American Journal of Botany 92(8): 1284–1293. 2005.

Duplication of floral regulatory genes in the Lamiales[1]

Jan E. Aagaard,[2,5] Richard G. Olmstead,[3] John H. Willis,[4] and Patrick C. Phillips[2]

[2]Center for Ecology and Evolutionary Biology, University of Oregon, Eugene, Oregon 97403 USA; [3]Biology Department, University of Washington, Seattle, Washington 98195 USA; and [4]Department of Biology, Duke University, Durham, North Carolina 27708 USA

Duplication of some floral regulatory genes has occurred repeatedly in angiosperms, whereas others are thought to be single-copy in most lineages. We selected three genes that interact in a pathway regulating floral development conserved among higher tricolpates (*LFY/FLO, UFO/FIM,* and *AP3/DEF*) and screened for copy number among families of Lamiales that are closely related to the model species *Antirrhinum majus*. We show that two of three genes have duplicated at least twice in the Lamiales. Phylogenetic analyses of paralogs suggest that an ancient whole genome duplication shared among many families of Lamiales occurred after the ancestor of these families diverged from the lineage leading to Veronicaceae (including the single-copy species *A. majus*). Duplication is consistent with previous patterns among angiosperm lineages for *AP3/DEF,* but this is the first report of functional duplicate copies of *LFY/FLO* outside of tetraploid species. We propose Lamiales taxa will be good models for understanding mechanisms of duplicate gene preservation and how floral regulatory genes may contribute to morphological diversity.

Key words: floral regulatory genes; gene duplication; Lamiales.

Morphological innovations are thought to frequently involve divergence among duplicated genes that comprise key genetic pathways regulating development (Raff, 1996; Carroll et al., 2001). Though most evidence comes from comparisons among animals, a similar pattern is beginning to emerge for flowering plants based primarily on studies of the MADS box family of transcriptional regulators (reviewed in Irish, 2003). Gene duplication has played a prominent role in the evolution of several families of MADS box genes. At least four distinct gene lineages predate land plants (>400 million years ago [mya]; reviewed in Theissen et al., 2000), and subsequent duplication has occurred within the seed plants (>200 mya; Kramer et al., 1998; reviewed in Theissen et al., 2000) and repeatedly within angiosperms (reviewed in Kramer et al., 2003). Floral homeotic genes belonging to the *AP3/DEF* clade of MADS box genes typify this trend. The ancestor of the orthologous genes *AP3* (Jack et al., 1992) and *DEF* (Sommer et al., 1990), which are necessary for development of petals and stamens in the model species *Arabidopsis thaliana* and *Antirrhinum majus*, respectively, duplicated within seed plants resulting in the ancestor of the *AP3/DEF* lineage as well as a second paralogous gene lineage also characterized in *A. thaliana* and *A. majus* (*PI/GLO*; Sundstrom et al., 1999; Aoki et al., 2004; Stellari et al., 2004). A later duplication of the ancestral *AP3/DEF* gene occurred near the base of the tricolpate (eudicot) lineage of angiosperms resulting in the *TM6* lineage (Kramer et al., 1998). Irish (2003) suggests duplication of several MADS box genes near the base of the tricolpates, including the ancestral *AP3/DEF* gene, may have contributed to key innovations such as the development of distinct petals and the fixation of whorled phyllotaxy.

In contrast to the radiation of the MADS box family, other floral regulatory genes appear to be quite conservative in copy number among angiosperms. For example, the orthologous genes *LFY* (Weigel et al., 1992) and *FLO* (Coen et al., 1990) are transcription factors characterized by a novel DNA-binding domain sufficient for specifying floral meristem identity in *A. thaliana* and *A. majus* and are derived from one of two paralogous genes found in gymnosperms, apparently the result of an ancient duplication within seed plants (Frohlich and Parker, 2000). *LFY/FLO* is thought to be a single copy gene in all extant angiosperms previously studied (Frohlich and Parker, 2000; reviewed in Cronk, 2001), with the notable exception of polyploid species (e.g., *Nicotiana tobaccum*, Kelly, 1995; *Zea mays*, Bomblies et al., 2003). Thus all *LFY/FLO*-like genes found in angiosperms appear to be orthologs of a single ancestral gene. Not surprisingly, given the apparent lack of duplication in the *LFY/FLO* gene lineage and its key position in the floral regulatory pathway (Parcy et al., 1998; Blazquez and Weigel, 2000), orthologs from *Z. mays* and tricolpates play similar roles in both meristem identity and regulating the expression of downstream genes (Bomblies et al., 2003).

