2004

Phylogenetic Relationships of Malvatheca (Bombacoideae and Malvoideae; Malvaceae sensu lato) as Inferred from Plastid DNA Sequences

David A. Baum
University of Wisconsin - Madison

Stacey DeWitt Smith
University of Nebraska - Lincoln, ssmith19@unl.edu

Alan Yen
Harvard University Herbaria,

William S. Alverson
The Field Museum (Chicago, Illinois)

Reto Nyffeler
Harvard University Herbaria,

See next page for additional authors

Follow this and additional works at: http://digitalcommons.unl.edu/bioscifacpub

Part of the Life Sciences Commons

Baum, David A.; Smith, Stacey DeWitt; Yen, Alan; Alverson, William S.; Nyffeler, Reto; Whitlock, Barbara A.; and Oldham, Rebecca L., "Phylogenetic Relationships of Malvatheca (Bombacoideae and Malvoideae; Malvaceae sensu lato) as Inferred from Plastid DNA Sequences" (2004). Faculty Publications in the Biological Sciences. 110.
http://digitalcommons.unl.edu/bioscifacpub/110

This Article is brought to you for free and open access by the Papers in the Biological Sciences at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications in the Biological Sciences by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Phylogenetic Relationships of Malvatheca (Bombacoideae and Malvoideae; Malvaceae sensu lato) as Inferred from Plastid DNA Sequences

David A. Baum, Stacey DeWitt Smith, Alan Yen, William S. Alverson, Reto Nyffeler, Barbara A. Whitlock, and Rebecca L. Oldham

Previous molecular phylogenetic analyses have revealed that elements of the former families Malvaceae sensu stricto and Bombacaceae together form a well-supported clade that has been named Malvatheca. Within Malvatheca, two major lineages have been observed; one, Bombacoideae, corresponds approximately to the palmate-leaved Bombacaceae, and the other, Malvoideae, includes the traditional Malvaceae (the mallows or Eumalvoideae). However, the composition of these two groups and their relationships to other elements of Malvatheca remain a source of uncertainty. Sequence data from two plastid regions, ndhF and trnK intron, from 34 exemplars of Malvatheca and six outgroups were analyzed. Parsimony, likelihood, and Bayesian analyses of the sequence data provided a well-resolved phylogeny except that relationships among lineages at the base of Malvatheca are poorly resolved. Nonetheless, a 6-bp insertion in matK suggests that Fremontodendraceae is sister to the remainder of Malvatheca. Our results suggest that the Malvoideae originated in the Neotropics and that a mangrove taxon dispersed across the Pacific from South America to Australasia and later radiated out of Australasia to give rise to the ca. 1700 living species of Eumalvoideae. Local clock analyses imply that the plastid genome underwent accelerated molecular evolution coincident with the dispersal out of the Americas and again with the radiation into the three major clades of Eumalvoideae.

Key words: biogeography; Bombacaceae; Eumalvoideae; local clock analysis; Malvatheca; matK; ndhF; phylogenetic nomenclature.

Molecular data accumulated during the last decade have greatly enhanced our knowledge of phylogenetic relationships within Malvaceae sensu lato (s.l.), a clade that includes members of four previously recognized families: Malvaceae sensu stricto (s.s.), Bombacaceae, Sterculiaceae, and Tiliaceae (Alverson et al., 1999; Bayer et al., 1999; Nyffeler and Baum, 2000; Whitlock et al., 2001; Pfeil et al., 2002). Analyses of the plastid genes atpB, rbcL, and ndhF revealed that elements of the former families, Malvaceae s.s and Bombacaceae, together form a well-supported clade (Alverson et al., 1999; Bayer et al., 1999), which has been named Malvatheca (Baum et al., 1998). Two major lineages were discerned within Malvatheca: Bombacoideae, corresponding approximately to the palmately leaved taxa of Bombacaceae, and Malvoideae, which includes Malvaceae s.s. However, the composition of Malvoideae and Bombacoideae and the relationships of certain problematic taxa within Malvatheca remain uncertain.

The largely subtropical and temperate mallow clade or Eumalvoideae contains a large portion of the species diversity of Malvatheca, with at least 1700 species, including such familiar groups as the herbaceous mallows, hibiscus, okra, and cotton. The Bombacoideae, in contrast, contain many fewer species (even if all taxa of uncertain affinities were included, there would still be fewer than 250 species) and are predominantly tropical trees. Indeed, the group has been touted as a classic case of a diverse temperate radiation emerging from a less diverse tropical source (Judd et al., 1994; Judd and Manchester, 1997). However, a full understanding of the evolutionary mechanisms responsible for the mallow radiation is contingent on a more fully resolved phylogeny than has been available to date.

