INVESTIGATING *PERESKIA* AND THE EARLIEST DIVERGENCES IN CACTACEAE

CHARLES BUTTERWORTH

Desert Botanical Garden, 1201 North Galvin Parkway,
Phoenix, Arizona 85008, USA; cbutterworth@dbg.org

ERIKA J EDWARDS

Department of Ecology & Evolutionary Biology, Brown University,
Providence, Rhode Island 02912, USA; erika_edwards@brown.edu

Abstract: The cacti are renowned desert plants of the New World. The are typically leafless, spine-bearing stem succulents. Species of *Pereskia* possess broad, regular leaves and have been viewed as being representative of ancestral cacti. A number of previous studies have attempted to resolve phylogenetic relationships within *Pereskia* and between *Pereskia* and other cacti. Here, we present the results of a joint analysis along with hypothesis testing using datasets previously published independently by each author. This study clearly shows a basal split in the cacti between a clade of Caribbean-basin-centered *Pereskia* species and all other cacti. Furthermore, hypothesis testing strongly supports the basal *Pereskia* hypothesis over an alternative hypothesis in which the Opuntioideae form the basal split in the cacti (although the latter hypothesis was not statistically rejected).

Key words: *Pereskia*, Cactaceae, cacti, phylogenetics, evolution

Introduction

Except for a single, epiphytic species, cacti are naturally restricted to the New World, where they represent a spectacular arid land–plant radiation. In the majority of cacti, persistent leaves are replaced by spines, and the green succulent stem has become the primary photosynthetic organ. A few genera do retain persistent leaves, however, and these have been treated as the most ancestral-like members of the family. Spines (developmentally modified leaves) are typically borne from areoles (specialized axillary buds that may bear spines, flowers, or fruit). Flower structure is predominated by receptacular epigyny; however, in the leafy genus *Pereskia* Miller there is variation in ovary position from superior to inferior ovaries.

Traditionally most cactus classifications recognized three subfamilies (Barthlott and Hunt 1993; Hunt and Taylor 1986, 1990). The Pereskioideae Schumann possess morphological features, such as persistent leaves and superior to inferior ovaries, that are considered primitive or relictual within the Cactaceae (Boke 1964; Gibson and Nobel 1986). For these reasons, members of the Pereskioideae are considered to be united by shared plesiomorphies rather than distinct synapomorphies. The Pereskioideae has traditionally been comprised of two genera, *Pereskia* and *Maihuenia* (Philippi ex F. A. C. Weber) Schumann, the latter being unusual due to its natural distribution in the extreme south of South America and its stem and leaf features (for instance, distribution of mucilage-bearing structures) unknown elsewhere in the family (Gibson 1977). The Opuntioideae Burnett includes prickly-pears, chollas, and other members that share distinct morphological and anatomical features, such as the presence of glochids (very short, fine, barbed spines), and seeds with a funicular envelope, usually forming a bony, white aril (Gibson and Nobel 1986). The Cactoideae are morphologically variable, from tall, columnar species to low-growing, globose cacti to epiphytic plants possessing flattened, almost leaf-like stems. There are no obvious morphological or anatomical synapomorphies uniting the Cactoideae, except possibly stem succulence and lack of persistent leaves and glochids, the latter being a derived trait in the Opuntioideae. Furthermore, the chloroplast *rpo*C1 intron is missing in all Cactoideae (Wallace and Cota 1996) with the exception of *Blossfeldia* (Butterworth 2006).

