Molecular Phylogenetics and Evolution
Vol. 13, No. 1, October, pp. 20–30, 1999
Article ID mpev.1999.0627, available online at http://www.idealibrary.com on IDEAL®

Potential Phylogenetic Utility of the Low-Copy Nuclear Gene *pistillata* in Dicotyledonous Plants: Comparison to nrDNA ITS and *trnL* Intron in *Sphaerocardamum* and Other Brassicaceae

C. Donovan Bailey and Jeff J. Doyle

L. H. Bailey Hortorium, 462 Mann Library, Cornell University, Ithaca, New York 14853-4301

Received May 11, 1998; revised October 27, 1998

We report the potential phylogenetic utility of DNA sequence data from the last 700 bp of a ca. 1-kb intron of the MADS-box gene *pistillata* from a sampling of *Sphaerocardamum* species and other Brassicaceae. These results are compared with nrDNA ITS and the chloroplast *trnL* intron for the same taxa to demonstrate the potential phylogenetic utility of this *pistillata* intron and to identify potential historically independent sequences for an ongoing study of relationships within *Sphaerocardamum.* Analyses of the DNA sequence data for Brassicaceae indicated that pairwise divergences and potentially informative characters were higher in the *pistillata* intron (0.6–30.8%, 284 characters) and ITS (0–24%, 94 characters) than in the chloroplast *trnL* intron (0–4.2%, 17 characters). A comparison of *Sphaerocardamum* sequences identified low divergences and numbers of informative characters for *trnL* intron (0–2.4%, 1 character) and nrDNA ITS (0–2.5%, 2 characters) and substantially more variation among the *pistillata* sequences (0.15–3.7%, 19 characters). Phylogenetic analyses of these *pistillata* sequences fully resolve ingroup relationships without character conflict. Results of *pistillata* PCR amplifications from a broader dicot sample showed that some primers may be useful in amplifying orthologous *pistillata* sequences. Ultimately this *pistillata* intron may be a valuable source of phylogenetic characters at lower taxonomic levels.

Key Words: **Brassicaceae; MADS-box; nrDNA ITS; *pistillata;* phylogeny; *Sphaerocardamum; trnL* intron.**

INTRODUCTION

Although sequences from chloroplast genomes have been routinely utilized in plant molecular phylogenetics (reviewed by Olmstead and Palmer, 1994) there is currently a limited selection of nuclear DNA sequences, coding or noncoding, from which researchers can hope to generate characters for resolving phylogenetic relationships among members of selected groups (reviewed by Soltis and Soltis, 1998). The delayed use of nuclear sequences has been a function of a number of complicating factors, including a relatively poor understanding of many nuclear genes which often exist in large multigene families (Clegg *et al.,* 1997). Sequences utilized from complex gene families can lead to problems with unidentified paralogy and incomplete concerted evolution (Sanderson and Doyle, 1992).

Noncoding DNA sequences are expected to evolve more rapidly than coding regions and so represent a potential source of characters for resolving relationships at lower taxonomic levels (e.g., Li, 1997). However, until recently, the only nuclear noncoding DNA sequences widely used to provide characters at lower taxonomic levels have been the nrDNA internal transcribed spacer region (nrDNA ITS). The nrDNA ITS cannot be expected to resolve lower level relationships under all circumstances, both because of potential paralogy problems (e.g., Wendel *et al.,* 1995; Buckler *et al.,* 1997; Campbell *et al.,* 1997) and because their combined sequence, ca. 400–500 bp for most angiosperms (Baldwin *et al.,* 1995), may lack sufficient characters for resolving relationships. In addition, researchers need alternative independent (unlinked) sequences to explore contrasting phylogenetic signals between chloroplast and nrDNA ITS sequences (e.g., Kim and Jansen, 1994; Soltis and Kuzoff, 1995; Wendel *et al.,* 1995; Schilling and Panero, 1996). Nuclear alternatives to the nrDNA ITS are being developed for use in plant phylogeny reconstruction (reviewed by Soltis and Soltis, 1998), including *waxy* introns (Mason-Gamer and Kellogg, 1996; Peralta *et al.,* 1997), argenine decarboxylase coding sequence (Galloway *et al.,* 1997), polygalacturonase (Atkinson *et al.,* 1997), glutamine synthetase introns and exons (Doyle and Doyle, 1997), *rpb 2* (Denton *et al.,* 1996), alcohol dehydrogenase (Morton *et al.,* 1996; Sang *et al.,* 1997), phytochrome (Mathews *et al.,* 1995), and histone H3 introns (Doyle *et al.,* 1996). However, a greater selection of nuclear sequences will provide researchers with increased possibilities of identifying sequences appropriate for particular phylogenetic questions.

1055-7903/99 $30.00

When selecting an alternative nuclear sequence for resolving relationships at lower taxonomic levels, several characteristics should be considered. First and foremost, researchers would like sequences that are known to be single or low copy number from a variety of taxa and that include one or more sizable noncoding regions flanked by conserved regions from which primers can be developed. Ideally these sequences should be single copy and not members of a gene family; however, if members of a gene family are to be used, their paralogs should be easily distinguishable from one another. These characteristics allow the design of conserved specific primer sequences, reduce paralogy concerns, and provide the potential for relatively large amounts of phylogenetic characters at lower taxonomic levels. Researchers would like to also consider sequences whose linkage relationships have been established or can be reasonably hypothesized. Unlinked loci may represent historically independent markers and thus their comparison can be a powerful tool for providing evidence of reticulate evolutionary events, a problem of particular concern when dealing with species-level relationships among plant taxa.

Members of the MADS-box nuclear gene family have been identified from a wide variety of plants (flowering and nonflowering) and they generally fit the criteria outlined above for potential use in phylogeny reconstruction. Phylogenetic analyses of the MADS-box subfamilies identify strong support for the paralogous groups (e.g., Theissen *et al.,* 1996; Münster *et al.,* 1997), suggesting that distinctions between these can be identified. This reduces, but not necessarily eliminates, the possibility of mistaken paralogy (Kramer *et al.,* 1998). Characterized MADS-box members maintain at least four of nine possible intron positions across all members of the gene family (Ma *et al.,* 1991; Doyle, 1994). These introns are typically flanked by conserved exons that provide the opportunity to develop universal primers (e.g., Coen *et al.,* 1990; Yanofsky *et al.,* 1990; Mandel *et al.,* 1992; Goto and Meyerowitz, 1994; Tandre *et al.,* 1995). Although MADS-box genes have been analyzed for a better understanding of the gene family (e.g., Doyle, 1994; Purugganan *et al.,* 1995; Tandre *et al.,* 1995; Theissen *et al.,* 1996; Münster *et al.,* 1997, Kramer *et al.,* 1998), there have been no published uses of these sequences to reconstruct taxic phylogenies.

The nuclear encoded MADS-box gene *pistillata* is single copy in *Arabidopsis thaliana* and contains several introns including one of nearly a kilobase (kb) (Goto and Meyerowitz, 1994), suggesting that it may be useful in phylogeny reconstruction. We present data utilizing the last 700 bp of the 1-kb first intron (IVS 1) from *pistillata.* Sequences were sampled from the small Brassicaceae genus *Sphaerocardamum,* the focus of our ongoing systematic studies (Bailey, unpubl.), and between this genus and other Brassicaceae to test the sequences' potential phylogenetic utility. These *pistillata* IVS 1 data are compared to sequences from nrDNA ITS and chloroplast *trnL* intron, for the same sample, to provide a comparison of *pistillata* IVS 1 to sequences of previously demonstrated utility. The latter sequences were chosen both due to their known utility, at lower phylogenetic levels, and due to their linkage relationships in *Arabidopsis* (ITS: chromosomes 2 and 4 [Meyerowitz, 1994]; *pistillata:* chromosome 5 [Goto and Meyerowitz, 1994]; and *trnL* intron: chloroplast). To the extent that linkage is conserved among members of the Brassicaceae (e.g., Teutonico and Osborn, 1994; Lagercrantz *et al.,* 1996), these loci may all be historically independent.

