A peer-reviewed version of this preprint was published in PeerJ on 20 July 2016.

View the peer-reviewed version (peerj.com/articles/2220), which is the preferred citable publication unless you specifically need to cite this preprint.

Roy T, Catlin NS, Garner DMG, Cantino PD, Scheen A, Lindqvist C. 2016. Evolutionary relationships within the lamioid tribe Synandreae (Lamiaceae) based on multiple low-copy nuclear loci. PeerJ 4:e2220 https://doi.org/10.7717/peerj.2220

Evolutionary relationships within the lamioid tribe Synandreae (Lamiaceae) based on multiple low-copy nuclear loci

Tilottama Roy, Nathan S Catlin, Drake MG Garner, Philip D Cantino, Anne-Cathrine Scheen, Charlotte Lindqvist

The subfamily Lamioideae (Lamiaceae) comprises ten tribes, of which only Stachydeae and Synandreae include New World members. Previous studies have investigated the phylogenetic relationships among the members of Synandreae based on plastid and nuclear ribosomal DNA loci. In an effort to re-examine the phylogenetic relationships within Synandreae, the current study incorporates data from four low-copy nuclear loci, *PHOT1*, *PHOT2*, *COR*, and *PPR*. Our results confirm previous studies based on chloroplast and nuclear ribosomal markers in supporting monophyly of tribe Synandreae, as well as sister relationships between *Brazoria* and *Warnockia*, and between that pair of genera and a monophyletic *Physostegia*. However, we observe incongruence in the relationships of *Macbridea* and *Synandra*. The placement of Synandreae within Lamioideae is poorly resolved and incongruent among different analyses, and the sister group of Synandreae remains enigmatic. Comparison of the colonization and migration patterns corroborates a single colonization of the New World by Synandreae during the mid-Miocene. This is in contrast to the only other lamioid tribe that includes New World members, Stachydeae, which colonized the New World at least twice—during the mid-Miocene and Pliocene. Edaphic conditions and intolerance of soil acidity may be factors that restricted the distribution of most genera of Synandreae to southeastern and south-central North America, whereas polyploidy could have increased the colonizing capability of the more wide-ranging genus, *Physostegia*.

Evolutionary relationships within the lamioid tribe Synandreae (Lamiaceae) based on multiple low-copy nuclear loci

Tilottama Roy[1,2], Nathan S. Catlin[1,3], Drake M. G. Garner[1], Philip D. Cantino[4], Anne-Cathrine Scheen[5], and Charlotte Lindqvist[1]

[1] Department of Biological Sciences, University at Buffalo (SUNY), Buffalo, NY, USA

[2] Current address: Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

[3] Current address: Department of Biology, University of Florida, Gainesville, FL, USA

[4]Department of Environmental and Plant Biology, Ohio University, Athens, OH, USA

[5] Museum of Archaeology, University of Stavanger, Stavanger, Norway

Corresponding Author:

Charlotte Lindqvist

Department of Biological Sciences, University at Buffalo (SUNY), Buffalo, NY, 14260, USA;

Email address: cl243@buffalo.edu

ABSTRACT

The subfamily Lamioideae (Lamiaceae) comprises ten tribes, of which only Stachydeae and Synandreae include New World members. Previous studies have investigated the phylogenetic relationships among the members of Synandreae based on plastid and nuclear ribosomal DNA loci. In an effort to re-examine the phylogenetic relationships within Synandreae, the current study incorporates data from four low-copy nuclear loci, *PHOT1*, *PHOT2*, *COR*, and *PPR*. Our results confirm previous studies based on chloroplast and nuclear ribosomal markers in supporting monophyly of tribe Synandreae, as well as sister relationships between *Brazoria* and *Warnockia*, and between that pair of genera and a monophyletic *Physostegia*. However, we observe incongruence in the relationships of *Macbridea* and *Synandra*. The placement of Synandreae within Lamioideae is poorly resolved and incongruent among different analyses, and the sister group of Synandreae remains enigmatic. Comparison of the colonization and migration patterns corroborates a single colonization of the New World by Synandreae during the mid-Miocene. This is in contrast to the only other lamioid tribe that includes New World members, Stachydeae, which colonized the New World at least twice—during the mid-Miocene and Pliocene. Edaphic conditions and intolerance of soil acidity may be factors that restricted the distribution of most genera of Synandreae to southeastern and south-central North America, whereas polyploidy could have increased the colonizing capability of the more wide-ranging genus, *Physostegia*.

Keywords Synandreae, *Brazoria*, *Macbridea*, *Physostegia*, *Synandra*, *Warnockia*, Biogeography, Phylogeny, Nuclear markers, North America, Stachydeae

INTRODUCTION

The angiosperm family Lamiaceae has a worldwide distribution, comprising ~7200 species in approximately 240 genera (Bentham, 1876; Harley et al., 2004). Lamiaceae is subdivided into seven subfamilies, of which Lamioideae, the second largest, exhibits an impressive ecological and taxonomic diversity (Scheen et al., 2010; Bendiksby et al., 2011; Roy and Lindqvist, 2015). Most members of Lamioideae have been classified into ten tribes, with the majority of the species inhabiting Eurasia and Africa. However, approximately 113 species are native to the New World, and they are members of just two tribes: Stachydeae and Synandreae (Scheen et al., 2010; Roy et al., 2013; 2015). Considerable molecular phylogenetic work has recently been performed in Stachydeae (Lindqvist and Albert 2002; Salmaki et al., 2013, Roy et al., 2013, 2015), and it has been suggested that the New World members of the genus *Stachys* colonized the Americas twice, first during the mid-Miocene and later during the early Pliocene (Roy et al., 2013; 2015). The focus of the current study is Synandreae, the other lamioid tribe represented in the New World, comprising five genera: *Synandra* Nutt., *Macbridea* Elliott ex Nutt., *Brazoria* Engelm & A. Gray, *Warnockia* M.W. Turner, and *Physostegia* Benth.

All five genera of Synandreae are herbs with relatively large flowers (for Lamiaceae), which are sessile or short-pedicellate in racemoid inflorescences. Corolla color ranges from white (*Macbridea alba*, *Synandra*, and some *Physostegia* species) to lavender (*Macbridea caroliniana*, *Brazoria*, *Warnockia*, and most *Physostegia* species). The anther thecae either narrow apically to a sharp point (*Synandra*) or bear one or more teeth along the suture. Monotypic *Synandra hispidula* ($2n$=18) is a biennial of mesic woodlands in the eastern United States, mostly in the Appalachian region (Harley et al., 2004). It differs from the rest of the tribe in having long-petiolate, cordate-ovate leaves. *Macbridea* ($2n$=18) comprises two species of

63 rhizomatous perennial herbs of wetlands and pine savannas in the southeastern United States
64 (Harley et al., 2004) (Fig. 1). *Macbridea* flowers are tightly packed into terminal and sub-
65 terminal capitate glomerules, unlike the elongate inflorescences of the other four genera, and its
66 three-lobed calyx is distinctive. *Brazoria* (2*n*=28) comprises three species of annuals of sandy
67 soils in eastern and central Texas (Fig. 1), with an erect and deeply bifid upper corolla lip
68 (Turner, 1996). Monotypic *Warnockia scutellarioides* (2*n*=20) is an annual of calcareous soils in
69 Texas, southern Oklahoma, and northwestern Mexico (Coahuila) (Turner, 1996) (Fig. 1).
70 *Physostegia* (2*n*=38 and 76), with 12 species of perennials, is the most widespread genus of
71 Synandreae, ranging from Northern Canada to Northern Mexico and growing in diverse habitats
72 and a wide range of soil conditions (Cantino, 1982). *Physostegia virginiana* is often grown as an
73 ornamental and has become naturalized in some areas. *Physostegia* is the only genus of
74 Synandreae with an actinomorphic, five-lobed calyx.

75 Bentham (1848) described subtribe Melittidinae ("Melittieae"), comprising the
76 monotypic European genus *Melittis* and the North American genera *Brazoria*, *Synandra*,
77 *Macbridea*, and *Physostegia*. Bentham (1876) and Briquet (1895-1897) added the Asian genus
78 *Chelonopsis* to this subtribe but transferred *Brazoria* to Scutellariinae and Prunellinae,
79 respectively. Cantino (1985a) and Abu-Asab and Cantino (1987) considered Melittidinae to
80 include *Brazoria*, and Turner (1996) segregated *Warnockia* from *Brazoria*. However,
81 morphological and karyological studies (Cantino 1982; 1985a) and investigation of leaf anatomy
82 (Abu-Asab and Cantino, 1987), palynology (Abu-Asab and Cantino, 1994), and pericarp
83 structures (Ryding, 1994) were unable to provide synapomorphies supporting the monophyly of
84 Melittidinae. Furthermore, molecular phylogenetic studies demonstrated the non-monophyly of
85 Melittidinae (Scheen et al., 2008; Scheen et al., 2010; Bendiksby et al., 2011; Salmaki et al.,

2013; Roy and Lindqvist, 2015). Scheen et al. (2008) found that, rather than grouping with the North American endemics, *Melittis melissophyllum* grouped with *Stachys*, and *Chelonopsis* grouped with the Asian genus *Gomphostemma*. These studies also demonstrated the monophyly of a group comprising the North American endemics (*Brazoria*, *Warnockia*, *Synandra*, *Macbridea*, and *Physostegia*). Since *Melittis* is not part of this clade, it could no longer be named Melittidinae and was instead named tribe Synandreae (Scheen et al., 2008). Since the study by Scheen et al. (2008) was based on chloroplast and nuclear ribosomal DNA markers, the goal of the current study is to investigate the phylogenetic relationships among the members of Synandreae based on low-copy nuclear markers.

With the availability of improved technologies and universal primers, there has been a shift from plastid and ribosomal loci towards the use of low-copy nuclear genes (Mort and Crawford, 2004) in investigations of interspecific phylogenetic relationships because they often have a higher rate of evolution, leading to higher resolution in species-level phylogenies. Furthermore, maternally inherited plastid DNA, as a single linkage group, can only provide the genealogical history of one parent and thus cannot provide any information on hybrid species histories. Although nuclear ribosomal DNA (ITS, ETS and 5S-NTS) is biparentally inherited, these data do not always provide reliable markers for the reconstruction of hybrid speciation and resolution of phylogenetic histories due to concerted evolution and homogenization (Wendel et al., 1995). Hence, the true evolutionary relationships among closely related taxa may be confounded. Also, in situations where speciation has taken place rapidly, as may be the case within Synandreae, genomic DNA may not have undergone enough divergence to resolve a phylogeny with only one locus (Seehausen, 2004). In such cases, multiple independent nuclear loci may provide the variability necessary to make a more accurate estimation of phylogenetic

relationships (Sang, 2002; Hughes et al., 2006). However, low-copy nuclear genes are not devoid of shortcomings. Some of the issues encountered while dealing with low-copy nuclear loci include presence of paralogous copies, incomplete lineage sorting, and gene tree/species tree incongruence due to hybridization and introgression. Hence, these factors should always be taken into account when drawing conclusions on evolutionary relationships.

In this study, we have analyzed data from four low-copy nuclear loci: two *PHOT* gene duplicates (*PHOT1* and *PHOT2*), *COR* (cold acclimation protein), and the *PPR* (pentatricopeptide repeat) region AT3G09060. The *PHOT* genes are responsible for encoding the blue and ultra-violet-A light receptor of plants involved in the process of phototropism (Christie et al., 1998), chloroplast relocation (Jarillo et al., 2001; Kagawa et al., 2001), and the regulation of stomatal openings (Kinoshita et al., 2001). Two *PHOT* loci are present in most angiosperms (*PHOT1* and *PHOT2*), resulting from a duplication event predating the divergence between monocots and tricolpates (Briggs et al., 2001). The two *PHOT* gene duplicates have accumulated a sufficiently large number of nucleotide substitutions since their divergence to be distinct from each other, which is important for overcoming orthology/paralogy issues when being utilized in phylogenetic analyses (Fitch, 1970). The two paralogs have been shown to be so variable that their intron regions are unalignable with each other and hence can be treated as two separate markers. Due to the presence of many small, relatively conserved exon regions, separated by variable introns, it has been suggested that the amount of information that can be collected from these loci is high relative to the effort that is applied to work with them (Yuan and Olmstead, 2008). Also, through the investigation of these two paralogs, the mode of intron evolution can be observed across closely related species, such as members of Synandreae. All these factors make the *PHOT* gene duplicates ideal for use in our current study. The *COR* locus also consists of

intron regions flanked by exons that provide conserved primer binding sites (EPIC markers; Curto et al., 2012; Thomson et al., 2008). Curto and colleagues (2012) have shown from their study of *Micromeria* (Lamiaceae) that this locus can be phylogenetically informative, providing a substantial amount of variation among closely related species. Lastly, the *PPR* gene family encodes a group of proteins with short helical repeats that are arranged in stacks, forming extended surfaces (Geddy and Brown, 2007; Barkan and Small, 2014). Previous studies (Yuan et al., 2009, 2010; Crowl et al., 2014) and our own study on Lamioideae (Roy and Lindqvist, 2015) have shown the *PPR* loci to be useful markers to reconstruct phylogenetic relationships involving rapidly radiating taxa. In addition to the low copy nuclear markers, we also incorporated chloroplast DNA (cpDNA) data from previous studies (Scheen et al., 2008; Scheen et al, 2010; Bendiksby et all, 2011) for four regions (*mat*K, *rps*16, *trn*L intron, and *trn*L-F spacer) to generate a more comprehensive multispecies coalescent tree.

