MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 2901–2915 doi: 10.1111/j.1365-294X.2011.05151.x

Recent vicariance and the origin of the rare, edaphically specialized Sandhills lily, *Lilium pyrophilum* (Liliaceae): evidence from phylogenetic and coalescent analyses

NORMAN A. DOUGLAS,* WADE A. WALL,* QIU-YUN (JENNY) XIANG,* WILLIAM A. HOFFMANN,* THOMAS R. WENTWORTH,* JANET B. GRAY† and MATTHEW G. HOHMANN‡

**Department of Plant Biology, PO Box 7612, North Carolina State University, Raleigh, NC 27695, USA, †Directorate of Public Works, Endangered Species Branch, United States Army, Fort Bragg, NC 28310, USA, ‡US Army Corps of Engineers, Engineer Research and Development Center, PO Box 9005, Champaign, IL 618262, USA*

Abstract

Establishing the phylogenetic and demographic history of rare plants improves our understanding of mechanisms that have led to their origin and can lead to valuable insights that inform conservation decisions. The Atlantic coastal plain of eastern North America harbours many rare and endemic species, yet their evolution is poorly understood. We investigate the rare Sandhills lily (*Lilium pyrophilum*), which is endemic to seepage slopes in a restricted area of the Atlantic coastal plain of eastern North America. Using phylogenetic evidence from chloroplast, nuclear internal transcribed spacer and two low-copy nuclear genes, we establish a close relationship between *L. pyrophilum* and the widespread Turk's cap lily, *L. superbum*. Isolation-with-migration and coalescent simulation analyses suggest that (i) the divergence between these two species falls in the late Pleistocene or Holocene and almost certainly post-dates the establishment of the edaphic conditions to which *L. pyrophilum* is presently restricted, (ii) vicariance is responsible for the present range disjunction between the two species, and that subsequent gene flow has been asymmetrical and (iii) *L. pyrophilum* harbours substantial genetic diversity in spite of its present rarity. This system provides an example of the role of edaphic specialization and climate change in promoting diversification in the Atlantic coastal plain.

Keywords: coalescence, divergence, edaphic, *Lilium*, Pleistocene, rarity

Received 18 November 2010; revision received 28 April 2011; accepted 5 May 2011

Introduction

Molecular studies of rare plant taxa usually aim to quantify the level and patterns of genetic diversity in a particular species (Karron 1987; Hamrick & Godt 1990; Ellstrand & Elam 1993; Gitzendanner & Soltis 2000). Phylogeographic studies, on the other hand, often focus on widespread species and try to discern continental-scale patterns (Taberlet *et al.* 1998; Brunsfeld *et al.* 2001; Soltis *et al.* 2006). However, the tools of phylogeography, particularly coalescent-based analyses that provide information about the age and historical demography of species (Knowles 2009), have only rarely been applied to investigate the history of rare species (Raduski *et al.* 2010; Whittall *et al.* 2010).

Of the 'seven forms of rarity' (Rabinowitz 1981), the most extreme describes taxa that have a narrow geographic range, require specific habitats and maintain only small local populations. Many edaphic endemics (plants restricted to soils with unusual physical or chemical properties) belong to this category. While the textbook examples of edaphic endemic plants are restricted to serpentine, various substrates support edaphic endemics, including guano, alkali, salt, and gypsum deposits, limestone, chalk, and granite outcrops, oligotrophic bogs and deep porous sands (Ornduff 1965; Axelrod 1972; Parsons 1976; Kruckeberg &

Correspondence: Norman A. Douglas, Fax: (919) 515 3436; E-mail: norman_douglas@ncsu.edu

Rabinowitz 1985; Kruckeberg 1986; Williamson & Bazeer 1997). Many aspects of the origin of edaphic endemic species are poorly understood (Rajakaruna 2004). For instance, such species often occur in close geographic proximity to their progenitor lineages (e.g. Baldwin 2005), yet it is not usually known whether or how strongly gene flow is interrupted. While taxa displaying edaphic endemic syndromes often show reduced genetic diversity compared with their close relatives (Godt & Hamrick 1993; Baskauf *et al.* 1994; Ayres & Ryan 1999; but see Raduski *et al.* 2010), this may reflect genetic drift due to lower population sizes or the effects of selection. Strong selection imposed by edaphically challenging soils could be sufficient to foster population divergence (Nosil *et al.* 2009; Freeland *et al.* 2010). Some edaphic endemics may represent vicariant populations isolated in narrow parts of formerly wider ranges and niches of their progenitors (e.g. Crawford *et al.* 1985), which may themselves be able to grow on the unusual substrate without being restricted to it.

Edaphic specialists (especially in bog and sand habitats, Sorrie & Weakley 2001) are an important component of the endemic-rich flora of the coastal plain of eastern North America. Few coastal plain endemics have been the subject of molecular analyses. Sand dune habitats in Florida apparently served as Pleistocene refugia for the genera *Dicerandra* and *Conradina* (Edwards *et al.* 2006; Oliveira *et al.* 2007), and in general, Florida has been proposed as a major Pleistocene refugium for many taxa in eastern North America (Soltis *et al.* 2006). Yet, recent phylogeographic work indicates that some coastal plain endemic species likely persisted in northerly latitudes throughout the Pleistocene. For instance, the Atlantic coastal plain endemic Pyxie Moss, *Pyxidanthera* (Diapensiaceae), shows long-term range stasis (Wall *et al.* 2010).

The Fall-Line Sandhills of North and South Carolina (which occur at the western boundary of the coastal plain) provide one of the clearest examples of the edaphic contribution to the botanical diversity of the Atlantic coastal plain. This region is comprised of rolling hills of open, fire-maintained longleaf pine (*Pinus palustris*) savanna dissected by numerous blackwater streams and wetlands, providing a diverse matrix of habitats that support at least eight endemic plants (and numerous near-endemics, Sorrie & Weakley 2001). In the core of the Sandhills region in southern North Carolina, the uppermost deposit is the Pinehurst formation, which is characterized by loose coarse-grained sands found along ridgetops. This formation was deposited in a tidal environment (J. Nickerson, North Carolina Geological Survey, personal communication) and may date to the Eocene (Cabe *et al.* 1992). Below the Pinehurst formation (and exposed along drainages and slopes throughout the region) lies the Cretaceous Middendorf formation, which is of deltaic origin and thus has more abundant clays (Sohl & Owens 1991). At the interface between these (and similar formations in the Carolinas and southeastern Virginia) occur Sandhills seep and streamhead pocosin ecotone communities. When kept open by frequent fires encroaching from the surrounding xeric pine savannas, these wetlands can support extremely high local species richness, among the highest values ever recorded in North America (>102 species per 1/100 ha, Schafale & Weakley 1990). The age of the formations implies that endemic species have potentially had a very long time to adapt to the unusual edaphic conditions.

In this study, we consider the Sandhills lily, *Lilium pyrophilum* (Liliaceae), a striking endemic of the Sandhills in the Carolinas and southeastern Virginia. Formally described only recently (Skinner & Sorrie 2002), specimens of this species were previously identified in herbaria as any of three similar species in the region (*L. superbum*, *L. michauxii* or *L. iridollae*) that share the distinctive 'Turk's cap' morphology, in which flowers are pendent with the tepals reflexed upward. Skinner & Sorrie (2002) identified three specific plant communities (Schafale & Weakley 1990; Sorrie *et al.* 2006) that support *L. pyrophilum*, including Sandhills seep and streamhead pocosin ecotones. The third, small stream swamps are affected by frequent flooding events in addition to seepage and rarely support *L. pyrophilum* (Sorrie *et al.* 2006).

Lilium pyrophilum is a very rare species. There are fewer than 75 historical and extant locations in North and South Carolina, and Virginia (North Carolina Natural Heritage Program 2007), and between 2007 and 2009, a survey of all known populations located <500 stems across 35 populations (W. Wall, unpublished data). Approximately half of the extant populations and a quarter of the individuals occur on Fort Bragg Military Reservation in North Carolina, where prescribed and ordnance-ignited fires maintain appropriate habitat.

In describing *L. pyrophilum* (Skinner & Sorrie 2002), the authors outlined three phylogenetic hypotheses concerning the origin of the species. First, they speculated that *L. pyrophilum* may represent a peripheral isolate of the Turk's cap lily, *L. superbum*, which it most resembles morphologically (albeit with significant differences, Skinner & Sorrie 2002). *Lilium superbum* is distributed throughout much of eastern North America (Fig. 1), and in contrast to the edaphically specialized *L. pyrophilum*, it is a generalist, occurring in rich woods and oligotrophic wetlands from high elevation to sea level. Especially in northern parts of its range (e.g. the Pine Barrens of New Jersey), it can be found in saturated

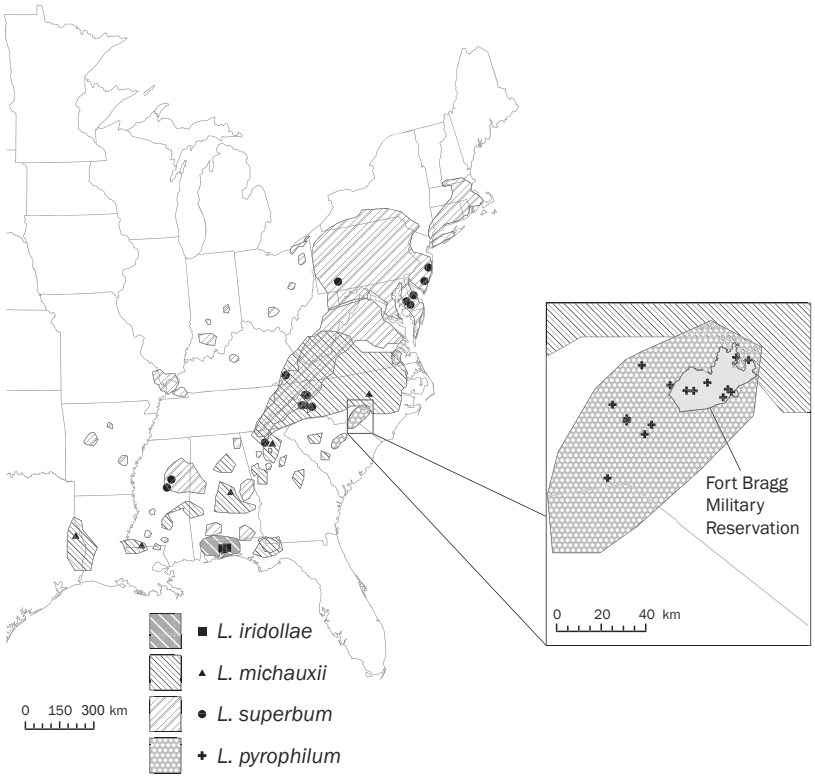

Fig. 1 Distribution of populations included in this study and geographic ranges of the four focal species.

sandy habitats not unlike those preferred by *L. pyrophilum*, but it is not restricted to them. However, it is essentially absent from the Piedmont and Atlantic coastal plain from the Carolinas southward. Thus, it is disjunct from *L. pyrophilum* by at least 150 km everywhere except in southeastern Virginia (Fig. 1) where the coastal plain narrows.

Second, they speculated that *L. pyrophilum* may represent a hybrid species, with the widespread Carolina lily (*L. michauxii*) and *L. superbum* as progenitors. Homoploid hybrid speciation has been implicated in the origin of other edaphic specialists, e.g. *Helianthus paradoxus* (Rieseberg *et al.* 1990) and Hawaiian *Scaevola* (Howarth & Baum 2005). Of the three potentially related species, *L. pyrophilum* resembles *L. michauxii* least, differing in leaf shape and producing fragrant flowers (Skinner 2002). While the range of *L. michauxii* does overlap the range of *L. pyrophilum* (Fig. 1), they occur in contrasting habitats, with *L. michauxii* favouring much drier sites. Notably, *L. michauxii* and *L. superbum* co-occur throughout much of their ranges (Fig. 1), yet natural hybrids are apparently rare (Skinner 2002).

Finally, Skinner and Sorrie suggested the possibility that *L. pyrophilum* may represent a disjunct population of the Pot-o'-gold or Panhandle lily (*L. iridollae*), a narrow endemic of wet pine savannas in northwestern Florida (where it is listed as endangered) and adjacent Alabama. This hypothesis emphasizes similar habitat requirements of the two species, but downplays consistent morphological differences (e.g. details of rhizome structure, Skinner 2002; Skinner & Sorrie 2002) and a range separation of over 700 km (Fig. 1).

In this study, we report the results of a molecular study focused on *L. pyrophilum* and its close relatives. First, we investigated the phylogeny of the eastern pendent species of *Lilium* to address whether *L. pyrophilum* represents a peripheral isolate of *L. superbum*, a hybrid between *L. superbum* and *L. michauxii*, or a disjunct population of *L. iridollae*. Second, we analysed the distribution of genetic variation within and among the taxa thought to be closely related to *L. pyrophilum* and used coalescent-based methods to explicitly evaluate the possible timing of the divergence of *L. pyrophilum*. Our results are interpreted in the context of the evolution of rare, edaphically specialized lineages in the Atlantic coastal plain.

Materials and methods

Sampling and molecular data

Samples were obtained from 50 populations spanning the geographic range of each of the four focal species (Fig. 1). We also sampled two populations of *Lilium*

canadense, another pendant species that lacks the Turk's cap morphology. Sampling information is provided in Table S1 (Supporting Information). Populations were located in the field based on documented occurrences from herbarium specimens, element occurrence records from state Natural Heritage Programs and communication with local botanists. We endeavoured to sample a similar number of populations of *L. superbum* and *L. michauxii* spanning the geographic range of each species. Our sampling of the rare *L. iridollae* was limited to two populations. In general, one individual was taken to represent each population. Genomic DNA was isolated from fresh or frozen leaves, using the CTAB method (Doyle & Doyle 1987). Nuclear ribosomal internal transcribed spacer ('ITS') sequences were obtained with primers ITS4 and ITS5a (White *et al.* 1990; Stanford *et al.* 2000). This locus was sequenced to facilitate comparison with abundant existing data available in GenBank to determine whether the species in this study form a monophyletic group. We screened eight chloroplast markers from Shaw *et al.* (2007); of these, three (the *atpI-atpH*, *psbD-trnT* and *rpl32-trnL* intergenic spacers) consistently amplified and contained variable sites. As the chloroplast behaves as a single nonrecombining locus, sequences of these three regions were concatenated, and this marker is hereafter referred to as 'CP'.

We developed single-copy nuclear markers for *Lilium*. In general, we screened EST or complete CDS sequences from *Lilium* against the *Oryza sativa* genomic sequence at GenBank using SPIDEY (Wheelan *et al.* 2001) with the 'divergent sequences' and 'use large intron sizes' options. Candidate sequences were downloaded and manually aligned in Se-Al (Rambaut 1996) using amino acid translations. Homologous sequences from GenBank were incorporated into the alignments. When we were confident of the positions of the introns in the rice genome, we then designed primers using Primer3 (Rozen & Skaletsky 2000), which were screened against DNA extracted from *L. longiflorum* and an Asiatic hybrid cultivar (which served as positive controls because nearly all of our candidate regions were based on sequences from these cultivated lilies) and the four taxa in our study. We were able to obtain single amplicons for relatively few of these regions even after extensive PCR optimization; it was often the case that primers would amplify nontarget regions or that introns would be small, invariant or missing entirely. The closely related *L. canadense* has a phenomenally large genome (1C = 47.90 pg, 46.9 Gbp; Zonneveld *et al.* 2005; Peruzzi *et al.* 2009), which may have contributed to the difficulty we encountered in obtaining single-copy nuclear sequences. However, we were able to design primers that amplified two novel regions. The first includes two introns between exons 8 and 10 of the *L. longiflorum* alkaline phytase gene, *LlAlp* ('AP', primers: AP8f, 5′-TCTCCTTGGGCTCTTTCTTG and AP10r, 5′-GAAAACCTCAAATGGGCAGAG), which is involved in phytic acid metabolism (Mehta *et al.* 2006). While GenBank contains sequences for two isoforms of this gene, our PCR experiments are consistent with these representing splice variants of a single locus. The second region corresponds to a region between exons 5 and 10 of the *AKT1*-like potassium channel *LilKT1* ('AKT', primers: AKT5f, 5′-AGAGACTCTTGATGCACT TCCTAAA and AKT10r, 5′-AAGAGAACAACA-CAACTTTCATTCC). This locus was more difficult to amplify, and we were unable to generate sequences for *L. iridollae*. Primers and PCR conditions for ITS and the chloroplast loci followed White *et al.* (1990) and Shaw *et al.* (2007). For AP and AKT, PCR contained 2.5 μL 10× PCR buffer, 1% BSA, 200 μM dNTPs, 2.5 mM $MgCl_2$, 4 μM of each primer and 0.5 U Taq DNA polymerase. Cycling conditions were 95 °C for 4 min, followed by 35 cycles of 95 °C for 30 s, 58 °C for 30 s, 72 °C for 2.5 min, and a final extension step of 72 °C for 4 min. Amplicons were cleaned with Antarctic Phosphatase and Exonuclease I (New England Biolabs, Ipswich, MA, USA). Sequencing was performed on an Applied Biosystems 3730 capillary sequencer (Foster City, CA, USA) using Big Dye chemistry. Chromatograms were edited in Sequencher 4.1.2 (Gene Codes Corporation, Ann Arbor, MI, USA). Heterozygous bases were easily identified in the chromatograms for the three nuclear regions and coded with standard IUPAC notation. Because of the low levels of divergence among our sequences, alignment was trivial and performed manually in Se-Al. The most likely haplotypic phases of AP and AKT genotype sequences were ascertained with a combination of cloning and the program PHASE 2.1 (Stephens *et al.* 2001; Stephens & Donnelly 2003) called by the 'Open/Unphase genotype' option in DnaSP v. 5 (Librado & Rozas 2009); the inferred alleles form the basis for all further analyses involving these loci. The preferred model of sequence evolution for each locus (ITS: TIM3ef + I + G; CP: K81uf + I; AP: TVM + I; AKT: TVM + I + G) was determined according to Akaike Information Criterion (AIC) in jModelTest (Posada 2008). Sampling details, genotype information and GenBank accession numbers are provided in Tables S1 and S2 (Supporting Information).

Phylogenetic analyses and descriptive population genetics

For the ITS analysis, 44 new sequences were aligned with 49 from GenBank to create a matrix of 93 sequences. Included were the four species in this study, plus 37 other taxa including the pendent eastern North

American species, *L. michiganense*, *L. canadense* and *L. grayi*, and eight others from *Lilium* section *Pseudolirium*, the monophyletic group of North American species (Nishikawa *et al.* 1999) to which all taxa in this study belong. Unweighted parsimony analysis for the ITS locus was accomplished using PAUP* 4.0b10 (Swofford 2002) using 100 random-addition sequence replicates with TBR branch swapping; owing to overall low sequence divergence, parsimony bootstrapping was conducted with 10^6 'fast' stepwise addition sequences (Soltis & Soltis 2003). Maximum-likelihood (ML) analysis for this locus was conducted in GARLI v. 1.0 (Zwickl 2006). Likelihood bootstrap values were obtained with 1000 replicate searches. The statistical parsimony haplotype network was computed for complete sequences of the three chloroplast regions, *atpI-atpH*, *psbD-trnT* and *rpl32-trnL* (38 sequences), using TCS (Clement *et al.* 2000). The nuclear loci (AP: 82 haplotypes; AKT: 62 haplotypes) have a more complicated evolutionary history than chloroplast sequences; thus, network analyses for the two were conducted using the geodesically pruned quasi-median network algorithm (Ayling & Brown 2008) as implemented in SplitsTree4 (Huson & Bryant 2006), which produces pruned networks that connect all sequences (including multistate characters) by at least one shortest path. ML trees (not shown) were inferred for these sequences as well; they were poorly resolved and showed few supported nodes. However, neither nuclear locus showed phylogenetic evidence of paralogy. For *L. michauxii*, *L. superbum* and *L. pyrophilum*, Arlequin v. 3.5 (Excoffier & Lischer 2010) was used to estimate haplotype richness, number of segregating sites, nucleotide diversity π (Nei 1987) and Watterson's (1975) population mutation parameter θ, for the chloroplast and single-copy nuclear loci.

Testing divergence between L. michauxii, L. pyrophilum *and* L. superbum

As our data include a single individual per 'population', we treated species as the main hierarchical level for the purposes of these analyses. Pairwise F_{ST} (Weir & Cockerham 1984) and the exact test of population differentiation (Raymond & Rousset 1995; Goudet *et al.* 1996) between *L. michauxii*, *L. superbum* and *L. pyrophilum* were calculated in Arlequin v. 3.5 (Excoffier & Lischer 2010), with individuals and species used as the hierarchical groupings. Significance was assessed with 10^3 permutations (F_{ST}) or 2×10^6 Markov chain steps (exact test).

The nature of the divergence between *L. superbum* and *L. pyrophilum* was further investigated using the isolation-with-migration model (Nielsen & Wakeley 2001), implemented in IMa2 (Hey & Nielsen 2007). The full model in the two-population case includes six parameters (divergence time, θ for the ancestral and two descendent populations and migration rates between the descendent populations). This model assumes no recombination within loci and free recombination between loci and that markers are selectively neutral. Thus, several recombination detection methods available in the program RDP3 (beta 40; Martin *et al.* 2005) were used to search for recombinant alleles. As selection or demographic changes can cause departures from neutral expectations, DnaSP v. 5 (Librado & Rozas 2009) was used to perform three different tests of neutrality: Tajima's D (Tajima 1989), Fay and Wu's H (Fay & Wu 2000) and R_2 (Ramos-Onsins & Rozas 2002). Critical values for these statistics were obtained using 10^5 coalescent simulations. The chloroplast data set showed no evidence of recombination; the AP and AKT data sets were filtered with IMgc Online (Woerner *et al.* 2007) to create data sets that were free of detectable recombination and infinite sites violations. Maximum priors for the IMa2 analysis were based on recommended starting values given in the program documentation and refined after preliminary exploratory runs. Priors ultimately selected were population mutation rates (for *L. pyrophilum*, *L. superbum* and ancestral population) θ_0, θ_1 and $\theta_2 = 47$, splitting time parameter $t = 3$ and population migration rate m_1 and $m_2 = 10$. Mutation rate priors (CP: 1.5×10^{-9}, AP & AKT: 6.03×10^{-9}) were specified based on values given by Gaut (1998). Seventy geometrically heated chains (using the heating parameters ha = 0.98, hb = 0.50) were run for 750 000 generations beyond a 150 000 generation burn-in and trees were sampled every 75 generations. This process was repeated 10 times using different random number seeds.

Because results from each replicate were similar, 10^5 trees were concatenated into a single run in load-trees mode and the 'test nested models' option was activated. This option evaluates the likelihood of 24 models simpler than the full isolation-with-migration model by constraining parameters (other than divergence time) and rejecting those that are significantly worse than the full model based on a likelihood ratio test. We also compared models using an information-theoretic method (Carstens *et al.* 2009), which allows the relative performance of nested and non-nested models to be compared using AIC. Compared with a hypothesis-testing approach, which simply identifies models that are rejected as significantly worse than the full model, the information-theoretic approach provides model weights that allow the relative performance of each of a given set of models, including the full model, to be compared directly with others given the data (Burnham &

Anderson 2002). We used the full model posterior probability and the 2(log-likelihood ratio) values, which IMa2 estimates for each model under the assumption that the model's posterior probability is proportional to its likelihood, to calculate the AIC for the full model and each nested model. Subsequently, Akaike weights and evidence ratios were calculated (Burnham & Anderson 2002; Carstens *et al.* 2009).

Conversion of the IMa2 parameter estimates from coalescent to demographic units was accomplished assuming a generation time of 20 years. This is arbitrary but conservative, based on what little is known about the natural history of these species. Germination and establishment is slow, taking two seasons, and plants need 7 years to reach flowering size. Year-to-year survivorship is relatively high (>0.95, Wade Wall, unpublished data). Using the equation $T = \alpha + [s/(1 - s)]$, where T = generation time, α = age of first reproduction and s = adult survivorship (Lande *et al.* 2003), we obtain a value of 26 years. Although estimates of survivorship could be too high, the Lande equation does not account for the fact that older plants are typically larger and more fecund than younger ones. In either case, our generation time should be considered a minimum estimate.

Because isolation is implicit in the isolation-with-migration framework, we tested this assumption with a series of coalescent simulations. Briefly, we estimated N_e for each locus using BEAST (Drummond & Rambaut 2007). Because only *L. pyrophilum* and *L. superbum* sequences were included, simpler ML models were utilized (CP: HKY, AP: TnN + I + G, AKT: K81uf + I). We then used Mesquite v. 2.73 (Maddison & Maddison 2010) to simulate 1000 data sets under each of several simple divergence models (using estimated substitution models for each locus). We treated each species as a population such that *L. superbum* had a N_e 3× that of *L. pyrophilum* (the total N_e corresponding to the value from BEAST). The two populations coalesced at times corresponding to 2.58 Ma (earliest Pleistocene), 126 ka (upper Pleistocene) or 18 ka (last glacial maximum). We then conducted parsimony searches using PAUP* 4.10b (Swofford 2002) on each simulated data set saving 1000 consensus trees. Slatkin and Maddison's s (i.e. the number of parsimony steps implied by a given topology treating source population as a character, Slatkin & Maddison 1989) was computed for each tree to create a null distribution for each locus and divergence time. This was compared with the value of s for the empirical data. When minimum empirical values for s were higher than 95% of the simulated values, we rejected the scenario. To evaluate the effect of the level of migration inferred by IMa2, we duplicated these analyses, but allowing migration. Because Mesquite only allows symmetrical migration, we specified a rate of 9.8×10^{-6} migrants per individual per generation, which corresponds to the estimated value of the parameter under the 'equal migration rate' nested model in IMa2. Finally, following Gugger *et al.* (2010), we evaluated the no-divergence scenario by simulating 1000 data sets per locus under a single population scenario. The resulting parsimony consensus trees were contained within the two-population model described previously, and the null distributions of s were calculated. In this case, the scenario was rejected if the maximum empirical values of s were lower than 95% of the simulated values. As coalescent parameter estimates based on single loci are highly sensitive to stochastic error (Edwards & Beerli 2000), these simulations were conducted for both the upper and lower 90% HPD estimates of N_e from BEAST.

Table 1 Genetic diversity and results of neutrality tests

	Lilium michauxii			*Lilium pyrophilum*			*Lilium superbum*		
Species locus	CP	AKT	AP	CP	AKT	AP	CP	AKT	AP
Individuals (haplotypes)	8 (8)	5 (10)	7 (14)	15 (15)	13 (26)	18 (36)	13 (13)	12 (24)	15 (30)
Aligned length (bp)	2361	1428	453	2360	1428	453	2361	1428	453
Segregating sites	7	10	13	7	24	8	9	30	18
Observed haplotypes	5	7	9	4	16	9	7	17	12
Nucleotide diversity π	0.0010	0.0033	0.0098	0.0008	0.0024	0.0016	0.0009	0.0040	0.0053
Watterson's theta θ	0.0011	0.0025	0.0090	0.0009	0.0044	0.0043	0.0012	0.0061	0.0100
Tajima's D	−0.4150	0.0487	0.3349	−0.4468	−1.7637*	−1.8536**	−1.0835	−1.2142	−1.6319*
Fay and Wu's H	1.7857	0.8000	2.2418	1.3429	−8.8862*	−2.8794*	−1.9615	−4.8333	0.6437
R_2	0.1577	0.2091	0.1597	0.1301	0.0625**	0.0495***	0.1105*	0.0828	0.0692*

Sampling represents the number of individuals and the number of haplotypes (for phased nuclear loci). Significance of neutrality tests was assessed with 10^5 coalescent simulations in DnaSP v. 5.1 (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

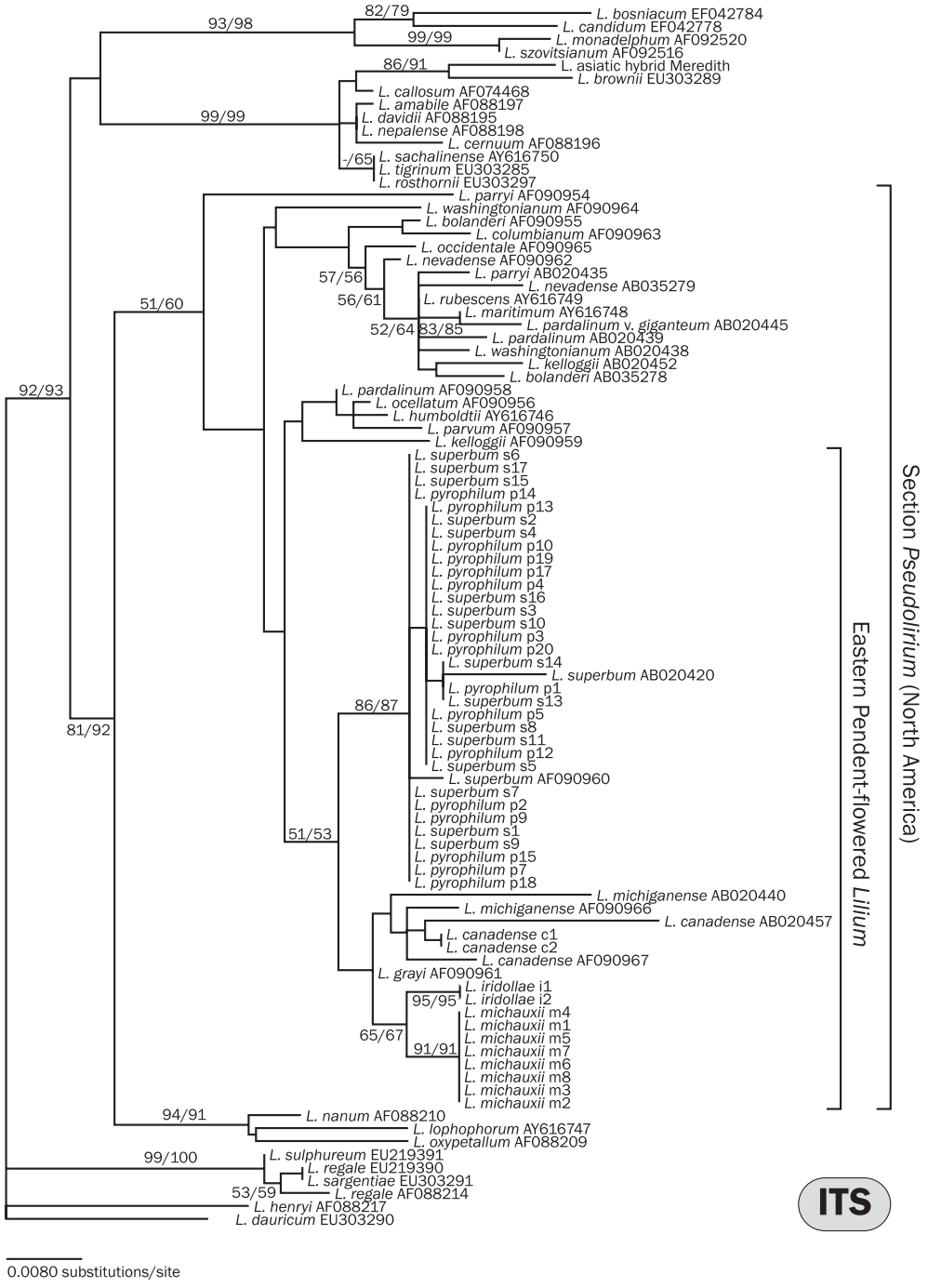

Fig. 2 Maximum-likelihood (ML) Phylogram of internal transcribed spacer sequences. Support values are ML bootstrap/Bayesian posterior probability.

