The Systematics and Evolution of Lantaneae (Verbenaceae), a Molecular Phylogenetic Approach

Patricia Lu-Irving

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Richard G. Olmstead, Chair

Harvey D. (Toby) Bradshaw

Caroline A. E. Stromberg

Program Authorized to Offer Degree:

Department of Biology

University of Washington

Abstract

The systematics and evolution of Lantaneae (Verbenaceae), a molecular phylogenetic approach

Patricia Lu-Irving

Chair of the Supervisory Committee:

Professor Richard G. Olmstead

Department of Biology

Lantaneae are a morphologically variable group of 300-400 species, representing the largest tribe within Verbenaceae. They are widespread and diverse in the new world tropics and subtropics; some members are native to Africa, and others, most notably the *Lantana camara* complex, have spread across the globe as noxious weeds. Complex patterns of morphological parallelism have hindered taxonomic efforts within Lantaneae, and previous molecular phylogenetic studies have failed to resolve relationships within the tribe. The lack of variability among loci commonly used to infer phylogeny at the species level in plants suggests that Lantaneae are

recently radiated. With growing interest in the taxonomy of this difficult group, and growing recognition of the worldwide ecological and economic impacts of *Lantana camara*, there is a clear need for a well resolved phylogenetic hypothesis for Lantaneae. Species-level phylogenetic reconstruction in taxonomically complex, recently radiated lineages is a major challenge in plant systematics, and represents an opportunity to test the limitations of the molecular methods that are currently prevalent in modern systematic biology. Here, I have taken a multi-locus approach to resolve the pattern of diversification among a broad representative sample of the morphological, taxonomic, and geographic diversity of Lantaneae, demonstrating the effectiveness of the PPR gene family as phylogenetic tools. The results reveal that major genera are not monophyletic, with *Lantana* species belonging to two main clades, derived within a background of *Lippia* species. The small African genus *Coelocarpum* is the sister group to the tribe. Different loci reconstruct the species of *Aloysia*, and its affiliated genera, differently: either in a paraphyletic grade to the *Lantana-Lippia* complex, or as its sister group. A species tree reconstruction supports the hypothesis of sister clades. Within the *Lantana-Lippia* complex, fleshy fruits have evolved four times independently from dry-fruited ancestors, and are associated with higher speciation rates. At a broad scale, there is no clear pattern suggesting that fleshy fruits confer a dispersal advantage over dry fruits. My results place the origin of core Lantaneae in the Miocene, in subtropical South America, with different lineages subsequently migrating independently throughout the neotropics, into North America, and twice to Africa.

University of Washington Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Patricia Lu-Irving

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Chair of the Supervisory Committee:

_____

Richard G. Olmstead

Reading Committee:

_____

Harvey D. (Toby) Bradshaw

_____

Caroline A. E. Stromberg

Date:_____

Contents

List of Figures

List of Tables

Acknowledgments

I am deeply grateful to Dick Olmstead for his guidance, patience, and support during the undertaking and completion of this thesis. In addition to an advisor, Dick is a friend, and an outstanding role model, and I thank him for providing the mentorship and opportunities that have encouraged my independence as a scientist. Additionally, I thank my committee members: Toby Bradshaw, Caroline Stromberg, and Roger Buick for their involvement and advice, and for sharing their diverse perspectives with me. In acknowledging my mentors, I include heartfelt thanks to Murray Henwood and Andrew Perkins for fostering my early growth as a botanist. This research was carried out in collaboration with several outstanding taxonomists: Nataly O'Leary, whose energy and enthusiasm have been inspirational, Roger Sanders, whose diligence and proficiency I can only aspire to emulate, in addition to Fátima Salimena and Tânia Silva, whose experience and kindness I am very grateful for.

Many colleagues provided advice and insightful discussion on this thesis: Rebecca Harris, Valerie Soza, Yaowu Yuan, John Chau, Ryan Miller, and the graduate students and faculty of the Biology department at the University of Washington, particularly Adam Leaché and members of his lab, as well as three anonymous reviewers of Chapter I. Additional colleagues were generous with their time, expertise, and resources in aid of this research: Loreta Freitas, Verônica Thode, and the members of the Freitas lab at the Universidade Federal de Rio Grande do Sul; Lyderson Viccini and his students at the Universidade Federal de Juiz de Fora; Doug Ewing at the University of Washington's Botany Greenhouse. I am indebted to

yet more colleagues and friends for their assistance and companionship in the field: Hannah Marx, Yaowu Yuan, Segundo Leiva, Nancy Refulio, Marybel Morales, Luke Ledwich, Verônica Thode, Nara Mota, Marcos Toledo, Joyce Maschinski, Matt McElroy, Marcos Caraballo, Gerson Feliz. For assistance and companionship in the lab and office, I thank Anna O'Brien, Ben Meersman, Laura Frost, Audrey Ragsac, and many past and present members of the Olmstead lab and the Biology department at the University of Washington, including several colleagues already named above.

Plant material was obtained with the assistance of scientists and staff at CESJ, FTG, ICN, JBSD, KEW, MERL, MO, NY, PRE, RSA, SI, TEX-LL, UPR, UPR-RP, US, WTU; in particular, I thank Dave Giblin, Tom Wendt, Elizabeth Retief, and Frank Axelrod for their extra help and generosity.

Financial support was provided by the National Science Foundation (DEBs 0542493, 1020369, 1120802), the Biology department at the University of Washington (Plant Systematics Fellowship, Giles Award, Denton Fellowship, travel awards), the Graduate and Professional Student Senate at the University of Washington (travel awards), the Botanical Society of America, the American Society of Plant Taxonomists, and the Society of Systematic Biologists (Graduate Student Research Awards).

Finally, I would like to add a personal note of thanks to my family, in particular my mother, Guiying Lu, my father, Chris Irving, and my husband, Luke Ledwich, for their love and care in all aspects of my life, including the undertaking and completion of this work.

CHAPTER I: Investigating the evolution of Lantaneae using multiple loci[1]

SUMMARY

Lantaneae is an example of a taxonomically problematic, widespread and recently radiated Neotropical lineage. Taxonomy in Lantaneae is difficult due to complex, overlapping patterns of shifts in morphological traits among members; monophyly of its traditional genera cannot be assumed without additional information from molecular data. We took a multi-locus approach to infer the Lantaneae phylogeny, resolving major clades among a broad representative sample that covers the morphological, taxonomic and geographic diversity of this group. Data from multiple, independent loci reveal individual gene trees that are incongruent with one another, with varying degrees of support. Without reliable, applicable methods to determine the sources of such incongruence, and to resolve it, we present the consensus between well-supported topologies among our data sets as the best estimate of Lantaneae phylogeny to date. According to this consensus tree, fleshy fruits in Lantaneae have been derived from dry fruits at least five times; taxonomic schemes separating genera based on fruit characteristics are artificial. Lantaneae have shifted into the Neotropics from the South American subtropics, and have colonized Africa in at least two separate long-distance dispersal events. This study provides a first pass at a broad Lantaneae phylogeny, but two important areas remain unresolved:

[1] This article was first published in the Botanical Journal of the Linnean Society: **Lu-Irving, P. and R. G. Olmstead. 2013.** Investigating the evolution of Lantaneae (Verbenaceae) using multiple loci. *Botanical Journal of the Linnean Society* **171**: 103-119.

the position of *Acantholippia* relative to *Aloysia* species, and species-level relationships within the *Lantana-Lippia* clade.

INTRODUCTION

The Neotropics are globally renowned as a region of remarkable floristic diversity. Much of its species richness is concentrated in large, endemic (or nearly endemic) lineages, such as Cactaceae, Bromeliaceae, and Bignoniaceae (Gentry, 1982). Within these, and other, characteristic Neotropical lineages, "problematic" taxa are common: plant groups in which traditional classifications are at odds with newly obtained molecular evidence. Examples include *Mammillaria* Haw. in Cactaceae (Butterworth & Wallace, 2004), subfamily Bromelioideae in Bromeliaceae (Schulte et al., 2008; Sass & Specht, 2010), *Tabebuia* Gomes ex DC. (Grose & Olmstead, 2007), and tribe Bignonieae (Lohmann, 2006) in Bignoniaceae.

With the increasing range of modern tools available to systematists, great progress has been made in the last several years in untangling the evolutionary histories of difficult taxa within important Neotropical families. Recent examples, in addition to those cited above, can be found in cycads (Gonzalez et al., 2008), palms (Eiserhardt et al., 2011; Ludeña et al., 2011), Fabaceae (Torke & Schaal, 2008), and Podostemaceae (Tippery et al., 2011). Each of the lineages studied in these examples has in common particular characteristics which make it problematic: it is species-rich and geographically widespread, classifications within the group are historically difficult, and previous broad, molecular phylogenetic studies fail to

resolve relationships within it. Here we present an additional example from our work in Lantaneae: a morphologically diverse group of several hundred species, representing the most species-rich tribe within Verbenaceae.

Background information

After recent recircumscription (Marx et al., 2010), the tribe Lantaneae is monophyletic, containing two major genera (*Lantana* L. and *Lippia* L.) and seven smaller genera. It is sister to the tribe Verbeneae (Marx et al., 2010; Yuan et al., 2009 b). The two principal genera of Lantaneae comprise about 75% of its species. *Lippia* contains about 200 species, and *Lantana* about 150 species (Atkins, 2004); however, some taxonomists consider *Lantana* to contain too many names (López-Palacios, 1991; Verdcourt, 1992; Santos, 2002), and to have as few as 55 species (Sanders, 2001). Other genera are smaller: *Aloysia* Palau (30 spp.), *Phyla* Lour. (five spp.; O'Leary & Mulgura, 2012), *Nashia* Millsp. (seven spp.), *Acantholippia* Griseb. (six spp.), *Coelocarpum* Balf. f. (five spp.), *Burroughsia* Moldenke (two spp.; Moldenke, 1940), and monotypic *Xeroaloysia* Tronc. (numbers of species from Atkins, 2004, unless otherwise attributed). Many members of Lantaneae are of ecological and ethnobotanical significance in their natural settings: e.g., *Acantholippia salsoloides* Griseb., which is a community dominant in the Altiplano, and used locally as a culinary herb. Others are of global economic and/or ecological importance, e.g., *Aloysia citriodora* Palau (lemon verbena), commonly cultivated for its medicinal and culinary uses, and *Lantana camara* (lantana), a popular ornamental and weed of global significance.

The evolutionary history of Lantaneae presents a difficult problem. The large number of species in Lantaneae encompass a great deal of morphological variation, ranging from herbs, to shrubs, to small trees, with a diverse spectrum of leaf morphologies and inflorescence architectures. Members of Lantaneae are found in many different habitats, from moist lowland forests, to the fire-prone Cerrado, to the dry Altiplano; each with accompanying morphological adaptations. Attempts to partition this wide range of variation according to generic and infrageneric boundaries traditionally rely heavily on fruit morphology (Chamisso, 1832; Schauer, 1847; Briquet, 1895, 1904; Moldenke, 1959; Troncoso, 1974). According to one scheme, species with schizocarpous fruit are assigned to *Lippia*, and species with fleshy drupes are placed in *Lantana* (Schauer, 1847; Troncoso, 1974). Alternatively, the number of mericarps or pyrenes per fruit has also been used to separate *Lantana* from *Lippia* (Chamisso, 1832; Silva, 1999). However, generic boundaries in Lantaneae are blurred by species that are difficult to assign unambiguously to genus, presumably due to convergence in these (and other) important diagnostic traits. These confounding morphological patterns are consistent with recent radiation, as are the short branch lengths in Lantaneae found by the molecular study of Marx et al. (2010).

Adding to the problems associated with describing the wide range of morphologies within Lantaneae, the tribe is also geographically wide-ranging. The origin of Lantaneae is in subtropical South America, and its center of diversity is in the

Neotropics (Atkins, 2004; Marx et al., 2010; Olmstead, 2013). Its native distribution spans the southern states of the USA, Mexico and Central America, the Caribbean, and South America; a few species also occur on the other side of a trans-Atlantic disjunction, in Africa and Madagascar. Some members, most notably of the *Lantana camara* L. species group, have been globally introduced as ornamentals and spread as weeds, apparently hybridizing with native species in some parts of the Neotropics (Sanders, 1987), and further confusing taxonomic efforts. Native African species are assigned to both *Lantana* and *Lippia*, suggesting at least two distinct colonization events.

There is a growing effort to address the troublesome classification schemes within Lantaneae, and to produce generic revisions (e.g., Silva, 1999; Salimena, 2002; Silva & Salimena, 2002; Santos, 2002; Sanders, 2001, 2006; Siedo, 2008; O'Leary & Mulgura, 2012). However, because Lantaneae are species-rich, geographically widespread, and recently radiated, these taxonomic efforts are hindered by the common problems that such a group presents. Their focus is often on specific geographic regions, usually defined by political boundaries, which may or may not be of biogeographic significance. Additionally, many taxonomic revisions focus on single genera, traditionally circumscribed, under the implicit assumption that generic boundaries are of evolutionary significance. There is a clear need for a broad, well-resolved phylogenetic hypothesis for Lantaneae, which has yet to be addressed in detail in a molecular phylogenetic study.

Phylogeny reconstruction using multiple, independent loci

Phylogenetic systematic studies in plants over the last three decades have made great use of sequence data from chloroplast DNA, and recent studies that sample very broadly across large Neotropical groups continue to rely on it (e.g., Lohmann, 2006; Olmstead et al., 2008, 2009; Marx et al., 2010; Givnish et al., 2011; Bárcenas et al., 2011). However, the chloroplast genome has a lower rate of molecular change compared to the nuclear genome, and individual chloroplast loci often are insufficiently variable to provide resolution between species in recently diversified groups (Small et al., 2004). The nuclear genome is an extensive source of variable DNA regions, and variable nuclear loci are often much richer sources of information for molecular phylogenetic studies in such groups (Small et al., 2004; Whittall et al., 2006; Steele et al., 2008). Additionally, hybridization and/or incomplete lineage sorting may be common among recently diverged species; their effects can only be exposed by multi-locus approaches. For example, the tribe Verbeneae has a complicated evolutionary history of chloroplast transfer, incomplete lineage sorting, and convergent character evolution, which was only revealed by molecular phylogenetic studies using multiple loci (Yuan & Olmstead, 2008 a, b; Yuan et al., 2009 b; O'Leary et al., 2009). As genomic resources and sequencing technologies continue to be developed, the information content of the nuclear genome has become increasingly accessible to and drawn upon by phylogenetic studies; the COSII genes in Solanaceae are one example of this (Levin et al., 2009).

Yuan et al. (2009 a) developed approaches to utilize the pentatricopeptide repeat

(PPR) gene family as a source of multiple nuclear loci suitable for use in phylogenetic studies, and optimized primers to amplify and sequence several of these loci in Verbenaceae (Yuan et al., 2009 b). PPR genes encode peptides with unusually high substitution rates. There are a large number of PPR loci, which are highly divergent from one another. The shared presence of many of these loci in such distantly related groups as Brassicaceae (*Arabidopsis thaliana* (L.) Heynh.) and Poaceae (rice, maize) suggests that the present diversity of PPR genes is due to ancient duplications (Yuan et al., 2009 a, b). Yuan et al. (2009 a) screened the genomes of *A. thaliana* and rice for intron-less PPR genes with a single orthologue in each, and published a list of over 100 of these (2009 a). The loci on this list are valuable as phylogenetic tools because they can be directly sequenced and easily and unambiguously aligned, problems caused by doubtful orthology are avoided, and they can potentially be developed for use in any plant group.

We took a multi-locus approach to reconstruct a Lantaneae phylogeny, in order to test monophyly of its genera, investigate the extent to which fruit characters are homoplasious, and seek evolutionary patterns in geographic distribution within the tribe. We collected DNA sequences across a broad sample of the tribe, from three PPR genes along with the nuclear ETS region and three chloroplast loci (*trn*T-L, *rpl*32-*trn*L, and *trn*Q-*rps*16). Two of the PPR loci used in this study were amplified using primers designed by Yuan et al. (AT1G09680 and AT5G39980; 2009 b); a third (AT3G25970) was selected from the original list of those with a single orthologue in *A. thaliana* and rice (Yuan et al., 2009 a) and new primers were

designed to amplify it.

MATERIALS AND METHODS

Sampling

Taxa were chosen to broadly represent the morphological and geographical variation found in Lantaneae. All genera belonging to the tribe were sampled (*Coelocarpum*, *Aloysia*, *Acantholippia*, *Xeroaloysia*, *Phyla*, *Burroughsia*, *Nashia*, *Lippia*, *Lantana*). Forty-seven Lantaneae species were chosen as the ingroup, and seven species from related lineages were chosen as outgroups. Voucher information and Genbank accession numbers for all taxa sampled are listed in Appendix 1.

DNA extraction, amplification and sequencing

DNA was extracted from dried leaf tissue that was either collected in the field and preserved in silica gel, or sampled from herbarium specimens. Extractions were carried out following a standard CTAB method (modified from Doyle & Doyle, 1987); DNA was purified by isopropanol precipitation, and some extractions were further purified using a DNA cleanup kit (Promega Corp.).

PCRs were performed in a Perkin-Elmer thermocycler, under the following general reaction conditions: 94^{o}C for two minutes, followed by 35 cycles of 94^{o}C for 30 seconds, 50^{o}C for 30 seconds, 72^{o}C for 1.5 – 2.5 minutes, followed by 72^{o}C for ten minutes. Universal primers were used to amplify the *trn*T-L (Taberlet et al., 1991), *rpl*32-*trn*L (Shaw et al., 2007), and *trn*Q-*rps*16 (Shaw et al., 2007) regions from the

chloroplast genome. The External Transcribed Spacer (ETS) region of the nuclear 18S/26S rDNA was amplified using the 18S-IGS primer of Baldwin & Markos (1998) with a custom primer designed to amplify ETS in Lamiales (ETS-B: 5'-ATA GAG CGC GTG AGT GGT G-3'). The AT1G09680 and AT5G39980 PPR genes (hereafter referred to as PPR 11 and PPR 123, from the order in which they are listed by Yuan et al., 2009 a) were amplified using primers optimized for use in Verbenaceae by Yuan et al. (2009 b). Primers specific to the AT3G25970 region (hereafter referred to as PPR 81; Yuan et al., 2009 a) in Verbenaceae were designed following the procedure outlined by Yuan et al. (2009 a); the following primers were successfully used to amplify a fragment of the coding sequence of approximately 1.2 kb in length: PPR 81-400f (5'-AGT GCR CTT TTW GAT ATG TAY GCA AAG TG-3') and PPR 81-1630r (5'-TCR ACT GCA CAT GCR TAA TKT TCC AT-3'). All PCR products were purified by PEG precipitation.

Cycle sequencing reactions were carried out in a Perkin-Elmer thermocycler using BigDye v.3.1 (Applied Biosystems Inc.), following a standard Applied Biosystems sequencing protocol. For all loci except ETS, internal sequencing primers were used in addition to PCR primers to obtain overlapping reads across fragments (Appendix 2). Products of sequencing reactions were purified by precipitation in sodium acetate and ethanol, or by passing through Sephadex G-50 columns. An Applied Biosystems genetic analyzer was used to generate raw sequence data; the reads were then edited and assembled using Sequencher (Gene Codes Corp.).

Phylogenetic analyses

Sequences were aligned using MAFFT version 6 online (Katoh et al., 2002); alignments were then inspected and manually adjusted where necessary. Sequence alignments for the three chloroplast loci (*trn*T-L, *rpl*32-*trn*L, *trn*Q-*rps*16) were concatenated and analyzed as a single data set. Alignments for nuclear loci were treated as separate data sets. Phylogenetic reconstructions were performed individually for each data set, and for a supermatrix consisting of data from all loci (chloroplast and nuclear) in concatenation. The supermatrix was treated as consisting of a single partition.

The suitability of different models of evolution to the data was assessed using jModeltest 0.1 (Posada, 2008). The GTR + I + Γ model was selected, and applied to all analyses. Phylogenetic reconstructions for individual data sets and supermatrices were carried out using both maximum likelihood and Bayesian approaches, as implemented in GARLI (version 2.0; Zwickl, 2006) and MrBayes (version 3.1.2; Ronquist & Huelsenbeck, 2003). Shimodaira-Hasegawa (SH) tests (Shimodaira & Hasegawa, 1999) were carried out to gauge the compatibility of the results of analyses of individual loci with one another. Tree likelihood scores were calculated and SH tests performed using PAUP* v.4b10 (Swofford, 2000), with RELL optimization and 5,000 replicates under the GTR + I + Γ model.

Maximum likelihood analyses used two replicate runs, which were run with the generation threshold for termination at 20,000 generations, and termination score

threshold 0.05. Bootstrapping was carried out with 100 replicates, with the generation threshold for termination lowered to 10,000 to facilitate faster analysis (as recommended in the GARLI manual, version 0.96).

Bayesian analyses used two replicate runs, each consisting of four chains, which were run for at least one million generations, and sampled every 1,000 generations. Convergence between runs was assessed by examining standard deviations of split frequencies, and by using AWTY (Wilgenbusch et al., 2004) to plot split frequencies over different runs. Analyses which had not converged after one million generations were run until convergence diagnostics indicated they had reached stationarity; up to 50 million generations. Longer MrBayes analyses were carried out using the NSF TeraGrid via the CIPRES portal (Miller et al., 2010). When summarizing consensus trees over all runs, the first 25% of sampled trees were considered burn-in, and discarded.

Fruit evolution and biogeography

A semi-strict (combinable component) consensus tree between trees inferred from different loci was constructed using PAUP* (Swofford, 2000); relationships that were not well-supported in individual trees (bootstrap value > 80% and posterior probability > 0.9) were considered unresolved and collapsed before creating the consensus. We used Mesquite v. 2.75 (Maddison & Maddison, 2011) to score taxonomically important fruit characters and geographic distributions and to map

them onto the consensus tree, and to infer the most parsimonious character states and distributions at ancestral nodes.

RESULTS

Data collection

Complete or nearly complete sequences of each target locus were obtained for the majority of taxa included in this study. Only the sequences of the PPR 81 locus for three taxa – *Lippia rehmannii* H. Pearson, *Lantana rugosa* Thunb., *Burroughsia fastigiata* (Brandegee) Moldenke – were not available; sequences for these taxa were treated as missing data in the phylogenetic analyses from all concatenated sequences, and not included in the individual analyses of the PPR 81 locus. A few other sequences were partial for some taxa, or included short regions of missing data (DNA from herbarium specimens was occasionally of poor quality, making amplification difficult). The ETS region for *Lippia origanoides* Kunth. was amplified and sequenced from a different DNA accession (individual) from that which provided sequences for other loci; ETS could not be sequenced directly from the original accession due to a length polymorphism. The sequences from ETS and from the PPR loci contained some single nucleotide allelic differences within individuals, which were scored as polymorphisms in alignments.

The total aligned sequence data gathered were 400 bp of ETS (all taxa), 1,180 bp of PPR 11 (except *Lippia lupulina* Cham.: 761 bp; *Lippia diamantinensis* Glaz.: 854 bp; *Lantana trifolia* L.: 753 bp; *Citharexylum montevidense* (Spreng) Moldenke: regions

of missing sequence totaling 253 bp), 1,059 bp of PPR 81 (except *Dipyrena glaberrima* (Gillies & Hook.) Hook.: 914 bp, *Lippia dulcis* Trevir.: 913 bp, *Lippia javanica* (Burm.f.) Spreng: 923 bp; sequences from *Lippia rehmannii*, *Lantana rugosa* and *Burroughsia fastigiata* were excluded), 1,047 bp of PPR 123 (except *Lippia lupulina*: 773 bp). Chloroplast loci were completely amplified and sequenced for all taxa (except *trn*Q-*rps*16 of *Phyla nodiflora* (L.) Greene, for which approximately 250 bp were missing from the 3' end). Chloroplast loci varied in length, from 626-698 bp for *trn*T-L fragments, 738-1,010 bp for *rpl*32-*trn*L fragments, and 1,065-1,652 bp for *trn*Q-*rps*16 fragments, and, in combination, provided 4335 bp of aligned sequence data. After alignment and concatenation, the supermatrix of all sequence data consisted of 8,734 aligned positions.

Phylogenetic analyses

The results of phylogenetic reconstructions from individual data sets are depicted in Figs. 1.1-1.3A; Fig. 1.3B shows the results of phylogenetic analysis of the supermatrix consisting of all data in concatenation. In SH tests, individual data sets all rejected each other's best likelihood trees with P = 0.000 (Appendix 3A). The combined tree was rejected with $P < 0.05$ by the chloroplast data, PPR 81, and PPR 123, but was not rejected by ETS (P = 0.118) and PPR 11 (P = 0.09).

Well-supported clades are consistent between the maximum likelihood and Bayesian analyses for each data set; relationships that are resolved differently by maximum likelihood and Bayesian analyses receive low support. Three out of five

gene trees place *Coelocarpum* in a sister relationship with the rest of Lantaneae, with good support; conflicting topologies receive poor support in the other two gene trees. Two well-supported clades of *Aloysia* species are present in all gene trees: the *Aloysia citriodora* clade, and the *Aloysia gratissima* clade, which includes *Xeroaloysia ovatifolia* (Moldenke) Tronc. However, there is conflict between gene trees about whether these two clades together form a clade (ETS and chloroplast trees do not feature this clade; all three PPR genes do). The tree inferred from chloroplast data places *Acantholippia salsoloides* as sister to the *A. citriodora* clade, with good support, but trees from the four nuclear loci place this species in various other relationships, with varying levels of support. The tree inferred from all loci in concatenation is consistent with the chloroplast gene tree with regards to the placement of *Coelocarpum*, the two *Aloysia* clades mentioned above, and *A. salsoloides*. *Acantholippia seriphioides* (A.Gray) Moldenke is consistently reconstructed in a well-supported sister relationship with a large clade comprising all sampled species of *Lantana* and *Lippia*. This large *Lantana-Lippia* clade also contains the sampled members of *Nashia*, *Burroughsia*, and *Phyla*, as well as one species of *Aloysia* (*Aloysia barbata* (Brandegee) Moldenke).

Fruit evolution and biogeography

The consensus between well-supported topologies of individual data sets is shown in Figs. 1.4 and 1.5. Fruit characters important in separating *Lantana* from *Lippia* are mapped in Fig. 1.4, together with parsimony reconstructions of ancestral states.

Geographic ranges of members of Lantaneae sampled in this study are mapped in Fig. 1.5 along with putative ancestral distributions inferred by parsimony.

DISCUSSION

These results provide the first phylogenetic hypotheses for Lantaneae which are broadly sampled and well resolved enough to reveal the major groups within the tribe. These major clades are consistent between gene trees, despite some points of incongruence in their relationships to one another, and the relationships among taxa within them. The monophyly of Lantaneae sensu Marx et al. (2010) is confirmed. The short branch lengths within the tribe, and particularly within the *Lantana-Lippia* clade, are consistent with a recent radiation. We find strong evidence for the non-monophyly of the major genera of Lantaneae. Species of *Lantana* and *Lippia* are interspersed throughout the *Lantana-Lippia* clade, while *Nashia*, *Burroughsia* and *Phyla* are nested within it, as is a lineage of *Aloysia* species. The rest of the *Aloysia* species sampled here are allied with *Acantholippia* species and *Xeroaloysia* in a paraphyletic grade to the *Lantana-Lippia* clade. Major taxonomic revisions are required in Lantaneae; in order to achieve monophyletic genera, *Lantana* and *Lippia* must either be fragmented into many smaller genera, or lumped into a single genus. Our phylogeny reveals multiple independent shifts in the fruit characteristics historically used to diagnose genera: fleshiness, and number of pyrenes; we also show that the African members of Lantaneae represent at least two independent colonization events. The finding that the *Lantana camara* species complex is not immediately related to most other *Lantana* species is of note to tropical

conservationists investigating biological means to control invasive *Lantana camara* populations.

Analyses of individual data sets

Areas of each individual tree which did not receive good support were sometimes reconstructed differently by the different methods of phylogenetic inference used here (indicated by dashed lines in Figs. 1.1-1.3). This is probably indicative of a lack of phylogenetic signal in the data in these areas.

The contrast between the relatively slow rate of change of the chloroplast genome and the higher substitution rates of the nuclear genome is evident in the branch lengths and resolution of the trees shown in Figs. 1.1-1.3A (note that the ETS tree is drawn to half the scale of the other trees). The concatenated chloroplast matrix was several times the length (aligned positions) of any other locus sequenced, but did not provide enough information to resolve relationships in the *Lantana-Lippia* clade (although deeper nodes in Lantaneae were resolved with confidence). This is consistent with our expectations and with findings in the sister group to Lantaneae, Verbeneae (Yuan & Olmstead, 2008 a, b). Chloroplast sequence would be needed in very great quantities compared with nuclear sequence in order to provide enough information to resolve relationships at the species level in Lantaneae.

While chloroplast data could not resolve relationships between closely related species, the rapidly-evolving nuclear ETS region failed to resolve many of the

deeper nodes with confidence. In contrast, sequences from PPR genes provided the greatest resolution over the whole tree. All nuclear loci sequenced for this study had polymorphic sites in some individuals, which did not affect direct sequencing (they were coded as polymorphisms in alignments). Some allelic variation is to be expected of nuclear loci, but would require the isolation of individual alleles via cloning in order to study in more detail.

Incongruence between loci

Trees reconstructed from different individual data sets differ in their topologies, and are not compatible with one another according to SH topology tests (Appendix 3A). However, most of the differences are in relationships which are not well supported, and are thus probably best explained by insufficent information and/or noise ("soft incongruence"; Seelanan et al., 1997). Our results also include a few instances of well supported incongruence between loci with respect to the placement of 1) *Dipyrena glaberrima* among the outgroups, 2) *Acantholippia salsoloides*, 3) *Lippia rhodocnemis*/*Lippia hermannioides*, 4) *Lippia aristata*.

Conflict between different loci over the placement of *Dipyrena glaberrima* has been previously reported (Marx et al., 2010), and, whereas it lies outside the scope of this study, the question of which topology best reflects the evolutionary history of this species remains open. The position of *Acantholippia salsoloides* relative to the two *Aloysia* clades will affect how *Acantholippia* and *Aloysia* are recircumscribed, and should be resolved before revision can take place. Given the generally poor

resolution of the backbone of the *Lantana-Lippia* clade, a future study using denser sampling and additional loci would be required to study the evolution of this group in detail, and the placement of *Lippia rhodocnemis* and *Lippia aristata* would be best addressed therein. The situation in which the position of a few lineages are in strongly supported conflict between gene trees was also found in Verbeneae, Lantaneae's sister tribe (Yuan & Olmstead, 2008 a, b; Yuan et al., 2009 b; O'Leary et al., 2009), and in the problematic Neotropical palm tribe Bactridinae (Eiserhardt et al., 2011; Ludeña et al., 2011). In these examples, the question of how the conflicting lineages are related to one another, and to other lineages within their respective tribes, also has yet to be resolved.

When phylogenetic signals between gene trees are in conflict, the pattern of species divergence is sometimes best represented by the combined phylogenetic signals; i.e., the best estimate of the species tree is provided by analyzing the conflicting loci in concatenation (the total evidence approach; Kluge, 1989). This approach provides a good approximation of the species tree under circumstances when stochastic error in the finite data partitions is the cause of incongruence (Olmstead & Sweere, 1994; Gadagkar et al., 2005), and is an attractive prospect when individual data sets do not provide enough information to resolve a tree. Combined analyses have been commonly performed in phylogenetic studies over the last 10-20 years (reviewed briefly by Edwards, 2009; recent examples in Neotropical plants include studies by Sass & Specht, 2010; Eiserhardt et al., 2011). However, analysis of combined data does not reliably reflect the species tree under other circumstances, such as when

conflicting evolutionary histories underlie individual genes due to incomplete lineage sorting, hybridization, or gene duplication and extinction (Maddison, 1997; Slowinski & Page, 1999; Kubatko & Degnan, 2007). Alternative approaches, most commonly assuming that incomplete lineage sorting is the cause of incongruence, rely on coalescent theory to infer the most likely species tree from a number of individual gene trees (e.g., Liu, 2008; Kubatko et al., 2009; Heled & Drummond, 2010; for review, see Knowles, 2009; Degnan & Rosenberg, 2009).

Unfortunately, no widely accessible method yet exists to tease apart the effects of incomplete lineage sorting from hybridization and gene duplication/extinction (but see Than & Nakhleh, 2009; Choi & Hey, 2011). Any of these mechanisms could be the cause of the incongruence seen among our Lantaneae data sets. It might even be the case that there is no single bifurcating tree that adequately describes the pattern of descent of the species of Lantaneae from their common ancestor; polytomy and reticulation may be characteristic of evolutionary history in difficult, recently diversified groups such as Lantaneae.

Although the tree inferred from our combined data is fully resolved with reasonable support (Fig. 1.3B), we do not assume that it necessarily corresponds with the Lantaneae species tree. Relationships that are in conflict between loci are often resolved in favor of the larger data sets, or of the majority of data sets; i.e. minority conflicting signals from individual loci are masked in the combined analysis, even though they may provide equally valid alternative estimates of phylogeny. We feel

that it is more conservative as well as more representative of our current understanding to leave unresolved any nodes where well-supported conflict exists. We thus consider the semi-strict consensus between well-supported topologies of individual gene trees to be the best current estimate of Lantaneae phylogeny.