The goals of this study included 1) assessing the monophyly of tribe Synandreae, 2) further clarifying relationships within Synandreae, 3) investigating the historical biogeography of Synandreae, including its introduction into the New World, and 4) comparing the migration and diversification patterns of Synandreae with those of tribe Stachydeae, the only other lamioid tribe with endemic New World species.

METHODS

Taxon sampling, DNA extraction, amplification, and sequencing

Leaf material was collected from specimens held at the following herbaria: BISH, C, GH, LL, O, TEX, UNA, UPS, and US (herbarium acronyms follow Holmgren et al., 1990). All taxon names in this study follow the "World checklist of Lamiaceae and Verbenaceae" (Govaerts et al., 2013).

DNA was extracted from silica-dried leaves or from herbarium specimen leaf fragments using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. DNA sequence data were collected for a total of 16 of the 19 species of Synandreae, including accessions from all five genera (Table 1). Furthermore, 22 additional lamioid species were selected based on previous studies (Scheen et al. 2010; Bendiksby et al. 2011). *Scutellaria hirta* was included as an outgroup since many studies have shown Scutellarioideae and Lamioideae to be closely related (Wagstaff et al., 1998, Scheen et al., 2010; Bendiksby et al., 2011; Li et al., 2012; Chen et al., 2014).

For amplification of the two *PHOT* loci, we used primers previously published by Yuan and Olmstead (2008). For the *PHOT1* locus, we utilized the primers 10F ('5'-ATTGGAGTSCAAYTAGATGGAAG-'3') and 12R ('5'-TCCACAAGTCCTCTGGTTTCT-'3'). For the *PHOT2* locus, due to difficulty in amplification of the entire locus, we amplified two separate fragments and treated them initially as two separate loci, labeling them *PHOT2A* and *PHOT2B*. For the amplification of *PHOT2A*, we utilized the primers 10F ('5'-GATGGAAGTGATMATKTGGAAC-'3') and 12R ('5'-AGCCCACAGGTCYTCTGGTCTC-'3'), whereas *PHOT2B* was amplified with primers 12F ('5'-GAGACCAGARGACCTGTGGGCT'-'3') and 14R ('5'- GATTTRTCCATTGCTTTCATGGC-'3'). The *COR* locus was amplified using the following primers previously published by Curto et al. (2012): forward primer ('5'-CTCGAATGTGTTCCTGCAG-'3') and reverse primer ('5'-CACATCCCTCTTAGTCCCATAC-'3'). Amplification and sequencing of *PPR* is described in Roy and Lindqvist (2015). All loci were amplified using a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA) using a touchdown method with the following thermocycling profile: hold for 10 min at 95◦C; 10 cycles of 1 min at 95◦C, 1 min at 60◦C and

decreasing the temperature by 1◦C every cycle, 1 min at 72◦C; followed by 35 cycles of 1 min at 94◦C, 1 min at 50◦C, 1 min at 72◦C; and a final extension of 1 min at 72◦C. In certain cases when this touchdown method failed to amplify our locus of interest, a modified touchdown method was used, where the annealing temperature started at 55◦C and decreased by 1◦C every cycle. PCR products were purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany) following the manufacturer's instructions.

All PCR products generated were further cloned using the Qiagen PCR cloning kit (Qiagen, Hilden, Germany) following the manufacturer's instructions, with the exception that we used 25 µL competent cells to transform each ligation reaction. Transformed clones were incubated overnight at 37◦C. Up to 12 positive clones were picked per individual, with the average number of clones varying between 2 and 4 per locus. PCR reactions were prepared in 25 µL volumes with the AmpliTaq DNA Polymerase buffer II kit (Applied Biosystems, Foster City, CA, USA) using 2.5 µL buffer, 2.5 µL MgCl2, 1.0 µL dNTP, 0.6 µL each of M13F and M13R primers and 0.2 µL of AmpliTaq polymerase. All PCR products were examined by gel electrophoresis on 1% agarose gels, and positive PCR amplified products were sequenced in one direction using SP6 or T7 primers at the University of Washington High Throughput Genomics Center, Seattle, USA.

Phylogenetic tree reconstruction

All sequences generated were edited and assembled in the program Sequencher v.4.7 (Gene Codes, Ann Arbor, Michigan, USA) and aligned with ClustalX v.2 (Larkin et al., 2007) or MAFFT (EMBL-EBI); the alignments were manually adjusted in BioEdit (Hall, 1999). Gaps were treated as missing, and indels were not coded. We evaluated evidence of recombination

using the Phi test (Bruen et al., 2006) in Splitstree v.4.13.1 (Huson, 1998). Initial Bayesian and maximum likelihood (ML) analyses were performed on the two *PHOT*2 regions, *PHOT*2*A* and *PHOT*2*B* (see above), separately, but since their topologies were compatible, the datasets generated from these two regions were concatenated in the program WINCLADA (Nixon, 1999) before running further phylogenetic analyses. Phylogenetic relationships were examined for the three loci, *PHOT1*, *PHOT2* and *COR*, separately using Bayesian inference conducted in either MrBayes v.3.1.2 or 3.2.2 (Huelsenbeck and Ronquist, 2001) using Cipress XSEDE (Miller et al., 2010). We used substitution models that best fit the data as determined by the Bayesian Information Criterion (BIC) using the program jModeltest v.1.1 (Posada, 2008). The HKY+G model was used for the *PHOT1* loci. The TPM1uf+G and TrN+G models were initially selected for *PHOT2*, and *COR*, respectively, but since these two models are not implemented in MrBayes 3.2.2, we utilized the next model (HKY+G) proposed with the highest score. We also conducted ML analyses using the RAxML Blackbox webserver (Stamatakis et al., 2008), or through the RAxML HPC Blackbox in the Cipress portal (Miller et al., 2010). We rooted the *COR* and *PHOT1* trees with *Scutellaria hirta* (not shown in figures), however, due to lack of sequence data for *S. hirta* for the *PHOT2* loci, we used *Gomphostemma javanicum* to root the *PHOT2* tree. Apart from analyzing our individual datasets, we also concatenated the three low-copy nuclear loci and conducted Bayesian and ML analyses on this dataset. For this purpose, we selected the GTR+G model and used *S. hirta* as the outgroup. Phylogenetic analyses of *PPR* alone are described in Roy and Lindqvist (2015).

Coalescence analysis, network analysis, and ancestral area reconstruction

We implemented a multispecies coalescence model within the BEAST v.1.8.0 software package (Drummond and Rambaut, 2007) to further explore phylogenetic signals from our nuclear loci. *BEAST applies Bayesian MCMC analysis of the sequence data, jointly exploring gene trees and species trees to estimate the species tree posterior distribution under the assumption of the coalescence model. For this purpose, we incorporated sequences from the low-copy nuclear loci *PPR* (Roy and Lindqvist, 2015), as well as from a concatenated dataset comprising sequences from the cpDNA regions *mat*K, *rps*16, *trn*L intron, and *trn*L-F spacer obtained from previously published studies (Scheen et al., 2008; Scheen et al., 2010; Bendiksby et al, 2011), for all the members of Synandreae. We pruned these datasets, keeping only members of Synandreae and the taxa common to all of the loci. The nuclear loci were treated as unlinked. A relaxed molecular clock model for all the loci and HKY+G models of nucleotide substitution were applied for the nuclear loci, and the GTR+G model for the cpDNA regions. The tree prior was set to exponential, and other priors were kept to default values. Analyses were done for 10 million generations sampling every 10,000 generations. A relative proportion of the posterior samples from each Markov chain were discarded as burn-in, and trees were summarized in TreeAnnotator v.1.8.0 (Drummond and Rambaut, 2007). The resulting trees were then visualized in FigTree v. 1.4.0 (Rambaut, 2008). We also implemented a phylogenetic network method to analyze signals of reticulate evolution and character conflicts in our datasets. The network was created with Neighbor-Net (Huson and Bryant, 2006) in SplitsTree v.4.13.1 (Huson, 1998) using uncorrected p-distances. For this purpose, we utilized *PHOT1* and *PPR*, the two nuclear loci that have the highest representation in the various lamioid tribes, generating a concatenated dataset in WINCLADA (Nixon, 1999). For distance calculations, we chose the most parameterized model available in SplitsTree v.4.13.1 with an HKY85 model, transitions: transversions weighted 2:1,

gamma model of rate heterogeneity, and base frequencies estimated empirically. For our ancestral area reconstruction, we used the program S-DIVA (Statistical Dispersal-Vicariance Analysis; Yu et al., 2010) as implemented in the program RASP v.2.1. The program implements a Bayesian approach to dispersal-vicariance analysis (DIVA; Ronquist, 1997), following the method suggested by Nylander et al. (2008), which estimates optimized areas over a set of trees and accounts for uncertainty in the phylogenetic estimate. The distribution range for the various species was selected based on present day distributions of the species of Synandreae according to information contained in the World Checklist of Lamiaceae and Verbenaceae (Goevarts et al., 2013). Following the geographical zones defined by Brummit (2001), six geographical areas were defined, and each included species of Synandreae assigned to one or more of these areas: A: southeastern US except Texas; B: east-central US; C: Texas; D: Mexico; E: southern Canada, North Dakota and Northwest Territories; F: western Canada and north-central US and G: Old World. The S-DIVA analysis was performed using the tree file generated after the burn-in period from the MrBayes run consisting of 1000 random trees utilizing the concatenated dataset mentioned above comprising of *PHOT1*, *PHOT2* and *COR*. Since this dataset had multiple cloned sequences for each species forming their own clades, we further trimmed this dataset by including only one representative sequence for each species to form our final dataset. This file was converted into a condensed tree file in RASP. We ran S-DIVA with the default settings, except maximum number of areas was set to 2. We did not select the "allow reconstruction" button, and this allowed the program to calculate the proportions of inferred alternative most-parsimonious ancestral ranges at each node in a tree accounting for topological as well as dispersal-vicariance uncertainties. We mapped the ancestral areas onto the 50% majority rule consensus tree derived from our Bayesian analysis of the above dataset.

RESULTS

Our complete datasets, including gaps, generated from our current study consisted of 564 characters for *PHOT1*, 1816 characters in the concatenated *PHOT2* dataset, and 352 characters for *COR*, totaling ~2.7 Kb characters (the raw alignment files in FASTA format for the three loci are provided in Supplemental Information). Our results indicated correlation within the overall topologies of the 50% majority rule Bayesian consensus trees and maximum likelihood (ML) trees for all these datasets (Figs. 1a, 1b, and 1c, respectively). DNA sequence data were collected for a total of 71 samples for the *PHOT1* locus, representing 34 species of Lamioideae, including 17 species of Synandreae. For the *PHOT2* locus, we generated a total of 51 sequences, comprising 25 lamioid species including 14 species representing all of the genera of Synandreae. For the *COR* locus, 64 sequences were included, representing 26 lamioid species and 15 species of Synandreae. The sampling for our three new datasets differs due to limitations in the availability of material and success with DNA extraction and amplification. However, based on the topological congruence in the overall placement of the various species, we expect that the few missing species will group with other members of their respective genera included in the analyses.