The *pistillata* locus on chromosome 5 in *A. thaliana* (Fig. 1) has a total length of approximately 2.3 kb, including five introns and six exons ranging in size from 997 to 72 bp and 381 to 30 bp, respectively (Goto and Meyerowitz, 1994). The results presented here compare the last 700 bp of the 997-bp IVS 1 of *pistillata* (located in the "I" coding region between the MADS and first K-box coding exons) to: (1) the ca. 460 bp of nrDNA ITS, the spacers between the 18S and 5.8S (ITS-1) and 5.8S and 26S (ITS-2) nrRNA genes, respectively (White *et al.,* 1990); and (2) the ca. 500-bp intron of chloroplast gene tRNA-*Leu* (*trnL* intron; Taberlet *et al.,* 1991; Fangan *et al.,* 1994).

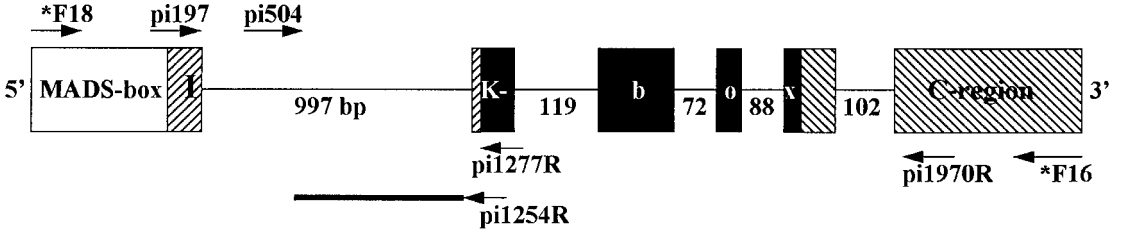

FIG. 1. Map of the 2.3-kb *Arabidopsis thaliana* transcribed *pistillata* locus (adapted from Goto and Meyerowitz, 1994). Boxes and lines represent exons and introns, respectively, and sizes of the five introns are given below each respective intron along with relative primer locations and orientations. The solid bar below the first intron identifies the position and approximate scale of the region analyzed in this study.

MATERIALS AND METHODS

Four of the eight currently recognized species of *Sphaerocardamum* (Rollins, 1984) and six additional Brassicaceae were included in this survey (Table 1). The four ingroup species were selected as representatives of most of the range of morphological diversity within the genus and five of the six outgroup taxa were selected as single species representatives of possible sister genera to *Sphaerocardamum* (I. Al-Shehbaz, R. Price, and S. O'Kane Jr., pers. comm.). Because sequences from two of the three loci were available for *A. thaliana* in GenBank, this taxon was also included (*trnL* intron, accession ATTRNLI: Fangan *et al.,* 1994; nrDNA ITS, ATU43225: O'Kane *et al.,* 1996).

DNAs were extracted from leaves of greenhouse-grown samples, field-collected and silica gel-dried samples, or herbarium specimens using the modified CTAB method of Doyle and Doyle (1990). Polymerase chain reactions (PCR) were conducted using either *Taq* DNA polymerase (BRL Life Technologies) or, for some difficult to amplify templates, using "Ready-to-go PCR beads" (Pharmacia Biotech). Amplifications were performed on a PTC-100 thermocycler (MJ Research Inc.). Each locus required different amplification conditions for optimal results. All amplifications began with a 4-min 94°C denaturation step, followed by 35 rounds of (1) 1 min 94°C denaturation; (2) 1 min annealing at 50°C (nrDNA ITS), 52°C (*pistillata*), or 55°C (*trnL* intron); and (3) a 72°C 1 min (nrDNA ITS) or 2 min (*pistillata* and *trnL* intron) extension. PCR products were separated on a 0.7% TBE agarose gel followed by band excision and isolation (QIAquick Gel Extraction Kit, QIAGEN Corp.) for direct sequencing. Sequencing steps were carried out via automated cycle sequencing using dRhodamine dye termination (ABI 377, Applied Biosystems Inc.; Cornell Biotechnology Resource Center) or ^{35}S manual cycle sequencing (Sequitherm EXCEL II kit; Epicentre Technologies). Sequencing results from a few templates identified apparent ambiguities due to length or substitutional polymorphism(s). These templates were subsequently cloned (TOPO-TA cloning kit, Invitrogen) and a minimum of two clones sequenced. Manufacturers' instructions were followed for all the kits.

PCR amplifications of the *trnL* intron region were conducted using primer c with either d or f as reverse primers (Taberlet *et al.,* 1991). Both strands of *trnL* intron were sequenced via automated cycle sequencing using each respective PCR primer as a sequencing primer, generally providing a 60% or greater confirmation.

ITS-1/5.8S/ITS-2 were amplified using ITS4 + ITS5 primers (White *et al.,* 1990). The antisense strand for ITS-1 and ITS-2 was sequenced via automated cycle sequencing utilizing the ITS4 primer. The sense stand of ITS-1 was 80% or more confirmed using the ITS5 primer via manual cycle sequencing (see Discussion for rationale).

Initial amplifications of the *pistillata* locus were carried out using the F18 + F16 primers (Fig. 1; Table 2) provided by E. M. Meyerowitz and E. Krizek (California Institute of Technology). These primers were designed against *A. thaliana* complementary DNA sequences (cDNA), with F18 developed as a general MADS-box primer and F16 as a *pistillata*-specific primer (E. Krizek, pers. comm.). Due to weak or unsuccessful amplifications with these primers and a need for sequencing primers located in or around IVS 1, a number

TABLE 1

Taxon Sampling, Voucher Information, and GenBank Accession Nos.

Taxon	Collection/herbarium	ITS	*trnL*	*Pistillata*
Arabidopsis thaliana (L.) Heynh.	C. D. Bailey 69/BH	—	—	AF055189 AF055190
Capsella bursa-pastoris (L.) Medikus	C. D. Bailey 1/BH	AF055196	AF055264	AF055181 AF055182 AF055183
Halimolobos jaegeri (Munz) Rollins	Tiehm & Moorefield 8542/BH	AF055201	AF055268	AF055191
Lepidium campestre (L.) R. Br.	C. D. Bailey 3/BH	AF055197	AF055265	AF055184
Lesquerella fendleri (A. Gray) S. Wats.	C. D. Bailey 43/BH & MEXU	AF055198 AF055199	AF055266	AF055185
Nerisyrenia linearifolia (S. Wats.) E. L. Greene	C. D. Bailey 56/BH & MEXU	AF055200	AF055267	AF055186 AF055187 AF055188
Sphaerocardamum macropetalum (Rollins) Rollins	C. D. Bailey 45/BH & MEXU	AF055192	AF055260	AF055176 AF055177
S. macrum (Standley) Rollins	C. D. Bailey 57/BH & MEXU	AF055194	AF055262	AF055179
S. nesliiforme Schauer	H. E. Moore 83349/BH	AF055195	AF055263	AF055180
S. stellatum (S. Wats.) Rollins	C. D. Bailey 47/BH & MEXU	AF055193	AF055261	AF055178

