Molecular Phylogenetics and Evolution 106 (2017) 192–208

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylogeny of Helieae (Gentianaceae): Resolving taxonomic chaos in a Neotropical clade

Maria Fernanda Calió [a,*], Katherine B. Lepis [b,*], José Rubens Pirani [a], Lena Struwe [b,c]

[a] Universidade de São Paulo, Instituto de Biociências, Departamento de Botânica, Rua do Matão, 277, CEP: 05508-090 São Paulo, SP, Brazil
[b] Rutgers University, Department of Plant Biology, 59 Dudley Road, New Brunswick, NJ 08901, USA
[c] Rutgers University, Department of Ecology, Evolution, and Natural Resources, 14 College Farm Road, New Brunswick, NJ 08901, USA

ARTICLE INFO

Article history:
Received 25 April 2016
Revised 19 August 2016
Accepted 14 September 2016
Available online 15 September 2016

Keywords:
5S-NTS
Gentianales
ITS
Morphology
matK
Systematics
Neotropics

ABSTRACT

The monophyletic and Neotropical tribe Helieae of the worldwide family Gentianaceae (Gentianales, Asterids, Angiospermae) is well known for its problematic generic classifications. An initial phylogenetic analysis of Helieae shed light onto the relationships between genera, and indicated that traditional generic limits did not correspond to monophyletic groups. In order to obtain a more thorough understanding of generic relationships within the group, we enhanced sampling within the so-called *Symbolanthus* clade and performed phylogenetic analyses from DNA sequences from one plastid region (matK) and two nuclear regions (ITS and 5S-NTS), plus 112 morphological characters, which were analyzed separately and in combination, using parsimony and Bayesian approaches. A total of 83 individuals representing 20 genera and 51 species of Helieae were sampled; 13 species were included in this study solely based on their morphological characters. Ancestral character reconstructions were performed to identify potential synapomorphies of clades and patterns of homoplasy in the morphological dataset. Our results demonstrate that *Prepusa* is sister to the remainder of Helieae. Furthermore, the *Macrocarpaea* clade, the *Irlbachia* clade and the *Symbolanthus* clade were also recovered. Within the *Symbolanthus* clade, our results confirm that *Calolisianthus* and *Chelonanthus* are not monophyletic, and also contest the monophyly of *Irlbachia* as currently circumscribed. Specifically, two species of *Calolisianthus* group with the type species of *Chelonanthus*, while the other *Calolisianthus* species are more closely related to *Tetrapollinia* and *Symbolanthus*. Moreover, the green-white-flowered *Chelonanthus* species and *Adenolisianthus* are undoubtedly related to *Helia* and several analyses support *Irlbachia pratensis* as more closely related to the lineage including the type species of *Chelonanthus* described above The addition of new characters and taxa led to higher confidence in the relative position of some clades, as well as provided further support for a new generic circumscription of *Calolisianthus*, *Chelonanthus*, and *Helia*. Even though several morphological characters traditionally used in the taxonomy of the group were shown to be homoplasious, most clades can be diagnosed by a combination of morphological character states.

1. Introduction

The taxonomic chaos associated with certain Neotropical angiosperms is legendary, and the tribe Helieae (Gentianaceae) is a prime example (Bentham and Hooker, 1876; Gilg, 1895; Grisebach, 1839, 1845; Kuntze, 1891; Lepis, 2009; Maas, 1985; Nilsson, 1970; Progel, 1865; Struwe et al., 2002, 1997). The taxonomic confusion that has characterized this group stems from the morphological diversity it exhibits, especially florally and palynologically. This is exemplified when comparing the trees and shrubs of *Macrocarpaea* and *Chorisepalum*, to the straggly vines of *Purdieanthus*, to the inconspicuous savanna herbs of *Tetrapollinia* (Struwe et al., 1999, 2002, 2009). The morphological diversity is further demonstrated by the various floral architectures and colors observed (Fig. 1). This floral diversity lends itself to a diverse expression of pollination syndromes with evidence supporting pollination by bees, moths, hummingbirds, and bats (Machado et al., 1998; Struwe et al., 2002; J.R. Grant, pers. comm.). However, even though members of the Helieae are morphologically variable, there are several potential synapomorphies that distinguish them from the rest of the family. These characters include bilamellate stigmas

* Corresponding authors at: Universidade Estadual de Campinas, Instituto de Biologia, Departamento de Biologia Vegetal, Rua Monteiro Lobato, 255, CEP: 13083-862 Campinas, SP, Brazil (M.F. Calió). Monmouth University, Biology Department, 400 Cedar Ave., West Long Branch, NJ 07764, USA (K.B. Lepis).
E-mail addresses: mfecalio@unicamp.br (M.F. Calió), klepis@monmouth.edu (K.B. Lepis).

http://dx.doi.org/10.1016/j.ympev.2016.09.013
1055-7903/

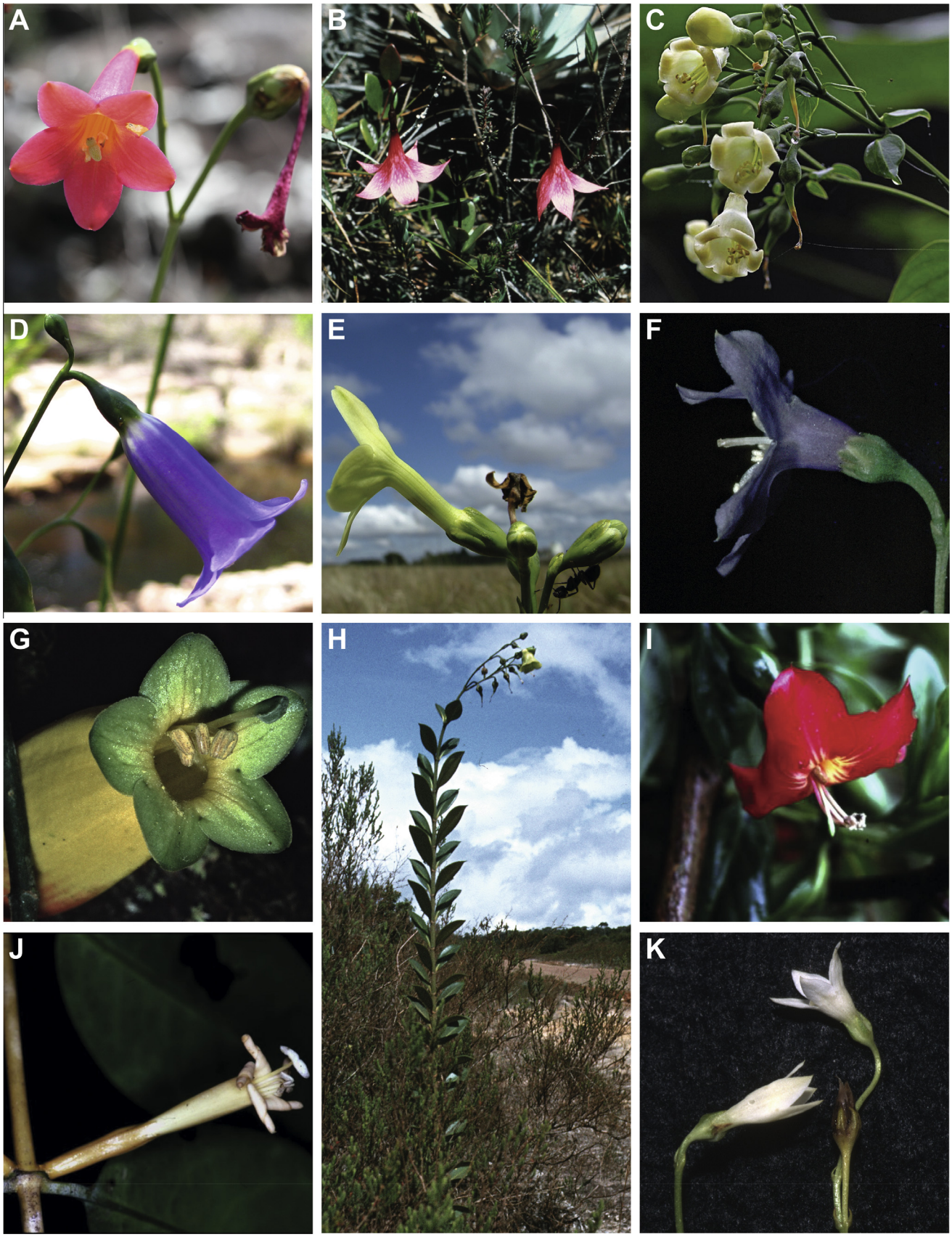

Fig. 1. Floral diversity in Helieae. (A) *Calolisianthus pedunculatus*. (B) *Celiantha bella*. (C) *Chelonanthus acutangulus*. (D) *Chelonanthus purpurascens*. (E) *Helia brevifolia*. (F) *Irbachia cardonae*. (G) *Lagenanthus princeps*. (H) *Rogersonanthus arboreus*. (I) *Symbolanthus australis*. (J) *Tachia guianensis*. (K) *Tetrapollinia caerulescens*. (photo credits: J. Lovo – A; S. Clamants – B; P. Hoell – C; V. Dittrich – D; M. Trovó – E; P.J.M. Maas – F, G, H, K; L. Dorr – I; C. Gracie – J).

atop long styles that become flat and twisted when dried, sessile ovaries with basal glandular areas, pollen with elaborate exine topologies, and dorsal glandular areas on each calyx lobe (Struwe et al., 2002, 2009).

The species of Helieae are also diverse in terms of their ecology and patterns of geographical distribution. They can be found in a wide variety of wet to semi-dry tropical habitats ranging from the summits of the *tepuis* in Venezuela and *páramos* of the Andes, the forested slopes and open *campos* of the Guayana and Brazilian Shields, to the seasonally inundated lowland forests and savannas of the river basins flowing toward the Atlantic Ocean. A few species reach into Mexico or onto some of the Caribbean Islands. Many of the common species and genera are well known to field botanists of the Neotropics, such as *Chelonanthus alatus* (Aubl.) Pulle, *Macrocarpaea glabra* Gilg, *Symbolanthus pulcherrimus* Gilg, and *Tachia grandiflora* Maguire & Weaver. There are also many local endemics and endangered species and genera in this tribe, for example, *Aripuana cullmaniorum* Struwe, Maas & V.A. Albert, *Prepusa*, *Roraimaea*, and *Senaea* (Calió et al., 2008; Struwe et al., 1997, 2008).

The problems associated with generic and tribal circumscriptions have persisted since the mid-1850 s. In fact, much of today's Helieae species have, historically, been placed in a variety of tribes, subtribes, and genera (for a detailed taxonomic history, see Lepis, 2009; Nilsson, 1970; Struwe et al., 1997; Supplemental material S1). Only with the advent of molecular phylogenetics has some of this confusion been resolved, and with the coupling these techniques with analyses of macro- and micro-morphological characters Helieae has, in its current form, begun to emerge (Struwe et al., 2002). Modern Helieae, a well-supported monophyletic group (Struwe et al., 2002, 2009; Struwe, 2014), is composed of at least two major evolutionary branches, the *Macrocarpaea* clade (including *Chorisepalum* Gleason & Wodehouse, *Macrocarpaea* (Griseb.) Gilg, *Tachia* Aubl., and *Zonanthus* Griseb.) and the *Symbolanthus* clade (*Adenolisianthus* (Progel) Gilg, *Aripuana* Struwe, Maas & V.A. Albert, *Calolisianthus* Gilg, *Chelonanthus* Gilg, *Helia* Mart., *Irlbachia pratensis* (Kunth) L. Cobb & Maas, *Lagenanthus* Gilg, *Lehmanniella* Gilg, *Rogersonanthus* Maguire & B.M. Boom, *Roraimaea* Struwe, S. Nilsson & V.A. Albert, *Purdieanthus* Gilg, *Sipapoantha* Maguire & B.M. Boom, *Symbolanthus* Gilg, *Tetrapollinia* Maguire & B.M. Boom, and probably *Yanomamua* Grant, Maas & Struwe). A third, smaller *Irlbachia* clade was also identified, containing all *Irlbachia* Mart. species except *Irlbachia pratensis* (positioned within the *Symbolanthus* clade). Analyses also consistently placed *Prepusa* Mart. and *Senaea* Taub. as sister to the rest of the tribe Helieae. Problems still persist with the presence of para- and polyphyletic genera and the uncertain placement of several genera, especially those represented only by morphological characters. Most of the uncertainty in generic delimitations and inter-generic relationships exists with the *Symbolanthus* clade, which was the main focus of the current study.

Attempts to sort out the taxonomic problems associated with Helieae are almost as variable as the diversity that characterizes this group (Supplemental material S1). Grisebach (1839, 1845) was the first to propose a family-wide classification for Gentianaceae. In his system, much of what we consider Helieae today was part of the tribe Lisyantheae which was composed of eight genera, some of which we recognize in modern-day Helieae (*Irlbachia*, *Prepusa*, *Tachia*). Many of the other Helieae genera were part of a large, broadly defined genus, *Lisyanthus*. *Lisyanthus* was divided into five sections, four of those bearing names we recognize as Helieae genera today (*Calolisyanthus* [=*Calolisianthus*], *Chelonanthus*, *Helia*, *Macrocarpaea*). Progel's (1865) Brazilian treatment of Gentianaceae followed that of Grisebach, but expanded the scope of *Lisianthus* (Grisebach changed the spelling in 1845) by adding a sixth section, *Adenolisianthus*. Bentham and Hooker (1876) followed suite and expanded the genus even further by adding the section *Symbolanthus* and by including *Irlbachia* as a section of *Lisianthus*. Defining the taxa somewhat differently from Grisebach and others, Kuntze (1891) also found it most appropriate to use a wide circumscription of the group, but did so under the name *Helia*. These widely circumscribed systems were predominantly based on the difficulty in definitively separating smaller taxa due to the occurrence of intermediates that blurred generic lines.