Among these three new datasets, the *PHOT1* phylogeny (Fig. 1a) is based on the most comprehensive sampling of taxa across most of the lamioid tribes. This dataset includes representative taxa from the tribes Pogostemoneae, Gomphostemmateae, Marrubieae, Leucadeae, Phlomideae, Stachydeae and Synandreae, and the unplaced genera *Galeopsis* and *Betonica*. Using *Scutellaria hirta* as the outgroup (not shown), the *PHOT1* phylogeny infers *Achyrospermum radicans* (Pogostemoneae) as sister to all other included Lamioideae (posterior probability value PP=1.00; bootstrap value BS=100). *Gomphostemma javanicum*

(Gomphostemmateae) is sister to the remaining lamioid tribes, but the latter clade is poorly supported. Although the inter-relationships of the remaining tribes, along with *Galeopsis* and *Betonica*, are unresolved or poorly supported, the tribes themselves, including Synandreae, are strongly supported as monophyletic. In the *PHOT*2 combined phylogeny (Fig. 1b), *Gomphostemma* was treated as the outgroup in the absence of *Scutellaria* or any representatives of Pogostemoneae. In this tree, *Acrotome* (Leucadeae) and *Ballota* (Marrubieae) form a clade, sister to the rest of the Lamioideae. Within the latter clade, members of Stachydeae are monophyletic (PP=1.00; BS=100) and group with a poorly supported clade comprising *Betonica*, Phlomideae, and a strongly supported Synandreae. In the *COR* tree (Fig. 1c), which used *Scutellaria hirta* as the outgroup (not shown), Gomphostemmateae emerges as sister to a trichotomy comprising the clades Phlomideae-Stachydeae (PP=0.98; BS<80), Marrubieae-Leucadeae (PP=0.99; BS<80) and Synandreae (PP=1.00; BS=95). Our concatenated dataset of the three loci *PHOT1*, *PHOT2*, and *COR* (Fig. 2) corroborates the *PHOT1* results in placing *Achyrospermum radicans* in the basal split of the lamioid tree (PP=1.00; BS=100). Members of Marrubieae form a strongly supported clade (PP=1.00; BS=99) sister to a monophyletic Leucadeae (PP=1.00; BS=99). The tribe Stachydeae is inferred to be paraphyletic, although this paraphyly is poorly supported in both the ML and Bayesian analyses.

The position of Synandreae within Lamioideae remains poorly resolved. It is inferred to be sister to Stachydeae based on *PHOT1* (Fig. 1a) sister to *Phlomis fruticosa* based on *PHOT2* (Fig. 1b), in a trichotomy with the Phlomideae-Stachydeae clade and the Marrubieae-Leucadeae clade based on *COR* (Fig. 1c), and as sister to a clade composed of some taxa from Stachydeae and Phlomideae in the concatenated tree (Fig. 2). All individual gene trees (Figs. 1a-c), as well as the phylogeny resulting from the concatenated dataset (Fig. 2), strongy support the monophyly of

315 Synandreae (PP=1.00 and BS=100 in *PHOT1*, *PHOT2,* and the concatenated dataset; PP=1.00
316 and BS=95 in the COR tree).

317 The five genera of Synandreae (*Synandra*, *Macbride*a, *Brazoria*, *Warnockia*, and
318 *Physostegia*) are each resolved as monophyletic in all trees, except in the *COR* analysis where
319 *Synandra* and *Macbridea* are unresolved with respect to each other. In phylogenies based on the
320 *PHOT1*, *PHOT2*, and concatenated datasets *Brazoria* and *Warnockia* are strongly supported as
321 sister groups (concatenated: PP=0.99; BS=90), and this clade is in turn sister to *Physostegia*
322 (concatenated: PP=1.00; BS=100). In the *COR* tree, *Brazoria*, *Warnockia,* and *Physostegia* form
323 a trichotomy, and the clade comprising these three genera receives only moderate support. All
324 the individual gene trees, as well as the concatenated dataset, strongly support the monophyly of
325 *Physostegia* (concatenated: PP=1.00; BS=100). Relationships among *Physostegia* species are
326 poorly resolved in the individual gene trees. In the phylogeny from the concatenated dataset,
327 relationships within *Physostegia* are better resolved but still poorly supported, with one
328 exception: all species of *Physostegia* are resolved into two main clades (clades A and B in Fig.
329 2), which are well supported in the Bayesian analyses, although not in the ML analyses. Clade A
330 (PP=0.93, BS=75) comprises *P. longisepala* (clone1), *P. ledinghamii*, *P. correlli, P. virginiana*,
331 *P. pulchella*, and *P. angustifolia*, whereas clade B (PP=0.94, BS=71) comprises *P. godfreyi*, *P.*
332 *digitalis*, *P. parviflora*, *P. leptophylla*, *P. longisepala* (clone 2), and *P. purpurea*.

333 The multispecies coalescence-based tree from the *BEAST analysis of all markers (Fig.
334 3), corroborates results from previous findings (Scheen et al., 2010; Bendiksby et al., 2011; Roy
335 and Lindqvist, 2015), as well as those from our individual gene trees (Figs. 1a-c) and
336 concatenated dataset (Fig. 2), supporting *Synandra* as sister to *Macbridea*, which together are

337 sister to the remaining Synandreae (PP= 1.00). *Warnockia* and *Brazoria* form a clade (PP=0.93),
338 which is sister to a robustly supported *Physostegia* (PP=1.00).

339 The neighbornet network analysis of the two loci *PHOT1* and *PPR* (Fig. 4) corroborates
340 the clustering of species into their respective tribes and an isolated phylogenetic position of
341 Synandreae separate from the remaining Lamioideae. Within Synandreae, *Synandra* and
342 *Macbridea* are close relatives and separate from its other members of which *Brazoria*, and
343 *Warnockia* are most closely related. No infrageneric phylogenetic structure is resolved among
344 the members of *Physostegia* included here.

345 The ancestral area reconstruction (not shown) infers southeastern US and Texas to be the
346 ancestral areas for the entire Synandreae clade, as well as for various subgroups of the
347 *Physostegia* clade.

348

349 DISCUSSION

350 Phylogenetic relationships among Synandreae and their position within Lamioideae were until
351 recently only investigated with cpDNA and nrDNA markers (Scheen et al., 2008; 2010;
352 Bendiksby et al., 2011). Our current study reconstructs evolutionary relationships in this group
353 based on multiple low-copy nuclear DNA markers. Although our results corroborate many of the
354 findings from previous research (Scheen et al., 2008; Scheen et al., 2010; Bendiksby et al.,
355 2011), we observe some instances of incongruence. Since low-copy loci are biparentally
356 inherited, there is a possibility that either the paternal or maternal gene copy in hybrid progeny
357 was randomly selected, resulting in conflicting patterns in the placement of some of the taxa in
358 the individual gene trees. Our phylogenetic network from the two loci *PHOT1* and *PPR* also
359 shows signatures of reticulation events throughout the phylogeny, including at the base where the

different tribes split (Fig. 4). As has been noted in previous studies, the signatures of ancestral gene flow that may have taken place in deep time could have eroded after a long history of divergence, and a substantially larger amount of data are required to precisely pinpoint those loci, which could have introgressed from one species to another (Leache et al., 2014).

Monophyly of tribe Synandreae: Chromosomal evolution and intergeneric relationships

All gene trees (Fig. 1a-c), as well as the tree from the concatenated dataset (Fig. 2), unanimously corroborate the monophyly of the New World tribe Synandreae, although its relationship with other lamioid members, and its sister group, remain enigmatic. This clade of North American (NA) endemics is distinguished from most other lamioid genera by the absence of thick-walled cells in the exocarp (Ryding, 1994). The five member genera—*Synandra*, *Macbridea*, *Brazoria*, *Warnockia*, and *Physostegia*—are also characterized by the presence of villous stamens (Harley et al., 2004) and by the anther thecae either narrowing apically to a sharp point (*Synandra*) or bearing one or more teeth along the suture (the other four genera), though it is not clear whether these two character states are homologous.

Our findings unanimously corroborate the monophyly of *Brazoria* and *Warnockia*, which together are sister to *Physostegia*, a relationship also found by Scheen et al. (2008). *Brazoria* and *Warnockia* were recently recognized as separate genera by Turner (1996), having long been treated as congeneric. *Brazoria*, *Warnockia*, and *Physostegia* share distinctive saclike idioblasts in the leaf mesophyll, a feature not found in *Synandra* and unknown elsewhere in the family (Abu-Asab and Cantino, 1987; Lersten and Curtis, 1998), thus an unambiguous synapomorphy.

The strongly supported sister-relationship between *Synandra* and *Macbridea*, which form a clade that is sister to the rest of Synandreae, was also encountered in a nuclear phylogeny based

on the *PPR* locus alone (Roy and Lindqvist, 2015), but not in studies based on cpDNA and nrDNA regions (Scheen et al., 2008, 2010; Bendiksby et al., 2011). In these latter studies, *Synandra* emerged as sister to the rest of Synandreae. There is non-molecular support for both phylogenetic hypotheses. Previous chromosomal studies (Cantino, 1985a) demonstrated that *Macbridea* and *Synandra* have the same chromosome number (2*n*=18). They also share a derived androecial character—the outer thecae of the anterior stamens are fused (for pictures of this feature in *Synandra*, see Cantino [1985b]). Chromosome numbers based on x=9 are uncommon in subfamily Lamioideae and may be a synapomorphy for a clade comprising *Synandra* and *Macbridea* (Cantino, 1985a). However, in leaf shape, texture, and indumentum, *Macbridea* is much more similar to *Brazoria*, *Warnockia*, and *Physostegia* than to *Synandra* (Cantino, 1982). The leaves in the former four genera are usually lanceolate to elliptical or oblanceolate (rarely ovate and never cordate), narrowing to a cuneate to rounded base, have a firm, semi-succulent texture, are glabrous or at most sparsely puberulent, and at least the upper leaves are sessile. In contrast, the leaves in *Synandra* are broadly ovate-cordate, membranaceous, villous, and petiolate below the inflorescence. Furthermore, Cantino (1990) suggested that absence of anomocytic stomata is a synapomorphy of a clade comprising *Macbridea*, *Brazoria* (including *Warnockia*), and *Physostegia*. It is thus evident that *Macbridea* shares conflicting sets of apparent synapomorphies with *Synandra*, on the one hand, and the *Brazoria-Warnockia-Physostegia* clade, on the other. A possible explanation for both this character distribution and the inconsistency between cpDNA and low-copy nuclear loci in the placement of *Macbridea* is a scenario involving ancient hybridization between the ancestors of these genera.

Synandra, Macbridea, Warnockia, *Brazoria*, and *Physostegia* are characterized by base chromosome numbers x=9 (2*n*=18), x=9 (2*n*=18), x=10 (2*n*=20), x=14 (2*n*=28), and x=19

406 (2*n*=38, 76), respectively (Cantino, 1985a). Although it has been suggested (Gill, 1981) that the
407 ancestral chromosome number in Lamiaceae is x=7, a base number of x=9 in the ancestor of
408 Synandreae could have evolved through aneuploid increase. Similarly, chromosome evolution
409 within Synandreae may have occurred through a series of aneuploidy events (Scheen et al., 2008)
410 from x=9 to x=10, x=14 and x=19 in the ancestors of *Warnockia, Brazoria*, and *Physostegia*,
411 respectively. Increasing chromosome numbers in these genera in comparison to *Synandra* and
412 *Macbridea* has been shown to be positively correlated with a decrease in the chromosome sizes
413 (Cantino, 1985a). Alternatively, the origin of the base chromosome number in *Physostegia* has
414 been posited to be a result of fusion of unreduced gametes (x=9 and x=10) or of polyploidization
415 and merger of normal gametes (Scheen et al., 2008). Hence, the chromosome number of 2*n*=38
416 in some *Physostegia* species may indicate tetraploidy, while species like *P. ledinghamii* and *P.*
417 *leptophylla* may be octoploids (2*n*=76; Cantino, 1985a). If this hypothesis is correct, *Warnockia*
418 is a good candidate to be one of the progenitors of *Physostegia*, based on its chromosome
419 number (2*n*=20) and overall morphological similarity. The other progenitor, with 2*n*=18, is most
420 likely extinct. One can hypothesize that this missing parent of *Physostegia* was the source of its
421 actinomorphic, 5-lobed calyx, a feature not found in any other extant genus of Synandreae.
422 *Macbridea* and *Synandra* would seem to be candidates for the missing parent based solely on
423 their chromosome number. However, there is no morphological evidence for a link between
424 *Synandra* and *Physostegia*. *Macbridea* and *Physostegia* do share a few character states that are
425 not found in *Warnockia*: a rhizomatous perennial habit, mid-stem leaves lacking capitate-
426 glandular hairs, and filaments roughly equal in length (Turner, 1996), suggesting that *Macbridea*
427 might be the other progenitor of *Physostegia*. However, all three of these character states are so

widespread in Lamioideae that they could easily be plesiomorphic in Synandreae and thus do not provide convincing evidence for a special relationship between *Macbridea* and *Physostegia*.