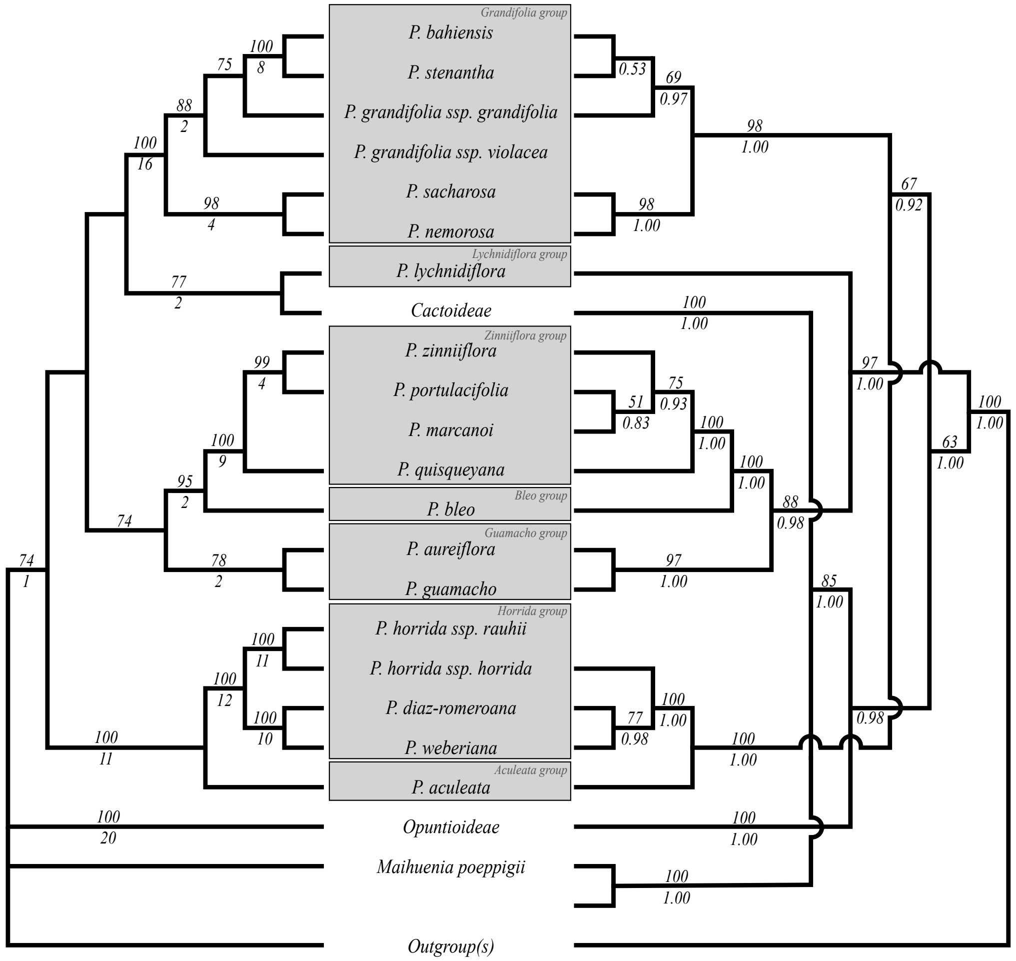

Figure 1. Phylogenetic relationships in Pereskia and basal Cactaceae. The cladogram on the left (redrawn from Butterworth and Wallace 2005) is the strict consensus of six most parsimonious trees for combined chloroplast *rpl*16 intron and *psb*A-*trn*H intergenic spacer sequence data. Bootstrap support ≥50% are shown above the branches, decay values are shown below the branches. The cladogram on the right (redrawn from Edwards and others 2005) is the Bayesian consensus tree for combined chloroplast *trn*K/*mat*K, *rbc*L, and *psb*A-*trn*H IGS, mitochondrial *cox*3, and nuclear *phy*C sequence data. Maximum parsimony bootstrap support ≥50% is shown above branches, Bayesian posterior probabilities are shown below branches. Shaded boxes denote species groups defined by Leuenberger (1986).

A number of studies have attempted to resolve subfamilial phylogenetic relationships in Cactaceae. Results of a cpDNA restriction-site survey of members of *Pereskia* (Pereskioideae) were summarized by Wallace (1995). These data, supplemented by chloroplast sequence data (Butterworth and Wallace 2005) suggest that Pereskioideae is not monophyletic for the following reasons: 1) sampled members of Cactoideae are embedded within *Pereskia*, and 2) *Maihuenia*, Opuntioideae, and the Cactoideae form an unresolved trichotomy.

Edwards and others (2005) presented a phylogeny based on sequence data from all three genomes (chloroplast, mitochondrion, nucleus) for *Pereskia* and representatives from Opuntioideae, Maihuenioideae, and Cactoideae. The maternally inherited genomes (chloroplast and mitochondrion) did not yield strong support in their placements' subfamilies. However, the nuclear component (phytochrome C) and the maternal data set both yielded a monophyletic Opuntioideae, which formed the sister group to a clade that included some *Pereskia* species, *Maihuenia*, and Cactoideae.

Here we present a summary of the two previous molecular (DNA) phylogenetic studies of *Pereskia* (Butterworth and Wallace 2005; Edwards and others 2005). We then present, for the first time, the analysis of combined DNA sequence data used in those studies.

