

Biogeography and Floral Evolution of Baobabs (*Adansonia*, Bombacaceae) as Inferred From Multiple Data Sets

DAVID A. BAUM,¹ RANDALL L. SMALL,² AND JONATHAN F. WENDEL²

¹Department of Organismic and Evolutionary Biology, Harvard University Herbaria, Cambridge, Massachusetts 02138, USA; E-mail: dbaum@oeb.harvard.edu

²Department of Botany, Bessey Hall, Iowa State University, Ames, Iowa 50011, USA

Abstract.—The phylogeny of baobab trees was analyzed using four data sets: chloroplast DNA restriction sites, sequences of the chloroplast *rpl16* intron, sequences of the internal transcribed spacer (ITS) region of nuclear ribosomal DNA, and morphology. We sampled each of the eight species of *Adansonia* plus three outgroup taxa from tribe Adansonieae. These data were analyzed singly and in combination using parsimony. ITS and morphology provided the greatest resolution and were largely concordant. The two chloroplast data sets showed concordance with one another but showed significant conflict with ITS and morphology. A possible explanation for the conflict is genealogical discordance within the Malagasy *Longitubae*, perhaps due to introgression events. A maximum-likelihood analysis of branching times shows that the dispersal between Africa and Australia occurred well after the fragmentation of Gondwana and therefore involved overwater dispersal. The phylogeny does not permit unambiguous reconstruction of floral evolution but suggests the plausible hypothesis that hawkmoth pollination was ancestral in *Adansonia* and that there were two parallel switches to pollination by mammals in the genus. [Biogeography, data set conflict, floral evolution, Gondwana, introgression, molecular clock, phylogeny.]

The eight species of baobab are distinctive trees growing in the old world tropics. One species (*A. digitata* L.) occurs throughout the drier parts of Africa, a second species is restricted to north-western Australia (*A. gibbosa* [A. Cunn.] D. Baum), and the remaining six species are endemic to Madagascar. Obtaining a phylogeny of the genus is of interest because of its unusual biogeography, its diverse floral biology, and its great ecological importance in many Malagasy communities (Perrier de la Bâthie, 1955; Baum, 1995a, 1995b, 1996).

Adansonia L. is almost certainly monophyletic based on a calyx that fully encloses the flower bud and a distinctive woody, indehiscent fruit found nowhere else in Bombacaceae. The current distribution of the baobabs has been explained either in terms of vicariance during the breakup of Gondwana (Maheshwari, 1971; Raven and Axelrod, 1972; Aubréville, 1975) or long-distance dispersal from Africa/Madagascar to Australia (Raven and Axelrod, 1974; Armstrong, 1983). A phylogenetic analysis of *Adansonia* could shed light on these

alternative biogeographic hypotheses if it can help date the divergence of the Australian species from the rest of the genus.

Within *Adansonia* there is a striking diversity in flower morphology, which has been the basis for the recognition of three sections (Hochreutiner, 1908; Baum, 1995a). *Adansonia* sect. *Adansonia* comprises just the type species, *A. digitata*, which has globose flower buds that hang down on long flower stalks (comprising both the pedicel and peduncle). This continental African species has long been known to be bat-pollinated (Porsch, 1935; Jaeger, 1945), but is also visited by bushbabies (Coe and Isaac, 1965). *Adansonia* sect. *Brevitubae* comprises two Malagasy species, *A. grandidieri* Baill. and *A. suarezensis* H. Perrier, which have ovoid flower buds borne on short erect stalks. Although fruit bats are likely pollinators of both species, nocturnal lemurs contribute significantly to the pollination of *A. grandidieri* (Baum, 1995b). *Adansonia* sect. *Longitubae* includes the remaining four Malagasy species and the Australian species. They have long, cylindrical

flower buds borne either erect or horizontally and are primarily pollinated by large hawkmoths (Baum, 1995b). In view of the great morphological variability within the genus, one might expect morphological data to be well suited to phylogenetic analysis. However, the majority of available characters are floral and, given the clear correlation between floral morphology and pollination in the genus, it would be inadvisable to trust morphology alone for phylogenetic analysis. Concerted convergence of suites of pollination-related characters could distort phylogenetic inference. Thus, in this study we used both molecular and morphological data to estimate the phylogeny of *Adansonia*.

MATERIAL AND METHODS

Terminal Taxa

Appendix 1 lists the taxa used in this analysis and voucher information. Four ingroup accessions and one outgroup could not be scored for restriction sites due to a paucity of DNA. As outgroups we used three genera of Adansonieae, the tribe that includes *Adansonia* (Bakhuizen van der Brink, 1924; Edlin, 1935; Hutchinson, 1967): *Pachira* Aubl. and *Pseudobombax* Dugand from the neotropics, and *Bombax* L. from the paleotropics. Higher-level molecular phylogenetic analyses based on *ndhF* (W. Alverson, B. Whitlock, R. Nyffeler, and D. Baum, unpublished data) confirm the close relationship between *Adansonia* and the outgroups we included here.

Morphology

The scoring of morphological data within *Adansonia* was conducted on fresh material observed in the field. Information on field sites and methods was published previously (Baum, 1995a, 1995b). The scoring of the three outgroup species was based on field observations, photographs, and herbarium specimens (Arnold Arboretum, Gray Herbarium). In

view of the lack of morphological variation within the named species of *Adansonia*, multiple accessions of a named species were scored identically for morphological data.

DNA Extraction

When fresh leaf tissue was available, genomic DNA was isolated following the protocol of Wallace and Cota (1996) with the omission of the cesium chloride density gradient centrifugation. Some additional samples for polymerase chain reaction (PCR) amplification were obtained from individual seeds using a CTAB extraction procedure (see Baum et al., 1994).

Chloroplast DNA Restriction Site Analysis

Approximately 1 μ g of total genomic DNA was digested with a battery of frequently cutting (4–5 bp recognition sites) restriction enzymes (*AluI*, *BstOI*, *DdeI*, *DpnII*, *HaeIII*, *Hinfi*, *HinPII*, *MspI*, *RsaI*, and *TaqAI*) according to the manufacturers' instructions. The digested DNAs were electrophoresed and transferred to nylon membranes for probing with radiolabeled *Petunia* or *Lactuca* cpDNA probes (Sytsma and Gottlieb, 1986; Jansen and Palmer, 1987). Together these probes cover approximately 95% of the large single-copy region, the entire small single-copy region, and approximately 50% of the inverted repeats of the chloroplast genome. Prehybridization (1–6 hr) and hybridization (12–16 hr) were carried out at 65°C in 6 \times or 8 \times SSC, 1% sodium dodecyl sulfate (SDS), and 0.25% nonfat dry milk. Membranes were washed and exposed to x-ray film.

Amplification and Sequencing of the rpl16 Intron

The chloroplast gene *rpl16* encodes a ribosomal protein that is interrupted by an intron in all angiosperms thus far investigated (Posno et al., 1986; Downie et al., 1996; Jordan et al., 1996; Kelchner

TABLE 1. Primers used in PCR and sequencing. All primers are given from 5' to 3'.

<i>rp116</i>	F71	GCTATGCTTAGTGTGTGACTCGTTG
	F220	CTGATTATGAGTTGTGAAGC
	R1516	CCCTTCATTCTTCTCTATGTTG
	R1661	CGTACCCATATTTTTCCACCACGAC
ITS	ITS.LEU	GTCCACTGAACCTTATCATTAG
	ITS2	GCTGCGTTCTTCATCGATGC
	ITS3B	GCATCGATGAAGAACGTAGC
	ITS4	TCCTTCCGCTTATTGATATGC

and Wendel, 1996). Primers F71 and R1661 (Jordan et al., 1996; Table 1) located in the flanking exons were used to amplify a 1.3-kb region including the entire intron. PCR amplification used standard protocols. The products were ligated into the pGEM-T vector (Promega) and one clone of each was sequenced using the Sequenase V 2.0 kit (Amersham). Sequencing utilized primers located in the vector (M13 universal and reverse) plus two internal primers, F220 and R1516 (Table 1). Together these four primers permitted the entire sequence to be determined from at least one strand of DNA. Additionally, PCR products of 11 of the samples were cleaned and then sequenced with primer F71 on an ABI Prism Model 377 automated DNA sequencing system (Iowa State University DNA Sequencing and Synthesis Facility). The resulting 400–500 bp of sequences was compared with the cloned sequences, revealing several autapomorphic substitutions in the clones that were not evident in the directly sequenced PCR products. These discrepancies are interpreted as errors arising from misincorporation during PCR. The probability of creating an apparent synapomorphy by *Taq* error must be extremely small, and thus we do not think that undetected PCR error is likely to affect phylogenetic inference. The sequence for *Bombax* was obtained by directly manually sequencing the cleaned PCR product with the Thermosequenase 33P-ddNTP terminator kit (Amersham).

Amplification and Sequencing of ITS

The internal transcribed spacer region includes the two internal transcribed spacers (ITS1, ITS2) and the 5.8S gene of the 18S/26S nuclear ribosomal DNA (rDNA) repeats. This region was amplified using primers ITS.LEU (L. E. Urbatsch, unpublished data) and ITS4 (White et al., 1990) in the presence of 10% dimethyl sulfoxide (DMSO). The gel-purified PCR product was sequenced using the ABI dyeprime kit and run on an ABI model 370A automated DNA sequencing system. Sequencing used the PCR primers plus ITS2 (White et al., 1990) and ITS3B (B. G. Baldwin, personal communication) as internal primers (Table 1). The sequencing strategy allowed the entire region to be sequenced in both directions except for close to the priming sites.

Phylogenetic Analysis

Sequences were aligned manually. Potentially informative indels that were located in regions of unambiguous alignment were scored and added to the matrix as extra gap characters (see Baum et al., 1994). Sequences were submitted to Genbank (accession numbers AF028521–AF028556).

Phylogenetic analysis of the data was primarily conducted using the computer program PAUP* (test versions 4.0d55–61 provided by D. L. Swofford). Unless otherwise stated, all analyses used the branch-and-bound algorithm to ensure that optimal trees were found.

Most parsimonious trees were generated independently for the four data sets, followed by analysis of (1) the cpDNA data (i.e., restriction sites and *rpl16*); (2) the nuclear data (i.e., ITS and morphology); (3) the molecular data (i.e., *rpl16*, restriction sites, and ITS); and (4) the global combination of the four data sets. These baseline analyses were conducted under equal weighting of all regions of the sequence, including gap characters. In the cpDNA, molecular, and global analyses, taxa not scored for restriction sites (*Bombax*, *A. grandidieri* b, *A. suarezensis* b, *A. madagascariensis* c, *A. za* c) were included in the analysis but were assigned as missing for all restriction site data. In the nuclear and global analyses the morphological data lines for each taxon were duplicated for those species that were multiply sampled.

Bootstrap support (Felsenstein, 1985) for each clade was estimated based on 100 bootstrap replications with each replicate being subject to simple-addition-sequence, tree-bisection-reconnection, heuristic searches. The decay index (DI; Bremer, 1988) for each clade was determined via one of three methods: (1) branch-and-bound searches for non-minimal trees ($DI < 4$), (2) converse-constraint branch-and-bound searches ($DI = 4-9$), or (3) 200 converse-constraint random-addition-sequence searches ($DI > 9$).

We explored the effect of deviation from the baseline assumptions, for example, different gain:loss (restriction sites) or transition-transversion (sequence data) weighting, treating all morphological characters as unordered, deleting gap and/or gapped characters (i.e., those with parts of sequences with a gap in any taxon). Only those explorations that suggested sensitivity of particular clades are reported in the results.

To examine the extent of conflict between the nuclear and chloroplast data we conducted the incongruence length difference (ILD) test (Farris et al., 1994) as implemented in PAUP* 4.0d60 using

simple addition-sequence, tree-bisection-reconnection, heuristic searches. The partitions we considered were cpDNA versus ITS, and cpDNA versus nuclear data with 500 replicates analyzed in each case.