Our phylogeny based on the concatenated dataset assembles all *Physostegia* species into two clades (labeled A and B in Fig. 2). Although we are aware of no morphological synapomorphies for either of these clades, previous morphological studies (Cantino, 1982) have suggested interspecific relationships that receive support in some of our analyses. For example, a close relationship between *P. pulchella* and *P. angustifolia* is indicated (within clade A in Fig. 2 and strongly supported in Fig. 3), corroborating Cantino's (1982) morphology-based studies. One of the two octoploid species, *P. leptophylla*, which was speculated to be a polyploid derivative of a hybrid between *P. purpurea* and *P. viriginiana* in previous studies (Cantino, 1982; Scheen et al., 2008), groups with both of these species in one of our analyses (Fig. 1a) and with *P. purpurea* in others (Figs. 1b, 1c, 2). However, our results provide only modest support for this hypothesis because *P. leptophylla* also groups with *P. longisepala* in three analyses (Fig. 1a, 1b, 2) and with *P. digitalis* and *P. parviflora* in the multi-locus coalescence-based analysis (Fig. 3). Cantino (1981) and Scheen et al. (2008) also hypothesized a hybrid origin for the other octoploid species, *P. ledinghamii*, involving *P. virginiana* and *P. parviflora* as parents. Although *P. ledinghamii* and *P. virginiana* group within the same clade (A) in the concatenated phylogeny (Fig. 2), our study does not support a close relationship among these three species. On the other hand, a close relationship is suggested between *P. ledinghamii* and *P. longisepala* (Fig. 1b and 2), a relationship also shown in Scheen et al.'s (2008) study, where these two species grouped closely in the 5S-NTS tree. This relationship, however, is not supported by cpDNA, morphology, or geographic distribution. It is also worth noting that a second *P. longisepala* clone groups with *P. leptophylla* (clade B in Fig. 2) It is possible that the different phylogenetic positions of these

two *P. longisepala* clones reflect paternal ancestries of the involved species, but further studies with greater number of clones and individuals are needed to support such a hypothesis.

Biogeography of Synandreae: Migration and diversification

Synandra and *Macbridea*, which together form a sister clade to the other three genera of Synandreae, are largely confined to southeastern USA, but the range of *Synandra* also extends north into northern West Virginia and central Ohio and Indiana (Cantino, 1985b). *Brazoria* and *Warnockia* are found in south-central US; *Brazoria* is endemic to the eastern half of Texas and *Warnockia* occurs mostly in central Texas, with a few outliers in eastern Texas, southern Oklahoma and Coahuila in northern Mexico (Turner, 1996). The most widespread genus, *Physostegia* with 12 species (Cantino, 1982), is extensively distributed across North America, stretching from northern Canada to northern Mexico. However, seven out of the 12 species occur in a region comprising southeastern Texas and southwestern Louisiana, suggesting that this area is the center of diversity for this genus (Cantino, 1982). Our ancestral area reconstruction (Figure not shown) supports a scenario in which southeastern US, including Texas, is the area where the most recent common ancestor (MRCA) of Synandreae most likely evolved. From this original center of diversity, migration and diversification took place northward and westward.

Roy and Lindqvist (2015) investigated the biogeography of the tribes of Lamioideae, and their fossil-based molecular dating suggests that the MRCA of Synandreae diversified in the New World (NW) from Old World (OW) relatives sometime during the Mid-Miocene epoch. Since Synandreae appear to be phylogenetically isolated from other lamioid groups, and no well-supported extant sister group of Synandreae has been determined (Scheen et al., 2008, 2010; Bendiksby et al., 2011; Roy and Lindqvist, 2015), it is likely that several lineages,

474 phylogenetically intercalated between Synandreae and the other extant Lamioideae, have
475 undergone extinction. The Miocene epoch was characterized by extreme climatic optima, with
476 major long-term cooling strongly affecting the distribution and establishment of modern
477 terrestrial biomes (Kurschner et al., 2008). Atmospheric carbon dioxide variations during the
478 Miocene led to changes in the structure and productivity of terrestrial biomes by affecting their
479 photosynthesis (Flower and Kennett, 1994). East Antarctic ice sheet growth and polar cooling
480 also had large effects on global carbon cycling and on the terrestrial biosphere, including
481 aridification of mid-latitude continental regions (Kurschner et al., 2008). Such cool-dry cycles of
482 the Miocene epoch could have caused the extinction of some of the closest OW relatives of
483 Synandreae. Numerous biogeographic studies have emphasized the origins and diversification
484 patterns of widely disjunct plant groups in the Northern Hemisphere (NH) (Tiffney and
485 Manchester, 2001; Wen, 1999, 2001; Donoghue et al., 2001; Donoghue and Smith, 2004), and
486 three different biogeographic patterns have been hypothesized for their current distributions. The
487 first pattern suggests that there was an extinction of a once-continuous Arcto-Tertiary, Tethyan
488 or boreal flora, giving rise to the current disjunct distributions of some genera (Mao et al., 2012;
489 Sun et al., 2001). The second pattern posits that a majority of genera showing disjunct
490 distributions had their origin on the Qinghai Tibetan Plateau (QTP) and adjacent regions in Asia,
491 later migrating to and colonizing other NH regions (including the Arctic), where they gave rise to
492 derivative species (Xu et al., 2010; Zhang and Fritsch, 2010; Zhang et al., 2009). The third
493 pattern assumes the origin of the groups in other regions of the world with subsequent
494 diversifications on the QTP after the arrival of their ancestors there (Liu et al., 2002; Tu et al.,
495 2010). The absence of a clear extant sister group of Synandreae, presumably due to extinction, is
496 most consistent with the first pattern.

497

498 **Comparison with Stachydeae: Exploring causes for the restricted distribution of most**
499 **genera of Synandreae**

500 Stachydeae and Synandreae, the only two lamioid tribes that include NW members,
501 independently colonized the NW via separate migratory events. Members of Stachydeae
502 (belonging to the genus *Stachys*) colonized the NW twice, once during the mid-Miocene and the
503 other during the Pliocene, whereas Synandreae colonized the NW only once during the mid-
504 Miocene (Roy et al., 2013). While the nearest OW ancestors of NW Stachydeae can be
505 confidently inferred, with African and East Asian *Stachys* species grouping closely with the
506 temperate NA and Hawaiian taxa (Lindqvist and Albert, 2002; Roy et al., 2013; 2015), the
507 closest extant OW relatives of Synandreae have not been determined. These two tribes contrast
508 sharply in their pattern of diversification within the NW. NW Stachydeae rapidly migrated and
509 radiated in different parts of temperate NA, Mesoamerica, and South America, and they also
510 successfully colonized and diversified in the Hawaiian archipelago, giving rise to one of the
511 largest clades endemic to the islands (Lindqvist and Albert, 2002; Roy et al., 2013; 2015).
512 Members of Synandreae, on the other hand, split into 19 species in five genera but did not spread
513 outside of North America, with most species restricted to the southeastern and south-central US.
514 The range of one species of *Brazoria* extends into northern Mexico, and one species of
515 *Physostegia* has reached northern Canada.

516 A number of factors, both biological and ecological, could have led to the disparities in
517 the colonization and diversification patterns of the members of these two groups of NW
518 endemics. Polyploidy seems to be one of the leading factors contributing to the widespread
519 distribution of NW Stachydeae (2*n*= 32-84), as well the genus *Physostegia* (2n=38, 76) within

Synandreae. Numerous studies have been performed on polyploid genome evolution, and these have shown that phenomena such as substantial intra-genomic rearrangement and altered gene regulatory relationships can lead to a certain degree of evolutionary flexibility, allowing for improved success in colonization and establishment in novel habitats (Levin, 2002; Soltis and Soltis, 2000; Wendel, 2000; Wendel and Doyle, 1998; Tate et al., 2005). The high-polyploid members of NW Stachydeae and the Hawaiian mints seem to have rapidly radiated and established themselves in novel habitats, carving out new niches, likely as a result of hybridization and polyploidization (Roy et al. 2013; 2015). This includes *Stachys* species derived from both the Miocene and Pliocene colonizations of the NW. We observe similar trends within the genus *Physostegia*, the only polyploid genus of Synandreae, which has been more successful in colonizing a broad geographic range within temperate NA than its diploid relatives, which have remained largely limited to warm-temperate southeastern and south-central NA.

Abiotic factors could also have played an important role in the current restricted distribution of Synandreae. Glacial climates were extremely variable, and it has been postulated that terrestrial organisms respond individually to climate changes (Huntley & Webb, 1989). A consensus opinion gleaned from palaeoecological studies is that individual species respond to changing environments through their geographical distributions (Webb, 1992). Glacial conditions have helped shape the modern distributions of most plant and animal species (Willis and Whittaker, 2000). Local flora and fauna during glaciations could have survived only within certain protected localities, referred to as “refugia” (Provan and Bennett, 2008). These refugia provided stable microclimates for species to persist. Southeastern US has been highlighted as a refugium for numerous other species (reviewed by Soltis et al., 2006). The geographic distribution of plant species in southeastern USA has been mainly shaped in an east to west

pattern by three geographical factors—the Apalachicola River discontinuity, the Tombigbee River discontinuity, and the Appalachian Mountains discontinuity—leading to endemism and climatically determined glacial refugia (Soltis et al., 2006), especially during the Pliocene and Pleistocene. Swenson and Howard (2005) cited instances of contact zones in Alabama, where closely related species or populations emerging from glacial refugia in Florida and eastern Texas/western Louisiana intermingled. However, due to differential tolerance of climatic and edaphic conditions, species emerging from these refugia became fragmented in their distributions, the less tolerant species thriving only within isolated pockets of favorable abiotic conditions. The spread of *Physostegia*, the most widespread genus of Synandreae, may be due in part to its ability to grow in a broad range of edaphic conditions. Cantino (1982) stated that this genus is tolerant of a wide range of soil acidity conditions. As a result of millions of years of weathering and acidification, southeastern NA is largely characterized by acidic, infertile soils leading to relatively small areas of rich, circum-neutral soils (Manthey et al., 2011). Hence, other genera of Synandreae, which are not as tolerant of acidic soil conditions, may have remained restricted to such pockets of fertile soil, resulting in their current, more limited, ranges. However, further studies are required to document and substantiate this hypothesis, and to investigate other possible causes, such as anthropogenic alterations of habitat conditions, loss of pollinators, and competition with invasive species, that may also have influenced the current restricted distributions of most species of Synandreae.

Supplemental Information

The raw alignment files in FASTA format for the three loci, *PHOT1*, *PHOT2*, and *COR*.

Additional Information and Declarations

Competing Interests

Charlotte Lindqvist is an Academic Editor for PeerJ. The authors declare that there are no competing interests.

Author Contributions

Tilottama Roy conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.

Nathan S. Catlin and Drake M. G. Garner performed the experiments, contributed reagents/materials.

Philip D. Cantino wrote the paper and reviewed drafts of the paper.

Anne-Cathrine Scheen conceived and designed the experiments, contributed reagents/materials, reviewed drafts of the paper.

Charlotte Lindqvist conceived and designed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.

DNA Deposition

The DNA sequences have been deposited in GenBank with the following accessions numbers: KT716867-KT717052

Funding

The funding for this work was provided by the College of Arts and Sciences, The State University of New York at Buffalo.

Acknowledgements

The authors thank the curators at the BISH, C, GH, LL, O, TEX, UNA, UPS, and US herbaria for permission to perform destructive work on herbarium material used as part of this study.

REFERENCES

Abu-Asab, M. S. and Cantino, P. D. 1987. Phylogenetic implications of leaf anatomy in subtribe Melittidinae (Labiatae) and related taxa. J. Arnold Arb. 68: 1-34.

Abu-Asab, M. S., and Cantino, P. D. 1994. Systematic implications of pollen morphology in subfamilies Lamioideae and Pogostemonoideae (Labiatae). *Annals of the Missouri Botanical Garden*. 653-686.

Bentham, G. 1876. Labiatae. In: Bentham, G., Hooker, J.D. (Eds.), Genera plantarum, vol. 12. Reeve, London, pp. 1160–1223.

Bendiksby, M., Thorbek, L., Scheen, A. C., Lindqvist, C., and Ryding, O. 2011. An updated phylogeny and classification of Lamiaceae subfamily Lamioideae. *Taxon* 60(2), 471-484.

Bentham, G. 1848. Labiatae. In: De Candolle, A. (Ed.), Prodromus systematis naturalis regni vegetabilis. Treuttel and Wurtz, Paris, pp. 27–603.

