Systematic Botany (2022), 47(2): pp. 457–466
ISSN 0363-6445 (print); ISSN 1548-2324 (online)

DOI 10.1600/036364422X16512564801605
Date of publication June 14, 2022

Molecular Systematics, Species Concepts, and Myrmecophytism in *Cecropia* (Cecropieae: Urticaceae): Insights from Restriction-Site Associated DNA

Erin L. Treiber,[1,7] Paul-Camilo Zalamea,[2,3] María Fernanda Torres,[4,5] Santiago Madriñán,[4] and George D. Weiblen[6]

[1]Department of Horticultural Sciences, University of Minnesota, 1970 Folwell Avenue, Saint Paul, Minnesota 55108, USA
[2]Department of Integrative Biology, University of South Florida, 4202 East Fowler Avenue, Tampa, Florida 33620, USA
[3]Smithsonian Tropical Research Institute, Apartado 0843-03092, Panama City, Republic of Panama
[4]Department of Biological Sciences, Universidad de los Andes, Apartado Aéreo 4976, Bogotá, Colombia
[5]Institute of Evolutionary Biology, University of Edinburgh, Ashworth Laboratories, Kings Buildings, West Mains Road, Edinburgh EH9 3JT, UK
[6]Department of Plant Biology and Bell Museum, University of Minnesota, 140 Gortner Laboratory, 1479 Gortner Avenue, Saint Paul, Minnesota 55108, USA
[7]Author for correspondence (treib020@umn.edu)

Communicating Editor: Jacob Landis

Abstract—*Cecropia* is a group of fast-growing pioneer trees that are important in forest regeneration and a common ant-plant mutualism in the Neotropics. To investigate the evolution of mutualism between *Cecropia* and associated ants, a phylogenetic framework is necessary. *Cecropia* species are difficult to distinguish morphologically and conventional genetic markers are insufficiently variable to resolve the phylogenetic relationships among species. Our study aimed to compare the phylogenetic utility of restriction site associated DNA (RAD) sequencing to prior work based on commonly sequenced gene regions. RAD sequence data resolved and supported species-level relationships better than previous studies. We identified a deeply divergent non-myrmecophytic clade including *C. sciadophylla* and African *Musanga*. Results from geographically widespread and morphologically heterogenous *C. obtusifolia* and *C. angustifolia* suggest that current synonymy has lumped phylogenetically divergent lineages. Reconstruction of ant associations on the highly supported *Cecropia* phylogeny inferred equal probability of the ancestor of *Cecropia* being myrmecophytic or not. More intensive genetic study is needed to refine species concepts in *Cecropia*.

Keywords—Ancestral state reconstruction, ant-plant associations, mutualism, pioneer trees, RADseq.

For decades, plant systematists have relied on direct sequencing of individual loci to resolve phylogenetic relationships (Soltis et al. 1992; Soltis 1998). However, it is both difficult to identify loci with sufficient molecular variation to resolve phylogenetic relationships and to find loci that may be consistently amplified and sequenced across the taxa of interest (Rubin et al. 2012). Limited variation in commonly used markers such as chloroplast spacers and nuclear ribosomal DNA for plant phylogenetics has been a particular problem in studies of large, woody, tropical genera. Lack of molecular variation is often blamed for poor phylogenetic resolution and clade support. Published phylogenies for many large, woody species such as *Ficus* L. (Rønsted et al. 2007), *Inga* Mill. (Richardson et al. 2001), *Macaranga* Thouars (Blattner 2001; Davies et al. 2001), and *Pouteria* Aubl. (Triono et al. 2007) demonstrate this lack of resolution and support. Smith and Donoghue (2008) showed slower rates of molecular evolution in woody plants than in herbs but the lack of resolution in species-level phylogenies in trees and shrubs may not only be due to slower rates of molecular evolution, but also due to the insufficient molecular variation in the few gene regions often sequenced in phylogenetic studies. Next generation sequencing provides megabases of data in a single sequencing run (Straub et al. 2012). Restriction-site-associated DNA sequencing (RADseq), which simultaneously samples multiple regions throughout the genome, has been proposed as an effective alternative to whole-genome sequencing for molecular phylogenetic studies (Davey et al. 2011; Rubin et al. 2012; Parchman et al. 2018). Many recent studies have used RADseq to address phylogenetic questions (Cavender-Bares et al. 2015; Eaton et al. 2015; Deng et al. 2018; Paetzold et al. 2019; Appelhans et al. 2020; Du et al. 2020; Penagos Zuluaga et al. 2021).The application of RADseq could provide the amount of sequence data and variation required for phylogenetic resolution that has been lacking in systematic studies of species-rich tropical genera. Here we examine the performance of RADseq applied to the woody tropical genus, *Cecropia* Loefl., which has been difficult to resolve with conventional loci (Treiber 2017) and lacks a well-supported molecular phylogeny (Gutiérrez-Valencia et al. 2017).

Cecropia (Urticaceae) consists of fast-growing and relatively short-lived dioecious trees with small, animal-dispersed seeds that germinate in canopy gaps or early in ecological succession after forest disturbance. There are 61 recognized species occupying a wide range of environments from southern Mexico to northern Argentina (Berg and Franco-Rosselli 2005). *Cecropia* species are commonly described as myrmecophytic, having a symbiotic relationship with aggressive, predatory ants including the genus *Azteca* Forel. *Cecropia* produces oil-rich Müllerian bodies at the base of the petioles in a cluster of dense hairs (trichilia) and smaller pearl bodies on abaxial leaf surfaces. These structures provide nutrition to ants and also to herbivores (Berg and Franco-Rosselli 2005; Dejean et al. 2012). Ecological studies have examined aspects of the mutualism including the value of ant defense to *Cecropia* (Vasconcelos and Casimiro 1997; Agrawal and Dubin-Thaler 1999; Fáveri and Vasconcelos 2004), the value of food rewards to ants (Sagers et al. 2000), and geographic variation in the host preferences of *Azteca* (Vieira et al. 2010). Approximately two thirds of *Cecropia* species are regularly ant-associated but the evolutionary history of the mutualism is not well understood (Janzen 1969, 1973a; Berg and Franco-Rosselli 2005). A recent study attempted to infer the evolution of ant-associated traits using ancestral state reconstruction but lacked a well-supported phylogeny (Gutiérrez-Valencia et al. 2017). Additional information is needed to gain insight

on how the mutualism originated and how dynamic it has been over time. Here we apply RADseq to *Cecropia* to re-evaluate the origin of myrmecophytism in light of other recent work.

Relationships among genera in the tribe Cecropieae were unknown until a phylogenetic study including all five genera of Cecropieae showed the African genus *Musanga* C.Sm. ex R.Br. to be embedded in *Cecropia* (Treiber et al. 2016). Alternatively, Gutiérrez-Valencia et al. (2017) placed *Musanga* sister to *Cecropia* but with low bootstrap support. Both studies found sequence variability at conventional loci (*rbcL*, *matK*, *trnL-trnF*, *trnH-psbA*, *psbK-psbI*, *G3PDH*, and *ndhF*) insufficient to resolve or support *Cecropia* phylogeny at the species level (Treiber et al. 2016; Gutiérrez-Valencia et al. 2017). Another problem with prior work was the untested assumption of species monophyly when combining sequences drawn from different populations in the same analysis.

The taxonomy of the genus and concepts of species are complicated by morphological variation in geographically widespread species sometimes exceeding variability among different species. Identification is extremely difficult in light of geographic and ecotypic variation, and Berg and Franco-Rosselli (2005) reduced at least 165 names to 61 recognized species in the only taxonomic revision of the genus. For example, *C. angustifolia,* a widely distributed species, varies greatly in indumentum, leaf venation, and inflorescence construction, which apparently correlate with annual precipitation and elevation. Berg and Franco-Rosselli (2005) placed 13 names under *C. angustifolia* representing different geographic regions and altitudinal ranges. Similarly, 14 taxa were synonymized under *C. obtusifolia*, which also exhibits substantial morphological variation and geographic range. In contrast, species such as *C. sciadophylla* and *C. membranacea* span the entire Amazon Basin but exhibit much less variability (Berg and Franco-Rosselli 2005; Zalamea et al. 2008, 2012).

It has also been suggested that hybridization among species complicates the taxonomy of *Cecropia* (Berg and Franco-Rosselli 2005). The large number of SNPs generated by RADseq has been used to test hypotheses of introgression (Green et al. 2010; Eaton and Ree 2013). Introgression resulting from a history of hybridization is common in plants and can also complicate the interpretation of phylogenetic analyses. A statistical parameter, the D-statistic, has been applied to detect evidence for introgression in Neanderthals and humans (Green et al. 2010), herbaceous plants (Eaton and Ree 2013), and trees (Eaton et al. 2015). This statistic estimates the frequency of incongruent SNP patterns, which is expected to be equal if they arise from stochastic processes (Eaton and Ree 2013). If a pattern of incongruence is more frequent than expected by chance, hybridization or introgression between species is inferred (Eaton and Ree 2013).