Butterworth CA and Wallace RS 2005. Published in December, 2005, this study added DNA sequence data from the chloroplast *rpl*16 intron and *psb*A-*trn*H IGS to the chloroplast restriction-site data of Wallace (1995). The strict consensus from maximum-parsimony analysis for the combined data set (Fig 1) gave greater phylogenetic resolution than that of Wallace (1995) and recovered several key *Pereskia* clades as outlined by Leuenberger (1986), including the Horrida, Zinniiflora, Guamacho, and Grandiflora groups. This study also placed several enigmatic taxa with good statistical support: *P. aculeata* Miller was united with the Horrida group and *P. bleo* with the Zinniiflora group. Relationships among the various subfamilies remained ambiguous, but importantly, there was no indication that *Maihuenia* should be united with *Pereskia*, and there was moderately strong support for sampled Cactoideae being nested within *Pereskia*.

Edwards E, Nyffeler R, Donoghue MJ 2005. This study, published in August 2005, not only utilized chloroplast *trn*K/*mat*K, *psb*A-*trn*H IGS, and *rbc*L, but also nuclear *phy*C and mitochondrial *cox*3 sequence data, that is, DNA sequence data from all three DNA compartments (chloroplast, mitochondrion, nucleus). The nuclear marker (*phy*C) provided the greatest statistical support to date for the deeper nodes within the phylogeny. The *Pereskia* clades outlined by Leuenberger (1986) and demonstrated by Butterworth and Wallace (2005) were also recovered in the Edwards study, yet a number of differences in arrangement of major clades (see Fig 1) can be seen. While *Pereskia* species from northern South America and the Caribbean (Guamacho, Bleo, and Zinniiflora groups), and species from the Andes region of South America (Horrida and Aculeata groups) form single clades in both studies, topological differences occur in phylogenetic placements of the Grandifolia and Lychnidiflora groups, which form a clade with the Cactoideae in the Butterworth and Wallace phylogeny. Also worth noting, both studies indicate a possible long-distance dispersal event due to the phylogenetic positioning of the Brazilian endemic *Pereskia aureiflora* Ritter with *P. guamacho* Weber from Venezuela. It is interesting that both these species are unique within *Pereskia* for their yellow-colored flowers, possibly related to their major pollinators.

In both studies *Pereskia* was shown to be paraphyletic, although there were differences regarding which other groups were nested within the genus. Butterworth and Wallace (2005) placed sampled members of the Opuntioideae and Maihuenioideae outside the clade containing *Pereskia*. However, the basal polychotomy does not allow inferences regarding the "ancestral" lineage of the Cactaceae. Sampled species of the Cactoideae were placed, phylogenetically within *Pereskia*, in a sister-group relationship with *P. lychnidiflora* De Candolle. Edwards and others (2005) also demonstrated a paraphyletic *Pereskia*, although their phylogeny indicated the presence, within *Pereskia*, of a clade containing members of the Cactoideae, Opuntioideae, and Maihuenioideae. Furthermore, according to Edwards and others, the fundamental split in the Cactaceae is between a lineage of northern *Pereskia* species and all other cacti termed the "Caulocacti" (in reference to the presence of stem stomata and delayed bark formation, two traits shared by most members of this clade). The caulocacti include the South American *Pereskia* species and the "core" cacti—Opuntioideae, Cactoideae, and Maihuenioideae. In general, statistical support for this phylogenetic hypothesis is the strongest of all previous studies of basal cactus phylogeny, but several nodes still remain quite tentative. Furthermore, most of the support for this hypothesis comes strictly from the *phy*C marker, as all sampled chloroplast and mitochondrion markers have not proved overly useful for resolving the basal cactus problem. A logical next step in studies of early cactus evolution will be to sequence additional gene regions, but we first thought it a useful exercise to perform a combined analysis of data from both studies.

Materials & Methods

Laboratory techniques (DNA extraction, PCR, sequencing, etc) are detailed in Butterworth and Wallace (2005) and Edwards and others (2005). DNA sequence data from both studies were combined using MacClade 4.08 (Maddison and Maddison 2005). Only taxa for which data from all three genomes were present were retained for analyses. The total data matrix consisted of 7192 characters for 28 taxa from the Cactaceae and five taxa from the Portulacaceae. The final data matrix had a total of 10.23% cells scored as missing data. These cells included indels and regions from taxa for which no sequence data was available.