Briggs, W. R., C. F. Beck, A. R. Cashmore, J. M. Christie, J. Hughes, J. A. Jarillo, and T. Kagawa. 2001. The phototropin family of photoreceptors. *Plant Cell* 13, 993 – 997.

Barkan, A., and Small, I. 2014. Pentatricopeptide repeat proteins in plants. *Annual Reviews in Plant Biology*. 65, 415-442.

Briquet, J.1895–1897. Labiatae. In: Engler, H.G.A., Prantl, K.A.E. (Eds.), Die natürlichen Pflanzenfamilien. Engelmann, W., Berlin pp. 183–375.

Bruen, T. C., Philippe, H., and Bryant, D. 2006. A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172(4), 2665-2681

Brummitt, R. K., Pando, F., Hollis, S., and Brummitt, N. A. 2001. World Geographical Scheme for Recording Plant Distributions. Available at http://www.nhm.ac.uk/hosted_sites/tdwg/TDWG_geo2.pdf.

Cantino, P. D. 1981. Change of status for *Physostegia virginiana* var. *ledinghamii* (Labiatae) and evidence for a hybrid origin. *Rhodora.* 83, 111–118.

Cantino, P. D. 1982. A monograph of the genus *Physostegia* (Labiatae). Contributions from the Gray Herbarium of Harvard University, 1-105.

Cantino, P.D. 1985a. Chromosome studies in subtribe Melittidinae (Labiatae) and systematic implications. *Systematic Biology*. 10, 1–6.

Cantino, P.D. 1985b. Facultative autogamy in *Synandra hispidula* (Labiatae). *Castanea*. 50, 105–111.

Cantino, P.D. 1990. The phylogenetic significance of stomata and trichomes in the Labiatae and Verbenaceae. *J. Arnold Arb.* 71, 323–370.

Chen, Y.-P., Li, B., Olmstead, R. G., Cantino, P. D., Liu, E.-D., and Xiang, C.-L. 2014. Phylogenetic placement of the enigmatic genus *Holocheila* (Lamiaceae) inferred from plastid DNA sequences. *Taxon* 63, 355-366.

Christie , J. M. , Reymond, P., Powell, G. K., Bernasconi, P., Raibekas, A. A., Liscum, E., and Briggs, W. R. 1998. *Arabidopsis* NPH1: A flavoprotein with the properties of a photoreceptor for phototropism. *Science* 282, 1698 – 1701 .

Crowl, A. A., E., Mavrodiev, G., Mansion, R., Haberle, A., Pistarino, G., Kamari, D., Phitos, T., Borsch, and N., Cellinese. 2014. Phylogeny of Campanuloideae (Campanulaceae) with emphasis on the utility of nuclear pentatricopeptide repeat (*PPR*) Genes. *PloS One* 9(4), e94199.

Donoghue, M. J., Bell, C. D., and Li, J. 2001. Phylogenetic patterns in Northern Hemisphere plant geography. *International Journal of Plant Sciences.* 162(S6), S41-S52.

638 Donoghue, M. J., and Smith, S. A. 2004. Patterns in the assembly of temperate forests around the
639 Northern Hemisphere. *Philosophical Transactions of the Royal Society of London, B:*
640 *Biological Sciences* 359(1450), 1633-1644.
641 Drummond, A.J., and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling
642 trees. *BMC Evolutionary Biology* 7, 214.
643 Curto, M.A., Puppo, P., Ferreira, D., Nogueira, M., and Meimberg, H. 2012. Development of
644 phylogenetic markers from single-copy nuclear genes for multi locus, species level
645 analyses in the mint family (Lamiaceae). *Molecular Phylogenetics and Evolution.* 63,
646 758-767.
647 Fitch, W. M. 1970. Distinguishing homologous from analogous proteins. *Systematic Zoology*. 19,
648 99 – 113.
649 Flower, B. P., and Kennett, J. P. 1994. The middle Miocene climatic transition: East Antarctic
650 ice sheet development, deep ocean circulation and global carbon cycling.
651 *Palaeogeography, Palaeoclimatolog, Palaeoecology.*108(3), 537-555.
652 Geddy, R., and Brown, G. G. 2007. Genes encoding pentatricopeptide repeat (*PPR*) proteins are
653 not conserved in location in plant genomes and may be subject to diversifying selection.
654 *BMC Genomics* 8(1), 130.
655 Gill, L.S., 1981. Chromosomal evolution and incidence of polyploidy in the Canadian Labiatae.
656 *International Review of Cytology*. 4, 331–339
657 Govaerts, R., Dransfield, J., Zona, S.F, Hodel, D.R., and Henderson, A. 2013. World checklist of
658 Lamiaceae and Verbenaceae. Facilitated by the Royal Botanic Gardens, Kew, Richmond.
659 Holmgren, P. K., Holmgren, N.H., and Barrett, L.C. (1990). Index Herbariorum. Part I. The
660 herbaria of the world. New York Botanical Garden Press, Bronx, New York, USA.

661 Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis
662 program for Windows 95/98/NT. In Nucleic acids symposium series (Vol. 41, pp. 95-98).
663 Harley, R. M., Atkins, S., Budantsev, A. L., Cantino, P.D., Conn, B. J., Grayer, R. Harley, M.
664 M., de Kok, R., Krestovskaya, T., Morales, R., Paton, A. J., Ryding, O., and Upson, T.
665 2004. Labiateae. In J. W. Kadereit (Ed.), VII Flowering Plants - Dicotyledons. Lamiales
666 (except Acanthaceae including Avicenniaceae) (Vol. 7). Berlin Heidelberg: Springer
667 Huelsenbeck, J. P. and Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees.
668 *Bioinformatics* 17, 754-755.
669 Hughes, C. E., Eastwood, R. J., and Bailey, C. D. 2006. From famine to feast? Selecting nuclear
670 DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical*
671 *Transactions of the Royal Society of London, B: Biological Sciences.* 361(1465), 211-
672 225.
673 Huson, D. H. 1998. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics*
674 14(1), 68-73.
675 Huson, D. H., and Bryant, D. 2006. Application of phylogenetic networks in evolutionary
676 studies. *Molecular Biology and Evolution*. 23(2), 254-267.
677 Huntley, B., and Webb III, T. 1989. Migration: species' response to climatic variations caused by
678 changes in the earth's orbit. *Journal of Biogeography.* 5-19.
679 Jarillo, J. A., H. Gabrys, J. Capel, J. M. Alonso, J. R. Ecker, and A. R. Cashmore. 2001.
680 Phototropin-related NPL1 controls chloroplast relocation induced by blue light. *Nature*
681 410, 952 – 954.

682 Kagawa, T., Sakai, T., Suetsugu, N., Oikawa, K., Ishiguro, S., Kato, T., and Tabata, S. 2001.
683 *Arabidopsis* NPL1: A phototropin homolog controlling the chloroplast high-light
684 avoidance response. *Science* 291, 2138 – 2141.
685 Kinoshita, T., Doi, M., Suetsugu, N., Kagawa, T., Wada, M., and Shimazaki, K. 2001. *PHOT*1
686 and *PHOT*2 mediate blue light regulation of stomatal opening. *Nature* 414: 656 – 660 .
687 Kürschner, W. M., Kvaček, Z., and Dilcher, D. L. 2008. The impact of Miocene atmospheric
688 carbon dioxide fluctuations on climate and the evolution of terrestrial ecosystems.
689 *Proceedings of the National Academy of Sciences*. 105(2), 449-453.
690 Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H.,
691 Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and
692 Higgins, D.G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947-2948.
693 Leache, A. D., Harris, R. B., Rannala, B., and Yang, Z. H. (2014). The influence of gene flow on
694 species tree estimation: a simulation study. *Systematic Biology.* 63, 17–30.
695 Lersten, N. R., and Curtis, J. D. 1998. Foliar idioblasts in *Physostegia virginiana* (Lamiaceae). *J.*
696 *Torrey Bot. Soc.* 125, 133–137.
697 Levin , D. A. 2002. The role of chromosomal change in plant evolution. Oxford University
698 Press, New York, New York, USA.
699 Li, B., Xu, W., Tu, T., Wang, Z., Olmstead, R. G., Peng, H., Francisco-Ortega, J., Cantino, P.D.,
700 and Zhang, D. 2012. Phylogenetic position of *Wenchengia* (Lamiaceae): a taxonomically
701 enigmatic and critically endangered genus. *Taxon* 61, 392-401.
702 Lindqvist, C., and Albert, V. A. 2002. Origin of the Hawaiian endemic mints within North
703 American *Stachys* (Lamiaceae). *American Journal of Botany*. 89, 1709-1724

704 Liu, J. Q., Gao, T. G., Chen, Z. D., Lu, A. M. 2002 Molecular phylogeny and biogeography of
705 the Qinghai-Tibet Plateau endemic *Nannoglottis* (Asteraceae). *Molecular Phylogenetics*
706 *and Evolution.* 23, 307–325.
707 Manthey, M., Fridley, J. D., and Peet, R. K. 2011. Niche expansion after competitor extinction?
708 A comparative assessment of habitat generalists and specialists in the tree floras of
709 south-eastern North America and south-eastern Europe. *Journal of Biogeography.* 38(5),
710 840-853
711 Mao, K., Milne, R. I., Zhang, L., Peng, Y., Liu, J., Thomas, P., and Renner, S. S. 2012.
712 Distribution of living Cupressaceae reflects the breakup of Pangea. *Proceedings of the*
713 *National Academy of Sciences.* 109(20), 7793-7798.
714 Miller, M. A., Pfeiffer, W., and Schwartz, T. 2010. "Creating the CIPRES Science Gateway for
715 inference of large phylogenetic trees" in Proceedings of the Gateway Computing
716 Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1 - 8.
717 Mort, M. E., and Crawford, D. J. 2004. The continuing search: low-copy nuclear sequences for
718 lower-level plant molecular phylogenetic studies. *Taxon* 53, 257–261.
719 Nixon, K. C. 1999. WinClada ver. 1.0000 Published by the author, Ithaca, NY, USA.
720 Nylander, J., Olsson, U., Alstrom, P., and Sanmartin, I. 2008. Accounting for Phylogenetic
721 Uncertainty in Biogeography: A Bayesian Approach to Dispersal-Vicariance Analysis of
722 the Thrushes (Aves: *Turdus*). *Systematic Biology* 57, 257 - 268.
723 Posada, D. 2008. jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution.*
724 25, 1253-1256.
725 Provan, J., and Bennett, K. D. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends*
726 *in Ecology and Evolution.* 23(10), 564-571.

727 Rambaut, A. 2008. FigTree v1.1.1: Tree figure drawing tool. Available:
728 http://tree.bio.ed.ac.uk/software/figtree/.
729 Ronquist, F. 1997. Dispersal-vicariance analysis: A new approach to the quantification of
730 historical biogeography. *Systematic Biology* 46, 195-203.
731 Roy, T., Chang, T., Lan, T., and Lindqvist, C. 2013. Phylogeny and biogeography of New World
732 Stachydeae (Lamiaceae) with emphasis on the origin and diversification of Hawaiian and
733 South American taxa. *Molecular Phylogenetics and Evolution.* 69, 535-551.
734 Roy, T., Cole, L., Chang, T., and Lindqvist, C. 2015. Untangling reticulate evolutionary
735 relationships among New World and Hawaiian mints (Stachydeae, Lamiaceae).
736 *Molecular Phylogenetics and Evolution.* 89, 46-62.
737 Roy, T. and Lindqvist, C. 2015 New insights into the evolutionary history of the subfamily
738 Lamioideae (Lamiaceae) based on a pentatricopeptide repeat (*PPR*) nuclear DNA region.
739 *American Journal of Botany*. 102 (10), 1-15.
740 Ryding, O. 1994. Pericarp structure in the subtribe Melittidinae (Lamiaceae-Lamioideae) and its
741 systematic implications. *Botanische Jahrbücher fur Systematik, Pflanzengeschichte und*
742 *Pflanzengeographie* 115, 547-555.
743 Salmaki, Y., Zarre, S., Ryding, O., Lindqvist, C., Bräuchler, C., Heubl, G., and Bendiksby, M.
744 2013. Molecular phylogeny of tribe Stachydeae (Lamiaceae subfamily Lamioideae). .
745 *Molecular Phylogenetics and Evolution.* 69(3), 535-551.
746 Sang, T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical*
747 *Reviews in Biochemistry and Molecular Biology*. 37(3), 121-147.