Results

Phylogenetic analyses

In the analysis of ITS data, overall support is quite weak at the level of intra- and interspecific relationships, with no significant (≥70%) bootstrap support for the monophyly of the North American section *Pseudolirium* or the eastern pendent-flowered species (Fig. 2). However, there is a relatively high level of support for the branch uniting two accessions of *Lilium iridollae*, for that uniting the eight samples of *L. michauxii*, and, finally, for the branch leading to the 32 samples of *L. pyrophilum* and *L. superbum*. Little divergence is evident among the

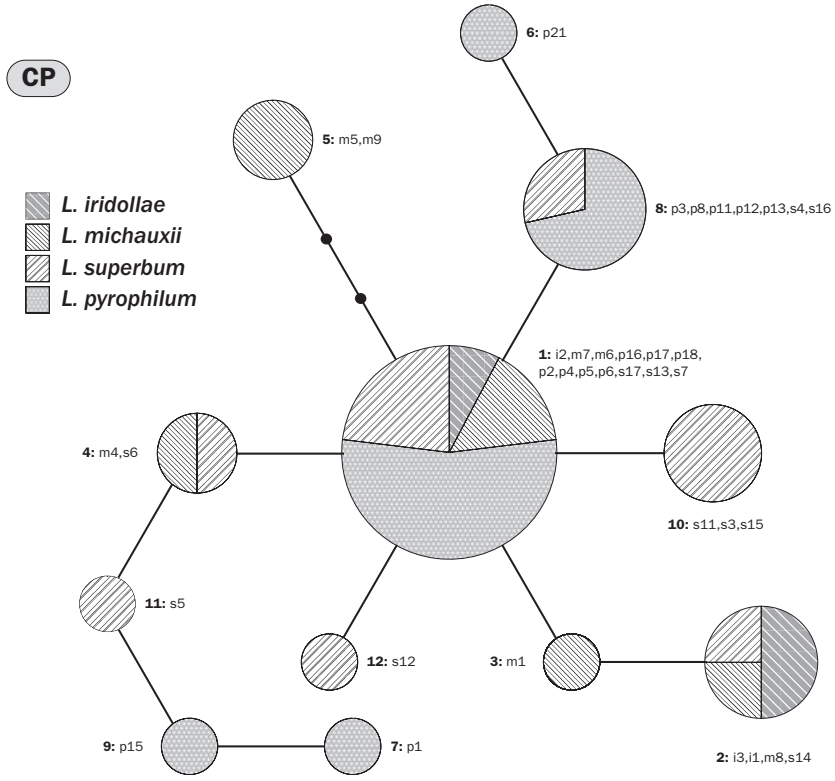

Fig. 3 Chloroplast haplotype network. Statistical parsimony network for CP haplotypes. Chart area reflects the frequency of the haplotype; each slice reflects the frequency at which each haplotype was found in each species. Haplotype numbers (bold) and sample abbreviations correspond to those in Tables S1 and S2 (Supporting Information). Edges represent mutations, black dots unsampled haplotypes.

accessions of each species (with the exception of the GenBank sequences for *L. superbum*, *L. canadense* and *L. michiganense*). The statistical parsimony network (Fig. 3) computed for the chloroplast data revealed a common haplotype (1) that was found in all four species, plus 11 less common types. Overall, four of the six non-singleton haplotypes occur in multiple species. Quasi-median networks produced for the AKT and AP loci (Fig. 4) showed that, while AP haplotype 8 is one mutational step from the nearest *L. michauxii* haplotype (m4a), most *L. michauxii* (and *L. iridollae* in AP) haplotypes are separated from a cloud of *L. pyrophilum* and *L. superbum* haplotypes, which are thoroughly intermixed and frequently shared. No haplotypes were shared between *L. pyrophilum* and *L. michauxii*.

Genetic diversity

Haplotype richness h, segregating sites S, nucleotide diversity π and Watterson's θ are given in Table 1. Nucleotide diversity is relatively low, with values between 0.0008 and 0.00978 substitutions per site, and average values for AP and AKT are nearly five times the value for the chloroplast data set.

Tests of neutrality

Departures from neutrality were detected in the nuclear data sets in *L. pyrophilum* and *L. superbum*, where there were significant negative estimates of Tajima's D and R_2. Fay and Wu's H is significant in *L. pyrophilum* only. Tajima's D is sensitive to both demographic expansion and selection, and R_2 is designed to detect population expansion (Ramos-Onsins & Rozas 2002). While Fay and Wu's H is most sensitive to recent positive selection, it may be sensitive to particular demographic conditions involving structured populations (Fay & Wu 2000). We believe these loci are unlikely to be under positive selection, because there is no obvious reason two loci should deviate from neutrality more strongly in *L. pyrophilum* than in the other two taxa. The chloroplast data also show some demographic expansion in *L. superbum* (weakly significant R_2) without a significantly negative D. Thus, while we cannot eliminate the possibility of some background selection in the nuclear data sets (which does not violate the assumptions of IMa2), it is more likely that demographic factors explain the significant values for these statistics.

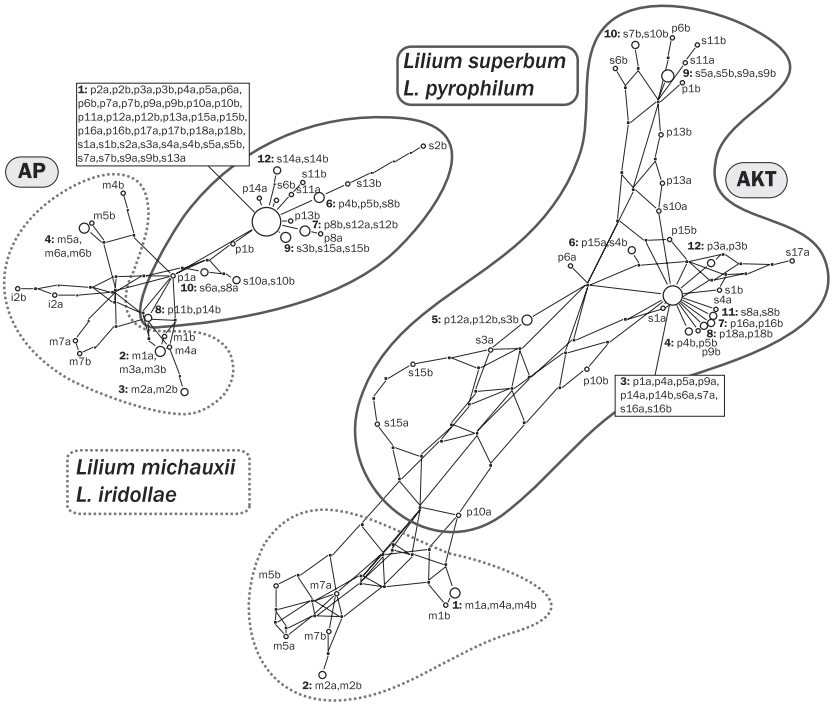

Fig. 4 Quasi-median joining networks for the nuclear loci AP and AKT. Network representations of the relationships between nuclear haplotypes (bold numbers and sample abbreviations correspond to Tables S1 and S2, in Supporting Information). In quasi-median-joining networks, each haplotype is connected to the others by at least one shortest path. Mutational steps are indicated by edges, and black dots represent potential unsampled haplotypes.

Differentiation of L. michauxii

Pairwise F_{ST} values (Table 2) revealed that *L. michauxii* was significantly divergent from *L. pyrophilum* and *L. superbum* for the AKT and AP data sets, whereas differentiation between *L. pyrophilum* and *L. superbum* was minimal and only significant in the AKT data set. No significant differentiation was detected among any of the three species for the CP data set. Conversely, all pairwise exact differentiation tests (Raymond & Rousset 1995) were significant for the two nuclear loci; for the cpDNA, a significant result was only obtained between *L. pyrophilum* and *L. michauxii*.

Divergence between L. pyrophilum *and* L. superbum

Under the isolation-with-migration model, estimates of the mutation parameter theta (θ) were *L. pyrophilum*: 3.736; *L. superbum*: 10.79; and ancestral population: 1.292, corresponding to effective population sizes (95% highest posterior density interval, abbreviated '95% HPD') of 11 400 (2800–29 700), 32 900 (12 800–86 900) and 3900 (0–14 400), respectively (Fig. 5a). The splitting time between *L. pyrophilum* and *L. superbum* was estimated as 0.7725 coalescent units, with the 95% HPD being 0.3435–2.405 (Fig. 5b). This estimate corresponds to a divergence time of 188 ka (95% HPD 84–586 ka) with the assumed mutation rates and generation time. The posterior distribution of splitting time did not reach zero (nor did it for much higher prior values in preliminary runs), so 95% HPD intervals should be interpreted with caution. The coalescent migration rate m from *L. superbum* into *L. pyrophilum* was highest at zero, while the converse was 1.915. Thus, population migration rates (2 NM, Hey & Nielsen 2004) are asymmetrical and quite high from *L. pyrophilum* into *L. superbum* (2 NM = 9.98, Fig. 5c). The model selection procedure (Table 3) preferred a model that holds the two species' population sizes equal and the migration rate from *L. superbum* to *L. pyrophilum* at zero (model weight $w = 0.32$). The next best model ($w = 0.22$) also fixed the *L. superbum* → *L. pyrophilum* migration rate at zero but allowed the population sizes to vary. The full model ($w = 0.19$) had the next highest weight, and the next three models differed in that they fixed the population sizes as above (model 4), held migration rates equal (model 5) and held the *L. pyrophilum* → *L. superbum* migration rate at zero (model 6). The six best models are assigned 95.6% of the total weight. The remaining 19 models had some combination of zero migration, and one or both of the population sizes

Table 2 Pairwise *Fst* and exact test of population differentiation

	Lilium michauxii	*Lilium pyrophilum*	*Lilium superbum*
L. michauxii		0.109/0.393***/0.625***	0.046/0.328***/0.567***
L. pyrophilum	*/***/***		0.007/0.021/0.057*
L. superbum	–/***/**	–/**/*	

Loci: CP/AP/AKT. Above diagonal, pairwise F_{ST}; below diagonal, exact test of differentiation (Goudet *et al*. 1996; Raymond & Rousset 1995). Significance assessed in Arlequin by either 10^3 permutations (F_{ST}) or 2×10^6 Markov chain steps (exact test); *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

equal to the ancestral population size. For the sake of comparison, likelihood ratio tests comparing each nested model to the full model rejected 20 of 24 nested models. The four that were not rejected, combined with the full model, represent 94.2% of the cumulative model weight from the information-theoretic analysis. Coalescent simulations under both the earliest Pleistocene (129 000 generations, 2.58 Ma) and upper Pleistocene (6300 generations, 126 ka) divergence scenarios were rejected (Table 4). However, divergence during the last glacial maximum (900 generations, 18 ka) was not rejected, and neither was the single population scenario under either the highest or lowest credible estimates for N_e. Inclusion of migration in these simulations did not qualitatively change the results.

Discussion

Three hypotheses

Our results do not favour two of the three hypotheses concerning the relationships of *Lilium pyrophilum* advanced by Skinner & Sorrie (2002). First, it is unlikely that *L. pyrophilum* simply represents a disjunct population of *L. iridollae*: the ITS phylogeny unambiguously allies *L. pyrophilum* with *L. superbum*, whereas *L. iridollae* is most closely related to *L. michauxii*. That *L. pyrophilum* and *L. iridollae* are independent only heightens the conservation concern of each of these rare species.

Second, the hypothesis that the species originated as a hybrid between *L. michauxii* and *L. superbum* is not supported by network analyses (Fig. 4). If *L. pyrophilum* represented a recent hybrid, single-copy nuclear loci should be related to both parental species. Instead, most *L. pyrophilum* and *L. superbum* haplotypes are closely related to each other (and many are shared), while they show less similarity to *L. michauxii*. The phylogenetic analysis of ITS sequences placed the *L. pyrophilum* samples with *L. superbum* sequences only, to the exclusion of the *L. michauxii* sequences.

Lilium pyrophilum appears to be a peripheral isolate of *L. superbum*. Our results indicate that the overall magnitude of divergence between the two lily species is very low and that the origin of *L. pyrophilum* is likely to have been very recent, i.e. during the latter Pleistocene or Holocene. Our estimated divergence date from the IMa2 analysis of 188 ka (Fig. 5b) would fall within the Illinoian glacial period. The minimum credible divergence time of 84 ka would seem to indicate that *L. pyrophilum* is in fact isolated from *L. superbum*. In spite of low F_{ST} values (Table 2), zero probability is assigned to the most recent divergence times in this analysis. The results of the simulation

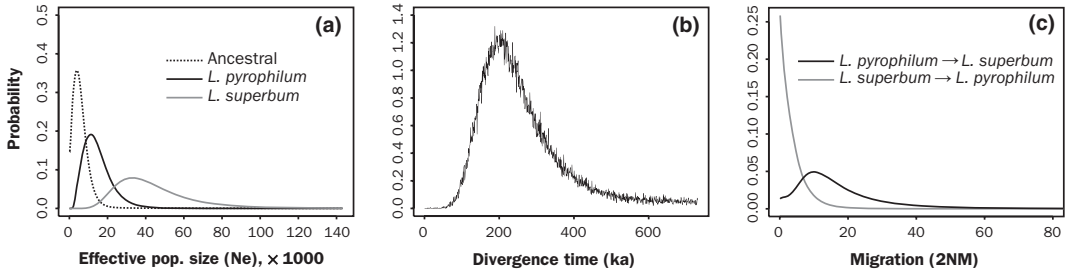

Fig. 5 Posterior probability distributions for IMa2 model parameters under the full model. (a) Effective population size for both species and the ancestral population. Both descendent taxa are inferred to have larger effective population size in this analysis. Estimated values for N_e are *Lilium pyrophilum*, 11 400 (95% HPD 2800–29 700); *L. superbum*, 32 900 (12 800–86 900); and ancestral, 3900 (0–14 400). (b) Divergence time. No probability is found for divergence times near zero; however, the distribution fails to reach zero at the upper end. The peak corresponds to a value of 188 (84–586) ka. (c) Migration rate. Highest probability for migration from *L. superbum* into *L. pyrophilum* is zero; there is, however, a higher probability of migration in the opposite direction (2 NM = 9.98).

Table 3 IMa2 analysis of nested models

Model description	log(*P*)	Terms	AIC	ΔAIC	Model weight	Cum. weight	d.f.	2LLR	*P*-value, LRT
θ (*pyrophilum*) = θ (*superbum*), *m* zero from *superbum* to *pyrophilum*	−4.442	3	14.884	0	0.301	0.301	2	2.986	0.2247
m zero from *superbum* to *pyrophilum*	−3.825	4	15.65	0.766	0.2052	0.5062	1	1.752	0.1856
Full IM model	−2.949	5	15.898	1.014	0.1813	0.6875	—	—	—
θ (*pyrophilum*) = θ (*superbum*)	−3.972	4	15.944	1.06	0.1772	0.8647	1	2.045	0.1527
Symmetrical migration	−4.803	4	17.606	2.722	0.0772	0.9419	1	3.707	0.0542
m zero from *pyrophilum* to *superbum*	−6.29	4	20.58	5.696	0.0174	0.9593	1	6.681	0.0097
θ (*pyrophilum*) = θ (ancestral), *m* zero from *superbum* to *pyrophilum*	−7.985	3	21.97	7.086	0.0087	0.968	2	10.07	0.0065
θ (*pyrophilum*) = θ (ancestral), *m* zero from *pyrophilum* to *superbum*	−8.116	3	22.232	7.348	0.0076	0.9757	2	10.33	0.0057
θ (*pyrophilum*) = θ (*superbum*), symmetrical migration	−8.408	3	22.816	7.932	0.0057	0.9814	2	10.92	0.0043
θ (*pyrophilum*) = θ (ancestral), symmetrical migration	−8.424	3	22.848	7.964	0.0056	0.987	2	10.95	0.0042
All θ equal, *m* zero from *superbum* to *pyrophilum*	−9.858	2	23.716	8.832	0.0036	0.9906	3	13.82	0.0032
θ (*pyrophilum*) = θ (ancestral)	−7.899	4	23.798	8.914	0.0035	0.9941	1	9.9	0.0017
θ (*superbum*) = θ (ancestral), *m* zero from *superbum* to *pyrophilum*	−9.192	3	24.384	9.5	0.0026	0.9967	2	12.49	0.0019
All θ equal	−9.858	3	25.716	10.832	0.0013	0.9981	2	13.82	0.001
θ (*superbum*) = θ (ancestral)	−9.192	4	26.384	11.5	0.001	0.999	1	12.49	0.0004
θ (*pyrophilum*) = θ (*superbum*), *m* zero from *pyrophilum* to *superbum*	−10.63	3	27.26	12.376	0.0006	0.9997	2	15.36	0.0005
θ (*superbum*) = θ (ancestral), symmetrical migration	−12.1	3	30.2	15.316	0.0001	0.9998	2	18.3	0.0001
All θ equal, symmetrical migration	−13.4	2	30.8	15.916	0.0001	0.9999	3	20.9	0.0001
θ (*superbum*) = θ (ancestral), zero migration	−14.26	2	32.52	17.636	0	0.9999	3	22.63	0
Zero migration	−13.35	3	32.7	17.816	0	1	2	20.8	0
θ (*superbum*) = θ (ancestral), *m* zero from *pyrophilum* to *superbum*	−14.26	3	34.52	19.636	0	1	2	22.63	0
All θ equal, *m* zero from *pyrophilum* to *superbum*	−18.52	2	41.04	26.156	0	1	3	31.13	0
θ (*pyrophilum*) = θ (*superbum*), zero migration	−24.86	2	53.72	38.836	0	1	3	43.83	0
θ (*pyrophilum*) = θ (ancestral), zero migration	−29.23	2	62.46	47.576	0	1	3	52.57	0
All θ equal, zero migration	−30.93	1	63.86	48.976	0	1	4	55.97	0

Models include the full IM model and 24 simpler nested models for the two-population case. Information-theoretic statistics, based on log(*P*), follow Burnham & Anderson (2002) and have been sorted by model weight. Models not rejected under traditional-likelihood ratio tests (LRT) are included in the 95% confidence set of models selected by AIC.

analysis lead us to interpret the IMa2 results with caution, however, because they reject divergence >6300 generations (126 ka) ago for each locus and fail to reject the scenarios with divergence at 900 generations (18 ka) and with no divergence (Table 4). The models tested in this approach, however, were simplified with respect to the full IMa2 model and treat each locus separately rather than jointly. Regardless of whether the IMa2 results or the coalescent simulation results are preferred, the isolation between the two taxa is not ancient. Mid- to late Pleistocene divergence times have been found in surprisingly few studies of plants (e.g. Strasburg & Rieseberg 2008; Bittkau & Comes 2009; Cooper *et al.* 2010).

Our results provide insight into the demographic patterns that have affected the two species. Deviations from neutral expectation indicate population expansion in both taxa (e.g. the average value for Tajima's *D* across three loci: *L. pyrophilum* = −1.35, *L. superbum* = −1.31, Table 1). This result is corroborated by the IMa2 analysis, which demonstrates modern effective population sizes higher than the ancestral, with the widespread *L. superbum* having a larger value (N_e 2.7 times that of the endemic *L. pyrophilum*, Fig. 5a). It is worth noting that the effective population size of *L. pyrophilum* (11 000 individuals) is surprisingly high considering the very small range of the species; in fact, our estimate of N_e is well in excess of the current census population size estimated by a recent inventory. Two factors may explain this discrepancy. First, our estimated generation time may be too low, which would cause us to overestimate effective population size (and underestimate divergence time). Second, agriculture, timber harvesting and fire suppression have

Table 4 Results of coalescent simulation study

	Marker		
Simulation model	AKT	AP	CP
Divergence time (in generations) without migration			
129 000	**0.000**	**0.000**	**0.000**
6300	**0.023**	**0.008**	**0.036**
900	0.992	0.379	0.394
Divergence time (in generations) with migration			
129 000	**0.001**	**0.001**	**0.006**
6300	**0.028**	**0.013**	**0.034**
900	0.839	0.438	0.422
No divergence			
High N_e	0.122	0.148	0.546
Low N_e	0.065	0.181	0.235

P-value for each model was obtained by comparison of either minimum (divergence) or maximum (no divergence) empirical *s* value (Slatkin & Maddison 1989) with simulated distributions of *s* under coalescent scenarios to test whether observed data were consistent with divergence times discussed in text. Simulations were based on assumed 20-year generation time.

dramatically transformed much of the landscape of the Sandhills over the past few hundred years, which may well have extirpated many populations. As these plants are long-lived outcrossers, too few generations may have elapsed for the impact of the current bottleneck to be fully reflected in the estimated N_e (Lande & Barrowclough 1987). Although our results suggest that the existing population has apparently been greatly reduced recently, much of the original genetic diversity remains and could be conserved, minimizing the impact of the present-day population bottleneck.

Gene flow is inferred from *L. pyrophilum* to *L. superbum*. Models including symmetrical migration are not weighted heavily compared with models that have zero or nearly zero gene flow from *L. superbum* to *L. pyrophilum* (Table 3). Presently, the two species are disjunct. However, the plants are visited by strong-flying pollinators, such as swallowtail butterflies and hummingbirds (Skinner 2002), and the seeds are adapted for wind dispersal. Why migration would be asymmetrical is unknown, but this could be explained by pollinator behaviour, dispersal or intrinsic barriers to gene flow.

Edaphic endemism in the Sandhills

The Sandhills pre-date the Pleistocene and may be substantially older, raising the possibility that some endemic taxa may have originated in the Pliocene or earlier and maintained populations in the region continuously. How might Pleistocene climate changes have affected the distribution of *Lilium* spp. in the coastal plain and effected the isolation of *L. pyrophilum*? While periods of severe climate change may eliminate edaphic endemics that are unable to migrate to areas with a suitable climate and substrate, edaphic endemics may in fact be likely to endure climate change in their geographic ranges. As their niches are defined more by soils than climate, they are likely to remain the best competitors on restrictive soils under a wide range of conditions. In fact, the degree of edaphic restriction exhibited by a species often varies with climate: populations may be widespread in environments with low competition and edaphically restricted in more favourable climates (Brooks 1987; Harrison *et al.* 2009).

The edaphic conditions that currently support populations of *L. pyrophilum* have probably been relatively stable, because the erosional process has no doubt continually exposed the interface between permeable and impermeable soils, creating seeps. Yet, the divergence between *L. pyrophilum* and *L. superbum* is comparatively recent. Genetic diversity of *L. pyrophilum*, while lower than that of *L. superbum*, is still high, making a vicariant scenario likely. Thus, the phenotypic divergence described by Skinner & Sorrie (2002) probably occurred in the context of large populations and substantial gene flow.

The combination of long-term persistence and recent divergence of *L. pyrophilum* indicates that this species descends from locally adapted populations that were stranded in the Sandhills as *L. superbum* retreated to higher elevations. It is not clear why the intervening Piedmont region supports neither taxon; however, many groups show a similar disjunction (Braun 1955; Sorrie & Weakley 2001). This study indicates that for these lilies, at least, the disjunction coincided with Pleistocene climate oscillations; this may apply to other taxa that share similar distributions. More in-depth studies of the *L. pyrophilum/L. superbum* system, using microsatellite markers, will quantify genetic structure within *L. pyrophilum*, and gene flow within and between *L. pyrophilum* and *L. superbum*. These more detailed analyses will improve estimates of divergence time and gene flow and identify populations of high conservation priority. Better understanding of this group will provide further insight into the role of edaphic specialization, possibly brought on by climate change, in promoting diversification.

Acknowledgements

We thank Fort Bragg Military Reservation and the Endangered Species Branch for logistic support and the Construction Engineering Research Laboratory (US Army Corps of Engineers Agreement #W9132T-07-2-0019) for funding. We also thank Xiang Liu, David Thomas, Patrick Zhou, Esther Ichugo, Matt

Cleary and Jacob Hilton for assistance with laboratory work and Marshall Wilson, Tom Phillips, Mac Alford, Rob Naczi, Gary Shurette, Viola Walker, Emil Devito, Heather Sullivan, Paul Manos, John Pogacnik, James Smith, Wayne Longbottom, Misty Buchanan and others for assistance locating *Lilium* populations.

References

Axelrod DI (1972) Edaphic aridity as a factor in angiosperm evolution. *The American Naturalist*, **106**, 311–320.

Ayling SC, Brown TA (2008) Novel methodology for construction and pruning of quasi-median networks. *BMC Bioinformatics*, **9**, 115.

Ayres DR, Ryan FJ (1999) Genetic diversity and structure of the narrow endemic *Wyethia reticulata* and its congener *W. bolanderi* (Asteraceae) using RAPD and allozyme techniques. *American Journal of Botany*, **86**, 344–353.

Baldwin BG (2005) Origin of the serpentine-endemic herb *Layia discoidea* from the widespread *L. glandulosa* (Compositae). *Evolution*, **59**, 2473–2479.

Baskauf CJ, McCauley DE, Eickmeier WG (1994) Genetic analysis of a rare and a widespread species of *Echinacea* (Asteraceae). *Evolution*, **48**, 180–188.

Bittkau C, Comes HP (2009) Molecular inference of a Late Pleistocene diversification shift in *Nigella* s. lat. (Ranunculaceae) resulting from increased speciation in the Aegean archipelago. *Journal of Biogeography*, **36**, 1346–1360.

Braun EL (1955) The phytogeography of unglaciated Eastern United States and its interpretation. *The Botanical Review*, **21**, 297–375.

Brooks RR (1987) *Serpentine and its Vegetation: A Multidisciplinary Approach*. Dioscorides Press, Portland, OR.

Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS (2001) Comparative phylogeography of northwestern North America: a synthesis. In: *Integrating Ecology and Evolution in a Spatial Context* (eds Silvertown J, Antonovics J), pp. 319–340. Blackwell Science, Oxford.

Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, New York.

Cabe S, Nickerson JG, Hoffman CW, Farrell K (1992) Cretaceous and Tertiary stratigraphy of the northern Sandhills area, North Carolina. In: *Geologic field guides to North Carolina and vicinity* (eds Dennison J, Stewart KG), Annual Meeting of the Southeastern Section of the Geological Society of America. Winston-Salem, NC: Geological Society of America, Boulder, CO.

Carstens BC, Stoute HN, Reid NM (2009) An information-theoretical approach to phylogeography. *Molecular Ecology*, **18**, 4270–4282.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Cooper EA, Whittall JB, Hodges SA, Nordborg M (2010) Genetic variation at nuclear loci fails to distinguish two morphologically distinct species of *Aquilegia*. *PLoS ONE*, **5**, e8655.

Crawford DJ, Ornduff R, Vasey MC (1985) Allozyme variation within and between *Lasthenia minor* and its derivative species, *L. maritima* (Asteraceae). *American Journal of Botany*, **72**, 1177–1184.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.

Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.

Edwards CE, Soltis DE, Soltis PS (2006) Molecular phylogeny of *Conradina* and other scrub mints (Lamiaceae) from the southeastern USA: evidence for hybridization in Pleistocene refugia? *Systematic Botany*, **31**, 193–207.

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics*, **24**, 217–242.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Fay JC, Wu CI (2000) Hitchhiking under positive Darwinian selection. *Genetics*, **155**, 1405.

Freeland JR, Biss P, Conrad KF, Silvertown J (2010) Selection pressures have caused genome-wide population differentiation of *Anthoxanthum odoratum* despite the potential for high gene flow. *Journal of Evolutionary Biology*, **23**, 776–782.

Gaut BS (1998) Molecular clocks and nucleotide substitution rates in higher plants. *Evolutionary Biology*, **30**, 93–120.

Gitzendanner MA, Soltis PS (2000) Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany*, **87**, 783–792.

Godt MJW, Hamrick JL (1993) Genetic diversity and population structure in *Tradescantia hirsuticaulis* (Commelinaceae). *American Journal of Botany*, **80**, 959–966.