To localize sources of conflict between the data sets, clades that were supported by either the cpDNA or nuclear data sets but were rejected by the other data set were coded as constraints. The length of the most parsimonious trees satisfying a constraint was determined using a branch-and-bound search. A statistical evaluation of the significance of the extra steps required beyond the unconstrained trees was conducted using a Wilcoxon sign-rank test (Templeton, 1983; Larson, 1994; Mason-Gamer and Kellogg, 1996) implemented in SYSTAT ver. 5.2.1. To correct for multiple tests a sequential Bonferonni analysis was conducted as described by Rice (1989).

Molecular Clock Analysis

To elucidate the biogeographic history of *Adansonia* it is necessary to attach a temporal dimension to the phylogeny. There are good reasons to treat molecular clock estimations of branching time with caution (see Sanderson, 1998). Nevertheless, we felt that there was a chance that the window of possible branching times for the Australian species, *A. gibbosa*, could exclude the possibility that this geographic disjunction occurred at or before the separation of Australia from Africa/Madagascar.

The analyses were conducted with the ITS data set alone because this is the data set with the most phylogenetic information. To reduce noise and facilitate searches we excluded gapped characters and sites that were ambiguous in any of the taxa. In addition we reduced the data set to one sequence for each species (using the "a" accession in each case).

The strategy we employed entailed the use of likelihood-ratio tests (Felsenstein, 1981; Goldman, 1993; Yang et al., 1995; Huelsenbeck and Rannala, 1997). We

explored three models of molecular evolution in PAUP*. The simplest model (F81) assumes equal rates of change across sites and equal probabilities of transitions and transversions (Felsenstein, 1981). The HKY model assumes equal rates of changes across sites, empirical base frequencies, and estimates the transition/transversion bias from the data (Hasegawa et al., 1985). The HKY- Γ is based on the HKY model but allows for variation in rate across sites using a discrete approximation to a gamma distribution with four rate categories and with the shape parameter, α , estimated by maximum likelihood. To select among these models of molecular evolution a likelihood-ratio test was conducted (Yang et al., 1995; Huelsenbeck and Rannala, 1997).

Having obtained optimal trees and associated likelihoods for the various models in the absence of a molecular clock, we repeated the analysis adding a molecular clock. The trees found without a clock (all models supported the same three optimal trees) were rooted and used as starting trees for tree-bisection-reconnection, heuristic searches under the molecular clock constraint. Each analysis found a single tree that corresponded to a rooted version of one of the trees found without the clock. However, the actual tree favored in each case differed. As a result, the likelihood ratio between models under a clock assumption may not follow a χ^2 distribution (see Goldman, 1993).

To test the assumption of clocklike evolution, the likelihoods of the clock and nonclock versions of the three models were compared. Because the clock trees were rooted versions of the nonclock trees, it is probably valid to use a χ^2 test on the likelihood ratios. The number of degrees of freedom for such a test with a fully bifurcating tree is $N - 2$ (Yang et al., 1995; Sanderson, 1998).

As an independent assessment of clock-like behavior, we conducted the 1D and 2D tests of Tajima (1993) as implemented using the computer program,

Tajima93 (T. Seelanan, Iowa State Univ.). Using *Pseudobombax* as the outgroup, all pairwise comparisons within *Adansonia* were analyzed for rate heterogeneity.

If a clock applies, then, under the preferred model of evolution with a molecular clock enforced, the branch lengths of the optimal tree are proportional to time. However, the branch length estimates have some variance. The standard error of the branch lengths were therefore obtained using the computer program PAML (Yang, 1996). The optimal tree was input as a user tree under the HKY- Γ model and standard errors of the branch lengths were obtained based on the curvature of the likelihood surface (Sanderson, 1998) using the getSE command (Yang, 1996).

Some previous studies have estimated the substitution rate for ITS (Savard et al., 1993; Suh et al., 1993; Sang et al., 1994; Wendel et al., 1995b). To obtain an approximate dating of branching events in the ITS tree we used the highest (5.3×10^{-9} substitutions/site/million years; Wendel et al., 1995b) and lowest (4.5×10^{-10} substitutions/site/million years; Suh et al., 1993) reported rates. The same data set was used as for the maximum-likelihood analysis, but with the 18S, 5.8S, and 28S coding regions excluded. Pairwise distances were determined in PAUP* 4.0d60 using maximum-likelihood estimation under the HKY model with shape parameter α and the transition/transversion ratio set to the values associated with the maximum-likelihood tree under a clock assumption.

RESULTS

Data Characteristics

The data matrices are shown in Appendices 2–6. The region of *rpl16* that was sequenced had an aligned length of 1,347 bp. All except the last 165 bp (which lack parsimony-informative sites) fall within the intron. Individual sequences within *Adansonia* ranged in length from 1,270 bp (*A. gibbosa* and *A. madagascariensis* c) to 1,341 bp (*A. suarezensis* a and b), whereas

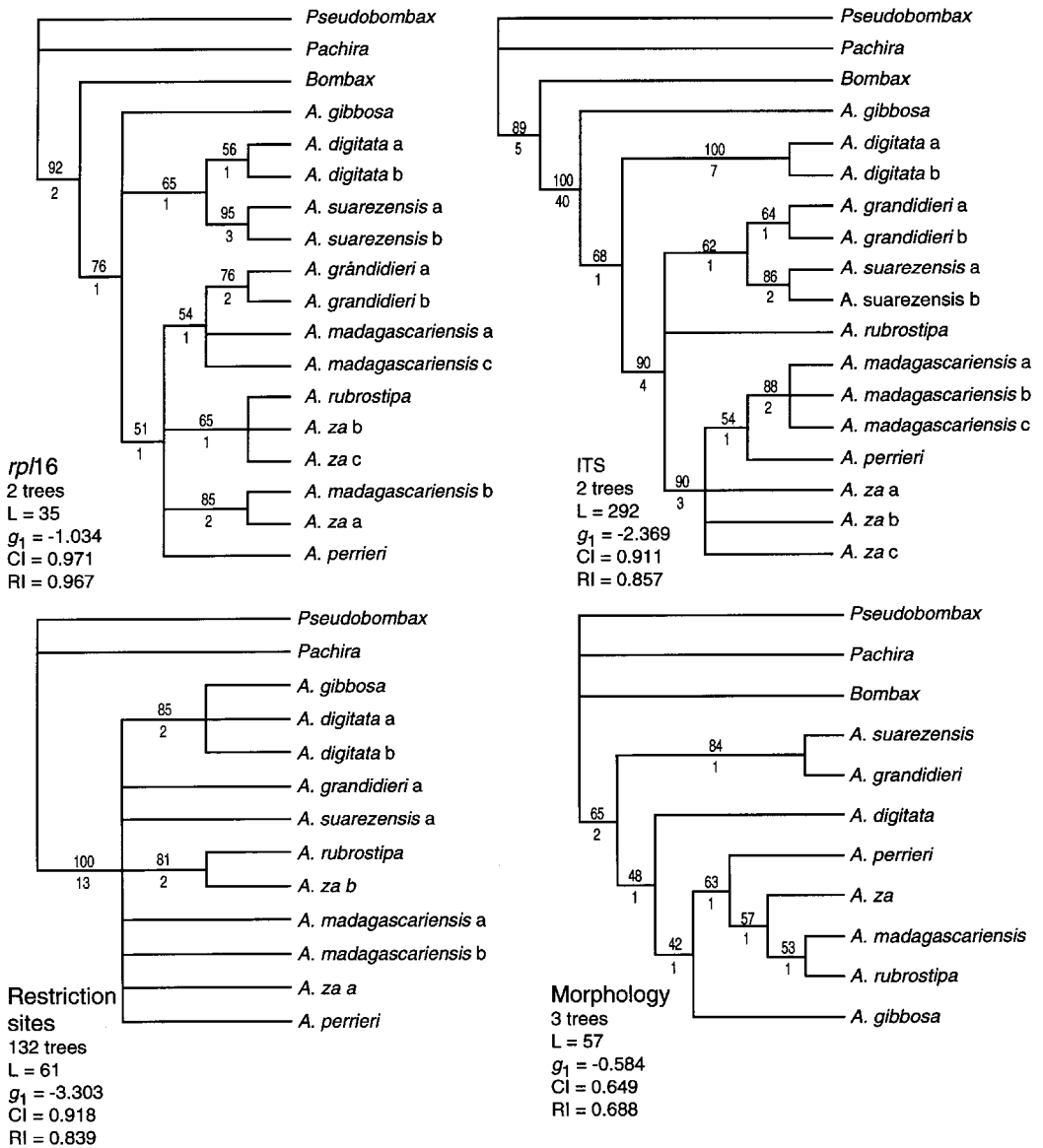


FIGURE 1. Strict consensus of the most parsimonious trees of *Adansonia* and outgroups for the four primary data sets. Bootstrap support is shown above branches; decay index is below. The number of equally parsimonious trees refers to the number of trees when branches whose maximum length is zero are collapsed.

outgroups sequences were shorter (down to 1,185 bp). Almost all this length variation is due to a single AT-rich hypervariable region. This region could not be aligned reliably due to repeated motifs and numerous indels and was, therefore,

excluded from all phylogenetic analyses. Outside the hypervariable region, few indels were required to achieve alignment and these were easy to determine. Four phylogenetically informative gap characters were scored.

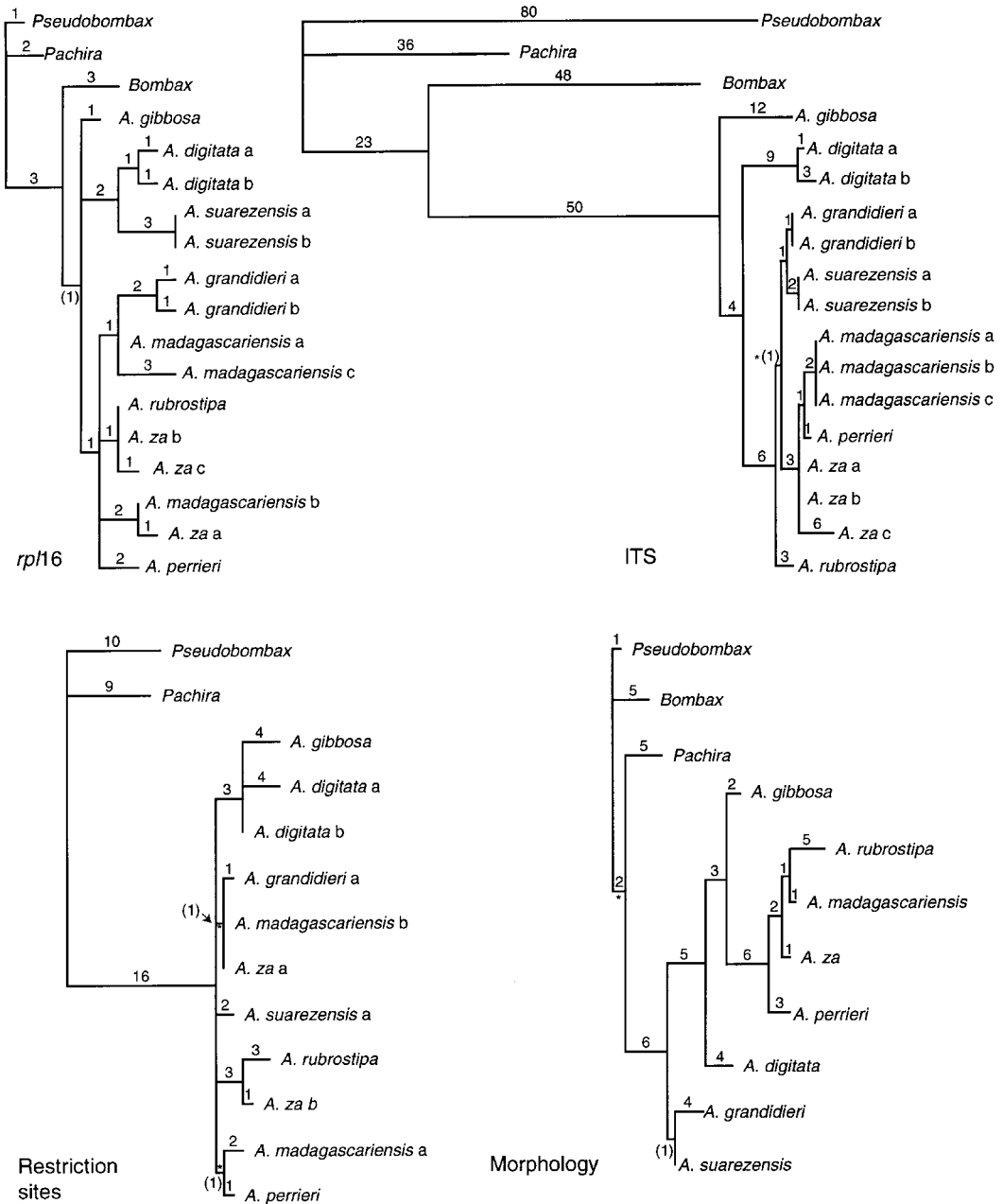


FIGURE 2. Phylograms of example trees for the four primary data sets. Branch lengths (ACCTRAN optimization) are given above branches. Branch lengths in parentheses have a minimum length of zero. Branches with asterisks are absent from the strict consensus of all equally parsimonious trees (Fig. 1).

rpl16 showed low levels of variation (up to 0.65% overall, 0.56% in the ingroup). The pairwise differences (ignoring the gaps), once the hyper-

variable region was excluded, ranged from four to eight sites when contrasting the ingroup and outgroup taxa, and zero to seven within the ingroup. Of the 28

variable positions, 11 were phylogenetically informative (9 when considering only the ingroup taxa).

