T., Lee, S., Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR protocols: A Guide to Methods and Applications. San Diego: Academic Press.

Wiens, J.J. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52, 528-538.

Wiens, J.J. 2005. Can incomplete taxa rescue phylogenetic analyses from long branch attraction? Syst. Biol. 54, 731-742.

Wiens, J.J., Moen, D.S. 2008. Missing data and the accuracy of Bayesian phylogenetics. J. Evol. Syst. 46(3), 307-314.

Yang, Z., Rannala, B. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. Mol. Biol. Evol. 14, 717-724.

Yoder, A.D., Irwin, J.A., and Payseur, B.A. 2001. Failure of the ILD to determine data combinability for slow loris phylogeny. Syst. Biol. 50, 408–42.

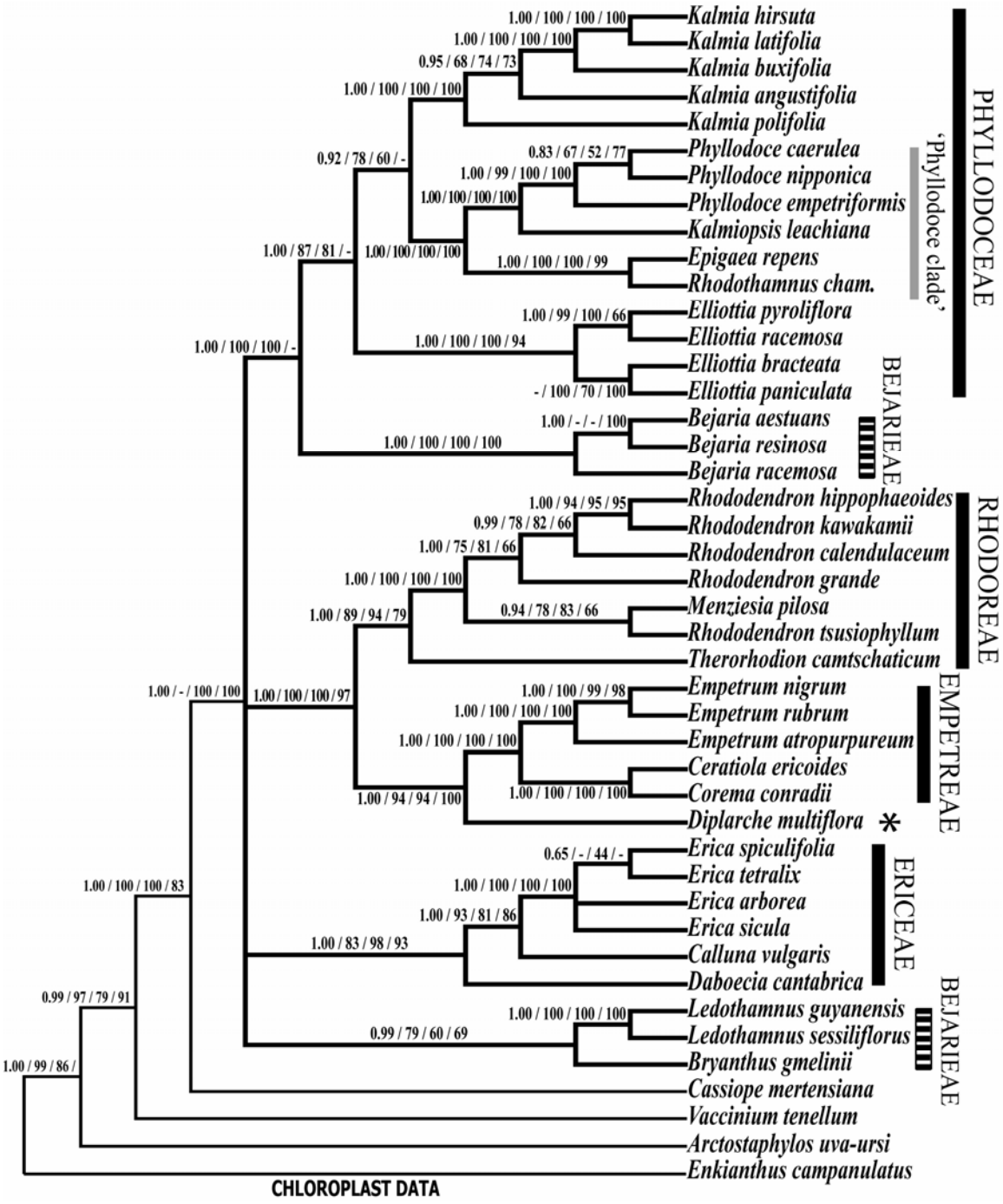


FIGURE 2.1. Total combined chloroplast data (*matK*, *ndhF*, *rbcL* and *trnS-G* spacer) analysis of the Ericoideae. Support values are to the left of nodes in the format (Bayesian posterior probability / ML bootstrap / MP bootstrap). Support derived from 20 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Tree length = 4263, CI = 0.5015, RI = 0.6160. Gray bar indicates the consistently resolved, four-genus 'Phyllodoce clade.' Hatched line indicates non-monophyly of Bejarieae sensu Kron et al. (2002). Placement of *Diplarche* sister to Empetreae is indicated by *****.

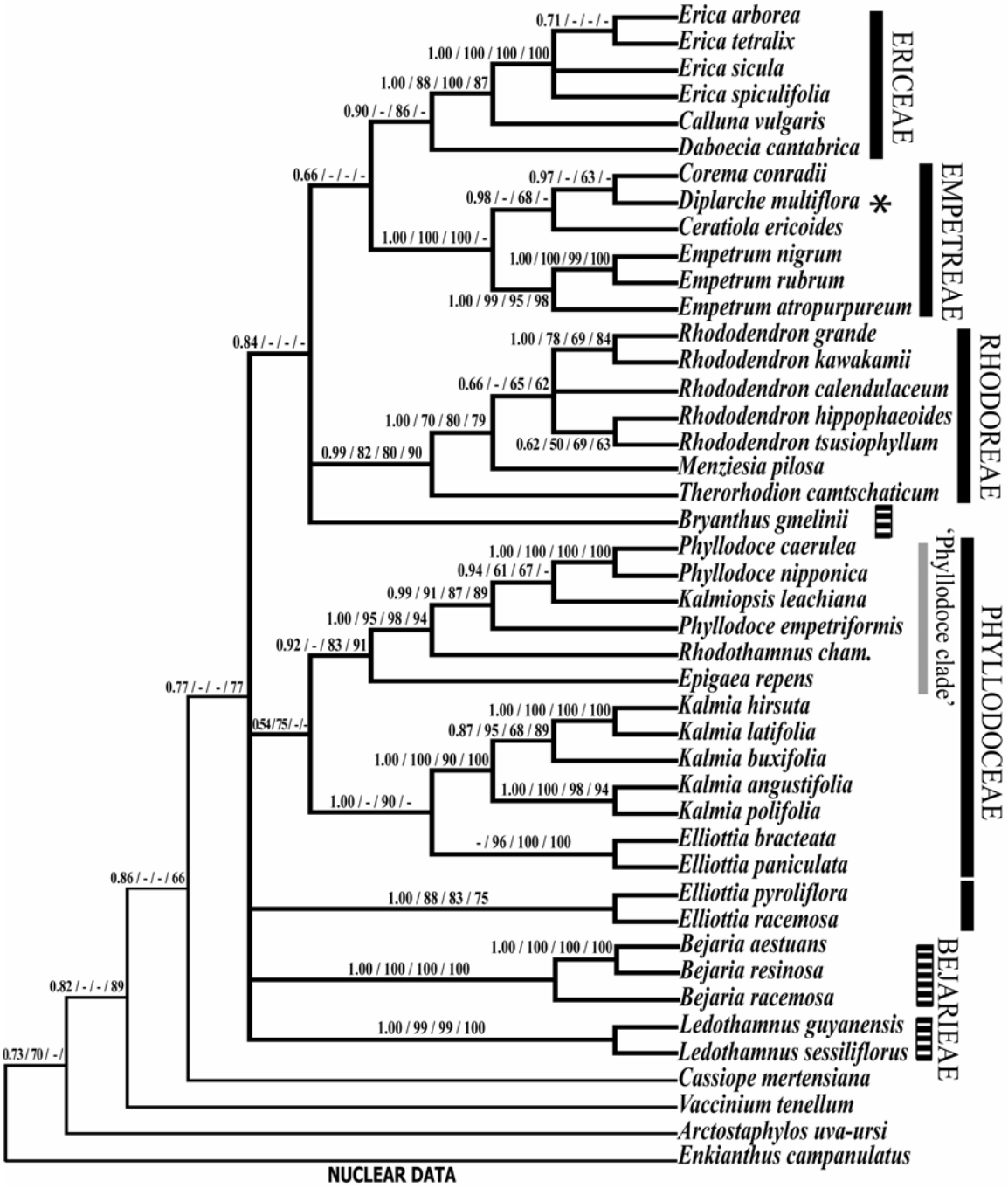


FIGURE 2.2. Total combined nuclear data (nrITS and *waxy*) analysis of the Ericoideae. Support values are to the left of nodes in the format (Bayesian posterior probability / ML bootstrap / MP bootstrap). Support derived from 20 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Tree length = 1581, CI = 0.5079, RI = 0.5948. Gray bar indicates the consistently resolved, four-genus 'Phyllodoce clade.' Hatched line indicates non-monophyly of Bejarieae sensu Kron et al. (2002). Placement of *Diplarche* sister to Empetreae is indicated by *****.

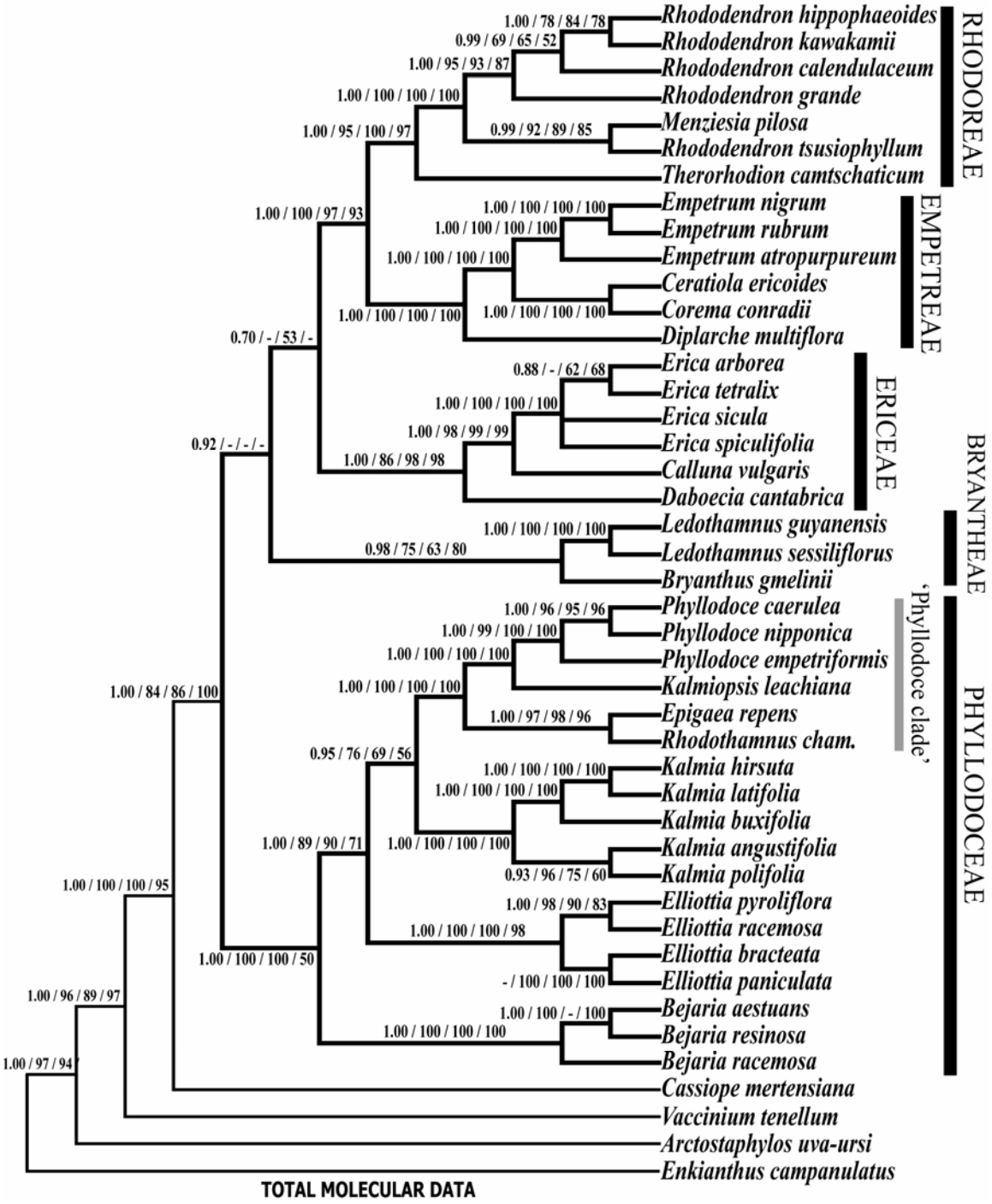


FIGURE 2.3. Total combined molecular data (nrITS, *waxy*, *matK*, *ndhF*, *rbcL* and *trnS-G* spacer) analysis of the Ericoideae. Support values are to the left of nodes in the format (Bayesian posterior probability / ML bootstrap / MP bootstrap). Support derived from 20 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Tree length = 5611, CI = 0.5004, RI = 0.6099. Tribes reflect the newly proposed classification of the Ericoideae.

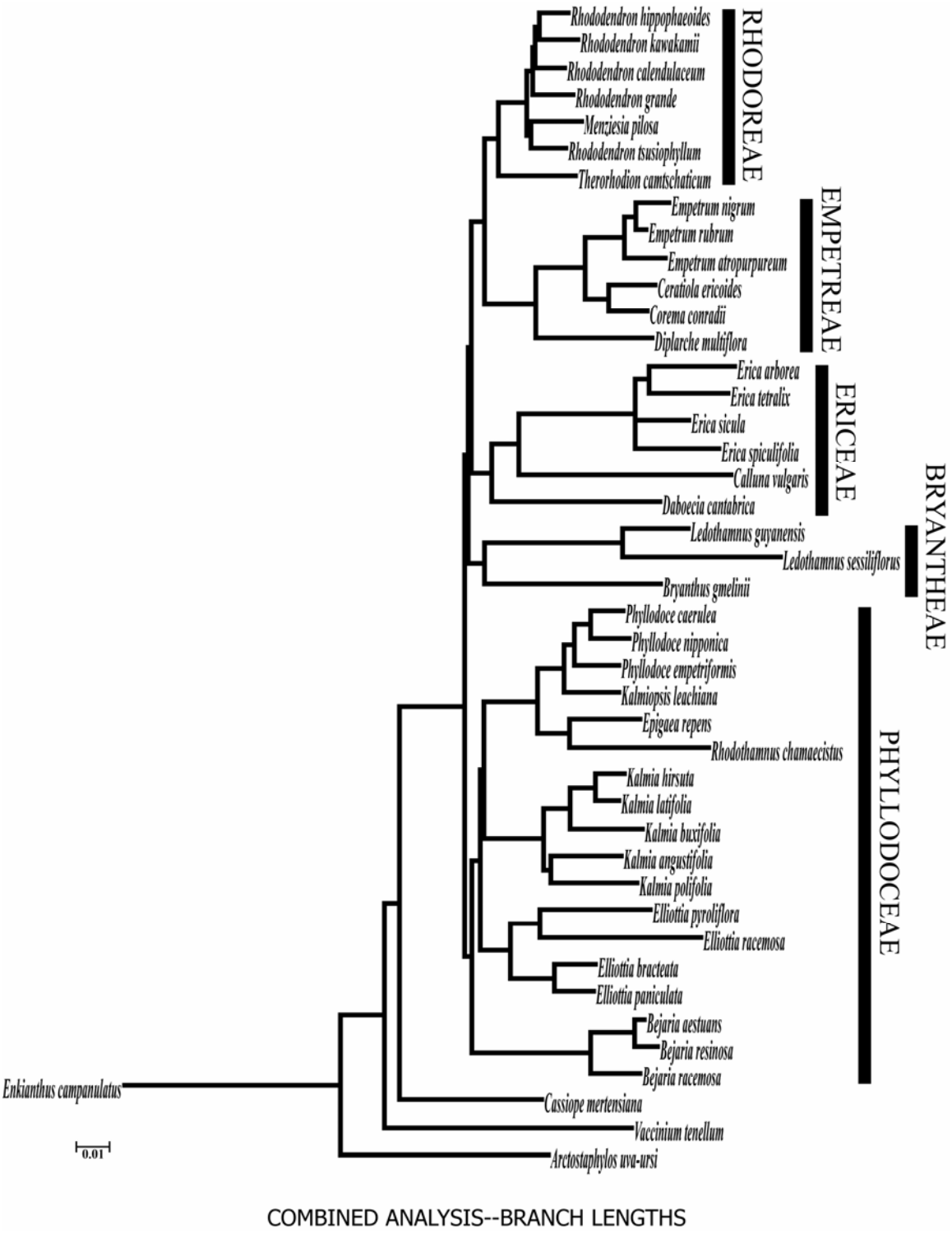


FIGURE 2.4. Phylogram of Ericoideae showing branch lengths resulting from Bayesian analysis of total combined molecular data.

TABLE 2.1. Abbreviated taxonomic history of the Ericoideae according to Copeland (1943), Cox (1948), Stevens (1971), Kron et al. (2002) and the current study.

	Copeland (1943)	**Cox (1948)**[2]	**Stevens (1971)**	**Kron et al., (2002)**	**Current Study**
Subfamily	Rhododendroideae	Rhododendroideae	Rhododendroideae	Ericoideae	Ericoideae
Phyllodoceae	*Rhodothamnus*	*Rhodothamnus*	*Rhodothamnus*	*Rhodothamnus*	*Rhodothamnus*
	Kalmiopsis	(incl. *Kalmiopsis*)	*Kalmiopsis*	*Kalmiopsis*	*Kalmiopsis*
	Phyllodoce	*Phyllodoce*	*Phyllodoce*	*Phyllodoce*	*Phyllodoce*
	Kalmia	*Kalmia*	*Kalmia*	*Kalmia*	*Kalmia* s.l.
	Loiseleuria	*Loiseleuria*	*Loiseleuria*	(incl. *Loiseleuria* &	*Epigaea*
	Leiophyllum[1]	*Leiophyllum*	*Leiophyllum*	*Leiophyllum*)	*Elliottia* s.l.
	Bryanthus	*Bryanthus*	*Bryanthus*	*Epigaea*	*Bejaria*
	Ledothamnus	*Ledothamnus*	*Ledothamnus*	*Elliottia* s.l.	
	Diplarche	*Diplarche*			
	Daboecia				
Rhodoreae	*Rhododendron*	*Rhododendron*	*Rhododendron*	*Rhododendron* s.l.	*Rhododendron* s.l.
	Hymenanthes	(incl. *Hymenanthes*)	(incl. *Azalea*	*Menziesia*	*Menziesia*
	Azalea	*Azalea*	& *Hymenanthes*)	*Therorhodion*	*Therorhodion*
	Tsusiophyllum	*Tsusiophyllum*	*Tsusiophyllum*	*Diplarche*	
	Azaleastrum	*Azaleastrum*	*Menziesia*		
	Ledum	*Ledum*	*Therorhodion*		
	Menziesia	*Menziesia*			
	Therorhodion				
Ericeae	N/A	N/A	Erica[3]	*Erica*[3]	*Erica*
				Calluna[3]	*Calluna*
				Daboecia[3]	*Daboecia*

	Copeland (1943)	Cox (1948)	Stevens (1971)	Kron et al. (2002)	Current Study
Empetreae	N/A	N/A	N/A	*Empetrum*[4]	*Empetrum*
				Ceratiola[4]	*Ceratiola*
				Corema[4]	*Corema*
					Diplarche
Bejarieae	*Bejaria*	*Bejaria*	*Bejaria*	*Bejaria*	N/A
				Ledothamnus	
				Bryanthus	
Cladothamneae	*Elliottia*	*Elliottia*	*Elliottia*	N/A	N/A
	(incl. *Tripetaleia*)	*Tripetaleia*	*Cladothamnus*		
	Cladothamnus	*Cladothamnus*	(incl. *Tripetaleia*)		
Daboecieae	N/A	*Daboecia*	*Daboecia*	N/A	N/A
Epigaeae	N/A	N/A	*Epigaea*	N/A	N/A
Diplarcheae	N/A	N/A	*Diplarche*	N/A	N/A
Calluneae[3]	N/A	N/A	*Calluna*[3]	N/A	N/A

[1] Copeland considered *Leiophyllum* to be more correctly named *Dendrium*, due to nomenclatural rules. [2] Cox did not examine *Therorhodion* due to lack of plant material, and did not discuss the placement of this taxon. [3] *Calluna* and *Erica* were not considered to be members of the subfamily Rhododendroideae at the time of each publication but instead were considered to be members of tribes of what was then called subfamily Ericoideae (not the modern Ericoideae). [4] Kron & Chase (1993) determined that members of Family Empetraceae were derived from within rhododendroid taxa, therefore it had not been included in previous studies of Ericoideae.

TABLE 2.2. Gene region, aligned length, number of potentially parsimony informative characters (and %), and number of missing taxa (and %) in a phylogenetic study of the Ericoideae.

Gene Region	**Aligned Length (bp)**	**Informative Characters**	**# Missing Taxa (%)**
rbcL	1288	162 (12.6%)	3 (6.8 %)
matK	1505	389 (25.8%)	1 (2.3 %)
ndhF	1909	532 (27.9%)	5 (11.4 %)
trnS-G-G	1489	313 (21.0%)	7 (15.9 %)
nrITS	688	235 (34.2%)	0
waxy	605	162 (26.8%)	10 (20.5 %)

APPENDIX 2.1. Taxa, voucher information and Genbank accessions for a phylogenetic study of the Ericoideae.

Species	**Voucher**	**nrITS**	***waxy***	***matK***	***trnS*-G-G**	***ndhF***	***rbcL***
Bejarieae							
Bejaria aestuans Mutis ex L. f	Luteyn 14175, NYBG	AF404817	DQ000589	**GU176669**	**GU176678**	DQ002362	**GU176638**
Bejaria racemosa Vent.	Kron 2070, NCU	U48604	DQ000594	U61327	**GU176679**	DQ002367	L12600
Bejaria resinosa Mutis ex L. f	Luteyn 14133, NY	**GU176622**	DQ000595	AF440412	**GU176680**	DQ002368	**GU176639**
Bryanthus gmelinii D. Don	Stevens s.n., WFU	U48612	**GU176650**	AF440413	**GU176681**	**GU176715**	AF419816
Ledothamnus guyanensis Meisn.	Picon & Williams 2910, WFU	**GU176623**	**GU176651**	AF440419	**GU176682**	**GU176716**	AF419827
Ledothamnus sessiliflorus N.E. Br.	Clement 2468A, NY	**GU176624**	**GU176652**		**GU176683**		**GU176640**
Empetreae							
Ceratiola ericoides Michx.	Kron 2069, WFU	AF519552	DQ000599	U61334	**GU176684**	**GU176717**	L12605
Corema conradii Torr. ex Loud.	Stevens s.n., HUH	AF519556	**GU176653**	AF440417		**GU176718**	AF419820
Empetrum atropurpureum Fern. & Wieg.	Chase 868, K	**GU176625**	DQ000601	U61355	**GU176685**	**GU176719**	**GU176641**
Empetrum nigrum L.	Hills, 89204, NCU	**GU176626**		**GU176670**		**GU176720**	AF419822
Empetrum rubrum Vahl	Chase 865, K	U48613	**GU176654**	U61342	**GU17686**	**GU176721**	**GU176642**
Ericeae							
Calluna vulgaris (L.) Hull 'Rannola'	1972-1443, RBGE	**GU176627**	**GU176655**	U61326	**GU176687**	**GU176722**	L12601
Daboecia cantabrica (Huds.) C. Koch	1975-1770, RBGE	AY520786	**GU176656**	U61349	**GU176688**	**GU176723**	L12611
Erica arborea L.	Small s.n., Heather Soc.	AY520788		AY517907			
Erica sicula Guss.	Chase 892, K	AY520804	**GU176657**	U61341	**GU176689**	**GU176724**	AF419823
Erica spiculifolia Reichb.	Chase 873, K	AY520785		U61337			AF419824
Erica tetralix L.	Anderberg 195-79, S	AY520806		U61340			AF419825
Phyllodoceae							
Elliottia bracteata Benth. & Hook. f	Chase 866, K	U48609	DQ000600	U61339	**GU176690**	**GU176725**	U49285
Elliottia paniculata Benth. & Hook. f	96RD00974FRBTU11	**GU176628**		**GU176671**	**GU176691**		**GU176643**
Elliottia pyroliflora (Bong.) Brim & Stev.	1934-009, RBGE	**GU176629**	**GU176658**	U61320		**GU176726**	**GU176644**
Elliottia racemosa Muhl. ex Elliott	1967-2632, RBGE	U48582		**GU176672**	**GU176692**	**GU176727**	L12615
Epigaea repens L.	Kron 162, WFU	U48611	**GU176659**	U61319	**GU176693**	**GU176728**	U49284

Kalmia angustifolia L.	Kron 1895, WFU	U48599	DQ000602	U61348	**GU176694**	**GU176729**	AF419826
Kalmia buxifolia (Berg.) Gift, Kron & Stevens	Gift s.n., HUH	U48581	**GU176660**	U61347	**GU176695**	**GU176730**	L12619
Kalmia hirsuta Walt.	Judd s.n., FLAS	U48601	**GU176661**	**GU176673**	**GU176696**	**GU176731**	**GU176645**
Kalmia latifolia L.	Kron 3020, WFU	U48600	**GU176662**	**GU176674**	**GU176697**	**GU176732**	U49294
Kalmia polifolia Wang.	Anderberg 325-89, S	U48597	**GU176663**	**GU183920**	**GU176698**	**GU176733**	U49289
Kalmiopsis leachiana (Hend.) Rehd.	Denton s.n.	U48608	DQ000603	U61323	**GU176699**	**GU176734**	U49290
Phyllodoce caerulea (L.) Bab.	1940-1013, RBGE	**GU176630**	DQ000604	U61318	**GU176700**	**GU176735**	AF419829
Phyllodoce empetriformis D. Don	Chase 871, K	U48607	DQ000605	U61333	**GU176701**	**GU176736**	U49291
Phyllodoce nipponica Makino	Anderberg 1756-77, S	U48606	DQ000606	U61325	**GU176702**	**GU176737**	U49292
Rhodothamnus chamaecistus Reichb.	Chase 877, K	U48605	DQ000607	U61321	**GU176703**	**GU176738**	U49287
Rhodoreae							
Diplarche multiflora Hook. f and Thomas	Suzuki et al. 8820561, HUH	**GU176631**	**GU176664**	AF440418		**GU176739**	AF419821
Menziesia pilosa Juss.	Anderberg 1360-65, S	AF393440	**GU176665**	U61351	**GU176704**	**GU176740**	U49293
Rhododendron calendulaceum (Michx.) Torr.	Kron s.n., WFU	**GU176632**	**GU176666**	**GU176675**	**GU176705**	**GU176741**	
Rhododendron grande Wight	1969-8606, RBGE	**GU176633**	EU669886	DQ002360	**GU176706**	DQ002383	**GU176646**
Rhododendron hippophaeoides Balf. f & Sm.	1932-1022, RBGE	**GU176634**	**GU176667**	U61353	**GU176707**	**GU176742**	L01949
Rhododendron kawakamii Hayata	79/026, RSF	**GU176635**		**GU176676**	**GU176708**	**GU176743**	
Rhododendron tsusiophyllum Tsugim.	76/353, RSF	**GU176636**		**GU176677**	**GU176709**	**GU176744**	**GU176647**
Therorhodion camtschaticum (Pall.) Sm.	73/054, RSF	**GU176637**	DQ000608	U61322	**GU176710**	DQ002382	AF419834
Outgroups							
Cassiope mertensiana G. Don	Anderberg 75-83, S	AF419798	DQ000598	U61346	**GU176711**	**GU176745**	L12603
Vaccinium tenellum Ait.	Kron & Powell s.n., WFU	AF382741		AF382818	**GU176712**	AF419769	**GU176648**
Arctostaphylos uva-ursi Spreng.	Anderberg 361-68, S	AF106811	**GU176668**	AF440411	**GU176713**	AJ236248	**GU176649**
Enkianthus campanulatus Nichols	Anderberg 14528, S	AF133752		U61344	**GU176714**	**GU176746**	L12616

Appendix 2.2. DNA regions, evolutionary models, primer sequences and PCR protocols for studies in the Ericoideae, Phyllodoceae and *Cassiope*.