Taxonomic implications

Marx et al. (2010) considered the assignment of *Coelocarpum* to Lantaneae to be discordant, given the major morphological differences between this genus and the other members of the tribe, but could not place it with confidence as a lineage separate from the rest of Lantaneae. Our results open the possibility of excluding *Coelocarpum* from Lantaneae, and confirm the monophyly of the tribe, whether *Coelocarpum* is included or not. However, none of the genera of Lantaneae that are represented here by more than one species are monophyletic. *Acantholippia* contains two distinct lineages, *Aloysia* contains at least two (the relationship of the *A. citriodora* clade to the *A. gratissima* clade should be considered equivocal, pending further investigation, and denser sampling, of these groups and of *Acantholippia*). *Lantana* species form two distinct clades, a *Lantana trifolia* clade and a *Lantana camara* clade. *Lippia* species are distributed throughout the *Lantana*-*Lippia* clade, and form the background from which *Nashia inaguensis* Millsp., *Burroughsia fastigiata*, *Phyla nodiflora*, *Aloysia barbata*, and the two *Lantana* clades are derived.

Our results show that assuming correspondence between traditional taxa and evolutionary lineages is not valid in Lantaneae, and should not be accepted

uncritically in other, difficult Neotropical groups. Generic revisions in Lantaneae should proceed carefully, contingent on thorough re-evaluation of the morphological characters which correspond with evolutionary lineages. Based on our results, *Lantana* and *Lippia* will either need to be fragmented, or lumped together with the smaller genera which nest within the *Lantana-Lippia* clade. In either scenario, genera will not be easy to define morphologically. We can identify no morphological characteristics that have not undergone multiple, parallel shifts among the major clades of Lantaneae. Taxonomic revisions within the tribe will probably involve re-circumscribing genera based on combinations of traits, rather than on one to a few diagnostic characters. Densely sampled molecular phylogenetic studies are needed to investigate each clade within Lantaneae, guided by the broad phylogenetic results published here, before reliable revisions can be made.

Fruit evolution

Classifications in Lantaneae have relied largely on fruit characteristics to separate its principal genera, *Lantana* and *Lippia*. Schauer (1847), followed by Troncoso (1974), assigned species with fleshy drupes to *Lantana*, and species with dry schizocarps to *Lippia*. Under this traditional scheme, *Lippia brasiliensis* (Link) T.R.S. Silva and *Lippia macrophylla* Cham. are placed in *Lantana* section *Sarcolippia*. More recent revisions (Silva, 1999) follow Chamisso (1832) by defining *Lippia* as species with divided fruits: grouping dry schizocarps together with dipyrenous drupes under *Lippia*, and limiting *Lantana* to include only species with monopyrenous drupes. This, more recent, scheme reassigns dipyrenous fleshy-fruited species such as *L.*

brasiliensis and *L. macrophylla* to *Lippia*. Our results show that both of these classification schemes are artificial, confounded by characters that have undergone multiple independent shifts in different lineages.

There have been at least five origins of a fleshy or leathery outer layer on the fruit in Lantaneae, four of them in the *Lantana-Lippia* clade (Fig. 1.4). Fleshy fruited lineages identified in our results are: 1) the *Lantana trifolia* clade, 2) the *Lantana camara* clade, 3) *Nashia*, 4) the clade corresponding to the traditional *Lantana* section *Sarcolippia* (represented here by *L. brasiliensis* and *L. macrophylla*), and 5) *Xeroaloysia*. Whether or not the common ancestor of *Coelocarpum* + Lantaneae had fleshy fruits is difficult to infer, due to the difficulty in placing fleshy-fruited *Dipyrena* relative to dry-fruited Verbeneae and Lantaneae. If *Dipyrena* is sister to Verbeneae + Lantaneae, it is most parsimonious to reconstruct a dry-fruited ancestor for Lantaneae and hypothesize that *Coelocarpum* represents another independent derivation of fleshy fruits (as shown in Fig. 1.4). If, however, *Dipyrena* is sister to Verbeneae (rather than to Verbeneae + Lantaneae), a fleshy-fruited ancestor for Lantaneae is the more parsimonious hypothesis.

Within the *Lantana-Lippia* clade, the independent derivation of fleshy drupes from dry schizocarps has resulted in dipyrenous fruits in two lineages (the *Sarcolippia* clade and *Nashia*), and monopyrenous fruits in two lineages (the *L. camara* clade, and the *L. trifolia* clade). In the *L. trifolia* clade, *Lippia aristata* Schauer represents a subsequent shift from monopyrenous fruits to dipyrenous fruits. The pattern of shifts

in fruit type (dry to fleshy), and subdivision (two mericarps to two pyrenes or to one pyrene; one pyrene to two pyrenes) in the *Lantana-Lippia* clade reveals a complex history of fruit evolution, which has had the consequence of misleading taxonomic efforts based on fruit characteristics.

Biogeographic patterns

Major clades in the Lantaneae phylogeny are geographically heterogeneous, suggesting that migration has been an important and common element in the evolution of Lantaneae (Fig. 1.5). Old World representatives of Lantaneae can be accounted for by at least three inter-continental colonization events. *Coelocarpum*, endemic to Madagascar and Socotra, is sister to the rest of Lantaneae, and represents one lineage which has dispersed to the Old World (Marx et al., 2010; Olmstead, 2013). Similar patterns of disjunction between sister lineages (with distributions in the New World and in Madagascar) are found in other families, e.g., *Tsoala* Bosser & D'Arcy in Solanaceae (Olmstead et al., 2008), and groups within Fabaceae (Lavin et al., 2000; 2004). The legumes are particularly well-studied examples, in which large shifts in geographic range belie a high degree of niche conservatism (Lavin et al., 2004). In addition to *Coelocarpum*, two long-distance dispersals from the Neotropics to Africa are inferred within the *Lantana-Lippia* clade: a lineage within a *Lippia* clade (represented here by *L. rehmannii* and *L. javanica*), and a lineage within a *Lantana* clade (represented here by *L. viburnoides* (Forssk.) Vahl and *L. rugosa*). This frequency of colonization of Africa seems high, given that Lantaneae is a young lineage, and that long distance dispersal between Africa and

South America has been found to be relatively infrequent in other lineages (Crisp et al., 2009).

Within the Americas, a geographic shift from temperate/subtropical regions into the tropics can be seen in Lantaneae (Fig. 1.5). *Aloysia* and *Acantholippia*, which form a paraphyletic grade at the base of the tribe, are distributed primarily in arid temperate regions of South America, extending north into the Andes. *Aloysia* has an amphi-tropical distribution with a secondary radiation in Mexico and the southwestern United States, which may be the result of long-distance dispersal (Lu-Irving & Olmstead, unpublished). Members of the *Lantana-Lippia* clade, derived from the grade of *Aloysia* and *Acantholippia* species, are found throughout the tropics. This suggests a general pattern of movement into the tropics from the arid temperate or subtropical regions of South America during the evolution of Lantaneae.

Members of Lantaneae mainly occur in dry to semi-arid habitats, and rarely in wet forest environments. For example, *Acantholippia seriphioides*, which is sister to the rest of the *Lantana-Lippia* clade, inhabits arid uplands in Argentina, while the next lineage to diverge consists of low or creeping suffrutescent herbs found in dry scrub and dry to mesic disturbed habitats. The majority of the rest of the clade are woody shrubs of open and disturbed habitats, forest edges, dry hills, and Cerrado. Occurrence records for the species of Lantaneae sampled here (Fig. 1.5 inset A) reveal geographic distributions that mostly exclude the Amazon or wet coastal forests, and correspond with the distribution of seasonally dry tropical forest and

chaco biomes as outlined by Pennington et al. (2009). The lineage corresponding with *Lantana* sect. *Sarcolippia* represents a shift to wetter and more closed forest environments, but, this shift notwithstanding, the overall biogeographical pattern in Lantaneae is one of niche conservatism. Verbeneae, Lantaneae's sister clade, generally occur in dry to semi-arid habitats in temperate zones, and are not diverse in the tropics. *Aloysia* species echo this pattern, and, in the colonization of the tropics represented by the *Lantana*-*Lippia* clade, the environmental preferences of most of these species reflect those of their ancestors. This is consistent with findings that biome shifts are uncommon among plant lineages (Crisp et al., 2009; but see also Simon et al., 2009).

There is no discernible correlation between fruit type (whether fleshy or dry) and biogeographic patterns. A more densely-sampled and fully-resolved phylogenetic hypothesis might reveal such a correlation, but, to date, if there is any consistent dispersal advantage possessed by fleshy-fruited species in Lantaneae, it is not apparent. In many dry-fruited species, segments of the hairy calyx persistently enclose the mericarp, facilitating ectozoochory, just as the fleshy fruits are adapted to endozoochory. The different dispersal strategies employed among members of Lantaneae have not been broadly studied, and are likely to be diverse in such a large and varied tribe.

Future prospects

With a broadly representative sample of Lantaneae, we have identified major clades within the tribe, and revealed the extent to which they do or do not correspond with accepted genera. With the evolutionary history of lineages within Lantaneae outlined here, future systematic studies can target specific groups for the dense sampling which will probably be necessary to elucidate relationships at the species level. Particularly important areas which have yet to be resolved are: 1) the relationship of *Acantholippia salsoloides* and its (unsampled) affiliates with *Aloysia* species (this will determine how these genera are redefined); and 2) species-level relationships within the *Lantana-Lippia* clade (these will reveal the patterns of trait and biogeographic evolution amongst these many species).

In Lantaneae, as in other problematic Neotropical groups (e.g., Bactridinae, Bromelioideae, and other examples cited above) a phylogenetic estimate using molecular data is essential as a basis for reliable taxonomic revisions and speculation on evolutionary history. The difficult taxonomy of such groups hints at the complex pattern of homoplasy which may exist in morphological characters used to define taxa. Shared ancestry among lineages cannot be unambiguously inferred from morphology alone. Molecular phylogenetic studies of difficult Neotropical groups should consider evidence from multiple, independent loci. If major points of departure between gene histories exist among the species under investigation, they can be discovered by taking a multi-locus approach. It is important to evaluate possible incongruence between gene trees, to avoid providing an inappropriate

interpretation of the species tree. Lineages that are species-rich and recently radiated may be particularly prone to the incongruence among phylogenetic signal from different loci that is due to incomplete lineage sorting and hybridization.

In recently radiated lineages, nucleotide variability between taxa is an important criterion when selecting loci from which to infer phylogeny. Resolving maternal relationships at the level of species is a valuable component of phylogenetic studies, but is likely to require large quantities of sequence data from rapidly evolving DNA regions in problematic, species-rich lineages. Individual chloroplast loci are unlikely to provide sufficient phylogenetic information in such groups. If a molecular systematic study is to be undertaken in a difficult group, such as Lantaneae, a period of extensive preliminary work should first be carried out in order to develop, evaluate and select the loci to provide data for it. We expect that the potential of the nuclear genome as a resource for phylogenetic information will be largely realised over the next decade. Growing access to complete genome sequences across a range of plant species will enable a variety of multi-locus approaches to be developed and applied in divergent groups of flowering plants. With continuing advances in sequencing technologies, we predict that large-scale sequencing approaches such as RAD tagging (Miller et al., 2007; Baird et al., 2008) and large-scale alignment of entire linkage groups will replace the use of sets of well-characterized loci for phylogenetic studies.

Most taxonomic and phylogenetic studies in large and geographically widespread plant groups are subject to a tradeoff between geographic and taxonomic comprehensiveness, and a related tradeoff between breadth and depth in treating the taxa in question. Broad molecular systematic studies across large groups guide the sampling of subsequent work focused on particular lineages within those groups. Our phylogenetic estimate for Lantaneae was guided by a previous, broader study of Verbenaceae (Marx et al., 2010), and, in turn, will provide a foundation for further efforts to revise genera and to elucidate patterns of trait evolution in greater detail at the species level, as well as to better understand patterns of migration and colonization among the Neotropical flora. Over the next decade, as phylogenetic data become more easily obtainable in larger quantities, the tradeoff between breadth and depth should become less limiting. We expect that large, data-rich studies which are both broadly and densely sampled will become more common. Moving forward, collaborative efforts will be needed to thoroughly represent species-rich and geographically widespread groups in molecular phylogenetic studies at a range of taxonomic levels. The development of collaborative networks across international boundaries will be an important task for systematists to undertake over the next ten years, as we pool our efforts and expertise to advance our understanding of evolution in problematic Neotropical plant groups.

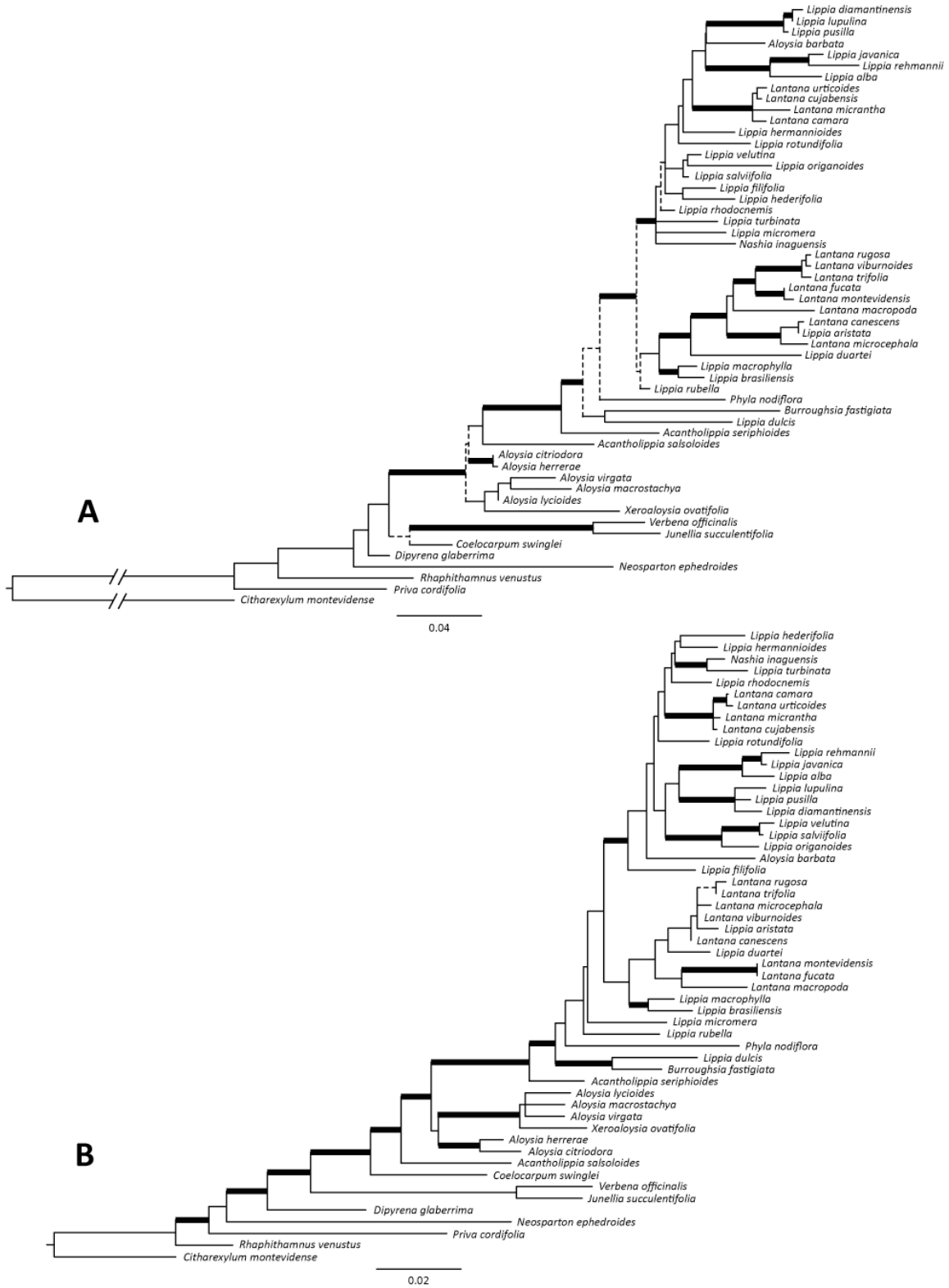


Figure 1.1 A) Maximum likelihood phylogeny inferred from DNA sequences from nuclear locus ETS (400 bp); B) PPR 11 (1,180 bp), for 47 Lantaneae species and seven outgroup species. Branches in bold are supported by greater than 80% of bootstrap replicates, and posterior probability values of higher than 0.9 in Bayesian analyses of the same data. Dashed lines indicate branches not present in the phylogeny inferred by Bayesian analysis.

Figure 1.2 A) Maximum likelihood phylogeny inferred from DNA sequences from nuclear locus PPR 81 (1,059 bp); B) PPR 123 (1,047 bp), for 44 Lantaneae species and seven outgroup species. Branches in bold are supported by greater than 80% of bootstrap replicates, and posterior probability values of higher than 0.9 in Bayesian analyses of the same data. Dashed lines indicate branches not present in the phylogeny inferred by Bayesian analysis.

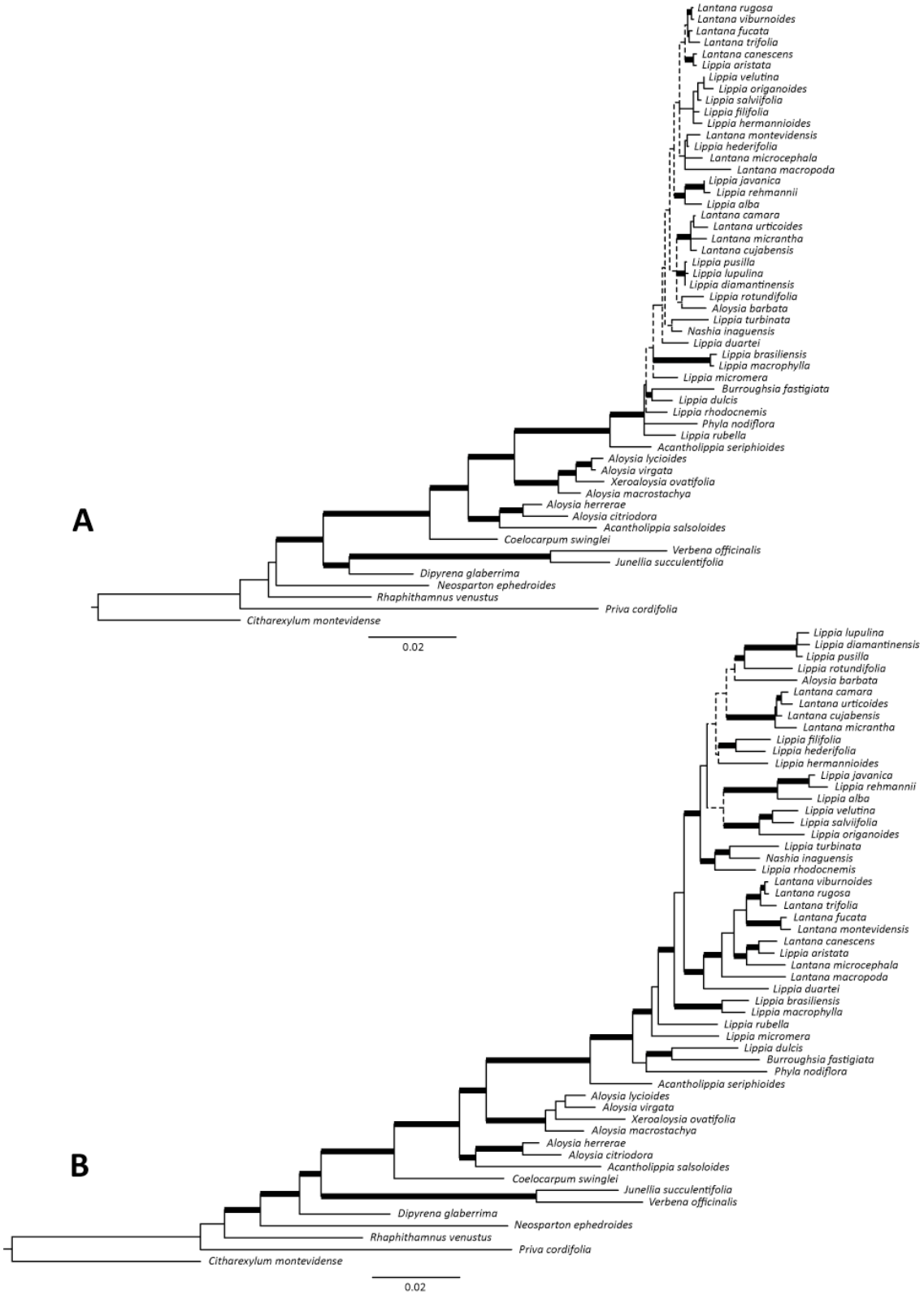


Figure 1.3 A) Maximum likelihood phylogeny inferred from DNA sequences from three chloroplast loci in combination (4,335 aligned positions); B) all DNA sequences in combination (8,734 aligned positions), for 47 Lantaneae species and seven outgroup species. Branches in bold are supported by greater than 80% of bootstrap replicates, and posterior probability values of higher than 0.9 in Bayesian analyses of the same data. Dashed lines indicate branches not present in the phylogeny inferred by Bayesian analysis.

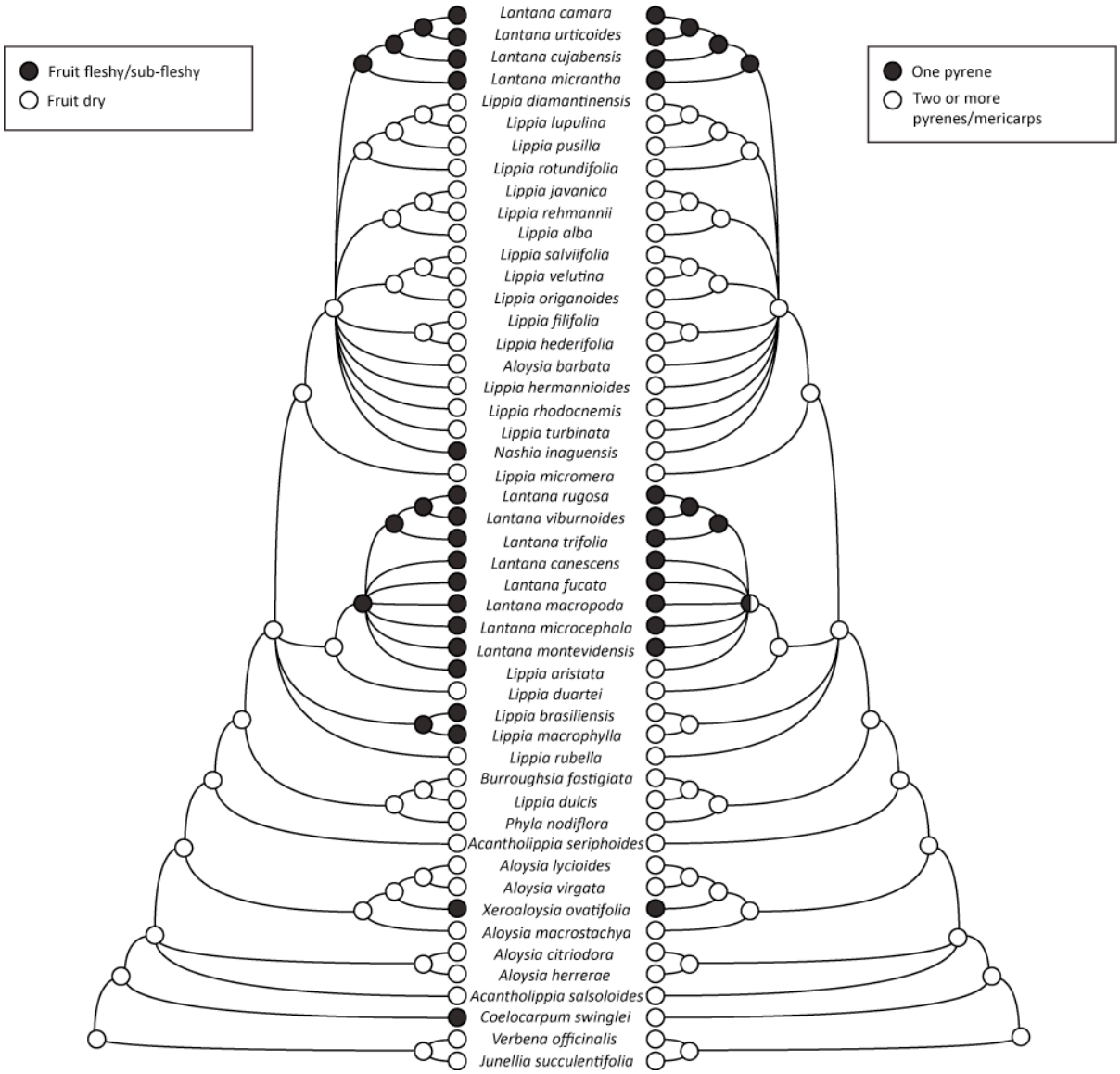


Figure 1.4 Semi-strict consensus between well supported topologies of individual phylogenies for Lantaneae, with fruit characters mapped as indicated (left: fruit type; right: number of pyrenes/mericarps). Character states at ancestral nodes are parsimony reconstructions.

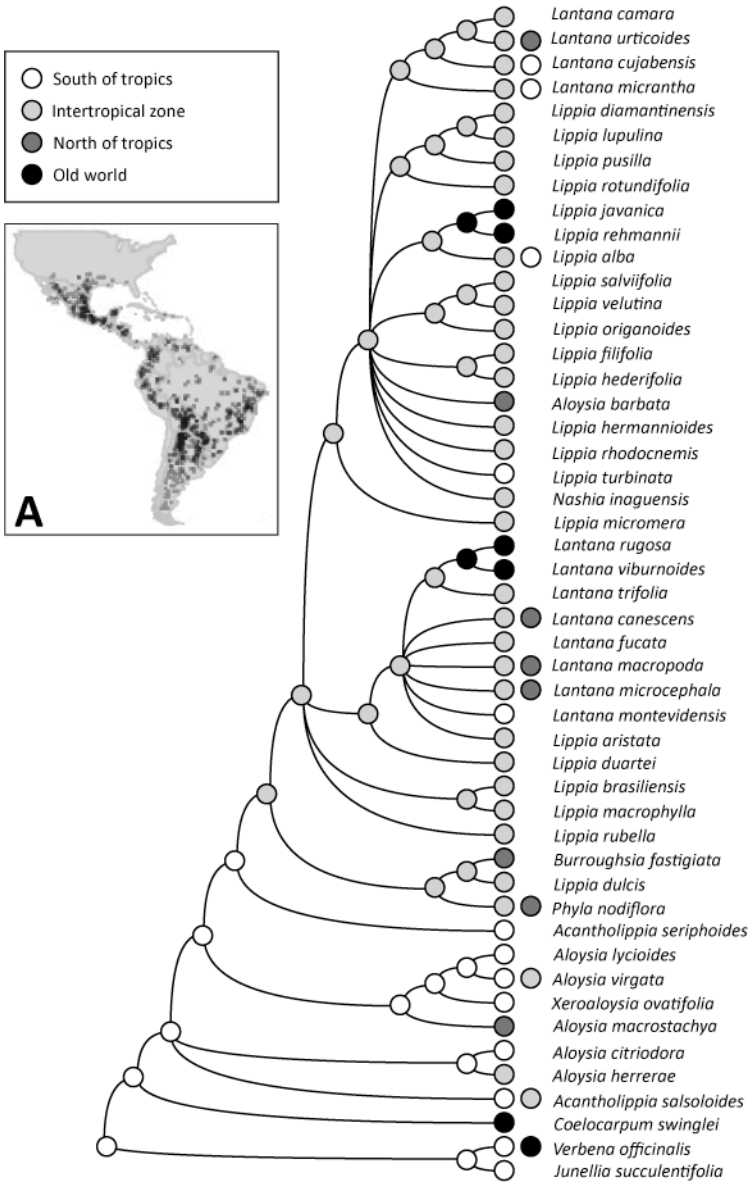


Figure 1.5 Semi-strict consensus between well supported topologies of individual phylogenies for Lantaneae, with geographic distributions mapped as indicated; species occurring in more than one coded geographic region are denoted with an additional circle. Distributions at ancestral nodes are parsimony reconstructions. Inset A. distribution of occurrence records for the species of Lantaneae included in this study (data from GBIF; records of globally invasive species and species with no georeferenced records omitted).

LITERATURE CITED

Atkins, S. 2004. Verbenaceae. In: Kadereit JW, ed. *The Families and Genera of Flowering Plants*. Berlin: Springer-Verlag, **7:** 449–468.

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.

Baldwin, B. and S. Markos. 1998. Phylogenetic utility of the External Transcribed Spacer (ETS) of 18S-26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution* **10**: 449-463.

Bárcenas, R. T., C. Yesson, and J. A. Hawkins. 2011. Molecular systematics of the Cactaceae. *Cladistics* **27**: 470-489.

Briquet, I. 1895. Verbenaceae. In: Engler A, Prantl K, eds. *Die Natürlichen Pflanzenfamilien.* Leipzig: W. Engelmann, **4**: 132-182.

Briquet, I. 1904. Verbenaceae. *Plantae Hassleriane. Bulletin de l'Herbier Boissier ser. 2* **4**: 1062–1066.

Butterworth, C. A. and R. S. Wallace. 2004. Phylogenetic studies of *Mammillaria* (Cactaceae) – insights from chloroplast sequence variation and hypothesis

testing using the parametric bootstrap. *American Journal of Botany* **91**:1086-1098.

Chamisso, A. 1832. De plantis in Expeditione Romanzoffiana observatis dicunt. Verbenaceae. *Linnaea* **7**: 105–128; 213–272.

Choi, S. C. and J. Hey. 2011. Joint inference of population assignment and demographic history. *Genetics* **189**: 561-577.

Crisp, M. D., M. T. K. Arroyo, L. G. Cook, M. A. Gandolfo, G. J. Jordan, M. S. McGlone, P. H. Weston, M. Westoby, P. Wilf, and P. Linder. 2009. Phylogenetic biome conservatism on a global scale. *Nature* **458**: 754-756.

Degnan, J. and N. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution* **24**: 332–340.

Doyle, J. J. and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19**: 11–15.

Edwards, S. V. 2009. Is a new and general theory of molecular systematics emerging? *Evolution* **63**: 1-19.

Eiserhardt, W. L., J-C. Pintaud, C. Asmussen-Lange, W. J. Hahn, R. Bernal, H. Balslev, and F. Borchsenius. 2011. Phylogeny and divergence times of Bactridinae (Arecaceae, Palmae) based on plastid and nuclear DNA sequences. *Taxon* **60**: 485-498.

Gadakgar, S. R., M. S. Rosenberg, S. Kumar. 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology* **304B**: 64–74.

Gentry, A. H. 1982. Neotropical Floristic Diversity: Phytogeographical Connections Between Central and South America, Pleistocene Climatic Fluctuations, or an Accident of the Andean Orogeny? *Annals of the Missouri Botanical Garden* **69**: 557-593.

Givnish, T.J., M. H. J. Barfuss, B. Van Ee, R. Riina, K. Schulte, R. Horres, P. A. Gonsiska, R. S. Jabaily, D. M. Crayn, J. A. C. Smith, K. Winter, G. K. Brown, T. M. Evans, B. K. Holst, H. Luther, W. Till, G. Zizka, P. E. Berry, and K. Sytsma. 2011. Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: insights from an eight-locus plastid phylogeny. *American Journal of Botany* **98**: 872-895.

Gonzalez, D., A. P. Vovides, and C. Barcenas. 2008. Phylogenetic relationships of the Neotropical genus Dioon (Cycadales, Zamiaceae) based on nuclear and chloroplast DNA sequence data. *Systematic Botany* **33**: 229-236.

Grose, S.O. and R. G. Olmstead. 2007. Evolution of a charismatic neotropical clade: molecular phylogeny of *Tabebuia* s.l., Crescentieae, and allied genera (Bignoniaceae). *Systematic Botany* **32:** 650-659.

Heled, J. and A. J. Drummond. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27**: 570–580.

Katoh, K., K. Misawa, K. Kuma, and T. Miyata. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**: 3059–3066.

Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology* **38**: 7–25.

Knowles, L. L. 2009. Estimating species trees: methods of phylogenetic analysis when there is incongruence across genes. *Systematic Biology* **58**: 463-467.

Kubatko, L. S., B. C. Carstens, and L. L. Knowles. 2009. STEM: species tree

estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* **25**: 971-973.

Kubatko, L. S. and J. H. Degnan. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology* **56**: 17-24.

Lavin, M., B. P. Schrire, G. Lewis, R. T. Pennington, A. Delgado-Salinas, M. Thulin, C. E. Hughes, A. B. Matos, and M. F. Wojciechowski. 2004. Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. *Philosophical Transactions of the Royal Society* B*: Biological Sciences* **359**: 1509-1522.

Lavin, M., M. Thulin, J-N. Labat, and R. T. Pennington. 2000. Africa, the odd man out: molecular biogeography of Dalbergioid legumes (Fabaceae) suggests otherwise. *Systematic Botany* **25**: 449-467.

Levin, R. A., A. P. Whelan, and J. S. Miller. 2009. The utility of nuclear conserved ortholog set II (COSII) genomic regions for species-level phylogenetic inference in *Lycium* (Solanaceae). *Molecular Phylogenetics and Evolution* **53**: 881-890.

Liu, L. 2008. BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* **24**: 2542-2543.

Lohmann, L. G. 2006. Untangling the phylogeny of Neotropical lianas (Bignonieae, Bignoniaceae). *American Journal of Botany* **93:** 304-318.

López-Palacios, S. 1991. Aspectos críticos de la sistemática de Verbenaceae de Venezuela. Addenda delenda et corrigenda a Verbenaceae Flora de Venezuela (1977). *Pittieria* **19**: 40-52.

Ludeña, B., N. Chabrillange, F. Aberlenc-Bertossi, H. Adam, J. W. Tregear, and J-C Pintaud. 2011. Phylogenetic utility of the nuclear genes AGAMOUS 1 and PHYTOCHROME B in palms (Arecaceae): an example within Bactridinae. *Annals of Botany* **108**: 1433-1444.