Goudet J, Raymond M, de-Meeus T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1933–1940.

Gugger PF, Sugita S, Cavender-Bares J (2010) Phylogeography of Douglas-fir based on mitochondrial and chloroplast DNA sequences: testing hypotheses from the fossil record. *Molecular Ecology*, **19**, 1877–1897.

Hamrick JL, Godt MJW (1990) Allozyme diversity in plant species. In: *Plant Population Genetics, Breeding, and Genetic Resources* (eds Brown AHD, Clegg MT, Kahler AL, Weir BS), pp. 43–63. Sinauer Associates, Sunderland, MA.

Harrison S, Damschen E, Going BM (2009) Climate gradients, climate change, and special edaphic floras. *Northeastern Naturalist*, **16**, 121–130.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences*, **104**, 2785–2790.

Howarth DG, Baum DA (2005) Genealogical evidence of homoploid hybrid speciation in an adaptive radiation of

Scaevola (Goodeniaceae) in the Hawaiian Islands. *Evolution*, **59**, 948–961.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.

Karron JD (1987) A comparison of levels of genetic polymorphism and self-compatibility in geographically restricted and widespread plant congeners. *Evolutionary Ecology*, **1**, 47–58.

Knowles LL (2009) Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 593–612.

Kruckeberg AR (1986) An essay: the stimulus of unusual geologies for plant speciation. *Systematic Botany*, **11**, 455–463.

Kruckeberg AR, Rabinowitz D (1985) Biological aspects of endemism in higher plants. *Annual Review of Ecology and Systematics*, **16**, 447–479.

Lande R, Barrowclough GF (1987) Effective population size, genetic variation, and their use in population management. In: *Viable Populations for Conservation* (ed. Soulé ME), pp. 87–124. Cambridge University Press, Cambridge.

Lande R, Engen S, Sæther B-E (2003) *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press, Oxford.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Maddison WP, Maddison DR (2010) *Mesquite: A Modular System for Evolutionary Analysis*. http://mesquiteproject.org.

Martin DP, Williamson C, Posada D (2005) RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics*, **21**, 260–262.

Mehta BD, Jog SP, Johnson SC, Murthy PPN (2006) Lily pollen alkaline phytase is a histidine phosphatase similar to mammalian multiple inositol polyphosphate phosphatase (MINPP). *Phytochemistry*, **67**, 1874–1886.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Nishikawa T, Okazaki K, Uchino T, Arakawa K, Nagamine T (1999) A molecular phylogeny of *Lilium* in the internal transcribed spacer region of Nuclear ribosomal DNA. *Journal of Molecular Evolution*, **49**, 238–249.

North Carolina Natural Heritage Program (2007) *North Carolina Natural Heritage Program Element Occurrence Database*. NC Department of Environment and Natural Resources, Raleigh, NC, 27699–1610.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375–402.

Oliveira LO, Huck RB, Gitzendanner MA, Judd WS, Soltis DE, Soltis PS (2007) Molecular phylogeny, biogeography, and systematics of *Dicerandra* (Lamiaceae), a genus endemic to the southeastern United States. *American Journal of Botany*, **94**, 1017–1027.

Ornduff R (1965) Ornithocoprophilous endemism in Pacific basin angiosperms. *Ecology*, **46**, 864–867.

Parsons RF (1976) Gypsophily in plants-a review. *American Midland Naturalist*, **96**, 1–20.

Peruzzi L, Leitch IJ, Caparelli KF (2009) Chromosome diversity and evolution in Liliaceae. *Annals of Botany*, **103**, 459–475.

Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Rabinowitz D (1981) Seven forms of rarity. In: *The Biological Aspects of Rare Plant Conservation* (ed. Synge H), pp. 205–217. Wiley, New York.

Raduski AR, Rieseberg LH, Strasburg JL (2010) Effective population size, gene flow, and species status in a narrow endemic sunflower, *Helianthus neglectus*, compared to its widespread sister species, *H. petiolaris*. *International Journal of Molecular Sciences*, **11**, 492–506.

Rajakaruna N (2004) The edaphic factor in the origin of plant species. *International Geology Review*, **46**, 471–478.

Rambaut A (1996) *Se-Al: Sequence Alignment Editor*. Available at http://evolve.zoo.ox.ac.uk/.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.

Rieseberg LH, Carter R, Zona S (1990) Molecular tests of the hypothesized hybrid origin of two diploid *Helianthus* species (Asteraceae). *Evolution*, **44**, 1498–1511.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, **132**, 365–386.

Schafale MP, Weakley AS (1990) *Classification of the Natural Communities of North Carolina, Third Approximation*. North Carolina Department of Environment, Health, and Natural Resources, Division of Parks and Recreation, Natural Heritage Program, Raleigh.

Shaw J, Lickey EB, Schilling EE, Small RL (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.

Skinner MW (2002) *Lilium*. In: *Flora of North America North of Mexico. Magnoliophyta: Liliidae: Liliales and Orchidales* (ed. Flora of North America Editorial Committee), pp. 172–197. Oxford University Press, Oxford.

Skinner MW, Sorrie BA (2002) Conservation and ecology of *Lilium pyrophilum*, a new species of Liliaceae from the Sandhills region of the Carolinas and Virginia, USA. *Novon*, **12**, 94–105.

Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*, **123**, 603–613.

Sohl NF, Owens JP (1991) Cretaceous stratigraphy of the Carolina Coastal Plain. In: *The Geology of the Carolinas: Carolina Geological Society Fiftieth Anniversary Volume* (eds Horton JW, Zullo VA, Carolina Geological Society), pp. 191–220. Univ. of Tennessee Press, Knoxville, TN.

Soltis PS, Soltis DE (2003) Applying the bootstrap in phylogeny reconstruction. *Statistical Science*, **18**, 256–267.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261–4293.

Sorrie BA, Weakley AS (2001) Coastal plain vascular plant endemics: phytogeographic patterns. *Castanea*, **66**, 50–82.

Sorrie BA, Gray JB, Crutchfield PJ (2006) The vascular flora of the longleaf pine ecosystem of Fort Bragg and Weymouth Woods, North Carolina. *Castanea*, **71**, 129–161.

Stanford AM, Harden R, Parks CR (2000) Phylogeny and biogeography of *Juglans* (Juglandaceae) based on matK and ITS sequence data. *American Journal of Botany*, **87**, 872–882.

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–1169.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.

Strasburg JL, Rieseberg LH (2008) Molecular demographic history of the annual sunflowers *Helianthus annuus* and *H. petiolaris*—large effective population sizes and rates of long-term gene flow. *Evolution*, **62**, 1936.

Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer, Sunderland, MA.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Wall WA, Douglas NA, Xiang QY, Hoffmann WA, Wentworth TR, Hohmann MG (2010) Evidence for range stasis during the latter Pleistocene for the Atlantic Coastal Plain endemic genus, *Pyxidanthera* Michaux. *Molecular Ecology*, **19**, 4302–4314.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wheelan SJ, Church DM, Ostell JM (2001) Spidey: a tool for mRNA-to-genomic alignments. *Genome Research*, **11**, 1952–1957.

White TJ, Bruns T, Lee S, Taylor J (1990) Amplifications and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications* (eds Innis M, Gelfand D, Sninsky J, White T), pp. 315–322. Academic Press, San Diego, CA.

Whittall JB, Syring J, Parks M *et al.* (2010) Finding a (pine) needle in a haystack: chloroplast genome sequence divergence in rare and widespread pines. *Molecular Ecology*, **19**, 100–114.

Williamson PS, Bazeer SK (1997) Self-incompatibility in *Abronia macrocarpa* (Nyctaginaceae). *Southwestern Naturalist*, **42**, 409–415.

Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**, 1851–1853.

Zonneveld BJM, Leitch IJ, Bennett MD (2005) First nuclear DNA amounts in more than 300 angiosperms. *Annals of Botany*, **96**, 229–244.

Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. http://code.google.com/p/garli/.

N.A.D. is an evolutionary biologist who focuses on phylogenetics and phylogeography of arid-adapted and edaphic endemic plants. W.A.W. is a plant ecologist interested in the interplay between population genetics and population dynamics of rare and endemic species. Q.Y.X.'s research explores the mechanisms underlying biodiversity patterns and morphological variation, with emphasis on the Cornales. W.A.H. is interested in ecology and conservation of savanna ecosystems. T.R.W. is a vegetation scientist with research interests in vegetation/environment relationships and biological diversity. J.B.G. is a botanist working on conservation of rare plants in Coastal Plain habitats. M.G.H. uses diverse approaches to inform rare species conservation.

Data accessibility

Sample and haplotype information is found in Table S1 (Supporting Information). DNA sequences: GenBank accessions JF829316–JF829423 (Table S2, in Supporting Information). ITS data and phylogenetic tree available at http://purl.org/phylo/treebase/phylows/study/TB2:S11519.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Sampling. Haplotype numbers correspond to sample labels in Figs 3 and 4, and to names accessioned in GenBank (Table S2, in Supporting Information). For AP and AKT sequences, integers identify haplotypes recovered more than once in this study and other identifiers refer to unique haplotypes

Table S2 GenBank accession numbers. Haplotype names correspond to samples in Table S1 (Supporting Information)

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 4302–4314 doi: 10.1111/j.1365-294X.2010.04793.x

Evidence for range stasis during the latter Pleistocene for the Atlantic Coastal Plain endemic genus, *Pyxidanthera* Michaux

WADE A. WALL,* NORMAN A. DOUGLAS,* QIU-YUN (JENNY) XIANG, * WILLIAM A. HOFFMANN,* THOMAS R. WENTWORTH* and MATTHEW G. HOHMANN†

**Department of Plant Biology, Box 7612, North Carolina State University, Raleigh, NC 27695, USA, †US Army Corps of Engineers, Engineer Research and Development Center, PO Box 9005, Champaign, IL 61826, USA*

Abstract

The general phylogeographical paradigm for eastern North America (ENA) is that many plant and animal species retreated into southern refugia during the last glacial period, then expanded northward after the last glacial maximum (LGM). However, some taxa of the Gulf and Atlantic Coastal Plain (GACP) demonstrate complex yet recurrent distributional patterns that cannot be explained by this model. For example, eight co-occurring endemic plant taxa with ranges from New York to South Carolina exhibit a large disjunction separating northern and southern populations by >300 km. *Pyxidanthera* (Diapensiaceae), a plant genus that exhibits this pattern, consists of two taxa recognized as either species or varieties. We investigated the taxonomy and phylogeography of *Pyxidanthera* using morphological data, cpDNA sequences, and amplified fragment length polymorphism markers. Morphological characters thought to be important in distinguishing *Pyxidanthera barbulata* and *P. brevifolia* demonstrate substantial overlap with no clear discontinuities. Genetic differentiation is minimal and diversity estimates for northern and southern populations of *Pxyidanthera* are similar, with no decrease in rare alleles in northern populations. In addition, the northern populations harbour several unique cpDNA haplotypes. *Pyxidanthera* appears to consist of one morphologically variable species that persisted in or near its present range at least through the latter Pleistocene, while the vicariance of the northern and southern populations may be comparatively recent. This work demonstrates that the refugial paradigm is not always appropriate and GACP endemic plants, in particular, may exhibit phylogeographical patterns qualitatively different from those of other ENA plant species.

Keywords: amplified fragment length polymorphism, cpDNA, Diapensiaceae, phylogeography, Pleistocene, refugium

Received 6 April 2010; revision received 28 June 2010; accepted 7 July 2010

Introduction

The alternating glacial and interglacial periods that characterized the Pleistocene had major impacts on the biogeography and genetic diversity of plant species in the Northern Hemisphere (Comes & Kadereit 1998; Hewitt 2000). The LGM, approximately 18 000 years BP, saw the Laurentide ice sheet reach its southern extent in eastern North America (ENA) (Ehlers & Gibbard 2004). The primary scenario describing plant species' ranges during and following the LGM in ENA includes (i) range contraction to southern refugia (Delcourt & Delcourt 1981) and (ii) subsequent recolonization of northern habitats after the retreat of the glaciers (Dorken & Barrett 2004). Previous studies have identified the resulting phylogeographical patterns of plant species in ENA and made inferences about possible refugia during the glacial maxima (Soltis *et al.* 2006). Several such patterns have been identified in the ranges

Correspondence: Wade A. Wall, Fax: (919) 515 3436;
E-mail: wade.wall@gmail.com

of ENA plant and animal species. These include east–west divisions between Gulf Coast and Atlantic Coast populations (Mylecraine *et al.* 2004), phylogeographical separation by river drainage systems (Church *et al.* 2003), and the identification of refugia closer to the glacial front, either in the southern Appalachians (Gonzales *et al.* 2008) or farther north (Magni *et al.* 2005; McLachlan *et al.* 2005).

Eastern North America has generally been divided into four physiogeographical regions: Interior Lowlands, Appalachian Highlands, Piedmont, and Gulf and Atlantic Coastal Plain (GACP) (Fenneman 1938). Compared to the other regions, the GACP is well defined geologically and floristically (Takhtajan 1986), but little is known about the phylogeography of the many widespread species that are endemic to the GACP. Previous phylogeographical studies of GACP plant species have generally focused on narrow endemics with small latitudinal ranges (Evans *et al.* 2000; Oliveira *et al.* 2007) or species with ranges that cross multiple physiographical regions (Morris *et al.* 2008; however, see Mylecraine *et al.* 2004). This is unfortunate; over 1300 species and 47 genera are endemic to the region (Sorrie & Weakley 2001), the second-highest concentration in the United States and only exceeded by the California Floristic Province (Flora of North America Editorial Committee 1993). Without taking into account the endemic species of the GACP and their distributional patterns, any attempt to understand the postglacial phylogeography of ENA is limited.

Sorrie & Weakley (2001) documented 27 different recurrent distributional patterns for endemic plant species of the GACP. One of the most interesting of these patterns is the disjunct distribution of eight taxa (*Calamovilfa brevipilis*, *Dichanthelium hirstii*, *Eupatorium resinosum*, *Gentiana autumnalis*, *Lobelia canbyi*, *Narthecium americanum*, *Pyxidanthera barbulata*, and *Rhynchospora pallida*) that occur in New York and New Jersey and eastern North and South Carolina, but not in the intervening areas of Maryland, Delaware, and most of Virginia. In addition to these eight taxa, numerous species exhibit the same disjunction between New Jersey and the southern GACP, but are more widespread in the southern part of the GACP. Common distributional patterns may be the result common biogeographical processes, but there is always the possibility of pseudocongruence (Hafner & Nadler 1990; Cunningham & Collins 1994).

We focus here on the genus *Pyxidanthera* Michaux as a case study to investigate the refugial paradigm in the GACP. *Pyxidanthera* is in Diapensiaceae, a small family with a circumboreal distribution, with some taxa extending southward into eastern Asia and ENA. The genus includes two recognized species; both are woody, winter-flowering, evergreen cushion plants. Populations of the more widespread *P. barbulata* occur on Long Island in New York, the Pine Barrens of New Jersey, several locations in southeastern Virginia, and the coastal plain of North Carolina and South Carolina. *P. brevifolia* has a more limited range; it has only been documented in six counties in the Sandhills region of North Carolina and South Carolina. *P. brevifolia*, currently under intensive study as a species at risk by the US Department of Defense, is considered vulnerable to extinction in North Carolina, with over 80% of the North Carolina populations confined to Fort Bragg Military Reservation, NC (Buchanan & Finnegan 2008). *P. brevifolia* is nearly restricted to xeric sandhill scrub communities within the long-leaf pine ecosystem (Schafale & Weakley 1990; Sorrie *et al*. 2006), one of the most imperiled ecosystems in North America, with approximately 2% of the historical area currently extant (Frost 2006). In addition to containing most of the remaining *P. brevifolia* populations, Fort Bragg Military Reservation is also one of the few places where the two species of *Pyxidanthera* co-occur. When sympatric, *P. barbulata* and *P. brevifolia* occupy nonoverlapping ecological habitats, with *P. barbulata* occupying wetter sites such as pocosin ecotones and *P. brevifolia* occurring on extremely xeric sand ridges.

Pyxidanthera was monotypic until 1929, at which time *P. brevifolia* was separated from sympatric populations of *P. barbulata* in the Sandhills region of North Carolina and upper South Carolina based on habitat differences, shorter leaves, and dense pubescence relative to the more widespread *P. barbulata* (Wells 1929). Differing ecological niches and morphological characters of *P. barbulata* and *P. brevifolia* led to a debate regarding the proper taxonomic status of the two taxa. In 1964, *P. brevifolia* was reduced to a variety of *P. barbulata* without comment (Ahles 1964). Afterwards, several studies investigated the appropriate taxonomic status of *P. brevifolia*. An embryological study (Reynolds 1966) concluded that both notable developmental similarities and differences existed between the two species and ultimately relied on the ecological and morphological differences to support the continued recognition of two species. Primack & Wyatt (1975) found correlation between leaf length and soil moisture of *P. brevifolia and P. barbulata* at a single site in South Carolina and concluded that the difference in leaf length between the two species is clinal, suggesting that *P. brevifolia* is simply a morphological variant of *P. barbulata*. More recently, an allozyme study – restricted to the populations from the southern range of the genus – found that the two species share similar levels of genetic diversity, with very little intertaxa genetic differentiation (Godt & Hamrick 1995). However, recent floras for the region

(Weakley 2008; Sorrie *et al.* 2009) have continued to recognize two species, emphasizing the morphological, ecological, and embryological differences between them.

In this study, we use cpDNA sequences, amplified fragment length polymorphism (AFLP) markers, and morphological measurements to investigate the taxonomy and phylogeography of both *P. barbulata* and *P. brevifolia* across the entire range of the genus. We attempt to determine whether clear morphological and genetic differences exist between the two species and whether the morphological, ecological, and embryological variation previously observed in the southern populations of *Pyxidanthera* (in the text, *Pyxidanthera* will refer to both *P. barbulata* and *P. brevifolia*) correlate with greater genetic diversity in the south. Using genetic data, we attempt to distinguish between two plausible phylogeographical scenarios. The first scenario represents typical refugial patterns described for numerous species in ENA; the genus *Pxyidanthera* was isolated in one or more southern refugia during the Pleistocene and subsequently recolonized northern areas after the LGM. Genetic patterns supporting this scenario would include reduced genetic diversity in northern populations (Hewitt 2000), recolonized areas containing only a subset of refugial population alleles (Broyles 1998), and putative refugia having a greater number of rare alleles, which may reflect historical processes better than genetic diversity estimates (Comps *et al.* 2001; Paun *et al.* 2008) Alternatively, the two species in *Pyxidanthera* could have persisted in their present ranges through the later Pleistocene rather than retreating into one or more glacial refugia. Genetic patterns that would suggest this second scenario include no reduction in genetic diversity or rare alleles in northern populations and the presence of alleles restricted to northern populations.

Materials and methods

Sampling and morphological measurements

We collected leaf tissue samples of 423 individuals from 29 *Pyxidanthera brevifolia* populations (defined as all *P. brevifolia* individuals that occurred within 0.75 km of each other) and 178 individuals from 14 *P. barbulata* populations, across the ranges of both species. A priori taxonomic identity was determined based on habitat differences, State Natural Heritage Program records, and geographical region (*P. brevifolia* is restricted to the Sandhills region of North and South Carolina). For each sample, we measured the longest leaf length and width and categorized the leaf pubescence into one of two categories: pubescence covering more than half of the leaf, and pubescence covering half or less of the leaf. We evaluated differences in leaf length and leaf width means between *P. barbulata* and *P. brevifolia* using *t*-tests and evaluated differences in leaf pubescence categories using a chi-square test.

Molecular methods

DNA was extracted from 319 *P. brevifolia* individuals across 17 populations and 157 *P. barbulata* individuals across 14 populations using the CTAB method with minor modifications (Doyle & Doyle 1987). After an initial screening of 16 cpDNA regions known to be highly polymorphic (Shaw *et al.* 2007), we amplified two polymorphic regions – *atpI-atpH* and *psbD-trnT*(GUU) – of 63 and 42 samples from 14 and 10 populations of *P. barbulata* and *P. brevifolia*, respectively, using universal primer pairs (Shaw *et al.* 2007). PCR conditions followed Shaw *et al.* (2005) in 12.5 μL solutions using the following protocol: 1 hold (5 min per 80 °C), 30 cycles [(1 min per 95 °C), (1 min per 50 °C), (4 min per 65 °C)], 1 hold (5 min per 65 °C). PCR products were cleaned prior to sequencing using Antarctic Phosphatase (0.5 Units), Exonuclease I (0.2 Units), and 1 μL 10× Antarctic Phosphatase buffer (New England BioLabs, Ipswich, MA, USA) at 37 °C (15 min) and 80 °C (15 min). We sequenced in the forward direction for the *atpH* and *psbD* regions using the Big Dye 3.1 kit (Applied Biosystems, USA) and analysed the products using an ABI 3730 DNA sequencer (Applied Biosystems, USA). We edited and aligned sequences using Sequencher 4.2.2 (Gene Codes Corporation, Ann Arbor, MI, USA) and MEGA version 4 (Tamura *et al.* 2007). Sequences were submitted to GenBank under accession numbers nos. HM564379-HM56491 (Table S1).

Amplified fragment length polymorphism markers are appropriate for examining low levels of genetic divergence within and between closely related taxa (Coart *et al.* 2002; McKinnon *et al.* 2008) and have been successfully used in phylogeographical studies (Meudt & Bayly 2008; Perez-Collazos *et al.* 2009). AFLP genotyping followed the multiplexing protocol described by Trybush *et al.* (2006), with the minor modification that the restriction and ligation steps were combined in a single reaction at a total volume of 10 μL. For the preamplification reaction, we used EcoRI+A and *Mse*I+C primers. Three selective primer pairs were chosen after a trial based on the number of reproducible polymorphic markers produced: Eco-ACC/*Mse*I-CAT (Hex), Eco-ATG/*Mse*I-CAT (FAM), and Eco-AGG/*Mse*I-CAT (NED). Selective amplification products were separated and analysed using an ABI 3730 DNA sequencer (Applied Biosystems, USA) and automatically scored with Genemarker version 1.8 (Softgenetics LCC, State College, PA, USA) using the default settings, with the

exception that we normalized the FAM-dyed markers and set the allele evaluation peak score to 'pass' if it was ≥1 (Holland *et al*. 2008). The reproducibility of the AFLP profiles was evaluated by running eight duplicate samples for each 96-well plate. Error rates between duplicates were calculated using a Euclidean distance measure (Bonin *et al*. 2004). To reduce the error rate, we removed bands with 10 or more errors when comparing duplicate samples (Zhang *et al*. 2010).

cpDNA data analysis

Because the chloroplast represents a single nonrecombining locus, sequences of the two sampled regions were concatenated. We recoded insertions or deletions (indels) that did not violate the assumptions of the infinite sites model (Kimura 1969) as identified by SNAP Workbench (Price & Carbone 2005). We performed three separate tests of neutrality to test for evidence of population expansion or selection in the cpDNA – Fu and Li's D* and F*(Fu & Li 1993) and Fu's F_S (Fu 1997) – using SNAP Workbench. Fu and Li's D* and F* neutrality tests are more powerful for detecting background selection, while Fu's F_s is more powerful for detecting population growth (Ramos-Onsins & Rozas 2002). We estimated an unrooted haplotype network using the haploNet function as implemented in the pegas package (Paradis 2009) in R (R Development Core Team 2009). This package implements the statistical parsimony method for network reconstruction (Templeton *et al*. 1992). We performed two analyses of molecular variance (AMOVA) using Arlequin version 3.01 (Excoffier *et al*. 2005) with the data set hierarchically partitioned by region and individual populations within regions (with two regions defined as New Jersey and New York, hereafter referred to as northern populations, and Virginia, North Carolina, and South Carolina, hereafter referred to as southern populations) and by species (*P. barbulata* and *P. brevifolia*) and populations within species. We estimated the nucleotide genetic diversity (π) (Nei 1987) for each population using DnaSP version 5 (Librado & Rozas 2009).

To test for phylogeographical structure in the data set, we compared two measures of genetic differentiation between populations – G_{ST}, based on haplotype frequency, and N_{ST}, by similarities between haplotype sequences – using PERMUT 2.0 (Pons & Petit 1996) with 1000 permutations. If N_{ST} is significantly greater than G_{ST}, it is taken as evidence of a phylogeographical signal in the data set. To test for isolation by distance (IBD), we performed Mantel tests using the R package vegan (Oksanen *et al*. 2009) between the log-transformed geographical distance matrix and the pairwise population N_{ST} matrix as calculated in DnaSP version 5 (Librado & Rozas 2009) for all populations and for the southern and northern populations separately.

We reconstructed the gene genealogy for the sampled chloroplast regions using Genetree version 9.0 (Bahlo & Griffiths 2000) as implemented in SNAP Workbench (Price & Carbone 2005). We estimated the population mutation rate (θ), using Watterson's method (1975) as calculated in Genetree for both geographical regions and used the average between the two regions as the starting θ. Because of the larger geographical area covered by the two species of *Pyxidanthera* in the southern populations, we assumed a model of unequal population sizes, with the southern population twice as large as the northern population, and nonexponential population growth. We performed ten independent simulations with different starting values of 10^6 iterations, selecting the rooted genealogy and mutation age estimates with the highest probability.

To simultaneously analyse the effects of incomplete lineage sorting and gene flow on the genetic structure of the northern and southern *Pyxidanthera*, we employed an isolation with migration model of population divergence (Nielsen & Wakeley 2001) implemented in the program IMa2 (Hey & Nielsen 2007). IMa2 estimates the following parameters based on the genetic data: θ for all populations (extant and ancestral), migration parameters (m) for gene flow between populations, and t, time in coalescent units since divergence of the extant populations. We performed three independent runs with ten chains each under an infinite sites model with a burn-in period of 150 000 steps. We sampled 500 000 genealogies , saving one genealogy every 100 steps. We evaluated proper mixing based on the absence of trends in plotted parameter estimates and congruence of parameter estimates between runs 100 000 of the 500 000 saved genealogies were combined to evaluate 24 models that were either nested within the full model or that constrained select parameters by setting them equal to each other (e.g. equal migration between populations). We compared the different model posterior probabilities using an information-theoretic approach recently extended to phylogeographical data (Carstens *et al*. 2009). Information theory statistics were calculated according to Burnham and Anderson (2002).

AFLP data analysis

We calculated the percentage of polymorphic loci (P%) and Nei's expected heterozygosity (Nei 1987) using AFLPsurv version 1.0 (Vekemans *et al*. 2002) and the 'frequency down weighted marker score' (DW) (Schönswetter & Tribsch 2005) using the R script AFLPdat (Ehrich 2006); several population genetic diversity

measures were included to ensure consistency between methods. DW is calculated by summing each occurrence of a particular marker in a population and dividing that value by the sum of the marker across all populations. For each population, these values are then averaged across all markers. Populations that have been isolated are expected to accumulate rare markers and thus their DW scores should be higher. We first removed populations that contained fewer than seven samples to minimize effects of low sample size (Bonin *et al.* 2007), leaving a total of 437 samples from 25 populations. We tested for effects of sample size on all the calculated genetic diversity estimates by regressing estimated diversity on sample size, and we also tested for correlation between all possible pairings of the included diversity measures. Populations were grouped according to taxonomic identity and region, and we tested for significant differences between diversity estimates using *t*-tests in R.

Population differentiation and structure were explored by first running an ordination using nonmetric multidimensional scaling (NMDS) to graphically display population pairwise genetic distances (D) (Nei 1972) in a reduced dimensional space using the R package labdsv (Roberts 2010). We included all populations regardless of sample size for the analysis. In addition, we explored population genetic structure using STRUCTURE 2.3.2.1 (Pritchard *et al.* 2000; Falush *et al.* 2007). For K1 through 9, we performed three runs with a burn-in length of 10 000 and post burn-in length of 25 000, assuming admixture and correlated allele frequencies. We determined the most likely number of populations by graphically analysing the model log likelihoods for each K. Because Ln *P*(D) did not increase monotonically to the optimal K (Herrera & Bazaga 2008), we did not use the methods of Evanno *et al.* (2005). Three analyses of molecular variance (AMOVA) were performed using Arlequin version 3.01 (Excoffier *et al.* 2005), with partitioning of the data following the chloroplast AMOVAs. To test for IBD, we performed a Mantel test between the population genetic distance matrix and the log-transformed geographical distance matrix using the R package vegan (Oksanen *et al.* 2009).

Results

Morphology

Pyxidanthera barbulata has significantly longer leaf lengths (6.3 mm vs. 4.5 mm, respectively, $P < 0.001$) and widths (1.9 mm vs. 1.3 mm, $P < 0.001$) compared to *P. brevifolia*, but there is considerable overlap between the two species in both traits (Fig. 1). The variation in leaf length is continuous between *P. barbulata* and *P. brevifolia* with no obvious break, certainly not at the 3.5–4 mm size suggested in taxonomic keys (Sorrie *et al.* 2009). There is a significant difference in leaf pubescence between the two species ($P < 0.001$). All *P. barbulata* had pubescence covering less than half of their leaves, but 49% of *P. brevifolia* also had pubescence covering more than half of their leaves. As with leaf length and width, there is considerable variation within taxa.