DNA Region	**Evolutionary Model (AICc)**	**Primer Sequence**	**PCR Protocol**
rbcL	GTR+I+G	*rbcL* 1F: ATGTCACCACAAACAGAAACTAAAGCAAGT	35 x (95°C,1:00—48°C, 2:00—72°C, 3:00)
		rbcL 1367R: CTTTCCAAATTTCACAAGCAGCAG	
matK	GTR+G	*matK* 710F: GTATCGCACTATGTWTCATTTGA	1) 35 x (94°C, 1:00—49°C, 2:00—72°C,3:00)
		matK 1100R: CGTGCTTGCATTTTTCATTGC	
		matK 650F: ATCCAAATAAATTTTTGGGG	
		matK 1295F: GCATTATGTTAGATATCGAGG	
		matK 1168R: ATTGAATGAATTGATCGTA	
		matK 1600R: CCTCGATACCTAACATAATGC	2) 97°C,5:00—35 x (95°C, 0:30--42°C, 1:00—72°C, 5:00)
		matK 2200R: TCTGTATAACCTCCACAAAG	
		matK 1350R: CCATTTATTCATCAAAAGAAACG	
		matK 1300F: GATGCCTCTTCTTTGCATT	
ndhF	GTR+G	*ndhF* 1F: ATGGAACAKACATATSAATATGC	94°C,1:00 + 35 x (94°C, 0:30—48°C,0:30—72°C, 1:00) + 72°C, 5:00
		ndhF 1318R: CGAAACATATAAAATGCRGTTAATCC	
		ndhF 972F: GTCTCAATTGGGTTATATGATG	
		ndhF 2110R: CCCCCTAYATATTTGATACGTTCTCC	
		ndhF 954F: GTCTCAATTGGGATAT	
		ndhF 1955R: CGATTATATGACCAATCATATA	
***trnS*-G-G**	GTR+I+G	*trnS* GCU2: AACTCGTACAACGGATTAGCAATC	80°C,5:00 + 30 x (95°C, 1:00–66°C, 4:00) + 66°C, 10:00)
		trnG UUC2: GAATCGAACCCGCATCGTTAG	
		trnS-G-F2: TGGATTCTTAGACAATG §	
nrITS	HKY+G	ITS 5F: GGAAGTAAAAGTCGTAACAAGG	35 x (97°C, 0:45—52°C, 0:45—72°C, 2:00)
		ITS 4R: TCCTCCGCTTATTGATATGC	
waxy	GTR+G	*waxy* ex9F: GATACCCAAGAGTGGAAYCC *waxy* ex11R: GTTCCATATCGCATRGCRTG	97°C,1:00 + 40 x (97°C, 1:00—56°C, 1:00—72°C, 0:45+4 sec/cycle)

CHAPTER 3

PHYLOGENETIC RELATIONSHIPS AND HISTORICAL BIOGEOGRAPHY OF THE PHYLLODOCEAE (ERICACEAE)

ABSTRACT

The tribe Phyllodoceae is comprised of seven genera (*Bejaria*, *Elliottia*, *Epigaea*, *Kalmia*, *Kalmiopsis*, *Phyllodoce* and *Rhodothamnus*). A clade within this tribe, comprised of all genera except *Bejaria*, was the subject of a phylogenetic study using six molecular markers. The goal of the study was to reconstruct a species-level phylogeny for the purpose of examining the evolution of morphological characters and historical biogeography. Twenty-nine of 31 Phyllodoceae species were included. Most nodes were strongly supported using four distinct analytical approaches. Morphological synapomorphies were identified for some clades within the group. An analysis of ancestral areas and vicariance was conducted in S-DIVA. Because reliably dated fossils are not available in this group, an experimental approach to node dating was attempted. First, relative dates were estimated using BEAST. Vicariance events found to have occurred in the same span of time were then constrained to absolute dates based upon biogeographical data, followed by a second analysis in BEAST to estimate absolute dates for other nodes. This approach provides a potential way to address to the obstacle of fossil-poor northern hemisphere taxa, but future analyses should test additional biogeographical hypotheses.

INTRODUCTION

The tribe Phyllodoceae Drude is one of the most morphologically diverse and taxonomically unstable groups within the subfamily Ericoideae. Nearly all descriptions of the tribe have diagnosed the group based upon suites of characters that are homoplasious within the broader Ericoideae, rather than recognizing any particular synapomorphic character. In addition, at least one taxon from all currently recognized tribes within Ericoideae have at some point been classified within the Phyllodoceae, illustrating the difficulty in determining relationships of these taxa based upon morphological evidence alone. Gillespie and Kron (2010) used molecular data to clarify relationships within Ericoideae and to propose a new classification that includes five tribes: Rhodoreae (*Menziesia*, *Rhododendron* s.l. and *Therorhodion*), Empetreae (*Corema*, *Ceratiola*, *Diplarche* and *Empetrum*), Ericeae (*Calluna*, *Daboecia* and *Erica*), Bryantheae (*Bryanthus* and *Ledothamnus*) and Phyllodoceae (*Bejaria*, *Elliottia*, *Epigaea*, *Kalmia*, *Kalmiopsis*, *Phyllodoce* and *Rhodothamnus*). According to Gillespie and Kron (2010), the Phyllodoceae are sister to a clade comprised of the other four tribes.

Genera removed from Phyllodoceae

Several genera were considered to be part of the Phyllodoceae by past authors, but have been removed by subsequent investigators. These include *Bryanthus*, *Daboecia*, *Diplarche*, and *Ledothamnus*. All four were included in the Phyllodoceae by Drude (1889) and Copeland (1943) because they possess actinomorphic, gamopetalous corollas and unwinged seeds. Cox (1948) thought *Daboecia* was sufficiently different to be placed in its own tribe, Daboecieae, and Stevens (1969; 1971) agreed. Kron (1997) and Kron et

al. (2002) placed *Daboecia* in a clade now known as the tribe Ericeae along with *Erica* and *Calluna*, based on molecular data, with moderate to strong support. Gillespie and Kron (2010) found strong support for the same relationship using additional molecular data.

Bryanthus and *Ledothamnus* have long been classified within the Phyllodoceae (Drude 1897; Copeland 1943; Cox 1948; Stevens, 1969; 1971) based upon a suite of morphological characters found in other members of the tribe. Kron et al. (2002) showed that neither genus grouped with the Phyllodoceae clade, but was unresolved within the Ericoideae. Gillespie and Kron (2010) found that *Bryanthus* and *Ledothamnus* are sister taxa, and are more closely related to the Ericeae, Empetreae and Rhodoreae than to the Phyllodoceae.

Diplarche has been taxonomically difficult. It has been placed in the ericoid tribes Rhodoreae (Kron et al. 2002) and Phyllodoceae (Drude 1897; Copeland 1943; Cox 1948), and in the Diapensiaceae (Airy Shaw, 1964). Its uncertain placement was due in large part to unusual characters such as a septifragal capsule, serrate leaf margins and epipetalous stamens. Gillespie and Kron (2010) found strong molecular support for its placement within the ericoid tribe Empetreae.

Genera transferred to Phyllodoceae

Three genera were not originally included within the Phyllodoceae by major treatments, *Bejaria*, *Elliottia* and *Epigaea*, but evidence now supports placement within this group. *Epigaea* was classified by Drude (1897) as part of the tribe Andromedeae (then subfamily Arbutoideae) because of the presence of loculicidally dehiscent capsules,

a character state rarely found in the current Phyllodoceae (*Bejaria*, *Elliottia*, *Epigaea*, *Kalmia*, *Kalmiopsis*, *Phyllodoce* and *Rhodothamnus*). Treatments by Copeland (1943) and Cox (1948) agreed with Drude and *Epigaea* was not included in their studies of the Ericoideae (then called Rhododendroideae). Watson et al. (1967) was the first to include *Epigaea* within the Ericoideae, closely related to *Kalmia*, *Phyllodoce* and *Rhodothamnus*; thus the first recognition of a 'core' group of Phyllodoceae, based on numerical taxonomic studies. Stevens (1969; 1971) agreed with this assessment based on characters shared with other Ericoideae such as viscin threads, axile placentation, glandular hairs, and elongate synergid cells. Stevens (1969; 1971) recognized *Epigaea* in its own tribe (Epigaeae) because of the dissimilarity of their tetracytic stomata, unusual hooked petiole vascular bundles, uniseriate hairs inside the corolla, and a fleshy placenta. *Epigaea* has three species distributed disjunctly on three continents. *Epigaea repens* is found broadly in the eastern United States, *E. asiatica* is found in Japan and *E. gaultherioides* is found in the Caucasus.

Bejaria was classified by Drude (1897) within the Ledeae, along with *Elliottia*, *Tripetaleia*, *Cladothamnus* (all now included within *Elliottia*) and *Ledum*. This classification was based on the presence of polypetalous corollas and long-winged seeds. Copeland (1943) placed *Bejaria* in its own tribe (Bejarieae) and placed *Cladothamnus*, *Tripetaleia* and *Elliottia* in the Cladothamneae. The separation of these genera into two tribes was based on *Bejaria* having spindle-shaped seeds and members of the Cladothamneae having ovoid seeds. Watson et al. (1967) maintained Bejarieae as a separate tribe as a result of phenetic studies indicating that *Bejaria* was isolated from the other genera studied. Stevens agreed with the rationale of Copeland (1943) and Watson et

al. (1967) and retained *Bejaria* in its own tribe, Bejarieae. *Bejaria* is comprised of fifteen species distributed in mountainous areas of South and Central America, in the Caribbean, and into Florida, USA. Members of the genus have rotate to tubular, seven-merous flowers and viscin threads associated with pollen. *Bejaria racemosa* is the only species in section Racemosae (Clemants, 1995), based on the presence of chartaceous leaves and rotate flowers. The other 14 species were classified within Section *Bejaria*. Bush and Kron (2008) found that *B. racemosa* is likely sister to the remaining species. Within section *Bejaria*, there is a variety of node support, from very weak to very strong, providing sufficient resolution to allow informed selection of representatives for the current study.

Stevens (1969; 1971) retained Copeland's tribe Cladothamneae because of the absence of an endothelium, thin walled testa, and a floral syndrome that functions to deposit pollen on the dorsal surface of pollinators, all characters unlike the other tribes. Hebda and Chinnappa (1980) found that all members of the Cladothamneae have chromosome numbers X= 11, a number which is otherwise unknown in the Ericoideae, and therefore they determined that *Tripetaleia* and *Cladothamnus* should be included within *Elliottia*. Kron (1997) and Kron et al. (2002) found that both *Bejaria* and *Elliottia* were likely closely related to the Phyllodoceae based on molecular data. The precise relationships and support among these groups varied depending upon the data used, and therefore both were included within a broadly circumscribed Phyllodoceae in recent studies. Gillespie and Kron (2010) found strong support that *Bejaria* and *Elliottia* are successive sister taxa to the 'core' Phyllodoceae, and both genera were formally included in the tribe.

Genera consistently classified as Phyllodoceae

Four genera have been classified within the Phyllodoceae in all treatments since its original description by Drude (1897). These include *Kalmia, Kalmiopsis*, (since its discovery in the 1930s), *Leiophyllum* (=*K. buxifolia*), *Loiseleuria* (=*K. procumbens*) (Kron and King, 1996; Gillespie and Kron, 2010) *Phyllodoce*, and *Rhodothamnus*. These genera, along with *Epigaea*, are here referred to as the 'core' Phyllodoceae. Recognition of these taxa as a group has been based on suites of characters. Drude (1897) found that these taxa had actinomorphic, mostly gamopetalous corollas and unwinged seeds. Copeland's (1943) Phyllodoceae were the same as Drude's (1897), but Copeland focused on floral and leaf anatomy. He described the Phyllodoceae as lacking viscin threads associated with pollen and bud scales, but having resorption tissue. Cox's (1948) wood anatomy studies resulted in a classification similar to Copeland (1943). Stevens' (1969, 1971) circumscription agreed largely with Drude (1897), except for Stevens' inclusion of *Epigaea* within the Phyllodoceae.

As currently circumscribed, *Kalmia* is a mostly North American genus with one species in the Caribbean and one whose range is circumboreal, with some populations at lower latitudes (*K. procumbens*). Ten species are recognized. Two of these (*K. hirsuta* and *K. ericoides*) were placed by some authors (e.g. Small, 1903; Britton and Wilson, 1920) in a separate genus, *Kalmiella*, but all recent treatments include them within *Kalmia* (e.g. Copeland, 1943; Southall and Hardin, 1974; Ebinger, 1974; Liu et al., 2009). Disagreements occur among some of these authors as to the rank of various taxa, particularly regarding *K. microphylla* and *K. polifolia* and the Cuban taxa, which are recognized as subspecies by some. Judd (1983) examined morphological variation among

the Cuban taxa and found insufficient morphological gaps with which to recognize three species, and instead recognized a broadly circumscribed *K. ericoides*. Kron and King (1996) found that *Leiophyllum buxifolium* and *Loiseleuria procumbens* were both phylogenetically nested within *Kalmia*. Therefore the genus as currently described includes *K. ericoides* (Cuba), *K. hirsuta* (S. Georgia and Florida, USA), *K. cuneata* (North Carolina, USA sandhills), *K. carolina* (mostly North Carolina, USA), *K. buxifolia* (disjunct in the southeastern U.S. and the New Jersey pine barrens), *K. microphylla* (Rockies, Pacific Northwest, USA), *K. latifolia* (widespread in the eastern US), *K. polifolia* (broadly northern North America), *K. angustifolia* (eastern/northeastern US), and *K. procumbens* (circumboreal).

Kalmia is a morphologically diverse genus, with species ranging from small prostrate subshrubs (e.g. *K. procumbens*) to small trees (e.g. *K. latifolia*). Leaves may be persistent or deciduous (*K. cuneata*), and alternate, opposite or whorled. Inflorescences can be axillary or terminal, and may be a raceme, panicle, umbel or fascicle. The ovary may be 2- or 5-locular. Several *Kalmia* are known to be toxic to livestock (*K. angustifolia*, *K. latifolia*, *K. microphylla*, and *K. polifolia*), and *K. latifolia* reportedly has been used extensively in folk medicine (Ebinger, 1974).

Jaynes (1968) found that in general, genetic barriers among species are strong, because many interspecific crosses yielded no viable offspring; however, questions of species boundaries exist particularly between *K. angustifolia*/*K. carolina* and *K. polifolia*/*K. microphylla*. There is no consensus on whether *K. angustifolia* and *K. carolina* should be considered species or varieties; Flora of North America (Liu et al., 2009) considers them varieties, whereas Weakley (2008) considers them separate species.

A few characters distinguish them, including the presence of a glandular calyx in *K. angustifolia*, densely abaxial leaf surfaces lacking glands in *K. carolina*, and aspects of stomata size and distribution. Their distributions are adjacent in eastern United States, with *K. carolina* having a smaller, more southern distribution. Species boundaries between *K. polifolia* and *K. microphylla* are apparently more complex. Morphologically, the two taxa are very similar and leaf morphology in *K. microphylla* is highly variable and overlapping with *K. polifolia*. Jaynes (1968) found that *K. polifolia* and *K. microphylla* interspecific crosses generated viable offspring with relative ease, suggesting that they should be recognized as a single species, but also found differences in their ability to cross with other *Kalmia* species. The two consistently differ in chromosome number across their ranges (*K. microphylla* is X=12, whereas *K. polifolia* is X=24). The most recent treatment of the entire genus, Flora of North America, (Liu et al., 2009), recognized the two as separate species.

Two species of *Rhodothamnus* are currently recognized. *Rhodothamnus chamaecistus* is distributed on rocky outcrops in the eastern European Alps, while *R. sessilifolius* is more narrowly distributed in the Artvin Province in northeastern Turkey. Both species are evergreen subshrubs that occupy rocky slopes where they encounter relatively little competition. Both have rotate corollas and leaves with ciliate margins.

Kalmiopsis was not described until 1930 (and therefore was not known to Drude). Since its discovery, *Kalmiopsis* has been included within the Phyllodoceae in nearly all treatments. (*K. leachiana* was originally described as a *Rhododendron* by Henderson (1931) but Rehder (1932) quickly transferred it to *Kalmiopsis*) All authors since this time (e.g. Stevens 1969, 1971; Kron, 1997; Kron et al., 2002, Gillespie and Kron, 2010) have

found *Kalmiopsis* to be grouped within the Phyllodoceae. Meinke and Kaye (2007) described a second, more narrowly distributed species, *K. fragrans*, and together they represent the only reported endemic vascular plant genus in Oregon. Recognition of a second species was made on the basis of numerous, but subtle, differences such as ovary color (*K. fragrans* = yellow to gold, *K. leachiana* = greenish-gold), the timing of pollen shedding (*K. fragrans* = delayed several hours past corolla expansion, *K. leachiana* = approximately at the same time as corolla expansion), and substrate type (*K. fragrans* = 'tuffaceous' sandstone, *K. leachiana* = broader range, including ultra-mafic rock).

Presently, *Phyllodoce* includes seven species, all of which are small, spreading shrubs that occur in tundra, alpine or subalpine meadows. All *Phyllodoce* have highly inrolled (ericoid) leaves. Good (1926) recognized two subgenera within *Phyllodoce*, and these correspond to corolla type. He placed those with campanulate corollas and relatively large corolla lobes in the subgenus *Parabryanthus*. These included *P. nipponica*, *P. empetriformis* and *P. breweri*. Good (1926) placed those species with urceolate corollas and relatively small corolla lobes within the subgenus *Eu-Phyllodoce* (*P. caerulea*, *P. glanduliflora* and *P. aleutica*). *Phyllodoce deflexa* was newly described by Yang (1990) as having affinities to *P. caerulea*. Because it also has an urceolate corolla, Good would have presumably included it within subgenus *Eu-Phyllodoce*. This genus exhibits an interesting geographical distribution. *Phyllodoce glanduliflora* and *P. aleutica* are both found in mountainous areas of northwestern North America, but *P. aleutica* extends into northeast Asia. *Phyllodoce breweri* is narrowly distributed in the Cascade Mountains of California, USA. *Phyllodoce empetriformis* is also found in northwestern North America south to northern California, and southern Idaho and

Wyoming, USA, but also has a disjunct distribution in the mountains of northern Arizona (Arizona Natural Heritage Program, 2009). *Phyllodoce caerulea* is distributed circumboreally. In North America, it occurs in Alaska, Northwest Territories and across eastern Canada and Greenland, southward to New England. On the Eurasian continent it is found in the northern provinces of China, Korea, Siberia and northern Mongolia, and Japan, northern Scandinavia, Scotland, and as far south as the Pyrenees (Vinogradova, 2001). *Phyllodoce deflexa* is known only from mountainous areas in Jilin Province, China, near the North Korean border (Wu et al., 2005). *Phyllodoce nipponica* is endemic to Japan.

The purpose of the current study is to use multiple nuclear and chloroplast markers to estimate phylogenetic relationships among species within the tribe Phyllodoceae, excluding *Bejaria*. Less emphasis is placed on *Bejaria*, because a phylogenetic study of this genus has recently been published (Bush and Kron, 2008). Subsequently, biogeographical analyses are conducted to ancestral areas analysis and to estimate node ages within the context of Northern Hemisphere biogeography.

MATERIALS AND METHODS

DNA was isolated from fresh, silica-dried or herbarium plant material using the Qiagen Plant Mini Kit (Qiagen, Valencia CA, USA) with modifications following Drábková et al. (2002). Methods pertaining to DNA accessions generated prior to this study are described in Gillespie and Kron (2010).

Thirty taxa representing the Phyllodoceae were included (Table 3.1). Taxon sampling was concentrated in the clade comprised of *Elliottia*, *Epigaea*, *Kalmia*,

Kalmiopsis, *Phyllodoce* and *Rhodothamnus*. The ingroup includes all seven species of *Phyllodoce*, both species of *Kalmiopsis*, both species of *Rhodothamnus*, all ten species of *Kalmia* (including *Leiophyllum* and *Loiseleuria*), all four *Elliottia*, and two of three species of *Epigaea*. One taxon (*Epigaea gaultherioides*) was not sampled because available herbarium material was degraded to a point that DNA extraction and/or PCR amplification were not successful, and living material could not be acquired. All four members of the sister group to this clade, *Elliottia*, were included, to aid in analysis of ancestral areas and estimation of node ages (see below), although relationships among these taxa have been established (Gillespie and Kron, 2010). Three members of the outer-most clade of Phyllodoceae, *Bejaria* (15 spp.), were included. Sampling within *Bejaria* here was limited because no additional taxa were available and no additional data collection beyond the three markers used by Bush and Kron (2008) was possible due to limited availability of plant material. Nine outgroup taxa were included. Two taxa from each additional tribe (Rhodoreae, Empetreae, Ericeae and Bryantheae) within the Ericoideae (Appendix 3.1) and one from *Cassiope*, which has been established as being sister to the Ericoideae (Kron, 1997; Kron and King, 1996; Kron et al., 2002; Gillespie and Kron, 2010). For each of these groups one representative from the basal-most node and one from a more recent node were included to represent each tribe in the clade sister to the Phyllodoceae. The exception to this is *Cassiope* which consists of a single genus whose interspecific relationships are poorly understood.

Six DNA regions (four chloroplast and two nuclear) were included in this study. These include chloroplast regions *rbcL, ndhF, matK,* and *trnS*^{GCU}*-trnG*^{UCC}*-trnG*^{UCC} intergenic spacer (*trnS-G-G*), and nuclear regions *waxy*/GBSS-1 (*waxy* exons 9-11) and

nuclear ribosomal Internal Transcribed Spacer (nrITS). PCR protocols and primer sequences are as reported in Gillespie and Kron (2010). Single gene, combined nuclear, combined chloroplast and total data matrices are deposited in TreeBase (www.treeBASE.org), under accession number 10491. All newly generated sequences are deposited in Genbank (www.ncbi.nlm.nih.gov) under numbers HM182001 - HM182049 and are listed in Appendix 3.1.

Amplified fragments were cleaned using Qiagen™ QIAquick Gel Isolation Kit (Qiagen, Valencia CA, USA). DNA was sequenced on an ABI 377 automated sequencer at the DNA Sequencing and Gene Analysis Laboratory at the Wake Forest University School of Medicine (Winston-Salem, NC) or at Nevada Genomics Center (Reno, NV). Sequences were minimally edited and manually aligned in Geneious 4.7.4 (Drummond et al., 2006).

Nuclear regions were cloned using the TOPO TA kit (Invitrogen, Carlsbad CA, USA) for representative taxa within the Ericaceae (see Gillespie and Kron, 2010). Five clones were sequenced for each of these taxa and no evidence of multiple copies of nuclear genes was found within the Phyllodoceae and relatives, therefore cloning was not generally carried out during this study.

Maximum Parsimony (MP) analyses were conducted using PAUP*4.0b10, (Swofford, 2002) with the following options: Parsimony-informative characters were unordered and equally weighted, gaps were treated as missing data, searches were heuristic with TBR branch swapping and 1000 random stepwise addition replicates. Relative clade support was assessed using bootstrap analysis (Felsenstein, 1985; Felsenstein, 1988) with the full bootstrap option in PAUP* (10,000 replicates).

Maximum Likelihood analyses were carried out using the RaxML online server (Stamatakis, 2006, Stamatakis et al., 2008) on the CIPRES web portal at San Diego Supercomputing Center, as well as using the PhyML online server (Guindon and Gascuel, 2003; Guindon et al., 2005) on the ATCG Montpellier bioinformatics platform. For RAxML, total molecular evidence, chloroplast, and nuclear datasets were run with genes as separate partitions. For PhyML, all three datasets (total, chloroplast and nuclear) were each run under a single GTR model (i.e. not partitioned by gene), as determined by the AICc criterion (Aikake, 1974) in MrAIC.pl 1.4.3 (Nylander, 2004). AICc was chosen as the criterion for model selection because it is not hierarchical in nature and also corrects for small sample sizes (approximately 40 and below). Bootstrap analysis (100 replicates) was conducted to determine node support. Both Maximum Likelihood strategies were used because they use fundamentally different bootstrapping strategies and are therefore essentially different approaches.

For Bayesian analyses, the data were partitioned by DNA region and evolutionary models were chosen using the AICc criterion in MrAIC.pl 1.4.3 (Nylander, 2004) (Appendix 3.1). Bayesian MCMC analyses (Yang and Rannala, 1997) as implemented in MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) consisted of 20 million generations with a burn-in of 25%. Clade support is reported as posterior probabilities.

Dispersal-vicariance analysis was carried out in S-DIVA (Statistical Dispersal-Vicariance Analysis, Yu et al., 2009, Yu et al., in press). S-DIVA builds on the earlier program DIVA 1.1 (Ronquist, 1997; 2001), which reconstructs ancestral areas of a clade within a parsimony framework. DIVA 1.1 penalizes reconstructions that require

dispersals and extinctions to explain present-day distributions, and therefore favors solutions requiring vicariance. S-DIVA incorporates the suggestion of Nylander (2008) to account for phylogenetic uncertainty that should increase the number of equally optimal reconstructions. Because DIVA favors vicariance, the root of a tree tends to reconstruct ancestral areas composed of many or most of the possible areas. Outgroups were coded for areas in order to minimize this artifact, following the suggestion of Ronquist (1996). The graphical output includes reconstructions that take into account both uncertainty resulting from widespread taxa or mutually exclusive distributions and phylogenetic uncertainty.

Estimation of node ages was conducted using the BEAST v1.5.3 software package (Drummond and Rambaut, 2007). Because no reliable fossils are available for calibration of nodes, an iterative approach to node dating in BEAST was used. The initial analysis calculated relative dates of all nodes in the phylogeny using a relaxed uncorrelated lognormal clock model (Drummond et al., 2006). The site model chosen was HKY + Γ. The tree prior incorporated a Yule speciation process, and the starting tree was randomly selected. The analysis was run for 20 million generations, sampled every 1000 generations. The Effective Sample Size (ESS) was inspected in Tracer v1.5.3 (Rambaut and Drummond, 2007) in order to ensure that a large enough sample was generated; all ESS values were above 200. The tree sample was compiled in TreeAnnotator v1.5.3 (Drummond and Rambaut, 2007) and node ages were visualized in FigTree v1.3.1 (Drummond and Rambaut, 2007).

The second analysis in BEAST was run with two nodes constrained based upon biogeographic evidence (see Results, below). The age estimate was constrained with a

lognormal probability distribution, which tightly constrains younger ages, but allows the older age limit to 'float'. All other parameters were left unchanged relative to the first BEAST analysis.

RESULTS

In all presentation of results, clade support values are presented in the following format: Bayesian PP / RAxML bootstrap / PhyML bootstrap / MP bootstrap. A dash (-) indicates support lower than 0.50 PP or 50% bootstrap. For the purpose of readability, any clade supported by exactly (1.00 / 100 / 100 / 100) is abbreviated 100x4. Gene regions, best evolutionary model, aligned length, and percent (%) missing taxa are reported in Table 3.1.

Chloroplast Data (Fig. 3.1.)