TABLE 2

***Pistillata* Primers**

Primer name	Primer sequence
Forward	
*F18 (1)	5′ ATG GGT AGA GGA AAG ATC GAG 3′
pi197	5′ G(CA)T ATG TTG GA(CT) CAA TAC CA 3′
pi504	5′ TAG ATG CAG ATC TAC AAA GGT 3′
Reverse	
pi1254R	3′ GG AGT CGT TAC TCT AAC TAT CCC 5′
pi1277R	3′ AGT TCT TTC TCT TAC TAT 5′
pi1970R	3′ ACG TTG GCT AAG TCG GTT 5′
*F16 (2005)	3′ TAC AGA AAC CAG TAG CTA ACT 5′

Note. The primers generated as part of this study begin with the prefix "pi" and are numbered according to their 5′ position on *Arabidopsis thaliana* sense strand relative to the start codon (designated position 1). Degenerate primer sites are indicated with parentheses around the nucleotide combination. Reverse primer sequences are given as antisense sequences; an "*" designates primers developed and provided by E. M. Meyerowitz and E. Krizek (*Arabidopsis* nucleotide position follows each name in parentheses).

of other primers were developed. Because no complete *pistillata* gene sequences were available, these primers (pi197, pi1254R, pi1277R, and pi1970R; Fig. 1; Table 2) were designed through comparison of putatively orthologous *pistillata* cDNA sequences available from GenBank: *Arabidopsis* (*pistillata:* accession no. D30807), *Nicotiana* (ntglo: X67959), *Petunia* (fbp1: M91190), and *Oryza* (MADS2: L37526; MADS4: L37527). An effort was made to make these primers as universal as possible; however, when differences were observed within the primer site between Brassicaceae and other taxa, nucleotides found in Brassicaceae were used. The pi504 primer was generated from within IVS 1 and was designed only from sequences of taxa obtained in this study (Table 1). Although many of the primers listed in Table 2 where used in early stages of the Brassicaceae portion of the project, the majority of sequences presented are from PCR products using pi504 + F16 or pi504 + pi1254R primer combinations. Both strands of these PCR products were sequenced via automated sequencing using the pi504 + pi1254R primers for sense and antisense stands, respectively, generally providing 70% or greater overlap. BLAST searches (Altschul *et al.,* 1990) using exon regions were conducted to support similarity to *pistillata.*

To test the utility of some of our primers across a wider taxonomic sample, PCR amplifications of *pistillata* sequences were attempted from *Asarum canadense* (Magnoliidae; voucher: L. Kelly 674, deposited at BH), *Betula jaquemontiana* (Hamamelididae; Doyle 1306, BH), *Glycine tabacina* (Rosidae), *Hintonia* (Asteridae; H. Ochoterena 174, BH), *Nicotiana tabacum* (Asteridae), *Petunia hybrida* (Asteridae), *Pisum sativum* (Rosidae), *Spinacia oleracea* (Caryophyllidae), and *Oxalis* sp. (Rosidae: Emshwiller 933, BH). Three primer combinations were attempted: (1) F18 + F16 primers with a 54°C annealing and 2-min extension using *Taq* polymerase, (2) pi504 + F16 primers with a 50°C annealing and 2-min extension using the "Ready-to-go PCR Beads" (Pharmacia), and (3) pi504 + 1254R primers using a 50°C annealing and 1-min extension with *Taq* polymerase.

Sequence alignments were conducted using the multiple sequence alignment program MALIGN (Wheeler and Gladstein, 1993). To assess the relative stability of the data to alignment perturbation, four different alignments were conducted on each dataset (gap/change cost ratios: 5/1, 3/1, 2/1, 1/1). Aligned DNA sequence data were edited in the matrix editor DADA (Nixon, 1993a), with individual substitutions and indels coded as single characters. Data were analyzed under equal weights and unordered states using 10 random addition sequences followed by tree bisection–reconnection (mult*N and max* commands) in the cladistics computer program NONA (Goloboff 1993, upgraded 1996). *Lepidium* was selected to root the trees based on preliminary *ndhF* phylogenetic analyses of Brassicaceae which identify *Lepidium* as a basal member of the clade containing *Sphaerocardamum* (R. Price, pers. comm.). One hundred strict consensus bootstrap replicates (Felsenstein, 1985) were calculated using Clados (Nixon, 1993b) in combination with NONA (using the same search strategy) and trees were observed with informative character state distributions using Clados. Pairwise divergences, pairwise transition/transversion ratios, and G + C content were calculated in PAUP* 4.0 (Swofford, unpubl.).

RESULTS

Amplifications of Pistillata

Initial PCR amplifications using the F18 + F16 primer combination on members of Brassicaceae generated weak products but were successful enough to allow cloning and partial sequencing of *Arabidopsis, Capsella, Nerisyrenia,* and *Sphaerocardamum macrum* products using the F18 and/or pi197 primers. These results provided MADS-box exon sequences similar enough to *pistillata* to suggest homology and allowed us to generate a conserved primer (pi504) from within the previously unsequenced IVS 1 for use in generating more reliable amplifications.

Amplification attempts of *pistillata* from a broader dicotyledonous sample (data not shown) using the F18 + F16 primer combination produced a band similar in size to the Brassicaceae size standard in *Pisum, Hintonia,* and *Spinacia;* however, *Pisum* and *Spinacia* produced additional bands. Using the pi504 and F16 primer combination produced bands similar in size (ca. 1700 bp) to our Brassicaceae standard from *Asarum,*

Betula, Glycine, and *Pisum;* however, aside from the *Pisum* sample, additional products were also amplified. Amplification utilizing the pi504 and pi1254R primer combination generated products similar in size to the known Brassicaceae standard (ca. 700 bp) in *Asarum, Betula, Hintonia, Nicotiana, Pisum,* and *Spinacia,* with *Asarum, Nicotiana,* and *Spinacia* producing several additional bands.

Sequence Variation in Pistillata

For *Arabidopsis, Capsella, Nerisyrenia, Sphaerocardamum macropetalum, S. macrum,* and *S. nesliiforme* apparent heterozygosity identified through direct sequencing necessitated cloning to produce clean sequences. For *Capsella* and *Nerisyrenia* three distinct clones were identified and for *Arabidopsis* and *S. macropetalum* two distinct clones were identified. All of these clones were included in the subsequent analyses.

The first ca. 50 bp of the IVS 1 was found to possess a small microsatellite repeat $(TC)_n$ in some of the taxa and was therefore not included. This repeat contains 10, 8, 4, and 8 copies in our accessions of *Capsella, Lepidium, Nerisyrenia,* and *S. macrum,* respectively, and was absent from *Lesquerella* sequences.

The 3′ portion of *pistillata* IVS 1 analyzed here displayed substantial length variation, with sequences ranging in size from 607 bp (*Lesquerella*) to 715 bp (*Capsella*). This length variation, which was concentrated at the 3′ end, along with reasonably high divergences (see below), created conflict among alignments generated under different cost ratios and resulted in alignments that varied in length from 768 to 824 bp. Topologies and relevant statistics for each locus (tree length, pairwise divergence, transition transversion ratio, etc.) are given from analyses of the 2/1 gap to change cost alignments. Topologies from analyses of the other alignments are discussed only if conflict was identified between the topology resulting from the 2/1 alignment and one or more of the alternative alignments. Alignments based on the 2/1 parameters were deposited in the EMBL alignment database (accessions: DS36177[*trnL* intron]; DS36178[*pistillata* IVS1]; and DS36179[nrDNA ITS]) and are available via anonymous FTP at FTP.EBI.AC.UK (directory: pub/database/embl/align). The alignments based on the alternative gap to change costs are available from the authors.

