Phylogenetic Analyses of New World Mints Using Nuclear and Chloroplast DNA Datasets

Nathan Catlin, Tilottama Roy and Charlotte Lindqvist

Department of Biological Sciences, State University of New York at Buffalo

Introduction

Lamiaceae is a distinct and ecologically and economically important family of flowering plants. Its second-largest subfamily, Lamioideae, displays a diversity and worldwide distribution that could be considered a microcosm for the entire family. Of the ten recognized tribes in Lamioideae[1,2] only two tribes include members endemic to North America, Synandreae and Stachydeae. These two tribes have previously been found to be distantly related, suggesting that North America has been invaded by at least two separate lineages of Lamioideae at different times during Earth history. A better understanding of the evolutionary and migration history of these lineages will help elucidate timing and routes of intercontinental migrations to the New World.

The focus of this study is to investigate phylogenetic relationships in Synandreae, which comprises five genera, *Synandra*, *Macbridea*, *Brazoria*, *Warnockia*, and *Physostegia*. *Synandra* and *Macbridea* are found in the South Eastern US, *Warnockia* and *Brazoria* are found in the South Central US, and *Physostegia* is distributed across North America. Phylogenetic relationships among these genera have previously been studied based on chloroplast markers[3]. In this study, we used nuclear low-copy markers. Chloroplast DNA is often complemented with nuclear DNA because chloroplast genomes are only maternally inherited, whereas nuclear DNA is biparentally inherited and may undergo genetic recombination. The use of low-copy nuclear loci may aid in the reconstruction of better resolved and more robust phylogenies. This will allow us to correctly place species within the tribe Synandreae. Also, we are investigating the phylogenetic position of Synandreae compared to other tribes within Lamioideae to elucidate its close relatives and biogeographical origins. Here, we reconstruct maternal lamioid relationships based on a large dataset of published chloroplast DNA sequences.

Materials and Methods

Nuclear Datasets

Lamioid mint DNA was previously extracted from silica dried leaves and herbarium samples. Polymerase chain reaction (PCR) was used to amplify four nuclear markers of interest: three regions of the phototropin gene, 1a (PHOT1A), 2a (PHOT2A), and 2b (PHOT2B), and the pentatricopeptide repeat region (PPR). Purified DNA from each successfully amplified region were cloned and sequenced. Sequences were aligned using the MAFFT multiple sequence alignment software[4] and further manually aligned in BioEdit[5]. Alignments were subjected to Bayesian inference in MrBayes v. 3.1.2[6] to reconstruct gene trees. A species tree was generated from the Species Tree Analysis Web Server (STRAW)[7] using the STAR method for species tree reconstruction, which uses average ranks of gene coalescence times to build species trees from a set of rooted gene trees. For this purpose we used rooted maximum likelihood trees from our individual datasets for 2 of our loci (PPR and Phot1A), built through the RAxML Blackbox web server[8].

Chloroplast Datasets

Previously published sequences of the *trnL* intron, *trnL-trnF* intergenic spacer region, *rps16* intron, and *matK* were selected from GenBank. Datasets for each gene region were aligned using the program MUSCLE[9] in the CIPRES[10] web portal. Sequences were then concatenated and subjected to maximum likelihood analysis built through the RAxML Blackbox web server[8].

Figure 1
Synandreae Genera
(A) - *Synandra*
(B) - *Macbridea*
(C) - *Physostegia*
(D) - *Warnockia*
(E) - *Brazoria*

- *Macbridea*
- *Synandra*
- *Warnockia*
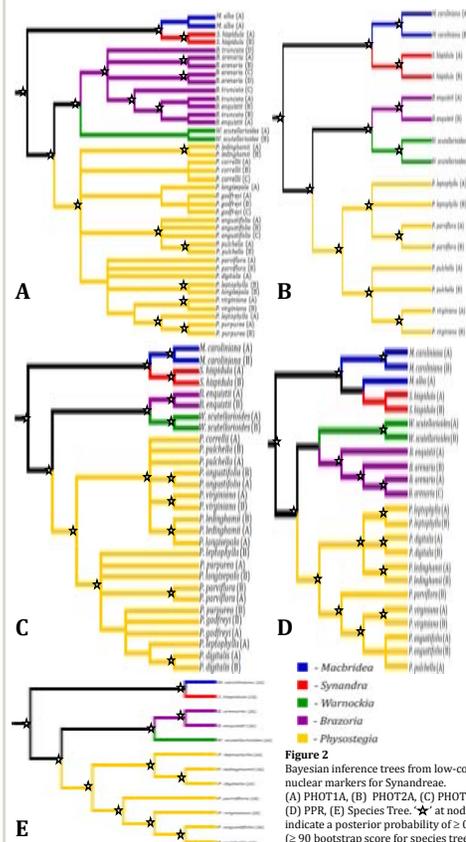- *Brazoria*
- *Physostegia*

Figure 2
Bayesian inference trees from low-copy nuclear markers for Synandreae.
(A) PHOT1A, (B) PHOT2A, (C) PHOT2B, (D) PPR, (E) Species Tree. '★' at nodes indicate a posterior probability of ≥ 0.9 (≥ 90 bootstrap score for species tree (E)).