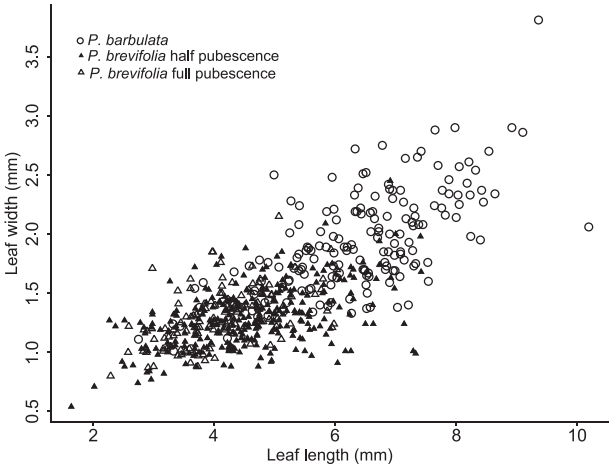

Fig. 1 Morphological variation in leaf length, leaf width, and pubescence of *Pyxidanthera barbulata* (circles) and *P. brevifolia* (triangles). Solid triangles represent *P. brevifolia* specimens that had pubescence for half or less than half of the leaf; open triangles represent *P. brevifolia* specimens with pubescence greater than half of the leaf. Although there are statistically significant differences between the two species for mean leaf length, leaf width, and pubescence, there is considerable overlap in the ranges of these traits between the two species.

cpDNA

The two sampled cpDNA regions for 105 individuals yielded 975 characters, of which 14 were polymorphic (Table S2). The data set included 12 substitutions and two indels that did not violate the infinite sites model. None of the three neutrality tests (Fu and Li's D* and F* and Fu's Fs) were significant ($P > 0.05$), indicating that there is no evidence of either population growth or background selection. A statistical parsimony haplotype tree identified 12 haplotypes (Fig. 2). The interior haplotypes of the network (H7, H11) are geographically widespread compared to the derived haplotypes, which tend to be both less frequent and geographically restricted (Fig. 2; Table S2). Results from Genetree indicate that H11 is the haplotype with the highest probability of being ancestral (average relative likelihood 69.5%); however, both of the interior haplotypes were

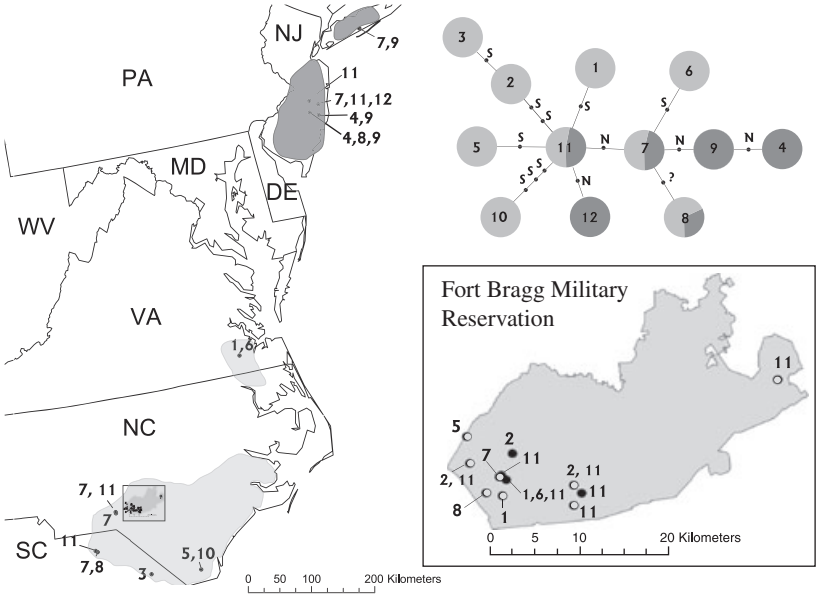

Fig. 2 Geographical distribution (shaded in grey) and statistical parsimony network for 12 haplotypes from 2 cpDNA regions of *Pyxidanthera*. State names are in bold abbreviations and numbers represent haplotypes from Table S2. Black dots in the haplotype network represent mutational steps; associated letters (S for South, N for North) represent the most likely (>95% probability) geographical origin of mutations inferred using Genetree 9.0. Light grey shading of haplotype network represents proportion of the associated haplotype comprised of southern individuals and darker grey shading represents proportion comprised of northern individuals. Inset map: Sampling of *Pyxidanthera* populations on Fort Bragg Military Reservation. *Pyxidanthera barbulata* populations are represented by closed circles and *P. brevifolia* populations are represented by open circles.

almost equally common in the northern and southern populations (Fig. 2). Two of the four haplotypes derived from H7 (H4 and H9) only occur in New Jersey and New York and had the highest probability of a northern origin, while the other two derived haplotypes (H6 and H8) occur in both northern and southern population. Five of the six haplotypes derived from the second interior haplotype, H11, only occur in southern populations and most likely are of southern origin. Only H12 has a higher probability of a northern origin; it is a private haplotype restricted to one northern population.

Region explains a small but statistically significant percentage of the genetic variation when used as the highest grouping variable in a hierarchical AMOVA (17.27%, $P < 0.05$), revealing that genetic variation is not evenly spread across the northern and southern populations (Table 1). Populations within regions explain most of the variation (56.58%, $P < 0.001$); within-population genetic differences and region

Table 1 Analyses of molecular variance (AMOVA) results for *Pyxidanthera barbulata* and *P. brevifolia* using cpDNA sequences and amplified fragment length polymorphism (AFLP) markers

	AFLP			cpDNA		
Source of variation	d.f.	Variance	% of variation	d.f.	Variance	% of variation
Grouped by species						
Between species	1	0.37	1.60**	1	−0.01	−1.71 NS
Among populations within species	23	1.85	7.87***	22	0.63	72.22***
Within populations	412	21.30	90.53***	81	0.26	29.49***
Grouped by region (North vs. South)						
Between regions	1	0.76	3.20**	1	0.17	17.27*
Among populations within regions	23	1.86	7.79***	22	0.56	56.58***
Within populations	412	21.30	89.01***	81	0.26	26.15***

***Indicates *P*-value <0.001, ** *P*-value <0.01, * *P*-value <0.05, and NS indicates nonsignificance of variation.

account for a smaller but still significant percentage of the variation (26.15%, $P < 0.001$). When species is used as the highest grouping variable, AMOVA results demonstrate significant genetic differences among populations (72.22%, $P < 0.001$), but not significant differences between the two species (0%, $P > 0.05$). The nucleotide genetic diversity (π) averages 0.0004 across all specimens with no significant differences between means for either regions or species ($P > 0.05$) (Table 2).

Geographically distant populations are not more differentiated from each other than populations in closer geographical proximity (Fig. 3). N_{ST} is significantly greater than G_{ST} (0.788 vs. 0.695, $P < 0.01$), indicating that there is a phylogeographical signal in the chloroplast data; in other words, haplotypes within populations are more similar to each other than expected. However, Mantel tests for IBD find no significant signal across the range of *Pyxidanthera* populations (Fig. 3, $R = 0.01$, $P = 0.39$). This pattern generally arises when genetic drift exerts more influence than gene flow at the regional scale (Hutchison & Templeton 1999). When northern and southern populations are analysed separately, there is no significant IBD in the northern populations ($R = -0.05$, $P = 0.47$) but there is marginally significant IBD in the southern populations, although the effect is weak ($R = 0.13$, $P = 0.049$).

Table 2 Genetic diversity indices for *Pyxidanthera barbulata* and *P. brevifolia* based on cpDNA sequences and AFLP markers

Population	Species	State	*N*	%P	DW	He	Π	Haplotypes
NC_1	*barbulata*	NC	14 (5)	41.0	12.99	0.15	0.0019	H1,H6,H11
NC_2	*barbulata*	NC	0 (5)	–	–	–	0.0000	H11
NC_3	*barbulata*	NC	0 (4)	–	–	–	0.0000	H2
NC_4	*barbulata*	NC	9 (4)	50.0	7.89	0.14	0.0000	H11
NC_6	*barbulata*	NC	19 (1)	26.5	7.40	0.08	0.0000	H7
NC_8	*barbulata*	NC	12 (4)	23.2	4.09	0.09	0.0005	H7,H11
NC_9	*barbulata*	NC	19 (6)	44.2	18.91	0.15	0.0013	H5,H10
NJ_CB	*barbulata*	NJ	0 (5)	–	–	–	0.0010	H7,H11,H12
NJ_CW	*barbulata*	NJ	17 (5)	34.2	14.41	0.11	0.0012	H4,H8,H9
NJ_WB	*barbulata*	NJ	15 (6)	27.1	9.48	0.10	0.0000	H11
NJ_WG	*barbulata*	NJ	18 (5)	38.7	18.69	0.12	0.0004	H4,H9
NY_1	*barbulata*	NY	0 (5)	–	–	–	0.0004	H7,H9
SC_1	*barbulata*	SC	8 (3)	50.3	12.67	0.16	0.0000	H3
VA_1	*barbulata*	VA	13 (5)	32.6	10.08	0.12	0.0019	H1,H6
SC_HP	*brevifolia*	SC	8 (5)	25.2	1.83	0.08	0.0004	H7,H8
SC_SL	*brevifolia*	SC	26 (5)	26.1	8.33	0.08	0.0000	H11
002A	*brevifolia*	NC	19 (5)	41.0	16.25	0.12	0.0000	H11
10	*brevifolia*	NC	17 (0)	47.4	23.21	0.15	–	–
20	*brevifolia*	NC	21 (5)	51.3	23.17	0.14	0.0008	H2,H11
24	*brevifolia*	NC	19 (0)	46.1	17.26	0.14	–	–
026D	*brevifolia*	NC	19 (1)	39.0	9.66	0.12	0.0000	H1
03_25	*brevifolia*	NC	30 (5)	36.1	15.18	0.11	0.0008	H2,H11
028E	*brevifolia*	NC	0 (4)	–	–	–	0.0000	H7
33	*brevifolia*	NC	33 (0)	31.3	22.40	0.11	–	–
038D	*brevifolia*	NC	15 (0)	41.9	13.65	0.14	–	–
057Y	*brevifolia*	NC	24 (3)	41.9	25.43	0.13	0.0000	H8
058B	*brevifolia*	NC	22 (0)	40.6	18.61	0.14	–	–
065N	*brevifolia*	NC	15 (5)	31.6	7.24	0.11	0.0000	H11
066A	*brevifolia*	NC	0 (4)	–	–	–	0.0000	H11
092B	*brevifolia*	NC	8 (0)	34.2	3.26	0.11	–	–
93_115	*brevifolia*	NC	17 (0)	43.2	20.41	0.14	–	–
Overall mean				37.8	13.70	0.12	0.0004	
barbulata				36.8	11.66	0.12	0.0006	
brevifolia				38.5	15.06	0.12	0.0002	
Northern				33.3	14.20	0.11	0.0006	
Southern				38.5	13.63	0.12	0.0004	

Amplified fragment length polymorphism (AFLP) genetic diversity indices were only calculated for populations with more than seven genotyped individuals (437 total specimens). %P represents the number of polymorphic loci, DW is a measure of rare alleles per population, and H*e* is a measure of expected heterozygosity based on the AFLP markers. π is a measure of cpDNA nucleotide diversity. *N* represents the number of specimens for each population for AFLP markers and cpDNA sequences (in parentheses).

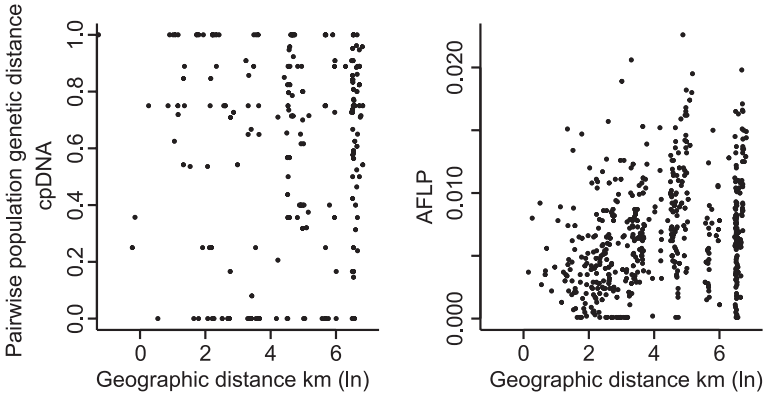

Fig. 3 Isolation by distance for cpDNA (left side) and amplified fragment length polymorphism (AFLP) (right side) markers for *Pyxidanthera barbulata* and *P. brevifolia* populations across the range of the genus *Pyxidanthera*. cpDNA data demonstrates no genetic isolation by geographical distance ($R = 0.01$, P-value = 0.39), while AFLP markers demonstrate weak but significant ($R = 0.27$, P-value = 0.02) isolation by distance at shorter distances with effects of genetic drift more evident at greater distances.

The highest posterior density for θ_{South} was higher than both $\theta_{Ancestral}$ and θ_{North}, although there is significant overlap between the 95% confidence intervals (Fig. S1, Supporting Information). The highest posterior density parameter estimate for migration from south to north is 2.09, while the estimate for migration from northern populations into southern populations is 0.01, indicating there has been gene flow between the two regions, with possibly greater migration from the southern populations into the northern. Past gene flow between the two populations is also supported by the model selection exercise; the worst-performing models constrained both migration parameters to 0 (Table S3). Time since divergence, t, was poorly estimated and failed to converge; this typically reflects a lack of a signal available in analyses that incorporate only a single locus with limited informative characters (J. Hey, personal communication).

AFLP

Three hundred and ten polymorphic bands were scored based on the three primer pairs. Each individual produced a unique AFLP profile, and the Euclidean error rate (based on 47 replicate pairs) was 4.2%, within the margin of acceptable error rates (Bonin *et al.* 2004). All genetic diversity indices were highly correlated, and sample size was not significantly correlated with any of the genetic diversity values (all $P > 0.05$). The population genetic diversity estimates for *P. barbulata* and *P. brevifolia* populations do not differ significantly for %P, DW, or H_e (all $P > 0.05$). In addition, there are no significant differences between regional genetic diversity estimates for percentage %P, H_e, or DW (all $P > 0.05$) (Table 1). The percentage of polymorphic loci (%P) ranges from 23.2% to 51.3%, with a mean of 37.8%, while Nei's population genetic diversity (H_e) ranges from 0.08 to 0.16 with a mean of 0.12 (Table 1).

Nonmetric multidimensional scaling ordination based on the population genetic distances (D) reveals no discrete grouping of populations based on either region or species (Fig. 4). Results from STRUCTURE also demonstrate little population genetic structure based on either geographical location or taxonomic identification (Fig. S2), and there was no graphical evidence for an optimal number of K distinct genetic groups (Fig. S3). The hierarchical AMOVAS grouped according to species (*P. barbulata* vs. *P. brevifolia*) and geographical region (North vs. South) found small but significant variation was explained by species (1.60%, $P < 0.01$) and region (3.20%, $P < 0.01$)(Table 1), while within-population variation remained high (90.53% and 89.01%, respectively, $P < 0.001$). There is evidence for a weak but significant effect of IBD in the AFLP data (Fig. 3; $R = 0.27$, $P = 0.02$). Genetic differentiation between populations increases with geographical distance, indicating low to moderate levels of short distance gene flow but little evidence of long-distance

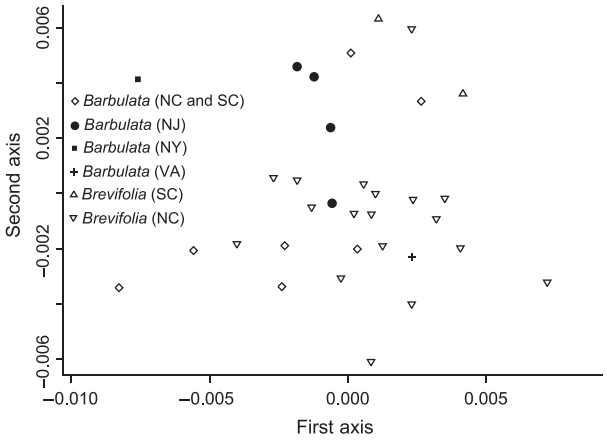

Fig. 4 Nonmetric multidimensional scaling ordination of *Pyxidanthera barbulata* and *P. brevifolia* population genetic distances (Nei's D) based on amplified fragment length polymorphism markers. In the legend, letters in parentheses represent US states. Little separation is evident among populations defined according to either taxonomic identity or geographical location.

gene flow; both gene flow and genetic drift influence the pattern depending on the geographical scale (Hutchison & Templeton 1999). At shorter distances, gene flow is dominant, increasing the correlation between genetic and geographical distances, while at greater distances genetic drift predominates.

Discussion

Taxonomy

There does not appear to be clear separation between *Pyxidanthera barbulata* and *P. brevifolia* based on either the morphological or genetic data (Figs 1 and 4). Although *P. brevifolia* in general has shorter and narrower leaves than the more widespread *P. barbulata*, there is significant overlap between the two species for both leaf length and width. Previously published work on the leaf morphology indicated that the differentiation between the two species was because of hydrological differences between the habitats that *P. barbulata* and *P. brevifolia* occupy (Primack & Wyatt 1975), with leaf length increasing continuously with increasing soil moisture. Although *P. brevifolia* individuals tend to be more pubescent than *P. barbulata* individuals (Fig. 1), there is significant variation in the pubescence of *P. brevifolia* both at the taxonomic level and within-populations (data not shown), with both glabrous and pubescent individuals represented in most populations. Interestingly, there are herbarium specimens of *P. barbulata* from xeric habitats of the Outer Coastal Plain of North Carolina that exhibit the shorter leaves of *P. brevifolia* specimens, but that are not pubescent; the extreme pubescence appears to be restricted to *P. brevifolia*.

Although several authors have suggested that *P. brevifolia* may represent a preadapted *P. barbulata* ecotype that moved into the Sandhills region from the Outer Coastal Plain (Wells & Shunk 1931; Primack & Wyatt 1975), the current study using cpDNA sequences and AFLP markers and a previous study using allozymes (Godt & Hamrick 1995) do not support this hypothesis. *P. barbulata* and *P. brevifolia* populations in the Sandhills are not genetically distinct from each other, with *P. barbulata* populations on Fort Bragg sharing cpDNA haplotypes with nearby *P. brevifolia* populations (Fig. 2 inset). In addition, there is no separation between *P. barbulata* and *P. brevifolia* populations in their AFLP profiles (Fig. 4). We cannot rule out the possibility that *P. brevifolia* is a recently derived ecotype of *P. barbulata*, restricted to the Sandhills, and that a few mutations have led to local adaptation, but this would need to have been recent enough that genetic differentiation is not apparent in AFLP profiles. Even though *P. brevifolia* appears to be an extreme morphological variant of *P. barbulata* associated with sandy, xeric sites, in our estimation, it warrants continued active management – specifically the regular prescribed fire schedule that Fort Bragg Military Reservation maintains – and further study because of its potentially critical role as an early season pollen and nectar provider and as a system for studying physiological adaptation to drought stress and phenotypic plasticity.

Phylogeography of the genus Pyxidanthera

Contrary to the well-documented trends of range contraction observed in many temperate plant species during the last glacial period in ENA, we found little evidence for either a southern refugium or range expansion following the LGM in the genus *Pyxidanthera*. Genetic diversity estimates for both the AFLP and cpDNA markers were not significantly different for northern and southern *P. barbulata* populations (Table 1), and northern populations contained several cpDNA haplotypes that did not occur in the southern populations (Fig. 2). More pointedly, estimates of the number of rare AFLP markers (DW), which may be more helpful in identifying refugial phylogeographical patterns (Paun *et al.* 2008), did not demonstrate significant differences between northern and southern populations. Finally, the two interior haplotypes – H7 and H11 – were widespread in both northern and southern populations with comparable frequencies (Fig. 2). These genetic patterns are contrary to what would be expected if there was a southern refugium for *Pyxidanthera* (Comps *et al.* 2001; Ikeda *et al.* 2008; Paun *et al.* 2008). Thus, it appears that the most likely scenario includes range stasis through the later Pleistocene. Furthermore, evidence of gene flow between geographically close populations suggests a possible explanation for low levels of genetic differentiation between northern and southern populations; these populations may not have been as geographically isolated in the recent past, with populations in the intervening area facilitating gene flow.

Several studies of tree species have also demonstrated the absence of typical refugial patterns (Palme *et al.* 2003; Maliouchenko *et al.* 2007), indicating that some species may have persisted closer to the ice sheet than previously thought. There is increasing evidence for 'cryptic refugia' in more northern latitudes for a number of mammal and plant species (Stewart & Lister 2001). Although mid-latitude refugia are possible, several alternatives have also been put forth. In the case of *Salix caprea*, which demonstrates little phylogeographical patterning, Palme *et al.* (2003) posit

high rates of dispersal, hybridization with other *Salix* species, and high mutation rates as possible reasons. These explanations are not very probable in the case of *Pyxidanthera*. *Pyxidanthera* seeds lack obvious morphological adaptations for dispersal, although ants have been observed transporting seeds (W. Wall, personal observation). Hybridization with other species is implausible, because *Pyxidanthera* is well differentiated from all other taxa within Diapensiaceae (Ronblom & Anderberg 2002). Although we have not estimated mutation rates, this alone would not generate the observed patterns.

That *P. barbulata* would persist, rather than retreat, during the climatic oscillations of the Pleistocene is consistent not only with the genetic data but also with our knowledge of Pleistocene habitats and the species' natural history. The GACP physiographical region, relative to more interior physiographical regions, may have been climatically buffered during the Pleistocene because of the moderating influence of the Atlantic Ocean (Rahmstorf 2002); moderation of climatic extremes could have allowed persistence closer to the ice sheet during glacial periods for some GACP species. Still, *P. barbulata* populations in New Jersey and New York would have experienced much colder conditions through much of the last glacial period (Jacobson *et al.* 1987; French *et al*. 2003, 2007). The vegetation community of the late Pleistocene in some of the areas of ENA does not have a modern analogue; most likely, it would have been a relatively open spruce (*Picea* spp.) forest with an herbaceous understory dominated by *Carex* spp. (Overpeck *et al*. 1992). The most important factors in determining the ecological niche of *P. barbulata* may be high light levels and an absence of competition, rather than temperature or moisture. The frequently burned habitats of the Sandhills of North and South Carolina and the Pine Barrens of New Jersey provide this habitat; it is conceivable that environments near the glacial boundary that lacked a dominant canopy cover during the last glacial period did as well. Finally, lower sea levels during glacial periods may have periodically increased available habitats for Atlantic Coastal Plain species such as *P. barbulata* on the exposed continental shelf (Hobbs 2004).

The present-day disjunction in the range of *P. barbulata* may be related to regional geomorphology. The Atlantic Coastal Plain is characterized by a series of alternating arches and embayments (Ward 1992); *Pyxidanthera* populations occur on the Cape Fear, Norfolk, and South New Jersey Arches, but are absent in the intervening Salisbury Embayment. The current disjunction in the range of the genus *Pyxidanthera* may be the result of oscillating sea levels that inundate embayment areas while arches remain above sea level (Bloom 1983; Sorrie & Weakley 2001). It is unlikely that long-distance gene flow between the northern and southern populations without intermediate populations would be high enough to prevent genetic differentiation. This suggests that the current vicariance between northern and southern populations may be recent and that during periods of relatively low sea levels, suitable habitat was exposed on the continental shelf, connecting northern and southern populations and allowing gene flow to minimize genetic differentiation.

The GACP floristic province contains the second-highest level of endemism in North America north of Mexico, yet the endemic plant species have been relatively understudied. Despite subtle topographic variation across the region, complex vegetation patterns exist and the biogeographical processes involved elude simple characterization. Although more phylogeographical studies of GACP endemic plant species are needed to determine whether the recent phylogeographical history of the genus *Pyxidanthera* is representative of multiple taxa or is simply an isolated case, it is apparent that the simple refugial model cannot account for the phylogeographical pattern in the genus *Pyxidanthera*. If similar phylogeographical patterns are found in similarly distributed GACP endemics, it would suggest a common mechanism was responsible and the remaining challenge would be to explain why only these taxa were thusly affected. Refugia are generally thought of as existing in the past; it could be the case that contemporary distributional patterns represent modern-day refugia for many Atlantic Coastal Plain endemic plant species.

Acknowledgments

We thank Fort Bragg Military Reservation and the Endangered Species Branch for logistic support and the Construction Engineering Research Laboratory (US Army Corps of Engineers Agreement #W9132T-07-2-0019) and the Native Plant Society for funding. We also thank the following individuals and laboratories for assistance: Consuelo Arellano, Ignazio Carbone, Matthew Cleary, Emile Devito, Janet Gray, Jacob Hilton, Xiang Liu, Bert Pittman, Bruce Sorrie, Andrew Walker, Alan Weakley, Steve Young and the Xiang and Hoffmann labs.

References

Ahles HE (1964) New combinations for some vascular plants of southeastern United States. *Journal of the Elisha Mitchell Scientific Society*, **80**, 172–173.

Bahlo M, Griffiths RC (2000) Inference from gene trees in a subdivided population. *Theoretical Population Biology*, **57**, 79–95.

Bloom AL (1983) Sea level and coastal changes. *Late-Quaternary Environments of the United States*, **2**, 42–51.

Bonin A, Bellemain E, Bronken Eidsen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261–3273.

Bonin A, Ehrich D, Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, **16**, 3737–3758.

Broyles SB (1998) Postglacial migration and the loss of allozyme variation in northern populations of Asclepias exaltata (Asclepiadaceae). *American Journal of Botany*, **85**, 1091–1097.

Buchanan M, Finnegan JT (2008) *Natural Heritage Program List of Rare Plant Species of North Carolina*, North Carolina Natural Heritage Program, Raleigh, NC.

Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, Verlag.

Carstens BC, Stoute HN, Reid NM (2009) An information-theoretical approach to phylogeography. *Molecular Ecology*, **18**, 4270–4282.

Church SA, Kraus JM, Mitchell JC, Church DR, Taylor DR (2003) Evidence for multiple Pleistocene refugia in the postglacial expansion of the eastern tiger salamander, Ambystoma tigrinum tigrinum. *Evolution*, **57**, 372–383.

Coart E, Lamote V, De Loose M, Van Bockstaele E, Lootens P, Roldán-Ruiz I (2002) AFLP markers demonstrate local genetic differentiation between two indigenous oak species [Quercus robur L. and Quercus petraea (Matt.) Liebl.] in Flemish populations. *Theoretical and Applied Genetics*, **105**, 431–439.

Comes HP, Kadereit JW (1998) The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, **3**, 432–438.

Comps B, Gomory D, Letouzey J, Thiébaut B, Petit RJ (2001) Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics*, **157**, 389–397.

Cunningham CW, Collins TM (1994) Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. In: *Molecular Ecology and Evolution: Approaches and Applications* (eds B Schierwater, B Streit, P Wagner, R DeSalle). pp. 405–433, Birkhauser Verlag, Basel, Switzerland.

Delcourt PA, Delcourt HR (1981) Vegetation maps for eastern North America: 40,000 yr BP to the present. In: *Geobotany II* (ed RC Romans). pp. 123–165, Plenum, New York.

Dorken ME, Barrett SCH (2004) Chloroplast haplotype variation among monoecious and dioecious populations of Sagittaria latifolia (Alismataceae) in eastern North America. *Molecular Ecology*, **13**, 2699–2707.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical bulletin*, **19**, 11–15.

Ehlers J, Gibbard PL (Eds.) (2004) *Quaternary Glaciations : Extent and Chronology*. Elsevier, Amsterdam.

Ehrich D (2006) AFLPdat: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes*, **6**, 603–604.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Evans MEK, Dolan RW, Menges ES, Gordan DR (2000) Genetic diversity and reproductive biology in Warea carteri (Brassicaceae), a narrowly endemic Florida scrub annual. *American Journal of Botany*, **87**, 372–381.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, **1**, 47–50.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Fenneman NM (1938) *Physiography of Eastern United States*. McGraw-Hill, New York.

Flora of North America Editorial Committee (1993) *Flora of North America North of Mexico*. Oxford University Press, New York.

French HM, Demitroff M, Forman SL (2003) Evidence for late-pleistocene permafrost in the New Jersey Pine Barrens (Latitude 39 degrees N), Eastern USA. *Permafrost and Periglacial Processes*, **14**, 259–274.

French HM, Demitroff M, Forman SL, Newell WL (2007) A chronology of Late-Pleistocene permafrost events in southern New Jersey, eastern USA. *Permafrost and Periglacial Processes*, **18**, 49–59.

Frost C (2006) History and future of the longleaf pine ecosystem. In: *The Longleaf Pine Ecosystem: Ecology, Silvaculture, and Restoration* (eds S Jose, EJ Jokela, DL Miller). pp. 9–42, Springer Science, New York.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Godt MJW, Hamrick JL (1995) Low levels of allozyme differentiation between Pyxidanthera (pyxie-moss) taxa (Diapensiaceae). *Plant Systematics and Evolution*, **195**, 159–168.

Gonzales E, Hamrick JL, Chang S (2008) Identification of glacial refugia in south-eastern North America by phylogeographical analyses of a forest understorey plant, Trillium cuneatum. *Journal of Biogeography*, **35**, 844–852.

Hafner MS, Nadler SA (1990) Cospeciation in host-parasite assemblages: comparative analysis of rates of evolution and timing of cospeciation events. *Systematic Biology*, **39**, 192–204.

Herrera CM, Bazaga P (2008) Population-genomic approach reveals adaptive floral divergence in discrete populations of a hawk moth-pollinated violet. *Molecular Ecology*, **17**, 5378–5390.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences*, **104**, 2785–2790.

Hobbs III CH (2004) Geological history of Chesapeake Bay, USA. *Quaternary Science Reviews*, **23**, 641–661.

Holland BR, Clarke AC, Meudt HM (2008) Optimizing automated AFLP scoring parameters to improve phylogenetic resolution. *Systematic Biology*, **57**, 347–366.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the

relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898–1914.

Ikeda H, Senni K, Fujii N, Setoguchi H (2008) Survival and genetic divergence of an arctic-alpine plant, Diapensia lapponica subsp obovata (Fr. Schm.) Hulten (Diapensiaceae), in the high mountains of central Japan during climatic oscillations. *Plant Systematics and Evolution*, **272**, 197–210.