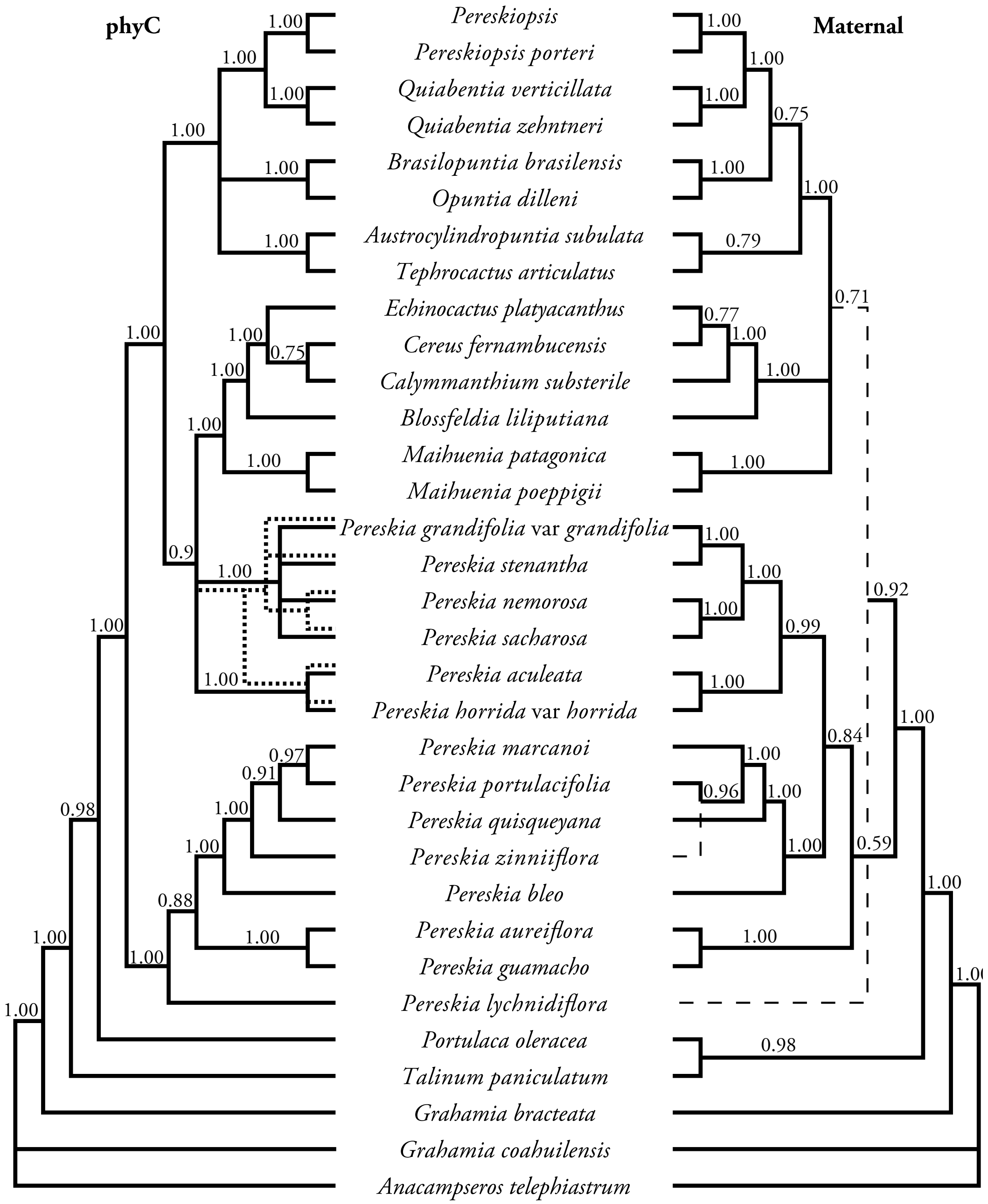

Figure 2. Bayesian consensus cladograms for the nuclear (*phy*C) and maternal datasets. Numbers above the branches are Bayesian posterior probabilities. The narrowly-dashed line for *phy*C represents the Maximum Likelihood topology. The widely-dashed line for the maternal dataset allows for direct comparison with the *phy*C cladogram without re-arranging taxon order.