For *pistillata* IVS 1 uncorrected pairwise divergences ranged from 0.6 to 30.8% and G + C contents ranged from 26.5 to 29.6% (Table 3). The single most parsimonious gene tree (Fig. 2a; L = 513, CI = 75, RI = 87) generated from the 2/1 alignment cost parameter (783 aligned bases, 284 potentially informative substitutions) placed *S. macrum* and *S. nesliiforme* in a clade sister to the *S. macropetalum* and *S. stellatum* clade, with *Halimolobos* and then *Capsella* successively sister to this clade. The remaining taxa formed a second clade in which *Lesquerella* and *Nerisyrenia* were sister and *Arabidopsis* basal. The topology presented in Fig. 2a was stable under analyses of all alignments, except that *Capsella* moved to a position sister to *Lesquerella* in the 5/1 gap/change cost alignment and *Arabidopsis* moved to the basal-most position in the 3/1 alignment. No length variation sites (indels) were scored as characters in any of the *pistillata* analyses since these were very sensitive to alignment and often difficult to delimit due to overlapping indels of differing lengths.

For these *pistillata* data an additional analysis was conducted which excluded substitution characters from the highly alignment-variable regions to test the relative importance of these characters to the reported topology. This analysis excluded 96 of the 284 potentially informative characters and resulted in the same

TABLE 3

Comparative Sequence Characteristics

Sequence	Aligned bases	GC content (%)	Pairwise divergence (%)	Potentially informative characters		Missing values
				Indels	Subst.	
TOTAL						
pistillata Intron	783	26.5–29.6	0.6–30.8	—	284	154
trnL intron	556	33.2–34.8	0–4.2	3 (& 1 inver.)	13	0
nrDNA Total	638	49.8–57.1	0–16.8	3	91	15
ITS1	268	48–58	0–24.2	0	58	—
ITS2	198	53–62	0–19.7	3	31	—
5.8S	172	52	0–2.5	0	2	—
INGR						
pistillata Intron	697	27.7–28.9	0.15–3.7	4	15	0
trnL Intron	505	34.4–34.8	0–2.4	0	1	0
nrDNA Total	625	56.9–57.1	0–2.5	0	2	0

Note. These data are taken from analyses based on alignments of the 2/1 gap to change cost alignments. The first six rows (TOTAL) refer to comparison of sequences from the complete dataset and the last three rows (INGR) compare the ingroup sequences (*Sphaerocardamum*) analyzed relative to the *Halimolobos* sequence. The final column lists the number of "missing values" (due to gaps in alignment) included in the phylogenetic analyses.

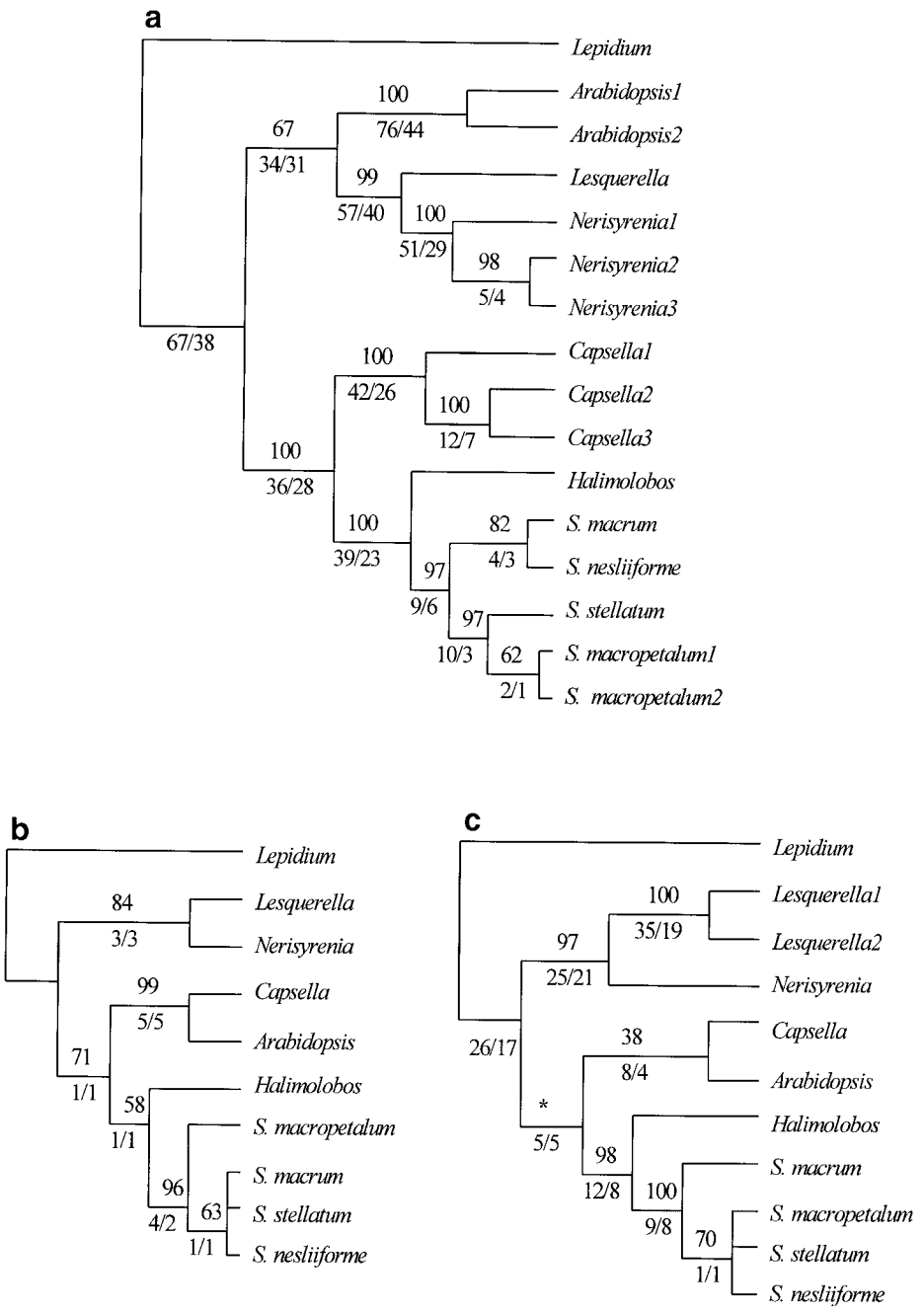

FIG. 2. Gene tree topologies. Bootstrap values for each clade are given on the top of the supporting branch and the total character state changes (acctran optimization) followed by the total that are nonhomoplaseous changes (total changes/nonhomoplaseous changes) are given below each branch. (a) Single most parsimonious gene tree generated for the complete *pistillata* IVS 1 dataset (L = 513, CI = 75, RI = 87). (b) Single most parsimonious gene tree generated for *trnL* intron dataset (L = 22, CI = 81, RI = 87). (c) One of two equally most parsimonious gene trees (L = 17.1, CI = 76, RI = 79) generated for the nrDNA ITS dataset. The clade not supported by all trees is identified by an "*."

single topology shown in Fig. 2a (L = 317, CI = 78, RI = 88).