Despite markedly different histories of duplication for *LFY/FLO* and *AP3/DEF,* these genes interact in a pathway essentially conserved among higher tricolpates (Ingram, 1995). Distinct from its role in floral meristem identity (Parcy et al., 1998), *LFY/FLO* positively regulates the expression of *AP3/DEF* in part through binding an upstream regulatory sequence of *AP3* (Lamb et al., 2002). A second gene, for which cognate homologs are well characterized in *A. thaliana* (*UFO*; Ingram et al., 1997) and *A. majus* (*FIM*; Simon et al., 1994), mediates *LFY/FLO* activity on *AP3/DEF* via ubiquitin-mediated degradation of transcriptional repressors of *AP3* (Zhao et al., 2001). Because of the conserved regulatory roles of *LFY/FLO, UFO/FIM,* and *AP3/DEF* during floral development, it is interesting to know whether duplication of one or more of these genes is correlated with duplication of other genes in the path-

[1] Manuscript received 13 September 2004; revision accepted 25 May 2005.
This work was supported by an NSF IGERT predoctoral fellowship and NSF dissertation improvement grant (DEB-0105176) to J. E. A., an NSF grant (DEB-0090313) to R. G. O., an NSF grant (DEB-0075704) to J. H. W., and an NIH grant to P. C. P. (GM54185).

[5] Author for correspondence (e-mail jaagaard@gs.washington.edu) present address: Department of Genome Sciences, University of Washington, Box 357730, Seattle, Washington 98195-7730 USA; phone 206-221-5124, fax 206)-685-7301.

way and if so, whether duplicate genes are retained jointly. Characterizing patterns of duplication is a first step toward the ultimate goal of understanding how the genes regulating floral development have contributed toward the diversification of floral morphology.

Previous work (Fishman et al., 2001) identified evidence of duplication for other floral regulatory genes in *Mimulus* species, a genus traditionally classified with the model species *A. majus* in the family Scrophulariaceae (Lamiales). Olmstead et al.'s (2001) molecular systematic study of taxa from Scrophulariaceae sensu lato and representatives of 15 other families of Lamiales led to a dramatic reorganization of the group. Their study found strong support for a monophyletic clade within a reclassified Veronicaceae that includes *A. majus* and suggested reclassification of other traditional Scrophulariaceae taxa to families such as Orobanchaceae. Significantly, the phylogenetic placement of *Mimulus* relative to other Lamiales families remains unclear despite additional molecular systematic studies (Beardsley and Olmstead, 2002). Given the dramatic reorganization of Scrophulariaceae (Olmstead et al., 2001), we undertook the present study to: (1) screen taxa historically classified with *A. majus* in Scrophulariaceae and taxa from related families for copy number of other floral regulatory genes including *LFY*/*FLO, UFO*/*FIM,* and *AP3*/*DEF*; (2) clarify when gene duplication occurred relative to the divergence among several lineages of the order Lamiales and *A. majus*; and (3) describe the patterns of gene duplication and retention among lineages for all three genes. By focusing on *LFY*/*FLO, UFO*/*FIM,* and *AP3*/*DEF* together, we examine duplication of multiple interacting members comprising a portion of a signaling pathway previously studied only individually (*LFY*/*FLO, AP3*/*DEF*) or for which little is known outside of model species (*UFO*/*FIM*).

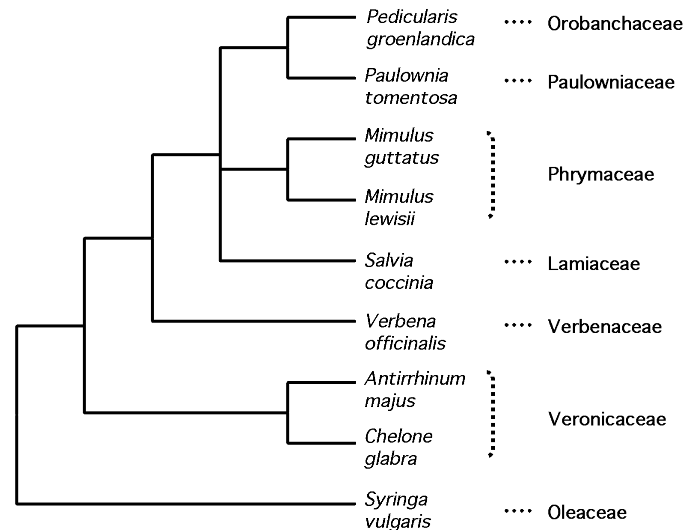

Fig. 1. Schematic depicting the phylogenetic relationships among species in the order Lamiales from which *LFY/FLO*-like, *UFO/FIM*-like, and *AP3/DEF*-like genes were cloned. Relationships among most taxa and familial classification are based on Olmstead et al. (2001). Familial classification of Phrymaceae is from Beardsley and Olmstead (2002).