748 Scheen, A. C., Lindqvist, C., Fossdal, C. G., and Albert, V. A. 2008. Molecular phylogenetics of
749 tribe Synandreae, a North American lineage of lamioid mints (Lamiaceae). *Cladistics* 24,
750 299-314.
751 Scheen, A. C., Bendiksby, M., Ryding, O., Mathiesen, C., Albert, V. A., and Lindqvist, C. 2010.
752 Molecular phylogenetics, character evolution, and suprageneric classification of
753 Lamioideae (Lamiaceae). *Annals of the Missouri Botanical Garden.* 97, 191-217
754 Seehausen, O. 2004. Hybridization and adaptive radiation. *Trends in Ecology and Evolution*.
755 19(4), 198-207.
756 Soltis , P. S. , and Soltis, D. E. 2000. The role of genetic and genomic attributes in the success of
757 polyploids. *Proceedings of the National Academy of Sciences.* 97, 7051 – 7057.
758 Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., and Soltis, P. S. 2006. Comparative
759 phylogeography of unglaciated eastern North America. *Molecular Ecology.*15(14), 4261-
760 4293.
761 Stamatakis, A., Hoover, P. and Rougemont, J. 2008. A Rapid Bootstrap Algorithm for the
762 RAxML Web Servers. *Systematic Biology.* 57, 758-771.
763 Sun, H., McLewin, W., and Fay, M. F. 2001. Molecular phylogeny of *Helleborus*
764 (Ranunculaceae), with an emphasis on the East Asian-Mediterranean disjunction. *Taxon*
765 50, 1001–1018.
766 Swenson, N.G. and Howard, D.J. 2005. Clustering of contact zones, hybrid zones, and
767 phylogeographic breaks in North America. *The American Naturalist*. 166(5), 581-591.
768 Tate , J. A. , Soltis, D. E., and Soltis, P. S. 2005. Polyploidy in plants. In T. R. Gregory [ed.]. *The*
769 *evolution of the genome*, 371 – 426. Elsevier Academic Press, San Diego, California,
770 USA.

771 Thomson, R. C., Shedlock, A. M., Edwards, S. V., and Shaffer, H. B. 2008. Developing markers
772 for multilocus phylogenetics in non-model organisms: a test case with turtles. *Molecular*
773 *Phylogenetics and Evolution.* 49(2), 514-525
774 Tiffney, B. H., and Manchester, S. R. 2001. The use of geological and paleontological evidence
775 in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary.
776 *International Journal of Plant Sciences.*162(S6), S3-S17.
777 Tu, T. Y., Volis, S., Dillon, M. O., Sun, H., Wen, J. 2010. Dispersals of Hyoscyameae and
778 Mandragoreae (Solanaceae) from the New World to Eurasia in the early Miocene and
779 their biogeographic diversification within Eurasia. *Molecular Phylogenetics and*
780 *Evolution.* 57, 1226–1237.
781 Turner, M.W. 1996. Systematic study of the genus *Brazoria* (Lamiaceae), and *Warnockia*
782 (Lamiaceae), a new genus from Texas. *Plant Systematics and Evolution*. 203, 65–82.
783 Wagstaff, S.J., Hickerson, L., Spangler, R., Reeves, P. A., and Olmstead, R. G. 1998. Phylogeny
784 in Labiatae s.l., inferred from cpDNA sequences. *Plant Systematics and Evolution* 209,
785 265-274.
786 Webb, T. and Bartlein, P. J. 1992. Global changes during the last 3 billion years: climatic
787 controls and biotic responses. *Annual Review of Ecology and Systematics*. 23, 141–173
788 Wendel, J. F., Schnabel, A., and Seelanan, T. 1995. Bidirectional interlocus concerted evolution
789 following allopolyploid speciation in cotton (*Gossypium)*. *Proceedings of the National*
790 *Academy of Sciences.* 92(1), 280-284.
791 Wendel, J. F. and Doyle, J. J. 1998. Phylogenetic incongruence: Window into genome history
792 and molecular evolution. Pages 265–296 *in* Molecular systematics of plants, II: DNA
793 sequencing (D. Soltis, P. Soltis, and J. Doyle, eds.) Kluwer Academic Publishers, Boston.

794 Wendel , J. F. 2000. Genome evolution in polyploids. *Plant Molecular Biology*. 42, 225 – 249.

795 Wen, J. 1999. Evolution of eastern Asian and eastern North American disjunct distributions in

796 flowering plants. *Annual Reviews in Ecology, Evolution and Systematics*. 421-455.

797 Wen, J. 2001. Evolution of eastern Asian–eastern North American biogeographic disjunctions: a

798 few additional issues. *International Journal of Plant Sciences*. 162(S6), S117-S122.

799 Willis, K. J., and Whittaker, R. J. 2000. The refugial debate. *Science* (Washington) 287(5457),

800 1406-1407.

801 Xu, T. T., Abbott, R. J., Milne, R. I., Mao, K., Du, F. K., Wu, G. L., Ciren, Z. X., Miehe, G., Liu,

802 J. Q. 2010. Phylogeography and allopatric divergence of cypress species (*Cupressus* L.)

803 in the Qinghai-Tibetan Plateau and adjacent regions. *BMC Evolutionary Biology.*10, 194.

804 Yu, Y., Harris, A. J., He, X. J. 2010. S-DIVA (statistical dispersal-vicariance analysis): a tool for

805 inferring biogeographic histories. *Molecular Phylogenetics and Evolution* 56(2), 848-

806 850.

807 Yuan, Y. W., and Olmstead, R. G. 2008. Evolution and phylogenetic utility of the *PHOT* gene

808 duplicates in the *Verbena* complex (Verbenaceae): dramatic intron size variation and

809 footprint of ancestral recombination. *American Journal of Botany.* 95(9), 1166-1176.

810 Yuan, Y. W., Liu, C., Marx, H. E., and Olmstead, R. G. 2009. The pentatricopeptide repeat

811 (*PPR*) gene family, a tremendous resource for plant phylogenetic studies. *New*

812 *Phytologist*. 182(1), 272-283.

813 Yuan, Y. W., Liu, C., Marx, H. E., and Olmstead, R. G. 2010. An empirical demonstration of

814 using pentatricopeptide repeat (*PPR*) genes as plant phylogenetic tools: Phylogeny of

815 Verbenaceae and the *Verbena* complex. *Molecular Phylogenetics and Evolution.* 54(1),

816 23-35.

817 Zhang, M.L., Kang, Y. and Yang, J.B. 2009. A preliminary molecular historical biogeography of
818 *Caragana* (Leguminosae) based on ITS and trnL-F DNA sequence data. *Journal of Arid*
819 *Environments*. 1, 64–73
820 Zhang, M. L. and Fritsch, P. W. 2010. Evolutionary response of *Caragana* (Fabaceae) to
821 Qinghai-Tibetan Plateau uplift and Asian interior aridification. *Plant Systematics and*
822 *Evolution*. 288, 191–199.

823
824
825
826
827
828
829
830
831
832
833
834
835
836

Table 1(on next page)

List of taxa and voucher information.

Abbreviations of herbaria follow the Index Herbariorum.

1 **Table 1**. List of taxa and voucher information. Herbaria abbreviations follow the Index Herbariorum.

Taxon names	Voucher information	Geographic distribution
Achyrospermum radicans Gürke in H.G. A. Engler	E. Farkas & T. Pocs 86604 (UPS)	Tanzania
Acrotome inflata Benth.	G.L. Maggs & L. Guarino 1072 (UPS)	Namibia
Acrotome pallescens Benth.	I. Ortendahl 105 (UPS)	Namibia
Ballota nigra L. subsp. *ruderalis* (Sw.) Briq.	M. Bendiksby & A.-C. Scheen (O)	Greece
Ballota pseudodictamnus (L.) Benth.	M. Bendiksby & A.-C. Scheen 0420 (O)	Greece
Betonica macrantha K. Koch	D. McNeal et al. 161 (C)	Georgia
Brazoria arenaria Lundell	M.W. Turner 25 (TEX)	USA
Brazoria enquistii M.W. Turner	M.W. Turner 61 (TEX)	Texas, USA
Brazoria truncata (Benth.) Engelm. & A.Gray var**.** *trunacata*	D.S. Corell 1605 (GH)	Texas, USA
Craniotome furcata (Link) Kuntze	O. Polunin et al. 5638 (UPS)	Nepal
Eremostachys labiosa Bunge.	V. Goloskokov s.n., 15 May 1963 (C)	CCCP, Kazakhstan
Eriophyton wallichii Benth.	Stainton et al. 7748 (UPS)	Nepal
Eurysolen gracilis Prain.	R. Geesink, P. Hiepko & C. Phengklai 8240 (C)	Thailand
Galeopsis pyrenaica Bartl.	P. Montserrat & al. 141487 (C)	Spain
Gomphostemma javanicum (Blume) Benth.	R.G. Olmstead 93-38	S. China to SE Asia
Leonotis nepetifolia (L.) R. Br. var. *nepetifolia*	R. Abdallah et al. 493 (UPS)	Tanzania

Leucas inflata Benth.	V. Goloskokov s.n., 15 May 1963 (C)	Ethiopia
Macbridea caroliniana (Walter) S.F.Blake	R.K. Godfrey & R.M. Tryon 741 (GH)	USA
Marrubium peregrinum L.	A. Strid 33875 (C)	Greece
Phlomis fruticosa L.	C. Mathiesen & J.M. Taylor 81 (National Collection of *Phlomis*, UK)	Sardegna (Italy) to Transcaucasus
Phlomis tuberosa L.	C. Mathiesen & J.M. Taylor 88 (National Collection of *Phlomis*, UK)	EC Europe to China and Mongolia
Phyllostegia kaalaensis St. John	S. Perlman 6117 (BISH)	Hawaii/O'ahu
Physostegia angustifolia Fernald	C.L. Lundell & A.A. Lundell 16031 (US)	Texas, USA
Physostegia correllii (Lundell) Shinners	D.S. Corell & I.M. Johnston 19427 (LL)	Texas, USA
Physostegia digitalis Small	P.D. Cantino 1076 (GH)	Texas, USA
Physostegia godfreyi P.D.Cantino	R.K. Godfrey 77073 (GH)	Florida, USA
Physostegia ledinghamii (Boivin) P.D.Cantino	V.L. Harms 34491 (GH)	Saskatchewan, Canada
Physostegia longisepala P.D.Cantino	L.E. Brown 13523 (TEX)	Texas, USA
Physostegia leptophylla Small	P.D. Cantino 1026 (GH)	Florida, USA
Physostegia parviflora Nutt. ex A.Gray	M. Mooar 13667 (GH)	Montana, USA
Physostegia pulchella Lundell	Wm.F. Mahler 8530 (GH)	Texas, USA
Physostegia virginiana (Walter) S.F.Blake	P.D. Cantino 1007 (GH)	Florida, USA
Scutellaria hirta Sm.	M. Bendiksby & A.-C. Scheen 0411 (O)	Greece

Stachys chamissonis Benth.	C. Lindqvist 10-02 (UB)	W. Canada to W. USA
Stachys sylvatica L.	C. Lindqvist & V.A. Albert 358 (UNA)	Macaronesia, Europe to W Himalaya (cultivar)
Stenogyne kamehamehae Wawra.	S. Permlan 6933 (BISH)	Hawaii
Synandra hispidula (Michx.) Baill.	V.E. McNeilus 97–143 (GH)	Tennessee, USA
Warnockia scutellarioides (Engelm. & A.Gray) M.W.Turner	M.W. Turner 67 (TEX)	Texas, USA