Analyses of chloroplast data reveal four main clades that have long branches (Fig. 3.4) leading to them. These include a clade comprised of *Bejaria* (100x4), *Elliottia* (1.00 / 99 / 100 / 91), *Kalmia* (100x4), and the '*Phyllodoce* clade' (100x4). *Bejaria* is strongly supported in most analyses (1.00 / 100 / 100 / -) as sister to the remaining Phyllodoceae, followed by *Elliottia* (1.00 / 80 / 65 / -). While the *Bejaria* clade itself is strongly supported in all analyses, the sister relationship of *B. aestuans* and *B. resinosa* is supported only by MP analyses. Within the *Elliottia* clade, *E. bracteata* is strongly supported as sister to *E. paniculata* (1.00 / 97 / 100 / 100) and *E. racemosa* is sister to *E. pyroliflora* with varying support (0.98 / 74 / 54 / 66). Within the *Kalmia* clade, there are four small clades, each with strong support: Clade 1 = [*K. hirsuta* + *K. ericoides*] + *K.*

latifolia. Clade 2 = *K. buxifolia* + *K. procumbens*. Clade 3 = [*K. angustifolia* + *K. carolina*] + *K. cuneata*. Clade 4 = *K. microphylla* + *K. polifolia*. The relationship of Clade 1 as sister to Clade 2 is strongly supported (100x4), but the relationship of this group to Clade 3 is low in all analyses (0.74 / - / 76 / 67). This eight-taxon clade is sister to Clade 4, which is strongly supported as a clade (100x4). Within the '*Phyllodoce*' clade, three genera are supported as monophyletic: *Phyllodoce* (1.00 / 100 / 100 / 99), *Kalmiopsis* (100x4), and *Rhodothamnus* (100x4). In the initial Bayesian and MP analyses of chloroplast data that included *Epigaea asiatica* and *E. repens*, *E. asiatica* was placed as sister to the *Phyllodoce* clade with no bootstrap or PP support, and a branch length longer than any other taxon in the analysis (not shown). In PhyML analyses, *E. asiatica* was placed as sister to the all Ericoideae with no bootstrap support. Given the long branch length, and the absence of evidence that *E. asiatica* is not properly placed within *Epigaea*, *E. asiatica* was excluded from further analyses of chloroplast data. Therefore, *Epigaea* is represented by a single taxon here, *E. repens*. *Epigaea repens* is strongly supported (1.00 / 100 / 99 / 98) as sister to a clade of both *Rhodothamnus* species (100x4). *Epigaea* + *Rhodothamnus* are strongly supported (100x4) as sister to a clade comprised of *Phyllodoce* and *Kalmiopsis*. *Phyllodoce caerulea* is sister to *P. deflexa* (0.89 / - / 50 / -), and these two are together sister to *P. nipponica* (1.00 / 92 / 84 / 62). *Phyllodoce empetriformis* is supported by only PhyML as sister to this clade (0.70 / - / 100 / -). *Phyllodoce aleutica* and *P. glanduliflora* is sister (1.00 / 99 / 100 / 100) and are sister to the clade including *P. empetriformis* and *P. caerulea* (1.00 / 100 / 100 / 99). *Phyllodoce breweri* is sister to all other *Phyllodoce*, but this relationship, and therefore the monophyly of *Phyllodoce*, is poorly supported (0.68 / 56 / 59 / 55).

Nuclear Data (Fig. 3.2.)

Analyses of nuclear data recognize the same four main clades as the chloroplast data, but the two largest clades have some differences in species relationships. *Bejaria* is supported as a clade by all analyses except MP (1.00 / - / 100 / 100). Nuclear data support the sister relationship of *B. aestuans* and *B. resinosa* better than chloroplast data (- / 100 / - / 100). *Ledothamnus* (Bryantheae) is supported as sister to *Bejaria* by the nuclear data (1.00 / 100 / 100 / -), a result in conflict with chloroplast data. *Bejaria* + *Ledothamnus* is supported (0.90 / - / - / -) by Bayesian analysis as sister to the remaining Phyllodoceae. *Elliottia* is poorly supported as a clade by nuclear data (0.58 / - / 50 / -), as is the sister relationship of *E. pyroliflora* and *E. racemosa* (0.69 / 50 / 56 / -). However, the sister relationship of *E. bracteata* and *E. paniculata* is strongly supported (1.00 / 99 / 100 / 100). Relationships within *Kalmia* are topologically different from the chloroplast data, and generally have lower support, especially at the deeper nodes. Within Clade 3, *K. angustifolia* is sister to *K. carolina* (100x4) and they are together sister to *K. cuneata* (0.66 / 66 / - / -). These three are weakly supported (0.69 / - / - / -) as sister to Clade 4, which is strongly supported as a clade by Bayesian analysis and weakly to moderately supported by the other analyses (0.98 / 82 / 62 / 62). This five-taxon clade is sister to Clade 2, but with only marginal Bayesian support (0.88 / - / - / -). Both nodes associated with Clade 1 are supported by all analyses (100x4), and Clade 1 is sister to all other *Kalmia*. The *Phyllodoce* clade is generally well supported (1.00 / 95 / 83 / 74). However, nuclear data places *Epigaea* as sister to the remaining taxa (1.00 / 98 / 99 / 94) rather than sister to *Rhodothamnus*, which is strongly supported (1.00 / 99 / 100 / 99). Here, *Epigaea*, and then *Rhodothamnus* are sister to a strongly supported (0.98 / 90 / 89 /

85) clade comprised of *Phyllodoce* and *Kalmiopsis*. Within this clade, *P. aleutica* is sister to *P. glanduliflora* (0.93 / 98 / 75 / 87). These two are weakly to moderately supported (0.64 / 85 / 50 / 61) and together are sister to a strongly supported (0.93 / 100 / 80 / 77) clade comprised of *P. caerulea* and *P. nipponica*. *Phyllodoce deflexa* is sister to this clade of four taxa, but with weak support (0.57 / - / - / -). *Kalmiopsis* is strongly supported as a clade (100x4), and is sister to the already-mentioned *Phyllodoce* with weak support (0.74 / - / - / 58). *Phyllodoce breweri* and *P. empetriformis* are sister to the other *Phyllodoce* and *Kalmiopsis*, making *Phyllodoce* potentially paraphyletic, and they are unresolved with respect to each other. Because of the lack of support along the backbone within the *Kalmiopsis* + *Phyllodoce* clade, there is essentially no resolution, except for clades comprised of the two *Kalmiopsis*, *P. caerulea* + *P. nipponica*, and *P. aleutica* + *P. glanduliflora*. Relationships among the *Phyllodoce* clade, the *Kalmia* clade and the *Elliottia* clade are unresolved, but are supported by Bayesian analysis as a monophyletic group (0.90 / - / - / -).

Total Molecular Data (Fig. 3.3.)

The combination of chloroplast and nuclear data provide the most resolution and support of any analyses. The Phyllodoceae are strongly supported as a clade (1.00 / 78 / 100 / -) by all analyses except MP. The four main clades (*Phyllodoce* clade, *Kalmia*, *Elliottia* and *Bejaria*) are all strongly supported in all analyses (100x4). *Bejaria* is sister to the remaining Phyllodoceae, and the sister relationship of *B. aestuans* and *B. resinosa* is supported in RAxML and MP analyses (- / 100 / - / 100). Within *Elliottia*, *E. bracteata* and *E. paniculata* are sisters (100x4), as are *E. pyroliflora* and *E. racemosa* (0.99 / 83 /

60 / 68). The topology within *Elliottia* is in agreement with analyses of both chloroplast and nuclear data. *Elliottia* is sister to the remaining Phyllodoceae excluding *Bejaria*, but with relatively weak support (0.73 / 72 / - / -). Within *Kalmia*, there are two clades. All nodes within these clades are strongly supported (100x4) and are in agreement with the chloroplast phylogeny, but not the nuclear phylogeny. Clade 3 is sister to Clade 4. Clade 3 and Clade 4 are each strongly supported (100x4), but their sister relationship is strongly supported only by Bayesian analysis (0.90 / 50 / 60 / -). This topology is in agreement with the nuclear topology, but not the chloroplast phylogeny. Within the *Phyllodoce* clade, *Epigaea* is sister to *Rhodothamnus* (1.00 / 99 / 99 / 96), and these two taxa are strongly supported (100x4) as sister to *Kalmiopsis* + *Phyllodoce*. This is in agreement with the chloroplast data, but not the nuclear data, where *Epigaea* and *Rhodothamnus* branch successively as the deepest nodes of the group. The two *Kalmiopsis* form a clade (100x4). *Phyllodoce* is supported as monophyletic and with stronger support than in the chloroplast analyses (0.81 / 71 / 65 / 63). *Phyllodoce breweri* is sister to the other *Phyllodoce* (1.00 / 100 / 99 / 100), followed by *P. empetriformis* (1.00 / 100 / 95 / 94). The remaining five taxa form two clades. *Phyllodoce aleutica* is sister to *P. glanduliflora* (100x4) and these are sister to a strongly supported clade of *P. nipponica*, *P. caerulea* and *P. deflexa* (1.00 / 97 / 91 / 71). *Phyllodoce caerulea* is sister to *P. deflexa* with low support (0.74 / 56 / 54 / -). The position of *P. empetriformis* is in conflict with both the nuclear and chloroplast data, and the position of *P. breweri* is potentially in conflict with the nuclear data, where it is placed 'outside' *Kalmiopsis* with poor support. A phylogram (Fig. 3.4) indicates very short branches as the basal most nodes within Ericoideae, in agreement with the broader analysis within the Ericoideae (Chapter 2/ Gillespie and

Kron, 2010). The phylogram also indicates short branches at the basal nodes within the *Phyllodoce* + *Kalmiopsis* clade and the *Kalmia* clade.

Ancestral Areas (Fig. 3.5.)

Reconstruction of the ancestral area of *Kalmiopsis* is straightforward. Both taxa are found in northwestern North America (C) and therefore the ancestor has the same distribution. Within *Phyllodoce*, only two areas must be considered; all taxa are found in either northwestern North America (C) or northeastern Asia (D), and all nodes have a single likely reconstruction. The ancestor of *P. aleutica* and *P. glanduliflora*, both of which are found in northwestern North America (C), occupied the same area. Although *P. caerulea* is distributed in five areas, its two nearest relatives, *P. deflexa* and *P. nipponica*, occur only in northeast Asia (D) and the ancestor of all three have an optimal reconstruction in northeast Asia (D). The ancestor of these five *Phyllodoce* occurred in both areas (C and D), implying a vicariance event associated with this node. *Phyllodoce empetriformis* and *P. breweri* are both found in northwestern North America (C) and the ancestor of all *Phyllodoce* is optimized to have occurred in northwestern North America, implying a range extension into northeastern Asia (D) by the ancestor of *P. caerulea*, *P. deflexa*, *P. nipponica*, *P. aleutica* and *P. glanduliflora*. A range expansion into southwestern North America (K) by *P. empetriformis* is also necessary. Because the basal-most node in both *Phyllodoce* and *Kalmiopsis* is optimized as northwestern North America (C), the ancestor of *Phyllodoce* + *Kalmiopsis* is also northwestern North American in origin.

The two *Rhodothamnus* species are both found in mountainous areas of southern Europe. Reconstruction of the ancestral node is also for southern Europe (H). Only one *Epigaea* was included in the analysis (*E. repens*) and its modern distribution is the northern and southern parts of eastern North America (A and B). Reconstruction of the ancestral node of *Epigaea* and *Rhodothamnus* is most likely all three areas (A, B and H). This reconstruction suggests a vicariance event at the point that *E. repens* and *Rhodothamnus* diverged. However, there is a less likely reconstruction for the common ancestor of *Epigaea* and *Rhodothamnus* of only southeastern North America and southern Europe (B and H). This reconstruction implies extinction of the *Rhodothamnus* ancestor from southeastern North America (B) and a range extension of *E. repens* from southeastern North America (B) into northeastern North America (A). The two *Epigaea* not included in the present analysis occur in northeast Asia (D) and in the Caucasus (H). The addition of these taxa would likely change this reconstruction, and this is examined more closely in the discussion.

The ancestral area of the *Phyllodoce* clade (*Kalmiopsis*, *Phyllodoce*, *Epigaea* and *Rhodothamnus*) is complicated by 1) the two possible reconstructions for the ancestor of *Epigaea repens* and *Rhodothamnus* and 2) the lack of distribution information in the analysis for *E. asiatica* and *E. gaultherioides*. The data presented here support four different reconstructions; two are somewhat less likely. These are 1) eastern North America (A and B) plus northwestern North America (C) plus southern Europe, and 2) eastern north America (A and B) plus northwestern North America. The two most likely reconstructions are 1) southeastern North America (B) and northwestern North America

(C), or 2) southeastern North America (B), northwestern North America (C) and southern Europe (H).

Nearly all nodes within *Kalmia* have a single optimal reconstruction. The common ancestor of *K. angustifolia* and *K. carolina* was likely widespread in Eastern North America (A and B), implying a vicariance event resulting in their present-day distributions. The reconstruction of the node subtending these two taxa plus *K. cuneata* is in southeastern North America, requiring a range extension from southeastern North America (B) to the northeast (A) by the ancestor of *K. angustifolia* and *K. carolina*. The best reconstruction of the ancestral node of *K. microphylla* and *K. polifolia* is the northwestern part of North America (C), requiring a range extension into northeastern North America (A) by *K. polifolia*. The ancestor of all five taxa was most likely distributed in southeastern North America (B) and northwest North America (D), requiring a vicariance event. The ancestor of *K. ericoides*, *K. hirsuta* and *K. latifolia* likely occurred only in southeastern North America (B), with a range extension by the ancestor of *K. hirsuta* and *K. ericoides* into the Caribbean (G). A vicariance event apparently resulted in the distribution of *K. ericoides* in Cuba (G) and *K. hirsuta* in the southeastern United States (B). Reconstruction for the ancestor of *K. procumbens* and *K. buxifolia* is not satisfactory using DIVA. Because *K. procumbens* has a circumboreal distribution and because the modern distributions of the two taxa overlap only in northeastern North America (A), virtually all possible reconstructions are equally likely. This would imply that *K. buxifolia* underwent a range expansion to the south and *K. procumbens* underwent long distance dispersal and/or range expansions around the boreal latitudes. Despite multiple equally probable reconstructions for *K. procumbens* + *K.*

buxifolia, the node uniting these five *Kalmia* (*K. ericoides*, *K. hirsuta*, *K. latifolia*, *K. procumbens*, and *K. buxifolia*), is reconstructed as southeastern North America (B) with a small probability that the ancestral area was both southeastern (B) and northeastern North America (A). The optimal reconstruction of the ancestor of *Kalmia* + the *Phyllodoce* clade is also southeastern North America (B).

The ancestor of *E. bracteata* and *E. paniculata* likely occurred in eastern Asia (D). The optimal reconstruction for the ancestor of *E. pyroliflora* and *E. racemosa* is in the southeast as (B) well as the northwest of North America (C). These two reconstructions are highly supported and no other reconstruction was indicated for either pair of species. Two equally likely reconstructions exist for the ancestor of all four *Elliottia*. One is all three areas (southeastern (B) and northwestern North America (C) and northeast Asia (D)), and the second is just two areas (C and D). If the former reconstruction explains the present-day distributions, then a vicariance event is implied to generate the distribution of *E. pyroliflora* + *E. racemosa* (B and C) and *E. paniculata* + *E. bracteata* (D). The ancestor of the *Phyllodoce* clade, *Kalmia* and *Elliottia* most likely occurred in southeastern North America.

The ancestral area for *Bejaria* may be southeastern North America (B) and mountainous Central and South America (I). Despite the sparse sampling in *Bejaria* for this study, the reconstruction may be reasonable since all areas within which *Bejaria* are distributed are represented. In this case, the common ancestor of all Phyllodoceae appears to have likely been distributed in southeastern North America.

Estimation of Relative Node Ages (Fig. 3.6)

The initial BEAST analysis estimated relative node ages. Inspection of the resulting chronogram revealed that two vicariance events involving the same areas appear to have occurred during the same relatively short span of time. These are the vicariance events involving the ancestor of *Elliottia racemosa* and *E. pyroliflora* and the common ancestor of *Kalmia angustifolia* and *K. polifolia*. Both events involve an ancestral range of Southeastern North America + Northwestern North America, which was then dissected into the two individual areas. Examination of climate data for central North America revealed that moderate cooling and aridification occurred in the central part of the continent beginning approximately 50mya and proceeding until approximately 35mya, when progressively more rapid cooling and drying continued, culminating in the glacial cycles associated with the Pleistocene, beginning approximately 1.6 mya (Graham, 1999; Tiffney and Manchester, 2001; Milne and Abbott, 2002). Therefore, it was hypothesized that these two seemingly identical vicariance events occurred together, sometime during the period of climate change that resulted in the cooling and aridification of central North America, most likely between 35 mya and 1.6 mya.

Estimation of Absolute Node Ages (Figs. 3.7A and 3.7B)

To examine whether this hypothesis can reasonably explain other vicariance events, the BEAST analysis was repeated, this time constraining the *Elliottia* and *Kalmia* vicariance events (see above) to 2 mya with a lognormal probability distribution reaching back to 35 mya. This prior allowed the age of the *Elliottia* and *Kalmia* nodes to 'drift' between 1.6 mya and 35 mya, but the tail is much longer toward the older time (and

consequently most of the probability area lies further back in time than 1.6 mya). The rationale for this approach is that the greatest impact on the vicariance events in question (cooling and drying of the central part of North America) was complete by the end of the Tertiary (1.6 mya) The vicariance event did not necessarily occur exactly at the end of the Tertiary, but may instead have occurred any time during the period of progressive cooling that proceeded most rapidly between 35 mya and 1.6 mya (Graham, 1999; Tiffney and Manchester, 2001; Milne and Abbott, 2002). Other parameters of the analysis were left unchanged relative to the initial BEAST analysis to determine relative node ages. The resulting chronogram (Figure 3.6A and 3.6B) indicated that vicariance events involving the ancestor of *Kalmia hirsuta* and *K. ericoides*, and the ancestor of *Elliottia pyroliflora* and *E. racemosa*, each occurred at approximately 4.25 mya, during the Pliocene. This analysis also estimated the date of the vicariance event proposed to have resulted in the two *Elliottia* clades to be at approximately 5.25 mya, earlier in the Pliocene. The Northwestern North America – Eastern Asia disjunction within *Phyllodoce* was estimated to have originated approximately 2.25 mya, during the late Pliocene. Two Quaternary variance events were proposed. One is the division of the ancestral range of *Kalmia hirsuta* + *K. ericoides* in the Southeastern United States and Cuba into its modern distribution. This event is estimated to have occurred approximately 0.5 mya. The other vicariance event within *Kalmia*, which resulted in the modern distributions of *K. angustifolia* and *K. carolina* is estimated to have occurred approximately 0.75 mya.

DISCUSSION

Phylogeny and Morphology

Within the genus *Phyllodoce*, *P. glanduliflora* + *P. aleutica* is the only strongly supported clade recovered in all analyses of nuclear, chloroplast and combined data. These taxa have been considered subspecies by some authors, and so this result is not unexpected. Two other taxa anticipated to be closely related, *P. caerulea* and *P. deflexa*, are sister taxa in the chloroplast data and combined data, but without strong support. In these analyses, the two are found in a clade with *P. nipponica* with strong support. The nuclear data places *P. deflexa* outside the clade containing *P. caerulea* and *P. nipponica*, sister to a clade containing these two taxa plus *P. aleutica* + *P. glanduliflora*, but this clade of four taxa has good support only from RAxML analyses. Therefore, *P. deflexa* is not strongly excluded from the clade comprised of *P. nipponica* and *P. caerulea*. Since the three taxa are well supported as a clade in chloroplast and combined analyses, it is reasonable to conclude that they are closely related. *Phyllodoce deflexa* is not sister to *P. nipponica* in any analysis, and therefore the two most reasonable topologies follow that of the combined analyses; that is, *P. nipponica* + [*P. caerulea* + *P. deflexa*] or the nuclear topology, *P. deflexa* [*P. caerulea* + *P. nipponica*]. The nuclear topology carries higher support. The position of *P. empetriformis* is also unclear when the chloroplast and nuclear trees are compared. In the chloroplast data, *P. empetriformis* is sister to [*P. caerulea* + *P. deflexa*] + *P. nipponica*. In the nuclear data, it is essentially unresolved within the genus. However, in the combined analysis, the position of *P. empetriformis* is very strongly supported as sister to the five-taxon clade comprised of *P. aleutica* + *P. glanduliflora* and *P. caerulea* + *P. deflexa* + *P. nipponica*.

Based on the combined data, several 'mapped' morphological characters are apparent that unite clades within the *Phyllodoce* clade. All *Phyllodoce* have sessile, ericoid leaves. Both petals and leaves are somewhat reflexed (but *Rhodothamnus* also have sessile leaves, and *Epigaea* have reflexed petals). *Phyllodoce* and *Kalmiopsis* share short, abrupt glandular hairs on the pedicel, to the exclusion of *Rhodothamnus* and *Epigaea*. Within *Phyllodoce*, several shifts in character state appear to occur after the divergence of *P. breweri*. These include a corymbose inflorescence, long glandular hairs on the pedicel, a densely tomentose abaxial leaf midrib, and included stamens. Viscin threads have not been reported in *P. caerulea*, *P. deflexa*, *P. nipponica*, *P. aleutica P. glanduliflora* and *P. empetriformis*, indicating a possible secondary loss of viscin threads after the divergence of *P. breweri*. Viscin threads are present in most Phyllodoceae. A shift from exserted stamens to included stamens occurs after the divergence of *P. empetriformis*. Four characters occur in the clade comprised of [[[*P. caerulea* + *P. deflexa*] + *P. nipponica*] + [*P. aleutica* + *P. glanduliflora*]]. These include valvate sepals rather than imbricate sepals, short ovoid sepals rather than elongate sepals, urceolate corollas and carpels that are longer than wide. *Phyllodoce nipponica* does not have these character states, however, and would therefore represent secondary reversals. It is worth noting that the results of the nuclear analysis exclude *P. nipponica* from the above group, although the support for this relationship is not particularly strong.

Four small, well supported clades are always recovered within *Kalmia*. The chloroplast topology is [4, [3, [2 + 1]]], whereas the nuclear is [1, [2, [3 + 4]]] and the total data is [[1+2], [3+4]]. That is, of the two clades recovered in the total analysis, one clade agrees with part of the nuclear topology and the other clade agrees with part of the

chloroplast topology. The conflict between the nuclear and chloroplast data is not strongly supported in any analysis. Further, it is apparent that although less nuclear data is included in this analysis, the chloroplast signal does not obscure the nuclear signal, since clade 3+4 is found in the nuclear and combined analyses, but not the chloroplast analyses.

In the *Kalmia* clade the deeper nodes appear to be supported primarily by molecular data and few distinctive floral characters distinguish species. Corolla color varies from whitish to deep pink even on the same individual. Inflorescences are also variable, and can be fascicles, racemes, panicles or solitary flowers; two types may co-occur on the same plant. Phyllotaxy is also not uniform within species and populations within a species may vary in this character. Some characters within *Kalmia* are autapomorphic (e.g. deciduous leaves and white corolla in *K. cuneata*, five stamens in *K. procumbens*), and some characters unite just two taxa (e.g. loss of anther pockets in *K. buxifolia* and *K. procumbens*, alternate leaves in *K. ericoides* and *K. hirsuta*). The sister taxa *K. buxifolia* and *K. procumbens* both lack the anther pockets and the tension-loaded pollen dispersal mechanism that characterizes the remaining members of the genus. Because the two species form a relatively derived clade within *Kalmia*, the absence of these two characters most likely represent secondary losses. These two species also have the only flowers in the genus whose petals are not fused nearly the entire length of the corolla. In *K. procumbens*, the petals are fused to approximately half the length of the corolla, and in *K. buxifolia*, they appear polypetalous or nearly so. It seems clear that the absence of spring-loaded anthers is a direct consequence of the absence of anther pockets.

However, it is unclear whether the absence of anther pockets is a direct consequence of the loss of petal fusion. Developmental analysis might clarify this issue.

Relationships within *Elliottia* are generally well-supported. A morphological synapomorphy for the clade is the presence of flattened anther filaments, a character not found in other Phyllodoceae (or Ericoideae). *Elliottia bracteata* and *E. paniculata* are sister taxa. Based on the results here, these two taxa are united by the presence of three stamens, three carpels, six stamens, and imbricate scales associated with the leaf buds. *Elliottia pyroliflora* and *E. racemosa* are sister taxa. *Elliottia pyroliflora* is found in the Pacific northwest coast of North America on volcanic substrates, and *E. racemosa* is found nearly exclusively along the Altamaha Grit, an unusual sandstone formation along the coastal plain of Georgia, USA (Harper, 1906). The presence of a single palisade layer unites these two taxa.

Historical Biogeography

Geographical distributions of extant taxa within the *Phyllodoce* clade are consistent with a remnant Tertiary flora. *Epigaea*'s three taxa are distributed on three continents. *Epigaea repens* is fairly widespread in eastern North America, *E. gaultherioides* is found in the Caucasus region, and *E. asiatica* is found in eastern Asia only. Both *Rhodothamnus* species are found only in the Caucasus. Clearly, the ancestor of these two species is inferred to have also occurred in the Caucasus. However interesting this distribution is, reconstructing the ancestral distribution of *Epigaea*, or *Epigaea* + *Rhodothamnus*, quickly becomes problematic, since two *Epigaea* species are not included. In Bayesian analyses that included *E. asiatica* (not shown), the two *Epigaea*

were resolved as sister, but posterior probability support was low for that clade, as well as for the *Epigaea* + *Rhodothamnus* clade. Since S-DIVA uses node support to determine support for alternative reconstructions, this diminished support resulted in many more equally likely reconstructions. Therefore the distribution examined using S-DIVA did not include two major disjunctions in this small genus. In S-DIVA, it is possible to include geographical distributions for missing taxa, by identifying a node subtending the likely phylogenetic position for the missing distributional data and considering the taxa to be sister to that node. Because a single *Epigaea* was included, it was not possible to have the distributions of *E. asiatica* or *E. gaultherioides* included in light of a phylogenetic topology, but they could be considered as a polytomy along with *E. repens*. That analysis in S-DIVA analysis suggested that the ancestor of all *Epigaea* most likely occurred in all four areas currently occupied by the three *Epigaea*: northeastern (A) and southeastern North America (B), northeast Asia (D) and the Caucasus (H), or alternately only in the Caucasus (H). The most likely reconstruction for the ancestral range of the entire *Epigaea* + *Rhodothamnus* clade is northwestern North America (C) and the Caucasus (H), although several other relatively unlikely reconstructions exist. The vicariance event resulting in the ancestral ranges of the three *Epigaea* species (Eastern North America, Caucasus and Eastern Asia) and the two *Rhodothamnus* species (Southern Europe / Caucasus) is estimated to have occurred approximately 3.75 mya, during the Pliocene. This date must be considered a very preliminary estimate, given that two species (*E. asiatica* and *E. gaultherioides*) are not included in the analysis. Because S-DIVA and BEAST analyses are sensitive to tree topology and sequence divergence respectively, it is anticipated that the addition of these two taxa would impact ancestral areas analysis and

node dating heavily. *Epigaea* is a small group and each species has a non-overlapping range on a different continent, relative to other members of the genus. The three *Epigaea* species are very similar in appearance, and so addition of these taxa would be valuable in examining whether the nodes are very old and morphological stasis has occurred, or whether they are relatively recent and insufficient time has passed for morphological divergence to have occurred.

The *Kalmiopsis* + *Phyllodoce* clade is somewhat more tractable. Both *Kalmiopsis* and all *Phyllodoce* except *P. caerulea*, *P. deflexa* and *P. nipponica* are found in the northwestern part of North America. *Phyllodoce nipponica* is found in Japan and surrounding areas, *P. caerulea* is distributed circumboreally and *P. deflexa* is found near the edge of the range of *P. caerulea* in northeastern China. Therefore the ancestor of *Kalmiopsis* and *Phyllodoce* most likely occurred in northwestern North America (C), where several divergence events occurred with no change in ancestral distribution, suggesting perhaps a series of sympatric speciation events. These events may coincide with the uplift of the western mountain ranges in North America, including the Cascades, which attained their maximum heights at approximately 3 mya (Graham, 1999). This is the same estimate acquired by BEAST for the origin of the *Kalmiopsis* + *Phyllodoce* clade. The ancestor of *P. nipponica*, *P. deflexa* and *P. caerulea* likely arose in eastern Asia, and *P. caerulea* expanded around the northern latitudes in both North America and Eurasia. The estimated date of the vicariance event resulting in Phyllodoce species in both Northwestern North America and Eastern Asia is approximately 2.25 mya, but the 95% confidence interval encompasses a window of time from approximately 3.5 mya to 1 mya. This interval is consistent with an ancestor of the clade comprised of *P. aleutica* and

P. nipponica occurring in both Northwestern North America and Eastern Asia as late as the Pliocene, followed by a disruption of that ancestral range by a rise in sea levels associated with interglacial periods during the Pleistocene (Graham, 1999).