Our study aimed to compare RADseq data to conventional gene sequences in estimating *Cecropia* phylogeny and to evaluate the current species concept. If morphological differences are predictive of genetic isolation, we expect morphologically homogeneous species to form monophyletic groups and morphologically heterogeneous to be polyphyletic. We also used the D-statistic to detect patterns of introgression among species in cases of polyphyletic taxa and conflicting estimates of phylogeny. Finally, we estimated probabilities for the ancestral state of myrmecophytism. We tested the hypothesis of a single origin in *Cecropia* followed by multiple losses or alternatively for multiple independent origins.

Materials and Methods

Sampling—We examined 47 collections representing 31 *Cecropia* species and four other members of the Cecropieae tribe (*Coussapoa* Aubl., *Musanga*, *Myrianthus* P.Beauv., and *Pourouma* Aubl.). Other members of Cecropieae were included to root the tree and *Musanga* was included to confirm its ingroup position, embedded in *Cecropia* as suggested previously (Treiber et al. 2016) versus its position sister to *Cecropia* (Gutiérrez-Valencia et al. 2017). Our sample included only half of the 61 recognized *Cecropia* species because DNA isolation from herbarium specimens was problematic, perhaps due to specialized metabolites and/or the age of the material. At least we covered the entire geographical range of the genus across six countries and the main centers of diversity in Amazon basin, Magdalena Valley, Pacific coastal lowlands, Andes mountains, and Central America (Appendix 1) (Zalamea et al. 2011, 2012). We also sampled locally endemic as well as widely distributed species (Berg and Franco-Rosselli 2005). To test hypotheses concerning morphological and genetic variation within species, we included multiple samples of four *Cecropia* species across their range. All four species are widespread with two being morphologically homogeneous throughout their range (*C. sciadophylla* and *C. membranacea*) and two being morphologically heterogeneous (*C. obtusifolia* and *C. angustifolia*) (Appendix 1).

DNA Preparation and Sequencing—Silica dried material collected in the field was used for DNA extractions, except for one sample for which we only had herbarium material. DNA was extracted using a modified CTAB method (Doyle and Doyle 1987) with a 2% CTAB buffer and extended incubations of 12 to 24 hrs. Each sample was extracted in duplicate to provide more material per sample for sequencing. Each sample for RADseq was required to contain 50 µL of high molecular weight DNA with no degradation or contaminating material at a concentration of 20 ng/µL. Each set of extractions was run with a negative control and was run on an agarose gel to look for degradation or contaminating material. Samples were sent to Floragenex Inc. (Eugene, OR) for RAD library preparation and sequencing. Libraries were prepared using the *PstI* restriction enzyme following the methods of Baird et al. (2008). The library was created from 95 pooled and barcoded samples sequenced on an Illumina Hi Seq 2000 to generate 100 bp single end reads. Samples were combined for each collection when demultiplexing the library.

Sequence Assembly—Sequences were demultiplexed using ea-utils (Aronesty 2011) with default settings, which allowed for one mismatch in the barcode sequence. The remaining steps of quality filtering and assembly of sequences into de novo loci were done using pyRAD v. 3.0.63 (Eaton 2014). Bases with a Phred score < 20 were converted into unknown base pairs (Ns) and reads with > 5 Ns were discarded. After filtering, reads were clustered within samples at thresholds between 82% and 98%. The default settings in pyRAD for max SNPs in a final locus were used (3, 6, 99, 99). Average parameter values estimated in pyRAD were used when making consensus base calls and clusters with minimum depth less than 5 were excluded. Additional analyses were also done with higher minimum depth and taxon coverage (15) to examine the effects of missing data. After removing loci containing more than two alleles in a sample as potential paralogs, consensus loci were clustered across samples using the same threshold used in the previous within sample clustering. Assembled loci were exported as a supermatrix with missing data converted to Ns for phylogenetic analysis.

Phylogenetic Analysis—Maximum likelihood analyses were performed on each assembled data set using RAxML version 8.2.4 (Stamatakis 2014) on the CIPRES Science Gateway (Miller et al. 2010). Bootstrap support was estimated from 300 replicate searches from random starting trees run using the GTR+Γ model of nucleotide substitution model.

Tests for Introgression—The D-statistic was used to test for evidence of introgression in *Cecropia* (Green et al. 2010; Eaton and Ree 2013). The D-statistic detects introgression between lineages based on the frequency of alleles that are discordant with a phylogenetic hypothesis. Although discordant site patterns often occur due to lineage sorting of ancestral polymorphisms, the different patterns occur at a mostly equal frequency due to the stochasticity of the process. The D-statistic calculates the asymmetry in the relative occurrence of the two discordant site patterns to test for introgression (Eaton et al. 2015).

We first tested the hypothesis that introgression is more common in species that are morphologically heterogeneous than in morphologically homogeneous species. We used pyRAD v. 3.0.63 to calculate the D-statistic using 1000 bootstrap replicate and significance was assessed with a p value less than 0.01 after Bonferroni correction for multiple testing. Both morphologically homogeneous species were monophyletic and the D-statistic was used to test for hybridizations with close relatives in the phylogeny. The location of *C. sciadophylla* sister to the rest of the species in

TABLE 1. Species in the groups P1, P2, and P3 for the D-statistic analyses ((P1,P2),P3)) testing for potential introgression in morphologically homogeneous species, morphologically heterogenous species, and the position of *C. herthae* in the phylogeny.

Test	P1	P2	P3	Outgroup
Morphologically homogenous species	*Musanga cecropioides*	*C. sciadophylla* Ecuador *C. sciadophylla* Brazil *C. sciadophylla* Amazonas *C. sciadophylla* Meta, CO	All other *Cecropia* species in phylogeny	*Coussapoa floccose* *Myrianthus arboreus* *Pourouma tomentosa*
Morphologically homogenous species	*C. latiloba*	*C. putumayonis*	*C. membranacea* Amazonas, CO *C. membranacea* Casanare, CO *C. membranacea* Ecuador	*C. sciadophylla* Ecuador *C. sciadophylla* Brazil *C. sciadophylla* Amazonas *C. sciadophylla* Meta, CO *Musanga cecropioides*
Morphologically heterogeneous species	*C. litoralis* *C. sararensis* *C. angustifolia* Boyaca, CO *C. reticulata* *C. engleriana* *C. metensis*	*C. herthae*	*C. obtusifolia* Costa Rica *C. obtusifolia* Panama *C. longipes* *C. mutisiana* *C. obtusifolia* Columbia	*C. sciadophylla* Ecuador *C. sciadophylla* Brazil *C. sciadophylla* Amazonas *C. sciadophylla* Meta, CO *Musanga cecropioides*
Position of *C. herthae*	*C. herthae*	*C. obtusifolia* Costa Rica *C. obtusifolia* Panama	*C. longipes* *C. mutisiana* *C. obtusifolia* Columbia	*C. sciadophylla* Ecuador *C. sciadophylla* Brazil *C. sciadophylla* Amazonas *C. sciadophylla* Meta, CO *Musanga cecropioides*

the phylogeny required the test to be performed against all other *Cecropia* samples (Table 1). Because morphologically heterogeneous samples were found in different parts of the phylogeny, we used the D-statistic to test for hybridization with species that were sister (or close relatives) where they were located in the phylogeny (Table 1).

Finally, we used the D-statistic to test for introgression for *C. herthae,* which changed position in different analyses. Although *C. herthae* had high bootstrap support in the reported phylogeny, the position changed in different analyses using different parameters in data processing (Table 2) so we used the D-statistic to test if it was potentially due to hybridization (Table 1).

Reconstruction of Ant Associations—CHARACTER MATRIX—We examined the literature to create a matrix of ant association as a discrete character for included species. Ant associations were scored as absent (0) if a species generally did not associate with ants or present (1) if a species has been recorded to have ant associations. Character-state designations were mainly based on Berg and Franco-Rosselli (2005) and data from specimen collections with additional sources used for confirmation (Longino 1989; Berg et al. 1990; Agrawal 1998; Dejean et al. 2012; Latteman et al. 2014).