Maddison, W. 1997. Gene trees in species trees. *Systematic Biology* **46**: 523–536.

Maddison, W. P. and D. R. Maddison. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75. http://mesquiteproject.org

Marx, H.E., N. O'Leary, Y-W. Yuan, P. Lu-Irving, D. C. Tank, M. E. Múlgura, and R. G. Olmstead. 2010. A molecular phylogeny and classification of Verbenaceae. *American Journal of Botany* **97**: 1647–1663.

Miller, M. A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway*

Computing Environments Workshop (GCE 2010). New Orleans, LA: Institute of Electrical and Electronics Engineers (IEEE), 1–8.

Miller, M. R., J. P. Dunham, A. Amores, W. A. Cresko, and E. A. Johnson. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* **17**: 240-248.

Moldenke, H. N. 1940. Novelties in the Avicenniaceae and Verbenaceae. *Phytologia* **1**: 411–412.

Moldenke, H. N. 1959. *A resume of the Verbenaceae, Avicenniaceae, Stilbaceae, Symphoremaceae, and Eriocaulaceae of the world as to valid taxa, geographic distribution and synonymy.* Yonkers: published by the author.

O'Leary, N. and M. E. Múlgura. 2012. A Taxonomic Revision of the Genus *Phyla* (Verbenaceae). *Annals of the Missouri Botanical Garden* **98**: 578-596.

O'Leary, N., Y. W. Yuan, A. Chemisquy, and R. G. Olmstead. 2009. Reassignment of species of paraphyletic *Junellia* s.l. to the new genus *Mulguraea* (Verbenaceae) and new circumscription of genus *Junellia*: molecular and morphological congruence. *Systematic Botany* **34**: 777-786.

Olmstead, R.G. 2013. Phylogeny and biogeography in Solanaceae, Verbenaceae, and Bignoniaceae: a comparison of continental and intercontinental diversification patterns. *Botanical Journal of the Linnean Society* **171**: 80–102.

Olmstead, R. G., L. Bohs, H. Abdel Migid, E. Santiago-Valentin, V. F. Garcia, and S. M. Collier. 2008. A molecular phylogeny of the Solanaceae. *Taxon* **57:** 1159-1181.

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

Olmstead, R.G., M. L. Zjhra, L. G. Lohmann, S. O. Grose, and A. J. Eckert. 2009. A molecular phylogeny and classification of Bignoniaceae. *American Journal of Botany* **96:** 1731-1743.

Pennington, R. T., M. Lavin, and a. Oliveira-Filho. 2009. Woody plant diversity, evolution, and ecology in the tropics: perspectives from seasonally dry tropical forests. *Annual Review of Ecology, Evolution, and Systematics* **40**: 437-457.

Posada, D. 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution* **25**: 1253–1256.

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

Salimena, F. R. G. 2002. New synonyms and typifications in *Lippia* sect. *Rhodolippia* (Verbenaceae). *Darwiniana* **40**: 121-125.

Sanders, R. W. 1987. Taxonomic significance of chromosome observations in Caribbean species of *Lantana* (Verbenaceae). *American Journal of Botany* **74**: 914-920.

Sanders, R. W. 2001. The genera of Verbenaceae in the southeastern United States. *Harvard Papers in Botany* **5**: 303-358.

Sanders, R. W. 2006. Taxonomy of *Lantana* sect. *Lantana* (Verbenaceae): I. Correct application of *Lantana camara* and associated names. *Sida* **22**: 381-421.

Santos, I. E. M. 2002. A taxonomic revision of *Lantana* sect. *Lantana* (Verbenaceae) in the Greater Antilles. *Willdenowia* **32**: 285-301.

Sass, C. and C. D. Specht. 2010. Phylogenetic estimation of the core Bromelioids with an emphasis on the genus *Aechmea* (Bromeliaceae). *Molecular Phylogenetics and Evolution* **55**: 559-571.

Schauer, J. C. 1847. Verbenaceae. In: De Candolle AP, ed. *Prodromus.* Lehre: Cramer, **11**: 522–700.

Schulte, K., M. H. J. Barfuss, and G. Zizka. 2008. Phylogeny of Bromelioideae (Bromeliaceae) inferred from nuclear and plastid DNA loci reveals the evolution of the tank habit within the subfamily. *Molecular Phylogenetics and Evolution* **51**: 327-339.

Seelanan, T., A. Schnabel, and J. F. Wendel. 1997. Congruence and consensus in the cotton tribe (Malvaceae). Systematic Botany 22: 259-290.

Shaw, J., E. B. Lickey, E. E. Schilling, and R. L. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in Angiosperms: The Tortoise and the Hare III. *American Journal of Botany* **94**: 275-288.

Shimodaira, H. and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* **16**: 1114.

Siedo, S. J. 2008. Systematics of *Aloysia* (Verbenaceae). Doctoral thesis, University of Texas at Austin.

Silva, T. R. S. 1999. Redelimitação e revisão taxonômica do genero *Lantana* L. (Verbenaceae) no Brasil. Doctoral thesis, Universidade de São Paulo.

Silva, T. R. S. and F. R. G. Salimena. 2002. Novas combinações e novos sinônimos em *Lippia* e *Lantana* (Verbenaceae). *Darwiniana* **40**: 57-59.

Simon, M. F., G. Grether, L. P. de Queiroz, C. Skema, R. T. Pennington, and C. E. Hughes. 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences of the United States of America* 106: 20359-20364.

Slowinski, J. B. and R. D. Page. 1999. How should species phylogenies be inferred from sequence data? *Systematic Biology* **48**: 814–825.

Small, R. L., R. C. Cronn, and J. F. Wendel. 2004. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* **17**: 145-170.

Steele, P.R., M. Guisinger-Bellian, R. Linder, and R. K. Jansen. 2008. Phylogenetic utility of 141 low-copy nuclear regions in taxa at different taxonomic levels in two distantly related families of rosids. *Molecular Phylogenetics and Evolution* **48**: 1013-1026.

Swofford, D. L. 2000. PAUP*: Phylogenetic Analysis Using Parsimony (*and other

methods). Version 4b.10. Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**: 1105-1109.

Than, C. and L. Nakhleh. 2009. Species tree inference by minimizing deep coalescences. *PLoS Computational Biology* **5**: e1000501.

Tippery, N. P., C. T. Philbrick, C. P. Bove, and D. H. Les. 2011. Systematics and phylogeny of Neotropical riverweeds (Podostemaceae: Podostemoideae). *Systematic Botany* **36**: 105-118.

Torke, B. M. and B. A. Schaal. 2008. Molecular phylogenetics of the species-rich Neotropical genus *Swartzia* (Leguminosae-Papilionoideae) and related genera of the Swartzioid clade. *American Journal of Botany* **95**: 215-228.

Troncoso, N. S. 1974. Los géneros de Verbenaceas de sudamerica extra-tropical (Argentina, Chile, Bolıvia, Paraguai, Uruguai y sur de Brasil). *Darwiniana* **18**: 295–412.

Verdcourt, B. 1992. Verbenaceae. In: Polhill RM, ed. *Flora of Tropical East Africa*. Rotterdam: A.A. Balkema.

Whittall, J. B., A. Medina-Marino, E. A. Zimmer, and S. A. Hodges. 2006. Generating single-copy nuclear gene data for a recent adaptive radiation. *Molecular Phylogenetics and Evolution* **39**: 124-134.

Wilgenbusch, J. C., D. L. Warren, and D. L. Swofford. 2004. AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.

Yuan, Y-W., C. Liu, H. E. Marx, and R. G. Olmstead. 2009 a. The pentatricopeptide repeat (PPR) gene family, a tremendous resource for plant phylogenetic studies. *New Phytologist* **182**: 272–283

Yuan, Y-W., C. Liu, H. E. Marx, and R. G. Olmstead. 2009 b. An empirical demonstration of using pentatricopeptide repeat (PPR) genes as plant phylogenetic tools: Phylogeny of Verbenaceae and the *Verbena* complex. *Molecular Phylogenetics and Evolution* **54**: 23–35.

Yuan, Y-W. and R. G. Olmstead. 2008 a. A species-level phylogenetic study of the *Verbena* complex (Verbenaceae) indicates two independent intergeneric chloroplast transfers. *Molecular Phylogenetics and Evolution* **48**: 23–33.

Yuan, Y-W. and R. G. Olmstead. 2008 b. Evolution and phylogenetic utility of the PHOT gene duplicates in the *Verbena* complex (Verbenaceae): dramatic intron size variation and footprint of ancestral recombination. *American Journal of Botany* **95**: 1166–1176.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin. http://garli.googlecode.com

CHAPTER II: Resolving the genera *Aloysia* and *Acantholippia* within Lantaneae

SUMMARY

Species belonging to the genera *Aloysia* and *Acantholippia* are difficult to place within Lantaneae due to gene tree incongruence and limited sampling in previous studies. We use an expanded sample of both genera, and DNA sequence data from six loci, to reveal that *Aloysia* and *Acantholippia* species occur in five consistently inferred, well-supported lineages. The precise relationships of these clades to one another are still enigmatic, due to gene tree incongruence. However, coalescent-based species tree inference supports the inclusion of most of *Acantholippia* in an expanded *Aloysia* sensu lato, with a 4-lobed calyx as its defining feature. Five new combinations are proposed to reflect this relationship: ***Aloysia salsoloides***, ***Aloysia deserticola***, ***Aloysia trifida***, ***Aloysia riojana***, ***Aloysia tarapacana***. Geographic range shifts from subtropical South America to North America have occurred at least twice in *Aloysia*. Shifts between determinate and indeterminate inflorescence arrangement have occurred at least twice independently. The elongate, racemose inflorescence which is characteristic of most of *Aloysia* is hypothesized to be derived from a condensed, spicate or capitate inflorescence.

INTRODUCTION

Species-level systematics can be challenging when the species under consideration have a tangled evolutionary history. If morphological traits are not true to lineages, and if evolutionary processes obscure phylogenetic inference from molecular data,

then satisfactory taxonomic schemes are difficult to achieve. This study focuses on resolving the phylogenetic relationships among a group of species in which morphological parallelisms have confounded traditional classification, and which have been difficult to resolve in previous molecular systematics studies, due to gene tree incongruence. We use expanded sampling and coalescent-based phylogenetic inference from multiple, independent loci to provide a basis for the revision of the genera *Aloysia* Palau and *Acantholippia* Grisebach.

Aloysia is a genus of 29 species of shrubs and small trees in tribe Lantaneae (Verbenaceae). Members of *Aloysia* are endemic to the New World, where they are mainly found in subtropical regions. The medicinal and culinary herb *Aloysia citrodora* ("lemon verbena"; the more commonly spelled "*Aloysia citriodora*" Ortega ex. Pers. is a later homonym) is cultivated worldwide. *Acantholippia* comprises seven species of shrubs, occurring in Argentina, Chile, and Bolivia, where they inhabit dry, open environments, including the Altiplano. The monotypic genus *Xeroaloysia* (Moldenke) Tronc. is segregated from *Aloysia* based on its unique fruit and inflorescence morphology.

The generic boundaries between *Aloysia* and *Lippia*, and between *Acantholippia* and *Lippia*, are historically somewhat blurred, with authorities such as Bentham & Hooker (1876) treating *Aloysia* and *Acantholippia* as part of *Lippia*, but later authorities such as Moldenke (1959) maintaining them as separate genera. Among the major defining features of both genera is a four-lobed calyx (where the calyces of

Lippia species are bifid or truncate), with the exception of some *Aloysia* species with bifid calyces. This has been interpreted as progressive reduction in the number of calyx teeth (from five, the condition in the rest of Verbenaceae; O'Leary et al., 2012). Additionally, *Aloysia* species characteristically possess loose, open inflorescences (racemes or panicles in which the rachis is visible and the floral bracts inconspicuous; Fig. 2.2), in contrast with the tightly condensed, capitate or spicate inflorescences of *Lippia*, which often feature relatively large, foliaceous or showy floral bracts. Again, there are exceptions, with condensed inflorescences occurring in some *Aloysia* species, and with a few *Lippia* species featuring rather loose inflorescences. *Acantholippia* has *Lippia*-like condensed inflorescences, but is recognized primarily by (in addition to a 4-lobed calyx) xerophytic adaptations such as spines and/or reduced leaves (Fig. 2.2); several species of *Lippia* and *Nashia* (another segregate from *Lippia*) found in dry habitats possess similar adaptations. Previous studies have suggested that traits traditionally used to characterize genera in Lantaneae do not define monophyletic groups (Marx et al., 2010; O'Leary et al., 2012; Lu-Irving & Olmstead, 2013). However, uncertainty in previous phylogenetic reconstructions means that the pattern of evolution of many traits within Lantaneae remains unclear.

Background Information

Palau (1784) erected the genus *Aloysia* as a note appended to a translation of Linnaeus' work, describing a single species, *Aloysia citrodora* (the obscurity of this publication has caused confusion over the authorship of *Aloysia*; Armada & Barra,

1992). Subsequently, *Aloysia* was treated as a subgenus or section within *Lippia* (e.g. Schauer, 1847; Bentham & Hooker, 1876; Briquet, 1897, 1904), but has most often been accepted as an independent genus (Chamisso, 1832; Moldenke, 1959; Troncoso, 1974; Atkins, 2004). Botta (1979) treated the Argentine species of *Aloysia*, but an unpublished thesis by Siedo (2006) is the most complete treatment to date, in which 30 species and 14 varieties are recognized across the geographic range of the group. New species have since been described (e.g., Wood, 2009), but the results of recent revision call for 29 species and eight varieties in *Aloysia* (O'Leary, unpublished), broadly similar to Siedo's (2006) treatment. Three widespread species, *Aloysia gratissima*, *Aloysia scorodonioides*, and *Aloysia virgata*, are particularly variable, and are circumscribed differently according to different treatments (Siedo, 2006; O'Leary, unpublished). *Aloysia* has a mainly amphitropical distribution, with a few Andean species occuring in the tropics. It is most diverse in South America, with 22 species occurring there, and seven endemic to North America (O'Leary, unpublished). One species, *A. gratissima*, is found in both North America and South America, with a disjunction in distribution across the tropics.

Xeroaloysia was separated from *Aloysia* by Troncoso (1960) on the basis of fruit anatomy. Fruits in *Aloysia* are typically dry schizocarps separating into two one-seeded units (cluses) at maturity, similar to fruits in *Lippia*. Troncoso observed that the fruits in *Aloysia ovatifolia* Moldenke were one-seeded drupes, and proposed the new genus *Xeroaloysia* to segregate this Argentine species from *Aloysia*.

Acantholippia was established by Grisebach in 1874. *Acantholippia* species were subsequently treated as belonging to *Lippia* (Bentham & Hooker, 1876; Briquet, 1897), but Moldenke (1959) and Troncoso (1974) both followed Grisebach in recognizing the genus as independent from *Lippia* based on the presence of albumen in the seeds, subactinomorphic corollas, and xerophytic adaptations such as spines and reduced leaves. The most recent comprehensive treatment of *Acantholippia* is that of Botta (1980). *Acantholippia* and *Aloysia* are both defined as having a four-lobed calyx, in contrast with the two-lobed (or unlobed) calyx characteristic of *Lippia*. Bentham & Hooker (1876) recognized this unifying trait when they grouped *Acantholippia* together with *Aloysia* in *Lippia* sect. *Aloysia*.

The most recent and complete phylogenetic treatment of Verbenaceae (Marx et al., 2010) found *Aloysia* to be non-monophyletic: *Aloysia* species formed two clades with *Xeroaloysia* and *Acantholippia* species nesting within them. Marx et al. (2010) were concerned with reconstructing broad relationships across the family based on chloroplast sequence data, and included only a limited sample of Lantaneae. They found that many traditionally recognized tribes, and some genera, especially among Lantaneae, were not monophyletic. However, they were unable to achieve good resolution within Lantaneae. With increased sampling, Lu-Irving & Olmstead (2013) confirmed the findings of Marx et al. (2010), and revealed a third distinct lineage of *Aloysia* species, derived within a clade of *Lantana* and *Lippia* species. However, the relationships between *Aloysia*, *Acantholippia*, and the rest of Lantaneae could not be resolved with confidence, and no taxonomic revisions were made.

The relationships inferred from chloroplast data by Marx et al. (2010) and Lu-Irving & Olmstead (2013) provided the basis for a detailed study of the evolution of morphological traits in Verbenaceae (O'Leary et al., 2012). The most important morphological characters found to vary among major groups in Lantaneae were the loss of the terminal unit in inflorescence arrangement (converting a determinate compound inflorescence to an indeterminate structure, or, the transition from heterothetic pleiobotrya to homothetic pleiobotrya sensu O'Leary et al., 2012), and reduction in number of calyx teeth. Because this was based on a chloroplast reconstruction, without taking conflicting signal from nuclear loci into account, a more complete phylogenetic study might prompt reinterpretation of the evolution of these traits.

Objectives

When different genes have different histories, efforts to obtain a correct phylogeny can be misled. Whereas gene trees are often implicitly assumed to reflect the species tree, this is not always the case (Maddison, 1997). Lantaneae have been shown to be a difficult group, with a tangled evolutionary history (Lu-Irving & Olmstead, 2013); a multi-locus approach is needed to resolve the phylogenetic positions of *Aloysia* and *Acantholippia*.

Here, we present a molecular phylogenetic study of Lantaneae focusing on *Aloysia* and its related genera, *Acantholippia* and *Xeroaloysia*. Our goal is to uncover the

extent to which generic revision is needed, and to provide a basis for that revision. We use a larger and broader sampling of *Aloysia* and *Acantholippia* than has been used previously, and DNA sequence data from six loci shown to be useful in phylogenetic studies in Lantaneae (Lu-Irving & Olmstead, 2013): high-copy nuclear rDNA locus ETS, two low-copy independent loci of the nuclear PPR gene family (PPR 81 and PPR 123; Yuan et al., 2009 a, 2009 b), and three intergenic chloroplast loci (*trn*T-L, *rpl*32-*trn*L, *trn*Q-*rps*16).

MATERIALS AND METHODS

Sampling

We sampled 45 accessions (individuals; Appendix 1) representing 21 of the 29 species accepted by the most recent treatment (O'Leary, unpublished). We use several synonymized names throughout this paper; synonymy according to Siedo (2006) and O'Leary (unpublished) is detailed in Appendix 3B. Four of the seven species of *Acantholippia* are sampled. One individual of *Xeroaloysia ovatifolia* is sampled. The species of *Aloysia* sampled represent the North American, Andean, and subtropical South American distribution of this genus. Fifteen species belonging to *Lantana*, *Lippia*, *Phyla* and *Nashia* were chosen to represent the *Lantana-Lippia* clade. Seven species representing the six lineages most closely related to Lantaneae (Marx et al., 2010) were chosen as the outgroup.

DNA Extraction, Amplification and Sequencing

DNA was extracted from dried leaf tissue. The source tissue was either collected in the field and preserved in silica gel, or sampled from herbarium specimens. Extractions were carried out following a standard CTAB method (modified from Doyle & Doyle, 1987); DNA was purified by precipitation in 100% isopropanol, and some extractions were further purified using a Promega DNA clean-up kit. Amplification of target loci was carried out by PCR, using equipment, primers and reaction conditions as described by Lu-Irving & Olmstead (2013). Amplification products were purified by PEG precipitation. Cycle sequencing reactions were carried out using standard Applied Biosystems sequencing reagents and protocols for dye terminator dideoxy sequencing. The internal sequencing primers used to obtain overlapping reads for each locus were those described by Lu-Irving & Olmstead (2013). Products of sequencing reactions were purified by precipitation in sodium acetate and ethanol, or by passing through Sephadex G-50 columns. Raw sequence data was generated using Applied Biosystems PRISM Genetic Analyzers, and processed using Sequencher (Gene Codes Corp.).

Alignment and Phylogenetic Inference

Sequences were aligned using MAFFT v.6 (Katoh et al., 2002), and minor adjustments were made manually, using SeAl v.2.0a11. Data from the six target loci were assembled into six data sets: ETS, PPR 81, PPR 123, concatenated chloroplast sequences, concatenated nuclear sequences, and all data in concatenation.

To determine the most appropriate model of evolution, each data set was evaluated using jModeltest v.0.1 (Posada, 2008), under both the AIC and BIC. The partition homogeneity test (PHT; Farris et al., 1995) as implemented in PAUP* v.4b.10 (Swofford, 2000) was carried out as a gauge of incongruence between data sets. Phylogeny was then inferred from each data set using the maximum likelihood (ML) criterion as implemented in GARLI v.2.0 (Zwickl, 2006), and Bayesian analysis as implemented in MrBayes v.3.2 (Ronquist & Huelsenbeck, 2003). Data sets consisting of concatenated loci were treated as single partitions. Shimodaira-Hasegawa (SH) tests of topology (Shimodaira & Hasegawa, 1999) were carried out using PAUP* to further assess the level of incongruence between data sets. Species tree reconstructions were carried out using the coalescence-based Bayesian approach implemented in *BEAST (via BEAST v.1.7.2; Heled & Drummond, 2010).

Maximum likelihood analyses in GARLI were carried out with termination conditions at 20,000 generations, and threshold score 0.05. Each analysis was run with two replicates. Bootstrapping was carried out with 1,000 replicates, with termination after 10,000 generations. Analyses in MrBayes used two replicate runs, each consisting of four chains, sampling every 1,000 generations. Convergence between runs was assessed by observing standard deviations of split frequencies of less than 0.01, and/or by examining plots of split frequencies between runs using AWTY (Wilgenbusch et al., 2004). If convergence diagnostics did not indicate stationarity after one million generations, analyses were allowed to continue up to 50 million

generations, with periodic monitoring, and were stopped after runs had converged. Processing power for longer MrBayes analyses was provided by the NSF TeraGrid via the CIPRES portal (Miller et al., 2010). A burn-in fraction of 25% was specified when summarizing trees.

For species tree analyses, four independent loci were specified – concatenated chloroplast sequences, ETS, PPR 11, and PPR 81. A large analysis including all taxa was run, and a smaller analysis using a reduced sample of taxa (ten species) was also run, to gauge robustness of the inferred topology to the quantity of input data. Because chloroplast capture through hybridization is not uncommon in plants, and is not a mechanism taken into account by the coalescent approach, *BEAST analyses were run both with and without the chloroplast data included. The chloroplast data were treated as an organellar (haploid) locus (with half the effective population size of a bi-parentally inherited locus), and other loci were treated as autosomal. The final analysis used an HKY model for all data sets, default speciation and clock models, and the priors for mean population size and birth rate were set to gamma distributions with shape=2 (additional test analyses were performed using more complex models and various priors). Replicate runs were performed for at least 100 million generations, sampling every 10,000; runs were considered converged when ESS values were less than 200, as assessed using Tracer v.1.5 (Rambaut & Drummond, 2007).

RESULTS

Sequences gathered for each DNA accession at each locus are to be lodged in GenBank (Appendix 1). Chloroplast loci varied in size amongst individuals, from 640–700 bp for *trnT-L*, 825–1,030 bp for *rpl32–trnL*, and 1,075–1,665 bp for *trn*Q-*rps*16. After alignment, the total number of aligned positions in each data set was: 514 for ETS, 1,221 for PPR 81, 1,325 for PPR 123, 4,266 for chloroplast data combined, 3,060 for nuclear data combined, and 7,326 for all data combined. Due to difficulty in amplifying and sequencing target regions from DNA extracted from herbarium specimens, a few sequences for target loci were partial, or missing from the final data sets. The proportion of all sequences that were partial or missing was less than 6%, and, with a few exceptions, were from accessions of species that were represented by another individual (Appendix 1). The total proportion of sites scored as missing data in the final data sets was approximately 20%, including gaps. The concatenated data matrix was lodged in TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S14117).

The models of evolution implemented for each data set were the best fit under both the AIC and BIC as indicated by jModeltest v.0.1 (Posada, 2008): SYM+Γ for ETS, GTR+Γ for PPR 81, HKY+ Γ for PPR 123, TVM+Γ for chloroplast, TVM+I+Γ for nuclear, and TVM+Γ for all. Partition homogeneity tests indicated significant differences (P = 0.01) between partitions (data sets). Replicate runs over all final Bayesian-based phylogenetic analyses reached stationarity, as determined by

comparing plots of split frequences (AWTY) and examining traces and ESS values (Tracer).

Summarized results of phylogenetic analyses of individual loci and chloroplast sequences are depicted in Fig. 2.1. These trees are largely resolved with support for major clades; topologies from ML and Bayesian analyses broadly congruent, with minor disagreements over poorly-supported nodes (Appendix 3B). The trees inferred from all data are fully resolved with strong support along the backbone of the ingroup; for the concatenated data set, ML and Bayesian analyses inferred identical topologies (Fig. 2.3: concatenated sequences; Fig. 2.4: coalescent species tree). The results of all analyses identify the same major clades, but reconstruct the relationships between and within them differently (Figs. 2.1, 2.3-2.4). Topology tests indicate significant incompatibility between the results of analyses of different data sets (Table 2.1). The results of species tree reconstructions were robust to varying the number of taxa and loci analyzed, and the same topology was inferred from the data using different models and priors (results not shown).

DISCUSSION

Five major clades are consistently inferred from all subsets of the data: 1) the majority of *Aloysia* species are grouped together in a clade which also includes *Xeroaloysia* (hereafter referred to as the *A. gratissima* clade; Fig. 2.3B, Fig. 2.4B); 2) the type species of *Aloysia*, *A. citrodora*, occurs in a clade of 3 species (hereafter referred to as the *A. citrodora* clade; Fig. 2.3C, Fig. 2.4C); 3) *Aloysia catamarcensis*

and *Aloysia polystachya* are each other's closest relatives (Fig. 2.3D, Fig. 2.4D); the type species of *Acantholippia*, *A. salsoloides*, is reconstructed in a sister relationship with *Acantholippia deserticola* (Fig. 2.3E, Fig. 2.4E); we find a well-supported clade of *Lippia* and *Lantana* species, including the small genera *Phyla* and *Nashia* (Fig. 2.3A, Fig. 2.4A), consistent with the results of previous studies (Marx et al., 2010; Lu-Irving & Olmstead, 2013). Three Mexican species of *Aloysia* form a clade nested within the *Lantana-Lippia* clade (the remaining North American endemics, *A. macrostachya* and *A. wrightii*, are sister species belonging to the *A. gratissima* clade). *Acantholippia seriphioides* is sister to the *Lantana-Lippia* clade (the *Lantana-Lippia* clade is hereafter described as including *A. seriphioides*, and the three *Aloysia* species that nest within it). *Acantholippia trifida* is positioned on its own, not as part of a larger clade.

Major Lineages of Aloysia *and* Acantholippia *species*

These results provide the first sufficiently representative sample of *Aloysia* and *Acantholippia* to allow us to identify and describe the evolutionary lineages to which these species belong.

THE *ALOYSIA GRATISSIMA* CLADE—Fig. 2.3B, Fig. 2.4B. This lineage includes the majority of *Aloysia* species, including *Xeroaloysia ovatifolia*. These species all have more or less elongate, racemose inflorescences, occurring in axillary arrangements (homothetic pleiobotrya sensu O'Leary et al., 2012). Two clades within the *A. gratissima* lineage, corresponding with geographic distribution, are consistently

recovered: a North American clade (two species: *A. macrostachya* and *A. wrightii*) and an Andean clade (*A. axillaris* and *A. peruviana*, together with Peruvian accessions of *A. scorodonioides*). The Andean clade and North American clade are reconstructed as sister to one another in the analysis of concatenated data, but this relationship is not found in the analyses of individual loci (Appendix 3B); they are not sister to one another in the species tree, but support for their positions is low (Fig. 2.4). A third clade, comprising subtropical South American species as well as most sampled individuals of *A. gratissima* and *A. scorodonioides*, is consistently inferred. We can, therefore, postulate a single distributional shift into the Andes, and at least two independent dispersals from South America to North America (the North American clade, and *A. gratissima*). It is unclear whether North American distributions are due to northward migration via the Andes, or to long-distance dispersal.

Individuals identified morphologically as *A. gratissima*, *A. scorodonioides*, and *A. virgata* do not form monophyletic lineages, confirming the suspicion that the boundaries of these species are not yet well understood. Branch lengths are short throughout the *A. gratissima* clade, indicative of recent radiation. A population-level approach to sampling, data gathering, and analysis may be required to gain insight into the identities and evolutionary histories of species belonging to this lineage.

THE *ALOYSIA CITRODORA* CLADE—Fig. 2.3C, Fig. 2.4C. This lineage includes the type species of *Aloysia*, *A. citrodora*, together with *A. herrerae*. A third species, *A.*

fiebrigii, morphologically similar to *A. herrerae*, is expected to belong to this clade. Inflorescences in these species are arranged in both axillary and terminal positions (heterothetic pleiobotrya sensu O'Leary et al., 2012). The inflorescence of *A. citrodora* is racemose, and in *A. herrerae* and *A. fiebrigii* it is more condensed, or spicate. These species are found naturally in allopatric distributions from Argentina and southern Bolivia (*A. citrodora* and *A. fiebrigii*) to southern Peru (*A. herrerae*), but *A. citrodora* is cultivated worldwide.

ALOYSIA POLYSTACHYA AND *ALOYSIA CATAMARCENSIS*—Fig. 2.3D, Fig. 2.4D. These species have only axillary inflorescences (homothetic pleiobotrya sensu O'Leary et al., 2012), which are condensed and spicate. Both occur in northern Argentina. Their geographic distributions include some overlap, but they are not suspected to form hybrids (Siedo, 2006). *Aloysia polystachya* and *Acantholippia salsoloides* are the only members of Lantaneae with alternate leaves.

ACANTHOLIPPIA SALSOLOIDES AND *ACANTHOLIPPIA DESERTICOLA*—Fig. 2.3E, Fig. 2.4E. These species have both axillary and terminal inflorescences (heterothetic pleiobotrya sensu O'Leary et al., 2012), condensed into spicate heads, and all occur in semi-arid to arid habitats in subtropical South America, near the borders between Argentina, Chile, and Bolivia. This lineage is predicted to include *Acantholippia tarapacana* and *Acantholippia riojana* in addition to the two species represented in our molecular data sets. All of these species have spiny branches.

ACANTHOLIPPIA TRIFIDA—This species is reconstructed as discrete from any other lineage. It is superficially similar to members of the *A. salsoloides* clade, but lacks spines, and its condensed inflorescences are axillary only (homothetic pleiobotrya sensu O'Leary et al., 2012). *Acantholippia trifida* is endemic to north-central Chile, ranging just across the border into Argentina.

ACANTHOLIPPIA SERIPHIOIDES—This species is consistently and confidently reconstructed in a sister relationship with the *Lantana*-*Lippia* clade. It possesses xerophytic adaptations in common with other species of *Acantholippia*, such as reduced leaves, but several characters unite it morphologically with the *Lantana*-*Lippia* clade: the inflorescence is spicate-capitate, and axillary only (homothetic pleiobotrya sensu O'Leary et al., 2012), and the calyx is bilabiate (Botta, 1980). *Acantholippia seriphioides* is widespread and abundant in dry habitats in southern Argentina, and is the only member of Lantaneae to occur naturally at such high latitudes.

ALOYSIA BARBATA AND RELATIVES—This lineage comprises five species, with condensed inflorescences featuring conspicuous floral bracts, and bifid calyces, indicative of their common ancestry with the rest of the *Lantana*-*Lippia* clade. It is unclear why these species have been considered members of *Aloysia*; the first so named was transferred from *Lippia* without accompanying justification by Moldenke (1940), who then described the remainder under *Aloysia*. All five are endemic to

Mexico; two (*A. nahuire* and its segregate, *A. coalcomana*) are each known only from single collections, and may be extinct (Siedo, 2006).

Gene Tree Incongruence and Species Tree Inference

We find incongruence between loci with regards to reconstructing the relationships between major clades. The chloroplast tree identifies the *A. gratissima* clade in a sister relationship with the *Lantana-Lippia* clade, with high confidence. Also inferred from chloroplast data is a strongly-supported clade consisting of the *A. citrodora* clade, *Aloysia catamarcensis* + *A. polystachya*, *Acantholippia salsoloides* + *A. deserticola*, and *Acantholippia trifida*. This lineage is placed sister to the rest of Lantaneae (excluding *Coelocarpum*), with high confidence. None of the analyses of individual nuclear loci recover these relationships. Trees inferred from individual nuclear loci disagree on the sister group of the *Lantana-Lippia* clade, with moderate support in each case. It is variously reconstructed as *Acantholippia salsoloides* + *A. deserticola* (ETS), a monophyletic group consisting of all other major clades (PPR 81), or *A. catamarcensis* + *A. polystachya* (PPR 123).

These strongly-supported, yet conflicting topologies suggest different phylogenetic histories among loci (rather than stochastic effects arising from data sampling as the only source of incongruence). The significant differences between data sets indicated by the PHT and SH tests are consistent with this interpretation. Inconsistency between nuclear and chloroplast regions may be due to chloroplast transfer between lineages, occurring when ancestral hybridization events are

followed by introgression, resulting in fixation of the captured chloroplast (reviewed by Rieseberg & Soltis, 1991; an example in Verbeneae is documented by Yuan et al., 2008 a, 2008 b). This might have occurred among the major lineages of *Aloysia* and *Acantholippia* species, but a more complicated hypothesis of incomplete lineage sorting and/or gene duplication, perhaps in addition to hybridization, cannot be ruled out (Pamilo & Nei, 1988; Maddison, 1997).