The chloroplast DNA restriction-site analysis found 56 variable sites (23 within the ingroup). Of these, 24 were phylogenetically informative (9 within the ingroup).

The ITS region that was analyzed had an aligned length of 787 bp. Individual sequences within *Adansonia* ranged in length from 771 to 772 bp. The outgroups had shorter sequences: 748 bp in *Pseudobombax* and *Bombax*; 753 in *Pachira*. Within *Adansonia* alignment was unambiguous, whereas between the four genera there were regions that were more problematic. Nonetheless, the alignment was sufficiently clear that we did not exclude any regions from phylogenetic analysis. Eleven phylogenetically informative gap characters were scored, of which only one was potentially informative within *Adansonia*.

ITS was more variable than *rpl16* (up to 18.20% overall, 4.00% in the ingroup). Within the ingroup, sequences differed at 0–29 nucleotide positions. Ingroup sequences differed from the outgroup sequences at 90–129 nucleotide positions. Of the 224 variable positions (66 within the ingroup), 85 were phylogenetically informative (23 within the ingroup).

Thirty variable morphological characters were scored (Appendices 5–6). The characters fell into the following categories: seed/seedling morphology (4), fruit (3), leaf morphology (5), tree architecture (1), flower morphology/pigmentation (16), phenology (1). Of these 30 characters, 24 were phylogenetically informative (19 within the ingroup).

As with any morphological data, there is the possibility of extensive nonindependent evolution due to either developmental interdependence or selection. For example, there is reason to suspect that the following suites of characters could be developmentally correlated: (1,2), (6,18), (6,19,20), (16,17), (8,19), (12,13,14). Likewise, features that are likely to evolve in response to pollinator

identity, whether by direct selection or by selection on a developmental correlate (i.e., 6–20, 24, and 27), have the potential for nonindependence. However, no two characters in the matrix must necessarily evolve simultaneously and no simple weighting scheme can be justified at this time. Therefore, while acknowledging the potential problems of nonindependence and the theoretical desirability of coding a priori hypotheses of nonindependence into character types and weights, our phylogenetic analyses treated all characters as equally weighted.

Phylogenetic Analysis

The strict consensus of the equally most parsimonious trees obtained when the four data sets are analyzed independently is shown in Figure 1 with representative phylograms shown in Figure 2. Strict consensus trees obtained from the cpDNA (restriction site + *rpl16*), nuclear (ITS + morphology), molecular (*rpl16* + restriction site + ITS), and global analyses (restriction site + *rpl16* + ITS + morphology) are shown in Figure 3.

The four primary data sets differ substantially in their topology and degree of resolution within *Adansonia*. The morphology and ITS trees disagree in the order of branching at the base of *Adansonia* and additionally on the relationships among the Malagasy *Longitubae* (*A. rubrostipa* Jum. & H. Perrier, *A. za* Baill., *A. madagascariensis* Baill., and *A. perrieri* Capuron). Nonetheless, they both support the *Brevitubae* (*A. grandidieri* and *A. suarezensis*) and are both consistent with the Malagasy *Longitubae* forming a clade. However, the *rpl16* phylogeny contradicts both of these results. In addition, most parsimonious trees for *rpl16* and restriction sites imply a lack exclusivity of some named species within *Adansonia*. However, the bootstrap support and DI of conflicting clades were generally low.

Sensitivity Analyses

One indirect means of evaluating the robustness of the most parsimonious tree

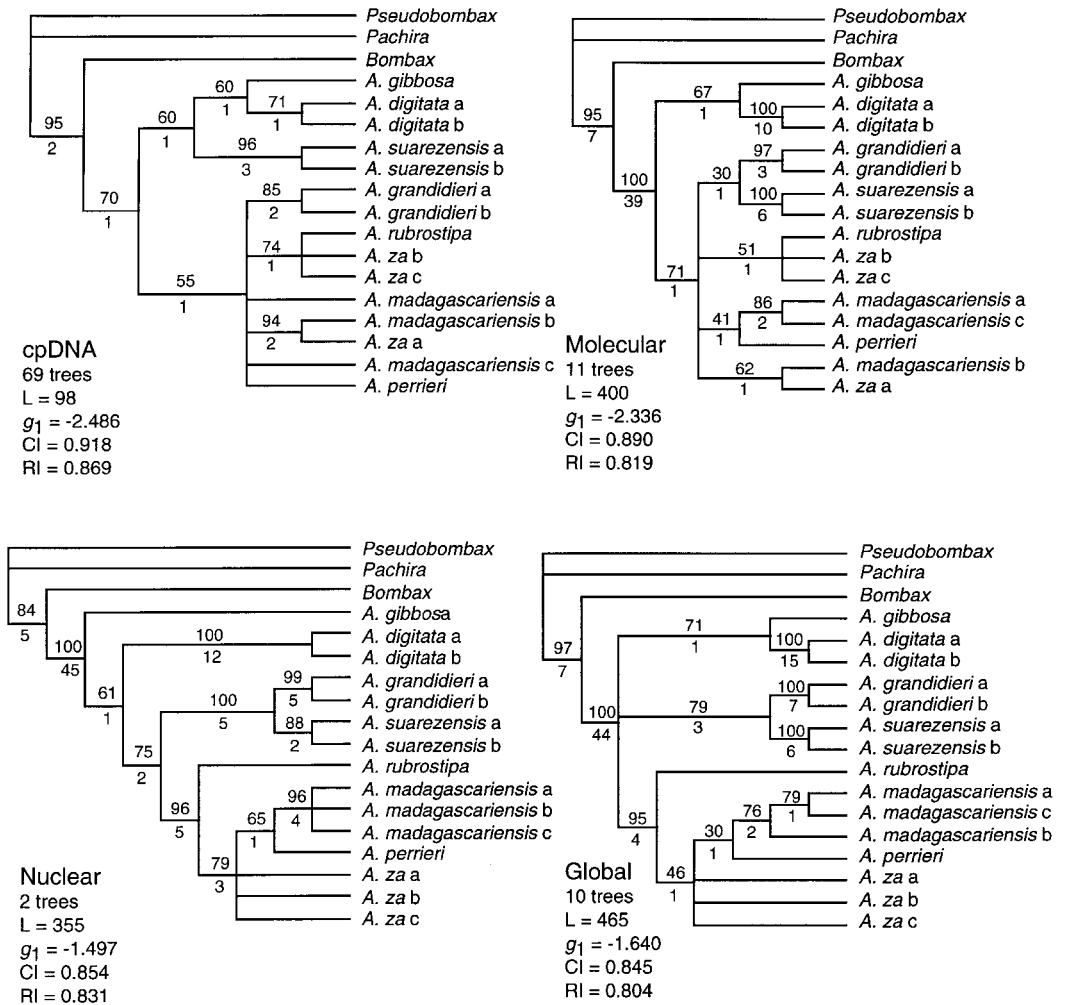


FIGURE 3. Strict consensus of the most parsimonious trees of *Adansonia* and outgroups for the four combined data sets. Bootstrap support is shown above branches; decay index is below. The number of equally parsimonious trees refers to the number of trees when branches whose maximum length is zero are collapsed.

associated with a given data set is to ask whether it changes when we alter our assumptions about character evolution, and when we subtract characters or taxa. Of the eight data sets analyzed, only phylogenies derived from the ITS and nuclear data were robust to all perturbations explored. Some specific results of the sensitivity analysis with the other data sets are worthy of comment.

The exclusion of gap characters from

the *rpl16* data resulted in the loss of two clades within *Adansonia* and the addition of two different clades. Examination of the data showed that this change is due to one multistate gap character that unites *A. suarezensis* and *A. digitata*. Given that this site has been subject to at least three insertion/deletion events, and that this clade gains no support from other data, we suspect that this gap character is misleading.

The morphological data cease to place *A. gibbosa* as sister to the Malagasy *Longitubae* when filament number is treated as unordered. This arises because *A. gibbosa* has filament number intermediate between the Malagasy *Longitubae* and *Brevitubae* and *A. digitata*. While this character can reasonably be scored as ordered, the breaking up of the continuous character, filament number, into discrete states was not without its problems (see Stevens, 1991; Gift and Stevens, 1997). Specifically, there is overlap in the range of filament numbers occurring within *A. gibbosa* and *A. perrieri*. We explored other scorings of this character and found that the position of *A. gibbosa* was not consistently resolved. Therefore, the morphological data cannot be said to support the relationship of *A. gibbosa* as sister to the Malagasy *Longitubae*, as also indicated by low bootstrap (42%) and DI (1).

The restriction site data has only two supported clades within *Adansonia* under the baseline conditions, and neither of these clades is robust to adopting a 3:1 gain:loss bias. One of these clades (*A. rubrostipa* + *A. za* b/c) is a point of conflict with the nuclear data that also appears on the *rpl16* tree. Use of a 3:1 gain:loss bias also removes this clade from some most-parsimonious cpDNA trees.

One of the two clades appearing when a 3:1 gain:loss bias was applied to the restriction-site data was also present in the *rpl16* tree (*A. za* a + *A. madagascariensis* b). This clade's bootstrap proportion goes up from 85% in the *rpl16* tree to 94% in the cpDNA tree (DI did not change), suggesting that there is some shared signal for this result in the *rpl16* and restriction-site data sets.

Statistical Analysis of Conflict

The ILD test using either the cpDNA versus ITS or cpDNA versus nuclear partitions returned a significant result ($P = 0.002$). If a 3:1 gain:loss weight is applied to the restriction site data, the conflict between the ITS and cpDNA par-

tion is less marked, but still significant ($P = 0.044$).

The chloroplast data suggested three patterns that are in conflict with the nuclear data sets: (1) the inclusion of *A. grandidieri* within the Malagasy *Longitubae* clade, (2) the *A. rubrostipa* + *A. za* b/c clade, and (3) the *A. madagascariensis* b + *A. za* a clade. Similarly, the nuclear data support four patterns that disagree with the chloroplast data sets: (4) the *A. grandidieri* + *A. suarezensis* clade, (5) the Malagasy *Longitubae* clade, (6) the monophyly of *A. madagascariensis* accessions, and (7) the *A. za* + *A. madagascariensis* + *A. perrieri* clade. The cost of forcing these constraints on the ITS, morphology, nuclear, and cpDNA data was determined (Table 2).