In addition to providing a framework for analysis of the mallow radiation, clarification of the relationships within Malvatheca is needed to evaluate published hypotheses regarding floral evolution and biogeography. Previous molecular data suggested the surprising hypothesis that the monothecate "half" anthers of many traditional Bombacaceae (e.g., Bombax, Adansonia) and traditional Malvaceae (e.g., Malva, Hibiscus), while looking very similar, could represent parallel evolution from an ancestor with sessile thecae (Alverson et al., 1999). However, support for this hypothesis was weak due to...
poor phylogenetic resolution near the base of Malvatheca. Similarly, biogeographic hypotheses need to be evaluated on a more inclusive phylogenetic data set than was previously considered. The Bombacoideae are predominantly neotropical whereas putatively early-branching Malvoideae have a predominantly Australasian/Asian distribution (Pfeil et al., 2002). One hypothesis to explain this pattern is that it is the product of Gondwanan disjunction (Pfeil et al., 2002), but this needs testing with a broader sampling of taxa.

The key to answering these evolutionary questions lies in determining the placement of a number of "wildcard" taxa that have uncertain relationships. The neotropical rainforest tree Pentaplatis (three species; traditionally placed in Tiliaceae; Williams and Standley, 1952) was found to be more closely related to Malvoideae using atpB and rbcL (Bayer et al., 1999), but this placement was very weak. The genus was tentatively assigned to Bombacoideae based on morphological data (Bayer et al., 1999), but then treated as incertae sedis by Bayer and Kubitzki (2003). The Matisieae clade, with about 80 neotropical tree species in three genera (Matisia, Quararibea, and Phragmatolea), was weakly supported as an early-branching lineage of the Malvoideae (Alverson et al., 1999). The Mexican and Californian Fremontodendron (three species) and its Mexican relative, Chiranthodendron (one species), were very weakly affiliated with the Malvoideae (Alverson et al., 1999) or Bombacoideae (Bayer et al., 1999). Ochroma (one neotropical species) was resolved either in the Bombacoideae (Bayer et al., 1999) or in a weakly supported clade with Patinoa (four neotropical species) at the base of Malvoideae (Alverson et al., 1999). Finally, two monotypic neotropical taxa, Septotheca and Uladendron, have not previously been studied phylogenetically and cannot easily be assigned to either Malvoideae or Bombacoideae based on morphological data.

The early-branching Malvoideae represent a second area in which further phylogenetic resolution is critical. Alverson et al. (1999) found that Campostemon, comprising two species of Australasian and Malesian mangroves that have traditionally been placed in Bombacaceae (e.g., Hutchinson, 1967), was well supported within Malvoideae, but outside Eumalvoideae. Similarly, Pfeil et al. (2002) showed that Radyera (two shrubby species, one Australian, one South African), Lagunaria (one small tree species from Australia and the Norfolk and Lord Howe Islands), and Howittia (one shrubby species from Australia) are also early branches of the core Malvoideae (= Malvaceae s.s.). However, apart from strong support for a Howittia-Lagunaria clade, the relationships among these taxa and the remainder of Malvoideae were very weakly supported.

In this study, we employed sequence data from the plastid genes ndhF and matK to examine phylogenetic relationships among the early-branching lineages of Malvatheca and Malvoideae and to study the group’s biogeographic history. We also explored the pattern of molecular evolution to assess the degree to which changes in rates of molecular evolution correlate with changes in life form and/or the burst of species diversification seen within Malvoideae.

MATERIALS AND METHODS

Sampling.—We generated six new ndhF sequences, 19 new matK sequences, which we combined with previously published sequences to create a complete data set for 40 taxa (Appendix 1; see Supplemental Data accompanying the online version of this article). Ingroup sampling included 31 genera known from previous studies to belong to Malvatheca along with two unsampled genera, Uladendron and Septotheca. Genera from Brownlowioideae, Dombeioideae, Helicterioideae, Sterculioideae, and Tilioideae were used as outgroups (see Alverson et al., 1999; Bayer et al., 1999).

DNA extraction and sequencing.—Total genomic DNA was extracted from fresh or silica-dried leaf material as described in Alverson et al. (1999). The ndhF gene was amplified as two overlapping fragments as described in Olmstead et al. (1993) and Alverson et al. (1999). Additionally, we used two primers, 13R (CCGAACATAAAAATGCAAGTTAACCC) and 10R (CCCCCTACATATTTGGATACCTCCTC), which are modified versions of 131R and 2110R (Olmstead et al., 1993), respectively. The matK region, comprising the trnK intron and matK coding sequence, was also amplified in two separate polymerase chain reactions (PCR) using primers modified from Johnson and Solis (1994) as described in Nyffeler et al. (in press).