Based predominantly on pollen characters, Gilg (1895) divided Gentianaceae into three tribes, one of which he named Helieae. Gilg's Helieae divided the broadly defined *Lisianthus* into a number of smaller genera by elevating several of its sections to generic status. This is the first time we see the tribe Helieae and the genera *Adenolisianthus*, *Calolisianthus*, *Chelonanthus*, *Lagenanthus*, *Lehmanniella*, and *Purdieanthus* accepted. This taxonomic system was widely used in the decades that followed, but not without considerable debate. For example, in the treatment for *The Flora of Venezuela*, Steyermark (1953) concluded that Gilg's system based on pollen differences produced an artificial segregation of genera that otherwise lacked consistent taxonomic differences, and retained the use of *Calolisianthus*, *Chelonanthus*, and *Irlbachia* as sections of the widely circumscribed *Lisianthus*. The debate over the use of widely circumscribed genera vs. more, smaller genera continued. This is exemplified by the work of Maguire, who revised and described many new species and genera in Helieae from his work on Gentianaceae for the *Flora of the Guayana Highlands* (Maguire, 1981, 1985; Maguire and Boom, 1989). Maguire's genus circumscriptions focused largely, but not exclusively, on pollen characters. On the other hand, the comprehensive analysis of morphological characters in Helieae by Maas (1985) took a broad approach by including most non-woody Helieae into a broadly circumscribed *Irlbachia* (included in his *Irlbachia* were species of today's *Adenolisianthus*, *Calolisianthus*, *Chelonanthus*, *Helia*, *Irlbachia*, *Rogersonanthus* and *Tetrapollinia).*

With the analysis of molecular characters, emerging phylogenies demonstrate that a broad, widely defined *Irlbachia* or *Lisianthus* genus does not represent a monophyletic evolutionary lineage (Struwe et al., 2002, 2009). To address the problems with generic delimitations within Helieae and more specifically, the *Symbolanthus* clade, we assembled a more complete species sampling, accounting for about 23% of the known species of the tribe Helieae. We analyzed a broad morphological character base, including characters from gross habit attributes to flowers, leaves, fruits and seed anatomy and palynology along with nuclear and plastid DNA sequences.

The goal of this study was to recover a robust phylogeny of the *Symbolanthus* clade that would permit us (1) to examine the monophyly of the currently recognized genera, (2) to examine the relationships between the species and the genera, (3) to examine morphological traits that can aid in characterization of particular clades, and (4) to provide the evidence needed to support a more stable taxonomic classification of species in the genera *Adenolisianthus*, *Calolisianthus*, *Chelonanthus*, and *Helia*.

2. Material and methods

2.1. Taxon sampling

Relationships between Helieae and the closely related tribes Potalieae and Gentianeae are still unclear (Struwe, 2014), so we chose *Halenia palmeri* A. Gray from the tribe Gentianeae (Gentianaceae) for rooting purposes. In this study, 82 terminals from the tribe Helieae were included, representing 20 genera and 51 species currently recognized in the tribe (87% of genera and ~23% of species). Within these 51 species, we included 10 species

that clearly did not belong to the *Symbolanthus* clade, i.e., *Chorisepalum ovatum* Gleason, *Cho. psychotrioides* Ewan, *Irlbachia cardonae* (Gleason) Maguire, *I. nemorosa* (Willd. ex Roem. & Schult.) Merr., *I. poeppigii* (Griseb.) L. Cobb & Maas, *I. pumila* (Benth.) Maguire, *Macrocarpaea domingensis* Urb. & Ekman, *M. rubra* Malme, *Prepusa montana* Mart., and *Tachia guianensis* Aubl. Therefore, in our final analyses we had 11 terminals functioning as outgroups to the taxa in the *Symbolanthus* clade. New molecular sampling focused mostly on *Calolisianthus*, *Chelonanthus* and *Helia*. Only three genera were not included in this study, due to lack of available material for sequencing and extreme endemicity and rarity: *Senaea* (2 species currently recognized), *Yanomamua* (1 sp.), and *Zonanthus* (1 sp.). *Senaea* is closely related to *Prepusa* (Calió et al., 2008), and poorly known *Zonanthus* from Cuba presumably is a relative of *Macrocarpaea* and *Tachia* (Struwe et al., 2009). *Yanomamua* was described (Grant et al., 2006) based on limited material, but morphological evidence suggests a possible affinity to genera within the *Symbolanthus* clade; nevertheless, its placement is still uncertain.

To account for interspecific variation, multiple accessions for the same species were used whenever material was available. The voucher information for the molecular study is listed in Table 1. Additionally, 13 species without available DNA data (due to lack of material or failed PCR) were included in the phylogenetic analyses using only morphological data, specifically: *Celiantha bella* Maguire & Steyerm., *Ce. chimantensis* (Steyerm. & Maguire) Maguire, *Ce. imthurniana* (Oliv.) Maguire, *Chelonanthus hamatus* Lepis, *Irlbachia cardonae*, *Lagenanthus princeps* (Lindl.) Gilg, *Neblinantha neblinae* Maguire, *Rogersonanthus arboreus* (Britton) Maguire & B.M. Boom, *Roraimaea aurantiaca* Struwe, S. Nilsson & V.A. Albert, *Ror. coccinea* (Steyerm. ex L. Struwe, S. Nilsson & V.A. Albert) Struwe, S. Nilsson & V.A. Albert, *Sipapoantha ostrina* Maguire & B.M. Boom, *Symbolanthus argyreus* (Maguire) Struwe & K.R. Gould, and *S. elisabethae* (M. R. Schomb.) Gilg.

2.2. Morphological characters

Morphological characters used in this analysis derive from Struwe et al. (2009), with some modification, i.e. the removal of phylogenetically uninformative characters for our set of study taxa. Character descriptions and coding have been posted as online Supplemental data on the journal's website (Supplemental material S2 and S3). Morphological characters were scored based on analysis of liquid preserved material or dried specimens from the following herbaria: AAU, ALCB, BHCB, BR, CHRB, CEN, CEPEC, CESJ, CGMS, COL, CTES, CVRD, ESA, ESAL, F, G, GUA, HB, HRB, HUCS, HUEFS, HUFU, HXBH, IAC, IAN, IBGE, INPA, IPA, K, LIL, LP, MBM, MBML, MG, MO, NY, PMSP, R, RB, RBR, RUSU, S, SJRP, SP, SPF, SPFR, SPSF, U, UB, UEC, UFG, UPCB, UPS, US, VIC, and W (abbreviations follow *Index Herbariorum*). Furthermore, the studies of Bouman et al. (2002), Struwe et al. (2002, 2009), and Nilsson (2002) were consulted for coding seed and pollen characters. If a species was polymorphic for a character, the species was scored with all applicable states for that character.

2.3. Molecular methods

Total genomic DNA was extracted from herbarium specimens or silica gel dried leaves with the DNeasy Plant Mini Kit (Qiagen), using ca. 1 cm^2 of dried leaf ground with mortar and pestle and liquid nitrogen or with a FastPrep FP120 machine (Bio 101; for 20 s on speed 4). The Qiagen manufacturer protocol was followed, except for one modification in the beginning of the procedure; instead of adding 4 μL RNAse A stock solution, we added 30 μL of β-mercaptoethanol. The optional centrifuge in step 4 was always performed. The DNA regions were amplified by polymerase chain reaction (PCR) either using a DNA Engine with Dual Alpha Unit PTC 200 (BioRad) or a GeneAmp® PCR System 9700 (Applied Biosystems). Two short regions of the plastid matK gene were amplified separately using the primers pairs 1198F/1581R and 1729F/2035R designed by Thiv et al. (1999); the entire nrITS region was amplified with the primers 18S1830 forward and 26S25 reverse developed by Nickrent et al. (1994). In some instances ITS1 and ITS2 had to be amplified separately with the internal primers ITS3 and ITS2 created by White et al. (1990). The PI and PII primers designed by Cox et al. (1992) were used to amplify the 5S-NTS nuclear region. PCR conditions for the amplification of all three regions were as follows: 94 °C for 2 min followed by 27 cycles of 94 °C for 1 min, 60 °C for 1 min, 72 °C for 1 min, then a final extension for 4 min at 72 °C. All regions amplified in 50 μl reactions contained 10 μL 5x GoTaq Buffer (Promega), 2 μL of bovine serum albumin (0.4% BSA), 1 μL of 10 mM dNTPs, 2 μL of each primer (10 mM), 1 μl of genomic DNA and water to adjust the volume. Varying concentrations of $MgCl_2$ and polymerase were used for each marker, as follows: for the amplification of the matK region we used 5 μL of 25 mM $MgCl_2$ and 0.4 μL of GoTaq polymerase (Promega); the ITS region was amplified with 1.5 μL of 25 mM $MgCl_2$ and 0.4 μL of GoTaq; and for the 5S-NTS region we used 3.5 μL of 25 mM $MgCl_2$, 0.2 μL of GoTaq polymerase, and 10 μL of betaine (5 M) was added in amplifying 5S-NTS. Alternatively, regions amplified in 25 μL reactions were prepared with half the concentration of the reagents described above and 1 μl of genomic DNA and water to adjust the volume. Following PCR cycling, the amplification was checked by agarose gel electrophoresis using 1–1.2% agarose gels. In some instances, the first PCR produced weak products, if so we performed a second PCR using 1 μl of the first PCR product as the template. Some of the 5S-NTS amplifications produced several bands, but only one with the target size. In these cases, we excised the bands from the gel as separate bands, resuspended in water or cleaned with the Illustra GFX™ DNA and Gel Band Purification Kit (GE Healthcare), and reamplified as described above to increase product yield and isolate each fragment. Amplified PCR products ready for sequencing were purified using the Illustra GFX™ DNA and Gel Band Purification Kit (GE Healthcare) or QIAquick spin columns (Qiagen), following manufacturers' protocols. Purified PCR products were cycle sequenced using the same primers used for amplification and Big Dye terminator reaction mix (Applied Biosystems). Sequencing reactions were then cleaned using Sephadex column purification and sequencing runs were carried out either by Centro de Estudos do Genoma Humano (Instituto de Biociências, Universidade de São Paulo) or by Rutgers University's sequencing facility at the Center for Biotechnology using automated fluorescent sequencing. Sequencing of 5S-NTS PCR products of excised bands that were not the target length invariably failed. All new sequences generated in this study have been deposited in GenBank (Table 1).

2.4. Alignment and phylogenetic analysis

DNA sequence alignments were produced for each region using ClustalX 1.81 (Thompson et al., 1997), and then the alignments were inspected and manually adjusted using MacClade 4.08 (Maddison and Maddison, 2005), following the criterion of similarity (Simmons, 2004). Indels were coded as binary characters according to the simple coding method of Simmons and Ochoterena (2000).

The maximum parsimony (MP) analyses were run in PAUP 4.0b10 (Swofford, 2003). Heuristic searches were conducted with 1000 replicates using random addition, tree-bisection-reconnection (TBR) branch swapping, with all characters unordered and equally weighted. To evaluate node support, bootstrap analyses (BS; Felsenstein, 1985) were performed using 1000

Table 1
Taxa studied, voucher information, and GenBank accession numbers. Sequences taken from GenBank are listed with the corresponding publication. Letters after voucher indicate the specimen on the phylogenetic trees.