ULTRAMETRIC TREE—To convert to an ultrametric tree for the ancestral reconstruction, we used three different smoothing algorithms from the ape v. 5 package in R Studio v. 1.0.136 (RStudio Team 2016). Because there were multiple samples of some species in the phylogeny, we removed duplicate samples before analyses. All samples of *C. sciadophylla* and *C. membranacea* were in the same clade, so we removed all but one sample to mimic the majority of species that only had one sample. However, *C. obtusifolia* and *C. angustifolia* samples were in different parts of the tree, so only samples that clustered together were removed. Also, two of the outgroup samples were randomly removed. We used time calibration in the model when determining branch lengths because pollen records were available for the region, but did not do a fully time calibrated tree due to the limited pollen data available. Using the chronos function, all analyses were calibrated using pollen data with a minimum root age of 65 MYA (Burnham and Graham 1999) based on the oldest angiosperm pollen found in the region and the node for *Coussapoa* with a minimum age of 8 MYA and a maximum age of 65 MYA between the oldest angiosperm fossil and *Coussapoa* pollen in the record (Burnham and Graham 1999). The three smoothing algorithms used were a strict clock, a relaxed clock, and penalized likelihood. The lambda value was also varied from each model from 0–1.5 to determine the best fit. The ultrametric tree with the highest likelihood was used for subsequent analyses.

ANCESTRAL RECONSTRUCTION—Maximum likelihood reconstructions of ant associations were done using the geiger v. 3.0–6 (Harmon et al. 2008) and phytools v. 0.4–56 (Revell 2012) packages in RStudio. Reconstructions were performed using the equal rates model (ER) and all rates different (ARD). Likelihood ratio tests were used to compare models.

RESULTS

Restriction Site Associated (RAD) Sequencing—Seven samples were excluded from the data analysis due to low reads and one due to apparent contamination by *Pourouma* DNA, resulting in a total of 39 samples (Appendix 1). Six matrices were assembled using different minimum depth, minimum coverage, and clustering thresholds that varied numbers of loci and phylogenetically informative sites (Table 2). The matrices ranged from 71,410 loci with a minimum coverage depth of five and clustering threshold of 0.98 to 24,371 loci with respective values of 15 and 0.82 (Table 2). Phylogenetically informative characters decreased from 299,910 to 26,561 with increasing minimum depth of coverage from five to 15 and clustering threshold from 0.82 to 0.98 (Table 2).

Matrices varying in stringency of clustering and minimum number of samples sharing a locus did not have much of an

TABLE 2. Numbers of loci in matrices, variable sites, and phylogenetically informative sites for runs with differing minimum depth for base calling (min. depth), the number of samples that must share a locus to be included (min. coverage), and the clustering threshold for matrix assembly for RADseq data.

Matrix number	Min. depth	Min. coverage	Clustering threshold	Number of final loci	Variable sites	Phylogentically informative sites
1	5	5	0.82	61,022	817,026	299,910
2	5	5	0.90	62,876	705,144	254,199
3	5	5	0.98	71,410	185,696	53,117
4	15	15	0.82	36,789	469,717	161,718
5	15	15	0.90	36,863	420,430	146,160
6	15	15	0.98	24,371	95,409	26,561

effect on inferred relationships. The data sets with a higher minimum depth for base calling prior to clustering (15) and a higher number of required samples per locus (15) had fewer informative sites, but yielded higher support values. As the clustering algorithms became more stringent, from 82–98%, the number of informative characters decreased as did support. Lower support values corresponded to reduced numbers of informative sites.

The resulting phylogenies had a consistent topology for the majority of the tree with only three areas having differences in topology: 1) *C. marginalis* and *C. goudotiana*, 2) *C. herthae*, and 3) the *C. gabrielis*/*C. telenitida*/*C. plicata* clade. The most highly supported phylogeny was from the matrix using a minimum depth of 15 and a clustering threshold of 0.82 (Fig. 1). The main difference in topology was the branching order of *C. marginalis* and *C. goudotiana* in how they split from the majority of other *Cecropia* species. The phylogeny was highly resolved and highly supported for the majority of clades. Nodes that had lower support values corresponded to the differences in topology between different matrices (see Treiber et al. 2022). *Cecropia herthae* was highly supported, but was not consistent in all topologies. As expected, *Musanga* was embedded within the *Cecropia* clade and was highly supported as sister to *C. sciadophylla*. This clade was also highly supported as sister to and deeply diverged from the remaining *Cecropia* species (Fig. 1). Morphologically homogenous

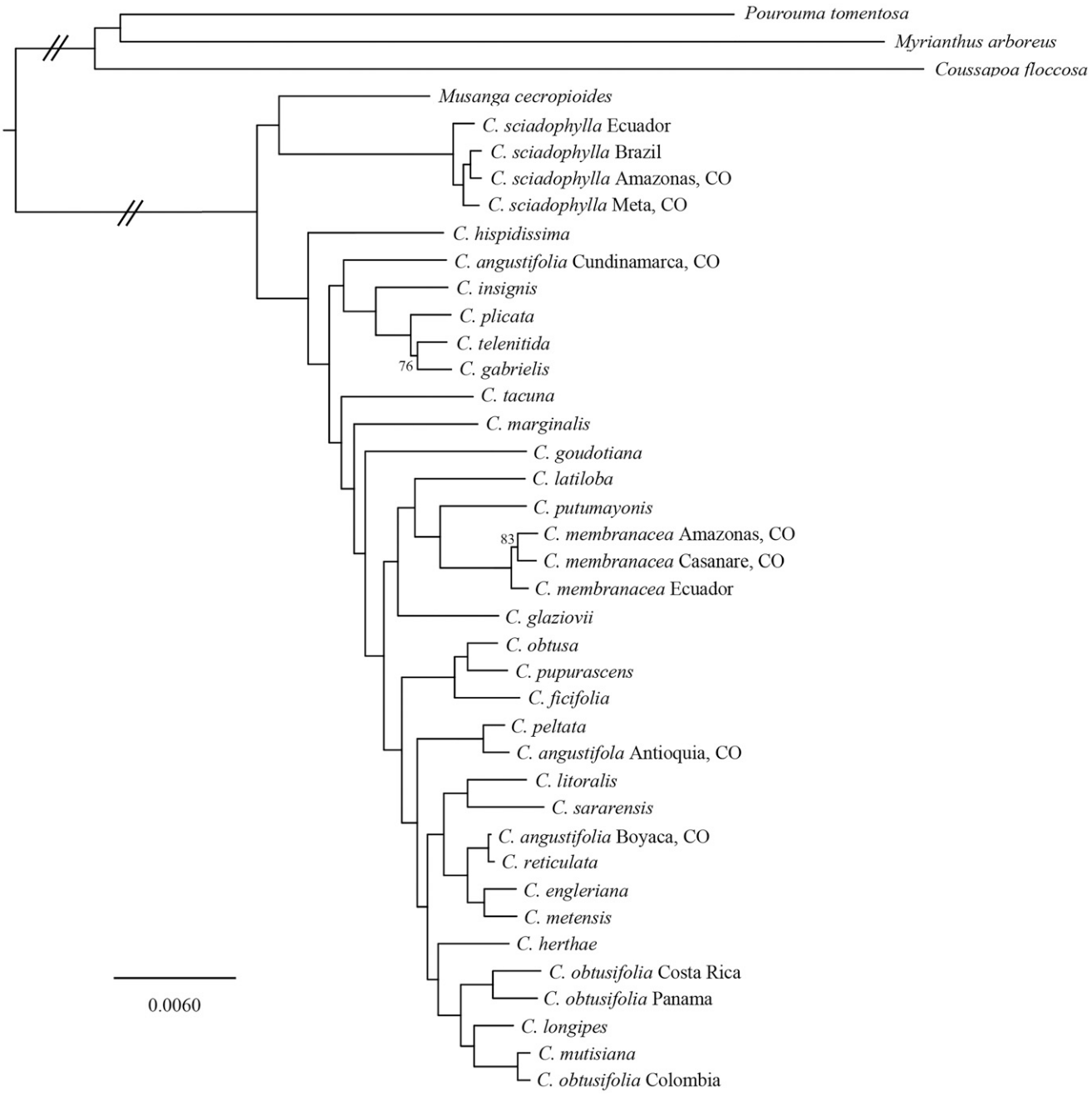

FIG. 1. Maximum likelihood phylogeny inferred from the concatenated RADseq data set with the highest minimum depth (15) and lowest clustering threshold (0.82). The phylogeny was rooted using other members of the Cecropieae tribe (*Coussapoa*, *Myrianthus*, and *Pourouma*). Bootstrap support was 100 except where indicated. The branches bearing double hatch marks have been truncated and are not proportional to the rest.

C. sciadophylla and *C. membranacea* were monophyletic whereas morphologically heterogenous *C. angustifolia* and *C. obtusifolia* were not. The three samples of *C. obtusifolia* were in a clade with *C. longipes* and *C. mutisiana* whereas samples of *C. angustifolia* samples were scattered across the phylogeny. One sample of *C. angustifolia* was in an earlier diverging clade including *C. insignis*/*C. plicata*/*C. telenitida*/*C. gabrielis*. The other two samples were in a later diverging clade including *C. peltata*/*C. litoralis*/*C. sararenis*/*C. reticulata*/*C. engleriana*/*C. metensis*.