Pistillata Transition to Transversion Ratios

Pairwise divergence and pairwise transition/transversion ratios were calculated and plotted against one another for *pistillata* IVS 1 (Fig. 3). In this small sample of taxa the transition/transversion ratios ranged from 0.25 to 3, with an average of 0.66. There was no great excess of transitions relative to transversions even in most comparisons of closely related sequences (Fig. 3). Linear regression (slope = −1.699) indicated a slight decrease in transitions relative to transversions with increasing divergence.

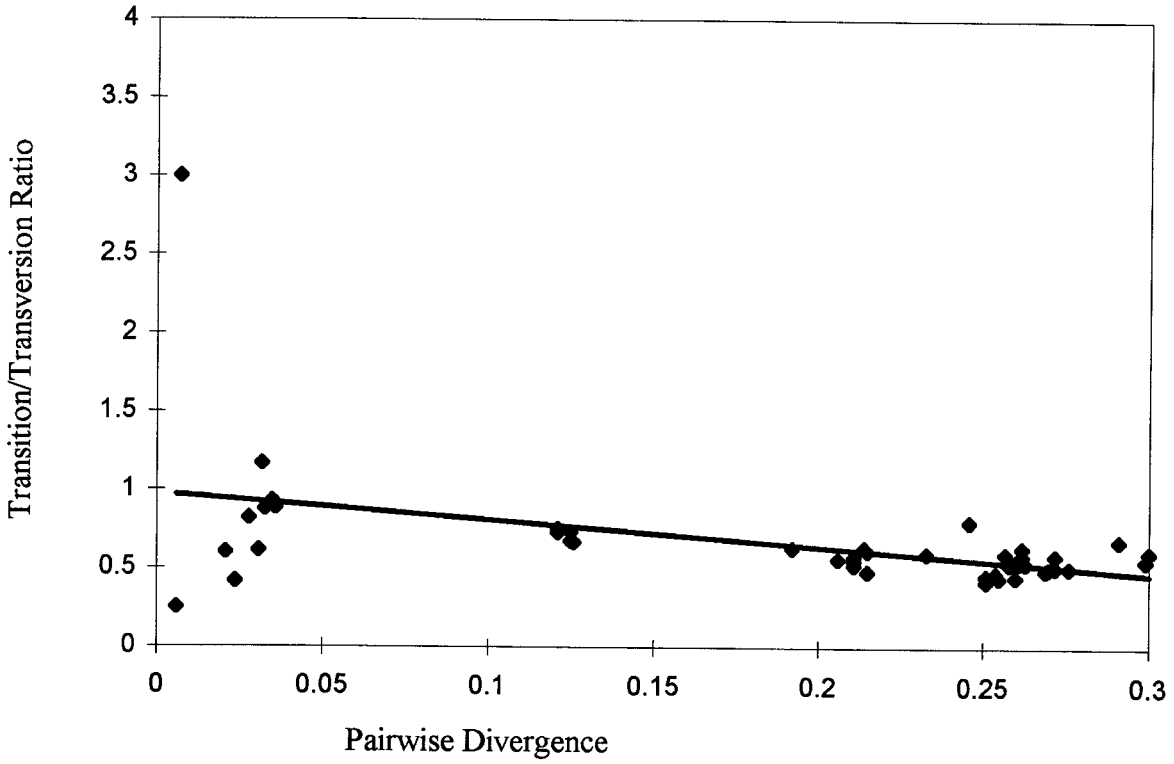

FIG. 3. *Pistillata* intron pairwise divergence vs transition/transversion ratio (slope = −1.699).

trnL Intron

Amplification and sequencing of the *trnL* intron resulted in the expected single product with sequences ranging in size from 495 bp (*Sphaerocardamum* spp.) to 533 bp (*Lesquerella*). Uncorrected pairwise divergences ranged from 0 to 4.2%, with an average G + C content of 33.2–34.8% (Table 3). Due to the moderate amount of length variation, alignments varied from 547 to 556 bp. Analyses on each of these identified the same single topology (Fig. 2b). Using the 2/1 gap/change ratio (556 aligned bases) the dataset generated a total of 13 potentially informative substitutions, three length variation characters, and one 6-bp putative inversion character. The single most parsimonious gene tree (Fig. 2b; L = 22, CI = 81, RI = 87) from parsimony analysis of this alignment resulted in a monophyletic *Sphaerocardamum* (*S. macropetalum* sister to an unresolved *S. macrum, S. nesliiforme,* and *S. stellatum* clade), with *Halimolobos* sister to this group, a clade with *Capsella* and *Arabidopsis* sister to this larger clade, and finally, a *Nerisyrenia* and *Lesquerella* clade sister to the overall group.

nrDNA ITS

Amplification and sequencing of the ITS-1, 5.8S, and ITS-2 from each taxon generated a single product that ranged in size from 594 to 626 bp. Only the *Lesquerella fendleri* sequence revealed polymorphism that required cloning to obtain unambiguous sequencing reads. For this species, two different ITS-1 regions were identified and included in subsequent analyses. For the total length of the sequenced region, pairwise divergences ranged from 0 to 16.8% and individual ITS-1, ITS-2, and 5.8S pairwise divergences ranged from 0 to 24.2, 0 to 19.7, and 0 to 2.5%, respectively (Table 3). G + C content for the total sequence ranged from 49.8 to 57.1% (Table 3). The overall number of aligned bases varied from 637 bp (5/1 gap to change ratio) to 639 bp (1/1 gap to change ratio). Analyses of the 2/1 gap to change cost ratio (91 potentially informative substitutions and 3 informative indels) produced two most parsimonious gene trees (L = 171, CI = 76, RI = 79); one tree is congruent with the tree generated by *trnL* (Fig. 2b), whereas the other is in conflict due to the reversed position of the *Capsella*/*Arabidopsis* clade relative to the *Lesquerella*/*Nerisyrenia* clade. The topology congruent with *trnL* intron is presented in Fig 2c. with an "*" indicating the node that collapses in the strict consensus of the two most parsimonious trees. The strict consensus topology was stable in analyses using the other alignment parameters except with respect to the clade containing *Arabidopsis* and *Capsella.* The *Arabidopsis*/*Capsella* clade (38% boot-

strap value) collapsed in a strict consensus of analyses based on all alignments.

Topology Comparison

The separate gene trees produced using the *trnL* intron, nrDNA ITS, and *pistillata* IVS 1 datasets, rooted with *Lepidium,* all identify: (1) a monophyletic *Sphaerocardamum,* (2) a *Halimolobos* sequence as most closely related to genes of *Sphaerocardamum,* and (3) *Nerisyrenia* and *Lesquerella* sequences as sister to one another. The differences among topologies supported by each dataset are found in the resolution within the ingroup and with respect to the positions of *Arabidopsis* and *Capsella.* At the ingroup level the single informative characters provided by *trnL* intron and ITS are in conflict with each other and both are in conflict with the more numerous *pistillata* characters.