Kalmia appears to have originated in southeastern North America. The earliest divergence within *Kalmia* resulted in two widespread taxa, at approximately 4.5 mya according to the BEAST analysis presented here. The ancestor likely occupied eastern North America, where its range expanded to include northeastern North America (*K. buxifolia*) and high latitudes circumboreally (*K. procumbens*). It is not possible using S-DIVA with the current level of phylogenetic and geographic detail to better understand how this circumboreal distribution arose. It is interesting to note that the circumboreal *Kalmia* generates a more complex reconstruction at more basal nodes within the *Kalmia* clade, yet the same result did not occur with the circumboreal *Phyllodoce* (*P. caerulea*) in the *Phyllodoce* clade. This widespread ancestral taxon also underwent divergence within southeastern North America (*K. latifolia* and *K. hirsuta*), and a vicariance event lead to speciation within Cuba (*K. ericoides*). This node was estimated to have occurred at approximately 0.5 mya, placing it during the Pleistocene. This is a reasonable estimate, because the proximity of Florida, USA and Cuba means that these two areas would be in contact, or very nearly so, when sea levels were lower during glacial maxima. Subsequent interglacials and increases in sea level would have divided Florida from Cuba, resulting in the modern distributions. It is not possible to assign a particular glacial maximum to this event without a more refined analysis, but the current analysis places the node in the Middle Quaternary (780 kya to 122 kya; Graham, 1999; Richmond and Fullerton, 1986). In the other *Kalmia* ancestor, a range expansion into northwestern North America was

followed by a vicariance event resulting in a northwestern taxon and a southeastern taxon. The northwestern taxon diverged into *K. polifolia* and *K. microphylla*. This is the node constrained to 1.6 mya in the second BEAST analysis, and it corresponds with the Pliocene/Pleistocene boundary recognized by Richmond and Fullerton (1986). The southeastern taxon gave rise to two southeastern taxa (*K. cuneata* and the ancestor of *K. carolina* and *K. angustifolia*), and a more recent vicariance event resulted in *K. angustifolia* in northeastern North America and *K. carolina* in North and South Carolina in the early to mid Quaternary, approximately 0.75 mya.

The ancestor of all *Elliottia* may have occurred in southeastern North America and northeastern Asia. An equally likely reconstruction indicates that the ancestor occurred in all areas occupied by extant *Elliottia* (northwestern and southeastern North America and northeastern Asia). A vicariance event may have led to the divergence of the two species pairs, or alternately dispersal into northwestern North America. Divergence within Japan resulted in *E. bracteata*, which occurs only at higher elevations and latitudes, and *E. paniculata*, which is more widespread at lower elevations.

The approach of first conducting an analysis of relative ages revealed two nodes for which an absolute date and a probability distribution on that date, as well as a specific biogeographical explanation could be proposed. Further analyses to estimate absolute dates based on these node constraints yielded reasonable dates for shallow nodes. However, it is highly likely that the Ericoideae and *Cassiope* originated earlier than 9 mya, given that their modern distributions are consistent with an origin earlier in the Tertiary. However, this analysis provides a case study for this approach, when fossil evidence is not available.

Summary

The 'core' Phyllodoceae (*Kalmia*, *Epigaea*, *Rhodothamnus*, *Kalmiopsis* and *Phyllodoce*) exhibit remarkable amounts of morphological diversity. Habit differences ranging from mat forming plants to small trees, inflorescence types ranging from solitary flowers to panicles and fascicles, and distributions ranging from a single region to all three northern continents illustrate the difficulty in determining phylogenetic and biogeographic patterns. In most cases, a morphological synapomorphy at deeper nodes is not apparent based on the molecular phylogeny, therefore the molecular phylogeny of this group is particularly useful for understanding evolutionary relationships.

Reconstruction of ancestral areas indicates a complex biogeographic history for this group, similar to northern hemisphere groups previously studied (e.g., Xiang et al., 1998; 2000; Wen et al., 1998). Extant members of each clade apparently arose through a combination of vicariance, long distance dispersal, extinction and range expansion events. Estimation of relative node ages revealed that two clades, *Kalmia* and *Elliottia*, may have been impacted by the same biogeographic event, namely the cooling and aridification of the central part of the North American continent, resulting in a Northwest-Southeast disjunction. A hypothesis about the absolute date of this disjunction was proposed in order to approximate the dates of other vicariance events. Absolute age estimates for shallower nodes within the Phyllodoceae appear to be reasonable, but as expected, age estimates on deeper nodes have large error bars and are likely to be underestimated. Greater taxon sampling and reliable age constraints for deeper nodes may help clarify these nodes.

ACKNOWLEDGEMENTS

The authors wish to thank the following institutions and individuals for their contribution to this research. Royal Botanic Garden Edinburgh (RBGE), Swedish Museum of Natural History (S), Harvard University Herbaria (HUH), Rhododendron Species Foundation (RSF), and Florida Museum of Natural History (FLAS) allowed use of herbarium or garden specimens. Jiang Hongyuan and Du Fengguo (Forestry College of Beihua University, China), Salih Terzioğlu (Karadeniz Technical University, Turkey), Richard Abbott (Florida Museum of Natural History), Arne Anderberg (Swedish Museum of Natural History), Robert Meinke and Aaron Liston (Oregon State University) provided plant material. Torsten Eriksson, Randy Small, A.J. Harris and Yu Yan provided technical support. Funding for this study was provided by NSF grant DEB-023043 and Wake Forest University.

LITERATURE CITED

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Automatic Control 19, 716–723.

Airy Shaw, H.K. 1964. Studies in the Ericales: XIV. The systematic position of the genus *Diplarche* Hook. f. & Thoms. Kew Bull. 17(3), 507-509.
Arizona Natural Heritage Program 2009

Britton, N.L. and Wilson, P. 1920. Descriptions of Cuban plants new to science. Mem. Torr. Bot. Club. 16(2), 57-118.

Bush, C.M. and Kron, K.A. 2008. A phylogeny of *Bejaria* (Ericaceae: Ericoideae) based on molecular data. J. Bot. Res. Inst. Texas. 2(2), 1193-1205.

Clements, S.E. 1995. Ericaceae subfamily Rhododendroideae. 6. Bejaria. In: J.L. Luteyn, ed. Ericaceae, Part II, The superior-ovaried genera. Flora Neotropica Monograph 66:54–106. New York Botanical Garden, New York.

Copeland, H.F. 1943. A study, anatomical and taxonomic, of the genera of Rhododendroideae. Am. Midl. Nat. 30(3), 533-625.

Cox, H.T. 1948. Studies in the comparative anatomy of the Ericales I-Subfamily Rhododendroideae. Am. Midl. Nat. 39(1), 220-245.

Drábková, L., Kirschner, J., Vlček, Č. 2002. Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens in Juncaceae. Plant Mol. Biol. Rep. 20, 161-175.

Drude, O. 1889. Ericaceae. Pp. 15–65 in A. Engler & K. Prantl (eds.), Die Natürlichen Pflanzenfamilien 4(1). W. Engelmann, Leipzig.

Drummond, A.J., Kearse, M., Heled, J., Moir, R., Thierer, T., Ashton, B., Wilson, A., and Stones-Havas, S. 2006. Geneious v2.5.

Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology 4: 699–710.

Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7: 214.

Ebinger, J.E. 1974. A systematic study of the genus *Kalmia* (Ericaceae). Rhodora 76, 315-398.

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. of Genet. 22, 521–565.

Gillespie, E. L. and Kron, K. A. 2010. Molecular phylogenetic relationships and a revised classification of the subfamily Ericoideae (Ericaceae). Mol. Phyl. Evol. doi:10.1016/j.ympev.2010.02.028

Good, R.D.O. 1926. The genera *Phyllodoce* and *Cassiope*. J. Bot. 64, 1-10.

Graham, A. 1999. Late Cretaceous and Cenozoic History of North American Vegetation North of Mexico. New York and Oxford. Oxford University Press. 350 pp.

Guindon, S., Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52(5), 696-704.

Guindon S., Lethiec, F., Duroux, P., Gascuel, O. 2005. PHYML Online—A web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. 1, 33 (Web Server issue):W557-9.

Harper, R.M. 1906. Some new or otherwise noteworthy plants from the coastal plain of Georgia. Bull. Torrey Bot. Club 33: 229-233.

Hebda, R.J., Chinnappa, C.C. 1980. Cytological evidence for generic limits in tribe Cladothamneae (Ericaceae). Can. J. Bot. 58, 1119-1120.

Henderson, L.F. 1931. New plants from Oregon. Rhodora. 34, 203-206.

Huelsenbeck, J.P., Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754-755.

Jaynes, R.A. 1968. Interspecific crosses in *Kalmia*. Am. J. Bot. 55(9), 1120-1125.

Judd, W.S. 1983. *Kalmia ericoides* revisited. Rhodora. 85, 45-54.

Kron, K.A. 1997. Phylogenetic relationships of Rhododendroideae (Ericaceae). Am. J. Bot. 84(7), 973-980.

Kron, K.A., King, J.M. 1996. Cladistic relationships of *Kalmia*, *Leiophyllum*, and *Loiseleuria* (Phyllodoceae, Ericaceae) based on *rbcL* and nrITS data. Syst. Bot. 21(1), 17-29.

Kron, K.A., Judd, W.S., Stevens, P.F., Crayn, D.M., Anderberg, A.A., Gadek, P.A., Quinn, C.J., Luteyn, J.L. 2002. Phylogenetic classification of Ericaceae: Molecular and morphological evidence. Bot. Rev. 68(3), 335-423.

Liu, S., Denford, K.E., Ebinger, J.E., Packer, J.G., and Tucker, G.C. 2009. *Kalmia*. In: Flora of North America Editorial Committee, eds. 1993+. Flora of North America North of Mexico. 15+ vols. New York and Oxford. Vol. 8, pp. 371, 372, 374, 375, 449, 480, 481 .

Meinke, R.J. and Kaye, T.N. 2007. *Kalmiopsis fragrans* (Ericaceae), a new distylous species from the southern Cascade Mountains of Oregon. J. Bot. Res. Inst. Texas 1(1), 9 – 19.

Milne, R.I. 2006. Northern hemisphere plant disjunctions: A window Tertiary land bridges and climate change? Ann. Bot. 98, 465-472.

Milne, R.I., and Abbott, R.J. 2002. The origin and evolution of Tertiary relict floras. Advances in Botanical Research 38: 281–314.

Nylander, J.A.A. 2004. MrAIC.pl. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nylander J.A.A., Olsson U, Alstrom P, Sanmartin I. 2008. Accounting for phylogenetic uncertainty in biogeography: a Bayesian approach to Dispersal–Vicariance Analysis of the thrushes (Aves: Turdus). Syst. Bio. 57: 257– 268.

Rambaut, A. and Drummond, A.J. 2007. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer

Rehder, A. 1932. New species, varieties and combinations from the herbarium and the collections of the Arnold Arboretum. J. Arnold Arbor. 13, 337–41.

Richmond, G.M. and Fullerton, D.S. 1986. Quaternary glaciations in the United States of America. In: V. Sibrava, D.Q. Bowen and G.M. Richmond (eds.), Quaternary Glaciations in the Northern Hemisphere. Report of the International Geological Correlation Programme Project 24. Pergamon Press, Oxford, U.K. pp 3-10.

Ronquist F. 1996. DIVA, version 1.1. Computer program and manual available from Uppsala University, http://www.ebc. uu.se/systzoo/research/diva/diva.html.

Ronquist F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Bio. 46: 195–203.

Ronquist, F. 2001. DIVAversion 1.2. Computer program for MacOS and Win32. Evolutionary Biology Centre, Uppsala University. Available at http://www.ebc.uu.se/systzoo/research/diva/diva.html.

Ronquist, F., and Huelsenbeck, J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Small, J.K. 1903. Flora of the southeastern United States, being descriptions of the seed-plants, ferns and fern-allies growing naturally in North Carolina, South Carolina, Georgia, Florida, Tennessee, Alabama, Mississippi, Arkansas, Louisiana, and in Oklahoma and Texas east of the one hundredth meridian. Published by the author, New York, NY.

Southall, R.M. and Hardin, J.W. 1974. A taxonomic revision of *Kalmia* (Ericaceae). Journal of the Elisha Mitchell Scientific Society. 90, 1-23.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22(21), 2688-2690.

Stamatakis, A., Hoover, P., and Rougemont, J. 2008. A Rapid Bootstrap Algorithm for the RAxML Web-Servers, Systematic Biology, 75(5), 758-771.

Stevens, P.F. 1969. Taxonomic studies in the Ericaceae. Ph.D. Thesis. University of Edinburgh.

Stevens, P.F. 1971. A classification of the Ericaceae: Subfamilies and tribes. Bot. J. Linn. Soc. 64, 1-53.

Swofford, D.L. 2002. Paup* 4.0b10: Phylogenetic analysis using parsimony (* and other methods). Sunderland: Sinauer Associates.

Tiffney B.H. and Manchester, S.R. 2001. The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the northern hemisphere tertiary. Int. J. of Plant Sciences 162: S3-S17.

Vinogradova, V.M. 2001. Ericaceae. In: Flora of Russia v. 5. 512 pp.

Watson, L., Williams, W.T. and Lance, G.N. 1967. A mixed-data approach to angiosperm taxonomy: The classification of Ericales. Proc. Linn. Soc. London 178: 25–35.

Weakley, A.S. 2008. Flora of the Carolinas, Virginia, Georgia, northern Florida, and surrounding areas. The University of North Carolina Press. Working Draft of 7 April 2008.

Wen, J., Shi, S., Jansen, R.K.and Zimmer, E.A. 1998. Phylogeny and biogeography of Aralia sect. Aralia (Araliaceae). Amer. J. Bot. 85: 866–875.

Wu, Z.Y., Raven, P.H., and Hong, D.Y., eds. 2005. Flora of China. Vol. 14 (Apiaceae through Ericaceae). Science Press, Beijing, and Missouri Botanical Garden Press, St. Louis.

Xiang Q-Y, Soltis D.E., Soltis P.S. 1998. The eastern Asian and eastern and western North America floristic disjunc- tion: congruent phylogenetic patterns in seven diverse genera. Molecular Phylogenetics and Evolution 10: 178– 190.

Xiang**,** Q.-Y., Soltis, D.E.,. Soltis, P.S, Manchester, S.R., and Crawford, D.J. 2000. Timing the eastern Asian-eastern North American floristic disjunction: molecular clock corroborates paleontological data. Molecular Phylogenetics and Evolution 15, 462-472.

Yang, H.P. 1990. A new species of *Phyllodoce* (Ericaceae) from China. Acta Phytotax. Sin. 28(5), 403.

Yang, Z.H. and Rannala, B. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. Mol. Biol. Evol. 14, 717–724.

Yu Y, AJ Harris, and X He. S-DIVA (Statistical Dispersal-Vicariance Analysis): a tool for inferring biogeographic histories. In press.

Yu Y, AJ Harris, and X He. (2009). A rough guide to S-DIVA. Available online at http://mnh.scu.edu.cn/S-DIVA. Accessed [3 January 2010].

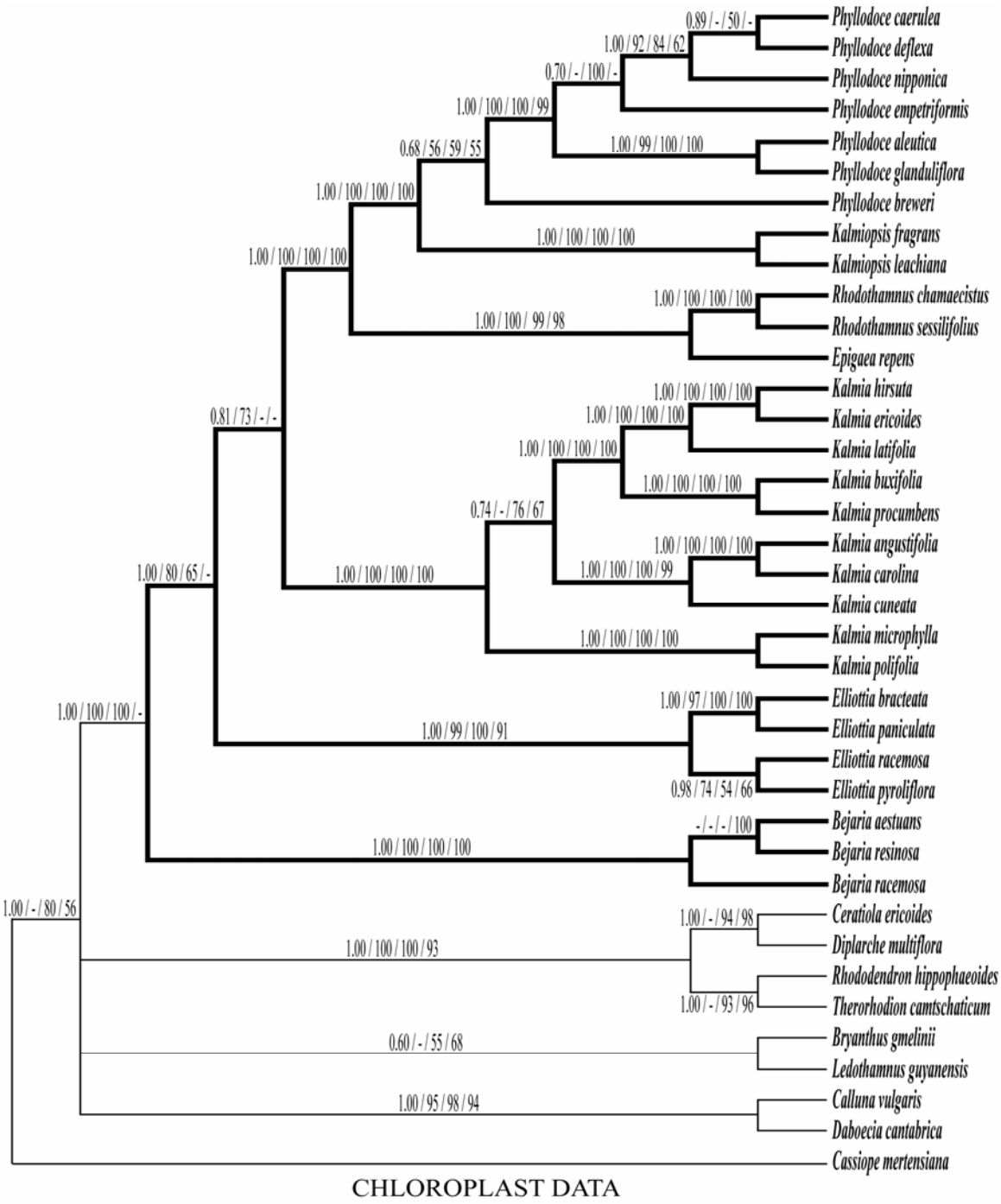


FIGURE 3.1. Total combined chloroplast data (*matK*, *ndhF*, *rbcL* and *trnS-G* spacer) analysis of the Phyllodoceae. Support values are to the left of nodes in the format (Bayesian posterior probability / RAxML bootstrap / PhyML bootstrap / MP bootstrap). Support derived from 20 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Tree length = 3400, CI = 0.5268, RI = 0.6546. Bold clade indicates the Phyllodoceae clade sensu Gillespie and Kron (2010).

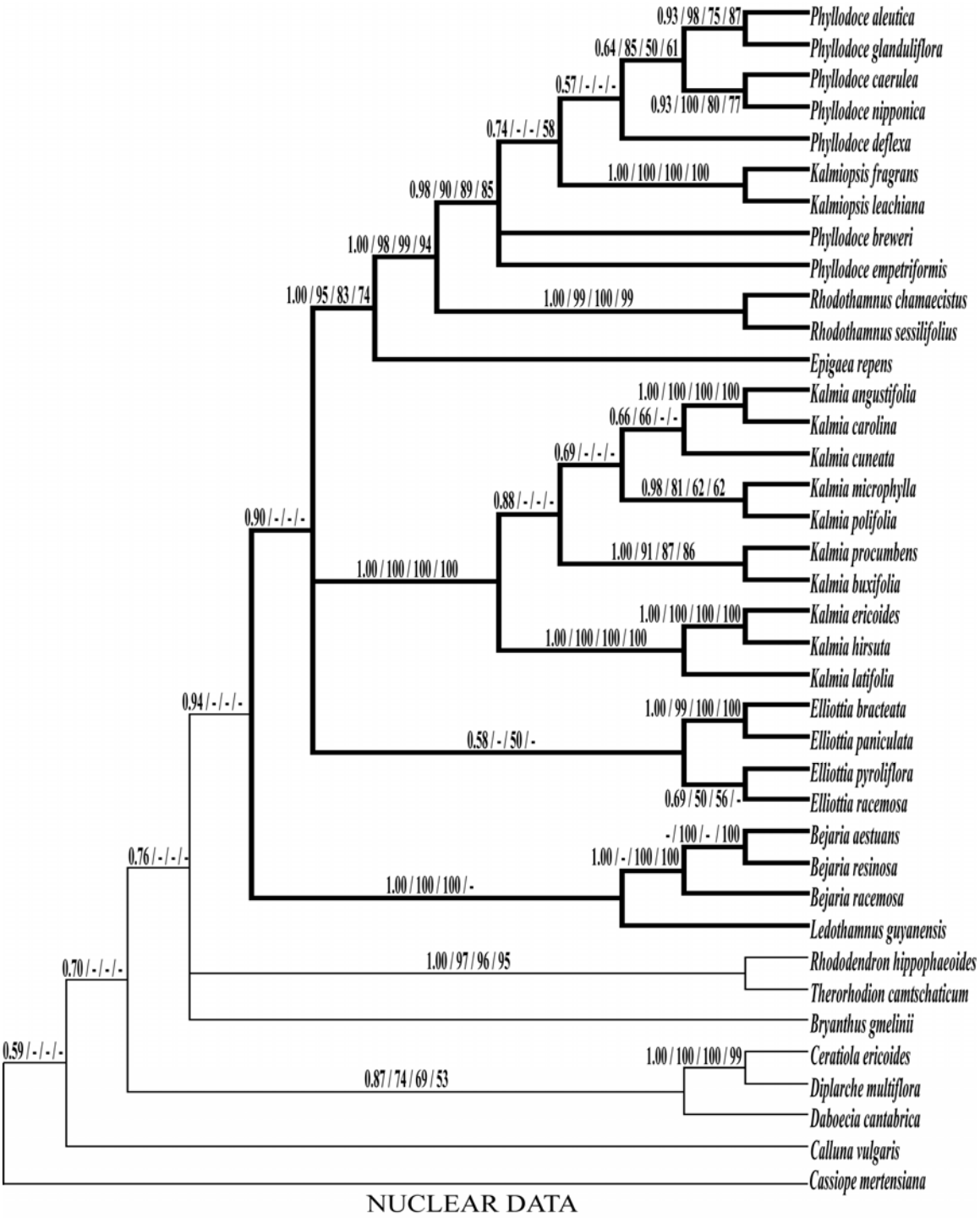


FIGURE 3.2. Total combined nuclear data (nrITS and *waxy*) analysis of the Phyllodoceae. Support values are to the left of nodes in the format (Bayesian posterior probability / RAxML bootstrap / PhyML bootstrap / MP bootstrap). Support derived from 20 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Tree length = 858, CI = 0.5431, RI = 0.6722. Bold clade indicates the Phyllodoceae, sensu Gillespie and Kron (2010).

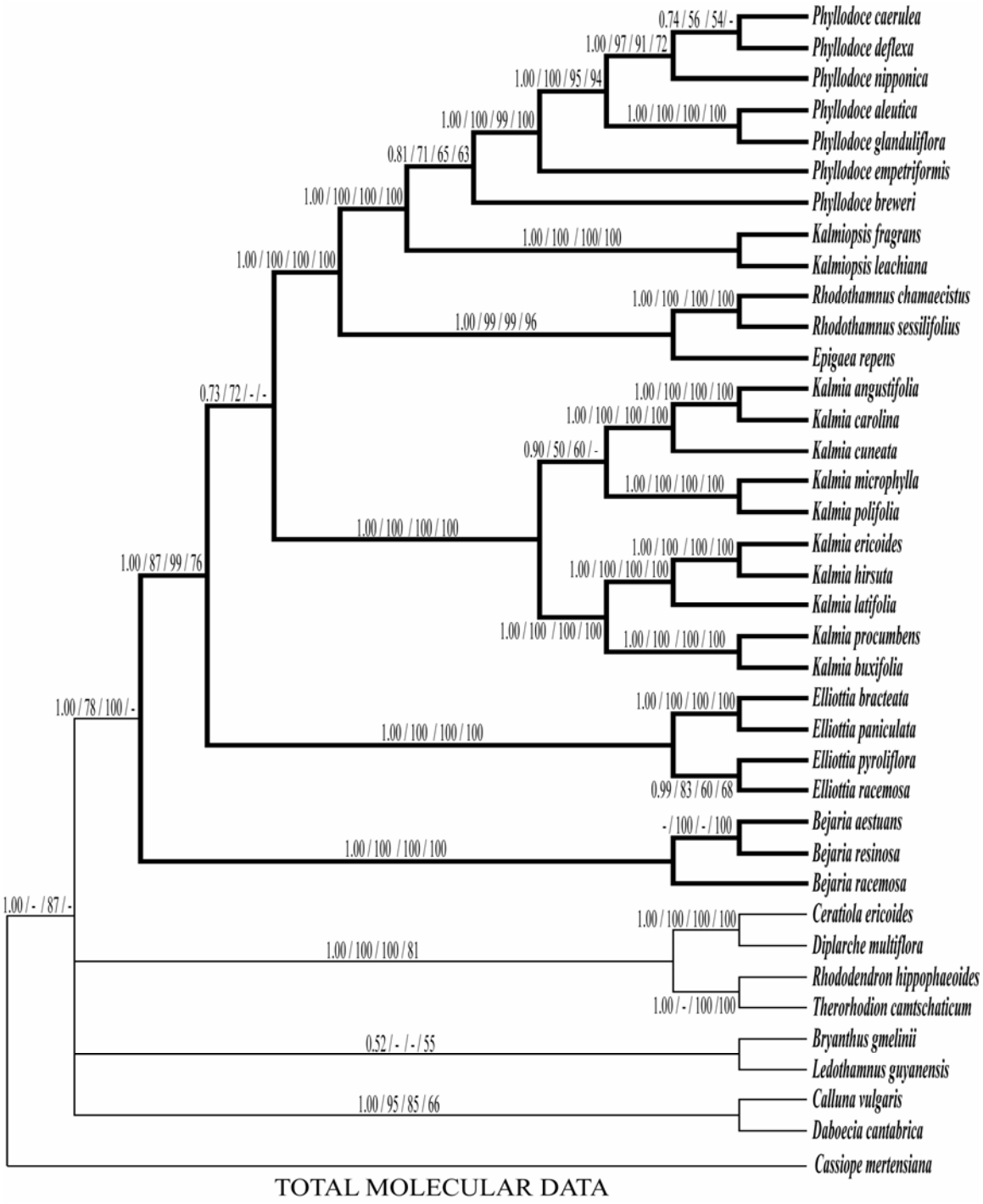


FIGURE 3.3. Total combined molecular data (nrITS, *waxy*, *matK*, *ndhF*, *rbcL* and *trnS-G* spacer) analysis of the Phyllodoceae. Support values are to the left of nodes in the format (Bayesian posterior probability / RAxML bootstrap / PhyML bootstrap / MP bootstrap). Support derived from 20 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Tree length = 4287, CI = 0.5265, RI = 0.6533. Bold clade indicates the Phyllodoceae sensu Gillespie and Kron (2010).