Jacobson GL, Webb III T, Grimm EC (1987) Patterns and rates of vegetation change during the deglaciation of eastern North America. In: *North America and Adjacent Oceans during the Last Deglaciation. The Geology of North America* (eds WF Ruddiman, HE Wright Jr). pp. 277–288, Geological Society of America, Boulder, Colorado.

Kimura M (1969) The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, **61**, 893–903.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Magni CR, Ducousso A, Caron H, Petit RJ, Kremer A (2005) Chloroplast DNA variation of Quercus rubra L in North America and comparison with other Fagaceae. North America and comparison with other Fagaceae. *Molecular ecology*, **14**, 513–524.

Maliouchenko O, Palme AE, Buonamici A, Vendramin GG, Lascoux M (2007) Comparative phylogeography and population structure of European Betula species, with particular focus on B. pendula and B. pubescens. *Journal of Biogeography*, **34**, 1601–1610.

McKinnon GE, Vaillancourt RE, Steane DA, Potts BM (2008) An AFLP marker approach to lower-level systematics in Eucalyptus (Myrtaceae). *American Journal of Botany*, **95**, 368–380.

McLachlan JS, Clark JS, Manos PS (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology*, **86**, 2088–2098.

Meudt HM, Bayly MJ (2008) Phylogeographic patterns in the Australasian genus Chionohebe (Veronica s.l., Plantaginaceae) based on AFLP and chloroplast DNA sequences. *Molecular Phylogenetics and Evolution*, **47**, 319–338.

Morris AB, Ickert-Bond SM, Brunson DB, Soltis DE, Soltis PS (2008) Phylogeographical structure and temporal complexity in American sweetgum (Liquidambar styraciflua; Altingiaceae). *Molecular Ecology*, **17**, 3889–3900.

Mylecraine KA, Kuser JE, Smouse PE, Zimmermann GL (2004) Geographic allozyme variation in Atlantic white-cedar, Chamaecyparis thyoides (Cupressaceae). *Canadian Journal of Forest Research*, **34**, 2443–2454.

Nei M (1972) Genetic distance between populations. *American Naturalist*, **106**, 283–292.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Oksanen J, Kindt R, Legendre P *et al.* (2009) *vegan: Community Ecology Package*.

Oliveira LO, Huck RB, Gitzendanner MA, Judd WS, Soltis DE, Soltis PS (2007) Molecular phylogeny, biogeography, and systematics of Dicerandra (Lamiaceae), a genus endemic to the southeastern United States. *American Journal of Botany*, **94**, 1017–1027.

Overpeck JT, Webb RS, Webb III T (1992) Mapping eastern North American vegetation change of the past 18 ka: No-analogs and the future. *Geology*, **20**, 1071.

Palme AE, Semerikov V, Lascoux M (2003) Absence of geographical structure of chloroplast DNA variation in sallow, Salix caprea L. *Heredity*, **91**, 465–474.

Paradis E (2009) pegas: Population and Evolutionary Genetics Analysis System. *R package version 0.3*.

Paun O, Schönswetter P, Winkler M, Intrabiodiv C, Tribsch A (2008) Historical divergence vs. contemporary gene flow: evolutionary history of the calcicole Ranunculus alpestris group (Ranunculaceae) in the European Alps and the Carpathians. *Molecular Ecology*, **17**, 4263–4275.

Perez-Collazos E, Sanchez-Gomez P, Jimenez JF, Catalan P (2009) The phylogeographical history of the Iberian steppe plant Ferula loscosii (Apiaceae): a test of the abundant-centre hypothesis. *Molecular Ecology*, **18**, 848–861.

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144**, 1237–1245.

Price EW, Carbone I (2005) SNAP: workbench management tool for evolutionary population genetic analysis. *Bioinformatics*, **21**, 402–404.

Primack RB, Wyatt R (1975) Variation and taxonomy of Pyxidanthera (Diapensiaceae). *Brittonia*, **27**, 115–118.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

R Development Core Team (2009) R: a Language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 (available from http://www.R-project.org).

Rahmstorf S (2002) Ocean circulation and climate during the past 120,000 years. *Nature*, **419**, 207–214.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Reynolds JD (1966) *Morphological* Studies in Diapensiaceae. The Embryology of Pyxidanthera Mich, PhD thesis, University of South Carolina.

Roberts DW (2010) *labdsv: Ordination and Multivariate Analysis for Ecology. R package version 1.4-1*.

Ronblom K, Anderberg AA (2002) Phylogeny of Diapensiaceae based on molecular data and morphology. *Systematic Botany*, **27**, 383–395.

Schafale MP, Weakley AS (1990) *Classification of the Natural Communities of North Carolina: Third Approximation*. North Carolina Natural Heritage Program, Raleigh, NC.

Schönswetter P, Tribsch A (2005) Vicariance and dispersal in the alpine perennial Bupleurum stellatum L. (Apiaceae). *Taxon*, **54**, 725–732.

Shaw J, Lickey EB, Beck JT *et al.* (2005) The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Shaw J, Lickey EB, Schill EE, Small RL (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.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261–4293.

Sorrie BA, Weakley AS (2001) Coastal plain vascular plant endemics: phytogeographic patterns. *Castanea*, **66**, 50–82.

Sorrie BA, Gray JB, Crutchfield PJ (2006) The vascular flora of the longleaf pine ecosystem of Fort Bragg and Weymouth Woods, North Carolina. *Castanea*, **71**, 129–161.

Sorrie BA, Weakley AS, Neesom GL (2009) *Pyxidanthera*. In: *Flora of North America North of Mexico Volume 3* (eds Flora of North America Editorial Committee). pp. 336–337, Oxford University Press, New York.

Stewart JR, Lister AM (2001) Crypitc refugia and the origins of the modern biota. *Trends in Ecology and Evolution*, **16**, 608–613.

Takhtajan A (1986) *Floristic Regions of the World*. University of California Press, Berkeley, CA.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Trybush S, Hanley S, Cho K *et al.* (2006) Getting the most out of fluorescent amplified fragment length polymorphism. *Canadian Journal of Botany*, **84**, 1347–1354.

Vekemans X, Beauwens T, Lemaire M, Roldán-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139–151.

Ward LW (1992) Tertiary molluscan assemblages from the Salisbury embayment of Virginia. *Virginia Journal of Science*, **43**, 85–100.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256.

Weakley AS (2008) *Flora of the Carolinas, Virginia, Georgia, northern Florida, and surrounding areas* (*Working Draft of 7 April 2008*).

Wells BW (1929) A new pixie from North Carolina. *Journal of the Elisha Mitchell Scientific Society*, **44**, 238–239.

Wells BW, Shunk IV (1931) The vegetation and habitat factors of the coarser sands of the North Carolina coastal plain: an ecological study. *Ecological Monographs*, **1**, 465–520.

Zhang YY, Zhang DY, Barrett SCH (2010) Genetic uniformity characterizes the invasive spread of water hyacinth (*Eichhornia crassipes*), a clonal aquatic plant. *Molecular Ecology*, **19**, 1774–1786.

W.A.W. is a plant ecologist interested in the interplay between population genetics and population dynamics of rare and endemic species. N.A.D. conducts research on phylogenetics and phylogeography of arid adapted and edaphic endemic plants. Q.Y.X.'s research explores the mechanisms underlying biodiversity patterns and morphological variation, with emphases on the Cornales. W.A.H. is interested in ecology and conservation of savanna ecosystems. T.R.W. is a vegetation scientist with research interests in vegetation/environment relationships and biological diversity. M.G.H. uses genetic approaches to inform rare species conservation.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Parameter estimates for θ (southern, northern, and ancestral populations), time since divergence, and migration (gene flow) between northern and southern populations of the genus *Pyxidanthera* based on results from IMa2.

Fig. S2 Population genetic structure for the genus *Pyxidanthera* as inferred from the program STRUCTURE for K2 through 9.

Fig. S3 Log likelihood [Ln *P*(D)] and standard deviation results from program STRUCTURE for K1 through 9.

Table S1 Chloroplast haplotype accession numbers as archived in Genbank for the *atpI-atpH* intergenic spacer region (partial sequence) and the *psbD-trnT* intergenic spacer region (partial sequence)

Table S2 Polymorphisms of the 12 cpDNA haplotypes based on the cpDNA regions *atpI-atpH* and *psbD-trnT* in the genus *Pyxidanthera*

Table S3 Summary of model statistics for the 24 IMa2 models.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

American Journal of Botany 94(5): 856–872. 2007.

MOLECULAR PHYLOGENY OF NYCTAGINACEAE: TAXONOMY, BIOGEOGRAPHY, AND CHARACTERS ASSOCIATED WITH A RADIATION OF XEROPHYTIC GENERA IN NORTH AMERICA[1]

NORMAN A. DOUGLAS[2] AND PAUL S. MANOS

Department of Biology, Duke University, Box 90338, Durham, North Carolina 27708 USA

The four o'clock family (Nyctaginaceae) has a number of genera with unusual morphological and ecological characters, several of which appear to have a "tendency" to evolve repeatedly in Nyctaginaceae. Despite this, the Nyctaginaceae have attracted little attention from botanists. To produce a phylogeny for the Nyctaginaceae, we sampled 51 species representing 25 genera (of 28–31) for three chloroplast loci (*ndhF*, *rps16*, *rpl16*, and nrITS) and included all genera from North America. Parsimony, likelihood, and Bayesian methods were used to reconstruct the phylogeny for the family. The family is neotropical in origin. A radiation of woody taxa unites *Pisonia* and *Pisoniella* with the difficult tropical genera *Neea* and *Guapira*, which also form a clade, though neither appears to be monophyletic. This group is sister to a clade containing *Bougainvillea*, *Belemia*, and *Phaeoptilum*. A dramatic radiation of genera occurred in the deserts of North America. The tribe Nyctagineae and its subtribes are paraphyletic, due to over-reliance on a few homoplasious characters, i.e., pollen morphology and involucre presence. Two notable characters associated with the desert radiation are cleistogamy and edaphic endemism on gypsum soils. We discuss evolutionary trends in these traits in light of available data about self-incompatibility and gypsum tolerance in Nyctaginaceae.

Key words: biogeography; cleistogamy; gypsophily; homoplasy; mating system; Nyctaginaceae; phylogeny; pollen morphology.

Nyctaginaceae Juss. is a family of 28–31 genera and 300–400 species, that contains the familiar cultivated four o'clocks (*Mirabilis jalapa*) and bougainvillea (*Bougainvillea* spp.). Nyctaginaceae has long been known to be one of the core groups of families of Caryophyllales (Centrospermae) on the basis of the presence of betalain pigments, free-central placentation, p-type sieve tube elements, and the presence of perisperm, as well as molecular evidence (Bittrich and Kühn, 1993; Bremer et al., 2003). Within this group, the modern consensus is that Nyctaginaceae are closely related to certain monocarpellate members of a paraphyletic Phytolaccaceae, especially subfam. Rivinoideae (Rodman et al., 1984; Rettig et al., 1992; Downie and Palmer, 1994; Behnke, 1997; Downie et al., 1997; Cuenoud et al., 2002), although *Sarcobatus* (Sarcobataceae) has also been implicated as a close relative of this group (Behnke, 1997; Cuenoud et al., 2002).

Nyctaginaceae have a uniseriate petaloid perianth, usually interpreted as sepalous in origin (Rohweder and Huber, 1974). In most taxa the lower part of the perianth is fleshy or coriaceous and encloses the superior ovary, giving it the appearance of an inferior ovary. This accessory fruit is persistent and accrescent around the mature achene. While technically a diclesium (Bogle, 1974; Spellenberg, 2003), it is typically referred to as an "anthocarp."

Most genera can be recognized on the basis of fruit structure alone. In *Boldoa*, *Cryptocarpus*, and *Salpianthus*, the perianth is persistent but not accrescent, and thus these taxa lack the anthocarp (Bittrich and Kühn, 1993). In *Andradea*, *Leucaster*, and *Reichenbachia*, the perianth is variously accrescent but is not expanded (Bittrich and Kühn, 1993). However, in the remaining genera the anthocarp completely encloses the fruit and takes many forms (Willson and Spellenberg, 1977; Bittrich and Kühn, 1993). In taxa in which anthocarps are ribbed, the 3–10 ribs can be elaborated into wings (*Phaeoptilum*, *Grajalesia, Tripterocalyx*, *Abronia*, and some *Colignonia*, *Acleisanthes*, and *Boerhavia*), covered by viscid glandular hairs or warts (*Pisonia*, *Pisoniella*, *Cyphomeris*, *Commicarpus*, and some *Boerhavia* and *Acleisanthes*), or unelaborated, to leave an essentially gravity-dispersed fruit (*Mirabilis*, *Anulocaulis*, *Nyctaginia*, and some *Colignonia* and *Boerhavia*). Fleshy anthocarps are probably bird-dispersed in *Neea* and *Guapira*. They are also found in *Okenia*, though this genus is geocarpic and the seeds generally germinate at the spot where they are "planted" by the maternal individual (N. Douglas, personal observation). The unusual anthocarps of *Allionia* are boat-shaped, with two rows of inward-pointing teeth lining the concave side, suggesting possible exozoochory or wind dispersal, though no observations on this are available. In herbaceous taxa, at least, species-level characters are often found in this structure (Willson and Spellenberg, 1977; Spellenberg, 2003).

The family was treated by Heimerl in *Die Natürlichen Pflanzenfamilien* (Heimerl, 1889, 1934) and by Standley in several papers (Standley, 1909, 1911, 1918, 1931a, b) by which time most of the currently recognized genera had been described. Standley (1931a) formally transferred *Oxybaphus* L'Hér. ex Willd., *Hesperonia* Standl., *Quamoclidion* Choisy, and *Allionella* Rydb. into *Mirabilis*, though this has been overlooked in some floras (e.g., Kearney and Peebles, 1960).

[1] Manuscript received 22 June 2006; revision accepted 28 February 2007.

The authors thank the members of the Manos Lab at Duke University, D. Stone, and E. Christine Davis for discussion and helpful comments on the manuscript; F. Kauff for translations from German; M. Hilda Flores Olvera and P. Hernández-Ledezma for assistance in the field; the curators of DUKE, NMC, and NY for permission to sample from their collections; R. Wilbur, Strybing Arboretum, and the Duke University Botany Greenhouses for providing plant material used in this study; and Grand Canyon National Park for permission to collect. The Department of Biology (Duke), the Consortium in Latin American & Caribbean Studies (Duke), and the National Science Foundation provided funding for this project. The authors especially thank R. Spellenberg, who provided invaluable assistance in many aspects of this project.

[2] Author for correspondence (e-mail: nad@duke.edu)

TABLE 1. Classification of Nyctaginaceae and estimates of species number.

Tribe Subtribe	Genus	Species number	Spellenberg (2003), if different	Distribution
Leucastereae	*Leucaster* Choisy	1		SA
	Reichenbachia Spreng.	2		SA
	Andradea Fr. Allemão	1		SA
	Ramisia Glaz. ex Baillon	1		SA
Boldoeae	*Boldoa* Cav. ex Lagasca	1		SA, CA
	Salpianthus Humb. & Bonpl.	1		CA
	Cryptocarpus H.B.K.	1		SA
Abronieae	*Abronia* Juss. (incl. *Tripterocalyx* Hook. ex Standl.)	33	24*	NA
Nyctagineae				
Colignoniinae	*Colignonia* Endl.	6		SA
	Pisoniella (Heimerl) Standl.	1		SA, CA
Boerhaviinae	*Boerhavia* L.	20	ca. 40	Pantropical/subtropical
	Anulocaulis Standl.	4–5	5	NA
	Cyphomeris Standl.	2		NA
	Commicarpus Standl.	25	30–35	Pantropical/subtropical
	Caribea Alain	1		Cuba
	Acleisanthes A. Gray	7	17**	NA
	Selinocarpus A. Gray (incl. *Ammocodon* Standl.)	10	—	NA, Africa
	Okenia Schldl. & Cham.	1–2		NA, CA
Nyctagininae	*Mirabilis* L.	54	ca. 60	NA, Asia
	Cuscatlania Standl.	1		CA
	Allionia L.	2		NA, CA, SA
	Nyctaginia Choisy	1		NA
Phaeoptilinae	*Phaeoptilum* Radlk.	1		Africa
Bougainvilleeae	*Bougainvillea* Comm. ex Juss.	18		SA
	Belemia Pires	1		SA
Pisonieae	*Pisonia* L.	40	10–50	Pantropical/subtropical
	Guapira Aubl.	70	10–50	SA, CA
	Neea Ruiz & Pavon	83		SA, CA
	Neeopsis Lundell	1		CA
	Cephalotomandra Karst. & Triana	1–3		SA
	Grajalesia Miranda	1		CA

Note: Classification scheme according to Bittrich and Kühn (1993) and estimates of species number. For those genera treated in *Flora of North America* (Spellenberg, 2003), species number reflects newly described species and taxonomic readjustments. SA = South America; CA = Central America; NA = North America. * As 20 spp. *Abronia* and four spp. *Tripterocalyx*. ** Including *Selinocarpus* & *Ammocodon* (Levin, 2002).

Heimerl (1934) synthesized the family as it was known, including in his classification genera that had been recently described by Standley (i.e., *Pisoniella*, *Cuscatlania*). He based his supergeneric classification on a combination of plant habit, indumentum, linear vs. capitate stigma, straight vs. curved embryo, sex distribution, pollen grain morphology, and the occurrence of bracts or involucre (Bittrich and Kühn, 1993; Heimerl, 1934). Bittrich and Kühn (1993) provided the most recent summary of the classification at the tribal and subtribal level (Table 1). Their treatment broadly followed that of Heimerl (1934), adjusting ranks and incorporating genera described after 1934, i.e., *Caribea*. It recognized six tribes, two of which, Pisonieae and Nyctagineae, contain the majority of genera and species (Table 1, Pisonieae: six genera, ca. 200 spp.; Nyctagineae: 14 genera, ca. 100 spp.).

Whereas the bulk of diversity of Pisonieae resides in three highly similar arborescent genera with poorly differentiated species, Nyctagineae sensu Bittrich and Kühn (1993) is a diverse, mainly herbaceous, group recognized largely on the basis of very large (100–200 μm in diameter), pantoporate pollen grains, among the largest known in angiosperms (Stevens, 2001). The original formulation of tribe Mirabileae subtribe Boerhaaviinae (Heimerl, 1934), the antecedent of tribe Nyctagineae, was partly diagnosed by the presence of pantoporate pollen grains. Of four currently recognized subtribes, the Nyctagininae comprises those taxa with involucres, which may be of connate or distinct bracts. In contrast, the largest subtribe, Boerhaviinae, is composed of eight genera united primarily by their lack of involucral bracts. Four of these (*Boerhavia*, *Anulocaulis*, *Cyphomeris*, *Commicarpus*) have occasionally been treated as a single *Boerhavia* (Fosberg, 1978). This seems merely to reflect a preference for fewer large genera, because the four segregate genera are as distinct from each other as any other given pair of genera in the herbaceous group. The others, *Caribea*, *Okenia*, *Acleisanthes*, and *Selinocarpus* (including *Ammocodon*), were placed in Boerhaviinae on the basis of pollen morphology and the absence of involucres subtending flowers or inflorescences (though small subtending bracts may be present). The remaining two subtribes, Colignoniiae and the monospecific Phaeoptilinae, have aberrant morphology compared to Nyctagininae and Boerhaviinae, for example, pollen grains in *Colignonia* and *Pisoniella* are dramatically smaller, and in *Phaeoptilum* they are pantocolpate. In *Pisoniella*, the embryo is straight, typical of Pisoneae, instead of a hooked embryo that encircles the perisperm as found in the remaining Nyctagineae (Bittrich and Kühn, 1993). Additionally, the shrubby, scandent or lianoid growth habits of these taxa are rare in the other subtribes, which are mostly perennial herbs. Though Heimerl placed *Colignonia* in a monogeneric tribe Colignonieae, Bittrich and Kühn (1993) include subtribe Colignoniinae (including *Pisoniella*) in tribe Nyctagineae, uniting all taxa with

pantoporate grains and *Phaeoptilum* (with its pantocolpate grains) in one tribe.

Two major centers of distribution have been noted for the Nyctaginaceae (Standley, 1909). The first is in the neotropics and Caribbean, characterized by arborescent genera such as *Neea*, *Guapira*, *Pisonia*, and *Bougainvillea*, as well as the herbaceous *Colignonia* and *Salpianthus*. The second is in arid western North America, where several herbaceous or suffrutescent genera are native, including *Boerhavia*, *Mirabilis*, *Abronia*, *Acleisanthes* sensu Levin (2002), and *Commicarpus*. A few genera are widespread in tropical and subtropical regions of the world (*Boerhavia*, *Commicarpus*, *Pisonia*): *Mirabilis* is present in North and South America with one species in Asia, and *Acleisanthes* contains the disjunct *A. somalensis* from Somalia. *Mirabilis* (*M. jalapa*, *M. oxybaphoides*) and *Bougainvillea* (*B. glabra*, *B. spectabilis*, *B. peruviana*, and numerous hybrid cultivars) are naturalized in many parts of the world. Only one genus is restricted to the Old World, the monospecific *Phaeoptilum* of southwestern Africa.

The first molecular phylogenetic study of Nyctaginaceae was presented by Levin (2000). The focus was on species in certain genera of tribe Nyctagineae sensu Bittrich and Kühn (1993), including genera in subtribes Nyctagininae (*Allionia*, *Mirabilis*) and Boerhaviinae (*Acleisanthes*, *Selinocarpus*, *Boerhavia*), as well as *Abronia* and *Pisonia*. The study justified the formal combination of *Acleisanthes*, *Selinocarpus*, and *Ammocodon* (Levin, 2002; Spellenberg and Poole, 2003), but due to limited sampling of genera, it was not possible to evaluate the monophyly of the subtribes of Nyctagineae (Levin, 2000). The *Flora of North America* treatment of Nyctaginaceae (Spellenberg, 2003), while not referring to tribal classification, reflected these and other taxonomic changes for the genera and species that occur in North America north of Mexico (Table 1).

In the herbaceous taxa of Nyctaginaceae found in the deserts of North America, several unusual characters occur with notable frequency. As indicated by the common name for the family, species in several genera (*Anulocaulis*, *Cyphomeris*, *Acleisanthes*, *Mirabilis*, *Abronia*, and *Tripterocalyx*) flower in the evening and are adapted to moth pollination (Baker, 1961; Grant, 1983; Grant and Grant, 1983; Hernández, 1990; Hodges, 1995; Levin et al., 2001). Internodal bands of viscid secretions, which may discourage aphid colonization (McClellan and Boecklen, 1993), are present in *Anulocaulis*, *Cyphomeris*, and some species of *Boerhavia*. As mentioned, anthocarp morphology is also variable, with wings and viscid glands being common modifications.

Because these characters are often polymorphic at the generic level, they would seem to represent evolutionary "tendencies." Sanderson (1991) discussed evolutionary tendencies in explicit phylogenetic terms: a tendency is a concentrated distribution of homoplasy within a tree. The main objection to the study of tendencies is the difficulty in defining the taxonomic scope at which they operate, in other words, it is "... biologically inappropriate [when investigating a hypothesized tendency] to include taxa that cannot under any circumstances exhibit the states of interest" (Sanderson, 1991, p. 357). Thus, when considering whether a character has a tendency to evolve, it is first necessary to evaluate the range of taxa in which it could potentially appear. In some cases, it may be possible to identify another character upon which the evolution of the character of interest is dependent. If this other trait is itself uniquely derived, its occurrence will define the group in which the tendency may conceivably exhibit itself. If the independent character is itself derived multiple times, then the problem is pushed back so that the challenge is first to explain the tendency for the *independent* character to evolve in the group.

In the case of tendencies in Nyctaginaceae, it is not immediately obvious what sorts of traits may be required to enable, for instance, a shift to nocturnal pollination or the development of viscid bands on stem internodes. There are two traits, however, that seem to have a tendency to evolve in Nyctaginaceae and that we can reasonably assume are contingent on other traits: the evolution of cleistogamy is improbable without prior self-compatibility, and lineages that specialize on gypsum are unlikely to have arisen from lineages with no latent or expressed gypsum tolerance.

Cleistogamous (closed, self-fertilizing) flowers are produced in addition to chasmogamous (open) flowers in four genera of Nyctaginaceae: *Acleisanthes*, *Cyphomeris*, *Nyctaginia*, and some *Mirabilis* (Cruden, 1973; Spellenberg and Delson, 1974; Fowler and Turner, 1977; Levin, 2002). Though species with cleistogamous flowers have evolved in a number of angiosperm families, only in much larger families, e.g., Poaceae, Fabaceae, and Malpighiaceae, is this trait found in as many genera (Lord, 1981). Despite a long awareness of this phenomenon generally (Darwin, 1884), the evolution of this character has only rarely been investigated with phylogenetic methods (Desfeux et al., 1996; Bell and Donoghue, 2003).

Second, as in many caryophyllid families, e.g., Amaranthaceae and Portulacaeae, there is a propensity in many Nyctaginaceae to be tolerant of, or specialists of, gypseous soils. Outcrops of gypsum (hydrous calcium sulfate) are quite common in arid North America, especially in the Chihuahuan Desert. These areas have a flora characterized by gypsophiles, which never occur on other substrates, and gypsum-tolerant species, which are found on both gypseous and nongypseous soils (Waterfall, 1946; Parsons, 1976; Meyer, 1986). In the United States and Mexico, Nyctaginaceae are well represented in gypsum communities (Parsons, 1976). At least 25 species in seven genera are known to occur on gypsum. Of these, roughly half are known gypsophiles, found only on gypsum soils (Johnston, 1941; Waterfall, 1946; Fowler and Turner, 1977; Turner, 1991, 1993; Spellenberg, 1993, 2003; Mahrt and Spellenberg, 1995; Harriman, 1999; Levin, 2002).

Although gypsum soils support a distinct flora, the evolution of gypsophily is not understood as well as other cases of edaphic endemism. Gypsum is not an inherently poor substrate for plants in the same way as soil with, for instance, toxic levels of heavy metals (Cockerell and Garcia, 1898; Johnston, 1941; Loomis, 1944; Parsons, 1976; Meyer, 1986; Oyonarte et al., 2002). Recent experimental work has pointed toward mechanical, rather than chemical, factors to explain the limited flora of gypsum soils: seedlings of nongypsophiles are unable to penetrate the hard crust typical of gypseous soils. This indicates that adaptations of gypsum-tolerant taxa primarily act to enhance survival in the establishment stage (Meyer, 1986; Meyer et al., 1992; Escudero et al., 1997, 1999, 2000; Romao and Escudero, 2005).

Edaphic-endemic species are sometimes found to be related to species that are merely tolerant: in the case of a serpentine endemic species of *Layia* (Asteraceae), certain populations of a non-endemic progenitor species were found to tolerate serpentine soils (Baldwin, 2005). Thus, even in the case of highly toxic soils, saltational speciation (Antonovics, 1971;

TABLE 2. Primer sequences used and original publication.

Region Primer name	Sequence	Reference
ITS		
ITS4	TCCTCCGCTTATTGATATGC	White et al., 1990
ITS5a	CCTTATCATTTAGAGGAAGGAG	Stanford et al., 2000
ndhF		
Nyct_ndhF1F	TGCCTGGATTATACCCTTCA	This study
NdhF972F	ATGTCTCAATTGGGTTATATGATG	Olmstead and Sweere, 1994
Nyct_ndhF13R	CAFCBGGATTACYGCATTT	This study
Nyct_ndhF22R	CTTGTAACGCCGAAACCATT	This study
Nyct_ndhF6F	AACGGGBAGTTTYGARTTTG	This study
Nyct_ndhF8R	AGTAGGCCCCTCCATAGCAT	This study
Nyct_ndhF14F	TCAATCGTTGCAATCCTTCT	This study
Nyct_ndhF16R	TTTCCGATTCATGAGGATATGA	This study
rps16		
rpsF	GTGGTAGAAAGCAACGTGCGA	Oxelman et al., 1996 (modified)
Rps2R	TCGGGATCGAACATCAATTGCAAC	Oxelman et al., 1996
rpl16		
F71	GCTATGCTTAGTGTGTGACTCGTTG	Jordan et al., 1996
R1661	CGTACCCATATTTTTCCACCACGAC	Jordan et al., 1996

Kruckeberg, 1986) is not required to explain edaphic endemism. These lines of evidence, and the fact that roughly half of the species of Nyctaginaceae found on gypsum are not restricted to it, make it reasonable to assume that an underlying ability to survive in gypsum soils is an early stage in the evolution of this type of edaphic endemism in Nyctaginaceae.

In principle, for both of these examples, the evolution of both the independent and contingent characters can be reconstructed on a phylogeny. With an understanding of the distribution of homoplasy in Nyctaginaceae, we will have a more robust framework for asking questions about character evolution and adaptation to xeric environments. In this phylogenetic study we comprehensively sample the genera of Nyctaginaceae, with the following goals: (1) to evaluate the existing classification of Bittrich and Kühn (1993), (2) to understand the biogeographic history of the family, and (3) to have a basis for understanding the evolutionary history of characters of historical taxonomic importance and the potential adaptive significance as manifested in their "tendency" to evolve repeatedly in lineages occurring in the deserts of North America.