Ingroup Comparison

Because resolution within *Sphaerocardamum* is our primary concern, we sought to eliminate alignment ambiguities caused by more divergent sequences. Therefore, the same multiple alignment parameter regime was implemented using only sequences from *Sphaerocardamum* and *Halimolobos.* Results from these analyses identified only one informative substitution character from the *trnL* intron (505 bp) and two from the ITS (625 bp) datasets, providing less than 0.25% informative characters across a total of 1130 aligned bases of these two commonly utilized sequences. The *pistillata* IVS 1 sequences provided 15 informative substitution characters and 4 stable indel characters (each 1–2 bp in length; a fifth indel was not scored because of overlapping gaps with different lengths) from 697 aligned bases for a total of 2.7% informative characters (Table 3). The results from the reduced *pistillata* dataset were stable to the different alignment parameters and phylogenetic analysis of this dataset identified a topology congruent with the full dataset (Fig. 2a), lacking any conflict among these 19 characters (L = 19, CI = 1.0, RI = 1.0).

DISCUSSION

These data identify uncorrected percentage divergences as much as seven times higher in the nuclear noncoding sequence nrDNA ITS (ITS-1, 0.0–24.2%) than in the chloroplast noncoding sequence *trnL* intron (0–4.2%). These findings are consistent with studies that have noted higher levels of sequence variation in nrDNA ITS relative to chloroplast noncoding regions (e.g., Gielly *et al.,* 1996; Small *et al.,* 1997). However, for *Sphaerocardamum* alone (whose study is our ultimate goal) neither of these commonly utilized sequences provided sufficient variation to warrant further use. Because our early results identified extremely low divergence levels among our ingroup ITS sequences, it was apparent that ITS was unlikely to provide information at our desired level. Due to this, we produced sequence confirmation only for the ITS-1 region, which has generally been noted to have diverged more than ITS-2 (e.g., Baldwin *et al.,* 1995; Gielly *et al.,* 1996; O'Kane *et al.,* 1996). Although the *trnL* intron was no more variable than nrDNA ITS, because the chloroplast genome is a single historical unit, it can be augmented with other chloroplast sequences to provide a better overall hypothesis for chloroplast haplotypes. For this reason (and because *trnL* intron may be used to help resolve relationships between the *Sphaerocardamum* and other family members) time and resources were expended to confirm double-stranded *trnL* intron sequences.

Our data indicate that *pistillata* IVS 1 was more variable than either nrDNA ITS or *trnL* intron, not only across the entire dataset but in the comparison of sequences from the ingroup. More importantly, this variation included informative characters and did not identify a clear sign of transition saturation with increasing divergence (Fig. 3), and phylogenetic analysis of the *pistillata* IVS 1 variation at the ingroup level lacked any character conflict. These results indicate that *pistillata* IVS 1 should permit the construction of robust gene trees for use in resolving ingroup relationships, our primary area of interest, and suggest that this intron may be useful for resolving relationships among closely related species in other groups of flowering plants.

Potential Problems

Working with large multigene families can be difficult due to the possibility of mistaking paralogous loci for orthologous loci. Analyses of vascular plant MADS-box genes indicate that *pistillata* sequences form a single well-supported clade (e.g., Münster *et al.,* 1997) relative to other MADS-box genes. However, a recent analysis of a somewhat larger sample of putative *pistillata* orthologues identified several possible *pistillata* duplication events within the angiosperms (Kramer *et al.,* 1998), suggesting that paralogy problems may be a concern when utilizing *pistillata.* In this study, we do not appear to have amplified any loci clearly paralogous to *pistillata.* BLAST searches on exon regions identified greatest similarity to *pistillata* sequences and each clone of *pistillata* IVS 1 was resolved on the tree together with the other clone(s) from each respective taxon. Our potentially troublesome finding of three distinct clones from *Capsella* ($2n = 32$) and *Nerisyrenia* ($2n = 36$) can be interpreted as fixed heterozygosity in these two putative tetraploids (Neuffer and Eschner, 1995; Rollins, 1993).

Although alignment was not particularly problematic for *pistillata* IVS 1 among members of the ingroup,

the large amount of apparent length variation from the more divergent sequences of the complete dataset created difficulties. Utilizing *pistillata* IVS 1 sequences among taxa that display relatively high divergences may be difficult. However, it should be noted that despite this length variation, and therefore large numbers of different possible alignments, the complete dataset showed relatively little homoplasy (CI = 75, RI = 87) and, with the exception of the placement of two taxa, produced the same topology regardless of the alignment parameters utilized.

Another potential problem with *pistillata* IVS 1 sequences is their high A + T content. Zhang and Hewitt (1997) noted two potential problems for the highly A + T-rich insect mitochondrial control region: (1) biased substitution toward A + T and (2) $(AT)_n$ microsatellite length variation. These can make initial homology assessment difficult. The *pistillata* sequences presented here shared relatively uniformly high A + T sequence bias. In addition, mapped substitutions from the ingroup *pistillata* dataset identified only 20 of the 38 total substitutions (autapomorphies and informative characters) as being either to an A or a T. These data do not support a seriously skewed substitution bias toward A + T (only 52.6%). In addition, a comparison of indel events across the same reduced dataset revealed 10 total indels, 7 associated with A or T mononucleotide repeats and 1 each for AC, AT, and TTA di/trinucleotide repeats. The four informative length variation characters from the ingroup *pistillata* dataset are all congruent with the phylogenetic hypothesis based on substitutions alone. Thus there appears to be neither serious substitution bias nor homoplaseous $(AT)_n$ length variation such as that noted by Zhang and Hewitt (1997).

Our results indicate that the *pistillata* intron contains information capable of being used to construct well-resolved gene trees using sequences from closely related species. This is an important prerequisite for phylogeny reconstruction at lower taxonomic levels. However, as with any single genetic locus, there is no guarantee that the gene tree will accurately reflect the history of the species from which the sequences are sampled. In particular, trans-specific polymorphisms could complicate interpretations of species relationships. Although the sampling in this study was not designed specifically to address this issue, our ongoing studies of *pistillata* IVS 1 in *Sphaerocardamum* have surveyed multiple accessions, and sequences from well defined species are coalescing within species (Bailey, unpubl. data). Of course this might not be the case in other genera.

Potential Broad-Scale Utility of Pistillata IVS 1

We have amplified *pistillata* sequences from a relatively broad range of Brassicaceae genera, demonstrating the potential value of *pistillata* IVS 1 for resolving lower level phylogenetic relationships. In addition, our results from the non-Brassicaceae amplifications suggest that at least three primer combinations, two utilizing a primer from within the large IVS 1 (pi504), may be widely useful. In nearly all successful amplifications, significant products were very similar in size to the known Brassicaceae sequences. In addition, another laboratory has used the F18 + F16 primer combination to amplify putative *pistillata* orthologues from members of the Asteraceae (J. Panero, pers. comm.). These results, in combination with the successful use of a primer from within the large IVS 1 (pi504), suggest that the loci orthologous to *pistillata* may possess a large IVS 1 with conserved flanking coding regions. Because most published *pistillata* orthologues are solely available as cDNAs, the only support for these hypotheses can be drawn from a comparison of *pistillata* to the *Antirrhinum* orthologue *globosa* (Troebner *et al.,* 1992). The comparison of *globosa* to *pistillata* (Fig. 4) reveals an intact MADS-box followed by an extra intron or shuffled section of the I coding region (relative to *pistillata*) found after the end of the MADS-box coding region. Based on this comparison, and on our amplification results, it appears that the structure of the *pistillata* transcription unit may be reasonably conserved.