Taxa	Voucher	Locality	matK	ITS	5S-NTS
Halenia palmeri A. Gray	Holmgren and Lowrey 8073 (NY)	Mexico, Durango	AJ388169, AJ388239 Struwe et al. (2002)	–	–
Halenia palmeri A. Gray	von Hagen 98/41 (HAL)	Mexico, Durango	–	AJ294632, AJ294692 Hagen and Kadereit, 2001	–
Adenolisianthus arboreus Gilg	Maguire 55601 (NY)	Venezuela, Amazonas	–	EU709784 Struwe et al. (2009)	–
Aripuana cullmaniorum Struwe, Maas & V.A. Albert	Ferreira 5906 (NY)	Brazil, Amazonas	AJ388140, AJ388209 Struwe et al. (2002)	EU709785 Struwe et al. (2009)	–
Calolisianthus amplissimus (Mart.) Gilg	Calió and Sasaki 66 (SPF) [a]	Brazil, São Paulo	KX904531, KX904566	KX904643	KX904652
Calolisianthus amplissimus (Mart.) Gilg	Pereira-Silva et al. 7411 (SPF) [b]	Brazil, Goiás	KX904532, KX904567	KX904644	KX904653
Calolisianthus amplissimus (Mart.) Gilg	Romero et al. 926 (SPF) [c]	Brazil, Minas Gerais	KX904533, KX904568	KX904645	–
Calolisianthus pedunculathus (Cham. & Schltdl.) Gilg	Calió et al. 15 (SPF) [a]	Brazil, Minas Gerais	KX904534, KX904569	KX904612	KX904654
Calolisianthus pedunculathus (Cham. & Schltdl.) Gilg	Farinaccio et al. 621 (SPF) [b]	Brazil, Paraná	KX904535, KX904570	KX904613	KX904655
Calolisianthus pedunculathus (Cham. & Schltdl.) Gilg	Loeuille et al. 339 (SPF) [c]	Brazil, Bahia	KX904536, KX904571	KX904614	KX904656
Calolisianthus pedunculathus (Cham. & Schltdl.) Gilg	Lovo et al. 145 (SPF) [d]	Brazil, Minas Gerais	–	KX904615	KX904657
Calolisianthus pedunculathus (Cham. & Schltdl.) Gilg	Mello-Silva and Forzza 2797 (SPF) [e]	Brazil, Bahia	KX904537, KX904572	KX904616	KX904658
Calolisianthus pedunculathus (Cham. & Schltdl.) Gilg	Pirani et al. 5393 (SPF) [f]	Brazil, Minas Gerais	KX904538, KX904573	KX904617	KX904659
Calolisianthus pedunculathus (Cham. & Schltdl.) Gilg	Trovó and Watanabe 361 (SPF) [g]	Brazil, Minas Gerais	KX904539, KX904574	KX904618	KX904660
Calolisianthus pendulus (Mart.) Gilg	Calió et al. 86 (SPF) [a]	Brazil, Minas Gerais	KX904540, KX904575	KX904620	KX904661
Calolisianthus pendulus (Mart.) Gilg	Mello-Silva and Ferreira 2848 (SPF) [b]	Brazil, Minas Gerais	KX904541, KX904576	KX904619	KX904662
Calolisianthus speciosus (Cham. & Schltdl.) Gilg	Calió et al. 51 (SPF) [a]	Brazil, Goiás	KX904542, KX904577	KX904621	KX904663
Calolisianthus speciosus (Cham. & Schltdl.) Gilg	Calió et al. 106 (SPF) [b]	Brazil, Minas Gerais	KX904543, KX904578	KX904623	KX904664
Calolisianthus speciosus (Cham. & Schltdl.) Gilg	Calió et al. 125 (SPF) [c]	Brazil, Bahia	KX904544, KX904579	KX904622	KX904665
Calolisianthus speciosus (Cham. & Schltdl.) Gilg	Trovó et al. 415 (SPF) [d]	Brazil, Minas Gerais	KX904545, KX904580	KX904624	KX904666
Calolisianthus sp. 1	Bezerra et al. 46 (SPF) [a]	Brazil, Minas Gerais	KX904546, KX904581	KX904625	KX904651
Calolisianthus sp. 1	Calió et al. 87 (SPF) [b]	Brazil, Minas Gerais	KX904547, KX904582	KX904626	KX904650
Calolisianthus sp. 2	Barbosa and Araújo 231 (SPF) [a]	Brazil, Goiás	KX904548, KX904583	KX904627	KX904648
Calolisianthus sp. 2	Pereira-Silva et al. 7097 (SPF) [b]	Brazil, Minas Gerais	–	KX904628	KX904649
Chelonanthus acutangulus (Ruiz et Pav.) Gilg	Gomez et al. 837 (MO) [a]	Colombia, Antioquia	–	KX904629	KX904667
Chelonanthus acutangulus (Ruiz et Pav.) Gilg	Hawkes et al. 5048 (MO) [b]	Bolivia, La Paz	KX904549, KX904584	KX904647 (ITS 2)	KX904668
Chelonanthus alatus (Aubl.) Pulle	Berry 5541 (NY) [a]	Venezuela, Bolívar	KX904550, KX904585	EU709790 Struwe et al. (2009)	KX904670
Chelonanthus alatus (Aubl.) Pulle	Maas 9316 (U) [b]	French Guiana	KX904551, KX904586	KX904610	KX904669
Chelonanthus albus (Spruce ex Progel) V.M. Badillo	Poole 2049 (NY)	Brazil, Amazonas	–	EU709789 Struwe et al. (2009)	KX904671
Chelonanthus angustifolius (Kunth) Gilg	Alencar 621 (US) [a]	Colombia, Santander	–	KX904603 (ITS 1)	KX904673
Chelonanthus angustifolius (Kunth) Gilg	Molina 18S364 (US) [b]	Brazil, Amazonas	KX904552, KX904587	KX904611	KX904672
Chelonanthus grandiflorus (Aubl.) Chodat et Hassl.	Kelloff et al. 598 (US) [a]	Guiana, Demerara-Mahaica	KX904553, KX904588	KX904630	KX904674
Chelonanthus grandiflorus (Aubl.) Chodat et Hassl.	Mori et al. 24801 (NY) [b]	French Guiana, Saül	KX904554, KX904589	EU709788 Struwe et al. (2009)	KX904675
Chelonanthus grandiflorus (Aubl.) Chodat et Hassl.	Mori et al. 25561 (NY) [c]	French Guiana	–	KX904646	–
Chelonanthus matogrossensis (J.M.G. Pers. et Maas) Struwe et V.A. Albert	Solomon 7880 (NY)	Bolívia, El Beni	–	KX904604 (ITS 1)	KX904676
Chelonanthus pterocaulis Lepis	Vargas 454 (NY)	Costa Rica, San Jose	–	KX904602 (ITS 1)	KX904677

Table 1 (*continued*)

Taxa	Voucher	Locality	matK	ITS	5S-NTS
Chelonanthus purpurascens (Aubl.) Struwe, S. Nilsson et V.A. Albert	Berry 5533 MO[a]	Venezuela, Bolívar	–	EU709791 Struwe et al. (2009)	–
Chelonanthus purpurascens (Aubl.) Struwe, S. Nilsson et V.A. Albert	Calió et al. 119 (SPF) [b]	Brazil, Bahia	KX904555, KX904590	KX904639	–
Chelonanthus purpurascens (Aubl.) Struwe, S. Nilsson et V.A. Albert	Harley 15889 (US) [c]	Brazil, Bahia	–	KX904638	KX904679
Chelonanthus purpurascens (Aubl.) Struwe, S. Nilsson et V.A. Albert	Maas 7456 (US) [d]	French Guiana, Upper Takutu-Upper Essequibo	–	KX904601 (ITS 2)	KX904678
Chelonanthus viridiflorus (Mart.) Gilg	Anderson 9743 (NY) [a]	Brazil, Mato Grosso	–	EU709792 Struwe et al. (2009)	–
Chelonanthus viridiflorus (Mart.) Gilg	Calió et al. 84 (CHRB) [b]	Brazil, Minas Gerais	–	KX904605 (ITS 1)	KX904681
Chelonanthus viridiflorus (Mart.) Gilg	Calió et al. 157 (SPF) [c]	Brazil, Minas Gerais	KX904556, KX904591	KX904608	–
Chelonanthus viridiflorus (Mart.) Gilg	Chatrou 321 (U) [d]	Bolivia, Santa Cruz	–	KX904607	KX904680
Chorisepalum ovatum Gleason	Maguire and Politi 27921 (NY)	Venezuela, Amazonas	AJ388147, AJ388216 Struwe et al. (2002)	–	–
Chorisepalum psychotrioides Ewan	Hankel 4267 (NY)	Venezuela, Bolivar	–	EU709793 Struwe et al. (2009)	–
Helia brevifolia Cham.	Calió et al. 168 (SPF) [a]	Brazil, São Paulo	KX904557, KX904592	KX904631	KX904682
Helia brevifolia Cham.	Serafim et al. 27 (SPF) [b]	Brazil, São Paulo	–	KX904632	KX904683
Helia brevifolia Cham.	Trovó et al. 316 (SPF) [c]	Brazil, São Paulo	KX904558, KX904593	KX904633	KX904684
Helia oblongifolia Mart.	Calió et al. 205 (SPF) [a]	Brazil, Minas Gerais	KX904559, KX904594	KX904634	KX904686
Helia oblongifolia Mart.	Harley et al. 25924 (NY) [b]	Brazil, Bahia	KX904560, KX904595	KX904635	KX904687
Helia oblongifolia Mart.	Harley et al. 25648 (NY) [c]	Brazil, Bahia	–	EU709794 Struwe et al. (2009)	–
Helia oblongifolia Mart.	Irwin et al. 21805 (NY) [d]	Brazil, Minas Gerais	–	KX904636	KX904685
Irlbachia nemorosa (Willd. ex Roem. & Schult.) Merr.	Klinge s.n. (NY)	Venezuela, Amazonas	–	EU709795 Struwe et al. (2009)	–
Irlbachia poeppigii (Griseb.) L. Cobb & Maas	Toedzia et al. 2278 (NY)	Brazil, Amazonas	–	EU709796 Struwe et al. (2009)	–
Irlbachia pratensis (Kunth) L. Cobb & Maas	Berry 7579 (NY)	Venezuela, Amazonas	KX904561, KX904596	EU709797 Struwe et al. (2009)	KX904688
Irlbachia pumila (Benth.) Maguire	Maas 6907 (U)	Brazil, Amazonas	–	EU709798 Struwe et al. (2009)	–
Lehmanniella splendens (Hook.) Ewan	Callejas 8575 (NY)	Colombia, Antioquia	AJ388172, AJ388242 Struwe et al. (2002)	–	
Macrocarpaea domingensis Urb. & Ekman	Maas 8395 (U)	Dominican Republic	AJ010523, AJ011452 Thiv et al. (1999)	EU709799 Struwe et al. (2009)	–
Macrocarpaea rubra Malme	Calió et al. 169	Brazil, São Paulo	KX904562, KX904597	KX904609	–
Neblinantha parvifolia Maguire	Maguire et al. 42384 (NY)	Brazil-Venezuela border, Amazonas	AJ388179, AJ388249 Struwe et al. (2002)	–	–
Prepusa montana Mart.	Calió et al. 116 (SPF)	Brazil, Bahia	KX904563, KX904598	KX904637	–
Purdieanthus pulcher (Hook.) Gilg	John 20673 (US)	Colombia, Boyacá	–	EU709800 Struwe et al. (2009)	–
Rogersonanthus quelchii (N. E.Br.) Maguire & B.M. Boom	Fiaschi and Plunkett 3192 (SPF)	Venezuela, Bolívar	KX904564, KX904599	KX904640	–
Symbolanthus australis Struwe	Dorr et al. 6691 (NY)	Bolivia, La Paz	–	EU709801 Struwe et al. (2009)	KX904689
Symbolanthus frigidus (Sw.) Struwe & K. Gould	Cooley 8211 (NY)	St. Vincent	AJ388198, AJ388268 Struwe et al. (2002)	EU709802 Struwe et al. (2009)	AY143370 Gould and Struwe (2004)
Symbolanthus nerioides (Griseb.) Ewan	Sobel and Strudwick 2164 (NY)	Venezuela, Mérida	–	KX904641	AY143377 Gould and Struwe (2004)
Symbolanthus pulcherrimus Gilg	Morales et al. 1600 (NY)	Costa Rica, Cartago	–	EU709803 Struwe et al. (2009)	AY143379 Gould and Struwe (2004)
Tachia guianensis Aubl.	Rova 1963 (NY)	French Guiana, Cayenne	AJ011433, AJ011461 Thiv et al. (1999)	–	–
Tachia guianensis Aubl.	Mori 23439 (NY)	French Guiana, Cayenne	–	DQ401419 Struwe et al. (2009)	–
Tetrapollinia caerulescens (Aubl.) Maguire & B.M. Boom	Calió et al. 154 (SPF) [a]	Brazil, Minas Gerais	KX904565, KX904600	KX904642	KX904690
Tetrapollinia caerulescens (Aubl.) Maguire & B.M. Boom	Irwin 34170 (NY) [b]	Brazil, Goiás	–	KX904606 (ITS 1)	KX904691

pseudoreplicates, each with 10 random taxon addition replicates and TBR branch swapping. The overall degree of homoplasy was estimated using consistency and retention indices (CI and RI). We consider here nodes with bootstrap support values of ⩾85% strongly supported, 75–84% as moderately and 50–74% as weakly supported.

Bayesian inference (BI) was performed with MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) on the CIPRES Science Gateway (Miller et al., 2010). The best-fit substitution model for each data matrix was first determined under the Akaike information criteria implemented in MrModeltest 2.3 (Nylander, 2004). The selected models were GTR + G for matK SYM + G for ITS, and HKY + G for 5S-NTS. The morphological matrix was analyzed with the standard discrete model (Lewis, 2001). The Bayesian analyses were run with four Markov chains, starting from random trees, for ten million generations and sampled every 100th generation. MCMC convergence was examined inspecting PSRF values for all model parameters, and using Tracer 1.5 (Rambaut and Drummond, 2009). After the first 20% generations were discarded as burn-in, a majority-rule consensus tree with posterior probabilities (PP) was constructed. Posterior probabilities were used to evaluate support for all nodes; we consider that clades with posterior probabilities above 0.95 are strongly supported.