2

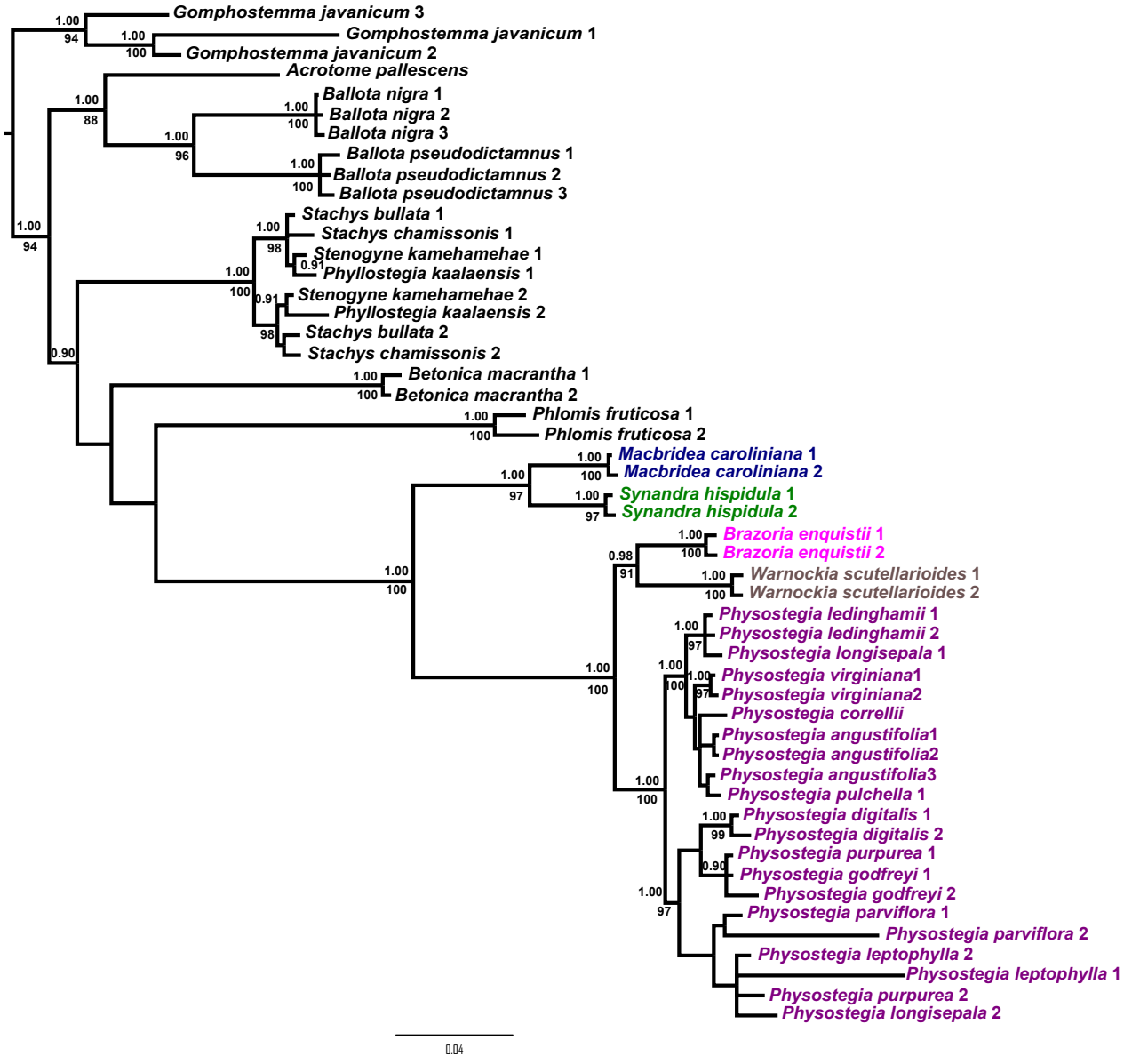
Figure 1(on next page)

Phylogenetic gene trees.

Bayesian 50% majority rule consensus trees obtained from analyses of (A) *PHOT1,* (B) *PHOT2*, and (C) *COR*, respectively. Bayesian posterior probability values ≥ 0.9 and maximum likelihood bootstrap support values ≥ 80 are shown above and below the nodes, respectively. Numbers following taxon names refer to different clones from PCR products.


A
Achyrospermum radicans 1
Achyrospermum radicans 2
Gomphostemma javanicum 1
Gomphostemma javanicum 2
Galeopsis pyrenaica 1
Galeopsis pyrenaica 2
Galeopsis pyrenaica 3
Leonotis nepetifolia 1
Leonotis nepetifolia 2
Acrotome inflata
Leucas inflata 1
Leucas inflata 2
Phlomis fruticosa 1
Phlomis fruticosa 2
Phlomis tuberosa 1
Phlomis tuberosa 2
Betonica macrantha 1
Marrubium peregrinum
Ballota nigra 1
Ballota nigra 2
Ballota pseudodictamnus 1
Ballota pseudodictamnus 2
Phyllostegia kaalaensis 1
Stachys chamissonis 2
Stenogyne kamehamehae 1
Stenogyne kamehamehae 2
Stachys bullata
Phyllostegia kaalaensis 2
Stachys chamissonis 1
Macbridea caroliniana 1
Macbridea caroliniana 2
Synandra hispidula 1
Synandra hispidula 2
Warnockia scutellarioides 1
Warnockia scutellarioides 2
Brazoria truncata 1
Brazoria arenaria 2
Brazoria arenaria 4
Brazoria arenaria 3
Brazoria arenaria 1
Brazoria truncata 3
Brazoria truncata 4
Brazoria enquistii 2
Brazoria truncata 2
Brazoria enquistii 1
Physostegia ledinghamii 1
Physostegia ledinghamii 2
Physostegia correllii 1
Physostegia correllii 2
Physostegia correllii 3
Physostegia longispela 1
Physostegia godfreyi 1
Physostegia godfreyi 2
Physostegia godfreyi 3
Physostegia angustifolia 1
Physostegia angustifolia 2
Physostegia angustifolia 3
Physostegia pulchella 1
Physostegia pulchella 2
Physostegia parviflora 1
Physostegia parviflora 2
Physostegia digitalis
Physostegia leptophylla 1
Physostegia longisepala 2
Physostegia virginiana 1
Physostegia virginiana 2
Physostegia leptophylla 2
Physostegia purpurea 1
Physostegia purpurea 2
0.07

B


Gomphostemma javanicum 3
Gomphostemma javanicum 1
Gomphostemma javanicum 2
Acrotome pallescens
Ballota nigra 1
Ballota nigra 2
Ballota nigra 3
Ballota pseudodictamnus 1
Ballota pseudodictamnus 2
Ballota pseudodictamnus 3
Stachys bullata 1
Stachys chamissonis 1
Stenogyne kamehamehae 1
Phyllostegia kaalaensis 1
Stenogyne kamehamehae 2
Phyllostegia kaalaensis 2
Stachys bullata 2
Stachys chamissonis 2
Betonica macrantha 1
Betonica macrantha 2
Phlomis fruticosa 1
Phlomis fruticosa 2
Macbridea caroliniana 1
Macbridea caroliniana 2
Synandra hispidula 1
Synandra hispidula 2
Brazoria enquistii 1
Brazoria enquistii 2
Warnockia scutellarioides 1
Warnockia scutellarioides 2
Physostegia ledinghamii 1
Physostegia ledinghamii 2
Physostegia longisepala 1
Physostegia virginiana1
Physostegia virginiana2
Physostegia correllii
Physostegia angustifolia1
Physostegia angustifolia2
Physostegia angustifolia3
Physostegia pulchella 1
Physostegia digitalis 1
Physostegia digitalis 2
Physostegia purpurea 1
Physostegia godfreyi 1
Physostegia godfreyi 2
Physostegia parviflora 1
Physostegia parviflora 2
Physostegia leptophylla 2
Physostegia leptophylla 1
Physostegia purpurea 2
Physostegia longisepala 2
0.04

C

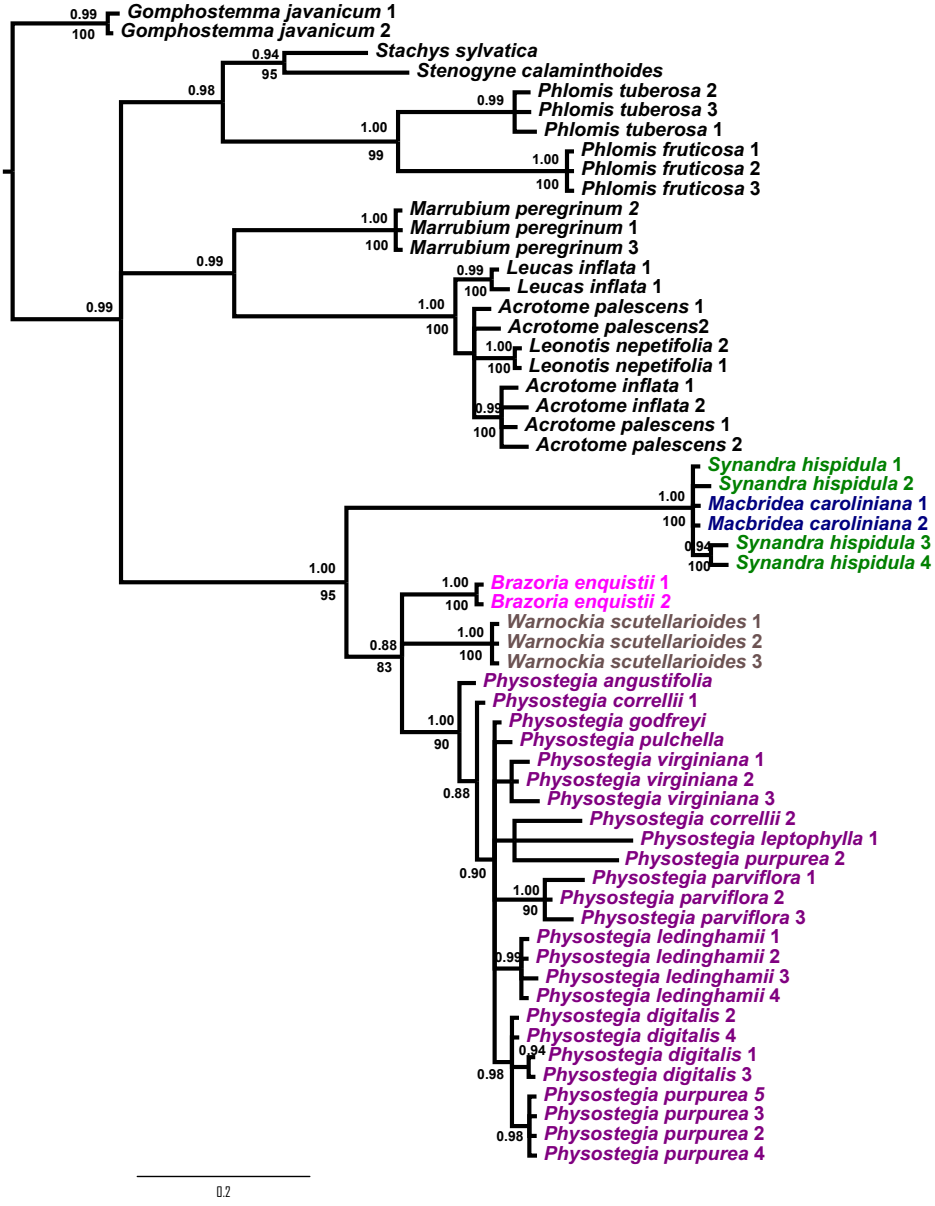

Gomphostemma javanicum 1
Gomphostemma javanicum 2
Stachys sylvatica
Stenogyne calaminthoides
Phlomis tuberosa 2
Phlomis tuberosa 3
Phlomis tuberosa 1
Phlomis fruticosa 1
Phlomis fruticosa 2
Phlomis fruticosa 3
Marrubium peregrinum 2
Marrubium peregrinum 1
Marrubium peregrinum 3
Leucas inflata 1
Leucas inflata 1
Acrotome palescens 1
Acrotome palescens2
Leonotis nepetifolia 2
Leonotis nepetifolia 1
Acrotome inflata 1
Acrotome inflata 2
Acrotome palescens 1
Acrotome palescens 2
Synandra hispidula 1
Synandra hispidula 2
Macbridea caroliniana 1
Macbridea caroliniana 2
Synandra hispidula 3
Synandra hispidula 4
Brazoria enquistii 1
Brazoria enquistii 2
Warnockia scutellarioides 1
Warnockia scutellarioides 2
Warnockia scutellarioides 3
Physostegia angustifolia
Physostegia correllii 1
Physostegia godfreyi
Physostegia pulchella
Physostegia virginiana 1
Physostegia virginiana 2
Physostegia virginiana 3
Physostegia correllii 2
Physostegia leptophylla 1
Physostegia purpurea 2
Physostegia parviflora 1
Physostegia parviflora 2
Physostegia parviflora 3
Physostegia ledinghamii 1
Physostegia ledinghamii 2
Physostegia ledinghamii 3
Physostegia ledinghamii 4
Physostegia digitalis 2
Physostegia digitalis 4
Physostegia digitalis 1
Physostegia digitalis 3
Physostegia purpurea 5
Physostegia purpurea 3
Physostegia purpurea 2
Physostegia purpurea 4
0.2

Figure 2(on next page)

Phylogenetic tree of concatenated nuclear loci.