Future studies utilizing *pistillata* or orthologous loci could sample the four downstream introns in addition to IVS 1. Although we have demonstrated that the large *pistillata* IVS 1 can provide information at lower taxonomic levels, it is possible that the downstream introns could display even greater divergence than IVS 1. This is suggested by the recent discovery that a MADS-box intervening sequence may contain a *cis*-acting transcription factor binding site (Sieburth and Meyerowitz, 1997). Sieburth and Meyerowitz have suggested that the large (2 kb) A + T-rich second intron of the *agamous* gene in *Arabidopsis* may contain one or more transcription factor binding sites which affect *agamous* regulation. The intron that contains these potential sites is positionally homologous with IVS 1 of *pistillata* (Ma *et al.,* 1991). If it is homologous, *pistillata* IVS 1 may contain these binding sites and be under selection pressure to maintain them. Therefore IVS 1 might be evolving at a slower rate than downstream introns lacking these sites. Amplification products generated using any of our forward primers in combination with either the pi1970R or the F16 primers amplifies the other *pistillata* introns.

These results identify *pistillata* IVS 1 as an excellent potential source of characters at lower taxonomic levels. *Pistillata* primer sites are potentially conserved, which may allow for use in other dicotyledonous and possibly monocotyledonous groups. These *pistillata* sequences do not appear to be complicated by paralogy

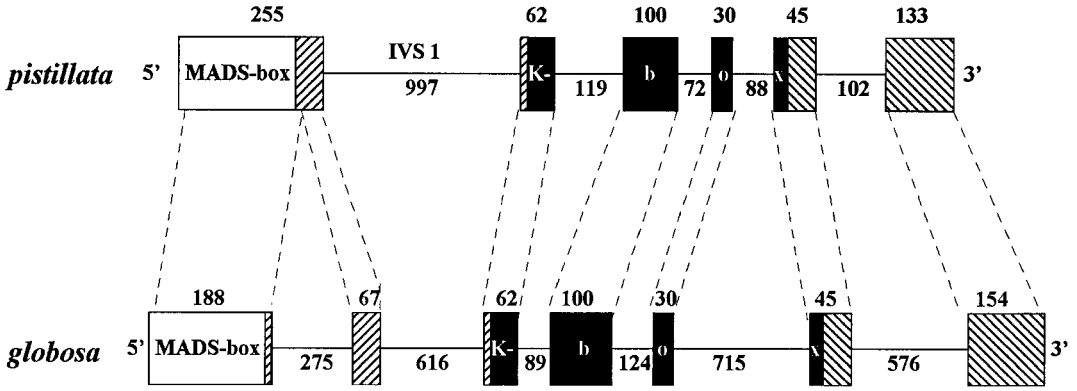

FIG. 4. *Pistillata* locus (adapted from Goto and Meyerowitz, 1994) comparison to the *globosa* locus in *Antirrhinum* (data from Troebner *et al.,* 1992). Exons are represented by boxes with their sizes (bp) designated above each respective region. Introns are represented by lines with their sizes designated below each respective region. Despite the much larger overall size (due to larger introns 5 and 6) and presence of an additional intron in the *globosa* gene, the structure of these two regions is generally conserved.

problems and ultimately *pistillata* IVS 1 sequences can provide more informative characters than the commonly utilized *trnL* intron or nrDNA ITS sequences.

ACKNOWLEDGMENTS

We thank E. Meyerowitz and E. Krizek for kindly providing primers which greatly assisted our progress in the early stages of this project, Jane Doyle for her assistance in the laboratory, and Helga Ochoterena for critical comments on the manuscript. We are grateful to Ihsan Al-Shehbaz, Bob Price, and Steve O'Kane for sharing their knowledge and unpublished results, aiding in our taxon sampling. We also thank Kevin Nixon for making current versions of DADA and Clados available and David Swofford for allowing us to use prerelease versions of PAUP* 4.0. This research was supported in part by an NSF Dissertation Improvement Grant DEB9701007, NSF Grants DEB9420215 and DEB 9614984, ASPT Graduate Student Research Grant, and a Mellon Foundation Graduate Research Grant.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment tool. *J. Mol. Biol.* **215:**403–410.

Atkinson, R. G., Cipriani, G., Whittaker, D. J., and Gardner, R. C. (1997). The allopolyploid origin of kiwifruit, *Actinidia deliciosa* (Actinidiaceae). *Plant Syst. Evol.* **205:** 111–124.

Baldwin, B. G., Sanderson, M. J., Porter, M. J., Wojciechowski, M. F., Campbell, C. S., and Donoghue, M. J. (1995). The ITS region of the nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Ann. Mo. Bot. Gard.* **82:** 247–277.

Buckler, E. S., IV, Ippolito, A., and Holtsford, T. P. (1997). The evolution of ribosomal DNA: Divergent paralogues and phylogenetic implications. *Genetics* **145:** 821–832.

Campbell, C. S., Wojciechowski, M. F., Baldwin, B. G., Alice, L. A., and Donoghue, M. J. (1997). Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). *Mol. Biol. Evol.* **14:** 81–90.

Clegg, M. T., Cummings, M. P., and Durbin, M. L. (1997). The evolution of plant nuclear genes. *Proc. Natl. Acad. Sci. USA* **94:**7791–7798.

Coen, E. S., Momero, J. M., Doyle, S., Elliot, R., Murphy, G., and Carpenter, R. (1990). *floricaula:* A homeotic gene required for flower development in *Antirrhinum majus. Cell* **63:** 1311–1322.

Denton, A. L., Hall, B. D., and McConaughy, B. L. (1996). *rpb2,* a nuclear gene for tracing angiosperm phylogeny. *Am. J. Bot.* **83S:** 150.

Doyle, J. J., and Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus* **12:** 13–15.

Doyle, J. J., Kanazin, V., and Shoemaker, R. C. (1996). Phylogenetic utility of *histone* H3 intron sequences in the perennial relatives of soybean (*Glycine:* Leguminosae). *Mol. Phylogenet. Evol.* **6:**438–447.

Doyle, J. J., and Doyle, J. L. (1997). The cytosolic glutamine synthetase gene family in Leguminosae: Gene phylogeny and evolution of its role in nodulation. *Am. J. Bot.* **84S:** 188.

Doyle, J. J. (1994). Evolution of a plant homeotic multigene family: Toward connecting molecular systematics and molecular developmental genetics. *Syst. Biol.* **43:** 307–328.

Fangan, B. M., Stedje, B., Stabbetorp, O. E., Jensen, E. S., and Jakobsen, K. S. (1994). A general approach for PCR-amplification and sequencing of chloroplast DNA from crude vascular plant algal tissue. *BioTechniques* **16:** 484–494.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39:** 783–791.

Galloway, G. L., Malmberg, R. L., and Price, R. A. (1997). Phylogenetic utility of the nuclear gene *argenine decarboxylase,* with an example from the Brassicaceae. *Am. J. Bot.* **84S:** 196.

Gielly, L., Yuan, Y. M., Kupfer, P., and Taberlet, P. (1996). Phylogenetic use of noncoding regions in the genus *Gentiana* L.: Chloroplast *trnL* (UAA) intron versus nuclear ribosomal internal transcribed spacer sequences. *Mol. Phylogenet. Evol.* **5:** 460–466.

Goloboff, P. (1993 Copyright; Upgraded 1996). Nona Computer Software version 1.5.1. Published by Author, NY.

Goto, K., and Meyerowitz, E. M. (1994). Function and regulation of the *Arabidopsis* floral homeotic gene *PISTILLATA. Genes Dev.* **8:** 1548–1560.

Kim, K. J., and Jansen, R. K. (1994). Comparison of phylogenetic hypotheses among different datasets in dwarf dandelions (*Krigia,* Asteraceae): Additional information from the internal transcribed spacer sequences of nuclear ribosomal DNA. *Plant Syst. Evol.* **190:** 157–185.