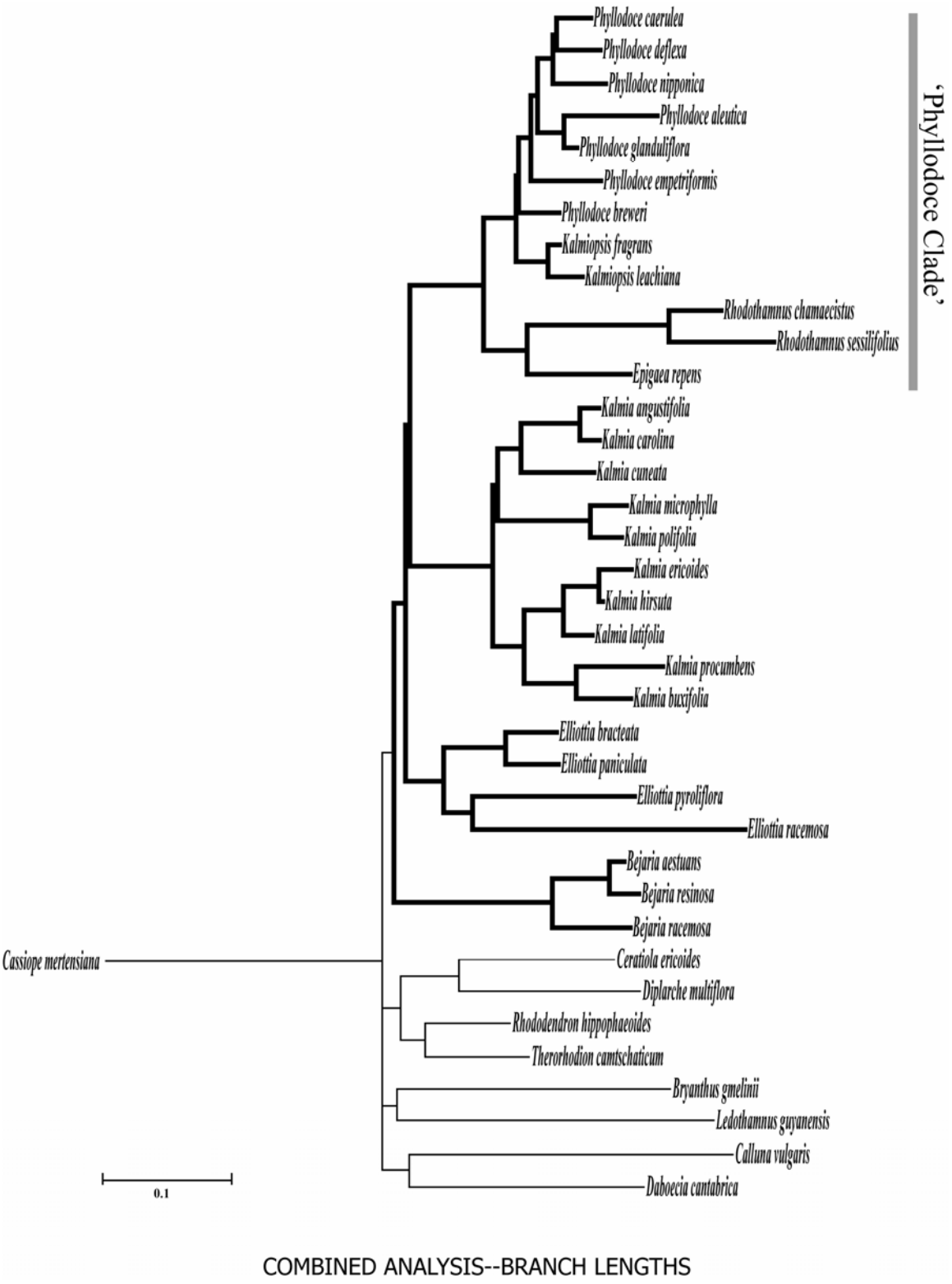


FIGURE 3.4. Phylogram of Phyllodoceae showing branch lengths resulting from Bayesian analysis of total combined molecular data.

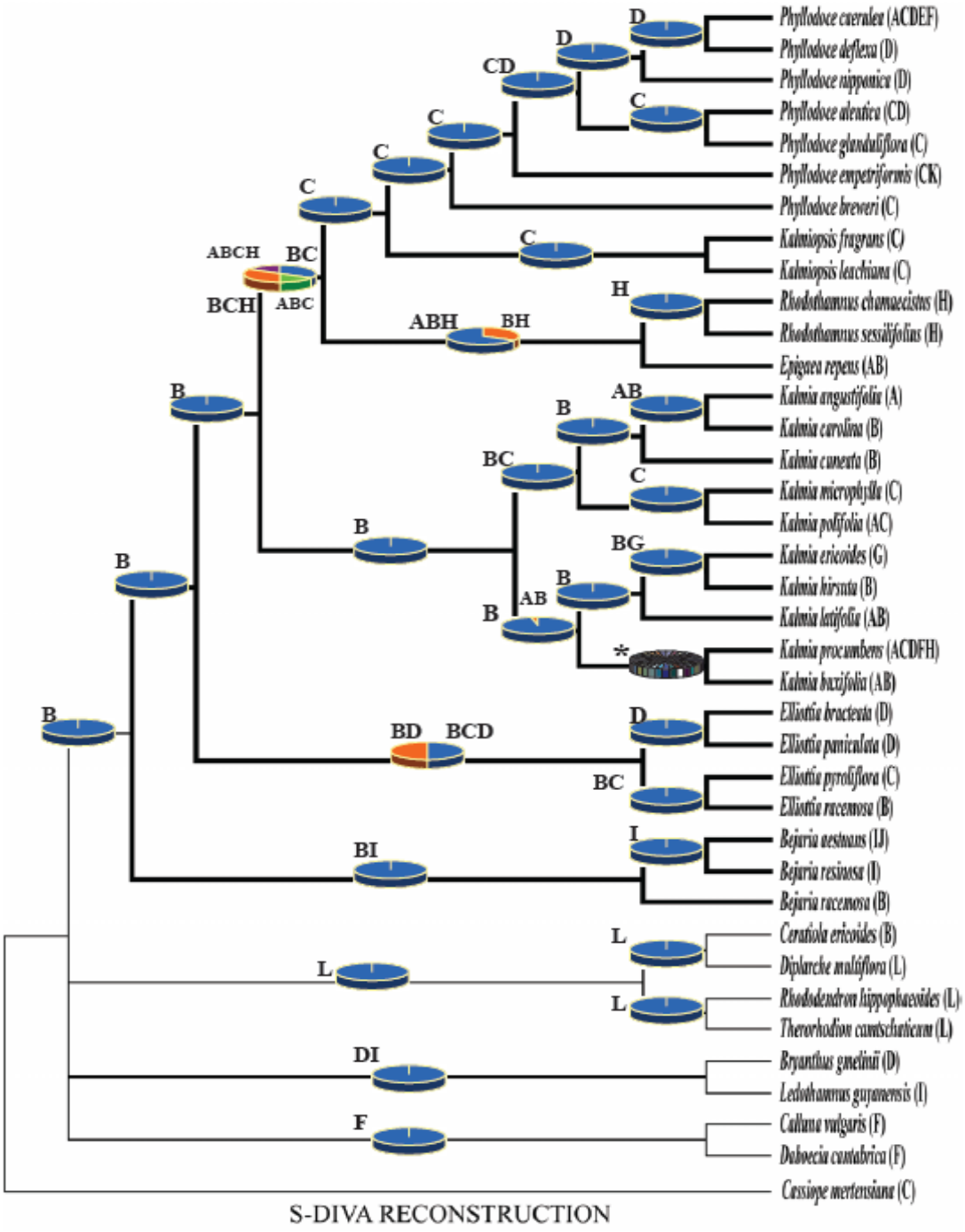


FIGURE 3.5. Reconstruction of ancestral areas of the Phyllodoceae in S-DIVA. Analysis of total combined molecular data included 10,000 Bayesian trees. Solid blue graph indicates a single likely ancestral area. A=Northeastern North America, B=Southeastern North America, C=Northwestern North America, D=Northeastern Asia, E=Sibera/North-central Eurasia, F= Northern Europe, G=Cuba, H= Southern Europe/European Alps/Caucasus, I=Northern South America, J=Central America, K=Southwestern North America, L=Southeatern Asia. * indicates that all possible combinations of areas A, B, C, D, F and H are equally likely.

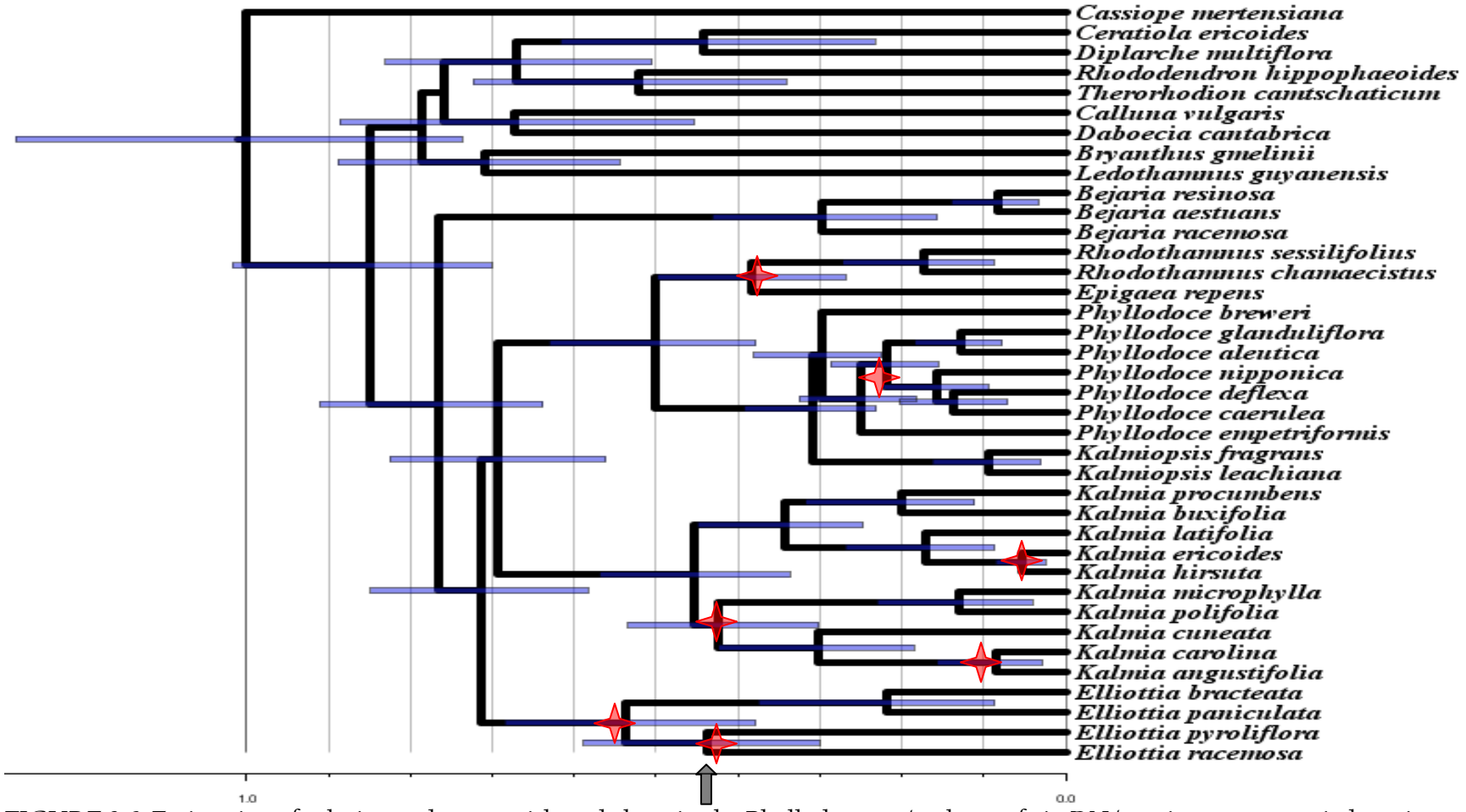


FIGURE 3.6. Estimation of relative node ages with node bars in the Phyllodoceae. Analyses of six DNA regions were carried out in BEAST for 20 million MCMC generations under an uncorrelated relaxed lognormal clock model. Scale is arbitrarily set from 0.0 to 1.0 for mean node ages. Node bars indicate 95% confidence intervals on mean relative ages. Stars (✦) indicate variance events reconstructed using S-DIVA. Arrow (⇨) indicates two vicariance events occurring sufficiently close together in time for further hypothesis testing.

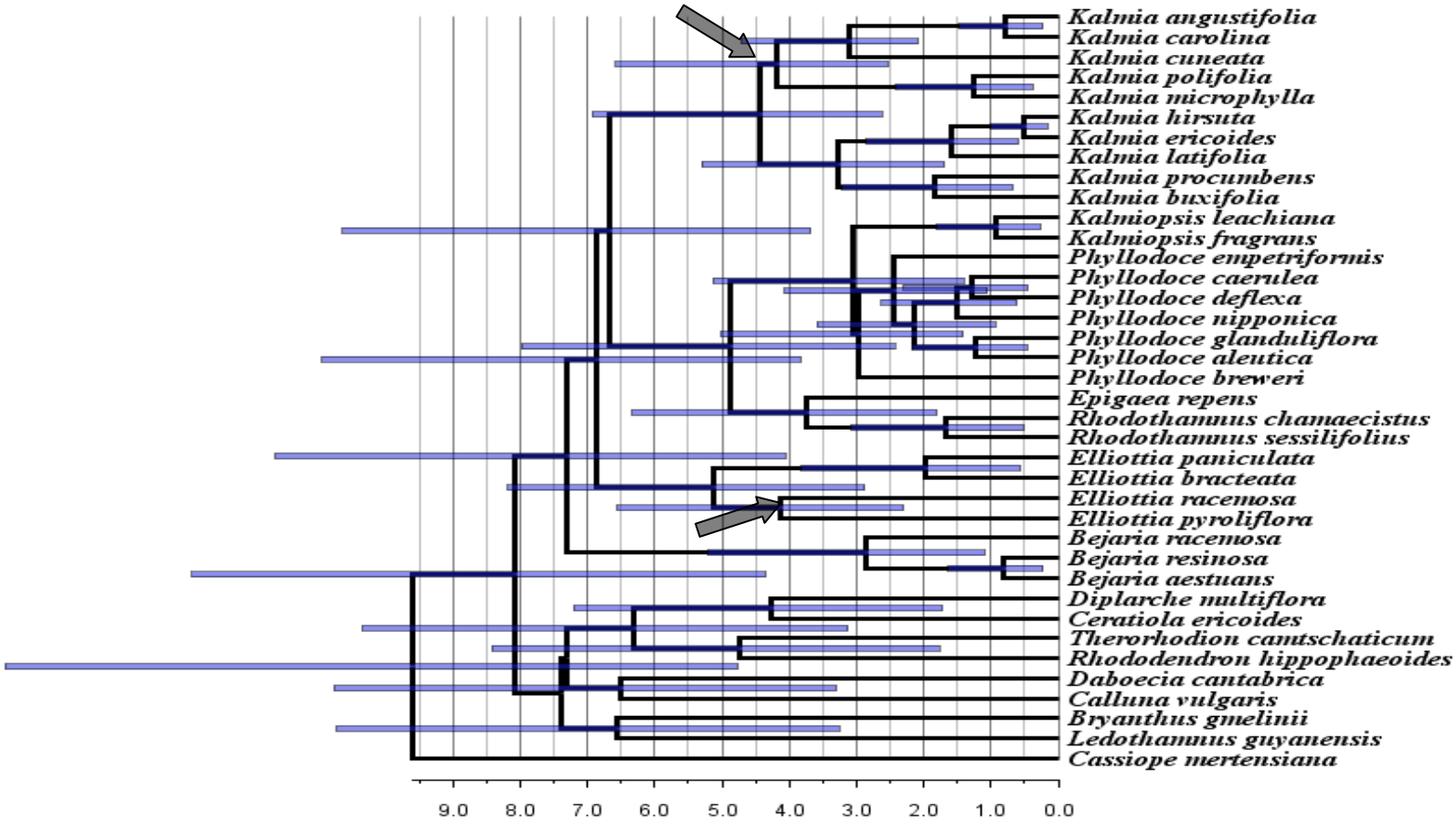


FIGURE 3.7A. Estimation of absolute node ages with node error bars in the Phyllodoceae. Analyses of six DNA regions were carried out in BEAST for 20 million MCMC generations under an uncorrelated relaxed lognormal clock model. Scale indicates absolute dates in millions of years before present. Node bars reflect 95% confidence intervals. Arrows ( ) indicate nodes constrained to 2mya with a lognormal probability distribution.

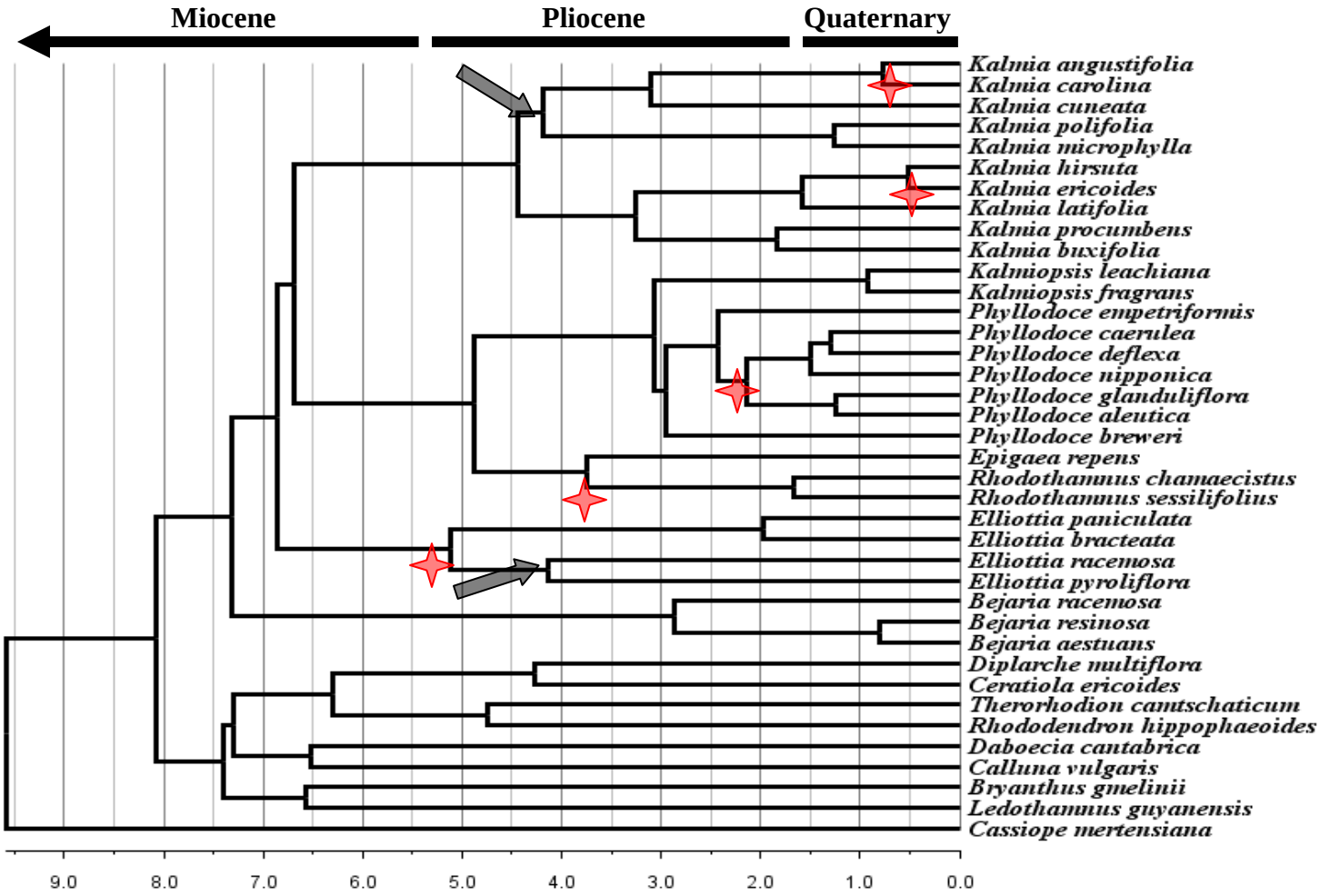


FIGURE 3.7B. Estimation of absolute mean node ages in the Phyllodoceae with node bars removed. Analyses of six DNA regions were carried out in BEAST for 20 million MCMC generations under an uncorrelated relaxed lognormal clock model. Scale indicates absolute dates in millions of years before present. Arrows (⇨) indicate nodes constrained to 2mya with a lognormal probability distribution. Stars (✦) indicate vicariance events for which absolute dates were estimated.

TABLE 3.1. Gene region, evolutionary model, aligned length, number of potentially parsimony informative characters (and %), and number of missing taxa (and %) in a phylogenetic study of the Phyllodoceae.

Gene Region	**Evolutionary Model (AICc)**	**Aligned Length (bp)**	**# Missing Taxa (%)**
rbcL	GTR+I+Γ	1251	2 (5.1%)
matK	GTR+ Γ	1528	3 (7.7%)
ndhF	GTR+ Γ	1965	2 (5.1%)
trnS-G-G	GTR+ Γ	1508	3 (7.7%)
nrITS	GTR+ Γ	665	0
waxy	HKY+ Γ	592	5 (12.8%)

Appendix 3.1. Taxa, voucher information and Genbank accesions for a phylogenetic study of the Phyllodoceae. Sequences newly generated for this study are in bold. **XXXXX** indicates sequences included in the analyses which have pending Genbank accession number assignment.

Species	**Voucher**	**nrITS**	***waxy***	***matK***	***trnS-G-G***	***ndhF***	***rbcL***
Phyllodoceae							
Phyllodoce aleutica (Spreng.) A. Heller	1932-0502, RBGE	**HM182008**	**HM182007**	**HM182019**	**HM182041**	**HM182028**	**HM182033**
Phyllodoce breweri (A. Gray) Maxim.	Gillespie 06-019, WFU	**HM182009**	**HM182001**	**HM182020**	**HM182042**	**HM182029**	**HM182034**
Phyllodoce caerulea (L.) Bab.	1940-1013, RBGE	GU176630	DQ000604	U61318	GU176700	GU176735	AF419829
Phyllodoce deflexa Ching ex H.P. Yang	Jian Hung Yuan s.n., WFU	**HM182010**			**HM182043**	**HM182030**	**HM182035**
Phyllodoce empetriformis D. Don	Chase 871, K	U48607	DQ000605	U61333	GU176701	GU176736	U49291
Phyllodoce glanduliflora (Hook.) Coville	1968-0179, RBGE	**HM182011**	**HM182002**	**HM182017**	**HM182044**	**HM182025**	**HM182036**
Phyllodoce nipponica Makino	Anderberg 1756-77, S	U48606	DQ000606	U61325	GU176702	GU176737	U49292
Kalmiopsis leachiana (Hend.) Rehd.	Denton s.n., WFU	U48608	DQ000603	U61323	GU176699	GU176734	U49290
Kalmiopsis fragrans Meinke and Kaye	Meinke s.n., WFU	**HM182012**	**HM182003**	**HM182021**	**HM182045**	**HM182031**	**HM182037**
Epigaea asiatica Maxim.	Takahashi & Murata 2943 HUH	**HM182013**			**HM182048**	**HM182027**	
Epigaea repens L.	Kron 162, WFU	U48611	GU176659	U61319	GU176693	GU176728	U49284
Rhodothamnus chamaecistus Reichb.	Chase 877, K	U48605	DQ000607	U61321	GU176703	GU176738	U49287
Rhodothamnus sessilifolius P.H. Davis	Terzioğlu s.n., WFU	**HM182014**			**HM182049**		
Kalmia angustifolia L.	Kron 1895, WFU	U48599	DQ000602	U61348	GU176694	GU176729	AF419826
Kalmia buxifolia (Berg.) Gift, Kron & Stevens	Gift s.n., HUH	U48581	GU176660	U61347	GU176695	GU176730	L12619
Kalmia carolina Small	Gillespie 07-004, WFU	**HM182015**	**HM182005**	**HM182018**	**HM182046**	**HM182032**	**HM182038**
Kalmia cuneata Michx.	Gillespie 07-003, WFU	KCU48603	**HM182005**	**HM182022**	**HM182047**	**HM182026**	**HM182039**
Kalmia ericoides C. Wright	Abbott 18854, FLAS	**HM182016**	**HM182006**	**HM182023**		**HM182024**	**HM182040**
Kalmia hirsuta Walt.	Judd s.n., FLAS	U48601	GU176661	GU176673	GU176696	GU176731	GU176645
Kalmia latifolia L.	Kron 3020, WFU	U48600	GU176662	GU176674	GU176697	GU176732	U49294
Kalmis microphylla A. Heller	Gillespie 06-020, WFU	**XXXXX**	**XXXXX**	**XXXXX**	**XXXXX**	**XXXXX**	**XXXXX**
Kalmia polifolia Wang.	Anderberg 325-89, S	U48597	GU176663	**XXXXX**	GU176698	GU176733	U49289
Kalmia procumbens Gift, Kron & Stevens	Gift, s.n., WFU	U48610	**XXXXX**	U61352	**XXXXX**	**XXXXX**	U49288
Elliottia bracteata Benth. & Hook. f	Chase 866, K	U48609	DQ000600	U61339	GU176690	GU176725	U49285
Elliottia paniculata Benth. & Hook. f	96RD00974FRBTU11	GU176628		GU176671	GU176691		GU176643
Elliottia pyroliflora (Bong.) Brim &	1934-009, RBGE	GU176629	GU176658	U61320		GU176726	GU176644

Stev.							
Elliottia racemosa Muhl. ex Elliott	1967-2632, RBGE	U48582		GU176672	GU176692	GU176727	L12615
Bejaria aestuans Mutis ex L. f	Luteyn 14175, NYBG	AF404817	DQ000589	GU176669	GU176678	DQ002362	GU176638
Bejaria racemosa Vent.	Kron 2070, NCU	U48604	DQ000594	U61327	GU176679	DQ002367	L12600
Bejaria resinosa Mutis ex L. f	Luteyn 14133, NY	GU176622	DQ000595	AF440412	GU176680	DQ002368	GU176639
Rhodoreae							
Rhododendron hippophaeoides Balf. f & Sm.	1932-1022, RBGE	GU176634	GU176667	U61353	GU176707	GU176742	L01949
Therorhodion camtschaticum (Pall.) Sm.	73/054, RSF	GU176637	DQ000608	U61322	GU176710	DQ002382	AF419834
Empetreae							
Ceratiola ericoides Michx.	Kron 2069, WFU	AF519552	DQ000599	U61334	GU176684	GU176717	L12605
Diplarche multiflora Hook. f and Thomas	Suzuki et al. 8820561, HUH	GU176631	GU176664	AF440418		GU176739	AF419821
Ericeae							
Calluna vulgaris (L.) Hull 'Rannola'	1972-1443, RBGE	GU176627	GU176655	U61326	GU176687	GU176722	L12601
Daboecia cantabrica (Huds.) C. Koch	1975-1770, RBGE	AY520786	GU176656	U61349	GU176688	GU176723	L12611
Bryantheae							
Bryanthus gmelinii D. Don	Stevens s.n., WFU	U48612	GU176650	AF440413	GU176681	GU176715	AF419816
Ledothamnus guyanensis Meisn.	Picon &Williams 2910, WFU	GU176623	GU176651	AF440419	GU176682	GU176716	AF419827
Cassiopoideae							
Cassiope mertensiana G. Don	Anderberg 75-83, S	AF419798	DQ000598	U61346	GU176711	GU176745	L12603

CHAPTER 4

PRELIMINARY PHYLOGENETIC ANALYSIS OF *CASSIOPE* AND EVALUATION OF THE EVOLUTION OF SOME MORPHOLOGICAL AND WOOD ANATOMY CHARACTERS

ABSTRACT

The genus *Cassiope* is comprised of approximately 15 species of dwarf alpine shrubs that exhibit a variety of leaf morphologies, despite having very similar flowers and habits. This clade has been previously shown to be strongly supported as sister to the subfamily Ericoideae, but relationships within the genus have never been the subject of a phylogenetic analysis. The goal of this study was to use three relatively 'fast' molecular markers from the chloroplast and nuclear genomes to generate a preliminary phylogenetic framework for future studies. The analyses generated a strongly supported Bayesian phylogeny, but other analyses were only weakly supported, and some instances of conflict among genomes were apparent. Future studies may find it helpful to add additional molecular markers that evolve relatively slowly. Despite this preliminary evidence, some evolutionary trends are apparent. Categorically varying characters including leaf form and stomatal distribution appear to be homoplasious but correlated. Continuously varying characters such as elevation and aspects of wood anatomy show compelling phylogenetic trends, which may be related to the fact they occur in alpine and subalpine habitats.