We did not use the ILD test (Farris et al., 1995) to assess incongruity between the separate datasets because its results have been shown to be misleading (e.g., Dolphin et al., 2000; Lee, 2001; Yoder et al., 2001). Instead, prior to analyzing our datasets in combined analyses, we conducted separate tree searches in the parsimony and Bayesian frameworks described above on each of the four datasets, aiming to detect potential problems for analyzing these datasets simultaneously. We used a conservative BS of more than 85% (Wiens, 1998) and PP of 0.95 as thresholds for identification of significant incongruent positions. Based on this criterion, because no incongruity was found, we carried on to perform the analyses with all datasets combined.

We constructed two matrices for the combined analyses, both containing data for all four datasets (morphology, matK, ITS, and 5S-NTS), but differing in the terminals included. The Total Combined matrix comprised all 83 sampled terminals, and the Partial Combined matrix consisted of 73 terminals. The terminals that were omitted from the Partial Combined analysis were seven species represented by morphological data only and that represented our sampling of entire genera (*Celiantha bella*, *C. chimanthensis*, *C. imthurniana*, *Lagenanthus princeps*, *Roraimaea aurantiaca*, *R. coccinea*, *Sipapoantha ostrina*). *Neblinantha neblinae* and *N. parvifolia* were represented by morphology and matK data, respectively, so they were omitted as well. *Lehmanniella splendens* was also omitted as it was represented only by morphology and matK data. With this procedure, we removed most of the terminals with larger amounts of missing data (over 90% of missing data), and those lacking congeners with more complete data sets. On the other hand, in the Partial Combined matrix, we kept five terminals that were represented only by morphological data, which positioned closely with their congeners in the Total Combined analyses topologies (*Chelonanthus hamatus*, *Irlbachia cardonae*, *Rogersonanthus arboreus*, *Symbolanthus argyreus*, and *S. elisabethae*). In this effort we hoped to reconstruct a better resolved topology and still maximize the number of terminals present. Additional information about the amount of missing data for each terminal in each data set is given in Supplemental material S4.

2.5. Character state reconstructions

The evolution of all morphological characters scored for this study was investigated. Reconstructions were performed with maximum parsimony using Mesquite 3.02 (Maddison and Maddison, 2014). Characters were treated as unordered and reconstructed onto the MP topology inferred from the Total Combined data (Fig. 3), but with multiple terminals representing a single species reduced to one (resulting in a tree with 52 terminals). We also replaced the polytomy involving *Symbolanthus*, *Rogersonanthus*, *Tetrapollinia* and *Calolisianthus* with a resolved clade representing the topology inferred from the BI from the Total Combined data (Supplemental material S9) and the Partial Combined data (Fig. 2). We understand that the resulting character state reconstructions will present some limitations in interpretation, not only due to this tree topology manipulation, but also due to limited sampling outside the tribe Helieae. However, these limitations do not preclude characterization of the main clades and individual lineages within the *Symbolanthus* clade, which is the focus of this study.

3. Results

3.1. Data matrices

In total, 112 morphological characters were scored for 83 terminals; 95 (84.8%) were parsimony informative. A total of 125 new DNA accessions from 52 individuals and 24 species were produced in this study, and deposited in GenBank (Table 1). Our analyses included these new accessions and also 31 previously published accessions (Table 1). The matK sequences were easily aligned due to the absence of insertions-deletions (indels). The matK matrix included 43 terminals (27 species) and 678 aligned nucleotides; 133 (19.6%) positions were variable and 57 (8.4%) of those were potentially parsimony informative. Alignments of the ITS and 5S-NTS regions were relatively straightforward, and indel characters were scored for inclusion in the analyses. The ITS matrix comprised 83 terminals (36 species), 638 aligned nucleotides and 16 coded indels; in total, 303 (46.3%) characters were variable and 220 (36.7%) were potentially parsimony informative. The 5S-NTS matrix included 47 terminals (24 species), 327 aligned nucleotides and 23 coded indels; 294 (84.0%) characters were variable and 268 (76.6%) were potentially parsimony informative. The Partial Combined matrix comprised 73 terminals (42 species) and 834 (46.7%) variable characters, 631 (35.3%) of which were potentially parsimony informative. The Total Combined matrix included 83 terminals (52 species) and 842 (46.9%) variable characters, of which 640 (35.7%) were potentially parsimony informative. Table 2 summarizes the properties of each dataset, and Supplemental material S4 gives information about the amount of missing data for each terminal in each data set. Alignments are available in Rutgers University's institutional repository, RUcore (http://dx.doi.org/10.7282/T3TF00KN).

3.2. Phylogenetic analyses

The parsimony and Bayesian approaches generally resulted in similar tree topologies, except for the Total Combined analyses (see below). The consensus topologies from the morphology and matK datasets were less resolved than the ones from nuclear DNA datasets and included only a few clades with support (Supplemental material S5 and S6). The analyses of ITS and 5S-NTS showed additionally resolved clades, and generally received stronger branch support values (Supplemental material S7 and S8). Analyses of the individual datasets were largely in topological agreement except for collapsing on poorly supported branches; therefore, we felt confident we could combine the datasets for the two simultaneous analyses.

MP searches of the Partial Combined dataset resulted in over 148,000 equally most parsimonious trees of 2048 steps (CI = 0.58, RI = 0.83). The parsimony strict consensus tree from the Partial

Combined analysis is presented in Fig. 2 with parsimony bootstrap (BS) values ⩾50% and Bayesian posterior probability (PP) values ⩾50% indicated above the branches; a dash indicates that the clade was not supported in that particular analysis. Helieae is monophyletic (BS 100, PP -), with *Prepusa montana* recovered as sister to all remaining species, which are divided in the three expected clades. One is comprised of *Macrocarpaea*, *Tachia* and *Chorisepalum* (BS 75, PP 0.96; the *Macrocarpaea* clade); the other includes all *Irlbachia* species except *I. pratensis* (BS 76, PP 1; the *Irlbachia* clade); and the larger one includes all other sampled genera (BS 71, PP 1; the *Symbolanthus* clade). Within the *Symbolanthus* clade, three main lineages are recovered in the Partial Combined analysis. Clade 1 includes *Aripuana*, *Purdieanthus*, *Irlbachia pratensis*, *Chelonanthus purpurascens*, and two species of *Calolisianthus* (*Ca. amplissimus* and *Ca.* sp. 2; BS 61, PP 1). In this clade, *Ch. purpurascens* and the two species of *Calolisianthus* are monophyletic (BS 99, PP 1), and sister to *I. pratensis* (BS 65, PP 1); these, in turn, are the sister group to the sister pair of *A. cullmaniorum* and *P. pulcher* (BS 73, PP 0.54). Clade 1 is positioned as sister to Clade 2 (BS 78, PP 1), which consists of the other 2 main clades, referred to as Clades 3 and 4. Within Clade 3 (BS 54, PP 0.99), *Rogersonanthus* forms a monophyletic group (BS 86, PP 1) that is either sister to *Symbolanthus* (MP topology) or included within *Symbolanthus* (BI topology, not shown). The clade composed of *Rogersonanthus* and *Symbolanthus* (BS 68, PP 0.99) is sister to a clade consisting of *Tetrapollinia* and the other four species of *Calolisianthus* (BS 91, PP 0.99); these four species of *Calolisianthus* form a strongly supported group (BS 100, PP 1). Clade 4 includes *Adenolisianthus*, *Helia* and all the *Chelonanthus* species, except *Ch. purpurascens* (BS 58, PP 1). *Helia* is recovered as monophyletic (BS 100, PP 1), but its exact relationships with *Adenolisianthus* and the species of *Chelonanthus* are unclear. Despite the large amount of missing data, the terminals represented by morphology data only, *Chelonanthus hamatus*, *I. cardonae*, *R. arboreus*, and *S. elisabethae*, were all positioned together with their congeners, for which molecular data was also available.

MP searches of the Total Combined dataset resulted in over 98,000 topologies of 2167 steps (CI = 0.56, RI = 0.81). The parsimony strict consensus topology of the Total Combined analysis is presented in Fig. 3 with parsimony bootstraps values ⩾50% and Bayesian posterior probabilities values ⩾50% indicated above the branches; a dash indicates that the clade was not supported in that particular analysis. The BI majority-rule consensus tree is given in Supplemental material S9. The *Macrocarpaea* (BS 73, PP 0.99), *Irlbachia* (BS 74, PP 0.90) and *Symbolanthus* (BS -, PP 0.85) clades were again recovered, however, the *Symbolanthus* clade was not strongly supported. Clades 1 (BS -, PP 0.84) and 2 (BS -, PP 0.79) were reconstructed in both MP and BI analyses, Clade 3 only emerged in the BI analysis (PP 0.58); its terminals emerged in a polytomy in the MP topology. Clade 4 was only recovered in the MP analysis, but not well supported (BS -); its terminals emerged in a polytomy in the BI topology. *Neblinantha* (BS 62, PP 0.99) was positioned as sister to the clade formed by the *Irlbachia* and *Symbolanthus* clades (MP) or in a polytomy with these 2 clades (BI); in both cases, the relationships were not strongly supported. *Celiantha* (BS 77, PP 0.95) was positioned in Clade 1, as sister to *I. pratensis* (MP), or within the *Irlbachia* clade (BI), and *Roraimaea* (BS 82, PP 1) was positioned as sister to *Aripuana* (MP) or within clade 2 (BI); none of these relationships were indeed supported. On the other hand, the positioning of *Sipapoantha* within Clade 2, and *Lagenanthus* and *Lehmanniella* emerging closely related to *Purdieanthus* (BS 84, PP 1) were similarly reconstructed in both approaches. Although several clades were not supported in the analyses of the Total Combined dataset, some groupings with strong support remained. For example, the clade composed of *Ch. purpurascens*, *Ca. amplissimus* and *Ca.* sp. 2 (BS 99, PP 0.98), *Helia* (BS 99, PP 0.98), and the clade with the other four species of *Calolisianthus* (BS 100, PP 0.99).

3.3. Character state reconstruction of key morphological and palynological characters

Most characters present CI and RI values greater than 0.5, but only eleven character states were mapped as clear synapomorphies (3, 12, 22, 30, 36, 43, 48, 72, 76, 101, 109; Supplemental material S2). Our results show that most morphological characters show some homoplasy, and that most of the clades within Helieae are defined by a suit of homoplasious character states. State reconstructions of six selected characters are given in Figs. 4 and 5. These are either characters used in previous taxonomic treatments (corolla basic color [ch. 39], pollen aggregation [ch. 80]) or promising new characters that might aid field characterization of groups (corolla persistence in fruit [ch. 47], corolla bud apex shape [ch. 49], corolla lobes with darkened tips [ch. 51], and corolla lobe margin [ch. 52]). Additional character state reconstructions that aid lineage identification can be found in Supplemental material S10.

The reconstruction of the evolution of corolla color (Fig. 4A) reveals that green, yellow, and white corollas are the plesiomorphic character state for Helieae. Our analyses show that the evolution of blue, purple, or pink corollas may have occurred early in the lineage leading to *Neblinantha*, the *Irlbachia* clade and the *Symbolanthus* clade. Red and orange corollas evolved from the blue, purple pink color in an ancestor to the clade including *Purdieanthus*, *Lagenanthus*, *Lehmanniella*, and *Roraimaea*, with *Aripuana* representing a reversal to a white corolla. Five additional reversals are suggested including the ancestor of clade 4 (*Adenolisianthus*, *Helia*, and white-green-flowered *Chelonanthus*).

Pollen grains released in tetrads are common for most species in the *Symbolanthus* clade (Fig. 4B), while monads are the ancestral condition for the *Macrocarpaea* clade, and polyads have evolved independently twice, once in the *Irlbachia* clade, and again in the clade containing *Irlbachia pratensis*, the three species of *Celiantha*, *Ch. purpurascens* and two *Calolisianthus* species.

The mapping of the evolution of corolla persistence in fruit indicates that a deciduous corolla evolved early in the history of Helieae, but reversal to persistence of the corolla in fruit occurred several times within the tribe, for example, once or twice in the clade containing *Irlbachia pratensis*, the three species of *Celiantha*, *Ch. purpurascens* and *Ca. amplissimus* and *Ca.* sp. 2, and also up to three times in the clade 4 (Fig. 4C).

Corolla bud apex tapering to a sharp point is the ancestral state for tribe Helieae, while buds with a rounded apex evolved at least six times within the tribe, including the ancestor to clade 4 (Fig. 4D)

Corolla lobes without darkened apices is a plesiomorphic trait for tribe Helieae. Corolla lobes with darkened apices may have evolved at least three times within the *Symbolanthus* clade, once in *Purdieanthus*, in the lineage containing *Ca. amplissimus* and *Ca.* sp. 2, and at least once in the lineage containing the rest of the *Symbolanthus* clade. Reversals to corolla lobes without darkened apices can be seen in the branches leading to the genera *Symbolanthus* and *Tetrapollinia* (Fig. 5A).

The optimization of the corolla lobe margin implies that entire or erose margin is the ancestral state for the tribe; the ciliate or papillose lobe margin has evolved at least seven times within the tribe, for example, in *Irlbachia* clade, in the *Ch. purpurascens*, *Ca. amplissimus* and *Ca.* sp. 2 lineage, and the *Rogersonanthus* lineage, and clade 4 (Fig. 5B).