Kramer, E. M., Dorit, R. L., and Irish, V. F. (1998). Molecular evolution of genes controlling petal and stamen development:

Duplication and divergence within *APETALA3* and *PISTILLATA* MADS-box gene lineages. *Genetics* **149:** 765–783.

Lagercrantz, U. J., Putterill, J., Coupland, G., and Lydiate, D. (1996). Comparative mapping in *Arabidopsis* and *Brassica,* fine scale genome collinearity and congruence of genes controlling flowering time. *Plant J.* **9:** 13–20.

Li, W. H. (1997). "Molecular Evolution," Sinauer, Sunderland, MA.

Ma, H., Yanofsky, M. F., and Meyerowitz, M. E. (1991). AGL1-AGL6, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. *Genes Dev.* **5:** 484–495.

Mandel, M. A., Gustafson-Brown, C., Savidge, B., and Yanofsky, M. F. (1992). Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1. Nature* **360:** 273–277.

Mason-Gamer, R. J., and Kellogg, E. A. (1996). Potential utility of the nuclear gene *waxy* for plant phylogenetic analyses. *Am. J. Bot.* **83S:** 178.

Mathews, S., Lavin, M., and Sharrock, R. A. (1995). Evolution of the *phytochrome* gene family and its utility for phylogenetic analyses of angiosperms. *Ann. Mo. Bot. Gard.* **82:** 296–321.

Meyerowitz, E. M. (1994). Structure and organization of the *Arabidopsis thaliana* nuclear genome. *In* "Arabidopsis" (E. Meyerowitz and C. Sommerville, Eds.), pp. 21–36. Cold Springs Harbor Laboratory Press, Cold Springs Harbor, NY.

Morton, B. R., Gaut, B., and Clegg, M. T. (1996). Evolution of alcohol dehydrogenase gene in the Palm and Grass Families. *Proc. Natl. Acad. Sci. USA* **93:** 11735–11739.

Münster, T., Pahnke, J., Di Rosa, A., Kim, J. T., Martin, W., Saedler, H., and Theissen, G. (1997). Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants. *Proc. Natl. Acad. Sci. USA* **94:** 2415–2420.

Neuffer, B., and Eschner, S. (1995). Life-history and ploidy levels in the genus *Capsella* (Brassicaceae). *Can. J. Bot.* **73:** 1354–1365.

Nixon, K. C. (1993a). DADA matrix editing program. Published by author, Trumansburg, New York.

Nixon, K. C. (1993b). Clados computer software version 1.4. Published by author, Trumansburg, New York.

O'Kane, S. L., Jr., Schaal, B. A., and Al-Shehbaz, I. A. (1996). The origins of *Arabidopsis suecica* (Brassicaceae), as indicated by nuclear rDNA sequences. *Syst. Bot.* **21:** 559–566.

Olmstead, R. G., and Palmer, J. D. (1994). Chloroplast DNA systematics: A review of methods and data analysis. *Am. J. Bot.* **81:** 1205–1224.

Peralta, I. E., Ballard, H., and Spooner, D. M. (1997). "Waxy" gene intron phylogeny of tomatoes, Solanum subsect. *Lycopersicum* (Solanaceae). *Am. J. Bot.* **84S:** 222.

Purugganan, M. D., Roundley, S. D., Schmidt, R. J., and Yanofsky, M. F. (1995). Molecular evolution of flower development: Diversification of the plant MADS-box regulatory gene family. *Genetics* **140:** 345–356.

Rollins, R. C. (1984). *Sphaerocardamum* (Cruciferae). *Contrib. Gray Herbarium* **213:**11–17.

Rollins, R. C. (1993). "The Continental Cruciferae of North America," Stanford Univ. Press, Stanford, CA.

Sang, T., Donoghue, M. J., and Zhang, D. (1997). Evolution of alcohol dehydrogenase genes in Peonies (*Paeonia*): Phylogenetic relationships of putative nonhybrid species. *Mol. Biol. Evol.* **14:** 994–1007.

Sanderson, M. J., and Doyle, J. J. (1992). Reconstruction of organismal and gene phylogenies from data on multigene families: Concerted evolution, homoplasy, and confidence. *Syst. Biol.* **41:** 4–17.

Schilling, E. E., and Panero, J. L. (1996). Phylogenetic reticulation in subtribe Helianthinae. *Am. J. Bot.* **83:** 939–948.

Sieburth, L. E., and Meyerowitz, E. M. (1997). Molecular dissection of the AGAMOUS control region shows that *cis* elements for spatial regulation are located intragenically. *Plant Cell* **9:** 355–365.

Small, R., Ryburn, J., Cronn, R., Seelanan, T., and Wendel, J. F. (1997). Molecular phylogenetics of recently diverged groups: Evidence from nuclear and plastid sequences. *Am. J. Bot.* **84S:** 231–232.

Soltis, D. E., and Kuzoff, R. K. (1995). Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). *Evolution* **49:** 727–742.

Soltis, D. E., and Soltis, P. S. (1998). Choosing an approach and an appropriate gene for phylogenetic analysis. *In* "Molecular Systematics of Plants, II, DNA Sequencing" (D. E. Soltis, P. S. Soltis, and J. J. Doyle, Eds.), pp. 1–42. Kluwer Academic, Dordrecht.

Swofford, D. L. (unpubl.) PAUP*: Phylogenetic Analysis Using Parsimony. Computer program distributed by the Illinois Natural History Survey, Champaign, IL.

Taberlet, P., Gielly, L., Pautou, G., and Bouvet, J. (1991). Universal primers for amplification of three noncoding regions of chloroplast DNA. *Plant Mol. Biol.* **17:** 1105–1109.

Tandre, K., Albert, V. A., Sundas, A., and Engstrom, P. (1995). Conifer homologues to genes that control floral development in angiosperms. *Plant Mol. Biol.* **27:** 69–78.

Teutonico, R. A., and Osborn, T. C. (1994). Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage maps of B. *napus, B. oleracea,* and *Arabidopsis thaliana. Theor. Appl. Genet.* **89:** 885–894.

Theissen, G., Kim, J. T., and Saedler, H. (1996). Classification and phylogeny of the MADS-Box multigene family suggest defined roles of MADS-Box gene subfamilies in the morphological evolution of Eukaryotes. *J. Mol. Evol.* **43:** 484–516.

Troebner, W., Ramirez, L., Motte, P. H. I., Huijser, P., Loennig, W. E., Saedler, H., Sommer, H., and Schwarz-Sommer, Z. (1992). *GLOBOSA:* A homeotic gene which interacts with *DEFICIENS* in the control of *Antirrhinum* floral organogenesis. *EMBO J.* **11:** 4693–4707.

Wendel, J. F., Schnabel, A., and Seelanan, T. (1995). An unusual ribosomal DNA sequence from *Gossypium gossypioides* reveals ancient, cryptic, intergenomic introgression. *Mol. Phylogenet. Evol.* **4:** 298–313.

Wheeler, W., and Gladstein, D. (1993). MALIGN DNA sequence alignment software version 1.99. Published by authors, New York.

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

Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A., and Meyerowitz, E. M. (1990). The protein encoded by *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature* **346:** 35–39.

Zhang, D. X., and Hewitt, G. M. (1997). Insect mitochondrial control region: A review of its structure, evolution, and usefulness in evolutionary studies. *Biochem. Syst. Ecol.* **25:** 99–120.