INTRODUCTION

Cassiope D. Don is a group of 15-17 species of dwarf alpine shrubs with small triangular, imbricate, decussate leaves. Most have an obvious abaxial groove or sulcus on the leaves, although a few have more or less planar leaves or thickened adaxially concave leaves. Flowers are generally campanulate and have pinkish to white corollas, are solitary and axillary, and emerge relatively close to the tips of branches. Most species have reddish-green or deep purple sepals. They occupy alpine meadows, lichen tundra, rocky *Rhododendron* thickets and similar areas. The distribution of *Cassiope* is circumboreal with a disjunction in the Himalayas. *Cassiope tetragona* is the only species that is known to be completely circumboreal in distribution, although the distributions of many *Cassiope* are rather poorly documented (Good 1926; Fang and Stevens, 2005; Wallace, 2009)

Cassiope is defined by several character states that are homoplasious in the broader Ericaceae. These include *Calluna*-type pith, the absence of vegetative bud scales, decussate leaves, leaf midrib fibers absent, single-flowered, axillary inflorescences, ericoid leaves, and anther appendages. A true morphological synapomorphy for the *Cassiope* clade appears to be the presence of a bisporic embryo sac (other Ericaceae have a *Polygonum*-type embryo sac) (Good 1926; Stevens 1969, 1971; Kron et al. 2002). *Cassiope* has been allied with the Gaultherieae, Andromedeae, Enkiantheae, and Oxydendreae (Hooker 1867; Drude 1897), as well as Calluneae (Watson 1964; Watson et al. 1967; Hagerup 1953) or in its own tribe, Cassiopeae (Stevens 1969, 1971). Most recently, Kron et al. (2002) and Gillespie and Kron (2010) demonstrated that *Cassiope* is sister to the Ericoideae clade.

Most character states shared between *Cassiope* and the Ericoideae are either highly variable within this clade, or are reasonably uniform throughout the Ericaceae. Therefore, there are few putative morphological synapomorphies for *Cassiope* + Ericoideae. Revolute leaves in bud and channeled leaves (in some taxa) are shared by both taxa, and are reasonably uncommon in the remaining Ericaceae even if not unique. However, longitudinally oriented stomata are by far the predominant orientation found in the Enkianthoideae, Monotropoideae, Styphelioideae and Vaccinioideae; one small genus in the Styphelioideae, *Lysinema*, has transversely oriented stomata (Stevens 1969; Watson 1962; Butterfass, 1987). Transversely oriented stomata are found only in *Cassiope* and the Ericoideae (Stevens 1969), and although the character state is not uniform across the clade, it may represent a synapomorphy for Ericoideae + *Cassiope*. Clearly, more work needs to be done to explore the distribution of this trait.

Cassiope is morphologically highly variable in several immediately noticeable characters. There are four basic leaf forms that exhibit morphological gaps. One species included in this study (*C. palpebrata*) has a planar/flat leaf with slightly thickened edges. Three species (*C. mertensiana, C. lycopodioides,* and *C. myosuroides*) have a leaf that is thickened and adaxially slightly concave, giving the appearance of being essentially flat on the 'top' or adaxial surface and rounded or convex on the 'bottom' (abaxial) surface. Most other species have a complex ericoid leaf architecture, where the adaxial surface is essentially flat or very slightly concave, but the abaxial surface has a prominent groove or sulcus that extends in most cases the entire length of the leaf. One species (*C. redowskii*) has a leaf that essentially forms a tube near the base. (Good 1926; Stevens 1969, 1971; Wallace 1986), and this leaf form is unlike any other in the genus.

The overall arrangement of hairs on the leaves is also highly variable and morphological gaps are not immediately clear. Most *Cassiope* species have uniseriate hairs, usually somewhat sparsely distributed on both surfaces of the leaf and along the leaf margins. Multiseriate hairs are common, and the arrangement and density of these hairs vary among species. A small number of *Cassiope* species have no multiseriate hairs (or nearly so) on the leaf surface or margin. Several have very stiff, straight multiseriate hairs on the margins that are as long as the leaf is wide, and these may be rather uniformly spread out around the leaf margin, or in loosely associated groups. A few species have multiseriate hairs that are distributed in small, regularly spaced groups, but appear somewhat curled and consequently appear as small 'balls' distributed around the leaf margins (Good, 1926; Stevens 1969). Still others appear to be multiseriate at their bases, but form a semi-transparent, stiff membrane-like margin shortly after emerging from the leaf margin. No literature exists that addresses whether these basic multiseriate units are derived from the same underlying developmental processes, whether the exact appearance and distribution of those units are under environmental or other influences, or whether the apparent different hair types represent truly different evolutionary trajectories. The apparent absence of morphological gaps makes them analytically challenging, but an understanding of evolutionary relationships may shed some light on general trends in leaf form, if not explicit hypotheses of hair type evolution.

Wallace (1986) documented significant diversity among 23 wood anatomy characters within *Cassiope*. Some of these characters related to aspects of vessel elements, tracheids, the nature of perforation plates, and the type of lateral wall pitting on vessels, among others. Wallace's study included all but two described *Cassiope* species

and therefore this dataset is a valuable resource in an otherwise rarely studied genus, providing an opportunity to evaluate wood anatomy characteristics in a phylogenetic context. This information can then be examined in terms of anatomical adaptations to an alpine environment.

Harrimanella Coville has been either included along with *Cassiope* in the tribe Cassiopeae, or included within *Cassiope* itself. An inferred close relationship to *Cassiope* has been based upon several similarities that are homoplasious, but relatively uncommon, within the Ericaceae, such as having revolute leaves in bud, ericoid leaves, more or less pendulous flowers, presence of stamen appendages, loculicidal capsules, and inflorescences of solitary flowers (Stevens 1969, 1971; Kron et al. 2002). Additionally, the overall appearance of the two genera is quite similar; both have whitish, campanulate, pendulous flowers, reddish-green calyces, and a short and mat-forming habit. Other characters do not support a close relationship between *Cassiope* and *Harrimanella*, including chromosome numbers (in *Cassiope*, X=13; in *Harrimanella*, X=18), flower position (axillary in *Cassiope*; terminal in *Harrimanella*), stomata distribution (*Cassiope* has stomata on the abaxial calyx, while *Harrimanella* has no floral stomata), pith type (*Cassiope* has *Calluna*-type pith and *Harrimanella* has homogenous pith), leaf characters (*Harrimanella* leaves are petiolate, whorled and needlelike, while *Cassiope* leaves are usually epetiolate, decussate and channeled abaxially). Many *Cassiope* have a complex indumentum comprised of multiseriate branched hairs, but *Harrimanella* has no indumentum or unicellular hairs only. *Cassiope* has axillary flowers that emerge near the tips of branches, whereas *Harrimanella* has truly terminal flowers (Kron et al., 2002; Stevens, 1969; 1970).

Given such a list of anatomical, cytological and morphological differences between *Harrimanella* and *Cassiope*, it was not wholly unexpected that *H. hypnoides* was resolved outside the *Cassiope* clade by Kron et al. (2002), but rather, it was resolved as sister to a clade comprised of Vaccinioideae + Styphelioideae. However, *H. stelleriana*, the only other *Harrimanella* species (and the type), was not included in the study. Evidence gathered by Hara (1958) suggested that leaf development of *H. stelleriana* and *Cassiope lycopodioides* might be similar, given that the only difference he discussed between them was that *C. lycopodioides* has a scarious leaf margin. Because questions about the morphological similarities of *Cassiope* and *Harrimanella* remain, it is important to include both species of *Harrimanella* in order to address their phylogenetic position convincingly.

The current study includes a preliminary phylogenetic analyses of 13 species within the genus *Cassiope* using three molecular markers. As part of this study, *Harrimanella hypnoides*, which has been resolved outside *Cassiope* in past studies, and *H. stelleriana*, which has not been included in any past study, are also included in order to better assess their phylogenetic position within the Ericaceae clade. The evolution of several morphological and anatomical characters is evaluated within the context of the current phylogenetic hypothesis.

METHODS

DNA was isolated from silica-dried or herbarium plant material using the Qiagen Plant Mini Kit (Qiagen, Valencia CA, USA) with modifications following Drábková et

al. (2002). Methods pertaining to DNA accessions generated prior to this study are described in Gillespie and Kron (2010).

Thirteen *Cassiope* species were included (Appendix 4.1). Two northwestern North American species (*C. mertensiana* and *C. lycopodioides*) were included. Eight Chinese taxa (*C. abbreviata*, *C. fastigiata*, *C. palpebrata*, *C. myosuroides*, *C. selaginoides*, *C. wardii*, *C. nana* and *C. pectinata*) were included. *Cassiope tetragona*, which is circumboreally distributed, was included, as well as *C. ericoides* from Kamtchatka. Both representatives of *Harrimanella* (*H. hypnoides* and *H. stelleriana*) were included. Three Chinese taxa (*C. membranifolia, C. argyrotricha*, and *C. fujianensis*) and *C. redowskii* from Siberia were not included because material was not available for DNA extraction.

Eleven additional taxa were included as outgroups and to accomplish a suitable taxon sampling to allow determination of the phylogenetic position of *Harrimanella*. Four members of the Ericoideae, the clade known to be sister to *Cassiope*, were included (*Daboecia cantabrica*, *Elliottia bracteata*, *Phyllodoce caerulea* and *Rhododendron hippophaeoides*). Six representatives of the Vaccinioideae were included (*Andromeda polifolia*, *Oxydendron arboreum Lyonia ligustrina* and *Vaccinium tenellum*, *V. uliginosum*, and *V. myrtillus*). *Arctostaphylos uva-ursi* and *Enkianthus campanulatus* were included to represent the Arbutoideae and Enkianthoideae, respectively.

Three DNA regions (one chloroplast and two nuclear) were included in this study. These include chloroplast regions $trnS^{GCU}$-$trnG^{UCC}$-$trnG^{UCC}$ intergenic spacer (*trnS-G-G*), and nuclear regions *waxy*/GBSS-1 (*waxy* exons 9-11) and nuclear ribosomal Internal Transcribed Spacer (nrITS). PCR protocols and primer sequences are as reported in

Gillespie and Kron (2010). Single gene, combined nuclear, combined chloroplast and total data matrices are deposited in TreeBase (www.treeBASE.org), under accession number 10491. All newly generated sequences are deposited in Genbank (www.ncbi.nlm.nih.gov) under numbers HM182050 - HM182100 and are listed in Appendix 4.1.

Amplified fragments were cleaned using Qiagen™ QIAquick Gel Isolation Kit (Qiagen, Valencia CA, USA). DNA was sequenced on an ABI 377 automated sequencer at the DNA Sequencing and Gene Analysis Laboratory at the Wake Forest University School of Medicine (Winston-Salem, NC) or at Nevada Genomics Center (Reno, NV). Sequences were minimally edited and manually aligned in Geneious 4.7.4 (Drummond et al., 2006).

Nuclear regions were cloned using the TOPO TA kit (Invitrogen, Carlsbad CA, USA) for representative taxa within the Ericaceae (see Gillespie and Kron, 2010). Five clones were sequenced for each of these taxa and no evidence of multiple copies of nuclear genes was found within *Cassiope* or *Harrimanella* generally. However, the nrITS sequence of *Harrimanella hypnoides* was virtually identical to those of *Cassiope*. Examination of the sequence trace did not reveal any evidence that multiple amplicons had been sequenced. Therefore PCR reactions using three different annealing temperatures (48°C, 50°C and 52°C) were conducted in order to attempt to amplify rare copies. These amplicons were pooled and cloned using the TOPO TA cloning kit. Ten clones were sequenced. All ten were identical to the original un-cloned sequence. Inclusion of this sequence appeared to introduce a large amount of phylogenetic 'noise' into the nuclear and total combined analyses, and therefore the *H. hypnoides* nrITS

sequence was removed from the final analyses. This issue is explored in the Discussion, but it is clear that further examination of nrITS should be pursued in order to clarify the evolutionary dynamics of this sequence in both *Cassiope* and *Harrimanella*.

Maximum Parsimony (MP) analyses were carried out using PAUP*4.0b10, (Swofford, 2002) with the following options: Parsimony-informative characters were unordered and equally weighted, gaps were treated as missing data, searches were heuristic with TBR branch swapping and 1000 random stepwise addition replicates. Relative clade support was assessed using bootstrap analysis (Felsenstein, 1985; Felsenstein, 1988) with the full bootstrap option in PAUP* (10,000 replicates).

Maximum Likelihood analyses were carried out using the RaxML online server (Stamatakis, 2006, Stamatakis et al., 2008) on the CIPRES web portal at San Diego Supercomputing Center, as well as using the PhyML online server (Guindon and Gascuel, 2003; Guindon et al., 2005) on the ATCG Montpellier bioinformatics platform. For RAxML, total molecular evidence, chloroplast, and nuclear datasets were run with genes as separate partitions. Rapid bootstrap analysis including 100 replicates was run to determine node support. For PhyML, all three datasets (total, chloroplast and nuclear) were each run under a single GTR model (i.e. not partitioned by gene), as determined by the AICc criterion (Aikake, 1974) in MrAIC.pl 1.4.3 (Nylander, 2004). AICc was chosen as the criterion for model selection because it is not hierarchical in nature and also corrects for small sample sizes (approximately 40 and below). Bootstrap analysis (100 replicates) was conducted to determine node support. Both Maximum Likelihood strategies were used although they both search for the tree with the best likelihood score,

because they use fundamentally different bootstrapping strategies and are therefore essentially different approaches.

For Bayesian analyses, the data were partitioned by DNA region and evolutionary models were chosen using the AICc criterion in MrAIC.pl 1.4.3 (Nylander, 2004) (Table 4.1). Bayesian MCMC analyses (Yang and Rannala, 1997) as implemented in MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) consisted of 20 million generations with a burn-in of 25%. Clade support is reported as posterior probabilities.

The Incongruence Length Difference test (Farris et al., 1995), implemented as the partition homogeneity test in PAUP*4.0b10 (Swofford, 2002), was conducted to determine whether nuclear and chloroplast data recovered sufficiently congruent phylogenies to be combined into a total evidence analysis. A heuristic search with 100 replicates was performed using TBR swapping and simple, stepwise addition of taxa.

Ancestral character states were determined using the combined DNA phylogenies in Mesquite 2.71 (Maddison and Maddison, 2009). Most categorical characters (specifically leaf type, hair form and leaf stomatal distribution) were coded from the literature (Good, 1926; Stevens, 1969, 1971; Flora of China) and from observations of herbarium material. Lowest reported elevation, highest reported elevation, and the mean reported elevation for each species were compiled from the literature (Wallace, 2009; Fang and Stevens, 2005). Continuously varying wood anatomy characters, mostly documented by Wallace (1986—from Table 2 in his publication, used with permission) were also mapped onto the combined molecular phylogeny. For the two taxa not included in the study of Wallace (1986), those values were extrapolated in Mesquite based upon

the values of the nearest relatives of those taxa. Those characters appearing to exhibit a moderate to strong phylogenetic pattern are shown in the results, Figures 4.4 – 4.18. Characters appearing to exhibit a weak pattern or no apparent phylogenetic pattern are not shown.

RESULTS

Chloroplast Data (Fig. 4.1)

Analyses of trnS-G-G spacer data strongly support the monophyly of *Cassiope* (1.00 / 75 / 100 / 100) to the exclusion of *Harrimanella*. The two *Harrimanella* are supported as a clade (1.00 / 93 / 92 / 93) which is resolved in a clade with representatives of the Vaccinioideae (1.00 / 82 / 100 / 100).

Within *Cassiope*, a clade comprised of *C. myosuroides*, *C. palpebrata*, *C. nana* and *C. pectinata* is also strongly supported (1.00 / 97 / 97 / 88). Within this clade, *C. myosuroides* is sister to *C. palpebrata* (1.00 / 99 / 99 / 99) and *C. nana* is sister to *C. pectinata* (1.00 / 100 / 100 / 100). A large clade comprised of all other *Cassiope* was resolved with good support from RAxML and MP, but poor support from Bayesian and PhyML (0.51 / 79 / 66 / 86). *Cassiope tetragona* is sister to the rest of this clade. Within this large clade, *C. ericoides* is strongly supported as sister to *C. lycopodioides* (1.00 / 100 / 97 / 98). This clade is strongly supported as sister to *C. dendrotricha* (1.00 / 100 / 100 / 99). These three taxa are supported by all analyses except RAxML as sister to *C. wardii* (1.00 / - / 93 / 95). Several more taxa are resolved as successive sister taxa to this clade with support from some, but usually not all, analyses. These successive sister taxa

are *C. abbreviata* (- / 96 / 85 / 75), then *C. selaginoides* (1.00 / 84 / - / 57), then *C. fastigiata* (0.64 / - / 88 / 82), then *C. mertensiana* (1.00 / - / - / 52).

Nuclear Data (Fig. 4.2)

Analyses of nuclear data (nrITS + waxy) generated a different topology than trnS-G-G chloroplast data. *Cassiope* is again strongly resolved as monophyletic (0.99 / 94 / 100 / 100) and is supported as sister to representatives of the Ericoideae (1.00 / 100 / 100 / 100). The two *Harrimanella* species form a clade (1.00 / 98 / 97 / 96) that is sister to *Cassiope* + Ericoideae (1.00 / 100 / 85 / 55).

Within *Cassiope*, three main clades are resolved. Within one clade, *C. pectinata* is sister to *C. selaginoides* (1.00 / 61 / 71 / 51) and these two taxa are strongly supported as sister to *C. palpebrata* (1.00 / 99 / 99 / 99). *Cassiope nana* is resolved with variable support as sister to these (1.00 / - / 60 / -), followed by *C. abbreviata*. These taxa are moderately to strongly supported as a clade by Bayesian and PhyML analyses, but not by RAxML and MP analyses (1.00 / 63 / 70 / 50). Within the next clade, *C. ericoides* is strongly supported as sister to *C. wardii* (1.00 / 100 / 99 / 95). *Cassiope fastigiata*, then *C. dendrotricha* are successive sister taxa, but both lack good support. The clade as a whole lacks strong support from nuclear data (0.87 / - / 65 / -). Bayesian support exists (1.00 / 66 / 52 / -) for a sister relationship of the '*pectinata*' clade and the '*ericoides*' clade. *Cassiope myosuroides* is resolved as sister to this large clade, but has moderate support from PhyML analysis only (0.84 / 58 / 82 / 51). Within the third clade, *C. lycopodioides* is supported by Bayesian analyses as sister to *C. mertensiana* (0.98 / 67 /

65 / 67) and these two are sister to *C. tetragona*. The clade as a whole lacks strong support (0.80 / 56 / 56 / 67).

Combined Molecular Data (Fig. 4.3)

The ILD test determined that the nuclear and chloroplast data partitions are in conflict (P=0.01). Strictly speaking, this indicates that the data cannot be combined for a total evidence analysis. However, questions about the applicability of the ILD test in phylogenetics have been raised by various researchers recently (e. g. Yoder et al., 2001; Barker and Lutzoni, 2002; Cunningham, 1997). These concerns stem in part from the likelihood of making a Type I error (i. e. failing to combine the datasets when they are not incongruent). Cunningham (1997) suggested that a more lenient P-value of between 0.01 and 0.001 would be more realistic value for the ILD test for use in phylogenetic studies, while Barker and Lutzoni (2002) argued that even these values were not appropriate to aid in determining whether to combine datasets. For two important reasons, data from the two genomes were combined. First, there is no consensus within the field of phylogenetics on the most useful way to determine whether data from different genomes should be combined, and therefore the concerns of Yoder et al (2001), Barker and Lutzoni (2002) and Cunningham (1997) regarding failure to combine data when they are actually 'combinable' in reality (a Type I error) is not inconsequential. Second, the evolutionary history of the organism reflects all genomes, whether or not there is statistical conflict. In this case, the decision was made to first examine the impact of combining the data, in order to determine if support increased for nodes where no

conflict occurred, or if weakly supported nodes in the chloroplast and nuclear genomes were resolved with stronger support in the combined analyses. Even if hybridization or lineage sorting has resulted in conflicting phylogenetic signals in the nuclear and chloroplast genomes, currently available methods of tree building are not well-suited for visualizing reticulate patterns of evolution. It is still desirable, therefore, to estimate the nearest relative of taxa that may have been impacted by non-cladogenetic events.

Combined analyses of three DNA regions resolves *Cassiope* as a clade, and strong support is recovered by all analyses except RAxML (1.00 / 65 / 100 / 100). *Harrimanella* is resolved in a clade (1.00 / 100 / 99 / 100) sister to representatives of the Vaccinioideae (0.99 / 73 / 100 / 67).

Cassiope tetragona and *C. mertensiana* are resolved as successive sister taxa to a large clade comprised of the remaining *Cassiope*, but the *C. mertensiana* node is poorly to moderately supported (0.78 / 78 / - / -). Within the large clade, two main clades are recovered. The first is strongly supported by most analyses (1.00 / 63 / 94 / 92). Within this clade, *C. pectinata* is strongly supported as sister to *C. nana* (1.00 / 100 / 100 / 100). *Cassiope palpebrata* is supported by most analyses as sister to *C. selaginoides* (1.00 / - / 80 / 80). These two pairs ([*C. nana* + *C. pectinata*] and [*C. palpebrata* + *C. selaginoides*]) are supported as a clade by Bayesian analysis (1.00 / - / 72 / 54), and this clade is sister to *C. myosuroides*. The second clade lacks strong support except by Bayesian analysis (1.00 / - / - / 50). Within this clade, *C. ericoides* is moderately supported as sister to *C. wardii* (1.00 / 85 / 75 / 77). This small clade is resolved in a polytomy with *C. dendrotricha* and *C. lycopodioides* (1.00 / - / 53 / 62). *Cassiope abbreviata* is sister to this clade, but with very low support (1.00 / - / - / 50), followed by

C. fastigata. A phylogram (Fig. 4.4) indicates a very short branch at the base of *Cassiope* coinciding with the only poorly supported node in the Bayesian analysis. Other branches in the analysis are relatively moderate in length.

Character Evolution (Figs. 4.5 through 4.10)

Mapping of morphological and anatomical characters indicated that some characters show apparent evolutionary trends. Those continuously varying anatomical characters that do not show a trend phylogenetically are not reported here. These characters included presence of opposite pits, presence of alternate pits, growth ring width, number of rays, ray height and width, tracheid length, vessel length and diameter, number of vessels per group, and conductivity, mesomorphy and vulnerability indices. These characters are either largely uniform across species of *Cassiope*, or the variability is distributed such that it appears more or less random. Should the preliminary phylogenetic analyses presented here be confirmed in future studies, the apparent randomly distributed character states could benefit from being re-evaluated based upon their potentially independent origins. These characters are not discussed further here.

All categorical characters examined and those continuously varying characters that do show some phylogenetic signal are shown in Figures 4.5 through 4.10. These included number of growth rings, vessel thickness, and reported elevation means.

Leaf form (Figure 4.5) is apparently homoplasious. Most leaves in *Cassiope* are ericoid in general architecture and this leaf form is found throughout the phylogram. The three species with concave leaves are not closely related in the combined analysis, and therefore a separate origin of these leaf types must be hypothesized. The origin of the

unique leaf type of *C. redowskii* is unclear since it was not included in this study. The presence of abaxial stomata (Figure 4.6) exhibits the same pattern as leaf type; the species with concave or planar leaves do not have abaxial stomata and therefore it appears that the absence of this character state represents multiple independent losses. All species with ericoid leaves have stomata within the abaxial groove. Therefore, leaf form is exactly correlated with leaf stomatal distribution based upon the current taxon sampling and combined molecular analysis.

Within *Cassiope*, there is a trend toward a decreasing number of growth rings (Figure 4.7). *Cassiope mertensiana* and *C. tetragona* both have a relatively high number of growth rings. Within the large clade of most *Cassiope* species, the clade that includes *C. selaginoides* and *C. myosuroides* also has relatively high numbers of growth rings and high vessel thickness, whereas the clade including *C. wardii* and *C. fastigiata* have much smaller numbers.

Although the elevation range of most *Cassiope* species is not especially narrow, the mean elevation reported for each species (Figure 4.8) also exhibits a pattern. In general, *C. mertensiana* and *C. tetragona* occur at lower average elevations than the remaining *Cassiope* species. The clade including *C. selaginoides* and *C. myosuroides* and their relatives generally occur at higher elevations than the clade that includes *C. wardii* and *C. fastigiata*. There is also an overall reduction in stem diameter (Figure 4.9) and an increase in vessel density (Figure 4.10). While mapping mean elevation as a character is a relatively coarse approach, wood anatomy is often somewhat correlated with habitat, including elevation, and these trends are discussed below.

DISCUSSION

Based on Bayesian, Maximum Likelihood (RAxML and PhyML) and Maximum Parsimony analyses of trnS-G-G chloroplast spacer data, two nuclear markers (nrITS and waxy) and combined molecular data, it is clear that the two *Harrimanella* species are properly excluded from *Cassiope*. This result is in agreement with Kron et al. (2002), who determined that *Harrimanella hypnoides* was resolved outside (and not sister to) the *Cassiope* clade. This study strongly suggests that both *Harrimanella* species form a clade, and that that clade is not closely related to *Cassiope*. Because the decision was made to remove the *Harrimanella hypnoides* nrITS sequence from the analysis, it would be valuable to examine additional specimens of this taxon in the future. The included specimen was wild-growing and therefore not likely the result of unintended hybridization similar to what might occur in a greenhouse setting. As a group, the *Cassiope* nrITS sequences are very similar (93.2% identity). The *H. hypnoides* nrITS sequence is not identical to any single *Cassiope* nrITS sequence, but the % identity between the *Cassiope* sequences and the *H. hypnoides* sequence is still high, at 91.4%. However, the nrITS sequences of *H. hypnoides* and *H. stelleriana* have just 75.1% identity. Therefore, there is no evidence of a clerical or other identification error leading to the mistaken labeling of a *Cassiope* sequence as a *Harrimanella* sequence. The nature of the *H. hypnoides* nrITS sequence cannot be addressed further with this dataset. Greater attention to this particular issue is warranted because if future studies continue to find that *Harrimanella hypnoides* possesses a *Cassiope*-like nrITS sequence, then hypotheses about the acquisition of this sequence by a relatively distantly related ericad can be tested. The nrITS sequence of *H. stelleriana* was unlike the *Cassiope* sequences, and no

analyses placed this taxon closely related to *Cassiope*, therefore if the presence of the *Cassiope*-like sequence in *H. hypnoides* is upheld upon analysis of individuals from across the range of the species, then it may reflect an introgression event after the divergence of the two *Harrimanella* species. If the *Cassiope*-like sequence is found throughout the range of *H. hypnoides*, then an introgression event would presumably have occurred relatively early in the evolutionary history of this species and would indicate that genus *Harrimanella* is of hybrid origin. However, because *waxy* and *trnS-G-G* spacer did not group *H. hypnoides* with *Cassiope*, it is likely that this phenomenon is limited to nrITS only. In this case, it may be that the *Cassiope*-like nrITS sequence in *H. hypnoides* is the only amplifiable copy remaining in modern populations, which could result from concerted evolution of an ancestral polymorphism among the nrITS 'population.'

Individual analyses of chloroplast and nuclear data do not agree, and therefore additional data from both genomes would be valuable, since it is possible that a single chloroplast marker does not evolve at a rate that provides optimal phylogenetic signal. Because there is so much apparent conflict in the two genomes, it is necessary to consider the impacts of combining the data into a combined analysis.

There are few truly identical clades that appear in more than one analysis, and support varies widely among nodes, genomes, and analytical methods. However, a few clades can be recognized on the basis of being consistently resolved together, irrespective of precise placement and support, which may provide a basis for future studies.