MATERIALS AND METHODS

Sampling—Fifty-one species representing 25 genera of Nyctaginaceae were sampled. Taxa, voucher information, and GenBank numbers are given in Appendix 1. Our sampling is nearly comprehensive at the generic level, with representative species of every genus except *Neeopsis*, *Cephalotomandra*, *Grajalesia*, *Cuscatlania*, *Boldoa*, and *Cryptocarpus*. The genera omitted are monotypic, rarely collected, and/or of dubious distinction. For example, *Boldoa purpurascens* is often included in *Salpianthus* (Pool, 2001). All tribes and subtribes recognized by Bittrich and Kühn (1993) are included. Because different taxa have been found to be sister to Nyctaginaceae (Rettig et al., 1992; Behnke, 1997; Downie et al., 1997; Cuenoud et al., 2002), outgroups were selected from both Phytolaccaceae and Sarcobataceae. More distantly related taxa in the "core Caryophyllales," i.e., Aizoaceae, Molluginaceae, and Stegnospermataceae (Cuenoud et al., 2002), were also included to enable us to test the monophyly of Nyctaginaceae and to identify which taxa are sister to the family. For four species, data were obtained from two different accessions, and for two, GenBank sequences were used for some loci. "*Phytolacca*" is a composite of one GenBank sequence from *P. acinosa* and three new sequences from *P. americana*.

Molecular data—Genomic DNA was extracted from fresh, silica-dried, or air-dried (herbarium) leaf tissue using either Qiagen DNAeasy Plant Mini Kits or a modified CTAB method (Doyle and Doyle, 1987). Internal transcribed spacer (ITS) sequences were obtained using primers ITS4 and ITS5a (White et al., 1990; Stanford et al., 2000), which amplifies ITS1, 5.8S, and ITS2. Chloroplast *ndhF* sequences were obtained as two overlapping fragments using primers Nyct-ndhF1, ndhF972, Nyct-ndhF13R, and Nyct-ndhF22R. With the exception of ndhF972 (Olmstead and Sweere, 1994), these were designed based on GenBank *ndhF* sequences for Nyctaginaceae and Phytolaccaceae. Many samples, especially those from herbarium materials, were recalcitrant to PCR of long (>1 kb) fragments due to DNA degradation; for these, four additional primers (Nyct-ndhF6F, Nyct-ndhF8R, Nyct-ndhF13F, and Nyct-ndhF16R) were designed, based on sequences for Nyctaginaceae and Phytolaccaceae, and used in conjunction with the aforementioned primers, so that the gene was amplified in four overlapping fragments. The chloroplast intron *rps16* was amplified using primers rpsF and rps2R (Oxelman et al., 1997), and *rpl16* was obtained using primers F71 and R1661 (Jordan et al., 1996). Primer sequences and references are given in Table 2. PCR products were cleaned with Qiaquick columns (Qiagen, Valencia, California, USA). Cycle sequencing was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit, and sequences were determined with an ABI 3700 DNA Analyzer (Applied Biosystems, Foster City, California, USA) in the Genetic Analysis facility in the Department of Biology at Duke University. Raw chromatograms were edited and assembled in Sequencher 4.1 (Gene Codes Corp., Ann Arbor, Michigan, USA). Sequence alignment was performed either by eye (*ndhF*) or in ClustalX (Thompson et al., 1997) (other regions) followed by manual adjustment in Se-Al (Rambaut, 1996). Across the entire data set, ITS1 and ITS2 were too variable to be confidently aligned, although the 5.8S region was highly conserved. Ambiguously aligned regions were excluded from further analyses of the entire data set, though they were used in analyses of more restricted taxon sets (see *Restricted analyses*).

Caribea littoralis Alain, a Cuban endemic, has been collected only once. The collection locality is in southeastern Cuba in a dry coastal habitat. The morphology of the plant is difficult to interpret because it is highly distinct from any other member of the Nyctaginaceae, and the leaves and flowers are highly reduced. Few details are clearly visible on the specimen, though the description appears to have been based on fresh material (Alain, 1960). Due to the age of the collection, only about 25% of an *ndhF* sequence was obtainable. This sequence was unique in our data set, and a BLAST search found that this sequence fragment was most similar to an existing *Bougainvillea ndhF* sequence (GenBank no. AF194825). Preliminary phylogenetic analysis (see *Data analysis*) placed this taxon as sister to either *Pisoniella* or *Belemia*. These last two are not closely related to each other, resulting in substantial loss of resolution in the clade including these taxa. Therefore, *Caribea* was excluded from all further analysis, and while this result confirms that this enigmatic taxon belongs in Nyctaginaceae, further study must await rediscovery of this species. Unfortunately, repeated attempts to relocate the population at the type locality

TABLE 3. Summary of sequence statistics by partition for the molecular matrix.

	Full analysis								
Partition	*ndhF* (entire)	*ndhF* (1st pos.)	*ndhF* (2nd pos.)	*ndhF* (3rd pos.)	*rps16*	*rpl16*	5.8S	Entire	Chloroplast
No. taxa (full matrix = 58)	54	54	54	54	51	42	55	55	55
Aligned length	2193	731	731	731	1237	1367	157	5505	4797
Analyzed length	2205	669	668	668	780	792	157	3734	3577
Constant	1348	474	532	342	520	552	136	2556	2420
Uninformative	272	83	69	120	110	139	5	526	521
Parsimony-informative	385	112	67	206	150	101	16	652	636
ML model	GTR+I+Γ	GTR+I+Γ	GTR+I+Γ	GTR+Γ	GTR+Γ	GTR+Γ	SYM+I+Γ	GTR+I+Γ	GTR+I+Γ

Note: Maximum-likelihood (ML) model estimated by ModelTest (Posada and Crandall, 1998): Full, Entire; Sensitivity, Entire; Restricted I & II, Entire. ML model for remaining partitions (used in Bayesian analyses "B2," "B4," and "B6," see text) estimated with MrModelTest (Nylander, 2004). Numbers in parentheses are number of informative characters gained from the inclusion of ITS1 and ITS2 in restricted analyses.

in Cuba have proved unsuccessful (D. Stone, Duke University, personal communication).

Data analysis—Initial maximum parsimony (MP), maximum likelihood (ML), and Bayesian analyses were performed for each of the four loci. The 5.8S, not surprisingly, had low variation and produced poorly resolved trees; however, examination of the support values for the topology favored by each locus revealed no supported nodes in conflict. Therefore, the data sets were combined for further analyses.

MP analysis was performed using PAUP* version 4.0b10 (Swofford, 2002). A heuristic search was performed, with 1000 replicates of 10 random-addition sequences, tree-bisection-reconnection (TBR) branch swapping, MAXTREES set to autoincrease, MULTREES = yes. Support was evaluated using 1000 bootstrap replicates of 10 random addition sequences, TBR branch swapping, MULTREES = YES.

For the ML analysis, the data set was first examined using ModelTest 2.0 (Posada and Crandall, 1998), which selected a complex model of evolution (GTR + I + Γ). Ten random-addition replicates (TBR, MAXTREES set to autoincrease, MULTREES = yes) were run in PAUP*. Maximum-likelihood bootstrap support values were obtained by 100 replicates of single random-addition sequences, TBR branch swapping, MULTREES = yes.

Bayesian analysis was performed using MrBayes 3.1 (Ronquist and Huelsenbeck, 2003). For exploring the effect of different models for different partitions of the data, best-fit models for each partition were estimated in MrModelTest (Nylander, 2004), which selects the best-fit model from those available in MrBayes. The partitions were as follows: 1, all loci together; 2, nuclear 5.8S; 3, all chloroplast loci; 4, *rpl16*; 5, *rps16*; 6, *ndhF*; and 7, 8, 9, first, second, and third positions of *ndhF*, respectively. The models selected by MrModelTest for each partition are given in Table 3. Bayesian searches were then performed on the entire data set using four partition/model combinations: "B1," single model for all partitions, (1); "B2," nuclear and chloroplast, (2 and 3); "B4," all loci, (2, 4, 5, and 6); and "B6," all loci with separate models for each codon position of *ndhF* (2, 4, 5, 7, 8, and 9). For each combination, we executed four independent runs of 1×10^6 generations each, sampling every 100th tree. After discarding trees from the burn-in (determined by visualizing the plateau in –lnL scores, approximately after 50 000 generations), we compared the posterior tree sets from each run by computing a 50% majority rule tree in PAUP*. No strongly supported topological differences (at posterior probability ≥95%) were found between the four runs of each model set. Therefore, the four posterior tree files for each set of models were combined into a single posterior tree file for purposes of assessing support values yielded by each set of models. These preliminary analyses were conducted including the partial *ndhF* sequence for *Caribea*; however, the B6 analysis was repeated without this sequence.

Sensitivity analyses—Due primarily to the inclusion of GenBank sequences for outgroup taxa and the failure of certain loci to amplify (mostly from herbarium material), approximately 17.7% of the data matrix was coded as "missing." The potential impact of this was investigated by deleting from the analysis 18 taxa (Appendix 1) for which one or more sequences were entirely missing and by combining sequences from *Bougainvillea glabra* and *B. infesta* into a composite operational taxonomic unit "*Bougainvillea*." "*Phytolacca*" and *Rivina humilis* were the only remaining outgroups in this analysis, which allowed us to examine the effect of including distant outgroups. The MP, ML, and corresponding bootstrap searches were performed with the same settings as in the analysis of the full matrix. The resulting trees were compared to the topology from the full analysis to see whether the exclusion of missing data led to a preferred topology that differed substantively from the topology or levels of support in the analysis of the full matrix.

Restricted analyses—To gain resolution within and between closely related genera, our selection of loci encompassed a large range of sequence variation. Because both the ITS1 and ITS2 regions had to be excluded from the analysis of the complete data set due to questionable alignment (though the highly conserved 5.8S region was kept in the full matrix), following the analysis of the full data set, two restricted data sets were constructed to allow us to increase the number of included characters (Table 3) by reducing the taxon sampling to two distinct clades found in the full analysis. These restricted data sets comprised all included nucleotide positions in the full data set, plus sites that were unalignable across the breadth of taxa included in the full data set, but that were alignable within each of the restricted sets of taxa. The first restricted analysis group was comprised of North American herbs representing all taxa in the sister group to *Allionia*, whereas the second corresponded to the Pisonieae, *Bougainvillea*, *Belemia*, and *Phaeoptilum* (the "B&P" clade from the full analysis). The MP, ML, and corresponding bootstrap analyses were performed in the same fashion as in the full matrix and sensitivity analyses, with the exception that the ML models were reestimated in ModelTest.

Character data—The historical taxonomic significance given to pollen morphology and involucral bracts led us to examine these characters in a phylogenetic context. Pollen data follows the scheme of Nowicke, who identified four types in Nyctaginaceae (Nowicke, 1968, 1970, 1975; Nowicke and Luikart, 1971; Reyes-Salas and Martínez-Hernández, 1982; Chavez et al., 1998). Pollen type was coded as a multistate, unordered character. In many cases, the exact species included in our study were not examined in the published studies. If there was no indication of within-genus pollen polymorphism, that pollen type was assigned to all species in this analysis. However, multiple pollen types were recorded within *Neea* and *Pisonia*. Thus, only *N. psychotrioides*, which was examined by Nowicke, was coded unambiguously; other species of *Neea* and *Pisonia* were coded as polymorphic (states "1&3" and "1&4," respectively) to reflect this uncertainty in the assignment of ancestral states. The presence of involucral bracts was scored as present/absent. If only small subtending bracteoles occur (common in many taxa), this character was coded as "absent," mirroring the usage of this character in defining subtribe Nyctagininae. The occurrence of cleistogamous flowers was scored based primarily on literature sources (Spellenberg and Delson, 1974; Bittrich and Kühn, 1993; Levin, 2002; Spellenberg, 2003). Gypsophilic taxa were identified in literature sources (Waterfall, 1946; Parsons, 1976; Fowler and Turner, 1977; Turner, 1991; Harriman, 1999; Levin, 2002; Spellenberg, 2003; N. Douglas, personal observation). Taxa were identified as full gypsophiles (recorded only from gypseous soils), gypsum tolerant (recorded from both gypseous and nongypseous soils), or nongypsophilic. Taxa that do not occur in areas with gypsum outcrops were considered to be nongypsophilic. It is unlikely that transitions to or from full gypsophily could evolve with no intermediate gypsum-tolerant step; therefore, this character was analyzed as both unordered and ordered, with two steps required between

TABLE 3. Extended.

Sensitivity analysis	Restricted analysis I	Restricted analysis II
Entire	Entire	Entire
39	19	15
5359	4887	4914
3396	4278	4082
2597	3883	3526
412	160	382
387	235 (122)	174 (76)
GTR+I+Γ	GTR+I+Γ	GTR+I+Γ

nongypsophily and full gypsophily. Parsimony ancestral states of all characters were reconstructed with the program Mesquite 1.6 (Maddison and Maddison, 2006). Those terminals that were not assigned a single state, and branches that were not unambiguously resolved, are depicted as "equivocal."

RESULTS

Data matrix—The entire data matrix (Table 3) had a length of 5505 bp, of which 1771 were excluded due to ambiguous alignment, mainly due to the presence of length variation in ITS1 and ITS2 and in the two chloroplast introns, *rpl16* and *rps16*. Of the remaining 3734 characters, 652 were parsimony informative.

Phylogenetic analysis of the complete dataset—The MP analysis resulted in 36 shortest trees (length: 2287, consistency index [CI]: 0.657, retention index: 0.809, rescaled CI: 0.531); however, the strict consensus (tree not shown) resolved all but two ingroup nodes. Thirty-nine nodes were supported with parsimony bootstrap values (MPBS) $\geq$70.

The best-fit model as determined by ModelTest (Table 3) using both a hiearchical liklihood ratio test (HLRT) and the Akaike information criterion (AIC) was a general-time-reversible model with a proportion of invariant sites and a gamma shape parameter (GTR+I+Γ). The ML search returned a single ML tree, which was nearly identical to the MP topology, except in the placement of the genus *Colignonia*. This taxon is placed as sister to the large clade containing *Acleisanthes* and *Boerhavia* in the MP analysis (MPBS = 80) and is not resolved with strong support in any ML or Bayesian analysis. Overall, 38 nodes in the ML analysis were supported with likelihood bootstrap values (MLBS) $\geq$70.

Models determined by MrModelTest for each data partition in the Bayesian analyses are given in Table 3. On the basis of our preliminary examination of partitioned models, the signal in the data set apparently is strong, and the topology is not contingent on model selection: the tree topologies produced by the Bayesian B1, B2, B4, and B6 searches were consistent. The principal difference between them is in the level of support for the topology, with 37, 39, 40, and 40 nodes, respectively, supported by posterior probabilities (PP) $\geq$95%. Deletion of *Caribea* led to the resolution, with support of two additional nodes in the repeated B6 search, for a total of 42 nodes supported at greater than 95% PP. The topology of this Bayesian B6 consensus tree is identical to the ML tree. All further Bayesian support values refer to the B6 analysis.

The Nyctaginaceae are supported as monophyletic by ML (MLBS = 71) and Bayesian (PP = 100) analyses (Fig. 1). Interestingly, in the MP bootstrap analysis of this matrix, the monophyly of the Nyctaginaceae is not supported. Despite the inclusion of several outgroups, no single sister lineage emerges with strong support.

Leucastereae, a tribe of four South American genera (*Andradea*, *Leucaster*, *Ramisia*, and *Reichenbachia*), is supported as the earliest branching lineage in Nyctaginaceae (Fig. 1) followed by Boldoeae, represented by *Salpianthus*. A clade containing largely neotropical trees and shrubs, and the African genus *Phaeoptilum*, receives support from MP and B6 analyses, though not from ML. We will refer to this group as the Bougainvilleeae and Pisonieae ("B&P") clade, (Fig. 2), recognizing that it also includes *Phaeoptilum* and *Pisoniella*, which are currently classified in Nyctagineae. *Bougainvillea*, *Belemia*, and *Phaeoptilum* form a clade within this group, which is sister to a clade containing the Pisonieae and the genus *Pisoniella*. Within the Pisonieae, *Neea* and *Guapira* together form a clade but neither genus appears to be monophyletic.

Strong support is found in all analyses for a clade including mostly North American xerophytic genera. For the purposes of this paper, we refer to it as the North American Xerophytic ("NAX" clade) (Fig. 2). The NAX clade is well defined by geography, habit, and habitat, but it has never been recognized formally. The earliest branch within this clade leads to *Acleisanthes* sensu Levin (2000). It is followed by a clade representing *Abronia* and *Tripterocalyx* (tribe Abronieae). Phylogenetic relationships of the remaining genera in the NAX clade are mostly well resolved, with the exception of low support values for the placement of *Commicarpus* and *Allionia*. Two pairs of genera in this clade are not resolved as monophyletic. *Anulocaulis* includes *Nyctaginia*, and *Boerhavia* includes *Okenia*, though support in both of these cases is weak or lacking. Examination of the branch lengths (Fig. 2) makes it clear that *Anulocaulis* and *Nyctaginia* are at least very closely related.

The position of *Colignonia* is not resolved in the ML and Bayesian analyses. The ML analysis resolves *Colignonia* sister to the B&P and NAX clades but with weak support. A position sister to only the NAX clade is supported in the MP analysis.

Sensitivity analyses—The deletion of taxa with significant missing data resulted in a matrix of 39 taxa with only 3.1% missing data, as compared to 58 taxa with 17.7% missing data in the full analysis (See Appendix 1). The MP/ML analyses of this matrix yielded trees (not shown) that had no well-supported nodes conflicting with the topology of the tree from the full matrix. The support for the monophyly of the Nyctaginaceae increased to 94/95 MPBS/MLBS, from −/71 in the analysis of the full data set. The high level of support found in this analysis for the monophyly of Nyctaginaceae indicates that the inclusion of many outgroups in the full matrix, including the quite distant *Stegnosperma*, may have affected the level of support in the MP analysis. Alternatively, high levels of missing data in the full data set may be responsible for low support values at this key node. Support for the placement of *Cyphomeris* decreased to 70/66 relative to the full analysis. *Commicarpus* and *Allionia* increased to 73/67 and 87/77, respectively; these nodes had not received strong support in any analysis of the full data set. The remainder of the comparable nodes were similarly supported between the full and sensitivity analyses.

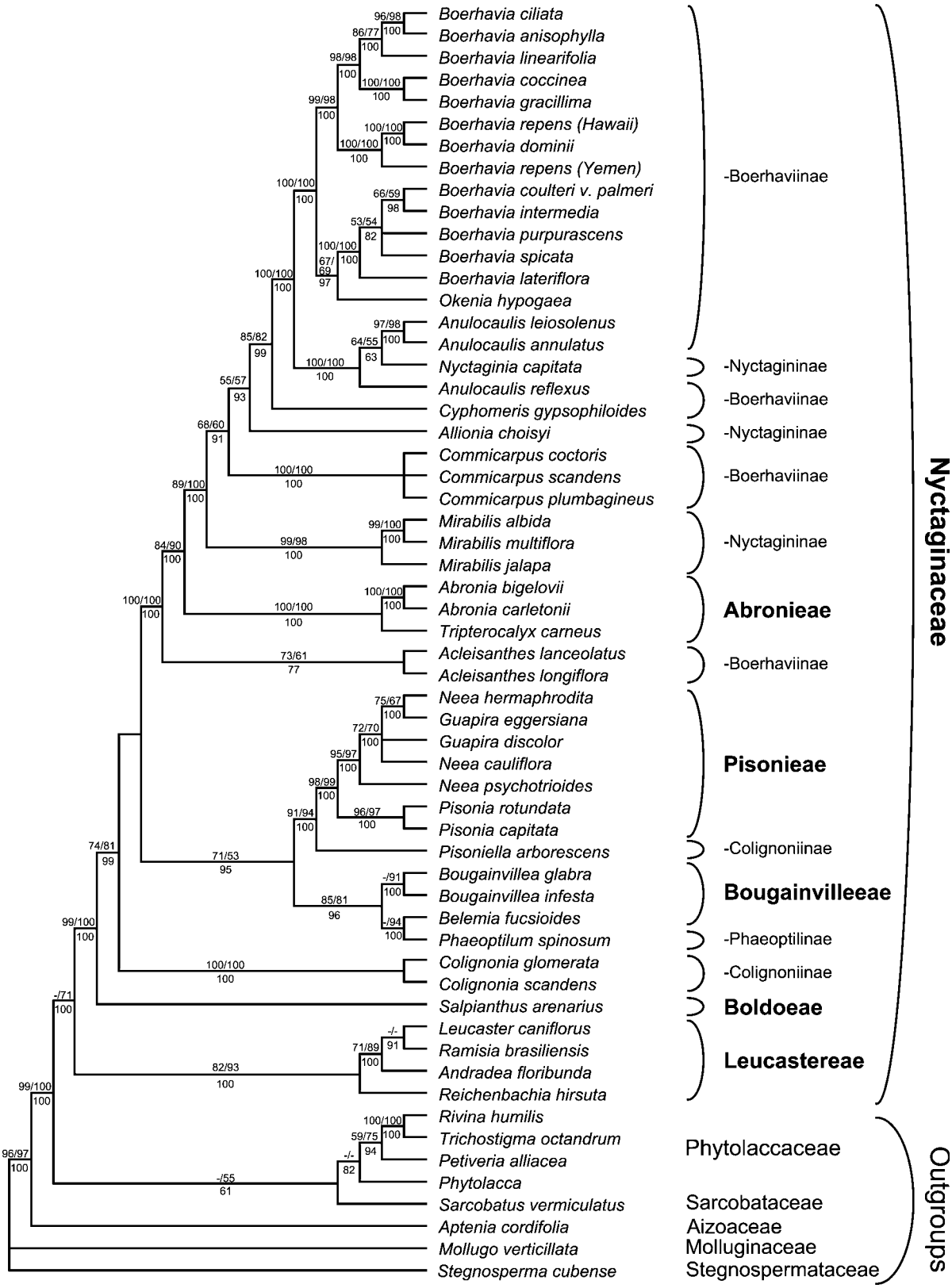

Fig. 1. Maximum-likelihood (ML) topology from the analysis of the entire data set. Parsimony bootstrap/ML bootstrap support values above branches, Bayesian posterior probability from the "B6" analysis below branches, "-" indicates bootstrap support value <50. tribes of Nyctaginaceae according to Bittrich and Kühn (1993) are in bold. "-" before unbold name signifies a subtribe of tribe Nyctagineae.

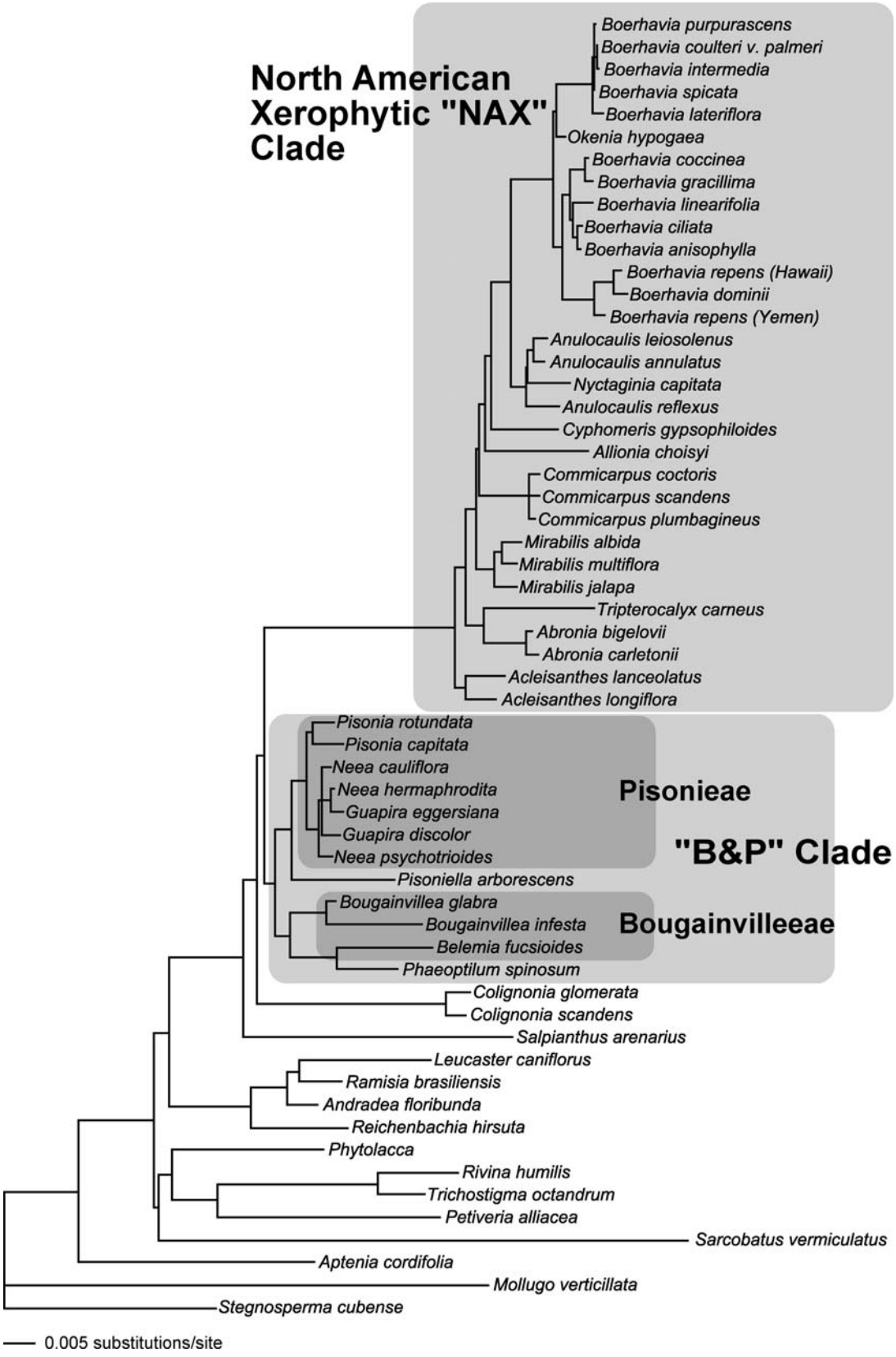

Fig. 2. Phylogram of the maximum-likelihood topology from Fig. 1. Major clades referred to in text are highlighted.

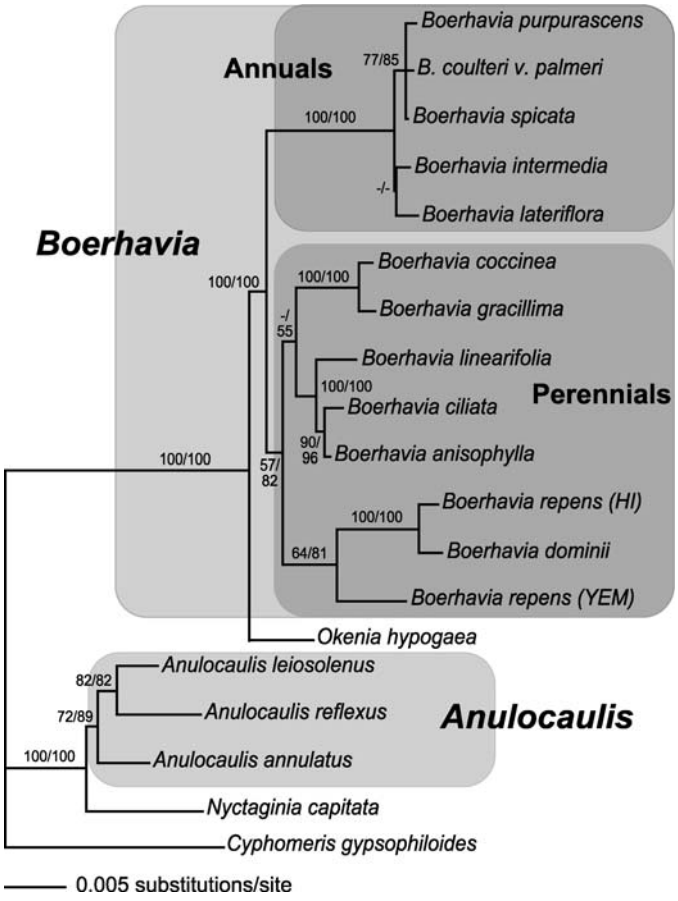

Fig. 3. Phylogram of the maximum-likelihood (ML) topology from the first restricted analysis. MP bootstrap/ML bootstrap support values are shown. *Anulocaulis* and *Boerhavia* are each supported as monophyletic.

Restricted analyses—For the two restricted analysis groups, 122 and 76 additional informative characters were gained with the inclusion of ITS1 and ITS2, respectively (Table 3). A small number of additional sites were gained from the chloroplast introns *rps16* and *rpl16* (<5 characters in either data set). ModelTest 3.7 selected a GTR+I+Γ model for each (Table 3) data set. For the first group (all taxa in the sister group to *Allionia*), MP and ML analyses produced a tree (Fig. 3) with improved resolution in the *Anulocaulis* + *Nyctaginia* clade. Though the placement of *A. annulatus* differs between the full matrix topology and the restricted analysis, the monophyly of *Anulocaulis* was well supported with bootstrap values of 72/89 MPBS/MLBS. Similarly, the full analysis resolves *Okenia* within a paraphyletic *Boerhavia* with low MP and ML bootstrap support, but 97% Bayesian posterior probability, yet the restricted analysis found *Boerhavia* strongly supported as monophyletic and sister to *Okenia* with high support (100/100). *Boerhavia* consists of two clades, corresponding to annual and perennial species, that were also found in the full analysis.

The restricted analysis of the B&P clade produced a tree (not shown) that did not conflict with the topology of this clade in the full matrix analysis. Support values were generally slightly lower, probably due to the concentration of missing data in this group and the lower number of additional characters from the ITS region. Support values remained high for the nodes uniting *Guapira eggersiana* and *Neea hermaphrodita*, for the placement of *G. discolor* in the clade sister to *N. psychotrioides*, and for the monophyly of *Neea* + *Guapira* (MPBS/MLBS bootstrap support of 64/73, 82/94, and 94/92, respectively).