Tests for Introgression—When *Musanga* and morphologically homogeneous *C. sciadophylla* were tested for evidence of introgression with the core *Cecropia* clade, a fifth of the comparisons involving African *Musanga* and core *Cecropia* were statistically significant but none of the tests involving *C. sciadophylla* and core *Cecropia* were significant (Fig. 2A). The remaining tests for introgression in morphologically heterogeneous species (*C. angustifolia* and *C. obtusifolia*) found little support for introgression, while in the homeogeneous species (*C. membranaceae*) the test detected introgression with *C. latiloba* (see Treiber et al. 2022).

With *C. herthae*, the test detected significant patterns of introgression with other species. For example, all tests were significant for introgression between *C. herthae* and the *C. longipes*/*C. mutisiana*/*C. obtusifolia* clade (Fig. 2C). When *C. herthae* was tested with the remaining large clade (Fig. 2B), approximately one third of the tests were significant and were equally distributed between individual members of the clade except for *C. sararensis*.

Ancestral Reconstruction of Myrmecophytism—Ancestral state reconstruction is reported for models assuming equal rates and all rates different (see Treiber et al. 2022). We were unable to reject the simplest model assuming equal rates of change for the character. The reconstruction of ant association (Fig. 3) had an equal probability of the ancestor being either myrmecophytic or non-myrmecophytic but the ancestor of the *C. sciadophylla*/*Musanga* clade had higher probability of being non-myrmecophytic. In the sister group to *C. sciadophylla*/*Musanga*, hereafter "core *Cecropia*", the probability of a myrmecophytic ancestor was 0.72 (Fig. 3). The probability of ancestral myrmecophytism was generally high in core *Cecropia* ancestor except near where non-myrmecophytic species were located in the phylogeny. The sister group to *C. tacuna*, including many recently diverged species, most probably had a myrmecophytic ancestor. An alternative coding of *C. sciadophylla* as myrmecophytic resulted in a similar pattern, but with slightly lower probabilities of myrmecophytism in the common ancestor of core *Cecropia*.

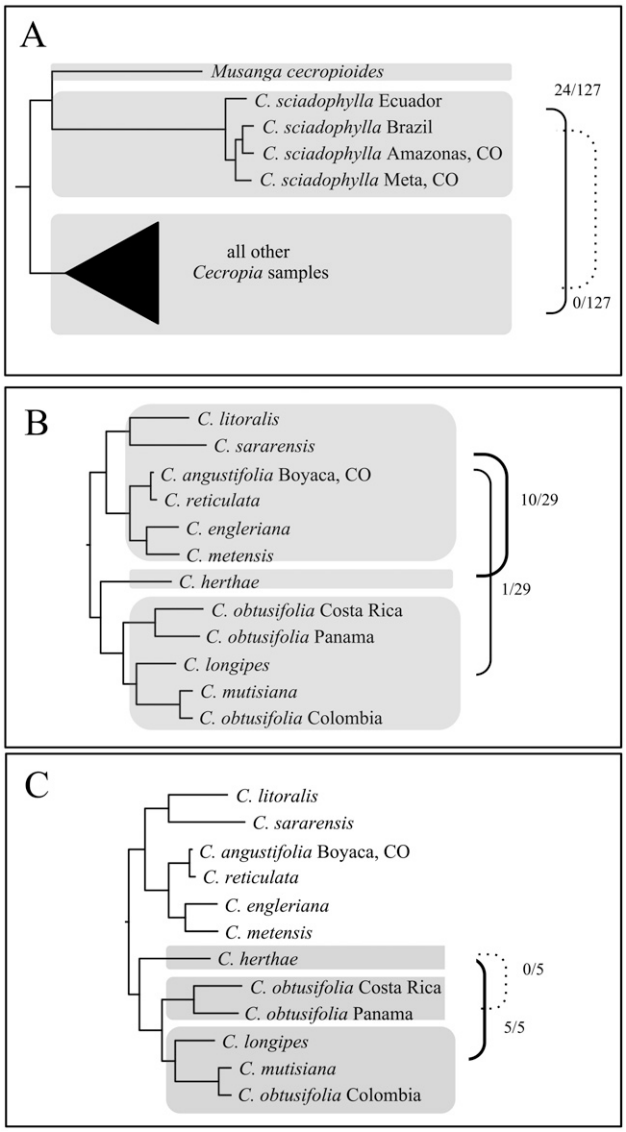

FIG. 2. Results of D-statistic tests using RADseq SNP data for *Cecropia*. Shaded regions indicate clades in which species tested. Brackets indicate tests between groups with solid lines for comparisons with significant results and dashed lines for comparisons with no significance. Brackets connecting clades show the number of significant tests (after Bonferroni correction) out of the total for each group. A) Tests *C. sciadophylla* agains all other Cecropia samples, B) tests *C. herthae* with all closely related clades, and C) tests *C. herthae* with a sister clade. Result of D-statistic in A is not supported by what we know of *Cecropia* geographic history, while B and C are supported by different phylogenetic analyses of RADseq data.

DISCUSSION

Compared to previous molecular systematic studies of *Cecropia*, the relatively larger volume of RADseq data was able to resolve *Cecropia* phylogeny at the species level with a highly supported backbone when a few genes were insufficient (Treiber et al. 2016; Gutiérrez-Valencia et al. 2017). Despite uncertainty about the ancestral state for myrmecophytism in the common ancestor of *Cecropia* owing to the deep divergence of the *C. sciadophylla*/*Musanga* clade from *Cecropia* s.s., we identified a major ancient split in the group with intriguing patterns of ant trait variation between the two clades. RADseq analyses strongly supporting *Musanga* along with *C. sciadophylla* as sister to the remaining *Cecropia* samples confirms the result of our molecular analysis of the tribe Cecropieae (Treiber et al. 2016). The previous study sampled rather few *Cecropia* species and adding half of all recognized *Cecropia* still supported the *Musanga*/*C. sciadophylla* clade. This result contradicts the only other molecular study based on seven loci that weakly supported *Musanga* as sister to *Cecropia* (Gutiérrez-Valencia et al. 2017). We did not sample a second species of *Musanga* but this taxon is an east African montane endemic that seems likely to be derived from the widespread Afrotropical *M. cecropioides*. *Musanga* is similar to *Cecropia* in both habitat and morphology, but lacks structures (i.e. trichilia) that are associated with the ant mutualism (Coombe and Hadfield 1962; Treiber et al. 2016; Gutiérrez-