- Pogostemoneae
- Gomphostemmateae
- Synandreae
- Stachydeae
- Paraphlomideae
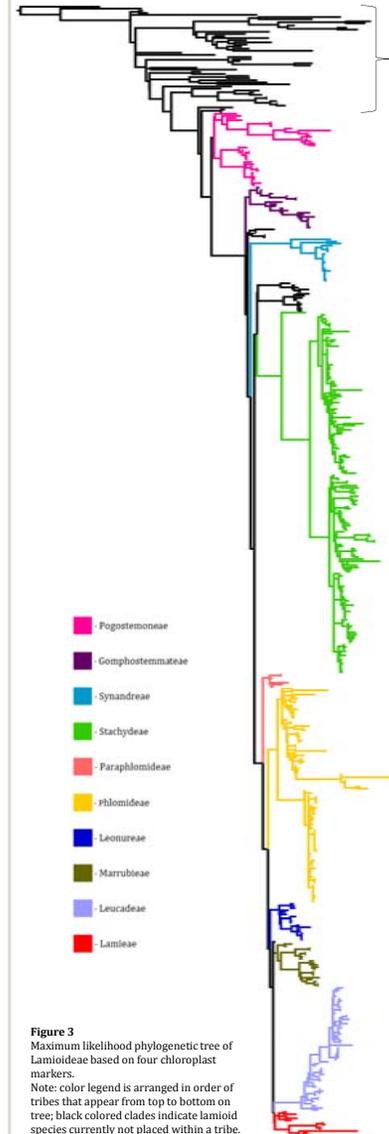- Phlomideae
- Leonureae
- Marrubieae
- Leucadeae
- Lamieae

Figure 3
Maximum likelihood phylogenetic tree of Lamioideae based on four chloroplast markers.
Note: color legend is arranged in order of tribes that appear from top to bottom on tree; black colored clades indicate lamioid species currently not placed within a tribe.

Results

From Fig. 2 (A, B, C, D) we see strong support values at each basal node for the four gene trees, supporting monophyly of the tribe Synandreae. We also observe robustly supported sister relationships between *Macbridea* and *Synandra*, *Warnockia* and *Brazoria*, respectively. The clade consisting of *Physostegia* species also resolves with strong support, and appears sister to the *Warnockia-Brazoria* clade. These results are also corroborated by our species tree (Fig. 2E). The chloroplast phylogeny of Lamioideae (Fig. 3) does not resolve any single close relative of the tribe Synandreae, but shows Synandreae as sister group to the rest of Lamioideae, excluding the tribes Pogostemoneae and Gomphostemmateae and the unplaced genus *Colquhounia*. Synandreae is still resolved as monophyletic, which is similar to what we observe from our individual nuclear gene trees and the species tree.

Discussion and Future Directions

Based on the four low-copy nuclear markers, gene phylogenies have been reconstructed for the North American tribe Synandreae (Fig. 2A-D). Each of the gene trees showed congruent relationships among the four genera, which is also demonstrated by the estimated species tree (Fig. 2E). These trees also show Synandreae as a monophyletic group, which is consistent with recent molecular phylogenetic studies for Synandreae based on chloroplast and nuclear ribosomal DNA sequences[3]. However, in contrast to the previous study where *Synandra* formed the sister group to all the other North American species, the low-copy nuclear data support a sister relationship between *Synandra* and *Macbridea*. A close relationship between these two genera is supported by their geographic distribution (endemic to SE US) and similar chromosome numbers (2n=18). The other genera in Synandreae have increasing number of chromosomes (2n=20, 28, 38, or 76). Such incongruence in phylogenetic placements between maternally inherited chloroplast and biparentally inherited nuclear sequence data may indicate hybrid speciation or incomplete lineage sorting.

Our analysis of a concatenated chloroplast data set for 542 species within Lamiaceae corroborated the monophyly and placement of each tribe within Lamioideae[1,2] (Fig. 3). The Synandreae holds a basal and distinct position in the subfamily as sister group to the majority of Lamioideae. We are currently augmenting this analysis with a phylogeny of the subfamily based on nuclear DNA sequences (in progress). In future studies, a species tree of low-copy nuclear markers for all sampled species may further elucidate the placement of tribes within Lamioideae. Also, a comparative analysis of migration and species divergence times for tribes Synandreae and Stachydeae would further unravel the relationships of their introduction to the New World and their subsequent diversifications. In addition, since there are about 1260 species and 63 genera currently recognized in the subfamily Lamioideae[1,2], a phylogeny that contains all species within the subfamily would give a complete representation of tribe and species placement within Lamioideae.

References

1. Scheen, A.C., Bendiksby, M., Ryding, O., Mathiesen, C., Albert, V.A., Lindqvist, C. 2010. Molecular phylogenetics, character evolution, and suprageneric classification of Lamioideae (Lamiaceae). Annals Missouri Bot. Garden 97: 191–217.
2. Bendiksby, M., Thorbek, L.B., Scheen, A.C., Lindqvist, C., Ryding, O. 2011. An updated phylogeny and classification of Lamiaceae subfamily Lamioideae. Taxon 60: 471–484.
3. Scheen, A.C., Lindqvist, C., Fossdal, C.G., Albert, V.A. 2008. Molecular phylogenetics of tribe Synandreae, a North American lineage of lamioid mints (Lamiaceae). Cladistics 24: 299–314.
4. Katoh, S. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30: 772-780.
5. Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Series 41: 95–98.
6. Huelsenbeck,Ronquist, F., Ronquist, J.P. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754-755.
7. Shaw, T.I., Ruan,Z., Glenn, T.C, Liu, L. 2013. STRAW: Species TRee Analysis Web server. Nucleic Acids Res. 41: W238–W241.
8. Stamatakis, A., Hoover, O., Rougemont, J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. Sys. Biol. 57:58-771.
9. Edgar, R. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res. 32: 1792-97.
10. Miller, M.A., Pfeiffer, W., and 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, LA, pp 1-8.

Acknowledgements