Character reconstructions—For each character reconstructed (Fig. 4), multiple state transitions are inferred. Tricolpate-spinulose pollen (Fig. 4a) appears to be the ancestral condition in the group, transitioning to a pantoporate-spinulose condition subsequent to the divergence of *Salpianthus* from the main lineage. The latter condition is found in nearly all members of the NAX clade, yet appears to predate that group. At least eight transitions among the four pollen types have occurred in the Nyctaginaceae. Considering the small number of *Neea* and *Pisonia* examined and the polymorphism exhibited by these genera, the number of transitions could be higher. Reconstruction of involucral bracts shows five gain/loss steps. This character is fixed within genera, thus this interpretation is likely to be affected only by the future inclusion of the remaining genera in the family. Only the inclusion of *Cuscatlania*, which has an involucre, could conceivably change the number of steps required. Cleistogamous flowers are uniquely derived in four genera. Gypsophily requires nine or 13 steps to explain, depending on whether it is considered to be an unordered or an ordered character. Reconstructions were performed only on the ML topology from the full analysis. Adjusting the positions of *Okenia* and *Nyctaginia* to reflect the topology from the restricted analysis (Fig. 3) results in the branches leading to *Nyctaginia* + *Anulocaulis* and *Nyctaginia* + *Anulocaulis* + *Okenia* + *Boerhavia* being resolved as nongypsophilic. Treating gypsophily as an unordered character has the same result. Otherwise, the alternative topology has no substantive effect on the conclusions we make regarding the degree of homoplasy shown by the remaining three characters shown in Fig. 4.

DISCUSSION

Phylogeny of Nyctaginaceae—The earliest branching lineage in Nyctaginaceae, the Leucastereae (Fig. 1), had been previously recognized as a natural group on the basis of arborescence, a stellate indumentum, and tricolpate pollen (Heimerl, 1934; Bittrich and Kühn, 1993). The Boldoeae, an herbaceous group native from the Galapagos to northwestern Mexico and the Caribbean, are represented in this study by *Salpianthus*. These two lineages had been predicted to be basal or outside of Nyctaginaceae on the basis of apparent pleisomorphies such as alternate leaves and bisexual flowers (Bittrich and Kühn, 1993). The anthocarp structure is absent in Leucastereae and Boldoeae, although the unexpanded perianth does persist around the fruit. Persistent tepals are also found in many Phytolaccaceae. However, the perianth consists of free tepals in most Phytolaccaceae and all of subfamily Rivinoideae (except *Hilleria*, in which three of four tepals are partially fused, (Rohwer, 1993)). In Nyctaginaceae, including Leucastereae and Boldoeae, tepals are fully connate.

Within the B&P clade (Fig. 2), *Phaeoptilum* is found to be sister to *Belemia*, rendering the Bougainvilleeae paraphyletic. The Pisonieae are found to be sister to *Pisoniella*, which had been included in that tribe by Heimerl (1934) but was removed

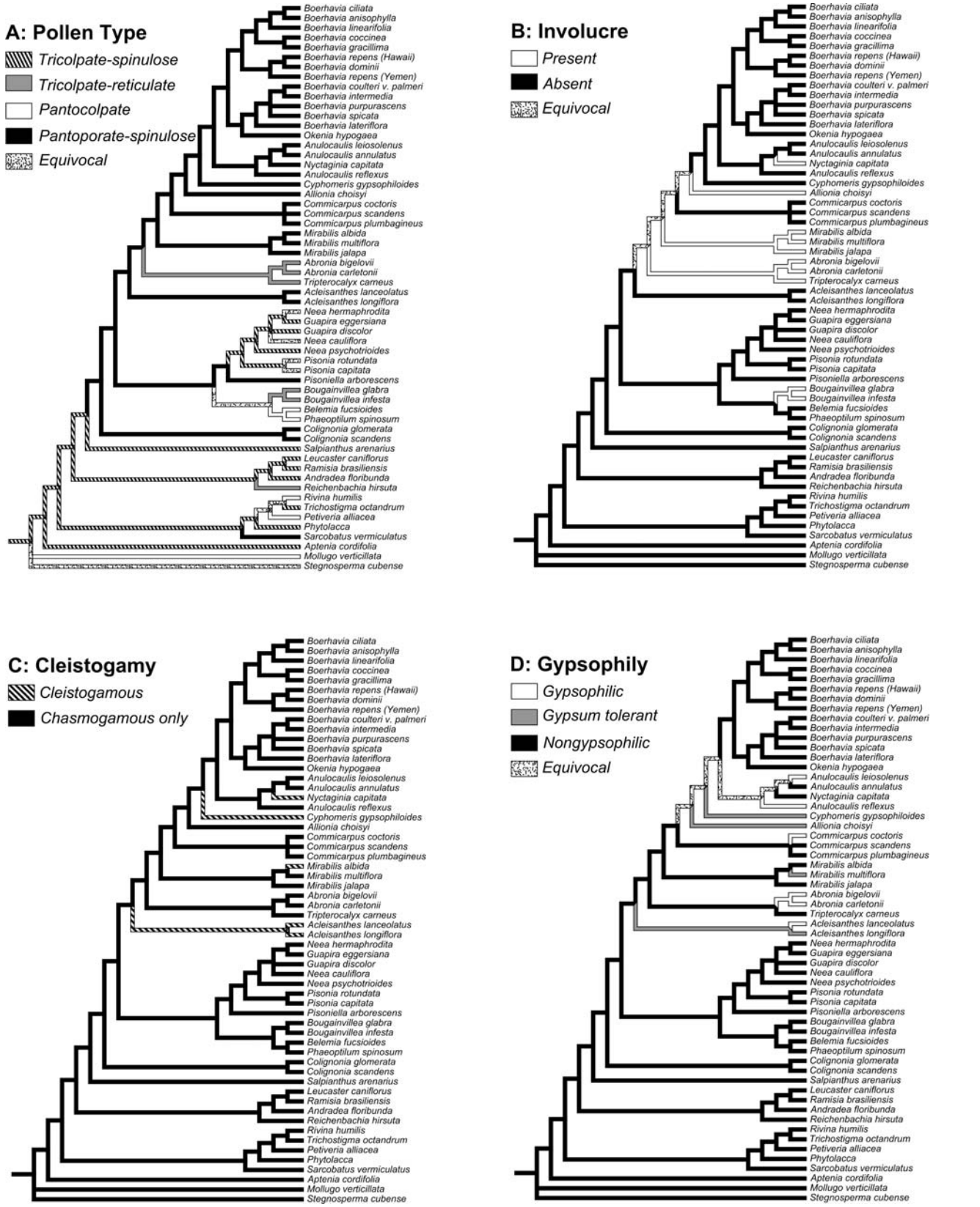

Fig. 4. Parsimony reconstruction of (A) pollen morphology, (B) involucre presence, (C) cleistogamous flowers, and (D) gypsophilic habit (based on ordered characters). See text for sources of characters used in reconstructions.

to subtribe Colignoniinae by Bittrich and Kühn (1993) following the suggestion of Bohlin (1988). The reasoning behind this move is mysterious, and in light of our results, it appears to have been unwarranted. *Pisoniella* possesses a straight embryo like other Pisonieae, and the large coriaceous anthocarps are provided with viscid glands along the ribs, much like those in *Pisonia* (Heimerl, 1934).

Within Pisonieae, *Neea* and *Guapira* form a clade (Fig. 2). These genera are distinguished primarily by whether the stamens are included (*Neea*) or exserted (*Guapira*). Our sampling is extremely limited in these two large genera, with only five accessions to represent ca. 150 species, though we were able to include accessions from geographically disparate locales. Neither genus forms a monophyletic group. This conclusion has been occasionally anticipated (e.g., Pool, 2001). It is unclear whether our sampling simply happened to include misclassified species in otherwise good genera, or whether this paraphyly is representative of *Neea* and *Guapira* generally. Much more intensive sampling is clearly needed to understand the relationships of the species in these genera, and it would be imprudent to attempt to reclassify them until a more detailed study is made including phylogenetic, morphological, and distributional data. Unfortunately, collections of these dioecious trees often do not include individuals of both sexes. Also, the tendency of many Pisonieae to oxidize when dried has left many descriptions lacking crucial information concerning the color of fruits. Therefore, the taxonomic literature is quite confused and species limits are known not much better than when Standley (1931a, p. 73) wrote that, "I know of few groups of plants in which specific differences are so unstable and so baffling ... particularly in *Neea*, *Torrubia* [= *Guapira*] and *Mirablis*, no single character seems to be constant." Finally, in this study we did not attempt to infer the ages of lineages, yet it appears that the branch lengths in the *Neea* + *Guapira* clade are comparatively short, especially considering that this clade can be expected to accommodate as many as 150 species (Fig. 2). A similar pattern has been noted in other radiations of neotropical trees, e.g., *Inga* (Fabaceae) (Richardson et al., 2001). If the pattern of relatively short branches inferred between species was upheld with the inclusion of a larger sample of taxa and more rapidly evolving markers, it would point to this clade as another example of rapid diversification in the neotropics.

Tribe Nyctagineae is broadly paraphyletic. As mentioned, *Pisoniella* and *Phaeoptilum* are not found in this study to be closest relatives of any other Nyctagineae. Based on pollen morphology, Bohlin (1988) has suggested that *Colignonia* (subtribe Colignoniinae) has affinities to the tribe Mirabileae of Heimerl (1934), which roughly corresponds to the tribe Nyctagineae and the NAX clade. *Colignonia* may in fact be sister to the NAX clade as suggested by the MP analysis or to the NAX + B&P clade as suggested by the ML analysis (Fig. 1). Tribe Nyctagineae also does not include *Abronia* or *Tripterocalyx* (tribe Abronieae, Fig. 1). There are certain characters of the Abronieae that are anomalous within the Nyctagineae (and the NAX clade) and that justified recognition at a higher taxonomic level, namely, tricolpate pollen and linear stigmas. The two genera in the tribe have long been thought to be a natural group and are often synonymized (Heimerl, 1934; Bittrich and Kühn, 1993), though most authors have maintained the two genera (Galloway, 1975; Spellenberg, 2003). The *Abronia* + *Tripterocalyx* clade is characterized by the combination of an umbellate inflorescence of salverform flowers with included stamens and style, an involucre, anthocarps with typically well-developed wings or lobes, and a mature embryo with a single cotyledon.

Anulocaulis and *Nyctaginia* are classified in different subtribes in the classification of Bittrich and Kühn (1993), presumably based on the presence of an involucre in *Nyctaginia*. Both genera are succulent perennial herbs, and the turbinate fruits with umbonate apices of *Nyctaginia capitata* strongly resemble those of *Anulocaulis eriosolenus*. They differ in many characters, including flower color (red-orange in *Nyctaginia* vs. white to pink in *Anulocaulis*) and flowering time (flowers of *Nyctaginia* are open during the day, while in *Anulocaulis* anthesis is at sunset or later and flowers wilt in the morning). While the full matrix ML tree (Fig. 1) indicates that *Anulocaulis* may not be monophyletic, this relationship is poorly supported (MPBS/MLBS/PP = 64/55/63). In the restricted MP and ML analysis (Fig. 3), however, a monophyletic *Anulocaulis* is more strongly supported (MPBS/MLBS = 72/89). Therefore, we see no compelling reason to question the taxonomic status of *Anulocaulis*.

Anulocaulis + *Nyctaginia* are sister to a strongly supported clade containing *Boerhavia* and *Okenia*. Like the previous instance, *Okenia* resolves within *Boerhavia* in the full matrix ML topology (Fig. 1), but support for this relationship is only moderately significant in the Bayesian analysis of the full data set (PP = 97) and weakly supported by MPBS and MLBS (67/69). Conversely, *Boerhavia* is strongly supported as a monophyletic group in the MP and ML analyses of the restricted data set (MPBS/MLBS = 100/100, Fig. 3). Vegetatively, *Okenia* strongly resembles most *Boerhavia* in its decumbent habit, and subequal opposite leaves with sinuate or undulate margins. The flowers of *Okenia*, though larger, are similar in color to some perennial *Boerhavia* from the Chihuahuan Desert. Finally, *Okenia* is annual, a condition found in one clade of *Boerhavia*. However, *Okenia* is strikingly different than *Boerhavia* in its unique reproductive biology: it produces aerial flowers, but the large, spongy fruits are geocarpic, with peduncles elongating greatly after fertilization and the fruits maturing several centimeters belowground. The relationship between these two genera is deserving of more study.

Biogeographical patterns—The basal lineages of Nyctaginaceae (Boldoeae, Leucastereae, *Colignonia*, Bougainvilleeae, and Pisonieae [including *Pisoniella*]) are fundamentally South American. Though some taxa have representatives or populations in (sub)tropical North America, (*Salpianthus*, *Neea*, *Guapira*, *Pisonia*, *Pisoniella*), their distributions all include the neotropics, and phylogenetically they are interspersed with neotropical endemics. The widespread tropical genus *Pisonia* possesses extremely viscid anthocarps, which aid dispersal, frequently by seabirds (Burger, 2005). The sole genus not native to the Americas is *Phaeoptilum*, endemic to arid southwestern Africa. This monospecific genus is closely related to *Belemia* and *Bougainvillea*, both from eastern and southern South America. *Phaeoptilum* is morphologically quite distinct from its sister taxon *Belemia*, though vegetatively it resembles the xeric-adapted *Bougainvillea spinosa*. The early Cretaceous date (130–90 Ma) for the opening of the south Atlantic (Smith et al., 1994) makes vicariance an unlikely explanation for this disjunction. Dispersal seems more likely, and while there is no specialized dispersal structure on the anthocarp of *Belemia*, both *Bougainvillea* and *Phaeoptilum*

have compelling (albeit different) adaptations for wind dispersal. *Phaeoptilum* produces winged anthocarps highly similar to those found in *Tripterocalyx* and some species of *Acleisanthes*. In *Bougainvillea*, most species display three showy bracts, each fused to a solitary flower. In fruit each involucral bract remains fused to a fruit and acts as a wing, the structure functioning as a unit of dispersal (Ridley, 1930).

The North American Xerophytic Clade has diversified in the deserts of the southwestern United States and northwestern Mexico. Every genus is confined to or has representatives in this region. Widespread taxa in this clade, namely *Commicarpus* and *Boerhavia*, possess glandular fruits, which have most likely aided bird-dispersal in a manner similar to that of *Pisonia*. Two red-flowered *Boerhavia*, *B. coccinea* and the similar *B. diffusa* are widespread in most tropical and subtropical areas. *Boerhavia diffusa* appears to have naturally dispersed from the Americas, though the confused taxonomy of this species and *B. coccinea* in regional floras makes this difficult to evaluate, and both of these species are frequently transported by human activity. The "repens" complex in *Boerhavia* (*B. repens* and related species) is widespread in coastal habitats throughout the tropical Pacific and Indian oceans to the Arabian Peninsula, along with *B. dominii* from Australia. Like the red-flowered perennial *Boerhavia* mentioned, these species also have viscid glandular anthocarps. *Okenia* is found in deep sand dune habitat along the Pacific and Caribbean coasts of Mexico and Central America, with a disjunct population in southern Florida. Other authors (Heimerl, 1934; Fowler and Turner, 1977; Thulin, 1994; Levin, 2002; Spellenberg and Poole, 2003) have discussed the remarkable disjunctions of *Acleisanthes somaliensis* and *Mirabilis himalaicus* from east Africa and southern Asia, respectively. These appear to be attributable to long-distance dispersal events, due to their derived position within otherwise exclusively American clades (Levin, 2000; N. Douglas, unpublished data).

Pollen and involucre evolution—Tribal and subtribal classifications (Table 1) of the Nyctaginaceae have relied heavily on a few characters, such as pollen morphology and the development of an involucre. However, divisions based on these characters are not supported by our results because these characters have a high degree of homoplasy among genera.

Parsimony reconstruction of pollen type across Nyctaginaceae (Fig. 4a) shows that substantial homoplasy exists (11 changes), involving three of the four types diagnosed by Nowicke (Nowicke, 1970, 1975; Nowicke and Luikart, 1971). Pantocolpate grains may constitute a synapomorphy for *Belemia* + *Phaeoptilum*. It has been noted that large, desiccation-resistant, pantoporate pollen grains, equipped with pore plates, were found primarily in the herbaceous desert taxa (Nowicke and Luikart, 1971). Specific correlations between large and/or polyaperturate grains and habitat in angiosperms have not been adequately investigated. In a study of ecological correlates of pollen morphology in a wide selection of angiosperms (Lee, 1978), there was an extremely weak correlation of pore number with width and with "temperature." According to our reconstructions, the origin of pantoporate-spinulose pollen predates the major radiation of desert taxa in the NAX clade. However, *Colignonia* and *Pisoniella* have much smaller grains than do the remaining taxa with pantoporate-spinulose pollen (*Colignonia* = 25–35 μM, *Pisoniella* = 30–37 μM, Nowicke and Luikart, 1971; N. Douglas, unpublished data). Therefore, it would seem best to consider grain size as a variable separate from grain shape and exine structure.

Within Nyctagineae, the subtribes Nyctagininae and Boerhaviinae were separated by the presence or absence of an involucre subtending the inflorescence. In subtribe Nyctagininae, the involucre of *Mirabilis* is comprised of fused bracts; the remaining genera possess involucres of distinct bracts. The involucre in *Bougainvillea* is distinctive; fruits of *Bougainvillea* retain a large involucral bract as discussed. Involucres have no known dispersal function in any of the other taxa; they likely serve merely to protect the flower buds and developing fruits or discourage nectar-robbing insects (Cruden, 1970). Parsimony reconstruction of this character on the molecular topology (Fig. 4b) indicates that, for involucres, there are at least five gain/loss steps in the family, four in the NAX clade, which contains the members of the Nyctagineae-Nyctagininae, Nyctagineae-Boerhaviinae, and Abronieae, reflecting the artificial nature of this classification. In this analysis, the character was treated in a very simplistic fashion, reflecting nothing more than taxonomic convention. Comparative developmental studies may shed light on deeper homologies or convergences, especially as they relate to the subtending bracts found in many genera. The selective benefits involved in the expression of this structure could be revealed by appropriate ecological investigations.

Self-compatibility and cleistogamy—The production of obligately selfing flowers is obviously contingent on the ability of plants to self-pollinate and produce fertile progeny. Our incomplete knowledge of reproductive systems in Nyctaginaceae means that an unambiguous reconstruction of self-compatibility is not currently possible. However, several studies have addressed mating systems in select Nyctaginaceae: sporophytic self-incompatibility (SI) is known in *Bougainvillea* (Zadoo et al., 1975; López and Galetto, 2002). Some *Mirabilis* (sect. *Quamoclidion*) and *Abronia macrocarpa* fail to set seed when self-pollinated (Cruden, 1973; Williamson et al., 1994), but the basis for incompatibility is not known in these genera. The Pisonieae are usually dioecious and are thus self-incompatible, although in these genera there are occasional monoecious or hermaphroditic species (e.g., *Pisonia brunoniana*) for which the mating system has not been studied (Sykes, 1987). Evidence suggests that many genera in the NAX clade are self-compatible: in addition to the production of cleistogamous flowers in four genera, *Boerhavia* and some *Mirabilis* are known to have a delayed self-pollination mechanism whereby the style curls and encounters the anthers as the flower wilts (Chaturvedi, 1989; Hernández, 1990; Spellenberg, 2000). Finally, flowers protected from pollinators have set viable seed in *Abronia umbellata* Lam. (McGlaughlin et al., 2002) and *Colignonia* (Bohlin, 1988).

Reasoning from these data, we can make certain inferences regarding the evolution of mating systems in Nyctaginaceae. Explanations for current distribution of mating systems family must incorporate one, or some combination of both, of the following scenarios. Which one is preferred depends on the likelihood of self-compatible lineages giving rise to lineages with an inability to self-fertilize, and the implications of either scenario are interesting.

One scenario, and the most parsimonious given our current knowledge, is that there have been at least three independent derivations of SI from a self-compatible ancestor. A single

change can account for the Pisonieae and *Bougainvillea*, one for the derived *Mirabilis* sect. *Quamoclidion*, and one for *Abronia macrocarpa*. It is often assumed that outcrossing species are not derived from selfing ancestors and that selfing lineages are an evolutionary "dead end" (Fisher, 1941; Stebbins, 1974; Lande and Schemske, 1985). In the case of Nyctaginaceae, however, the question is whether it is possible that *self-incompatible* species have arisen from *self-compatible* ancestors. It would seem that populations making this transition would be subject to most of the forces that affect the balance of selfing and outcrossing in self-compatible populations. A recent study of *s*-locus polymorphism in Solanaceae (Igic et al., 2006) has shown that losses of SI are irreversible in that family. The "cost" of developing the complex genetic systems necessary for SI would be added to the transmission advantage of alleles promoting self-fertilization (Uyenoyama et al., 1993); these factors must count against a hypothesis of multiple transitions to SI in one family.

Conversely, if we assume that SI is ancestral and has been lost repeatedly, transitions from SI to self-compatibility have occurred a minimum of six times (in *Colignonia*, *Acleisanthes*, some *Abronia*, two or more times in *Mirabilis*, and finally in the clade sister to *Mirabilis*). This represents a doubling of the number of evolutionary steps required to explain the distribution of known Nyctaginaceae mating systems. Other authors have discussed the merits of parsimony weighting schemes or maximum-likelihood approaches to testing the irreversibility of selfing (Barrett et al., 1996; Bena et al., 1998; Takebayashi and Morrell, 2001). In these cases, it may not be possible to escape a circular argument employing *only* phylogenetic evidence, because a weighting scheme favoring losses of SI assumes the conclusion. In Solanaceae (Igic et al., 2006), evidence of ancient polymorphism at the incompatibility locus itself was required to demonstrate the irreversibility of the loss of SI. In our case, the most convincing resolution will come when SI is characterized in *Mirabilis* sect. *Quamoclidion* and SI *Abronia*. If in these taxa and any others that may yet be discovered to be self-incompatible the genetic basis for SI can be identified, homology could be assessed and the ancestral functionality of the underlying mechanism could be tested.

Assuming the derived state is self-compatibility, of these six lineages, three have given rise to cleistogamous/chasmogamous lineages, and four gains of cleistogamy are required to explain the distribution of the character in Nyctaginaceae (Fig. 4c). Interestingly, the cleistogamous genera are all perennial, which should be less susceptible to selection pressure for reproductive assurance than annuals (Barrett et al., 1996). Alternatively, cleistogamous flowers can function to maximize seed set when resources, rather than pollinators, are limiting (Schemske, 1978). These hypotheses are both applicable to the cleistogamous Nyctaginaceae, though distinguishing between them may be difficult, because pollinators in desert environments tend to be scarce when water is scarce. Spellenberg and Delson (1974) found that *Acleisanthes* (*Ammocodon*) *chenopodioides*, with a generalized flower morphology and a diurnal pollinator fauna, produced roughly equal numbers of seeds from cleistogamous and chasmogamous flowers, and did not have a strong seasonal pattern in the production of cleistogamous flowers. In contrast, *Acleisanthes longiflora*, a species with large, specialized hawkmoth-pollinated flowers, produced the majority of a season's seeds from cleistogamous flowers produced preferentially in the dry early summer when sphingid moths are less active. This may suggest that cleistogamy in this genus is insurance against reproductive failure due to the absence of pollinators in some years.

Gypsophily—Parsimony reconstruction of gypsophily in Nyctaginaceae (Fig. 4d) indicates that gypsophiles and gypsum-tolerant species are widely dispersed in the NAX clade. With the current sampling, the ancestor of this clade is inferred to be nongypsophilic (whether or not the character is considered "ordered"), indicating that gypsum tolerance is derived multiple times. This conclusion is tenuous for two reasons. First, gypsum outcrops are common in the Chihuahuan Desert but less so in other parts of the ranges of the NAX genera. We are unable to rule out the possibility that taxa coded in this analysis as "nongypsophilic" are actually gypsum-tolerant, but simply do not occur in areas with gypsum soils.

Second, there are two *Mirabilis* [*M. nesomii* Turner and *M. linearis* (Pursh) Heimerl] which are gypsophilic (Turner, 1991) and gypsum-tolerant (R. Spellenberg, New Mexico State University, personal communication), respectively. These species, both in section *Oxybaphus*, are close relatives of the oxybaphoid *M. albida*, a nongypsophile included in this study. It is possible to add gypsophilic taxa as sisters to *M. albida* on our topology, so that the resolution of the ancestor of the NAX clade becomes equivocal, with ACCTRAN reconstruction as gypsum-tolerant, and DELTRAN as nongypsophilic. The same reconstruction would be made for the ancestors of *Commicarpus* and *Abronia* + *Tripterocalyx*. The sensitivity of the reconstruction at these key nodes to sampling artifacts indicates that in order to reconstruct the history of gypsophily in this clade, it will be necessary to undertake more intensive phylogenetic sampling at the species level, investigating an appropriate sample of nongypsophilic taxa closely related to known gypsophiles.

Even if we cannot know the gypsum tolerance of the ancestor of the NAX clade based on existing data, it is evident that there are at least four instances of strong gypsophily evolving in the family. It would be profitable to investigate the ecology of these gypsophytes and their relatives in the NAX clade. An experimental approach investigating whether or not seedlings of nongypsphiles have the latent ability to establish on gypseous crusts would disentangle the expression of gypsum tolerance from biogeographic complications, clarify the phylogenetic distribution of gypsum tolerance and perhaps reveal the nature of the adaptation(s) involved.

It is possible that establishment on gypsum is facilitated by some sort of modification to the radicle. Alternatively, because germination in a desert environment is always risky, adaptations to gypsum soils may differ little from germination strategies of desert taxa generally. Possible strategies could serve to optimize the timing of germination, minimize the risk of all seedlings perishing or increase the length of time a seedling has to establish itself. These could include high germination rate at low temperatures and various forms of bet-hedging, such as seed heteromorphism and variable seed dormancy (Escudero et al., 1997). The production of mucilage upon wetting by the seed coat presumably increases the local availability of water and upon drying, anchors the seed (Romao and Escudero, 2005). Some of these traits are known in Nyctaginaceae. For instance, production of mucilage by the anthocarp is common in both gypsophilic and nongypsophilic taxa in the NAX clade (Spellenberg, 2003), and fruit/seed heteromorphism is known in *Abronia* and *Tripterocalyx* (Wilson, 1974).

Understanding when in their history Nyctaginaceae became gypsum-tolerant will clarify whether homoplasy is best explained by answering the question "how do species become gypsum-tolerant?" or "why are certain species found only on gypsum?" If it turned out that gypsum tolerance was ancestral in the NAX clade, then experiments may reveal the reasons full gypsophiles do not occur on more typical soils.

The tendency of Nyctaginaceae to evolve cleistogamy and gypsophily has been shown to the extent that we have demonstrated that the high level of homoplasy for these traits is restricted to the NAX clade. In neither case are we able to conclusively identify the largest group capable of evolving the trait. Largely due to the phylogenetic position of *Acleisanthes* (with gypsophilic, cleistogamous species), we infer that it is possible that the ancestor of the entire NAX clade was predisposed to evolve these traits. In the case of cleistogamy, the topology indicates either that SI mechanisms develop easily in Nyctaginaceae, or that once self-compatibility emerges, there is a high chance of cleistogamy following. If the latter situation is correct, the explanation for the large number of cleistogamous species in the NAX clade must ultimately rely on explaining the frequent loss of SI, though the proximate cause is more likely related to resource or pollinator limitation in xeric environments. With gypsophily, it remains to be seen what trait(s) allow for tolerance of gypsum soils and when they evolved and what factors act exclude to gypsophiles from nongypsum soils.

The present study is the first to provide a comprehensive genus-level examination of the phylogeny of Nyctaginaceae. Though sampling of *Caribea*, *Cuscatlania*, *Cephalotomandra*, *Grajalesia*, and *Neeopsis* would be desirable, the current level of sampling is sufficient to draw several useful conclusions with bearing on future studies of the family. Aside from providing a framework for future taxonomic revisions, it raises interesting evolutionary questions regarding biogeography, reproductive biology, and edaphic endemism. To a degree, this work may be considered a case study into the practical issues that may arise in an investigation of tendencies in character evolution. New insights will be gained with a combination of phylogenetic work at finer taxonomic scales and experimental data to better understand the natural history of individual species, especially those in the xerophytic clade.

LITERATURE CITED

ALAIN, H. 1960. Novedades en la flora cubana, XIII. *Candollea* 17: 113.

ANTONOVICS, J. 1971. Heavy metal tolerance in plants. *Advances in Ecological Research* 7: 1–85.

BAKER, H. G. 1961. The adaptation of flowering plants to nocturnal and crepuscular pollinators. *Quarterly Review of Biology* 36: 64–73.

BALDWIN, B. G. 2005. Origin of the serpentine-endemic herb *Layia discoidea* from the widespread *Layia glandulosa* (Compositae). *Evolution* 59: 2473–2479.

BARRETT, S. C. H., L. D. HARDER, AND A. C. WORLEY. 1996. The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 351: 1271–1280.

BEHNKE, H. D. 1997. Sarcobataceae—a new family of Caryophyllales. *Taxon* 46: 495–507.

BELL, C. D., AND M. J. DONOGHUE. 2003. Phylogeny and biogeography of Morinaceae (Dipsacales) based on nuclear and chloroplast DNA sequences. *Organisms Diversity & Evolution* 3: 227–237.

BENA, G., B. LEJEUNE, J.-M. PROSPERI, AND I. OLIVIERI. 1998. Molecular phylogenetic approach for studying life-history evolution: the ambiguous example of the genus *Medicago* L. *Proceedings of the Royal Society of London, B, Biological Sciences* 265: 1141–1151.

BITTRICH, V., AND U. KÜHN. 1993. Nyctaginaceae. *In* K. Kubitzki, J. G. Rohwer, and V. Bittrich [eds.], The families and genera of flowering plants, 473–486. Springer-Verlag, Berlin, Germany.