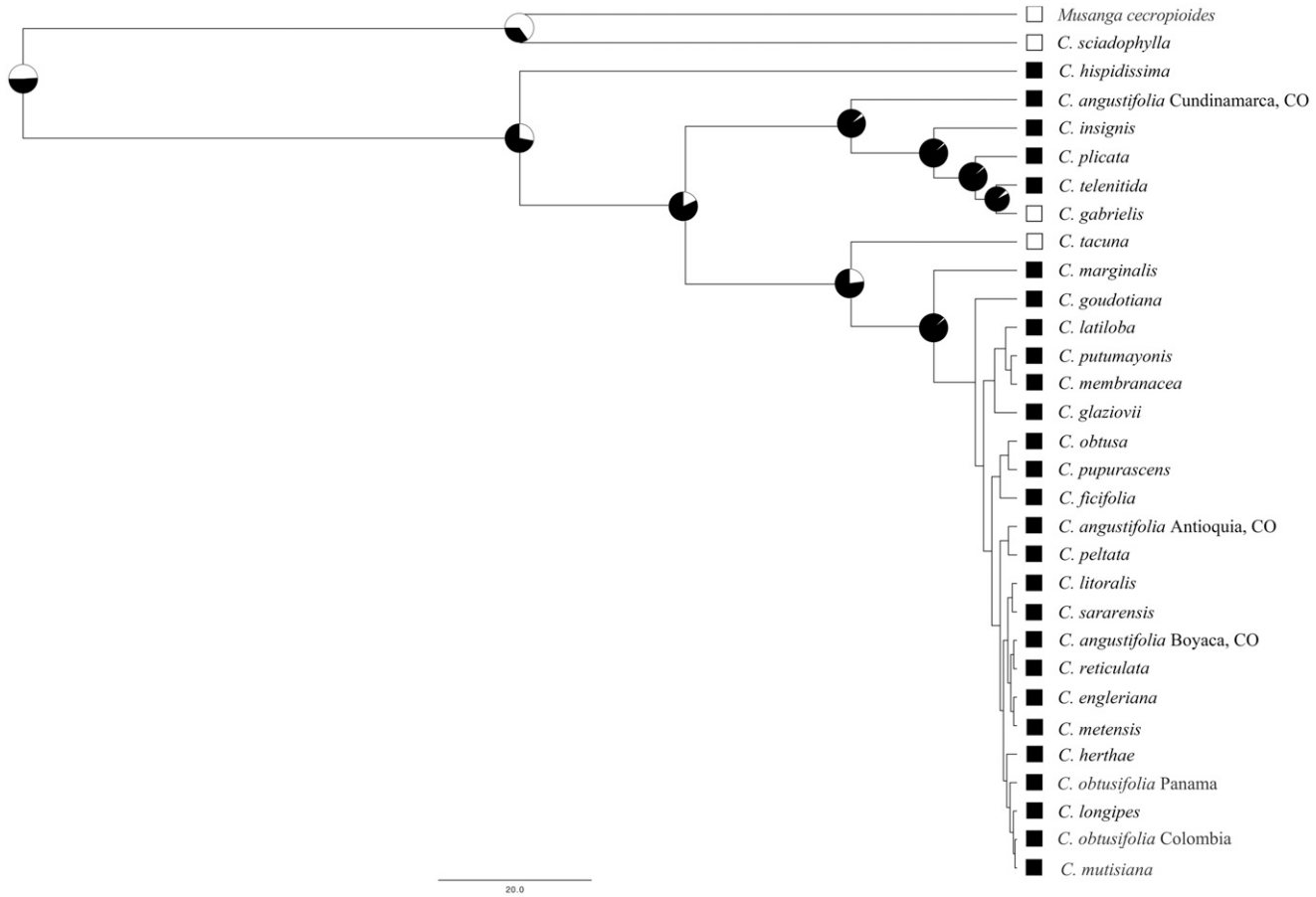

FIG. 3. Maximum likelihood ancestral state reconstruction of ant associations in *Cecropia* based on an equal rates model of evolutionary transitions. Pie charts at nodes represent the probability of that ancestor lacking ant associations (white) or with ant associations present (black). Nodes without circles had absolute probability of myrmecophytism. The squares at the tips of the branches indicate the state coded for species in the same format as above. Non-myrmecophytic *Coussapoa* (not shown) was the outgroup.

Valencia et al. 2017). *Cecropia sciadophylla* is one of the few species of *Cecropia* that always lack trichilia (Berg and Franco-Rosselli 2005). Janzen and McKey (1977) stated that if *Musanga* occurred within the range of *Cecropia* it would have been placed in the genus. The sister relationship of *C. sciadophylla* and *Musanga* suggests the need for taxonomic revision to either transfer *C. sciadophylla* to *Musanga* or to synonymize *Musanga* with *Cecropia* if *Cecropia* is to remain monophyletic. More complete sampling of *Musanga* and *Cecropia* is needed to choose among these options.

Berg and Franco-Rosselli (2005) acknowledged that few groupings could be recognized within *Cecropia* but they did define three, based on shared morphological characteristics and geography. With molecular data for many of the species they recognized, we can begin to compare how their groupings compare with phylogenetic relatedness. Berg and Franco-Rosselli (2005) recognized two groups based on morphology, the *C. peltata*-group and the *C. telenitida*-group (Table 3). Our analysis placed four members of the *C. peltata*-group into a clade, but this clade also included *C. angustifolia* from Antioquia, Colombia and *C. reticulata*. Also, *C. peltata*, for which the group was named, was part of an even larger clade including four species not grouped with it by Berg and Franco-Rosselli (2005). Our study also lacked four species of the *C. peltata*-group, so further sampling is needed to assess its taxonomic validity. We sampled two out of seven species in the *C. telenitida*-group (*C. gabrielis* and *C. telenitida*) and they were sister in the phylogeny, but embedded in a larger clade including *C. angustifolia* from Cundinamarca, Colombia, *C. insignis*, and *C. plicata*. Although Berg and Franco-Rosselli (2005) did not include *C. plicata* and *C. insignis* with the *C. telenitida*-group, our findings suggest that they might be associated with it.

Berg and Franco-Rosselli (2005) also recognized a group of seven species based on their presence in the Guyana region and broad leaf segments (Table 3), and the three species that we sampled from the group (*C. ficifolia*, *C. obtusa*, and *C. purpurascens*) formed a clade. The Guyana Shield is known for its distinctive weathered soils and high level of plant endemism (Berry and Riina 2005). It is not uncommon to find habitat specialists in this region with narrow geographic ranges or occasionally in similar habitats in neighboring regions (Berry and Riina 2005). Three species restricted to the Guyana region (*C. granvilleana*, *C. kavanayensis*, and *C. angulata*) have yet to be sequenced. Although a more complete sampling is needed, it appears that geography and morphological similarity may predict a modest degree of relatedness in *Cecropia*.

Our exploration of the *Cecropia* species concept supported our expectations that morphologically homogenous samples of currently circumscribed species would be monophyletic while morphologically heterogeneous species would be polyphyletic. As we predicted, both morphologically

TABLE 3. Three main groupings and distinctive features of a subset of *Cecropia* species according to Berg and Franco-Rosselli (2005). Species in bold were included in this study.

	C. peltata group	*C. telenitida* group	Guyana group
Features	Peltate stigma	Leafy twigs and the upper leaf surface glabrous and the outer surface of the stipules and spathes are either glabrous or villous	Largely associated with the Guyana region and with relatively few and often broad lamina segments
Species	*C. concolor*	*C. albicans*	*C. angulata*
	C. engleriana	*C. bullata*	***C. ficifolia***
	C. granvilleana	***C. gabrielis***	*C. distachya*
	C. litoralis	*C. maxima*	*C. kavanayensis*
	C. metensis	*C. telealba*	***C. obtusa***
	C. pachystachya	***C. telenitida***	***C. purpurascens***
	C. peltata		
	C. sararensis		
	C. schreberiana		

homogeneous species, *C. membranacea* and *C. sciadophylla*, were monophyletic whereas morphologically heterogeneous species *C. obtusifolia* and *C. angustifolia* were polyphyletic. The three samples of *C. obtusifolia* were localized in part of the tree, although with other species (*C. longipes* and *C. multisiana*), while samples of *C. angustifolia* were scattered among three divergent branches.

This finding and another recent study (Santos et al. 2020) suggest that more population genetic studies are needed to improve the *Cecropia* species concept. Santos et al. (2020) sequenced multiple samples of *C. pachystachya* and suggested the reinstatement of four synonyms. Such evidence renders problematic the pooling of genetic data from multiple individuals, especially from different geographic regions, for a given species when inferring phylogeny, as in Gutiérrez-Valencia et al. (2017). Future work ought to consider more intensive sampling of populations, particularly for taxa with numerous synonyms. Penagos Zuluaga et al. (2021) offer an example with similar data and more complete sampling to propose a phylogenetic classification of a difficult species complex in Supraocotea (Lauraceae). In our case, we sampled only a single individual of *C. peltata* while Berg and Franco-Rosselli (2005) recognized 11 synonyms. This species shares the same extent of morphological heterogeneity, geographical range, and synonymy as the two species that we found to be polyphyletic.

Our specimen of *C. angustifolia* from Boyaca, Colombia was collected near the type locality for the synonym *C. moniquerana* Cuatrecasas and a distinct phylogenetic position from the other samples would favor reinstating *C. moniquerana*. However, a formal taxonomic change seems to us premature in the presence of 12 other synonyms and only two other samples of the species complex in our study. We choose instead to speculate about how physical geography might result in reproductive isolation. The *C. angustifolia* complex is naturally distributed in the montane zone (~800–2400 m) from Mexico to Bolivia, where discontinuities among mountain ranges in suitable habitat could interrupt gene flow and contribute to allopatric speciation. Ecotypic variation in the complex along the precipitation gradient between wet and dry montane regions also deserves further genetic study to disentangle locally adapted populations from divergent taxa. Compared to *C. angustifolia,* the morphologically homogenous *C. sciadophylla* is equally widespread but occurs in low elevation forest (sea level to ~1000 m) across the Amazon basin, with a relatively more uniform habitat and uninterrupted distribution. Beyond these environmental geographic explanations, we considered that gene flow among species might also contribute to morphological heterogeneity.

We hypothesized that heterogeneous species would show more evidence of introgression than homogeneous species, but our results did not support this contention. Tests with the D-statistic including homogeneous *C. membranacea* and close relatives were significant with *C. latiloba* and each *C. membranacea* individual. No tests involving individuals of morphologically heterogeneous *C. obtusifolia* and their closest relatives detected patterns of introgression whereas tests for two out of three *C. angustifolia* individuals were significant (Fig. S1, Treiber et al. 2022). D-tests also investigated the unstable placement of *C. herthae* in phylogenies based on different matrices. Significant deviations in allele patterns were frequently detected between *C. herthae* and close relatives (Fig. 2B) despite strong support for its position as sister to the *C. obtusifolia/C. longipes/C. mutisiana* clade in most analyses. We conclude that this taxon, its close relatives, and *C. angustifolia* are candidates for more detailed population genetic studies of potential hybridization and introgression.