Bayesian 50% majority rule consensus tree obtained from analyses of the concatenated dataset. Nodes supported by Bayesian posterior probability values (PP) ≥0.9 and maximum likelihood bootstrap support (BS) ≥ 80 are labeled with pink dots. The green stars represent two nodes (clades A and B) discussed in the text, which have a PP ≥0.9 but a BS <80. Numbers following taxon names refer to different clones from PCR products.

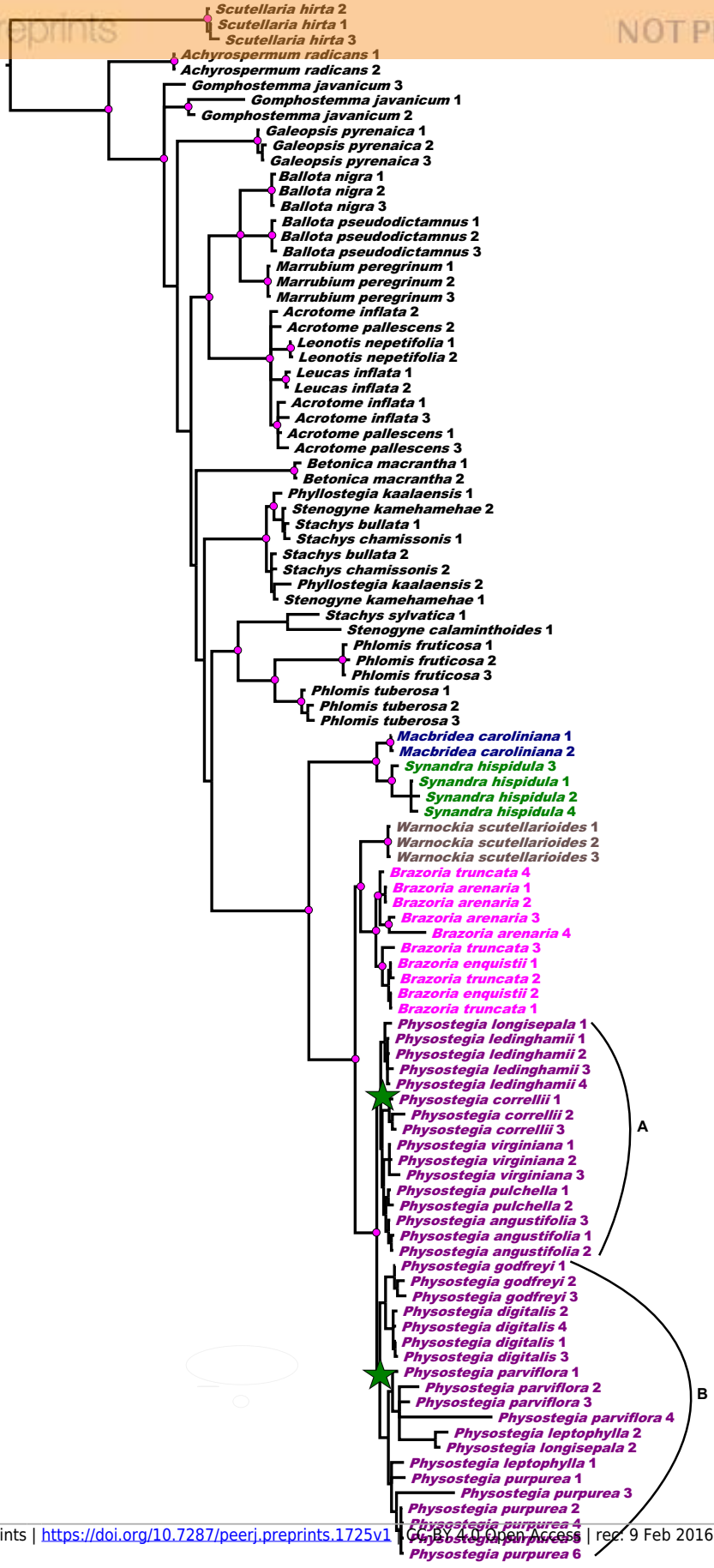


Figure 3(on next page)

Multi-locus coalescent tree.

The coalescence-based tree is inferred from a *BEAST analysis of the nuclear loci, as well as the concatenated chloroplast DNA data set. Only nodes with Bayesian posterior probability values ≥0.8 are labeled.

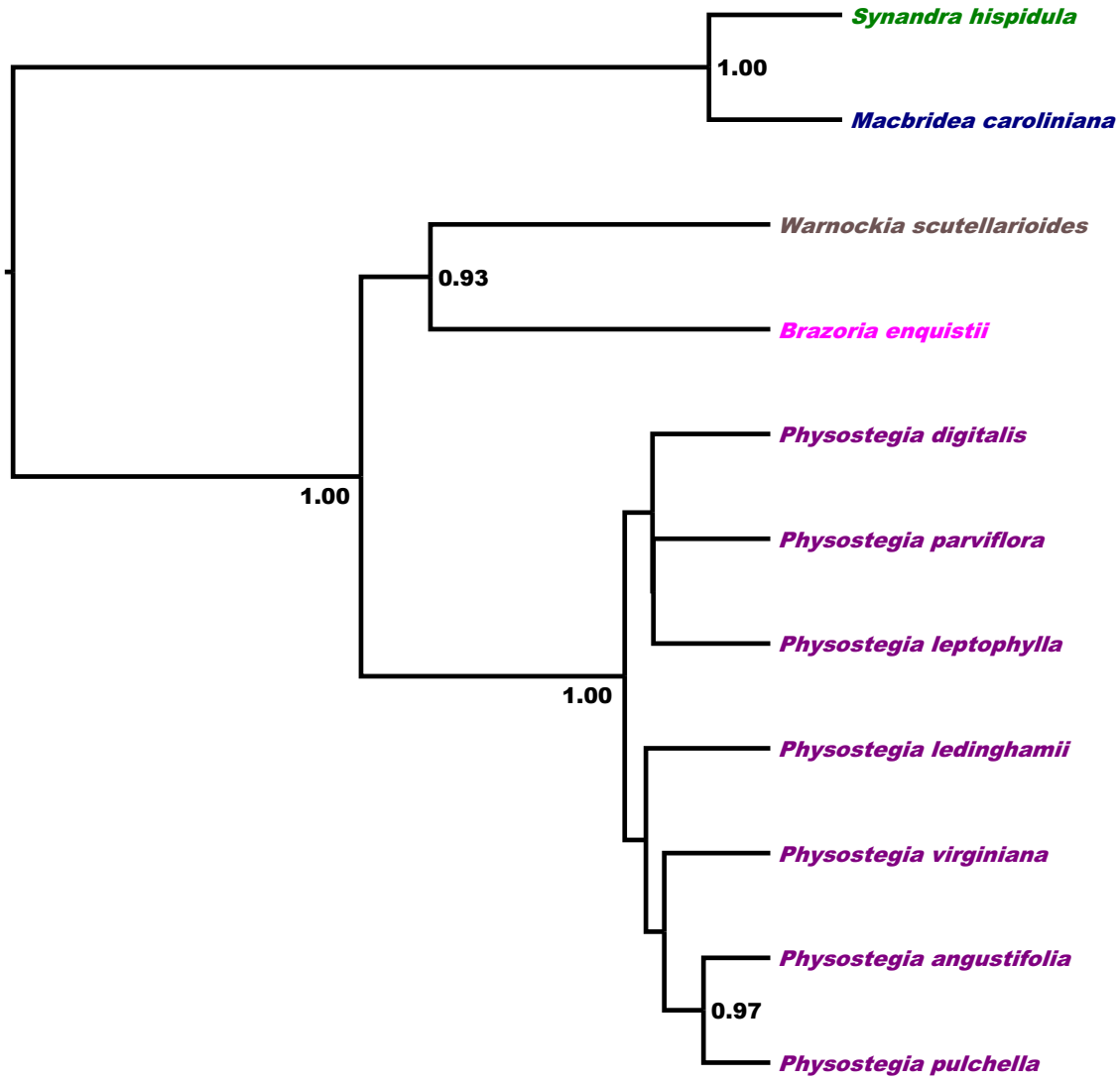
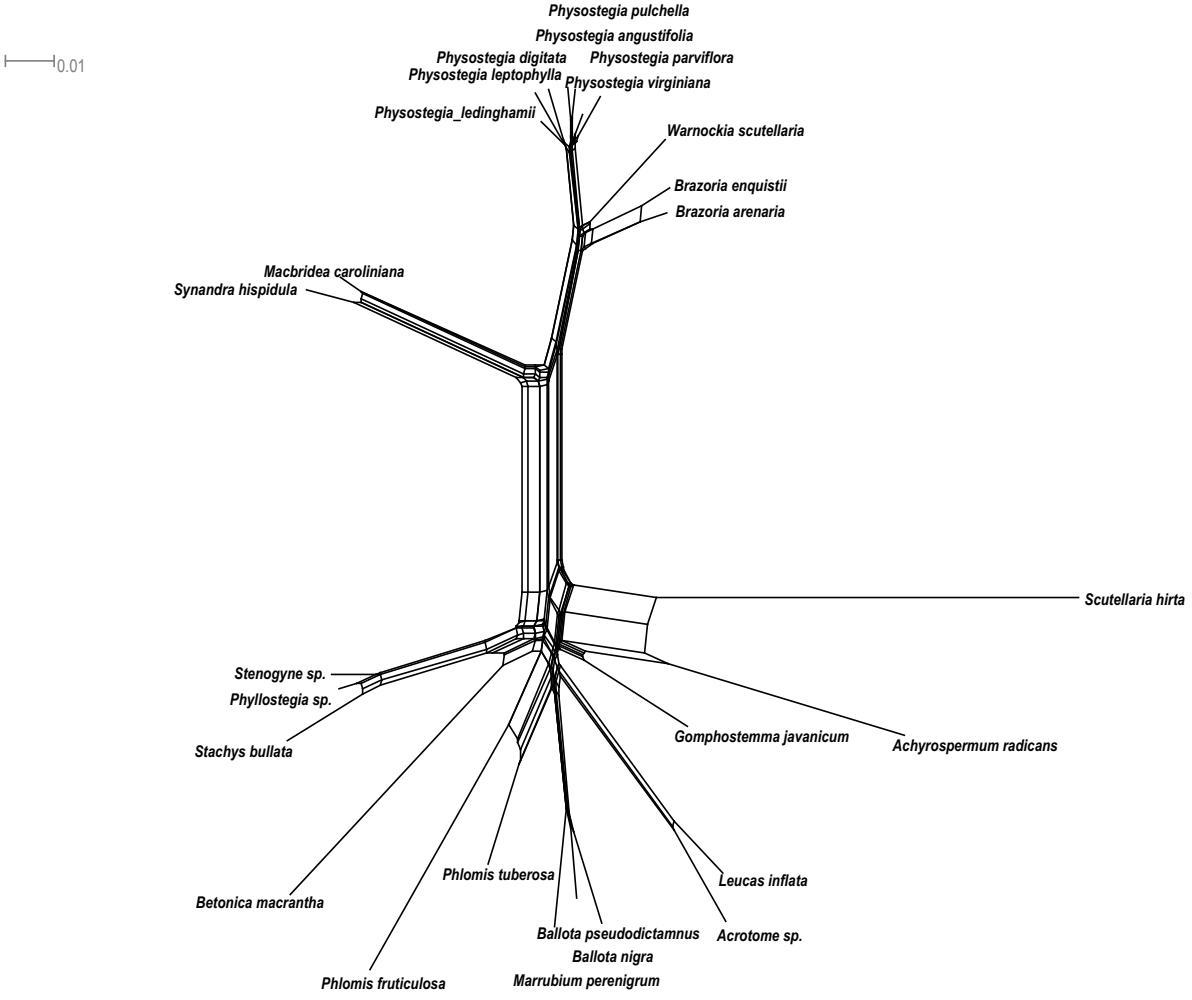


Figure 4(on next page)

Phylogenetic network.

NeighborNet analysis of the concatenated data set for *PHOT1* and *PPR* loci.


0.01
Physostegia pulchella
Physostegia angustifolia
Physostegia digitata
Physostegia parviflora
Physostegia leptophylla
Physostegia virginiana
Physostegia_ledinghamii
Warnockia scutellaria
Brazoria enquistii
Brazoria arenaria
Macbridea caroliniana
Synandra hispidula
Scutellaria hirta
Stenogyne sp.
Phyllostegia sp.
Stachys bullata
Gomphostemma javanicum
Achyrospermum radicans
Phlomis tuberosa
Leucas inflata
Betonica macrantha
Ballota pseudodictamnus
Acrotome sp.
Ballota nigra
Phlomis fruticulosa
Marrubium perenigrum