Bayesian analyses. Models of sequence evolution were estimated using MrModeltest var2 (Nylander 2004) implemented in PAUP* (Swofford 2002). Selected models were F81 (Felsenstein 1981) for *rpl*16; HKY (Hasegawa and others 1985) for *rbc*L and *phy*C; and GTR (Tavare 1986) for *psb*A-*trn*H, *trn*K/*mat*K, and *cox*3. Additionally, all regions with the exception of *cox*3, which utilized the proportion of invariable-sites model, also utilized a gamma-distributed rates model (Yang 1993). Phylogenetic reconstruction was undertaken using MrBayes var 3.1.2 (Ronquist and Huelsenbeck 2003). Incongruence Length Difference (ILD) tests were implemented in PAUP* to test congruence between the three genomes. With all taxa included, the ILD test rejected the null hypothesis of congruence with a p–value of

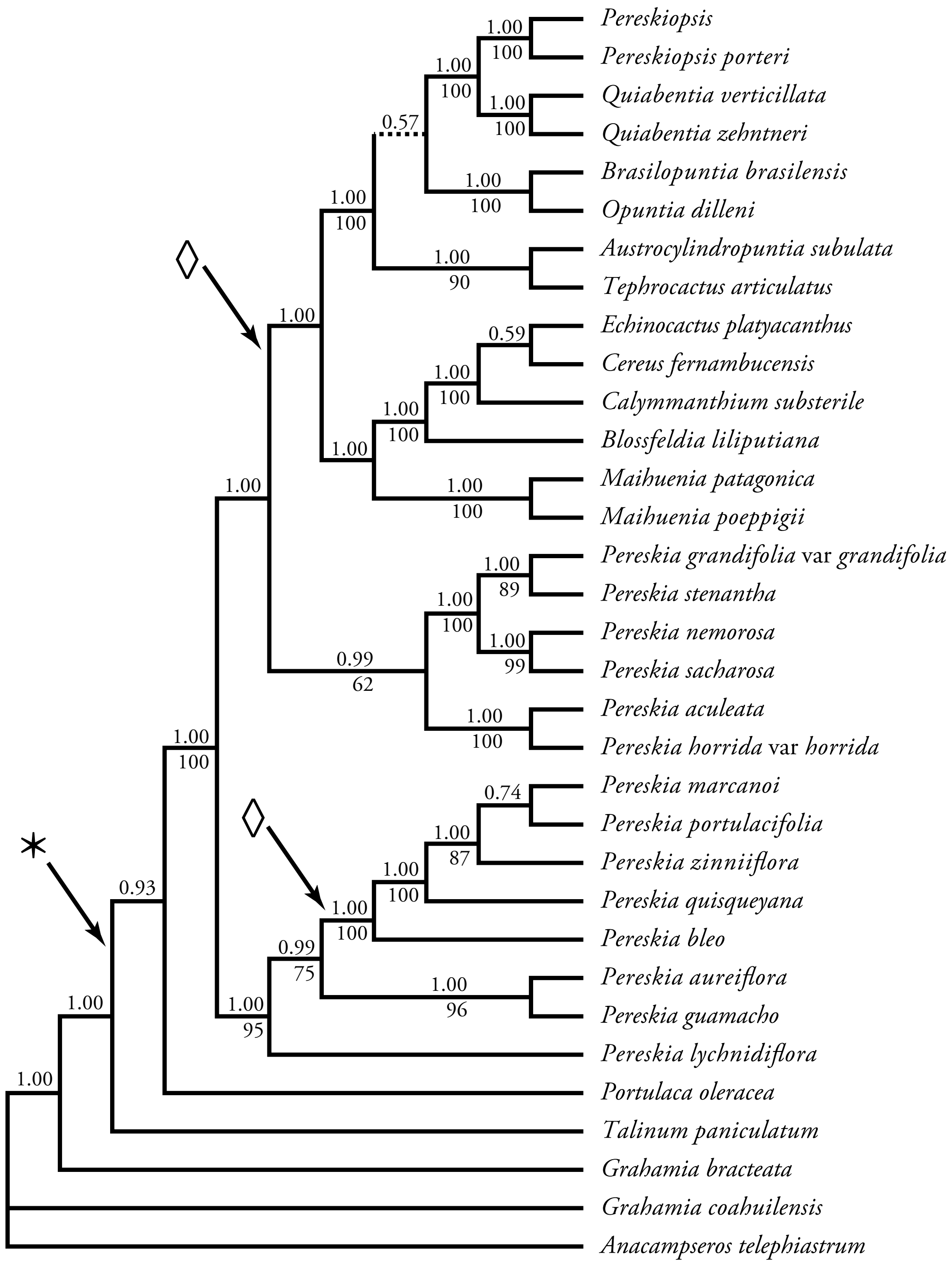

Figure 3. Bayesian consensus cladogram for the combined analysis. Bayesian posterior probabilities are shown above the branches and maximum parsimony bootstrap above 50% are shown below the branches. Dashed line represents branch collapsed in Maximum Likelihood cladogram. * represents change from superior to half inferior ovaries; ◊ represents change from half-inferior to fully inferior ovaries.