The only cases in which a single clade is rejected at the $P < 0.05$ level involve the rejection of constraint 1 (monophyly of Malagasy *Longitubae* plus *A. grandidieri*) by the morphology and nuclear data, and rejection of constraint 3 (monophyly of *A. madagascariensis* b plus *A. za* a) by the nuclear data (Table 2). However, if one corrects these results for the fact that multiple tests were conducted, these three Wilcoxon sign-rank tests cease to be judged significant (Table 2).

There are several cases where significant conflict is observed when multiple constraints are simultaneously imposed on a data set. For example, ITS can reject constraints 2 + 3 and 1 + 2 + 3, whereas cpDNA conflicts with constraints 6 + 7 and 4 + 5 + 6 + 7. However, the significance of these results disappears after Bonferonni correction (Table 2). Indeed, the only conflict to be judged significant after correction for multiple tests is the rejection of constraints 1 + 2 + 3 by the morphological and nuclear data.

Applying a 3:1 gain:loss bias to the restriction site data results in one conflicting clade (*A. rubrostipa* + *A. za* b/c) disappearing from the cpDNA tree, and a second (*A. za* a + *A. madagascariensis* b) being strengthened (see earlier discussion). To see the effect of this

TABLE 2. Wilcoxon sign-rank tests of the cost of imposing constraints suggested by another data partition. Cost is the length of the most-parsimonious tree under the constraint minus the length of the unconstrained analyses. Gain and Loss refer to the number of characters having extra steps on the constrained or unconstrained tree, respectively. N = number of ranks. The test statistic (Ts) is the sum of the ranks of the losses for which two-tailed P values are given based on a normal approximation.

Data set	Constraint ^a	Cost	Gain	Loss	N	Ts	$P(2\text{-tailed})^b$
ITS	1	1	1	0	1	0	0.31731
	2	3	3	0	2	0	0.08326
	3	3	3	0	2	0	0.08326
	2 + 3	6	6	0	6	0	0.01431
	1 + 2 + 3	7	7	0	7	0	0.00815
Morphology	1	8	9	1	10	5.5	0.01141
	2	2	2	0	2	0	0.31731
	3	3	3	0	3	0	0.08326
	2 + 3	4	4	0	4	0	0.04550
	1 + 2 + 3	12	13	1	14	7.5	0.00134*
Nuclear	1	5	5	0	5	0	0.02535
	2	5	7	2	9	10	0.09558
	3	6	6	0	6	0	0.01431
	2 + 3	10	12	2	13	14	0.01242
	1 + 2 + 3	15	17	2	18	18	0.00106*
cpDNA	4	2	3	1	4	2.5	0.31731
	5	1	2	1	3	2	0.56370
	6	3	4	1	5	3	0.17971
	7	1	2	1	3	2	0.56370
	6 + 7	7	8	1	9	5	0.01963
	4 + 5 + 6 + 7	9	10	1	10	5	0.01255

^a 1 = ((GRA, RUB, MAD, ZA, PER)); 2 = ((RUB, Zab, ZAc)); 3 = ((MADb, ZAa)); 4 = ((GRA, SUA)); 5 = ((RUB, MAD, ZA, PER)); 6 = ((MADa, MADb, MADc)); 7 = ((MAD, ZA, PER)). GRA = *A. grandidieri*, SUA = *A. suarezensis*, DIG = *A. digitata*, GIB = *A. gibbosa*, RUB = *A. rubrostipa*, ZA = *A. za*, MAD = *A. madagascariensis*, PER = *A. perrieri*. Multiple accessions from the same species were in the same clade unless indicated otherwise.

^b Significance shown as * $P < 0.05$ level after sequential Bonferroni correction.

weighting scheme on data set conflict we imposed constraint 4 + 5 + 6 + 7 on the cpDNA data under a 3:1 gain:loss bias. The strength of rejection of the constraint was somewhat higher than the unweighted test (Cost = 9, $N = 8$, Ts = 0, $P = 0.0066$), but is still nonsignificant after Bonferroni correction.

TABLE 3. Comparison of likelihood scores for different models of evolution. In each comparison, model 1 is the simpler model and represents a special case of model 2. The likelihood-ratio test is based on twice the difference between the log-likelihoods for the two models ($-2 \log L1/L2$). The P values are based on a chi-square distribution with either 1 or 9 degrees of freedom: ** $P < 0.005$. For other values $P > 0.05$.

Model 1	Model 2	$-\log L1$	$-\log L2$	$-\log L1/L2$
Models of evolution (df = 1)				
F81 (nc)	HKY (nc)	1957.07	1932.59	48.96**
HKY (nc)	HKY- Γ (nc)	1932.59	1910.28	44.62**
F81 (c)	HKY (c)	1963.78	1939.11	49.34**
HKY (c)	HKY- Γ (c)	1939.11	1915.13	47.96**
Clock vs. nonclock (df = 9)				
F81	F81	1963.78	1957.07	13.42
HKY	HKY	1939.11	1932.59	13.04
HKY- Γ	HKY- Γ	1915.13	1910.28	9.70

Molecular Clock Analyses

The ITS data set (minus sites with gaps or ambiguous states) for one exemplar per species was analyzed using maximum likelihood. For all the nonclock models the same three unrooted trees were optimal. These corresponded to a pruned version of the parsimony ITS tree (Fig. 1). All of the clock models favored a single tree and in each case it was a rooted version of one of the three unrooted trees. The F81 and HKY models favored a tree with *Bombax* sister to *Adansonia* whereas the HKY- Γ favored a tree with *Pachira* sister to *Adansonia*.

As summarized in Table 3, the more complex models had significantly higher likelihoods whether the comparisons were made with or without molecular clock assumptions. Therefore, the ITS sequences are most compatible with an HKY- Γ model of evolution. Figure 4 shows the optimal rooted tree as estimated by the HKY- Γ model with a molecular clock assumed.

Comparing the clock and nonclock versions of all three models, a molecular clock could not be rejected (Table 3). Similarly, application of either the 1D or 2D tests of Tajima (1993) failed to detect significant rate heterogeneity. Therefore, the branch lengths (and associated errors) for the HKY- Γ model under a molecular clock can be taken as being approximately proportional to time.

TABLE 4. Estimates of timing of events a, b, c, and d (Fig. 4) under different calibrations of the ITS data. Row 1 gives the branch length in likelihood units \pm two standard errors. In rows 2–5 those branch lengths are converted into times by fixing one branching event. Row 6 gives the mean pairwise differences (after HKY- Γ correction for multiple hits) for each node \pm two standard deviations. In rows 7–8 those distances are converted into time estimates assuming a rate of 5.3×10^{-9} (row 7) or 4.5×10^{-10} (row 8) substitutions/site/million years.

	a	b	c	d
1. Likelihood	8547 \pm 2050	1599 \pm 586	1442 \pm 578	582 \pm 0
2. Time (mbp)	535 \pm 128	100 ^a	90 \pm 36	36 \pm 0
3. Time (mbp)	90 ^a	17 \pm 6	15 \pm 6	6 \pm 0
4. Time (mbp)	58 ^a	11 \pm 4	10 \pm 4	4 \pm 0
5. Time (mbp)	36 ^a	7 \pm 2	6 \pm 2	2 \pm 0
6. Distance	0.22 \pm 0.00	0.04 \pm 0.01	0.04 \pm 0.01	0.01 \pm 0.01
7. Time (mbp)	21.1 \pm 0.08	3.81 \pm 0.16	3.57 \pm 0.63	1.24 \pm 0.12
8. Time (mbp)	248.5 \pm 9.2	44.9 \pm 19.3	42.1 \pm 14.8	14.6 \pm 14.0

^a Fixed branching event.

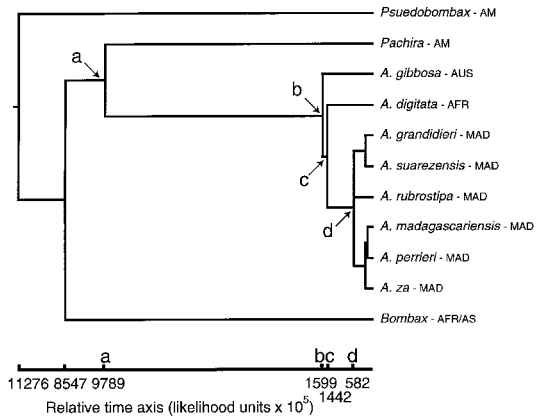


FIGURE 4. Phylogram of the maximum likelihood ITS tree (HKY- Γ model) for *Adansonia* and outgroups under the assumption of a molecular clock. Branch lengths as estimated by PAML. Four key branching events are labeled a–d. These and other branching events are dated on the relative time axis. The geographic distribution of the terminal taxa are given: AM = Americas, AUS = Australia, AFR = Africa, MAD = Madagascar, AS = Asia.

From the point of view of testing the hypothesis of *A. gibbosa* having arrived in Australia via Gondwanan disjunction, the key issue is the date of the origin of the stem lineage of *Adansonia* (node a, Fig. 4) relative to the divergence of the Australian species (node b, Fig. 4). Additionally, it is interesting to determine when *Adansonia* might have radiated into a Malagasy and African lineage (nodes c and d, Fig. 4). Table 4 converts these

branch lengths into branching times (with confidence intervals) using several calibration points (see below).

Table 4 provides the pairwise distances between species across nodes a–d after correcting for multiple substitutions with the HKY- Γ model. Within *Adansonia* the estimated distances did not vary significantly when different models were used to correct for multiple hits. However, the distances associated with node a were sensitive to the choice of model. For example, use of the Kimura two-parameter model gave a distance of 0.176 substitutions/site, as compared to 0.224 substitutions/site for the HKY- Γ model. However, this effect is small relative to the more than 10-fold difference between the high and low substitution rates. However, the sensitivity to evolutionary model suggests special caution when interpreting the estimated ages of node a.

DISCUSSION

Data Set Conflict and Genealogical Discordance

In recent years there has been much debate over how to proceed when multiple data sets for the same taxa give conflicting estimates of phylogeny (reviewed by de Queiroz et al., 1995; Huelsenbeck et al., 1996). When different data sets have tracked the same underlying genealogy, the best estimate of that phylogeny is usually derived from a combined analysis of all the data (Barrett et al., 1991). However, if data sets have tracked different histories—that is, they show genealogical discordance (Avice and Ball, 1990), for example due to lineage sorting or introgression (see review by Wendel and Doyle, 1998)—a combined analysis can result in a tree that does not represent any one genealogy but a composite picture of several. Thus, when genealogical discordance applies, comparison of individual gene trees may provide a clearer picture of organismic history (Wendel and Doyle, 1998).

How are we to determine if different data sets have indeed tracked different

histories? An ever-broadening array of statistical tests has been proposed to ask whether the observed disagreements could have occurred by chance if the partitions were drawn from the same genealogy and evolutionary process (reviewed by de Queiroz et al., 1995; Wendel and Doyle, 1998). However, significant conflict can arise for reasons other than genealogical discordance, for example, because a data set has been affected by convergence of two lineages for a suite of characters or contains long branches that attract each other (Wendel and Doyle, 1998). Thus, despite the availability of statistical tests, determination of whether phenomena such as introgression and lineage sorting have taken place depends on careful, somewhat subjective, evaluation of the available evidence.

For these data, disagreement is most obvious in a comparison of the cpDNA and nuclear trees (Fig. 2). These partitions return a significant ILD, and each statistically rejects a combinations of clades supported by the other tree (Table 2). Thus, the disagreements seen in the most parsimonious trees do not seem to arise simply from sampling error. To aid in evaluating whether this conflict is due to genealogical discordance it is useful to try and determine if it is localized on the tree, and if so where.