Test results for *C. sciadophylla*, *Musanga*, and the core *Cecropia* clade highlight a problem with applying the D-statistic to deeply divergent lineages. No tests including *C. sciadophylla* and the rest of *Cecropia* were significant, whereas one-fifth of the tests with African *Musanga* and neotropical *Cecropia* pointed to a history of gene flow (Fig. 2A). This seems highly unlikely given that *Musanga* and the core *Cecropia* appear to have been geographically isolated by the Atlantic Ocean and RADseq branch lengths (Fig. 1) suggest that their separation occurred several million years ago. The D-statistic might be sensitive to the extent of genetic divergence and could yield type I error in the case of deeply diverged clades. Eaton et al. (2015) demonstrated the difficulty of detecting introgression over deep evolutionary time scales when studying oaks. Missing taxa also complicate the inference of introgression. Choosing taxa to include in a D-test the absence of a complete *Cecropia* phylogeny means that we might have missed the most closely related, and therefore more appropriate species, for estimating gene flow. These considerations limit the strength of our conclusions about introgression until we have a more complete sampling of the genus and a better understanding of the D-statistic. We suggest that future studies consider using HyDe (Blischak et al. 2018) as an alternative to the D-statistic.

We used our RADseq phylogeny to investigate the ancestral state of myrmecophytism in the genus. Sampling approximately half of the recognized species in the genus, we closely approximated the genus-wide ratio of character states for ant association (~16%). While further study of *Cecropia* is needed, we have established a major ancient split in the group with intriguing patterns of ant trait variation between the two clades. We recognize that inferences of ancestral states could change depending on how the species we did not sample are related to those included in our analysis. For instance, if even a few myrmecophytic species joined the *Musanga*/*C. sciadophylla* clade, we would most probably infer a myrmecophytic ancestor for the entire clade. At present, it appears that myrmecophytism evolved once and was lost several times. Based on morphology we predict that at least some ant less species, such as *C. hololeuca* Miq. (not included in this study), which lack trichilia and is the only *Cecropia* s.l. other than *Musanga* to lack a spathe enclosing the emerging inflorescence (Berg and Franco-Rosselli 2005; Wheeler 1942); belong to the group with *C. sciadophylla*/*Musanga*. Gutiérrez-Valencia et al. (2017) suggested that *C. sciadophylla* and *C. hololeuca* are closely related but the relationship was not highly supported. If *C. hololeuca* is indeed a member of the *C. sciadophylla*/*Musanga* clade, it would strengthen the case for myrmecophytism to have evolved in the sister group, *Cecropia* s.s.

Coding of myrmecophytism as a binary character can be challenging because ant associations sometimes vary within species. Species such as *C. angustifolia* have ranges extending to 2000–2400 m, above the habitable zone for ants, and may therefore lack ant associations in part of their altitudinal range (Janzen 1973a, 1973b). However, ant-associated traits like trichilia are still present. Host associations also vary within species such as *C. hispidissima* that usually, but not exclusively, host *Pachycondyla* Smith rather than *Azteca*. It is commonly observed that species associated with *Azteca* may be inhabited by other ant genera (Wheeler 1942; Berg and Franco-Rosselli 2005). Understanding this variability and its causes requires intensive sampling throughout the geographic range of many species and experimental work that was beyond the scope of this study.

To better understand the evolution of myrmecophytism in *Cecropia*, population genetic and phylogenetic studies of the associated ants are needed. *Cecropia* mainly associate with the genus *Azteca*, which has not received broad molecular phylogenetic study and morphological species concepts were thought to be problematic because of conflicting and homoplasious character states (Ayala et al. 1996). The only phylogenetic analysis of *Azteca* including eight myrmecophytic lineages and a single gene suggested that *Cecropia*-inhabiting *Azteca* are not monophyletic (Ayala et al. 1996).

Azteca species typically build large carton nests suspended from tree branches (Wheeler 1910, 1942; Lucas et al. 2017). These nests are reduced and have varied construction when built in *Cecropia* (Marting et al. 2018). For example, some species build spindle-shaped carton nests around the trunks of the trees that deform the trunk and numerous exit holes are observed from domatia in the vicinity of the nest (Berg et al. 1990; Longino 1991a, 1991b; Ayala et al. 1996). Other *Azteca* create a cylindrical carton nest that does not deform the trunk and exit holes from domatia are located at a distance from the carton (Longino 1991a, 1991b; Ayala et al. 1996; Lucas et al. 2019). Gaining better insight into the evolutionary history of myrmecophytism in *Cecropia* will require deeper investigation of the phylogeography and host associations of the ants. Considering the potential for repeated colonization of a given plant lineage by different ant lineages and subsequent breakdown of mutualism (Weiblen and Treiber 2015), a comparative phylogeography is needed. Recent examples of insight gained from this approach to understanding the evolution of other ant-plant mutualisms include Torres Jimenez et al. (2021) and Chomicki and Renner (2016).

In conclusion, our preliminary examination of the origin and loss of myrmecophytism in *Cecropia* suggests a potential deep split between myrmecophytic and non-myrmecophytic lineages but the ambiguity of the ancestral state reiterates the need for more sampling of the group. The resolution and support provided by RADseq suggest that this approach might be sufficient to more thoroughly estimate phylogeny with the remainder of accepted names and synonyms that were not sampled in this study. However, weaknesses of RADseq include potential sensitivity to unbalanced sampling designs (such as unequal numbers of samples per species) and the inability to associate the allelic data with annotated gene regions. In future, it will be important to leverage annotated genomes, new sequencing technologies, and more intensive sampling to improve species concepts and phylogenetic knowledge for *Cecropia*. We anticipate that such an approach may also hold promise for studies of large, tropical woody genera where phylogenetic analysis has proven difficult.

Acknowledgments

We would like to thank C. Sagers for encouraging and supporting our work with *Cecropia*. We also thank J. Bevington, P. Barriga, A. Gaglioti, and E. Kami for fresh plant material and the Paris Herbarium and the Missouri Botanical Garden Herbarium for access to herbarium specimens. We would like to acknowledge A. Guensberg and G. Kandlikar for laboratory assistance. This manuscript was improved by valuable comments from J. W. Dalling. The work was supported by a US National Science Foundation Grant to GW and SM (DEB-1132916). Support for ET was provided by the Bernard and Jean Phinney Graduate Fellowship in Plant Molecular Biology from the Department of Plant Biology, the Joyce Davenport Summer Fellowship the Bell Museum of Natural History, the College of Biological Sciences Excellence Fellowship from the University of Minnesota. GW was supported by the Sanford Chair at Universidad de los Andes. PCZ acknowledges support from the University of South Florida.

Author Contributions

Erin L Treiber made substantial contribution to the concept and design of the study, acquisitions of data, analysis and interpretation of data, and led the writing of the manuscript. Paul-Camilo Zalamea was involved in acquisition of plant material for analyses, revising the manuscript critically for important intellectual content, and drafting some sections of the manuscript. María Fernanda Torres made contribution to the concept of the study, was key in collection of material, and supported the data analyses. Santiago Madriñán made substantial contribution to the concept and design of the study and was involved in revising the manuscript. George D. Weiblen made substantial contribution to the concept and design of the study, critically reviewing analysis and interpretation of data, and was involved in revising the manuscript.

Literature Cited

Agrawal, A. 1998. Leaf damage and associated cues induce aggressive ant recruitment in a neotropical ant-plant. *Ecology* 79: 2100–2112.

Agrawal, A. and B. Dubin-Thaler. 1999. Induced responses to herbivory in the Neotropical ant-plant association between *Azteca* ants and *Cecropia* trees: Response of ants to potential inducing cues. *Behavioral Ecology and Sociobiology* 45: 47–54.

Appelhans, M. S., C. Paetzold, K. R. Wood, and W. L. Wagner. 2020. RADseq resolves the phylogeny of Hawaiian *Myrsine* (Primulaceae)

and provides evidence for hybridization. *Journal of Systematics and Evolution* 58: 823–840.

Aronesty, E. 2011. ea-utils: Command-line tools for processing biological sequencing data. https://github.com/ExpressionAnalysis/ea-utils.

Ayala, F., J. Wetterer, J. Longino, and D. Hartl. 1996. Molecular phylogeny of *Azteca* ants (Hymenoptera: Formicidae) and the colonization of *Cecropia* trees. *Molecular Phylogenetics and Evolution* 5: 423–428.

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, W. A. Cresko, and E. A. Johnson. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3: e3376.

Berg, C. C. and P. Franco-Rosselli. 2005. Cecropia. *Flora Neotropica, Monograph 94.* New York: The New York Botanical Garden.