In the first instance, *Cassiope wardii*, *C. ericoides* and *C. dendrotricha* are closely related in nuclear, chloroplast and combined analyses. Chloroplast data places *C.*

lycopodioides close to these taxa, whereas nuclear data places *C. fastigiata* close to them. In combined analyses, both taxa appear in a clade with this small, consistent clade. In chloroplast and combined analyses, *C. lycopodioides* also appears within the same clade. *Cassiope abbreviata* is placed close to the *C. wardii* clade by chloroplast and combined data, but nuclear data places it with the *C. nana* clade (see next). In the second instance, *Cassiope nana*, *C. pectinata* and *C. palpebrata* comprise three of four taxa in a clade, in all three analyses. Chloroplast data places *C. myosuroides* closest to them, whereas nuclear data places *C. selaginoides* closest to them. Combined analyses places *C. myosuroides* closest to them, with *C. selaginoides* sister to that clade. In the third instance, *Cassiope mertensiana* and *C. tetragona* are resolved as either a clade along with *C. lycopodioides* (nuclear) or a grade (chloroplast and combined). Nuclear and combined data places this group sister to the remaining *Cassiope* species, whereas in the chloroplast analyses, this clade/grade is nested somewhat within *Cassiope*. Therefore three taxa seem to be responsible for much of the phylogenetic 'noise': *Cassiope lycopodioides*, *C. abbreviata* and *C. selaginoides*. The positions of these taxa change most radically among analyses of different data partitions. These taxa may be influencing the positions of other taxa to a degree that diminishes support and resolution (e.g. *Cassiope fastigiata*, none of whose different placements are strongly supported). Inclusion of multiple accessions of each of these species, from across the geographic range, would be valuable in order to begin assessing whether any or all of them may be of hybrid origin, or whether species boundaries among taxa are simply unclear. Documentation of putative hybrid origin would require additional data, since the current study included a single chloroplast region. Although this region (*trnS-G-G* spacer) indicates a different placement for these

particular taxa in different analyses, it is possible that additional evidence from the chloroplast genome will generate a more stable placement. It is also possible that species boundaries among some of these taxa are unclear if they are not yet completely reproductively isolated, either because of incipient speciation, or because of secondary contact following relatively brief isolation. Detailed population-level studies measuring gene flow, perhaps in combination with absolute dating of nodes, would be valuable to assess which mechanism might be responsible for these 'fuzzy' species boundaries, and therefore the poor resolution of some *Cassiope* species.

In general, *Cassiope* species conform to the morphologies expected in alpine plants exposed to high light and high winds (Forsaith, 1920; Carlquist, 1977; Codignola et al., 2008; Dickison, 2000; Körner, 2003). Most *Cassiope* species are covered, some very densely, in hairs. This is thought to increase light reflectance, and/or reduce gas exchange across the interface of the leaf and the surrounding air. All *Cassiope* species examined so far have a rather thick, dark-staining cuticle on at least one leaf surface. Species with ericoid leaves tend to have a thicker cuticle on the abaxial surface, which faces 'outward' because of the appression of leaves to the stem, whereas their adaxial leaf surfaces, which more or less face the stem, tend to have thinner cuticles. There also appears to be a similar phylogenetic patten with regard to vessel density, stem diameter and elevation. Despite fairly crude estimates of mean elevation, it is apparent that as elevation increases among species of Cassiope, vessel density also increases, but stem diameter decreases. These also conforms to the alpine plant 'syndrome' because densely packed vessels decrease the likelihood of air embolism. However, while *Cassiope* species appear to have tightly packed palisade layers, they do not appear to have multiple

layers of palisade cells, as is reported generally for alpine plants (Dickison, 2000) . Additionally, the mesophyll layer in most of the ericoid *Cassiope*, is quite loosely arranged. In these respects, *Cassiope* species do not conform to the anticipated adaptations of alpine plants.

Despite the preliminary nature of this project, an attempt was made to detect evolutionary trends within the genus. The tight correlation between leaf form and leaf stomatal distribution is clear for the taxa included in this study. It is widely thought that the ericoid leaf form has evolved as an adaptation to environments where plants are exposed to high light and/or wind. In these environments, dessication is a regular problem faced by plants. It is thought that the restriction of abaxial stomata to within the abaxial groove may protect the plant from excessive dessication. *Cassiope* species with concave or planar leaves do not have abaxial leaf stomata, which is in agreement with this hypothesis. It was noted that while some morphological characters, such as leaf form and hair type, did not lend themselves to forming hypotheses about evolution within *Cassiope*, anatomical characters may have some explanatory power. Examination of the wood anatomy data collected by Wallace (1986) within a phylogenetic context provides a hint of the anatomical specialization that may have accompanied diversification within the genus.

The geographic ranges of most *Cassiope* species are poorly documented, particularly in China, where they are often described from a single locality, or a very small number of populations. Given that several new species have been described relatively recently (Hsu 1982; Fang 1999), it is highly likely that the full morphological, anatomical and genetic diversity of the genus is as yet undocumented.

ACKNOWLEDGEMENTS

The authors thank the following institutions and individuals for their assistance. The following herbaria provided specimens for study: Swedish Museum of Natural History Herbarium (S), Royal Botanic Garden Edinburgh (RBGE), University of Alaska Museum of the North (ALA), Harvard University Herbaria (HUH) and the University of British Columbia Botanical Garden. Torsten Eriksson provided technical advice. Rick Ree (Field Museum), Jun Wen (Smithsonian Institution), David Boufford (Harvard University), Peter Fritsch (California Academy of Sciences), and Bruce Bartholemew (California Academy of Sciences) generously provided plant material from the Gaoligonshan Biodiversity Survey collections. Gary Wallace (Rancho Santa Ana Botanic Garden) provided guidance early in the project, and Dr. Wallace and Rancho Santa Ana Botanic Garden permitted the use of his wood anatomy data.

LITERATURE CITED

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Automatic Control 19, 716–723.

Barker, F.K., Lutzoni, F. 2002. The utility of the incongruence length difference test. Syst. Biol. 51, 625–637.

Bliss, L.C. 1971. Arctic and alpine plant life cycles. Ann. Rev. of Ecol. and Syst. 2, 405-438.

Butterfass, T. 1987. The transverse orientation of stomata. Bot. Rev. 53, 415-441.

Codignola, A., Maffei, M., Fusconi, A., and Fieschi, M. 2008. Leaf anatomy of alpine plants as related to altitudinal stress. Nordic J. Bot. 7(6), 673-685.

Cunningham, C.W. 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. 14, 733–740.

Dickison, W.G. 2000. Integrated Plant Anatomy. Harcourt Academic Press, San Diego. 533 pages.

Drábková, L., Kirschner, J., Vlček, Č. 2002. Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens in Juncaceae. Plant Mol. Biol. Rep. 20, 161-175.

Drude, O. 1889. Ericaceae. Pp. 15–65 in A. Engler & K. Prantl (eds.), Die Natürlichen Pflanzenfamilien 4(1). W. Engelmann, Leipzig.

Drummond, A.J., Kearse, M., Heled, J., Moir, R., Thierer, T., Ashton, B., Wilson, A., and Stones-Havas, S. 2006. Geneious v2.5.

Fang, R-C. 1999. New taxa of Ericaceae from China. Novon 9(2), 162-178.

Fang, R-C. and Stevens, P.F. 2005. *Phyllodoce*. In: Flora of China. Wu, Z.Y., Raven, P.H., and Hong, D.Y., eds. 2005. Flora of China. Vol. 14 (Apiaceae through Ericaceae). Science Press, Beijing, and Missouri Botanical Garden Press, St. Louis.

Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C. 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. of Genet. 22, 521–565.

Forsaith, C.C. 1920. Anatomical reduction in some alpine plants. Ecology. 1(2), 124-135.
Gillespie, E. L. and Kron, K. A. 2010. Molecular phylogenetic relationships and a revised classification of the subfamily Ericoideae (Ericaceae). Mol. Phyl. Evol. doi:10.1016/j.ympev.2010.02.028

Good, R.D.O. 1926. The genera *Phyllodoce* and *Cassiope*. J. Bot. 64, 1-10.

Guindon, S., Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52(5), 696-704.
Guindon S., Lethiec, F., Duroux, P., Gascuel, O. 2005. PHYML Online—A web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. 1, 33 (Web Server issue):W557-9.

Hagerup, O. 1953. The morphology and systematics of the leaves in Ericales. Phytomorphology 3, 459-464.

Hooker, J.D. 1876. in: Bentham and Hooker. Genera Plantarum 2(2). Reeve and Co., London.

Hsu, T. 1982. Three new taxa of the genus *Cassiope* D. Don from Yunnan. Acta Botanica Yunnanica 4(3), 263-266.

Huelsenbeck, J.P., Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754-755.

Körner, C. 2003. Alpine plant life: Functional ecology of high mountain ecosystems. Berlin, Springer-Verlag. 343 pp.

Kron, K.A., Judd, W.S., Stevens, P.F., Crayn, D.M., Anderberg, A.A., Gadek, P.A., Quinn, C.J., Luteyn, J.L. 2002. Phylogenetic classification of Ericaceae: Molecular and morphological evidence. Bot. Rev. 68(3), 335-423.

Maddison, W.P. and Maddison, D.R. 2009. Mesquite: a modular system for evolutionary analysis. Version 2.72 http://mesquiteproject.org

Nylander, J.A.A. 2004. MrAIC.pl. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Ronquist, F., and Huelsenbeck, J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22(21), 2688-2690.

Stamatakis, A., Hoover, P., and Rougemont, J. 2008. A Rapid Bootstrap Algorithm for the RAxML Web-Servers, Systematic Biology, 75(5), 758-771.

Stevens, P.F. 1969. Taxonomic studies in the Ericaceae. Ph.D. Thesis. University of Edinburgh.

Stevens, P.F. 1971. A classification of the Ericaceae: Subfamilies and tribes. Bot. J. Linn. Soc. 64, 1-53.

Swofford, D.L. 2002. Paup* 4.0b10: Phylogenetic analysis using parsimony (* and other methods). Sunderland: Sinauer Associates.

Wallace, G.D. 1986. Wood anatomy of *Cassiope* (Ericaceae). Aliso 11(3), 393-415.

Wallace, G.D. 2009. *Phyllodoce*. In: Flora of North America Editorial Committee, eds. 1993+. Flora of North America North of Mexico. 15+ vols. New York and Oxford. Vol. 8, pp. 376, 443, 446 – 448.

Watson, L. 1962. Taxonomic significance of stomatal distribution and morphology in the Epacridaceae. New Phytologist 61, 36-40.

Watson, L. 1964. The taxonomic significance of certain anatomical observations on Ericaceae- The Ericoideae, *Calluna* and *Cassiope*. New Phytologist 63, 274-280.

Watson, L., Williams, W.T. and Lance, G.N. 1967. A mixed-data approach to angiosperm taxonomy: The classification of Ericales. Proc. Linn. Soc. London 178: 25–35.

Yoder, A.D., Irwin, J.A., and Payseur, B.A. 2001. Failure of the ILD to determine data combinability for slow loris phylogeny. Syst. Biol. 50, 408–424.

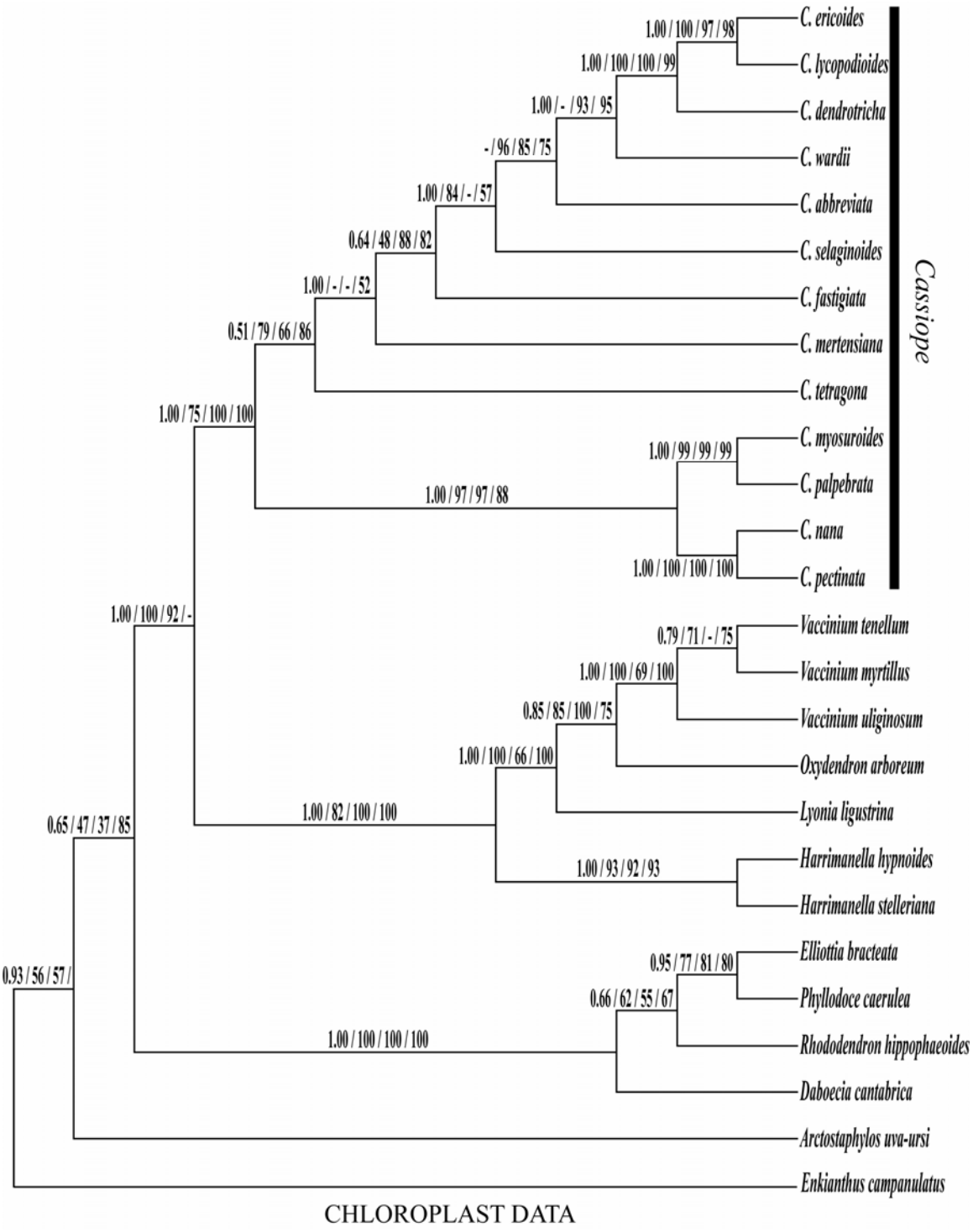


FIGURE 4.1. Chloroplast data (*trnS-G* spacer) analysis of *Cassiope*. Support values are to the left of nodes in the format (Bayesian posterior probability / ML bootstrap / MP bootstrap). Support derived from 20 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Tree length = 576, CI = 0.6510, RI = 0.7815.

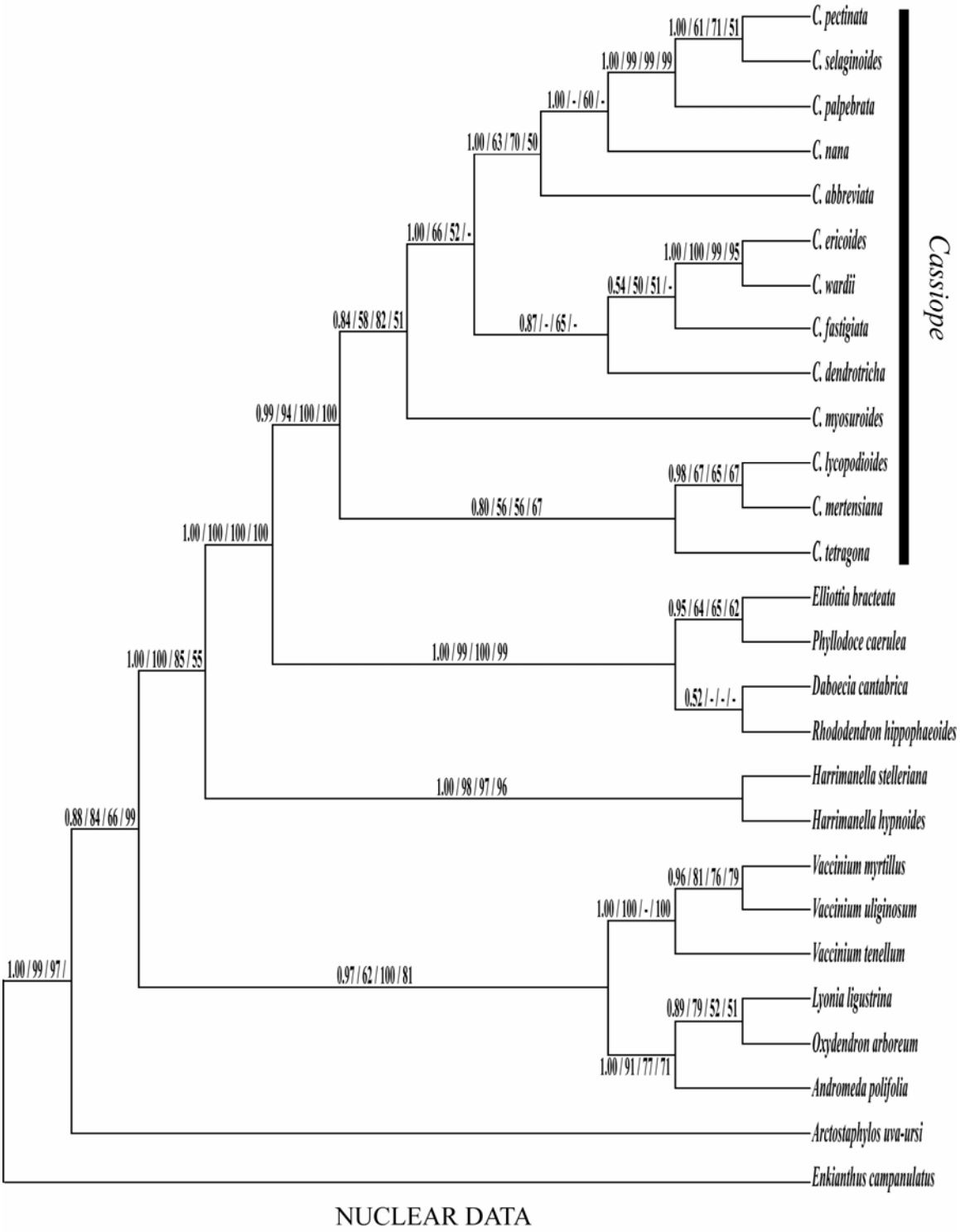


FIGURE 4.2. Total combined nuclear data (nrITS and *waxy*) analysis of *Cassiope*. Support values are to the left of nodes in the format (Bayesian posterior probability / ML bootstrap / MP bootstrap). Support derived from 20 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Tree length = 940, CI = 0.5926, RI = 0.7384.

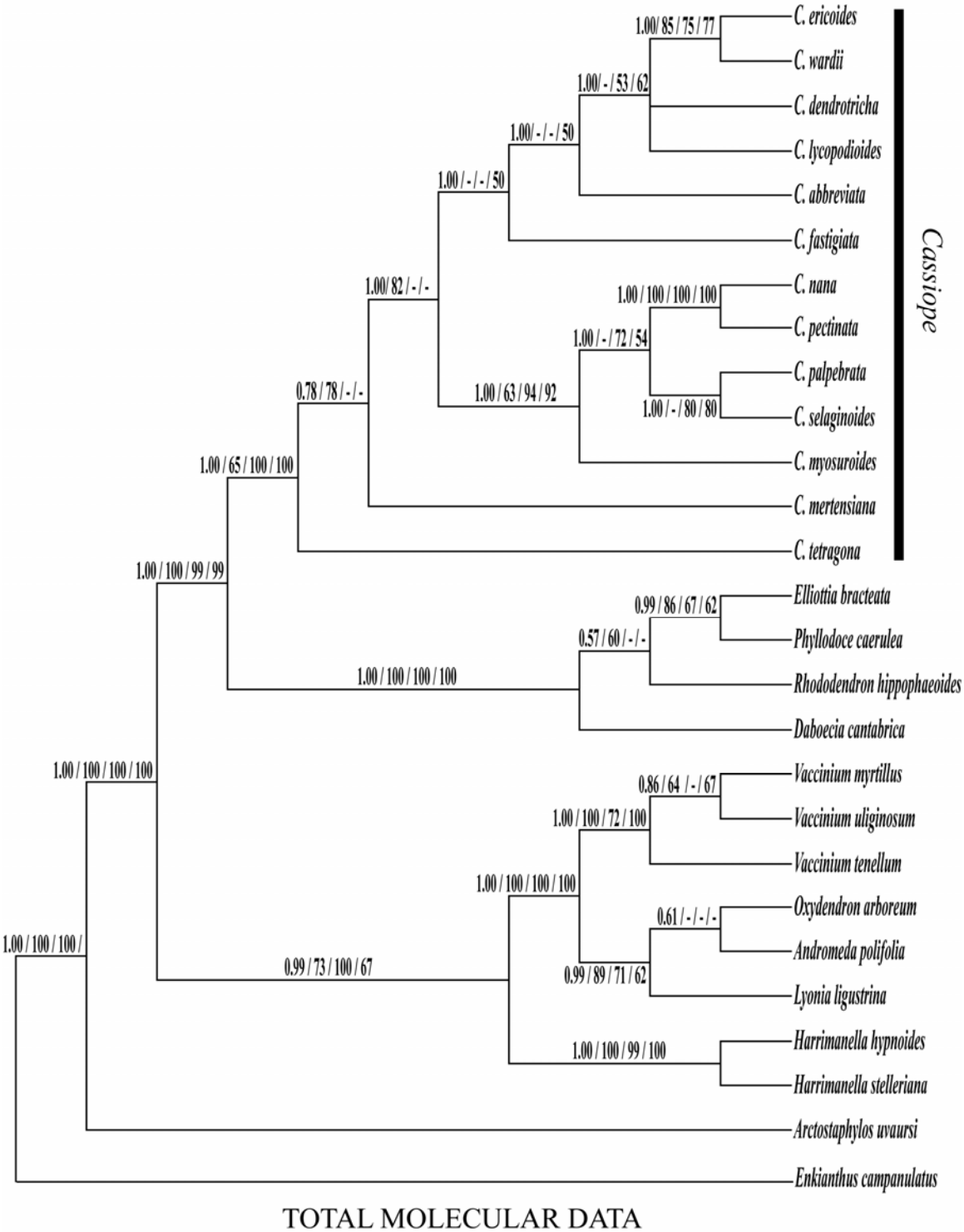


FIGURE 4.3. Total combined molecular data (*trnS-G* spacer, nrITS and *waxy*) analysis of *Cassiope*. Support values are to the left of nodes in the format (Bayesian posterior probability / ML bootstrap / MP bootstrap). Support derived from 20 million Bayesian MCMC generations, 100 ML bootstrap replicates and 10,000 MP bootstrap replicates. Tree length = 1594, CI = 0.5847, RI = 0.7223.

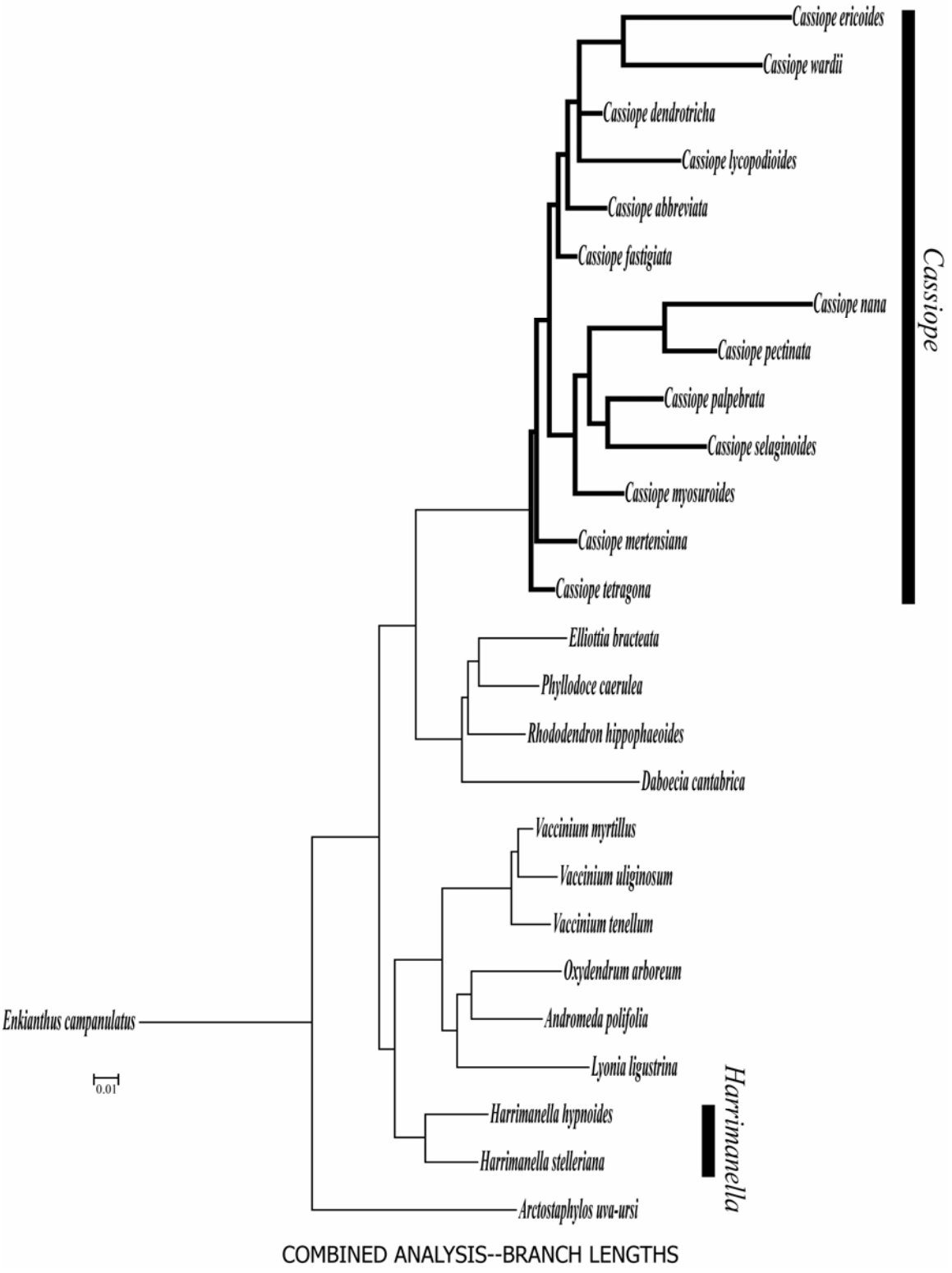


FIGURE 4.4. Phylogram of *Cassiope* showing branch lengths resulting from Bayesian analysis of total combined molecular data.

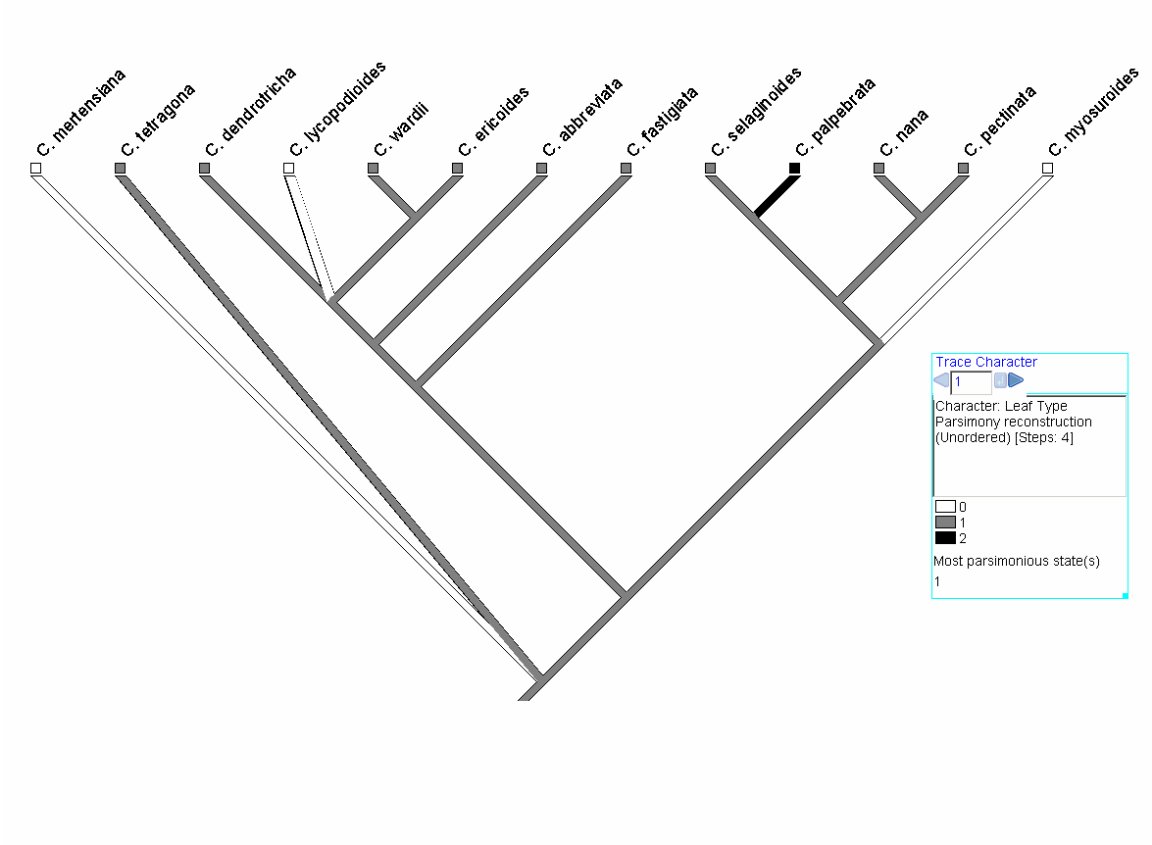


FIGURE 4.5. Parsimony reconstruction of leaf form in *Cassiope* traced onto the combined molecular phylogeny. 0 = concave leaf; 1 = ericoid leaf; 2 = planar leaf.