All the data sets except *rpl16* and cpDNA are consistent with there being four main branches at the base of *Adansonia*: *A. gibbosa*; *A. digitata*; the *Brevitubae* (*A. grandidieri* and *A. suarezensis*), and the Malagasy *Longitubae* (*A. rubrostipa*, *A. za*, *A. madagascariensis*, and *A. perrieri*). The *rpl16* and cpDNA data differ in that they contradict *Brevitubae* being a clade. However, these data sets cannot statistically reject this clade even before Bonferonni correction (Table 2) and, furthermore, the relevant nodes have low bootstraps and DIs (Figs. 1, 3). Therefore, we believe that the data as a whole support the monophyly (where relevant) and basal position of these four clades, with the *rpl16* and cpDNA results being due to sampling error.

The relationships among the four basal branches vary between data sets and are generally weak. The strongest result is the sister-group relationship between the *Brevitubae* and Malagasy *Longitubae*, which appears on the ITS (bootstrap = 90% ; DI = 4), nuclear (bootstrap = 73% ; DI = 2), and molecular trees (bootstrap = 71% ; DI = 1). However, while this result may be strong compared to other implied relationships among the basal branches (bootstrap = 42–85%, DI = 1–2) it must be treated with suspicion given its absence from the global analysis. Hence, the extensive disagreement between data sets over the arrangement at the base of *Adansonia* is probably due to sampling error rather than genealogical discordance or the presence of non-phylogenetic signal.

In contrast to the lack of significant conflict regarding basal relationship, there is evidence of such conflict within the Malagasy *Longitubae*. ITS and nuclear data support two clades that contradict the cpDNA tree: the clade of *A. madagascariensis* accessions (bootstrap = 88–96% ; DI = 2–4), and *A. madagascariensis* + *A. za* + *A. perrieri* (bootstrap = 79–90% ; DI = 3). The clades supported by cpDNA that contradict these two clades were also strongly supported: *A. madagascariensis* b + *A. za* a (bootstrap 85–94%, DI = 2), and *A. rubrostipa* + *A. za* b/c (bootstrap = 74–81%, DI = 1–2). Forcing both clades simultaneously onto the opposing data set caused a marked increase in tree length that would be judged statistically significant were it not for the need to conduct Bonferroni correction for multiple tests (Table 2).

Although the conflict observed within the Malagasy *Longitubae* cannot easily be ascribed to sampling error, it is an open question whether this conflict is due to genealogical discordance. It is noteworthy that *A. madagascariensis* b and *A. za* a are accessions from the same area of northwest Madagascar. Similarly, *A. rubrostipa*, *A. za* b, and *A. za* c are all from southern Madagascar. Thus, insofar as one wished to explain the discordance by

introgression, the geography is consistent with that scenario. Introgression of the chloroplast genome, resulting in conflict with nuclear data, has been reported in other groups (e.g., Rieseberg and Soltis, 1991; Whittemore and Schaal, 1991; Rieseberg and Wendel, 1993; Wolfe and Elisens, 1995). However, it should be noted that there is no evidence of hybridization in extant baobabs, and that sympatric species of the Malagasy *Longitubae* have nonoverlapping flowering seasons (Baum, 1995b). Thus, even if one were to assume that genealogical discordance applied, it would remain unclear whether it was due to introgression, lineage sorting, or both.

The Best Estimate of the Population Phylogeny

For the remainder of this article we focus on the nuclear tree (Fig. 3) for the purposes of studying character evolution and biogeography. We select this tree because it appears to provide a robust phylogeny (as judged by bootstraps, DIs, and sensitivity analysis). Nonetheless, it is important to note certain potential problems with ITS, the dominant information within the nuclear data set. In polyploids and many diploids, ITS occurs in multiple unlinked loci, each comprising numerous tandem repeats. Although this has not been specifically determined for *Adansonia*, it has been shown for many plant species including cultivated cotton (Crane et al., 1993), a member of the closely related Malvaceae. As a result of this genetic architecture, a lineage can maintain distinct copy types through cladogenetic events (e.g., Suh et al., 1993; Buckler et al., 1997) and then undergo concerted evolution to become fixed for one or another copy type (e.g., Wendel et al., 1995a). This can result in an ITS tree that is robust but is misleading with respect to populational history (Wendel et al., 1995a, 1995b; Buckler et al., 1997). These problems might be implicated in *Adansonia* because of the high base chromosome number ($2n = 88$; Baum and

Oginuma, 1994) and the occurrence of one tetraploid species, *A. digitata* ($2n = 160$; Baum and Oginuma, 1994). Thus, it is conceivable that ITS could have failed to track the same history as the rest of the nuclear genome. Nonetheless, while bearing these caveats in mind, we consider the nuclear data to provide a reasonable working hypothesis of organismic relationships at this time. It should be noted that the possibility of discrepancies within the Malagasy *Longitubae* does not affect the inferences of biogeographic history or floral evolution, described later, because all members of this clade are endemic to Madagascar and possess elongated, red or yellow hawkmoth-pollinated flowers.

Taxonomic Implications

Adansonia grandidieri, *A. digitata*, and *A. suarezensis* were each multiply sampled. In each case the accessions formed clades on both the ITS trees (bootstraps from 59% for *A. grandidieri* to 100% for *A. digitata*) and the cpDNA trees (bootstraps from 71% in *A. digitata* to 96% in *A. suarezensis*). Thus, these data support exclusivity (*sensu de Queiroz and Donoghue, 1990; Baum and Shaw, 1995*) for these three species.

Adansonia madagascariensis accessions formed a clade according to ITS (bootstrap = 89%) but were nonmonophyletic on the cpDNA and molecular trees. A clade of *A. za* accessions was not supported by any of the trees and was positively contradicted by cpDNA data. Therefore, despite the few informative characters and the resulting lack of statistical power, these data call into doubt the status of *A. za* and *A. madagascariensis* as genealogical species (Baum, 1995a).

The phylogenetic results are compatible with the subgeneric classification (Hochreutiner, 1908; Baum, 1995a) in that sect. *Brevitubae* is monophyletic and *A. digitata* (sect. *Adansonia*) is isolated. However, the monophyly of sect. *Longitubae* (type species = *A. madagascariensis*)

is compromised because the Australian *A. gibbosa* does not form a clade with the Malagasy *Longitubae*. Therefore, if these data are corroborated, some taxonomic realignment at the sectional level may be needed.

Insofar as the choice of taxa constitutes a test of the monophyly of *Adansonia*, the results uniformly and strongly support it with a DI as high as 45.

Biogeography

Given that Madagascar is the current center of diversity for the genus, it has generally been assumed that this geographic region was its center of origin (Aubréville, 1975; Armstrong, 1983). While not yet definitive, the suggestion of monophyly of the Malagasy species renders this inference no more parsimonious than a center of origin in either Australia or continental Africa.

Irrespective of the direction of dispersal, a key question is whether these dispersal events were overland before the fragmentation of Gondwana or entailed later, transoceanic dispersal. One approach to answering this question is to estimate branching times for the radiation of *Adansonia* relative to branching events deeper in the tree. For Gondwanan fragmentation to be invoked, the Australian lineage would have to have diverged by more than 100 million years before present (MaBP). In fact this is a conservative estimate given that land connections between Australia and Africa/Madagascar were probably broken by 120 MaBP (Rabinowitz et al., 1983; Wilford and Brown, 1994). Fixing the deepest divergence within *Adansonia* to 100 MaBP, the best estimate of the time of divergence between *Adansonia* and its nearest relatives is 535 MaBP (Table 4). If we assume an error of two standard errors around the branch-length estimate, the latest branching time for the divergence between *Adansonia* and *Pachira* is 407 MaBP. This is unrealistically early (before the inferred origin of seed plants) and therefore shows that the

distribution of *Adansonia* cannot have been due to Gondwanan disjunction but must have involved transoceanic dispersal. Such long-distance dispersal is plausible given that the tough fruit of many species are, at least partly, water dispersed (Wickens, 1983; Baum, 1995a).

Given the apparent clocklike evolution of ITS, it ought to be possible to date dispersal events within *Adansonia* once the molecular clock has been calibrated (Table 4). The temporal and geographic distribution of fossil pollen is a potential source of such calibration. The earliest pollen assigned definitively to Bombacaceae is *Bombacacidites* Couper from the Lower Maestrichtian (Upper Cretaceous) of North America (Wolfe, 1975, 1976). However, this pollen type does not place a lower bound on diversification of Adansonieae because one cannot rule out this pollen being on the lineage leading to *Bombax* (Fuchs, 1967; Wolfe, 1975; Nilsson and Robyns, 1986). Instead, the earliest possible date for the diversification of Adansonieae genera is defined by the Lower Campanian pollen, which Wolfe (1975, 1976) interpreted as a progenitor of *Bombacacidites*. If this fossil is indeed on the lineage leading to *Bombacacidites*, then its lack of ornamentation and small size would place it before (perhaps much before) any of the extant genera had begun to diversify (Wolfe, 1975). Thus, the earliest that one could date the origin of the stem lineage of *Adansonia* is the Lower Campanian, ca. 90 MaBP. Using this as a calibration point, the earliest date of the deepest split within *Adansonia* is 17 ± 6 MaBP (Table 4). At this time, Australia had nearly attained its current position, encompassing an extensive tropical zone (Quilty, 1994) with suitable habitats for baobabs.

While 90 MaBP is the earliest possible date for node a (Fig. 4), circumstantial evidence points to a later date. *Bombacacidites* pollen and other Bombacaceous pollens appear in the Caribbean by the Paleocene, in Africa by the Paleocene–Eocene border

(Germeraad et al., 1968), and in Australia by the early Eocene (Stover and Partridge, 1973). By the end of the Eocene, other Bombacaceae pollen types (probably Adansonieae) are reported (Wolfe, 1976). Thus, it is most likely that the main diversification of Adansonieae in Africa occurred during the Eocene, 36–58 MaBP (Kruttsch, 1989). This suggests that the deep splits within *Adansonia* occurred 2–15 MaBP (Table 4).

The great range in reported substitution rates for ITS means that the dating events can only be estimated approximately. However, even under the slower calibration, node b is dated well after the breakup of Gondwana (Table 4).

Based on comparison with the paleontological dates, it seems that the rate of ITS evolution within *Adansonia* falls between the two rates used. *Adansonia* is closely related to *Gossypium*, the basis for the fast rate, but Malvaceae (including *Gossypium*) has a higher rate of *ndhF* evolution than does the Adansonieae (W. Alverson, B. Whitlock, R. Nyffeler, and D. Baum, unpublished data). This may be explained based on the longer generation time of baobab trees, with trees living over 1,000 years (Swart, 1963), as compared to the herbaceous or shrubby Malvaceae. If there is a generation time effect on the rate of molecular evolution in plants (e.g., Gaut et al., 1992), the best estimates of the branching events are probably somewhat earlier than inferred based on the *Gossypium* calibration, and therefore accord well with the paleontological data.