The PCR products were purified using the Qiagen PCR purification kit (Qiagen, Valencia, California, USA) or AMPure beads (Agencourt Bioscience, Beverly, Massachusetts, USA) and cycle-sequenced (Big Dye version 3.1; Applied Biosystems, Foster City, California, USA) using the manufacturers’ protocols. All PCR products were sequenced in both directions. Sequences were edited and assembled in Sequencer 4.1 (Gene Code, Ann Arbor, Michigan, USA), imported into MacClade (Maddison and Maddison, 2002), and aligned manually.

Phylogenetic analysis.—We estimated incongruence between ndhF and matK with the incongruence length difference (ILD) test (Farris et al., 1994), implemented as the partition homogeneity test in PAUP* version 4.0b10 (Swofford, 2002) using simple taxon addition tree bisection-reconnection (TBR) searches holding 10 trees at each step, and with maxtrees set to 100. Maximum parsimony (MP) and maximum likelihood (ML) analyses were performed in PAUP* 4.0b10 (Swofford, 2002). The MP heuristic searches used 100 random taxon addition replicates (holding one tree at each step) and TBR branch swapping. All characters were equally weighted, and gaps were treated as missing data. To estimate clade support, we obtained bootstrap percentages (BP) for each clade using 10000 bootstrap replicates with simple taxon addition (holding one tree at each step) and TBR branch swapping. Templeton tests (Templeton, 1983) as implemented in PAUP* 4.0b10 (Swofford, 2002) were used to explore alternative topologies in an MP framework. These tests are described in more detail in the results section.

ML searches with sub-tree pruning regrafting (SPR) branch-swapping were initially conducted under a Jukes-Cantor (JC) model using maximum parsimony trees as starting trees. The ML tree scores were calculated on the JC trees for the following models (in order of increasing complexity): K2P HKY, HKY + G, GTR + G (Swofford et al., 1996, and references therein). Likelihood ratio tests were used to compare the fit of the models and to choose the most appropriate (Huelsenbeck and Rannala, 1997). Heuristic searches were completed with the best model using the JC tree as the starting tree and SPR branch swapping.

Bayesian phylogenetic analysis was conducted on the combined data set in MrBayes version 3.0 (Huelsenbeck and Ronquist, 2001). We allowed as many model parameters as suggested by hierarchical likelihood tests (as previously discussed) and created four data partitions: the trnK intron, and the first, second, and third positions of the coding regions. Each partition was considered independent with different model parameters. We conducted three Markov chain Monte Carlo (MCMC) analyses, each composed of four linked chains (sequential heat = 0.2) run for 5000000 generations, sampling every 100 generations. The burn-in period was estimated by visual examination of a likelihood-by-generation plot. We compared the three majority rule consensus trees and then pooled the posterior distributions to obtain our best estimates of clade posterior probabilities (PP).

It is commonly noted that very short branches may receive inflated posterior support. One reason for this is because the prior probability density placed on a hard polytomy (a zero length branch) is zero (P. Lewis, University of Connecticut, personal communication). To evaluate whether suspect branches could indeed represent hard polytomies, we used likelihood ratio tests to determine whether branches appearing in the Bayesian majority-rule consensus
tree are significantly longer than zero length (Felsenstein, 1988). We estimated parameters and obtained the likelihood score of the Bayesian majority rule consensus tree topology under a GTR + $\Gamma$ model. We then collapsed each branch in turn and determined whether the resultant tree had a likelihood score significantly different from the fully resolved tree. Twice the likelihood ratio was assumed to fit a chi-square distribution with one degree of freedom (because one branch, i.e., parameter, was constrained to a length of zero).

Based on ndhF data, Malvoideae have been shown to have an increased rate of molecular evolution relative to Bombacoideae (Alverson et al., 1999). We wished to quantify this acceleration and more accurately ascertain the branch(es) on which the rate of molecular evolution changed. The strategy we chose involved the use of “local clock” models available in PAML version 3.13 (Yang, 2002) using the ML topology with outgroups pruned and using the model of molecular evolution suggested by likelihood ratio tests (GTR + $\Gamma$). We first assumed a two-rate model with core Bombacoideae and core Malvoideae having different rates of molecular evolution with a shift in rate mapped to one branch of the phylogeny between these two clades. We calculated the likelihood for all possible rate-change points (internodes assigned greater than zero length) and selected that which gave the highest likelihood score. Second we repeated the above procedure but allowed for three rates (two changes in rate), one for core Malvoideae, one for “intermediate” taxa, and one for core Bombacoideae. Because the two-rate model suggested a rate change just below the Eumalvoideae, we explored all possible branches for the second rate change between that point and the core Bombacoideae. Again, we selected the rate-change point that maximized the likelihood of the data. The two-rate and three-rate models were compared with the one-rate (clock) and free (non-clock) models using likelihood ratio tests. These latter tests should be treated cautiously since local clock rates categories were assigned, at least partly, a posteriori (i.e., assignments were influenced by the data).

For reasons discussed in the results, we suspected that the inclusion of the neotropical rainforest tree taxon, Uladendron, within a clade of Australasian mangroves might be erroneous. We wished, therefore, to repeat the three-rate local model on the ML tree obtained under the constraint that Uladendron not be a member of the Australasia grade (or Eumalvoideae). We optimized parameters of the GTR + $\Gamma$ model on one of the 123 MP trees under this constraint and then started a constrained, NNI heuristic search, using the 123 MP trees as starting trees. The single ML tree found was then used for local clock analyses in PAML under the three-rate model.

RESULTS

Phylogenetic analyses—The ndhF and matK data sets have aligned lengths of 2156 and 2755 bases and 167 and 249 parsimony-informative sites, respectively. Separate analysis of ndhF and matK revealed no strongly supported contradictory clades. A partition homogeneity test failed to reject the null hypothesis of no discordance ($P = 0.07$), providing no evidence of meaningful conflict between these partitions.

 Parsimony analyses of the combined data set resulted in 356 MP trees of length 847 steps with a consistency index (CI) of 0.613 and a retention index (RI) of 0.758 (Fig. 1A). Likelihood ratio tests indicated that GTR + $\Gamma$ was the best fitting likelihood model for the combined data. Maximum likelihood (ML) analyses under this model (Fig. 1B) produced a topology that was very similar to that obtained by MP, at least for those branches with moderate bootstrap support.

 For reasons discussed in the results, we suspected that the inclusion of the neotropical rainforest tree taxon, Uladendron, within a clade of Australasian mangroves might be erroneous. We wished, therefore, to repeat the three-rate local model on the ML tree obtained under the constraint that Uladendron not be a member of the Australasia grade (or Eumalvoideae). We optimized parameters of the GTR + $\Gamma$ model on one of the 123 MP trees under this constraint and then started a constrained, NNI heuristic search, using the 123 MP trees as starting trees. The single ML tree found was then used for local clock analyses in PAML under the three-rate model.

For reasons discussed in the results, we suspected that the inclusion of the neotropical rainforest tree taxon, Uladendron, within a clade of Australasian mangroves might be erroneous. We wished, therefore, to repeat the three-rate local model on the ML tree obtained under the constraint that Uladendron not be a member of the Australasia grade (or Eumalvoideae). We optimized parameters of the GTR + $\Gamma$ model on one of the 123 MP trees under this constraint and then started a constrained, NNI heuristic search, using the 123 MP trees as starting trees. The single ML tree found was then used for local clock analyses in PAML under the three-rate model.
Fig. 2. Bayesian analysis of combined data using four data partitions and the GTR + Γ model of evolution. Three separate analyses of 5,000,000 generations were run (each comprising four Metropolis-coupled chains). The 40,000 trees retained from each analysis were pooled to generate the majority rule consensus tree shown. The posterior probability (PP) of each clade is shown above the branches. Branches that were found not to be significantly longer than zero length by a likelihood ratio test are indicated by parentheses around the corresponding posterior probabilities. Clade names used in the text are indicated to the right.

and used the remaining 40,000 trees to estimate the posterior distribution. All three runs yielded identical 50% majority rule consensus trees. Rounded clade posterior probabilities were very similar between the three runs, differing by an average of only 0.16% and never differing in the ingroup by more than 1%. This result suggests good mixing during the MCMC analysis and, therefore, we combined the 40,000 post burn-in trees from each run into a single pool of 120,000 that was used as a basis for the calculation of a majority rule consensus tree (Fig. 2).

Bayesian analysis failed to provide strong resolution of the earliest divergences within Malvatheca: five lineages are supported (Ochroma-Patinoa, Septotheca, Chiranthodendron-Fremontodendron, Malvoideae, and Bombacoideae) but these appear in all possible arrangements within the posterior distribution with approximately equal posterior probabilities. Thus, the sequence data fail to resolve this region of the phylogeny. Likewise, the relationships among five early branching lineages of Malvoideae (Radyera, Camptostemon, Uladen-dron, Howittia-Lagunaria, and Eumalvoideae) are not strongly resolved. However, in this case, three resolutions of Radyera account for 99.9% of the posterior: sister to all core Malvoideae except Pentaplaris (43.4%), in a clade with Uladen-dron, Howittia, Lagunaria, and Camptostemon (29.3%), or sister to the Eumalvoideae (27.2%). Among trees with the latter arrangement, it is noteworthy (see Discussion) that 57 trees (0.17%) have Uladen-dron situated as sister to all core Malvoideae except Pentaplaris.

**Indel data**—We examined the data for parsimony-informative indel characters that are not alignment sensitive. Seven such informative indels were found in matK and four in ndhF. Only two of these indels are homoplasious on the optimal trees (one appears to unite Septotheca and Abutilon, the other Hibiscus costatus, Decaschistia, and Abutilon). All but one of the non-homoplasious indels support branches in the Eumalvoideae. One indel supports the monophyly of each of the three main clades: Gossypieae (Hampea, Thespesia, Gossypium), Malveae (Lavatera, Malope, Abutilon), and Hibisceae s.s. (Hibiscus, Pavonia, Decaschistia, and Malvaviscus). Additionally, four indels support the Decaschistia-Hibiscus costatus clade and one the Pavonia-Malvaviscus clade. However, the most interesting indel character is a 6-bp/two-residue insertion in matK that is shared by all Malvatheca except Fremontodendreae. This insertion is in-frame and is situated in a relatively conserved part of the coding sequence. The insertion itself is conserved: all but three taxa have the amino acid sequence Tyr-Glu, and the three exceptions (Abutilon, Howittia, and Quararibea) each have at least one of the two residues conserved. Given that no other homoplasious indels support any alternative positions for Fremontodendreae, the matK indel
suggests that Fremontendorae is sister to the rest of Malvateca.

**Topology tests**—Templeton tests were used to test certain phylogenetic hypotheses. First, we were interested to know whether the data can reject the neotropical taxon *Uladendron* being sister to the Eumalvoideae plus the Australasian taxa (*Camptostemon, Howittia, Lagunaria, Radyera*). This relationship would be the most biogeographically parsimonious because *Uladendron*, like *Pentaplaris* and Matisieae, has a neotropical distribution. Therefore, we searched for MP trees under the constraint that Eumalvoideae forms a clade with the Australasian taxa but not *Uladendron*. The shortest tree compatible with this constraint is only one step longer than the unconstrained tree, which is judged not to be significant using a Templeton test (P = 0.56–0.66). We also tested the possibility that the taxa within Malvateca that possess stalked monothecate anthers (*Gossypium, Hampea, Theesperia, Latavera, Malope, Abutilon, Pavonia, Malaviscus, Hibiscus, Deccachistia, Alyogyne, Uladendron, Radyera, Lagunaria, Howittia, Camptostemon, Pentaplaris, Adansonia, Ceiba, Pachira, Bombax, Catostemma, Scieronema*) form a clade. Such a relationship would suggest a single origin of this anther morphology as opposed to the separate origins hypothesized by Alverson et al. (1999); however, the combined data set significantly rejected this topology (P = 0.002).

**Rates of molecular evolution**—Local clock analyses allowed us to identify the points along the tree's backbone where a rate change would maximize the likelihood of the data. Using a two-rate GTR + \Gamma model, the highest likelihood was obtained when the shift to a different rate of evolution was placed on the branch leading to Eumalvoideae (branch “E” in Fig. 1B). Under this model, the rate of evolution of Eumalvoideae was 4.09 times as fast as the rate of evolution in the remaining taxa (Table 1).

For the three-rate model, the optimal branch for the second change in rate was on the branch above *Pentaplaris* (branch “F” in Fig. 1B). In this case, the rate difference between Eumalvoideae and Bombacoideae was 9.66-fold, with the intervening taxa having an intermediate rate, 3.55-fold faster than Bombacoideae (Table 2).

Likelihood ratio tests showed that the three-rate model explains the data significantly better than the two-rate model (P < 0.001) and the one-rate (strict clock) model (P < 0.001). However, the data have a significantly higher likelihood under a non-clock model than under any of the local clock models (P < 0.001). This results shows that there are additional sources of lineage-to-lineage rate heterogeneity besides the tendency for faster rates of molecular evolution in eumalvooids than in other ingroup taxa. To evaluate whether this additional rate heterogeneity is concentrated in the slow, intermediate, or fast-evolving taxa we conducted a likelihood ratio test of clock-like evolution of each of these sets of taxa in isolation (with remaining taxa excluded). In all cases clock-like evolution was rejected (P < 0.001). Dividing the log-likelihood ratio of clock vs. non-clock models by the degrees of freedom (number of taxa minus two) provides a crude index of the degree of deviation from a clock. The intermediate-rate taxa have greater deviation from a clock (6.16) than either the slow- (1.48) or fast-rate (2.62) taxa. The strongly non-clock like evolution of the medium rate taxa could reflect either a gradual shift in rate, rate variation due to other factors, or topological errors in the ML tree.

One obvious topological error to consider is that the neotropical *Uladendron* could actually be sister to all core Malvoideae except *Pentaplaris*, rather than branching from within a grade of Australasian taxa. We repeated the three-rate local clock analyses with the ML tree under the constraint that *Uladendron* falls outside of the Australasian grade (Fig. 3). On this constrained topology, a higher likelihood score was obtained when *Uladendron* was assigned the slower rate than when it was assigned the intermediate rate typical of the Australasian grade (Table 2). Furthermore, under the local clock model, the constrained tree with *Uladendron* assigned the slowest rate has a higher overall likelihood score (–lnL = 14 999.26) than the unconstrained tree (–lnL = 15 000.67; Table 2). This means that, under a local clock model, the data favor the constrained over the unconstrained topology and, hence, support a topology in which *Uladendron* is sister to all core Malvoideae except *Pentaplaris*.

Using the three-rate model on the constrained topology, the relative rates of the slow, medium, and fast-evolving taxa are similar to those obtained on the unconstrained topology (Table 1).
We will here use the name Eumalvoideae to refer to the smallest clade that includes the type species of the three traditional tribes, Hibisceae, Malvae, and Gossypieae. This definition differs from the phylogenetic definition given previously (Baum et al., 1998). We prefer this modified definition because the previous one listed Lagunaria as a specifier and thus has ambiguous content based on current knowledge. Additionally, it seems most sensible to retain the name Eumalvoideae for the diverse core of the former Malvaceae rather than that core plus a partial subset of the early-diverging Australasian grade. We therefore propose redefining “Eumalvoideae.” Such a move is still permissible under the PhyloCode (www.ohiou.edu/phylocode) and is desirable because it avoids the need to coin a new name for this important clade.

The major lineages of Malvatheca—Our analyses of ndhF and matK confirm the overall pattern of relationships inferred for Malvatheca based on ndhF (Alverson et al., 1999) and rbcL + atpB data sets (Bayer et al., 1999). Specifically, our analysis confirms the monophyly of Malvatheca (BP = 98%; PP = 100%) and the existence of two major lineages: core Malvoideae (BP = 100%; PP = 100%) and core Bombacoideae (BP = 99%; PP = 100%). Core Malvoideae includes all traditional members of Malvaceae s.s. plus several taxa that have sometimes been assigned to different families. Our finding that Malvoideae includes Camptostemon (traditionally Bombacaceae) is consistent with prior studies (Alverson et al., 1999; Nyffeler and Baum, 2000). The inclusion of Pentaplaris in this clade was previously noted by Bayer et al. (1999), but was judged to be a possible artifact. Our study is the first to show that Uladendron falls within core Malvoideae.

Previous analysis of ndhF had suggested that Matisieae are sister to core Malvoideae (Alverson et al., 1999), but this result was only weakly supported (BP = 54%). This relationship was confirmed and strengthened by our analysis (BP = 83%; PP = 100%). Similarly, previous analysis of ndhF had only weakly supported a sister-group relationship between Ochroma and Patinoa (Alverson et al., 1999: BP = 54%), whereas our analysis provides much greater support (BP = 80%; PP = 96%).

Analyses of the sequence data failed to resolve the earliest branching events within Malvatheca. Specifically, there is no clear resolution of the relationships among five clades: core Malvoideae-Matisieae, core Bombacoideae, Ochroma-Patinoa, Fremontodendreae (Chiranthodendron-Fremontodendron), and Septotheca. Nonetheless a single, non-homoplastic, six base-pair deletion in a conserved region of matK suggests that Fremontodendreae are sister to the rest of Malvatheca. This arrangement occurs in 12.9% of trees in the Bayesian posterior distribution. Given the strength of this structural character and the lack of any contradiction from sequence data, we conclude that Fremontodendreae is sister to the remainder of Malvatheca.

Fremontodendreae are a small group comprising the three species of *Fremontodendron* and the monotypic *Chiranthodendron*. The clade has posed problems for taxonomists who most commonly placed it in Sterculiaceae (e.g., Hutchison, 1967) although its pollen resembles certain core Bombacoideae (Erdtman, 1952). That this clade should occupy a deep divergence within Malvatheca is consistent with biogeographic considerations: Fremontodendreae are centered in North America as contrasted with Bombacoideae and Malvoideae, both of which appear to have a South American center of origin (as discussed later).
**Phylogeny and biogeography of Bombacoideae**

The monophyly of core Bombacoideae is well supported by molecular data and by the presence of palmately compound leaves (of the non-peltate form; Kim et al., 2003) in almost all taxa (Alverson et al., 1999). Within core Bombacoideae, our data support a Catasetumma-Scleronema clade and place a clade comprising Gyranthera and Huberodendron as sister to the remainder of core Bombacoideae. We did not include *Ber- noullia* in this study, but based on prior analysis of ndhF (Alverson et al., 1999) and RFLP data (W. S. Alverson, unpublished data) it probably forms a clade with Gyranthera and Huberodendron. The flowers of Gyranthera, Huberodendron, and *Bernoullia* have elongated, sessile thecae, whereas all other members of core Bombacoideae have individually stalked, monothecate androecial units (except *Celba* and *Spirotheca*, which typically bear one or two pairs of thecae on each of five filamentous lobes at the summit of the staminal column). Given that *Chiranthodendron, Matisiaceae, Ochroma, Patinoa,* and *Seiptotheca* also have sessile thecae, our analysis suggests that sessile thecae are plesiomorphic for Bombacoideae and that stalked, monothecate units are a synapomorphy of a clade that includes Catasetumma, Scleronema, Adansonia, Celba, Bombax, and Pachira. Although not sampled here, morphological data and other molecular data (Alverson et al., 1999) suggest that this clade also includes six additional genera (Aguaria, Cavanillesia, Eriotheca, Neobuchia, Pseudobombax, and *Spirotheca*) and a total of approximately 160 species (Bayer and Kubitzki, 2003).

Resolution within Bombacoideae is weak given the taxon sampling used here, but adequate to show that the Old World taxa, *Bombax* (eight spp.) and *Adansonia* (eight spp.) are embedded in this otherwise neotropical clade. In order to evaluate whether this disjunction could be due to Gondwanan vicariance we compared the levels of sequence divergence with those reported for Alzateaceae (Myrtales), one of the very few tropical plant groups for which an Old World-New World disjunction has been attributed to Gondwanan vicariance (Conti et al., 2002, 2004; Schonenberger and Conti, 2003), although Moyle (2004) recently argued that, even in this case, dispersal rather than vicariance was likely responsible. The pairwise sequence divergences (using Hasegawa-Kishino-Yano distances) between *Alzatea* and its African sister-group averaged 0.035 (range 0.027–0.042) for matK and 0.038 (range 0.029–0.047) for ndhF. For the identical regions of these genes, the distance between *Bombax* or *Adansonia* (Old World) to *Celba* or *Pachira* (New World) averaged 0.012 (range 0.007–0.019) for matK and 0.013 (range 0.010–0.016) for ndhF. With distances less than half those seen for Alzateaceae, Gondwanan disjunction seems unlikely to explain the disjunctions seen in Bombacoideae.

Given that Gondwanan vicariance is unlikely, we infer that Bombacoideae originated in the Neotropics and later became dispersed to the Old World, either by sweepstakes dispersal or boreotropical migration. Further, because *Bombax* and *Adansonia* do not appear to be sister to each other, nor to the pa-leotropical species of *Pachira* that are sometimes segregated as *Rhodogynaphalol* (K. C. Walsh and D. Baum, unpublished data), three neotropical to boreotropical dispersal events are implied. Two additional cases in which single species occur in both the Neotropics and Paleotropics (*Celba pentandra* and *Pachira glabra*) most likely represent human dispersal (Robyns, 1960; Baker, 1965, 1983).

**An Australasian origin of Eumalvoideae**

—Pfeil et al. (2002) suggested that a widely distributed malvoid ancestor diverged into two lineages, the New World Matiseae and an Old World lineage corresponding to the traditional Malvaeae s.s., and that this divergence corresponded with the break-up of Gondwana. Under this scenario, the Old World lineage gave rise to Australasian taxa, such as *Camptostemon, Howittia, Lagunaria,* and *Radysya,* along with the Eumalvoideae, which then experienced an extensive global radiation. By including a broader sampling of early-diverging Malvoideae and with additional sampling among non-malvoid Malvatheca, we were able to shed further light on the biogeographic context of the mallow radiation and test the Gondwana disjunction hypothesis.

The neotropical Matiseae and *Pentaplaris* are well supported as early-branching members of Malvoideae. Furthermore, given that Bombacoideae (see Phylogeny and biogeography of Bombacoideae), Fremontodendraceae, *Ochroma, Patinoa,* and *Seiptotheca* are centered in the New World, the Eumalvoideae clearly arise from a paraphyletic grade that comprises at least four distinct New World lineages. Although this topology does not absolutely refute the Gondwana vicariance hypothesis, combined with other data it renders it less probable than a dispersalist explanation. First, most cases of a southern Gondwanan distribution (e.g., *Nothofagus*, Manos, 1997; *Araucaria*, Sequeira and Farrell, 2001) involve south-temperate groups, whereas Malvatheca are fundamentally tropical in origin. Second, vicariance is most likely to result in disjunct sister clades, whereas, without invoking additional extinction events, dispersal is more likely to result in the pattern we observed: a clade with one center of origin nested within a grade of lineages from the other biogeographic zone. Third, the palynological evidence (summarized in Pfeil et al., 2002) suggests that Malvoideae are a relatively young group, with no definitive pollen records prior to the late Eocene. Although some Australian/South American connections are suggested to have persisted via Antarctica until the early Oligocene (Dingle and Lavelle, 2000), it is unclear if these represented suitable migration routes for tropical-adapted taxa. Thus, overall, the biogeographic data suggest the need to consider alternative scenarios to Gondwana disjunction to explain the Australasian origins of Eumalvoideae.

It is noteworthy that the early-branching Australasian taxa (Howittia, Radyera, Camptostemon, Lagunaria, and Alyogyne) are mangroves or small sublittoral trees. Such taxa are typically well adapted to oceanic dispersal and often have extremely broad geographic ranges. This is well illustrated by *Radyera* with disjunct species in South Africa and Australia, and *Lagunaria,* which is present on the isolated Norfolk and Lord Howe Islands. This raises the possibility that a South American lineage acquired the mangrove habit and then migrated over water to Australasia and/or Southeast Asia where it radiated into a grade of littoral forms. Sometime later, this nexus gave rise to the major radiations of Malvaeae, Glossypieae, and Hibisceae. Spiny pollen occur in the South American *Pentaplaris* (Bayer and Dorr, 1999) and *Uladendron* (Marciano-Berti, 1971), thus, pending further examination of the palynomorphs in question, it is plausible that “Hibiscus”-type pollen reported from South America 44–39 million years ago (Mya) (Muller, 1981) could provide an approximate date for the origin of the ancestral mangrove (but see Pfeil et al., 2002). In that case the first records in Australia (37–35 Mya; Muller, 1981) would provide a minimum date for the dispersal
of this lineage across the Pacific. Future analysis using molecular dating methods could clarify the temporal context for the diversification of Malvatheca and Eumalvoideae.

The fly in the ointment is *Uladendron*, a neotropical rainforest tree endemic to a small area of forest in central Venezuela, that appears on the unconstrained ML and MP trees to be more closely related to Eumalvoideae than at least *Radyera*, one of the Australasian taxa. Such a topology would tend to suggest either two of the origins of the mangrove habit and two dispersals westward across the Pacific or one migration westward but then a back-migration and reestablishment of the rainforest tree habit. However, such unparsimonious scenarios need not be invoked if, in fact, *Uladendron* branches off after *Pentaplaris* as sister to the rest of core Malvoideae. Indeed, such a position is only one step longer than the most-parsimonious topology and cannot be rejected by a Templeton test. Likewise, examination of the Bayesian posterior distribution shows that 0.05% of the trees place *Uladendron* sister to all core Malvoideae except *Pentaplaris*. The small amount of the posterior distribution showing this topology would constitute rejection of this position for *Uladendron* at the $P = 0.005$ level if a uniform prior were assumed on all tree topologies. However, in this case, one can validly attach a higher prior probability to trees that do not have the neotropical *Uladendron* embedded within the Australasian grade and, as a result, the observed posterior probability is not necessarily adequate to reject the alternate position of *Uladendron*. Even more telling, under a three-rate local clock model, trees that embed *Uladendron* in the Australasian grade actually have lower likelihood than trees that place *Uladendron* sister to all core Malvoideae except *Pentaplaris*. Consequently, our data are compatible with *Uladendron* being the closest living relative of a mangrove taxon that migrated to Australasia and, hence, to a single trans-Pacific migration event.

**Rates of molecular evolution**—Simple observation of phylogenograms was sufficient to suggest accelerated molecular evolution within the Malvoideae relative to Bombacoideae (Alverson et al., 1999). Here we were able to quantify the changes in rate and to localize them to particular branches. If one assumes a two-rate model, the most significant change in rate occurred on the branch that subtends Eumalvoideae (Fig. 1B, Branch E). If one allows for two rate changes (three rates) within Malvatheca, as supported by likelihood ratio tests, the second rate change is placed on the branch above *Pentaplaris* (and *Uladendron* on the constrained tree), assigning all the Australasian taxa to an intermediate rate category. According to the three-rate model, the Eumalvoideae have experienced a greater than nine-fold faster rate of molecular evolution than Bombacoideae, with the intermediate taxa having an intermediate rate (ca. 3.5-fold faster).

Several potential mechanisms have been proposed to explain accelerated molecular evolution, including changes in mutation rate, generation time, effective population size, and the rate of species diversification (Barraclough and Savolainen, 2001). Under the maximum likelihood three-rate model (on trees with *Uladendron* constrained to branch just above *Pentaplaris*), the first rate acceleration coincides exactly with the inferred transition from large rainforest trees to a shrubby, possibly mangrove habit. This transition probably resulted in a change in generation time and population size but apparently did not involve an accelerated rate of species diversification because lineages in the mangrove grade include few species.

In contrast, the second rate jump coincides with the Eumalvoideae radiation making a change in diversification rate a plausible causal factor. Nonetheless, because the Eumalvoideae tend towards an even shrubbier and in some cases, a herbaceous and/or weedy habit, changes in life history (including population size) could have played a role in both the increased rates of diversification and molecular evolution. Further work incorporating additional genes and taxa may shed light on the mechanisms that have contributed to the mallow species radiation and to the group’s accelerated molecular evolution.

**LITERATURE CITED**