In cases of incongruence between phylogenetic estimates from independent loci, two approaches to infer the species tree are commonly employed. Concatenation of sequences from different loci into a supermatrix, analyzed as a single data set, is one approach (the "total evidence" argument of Kluge, 1989), and may be preferred when differences among gene trees derive only from stochastic sampling effects (Olmstead & Sweere, 1994; Gadakgar et al., 2005). An alternative approach, which has become popular over the last decade, is to consider each gene tree as a data point from which a species tree may be inferred (Doyle, 1992; Maddison, 1997; Slowinski & Page, 1999). The most well-developed computational tools to do this are based on coalescent theory (reviewed by Degnan & Rosenberg, 2009), and assume that incongruence between genes is due to lineage sorting effects, as might be expected when ancestral population sizes are large and branch lengths are short (Pamilo & Nei, 1988). Coalescent-based approaches explicitly account for potentially different phylogenetic histories between loci; phylogenetic inference based on the coalescent has been shown to recover the species tree more reliably than concatenation (Edwards et al., 2007; Leaché & Rannala, 2011).

Here we have explored both a concatenation and a coalescent approach. The combined analysis of all data echoes the chloroplast tree, finding strong support for a sister relationship between the *A. gratissima* clade and the *Lantana-Lippia* clade, and strong support for a third monophyletic group, comprising the remainder of the major lineages, as sister to both, with high confidence. There is a lack of signal for any of these relationships among individual nuclear loci, and also in the combined nuclear data. Given the relatively large quantity of chloroplast data, and its strong phylogenetic signal, it seems likely that the chloroplast gene history is masking the conflicting histories of the nuclear loci in the combined analysis. In contrast, the tree inferred from all data using *BEAST strongly supports a sister relationship between the *Lantana-Lippia* clade and a monophyletic group consisting of *Aloysia* and *Acantholippia* lineages. This result is consistent with the topology of one nuclear gene tree (PPR 81), implying that the phylogenetic history of this locus is the same as the species tree. Neither analysis of all data (concatenated or coalescent) reconstructs shallower relationships between major lineages with high confidence (Fig. 2.3C-E, Fig. 2.4C-E).

Patterns of trait evolution

Consideration of morphological trait evolution in light of these phylogenetic results might yield further insight into the relationships among major clades of *Aloysia* and *Acantholippia* species. O'Leary et al. (2012) identified two traits which varied in potentially informative ways among major clades within Lantaneae: the presence or

absence of a terminal unit in the arrangement of inflorescences, resulting in either determinate or indeterminate compound structures (heterothetic vs. homothetic pleiobotrya), and the number of calyx lobes.

The homothetic pleiobotrya sensu O'Leary et al. (2012) is found in the *Lantana-Lippia* clade, the *A. gratissima* clade, *A. catamarcensis* + *A. polystachya*, and *Acantholippia trifida*. This pattern was interpreted as resulting from two parallel losses of the terminal inflorescence, based chloroplast topology and limited sampling, where one shift from heterothetic to homothetic pleiobotrya is interpreted as a synapomorphy for the *A. gratissima* clade + *Lantana-Lippia* clade (O'Leary et al., 2012). Our results, based on increased data and sampling, suggest two losses and one subsequent gain of the terminal inflorescence. This is the most parsimonious reconstruction in both analyses of all data (concatenated and coalescent).

The number of calyx teeth has traditionally been used to separate *Aloysia* and *Acantholippia* (with 4-lobed calyces) from members of the *Lantana-Lippia* clade (with bifid or truncate calyces). This was interpreted by O'Leary et al. (2012), based on a chloroplast phylogeny, as a progressive reduction in the number of calyx teeth from five in the rest of Verbenaceae, to four in Lantaneae, to two in the *Lantana-Lippia* clade. Our findings prompt re-interpretation of the evolution of this trait. Close examination of the morphology of *Acantholippia seriphioides* reveals that the calyx is bilobed, with each lobe only minutely 2-toothed (Botta, 1980). This suggests

homology with the 2-lobed calyx characterizing the rest of the *Lantana-Lippia* clade, rather than with the equally 4-fid calyces of the other species of *Acantholippia*, to which this species is unrelated. Thus, according to the species tree topology, the *Aloysia* + *Acantholippia* clade is characterized by the synapomorphy of an equally 4-fid calyx (with one exception, *A. dusenii*, representing an independent shift to a bilobed calyx).

Based on the results presented here, the condensed, spicate or capitate inflorescence found in *Acantholippia* species, *Aloysia polystachya* + *Aloysia catamarcensis*, and the *Lantana-Lippia* clade is most parsimoniously interpreted as representative of the ancestral condition in core Lantaneae (excluding *Coelocarpum*). Both of our combined data analyses suggest that the loose, racemose inflorescence characteristic of *Aloysia* as traditionally circumscribed is derived twice independently: in the *A. gratissima* clade, and in *A. citrodora*.

Taxonomic Recommendations

Aloysia and *Acantholippia* are not monophyletic, requiring revision. *Xeroaloysia ovatifolia* nests within a clade of *Aloysia* species, and can thus not be maintained in its own genus (without fragmenting *Aloysia*). Interpreting gene tree incongruence with the intent to realign generic boundaries to coincide with monophyletic groups is challenging. To produce a revision that best reflects what is known about the phylogeny of these genera, we outline and discuss three potential approaches:

1) Discount the potential problems caused by incompatible gene histories, and accept the tree inferred from concatenated loci as the best estimate of the species tree. Recognizing the three major lineages reconstructed by the chloroplast tree would require the absorption of most *Acantholippia* species into *Aloysia*, and the transfer of the majority of *Aloysia* species (those belonging to the *A. gratissima* clade) into *Xeroaloysia*. *Acantholippia seriphioides* and the Mexican *Aloysia* species nested within the *Lantana-Lippia* clade would require new names or combinations, pending a detailed revision of *Lantana* and *Lippia*. This scheme would require around 25 new combinations (not including *Acantholippia seriphioides* and the *Aloysia* species nesting within the *Lantana-Lippia* clade).

This is inadvisable because the relationships between lineages inferred on the combined tree are only compatible with the chloroplast gene tree, and it is apparent that the chloroplast genome and the nuclear regions sampled here have different phylogenetic histories. For this reason, it cannot be assumed that the tree inferred from concatenated data is a good estimate of the species tree. Furthermore, diagnostic morphological traits to discriminate the newly circumscribed *Aloysia* and *Xeroaloysia* are lacking.

2) Circumscribe genera to match only the well supported monophyletic groups consistently inferred among all independent loci. This would result in a much-reduced *Aloysia* and *Acantholippia*, while requiring the species belonging to the *Aloysia gratissima* clade to be transferred to *Xeroaloysia*, as described above. It

would require a new genus to be erected for *A. catamarcensis* + *A. polystachya* and another new genus for *Acantholippia trifida*. *Acantholippia seriphioides* and the Mexican *Aloysia* species nested within the *Lantana-Lippia* clade would require new names or combinations, pending a detailed revision of *Lantana* and *Lippia*. This scheme would require two new genera, and around 25 new combinations (not including *Acantholippia seriphioides* and the *Aloysia* species nesting within the *Lantana-Lippia* clade).

As with the previous solution, there is the problem of distinguishing the recircumscribed genera morphologically. Morphological traits simply to do not provide good indicators of evolutionary relationships amongst these species, with variation being either homoplastic or uninformative amongst the major lineages outlined above. Furthermore, it is our opinion that splitting the species of *Aloysia* and *Acantholippia* amongst five genera would be a poor representation of their close affiliation with one another. Another potential problem with this plan is that the evolutionary relationships of species not represented in our phylogenetic analyses might be other than predicted, which would result in a need for additional revisions in the future.

3) Accept the results of the *BEAST analysis as the best estimate of the species tree. According to this phylogenetic reconstruction, most *Aloysia* and *Acantholippia* species belong to a monophyletic lineage sister to the *Lantana-Lippia* clade. This prompts the absorption of *Acantholippia* and *Xeroaloysia* into *Aloysia*, leaving the

majority of names in *Aloysia* unchanged. *Acantholippia seriphioides* and the Mexican *Aloysia* species nested within the *Lantana-Lippia* clade would require new names or combinations, pending a detailed revision of *Lantana* and *Lippia*. This scheme would require five new combinations (not including *Acantholippia seriphioides* and the *Aloysia* species nesting within the *Lantana-Lippia* clade).

This is, in our opinion, the best solution. We consider the coalescent approach to provide the best estimate of the species tree, for reasons argued above. The monophyletic lineage comprising most of *Aloysia* (including *Xeroaloysia*) and *Acantholippia* reconstructed in species tree analyses is strongly supported (Fig. 2.4), and robust to varying the models, taxa, and loci analysed (results not shown). The expanded *Aloysia* can be recognized, and distinguished from the *Lantana-Lippia* clade, by the morphological synapomorphy of the 4-lobed calyx. *Acantholippia seriphioides* should be excluded from *Aloysia* s.l., as should the North American *Aloysia* species nested within the *Lantana-Lippia* clade. These species could be transferred to *Lippia*, but this would be premature because *Lippia* and its affiliated genera are not monophyletic, and will themselves need extensive revision. We defer the creation of new combinations for these species until a detailed phylogenetic study of the *Lantana-Lippia* complex is completed.

Based on the results and arguments presented here, we propose expanding the definition of *Aloysia* to include all members of Lantaneae with 4-lobed calyces. These include all the species currently described under *Aloysia* (except the North

American species with 2-lobed calyces), *Xeroaloysia ovatifolia*, and all but one of the species of *Acantholippia* (excluding *Acantholippia seriphioides*, but including the type, *A. salsoloides*). The following five new combinations and one new accepted taxon name are proposed at this time:

ALOYSIA OVATIFOLIA Moldenke, Lilloa 5: 379. 1940. *Xeroaloysia ovatifolia* (Moldenke) Troncoso, Darwiniana 12: 51. 1960.

Aloysia salsoloides (Grisebach) Lu-Irving and O'Leary **comb. nov.** *Acantholippia salsoloides* Grisebach, Pl. lorentz.: 196. 1874. *Lippia salsoloides* (Grisebach) Briquet, Nat. Pflanzenfam. 4 (3a): 152. 1897.

Aloysia deserticola (Philippi) Lu-Irving and O'Leary **comb. nov.** *Acantholippia deserticola* (Philippi) Moldenke, Lilloa 5: 370. 1940. *Lippia deserticola* Philippi, Anales Univ. Chile 59: 262. 1881.

Aloysia trifida (Gay) Lu-Irving and O'Leary **comb. nov.** *Acantholippia trifida* (Gay) Moldenke, Lilloa 5(2): 371. 1940. *Lippia trifida* Gay, Fl. Chil. 5: 29. 1849.

Aloysia riojana (Moldenke) Lu-Irving and O'Leary **comb. nov.** *Acantholippia riojana* Moldenke, Phytologia 3 (3): 106, 1949.

Aloysia tarapacana (Botta) Lu-Irving and O'Leary **comb. nov.** *Acantholippia tarapacana* Botta, Hickenia 1: 197. 1979.

Table 2.1. The results of SH test comparisons between trees inferred from different data sets.

Data set	Tree					
	ETS	**PPR 81**	**PPR 123**	**Combined cp**	**Combined nr**	**All combined**
ETS	(best)	0	0	0	0.125	0
PPR 81	0	(best)	0	0	0.190	0.044
PPR 123	0	0	(best)	0	0.006	0
Combined cp	0	0	0	(best)	0	0.008
Combined nr	0	0.015	0	0	(best)	0.073
All combined	0	0	0	0	0.002	(best)

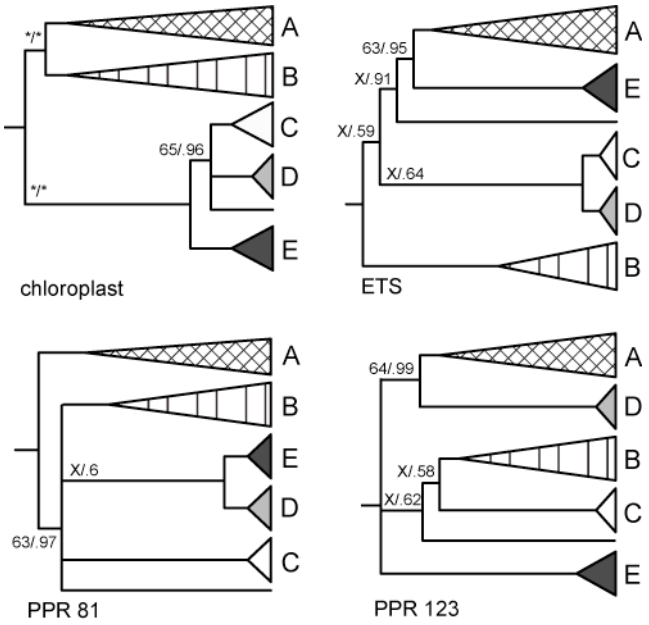


Figure 2.1 Schematic summarizing the results of phylogenetic analyses of individual loci, showing conflicting positions of major lineages. A. *Lantana-Lippia* clade. B. *Aloysia gratissima* clade. C. *Aloysia citrodora* clade. D. *Aloysia catamarcensis* + *Aloysia polystachya*. E. *Acantholippia salsoloides* + *Acantholippia deserticola*. Single tip represents *Acantholippia trifida*. Support values for the arrangement of major clades are ML bootstrap values/Bayesian posterior probabilities greater than 50%/0.50. Stars (*) denote 100% support, Xs denote bootstrap values below 50%. Phylogenetic reconstructions from individual loci are shown in detail in Appendix 3B.

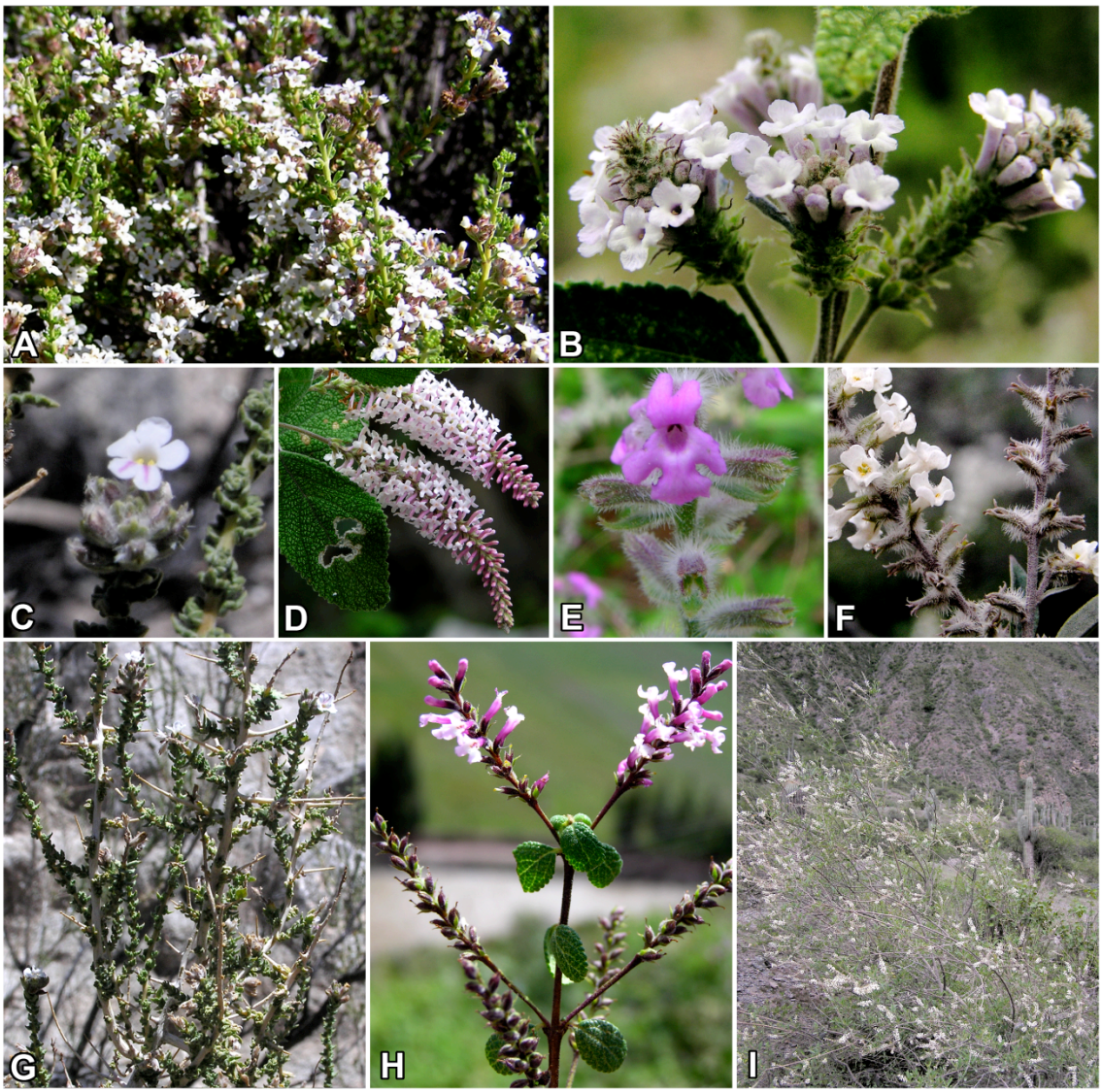


Figure 2.2 Selected species of *Aloysia* and *Acantholippia*. A. *Acantholippia seriphioides*. B. *Aloysia catamarcensis*. C. *Acantholippia salsoloides*, inflorescence. D. *Aloysia* aff. *scorodonioides*, inflorescence. E. *Aloysia macrostachya*. F. *Aloysia citrodora*, inflorescence. G. *Acantholippia salsoloides*, habit. H. *Aloysia scorodonioides* var. *hypoleuca*, inflorescence arrangement. I. *Aloysia citrodora*, habit.

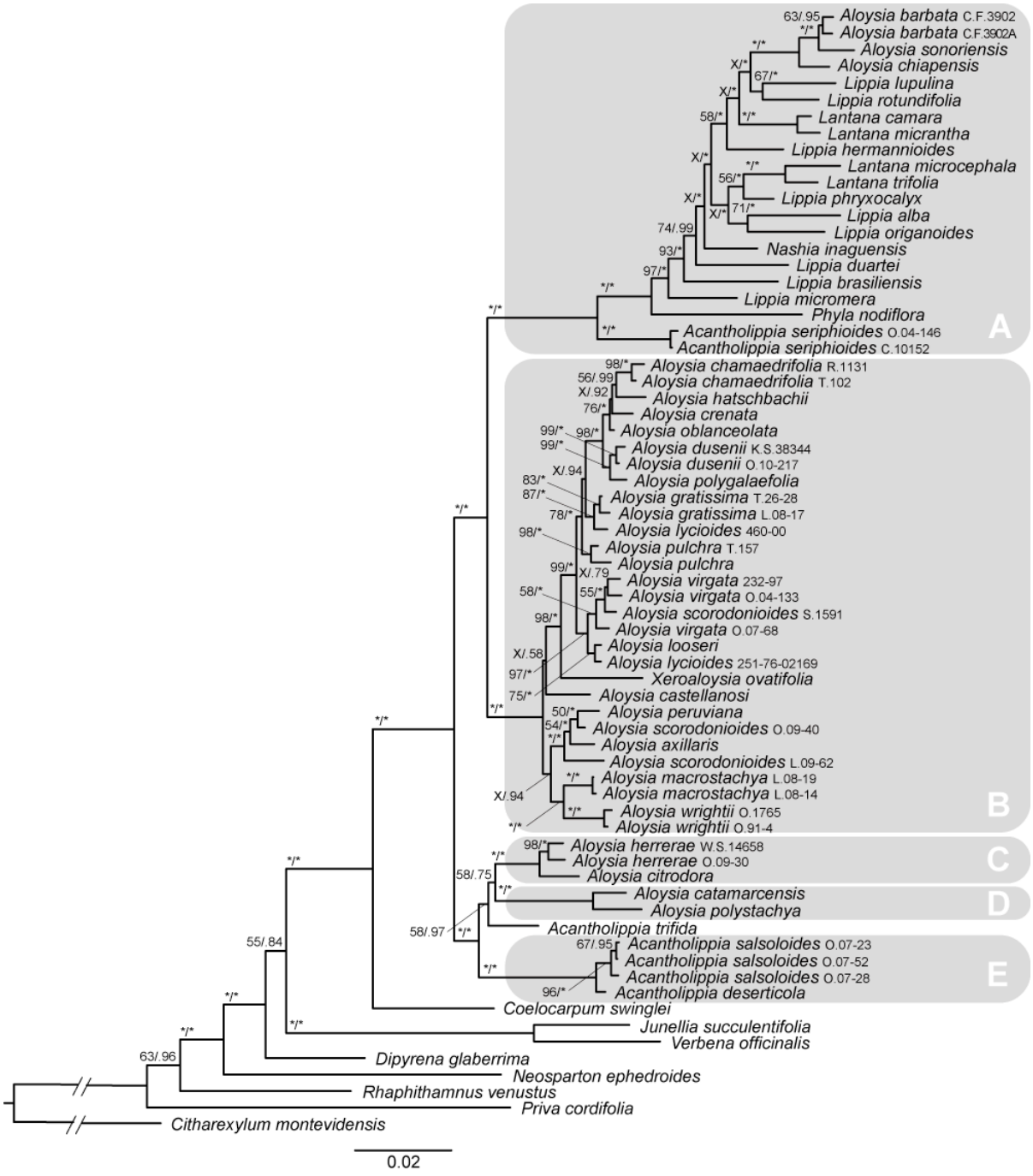


Figure 2.3 Phylogeny inferred from 7,326 aligned positions of DNA sequence data from 3 chloroplast and 3 nuclear loci in combination. Topology inferred by ML and Bayesian analyses, branch lengths inferred by Bayesian analysis. Branches are labeled with ML bootstrap values/Bayesian posterior probabilities greater than 50%/0.50. Stars (*) denote 100% support, Xs denote bootstrap values below 50%. A. *Lantana-Lippia* clade. B. *Aloysia gratissima* clade. C. *Aloysia citrodora* clade. D. *Aloysia catamarcensis* + *Aloysia polystachya*. E. *Acantholippia salsoloides* + *Acantholippia deserticola*.

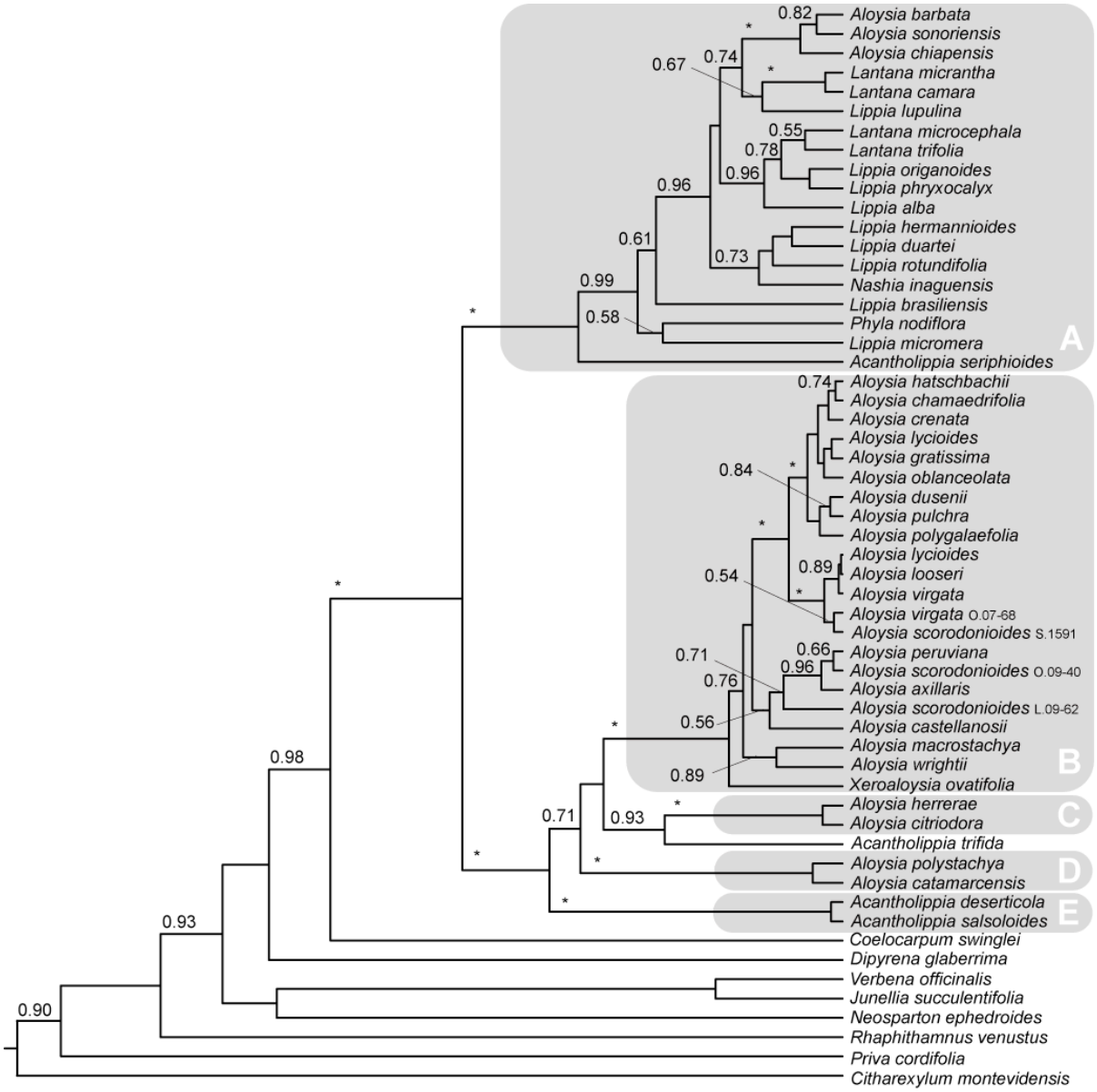


Figure 2.4 Maximum clade credibility tree inferred using *BEAST, from 3 combined chloroplast loci and 3 individual nuclear loci. Branches are labeled with posterior probabilities greater than 0.50, rounded to two decimal places; stars (*) denote posterior probabilities of 1. A. *Lantana-Lippia* clade. B. *Aloysia gratissima* clade. C. *Aloysia citrodora* clade. D. *Aloysia catamarcensis* + *Aloysia polystachya*. E. *Acantholippia salsoloides* + *Acantholippia deserticola*.

LITERATURE CITED

Armada, J. and A. Barra. 1992. On *Aloysia* Palau (Verbenaceae). *Taxon* **41**: 88–90.

Atkins, S. 2004. Verbenaceae. In: Kadereit JW, ed. *The Families and Genera of Flowering Plants*. Berlin: Springer-Verlag, **7:** 449–468.

Bentham, G. and J. D. Hooker. 1876. Verbenaceae. In: *Genera Plantarum*. London: A. Black, **2**: 1131–1160.

Botta, S. M. 1979. Las especies argentinas del género *Aloysia*. *Darwiniana* **22**: 67–108.

Botta, S. M. 1980. Las especies del género *Acantholippia* (Verbenaceae). *Darwiniana* **22**: 511–532.

Briquet, I. 1897. Verbenaceae. In: Engler A, Prantl K, eds. *Die Natürlichen Pflanzenfamilien.* Leipzig: W. Engelmann, **4**: 132–182.

Briquet, I. 1904. Verbenaceae. *Plantae Hassleriane. Bulletin de l'Herbier Boissier ser. 2* **4**: 1062–1066.

Chamisso, A. 1832. De plantis in Expeditione Romanzoffiana observatis dicunt.

Verbenaceae. *Linnaea* **7**: 105–128; 213–272.

Degnan, J. and N. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution* **24**: 332–340.

Doyle, J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany* **17**: 144–163.

Doyle, J. J. and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19**: 11–15.

Edwards, S., L. Liu, and D. Pearl. 2007. High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences* **104**: 5936–5941.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1995. Testing significance of incongruence. *Cladistics* 10: 315–319.

Gadakgar, S. R., M. S. Rosenberg, S. Kumar. 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology* **304B**: 64–74.

Heled, J. and A. J. Drummond. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27**: 570–580.

Katoh, K., K. Misawa, K. Kuma, and T. Miyata. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**: 3059–3066.

Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology* **38**: 7–25.

Leaché, A. and B. Rannala. 2011. The accuracy of species tree estimation under simulation: a comparison of methods. *Systematic Biology* **60**: 126–137.

Lu-Irving, P. and R. G. Olmstead. 2013. Investigating the evolution of Lantaneae (Verbenaceae) using multiple loci. *Botanical Journal of the Linnean Society* **171**: 103–119.

Maddison, W. 1997. Gene trees in species trees. *Systematic Biology* **46**: 523–536.

Marx, H.E., N. O'Leary, Y-W. Yuan, P. Lu-Irving, D. C. Tank, M. E. Múlgura, and R. G. Olmstead. 2010. A molecular phylogeny and classification of Verbenaceae. *American Journal of Botany* **97**: 1647–1663.

Miller, M. A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE 2010)*. New Orleans, LA: Institute of Electrical and Electronics Engineers (IEEE), 1–8.

Moldenke, H. N. 1940. Novelties in the Avicenniaceae and Verbenaceae. *Phytologia* **1**: 411–412.

Moldenke, H. N. 1959. *A resume of the Verbenaceae, Avicenniaceae, Stilbaceae, Symphoremaceae, and Eriocaulaceae of the world as to valid taxa, geographic distribution and synonymy.* Yonkers: published by the author.

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

Palau, A. 1784. In: *Parte Práctica de Botánica*. Madrid: Imprenta Real, **1**: 767–771.

Pamilo, P. and M. Nei. 1988. Relationships between gene trees and species trees. *Molecular Biology and Evolution* **5**: 568-583.

Posada, D. 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution* **25**: 1253–1256.

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

Rieseberg, L. and D. Soltis. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* **5**: 65–84.

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

Schauer, J. C. 1847. Verbenaceae. In: De Candolle AP, ed. *Prodromus.* Lehre: Cramer, **11**: 522–700.

Siedo, S. J. 2006. Systematics of *Aloysia* (Verbenaceae). Doctoral thesis, University of Texas at Austin.

Slowinski, J. B. and R. D. Page. 1999. How should species phylogenies be inferred from sequence data? *Systematic Biology* **48**: 814–825.

Swofford, D. L. 2000. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4b.10. Sinauer Associates, Sunderland, Massachusetts.

Troncoso, N. S. 1960. *Xeroaloysia*, un nuevo género argentino de Verbenáceas. *Darwiniana* **12**: 48–57.

Troncoso, N. S. 1974. Los géneros de Verbenaceas de sudamerica extra-tropical (Argentina, Chile, Bolıvia, Paraguai, Uruguai y sur de Brasil). *Darwiniana* **18**: 295–412.

Wilgenbusch, J. C., D. L. Warren, and D. L. Swofford. 2004. AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.

Wood, J. R. I. 2009. *Aloysia axillaris* (Verbenaceae), a new species, with notes on the genus in Bolivia. *Kew Bulletin* **64**: 513–523.

Yuan, Y-W. and R. G. Olmstead. 2008 a. A species-level phylogenetic study of the *Verbena* complex (Verbenaceae) indicates two independent intergeneric chloroplast transfers. *Molecular Phylogenetics and Evolution* **48**: 23–33.

Yuan, Y-W. and R. G. Olmstead. 2008 b. Evolution and phylogenetic utility of the PHOT gene duplicates in the *Verbena* complex (Verbenaceae): dramatic intron size variation and footprint of ancestral recombination. *American Journal of Botany* **95**: 1166–1176.

Yuan, Y-W., C. Liu, H. E. Marx, and R. G. Olmstead. 2009 a. The pentatricopeptide repeat (PPR) gene family, a tremendous resource for plant phylogenetic studies. *New Phytologist* **182**: 272–283

Yuan, Y-W., C. Liu, H. E. Marx, and R. G. Olmstead. 2009 b. An empirical demonstration of using pentatricopeptide repeat (PPR) genes as plant phylogenetic tools: Phylogeny of Verbenaceae and the *Verbena* complex. *Molecular Phylogenetics and Evolution* **54**: 23–35.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin. http://garli.googlecode.com

CHAPTER III: Phylogeny, fruit evolution, and diversification rates in *Lantana* and *Lippia*

SUMMARY

Fleshy fruits may be a key innovation in angiosperms, but the circumstances under which they are correlated with increased diversification rates are unclear. Evidence from detailed empirical studies is needed to elucidate the evolutionary patterns linking fleshy fruits with enhanced diversification. The *Lantana-Lippia* clade comprises approximately 230 species of woody neotropical shrubs and trees, within which fleshy fruits have been derived from dry fruits four times independently. With a well-resolved phylogeny for this lineage, we can test for a relationship between fleshy fruits and 1) higher diversification rates, and 2) larger geographic ranges. Phylogenetic reconstruction in the *Lantana-Lippia* clade is challenging; previous studies have found insufficient variability among chloroplast loci and incongruence between nuclear loci sequenced as data sources. We sequenced several low copy nuclear loci (PPR genes) to resolve the phylogeny of a representative sample of 71 species. We found that there was incongruence between phylogenetic reconstructions from different loci, as expected, but that the topologies of the coalescent ("species") tree and the tree inferred from concatenated data were broadly similar. The concatenated data were used to infer an ultrametric tree, using three internal calibration points, which reconstructed a Miocene origin for the *Lantana-Lippia* clade, and for core Lantaneae. Speciation rates in fleshy-fruited taxa were found to be significantly higher than in dry-fruited taxa, but there is no clear

pattern at a broad scale to indicate that fleshy fruits may be linked with an increased capacity for geographic range expansion.