Floral Evolution

Figure 5 summarizes the phylogeny of *Adansonia* that we will use for the purposes of discussion. At the time that this work was initiated we hoped that a phylogeny of *Adansonia* would lead to a clarification of the group's floral evolution. In particular, it seemed possible that we could determine whether mammal pollination (*Brevitubae* and *A. digitata*) or hawkmoth pollination (Malagasy *Longi-*

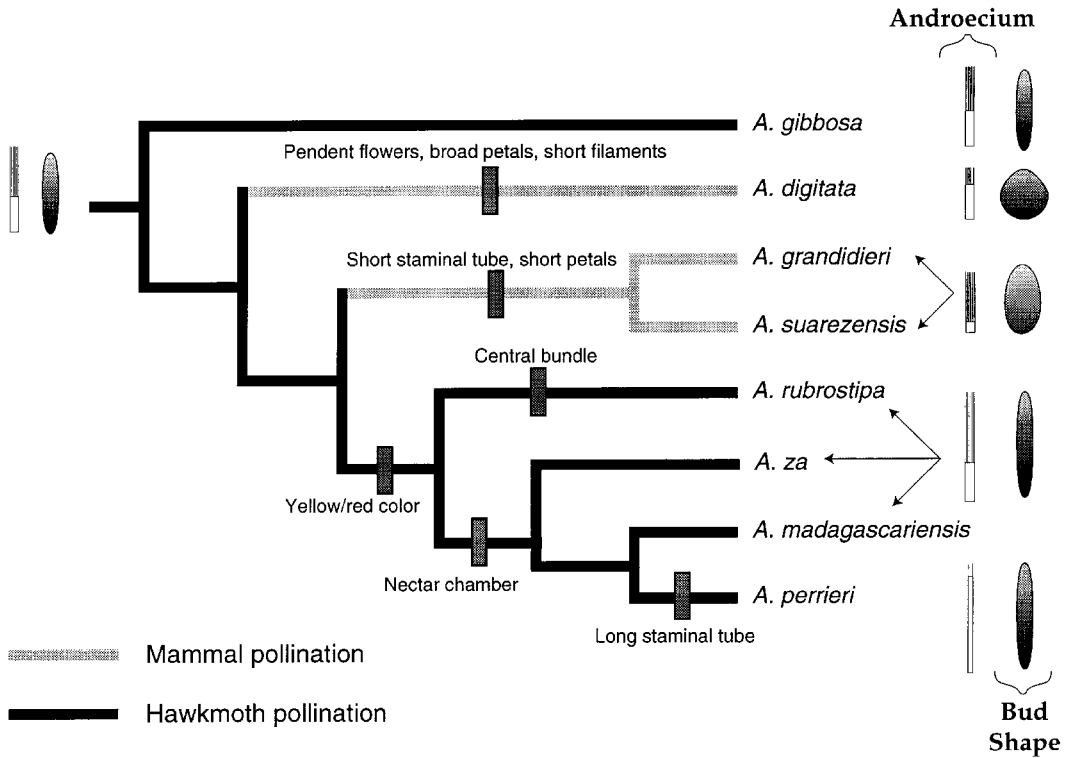


FIGURE 5. Hypothesis for pollination and floral evolution in *Adansonia* under the assumption that hawkmoth pollination is plesiomorphic. The relative length of the staminal column (white) and filaments (gray) and the bud shape are shown schematically for each extant species. The ancestral state for these characters under parsimony is shown at the root. Some floral characters are mapped onto this phylogeny based on parsimony.

tubae and *A. gibbosa*) is plesiomorphic. Unfortunately, the topology obtained is such that simple reconstruction of pollination system is equivocal.

Outgroup comparison has the potential to resolve the polarity of pollination systems. *Bombax* is pollinated by birds (Werth, 1915) or bats (e.g., Kaisila, 1966). *Pachira* (sensu Alverson, 1994) and *Pseudobombax* include species that have been reported to be pollinated by hawkmoths (Haber and Frankie, 1989), bats (Heithaus et al., 1974), and nonflying mammals (Gribel, 1988). Consideration of other Adansonieae genera complicates the matter still further as they include the aforementioned pollinators plus bees (Oliveira et al., 1992) and butterflies (Gibbs et al., 1988). Thus, without a phylogeny for Adansonieae the existing pol-

lination data are too variable to permit polarity determination within *Adansonia*.

While acknowledging the uncertainty in the phylogenetic inference, we think it is valuable to present one plausible hypothesis for floral evolution in *Adansonia*. Figure 4 shows, schematically, the androecial morphology and flower buds of the extant species. It is noteworthy that *A. gibbosa*, *A. rubrostipa*, *A. za*, and *A. madagascariensis* have similar flower buds and have androecia with similar length:width ratios and similar degrees of staminal fusion. Given the nuclear tree and the position of these taxa on the tree, we hypothesize that they manifest the plesiomorphic states for these characters (Fig. 4). This scenario suggests that the ancestral baobab had an elongated flower, which, in turn, leads

us to suspect that hawkmoth pollination rather than mammal pollination was ancestral.

Although *A. digitata*, *A. grandidieri*, and *A. suarezensis* are pollinated by a similar suite of floral visitors, dominated by fruit bats, they differ greatly in their floral morphology (Baum, 1995b). Floral organ development is distinct (Baum, unpublished data), as is the means by which flowers are presented to bats (pendulous below the crown in *A. digitata*, erect on the top of the crown in *Breuitubae*) and the means by which nectar is held in the flower. In view of the different means by which these two lineages have come to fit bat pollination, we infer that they independently evolved mammal pollination from hawkmoth pollination (Fig. 4). This scenario is, however, built upon numerous assumptions (e.g., the nuclear topology is correct, the ancestral flower was elongated, the ancestor was hawkmoth pollinated, etc.), any one of which could be flawed. Nonetheless, we offer this hypothesis in the hope that it may fuel further phylogenetic and ecological research aimed at explaining the evolutionary forces that have led to the baobabs' striking floral diversity.

ACKNOWLEDGMENTS

This work was supported by grants from the National Science Foundation to D.A.B. and J.F.W. and an Iowa State University Professional Advancement grant to R.L.S. We are grateful to the colleagues who provided material (see Appendix 1) and to Jason Hill for technical assistance. David Cannatella, Richard Olmstead, William Alverson, and two anonymous reviewers provided useful comments. We are grateful to John Huelsenbeck and Ziheng Yang for advice on likelihood analyses and David Swofford for making test versions of PAUP* available.

REFERENCES

ALVERSON, W. S. 1994. New species and combinations of *Catostemma* and *Pachira* (Bombacaceae) from the Venezuelan Guayana. *Novon* 4:3-8.

ARMSTRONG, P. 1983. The disjunct distribution of the genus *Adansonia*. *Natl. Geog. J. India* 29:142-163.

AUBRÉVILLE, A. 1975. Essais de géophylétique de Bombacacées. *Adansonia*, ser. 2 15:57-64.

AVISE, J. C., AND R. M. BALL, JR. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxf. Surv. Evol. Biol.* 7:45-67.

BAKHUIZEN VAN DER BRINK, R. C. 1924. Revisio Bombacacearum. *Bull. Jard. Bot. Buitenzorg Ser. 3* 6:161-240.

BARRETT, M., M. J. DONOGHUE, AND E. SOBER. 1991. Against consensus. *Syst. Zool.* 40:486-493.

BAUM, D. A. 1995a. A systematic revision of *Adansonia* (Bombacaceae). *Ann. Mo. Bot. Gard.* 82:440-470.

BAUM, D. A. 1995b. The comparative pollination and floral biology of baobabs (*Adansonia*—Bombacaceae). *Ann. Mo. Bot. Gard.* 82:322-348.

BAUM, D. A. 1996. The ecology and conservation of the baobabs of Madagascar. *Primate Rep.* 46:311-327.

BAUM, D. A., AND K. OGINUMA. 1994. A review of chromosome numbers in Bombacaceae with new counts for *Adansonia*. *Taxon* 43:11-20.

BAUM, D. A., AND K. L. SHAW. 1995. Genealogical perspective on the species problem. Pages 269-303 in *Experimental and molecular approaches to plant biosystematics* (P. C. Hoch and A. G. Stephenson, eds.). Missouri Botanical Garden, St. Louis, Missouri.

BAUM, D. A., K. J. SYTSMAN, AND P. C. HOCH. 1994. A phylogenetic analysis of *Epilobium* (Onagraceae) based on nuclear ribosomal DNA sequences. *Syst. Bot.* 19:363-388.

BUCKLER, E. A. IV, A. IPPOLITO, AND T. P. HOLTSFORD. 1997. The evolution of ribosomal DNA: Divergent paralogues and phylogenetic implications. *Genetics* 145:821-832.

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42:795-803.

COE, M. J., AND F. M. ISAAC. 1965. Pollination of the baobab (*Adansonia digitata* L.) by the lesser bush baby (*Galago crassicaudatus* E. Geoffroy). *East Afr. Wildl.* 3:123-124.

CRANE, C. F., H. J. PRICE, D. M. STELLY, D. G. CZESCHIN, AND T. D. MCKNIGHT. 1993. Identification of a homeologous chromosome pair by in situ DNA hybridization to ribosomal RNA loci in meiotic chromosomes of cotton (*Gossypium hirsutum*). *Genome* 36:1015-1022.

DE QUEIROZ, K., AND M. J. DONOGHUE. 1990. Phylogenetic systematics or Nelson's version of cladistics? *Cladistics* 6:61-75.

DE QUEIROZ, A., M. J. DONOGHUE, AND J. KIM. 1995. Separate versus combined analysis of phylogenetic evidence. *Annu. Rev. Ecol. Syst.* 26:657-681.

DOWNIE, S. R., D. S. KATZ-DOWNIE, AND K.-J. CHO. 1996. Phylogenetic analysis of Apiaceae subfamily Apioideae using nucleotide sequences from the

- chloroplast *rpoC1* intron. *Mol. Phylogenet. Evol.* 6:1–18.
- EDLIN, H. L. 1935. A critical revision of certain taxonomic groups of the Malvales. *New Phytol.* 34:1–20, 122–143.
- PARIS, J. S., M. KÄLLERSJÖ A. G. KLUGE, AND C. BULT. 1994. Testing significance of incongruence. *Cladistics* 10:315–319.
- FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* 17:368–376.
- FELSENSTEIN, J. 1985. Confidence limits of phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.
- FUCHS, H. P. 1967. Pollen morphology of the family Bombacaceae. *Rev. Palaeobot. Palynol.* 3:119–132.
- GAUT, B., S. V. MUSE, W. D. CLARK, AND M. CLESS. 1992. Relative rates of nucleotide substitution at the *rbcL* locus of monocotyledonous plants. *J. Mol. Evol.* 35:292–303.
- GERMERAAD, J. H., C. A. HOPPING, AND J. MULLER. 1968. Palynology of Tertiary sediments from tropical areas. *Rev. Palaeobot. Palynol.* 6:189–348.
- GIBBS, P. E., J. SEMIR, AND N. D. DA CRUZ. 1988. A proposal to unite the genera *Chorisia* Kunth and *Ceiba* Miller (Bombacaceae). *Notes R. Bot. Gard. Edinb.* 45:125–136.
- GIFT, N., AND P. F. STEVENS. 1997. Vagaries in the delimitation of character states in quantitative variation.—An experimental study. *Syst. Biol.* 46:112–125.
- GOLDMAN, N. 1993. Statistical tests of models of DNA substitution. *J. Mol. Evol.* 36:182–198.
- GRIBEL, R. 1988. Visits of *Caluromys lanatus* (Didelphidae) to flowers of *Pseudobombax tomentosum* (Bombacaceae): A probable case of pollination by marsupials in central Brazil. *Biotropica* 20:344–347.
- HABER, W. A., AND G. W. FRANKIE. 1989. A tropical hawkmoth community: Costa Rican dry forest Sphingidae. *Biotropica* 21:155–172.
- HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating of the human–ape splitting by a molecular clock. *J. Mol. Evol.* 22:160–174.
- HEITHAUS, P., P. OPLER, AND H. G. BAKER. 1974. Bat activity and pollination of *Bauhinia pauletia*. *Ecology* 55:412–419.
- HOCHREUTNER, B. P. G. 1908. Un nouveau baobab—Revision du genre *Adansonia*. *Ann. Conserv. Jard. Bot. Geneve* 1908:136–143.
- HUELSENBECK, J. P., J. J. BULL, AND C. W. CUNNINGHAM. 1996. Combining data in phylogenetic analysis. *Trends Ecol. Evol.* 11:152–157.
- HUELSENBECK, J. P., AND B. RANNALA. 1997. Phylogenetic methods come of age: Testing hypotheses in an evolutionary context. *Science* 276:227–232.
- HUTCHINSON, J. 1967. The genera of flowering plants (Angiospermae), vol. 2. Oxford Univ. Press, Oxford, England.
- JAEGER, P. 1945. Evanouissement et pollinisation de la fleur du baobab. *C. R. Acad. Sci. Paris* 220:369–371.
- JANSEN, R. K., AND J. D. PALMER. 1987. Chloroplast DNA from lettuce and *Barnadesia* (Asteraceae): Structure, gene localization, and characterization of a large inversion. *Curr. Gen.* 11:553–564.
- 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). *Am. J. Bot.* 83:430–439.
- KAISILA, J. 1966. The Egyptian fruit-bat, *Rousettus aegyptiacus* Geoffr. (Megachiroptera, Pteropodidae) visiting flowers of *Bombax malabaricum*. *Ann. Zool. Fenn.* 3:1–3.
- KELCHINER, S. A., AND J. F. WENDEL. 1996. Hairpins create minute inversions in non-coding regions of chloroplast DNA. *Curr. Gen.* 30:259–262.
- KRUTZSCH, W. 1989. Paleogeography and historical phytogeography (paleochorology) in the Neophyticum. *Pl. Syst. Evol.* 162:5–61.
- LARSON, A. 1994. The comparison of morphological and molecular data in phylogenetic systematics. Pages 371–390 in *Molecular ecology and evolution: Approaches and applications* (B. Schierwater, B. Streit, G. P. Wagner, and R. DeSalle, eds.). Birkhäuser Verlag, Basel, Switzerland.
- MAHESHAWARI, J. K. 1971. The baobab tree: Disjunctive distribution and conservation. *Biol. Conserv.* 4:57–60.
- MASON-GAMER, R. J., AND E. A. KELLOGG. 1996. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). *Syst. Biol.* 45:524–545.
- NILSSON, S., AND A. ROBYNS. 1986. Bombacaceae Kunth. *World Poll. Spore Fl.* 14:1–59.
- OLIVEIRA, P. E., P. E. GIBBS, A. A. BARBOSA, AND S. TALAVERA. 1992. Contrasting breeding systems in two *Eriotheca* (Bombacaceae) species of the Brazilian cerrados. *Pl. Syst. Evol.* 179:207–219.
- PERRIER DE LA BÂTHIE, H. 1955. *Adansonia* L. Pages 2–17 in *Flore de Madagascar et des Comores*, 130e famille, Bombacacées (H. Humbert, ed.). Botanical Society of France, Paris.
- PORSCH, O. 1935. Zur Blütenbiologie de Affenbrotbaumes. *Osterr. Bot. Z.* 84:219–224.
- POSNO, M., A. VAN VLIET, AND G. S. P. GROTT. 1986. The gene for *Spirodela oligorhiza* chloroplast ribosomal protein homologous to *E. coli* ribosomal protein L16 is split by a large intron near its 5' end: Structure and expression. *Nucleic Acids Res.* 14:3181–3195.
- QUILTY, P. G. 1994. The background: 144 million years of Australian palaeoclimate and palaeogeography. Pages 14–43 in *History of the Australian vegetation* (R. S. Hill, ed.). Cambridge Univ., Cambridge, England.
- RABINOWITZ, P. D., M. F. COFFIN, AND D. FALVEY. 1983. The separation of Madagascar and Africa. *Science* 220:67–69.
- RAVEN, P. H., AND D. I. AXELROD. 1972. Plate tectonics and Australian paleobiogeography. *Science* 176:1379–1386.
- RAVEN, P. H., AND D. I. AXELROD. 1974. Angiosperm biogeography and past continental movements. *Ann. Mo. Bot. Gard.* 61:539–673.