0.01. However, maximum parsimony analyses of the individual genomes showed that the main topological incongruence was due to the placement of the *Pereskia* species from Brazil and the Andes region of South America. When these taxa were removed, the ILD test showed significant congruence between all genomes ($p = 0.08$). Three dataset combinations were analyzed (all regions, chloroplast + *cox*3, and *phy*C) in partitioned analyses so that each region could be run under their specific model. Analyses were run using the MrBayes default prior parameters in Markov-chain Monte Carlo (MCMC) searches using four incrementally heated chains for 1,000,000 generations sampled every 100 generations. Following the MCMC analyses, posterior parameter distributions were viewed using Tracer v1.4 (Rambaut and Drummond 2007), which indicated that exclusion of the first 100 trees (10,000 generations) was more than adequate to ensure stationarity.

Hypothesis Testing. Different hypotheses of Cactaceae evolution were undertaken using constraint trees and Maximum Likelihood (ML) analyses using PAUP*. Suitable models of sequence evolution were K81 + I + G (Kimura 1981) for the complete dataset, TVM + I + G for the maternal dataset, and HKY + G (Hasegawa and others 1985) for nuclear dataset were estimated using ModelTest (Posada and Crandall 1998). These models were then used in maximum likelihood analyses on the unconstrained datasets. The same analyses and models were then repeated using topological constraints for the hypotheses being tested: 1. "basal" *Pereskia*; 2. "basal" Opuntioideae; 3. "basal" *Maihuenia*. The ML trees from each of these studies were compared to their unconstrained counterparts using the Shimodaira-Hasegawa Test (Shimodaira and Hasegawa 1999) as implemented in PAUP* using full optimization and 1000 bootstrap replications.

Results

The Bayesian consensus trees from all three datasets (maternal, nuclear, combined) supported monophyly of the Cactaceae as well as subfamilies Opuntioideae and Cactoideae (Figs 2, 3). Furthermore, all data analyses supported the hypothesis of *Pereskia* representing the most ancestral group within the family. However, there is incongruence between the maternally inherited datasets and *phy*C, particularly with regard to the placement of the species from Brazil and the Andes region of South America (*P. aculeata*, *P. horrida*, *P. grandifolia*, *P. stenantha*, *P. nemorosa*, *P. sacharosa*).

Nuclear Data. The earliest divergence within the Cactaceae is between a clade containing *Pereskia* species from the Caribbean (*P. bleo*, *P. quisqueyana*, *P. zinniiflora*, *P. portulacifolia*, *P. marcanoi*) as well as *P. lychnidiflora*, *P. aureiflora*, and *P. guamacho*, and the remaining sampled members of the Cactaceae (Fig 2). Within the remaining cacti, a divergence clearly reveals a monophyletic Opuntioideae and a second clade containing a trichotomy involving *P. aculeata* and *P. horrida*; the southern South American species (*P. grandifolia*, *P. stenantha*, *P. nemorosa*, *P. sacharosa*) in an unresolved polychotomy; and a final clade containing *Maihuenia* in a sister-group relationship to the Cactoideae.

Maternal data (chloroplast, mitochondrion). With the exception of a single species (*P. lychnidiflora*), the major early divergence within the Cactaceae is between a clade containing species of *Pereskia* and all other cacti (Fig 2). Furthermore, *P. lychnidiflora* is not embedded within the remaining cacti; rather, it forms the sister-group to the other Cactaceae. Within the other cacti, there is an unresolved trichotomy involving *Maihuenia*, Opuntioideae, and Cactoideae.

Combined data. The Bayesian consensus cladogram (Fig 3) for the combined dataset is topologically highly congruent with the *phy*C cladogram, differing only by being fully resolved. Furthermore, all of the deep nodes within the cacti are supported with high Bayesian posterior probabilities (that is, 99 or 100%).

Hypothesis testing. With the exception of the combined and nuclear datasets for the "basal" *Maihuenia* hypothesis, none of the hypotheses of early cactus evolution were fully rejected using the SH-Test (Table 1). A comparison of the p-values for "basal" Opuntioideae and "basal" *Pereskia* show that the latter hypothesis has better statistical support than the former. However, it is important to note that both these hypotheses were not statistically rejected.