INTRODUCTION

Across the tree of life, some branches are more species-rich than others. The contrast between remarkably diverse lineages and lineages containing only a few species is readily observed by all students of biodiversity. This pattern has been attributed to various processes, e.g., differences in diversification rates (Magallón & Sanderson, 2001), the accumulation of more species in older lineages (McPeek & Brown, 2007), ecological constraints on lineages (Rabosky, 2009), and geographic effects on diversification (Pigot et al., 2010). Factors influencing diversification rates in angiosperms have been much discussed, and ecological characteristics are often implicated (Davies & Barraclough, 2007; Crepet & Niklas, 2009; Magallón & Castillo, 2009; Vamosi & Vamosi, 2010, 2011). Traits which lead to elevated rates of diversification are referred to as key innovations (Heard & Hauser, 1995), and there are well-documented examples of these in angiosperms: from traits such as nectar spurs in *Aquilegia* (Hodges & Arnold, 1995; Hodges, 1997), to ecological strategies which affect suites of traits, such as animal pollination (Eriksson & Bremer, 1992; Dodd et al., 1999).

Fleshy fruit (endozoochory) has been suggested to be a key innovation in angiosperms (Regal, 1977; Tiffney, 1984). The dispersal of seed by animals is potentially advantageous over inanimate mechanisms: dispersal distance may be

increased, especially for larger seeds, and dispersal may be more targeted to appropriate environments for germination and growth. This can theoretically result in decreased risk of extinction (through competitive advantage and maximized geographic range size), and greater opportunity for speciation (through disperser specialization and greater potential for geographic range expansion; Howe & Smallwood, 1982; Levin et al., 2003; Cousens et al., 2008). Extensive geographic distributions in angiosperms have been positively correlated with increased diversification, though the specific role of dispersal strategy is unclear (Vamosi & Vamosi, 2010, 2011).

Early tests of the hypothesis of fleshy fruit as a key innovation found no general correlation with increased diversity at broad taxonomic scales (Herrera, 1989; Eriksson & Bremer, 1992). However, when growth form and ecological niche were considered in combination with dispersal mode, fleshy fruit was generally linked with elevated diversification rate in woody plants, and/or plants inhabiting closed forests (Eriksson & Bremer, 1991; Tiffney & Mazer, 1995; Smith, 2001; Biffin et al., 2010). This fits with observations that trees and shrubs in tropical forests are predominantly fleshy-fruited (Gentry, 1982; Howe & Smallwood, 1982; Fleming & Kress, 2011; Knörr et al., 2012). Most previous studies of the effect of fleshy fruit on diversification rates have attempted to detect general patterns over a broad taxonomic range. With the understanding that fleshy fruits are only conditionally correlated with increased diversity, finer-scale studies are needed to delineate the range of circumstances under which fleshy fruits might stimulate diversification. In this paper, we present

such a case study: the evolution of fleshy fruit and its relationship with speciation rate in the Neotropical *Lantana-Lippia* complex.

The *Lantana-Lippia* complex is a morphologically diverse lineage of approximately 230 species, comprising the majority of tribe Lantaneae, representing the largest radiation within Verbenaceae (Marx et al., 2010; Lu-Irving & Olmstead, 2013). These species are mainly aromatic shrubs, widely distributed between 45 degrees north and south latitude: most diverse in the New World tropics, but with some species in southern and eastern Africa, and a few invasive species occurring globally. This clade is one of many plant groups in which a contrast between biotic and abiotic dispersal exists; in fact, this is a traditional distinction between principal genera *Lantana* (fleshy, bird-dispersed fruits) and *Lippia* (dry fruits lacking apparent dispersal agent; Schauer, 1847). Neither *Lantana* nor *Lippia* are monophyletic, with fleshy fruits derived multiple times independently from dry fruits (Lu-Irving & Olmstead, 2013).

Core Lantaneae contains six genera after recent revision (Lu-Irving et al., submitted), and is a monophyletic lineage sister to the small Old-World genus *Coelocarpum* (five species). It comprises two sister clades: *Aloysia* sensu lato (including *Acantholippia* and *Xeroaloysia*; Lu-Irving et al., submitted), and the *Lantana-Lippia* complex. *Aloysia* s.l. (37 species) is restricted to the New World, whereas both *Lantana* and *Lippia* are represented in Africa. Within the *Lantana-Lippia* complex, *Acantholippia seriphioides* is sister to the rest, which form a well-

supported monophyletic group (the core *Lantana-Lippia* clade). Four independent shifts from dry fruits to fleshy fruits have been reconstructed in the core *Lantana-Lippia* clade (Lu-Irving & Olmstead, 2013).

In addition to *Lantana* and *Lippia*, the complex includes three genera segregated from *Lippia*: *Nashia*, *Phyla*, and *Burroughsia*. Sections have been described in *Lantana*, including sect. *Lantana* (sensu Sanders, 2006; equivalent to Chamisso's sect. *Camara*, 1832), sect. *Callioreas* (Chamisso, 1832), sect. *Sarcolippia* (Schauer, 1847; transferred to *Lippia* by Dos Santos Silva & Salimena, 2002), and sect. *Rhytidocamara* (Briquet, 1904). Troncoso (1974) provided an infrageneric classification for *Lippia*, including sections *Lippia*, *Dipterocalyx*, *Dioicolippia*, *Rhodolippia*, *Zapania*, *Pseudoaloysia*, and *Goniostachyum*. The delineation of taxa in the *Lantana-Lippia* clade has been much-revised, and remains difficult due to complex morphological patterns of parallelism and intermediacy. Phylogenetic hypotheses within the *Lantana-Lippia* clade are poorly-resolved, primarily due to limited sampling, and because sequence data analyzed to date have not been variable enough to provide resolution at the species level (Lu-Irving & Olmstead, 2013). Diversification in this group is recent, and possibly characterized by rapid divergences among large populations, resulting in incomplete lineage sorting and potential for gene flow (hybridization).

With a well-resolved phylogenetic hypothesis, the *Lantana-Lippia* complex could provide a test case in which to explore the relationship between diversification and

dispersal strategy (fleshy fruit versus dry fruit). With multiple independent origins of fleshy fruit within a well-supported monophyletic group, factors which might confound comparison of the effect of fruit type on speciation rates between groups are minimized. *Lantana* (fleshy fruits) and *Lippia* (dry fruits) comprise roughly equal numbers of species (approximately 100 in *Lantana*, and 120 in *Lippia*). All are woody trees, shrubs, or sub-shrubs (the genus *Phyla*, comprising five species, is a dry-fruited lineage of herbaceous perennials). All share the same general habitat preference: open, dry environments with frequent disturbance (such as seasonally dry forests and tree savanna), with the exception of some fleshy-fruited species found in moist forests (*Lantana* sect. *Sarcolippia*). *Lantana* and *Lippia* are both widely distributed within the same latitudinal limits, and both are represented in the Old World. In a phylogenetic context, this might provide a basis for investigation into the relationship between fruit type and the evolution of geographic range.

To resolve the poorly-understood relationships among the members of the *Lantana-Lippia* complex, we screened multiple, independent nuclear loci for variability in this group, and sequenced seven of the most informative loci across a broad, representative sample. We aimed to produce a good estimate of the phylogenetic history of this lineage in order to investigate whether the evolution of fleshy fruit is correlated with increased speciation rates. Our goal was to provide an empirical contribution toward the elucidation of patterns of diversification in flowering plants. Additionally, resolving a phylogeny for the *Lantana-Lippia* complex will provide an essential foundation for the revision of generic limits in Lantaneae.

MATERIALS AND METHODS

Data collection

Seventy-one species belonging to the *Lantana-Lippia* complex were chosen to represent the taxonomic and geographic diversity the group. Several widespread, variable species were represented by 2-3 accessions, to provide indicators of intraspecific sequence variation, and to test monophyly of these species. A total of 78 accessions formed the ingroup, and three species of *Aloysia* were used as the outgroup (Appendix 1).

Leaf tissue was collected from dried specimens or from living plants; DNA was extracted following a standard modified CTAB protocol (Doyle & Doyle, 1987), and purified by isopropanol precipitation. All PCR and sequencing reactions were carried out according to standard protocols, as described by Lu-Irving and Olmstead (2013).

Preliminary evaluations of chloroplast sequences showed insufficient variation in these data to resolve phylogenetic relationships in the *Lantana-Lippia* complex, consistent with previous studies (Appendix 3C; Marx et al., 2010; Lu-Irving & Olmstead, 2013), so we focused on gathering nuclear sequence data. Ten nuclear loci were screened for variability among *Lantana* and *Lippia* species: each locus was amplified and sequenced in four representatives of the *Lantana-Lippia* clade (*Lantana trifolia*, *Lantana depressa*, *Lantana ferreyrae*, *Lippia dulcis*). Pairwise distances between each representative species were then measured (Appendix 3C), and the seven most variable loci on average were selected to provide data for this

study: ETS, ITS, PPR loci 11, 81, 90, 97, 123 (PPR loci 24 and 47 were less variable, and the PHOT II intronic region, though informative, could not be direct-sequenced due to allelic variation in length). Primers used to amplify and sequence ITS were universal primers (ITS 4 and ITS 5; White et al., 1990), a custom forward primer was substituted in a few cases in which universal primers amplified fungal ITS sequences; ETS primers were those described by Lu-Irving & Olmstead (2013). The PPR loci were amplified and sequenced using previously published primers (Yuan et al., 2009 b; Lu-Irving & Olmstead, 2013), and primers developed to target additional loci, following the general procedure outlined by Yuan et al. (2009 b; B. Meersman & A. O'Brien, unpublished data). Sequences of primers used in this study are listed in Appendix 2.

Phylogenetic analyses

Sequence data from each locus was aligned using MAFFT v.7 (Katoh & Standley, 2013), with minor manual adjustments, and assembled into individual data sets. Model testing for alignments representing seven individual loci was conducted using 24 models of nucleotide evolution, as implemented in jModeltest v.2.31 (Darriba et al., 2012). Each locus was analyzed separately, and an analysis of the concatenated data from all seven loci was also performed. These analyses were carried out using MrBayes v.3.2.1 on XSEDE via the CIPRES Science Gateway (Ronquist & Huelsenbeck, 2003; Miller et al., 2010). Data from individual loci were analyzed using the model indicated as the best fit under the BIC criterion as implemented in jModeltest v.2.3.1. The concatenated data were partitioned into individual loci

(character sets), with substitution models specified respectively. Each analysis consisted of two runs of four chains, run for 10 million generations. Convergence was assessed by examining the standard deviations of split frequencies between runs, and by using AWTY to plot comparisons of split frequences between runs (Nylander et al., 2008). A burn-in fraction of 25% was specified when summarizing trees.

A species tree from the combined data was inferred using *BEAST (Heled & Drummond, 2010) as implemented in BEAST v. 1.7.5 (Drummond et al., 2012), with each of the seven loci treated as independent, with unlinked substitution, clock and tree model estimates. The tree was inferred under a Yule speciation model with piecewise constant population sizes, using a strict clock; an HKY model of nucleotide substitution without rate variation between sites was specified for all loci (specifying more parameter-rich tree, clock and nucleotide models resulted in parameter estimates failing to converge). Clock rates were estimated from that of PPR 81, which was assigned a starting value of 1.0; an exponential prior distribution with mean 10 was specified for clock rates. The Yule speciation rate and population size parameters were assigned gamma-distributed priors with shape 2. The MCMC was run for 500 million generations; convergence was assessed by examining logged states using Tracer v.1.5 (Rambaut & Drummond, 2009). A burn-in fraction of 25% was specified when summarizing trees.

Diversification rates

Trait-dependent speciation rates were estimated using the maximum likelihood and MCMC approaches implemented in BiSSE (Maddison et al., 2007; FitzJohn, 2012). Because there is no reliable fossil record for Lantaneae (apart from pollen in Holocene deposits; Dupont et al., 2008), other ways of calibrating node ages were explored in order to infer an ultrametric tree for diversification rate estimation.

Chloroplast data from Marx et al. (2010) were used to infer node ages within Verbenaceae, using two fossil calibration points: *Petrea* from 37-34 ma (MacGinitie, 1953; used to specify the crown node age for the family), and *Verbena* from 10-5 ma (Farlow et al., 2001; used to specify a crown node age for the *Verbena*-*Glandularia* clade). Because this data set included species belonging to other families as outgroups, secondary calibration using divergence dates among asterid lineages (Bremer et al., 2004) was possible. In accordance with the findings of Bremer et al. (2004), we assigned a date of 75 ma to the stem node of the Scrophulariaceae. A time-calibrated tree for the *Lantana*-*Lippia* clade, using the sequence data gathered as part of this study, was then inferred.

To infer an ultrametric tree for the *Lantana*-*Lippia* complex, three internal calibration points were used. Two were secondary calibrations from the Verbenaceae dated tree: the crown node age of the core *Lippia*-*Lantana* clade (not including *Acantholippia seriphioides*), and the crown node age of *Phyla*. The third calibration was a maximum crown node age of 8-4 ma for the earliest-diverging Cerrado

endemic lineage within the *Lantana*-*Lippia* clade (comprising *Lippia hederifolia*, *Lippia filifolia*, and *Lippia florida*). This corresponds with the earliest origin of grassland biomes in South America (Simon et al., 2009), following the reasoning that the Cerrado environmental niche could not have existed before the origin of savanna.

Time-calibrated trees were inferred using BEAST v.1.7.5 (Drummond et al., 2012), specifying Yule speciation models, uncorrelated lognormal relaxed clock models, and nucleotide substitution models as indicated by model testing. The mean clock rate was estimated using a uniform prior. Analyses were run for 100 million generations; convergence was assessed using Tracer v.1.5 (Rambaut & Drummond, 2009). A burn-in fraction of 25% was specified when summarizing trees.

The BiSSE analysis was performed using the Diversitree package in R v.3.0 (FitzJohn, 2012). Each species was scored for fruit type (dry or fleshy); because sampling was not complete, the sampled proportion of species with each fruit type was specified (dry: 0.3, fleshy: 0.28), assuming random sampling. Maximum likelihood estimates of trait-dependent speciation rates were obtained, as were posterior distributions of speciation rates estimated via an MCMC process.

RESULTS

Sequence data were collected for 94% of cells in the data matrix (81 taxa by seven loci; Table 3.1); approximately 11% of states in the final (combined) analyses were

scored as missing (including gaps). Sequences gathered were deposited in GenBank (Appendix 1).

Alignment lengths and details of models inferred for each alignment are summarized in Table 3.1. Phylogenetic analyses conducted using MrBayes reached convergence within 10 million generations, as indicated by distributions of split frequences between replicate runs. Gene trees for individual loci were largely well-resolved at the level of major clades; these loci were sufficiently informative to infer phylogenetic history at this level in Lantaneae (Appendix 3C). There were some well-supported differences between individual gene trees, possibly indicating that the sampled loci have different phylogenetic histories (some or all gene trees are not representative of the species tree; Maddison, 1997).

The analysis of the concatenated matrix of all data converged within 10 million generations, and resulted in a fully-resolved tree with the exception of the first node within the core *Lippia-Lantana* clade (Fig. 3.1). Most of the topology was supported by high posterior probability values. The species tree inferred from all data using the coalescent-based approach implemented in *BEAST was fully resolved, with high posterior probabilities for major clades, but less confidence in lower-order branches (Fig. 3.2). Almost all parameter estimates had converged by 250 million generations; three with ESS values below 200 after 500 million generations showed more-or-less flat traces, and higher ESS values after increasing the burn-in fraction. The species tree and tree inferred from concatenated data had similar topologies, both

reconstructing the same major clades. The most noticeable difference was that the tree from concatenated data placed some species in grades, whereas the coalescent tree grouped them in monophyletic lineages; e.g., positions of *Lippia rubella*, *Lippia micromera*, *Lippia lasiocalycina* and the Sarcolippia clade, relative to the Callioreas clade.

Analyses of divergence timing in BEAST converged on parameter and tree estimates within 50 million generations, except for a few parameters in the analysis of chloroplast data across Verbenaceae, for which 100 million generations were run, and a burn-in fraction of 35% was specified, in order to achieve flat traces and ESS values above 200. The trees inferred were fully resolved, and congruent with the topologies inferred from other analyses (Figs. 3.1-3.2; Marx et al., 2010). The analysis of Verbenaceae reconstructed the crown node of Lantaneae (including *Coelocarpum*) at approximately 19 (±4) ma, and the crown node age of the core *Lantana-Lippia* clade (excluding *Acantholippia seriphioides*) at approximately 9 (±4) ma (Appendix 3C). The analysis of the *Lantana-Lippia* clade resulted in a root height estimate of 14 (±4) ma for Lantaneae, consistent with the estimate from analysis of Verbenaceae (Fig. 3.3). This sample did not include *Coelocarpum*, so a younger estimate of the root node was to be expected. If the crown node age of the core *Lantana-Lippia* clade is specified to be 9 ma (the estimate from the time-calibrated tree for Verbenaceae), the rate of ITS evolution estimated from the maximum likelihood branch lengths of the ITS gene tree is 0.007 (±0.002) substitutions per site per million years. This is within the range of expectation based on published rates of

ITS evolution in angiosperms, although higher than the average rate (Kay et al., 2006).

Using the time-calibrated tree inferred for the *Lantana-Lippia* clade, the analysis of trait-dependent speciation rate found higher rates associated with fleshy fruit compared with dry fruit: $\lambda = 1.12$ for fleshy fruit, and $\lambda = 0.39$ for dry fruit (maximum likelihood estimates). The log-likelihood of these estimates was significantly higher than that of an equal-rates model, according to a chi-squared test (Table 3.2). The posterior distributions of speciation rate estimates within the 95% confidence interval from MCMC analysis (1,000 steps) are depicted in Fig. 3.4.

DISCUSSION

The nuclear loci sequenced here were successful in resolving phylogenetic relationships among the species of the *Lantana-Lippia* clade. The rDNA spacers (ITS and ETS) were among the most variable loci tested to provide data for this study, supporting their ubiquitous use in species-level phylogenetic inference. The PPR loci provided useful additional independent data sources, fulfilling their promise as phylogenetic tools as predicted by Yuan et al. (2009 a, b). We expect that phylogenetic studies which rely on targeting specific loci as data sources will continue to find the rDNA loci useful, but the value of many available PPR loci as potential sources of data will be especially valuable for species tree estimation.

Some differences in phylogenetic reconstruction from different loci were observed (Appendix 3C), which was to be expected in a difficult, recent radiation such as the

Lantana-Lippia complex. In some cases, these were well-supported differences, which may reflect different phylogenetic histories among loci, owing to the effects of lineage sorting, hybridization, and/or gene birth and death. Whereas the tree topology inferred using a total evidence approach (concatenating all sequence data) is supported by high confidence values, the species tree inferred using coalescent methods, which assume that incongruence between data sets is due to lineage sorting effects, has low posterior probabilities for many nodes. We interpret this as indicative of uncertainty in the branching order of the true phylogeny (species tree), in which diversification occurred rapidly.

We consider the trees inferred from combined data from the seven loci sequenced here to be generally representative of the phylogenetic history of the *Lantana-Lippia* clade, with the caveat that some nodes should be considered equivocal despite receiving high support in the concatenated tree. Until methods are developed that permit the effects of hybridization to be teased apart from lineage sorting in phylogenetic inference, and systems to infer reticulations are more robust, tested, and widely used, phylogenetic estimates in groups such as the *Lantana-Lippia* complex should be interpreted with full acknowledgment of potential uncertainty.

Despite the uncertainty described above, the same major lineages within the *Lantana-Lippia* complex are consistently and confidently inferred, in this and in previous studies (Lu-Irving & Olmstead, 2013). We consider these to be good monophyletic groups: 1) a clade corresponding with the genus *Phyla* sensu O'Leary

& Múlgura (2012), 2) a clade corresponding with *Lantana* section *Sarcolippia* (sensu Schauer, 1847), 3) a clade comprising *Lantana* sections *Rhytidocamara* (sensu Briquet, 1904) and *Callioreas* (sensu Chamisso, 1832), and 4) a clade corresponding with *Lantana* sect. *Lantana* sensu Sanders (2006); hereafter referred to as the *Camara* clade (from Chamisso, 1832). These clades are derived from within a background of *Lippia* species, which are variously reconstructed relative to major clades depending on the data source and method of inference. With the expanded sampling in this study, relative to previous work, additional well-supported clades of *Lippia* species are revealed: two lineages of Cerrado endemics (*L. hederifolia*, *L. filifolia*, *L. florida*; and *L. lupulina*, *L. diamantinensis*, *L. pusilla*), and a lineage corresponding with *Lippia* section *Goniostachyum*. No other *Lippia* sections are monophyletic groups. Other clades of *Lippia* species contain mixtures of species from different sections.

Diversification rates: fleshy fruit and biogeography

Our results suggest that the core Lantaneae (excluding *Coelocarpum*) originally radiated between 10 and 18 ma, with much of the diversification in the *Lantana*-*Lippia* clade taking place within the last eight million years. The error bars on node ages are large, however (± approximately 4 ma), and the lack of precision in divergence time estimation might be a reflection of the topological conflicts between loci. Dated species tree approaches are now possible (Drummond et al., 2012) and might be explored to provide another perspective on divergence dating in Lantaneae. Additional calibration points might also result in better-informed

reconstructions, but the fossil record of Verbenaceae is poor, and that of Lantaneae is almost non-existent. For analysis of trait-dependent diversification rate, however, the relative timing of divergences is the critical factor; precision of absolute dates is less important.

We found a significant increase in speciation rate associated with the evolution of fleshy fruits in the *Lantana-Lippia* clade, consistent with previous studies which have found a correlation between diversity and fleshy fruit in lineages of woody perennials (Eriksson & Bremer, 1991; Tiffney & Mazer, 1995; Biffin et al., 2010). One explanation for this pattern might be the selective advantage of larger seeds and more effective dispersal to openings (Eriksson et al., 2000; Bolmgren & Eriksson, 2005). However, most species of the *Lantana-Lippia* clade are not inhabitants of closed forest environments, preferring instead dry, open habitats. So, in the case of *Lantana*, the origin and subsequent proliferation of fleshy-fruited species is not readily attributable to the selective advantage of increased seed size under light-limiting conditions (Eriksson et al., 2000). The Sarcolippia clade, however, represents a shift into moist, closed forest habitat, coincident with a shift to fleshy fruits.

One possible explanation for the increased evolutionary success of fleshy-fruited lineages in this case might be the increased dispersability resulting from employing animal vectors. Another interpretation is increased specialization in the species of animal mutualists thus employed, but what little is known about the animals that feed

on the fleshy fruits of *Lantana* species suggests generalism (Day et al., 2003). Increased dispersability has been implicated as a potential driver of increases in diversification (e.g., by Vamosi & Vamosi, 2010; Fleming & Kress, 2011); by employing animal dispersers, the species of *Lantana* may have maximized their access to new geographic areas, increasing allopatric separation between populations, promoting local adaptation, and becoming exposed to potential new niche space.

Examining the biogeography of the *Lantana-Lippia* clade in light of the phylogeny reveals a high degree of geographic heterogeneity within clades; i.e., the pattern of co-distributed species of *Lantana* and *Lippia* is one of phylogenetic overdispersion. The full latitudinal range of tribe Lantaneae lies within 45 degrees north and south of the equator; members of both fleshy-fruited and dry fruited lineages can be found throughout the extent of this range (as determined by consulting floristic treatments and biodiversity occurrence records via Global Biodiversity Data Facility). Both fleshy-fruited and dry-fruited lineages have colonized Africa, once each. The Camara clade occupies the full extent of the latitudinal range of Lantaneae, with plants identified as *Lantana camara* occurring from 50 degrees north to 50 degrees south latitude (this range is almost as extensive when invasive *L. camara* is not considered). The distribution of the *Callioreas* clade likewise extends from almost 50 degrees north latitude to 50 degrees south. The sister clades to these lineages are relatively species-poor (the sister to the fleshy-fruited Callioreas clade is here identified as a single species, *Lippia duartei*), and, unsurprisingly, have smaller

geographic ranges. The other fleshy-fruited lineages have smaller ranges, and fewer species, with *Nashia* confined to the West Indies, and the *Sarcolippia* clade found in Brazil and its neighboring countries. Dry-fruited lineages with crown ages comparable to those in the *Camara* and *Callioreas* clades, such as *Phyla*, and the lineage including *Lippia rotundifolia* and *Lippia javanica*, are also distributed throughout the full latitudinal range of Lantaneae. There is no clear qualitative pattern to suggest a major difference in the dispersability of fleshy-fruited vs. dry-fruited lineages, at this phylogenetic scale.

What is known about the habitat preference, ecology, and distribution of the *Lantana* and *Lippia* species studied here does not immediately suggest a mechanism for the difference in speciation rate beween dry-fruited and fleshy-fruited species. Discernible patterns in ecology or geographic range size in relation to fruit type in the *Lantana*-*Lippia* complex might emerge with increased sampling, or finer-level study of ecological traits.

Taxonomic implications

No previously proposed combinations of species under *Lantana*, *Lippia*, or their segregate genera have aligned well with monophyletic groups. Revising these genera under the ICBN will require that they be either lumped, or fragmented. Neither strategy satisfactorily reflects the evolutionary history of the *Lantana*-*Lippia* complex: a single genus fails to evoke its extensive morphological diversity, and

splitting the clade into a large number of genera runs into problems presented by the extensive parallelisms which characterize this diversity.

Under the constraints of the International Code of Botanical Nomenclature, we would favor lumping; uniting all the species of the *Lantana-Lippia* complex under a single name would better reflect the close relationship they share than dividing them under many names. The names *Lantana* and *Lippia* were both created by Linnaeus in the Species Plantarum (1753), taking priority over any other names applied to members of the *Lantana-Lippia* complex. Neither of these names has been conserved, and neither has been used to describe all the species of both, so the names *Lantana* and *Lippia* are equally available for the new, united genus. Our choice would be *Lantana*: it is the type genus of tribe Lantaneae, it precedes *Lippia* both alphabetically, and in order of appearance in the Species Plantarum, and *Lantana* is the more widely known name (probably due to the global impacts of the invasive taxa), despite including fewer species than *Lippia*.

We do not advise fragmenting the *Lantana-Lippia* clade, but even if this strategy were to be followed, it would be inadvisable to proceed until a well-supported phylogenetic hypothesis based on more extensive data can be resolved for the majority of its species. The results presented here are based on representative sampling, and do not fully resolve the tree; relationships of species that are reconstructed with low confidence, or not sampled, may turn out to be other than predicted.

The option exists under the PhyloCode (de Queiroz & Donoghue, 2010) to preserve the names *Lippia*, *Lantana*, *Nashia*, and *Phyla* as clade names; this would involve naming the clade corresponding with the *Lantana-Lippia* complex *Lippia*, and then naming sub-clades within it (de Queiroz & Gauthier, 1992; 1994). Under this scheme, the species of *Phyla* and *Nashia* would keep their names, the Camara clade would be named *Lantana*, and the name *Callioreas* could be applied to the Callioreas clade. *Sarcolippia*, as another distinct lineage of fleshy-fruited species, distinguished by fruits splitting at maturity into two halves each containing one pyrene, could also be named as a sub-clade within *Lippia*. In our opinion, this strategy satisfactorily reflects the different lineages within the *Lantana-Lippia* complex, while retaining the identity of the group as a whole, without the expectation of equivalence between ranks. A taxonomically-focused paper, in which these concepts are discussed and formal revisions are made, is planned for a later date.

Conclusions

The species of the *Lantana-Lippia* complex are closely related and recently diversified, belying their remarkable morphological diversity and wide geographic distribution. The close relationships between them can be resolved using DNA sequence data of sufficient variability, and in sufficient quantity, but care should be taken in interpreting the results, due to the possible confounding effects of gene tree/species tree incongruence. Fleshy fruits are associated with increased speciation rates in the *Lantana-Lippia* complex, but a candidate process that might

cause this pattern is not immediately apparent. All major lineages of the *Lantana-Lippia* complex are distributed throughout the extent of the group's geographic range, implying equivalent dispersability over evolutionary timescales, regardless of fruit type. Revising generic boundaries within the *Lantana-Lippia* complex will not be straightforward; we recommend either absorbing all its species into an expanded *Lantana*, and/or exploring a rank-free classification scheme.

Table 3.1 Summary of sequence data collected as part of this study: alignment dimensions for each of seven loci, assembled into individual data sets, and best-fit models for each data set.

	ETS	**ITS**	**PPR 11**	**PPR 81**	**PPR 90**	**PPR 97**	**PPR 123**
Length	480	762	1277	1162	986	747	1122
Accessions	81	80	80	76	76	63	76
Model	GTR+Γ	GTR+I+Γ	GTR+Γ	GTR+I+Γ	HKY+I+Γ	HKY+I+Γ	HKY+I+Γ

Table 3.2 Results of BiSSE maximum likelihood estimation of trait-dependent speciation rate (λ), and comparison with equal rates model.

	df	**lnL**	**AIC**	**χ2**	**significance**
Full model	6	-194.78	401.55		
Equal rates	5	-197.56	405.11	5.5587	0.018

Figure 3.1 Phylogenetic tree inferred from concatenated sequence data from seven nuclear loci (6,536 aligned positions). Posterior probability values greater than 0.5 are shown above branches.

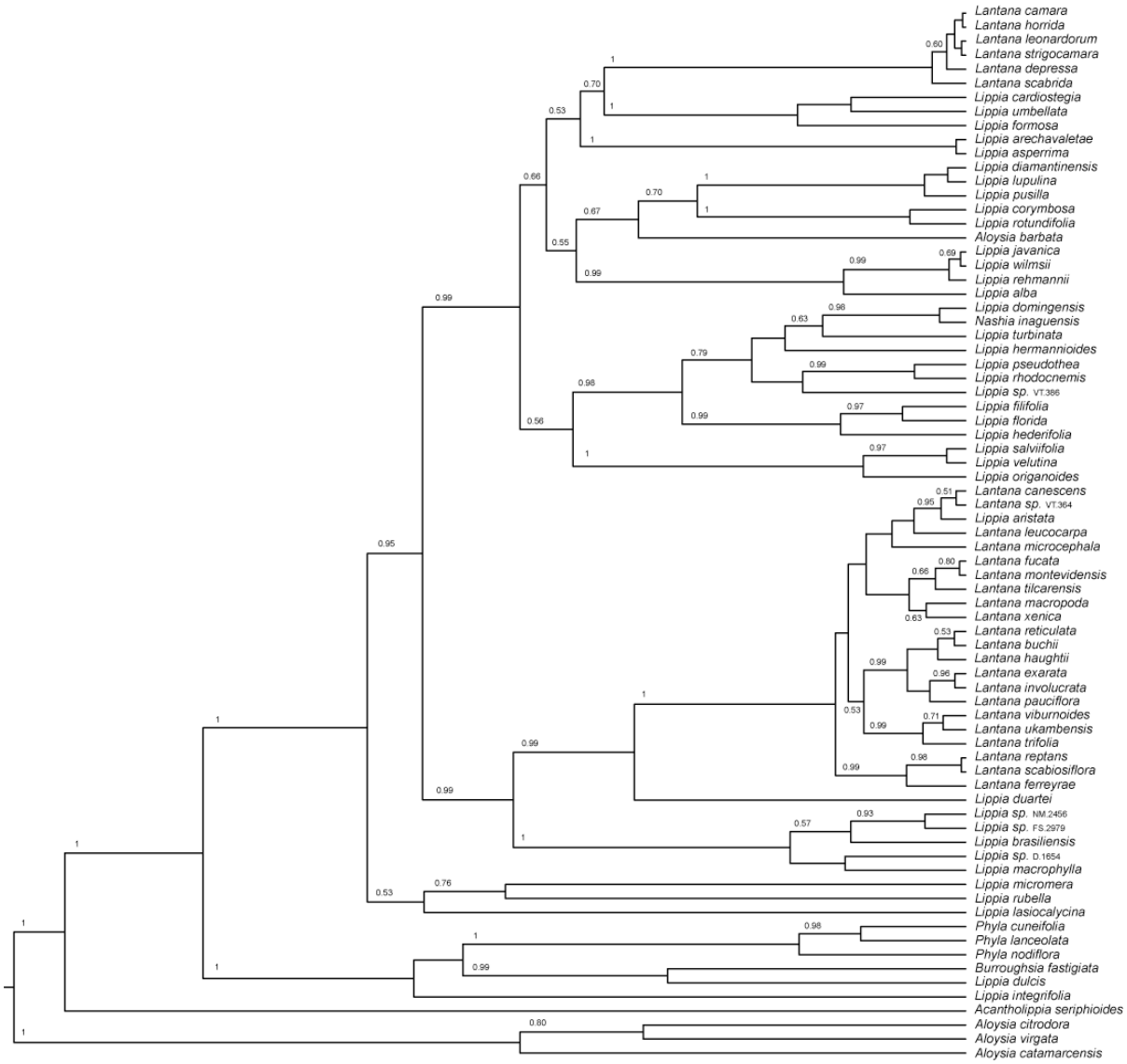


Figure 3.2 Species tree (accounting for incongruence between gene trees using coalescent theory) inferred from sequence data from seven nuclear loci in combination. Posterior probability values greater than 0.5 are shown above branches.