4. Discussion

In this study, we investigated the systematics of the Helieae, with particular focus on the *Symbolanthus* clade, using phylogenetic reconstruction of molecular and morphological data. Individ-

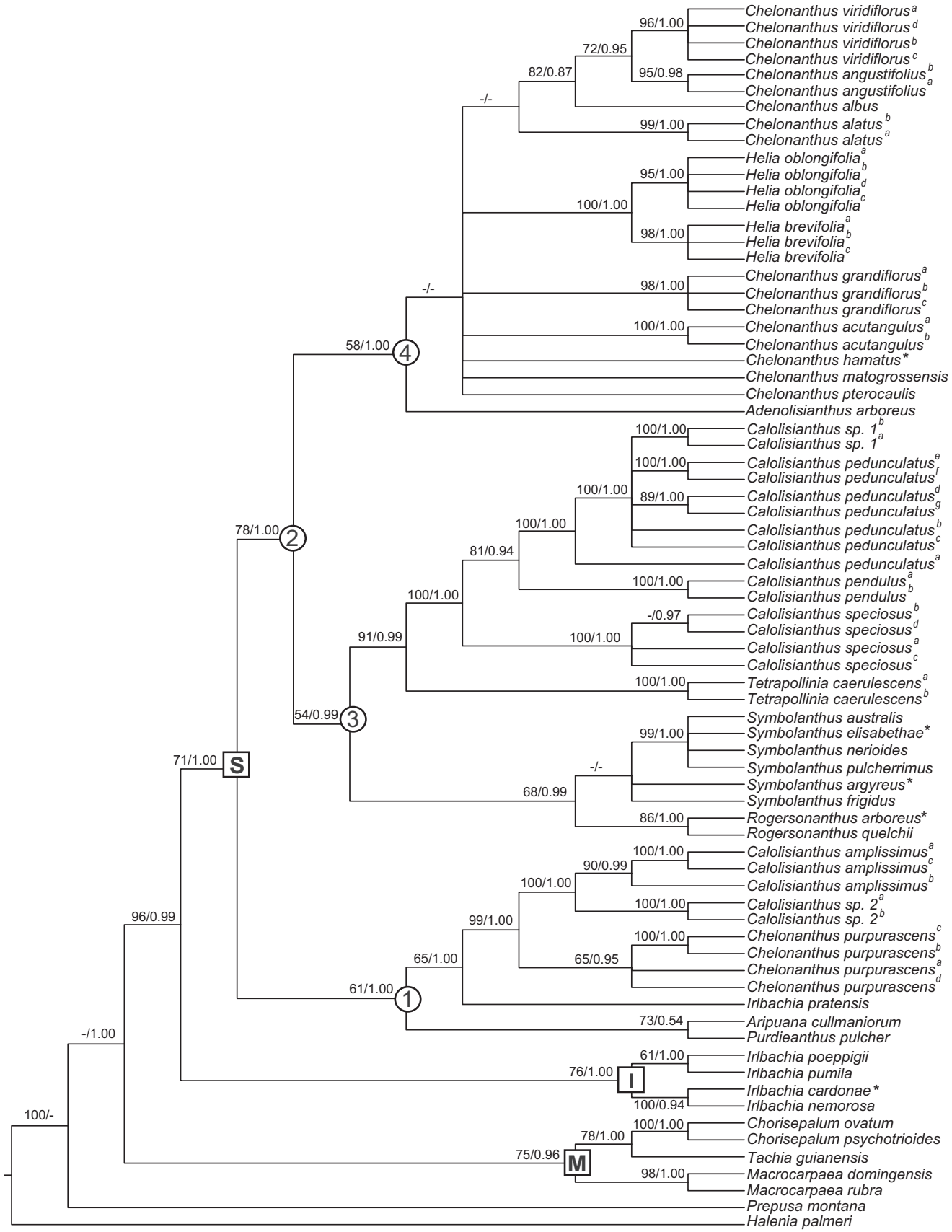

Fig. 2. Strict consensus tree from parsimony analysis of the Partial Combined dataset. Values above branches indicate parsimony bootstrap (BS)/Bayesian posterior probability (PP). A dash indicates that the clade was not supported. Labeled nodes are discussed in the text. The asterisks indicate taxa included in the analysis with morphological data only. Superscript letters close to taxa names indicate the voucher on Table 1.

Table 2
Sampling, matrix values and parsimony statistics of the separate and combined data sets. Missing data indicates the terminals (absolute number and percentage) with missing data for each data set: up to 5 (⩽5), over 5–50% (5–50), over 50–90% (50–90), over 90% (>90); an asterisk indicates the number of terminals that were not included in the analyses of the individual data sets. Variable characters and Parsimony informative characters include the indel characters.

Matrices	Terminals included	Missing data				Size	Indel characters	Variable characters	Parsimony informative characters	Most parsimonious trees	Tree length	CI	RI
		<5	5–50	50–90	>90								
Morphology	83	10 (12.0%)	73 (88.0%)	0	0	112	0	112	95 (84.8%)	1795	415	0.32	0.72
matK	43	43 (100%)	0	0	40*	678	0	133 (19.6%)	57 (8.4%)	2348	191	0.76	0.83
ITS	67	55 (82.1%)	5 (7.5%)	7 (10.4%)	16*	638	16	303 (46.3%)	220 (36.7%)	253028	660	0.63	0.84
5S-NTS	47	42 (89.4%)	5 (10.6%)	0	36*	327	23	294 (84.0%)	268 (76.6%)	3513	834	0.62	0.87
Partial Combined	73	29 (39.7%)	21 (30.1%)	20 (26.6%)	3 (4.1%)	1755	39	834 (46.7%)	631 (35.3%)	148113	2048	0.58	0.83
Total Combined	83	29 (34.9%)	21 (26.5%)	20 (22.9%)	13 (15.7%)	1755	39	842 (46.9%)	640 (35.7%)	98561	2167	0.56	0.81

ual markers and morphological datasets were analyzed separately and in combination. The consensus topologies from the morphology dataset were not well-resolved and the clades received little or no support, despite its high number of parsimony informative characters (84.8%). The low CI (0.32), however, indicates a high amount of homoplasy in this dataset, and to some extent partially explains the historical difficulties in circumscribing taxa based solely on morphological data. The consensus topologies from the matK dataset only included a few supported clades, which agreed with the low proportion of parsimony informative sites found within this dataset (8.4%). The analyses of ITS and 5S-NTS included more resolved clades, generally with stronger branch support values, although both datasets differ greatly in the percentages of informative characters (36.7% and 76.6%, respectively).

Analyses of the Partial Combined dataset (all data partitions and 73 terminals) resulted in moderately-supported trees, while the Total Combined dataset (all data partitions and 83 terminals) gave less resolved topologies, lacking support for several clades. Although both topologies differ only in ten taxa, the inclusion of these additional taxa increased the total amount of missing data from 30.5 to 37.3%, when comparing the Partial and the Total Combined datasets, respectively. However, it is not just the amount of missing data present in the Total Combined analyses that explains its lower support; the distribution of missing data cells is likely very important in this scenario, as demonstrated in several studies focusing on the relative importance of missing data in phylogenetic studies (Wiens, 2003; Wiens and Morrill, 2011). In this larger dataset, more species lacking nuclear DNA data, which were much more informative than the chloroplast data, were included. However, it is interesting to note that four species that totally lacked DNA data emerged intermingled with their congeners represented by more complete data (*Chelonanthus hamatus*, *Irlbachia cardonae*, *Rogersonanthus arboreus*, and *Symbolanthus elisabethae*). Comparisons between the topologies of the combined analyses show trees with congruent, or at least not conflicting, overall topologies.

The taxonomic chaos within the Helieae has been known for at least a few decades (see Struwe et al. (2009) and references therein). The present work, in addition to supporting previously suggested changes, provides further evidence for new circumscriptions for *Adenolisianthus*, *Calolisianthus*, *Chelonanthus* and *Helia*, which, in some cases, are very different from the traditional and current definitions. We characterize here the three main groups resolved within the *Symbolanthus* clade, and discuss the inferred systematic relationships. Unless otherwise indicated, the phylogenetic hypotheses referred to are those from the Partial (Fig. 2) and Total Combined data sets (Fig. 3, Supplemental material S9).

4.1. Clade 1

4.1.1. Aripuana, Lagenanthus, Lehmanniella, Purdieanthus, Roraimaea

In our analysis three major lineages emerged within the *Symbolanthus* clade. In the Total Combined topology (Fig. 3), Clade 1 contains two main subclades, albeit with some branches not well supported. One subclade is composed of *Aripuana cullmaniorum*, *Lagenanthus princeps*, *Lehmanniella splendens*, *Purdieanthus pulcher*, and the two species of *Roraimaea*. This particular lineage received only low to moderate support (Figs. 2 and 3), but this was most likely due to the large amount of unknown character states representing four of the six terminals included by morphological data only. However, all six genera share long tubular flowers that are red-orange in color with the exception of *Aripuana*, which has white corollas. Outside this clade, the only species within Helieae with red corollas is *Calolisianthus pedunculatus* (Clade 3).

The close relationship between *Lagenanthus*, *Lehmanniella and Purdieanthus* was also supported by Struwe et al. (2009), who defined the three as separate genera, and Simonis (in Maas (1985)), who circumscribed the three taxa under *Lehmanniella*. Unfortunately, the increased sampling of the current study was not able to include *Lehmanniella huanucensis* J.E. Simonis *ex* P.J.M. Maas, and therefore cannot confirm the monophyletic status of that genus. It remains to be seen if the evolutionary history of the three genera is most accurately reflected by their inclusion into *Lehmanniella sensu* Simonis, or by retaining separate genera as viewed by Struwe et al. (2002, 2009). As mentioned above the three genera share very similar floral morphologies, and the similarities in seed characters supports a close relationship (Bouman et al., 2002), but their pollen characters (e.g., reticulum size; ch. 96) are distinct and do not support a broadly defined *Lehmanniella* (Nilsson, 2002).

The monophyly of *Roraimaea* was well supported in this study (Fig. 3) and congruent with the analyses of Struwe et al. (2009). The sister relationship between *Aripuana* and *Roraimaea*, which was suggested by Struwe et al. (2009), was supported by the Total Combined parsimony topology from the current study, but that relationship received no branch support. The Bayesian topology (Supplemental material S9) placed *Roraimaea* in Clade 2 as part of a large polytomy, including the *Chelonanthus* species with green-white colored corollas and *Helia*. As stated by Struwe et al. (2009), *Aripuana* and *Roraimaea* share several potential synapomorphies such as actinomorphic stamens and style and erect fruit that apically dehisce (ch. 54, 70, 73, 74). These traits can be found elsewhere in the tribe, but are not typical. Molecular data was not available for either species of *Roraimaea*; obtaining DNA data from

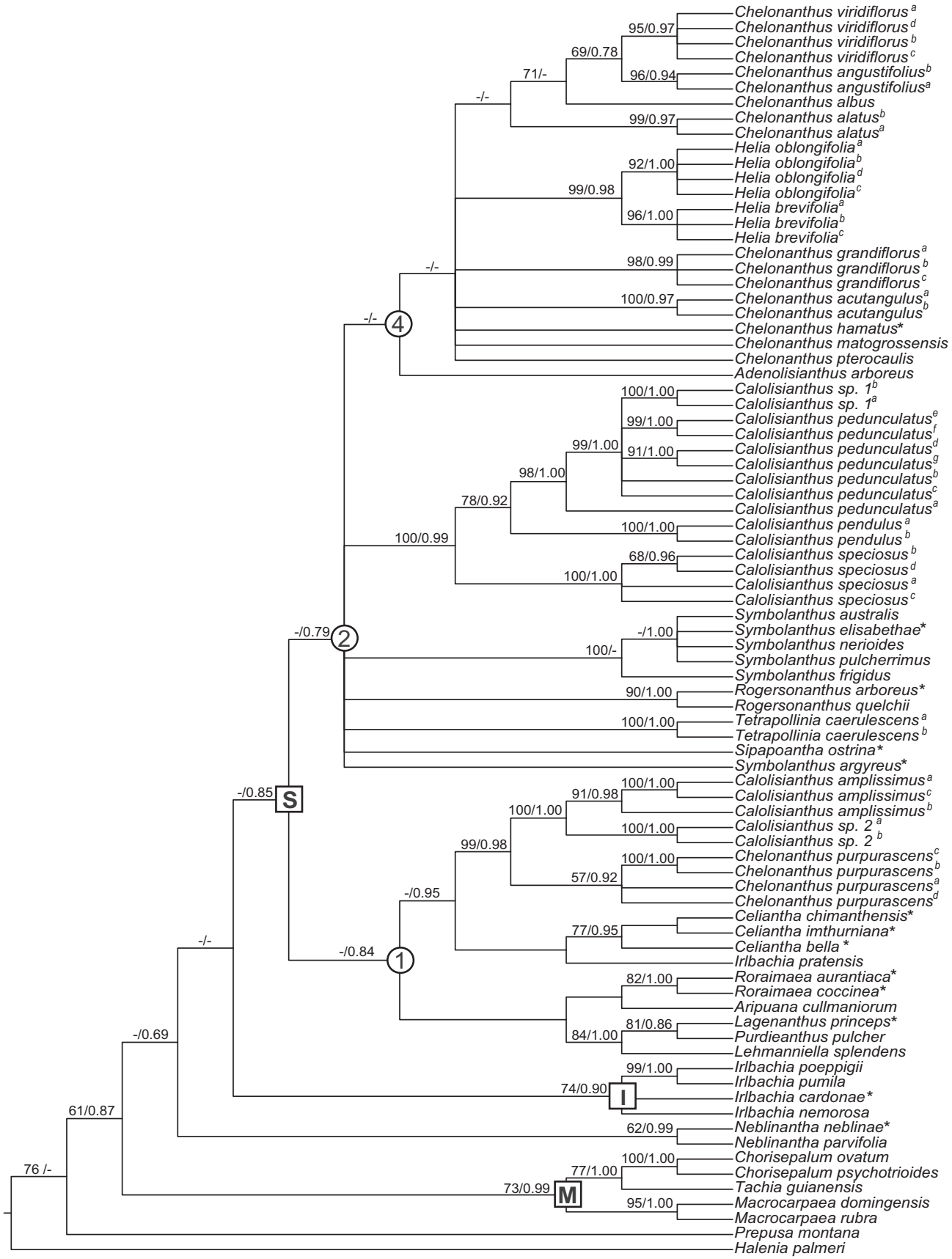

Fig. 3. Strict consensus tree from parsimony analysis of the Total Combined dataset. Values above branches indicate parsimony bootstrap (BS)/ Bayesian posterior probability (PP). A dash indicates that the clade was not supported. Labeled nodes are discussed in the text. The asterisks indicate taxa included in the analysis with morphological data only. Superscript letters close to taxa names indicate the voucher in Table 1.