- RICE, W. R. 1989. Analyzing tables of statistical tests. *Evolution* 43:223–225.
- RIESEBERG, L. H., AND D. E. SOLTIS. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol. Trends Pl.* 5:65–84.
- RIESEBERG, L. H., AND J. F. WENDEL. 1993. Introgression and its consequences. Pages 70–109 in *Hybrid zones and the evolutionary process* (R. Harrison, ed.). Oxford Univ. Press, New York.
- SANDERSON, M. J. 1998. Estimating rate and time in molecular phylogenies: Beyond the molecular clock? Pages 242–264 in *Molecular systematics of plants*, 2nd edition (P. S. Soltis, D. E. Soltis, and J. J. Doyle, eds.). Chapman and Hall, New York.
- SANG, T., D. J. CRAWFORD, S.-C. KIM, AND T. F. STUESSY. 1994. Radiation of the endemic genus *Dendroseris* (Asteraceae) on the Juan Fernandez Islands: Evidence from sequences of the ITS regions of nuclear ribosomal DNA. *Am. J. Bot.* 81:1494–1501.
- SAVARD L., M. MICHAUD, AND J. BOUSQUET. 1993. Genetic diversity and phylogenetic relationships between birches and alders using ITS, 18S rRNA, and *rbcL* gene sequences. *Mol. Phylogenet. Evol.* 2:112–118.
- STEVENS, P. F. 1991. Character states, morphological variation, and phylogenetic analysis: A review. *Syst. Bot.* 16:553–583.
- STOVER, L. E., AND A. D. PARTRIDGE. 1973. Tertiary and Late Cretaceous spores and pollen from the Gippsland Basin, southeastern Australia. *Proc. R. Soc. Vic.* 85:237–286.
- SUH, Y., L. B. THIEN, H. E. REEVE, AND E. A. ZIMMER. 1993. Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal DNA in Winteraceae. *Am. J. Bot.* 80:1042–1055.
- SWART, E. R. 1963. Age of the baobab tree. *Nature* 198:708–709.
- SYTSMA, K. J., AND L. D. GOTTLIEB. 1986. Chloroplast DNA evolution and phylogenetic relationships in *Clarkia* sect. *Peripetasma* (Onagraceae). *Evolution* 40:1248–1261.
- TAJIMA, F. 1993. Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* 135:599–607.
- TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 37:221–244.
- WALLACE, R. A., AND J. H. COTA. 1996. An intron loss in the chloroplast gene *rpoC1* supports a monophyletic origin for the subfamily Cactoideae of the Cactaceae. *Curr. Gen.* 29:275–281.
- WENDEL, J. F., AND J. J. DOYLE. 1998. Phylogenetic incongruence: Window into genome history and molecular evolution. Pages 265–296 in *Molecular systematics of plants*, 2nd edition (P. S. Soltis, D. E. Soltis, and J. J. Doyle, eds.). Chapman and Hall, New York.
- WENDEL, J. F., A. SCHNABEL, AND T. SEELANAN. 1995a. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proc. Natl. Acad. Sci. USA* 92:280–284.
- WENDEL, J. F., A. SCHNABEL, AND T. SEELANAN. 1995b. An unusual ribosomal sequence from *Gossypium gossypoides* reveals ancient, cryptic, intergenomic introgression. *Mol. Phylogenet. Evol.* 4:298–313.
- WERTH, E. 1915. Kurzer Überblick über die Gesamtfrage der Ornithophilie. *Bot. Jahrb. Syst.* 53:314–378.
- WHITE, T. J., T. BIRNS, S. LEE, AND J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pages 315–322 in *PCR protocols: A guide to methods and applications* (M. Innis, D. Gelfand, J. Sninsky, and T. White, eds.). Academic Press, San Diego.
- WHITTEMORE, A. T., AND B. A. SCHAAL. 1991. Interspecific gene flow in sympatric oaks. *Proc. Natl. Acad. Sci. USA* 88:2540–2544.
- WICKENS, G. E. 1983. The baobab: Africa's upside-down tree. *Kew Bull.* 37:173–209.
- WILFORD, G. E., AND P. J. BROWN. 1994. Maps of late Mesozoic-Cenozoic Gondwana break-up: Some palaeogeographical implications. Pages 5–43 in *History of the Australian vegetation* (R. S. Hill, ed.). Cambridge Univ. Press, Cambridge, England.
- WOLFE, A. D., AND W. J. ELISENS. 1995. Evidence of chloroplast capture and pollen-mediated gene flow in *Penstemon* sect. *Peltanthera* (Scrophulariaceae). *Syst. Bot.* 20:395–412.
- WOLFE, J. A. 1975. Some aspects of plant geography of the northern hemisphere during the Late Cretaceous and Tertiary. *Ann. Mo. Bot. Gard.* 62:264–279.
- WOLFE, J. A. 1976. Stratigraphic distributions of some pollen types from the Campanian and Lower Maestrichtian rocks (Upper Cretaceous) of the Middle Atlantic states. *Geol. Surv. Prof. Pap.* 977:1–18.
- YANG, Z. 1996. PAML: Phylogenetic analysis by maximum likelihood, version 1.2. Department of Integrative Biology, Univ. California, Berkeley.
- YANG, Z., N. GOLDMAN, AND A. FRIDAY. 1995. Maximum likelihood trees from DNA sequences: A peculiar statistical estimation problem. *Syst. Biol.* 44:384–399.

Received 19 August 1997; accepted 2 October 1997
Associate Editor: D. Cannatella

APPENDIX 1

Material used for DNA extractions. Ingroup taxonomy based on Baum (1995a). Specimens at MO unless otherwise stated. Population vouchers refer to specimens of adult trees from the same population in which seeds were collected.

Genus Species	Tissue ^a	Collection
<i>Adansonia</i>		
<i>A. grandidieri</i> a	L	Baum 347 (population voucher: Baum 345); Madagascar, Prov. Toliara, Morondava
<i>A. grandidieri</i> b	S	Baum 345 (population voucher); Madagascar, Prov. Toliara, Morondava
<i>A. suarezensis</i> a	L	Baum 348; Madagascar, Prov. Antsiranana, Montagnes des Français
<i>A. suarezensis</i> b	S	no voucher; Madagascar, Prov. Antsiranana, Cap Diego, ca. 1 km S of village on beach
<i>A. digitata</i> a	L	Small s.n. (ISC); cultivated, Iowa State Univ. Greenhouse, seed source unknown
<i>A. digitata</i> b	L	Baum 349; Burkina-Faso. Pobé Mengao/Touflé, 35 km from Djibo in an area west of Titao and 40 km NE of Quschigouya (collected by A. J rgensen)
<i>A. gibbosa</i>	L	Wendel s.n. (ISC); northwest Australia
<i>A. rubrostipa</i>	L	Baum 354 (population voucher: Baum 313); Madagascar, Prov. Toliara, Kirindy forest near Morondava
<i>A. madagascariensis</i> a	L	Baum 355 (population voucher: Baum 338); Madagascar, Prov. Antsiranana, 14 km S of Ambanja
<i>A. madagascariensis</i> b/c	S	Baum 338; (population voucher); Madagascar, Prov. Antsiranana, 14 km S of Ambanja
<i>A. za</i> a	L	Baum 336 (population voucher); Madagascar, Prov. Antsiranana, Manongarivo, Ambondrona, near Ambaibo
<i>A. za</i> b	L	Baum 357; Madagascar, Prov. Toliara, without further locality (seed provided by S. Taylor, Missouri Botanical Garden)
<i>A. za</i> c	S	No voucher; Madagascar, Prov. Toliara, Manombovo, 38 km E of Beloha on the road to Tsihoumbé (collected by D. Baum)
<i>A. perrieri</i>	L	Baum 359; Madagascar, Prov. Antsiranana, Montagne d'Ambre
<i>Pseudobombax</i>		
<i>P. marginatum</i>	L	Small s.n. (ISC); Bolivia (collected by T. Killeen)
<i>Pachira</i>		
<i>P. aquatica</i>	L	Alverson s.n. (WIS); Venezuela
<i>Bombax</i>		
<i>B. buonopozense</i>	L	Alverson s.n. (WIS); Pacific Tropical Garden accession 770474001, originally from Nigeria

^a L = leaf tissue of a cultivated seedling; S = single seed.