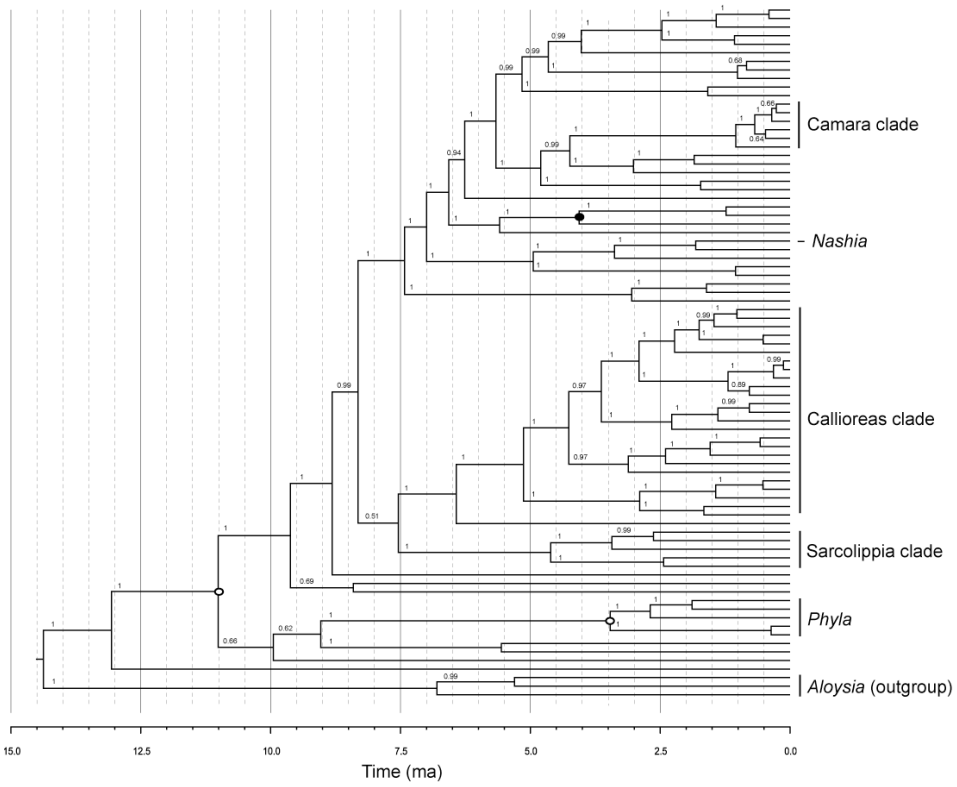


Figure 3.3 Time-calibrated (ultrametric) tree inferred from 6,536 aligned positions of DNA sequence data from seven nuclear loci in concatenation. The three calibration points used are indicated with circles: white circles are secondary calibrations from analysis of divergence timing across Verbenaceae; the black circle corresponds with the maximum age set for the Cerrado-endemic *L. hederifolia* lineage.

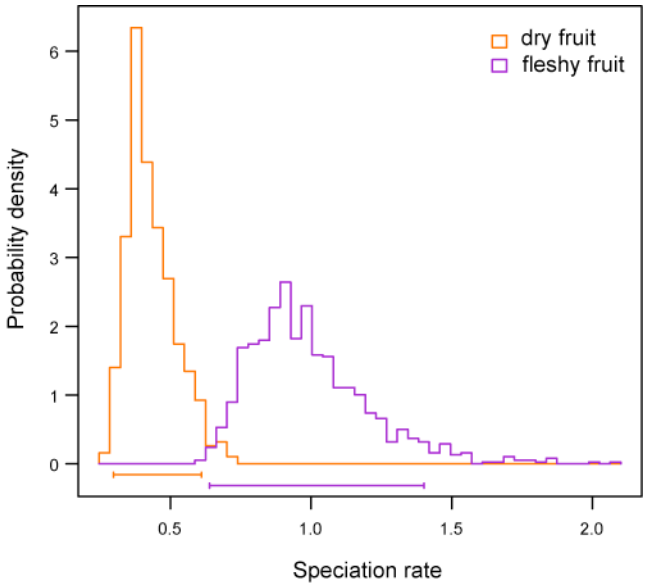


Figure 3.4 Graph depicting 95% confidence intervals of estimates of speciation rate in dry-fruited and fleshy-fruited species.

LITERATURE CITED

Bremer, K., E. M. Friis, and B. Bremer. 2004. Molecular phylogenetic dating of Asterid flowering plants shows early Cretaceous diversification. *Systematic Biology* **53**: 496-505.

Briquet, I. 1904. Verbenaceae. *Plantae Hassleriane. Bulletin de l'Herbier Boissier ser. 2* **4**: 1062–1066.

Cantino, P. D. and K. de Queiroz. 2010. International Code of Phylogenetic Nomenclature. Version 4c. http://www.ohio.edu/phylocode/preface.html

Chamisso, A. 1832. De plantis in Expeditione Romanzoffiana observatis dicunt. Verbenaceae. *Linnaea* **7**: 105–128; 213–272.

Cousens, R., C. Dytham, and R. Law. 2008. Dispersal in Plants: A Population Perspective. Oxford University Press, New York, New York, USA.

Crepet, W. L. and K. J. Niklas. 2009. Darwin's second "abominable mystery": why are there so many angiosperm species? *American Journal of Botany* **96**: 366-381.

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772.

Day, M. D., C. J. Wiley, J. Playford, and M. P. Zalucki. 2003. Lantana: current management status and future prospects. *ACIAR Monograph Series* **102**.

Davies, T. J. and T. G. Barraclough. 2007. The Diversification of Flowering Plants through Time and Space: Key Innovations, Climate and Chance. *In* T. R. Hodkinson and J. A. N. Parnell, Reconstructing the tree of life: taxonomy and systematics of species rich taxa, 149-163. CRC Press, Boca Raton, Florida, USA.

de Queiroz, K. and J. Gauthier. 1992. Phylogenetic taxonomy. *Annual Review of Ecology and Systematics* **23**: 449-480.

de Queiroz, K. and J. Gauthier. 1994. Toward a phylogenetic system of biological nomenclature. *Trends in Ecology and Evolution* **9**: 27-31.

Dodd, M. E., J. Silvertown, and M. W. Chase. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution* **53**: 732-744.

Dos Santos Silva, T. R. and F. R. G. Salimena. 2002. Novas combinações e novos sinônimos em *Lippia* e *Lantana* (Verbenaceae). *Darwiniana* **40**: 57-59.

Doyle, J. J. and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19**: 11–15.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology And Evolution* **29**: 1969-1973.

Dupont, L. M., H. Behling, and J-H Kim. 2008. Thirty thousand years of vegetation development and climate change in Angola (Ocean Drilling Program Site 1078). *Climate of the Past* **4**: 107-124.

Eriksson, O. and B. Bremer. 1991. Fruit characteristics, life forms, and species richness in the plant family Rubiaceae. *American Naturalist* **138**: 751-761.

Eriksson, O. and B. Bremer. 1992. Pollination systems, dispersal modes, life forms, and diversification rates in angiosperm families. *Evolution* **46**: 258-266.

Eriksson, O., E. M. Friis, and P. Löfgren. 2000. Seed size, fruit size, and dispersal systems in angiosperms from the early Cretaceous to the late Tertiary. *American Naturalist* **156**: 47-58.

FitzJohn, R. G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* **3**: 1084-1092.

Fleming, T. H., and J. Kress. 2011. A brief history of fruits and frugivores. *Acta Oecologica* **37**: 521-530.

Heard, S. B. and D. L. Hauser. 1995. Key evolutionary innovations and their ecological mechanisms. *Historical Biology* **10**: 151-173.

Heled, J. and A. J. Drummond. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27**: 570–580.

Herrera, C. M. 1989. Seed *dispersal by animals*: a role in angiosperm diversification? *American Naturalist* **133**: 309-322.

Hodges, S. A. 1997. Floral nectar spurs and diversification. *International Journal of Plant Sciences* **158**: S81-S88.

Hodges, S. A. and M. L. Arnold. 1995. Spurring plant diversification: are floral nectar spurs a key innovation? *Proceedings of the Royal Society of London Series B: Biological Sciences* **262**: 343-348.

Howe, H. F. and J. Smallwood. 1982. Ecology of seed dispersal. *Annual Review of Ecology, Evolution, and Systematics* **13**: 201-228.

Katoh, K. and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772-780.

Knörr, U. C., J. Kovar-Eder, P. Mazouch, and A. Roth-Nebelsick. 2012. Fruit dispersal ecology of woody taxa in temperate to tropical forests of China and Japan. *Palaios* **27**: 523-540.

Levin, S. A., H. C. Muller-Landau, R. Nathan, and J. Chave. 2003. The ecology and evolution of seed dispersal. *Annual Review of Ecology, Evolution, and Systematics* **34**: 575-604.

Linnaeus, C. 1753. Species Plantarum. Laurentii Salvii, Stockholm, Sweden.

Lu-Irving, P. and R. G. Olmstead. 2013. Investigating the evolution of Lantaneae (Verbenaceae) using multiple loci. *Botanical Journal of the Linnean Society* **171**: 103–119.

MacGinitie, H. D. 1953. Fossil plants of the Florissant beds of Colorado. *Carnegie Institute of Washington, Contributions to Paleontology* **599**: 1-198.

Maddison, W. 1997. Gene trees in species trees. *Systematic Biology* **46**: 523–536.

Maddison, W. P., P. E. Midford, and S. P. Otto. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* **56**: 701-710.

Magallón, S. and A. Castillo. 2009. Angiosperm diversification through time. *American Journal of Botany* **96**: 349-365.

Magallón, S. and M. J. Sanderson. 2001. Absolute diversification rates in angiosperm clades. *Evolution* **55**: 1762-1780.

McPeek, M. A. and J. M. Brown. 2007. Clade age and not diversification rate explains species richness among animal taxa. *American Naturalist* **169**: E97-E106.

Nylander, J. A., J. C. Wilgenbusch, D. L. Warren, and D. L. Swofford. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* **24**: 581-583.

O'Leary, N. and M. E. Múlgura. 2012. A Taxonomic Revision of the Genus *Phyla* (Verbenaceae). *Annals of the Missouri Botanical Garden* **98**: 578-596.

Pigot, A. L., A. B. Phillimore, I. P. Owens, and C. D. Orme. 2010. The shape and temporal dynamics of phylogenetic trees arising from geographic speciation. *Systematic Biology* **59**: 660-673.

Rabosky, D. L. 2009. Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecology Letters* **12**: 735-743.

Rambaut, A. and A. J. Drummond. 2009. Tracer v.1.5, available from http://beast.bio.ed.ac.uk/Tracer.

Regal, P. J. 1977. Ecology and evolution of flowering plant dominance. *Science* **196**: 622-629.

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

Sanders, R. W. 2006. Taxonomy of *Lantana* sect. *Lantana* (Verbenaceae): I. Correct application of *Lantana camara* and associated names. *Sida* **22**: 381-421.

Schauer, J. C. 1847. Verbenaceae. In: De Candolle AP, ed. *Prodromus.* Lehre: Cramer, **11**: 522–700.

Simon, M. F., G. Grether, L. P. de Queiroz, C. Skema, R. T. Pennington, and C. E. Hughes. 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences of the United States of America* 106: 20359-20364.

Smith, J. F. 2001. High species diversity in fleshy-fruited tropical understory plants. *American Naturalist* **157**: 646-653.

Tiffney, B. H. 1984. Seed size, dispersal syndromes, and the rise of angiosperms: evidence and hypothesis. *Annals of the Missouri Botanical Garden* **71**: 551-576.

Tiffney, B. H. and S. J. Mazer. 1995. Angiosperm growth habit, dispersal and diversification reconsidered. Evolutionary Ecology **9**: 93-117.

Troncoso, N. S. 1974. Los géneros de Verbenaceas de sudamerica extra-tropical (Argentina, Chile, Bolıvia, Paraguai, Uruguai y sur de Brasil). *Darwiniana* **18**: 295–412.

Vamosi, J. C. and S. M. Vamosi. 2010. Key innovations within a geographical context in flowering plants: towards resolving Darwin's abominable mystery. *Ecology Letters* **13**: 1270-1279.

Vamosi, J. C. and S. M. Vamosi. 2011. Factors influencing diversification in angiosperms: at the crossroads of intrinsic and extrinsic traits. *American Journal of Botany* **98**: 460-471.

White, T. J., T. Bruns, S. Lee, and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, PCR Protocols: a guide to methods and applications, 315–322. Academic Press, New York, New York, USA.

Yuan, Y-W., C. Liu, H. E. Marx, and R. G. Olmstead. 2009 a. The pentatricopeptide repeat (PPR) gene family, a tremendous resource for plant phylogenetic studies. *New Phytologist* **182**: 272–283

Yuan, Y-W., C. Liu, H. E. Marx, and R. G. Olmstead. 2009 b. An empirical demonstration of using pentatricopeptide repeat (PPR) genes as plant phylogenetic tools: Phylogeny of Verbenaceae and the *Verbena* complex. *Molecular Phylogenetics and Evolution* **54**: 23–35.

Yuan, Y-W. and R. G. Olmstead. 2008. Evolution and phylogenetic utility of the PHOT gene duplicates in the *Verbena* complex (Verbenaceae): dramatic

intron size variation and footprint of ancestral recombination. *American Journal of Botany* **95**: 1166–1176.

Concluding remarks

Lantaneae are a complicated lineage, presenting many challenges to a plant systematist: the number and limits of species are difficult to define, as are the limits of genera. The range of diversity in ecologically important traits invites the application of phylogenetic approaches to understand their evolutionary patterns, but the species richness of the tribe precludes the population-level sampling which would be the best approach to untangle its difficult phylogenetic history. When I began my studies of Lantaneae, it was with the naïve attitude that I would solve all the problems it presented with the straightforward application of targeted (Sanger) sequencing approaches to phylogenetic reconstruction. Six years later, rather more humble (but also rather more educated), I hope that I have achieved some measure of my starting aspirations: a meaningful contribution to the ongoing efforts to sort out the systematics of this mysterious, beautiful, frustrating and fascinating lineage of plants.

Good justification now exists for taxonomic revisions which would result in a stable, phylogenetically-based classification scheme. According to the proposed scheme, in compliance with the International Code of Botanical Nomenclature, tribe Lantaneae consists of two widespread genera: *Aloysia* (37 species) and *Lantana* (230 species), each encompassing a wide range of morphological variation. I intend to complete these changes following the submission of this thesis, and expect that the revised scheme will come into acceptance over the next few years. The evolution of the array of morphological diversity in Lantaneae was just as complex as might have

been expected, given the difficulty of its taxonomy. Almost every trait used as an indicator of evolutionary affinity has undergone multiple shifts: the presence or absence of a terminal unit in the inflorescence, a fleshy outer layer on the fruit, whether the mature fruit splits into two single-seeded halves. Other traits vary on a continuous spectrum, labile within and between lineages: the density of the inflorescence, ranging from a loose raceme to a tight head, the prominence of the floral bracts, from inconspicuous to large and showy, the color of the corollas, from white, to yellow through red, to lavender.

Clearly, there is great potential for more extensive study of Lantaneae from an evolutionary perspective. I hope to have the opportunity to build on the work described in this thesis, to continue to unravel the mysteries that this lineage presents. With the growing accessibility of next-generation sequencing and the large quantities of data that it makes available, this is an exciting time to be a student of systematics and evolution. I am grateful for everything that I have learned so far, and look forward to the lessons of the future.

Appendices

Appendix 1. Sample DNA accession and voucher information.

Accession	Species	Voucher	Origin	*trn*T-L	*rpl*32-*trn*L	*trn*Q-*rps*16	ETS	ITS	PPR 11	PPR 81	PPR 90	PPR 97	PPR 123
07-01	*Acantholippia deserticola*	Biurrun 4963; SI	Argentina	pending	pending	pending	pending	-	-	-	-	-	pending
07-64	*Acantholippia salsoloides*	Olmstead 07-23; WTU	Argentina	pending	pending	pending	pending	-	-	pending	-	-	pending
07-65	*Acantholippia salsoloides*	Olmstead 07-28; WTU	Argentina	JX966953	JX966845	JX966899	JX966792	-	JX966650	JX966695	-	-	JX966746
07-66	*Acantholippia salsoloides*	Olmstead 07-52; WTU	Argentina	pending	pending	pending	pending	-	-	pending	-	-	pending
06-73	*Acantholippia seriphioides*	Olmstead 04-146; WTU	Argentina	JX966954	JX966846	JX966900	JX966793	pending	JX966651	JX966696	pending	pending	JX966747
10-56	*Acantholippia seriphioides*	Correa 10152; SI	Argentina	pending	pending	pending	pending	-	-	-	-	-	-
06-84	*Acantholippia trifida*	Biurrun 7706; SI	Argentina	pending	pending	pending	pending	-	-	pending	-	-	pending
11-107	*Aloysia axillaris*	Wood & Atahuachi 21575; KEW	Bolivia	pending	pending	pending	pending	-	-	pending	-	-	pending
10-62	*Aloysia barbata*	Carter & Ferris 3902; US	Mexico	pending	pending	pending	pending	-	-	pending	-	-	pending
11-102	*Aloysia barbata*	Carter & Ferris 3902A; TEX	Mexico	JX966955	JX966847	JX966901	JX966794	pending	JX966652	JX966697	pending	-	JX966748
10-58	*Aloysia castellanosii*	Ferriencia 41191; MERL	Argentina	pending	pending	pending	pending	-	-	pending	-	-	pending
07-90	*Aloysia catamarcensis*	Olmstead 07-82; WTU	Argentina	pending	pending	pending	pending	pending	pending	pending	pending	pending	pending
08-209	*Aloysia chamaedryfolia*	H. Rimpler 1131; FB	Brazil	pending	pending	pending	pending	-	-	pending	-	-	pending
10-290	*Aloysia chamaedryfolia*	Thode 102; ICN	Brazil	pending	pending	-	pending	-	-	pending	-	-	pending
11-104	*Aloysia chiapensis*	Martinez 932; TEX	Mexico	pending	pending	pending	pending	-	-	pending	-	-	pending
07-61	*Aloysia citrodora*	Olmstead 07-13; WTU	Argentina	JX966956	JX966848	JX966902	JX966795	pending	JX966653	JX966698	pending	pending	JX966749
07-02	*Aloysia crenata*	Cabrera 29106; SI	Argentina	pending	pending	pending	pending	-	-	pending	-	-	-
08-217	*Aloysia dusenii*	Krapovickas & Schinini 38344; TEX	Brazil	pending	pending	pending	pending	-	-	pending	-	-	pending
10-195	*Aloysia dusenii*	Olmstead 10-217; WTU	Brazil	pending	pending	pending	pending	-	-	pending	-	-	pending
04-64	*Aloysia gratissima*	Valencia BG 460-00; VAL	cultivated	JX966958	JX966850	JX966904	JX966797	-	JX966655	JX966700	-	-	JX966751
08-199	*Aloysia gratissima*	Lu-Irving 08-17; WTU	Texas	pending	pending	pending	pending	-	-	pending	-	-	pending
10-106	*Aloysia gratissima*	Turner 26-28; TEX	Texas	pending	pending	pending	pending	-	-	pending	-	-	pending
10-61	*Aloysia hatschbachii*	Hatschbach 51897; US	Brazil	pending	pending	-	pending	-	-	pending	-	-	pending
09-93	*Aloysia herrerae*	Olmstead 09-30; WTU	Peru	JX966957	JX966849	JX966903	JX966796	-	JX966654	JX966699	-	-	JX966750
11-109	*Aloysia herrerae*	Wood & Serrano 14658; KEW	Bolivia	pending	pending	pending	pending	-	-	pending	-	-	pending
10-122	*Aloysia looseri*	Roig 9847; MERL	Ecuador	pending	pending	pending	pending	-	-	-	-	-	pending
92-199	*Aloysia lycioides*	Kew BG 251-76-02169; KEW	cultivated	pending	pending	pending	pending	-	-	pending	-	-	pending
08-200	*Aloysia macrostachya*	Lu-Irving 08-19; WTU	Texas	pending	pending	pending	pending	-	-	pending	-	-	pending
08-205	*Aloysia macrostachya*	Lu-Irving 08-14; WTU	Texas	JX966959	JX966851	JX966905	JX966798	-	JX966656	JX966701	-	-	JX966752
10-312	*Aloysia oblanceolata*	Thode 96; ICN	Brazil	pending	pending	pending	pending	-	-	pending	-	-	pending

09-95	*Aloysia peruviana*	Olmstead 09-45; WTU	Peru	pending	pending	pending	pending	-	-	pending	-	-	pending
10-293	*Aloysia polygalifolia*	Thode 398; ICN	Brazil	pending	pending	pending	pending	-	-	pending	-	-	pending
10-277	*Aloysia polystachya*	Kranz 817; CESJ	Brazil	pending	pending	pending	pending	-	-	pending	-	-	pending
07-49	*Aloysia pulchra*	Olmstead 04-129; WTU	Argentina	pending	pending	pending	pending	-	-	pending	-	-	pending
10-311	*Aloysia pulchra*	Thode 157; ICN	Brazil	pending	pending	pending	pending	-	-	pending	-	-	pending
07-3	*Aloysia scorodonioides*	Saravia 1591; SI	Argentina	pending	pending	pending	pending	-	-	-	-	-	pending
09-94	*Aloysia scorodonioides*	Olmstead 09-40; WTU	Peru	pending	pending	pending	pending	-	-	pending	-	-	pending
10-128	*Aloysia scorodonioides*	Lu-Irving 09-62; WTU	Peru	pending	pending	pending	pending	-	-	pending	-	-	pending
11-105	*Aloysia sonoriensis*	Reichenbacher 85-1108; TEX	Mexico	pending	pending	pending	pending	-	-	pending	-	-	pending
04-63	*Aloysia virgata*	Valencia BG 232-97; VAL	cultivated	JX966960	JX966852	JX966906	JX966799	-	JX966657	JX966702	-	-	JX966753
05-12	*Aloysia virgata*	Olmstead 04-133; WTU	Argentina	pending	pending	pending	pending	pending	pending	pending	-	pending	pending
07-70	*Aloysia virgata*	Olmstead 07-68; WTU	Argentina	pending	pending	pending	pending	-	-	pending	-	-	pending
08-206	*Aloysia wrightii*	Ocampo 1765; WTU	cultivated	pending	pending	pending	pending	-	-	pending	-	-	pending
08-210	*Aloysia wrightii*	Olmstead 91-4; WTU	Arizona	pending	pending	pending	pending	-	-	pending	-	-	pending
10-55	*Burroughsia fastigiata*	Sikes & Babcock 294; TEX	Mexico	JX966961	JX966853	JX966907	JX966800	pending	JX966658	-	-	pending	JX966754
08-169	*Citharexylum montevidense*	Olmstead 04-102; WTU	Argentina	JX966962	JX966854	JX966908	JX966801	-	FJ549107	JX966703	-	-	FJ549285
08-361	*Coelocarpum swinglei*	Phillipson 3443; MO	Madagascar	JX966963	JX966855	JX966909	JX966802	-	JX966659	JX966704	-	-	JX966755
06-77	*Dipyrena glaberrima*	Olmstead 04-179; WTU	Argentina	JX966964	JX966856	JX966910	JX966803	-	FJ549099	JX966705	-	-	FJ549277
10-29	*Junellia succulentifolia*	Olmstead 10-1; WTU	Argentina	JX966965	JX966857	JX966911	JX966804	-	JX966660	JX966706	-	-	JX966756
13-26	*Lantana buchii*	Lu-Irving 12-107; WTU	Dominican Republic	-	-	-	pending	pending	pending	pending	pending	pending	pending
11-114	*Lantana camara*	Lu-Irving 12-1; WTU	cultivated	JX966966	JX966858	JX966912	JX966805	-	JX966661	JX966707	-	-	JX966757
12-176	*Lantana camara*	Lu-Irving 12-37; WTU	Puerto Rico	-	-	-	pending	pending	pending	pending	pending	pending	pending
07-58	*Lantana canescens*	Olmstead 07-06; WTU	Argentina	JX966967	JX966859	JX966913	JX966806	pending	FJ549096	JX966708	pending	pending	FJ549274
08-202	*Lantana canescens*	Lu-Irving 08-7; WTU	cultivated	-	-	-	pending	pending	pending	pending	pending	-	pending
10-227	*Lantana cujabensis*	Lu-Irving 10-19; WTU	Brazil	JX966968	JX966860	JX966914	JX966807	-	JX966662	JX966709	-	-	JX966758
12-62	*Lantana depressa*	Lu-Irving 12-1; WTU	Florida	-	-	-	pending	pending	pending	pending	pending	pending	pending
12-188	*Lantana exarata*	Lu-Irving 12-49; WTU	Puerto Rico	-	-	-	pending	pending	pending	pending	pending	pending	pending
12-65	*Lantana ferreyrae*	Lu-Irving s.n.; WTU	Peru	-	-	-	pending	pending	pending	pending	pending	pending	pending
10-169	*Lantana fucata*	Salimena 2952; CESJ	Brazil	JX966969	JX966861	JX966915	JX966808	pending	JX966663	JX966710	pending	-	JX966759
13-62	*Lantana haughtii*	Lu-Irving 09-34; WTU	Peru	-	-	-	pending	pending	pending	pending	pending	pending	pending
12-200	*Lantana horrida*	Lu-Irving 12-61; WTU	Dominican Republic	-	-	-	pending	pending	pending	pending	pending	pending	pending
12-90	*Lantana involucrata*	Lu-Irving 12-13; WTU	Florida	-	-	-	pending	pending	pending	pending	pending	pending	pending
13-21	*Lantana leonardorum*	Lu-Irving 12-102; WTU	Dominican Republic	-	-	-	pending	pending	pending	pending	pending	pending	pending
12-209	*Lantana leucocarpa*	Lu-Irving 12-70; WTU	Dominican Republic	-	-	-	pending	pending	pending	pending	pending	-	pending
08-222	*Lantana macropoda*	Nesom & Mayfield 7355; TEX	Mexico	JX966971	JX966863	JX966917	JX966810	pending	JX966665	JX966712	-	-	JX966761

07-59	*Lantana micrantha*	Olmstead 07-8; WTU	Argentina	JX966972	JX966864	JX966918	JX966811	-	JX966666	JX966713	-	-	JX966762
08-202	*Lantana microcephala*	Lu-Irving 08-7; WTU	cultivated	JX966973	JX966865	JX966919	JX966812	-	JX966667	JX966714	-	-	JX966763
08-203	*Lantana montevidensis*	Lu-Irving 08-15; WTU	Texas	JX966974	JX966866	JX966920	JX966813	pending	JX966668	JX966715	pending	-	JX966764
10-191	*Lantana montevidensis*	Olmstead 10-203; WTU	Brazil	-	-	-	pending	pending	pending	pending	pending	pending	pending
13-25	*Lantana pauciflora*	Lu-Irving 12-106; WTU	Dominican Republic	-	-	-	pending	pending	pending	pending	pending	pending	pending
13-60	*Lantana reptans*	Lu-Irving 09-14; WTU	Peru	-	-	-	pending	pending	pending	pending	pending	pending	pending
12-205	*Lantana reticulata*	Lu-Irving 12-66; WTU	Dominican Republic	-	-	-	pending	pending	pending	pending	pending	-	pending
10-51	*Lantana rugosa*	Lu-Irving 08-25; WTU	South Africa	JX966975	JX966867	JX966921	JX966814	-	JX966669	-	-	-	JX966765
13-59	*Lantana scabiosiflora*	Lu-Irving 09-1; WTU	Peru	-	-	-	pending	pending	pending	pending	pending	pending	pending
13-8	*Lantana scabrida*	Lu-Irving 12-89; WTU	Dominican Republic	-	-	-	pending	pending	pending	pending	pending	pending	pending
10-206	*Lantana* sp.	Salimena 2979; WTU	Brazil	-	-	-	pending	pending	pending	pending	pending	-	pending
10-209	*Lantana* sp.	Thode 364; ICN	Brazil	-	-	-	pending	-	pending	pending	pending	-	pending
12-99	*Lantana strigocamara*	Lu-Irving 12-22; WTU	Puerto Rico	-	-	-	pending	pending	pending	pending	pending	pending	pending
07-62	*Lantana tilcarensis*	Olmstead 07-18; WTU	Argentina	-	-	-	pending	pending	pending	pending	pending	pending	pending
13-46	*Lantana trifolia*	Lu-Irving s.n.; WTU	Peru	-	-	-	pending	pending	pending	pending	pending	pending	pending
13-9	*Lantana trifolia*	Lu-Irving 12-90; WTU	Dominican Republic	-	-	-	pending	pending	pending	pending	pending	pending	pending
97-36	*Lantana trifolia*	Olmstead 96-98; WTU	cultivated	JX966976	JX966868	JX966922	JX966815	pending	JX966670	JX966716	pending	pending	JX966766
10-46	*Lantana ukambensis*	Mawi 80; MO	Tanzania	-	-	-	pending	pending	pending	-	pending	pending	-
08-196	*Lantana urticoides*	Lu-Irving 08-2; WTU	Texas	JX966970	JX966862	JX966916	JX966809	-	JX966664	JX966711	-	-	JX966760
08-257	*Lantana viburnoides*	Miyazaki 991013R29; TEX	Saudi Arabia	JX966977	JX966869	JX966923	JX966816	pending	JX966671	JX966717	pending	pending	JX966767
08-204	*Lantana xenica*	Soza 1838; WTU	Argentina	-	-	-	pending	pending	pending	pending	pending	pending	pending
04-29	*Lippia alba*	Fairchild BG 37139; FTG	cultivated	JX966978	JX966870	JX966924	JX966817	pending	JX966672	JX966718	pending	pending	JX966768
06-68	*Lippia alba*	Olmstead 04-110; WTU	Argentina	-	-	-	pending	pending	pending	pending	-	pending	pending
10-308	*Lippia arechavaletae*	Thode 54; ICN	Brazil	-	-	-	pending	pending	pending	-	pending	pending	-
10-167	*Lippia aristata*	Lu-Irving 10-5; WTU	Brazil	JX966979	JX966871	JX966925	JX966818	pending	JX966673	JX966719	pending	-	JX966769
06-74	*Lippia asperrima*	Olmstead 04-140; WTU	Argentina	-	-	-	pending	pending	pending	pending	pending	pending	-
10-163	*Lippia brasiliensis*	Lu-Irving 10-17; WTU	Brazil	JX966980	JX966872	JX966926	JX966819	pending	JX966674	JX966720	pending	pending	JX966770
04-36	*Lippia cardiostegia*	Grose 144; WTU	Nicaragua	-	-	-	pending	pending	pending	pending	-	-	pending
10-162	*Lippia corymbosa*	Lu-Irving 10-13; WTU	Brazil	-	-	-	pending	pending	pending	pending	pending	pending	pending
10-170	*Lippia diamantinensis*	Salimena 2943; CESJ	Brazil	JX966981	JX966873	JX966927	JX966820	pending	JX966675	JX966721	pending	-	JX966771
12-218	*Lippia domingensis*	Lu-Irving 12-80; WTU	Dominican Republic				pending	pending	pending	pending	pending	pending	pending
10-204	*Lippia duartei*	Lu-Irving 10-11; WTU	Brazil	JX966982	JX966874	JX966928	JX966821	pending	JX966676	JX966722	pending	pending	JX966772
13-45	*Lippia dulcis*	Lu-Irving 13-2; WTU	cultivated	-	-	-	-	pending	-	-	pending	pending	-
99-45	*Lippia dulcis*	Olmstead 98-56; WTU	cultivated	JX966983	JX966875	JX966929	JX966822	-	FJ549095	JX966723	-	-	FJ549273
10-153	*Lippia filifolia*	Thode 352; WTU	Brazil	JX966984	JX966876	JX966930	JX966823	pending	JX966677	JX966724	pending	pending	JX966773
10-173	*Lippia florida*	Salimena 2945; CESJ	Brazil	-	-	-	pending	pending	pending	pending	pending	pending	pending

08-207	*Lippia formosa*	Ocampo 1764; WTU	cultivated	-	-	-	pending	pending	pending	pending	pending	-	pending
10-175	*Lippia hederifolia*	Lu-Irving 10-14; WTU	Brazil	JX966985	JX966877	JX966931	JX966824	pending	JX966678	JX966725	pending	pending	JX966774
10-168	*Lippia hermannioides*	Thode 389; WTU	Brazil	JX966986	JX966878	JX966932	JX966825	pending	JX966679	JX966726	pending	pending	JX966775
07-87	*Lippia integrifolia*	Olmstead 07-78; WTU	Argentina	-	-	-	pending	pending	pending	pending	pending	pending	pending
11-155	*Lippia javanica*	Lu-Irving 12-1A; WTU	South Africa	JX966987	JX966879	JX966933	JX966826	pending	JX966680	JX966727	pending	pending	JX966776
10-207	*Lippia lasiocalycina*	Thode 363; WTU	Brazil	-	-	-	pending	pending	pending	pending	pending	pending	pending
10-171	*Lippia lupulina*	Salimena 2941; CESJ	Brazil	JX966988	JX966880	JX966934	JX966827	pending	JX966681	JX966728	pending	-	JX966777
10-259	*Lippia macrophylla*	Thomas 13474; CESJ	Brazil	JX966989	JX966881	JX966935	JX966828	pending	JX966682	JX966729	pending	-	JX966778
92-225	*Lippia micromera*	Olmstead 92-225; WTU	cultivated	JX966990	JX966882	JX966936	JX966829	pending	JX966683	JX966730	pending	pending	JX966779
10-148	*Lippia origanoides*	Lu-Irving 10-18; WTU	cultivated	JX966991	JX966883	JX966937	pending	pending	JX966684	JX966731	pending	pending	JX966780
92-210	*Lippia origanoides*	Olmstead 92-210; WTU	cultivated	-	-	-	JX966830	-	-	-	-	-	-
10-63	*Lippia phryxocalyx*	Eiten 4506; US	Brazil	pending	pending	pending	pending	-	-	-	-	-	-
10-172	*Lippia pseudothea*	Salimena 2940; CESJ	Brazil	-	-	-	pending	pending	pending	pending	pending	pending	pending
10-197	*Lippia pusilla*	Thode 337; ICN	Brazil	JX966992	JX966884	JX966938	JX966831	pending	JX966685	JX966732	pending	pending	JX966781
10-53	*Lippia rehmannii*	Lu-Irving 08-20; WTU	South Africa	JX966993	JX966885	JX966939	JX966832	-	JX966686	-	-	-	JX966782
12-63	*Lippia rehmannii*	Lu-Irving 13-1; WTU	South Africa	-	-	-	pending	pending	pending	-	pending	pending	pending
10-161	*Lippia rhodocnemis*	Lu-Irving 10-6; WTU	Brazil	JX966994	JX966886	JX966940	JX966833	pending	JX966687	JX966733	pending	pending	JX966783
10-151	*Lippia rotundifolia*	Salimena 2958; CESJ	Brazil	JX966995	JX966887	JX966941	JX966834	pending	JX966688	JX966734	pending	pending	JX966784
10-152	*Lippia rubella*	Lu-Irving 10-3; WTU	Brazil	JX966996	JX966888	JX966942	JX966835	pending	JX966689	JX966735	pending	pending	JX966785
10-150	*Lippia salviifolia*	Salimena 2975; WTU	Brazil	JX966997	JX966889	JX966943	JX966836	pending	JX966690	JX966736	pending	pending	JX966786
10-201	*Lippia* sp.	Thode 386; ICN	Brazil	-	-	-	pending	pending	pending	-	pending	-	-
10-248	*Lippia* sp.	Dittrich 1654; CESJ	Brazil	-	-	-	pending	pending	pending	pending	pending	pending	pending
12-76	*Lippia* sp.	Mota 2456; BHCB	Brazil	-	-	-	pending	pending	pending	pending	pending	-	pending
07-85	*Lippia turbinata*	Olmstead 07-74; WTU	Argentina	JX966998	JX966890	JX966944	JX966837	pending	JX966691	JX966737	pending	pending	JX966787
08-243	*Lippia umbellata*	Van Devender 06-194; TEX	Mexico	-	-	-	pending	pending	pending	pending	pending	-	-
10-166	*Lippia velutina*	Lu-Irving 10-16; WTU	Brazil	JX966999	JX966891	JX966945	JX966838	pending	JX966692	JX966738	pending	pending	JX966788
12-64	*Lippia wilmsii*	Lu-Irving 12-111; WTU	South Africa	-	-	-	pending	pending	pending	pending	pending	pending	pending
10-50	*Nashia inaguensis*	Lu-Irving s.n.; WTU	cultivated	JX967000	JX966892	JX966946	JX966839	pending	JX966693	JX966739	pending	pending	JX966789
07-86	*Neosparton ephedroides*	Olmstead 07-77; WTU	Argentina	JX967001	JX966893	JX966947	JX966840	-	FJ549101	JX966740	-	-	FJ549279
08-218	*Phyla cuneifolia*	Olmstead 92-134; WTU	Colorado	-	-	-	pending	pending	-	pending	pending	pending	pending
08-195	*Phyla lanceolata*	Lu-Irving 08-16; WTU	Texas	-	-	-	pending	pending	pending	pending	pending	pending	pending
06-76	*Phyla nodiflora*	Olmstead 04-159; WTU	Argentina	-	-	-	pending	pending	pending	pending	pending	pending	pending
07-68	*Phyla nodiflora*	Olmstead 07-65; WTU	Argentina	JX967002	JX966894	JX966948	JX966841	pending	JX966694	JX966741	pending	pending	JX966790
08-194	*Phyla nodiflora*	Lu-Irving 08-4; WTU	Texas	-	-	-	pending	pending	pending	pending	pending	pending	pending
93-105	*Priva cordifolia*	Vos 391; NU	South Africa	JX967003	JX966895	JX966949	JX966842	-	FJ549103	JX966742	-	-	FJ549281
06-39	*Rhaphithamnus venustus*	Stuessy 11855; OS	Chile	JX967004	JX966896	JX966950	JX966843	-	FJ549104	JX966743	-	-	FJ549282
03-101	*Verbena officinalis*	Olmstead 03-156; WTU	cultivated	EF571525	JX966897	JX966951	FJ867561	-	FJ549074	JX966744	-	-	FJ549252
06-79	*Xeroaloysia ovatifolia*	Olmstead 04-184; WTU	Argentina	JX967005	JX966898	JX966952	JX966844	-	FJ549097	JX966745	-	-	JX966791

Appendix 2. Primer sequences.