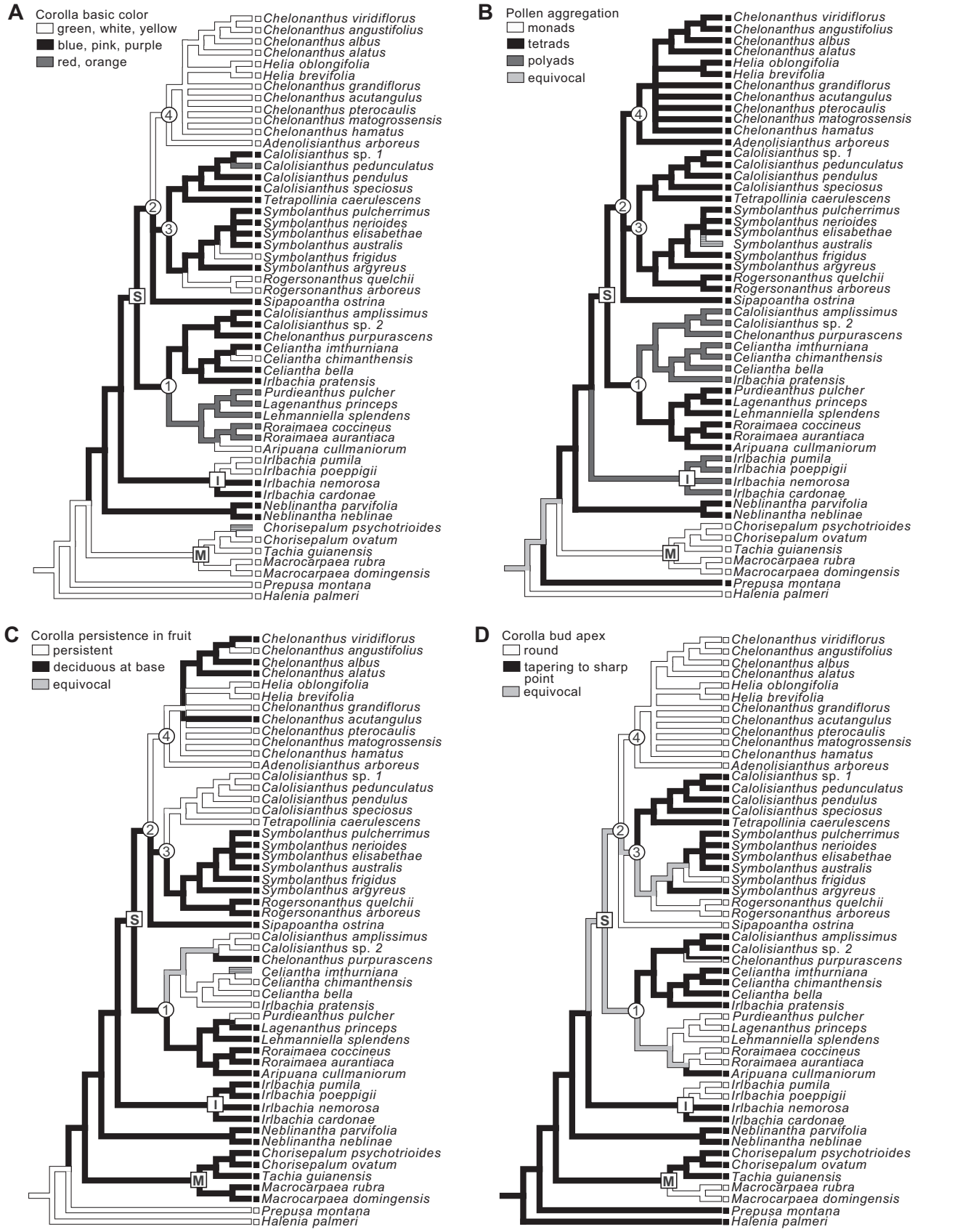

Fig. 4. Ancestral character reconstructions for selected morphological characters based on the parsimony tree inference of the Total Combined dataset (taxa with multiple assessments were removed and Clade 3 from the Bayesian Total Combined and Partial Combined analyses imposed on the polytomies involving the appropriate taxa). (A) Corolla basic color. (B) Pollen aggregation when released. (C) Corolla persistence in fruit. (D) Corolla bud apex shape. The color gray with stripes indicates terminals with missing data.

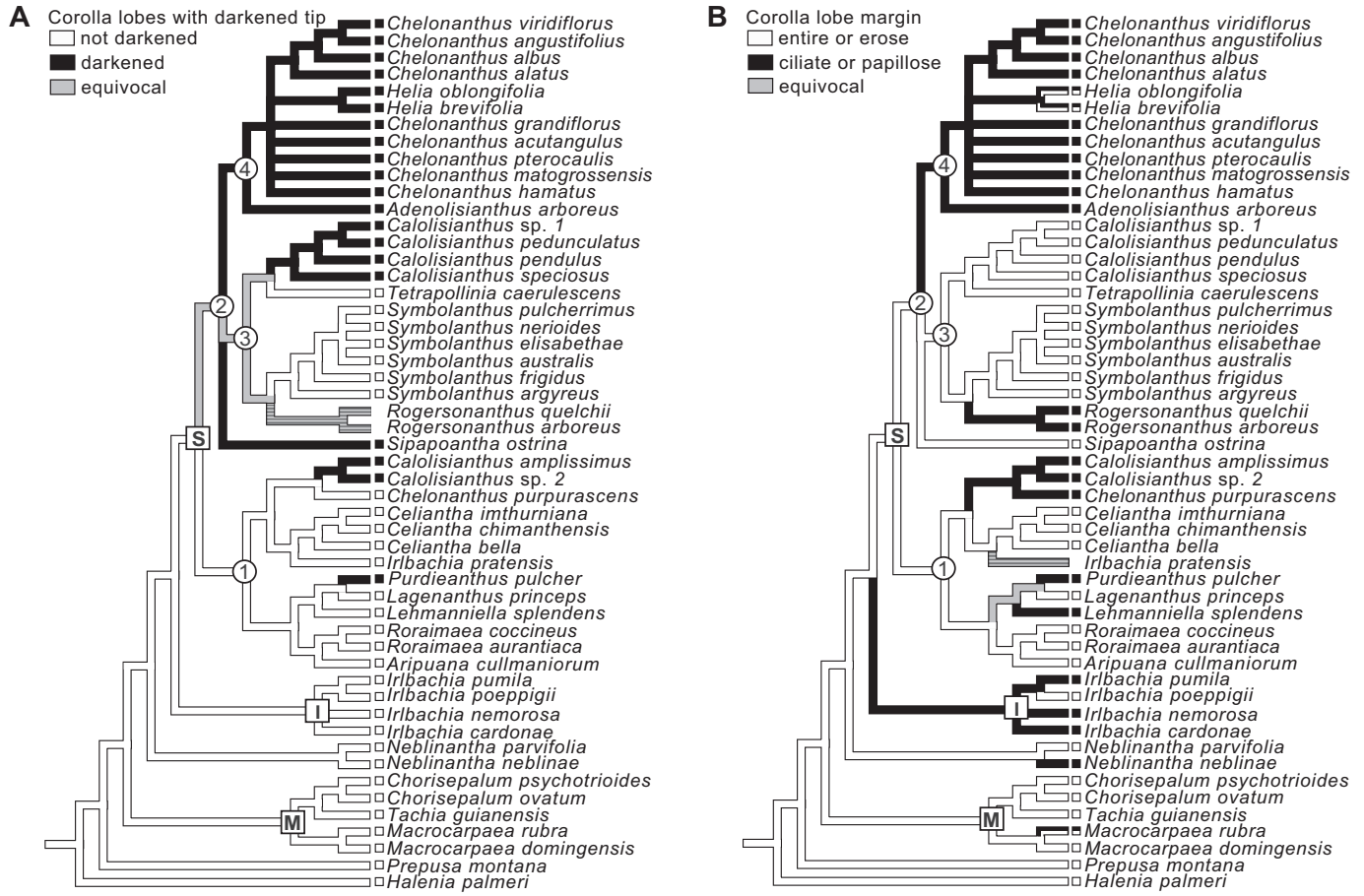

Fig. 5. Ancestral character reconstructions for selected morphological characters based on the parsimony tree inference of the Total Combined dataset (taxa with multiple assessments were removed and Clade 3 from the Bayesian Total Combined and Partial Combined analyses imposed on the polytomies involving the appropriate taxa). (A) Corolla lobes darkness at apex. (B) Corolla lobes margin. The color gray with stripes indicates terminals with missing data.

this genus would be helpful in validating the potential sister relationship between the two genera.

4.1.2. Calolisianthus, Celiantha, Chelonanthus purpurascens, Irbachia pratensis

The other subclade of Clade 1 received strong support using the Bayesian inference in both Partial and Total Combined analyses, but was not well supported using parsimony (Figs. 2 and 3). Based on the Total Combined parsimony topology (Fig. 3), *Irlbachia pratensis* and the three species of *Celiantha* were positioned as sister to the subclade containing *Chelonanthus purpurascens*, and two species of *Calolisianthus* (the other four species of *Calolisianthus* emerged in Clade 4). The sister relationship of *Celiantha* and *Irlbachia pratensis* was not well supported nor was the inclusion of *Celiantha* within this subclade in general. The Total Combined Bayesian topology (Supplemental material S9) that strongly supported this subclade did so with the exclusion of *Celiantha*. In that analysis, this genus was weakly placed as sister to the *Irlbachia* clade. Struwe et al. (2009) also had trouble definitively placing *Celiantha*, but concluded the taxon may represent a basal lineage, along with *Prepusa* and *Senaea* as sisters to the rest the tribe. That position was not supported here. *Celiantha* lacks a glandular area on the calyx lobes and medially dehiscent capsules (ch. 33, 74), traits considered typical of Helieae, but aside from those traits, *Celiantha* has a flower morphology consistent with the tribe (Struwe et al., 2009). The continued inability to place this genus with any certainty stresses the need for molecular data representing this group. Nonetheless, all genera in this subclade are characterized by pollen shed as polyads (ch. 80). This attribute is also typical of the *Irlbachia* clade, but absent in other members of the Helieae.

Although the current study supports the close relationship between *Chelonanthus purpurascens* and *Irlbachia pratensis* demonstrated before (Struwe et al., 2002, 2009), our sampling efforts were unable to add accessions from other *I. pratensis* specimens to the data previously analyzed. Therefore we cannot improve on the phylogenetic understanding of this species, but will provide below morphological evidence that supports this relationship. Lepis (2009) previously refuted the monophyly of *Calolisianthus*; a conclusion strongly supported here with complete sampling of all species in the three genera mentioned. The relationship between *Ch. purpurascens*, *Ca. amplissimus* and *Ca.* sp. 2 received strong support in both combined analyses (Figs. 2 and 3), and in some of the analyses of individual datasets (Supplemental material S6 and S7). *Irlbachia pratensis* as sister to *Ch. purpurascens*, *Ca. amplissimus* and *Ca.* sp. 2 was strongly supported in the Bayesian analysis of the Partial and Total Combined datasets (Fig. 2, Supplemental material S9), but this relationship is not as clear using parsimony. There are no morphological traits exclusive to this subclade, but some attributes support a close relationship between the two species of *Calolisianthus*, *Ch. purpurascens* and *Irbachia pratensis* and a more distant relationship between *Ca. amplissimus*/*Ca.* sp. 2 and the other species of *Calolisianthus*. For example, cilliate or pappillose calyx and corolla lobe margins (ch. 35 and 52) are characteristic of *Ch. purpurascens*, *Ca. amplissimus*, *Ca.* sp. 2 and *Irlbachia pratensis*, as well as other species of Helieae; however, the clade containing the other four species of *Calolisianthus* are characterized by entire to erose calyx and corolla lobe margins. In

addition, the four species of *Calolisianthus* in Clade 3 have filaments that are filiform in cross-section (ch. 63), while the filaments of *Ca. amplissimus* and *Ca.* sp. 2 (as well as the rest of their subclade) have filaments that are flattened. The polyad pollen grains (ch. 80) with exine consisting of muri forming loops (ch. 83) that *Ca. amplissimus* and *Ca.* sp. 2 share with *Ch. purpurascens* and *I. pratensis* also suggest the need to segregate the two *Calolisianthus* species from the other four that shed their pollen as tetrads and lack such loops. The only other genus in Helieae with muri that form loops is the tetrad bearing *Neblinantha*.