Berg, C. C., R. W. A. P. Akkermans, and E. C. H. Heusden. 1990. *Cecropiaceae:* Coussapoa *and* Pourouma, *with an Introduction to the Family.* New York: The New York Botanical Garden.

Berry, P. E. and R. Riina. 2005. Insights into the diversity of the Pantepui Flora and the biogeographic complexity of the Guayana Shield. *Biologiske Skrifter* 55: 145–167.

Blattner, F. 2001. Molecular analysis of phylogenetic relationships among myrmecophytic *Macaranga* species (Euphorbiaceae). *Molecular Phylogenetics and Evolution* 19: 331–344.

Blischak, P. D., J. Chifman, A. D. Wolfe, and L. S. Kubatko. 2018. HyDe: A Python package for genome-scale hybridization detection. *Systematic Biology* 67: 821–829.

Burnham, R. J. and A. Graham. 1999. The history of neotropical vegetation: New developments and status. *Annals of the Missouri Botanical Garden* 86: 546–589.

Cavender-Bares, J., A. González-Rodríguez, D. A. R. Eaton, A. A. L. Hipp, A. Beulke, and P. S. Manos. 2015. Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. *Molecular Ecology* 24: 3668–3687.

Chomicki, G. and S. S. Renner. 2016. Evolutionary relationships and biogeography of the ant-epiphytic genus *Squamellaria* (Rubiaceae: Psychotrieae) and their taxonomic implications. *PLoS One* 11: e0151317.

Coombe, D. E. and W. Hadfield. 1962. An analysis of the growth of *Musanga cecropioides*. *Journal of Ecology* 50: 221–234.

Davey, J. W., P. A. Hohenlohe, P. D. Etter, J. Q. Boone, J. M. Catchen, and M. L. Blaxter. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews. Genetics* 12: 499–510.

Davies, S. J., S. K. Y. Lum, R. Chan, and L. K. Wang. 2001. Evolution of myrmecophytism in western Malesian *Macaranga* (Euphorbiaceae). *Evolution* 55: 1542–1559.

Dejean, A., F. Petitclerc, O. Roux, J. Orivel, and C. Leroy. 2012. Does exogenic food benefit both partners in an ant-plant mutualism? The case of *Cecropia obtusa* and its guest *Azteca* plant-ants. *Comptes Rendus Biologies* 335: 214–219.

Deng, M., X.-L. Jiang, A. L. Hipp, P. S. Manos, and M. Hahn. 2018. Phylogeny and biogeography of East Asian evergreen oaks (*Quercus* section *Cyclobalanopsis*; Fagaceae): Insights into the Cenozoic history of evergreen broad-leaved forests in subtropical Asia. *Molecular Phylogenetics and Evolution* 119: 170–181.

Doyle, J. J. and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 97: 11–15.

Du, Z.-Y., A. Harris, and Q.-Y. Xiang. 2020. Corrigendum to "Phylogenomics, co-evolution of ecological niche and morphology, and historical biogeography of buckeyes, horsechestnuts, and their relatives (Hippocastaneae, Sapindaceae) and the value of RAD-Seq for deep evolutionary inferences back to the Late Cretaceous.". *Molecular Phylogenetics and Evolution* 150: 106889.

Eaton, D. A. R. 2014. Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30: 1844–1849.

Eaton, D. A. R., A. L. Hipp, A. González-Rodríguez, and J. Cavender-Bares. 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution* 69: 2587–2601.

Eaton, D. A. R. and R. H. Ree. 2013. Inferring phylogeny and introgression using RADseq data: An example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology* 62: 689–706.

Fáveri, S. and H. L. Vasconcelos. 2004. The *Azteca-Cecropia* association: Are ants always necessary for their host plants? *Biotropica* 36: 641–646.

Green, R. E., J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, et al. 2010. A draft sequence of the Neandertal genome. *Science* 328: 710–722.

Gutiérrez-Valencia, J., G. Chomicki, and S. S. Renner. 2017. Recurrent breakdowns of mutualisms with ants in the neotropical ant-plant genus *Cecropia* (Urticaceae). *Molecular Phylogenetics and Evolution* 111: 196–205.

Harmon, L., J. Weir, C. Brock, R. Glor, and W. Challenger. 2008. GEIGER: Investigating evolutionary radiations. *Bioinformatics* 24: 129–131.

Janzen, D. H. 1969. Allelopathy by myrmecophytes: The ant *Azteca* as an allelopathic agent of *Cecropia*. *Ecology* 50: 147–153.

Janzen, D. H. 1973a. Dissolution of mutualism between *Cecropia* and its *Azteca* ants. *Biotropica* 5: 15–28.

Janzen, D. H. 1973b. Sweep samples of tropical foliage insects: Effects of seasons, vegetation types, elevation, time of day, and insularity. *Ecology* 54: 687–708.

Janzen, D. H. and D. McKey. 1977. *Musanga cecropiodes* is a *Cecropia* without its ants. *Biotropica* 9: 57.

Latteman, T. A., J. E. Mead, M. A. Duvall, C. C. Bunting, and J. M. Bevington. 2014. Differences in anti-herbivore defenses in non-myrmecophyte and myrmecophyte *Cecropia* trees. *Biotropica* 46: 652–656.

Longino, J. T. 1989. Taxonomy of the *Cecropia*-inhabiting ants in the *Azteca alfari* species group (Hymenoptera: Formicidae): Evidence for two broadly sympatric species. *Journal of Natural History* 25: 1571–1602.

Longino, J. T. 1991a. *Azteca* ants in *Cecropia* trees: Taxonomy, colony structure, and behaviour. Pp. 271–288 in *Ant-Plant Interactions*, eds. C. R. Huxley and D. F. Cutler. Oxford: Oxford University Press.

Longino, J. T. 1991b. Taxonomy of the *Cecropia*-inhabiting *Azteca* ants. *Journal of Natural History* 25: 1571–1602.

Lucas, J., B. Bill, B. Stevenson, and M. Kaspari. 2017. The microbiome of the ant-built home: The microbial communities of a tropical arboreal ant and its nest. *Ecosphere* 8: e01639.

Lucas, J. M., A. A. Madden, C. A. Penick, M. J. Epps, P. R. Marting, J. L. Stevens, D. J. Fergus, R. R. Dunn, and E. K. Meineke. 2019. *Azteca* ants maintain unique microbiomes across functionally distinct nest chambers. *Proceedings. Biological Sciences* 286: 20191026.

Marting, P. R., N. M. Kallman, W. T. Wcislo, and S. C. Pratt. 2018. Ant-plant sociometry in the *Azteca-Cecropia* mutualism. *Scientific Reports* 8: 1–8.

Miller, M. A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Pp. 1–8 in *Proceedings of the Gateway Computing Environments Workshop (GCE)*. New Orleans: Gateway Computing.

Paetzold, C., K. R. Wood, D. A. R. Eaton, W. L. Wagner, and M. S. Appelhans. 2019. Phylogeny of Hawaiian *Melicope* (Rutaceae): RAD-seq resolves species relationships and reveals ancient introgression. *Frontiers in Plant Science* 10: 1–16.

Parchman, T. L., J. P. Jahner, K. A. Uckele, L. M. Galland, and A. J. Eckert. 2018. RADseq approaches and applications for forest tree genetics. *Tree Genetics & Genomes* 14: 39.

Penagos Zuluaga, J. C., H. Werff, B. Park, D. A. R. Eaton, L. S. Comita, S. A. Queenborough, and M. J. Donoghue. 2021. Resolved phylogenetic relationships in the *Ocotea* complex (*Supraocotea*) facilitate phylogenetic classification and studies of character evolution. *American Journal of Botany* 108: 664–679.

Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.

Richardson, J. E., R. T. Pennington, T. D. Pennington, and P. M. Hollingsworth. 2001. Rapid diversification of a species-rich genus of Neotropical rain forest trees. *Science* 293: 2242–2245.

Rønsted, N., G. Salvo, and V. Savolainen. 2007. Biogeographical and phylogenetic origins of African fig species (*Ficus* section *Galoglychia*). *Molecular Phylogenetics and Evolution* 43: 190–201.

RStudio Team. 2016. RStudio: Integrated Development for R, version 1.1.423. http://www.rstudio.com/.

Rubin, B. E. R., R. H. Ree, and C. S. Moreau. 2012. Inferring phylogenies from RAD sequence data. *PLoS One* 7: e33394.

Sagers, C., S. Ginger, and R. Evans. 2000. Carbon and nitrogen isotopes trace nutrient exchange in an ant-plant mutualism. *Oecologia* 123: 582–586.

Santos, J. D. O., F. L. Zchonski, L. Pilati, A. L. Gaglioti, S. Romaniuc-Neto, and P. R. Da-Silva. 2020. Morphological and DNA analyses suggest the reinstatement of four synonymized *Cecropia* species. *Tree Genetics & Genomes* 16: 51.