Locus	Primer	Use	Sequence (5'-3')	Reference/Description
ETS	ETSB	PCR/Sequencing	ATAGAGCGCGTGAGTGGTG	Lu-Irving & Olmstead, 2013
	18SIGS	PCR/Sequencing	GAGACAAGCATATGACTACTGGCAGGATCAACCAG	Baldwin & Markos, 1998
ITS	ITS4	PCR/Sequencing	TCCTCCGCTTATTGATATGC	White et al., 1990
	ITS5	PCR/Sequencing	GGAAGGAGAAGTCGTAACAAGG	R. Olmstead, unpublished
	ITS.LL.F	PCR/Sequencing	ATCCCGCCTGACCTGGGGTCG	Designed for *Lantana-Lippia* clade
PPR 11	320F	PCR/Sequencing	TCTTCTCTTTCTTCACATGGCT	Yuan et al., 2009 b
	850F	Sequencing	GTTAGTTTCAATACTTTGATGAA	Yuan et al., 2009 b
	850R	Sequencing	TTCATCAAAGTATTGAAACTAAC	Yuan et al., 2009 b
	1110F	Sequencing	GATTTGGCWATGGARATTTA	Y-W. Yuan, unpublished
	1300R	Sequencing	TCCARATCTCCYTCCTTACAA	Yuan et al., 2009 b
	1590R	PCR/Sequencing	TAACCGTTCATAAGCACATTGTA	Yuan et al., 2009 b
PPR 81	81.LL.F	PCR/Sequencing	GCAAAGTGCAGAARAGTTGA	Designed for *Lantana-Lippia* clade
	81.LL.R	PCR/Sequencing	CCAATGTGRCTACATGCAGT	Designed for *Lantana-Lippia* clade
	400F	PCR & sequencing	AGT GCR CTT TTW GAT ATG TAY GCA AAG TG	Lu-Irving & Olmstead, 2013
	1630R	PCR & sequencing	TCR ACT GCA CAT GCR TAA TKT TCC AT	Lu-Irving & Olmstead, 2013
	910F	Sequencing	TGG AAA TGG ATG CYT AYA CRT	Lu-Irving & Olmstead, 2013
	1340R	Sequencing	GTR TAR GCA TCC ATT TCC AWC C	Lu-Irving & Olmstead, 2013
PPR 90	313F	PCR/Sequencing	TCTGTTRTTAAACTCGGCTATGATTC	B. Meersman et al., unpublished
	613F	Sequencing	GGRAAGSAAGTTCATGGSTATA	B. Meersman et al., unpublished
	1073R	Sequencing	TATAACCAGYRAGCATRGCATTCCA	B. Meersman et al., unpublished
	1346R	PCR/Sequencing	TATCTTTRCTCTCCATRKTGTGAAA	B. Meersman et al., unpublished
PPR 97	781F	PCR/Sequencing	CTTGTRGATTTGGGTGCWARGTGGTT	B. Meersman et al., unpublished
	1585R	PCR/Sequencing	TTTTTCACATAAGCWGTYACAAGAAT	B. Meersman et al., unpublished
PPR 123	123.LL.F	PCR/Sequencing	GTGCCTGGGGATTTGGTTCTGTA	Designed for *Lantana-Lippia* clade
	LL.825F	Sequencing	GTGTTTGGAAAGGCTAAGC	Lu-Irving & Olmstead, 2013
	1030R	Sequencing	GCCCATAMACATCKATCATTAT	Yuan et al., 2009 b
	1890R	PCR/Sequencing	AGACTCAGCATCTGRAAATGAAC	Yuan et al., 2009 b
	550F	PCR & sequencing	CAC GGR CTG TTC GAC GAA ATG CG	Yuan et al., 2009 b
	1370F	Sequencing	AAG TTA GAT AGA GCA GCC ATG C	Yuan et al., 2009 b
	1620R	Sequencing	AAG ACC GTT ATR TCC TTG ACC TC	Yuan et al., 2009 b
***trn*T-L**	tabA	PCR & sequencing	CAT TAC AAA TGC GAT GCT CT	Taberlet et al., 1991

	tabB	PCR & sequencing	TCT ACC GAT TTC GCC ATA TC	Taberlet et al., 1991
	TL-1R	Sequencing	TAT AGC GAT CTG GGA TTT CG	Lu-Irving & Olmstead, 2013
	TL-2F	Sequencing	GTT TCT CTT ACT GCC ATT TTC CC	Lu-Irving & Olmstead, 2013
***rpl*32-*trn*L**	trnL(UAG)	PCR & sequencing	CTG CTT CCT AAG AGC AGC GT	Shaw et al., 2007
	*rpl*32	PCR & sequencing	CAG TTC CAA AAA AAC GTA CTT C	Shaw et al., 2007
	L32-1F	Sequencing	CCC ATC AAC CTA TTT GTT A	Lu-Irving & Olmstead, 2013
	L32-2R	Sequencing	CCC AAA AAT CAA TTT GAT CRT TGA C	Lu-Irving & Olmstead, 2013
***trn*Q-*rps*16**	*trn*Q	PCR & sequencing	GCG TGG CCA AGY GGT AAG GC	Shaw et al., 2007
	*rps*16	PCR & sequencing	GTT GCT TTY TAC CAC ATC GTT T	Shaw et al., 2007
	400F	Sequencing	GAT GGT ATG TAG CGT TCT ATT TCA ATG	Lu-Irving & Olmstead, 2013
	1000F	Sequencing	CTA TCC AAA CAG GAA CCA CCC AA	Lu-Irving & Olmstead, 2013

Appendix 3A. Supplementary material to Chapter I.

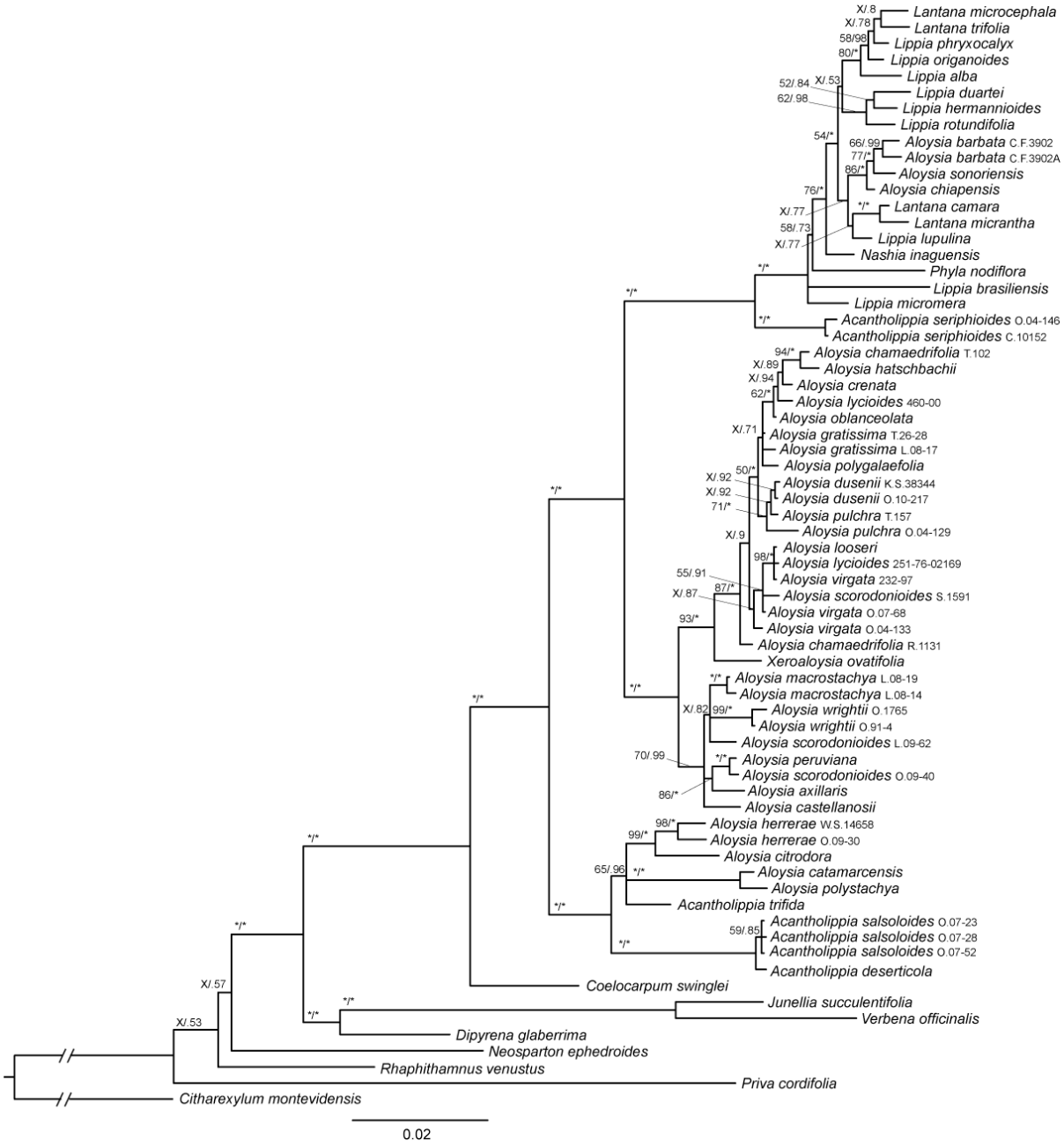
Results of SH tests (P values) on trees inferred from different data sets.

	Chloroplast tree	**ETS tree**	**PPR 11 tree**	**PPR 81 tree**	**PPR 123 tree**	**Combined tree**
Chloroplast data	(best)	0.000	0.000	0.000	0.000	0.003
ETS data	0.000	(best)	0.000	0.000	0.000	0.118
PPR 11 data	0.000	0.000	(best)	0.000	0.000	0.09
PPR 81 data	0.000	0.000	0.000	(best)	0.000	0.000
PPR 123 data	0.000	0.000	0.000	0.000	(best)	0.000
Combined data	0.000	0.000	0.000	0.000	0.000	(best)

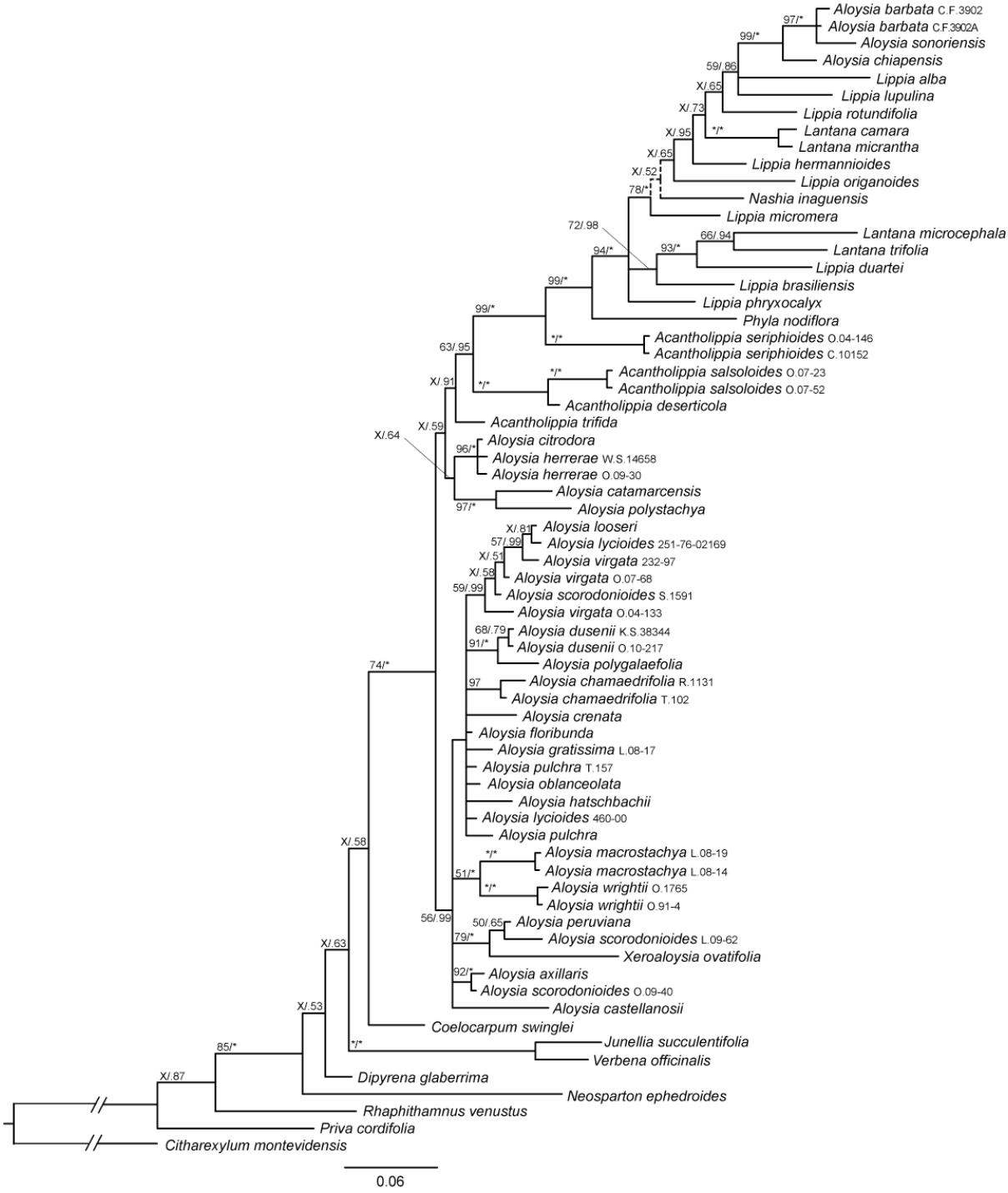
Appendix 3B. Supplementary material to Chapter II.

Status of species names used in this study, according to different taxonomic treatments of *Aloysia*.

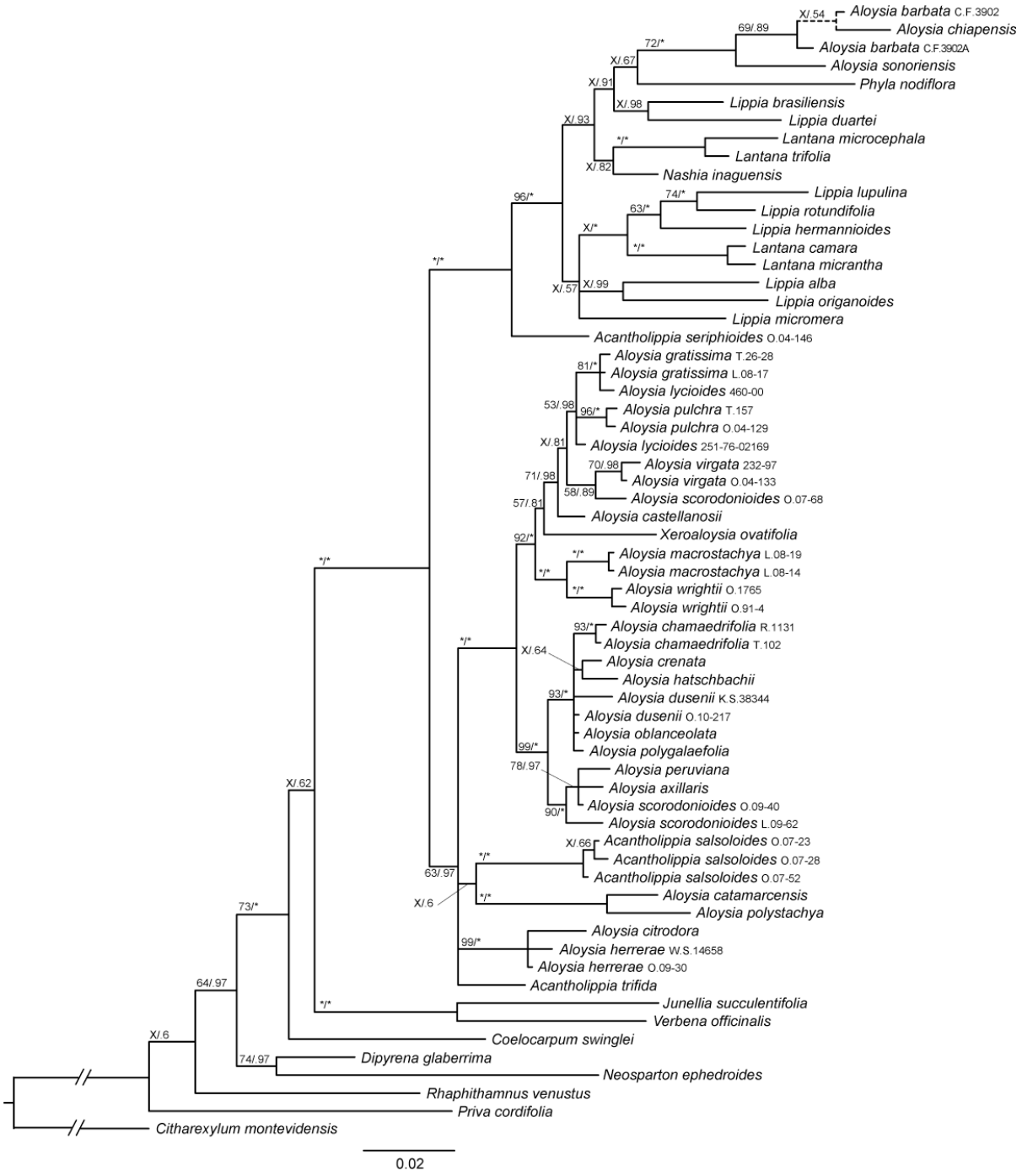
Species and authority	**O'Leary et al., unpublished**	**Siedo, 2006**
Aloysia axillaris J.R.I.Wood	*A. scorodonioides*	[not included]
Aloysia barbata (Brandegee) Moldenke	accepted, not treated	accepted
Aloysia castellanosii Moldenke	accepted	accepted
Aloysia catamarcensis Moldenke	accepted	accepted
Aloysia chamaedryfolia Cham.	accepted	accepted
Aloysia chiapensis Moldenke	accepted, not treated	accepted
Aloysia citrodora Palau	accepted	accepted
Aloysia crenata Moldenke	accepted	accepted
Aloysia dusenii Moldenke	accepted	accepted
Aloysia gratissima (Gillies & Hook.) Tronc.	accepted	accepted
Aloysia hatschbachii Moldenke	accepted	accepted
Aloysia herrerae Moldenke	accepted	accepted
Aloysia looseri Moldenke	*A. gratissima*	*A. virgata*
Aloysia lycioides Cham.	*A. gratissima*	accepted
Aloysia macrostachya (Torr.) Moldenke	accepted, not treated	accepted
Aloysia oblanceolata Moldenke	accepted	accepted
Aloysia peruviana (Turcz.) Moldenke	accepted	accepted
Aloysia polygalifolia Cham.	accepted	accepted
Aloysia polystachya (Griseb.) Moldenke	accepted	accepted
Aloysia pulchra (Briq.) Moldenke	accepted	*A. lycioides*
Aloysia scorodonioides (Kunth) Cham.	accepted	accepted
Aloysia sonoriensis Moldenke	accepted, not treated	accepted
Aloysia virgata (Ruiz & Pav.) Juss.	accepted	accepted
Aloysia wrightii A.Heller	accepted, not treated	accepted



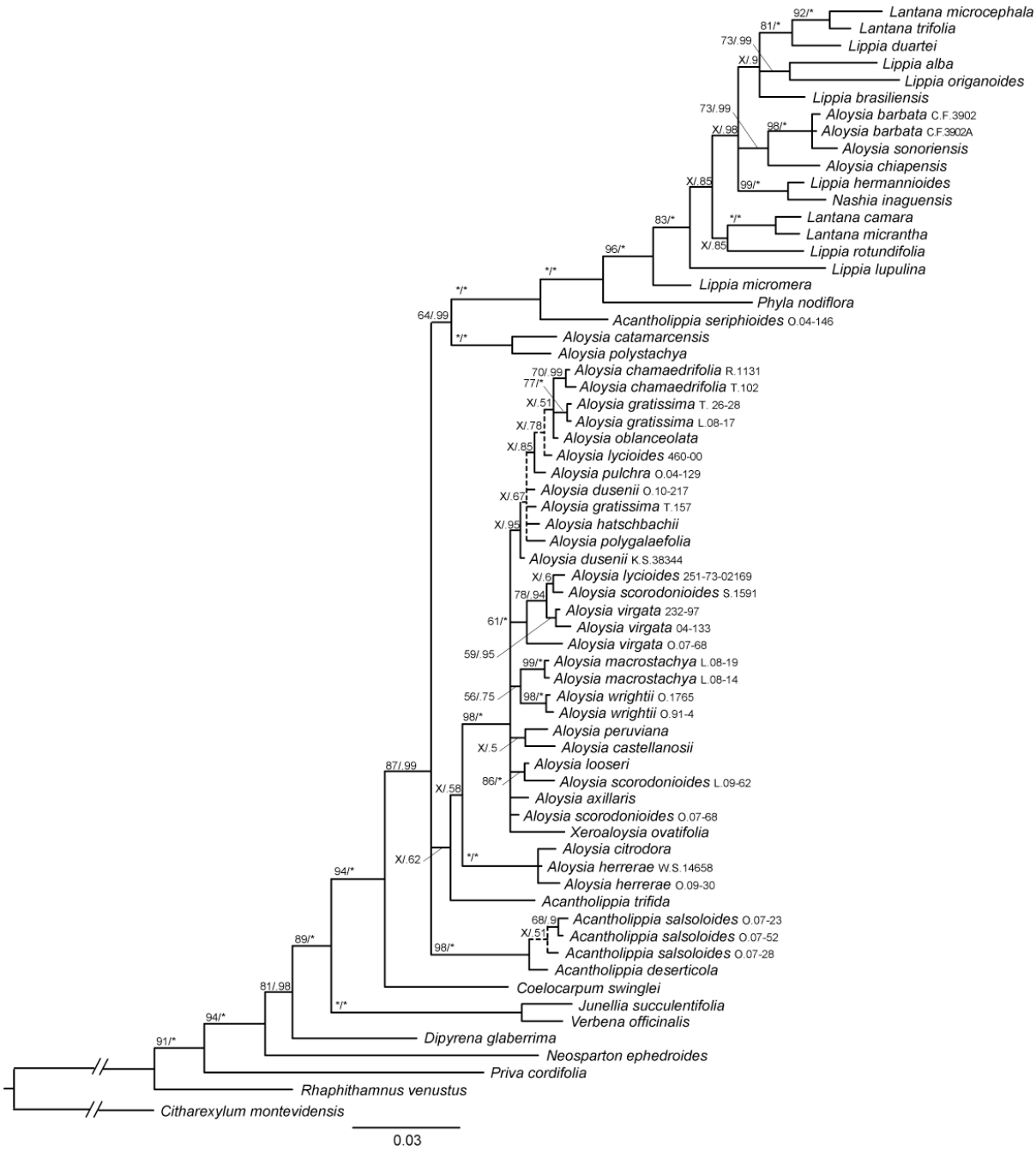
Phylogeny inferred from 4,266 aligned positions of DNA sequence data from 3 chloroplast loci in combination. Topology inferred by ML and Bayesian analyses, branch lengths inferred by Bayesian analysis. Branches are labeled with ML bootstrap values/Bayesian posterior probabilities greater than 50%/0.50. Stars (*) denote 100% support, Xs denote bootstrap values below 50%.



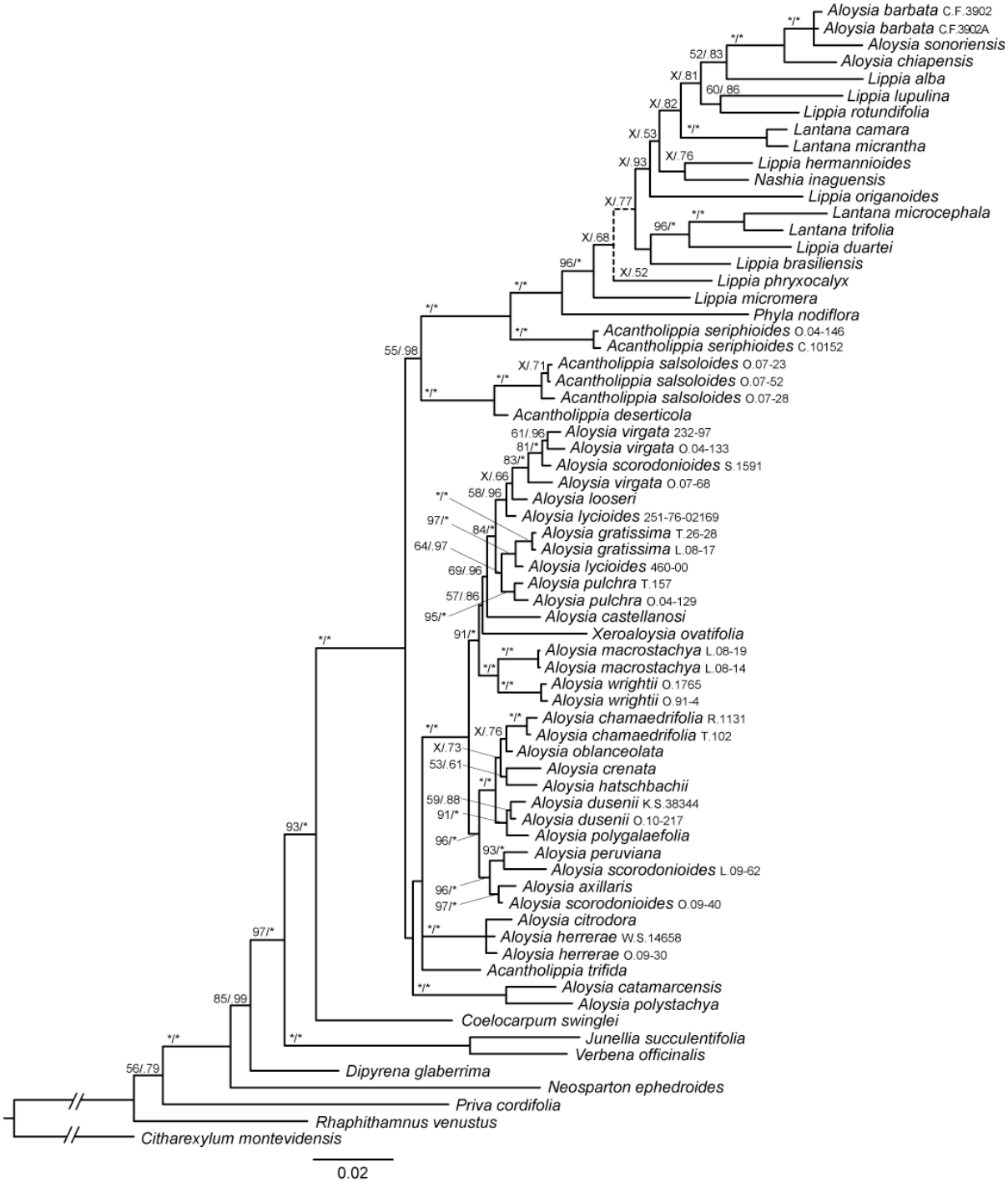
Phylogeny inferred from DNA sequence from nuclear region ETS (514 aligned positions). Topology inferred by ML and Bayesian analyses, branch lengths inferred by Bayesian analysis. Branches are labeled with ML bootstrap values/Bayesian posterior probabilities greater than 50%/0.50. Stars (*) denote 100% support, Xs denote bootstrap values below 50%. Dashed lines indicate disagreement between ML and Bayesian analyses; topology inferred from Bayesian analysis is shown.



Phylogeny inferred from DNA sequence from nuclear region PPR 81 (1,221 aligned positions). Topology inferred by ML and Bayesian analyses, branch lengths inferred by Bayesian analysis. Branches are labeled with ML bootstrap values/Bayesian posterior probabilities greater than 50%/0.50. Stars (*) denote 100% support, Xs denote bootstrap values below 50%. Dashed lines indicate disagreement between ML and Bayesian analyses; topology inferred from Bayesian analysis is shown.