The monophyly of *Chelonanthus* had previously been refuted by analyses depicting two lineages; one positioning *Ch. purpurascens* as closely allied with *Calolisianthus amplissimus* and *Irlbachia pratensis* and the lineage containing the rest of *Chelonanthus* closely related to *Helia* and *Adenolisianthus* (Lepis, 2009; Struwe et al., 2002, 2009). The current study, which sampled all recognized species of *Chelonanthus*, including two newly described species (*Chelonanthus hamatus* and *Ch. pterocaulis*; Lepis et al., 2014), also supports this split. As with the case of *Calolisianthus*, the pollen of *Chelonanthus purpurascens* (polyads with exine ornamentation forming loops) is very different from the other species of *Chelonanthus* (pollen grains united into tetrads and ornamentation lacking looped processes). Additionally, the rest of *Chelonanthus* has an exine that is thickest at the equator (ch. 84), while *Ch. purpurascens* (and the rest of its subclade) lack this thickening. Other macroscopic characters that support this split include the blue to purple corollas (ch. 39) that emerge from buds with a tapering point (ch. 49) found in *Ch. purpurascens* compared to the green, white to yellow corollas that emerge from flower buds with a blunt apex that is found in the rest of the genus (hence forth referred to as green-white-flowered *Chelonanthus*). Also the corolla lobes in the green-white-flowered *Chelonanthus* have a darkened spot at each apex (ch. 51), which is lacking in *Ch. purpurascens*. Potential synapomorphic characters shared between green-white-flowered *Chelonanthus* and the rest of Clade 4 will be discussed in Section 4.3.

Chelonanthus purpurascens is the type species of *Chelonanthus*. The phylogenetic relationship between *Ch. purpurascens*, *Ca. amplissimus* and *Ca.* sp. 2, along with the clear morphological distinction between these three species and the other, distantly related species of *Calolisianthus* and the green-white-flowered *Chelonanthus* form the basis for the formal transference of *Ca. amplissimus* and *Ca.* sp. 2 to *Chelonanthus* and the inclusion of the green-white-flowered *Chelonanthus* and *Adenolisianthus* in *Helia* (more on that in Section 4.3). New taxonomic combinations and synonymies are currently being prepared.

4.2. Clade 3

Clade 3, composed of four species of *Calolisianthus* (*Ca. pedunculatus*, *Ca. pendulus*, *Ca. speciosus* and *Ca.* sp. 1), the two species of *Rogersonanthus*, six species of *Symbolanthus* and *Tetrapollinia caerulescens*, was recovered in the Partial Combined analyses using parsimony and Bayesian inferences (Fig. 2), but in the Total Combined analyses this clade was recovered only on the Bayesian topology (Supplemental material S9). The parsimony topology (Fig. 3) depicts the four genera in a polytomy that also includes *Sipapoantha ostrina* and Clade 4. Therefore, discussions about this clade are focused on the three topologies that support it as a distinct lineage. Ancestral reconstruction of morphological characters resulted in equivocal or uncertain reconstruction at the base of this clade (e.g., Figs. 4C, 5A, Supplementay material S10 – ch. 46, 49, 95); however, a few attributes are shared between their members, and seem to be good for a general characterization of the clade. These attributes include blue-, purple-, pink-colored corollas (ch. 39) that are fleshy or coriaceous in texture (ch. 46), corolla lobe margins that are entire to erose (ch. 52), and filaments that are cylindrical (filiform) in shape (ch. 63).

4.2.1. Rogersonanthus and Symbolanthus

Prior to this study, *Rogersonanthus* has been included in phylogenetic analyses based solely on morphological data and emerged nested among *Adenolisianthus*, *Helia*, and the green-white-flowered *Chelonanthus* (Struwe et al., 2009). The current study represented *Rogersonanthus* with molecular data (ITS and matK) in addition to the morphological characters. This combination, for the first time, suggests a close relationship between *Rogersonanthus* and *Symbolanthus* (Fig. 2). Gould and Struwe (2004) distinguished *Symbolanthus* from the rest of the tribe Helieae based on the presence of a staminal corona or corona-like structure found within the corolla (ch. 58) and *Rogersonanthus* shares this trait with *Symbolanthus*. The only other known species in the family to have this structure is the newly described species, *Chelonanthus hamatus* (Lepis et al., 2014). Although Nilsson (2002) thought a relationship between *Rogersonanthus* and *Symbolanthus* justified based on pollen characters, *Rogersonanthus* has been characterized as having *Chelonoides*-type pollen, the pollen type found in the green-white-flowered *Chelonanthus* and *Adenolisianthus*. *Rogersonanthus* shares a woody habit (ch. 1) with most of *Symbolanthus* (except *S. argyreus* and *S. frigidus*), whereas this trait is nearly absent from Clade 4 (*Adenolisianthus* the exception). *Rogersonanthus* also shares leaf margins that are chartaceous or coriaceous (ch. 11) with the rest of *Symbolanthus* (*S. argyreus* and *S. frigidus* the exception), while species in Clade 4 all have leaf margins that are membranaceous and hyaline.

As suggested by the analysis based solely on morphology (Supplemental material S5), *Rogersonanthus* and Clade 4 share characters that are not found in *Symbolanthus*. One such character is corolla color, which is white, green to yellow in *Rogersonanthus* and Clade 4, but is blue, pink to purple in *Symbolanthus* (except *S. frigidus* with green to yellow corollas). *Rogersonanthus* and Clade 4 share corolla lobe margins that are ciliate or papillose (ch. 52), while in *Symbolanthus* the margins are fringed. Additionally, *Rogersonanthus* shares a filament that is flattened in cross-section (ch. 63) with Clade 4, while *Symbolanthus* and the rest of Clade 3 have filiform filaments (except *S. argyreus*). Lastly, *Rogersonanthus* and Clade 4 have medially dehiscent capsules (ch. 67), a trait typical to the tribe, but in contrast *Symbolanthus* has apically dehiscent capsules (medially in *S. frigidus*).

Clearly *Rogersonanthus* shares traits with *Symbolanthus* and Clade 4 and with most of the character comparisons there are exceptions that usually involve *S. frigidus* and *S. argyreus* as outliers from *Symbolanthus* and *Adenolisianthus* as the outlier from Clade 4. When Gould and Struwe (2004) included *S. frigidus* and *S. argyreus* in *Symbolanthus* (previously *Wurdackanthus*) they recovered *S. frigidus* as the sister to the remainder of the genus, and our data suggests *Adenolisianthus* as a possible sister to the rest of Clade 4. The inclusion of the more rapidly evolving 5S-NTS region for *Rogersonanthus* may be helpful in clarifying the evolutionary history of this genus.

4.2.2. Calolisianthus and Tetrapollinia

Our results revealed that *Calolisianthus* as currently circumscribed is not monophyletic; two species emerged within Clade 1 (see Section 4.1 for further discussion), while the other four species are strongly supported as members of Clade 3. These four species formed a clade when the Partial and the Total Combined matrices were analyzed, but also in the trees resulting from the analyses of individual datasets (molecular and morphological). These four species of *Calolisianthus* as a separate lineage from the *Calolisianthus* species allied with *Ch. purpurascens* is strongly supported in all analyses (Figs. 2 and 3; Supplemental material S9). The morphological characters that are shared between these four species from

Clade 3 are prominent, raised veins on the abaxial side of the leaves (ch. 20), corolla lobes with darkened tips (ch. 51), pollen with an exine that is thickest at the polar areas (ch. 85), and a reticulum that has the coarsest mesh pattern near the equator (ch. 96). In addition, field characteristics that segregate the four species that emerged in Clade 3 from those of Clade 1, respectively, are subshrubs to suffructescent herbs (vs. herbaceous and never woody at base), the ribbed to slightly winged stems (vs. strongly winged stems), and coriaceous leaves (vs. membranous leaves; these characters were not included in the phylogenetic analyses due to our inability to consistently code them for all herbarium material). The taxonomic revision of the newly circumscribed *Calolisianthus*, including only these four species that emerged within Clade 3, is being produced.

The sister relationship between the four species of *Calolisianthus* and *Tetrapollinia* is well supported in the Partial Combined analyses (Fig. 2) and congruent with previous findings (Struwe et al., 2009). In addition to the characters that unite Clade 3, the *Tetrapollinia/Calolisianthus* subclade shares corollas that persist in fruit (ch. 47); this trait is found elsewhere in the *Symbolanthus* clade, but is unique within Clade 3. Unlike Clade 3 (except *Sipapoantha*), the *Tetrapollinia/Calolisianthus* subclade lacks a thickening of the exine near the equator (ch. 84). Morphology that supports *Tetrapollinia* as a separate genus include the unique pollen grains ornamented with spines that inspired Maguire and Boom (1989) to distinguish this taxon as its own genus from *Irlbachia* (Struwe et al., 2002). *Tetrapollinia* is also unique in lacking the tribal trait of a nectariferous disk at the base of ovary. Additionally, the texture of the corolla in *Tetrapollinia* is thin in contrast to the thick and coriaceous corolla found in *Calolisianthus* (ch. 46). *Tetrapollinia* also lacks the darkend spot on the apex of each *Calolisianthus* corolla lobe (although this trait varies throughout the tribe; ch. 51). Lastly, like *Symbolanthus*, *Tetrapollinia* has fruit that open apically compared to the medially dehiscing fruit of *Calolisianthus* (and much of the rest of the tribe). The close relationship between the four species of *Calolisianthus* and *Tetrapollinia* is well supported, but it is clear the generic distinctions between the two should remain.

4.2.3. Sipapoantha

Sipapoantha ostrina was included in this study based solely on morphological data and placed within the *Symbolanthus* subclade with moderate support and in Clade 2 as part of a polytomy, but not well supported (Fig. 3, Supplemental material S9). *Sipapoantha ostrina* is one of two species in this rare genus endemic to the Guayana Highlands that is represented in the scientific record by only a handful of herbarium specimens. *Sipapoantha* is characterized by coriaceous leaves with strongly revolute leaf margins that are sulfur yellow when dry and blue corollas (Lepis et al., 2011). The second species, *Sipapoantha obtusisepala* Lepis, Maas and Struwe, was not included in this study because of the exceedingly large amount of unknown character states as it is known from a single herbarium sample that lacks flowers (Lepis et al., 2011). Struwe et al. (2009) study tentatively placed *Sipapoantha ostrina* as part of a polytomy including the *Macrocarpaea* and *Symbolanthus* subclades and discussed the possibility of the genus residing one step up in Helieae from the basal genus *Prepusa*, but our results do not support this scenario. The species of *Sipapoantha* exist in an understudied region of the Neotropics (Lepis et al., 2011) and more information about this genus is needed to determine its role in the evolutionary history of Helieae.

4.3. Clade 4: Adenolisianthus, Chelonanthus, Helia

Previous studies supported a close relationship between *Adenolisianthus*, *Helia*, and the green-white-flowered *Chelonanthus* (Lepis, 2009; Struwe et al., 2002, 2009). The current study employed a broader sampling scheme that included all recognized species within the three genera, multiple accessions representing the variability that exists in the more common widespread taxa and additional data partitions; the results were congruent with previous findings. Clade 4 was strongly supported using the Bayesian inference in the Partial Combined analysis, while the parsimony topology revealed the group, the branch support was weak (Fig. 2). Similarly, in the Total Combined analyses parsimony revealed this lineage, but without branch support (Fig. 3) and the Bayesian topology divided this group among a large polytomy involving much of Clade 2 (Supplemental material S9). As with most other clades in Helieae, potential synapomorphies exists, but they are not traits unique to this lineage (Struwe et al., 2009). You can however, use the following list of traits in tandem as field characters to identify this clade. These include corollas that are green, white to yellow in color (ch. 39), with a dark spot on the apex of each corolla lobe (ch. 51), corolla lobe margins that are ciliate or papillose (ch. 52; *Helia* can be dimorphic with margins that are also entire), flower buds that are blunt or rounded at the apex (ch. 49), and membranaceous and hyaline leaf margins (ch. 11). Although not well suited for field identification, pollen exine that has a coarser reticulum around the equator of each pollen grain (ch. 96), and muri that is unevenly thickend (ch. 95) also help to characterize this clade. Generally speaking, pollen characters support the combination of these three genera under a single name. For instance, Nilsson (2002) considered the pollen of the green-white-flowered *Chelonanthus* (Chelonoides-type) and *Adenolisianthus* to be similar enough to reduce *Adenolisianthus* to a synonym of *Chelonanthus* and others came to similar conclusions (Maas, 1985; Maguire and Boom, 1989). *Helia* also has pollen very similar to the Chelonoides-type and, although distinct, was considered by Nilsson (1970, 2002) to be within the realm of variation observed for that pollen type.

The relationships within Clade 4 are not clearly resolved, but our analyses weakly suggest *Adenolisianthus* as sister to the rest of the clade (Figs. 2 and 3). Several traits segregate *Adenolisianthus* from the rest. One obvious difference is the woody habit of *Adenolisianthus* while all other species in Clade 4 are small herbs or larger herbaceous plants that are suffrutescent at the base (ch. 1). Struwe et al. (2009) discusses the evolution of the woody habit within Gentianaceae and concluded the herbaceous state was plesiomorphic for the family and the tribe with woodiness evolving at least nine separate times throughout Helieae's history. Additional characters that can be used to differentiate *Adenolisianthus* include a lack in prominence or raised veins on the abaxial side of the leaf (ch. 20). All other species in Clade 4 have raised veins, although *Helia* is dimorphic. Based on our data, raised veins on the underside of the leaf appears to be a derived trait that has evolved several times throughout the tribe (Supplemental material S10, ch. 20). *Adenolisianthus* is also the only taxon in Clade 4 with filaments that are winged at the base (ch. 59). Molecular data for *Adenolisianthus* was only represented by ITS and perhaps the inclusion of a more quickly evolving region like 5S-NTS as well as the reduction in the number of unknown character states that would result could help solidify *Adenolisiathus'* placement within Clade 4.

The type species of *Chelonanthus* is *Ch. purpurascens*, which is not closely related to the species in Clade 4 as discussed in Section 4.1. This means that all species of *Chelonanthus* that belong to Clade 4 need to be renamed. Considering that the oldest genus name available for identifying this particular lineage is *Helia*, we are preparing the taxonomic revision of *Helia* under a broader circumscription that will include the species that are currently recognized as *Helia*, the green-white-flowered species of *Chelonanthus* and *Adenolisianthus arboreus* (i.e., Clade 4). New combinations under the name *Helia* are necessary for just a few, recently described species because several species of the tribe Helieae had been previously treated under the name *Helia* by Kuntze (1891).

5. Conclusions

This study greatly advances the understanding of the phylogenetic history of Helieae and in particular the *Symbolanthus* clade. With increased taxon sampling and additional data partitions, support for novel phylogenetic relationships are indicated, such as the clade formed by *Ch. purpuracens*, *Ca. amplissimus*, and *Ca.* sp. *2* and the non-monophyly of *Calolisianthus*. We also provide additional support for the polyphyly of *Chelonanthus*, as well as for the close relationship between *Adenolisianthus*, *Helia* and the green-white-flowered *Chelonanthus*. Some morphological and palynological characters that had been previously used for classification, such as flower color and pollen aggregation, are indeed useful for characterizing the main clades. Further work based on broader sampling (specially within the genus *Symbolanthus*), more molecular, morphological, anatomical and palynological data is still needed. However, the evidence already gathered provides support for a more stable taxonomic classification for the genera listed above and new taxonomic combinations and synonymies are currently being prepared.

Acknowledgements

We acknowledge funding provided by FAPESP (grant #2003/10918-3) and by IAPT Research Grant Program 2006 to MFC, by Rutgers University, Department of Plant Biology and Pathology Travel Grants 2003 & 2005 and Rutgers University, Graduate School of New Brunswick Travel Grants 2003 & 2005 to KL, by CNPq (305776/2010-0) to JRP, and by USDA/Hatch NJ17112, USDA-Hatch NJ17145, and National Science Foundation NSF-DEB 317612 to LS. This study is part of the two first authors PhD theses through the Departamento de Botânica – IBUSP (MFC) and through Department of Plant Biology and Pathology – Rutgers University (KL). We thank all herbaria cited for access to material and information; colleagues who have provided new specimens, silica gel dried leaves, and photos for this study; Dr. Mariana Cabral de Oliveira for granting access to the molecular laboratory in the "Laboratório de Algas Marinhas Édison José de Paula" (IBUSP) and for laboratory assistance; Paul J.M. Maas for access to research material and unpublished manuscripts. We also thank the members of both theses committees and the reviewers for suggestions and constructive comments that greatly improved this manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2016.09.013.

References

Bentham, G., 1876. Gentianeae. In: Bentham, G., Hooker, J. (Eds.), Genera Plantarum, vol. 2(2). L. Reeve & Co., Williams & Norgate, London, pp. 799–820.
Bouman, F., Cobb, L., Devente, N., Goethals, V., Maas, P.J.M., Smets, E., 2002. The seeds of Gentianaceae. In: Struwe, L., Albert, V.A. (Eds.), Gentianaceae – Systematics and Natural History. Cambridge University Press, Cambridge, pp. 498–572.
Calió, M.F., Pirani, J.R., Struwe, L., 2008. Morphology-based phylogeny and revision of *Prepusa* and *Senaea* (Gentianaceae: Helieae): rare endemics from eastern Brazil. Kew. Bull. 63, 169–191.
Cox, A.V., Bennet, M.D., Dyer, T.A., 1992. Use of the polymerase chain reaction to detect spacer size heterogeneity in plant 5S-rRNA gene clusters and to locate such clusters in wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 83, 684–690.
Dolphin, K., Belshaw, R., Orme, C.D.L., Quicke, D.L.J., 2000. Noise and incongruence: interpreting results of the incongruence length difference test. Mol. Phylogenet. Evol. 17, 401–406.
Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1995. Testing significance of incongruence. Cladistics 10, 315–319.
Felsenstein, J., 1985. Confidence-limits on phylogenies – an approach using the bootstrap. Evolution 39, 753–779.
Gilg, E., 1895. Gentianaceae. In: Engler, A., Prantl, K. (Eds.), Die Natürlichen Pflanzenfamilien, vol. 4(2). Verlag von Wilhelm Engelmann, Leipizig, pp. 50–108.
Gould, K.R., Struwe, L., 2004. Phylogeny and evolution of *Symbolanthus* and *Wurdackanthus* (Gentianaceae – Helieae) in the Guayana Highlands and Andes, based on ribosomal 5S-NTS sequences. Ann. MO. Bot. Garden 91, 438–446.
Grant, J.R., Maas, P.J.M., Struwe, L., 2006. *Yanomamua araca* (Gentianaceae), a new genus and species from Serra do Araçá, an outlier of the Guayana Region in Amazonas state, Brazil. Harv. Pap. Bot. 11, 29–37.
Grisebach, A.H.R., 1839. Genera et Species Gentianearum. J.G. Cotta, Stuttgart and Tübingen.
Grisebach, A.H.R., 1845. Gentianaceae. In: de Candolle, A. (Ed.), Prodromus systematis naturalis regni vegetabilis, vol. 9. Fortin, Masson, et Sociorum, Paris, pp. 39–141.
Hagen, K.B. von, Kadereit, J.W., 2001. The phylogeny of *Gentianella* (Gentianaceae) and its colonization of the southern hemisphere as revealed by nuclear and chloroplast DNA sequence variation. Org. Div. Evol. 1, 61–79.
Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.
Kuntze, C.E.O., 1891. Revisio Generum Plantarum, vol. 2. Arthur Felix, Leipzig.
Lee, M.S.Y., 2001. Uninformative characters and apparent conflict between molecules and morphology. Mol. Phylogenet. Evol. 18, 676–680.
Lepis, K.B., 2009. Evolution and Systematics of Chelonanthus (Gentianaceae) Ph.D. dissertation. Rutgers University, New Brunswick.
Lepis, K.B., Maas, P.J.M., Struwe, L., 2011. A new species of *Sipapoantha* (Gentianaceae: Helieae) from northern Brazil. Blumea 56, 28–32.
Lepis, K.B., Palmer, M., Cobb, L., 2014. Chelonanthus Treatment. In: de Oliveira, S.M. (Ed.), Flora of the Guianas, 139. Gentianaceae, Series A: Fascicle 30. Royal Botanic Garden, Kew, Richmond, U.K, pp. 7–22.
Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925.
Maas, P.J.M., 1985. Nomenclatural notes on neotropical Lisyantheae (Gentianaceae). Proc. Kon. Ned. Akad. Wetensch. Ser. C 88, 405–412.
Machado, I.C.S., Sazima, I., Sazima, M., 1998. Bat pollination of the terrestrial herb *Irlbachia alata* (Gentianaceae) in northeastern Brazil. Plant Syst. Evol. 209, 231–237.
Maddison, D.R., Maddison, W.P., 2005. MacClade 4: Analysis of phylogeny and character evolution. Version 4.08a. <http://macclade.org>.
Maddison, W.P., Maddison, D.R., 2014. Mesquite: a modular system for evolutionary analysis. Version 3.02 <http://mesquiteproject.org>.
Maguire, B., 1981. Gentianaceae. In: Maguire, B. et al. (Eds.), The Botany of the Guayana Highland Part XIII, Mem. New York Bot. Gard., vol. 51, pp. 2–56.
Maguire, B., 1985. Gentianaceae – Part 2. Phytologia 57, 311–312.
Maguire, B., Boom, B.M., 1989. Gentianaceae (Part 3). In: Maguire, B. et al. (Eds.), The Botany of the Guayana Highland Part XI, Mem. New York Bot. Gard., vol. 32, pp. 330–388.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, Louisiana, pp. 1–8.
Nickrent, D.L., Schuette, K.P., Starr, E.D., 1994. A molecular phylogeny of *Arceuthobium* (Viscaceae) based on nuclear ribossomal DNA internal transcribed spacer sequences. Am. J. Bot. 81, 1149–1160.
Nilsson, S., 1970. Pollen morphological contributions to the taxonomy of *Lisianthus* L. s. lat. (Gentianaceae). Svensk Bot. Tidskr. Foreningen 64, 1–43.
Nilsson, S., 2002. Gentianaceae – a review of palynology. In: Struwe, L., Albert, V.A. (Eds.), Gentianaceae – Systematics and Natural History. Cambridge University Press, Cambridge, pp. 377–497.
Nylander, J.A.A., 2004. MrModeltest Versão 2.2. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University. Available at: <http://www.abc.se/~nylander/>.
Progel, A., 1865. Gentianaceae. In: von Martius, C.F.P. (Ed.), Flora Brasiliensis, vol. 6 (1). Frid. Fleischer, Munich, pp. 197–248.
Rambaut, A., Drummond A.J., 2009. Tracer v.1.5. Available at: <http://beast.bio.ed.ac.uk/Tracer>.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed model. Bioinformatics 19, 1572–1574.
Simmons, M.P., 2004. Independence of alignment and tree search. Mol. Phylogenet. Evol. 31, 874–879.
Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49, 369–381.
Steyermark, J.A. et al., 1953. Gentianaceae. Contributions to the flora of Venezuela. Fieldiana 28 (3), 496–499.
Struwe, L., Maas, P.J.M., Albert, V.A., 1997. *Aripuana cullmaniorum*, a new genus and species of Gentianaceae from white sands of southeastern Amazonas, Brazil. Harv. Pap. Bot. 2, 235–253.
Struwe, L., Maas, P.J.M., Pihlar, O., Albert, V.A., 1999. Gentianaceae. In: Berry, P.E. et al. (Eds.), Flora of the Venezuelan Guayana, vol. 5. Missouri Botanical Garden, St. Louis, pp. 474–542.
Struwe, L., Kadereit, J., Klackenberg, J., Nilsson, S., Thiv, M., von Hagen, K.B., Albert, V. A., 2002. Systematics, character evolution, and biogeography of Gentianaceae, including a new tribal and subtribal classification. In: Gentianaceae –

Systematics and Natural History. Cambridge University Press, Cambridge, pp. 21–309.
Struwe, L., Albert, V.A., Calió, M.F., Frasier, C., Lepis, K.B., Mathews, K.G., Grant, J.R., 2009. Evolutionary patterns in Neotropical Helieae (Gentianaceae): evidence from morphology, chloroplast and nuclear DNA sequences. Taxon 58, 479–499.
Struwe, L., Nilsson, S., Albert, V.A., 2008. *Roraimaea* (Gentianaceae: Helieae) – a new gentian genus from white sand campinas and Cerro de la Neblina of Brazil and Venezuela. Harv. Pap. Bot. 13, 35–45.
Struwe, L., 2014. Classification and Evolution of the Family Gentianaceae. In: Rybczynski, J.J. et al. (Eds.), The Gentianaceae – Characterization and Ecology, vol. 1. Springer-Verlag, Berlin, pp. 13–36.
Swofford, D.L., 2003. PAUP*: Phylogenetic Analysis Using Parsimony* and Other Methods. Version 4.0b10. Sinauer, Sunderland, USA.
Thiv, M., Struwe, L., Kadereit, J.W., 1999. The phylogenetic relationships of the Canarian laurel forest endemic *Ixanthus viscosus* (Aiton) Griseb. (Gentianaceae): evidence from *mat*K and ITS sequences, and floral morphology and anatomy. Plant Syst. Evol. 218, 299–317.
Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.
White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A. et al. (Eds.), PCR Protocols: A Guide to Methods and Amplifications. Academic Press, San Diego, pp. 315–322.
Wiens, J.J., 1998. Combining data sets with different phylogenetic histories. Syst. Biol. 47, 568–581.
Wiens, J.J., 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52, 528–538.
Wiens, J.J., Morrill, M.C., 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. Syst. Biol. 60, 719–731.
Yoder, A.D., Irwin, J.A., Payseur, B.A., 2001. Failure of the ILD to determine data combinability for slow loris phylogeny. Syst. Biol. 50, 408–424.