BOGLE, A. L. 1974. The genera of Nyctaginaceae in the southeastern United States. *Journal of the Arnold Arboretum* 55: 1–37.

BOHLIN, J. E. 1988. A monograph of the genus *Colignonia* (Nyctaginaceae). *Nordic Journal of Botany* 8: 231–252.

BREMER, B., K. BREMER, M. W. CHASE, J. L. REVEAL, D. E. SOLTIS, P. S. SOLTIS, P. F. STEVENS, A. A. ANDERBERG, M. F. FAY, P. GOLDBLATT, W. S. JUDD, M. KÄLLERSJÖ, J. KAREHED, K. A. KRON, J. LUNDBERG, D. L. NICKRENT, R. G. OLMSTEAD, B. OXELMAN, J. C. PIRES, J. E. RODMAN, P. J. RUDALL, V. SAVOLAINEN, K. J. SYTSMA, M. VAN DER BANK, K. WURDACK, J. Q. Y. XIANG, AND S. ZMARZTY. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141: 399–436.

BURGER, A. E. 2005. Dispersal and germination of seeds of *Pisonia grandis*, an Indo-Pacific tropical tree associated with insular seabird colonies. *Journal of Tropical Ecology* 21: 263–271.

CHATURVEDI, S. K. 1989. A new device of self pollination in *Boerhaavia diffusa* L. Nyctaginaceae. *Beitraege zur Biologie der Pflanzen* 64: 55–58.

CHAVEZ, R. P., R. F. NAVA, E. GRAFSTROM, M. VON PFALER, AND S. NILSSON. 1998. On the pollen of *Salpianthus* (Nyctaginaceae)—a morphological and image analysis approach. *Grana* 37: 352–357.

COCKERELL, T. D., AND F. GARCIA. 1898. Preliminary note on the growth of plants in gypsum. *Science* 8: 119–121.

CRUDEN, R. W. 1970. Hawkmoth pollination of *Mirabilis multiflora* Nyctaginaceae. *Bulletin of the Torrey Botanical Club* 97: 89–91.

CRUDEN, R. W. 1973. Reproductive biology of weedy and cultivated *Mirabilis* Nyctaginaceae. *American Journal of Botany* 60: 802–809.

CUENOUD, P., V. SAVOLAINEN, L. W. CHATROU, M. POWELL, R. J. GRAYER, AND M. W. CHASE. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. *American Journal of Botany* 89: 132–144.

DARWIN, C. 1884. The different forms of flowers on plants of the same species. J. Murray, London, UK.

DESFEUX, C., S. MAURICE, J. P. HENRY, B. LEJEUNE, AND P. H. GOUYON. 1996. Evolution of reproductive systems in the genus *Silene*. *Proceedings of the Royal Society of London, B, Biological Sciences* 263: 409–414.

DOWNIE, S. R., D. S. KATZ-DOWNIE, AND K. J. CHO. 1997. Relationships in the Caryophyllales as suggested by phylogenetic analyses of partial chloroplast DNA ORF2280 homolog sequences. *American Journal of Botany* 84: 253–273.

DOWNIE, S. R., AND J. D. PALMER. 1994. A chloroplast DNA phylogeny of the Caryophyllales based on structural and inverted repeat restriction site variation. *Systematic Botany* 19: 236–252.

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

ESCUDERO, A., L. F. CARNES, AND F. PEREZ-GARCIA. 1997. Seed germination of gypsophytes and gypsovags in semi-arid central Spain. *Journal of Arid Environments* 36: 487–497.

ESCUDERO, A., J. M. IRIONDO, J. M. OLANO, A. RUBIO, AND R. C. SOMOLINOS. 2000. Factors affecting establishment of a gypsophyte: the case of *Lepidium subulatum* (Brassicaceae). *American Journal of Botany* 87: 861–871.

ESCUDERO, A., R. C. SOMOLINOS, J. M. OLANO, AND A. RUBIO. 1999. Factors controlling the establishment of *Helianthemum squamatum*, an endemic gypsophile of semi-arid Spain. *Journal of Ecology* 87: 290–302.

FISHER, R. A. 1941. Average excess and average effect of a gene substitution. *Annals of Eugenics* 11: 53–63.

FOSBERG, F. R. 1978. Studies in the genus *Boerhavia* Nyctaginaceae, parts 1–5. *Smithsonian Contributions to Botany*: 1–20.

FOWLER, B. A., AND B. L. TURNER. 1977. Taxonomy of *Selinocarpus* and *Ammocodon* Nyctaginaceae. *Phytologia* 37: 177–208.

GALLOWAY, L. A. 1975. Systematics of the North American desert species of *Abronia* and *Tripterocalyx* Nyctaginaceae. *Brittonia* 27: 328–347.

GRANT, V. 1983. The systematic and geographical distribution of hawk moth flowers in the temperate North American flora. *Botanical Gazette* 144: 439–449.

GRANT, V., AND K. A. GRANT. 1983. Hawkmoth pollination of *Mirabilis longiflora* (Nyctaginaceae). *Proceedings of the National Academy of Sciences, USA, Biological Sciences* 80: 1298–1299.

HARRIMAN, N. A. 1999. Synopsis of New World *Commicarpus* (Nyctaginaceae). *SIDA Contributions to Botany* 18: 679–684.

HEIMERL, A. 1889. Nyctaginaceae. *In* A. Engler and K. Prantl [eds.], Die Natürlichen Pflanzenfamilien, 14–32. Engelmann, Leipzig, Germany.

HEIMERL, A. 1934. Nyctaginaceae. *In* A. Engler and K. Prantl [eds.], Die Natürlichen Pflanzenfamilien, 86–134. Engelmann, Leipzig, Germany.

HERNÁNDEZ, H. 1990. Autopolinización en *Mirabilis longiflora* L. (Nyctaginaceae). *Acta Botanica Mexicana* 12: 25–30.

HODGES, S. A. 1995. The influence of nectar production on hawkmoth behavior, self-pollination, and seed production in *Mirabilis multiflora* (Nyctaginaceae). *American Journal of Botany* 82: 197–204.

IGIC, B., L. BOHS, AND J. R. KOHN. 2006. Ancient polymorphism reveals unidirectional breeding system shifts. *Proceedings of the National Academy of Sciences, USA* 103: 1359–1363.

JOHNSTON, I. M. 1941. Gypsophily among Mexican desert plants. *Journal of the Arnold Arboretum* 22: 145–170.

JORDAN, W. C., M. W. COURTNEY, AND J. E. NEIGEL. 1996. Low levels of intraspecific genetic variation at a rapidly evolving chloroplast DNA locus in North American duckweeds (Lemnaceae). *American Journal of Botany* 83: 430–439.

KEARNEY, T. H., AND R. H. PEEBLES. 1960. Nyctaginaceae. *In* T. H. Kearney and R. H. Peebles [eds.], Arizona flora, 270–279. University of California Press, Berkeley, California, USA.

KRUCKEBERG, A. R. 1986. The stimulus of unusual geologies for plant speciation—an essay. *Systematic Botany* 11: 455–463.

LANDE, R., AND D. SCHEMSKE. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* 39: 24–40.

LEE, S. 1978. A factor analysis study of the functional significance of angiosperm pollen. *Systematic Botany* 3: 1–19.

LEVIN, R. A. 2000. Phylogenetic relationships within Nyctaginaceae tribe Nyctagineae: evidence from nuclear and chloroplast genomes. *Systematic Botany* 25: 738–750.

LEVIN, R. A. 2002. Taxonomic status of *Acleisanthes*, *Selinocarpus*, and *Ammocodon* (Nyctaginaceae). *Novon* 12: 58–63.

LEVIN, R. A., R. A. RAGUSO, AND L. A. MCDADE. 2001. Fragrance chemistry and pollinator affinities in Nyctaginaceae. *Phytochemistry* 58: 429–440.

LOOMIS, W. E. 1944. Effect of heavy applications of gypsum on plant growth. *Plant Physiology* 19: 706–708.

LÓPEZ, H. A., AND L. GALETTO. 2002. Flower structure and reproductive biology of *Bougainvillea stipitata* (Nyctaginaceae). *Plant Biology* 4: 508–514.

LORD, E. M. 1981. Cleistogamy: a tool for the study of floral morphogenesis, function, and evolution. *Botanical Review* 47: 421–449.

MADDISON, W. P., AND D. R. MADDISON. 2006. Mesquite: a modular system for evolutionary analysis, version 1.1. Website: http://mesquiteproject.org.

MAHRT, M., AND R. SPELLENBERG. 1995. Taxonomy of *Cyphomeris* (Nyctaginaceae) based on multivariate analysis of geographic variation. *SIDA Contributions to Botany* 16: 679–697.

MCCLELLAN, Y., AND W.-J. BOECKLEN. 1993. Plant mediation of ant–herbivore associations: the role of sticky rings formed by *Boerhavia spicata*. *Coenoses* 8: 15–20.

MCGLAUGHLIN, M., K. KAROLY, AND T. KAYE. 2002. Genetic variation and its relationship to population size in reintroduced populations of pink sand verbena, *Abronia umbellata* subsp. *breviflora* (Nyctaginaceae). *Conservation Genetics* 3: 411–420.

MEYER, S. E. 1986. The ecology of gypsophile endemism in the eastern Mojave Desert. *Ecology* 67: 1303–1313.

MEYER, S. E., E. GARCÍA-MOYA, AND L. D. LAGUNES-ESPINOZA. 1992. Topographic and soil surface effects on gypsophile plant community patterns in central Mexico. *Journal of Vegetation Science* 3: 429–438.

NOWICKE, J. 1968. Palynotaxonomic study of the Phytolaccaceae. *Annals of the Missouri Botanical Garden* 55: 294–364.

NOWICKE, J. W. 1970. Pollen morphology in the Nyctaginaceae. Part 1: Nyctagineae, Mirabileae. *Grana* 10: 79–88.

NOWICKE, J. W. 1975. Pollen morphology in the order Centrospermae. *Grana* 15: 51–78.

NOWICKE, J. W., AND T. J. LUIKART. 1971. Pollen morphology of the Nyctaginaceae. Part 2: Colignonieae, Boldoeae, and Leucastereae. *Grana* 11: 145–150.

NYLANDER, J. A. A. 2004. MrModeltest, version 2. Program distributed by the author, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden. Website: http://www.abc.se/~nylander/mrmodeltest2/mrmodeltest2.html.

OLMSTEAD, R. G., AND J. A. SWEERE. 1994. Combining data in phylogenetic systematics—an empirical approach using 3 molecular data sets in the Solanaceae. *Systematic Biology* 43: 467–481.

OXELMAN, B., M. LIDEN, AND D. BERGLUND. 1997. Chloroplast *rps*16 intron phylogeny of the tribe Sileneae (Caryophyllaceae). *Plant Systematics and Evolution* 206: 393–410.

OYONARTE, C., G. SANCHEZ, M. URRESTARAZU, AND J. J. ALVARADO. 2002. A comparison of chemical properties between gypsophile and non-gypsophile plant rhizospheres. *Arid Land Research and Management* 16: 47–54.

PARSONS, R. F. 1976. Gypsophily in plants—a review. *American Midland Naturalist* 96: 1–20.

POOL, A. 2001. Nyctaginaceae. *In* W. D. Stevens, C. Ulloa U., A. Pool, and O. M. Montie [eds.], Flora de Nicaraugua, 1581–1592. Missouri Botanical Garden Press, St. Louis, Missouri, USA.

POSADA, D., AND K. A. CRANDALL. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

RAMBAUT, A. 1996. Se-Al: sequence alignment editor. Available at http://evolve.zoo.ox.ac.uk/.

RETTIG, J. H., H. D. WILSON, AND J. R. MANHART. 1992. Phylogeny of the Caryophyllales—gene sequence data. *Taxon* 41: 201–209.

REYES-SALAS, M., AND E. MARTÍNEZ-HERNÁNDEZ. 1982. Palynological catalog for the flora of Veracruz, Mexico. 8. Nyctaginaceae family. *Biotica (Mexico)* 7: 423–456.

RICHARDSON, J. E., R. T. PENNINGTON, T. D. PENNINGTON, AND P. M. HOLLINGSWORTH. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science* 293: 2242–2245.

RIDLEY, H. N. 1930. The dispersal of plants throughout the world. L. Reeve & Co., Ashford, UK.

RODMAN, J. E., M. K. OLIVER, R. R. NAKAMURA, J. U. MCCLAMMER JR., AND A. H. BLEDSOE. 1984. A taxonomic analysis and revised classification of Centrospermae. *Systematic Botany* 9: 297–323.

ROHWEDER, O., AND K. HUBER. 1974. Centrospermen-Studien. 7. Beobachtungen und Anmerkungen zur Morphologie und Entwicklungsgeschichte einiger Nyctaginaceen. *Botanishe Jahrbücher* 94: 327–359.

ROHWER, J. G. 1993. Phytolaccaceae. *In* K. Kubitski, J. G. Rohwer, and V. Bittrich [eds.], The families and genera of flowering plants, 506–515. Springer-Verlag, Berlin, Germany.

ROMAO, R. L., AND A. ESCUDERO. 2005. Gypsum physical soil crusts and the existence of gypsophytes in semi-arid central Spain. *Plant Ecology* 181: 127–137.

RONQUIST, F., AND J. P. HUELSENBECK. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

SANDERSON, M. J. 1991. In search of homoplastic tendencies—statistical inference of topological patterns in homoplasy. *Evolution* 45: 351–358.

SCHEMSKE, D. 1978. Evolution of reproductive characters in *Impatiens* (Balsaminaceae): the significance of cleistogamy and chasmogamy. *Ecology* 59: 596–613.

SMITH, A. G., D. G. SMITH, AND B. M. FUNNELL. 1994. Atlas of Mesozoic and Cenozoic coastlines. Cambridge University Press, Cambridge, UK.

SPELLENBERG, R. 1993. Taxonomy of *Anulocaulis* (Nyctaginaceae). *SIDA Contributions to Botany* 15: 373–389.

SPELLENBERG, R. 2000. Blooming "behavior" in five species of *Boerhavia* (Nyctaginaceae). *SIDA Contributions to Botany* 19: 311–323.

SPELLENBERG, R. 2003. Nyctaginaceae. *In* F. o. N. A. E. Committee [ed.], Flora of North America north of Mexico, 14–74. Oxford University Press, New York, New York, USA.

SPELLENBERG, R., AND R. K. DELSON. 1974. Aspects of reproduction in Chihuahuan Desert Nyctaginaceae. Transactions of the Symposium on the Biological Resources of the Chihuahuan Desert Region, United States and Mexico, 273–287. Sul Ross State University, Alpine, Texas, USA.

SPELLENBERG, R., AND J. M. POOLE. 2003. Nomenclatural adjustments and comments in *Abronia* and *Acleisanthes* (Nyctaginaceae). *SIDA Contributions to Botany* 20: 885–889.

STANDLEY, P. C. 1909. The Allioniaceae of the United States with notes on Mexican species. *Contributions from the United States National Herbarium* 12: 303–389.

STANDLEY, P. C. 1911. The Allioniaceae of Mexico and Central America. *Contributions from the United States National Herbarium* 13: 377–430.

STANDLEY, P. C. 1918. Allioniaceae. *In* N. L. Britton [ed.], North American Flora, 171–254. New York Botanical Garden Press, Bronx, New York, USA.

STANDLEY, P. C. 1931a. Studies of American plants: Nyctaginaceae. *Field Museum Botanical Series* 8: 304–311.

STANDLEY, P. C. 1931b. The Nyctaginaceae of northwestern South America. *Field Museum Botanical Series* 11: 71–126.

STANFORD, A. M., R. HARDEN, AND C. R. PARKS. 2000. Phylogeny and biogeography of Juglans (Juglandaceae) based on *matK* and ITS sequence data. *American Journal of Botany* 87: 872–882.

STEBBINS, G. L. 1974. Flowering plants. Evolution above the species level. Belknap Press of Harvard University Press, Cambridge, Massachusetts, USA.

STEVENS, P. F. 2001-onwards. Nyctaginaceae. Angiosperm phylogeny website http://www.mobot.org/MOBOT/research/APweb/.

SWOFFORD, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, Massachusetts, USA.

SYKES, W. R. 1987. The parapara, *Pisonia brunoniana* (Nyctaginaceae). *New Zealand Journal of Botany* 25: 459–466.

TAKEBAYASHI, N., AND P. L. MORRELL. 2001. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *American Journal of Botany* 88: 1143–1150.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, AND D. G. HIGGINS. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25: 4876–4882.

THULIN, M. 1994. Aspects of disjunct distributions and endemism in the arid parts of the Horn of Africa, particularly Somalia. Proceedings of the XIIIth plenary meeting AETFAT, 2: 1105–1119. National Herbarium and Botanic Gardens of Malawi, Zomba, Malawi.

TURNER, B. L. 1991. A new gypsophilic species of *Mirabilis* (Nyctaginaceae) from Nuevo León, Mexico. *Phytologia* 70: 44–46.

TURNER, B. L. 1993. A new species of *Anulocaulis* (Nyctaginaceae) from southern Coahuila, Mexico. *SIDA Contributions to Botany* 15: 613–615.

UYENOYAMA, M. K., K. E. HOLSINGER, AND D. M. WALLER. 1993. Ecological and genetic factors directing the evolution of self-fertilization. Oxford University Press, Oxford, UK.

WATERFALL, U. T. 1946. Observations on the gypsum flora of southwestern Texas and adjacent New Mexico. *American Midland Naturalist* 36: 456–466.

WHITE, T. J., T. BRUNS, S. LEE, AND J. TAYLOR. 1990. Amplifications and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* M. Innis, D. Gelfand, J. Sninsky, and T. White [eds.], PCR protocols: a guide to methods and applications, 315–322. Academic Press, San Diego, California, USA.

WILLIAMSON, P. S., L. MULIANI, AND G. K. JANSSEN. 1994. Pollination biology of *Abronia macrocarpa* (Nyctaginaceae), an endangered Texas species. *Southwestern Naturalist* 39: 336–341.

WILLSON, J., AND R. SPELLENBERG. 1977. Observations on anthocarp anatomy in the subtribe Mirabilinae (Nyctaginaceae). *Madrono* 24: 104–111.

WILSON, R. C. 1974. *Abronia.* Part 2. Anthocarp polymorphism and anatomy for 9 *Abronia* species found in California. *Aliso* 8: 113–128.

ZADOO, S. N., R. P. ROY, AND T. N. KHOSHOO. 1975. Cytogenetics of cultivated bougainvilleas. Part 2. Pollination mechanism and breeding system. *Proceedings of the Indian National Science Academy, Part B, Biological Sciences* 41: 498–502.

APPENDIX 1. Taxa, GenBank accession numbers, and voucher information used in this study. Regions not sampled are indicated by a dash. Cultivated plants were obtained from the following sources: DUBG = Duke University Botany Greenhouses, Durham; STRYB = Strybing Arboretum, San Francisco. Vouchers are deposited at the following herbaria: DUKE = Duke University, NMC = New Mexico State University, NY = New York Botanical Garden. † Accession included in "sensitivity" analysis. If sequences were downloaded from GenBank, then no voucher information is given.

Species—GenBank accession numbers: ITS, *ndhF*, *rpl16*, *rps16*, *Voucher specimen*, Locality, Year, Herbarium

Abronia bigelovii Heimerl†—EF079455, EF079510, EF079564, EF079606, *Douglas 2088,* New Mexico, USA, 2001, DUKE; ***Abronia carletonii*** J.M. Coult. & Fisher†—EF079456, EF079511, EF079565, EF079607, *Douglas 2091,* New Mexico, USA, 2001, DUKE; ***Acleisanthes lanceolatus*** (Wooton) R.A. Levin†—EF079454, EF079509, EF079563, EF079605, *Douglas 2072,* New Mexico, USA, 2001, DUKE; ***Acleisanthes longiflora*** A. Gray—EF079457, EF079512, —, EF079608, *Douglas 2098,* New Mexico, USA, 2001, DUKE; ***Allionia choisyi*** Standl.†—EF079467, EF079519, EF079574, EF079618, *Douglas 2187,* Coahuila, Mexico, 2002, DUKE; ***Andradea floribunda*** Allemão—EF079491, EF079545, —, EF079639, *Amorim 2294,* Brazil, 1998, NY; ***Anulocaulis annulatus*** (Coville) Standl.†—EF079503, EF079557, EF079599, EF079650, *Spellenberg 3162,* California, USA, 1993, NMC; ***Anulocaulis leiosolenus*** (Torr.) Standley ***v. leiosolenus*** Spellenberg—EF079464, EF079517, —, EF079615, *Douglas 2122,* Arizona, USA, 2002, DUKE; ***Anulocaulis reflexus*** I.M. Johnst.†— EF079468, EF079520, —, —, *Douglas 2192,* Chihuahua, Mexico, 2002, DUKE; ***Anulocaulis reflexus*** I.M. Johnst.†— —, —, EF079586, EF079629, *Spellenberg 10739,* Chihuahua, Mexico, 1990, NMC; ***Belemia fucsioides*** Pires—EF079488, EF079542, —, —, *Belem 3796,* Brazil, 1968, NY; ***Boerhavia anisophylla*** Torr.†— EF079469, EF079521, EF079575, EF079619, *Douglas 2194,* Durango, Mexico, 2002, DUKE; ***Boerhavia ciliata*** Brandegee— EF079465, —, EF079572, EF079616, *Douglas 2145,* Texas, USA, 2002, DUKE; ***Boerhavia coccinea*** Mill.†— EF079472, EF079525, EF079579, EF079622, *Spellenberg 13275,* Arizona, USA, 2001, DUKE; ***Boerhavia coulteri*** (S. Wats.) ***v. palmeri*** Spellenb.†— EF079471, EF079524, EF079578, EF079621, *Spellenberg 13273,* Arizona, USA, 2001, DUKE; ***Boerhavia dominii*** Meikle & Hewson†—EF079487, EF079540, EF079594, EF079638, *Smyth 42,* Australia, 1997, MO; ***Boerhavia gracillima*** Heimerl†—EF079479, EF079533, EF079587, EF079630, *Spellenberg 12447 ,* Texas, USA, 1997, NMC; ***Boerhavia intermedia*** M.E. Jones†—EF079474, EF079527, EF079581, EF079624, *Spellenberg 13279,* Arizona, USA, 2001, DUKE; ***Boerhavia lateriflora*** Standl.†—EF079466, EF079518, EF079573, EF079617, *Douglas 2161,* Sonora, Mexico, 2002, DUKE; ***Boerhavia linearifolia*** A. Gray†—EF079459, EF079514, EF079567, EF079610, *Douglas 2102,* New Mexico, USA, 2001, DUKE; ***Boerhavia purpurascens*** A. Gray†—EF079470, EF079523, EF079577, EF079620, *Spellenberg 13261,* Arizona, USA, 2001, DUKE;

Boerhavia repens L.†—EF079480, EF079534, EF079588, EF079631, *Spellenberg 7183,* Sana, Yemen , 1983, NMC; ***Boerhavia repens*** L. †—EF079477, EF079531, EF079584, EF079627, *Rose 2,* Oahu, Hawaii, USA, 2001, DUKE; ***Boerhavia spicata*** Choisy†—EF079473, EF079526, EF079580, EF079623, *Spellenberg 13276,* Arizona, USA, 2001, DUKE; ***Bougainvillea glabra*** Choisy†—EF079463, —, EF079571, EF079614, *Douglas 2121,* North Carolina, USA (DUBG), 2002, DUKE; ***Bougainvillea infesta*** Griseb.—EF079498, EF079551, —, EF079644, *Nee 51442,* Bolivia, 2000, NY; ***Caribea litoralis*** Alain— —, EF079530, —, —, *A. H. Liogier 7013,* Cuba, 1959, NY; ***Colignonia glomerata*** Griseb.—EF079495, EF079549, —, EF079642, *Nee 52523,* Bolivia, 2003, NY; ***Colignonia scandens*** Benth.†—EF079502, EF079556, EF079598, EF079648, *Grantham 63,* Lojas, Ecuador (STRYB), 2003, DUKE; ***Commicarpus coctoris*** N.A. Harriman†—EF079481, EF079535, EF079589, EF079632, *Spellenberg 12883,* Oaxaca, Mexico, 1998, NMC; ***Commicarpus plumbagineus*** (Cav.) Standl.†—EF079504, EF079558, EF079600, EF079651, *Spellenberg 7374,* Ta'izz, Yemen, 1983, NMC; ***Commicarpus scandens*** (L.) Standl.†—EF079482, EF079536, EF079590, EF079633, *Spellenberg 12887,* Puebla, Mexico, 1998, NMC; ***Cyphomeris gypsophiloides*** (M. Martens & Galeotti) Standl.†—EF079458, EF079513, EF079566, EF079609, *Douglas 2100,* New Mexico, USA, 2001, DUKE; ***Guapira discolor*** (Spreng.) Little†—EF079476, EF079529, EF079583, EF079626, *Spellenberg 13294,* Florida, USA, 2001, DUKE; ***Guapira eggersiana*** (Heimerl) Lundell—EF079496, EF079550, —, EF079643, *Mori 25542/40,* French Guiana, 2003, NY; ***Leucaster caniflorus*** (Mart.) Choisy— —, EF079541, —, —, *Pirani 3602,* Brazil, 1995, NY; ***Leucaster caniflorus*** (Mart.) Choisy—EF079497, —, —, —, *Hatschbach 50421,* Brazil, 1993, NY; ***Mirabilis albida*** (Walter) Heimerl†—EF079451, EF079506, EF079560, EF079602, *Douglas 2035,* Arizona, USA, 2001, DUKE; ***Mirabilis jalapa*** L.†—EF079461, EF079515, EF079569, EF079612, *Douglas 2119,* North Carolina, USA (DUBG), 2002, DUKE; ***Mirabilis multiflora*** (Torr.) A. Gray†—EF079452, EF079507, EF079561, EF079603, *Douglas 2037,* Arizona, USA, 2001, DUKE; ***Neea cauliflora*** Heimerl—EF079493, EF079547, —, —, *Schanke S15106,* Peru, 2002, NY; ***Neea hermaphrodita*** S. Moore—EF079489, EF079543, —, —, *Nee 51426,* Bolivia, 2000, NY; ***Neea psychotrioides*** Donn. Sm.†—EF079505, EF079559, EF079601, EF079652, *Wilbur 63654,* Heredia, Costa Rica, 1995, DUKE; ***Nyctaginia capitata*** Choisy†—EF079478, EF079532, EF079585, EF079628, *McIntosh 2049,* New Mexico, USA, 1992, NMC; ***Okenia hypogaea*** Schltdl. & Cham.†—EF079483, —, —, EF079634, *TR & RK Van Devender 92–1069,* Sonora, Mexico, 1992, NMC; ***Okenia hypogaea*** Schltdl. & Cham.†——, EF079522, EF079576, —, *Douglas 2206,* Veracruz, Mexico, 2002, DUKE; ***Phaeoptilum spinosum*** Radlk.†—EF079490, EF079544, —, —, *Seydel 4077,* Namibia, 1964, NY; ***Pisonia capitata*** (S. Watson) Standl.†—EF079484, EF079537, EF079591, EF079635, *AL Reina G. (2000—193),* Sonora, Mexico, 2000, NMC; ***Pisonia rotundata*** Griseb. †—EF079475, EF079528, EF079582, EF079625, *Spellenberg 13293,* Florida, USA, 2001, DUKE; ***Pisoniella arborescens*** (Lag. & Rodr.) Standl.†—EF079485, —, EF079592, EF079636, *LeDuc 231,* Oaxaca, Mexico, 1992, NMC; ***Pisoniella arborescens*** (Lag. & Rodr.) Standl. †— —, EF079539, —, —, *Anderson 13522,* Oaxaca, Mexico, 1988, NY; ***Ramisia brasiliensis*** Oliv.— EF079492, EF079546, —, EF079640, *Jardim 1507,* Brazil, 1998, NY; ***Reichenbachia hirsuta*** Spreng.†—EF079494, EF079548, EF079595, EF079641, *Nee 51972,* Bolivia, 2002, NY; ***Salpianthus arenarius*** Humb. & Bonpl.†—EF079486, EF079538, EF079593, EF079637, *Spellenberg 12903,* Michoacan, Mexico, 1999, NMC; ***Tripterocalyx carneus*** (Greene) L. A. Galloway†—EF079453, EF079508, EF079562, EF079604, *Douglas 2060,* New Mexico, USA, 2001, DUKE; ***Outgroups*: *Aptenia cordifolia*** (L. f.) Schwantes— —, AF194824, —, —, *;* ***Mollugo verticillata*** L.— —, AF194827, —, —, *;* ***Mollugo verticillata*** L.— —, —, —, EF079649, *Wilbur 77788,* North Carolina, USA, 2004, DUKE; ***Petiveria alliacea*** L.—EF079499, EF079552, —, —, *AL Reina G. 98–2048,* Sonora, Mexico, 1998, NY; ***Phytolacca americana*** Roxb.—EF079460, —, EF079568, EF079611, *Douglas 2118,* North Carolina, USA, 2002, DUKE; ***Phytolacca acinosa*** L— —, AF194828, —, —, *;* ***Rivina humilis*** L.†—EF079462, EF079516, EF079570, EF079613, *Douglas 2120,* North Carolina, USA (DUBG), 2002, DUKE; ***Sarcobatus vermiculatus*** (Hook.) Torr. †—EF079501, EF079555, EF079597, EF079647, *Spellenberg 13312,* Nevada, USA, 2002, DUKE; ***Stegnosperma cubense*** A. Rich.—EF079500, EF079554, EF079596, EF079646, *Salas–M. 2649,* Oaxaca, Mexico, 1999, NY; ***Trichostigma octandrum*** (L.) H. Walter— —, EF079553, —, EF079645, *Acevedo–Rodriguez 5447* Virgin Islands, USA 1993, NY.