Comparison between this and previous studies. In essence, combination of the sequence data from Butterworth and Wallace (2005) and Edwards and others (2005) is equivalent to adding chloroplast *rpl*16 data to the latter dataset. In terms of topology, the combined dataset in this paper (Fig 3) and Edwards and others (2005) are highly congruent.

Discussion

Given that the Bayesian consensus cladogram (Fig 3) in this study is topologically highly congruent with that in Edwards and others (2005) it is unnecessary to include a detailed discussion here. However, a number of issues pertinent to early cactus evolution can be addressed based upon the results of the study presented here.

"Basal" relationships in the Cactaceae. The phylogeny presented here, as well as results from the hypothesis testing, allow a number of scenarios of cactus evolution to be scrutinized. The traditional point of view that the earliest members of the Cactaceae were *Pereskia*-like (Barthlott and Hunt 1993; Britton and Rose 1919; Cronquist 1981; Gibson and Nobel 1986) seems to be borne out in this study. The

Table 1. Results from hypothesis testing using Maximum Likelihood and the Shimodaira-Hasegawa Test. Maternal data incorporates sequence data from the chloroplast *rbc*L, *rpl*16 intron, *psb*A-*trn*H IGS, *trn*K/*mat*K, and the mitochondrial *cox*3. Hypotheses rejected below the 5% value are marked with an asterisk (*).

	Basal Opuntioideae	Basal *Pereskia*	Basal *Maihuenia*
Combined Data	0.067	1.00	0.047*
Maternal Data	0.335	1.00	0.418
Nuclear (*phyC*)	0.081	0.224	0.016*

Bayesian consensus cladogram (Fig 3) clearly shows that *Pereskia* is paraphyletic, with *Maihuenia*, Opuntioideae, and Cactoideae embedded within.

Griffith (2004) discussed an alternative hypothesis in which a leafless stem succulent of the Opuntioideae represented the ancestral form of Cactaceae. However, the cladograms presented in this study do not support a "basal" position for the Opuntioideae (also demonstrated by a parametric bootstrapping test in Edwards and others 2005). It is important to note, however, that the results of the hypothesis testing did not statistically reject the hypothesis of a "basal" Opuntioideae. Griffith presents a good discussion of problems based on the conceptual model of the phylogeny of the Cactaceae (Griffith 2004) including ovary position, which he suggests is difficult to polarize in the Portulacaceae/Cactaceae group due to the presence of half-inferior ovaries in *Portulaca*. However, based on the Bayesian cladogram (Fig 3), we can easily hypothesize an evolutionary transition from superior ovaried ancestors via the half-inferior state to the convergent acquisition of inferior ovaries in the *Pereskia bleo* clade and the Opuntioideae/Cactoideae clade.

The presence of persistent broad leaves in *Pereskia* also puzzled Griffith (2004). He argued that because Portulacaceous relatives of cacti (namely *Anacampseros* and *Talinum*) possessed small succulent leaves lacking strong venation, the leaves of *Pereskia* could have evolved to superficially resemble those of typical dicotyledonous plants. However, until we have confirmed an accurate placement of the Cactaceae within the Portulacaceae (see Nyffeler 2006; Nyffeler and others, this volume pp 26–36) we cannot be sure of the polarity of leaf form in this group of plants. Furthermore, it is even plausible that *Pereskia* might have retained an ancestral leaf form while the occurrence of small succulent leaves happened in numerous independent lineages in the Portulacaceae/Cactaceae as the result of convergent evolution. The same argument could also apply to wood characters of *Pereskia*, which are considered to be highly relictual within the family (Mauseth and Landrum 1997). Griffith (2004) points out that woody members of the Caryophyllales usually possess some form of anomalous secondary growth or stem structure (Cronquist 1981, 1988).

The phylogenies presented in this paper support the hypothesis of *Pereskia* representing the earliest members of the cacti. This conclusion supports earlier hypotheses of cactus evolution following a course from ancestral woody shrubs with typical dicotyledonous leaves in *Pereskia* to leafless stem succulents encountered in the Opuntioideae and Cactoideae (Barthlott and Hunt 1993; Britton and Rose 1919; Cronquist 1981; Gibson and Nobel 1986).