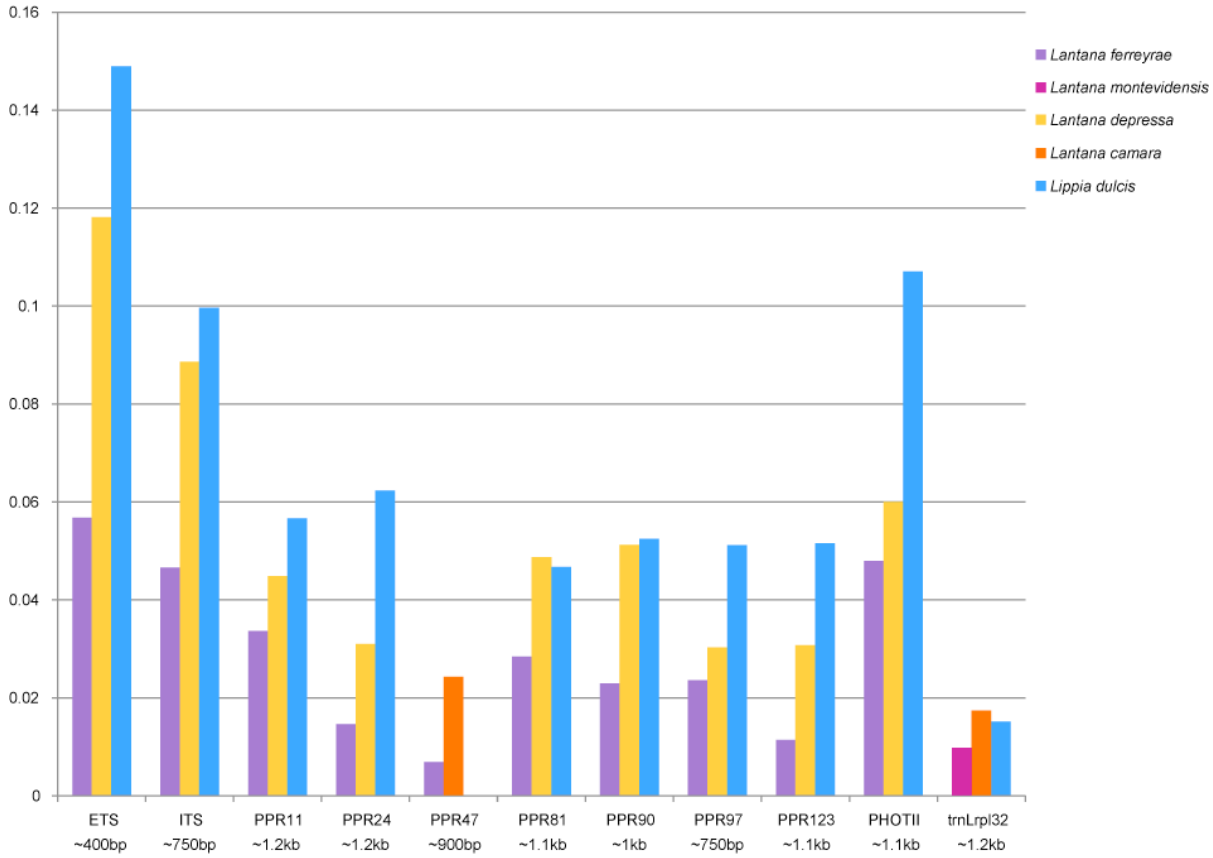


Phylogeny inferred from DNA sequence from nuclear region PPR 123 (1,325 aligned positions). Topology inferred by ML and Bayesian analyses, branch lengths inferred by Bayesian analysis. Branches are labeled with ML bootstrap values/Bayesian posterior probabilities greater than 50%/0.50. Stars (*) denote 100% support, Xs denote bootstrap values below 50%. Dashed lines indicate disagreement between ML and Bayesian analyses; topology inferred from Bayesian analysis is shown.

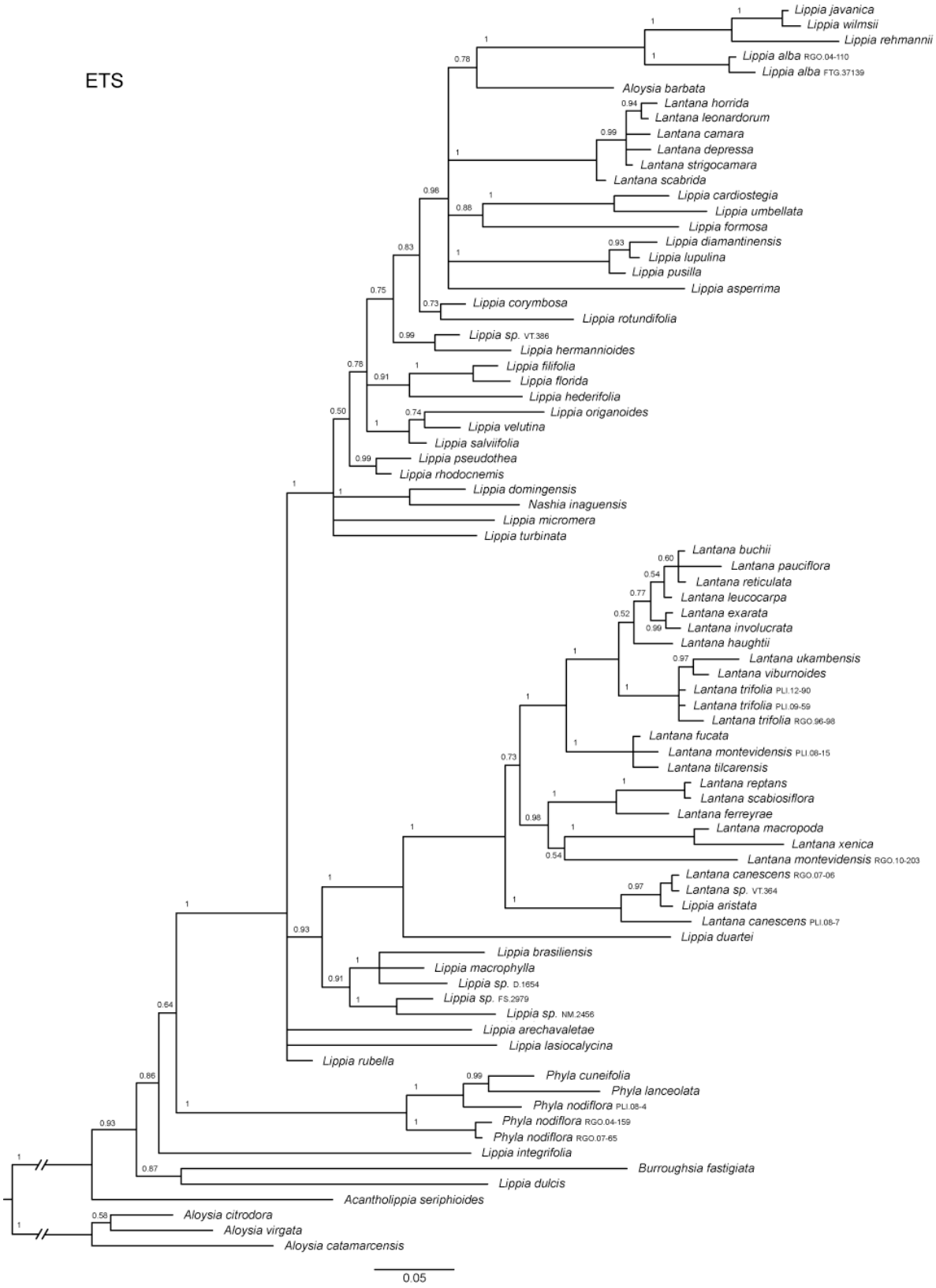


Phylogeny inferred from 3,060 aligned positions of DNA sequence data from 3 nuclear loci in combination. Topology inferred by ML and Bayesian analyses, branch lengths inferred by Bayesian analysis. Branches are labeled with ML bootstrap values/Bayesian posterior probabilities greater than 50%/0.50. Stars (*) denote 100% support, Xs denote bootstrap values below 50%. Dashed lines indicate disagreement between ML and Bayesian analyses; topology inferred from Bayesian analysis is shown.

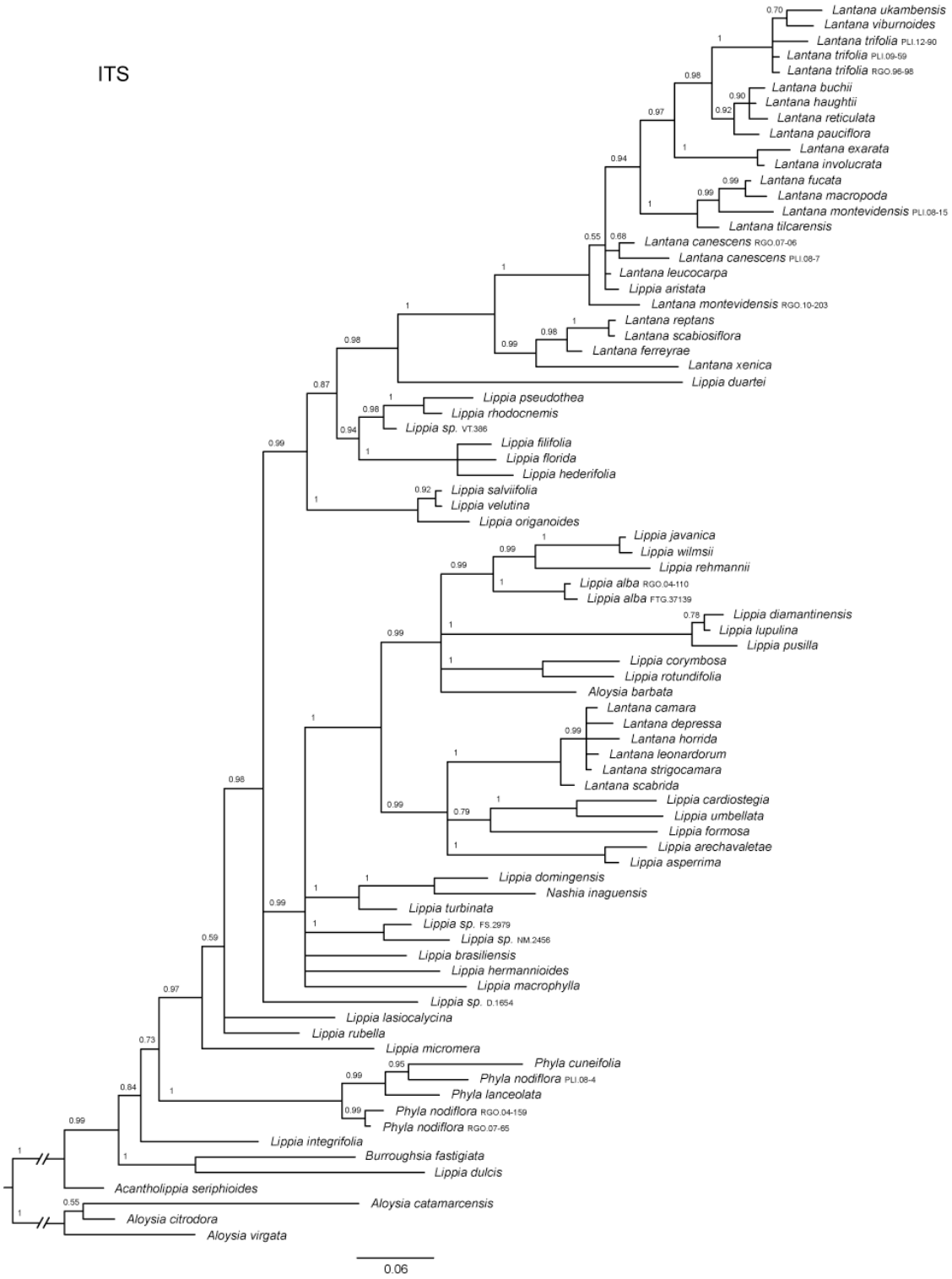
Appendix 3C. Supplementary material to Chapter III.



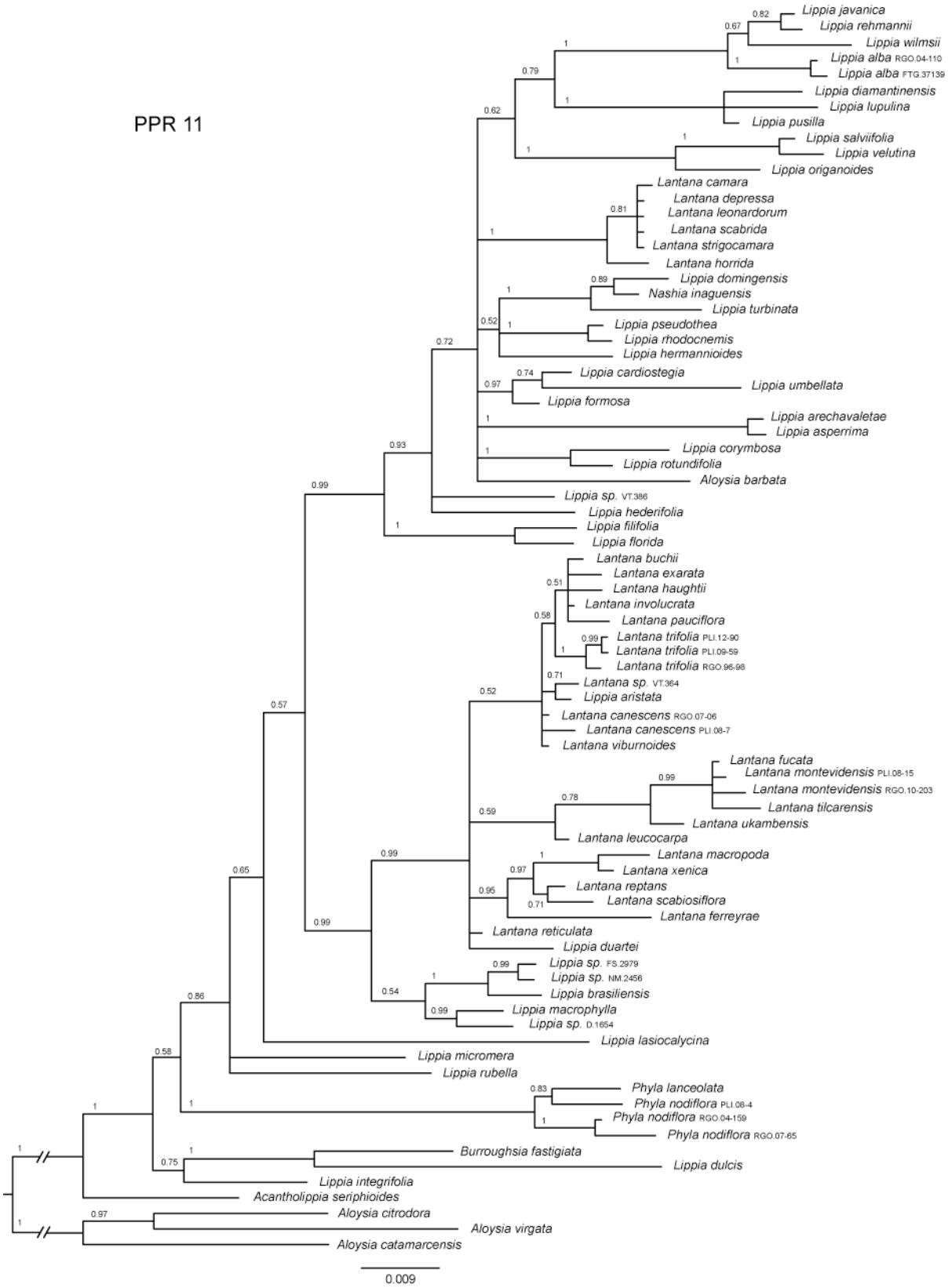
Pairwise distances between representative members of the *Lantana-Lippia* clade, used to gauge variability of loci to select data sources for phylogenetic analysis.



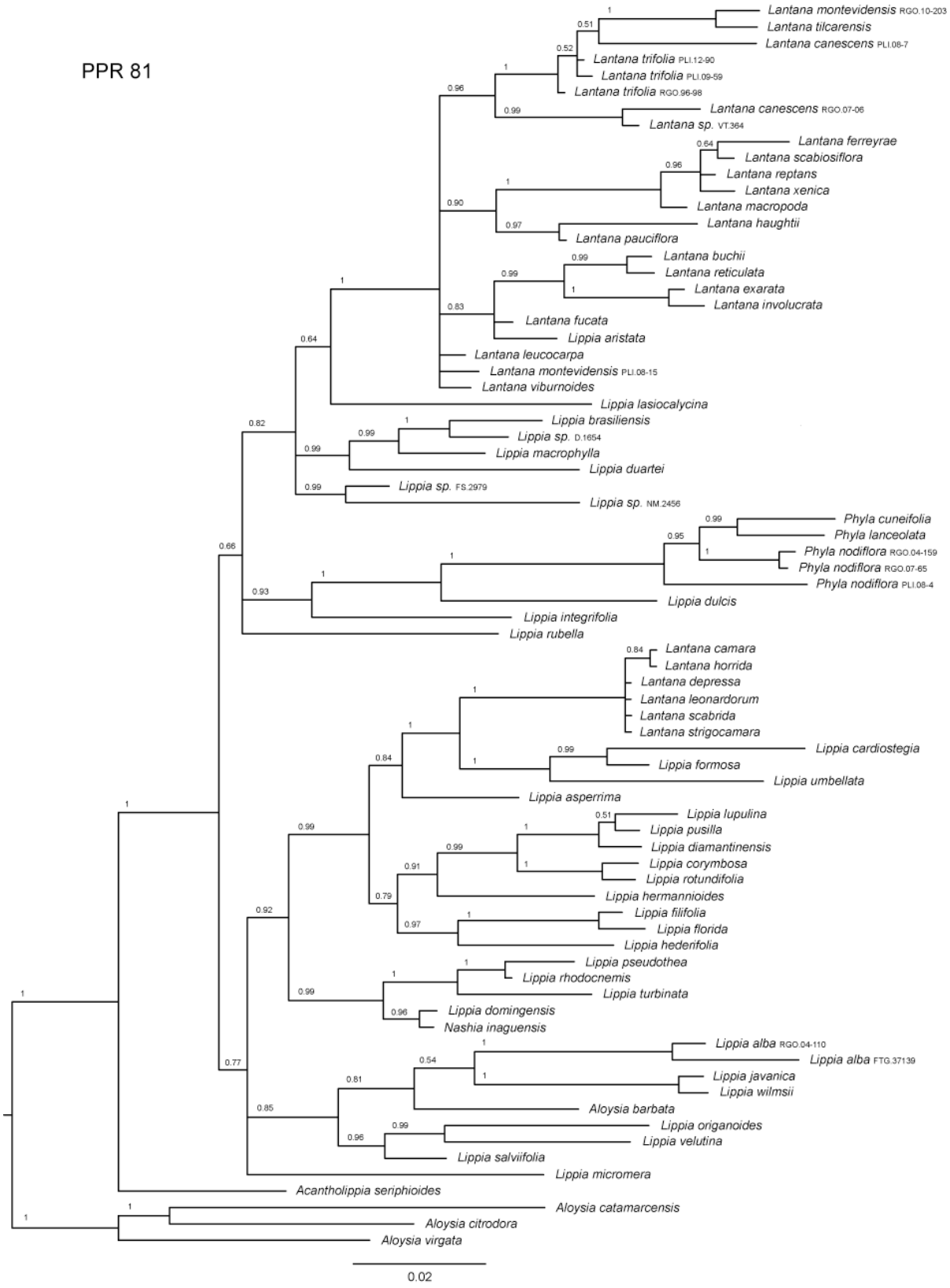
Gene tree inferred from 480 aligned positions from ETS. Posterior probability values greater than 0.5 are shown.



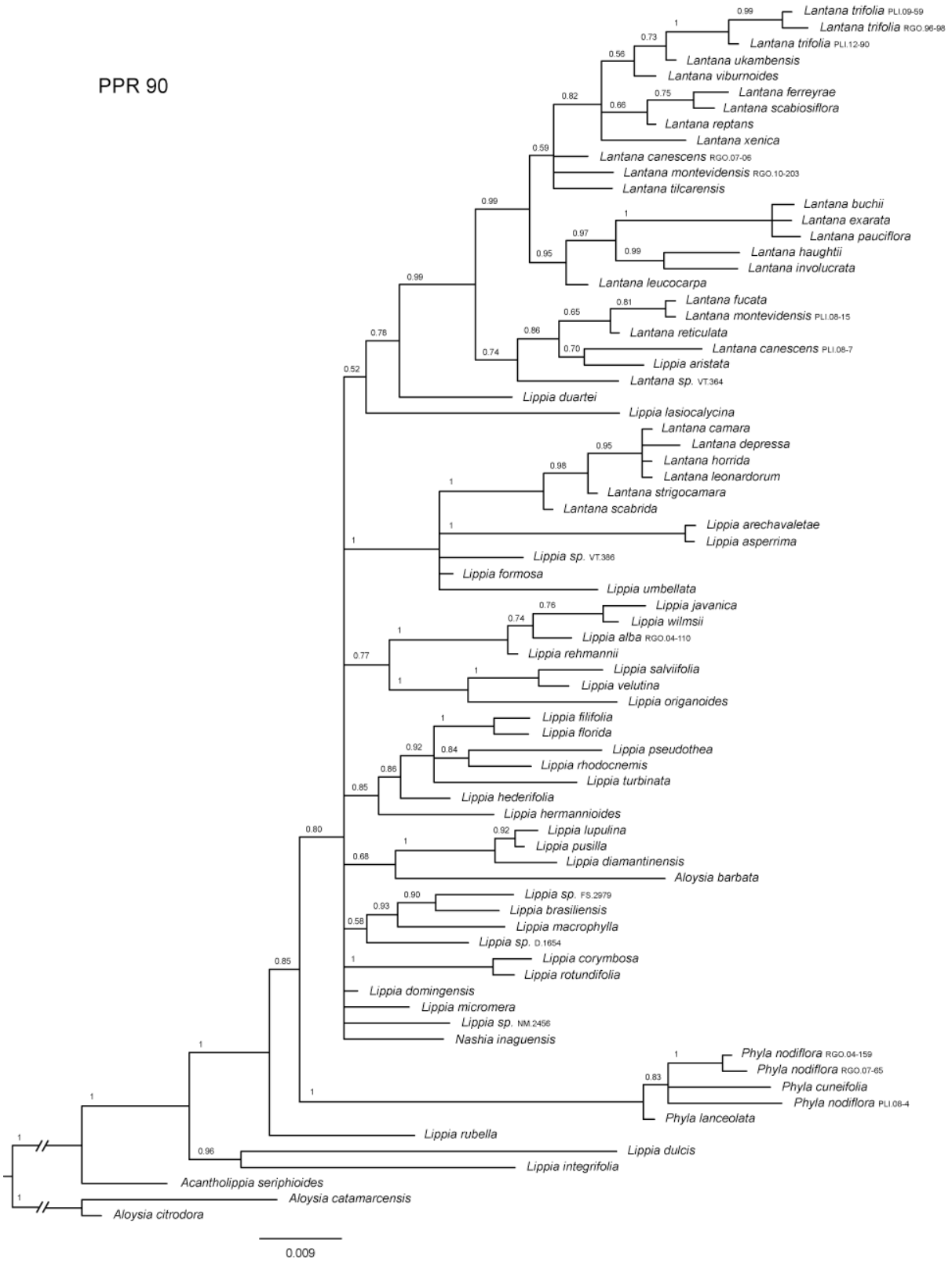
Gene tree inferred from 762 aligned positions from ITS. Posterior probability values greater than 0.5 are shown.



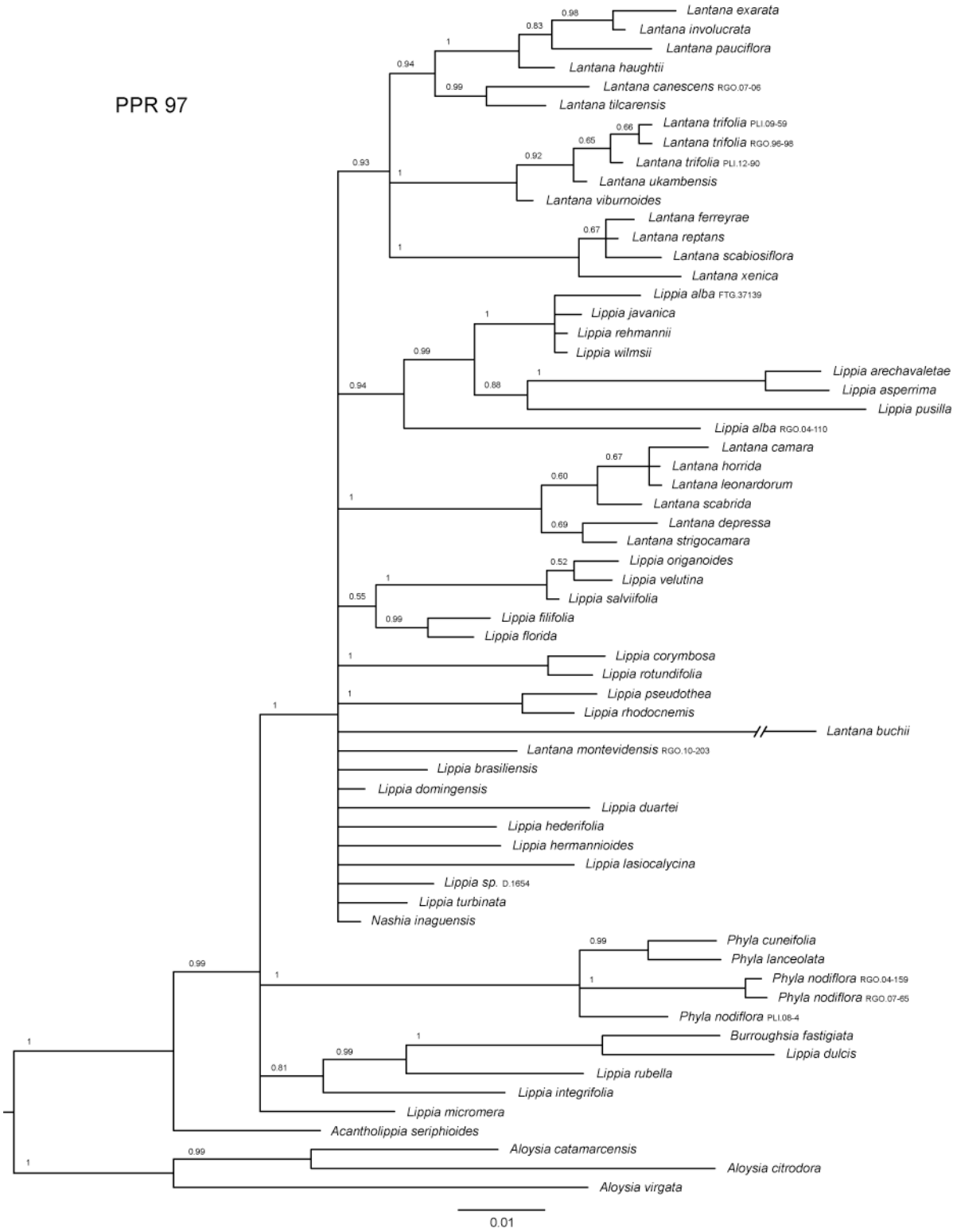
Gene tree inferred from 1,277 aligned positions from PPR 11. Posterior probability values greater than 0.5 are shown.



Gene tree inferred from 1,162 aligned positions from PPR 81. Posterior probability values greater than 0.5 are shown.

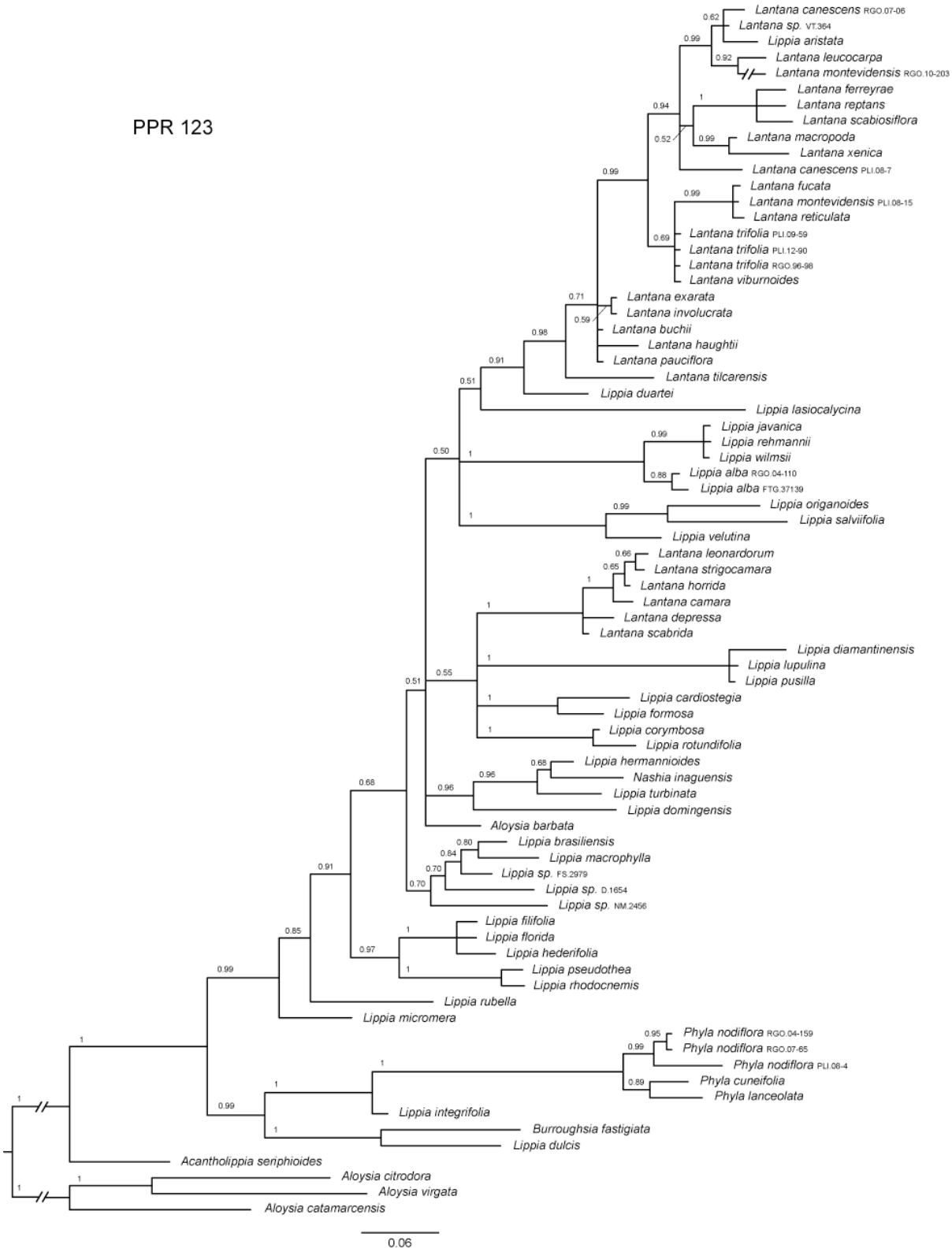


Gene tree inferred from 986 aligned positions from PPR 90. Posterior probability values greater than 0.5 are shown.

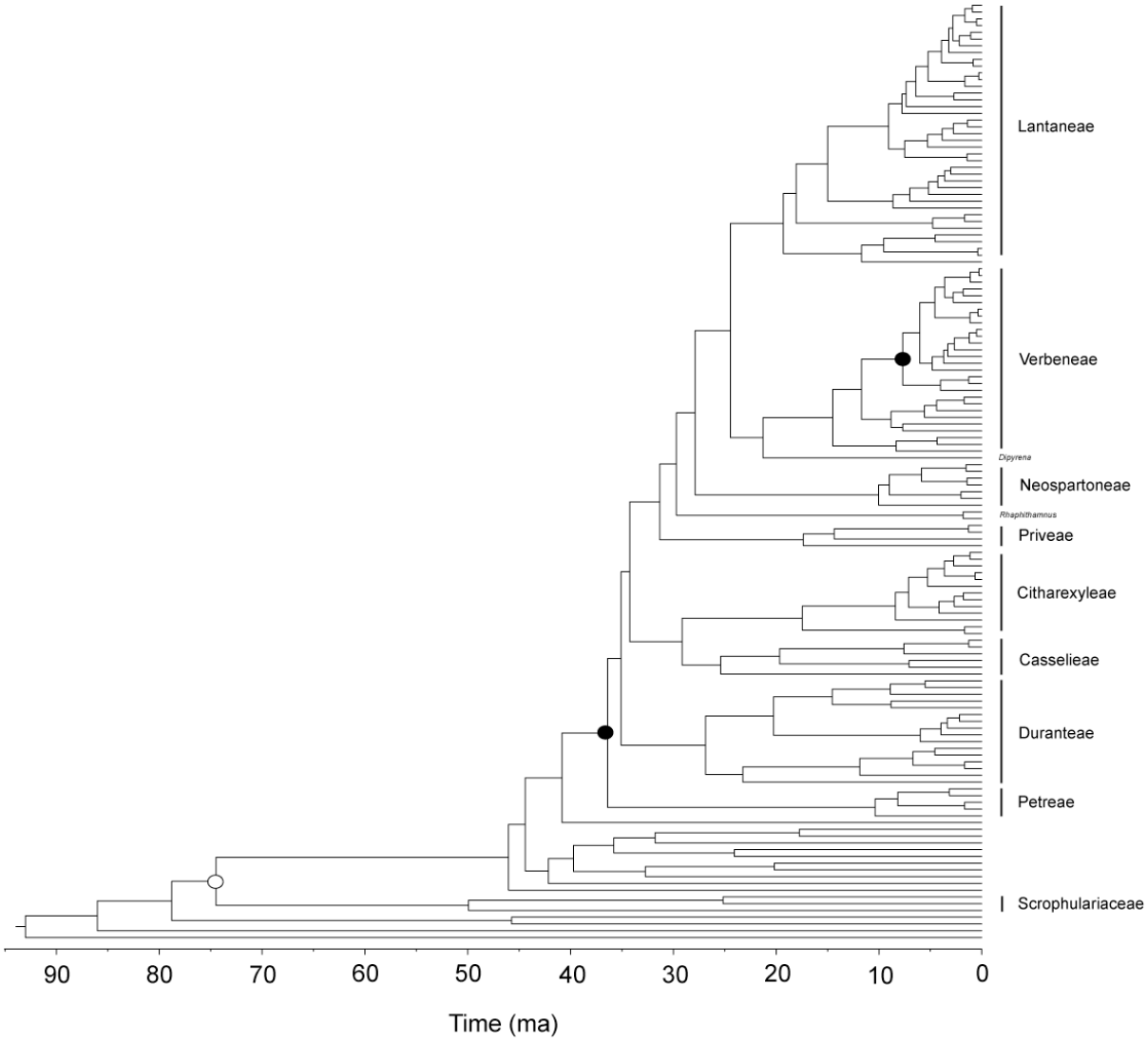


Gene tree inferred from 747 aligned positions from PPR 97. Posterior probability values greater than 0.5 are shown.

PPR 123



Gene tree inferred from 1,122 aligned positions from PPR 123. Posterior probability values greater than 0.5 are shown.



Time-calibrated tree inferred for Verbenaceae, using chloroplast data from Marx et al., 2010. The three calibration points used are indicated with circles: