Paralogy and Orthology in the Malvaceae *rpb2* Gene Family: Investigation of Gene Duplication in *Hibiscus*

B. E. Pfeil,†[1] C. L. Brubaker,* L. A. Craven,* and M. D. Crisp†*

*CSIRO Plant Industry, Canberra, Australia; †Australian National University, School of Botany and Zoology, Canberra, Australia

A sample of the second largest subunit of low-copy nuclear RNA polymerase II (*rpb2*) sequences from Malvaceae subfamily Malvoideae suggests that *rpb2* has been duplicated early in the subfamily's history. *Hibiscus* and related taxa possess two *rpb2* genes, both of which produce congruent phylogenetic patterns that are largely concordant with cpDNA topologies. No evidence of functional divergence or disruption was found among duplicated copies, suggesting that long-term maintenance of duplicated copies of *rpb2* is usual in this lineage. Therefore, this gene may be suitable for the potential diagnosis of relatively old polyploid events. One probable pseudogene was found in *Radyera farragei* and a single chimeric sequence was recovered from *Howittia trilocularis*, suggesting that the *rpb2* locus is not as prone to evolutionary processes that can confound phylogenetic inferences based on nDNA sequences. The pattern of relationships among *rpb2* sequences, coupled with chromosome number information and Southern hybridization data, suggests that an early polyploid event was not the cause of the duplication, despite independent evidence of paleopolyploidy in some members of Malvoideae. *Rpb2* exons and introns together are suitable for phylogenetic analysis, producing well-resolved and well-supported results that were robust to model permutation and congruent with previous studies of subfamily Malvoideae using cpDNA characters.

Introduction

Malvaceae s.l. subfamily Malvoideae (= Malvaceae s.s [*sensu* Bayer et al. {1999}]) is a mostly tropical group of plants comprising three main species-rich lineages: tribes Gossypieae and Malveae and a group of taxa nested within the large genus *Hibiscus* (Pfeil et al. 2002). The subfamily shows a large diversity of chromosome numbers (Bates 1968; Bates and Blanchard 1970; Fryxell 1988). Base chromosome numbers in the tribe Gossypieae have a limited range (n = 10 to 13 [Fryxell 1968, and references therein]), whereas those in *Hibiscus* and related genera (hereafter, *Hibiscus* s.l. [see Pfeil et al. {2002}]) are more variable (n = 12 to ~144, commonly n = 14 or 18 [Fryxell 1968, 1988, and references therein]). Malveae also displays a wide chromosome number diversity and contains the lowest haploid numbers seen in the subfamily (n = 5 to 21 [Bates 1968; Bates and Blanchard 1970]). However, the high chromosome numbers seen in generic groups in Malveae are often exact multiples of, or additions of, the base numbers found in those groups (e.g., the *Sida* alliance [*sensu* Fryxell {1988}] has n = 7, 8, 16, and 21). This suggests that several independent rounds of ploidy increase have taken place from these low base numbers subsequent to the divergence of the generic groups.

A recent allopolyploid event has taken place 0.5 to 2 MYA in *Gossypium* and has been the focus of intensive investigation (Wendel 1989; Reinisch et al. 1994; Small et al. 1998; Brubaker, Paterson, and Wendel 1999; Cronn, Small, and Wendel 1999). Although *Hibiscus* also contains known polyploids, those best understood (in *Hibiscus* section *Furcaria*) have not received the same attention (Wilson 1994). However, "diploid" species in *Gossypium* appear to have undergone an early ploidy increase, which may be as old as 20 to 40 Myr, probably predating the divergence of genera in this tribe (Reinisch et al. 1994; Seelanan, Schnabel, and Wendel 1997; Brubaker, Paterson and Wendel 1999). The evidence for this rests on three observations: (1) high haploid chromosome numbers (13 in diploid *Gossypium* and 12 to 13 in most members of the tribe [Webber 1934; Fryxell 1968]), (2) mapping studies in which diploid and allopolyploid subgenomes contain nested duplications (Reinisch et al. 1994; Brubaker, Paterson, and Wendel 1999), and (3) a BrdU-Hoechst-Giemsa banding analysis that revealed similarity between six nonhomologous pairs of diploid chromosome pairs (i.e., suggesting an ancestral haploid chromosome number of seven) in two Gossypium A-type genomes (Muravenko et al. 1998).

As part of an ongoing study of the phylogeny of Malvaceae subfamily Malvoideae, we wished to test hypotheses of allopolyploidy in *Hibiscus* using a low-copy nuclear gene. However, characterization of low-copy nuclear markers is advisable to determine orthologous versus paralogous relationships among sequences because of the possibility of gene duplication followed by selective gene silencing or loss (Wendel and Doyle 1998). Gene duplication cause by polyploidy may be prevalent in plant nuclear gene families, given the high estimates of polyploidy in some plant groups (e.g., angiosperms [Masterson 1994]), although it is not the only cause of duplication (Wendel 2000, and references therein). Concerted evolution among duplicated loci may obscure comparable phylogenetic patterns among gene lineages produced by gene duplication and is another process that needs to be taken into account when examining data from nuclear regions.

This paper reports the results of an initial survey of the main lineages in Malvaceae subfamily Malvoideae for the gene that encodes the second largest subunit of RNA polymerase II. This enzyme is responsible for the synthesis of pre-mRNAs, and its second largest subunit (*rpb2*) in *Saccharomyces cerevisiae* (yeast) contains two known

[1] Present address: Department of Plant Biology, Cornell University, Ithaca, New York.

Key words: gene duplication, RNA polymerase II, *Hibiscus*, phylogenetics, evolution, low-copy nuclear DNA.

E-mail: bep27@cornell.edu.

Mol. Biol. Evol. 21(7):1428–1437. 2004
doi:10.1093/molbev/msh144
Advance Access publication April 14, 2004

Table 1
Taxa Used in This Study

Species	Voucher	*rpb2* Sequence GenBank Number[a]
Ingroup		
Abutilon fraseri (Hook.) Walp.	BEP[b] 262	AY463782
Alyogyne cravenii Fryxell	Fryxell, LAC[c], and Stewart 4870	AY463746, 463762–463768
A. huegelii (Endl.) Fryxell	ANBG[d] 9800039c	AY463717–463722
A. pinoniana (Gaudich.) Fryxell	LAC and BEP 10339	AY463723–463728
Gossypium sturtianum J. H. Willis	Gos[e]-5071	AY463706–463710
Hibiscus macrophyllus Roxb.	LAC 10202	AY463754–463756, 463787, 463788, 463793
H. pedunculatus L. f.	PI[f] 364903	AY463747, 463757–463761
H. surattensis L.	PI 585138	AY463753, 463785, 463786
H. trionum L.	PI 500697	AY463742–463745, 463783, 463784
Howittia trilocularis F. Muell.	ANBG 08910071	AY463731–463733, 463735, 463736
Lagunaria patersonia (Andrews) G. Don	ANBG s.n.	AY463734, 463737–463741
Lavatera arborea L.	Slee 2395	AY463775
Malva neglecta Wallr.	Local weed (n.v.[g])	AY463729–463730
Modiola caroliniana G. Don	CPI[h] 142480	AY463776–463781
Pavonia hastata Cav.	Purchased (n.v.)	AY463749–463750, 463789–463792
Radyera farragei (F. Muell.) Fryxell and S. H. Hashmi	Fryxell, Craven, and Stewart 4462	AY463795–463796
Sida acuta Burm. f.	BEP 327	AY463769–463774
Thespesia thepesioides (R. Br. ex Benth.) Fryxell	n.v. (same DNA as Seelanan, Schnabel, and Wendel [1997])	AY463711–463716
Urena lobata L.	CLB[i] 1451	AY463748, 463751, 463752, 463794
Outgroup[j]		
Fremontodendron californicum (Torrey) Cov. × *mexicanum* Davidson	BEP 339	AY463701–463705
Melhania sp.	CS[k] 24	AY463698–463700

NOTE.—All vouchers were deposited at the Australian National Herbarium (CANB), except living collections which are part of the Australian National Botanical Gardens (ANBG).

[a] GenBank numbers in brackets have been previously published.

[b] BEP = B.E. Pfeil.

[c] LAC = L.A. Craven.

[d] ANBG = Australian National Botanic Gardens living collections.

[e] Gos = CSIRO *Gossypium* germplasm collection.

[f] PI = U.S.D.A. Plant Introduction.

[g] n.v. = Not vouchered.

[h] CPI = CSIRO Plant Introduction.

[i] CLB = C.L. Brubaker.

[j] Outgroup GenBank numbers for the exon-only analysis are shown in figure 1.

[k] CS = C.L. Brubaker DNA number.

functional domains: a purine nucleotide-binding domain, and a zinc-complexing domain (Sweetser, Nonet, and Young 1987). This gene is found in a single copy in the haploid yeast genome and encodes a polypeptide that contains several blocks of amino acid sequences that are relatively conserved between yeast and the product of the *rpoB* gene, a prokaryotic RNA polymerase subunit found in *Escherichia coli* (Sweetser, Nonet, and Young 1987). At least some of these conserved blocks overlap blocks of amino acids shared among eukaryotes (the latter defined in Denton, McConaughy, and Hall [1998]).

rpb2 has been used in four phylogenetic studies of plants (Denton, McConaughy, and Hall 1998; Oxelman and Bremer 2000; Popp and Oxelman 2001; Oxelman et al. 2004). Denton, McConaughy, and Hall (1998) found no evidence for the presence of multiple copies in the 11 species they sampled. However, Oxelman and Bremer (2000) reported the presence of two copies in Gentianales, one of which did not contain introns in the portion of the gene that they studied. Oxelman and Bremer (2000) also cited unpublished research that found two copies in *Lycopersicon* and *Rhododendron*. Popp and Oxelman (2001) demonstrated the utility of *rpb2* in the diagnosis of a relatively recent polyploid and did not find any recombinant sequences in the portion of *rpb2* used in that study. Oxelman et al. (2004) found that the duplication of copies in asterids appears to have occurred near the origin of the core eudicot group, inferring that the gene must have been lost in the rosids they investigated.

The aims of this study are to (1) determine whether *rpb2* has been duplicated in Malvaceae, (2) if duplicated, to document the patterns of gene evolution and the relationships among and between copies of this gene, (3) determine whether this gene is suitable for the diagnosis of polyploid events in Malvaceae, both relatively recent and ancient, and (4) determine whether the rate of sequence change in the 700-bp part of *rpb2* examined here is suitable for phylogenetic inference in Malvaceae subfamily Malvoideae, a group that appears to have radiated about 40 MYA.

Materials and Methods

Taxon Sampling

The taxa sampled (table 1) represent the primary Malvoideae lineages identified from analysis of chloroplast

DNA (cpDNA) sequences (*Alyogyne*, Gossypieae, Malveae, *Hibiscus* s.l., *Howittia+Lagunaria*, and *Radyera* [Pfeil et al. 2002]). We used a close outgroup, *Fremontodendron* (in subfamily Bombacoideae [Bayer and Kubitzki 2002]), which, although closely related to subfamily Malvoideae, is outside it according to several cpDNA-based analyses (Alverson et al. 1999; Bayer et al. 1999; Pfeil et al. 2002). We also used a more distant outgroup, *Melhania* (in subfamily Dombeyoideae [Bayer and Kubitzki 2002]). Although this taxon has not been sampled in previous cpDNA analyses, other members of its subfamily have been (e.g., *Dombeya*) and are more distantly related to subfamily Malvoideae than is *Fremontodendron* (Alverson et al. 1999; Bayer et al. 1999).

Rpb2 Cloning and Sequencing

PCR, using the plant-specific *rpb2* primers P6F and P7R (Denton, McConaughy, and Hall 1998), produced an approximately 1300-bp band. Direct sequencing of this band produced readable sequences for diploid *Gossypium* but not for several species of *Hibiscus*. The *Gossypium* genomic DNA sequences were compared with cDNA sequences from *Arabidopsis* (GenBank accession number Z19120) to design an internal forward-reading primer (situated approximately half way between P6F and P7R). This primer, designated P6FB (with the 5′ to 3′ sequence ACA CTG AAG TTG GTG TTG TTC G/T), begins at nucleotide 1802 in *Arabidopsis* (GenBank accession number Z19120) and amplifies an approximately 700-bp product. It was used in all subsequent PCR reactions in combination with P7R. Primers P6FB and P7R capture three complete introns, all of two exons, and the 5′ and 3′ ends of two flanking exons. PCR conditions were 92°C for 1 min, then five cycles at 55°C (30 s to 1 min), 72°C (50 s to 1 min), and 92°C (1 min); then five cycles with a touchdown from 55 to 53°C (−0.4°C each cycle), 72°C (1 min adding 2 s each cycle), and 92°C (1 min); then 20 cycles at 53°C (1 min), 72°C (70 s), and 92°C (1 min). Finally a 5 min extension step at 72°C was done.

PCR products were cleaned using a Microcon PCR filter (Amicon) to increase ligation efficiency, then cloned using a T-A cloning kit with JM109 competent cells according to the manufacturer's protocol (pGEM-T vector system, Promega). After colonies were grown on a nutrient medium containing XGAL and IPTG, blue/white screening of the colonies was done. Positive colonies were harvested and mixed with 15 μl TE and replicated on an archival LB-ampicillin plate. Two microliters of the TE colony suspension were then utilized as a PCR template using the universal M13 forward and reverse primers for PCR screening of the insert size and subsequent sequencing of colonies with the inserted region. Up to six clones were forward and reverse sequenced to control for error in *Taq* polymerase (PerkinElmer) replication of the target sequence and to identify allelic and homeologous versions of the gene. Whereas most DNA was extracted from single individuals, some DNA was extracted from multiple progeny (because of limited material in some cases). Therefore, multiple alleles belonging to either heterozygous individuals or multiple individuals were occasionally detected.

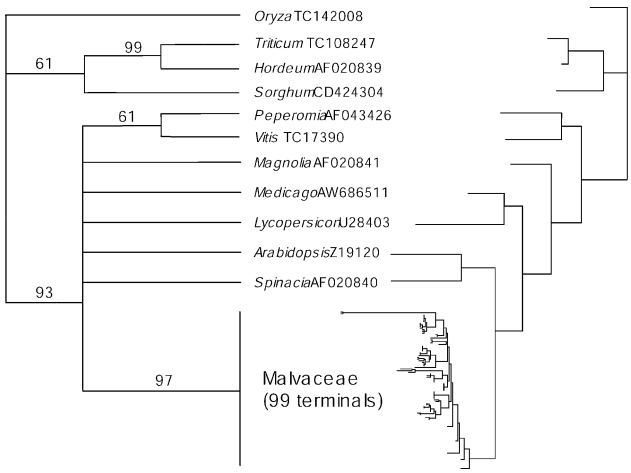

FIG. 1.—Bootstrap consensus (left) and one most-parsimonious phylogram (right) of 385 bp of *rpb2* exon sequences from individual clones (in Malvaceae) and other sequences (from GenBank). Bootstrap percentages above 50% are shown above branches (left). The trees are unrooted.

To avoid making a priori orthology and paralogy assessments, individual clones rather than consensus sequences have been used in these analyses; however, this means that *Taq* errors may be present in some sequences. *Taq* polymerase error in PCR occurs at a frequency of 0.27 to 0.85×10^{-4} point mutation errors per bp per cycle; this translates to an expectation that 0 to 2 nucleotide differences between clones from the same PCR reaction may be attributed to *Taq* error and sequences that differ more substantially can reasonably be inferred to arise from alleles, orthologs, or paralogs. Furthermore, because *Taq* errors that cause substitutions are mostly random (although indels may not be [Bracho, Moya, and Barrio 1998]), it is unlikely that any two sequences would share identical *Taq* errors to create false synapomorphies.

Alignment and Analysis

The first phylogenetic analysis was done to place the Malvaceae *rpb2* sequences in a broader context and to test the monophyly of these sequences. Sequences from GenBank were added to those gathered here (fig. 1). Exons alone were used because introns could not be reliably aligned between sequences from Malvaceae and other families.

rpb2 exon sequences were aligned by eye. The 385-bp alignment required the inference of only a single 3-bp indel. These sequences were analyzed under maximum parsimony using PAUP* 4.10b (Swofford 1998), with 100 random addition sequence (RAS) replicates, keeping 100 trees per replicate and swapping to completion with a 10,000-maxtrees buffer. Bootstrap (BS) percentages are based on 200 BS replicates, with 10 RAS replicates, keeping 10 trees per RAS replicate. Analysis under differing models of evolution was not conducted, because the results clearly indicated that Malvaceae *rpb2* exon sequences are more closely related to each other than to other plant *rpb2* sequences.

Distances from each ingroup sequence to an outgroup (*Spinacia*) were calculated using the *p*-distance function in PAUP* (Swofford 1998) to detect rate heterogeneity that may be associated with loss of function in some gene copies. In-frame codons (relative to the known reading frame in *Arabidopsis*) were examined in the ingroup sequences to check for the presence of premature stop codons, again to check for loss of function. Amino acid substitutions shared by several clones were examined among those taxa with more than one *rpb2* copy. The types of substitution and location of changes may indicate whether functional divergence or functional disruption may be occurring (if significant amino acid substitutions occur in functional domains) or whether long-term maintenance of multiple copies with the same function has occurred in these taxa (if no significant changes occur, or changes are confined to regions that may not be critical to function).

The second phylogenetic analysis conducted used exons and introns (810 bp) solely from Malvaceae taxa (with *Radyera* excluded [see below]), with 54 indels included at the end of the matrix. Known mechanisms of mutation were considered when coding indel characters (see Kelchner [2000]). Analyses under maximum-parsimony and BS resampling, using *Melhania* as the outgroup, were conducted as above. Although *Melhania* was not sister to all other Malvaceae in the exon-only analysis (details not shown), there is no meaningful BS support for this result. The taxonomic position of *Melhania* (as mentioned above) suggests that it is an appropriate outgroup choice. Midpoint rooting of trees based on exons and introns (while not completely reliable) is nonetheless consistent with this assumption. An additional analysis using likelihood allowing for separate patterns of evolution for coding versus noncoding regions versus indels (with the GTR+gamma model for the sequences, with a single substitution class+gamma for the indels, and with separate parameter estimates for each of the partitions) was conducted using Bayesian methods (Huelsenbeck et al. 2001, and references therein), implemented in Mr Bayes version 3.0 (Huelsenbeck and Ronquist 2001) with flat prior probabilities. 500,000 generations in four chains were run, sampling every 10 generations. After examining where the likelihood scores stabilized, the first 5,000 samples were discarded as the burn-in phase. This analysis was done to see whether the clades found in the parsimony analysis were robust to a different model of sequence evolution. The indels were analyzed with only a single substitution class (they were coded using nucleotides in the matrix), as the relative probabilities of indel changes are not as well understood as nucleotide substitutions. GTR was arbitrarily chosen for analysis of the sequences. Among site rate heterogeneity is a common feature of molecular data, both of introns and coding regions (personal observation), therefore gamma was included in the model for each partition in the Bayesian analysis.

During the initial alignment, a chimeric sequence of *Howittia trilocularis* was found by eye and confirmed using separate analyses of the first 290 bp versus the remainder of the sequence, without indels encoded. This sequence was excluded from subsequent analyses.

Southern Hybridizations

DNA (10 μg) from taxa representing each major clade (*Gossypium hirsutum* [tetraploid n = 26], *G. arboreum* [diploid n = 13], and *G. hirsutum* × *barbadense* [tetraploid n = 26], *Alyogyne huegelii* [n = 32], *Alyogyne cravenii* [n = unknown], *Modiola caroliniana* [diploid n = 9], *Malva neglecta* [hexaploid n = 21], *Radyera farragei* [n = 18], *Fioria vitifolia* [n = 16 and 17] and *Hibiscus pentaphyllus* [n = 18]) were digested with four restriction enzymes (BamHI, NcoI, SacI, and XhoI) and transferred to a positively charged nylon membrane. The membrane was probed with a mixture of *rpb2* PCR fragments amplified from the taxa on the membrane and washed at high stringency to maintain specificity for *rpb2* and no other RNA polymerases.

Synonymous and Nonsynonymous Rates from Ingroup-Outgroup Comparisons

Synonymous (Ks) and nonsynonymous (Ka) rates of substitutions were estimated by the weighted method of Yang and Nielsen implemented in PAML version 3.0 (Yang 2000) by pairwise comparisons of Malvoideae sequences to a *Melhania* sequence. Only one clone per paralog comparison was made so as not to bias the comparison toward those taxa with larger numbers of clones per paralog.

Genes maintained by purifying selection are expected to show Ka/Ks less than 1, whereas noncoding sequences (i.e., pseudogenes) are expected to have Ka/Ks ratios that approach 1. Ka/Ks ratios were examined to identify clones that were violating expected Ka/Ks ratios for coding sequences. As this method uses an outgroup comparison, any differences seen in the Ka/Ks ratio should reflect differences since the divergence of the paralogs.

Results

Malvaceae rpb2 Sequences Are Monophyletic

Of the approximately 700 bp of *rpb2* sequenced, 385 bp were partial or whole exons. The exons from Malvaceae species alone contained 33.5% parsimony informative sites and 39.5% total variable sites.

A bootstrap consensus tree (fig. 1 [left]) shows good support for the monophyly of all the Malvaceae sequences (BS 97%). A representative phylogram (one of the most parsimonious trees) illustrates the sequence similarity among Malvaceae *rpb2* exons relative to other flowering plants (fig. 1 [right]).

Some Malvaceae *rpb2* Sequences May Originate from Pseudogenes

The distribution of *p*-distances between Malvaceae and *Spinacia* exon sequences identified two sequences from *Radyera* that are evolving at a much faster rate than are other Malvaceae (fig. 2). These two sequences also contain premature in-frame stop codons; all the other sequences had open reading frames, except three *Alyogyne huegelii* sequences (discussed below). A BlastN search (Altschul et al. 1997) on GenBank nevertheless found that

the *Radyera farragei* sequences still matched *rpb2* sequences from other angiosperms more closely than any other region ($E = 2 \times 10^{-7}$ for several parts of *rpb2* cDNA from *Spinacia*), ruling out unintentional amplification of other members of the RNA polymerase gene family. Together, these results indicate that the *rpb2* sequences recovered in *Radyera* represent a gene that has lost its function and become a pseudogene.

Three clones of *Alyogyne huegelii* (2, 4, and 5) contain a 14-bp deletion relative to the other clones from that species and to all other sequences in the study (except *Radyera*, where the difference is 17 bp). This deletion caused a frame shift in these sequences, with the presence of numerous downstream premature stop codons. Without this deletion, and aligned with the rest of the matrix, these sequences would not contain premature stop codons. The other clones from this species (1, 3, and 6) do not contain this deletion, and the sequence differences among clones with and without the deletion are otherwise unremarkable. No obvious rate shift is present in *A. huegelii* clones 2, 4, and 5 compared with all other exon sequences examined. It is not clear whether this deletion is indicative of a copy of *rpb2* that has lost its function more recently than the *Radyera* sequences or whether this deletion is a PCR artifact.

One *Howittia trilocularis* Clone Is Chimeric

A single *Howittia trilocularis* clone (clone 3) occupied two well-supported alternative positions in separate parsimony analyses of the first 290 bp versus the remaining 379 bp (not shown; BS support greater than 98% in each case). This sequence clearly combines the first part of a sequence very similar to clones 1, 6, and 7 (up to and including position 242) with the second part of a sequence very similar to clones 10 and 12 (from position 299). The intervening sequence does not contain synapomorphies for either group of clones, so the exact position of the recombination is unclear, although it is contained within the second exon. No autapomorphic A/T insertion was observed within the intervening sequence, which is often, but not always, observed if fragmented DNA was causing PCR mediated recombination (Pääbo, Irwin, and Wilson 1990). However, template mixtures can show PCR-mediated recombination without autapomorphic A/T insertions (e.g., Popp and Oxelman [2001]). Therefore, the chimeric sequence pattern observed in *Howittia* clone 3 may be caused by either intergenic recombination (forming a new locus) or PCR-mediated recombination.

The *rpb2* Locus Is Duplicated in Some Malvaceae

The entire *rpb2* fragments could be aligned among the Malvaceae representatives and the two closely related outgroups (*Melhania* and *Fremontodendron*). The final alignment was 810 bp in length, and to this, 54 indel characters were appended before analysis. This alignment had 37.5% parsimony-informative sites and 43.3% total-variable sites (excluding *Melhania* and *Fremontodendron*, 34.6% and 39.5% were parsimony-informative and total-variable sites, respectively). The number of parsimony-informative sites differed markedly between exons and introns (with coded indels in the latter). The Malvaceae exons (excluding *Radyera* sequences) contained 21.3% informative sites, whereas the introns and coded indels contained 50.5% informative sites.

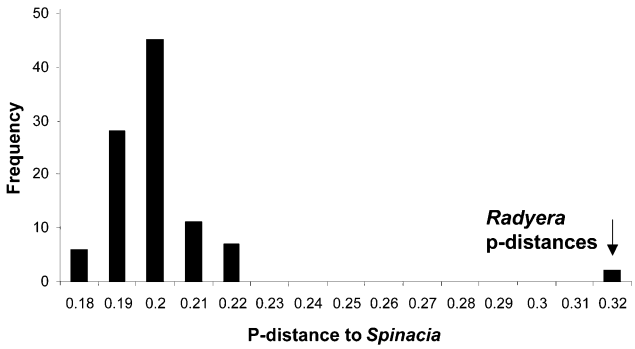

FIG. 2.—Distribution of uncorrected *p*-distances from each ingroup exon sequence to one of the outgroup sequences (*Spinacia*). *p*-Distances calculated using PAUP* (Swofford 1998).

The exon and intron parsimony analysis found 1,915 shortest trees of length 851 steps with high internal consistency (CI = 0.65 or 0.62 with or without uninformative characters, and RI = 0.90), the strict consensus of which is shown in figure 3. Bootstrap and Bayesian posterior probabilities (PP) are also shown in figure 3 and are approximately correlated. Only one node (grouping *Alyogyne cravenii* 3 and *Hibiscus pedunculatus* 5, with PP 85%) not found in the parsimony strict consensus was present in the Bayesian posterior probability majority rule tree (latter tree not shown). Otherwise, results differ only in the degree of support found under the alternative models. The results found in the parsimony analysis, therefore, appear to be robust to this model permutation.

Several taxa had distinctive sequence types that were not sister to each other (including all *Hibiscus* species sampled) and formed two well-supported clades with similar internal phylogenetic pattern (clades 1 and 2 in figure 3). Included is *Alyogyne cravenii*, which is now thought to belong in *Hibiscus* rather than *Alyogyne* (Pfeil and Craven, unpublished data). This is consistent with the duplication of the *rpb2* locus in these taxa. That both of the duplicated *rpb2* loci are unambiguously nested within the Malvaceae (BS 97% [fig. 1]) demonstrates that duplication (regardless of mechanism) postdates the divergence of Malvales from the other angiosperms and is not linked to the other known plant *rpb2* duplications (Oxelman and Bremer 2000; Popp and Oxelman 2001; Oxelman et al. 2004).

The *Lagunaria* and *Howittia* sequences also formed two clades, each of which contains sequences of *Lagunaria* sister to those of *Howittia*, demonstrating that *rpb2* is duplicated in these taxa as well.

Sequences from the tribe Gossypieae (*Gossypium* and *Thespesia*), tribe Malveae (*Abutilon*, *Lavatera*, *Malva*, *Modiola*, and *Sida*), and *Alyogyne* (*A. huegelii* and *A. pinoniana*) did not display multiple sequence types falling into different well-supported clades. In fact, these taxa formed a clade in the strict consensus tree (clade 3 [fig. 3])

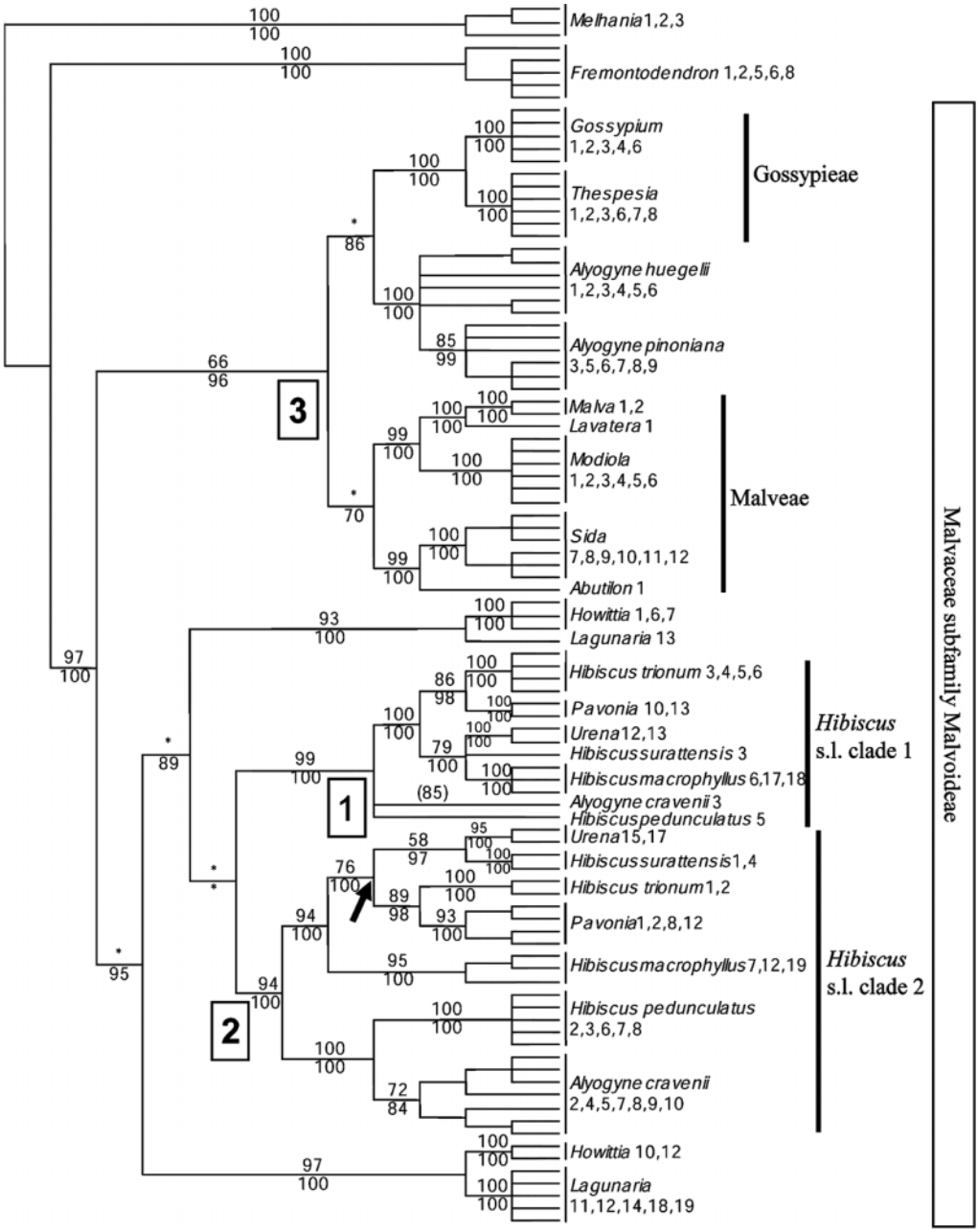

FIG. 3.—Maximum parsimony strict consensus of 1915 trees after analysis of 810 bp (aligned) of exon+intron Malvaceae *rpb2* sequences with an additional 54 coded indels, rooted using *Melhania* as the outgroup. Bootstrap percentages above 50% are shown above branches and posterior probabilities above 50% are shown below branches (asterisks mark where these values fall below 50%). Clade 1 = *Hibiscus* s.l. *rpb2* copy 1. Clade 2 = *Hibiscus* s.l. *rpb2* copy 2. Clade 3 = taxa with only a single *rpb2* copy recovered using PCR methods (Gossypieae, Malveaem, and *Alyogyne*). Terminal clones from the one species are grouped by a vertical line adjacent to the taxon name, which is followed by clone numbers. *Hibisucs* and *Alyogyne* epithets are included (as these are the only genera with multiple species). Other genera are represented by a single species, and their epithets can be found in table 1. *Fremontodendron* is placed in subfamily Bombacoideae, whereas *Melhania* is placed in subfamily Dombeyoideae (both *sensu* Bayer et al. [1999]).

to the exclusion of the remainder of Malvaceae subfamily Malvoideae taxa (although with less than 80% BS support, but greater than 95% PP).

The relationships found within clades 1 and 2 are nearly identical to those found in cpDNA (Pfeil et al. 2002, and unpublished data). One minor exception is marked by the arrow in figure 3. *Hibiscus macrophyllus* was expected to be sister to *Urena*+*H. surattensis* based on cpDNA, but a clade of *H. trionum*+*Pavonia* occupied that position.

Amino Acid changes

Nine amino acid changes among the taxa containing duplicated copies of *rpb2* (*Lagunaria*, *Howittia*, and *Hibiscus* s.l.) occurred in more than one clone. Five of

these amino acid positions showed unambiguous change among either the *Lagunaria* and *Howittia* copies (position 86) or among the *Hibiscus* s.l. copies (positions 43, 46, 50, 54, and 86). Comparison of yeast and *Arabidopsis* amino acid sequences revealed that the position of these changes is within a less conserved portion of the *rpb2* gene (a portion where amino acid similarity is 48% or less [not shown]). Of the six copy-specific changes at five sites, only two showed a shift between amino acid classes: position 43 changes from basic histidine to hydrophobic leucine, and position 46 changes from polar glutamine to basic histidine. No changes from polar to hydrophobic or from acidic to basic (and vice versa) were observed.

Southern Hybridizations

When genomic DNA was probed with the *rpb2* fragment sequenced here, multiple bands consistent with multiple loci were seen in several species from the *Hibiscus* s.l. lineage (*Fioria vitifolia*, n = 16 to 17; *H. pentaphyllus*, n = 18; and *A. cravenii*, n = unknown), although the actual locus number could not be unambiguously determined. *Radyera* also appears to have more than one *rpb2* locus, despite only a single apparently nonfunctional copy being recovered by PCR methods. It is extremely unlikely that *Radyera* would not have a functional *rpb2* gene, as this gene is thought necessary for cell viability (Sweetser, Nonet, and Young 1987).

The Southern data do support the inference of a single *rpb2* locus for the Gossypieae and for *Alyogyne*. Diploid (n = 13) and tetraploid (n = 26) *Gossypium* species contain one and two fragments, respectively, consistent with the presence of a single *rpb2* per genome. *Alyogyne huegelii* (n = 32) appears to have two *rpb2* loci, consistent with the assumption that it is a tetraploid.

The observations in two Malveae species are more complex. Diploid *Modiola* (n = 9) and hexaploid *Malva* (n = 21) appear to have multiple *rpb2* loci, although only a single sequence type was detected by PCR and cloning.

Ka/Ks ratios

With respect to the two clades formed by *rpb2* duplication in *Hibiscus* s.l., the means of the pairwise comparisons to *Melhania* of Ka for clades 1 and 2 were 0.02 and 0.05, respectively (with no overlap in their ranges): However, it is difficult to obtain valid *P*-values for tests of significance with this type of data (S.V. Muse, personal communication). The means of Ks for clades 1 and 2 were 0.40 and 0.38, respectively, and showed a large overlap in their ranges. Mean Ka/Ks ratios for clades 1 and 2 were 0.06 and 0.11, respectively, and showed little overlap in their ranges. It appears that substitutions causing amino acid changes may have occurred more quickly in clade 2 than in clade 1, but the magnitude of these changes, particularly as reflected in Ka/Ks (despite some difference among clades) indicates that any shift in selective constraint is minimal between the copies of *rpb2* maintained in *Hibiscus* s.l.

Pairwise comparisons of *Radyera* sequences to *Melhania* reveal that the Ka/Ks ratio is nearly an order of magnitude higher than the mean Ka/Ks for all other ingroup pairwise comparisons to *Melhania* (0.70 and 0.08, respectively). However, Ks did not differ so drastically. The mean of *Radyera* Ks and all other ingroup Ks were 0.35 and 0.37, respectively. Although these differences could not be tested statistically, it appears that nonsynonymous changes are the main cause of the long *Radyera* branch and Ka/Ks difference seen here. This observation is consistent with the *Radyera* sequence originating from a pseudogene, because a relaxation of selective constraint could cause an increase in the nonsynonymous rate.

Discussion

Duplication and Long-Term Maintenance of *rpb2* in *Hibiscus* s.l.

The partitioning of multiple sequence types from several species into two lineages, each of which displays similar phylogenetic pattern, indicates that *rpb2* is duplicated in the common ancestor of the *Hibiscus* s.l. lineage. Initial Southern hybridizations confirm the presence of more than one *rpb2* locus in these taxa. Both loci appear to be functional, insofar as neither in-frame premature stop codons nor accelerated relative rates of evolution were found in the coding sequence examined.

Lineage-wide duplications of *rpb2* in plants (rather than single-species duplications caused by recent polyploidy) have only been reported for asterid angiosperms (Oxelman and Bremer 2000; Oxelman et al. 2004). However, the position of the duplication event in those taxa appears to be near the origin of the core eudicot group, indicating that most rosid angiosperms examined have lost one copy of *rpb2* (Oxelman et al. 2004). In contrast, the duplication in Malvaceae may be much more recent and appears to be confined to subfamily Malvoideae, although whether it involves the tribes Gossypieae and Malveae and the genus *Alyogyne* is uncertain because of the limited support found for several key nodes after the divergence of members of the subfamily (but see below). The age of the duplication would appear to fall between 44 to 39 and 31 to 25 Myr (between the estimated age of the subfamily [Pfeil et al. 2002] and the estimated age of the youngest duplicated lineage pair [Pfeil, unpublished data]).

The expectation for duplicated loci is one of several fates: long-term maintenance of the same or similar function (although developmental or organ-specific expression may eventuate [Adams et al. 2003]), divergence in function, silencing or loss of one copy, or intralocus or interlocus gene conversion. The last process can generate differences from expected phylogenetic relationships (Wendel 2000). However, there is no evidence for any gene conversion in the duplicated *rpb2* lineages found here, because the phylogenetic patterns within each copy in *Hibiscus* s.l. are largely concordant with each other and with cpDNA results (Pfeil et al. 2002). Only *Radyera farragei* and *Alyogyne huegelii* may have experienced some silencing of *rpb2* duplicates.

The occurrence of either functional divergence or gene silencing is difficult to ascertain without full exonic sequences but does not appear to be occurring (except for

Radyera farragei and *Alyogyne huegelii*), because little difference was observed between copies in the exon sequences generated here (paralog *p*-distances averaged 0.068). This is despite considerable difference among copies in their intron sequences (paralog *p*-distances averaged 0.157; exons and introns differed significantly, $P < 0.05$). Although copy-specific amino acid changes were found, these changes were not very numerous, are confined to parts of the gene that are not highly conserved, and only two of these changes resulted in a replacement amino acid from a different biochemical class. The Ka/Ks difference seen between copies may be explained by these amino acid substitutions (the range of Ka values alone does not overlap between copies), but given the small size of the difference in Ka/Ks and the position of these changes, there is no compelling evidence for either functional divergence or functional disruption.

The remaining process (locus loss) may produce gaps in the phylogeny where a sequence type is expected but was not seen in the *Hibiscus* s.l. lineage displaying duplications here (although further sampling underway has not recovered all expected sequences [Pfeil, unpublished data]). Long-term maintenance of some duplicated *rpb2* paralog appears to be occurring in this lineage.

When Did the *rpb2* Duplication Occur?

The phylogenetic, chromosome count, and Southern hybridization data suggest two key questions with regard to the origin of the *rpb2* duplicates found in *Hibiscus* s.l.: (1) At what point in the history of subfamily Malvoideae did the duplication occur (relative to taxon divergences)? (2) Was the duplication caused by a polyploid event? Given the weak nodes in the *rpb2* phylogeny among the main lineages in Malvoideae, the first question will be reduced to whether the *rpb2* duplication included the Gossypieae-*Alyogyne*-Malveae (GAM) lineage (i.e., all of Malvoideae) or not (i.e., only *Hibiscus* s.l. ± the *Lagunaria/Howittia* lineage). It should be noted that the support for the monophyly of Malvoideae (that includes both *Hibiscus* s.l. *rpb2* clades) excluding *Fremontodendron*, which is consistent with cpDNA results, must place the *rpb2* duplication after Malvoideae diverged from *Fremontodendron*.

The *rpb2* phylogenetic data suggest that the GAM clade does not ancestrally possess duplicate *rpb2* copies. However, Southern data is not entirely consistent with this result. Diploid *Gossypium* is clearly in possession of only a single *rpb2* locus, whereas two loci are present in *Alyogyne huegelii*. If *A. huegelii* is tetraploid (as its chromosome count of n = 32 would indicate), the common ancestor of these taxa most likely possessed a single *rpb2* locus (*Alyogyne* and Gossypieae appear to be more closely related than either is to Malveae). However, both Malveae representatives, including the presumably diploid *Modiola* (n = 9), contain more than one *rpb2* locus. Thus, the ancestral state in the GAM clade is equivocal, possessing either one or several *rpb2* loci.

Given the lack of supported resolution among the main Malvoideae lineages, the ancestral locus number of *rpb2* is difficult to infer, but two hypotheses present themselves: (1) The common ancestor of all Malvoideae possessed at least two *rpb2* loci, with a loss in the Gossypieae/*Alyogyne* ancestor. (2) The Malvoideae ancestor had just one *rpb2* locus, with duplication(s) taking place after the GAM lineage diverged from other Malvoideae (in non-GAM lineages) and within the Malveae (after it diverged from other GAM lineage members). If all of the non-GAM lineages that possess duplicated *rpb2* (*Hibiscus* s.l., *Lagunaria/Howittia*, and *Radyera*) are most closely related and share a duplication of *rpb2*, then these hypotheses are equivocal in terms of the number of evolutionary events required to explain the observations. Although *rpb2* is unresolved with regard to the latter three lineages, chloroplast and limited morphological data do not support the monophyly of these non-GAM lineages (Pfeil et al. 2002), increasing the number of events required under the second hypothesis. Although this issue is not settled, it appears that the first hypothesis, that the common ancestor of all Malvoideae possessed at least two *rpb2* loci with subsequent losses in *Gossypieae/Alyogyne*, is more consistent with the data in hand.

It is unclear whether the duplication of *rpb2* is the result of genome-wide polyploidy or a smaller scale segmental or single-gene duplication. Given the phylogeny here and that based on chloroplast data, the ancestral chromosome number of Malvoideae may be between n = 12 and n = 18 (assuming base numbers of *Hibiscus* s.l., n = 12 to 18; Gossypieae, n = 12 to 13; *Alyogyne*, n = 16; Malveae, n = 7 to 8; *Lagunaria/Howittia*, n = unknown; and *Radyera* n = 18). This number is high enough to hypothesize that Malvoideae may share a polyploid event that duplicated *rpb2*, but comparisons with other closely related Malvaceae complicate this picture. *Fremontodendron*, the closest relative sampled here, is n = 20 (Lenz 1950) or n = 49 (Lloyd 1965), while some of the Matisieae (which are closer to Malvoideae than *Fremontodendron* in cpDNA data) are n = 44 (*Quararibea aurantiocalyx* [Oginuma, Alverson, and Baum 1999]). Bombacoideae generally fall between n = 43 and 48 (Oginuma, Alverson, and Baum 1999). If the ancestral chromosome number was n = 43 to 49, then two independent reductions would have taken place, one in *Fremontodendron* and one in Malvoideae. If the ancestral state was n = 12 to 20, three independent increases to double or more chromosome number would be inferred. Although this is a simplification of chromosome evolutionary possibilities, the main conclusion is that there is no evidence that the ancestral chromosome number in Malvoideae has doubled after divergence from *Fremontodendron*, which would be expected if Malvoideae shared a polyploid history. Therefore, it appears that *rpb2* duplicated independently rather than being the result of a polyploid event. It remains unclear how the evidence for paleopolyploidy in "diploid" (n = 13) *Gossypium* species (see *Introduction*) will be reconciled with this conclusion. Perhaps that evidence is the result of a much older event—one that apparently predates the Malvaceae sampled here.

Pseudogene Found in *Radyera*

The accelerated rate of exon change coupled with the presence of premature in-frame stop codons indicated that

the *rpb2* sequences found in *Radyera* represent a pseudogene. The dramatic increase in Ka in *Radyera* relative to other Malvoideae, despite very similar Ks, is consistent with this hypothesis. No functional copy was recovered from this taxon. Some sequences found in *Alyogyne huegelii* may also represent a pseudogene, although this is not clear. If so, this may be an example of a relatively recent loss of function.

Recombinant Sequences Rare in Malvaceae *rpb2*

In the absence of any information regarding locus number for *Howittia*, it is impossible to rule out either intergenic or PCR-mediated recombination in this species. Either process, if common, can make a phylogenetic study using such a region difficult. Whichever phenomenon explains the *Howittia* recombinant clone, this observation is rare in this region and in the species investigated. Of 99 clones examined only this one case was found. Sampling of additional species has increased the number of clones to 172 with no additional cases of recombination detected (Pfeil, unpublished data). It may be that this portion of *rpb2* is not prone to recombination events in Malvaceae, perhaps because of the time since the paralogs diverged and possibly chromosomal location. Very similar multiple copies of target sequences present in the same PCR reaction are expected to generate PCR-mediated recombination (Cronn et al. 2002), whereas chromosomal position may influence the amount of interlocus interactions among nonhomologous chromosomes (Wendel 2000). The first process may also be able to operate in *rpb2* in Malvaceae because the exon sequences contain stretches of high conservation that could act as priming sites for an incomplete PCR product. A stretch of such sequence lies in between the halves of the recombined sequence found here, and, therefore, PCR mediated recombination still offers the most reasonable explanation for the origin of this sequence. It is unclear whether this low level of recombination will be a general property of *rpb2*, but if so, this region has great potential for systematic studies of ancient polyploids where the possession of duplicated nuclear genes is very common.

Phylogenetic Utility of *rpb2* in Malvaceae Subfamily Malvoideae

Whereas exons alone provided little phylogenetic information within the subfamily, the combined exon and intron data produced an internally consistent set of characters that gave rise to a phylogenetic hypothesis that was (1) mostly well-resolved and well-supported, (2) largely consistent with cpDNA results, and (3) robust to analysis using alternative models of evolution. Where *rpb2* data fail to robustly resolve relationships in subfamily Malvoideae (between Gossypieae, Malveae, *Alyogyne*, *Hibiscus* s.l., and *Howittia*/*Lagunaria*), other data have also failed (Pfeil et al. 2002).

Conclusion

Low-copy nuclear DNA regions have great potential for phylogenetic analysis at all levels of divergence and may allow for the detection of interesting processes such as ancient polyploidy. However, these regions present technical and analytical challenges. Studies using low-copy nDNA need to (1) investigate the presence of multiple copies, (2) establish orthology versus paralogy, and (3) consider PCR and *Taq* artifacts.

Duplication of *rpb2* was found in several species of Malvaceae and a single-gene or segmental duplication appears the more likely cause than ancient polyploidy. Long-term maintenance of duplicated copies and low levels of recombination (either interlocus or PCR induced) were found in some Malvaceae *rpb2* lineages, in particular *Hibiscus* s.l.. These findings, coupled with well-resolved and well-supported phylogenetic hypotheses congruent with cpDNA results, demonstrate the utility of *rpb2* in Malvaceae systematics and potentially for phylogenetic studies in other groups.

Acknowledgments

The authors thank L. M. Broadhurst, A. H. D. Brown, L. G. Cook, S. V. Muse, B. Oxelman, and an anonymous reviewer for comments on the manuscript and the Australian Biological Resources Study for funding support. This study was undertaken while B.E.P. was supported by an Australian Postgraduate Award.

Literature Cited

Adams, K. L., R. C. Cronn, R. Percifield, and J. F. Wendel. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-spcific reciprocal silencing. Proc. Natl. Acad. Sci. USA **100**:4659–4654.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389–3402.

Alverson, W. S., B. A. Whitlock, R. Nyffeler, C. Bayer, and D. A. Baum. 1999. Phylogeny of the core Malvales: evidence from *ndhF* sequence data. Am. J. Bot. **86**:1474–1486.

Bates, D. M. 1968. Generic relationships in the Malvaceae, tribe Malveae. Gentes Herbarum **10**:117–135.

Bates, D. M., and O. J. Blanchard. 1970. Chromosome numbers in the Malvales. II. New or otherwise noteworthy counts relevant to classification in the Malvaceae, tribe Malveae. Am. J. Bot. **57**:927–934.

Bayer, C., M. F. Fay, A. Y. de Bruijn, V. Savolainen, C. M. Morton, K. Kubritzki, W. S. Alverson, and M. W. Chase. 1999. Support for an expanded family concept of Malvaceae within a recircumscribed order Malvales: a combined analysis of plastid *atpB* and *rbcL* DNA sequences. Bot. J. Linn. Soc. **129**:267–303.

Bayer, C., and K. Kubitzki. 2002. Malvaceae. Pp. 225–311 *in* K. Kubitzki and C. Bayer, eds. The families and genra of vascular plants. Vol. 5. Springer, Berlin.

Bracho, M. A., A. Moya, and E. Barrio. 1998. Contribution of *Taq* polymerase-induced errors to the estimation of RNA virus diversity. J. Gen. Virol. **79**:2921–2928.

Brubaker, C. L., A. H. Paterson, and J. F. Wendel. 1999. Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. Genome **42**:184–203.

Cronn, R. C., M. Cedroni, T. Haselkorn, C. Grover, and J. F. Wendel. 2002. PCR-mediated recombination in amplification

products derived from polyploid cotton. Theoret. Appl. Genet. **104**:482–489.

Cronn, R. C., R. L. Small, and J. F. Wendel. 1999. Duplicated genes evolve independently after polyploid formation in cotton. Proc. Natl. Acad. Sci. USA **96**:14406–14411.

Denton, A. L., B. L. McConaughy, and B. D. Hall. 1998. Usefulness of RNA polymerase II coding sequences for estimation of green plant phylogeny. Mol. Biol. Evol. **15**: 1082–1085.

Fryxell, P. A. 1968. A redefinition of the tribe Gossypieae. Bot. Gaz. **129**:296–308.

———. 1988. Malvaceae of Mexico. Syst Bot. Monogr. **25**: 1–522.

Huelsenbeck, J. P. and F. Ronquist. 2001. Mr Bayes. Version 2. 01

Huelsenbeck, J. P., F. Ronquist, R. Nielsen, and J. P. Bollback. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science **294**:2310–2314.

Kelchner, S. A. 2000. The evolution of noncoding chloroplast DNA and its application in plant systematics. Ann. MO Bot. Gard. **87**:482–498.

Lenz, L. W. 1950. Chromosome numbers of some western American plants. Aliso **2**:317–318.

Lloyd, R. M. 1965. A new species of *Fremontodendron* (Sterculiaceae) from the Dierra Nevada foothills, California. Brittonia **17**:382–384.

Masterson, J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science **264**:421–423.

Muravenko, O. V., A. R. Fedotov, E. O. Punina, L. I. Fedorova, V. G. Grif, and A. V. Zelenin. 1998. Comparison of chromosome BrdU-Hoechst-Giesma banding patterns of the A_1 and $(AD)_2$ genomes of cotton. Genome **41**:616–625.

Oginuma, K., W. S. Alverson, and D. A. Baum 1999. A cytological study of three genera of neotropical Bombacaceae (clades Bombacoideae and Malvoideae). Acta Phytotax. Geobot. **50**:173–178.

Oxelman, B., and B. Bremer. 2000. Discovery of paralogous nuclear gene sequences coding for the second-largest subunit of RNA polymerase II (rpb2) and their phylogenetic utility in Gentianales of the Asterids. Mol. Biol. Evol. **17**:1131–1145.

Oxelman, B., N. Yoshikawa, B. L. McConaughy, J. Luo, A. L. Denton, and B. D. Hall. 2004. Rpb2 gene phylogeny in flowering plants, with particular emphasis on asterids. Mol. Phylogenet. Evol. (in press).

Pääbo, S., D. M. Irwin, and A. C. Wilson. 1990. DNA damage promotes jumping between templates during enzymatic amplification. J. Biol. Chem. **265**:4718–4721.

Pfeil, B. E., C. L. Brubaker, L. A. Craven, and M. D. Crisp. 2002. Phylogeny of *Hibiscus* and the tribe Hibisceae (Malvaceae) using chloroplast DNA sequences of *ndhF* and the *rpl16* intron. Syst. Bot. **27**:333–350.

Popp, M., and B. Oxelman. 2001. Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoeologous nuclear DNA sequences. Mol. Phylogenet. Evol. **20**:474–481.

Reinisch, A. J., J.-M. Dong, C. L. Brubaker, D. M. Stelly, J. F. Wendel, and A. H. Paterson. 1994. A detailed RFLP map of cotton, *Gossypium hirsutum* × *Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. Genetics **138**:829–847.

Seelanan, T., A. Schnabel, and J. F. Wendel. 1997. Congruence and consensus in the cotton tribe (Malvaceae). Syst. Bot. **22**:259–290.

Small, R. L., J. A. Ryburn, R. C. Cronn, T. Seelanan, and J. F. Wendel. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear ADH sequences for phylogeny reconstruction in a recently diverged plant group. Am. J. Bot. **85**:1301–1315.

Sweetser, D., M. Nonet, and R. A. Young. 1987. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc. Natl. Acad. Sci. USA **84**:1192–1196.

Swofford, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.10b. Sinauer Associates, Sunderland, Mass.

Webber, J. M. 1934. Chromosome number and meiotic behaviour in *Gossypium*. J. Agric. Res. **49**:223–237.

Wendel, J. F. 1989. New World tetraploid cottons contain Old World cytoplasm. Proc. Natl. Acad. Sci. USA **86**:4132–4136.

———. 2000. Genome evolution in polyploids. Plant Mol. Biol. **42**:225–249.

Wendel, J. F. and J. J. Doyle. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. Pp. 265–296 *in* D. E. Soltis, P. S. Soltis, and J. J. Doyle, eds. Molecular systematics of plants II: DNA sequencing. Kluwer Academic Publishers, Norwell, Mass.

Wilson, F. D. 1994. The genome biogeography of *Hibiscus* L. section *Furcaria* DC. Genet. Resources Crop Evol. **41**:13–25.

Yang, Z. 2000. Phylogenetic analysis by maximum likelihood (PAML). Version 3.0. University College, London.

Spencer Muse, Associate Editor

Accepted March 23, 2004