Pseudo AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 Pachir AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 Bombax AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A gib AAATCTCAAAGCAAGA - CCGAGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A diga AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A digb AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A graa AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATCTTTTTTAGTCATTTTTTTTT
 A suab AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A suaa AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A rub AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A mada AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A madb AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A zaa AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A zab AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A zac AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT
 A per AAATCTCAAAGCAAGAATCCAGGATTCATCTATTATCCMTAACTACCTGGTATATCTTAAGAATAAAATAAAATATTTTTTAGTCATTTTTTTTT

Pseudo -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 Pachir -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 Bombax -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A gib -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A diga -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A digb TCGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A graa -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A grab -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A suaa -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A suab -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A rub -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A mada -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A madb -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A mada -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A zaa -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A zab -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A zac -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT
 A per -CGCGAGGAGCTGGATGAGAAGAACTCTCATGTCCAGTCTGTAGTAGAGATGGAATGTATAAACCAACTCAACTATAACCCAAAAGAACCCAGATTT

Pseudo CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 Pachir CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 Bombax CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A gib CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A diga CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A digb CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A graa CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A grab CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A suab CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A suaa CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A rub CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A mada CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A madb CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A mada CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A zaa CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A zab CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A zac CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA
 A per CGTAAACAACATAGAGGAAGAATGAAAGGAATATCTTATCGAGGTAATCGCATTTGTTTCGGTAGATATGCTCTTCAAGCACTTGAACCCGCTTGGATTA

Pseudo ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0000
 Pachir ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0000
 Bombax ATCTAGACAAATAGAA???????????????????????????? 0010
 A gib ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0011
 A diga ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 1012
 A digb ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 1013
 A graa ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 1110
 A grab ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 1110
 A suaa ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 1012
 A suab ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 1012
 A rub ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0010
 A mada ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0010
 A madb ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0010
 A mada ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0010
 A zaa ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0010
 A zab ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0010
 A zac ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0010
 A per ATCTAGACAAATAGAAGCGGGCGCCGCAATGACACGAAATGTAC 0010

APPENDIX 3

Aligned restriction site matrix.

Pseudo 0000701001110003:0010010100601101100112000000000000111
 Pachir 00:000000000:0030001000001000000000100100100000000001
 A gib 1101101011101100110110100011001000000000100001010010000
 A diga 1101101011101101101101000110010000000000000000000000
 A digb 11011010021101200110110000110100000000000000000000000
 A graa 111010100111010001101100001110010000000000000000000000
 A suaa 110010100111010101101100000110010000000000000000000000
 A rub 11101110011101000:11011000000110010000000011000100010000
 A mada 1100101001110100011011000001100100110000100011000000000
 A madb 111010100111010001101100000110010000000000000000000000
 A zaa 111010100111010001101100000110010000000000000000000000
 A zab 110011100111010001101100000110010000000000000000000000
 A zac 110011100111010001101100000110010000000000000000000000
 A per 1100101001110100011011000001100100100000100001000100000

APPENDIX 4

Aligned sequences of ITS. A single line over the sequence indicates the ribosomal 18S, 5.8S, and 26S genes. Potentially informative gaps are numbered and scored at the end of the matrix. Pseudobombax is scored as 0/1 for gap character 9.

Table with columns for species (Pseudo Pachir Bombax A gib A dig A graa A suaa A suab A rub A mada A nade A zaa A zab A zac A per) and rows for aligned DNA sequences. Includes alignment markers (18S, 5.8S, 26S) and a character matrix at the bottom.

5.8S

Pseudo ACGCCGCGCTGGGTGTACAGCATAGTGCACCCCTCCAA--CCCTAGCCCTTGGGGCTTCGGATGAGGTGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 Pachir ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----CTAGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 Bombax ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----CTAGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A gib ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A diga ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A digb ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A graa ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A grab ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A suaa ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A suab ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A rub ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A mada ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A madb ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A madc ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A zaa ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A zab ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A zac ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG
 A per ACCTGTGCCGGGTGTACAGCATAGTGCACCCCTCCGGAACCATACCCCAAGGGG-----TACGGAARAGGCGGGGGCGGAAAATGGCTCCGGTCGGCTCCG

10

11

Pseudo CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 Pachir CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 Bombax CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A gib CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A diga CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A digb CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A graa CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A grab CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A suaa CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A suab CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A rub CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A mada CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A madb CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A madc CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A zaa CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A zab CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A zac CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG
 A per CGCTCGCGGTGGCTTAAAAATCGAGTCTTCGGCAACGGCAGCGGCGGACGATCGGTGGTGTATGCCTCG-----TCC--CCGTCGCGTCCGCG

268

Pseudo TCTTCGCTCCCG--GAAACCCATCGAGA--CCCTG-CTAGCGTTCGGCAAGCGCTCGCACCCAGTTCAGCGGGGATACCG????????????
 Pachir TTTTCGTCGCCCG--GGACTCATCGAGAGACCCCTT--CCGGCGTTCGCACACCCAGTGCAGCCGATCGCGACCCAGGTGAGCGGGGATAC
 Bombax TTTAG?TCGKCU--GGACCAC?GAGA--CCCTA-CT?????TCGACAGAGG????????????????????????????????
 A gib TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTTGTAGCGTTCGCACAAGCGTTCGACCGGTCGCGACCCAGGTGAGCGGGGATAC
 A diga TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTG-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A digb TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A graa TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A grab TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A suaa TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A suab TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A rub TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A mada TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A madb TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A madc TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A zaa TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A zab TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC
 A zac TTTTCGTCGCCCG--GA--CCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCNACC????????????????
 A per TTTTCGTCGCCCG--GGACCATPCGAGA--CCCTT-CTAGCGTTCGGCAAGCGCTCGCACCCATCGCGACCCAGGTGAGCGGGGATAC

Pseudo	0	0	1	0	3	0	0	0	0	0	0	0
Pachir	0	0	1	0	3	0	0	0	0	0	0	0
Bombax	0	1	1	0	3	0	0	0	0	0	0	0
A gib	1	1	0	1	1	1	1	1	1	1	1	1
A diga	1	1	0	1	2	1	1	1	1	1	1	1
A digb	1	1	0	1	2	1	1	1	1	1	1	1
A graa	1	1	0	1	1	1	1	1	2	1	1	1
A grab	1	1	0	1	1	1	1	1	2	1	1	1
A suaa	1	1	0	1	1	1	1	1	2	1	1	1
A suab	1	1	0	1	1	1	1	1	2	1	1	1
A rub	1	1	0	1	1	1	1	1	2	1	1	1
A mada	1	1	0	1	1	1	1	1	2	1	1	1
A madb	1	1	0	1	1	1	1	1	2	1	1	1
A madc	1	1	0	1	1	1	1	1	2	1	1	1
A zaa	1	1	0	1	1	1	1	1	2	1	1	1
A zab	1	1	0	1	1	1	1	1	2	1	1	1
A zac	1	1	0	1	1	1	1	1	2	1	1	1
A per	1	1	0	1	1	1	1	1	2	1	1	1

APPENDIX 5

Morphological characters and character states. Multistate characters are labeled as either ordered or unordered.

1. Germination type. 0 = cleistocotylar (the cotyledons retained at soil level within the testa); 1 = phanerocotylar (the cotyledons released from testa and raised clear of the soil level).
2. Seed shape. 0 = regularly or irregularly globose; 1 = reniform (with a distinct sinus on one side).
3. Seed dimensions. 0 = unflattened (width laterally and dorsoventrally approximately equal); 1 = laterally flattened (lateral width approximately half the dorsoventral width).
4. Pericarp thickness 0 = relatively thin and fragile (<1 cm); 1 = relatively thick and woody (>1 cm). In view of the distinctly different fruit morphology in the outgroups we scored them as absent for this character.
5. Fruit shape. 0 = globose (length:width ca. 1); 1 = ovoid-oblong (length:width >1.5).
6. Flower bud shape. 0 = globose just prior to anthesis (length:width ca. 1); 1 = ovoid just prior to anthesis (length:width = 1.5-3); 2 = cylindrical just prior to anthesis (length:width >5). Treated as ordered.
7. Calyx abscission. 0 = calyx persists after the corolla and androecium falls; 1 = calyx abscises in a unit with the corolla and androecium.
8. Style abscission. 0 = style abscises with the androecium; 1 = style persists after the androecium falls.
9. Calyx length. 0 = truncate (calyx length:corolla length <0.2); 1 = long (calyx length:corolla length >0.9).
10. Crown form. 0 = main branches generally ascending or irregular, crown in mature trees usually rounded or irregular; 1 = main branches horizontal, often whorled, crown in mature trees distinctly flat-topped.
11. Flowering season. 0 = early wet-season (Nov.-Feb. in Madagascar); 1 = late wet-season (Mar.-Apr. in Madagascar); 2 = dry season (Jun.-Sept. in Madagascar). Treated as unordered.
12. Inner calyx color. 0 = pale cream or greenish; 1 = red.
13. Petal color. 0 = white or pale cream; 1 = yellow perhaps with a red medial streak; 2 = dark red throughout. Treated as unordered.
14. Style color. 0 = white or pale cream; 1 = dark red at least for the top 1 cm.
15. Calyx nectar chamber. 0 = calyx base cylindrical; 1 = annular inflation of the calyx around the nectariferous tissue.
16. Petal length to width ratio. 0 = approximately 1:1; 1 = 2-5:1; 2 = >10:1. Treated as ordered.
17. Petal widest point. 0 = Proximal to mid-point; 1 = Distal to midpoint.
18. Filament number. 0 = 90-200; 1 = 200-350; 2 = >500. Treated as ordered.
19. Staminal tube length. 0 = approximately equal in length to the ovary; 1 = at least twice as long as ovary.
20. Filament length. 0 = shorter than the staminal tube; 1 = longer than staminal tube.
21. Outer calyx color. 0 = green; 1 = brown.
22. Leaflet color. 0 = bright to dark green; 1 = pale bluish green.
23. Leaflet shape. 0 = narrow (narrowly elliptical/lanceolate); 1 = broad (oblong, obovate, or broadly elliptical).
24. Pedicel/peduncle orientation. 0 = erect or horizontal; 1 = pendulous.
25. Fruit dehiscence. 0 = indehiscent. 1 = irregular cracking; 2 = regular lines of dehiscence. Treated as unordered.
26. Leaf margin in adult leaves. 0 = entire; 1 = serrate.
27. Staminal bundles. 0 = none (filaments free from top of staminal tube); 1 = central bundle (central filaments fused around style); 2 = phalanges (filaments grouped into clusters). Treated as unordered.
28. Stipule persistence on reproductive branches. 0 = caducous; 1 = persistent even after leaf abscission.
29. Leaflet articulation. 0 = present; 1 = absent.
30. Number of simple leaves produced by seedling. 0 = 0-1; 1 = >2.

APPENDIX 6

Morphological matrix. Polymorphisms/uncertainties that do not include all the character states are underlined.

Pseudo	1	0	0	?	1	<u>12</u>	0	1	0	1	?	0	0	0	0	1	1	<u>12</u>	1	1	0	0	1	0	2	0	2	0	1	?	
Pachir	0	0	0	?	0	2	0	1	0	1	?	0	1	1	0	<u>12</u>	0	1	1	1	?	0	1	0	0	0	2	0	0	0	
Bombax	1	0	0	?	1	1	0	0	1	2	0	2	1	0	0	1	1	1	0	1	0	0	1	0	2	0	2	0	0	0	
A gib	1	1	1	0	?	2	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	1	0	1	0	0	0	0	1		
A dig	1	1	1	1	?	0	1	1	1	0	?	0	0	0	0	0	1	2	1	0	0	0	1	1	0	0	0	0	0	1	
A gra	0	0	1	0	0	1	1	1	1	1	2	0	0	0	0	1	2	0	1	1	1	0	0	0	0	0	0	0	1		
A sua	0	0	1	0	1	1	1	1	1	1	2	0	0	0	0	1	2	0	1	0	0	1	0	0	0	0	0	0	0	1	
A rub	1	1	1	0	0	2	0	0	1	0	0	1	1	1	0	2	0	0	1	1	0	0	0	0	0	1	1	0	0	1	
A mad	1	1	1	1	0	2	0	0	1	0	1	1	<u>12</u>	1	1	2	0	0	1	1	0	0	1	0	0	0	0	0	0	1	
A za	1	1	1	1	1	2	0	1	1	0	0	1	1	1	1	2	0	0	1	1	0	0	1	0	0	0	0	0	0	1	
A per	1	1	1	1	1	2	?	0	1	0	0	?	1	1	1	1	1	0	1	0	0	0	0	1	0	0	0	0	1	0	0