Smith, S. A. and M. J. Donoghue. 2008. Rates of molecular evolution are linked to life history in flowering plants. *Science* 322: 86–89.

Soltis, D. E. 1998. *Molecular Systematics of Plants II: DNA Sequencing.* New York: Springer US.

Soltis, P. S., D. E. Soltis, and J. J. Doyle. 1992. *Molecular Systematics of Plants*. New York: Springer US.

Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.

Straub, S. C. K., M. Parks, K. Weitemier, M. Fishbein, R. C. Cronn, and A. Liston. 2012. Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American Journal of Botany* 99: 349–364.

Torres Jimenez, M. F., G. N. Stone, A. Sanchez, and J. E. Richardson. 2021. Comparative phylogeography of an ant-plant mutualism: An encounter in the Andes. *Global and Planetary Change* 205: 103598.

Treiber, E. L. 2017. *Phylogenetics of Cecropieae (Urticaceae) and the Evolution of an Ant-Plant Mutualism in* Cecropia. Ph.D. dissertation. Minneapolis: University of Minnesota.

Treiber, E. L., A. L. Gaglioti, S. Romaniuc-Neto, S. Madriñán, and G. D. Weiblen. 2016. Phylogeny of the Cecropieae (Urticaceae) and the evolution of an ant-plant mutualism. *Systematic Botany* 41: 56–66.

Treiber, E. L., P.-C. Zalamea, M. F. Torres, S. Madriñán, and G. D. Weiblen. 2022. Data from: Molecular systematics, species concepts, and myrmecophytism in *Cecropia* (Cecropieae: Urticaceae): Insights from restriction-site associated DNA. Dryad Digital Repository. https://doi.org/10.5061/dryad.kwh70rz2x.

Triono, E. L., A. Brown, J. West, and M. Crisp. 2007. A phylogeny of *Pouteria* (Sapotaceae) from Malesia and Australasia. *Australian Systematic Botany* 29: 107–118.

Vasconcelos, H. and A. Casimiro. 1997. Influence of *Azteca alfari* ants on the exploitation of *Cecropia* trees by a leaf-cutting ant. *Biotropica* 29: 84–92.

Vieira, A. S., O. Faccenda, W. F. Antonialli-Junior, and W. D. Fernandes. 2010. Nest structure and occurrence of three species of *Azteca* (Hymenoptera, Formicidae) in *Cecropia pachystachya* (Urticaceae) in non-floodable and floodable pantanal areas. *Revista Brasileira de Entomologia* 54: 441–445.

Weiblen, G. D. and E. L. Treiber. 2015. Evolutionary origins and diversification of mutualism. Pp. 36–56 in *Mutualism*, ed. J. L. Bronstein. Oxford: Oxford University Press.

Wheeler, W. M. 1910. *Ants: Their Structure, Development and Behavior*. New York: Columbia University Press.

Wheeler, W. M. 1942. *Studies of Neotropical Ant-Plants and Their Ants*. Cambridge: Harvard College.

Zalamea, P. C., P. Stevenson, S. Madrinan, P. Aubert, and P. Heuret. 2008. Growth pattern and age determination for *Cecropia sciadophylla* (Urticaceae). *American Journal of Botany* 95: 263–271.

Zalamea, P. C., F. Munoz, P. R. Stevenson, C. E. T. Paine, C. Sarmiento, D. Sabatier, and P. Heuret. 2011. Continental-scale patterns of *Cecropia* reproductive phenology: evidence from herbarium specimens. *Proceedings. Biological Sciences* 278: 2437–2445.

Zalamea, P. C., P. Heuret, C. Sarmiento, M. Rodriguez, A. Berthouly, S. Guitet, E. Nicolini, C. Delnatte, D. Barthélémy, and P. R. Stevenson. 2012. The genus *Cecropia*: a biological clock to estimate the age of recently disturbed areas in the Neotropics. *PLoS One* 7: e42643.

APPENDIX 1. Summary of RAD sequencing data for specimens included in phylogenetic analyses including: species, location, collector information, and loci per sample for the largest data set (RADseq5) and the smallest (RADseq15) used for analyses.

Cecropia angustifolia Trécul, Antioquia, Colombia, *Torres 81* (ANDES), 37544, 21342; *Cecropia angustifolia* Trécul, Cundinamarca, Colombia, *Treiber 01* (ANDES), 40237, 27795; *Cecropia angustifolia* Trécul, Boyaca, Colombia, *Zalamea 48* (ANDES), 43298, 32258; *Cecropia engleriana* Snethl., Ecuador, *Barriga & Alvia 2009-009* (MIN), 42982, 32628; *Cecropia ficifolia* Warb. & Snethl., Peru, *Barriga & Bellota 2009-168* (MIN), 31185, 15220; *Cecropia gabrielis* Cuatrec., Colombia, *Treiber 26* (ANDES), 32078, 15360; *Cecropia glaziovii* Snethl., Brazil, *Gaglioti 156* (SP), 41835, 33857; *Cecropia goudotiana* Trécul, Colombia, *Treiber 10* (ANDES), 41585, 33136; *Cecropia herthae* Diels, Ecuador, *Barriga 2009-091* (SP), 45504, 34634; *Cecropia hispidissima* Cuatrec., Colombia, *Treiber 49* (ANDES), 37843, 24837; *Cecropia insignis* Liebm., Colombia, *Zalamea 70* (ANDES), 37880, 23503; *Cecropia latiloba* Miq., Ecuador, *Barriga & Alvia 2009-052* (MIN), 41914, 32670; *Cecropia litoralis* Snethl., Ecuador, *Barriga 2009-039* (MIN), 41408, 31403; *Cecropia longipes* Pittier, Colombia, *Torres 87* (ANDES), 22458, 8282; *Cecropia marginalis* Cuatrec., Ecuador, *Barriga 2009-004* (MIN), 43015, 32993; *Cecropia membranacea* Trécul, Ecuador, *Barriga 2009-001* (MIN), 42802, 33683; *Cecropia membranacea* Trécul, Amazonas, Colombia, *Torres 23* (ANDES),38314, 25225; *Cecropia membranacea* Trécul, Casanare, Colombia, *Zalamea 54* (ANDES), 41791, 33076; *Cecropia metensis* Cuatrec., Colombia, *Zalamea 52* (ANDES), 39019, 26330; *Cecropia mutisiana* Mildbr. ex Cuatrec., Colombia, *Zalamea 73* (ANDES), 43820, 34298; *Cecropia obtusa* Trécul, Brazil, *Gaglioti 159* (SP), 39435, 29640; *Cecropia obtusifolia* Bertol., Costa Rica, *Barriga 2010-010* (MIN), 33754, 16605; *Cecropia obtusifolia* Bertol., Colombia, *Treiber 02* (ANDES), 42079, 30869; *Cecropia obtusifolia* Bertol., Panama, *Weiblen 3004* (MIN), 46780, 34821; *Cecropia peltata* L., Colombia, *Treiber 07* (ANDES), 38756, 25076; *Cecropia plicata* Cuatrec., Colombia, *Torres 104* (ANDES), 12857, 1666; *Cecropia purpurascens* C.C.Berg, Brazil, *Gaglioti 174* (SP), 25829, 9674; *Cecropia putumayonis* Cuatrec., Ecuador, *Barriga & Alvia 2009-010* (MIN), 42249, 33083; *Cecropia reticulata* Cuatrec., Colombia, *Torres 78* (ANDES), 37334, 20260; *Cecropia sararensis* Cuatrec., Colombia, *Torres 37* (ANDES), 17584, 5315; *Cecropia sciadophylla* Mart., Ecuador, *Barriga & Alvia 2009-090* (MIN), 37446, 25472; *Cecropia sciadophylla* Mart., Brazil, *Gaglioti 124* (SP), 39467, 27043; *Cecropia sciadophylla* Mart., Amazonas, Colombia, *Torres 30* (ANDES), 39031, 28311; *Cecropia sciadophylla* Mart., Meta, Colombia, *Zalamea 57* (ANDES), 37438, 23956; *Cecropia tacuna* C.C.Berg & P.Franco, Peru, *Bevington 64* (MIN), 41549, 33572; *Cecropia telenitida* (Cuatrec.), Colombia, *Torres 69* (ANDES), 39686, 26951; *Coussapoa floccosa* Akkermans & C.C.Berg, Brazil, *Gaglioti 104* (SP), 37446, 8901; *Musanga cecropioides* R. Br. ex Tedlie, Guinea, *Cabezas 114* (P), 39435, 31477; *Myrianthus arboreus* P.Beauv, Republic of the Congo, *Kami 242* (SP), 21739, 16227; *Pourouma tomentosa* Mart. ex Miq., Brazil, *Gaglioti 139* (SP), 18124, 6704.