However, understanding early evolutionary events with certainty in the Cactaceae will be difficult. The patterns of DNA sequence divergence in molecular data collected thus far suggest rapid radiation in early cactus history. Uncovering the sequence of divergence events during this time (and thus, the relationships among the major cactus clades) may require an extremely large amount of data. Equally challenging will be accurate character reconstruction—as there is no fossil record for cacti—and it is likely that many lineages that might have represented intermediate steps in the transition to the typical, leafless cactus have since gone extinct. In spite of these obstacles, we have already gained much clarity on the subject in recent years using the handful of extant taxa that are available, and we are certain that additional efforts will yield further insight into the early evolution of this important plant lineage.

Acknowledgements

The authors would like to thank the following people and institutions for their assistance and support: Wendy Applequist, Miriam Diaz, Urs Eggli, Beat Leuenberger, the Desert Botanical Garden (Phoenix, Arizona, USA), the Berlin Botanical Garden (Berlin-Dahlem, Germany), the Sukkulenten-Sammlung (Zürich, Switzerland), and the Jardín Botánico Nacional (Santo Domingo, Dominican Republic).

Literature cited

Barthlott W, Hunt D. 1993. Cactaceae. pp 161–197 *In* The families and genera of vascular plants (K Kubitzki, ed) Springer-Verlag, Berlin.

Boke NH. 1964. The cactus gynoecium: a new interpretation. *American Journal of Botany* 51: 598–610.

Britton NL, Rose JN. 1919. *The Cactaceae*. Carnegie Institution, Washington.

Butterworth CA. 2006. Resolving "Nyffeler's Puzzle"—the intriguing taxonomic position of *Blossfeldia*. *Haseltonia* 12: 3–10.

Butterworth CA, Wallace RS. 2005. Molecular phylogenetics of the leafy cactus genus *Pereskia* (Pereskioideae). *Systematic Botany* 30: 800–808.

Cronquist A. 1981. *An integrated system of classification of flowering plants*. Columbia University Press.

Cronquist A. 1988. *The evolution and classification of flowering plants*. New York Botanical Garden, Bronx.

Edwards EJ, Nyffeler R, Donoghue MJ. 2005. Basal cactus phylogeny: implications of *Pereskia* (Cactaceae) paraphyly for the transition to the cactus life form. *American Journal of Botany* 92: 1177–1188.

Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17: 368–376.

Gibson AC. 1977. Vegetative anatomy of *Maihuenia* (Cactaceae) with some theoretical discussions of ontogenetic changes in xylem cell types. *Bulletin of the Torrey Botanical Club* 104: 35–48.

Gibson AC, Nobel PS. 1986. *The cactus primer*. Harvard University Press, Cambridge.

Griffith MP. 2004. What did the first cactus look like? An attempt to reconcile the morphological and molecular evidence. *Taxon* 53: 493–499.

Hasegawa M, Kishino H, Yano T. 1985. Dating the human-ape split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 22: 160–174.

Hunt D, Taylor N. 1986. The genera of the Cactaceae: towards a new consensus. *Bradleya* 4: 65–78.

Hunt D, Taylor N. 1990. The genera of Cactaceae: progress towards consensus. *Bradleya* 8: 85–107.

Kimura M. 1981. Estimation of the evolutionary distances between homologous nucleotide sequences. *Proceedings of the American Academy of Arts and Sciences* 78: 454–458.

Leuenberger BE. 1986. *Pereskia* (Cactaceae). *Memoirs of The New York Botanical Garden* 41: 1–141.

Maddison D, Maddison W. 2005. MacClade, version 4.08. Sinauer.

Mauseth JD, Landrum JV. 1997. Relictual vegetative anatomical characters in Cactaceae: the genus *Pereskia*. *Journal of Plant Research* 110: 55–64.

Nyffeler R. 2006. The closest relatives of cacti: insights from phylogenetic analyses of chloroplast and mitochondrial sequences with special emphasis on relationships in the tribe Anacampseroteae. *American Journal of Botany* 94: 89–101.

Nylander JA. 2004. MrModeltest, version 2. Evolutionary Biology Centre, Uppsala University.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

Rambaut A, Drummond AJ. 2007. Tracer. University of Edinburgh.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* 16: 1114–1116.

Swofford DC. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates.

Tavare S. 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. *Lectures in Mathematics and Life Sciences* 17: 57–86, 368–376.

Wallace RS. 1995. Molecular systematic study of the Cactaceae: Using chloroplast DNA variation to elucidate Cactus phylogeny. *Bradleya* 13: 1–12.

Wallace RS, Cota JH. 1996. An intron loss in the chloroplast gene *rpoC1* supports a monophyletic origin for the subfamily Cactoideae. *Current Genetics* 29: 275–281.