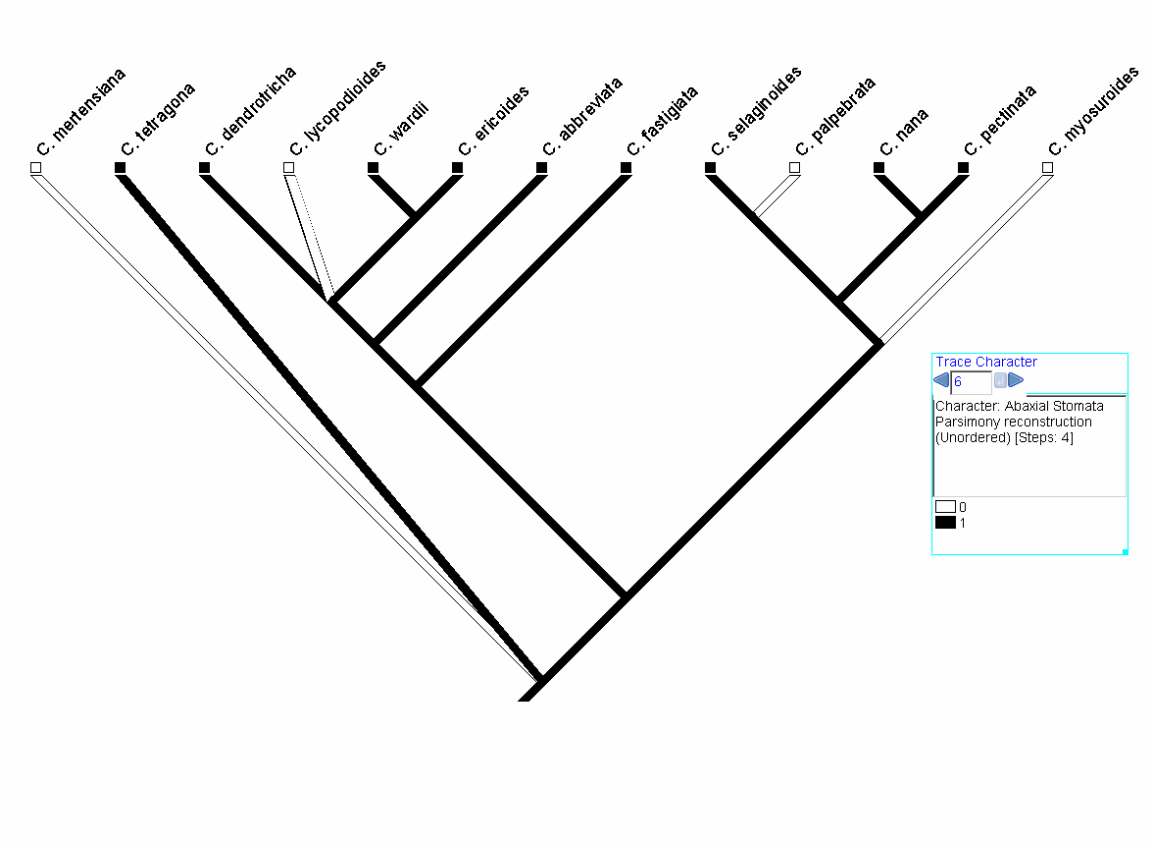


FIGURE 4.6. Parsimony reconstruction of abaxial leaf stomata in *Cassiope* traced onto the combined molecular phylogeny. 0 = absent; 1 = present.

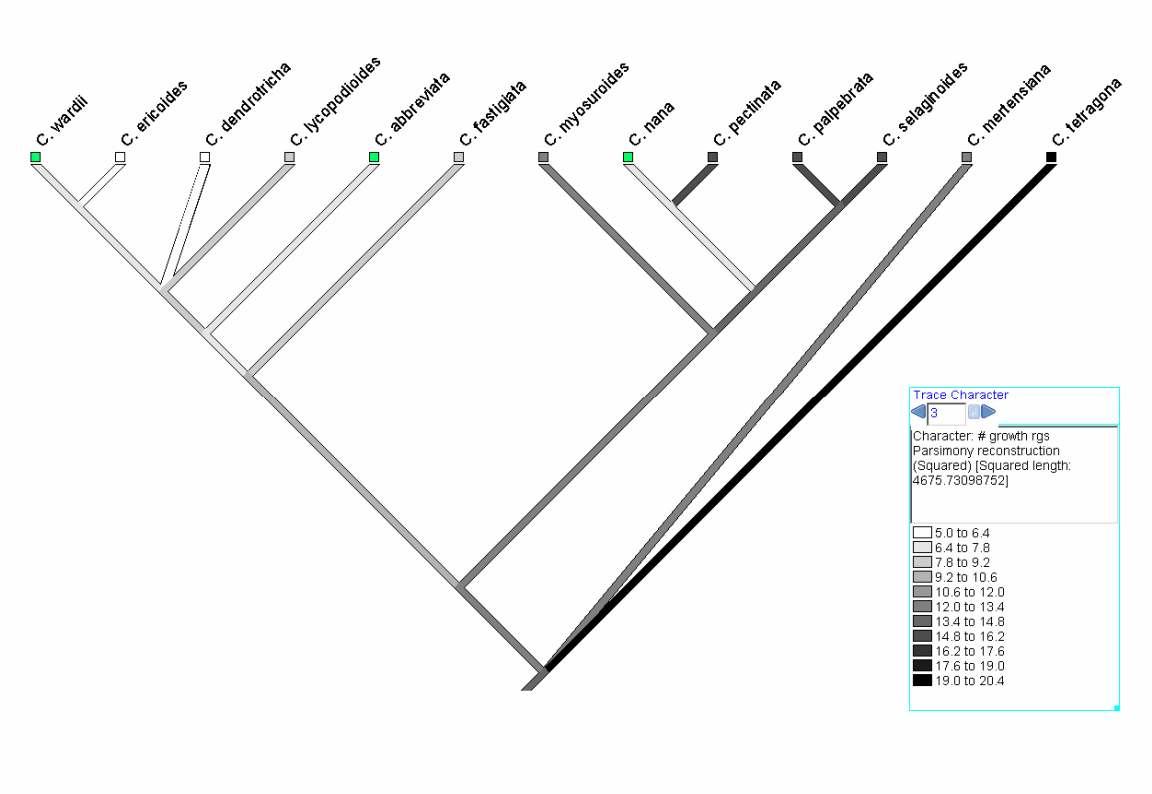


FIGURE 4.7. Parsimony reconstruction of number of growth rings in *Cassiope* traced onto the combined molecular phylogeny.

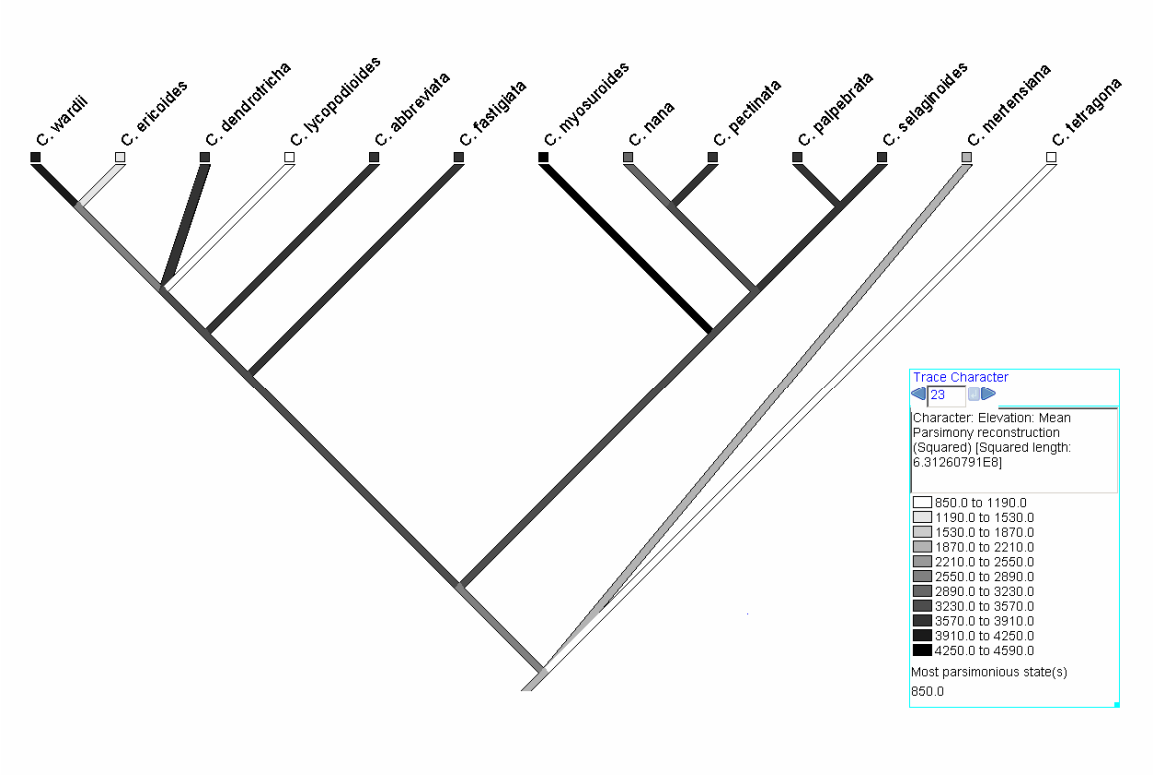


FIGURE 4.8. Parsimony reconstruction of mean elevation in *Cassiope* traced onto the combined molecular phylogeny.

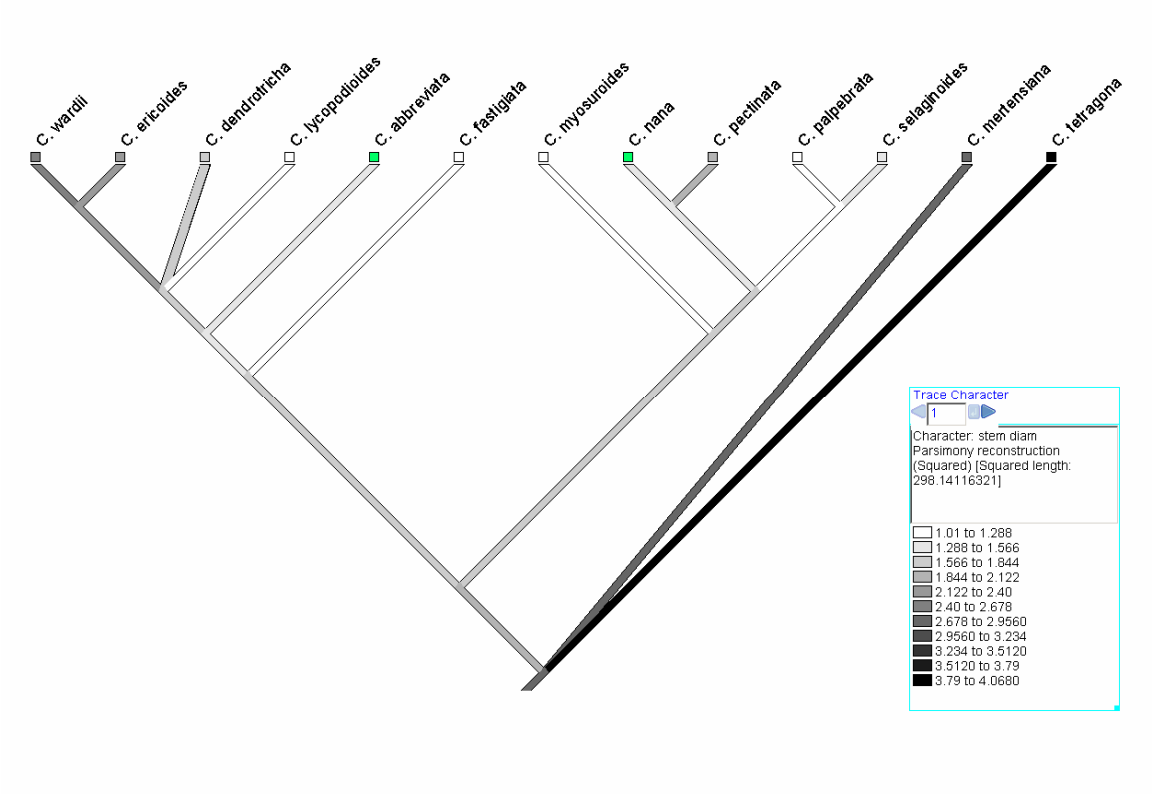


FIGURE 4.9. Parsimony reconstruction of stem diameter in *Cassiope* traced onto the combined molecular phylogeny.

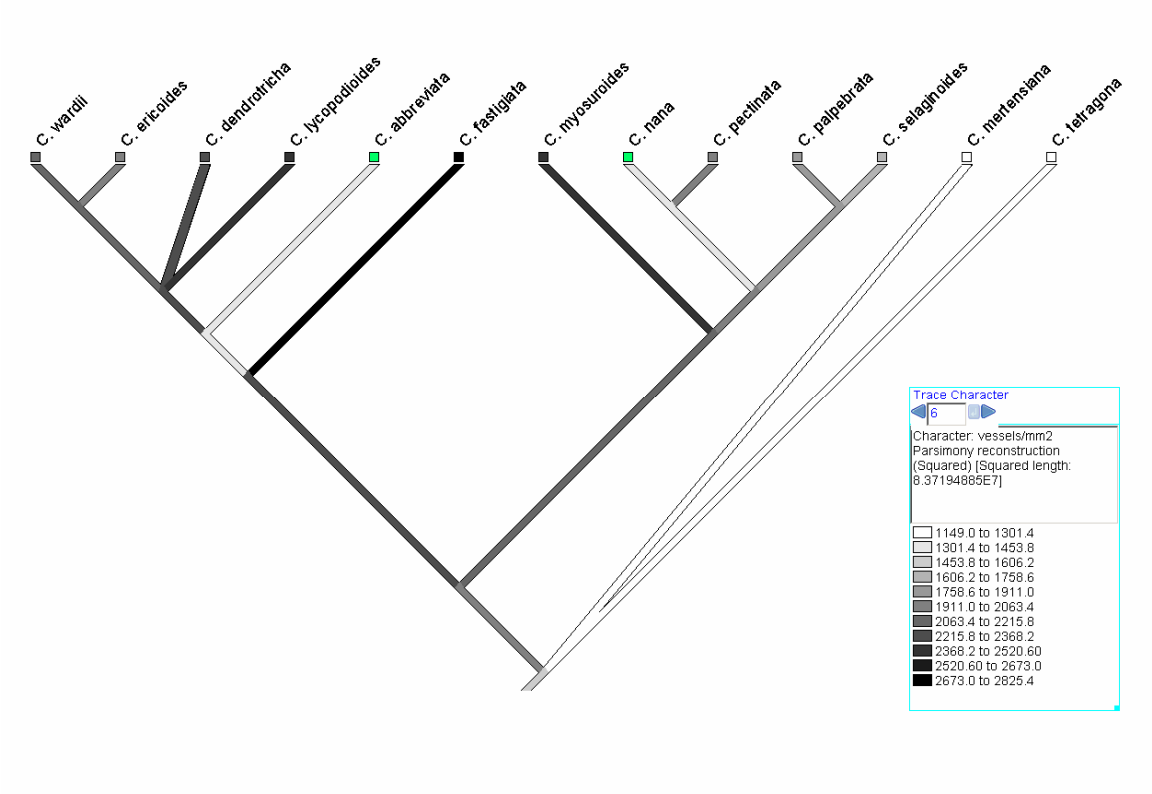


FIGURE 4.10. Parsimony reconstruction of vessel density (vessels per mm^2) in *Cassiope* traced onto the combined molecular phylogeny.

TABLE 4.1. Gene region, evolutionary model, aligned length, number of potentially parsimony informative characters (and %), and number of missing taxa (and %) in a phylogenetic study of *Cassiope*

Gene Region	**Evolutionary Model**	**Aligned Length (bp)**	**Informative Characters**	**# Missing Taxa (%)**
trnS-G-G	GTR+ Γ	1513	284 (18.8%)	1 (3.7%)
nrITS	SYM+Γ	635	190 (29.9%)	0
waxy	HKY+ Γ	593	170 (28.7%)	1 (3.7%)

Appendix 4.1. Taxa, voucher information and Genbank accession numbers for a phylogenetic study of *Cassiope*. Newly generated sequences for this study are in bold.

Species	**Voucher**	**nrITS**	***waxy***	***trnS-G-G***
Cassiope mertensiana G. Don	Anderberg 75-83, S	AF419798	DQ000598	GU176711
Cassiope fastigiata D. Don	1983, RBGE	**HM182073**	**HM182053**	**HM182089**
Cassiope lycopodioides D. Don	Studebaker 06-293, ALA	**HM182074**	**HM182054**	**HM182090**
Cassiope abbreviata Handel-Mazzetti	Gaoligongshan Biodiversity Survey #39745	**HM182070**	**HM182050**	**HM182086**
Cassiope wardii C. Marquand	D. Boufford et al. 30132 HUH	**HM182081**	**HM182061**	**HM182097**
Cassiope dendrotricha Handel-Mazzetti	Gaoligongshan Biodiversity Survey #690	**HM182071**	**HM182051**	**HM182087**
Cassiope ericoides (Pallas) D. Don	Zharkevich 1676, HUH	**HM182072**	**HM182052**	**HM182088**
Cassiope myosuroides W.W. Sm.	E00286046, RBGE	**HM182075**	**HM182055**	**HM182091**
Cassiope selaginoides Hook. F & Thompson	Gaoligongshan Biodiversity Survey #350	**HM182079**	**HM182059**	**HM182095**
Cassiope nana T.Z. Hsu	Gaoligongshan Biodiversity Survey #28588	**HM182076**	**HM182056**	**HM182092**
Cassiope palpebrata W.W. Sm.	Gaoligongshan Biodiversity Survey #32171	**HM182077**	**HM182057**	**HM182093**
Cassiope pectinata Stapf	Gaoligongshan Biodiversity Survey #26520	**HM182078**	**HM182058**	**HM182094**
Cassiope tetragona (L.) D. Don	Anderberg BN9253 S	**HM182080**	**HM182060**	**HM182096**
Harrimanella stelleriana Coville	Parker&Studebaker 16655 HUH	**HM182082**	**HM182063**	**HM182085**
Harrimanella hypnoides (L.) Coville	Anderberg BN9252, S	**HM182069**	**HM182062**	**HM182098**
Elliottia bracteata Benth. & Hook. f	Chase 866, K	U48609	DQ000600	GU176690
Phyllodoce caerulea (L.) Bab.	1940-1013, RBGE	GU176630	DQ000604	GU176700
Daboecia cantabrica (Huds.) C. Koch	1975-1770, RBGE	AY520786	GU176656	GU176688
Rhododendron hippophaeoides Balf. f & Sm.	1932-1022, RBGE	GU176634	GU176667	GU176707
Vaccinium tenellum Ait.	Kron & Powell s.n., WFU	AF382741	**HM182065**	GU176712
Vaccinium myrtillus L.	Anderberg s.n., WFU	AF382732	**HM182064**	DQ073200
Vaccinium uliginosum L.	VanderKloet s.n. WFU	AF419788	**HM182066**	DQ073187
Andromeda polifolia L.	1976-6099, RBGK	AF358872	**HM182067**	
Lyonia ligustrina DeCandolle	Kron s.n., WFU	**HM182083**		**HM182099**
Oxydendrum arboreum DeCandolle	Kron s.n., WFU	**HM182084**	**HM182068**	**HM182100**
Arctostaphylos uva-ursi Spreng.	Anderberg 361-68, S	AF106811	GU176668	GU176713
Enkianthus campanulatus Nichols	Anderberg 14528, S	AF133752		GU176714

CHAPTER 5

CONCLUSIONS

This study resulted in important contributions to our understanding of evolution within the Ericaceae by clarifying membership of tribes within the Ericoideae clade, the relationships of those tribes to each other, and relationships within *Cassiope*. Several well-supported clades within the Ericoideae had been previously identified, but the inter-relationships of these clades remained unclear. Additionally, the tribe Phyllodoceae could not be resolved as a clade at all. This lack of resolution has been an obstacle to more detailed studies within the subfamily. This study found that most genera previously recently classified within the Phyllodoceae are in fact closely related. These include *Epigaea*, *Kalmia*, *Kalmiopsis*, *Phyllodoce*, and *Rhodothamnus*. These genera comprise a 'core' Phyllodoceae. *Elliottia*, which had been classified in its own tribe (Cladothamneae) or with the rhododendrons (Rhodoreae), is sister to this core group. *Bejaria*, which had been classified within several different tribes (Bejarieae, Phyllodoceae, and Rhodoreae), is sister to the 'core' Phyllodoceae + *Elliottia*. Based upon these relationships, all of these genera were grouped together within the Phyllodoceae. The genera long thought to be closely related to members of the Phyllodoceae, *Bryanthus* and *Ledothamnus*, were found to be more distantly related, and were placed together in a new tribe, Bryantheae. The Phyllodoceae was found to be sister to a clade comprised of the other ericoid tribes (Bryantheae, Empetreae, Ericeae, and Rhodoreae). The recognition of a strongly supported Phyllodoceae permitted a more detailed study of this clade.

Phylogenetic analyses of the 'core' Phyllodoceae demonstrated that a group of four genera are closely related. These are *Epigaea* + *Rhodothamnus* and *Kalmiopsis* +

Phyllodoce. *Kalmia* is sister to these, followed by *Elliottia*. Relationships among species within this clade were generally strongly supported, which permitted an examination of morphology and historical biogeography. The challenges associated with the absence of a reliable fossil record and the complex timing of Northern Hemisphere biogeographical events presented an obstacle in this study. To attempt to circumvent this problem, a combination of methodologies was employed and the utility of this approach was examined. Ancestral areas of nodes were calculated, taking into account tree topology as well as Bayesian support for those nodes. Relative dating of these nodes was conducted, and nodes that could have been impacted by the same biogeographical history were identified. An explicit hypothesis regarding the age of these nodes was proposed, and the absolute ages of nodes within the clade were calculated based upon this hypothesis. It was found that absolute dates of shallower nodes fit reasonably well within established ideas about the biogeographical history of the Northern Hemisphere. However, deeper nodes contain much more error and are probably underestimated in the analysis presented here. This iterative approach appears to be a workable solution to a common problem in studies of Northern Hemisphere plant biogeography, however further refinement of model parameters and node constraints would be valuable.

A preliminary phylogenetic study of the genus *Cassiope* was undertaken because of its close relationship to the Ericoideae. Bayesian analysis was able to resolve most relationships with strong support. This study determined that the genus *Harrimanella*, previously thought to be closely related to (or nested within) *Cassiope*, was more distantly related, and was excluded from further analysis of *Cassiope*. However, the discovery of a *Cassiope*-like nrITS sequence in *H. hypnoides* will require more detailed

examination in a future study. The analysis of *Cassiope* presented here found no phylogenetic pattern in leaf type, the most immediately obvious source of morphological variation in the genus. Wood anatomy characters show some phylogenetic trends, and these trends generally agree with what would be expected in alpine plants. Therefore, *Cassiope* would be a good clade within which to further examine adaptations to demanding environments within a phylogenetic framework.

When viewed from a broad perspective, this analysis was generally one of 'clade discovery.' In evaluating the strategies used in this project, it is clear that a minimum of several genes are necessary to provide an appropriate level of resolution and support in the clades studied here. This expectation can be reasonably extended to other groups within the Ericaceae that exhibit similar levels of divergence. The six genes used in the Ericoideae and Phyllodoceae studies resulted in phylogenies that are very likely to remain stable in the future. The three genes used in the *Cassiope* study provided some information about phylogenetic relationships, but those relationships are not as confidently resolved. Given that patterns in wood anatomy are evident even in this preliminary study, it is expected that although additional data may refine the tree topology somewhat, this genus is well on its way to a stable phylogeny.

It is also clear that multiple analytical approaches are valuable in order to generate a robust phylogeny. While some nodes were supported by all analyses (particularly in the Ericoideae and Phyllodoceae studies), other nodes were supported by just one or two analyses. Occasionally there was conflict, or radically different levels of support, at different nodes. Therefore, to rely on a single analysis, or analyses based upon very similar theory, may mislead reconstruction of a phylogeny. Examination of the

peculiarities of any given dataset with regard to the inner workings of available analytical methods is usually beyond the scope of studies interested in phylogeny reconstruction (in contrast to studies of molecular evolution, for example). There is at least a minimal 'black box' effect in most phylogenetic studies because simulations used to demonstrate algorithm behavior are not based upon 'real' data. This study clearly demonstrates the value of employing multiple analytical frameworks in order to mitigate for this largely unavoidable issue.

While the phylogeny of *Cassiope* must be viewed as preliminary, it is already clear that questions about the developmental control of leaf form and the extent of variability in morphological and anatomical characters should be addressed. The geographic ranges of species also need to be documented in greater detail as a complement to any study of morphological variation, particularly in the Chinese taxa where several species of *Cassiope* are often found in proximity.

The phylogeny of the Phyllodoceae is very robust and provides a solid framework for a variety of future studies. More detailed examination can now be conducted on the evolution of the ericoid leaf form, developmental aspects of flower architecture, evolution of pollination mechanisms, and evolution of ecological niches. The establishment of species-level relationships within the Phyllodoceae will now allow examination of complex species boundaries among several sets of sister taxa, and examination of gene flow and population dynamics of reported naturally occurring hybrid taxa, among other questions. An understanding of species-level relationships of widespread taxa such as *Kalmia* and *Phyllodoce* provide an opportunity to examine phylogeographic patterns on a continental scale.

This study has generated an incremental improvement in our understanding of how the application of multiple molecular datasets, coupled with robust methods of analysis, can be used to conduct vigorous searches for phylogenetic structure in what have historically been recognized as “difficult” lineages. As in most intensive studies, more questions have been generated than have been answered. Hopefully, the research presented here will continue over time to prompt investigations into the biology of this diverse group of flowering plants.

VITA

Emily Laura Gillespie was born in Sylva, North Carolina on March 9, 1973 to James and Judith Gillespie. She lived at the boundary of the Great Smoky Mountains National Park and the Cherokee Indian Reservation until moving to Asheville, North Carolina in 1984. She graduated from Erwin High School in 1991 as an accomplished musician and pursued studies in music early in college. Eventually she discovered that biology was her true calling. At the University of North Carolina at Asheville, she studied the helminth parasites of mantled howling monkeys in Costa Rica, and earned a B.A. in Biology in 2000. She attended graduate school at Appalachian State University in Boone, North Carolina, where she discovered plant taxonomy. She studied the phylogeography and morphological variation in a widespread North American sedge, *Carex eburnea*, and completed her M.S. in Biology in 2005. She taught 6th grade science for a year in Winston-Salem, North Carolina before continuing as a doctoral student at Wake Forest University, where she earned her Ph.D. in 2010, studying evolution of the laurel and mountain heather clade (tribe Phyllodoceae) and the moss heather clade (*Cassiope*) within the plant family Ericaceae.