MATERIALS AND METHODS

Identifying floral regulatory genes in Lamiales—We selected nine species in the order Lamiales based on classifications from recent phylogenetic studies (Olmstead et al., 2001; Beardsley and Olmstead, 2002; Fig. 1, Table 1) to bracket the phylogenetic placement of the model species *Antirrhinum majus* as well as *Mimulus* from which duplicates of floral regulatory genes were found previously (Fishman et al., 2001). Selected taxa include *Syringa vulgaris* (Oleaceae), *Chelone glabra* and *Antirrhinum majus* (Veronicaceae), *Verbena officinalis* (Verbenaceae), *Salvia coccinia* (Lamiaceae), *Mimulus guttatus* and *Mimulus lewisii* (Phrymaceae), *Paulownia tomentosa* (Paulowniaceae), and *Pedicularis groenlandica* (Orobanchaceae; Table 1). Source material was obtained from the University of Washington medicinal herb gardens (*Syringa vulgaris*, *Chelone glabra*, *Verbena officinalis*, *Salvia coccinia*, and *Paulownia tomentosa*), Fishman et al. (2001; *Mimulus guttatus*), Bradshaw et al. (1995; *Mimulus lewisii*), John Innes Centre (Norfolk, UK; *Antirrhinum majus*) or was wild collected (*Pedicularis groenlandica*; J. Aagaard 2003–1, WTU). For each species, genomic DNA was prepared from leaf material of a single plant using the cetyltrimethylammonium bromide (CTAB) extraction method of Kelly and Willis (1998). Total RNA was prepared from whole flower buds of the same plants across a range of developmental stages using the RNeasy Plant Mini kit (Qiagen, Valencia, California, USA) and used to make 5′- and 3′-RACE-Ready cDNA pools (Clontech, Palo Alto, California, USA).

Homologs of *LFY*/*FLO* were cloned from all nine taxa within Lamiales using a two-step approach that (1) screened for duplicate copies based on length polymorphism in introns from genomic DNA and (2) tested for expression of genes while also extending coding sequence using 5′- and 3′-RACE from floral bud cDNA pools. (1) Two primer sets nested within conserved exon domains spanning the first (5′ primer 5′-ATGAGGGATGAGGAGCTTGATSANATGATGRA-3′, 3′ primer 5′-GCTCCGTCACGATAAANGGRTGYT-3′) and second (5′ primer 5′-CGGCAGCGGGAGCAYCCNTTYAT-3′, 3′ primer 5′-GCGTTGAAGATCGCRTCDATRTCC-3′) introns were designed from an alignment of *LFY/FLO, Petunia hybrida ALF,* and *Solanum lycopersicon FA* (Table 1). Polymerase chain reaction was performed (25 μL reaction volume) using 20 ηg genomic DNA, 50 mM KCl, 10 mM Tris pH 8.4, 2.5 mM $MgCl_2$, 0.06% BSA, 0.2 mM dNTPs, 5 ρmol of each primer, and 1 unit Promega *Taq* polymerase (Madison, Wisconsin, USA). Thermal cycler conditions included an initial denaturation step of 2 min at 94°C followed by 37 cycles of 20 s at 94°C, 20 s at 58°C, 90 s at 72°C, and a final extension at 72°C for 5 min. The PCR products were run on 2% Tris-borate-EDTA (TBE) gels, all visible bands excised and purified using the GENECLEAN kit (Qbiogene, Carlsbad, California, USA), and cloned using the TOPO TA cloning kit (Invitrogen, Carlsbad, California, USA). A minimum of four clones of each PCR fragment was sequenced using automated fluorescent sequencing methods (Applied Biosystems, Foster City, California, USA). (2) 3′-RACE primers were designed from sequence of all genomic clones of *LFY/FLO*-like sequences spanning the first intron, and 3′-RACE PCR carried out using the SMART RACE cDNA amplification kit (Clontech, Palo Alto, California, USA). 3′-RACE products were cloned directly and sequenced (four or more clones) as above for genomic PCR fragments. *LFY/FLO*-like 3′-RACE sequences were compared with the coding region of genomic fragments spanning the second intron cloned earlier; for unique *LFY/FLO*-like genomic sequences spanning intron 2, 5′- and 3′-RACE was used to extend coding sequence of clones as before for intron 1 genomic clones.

Homologs of *UFO*/*FIM* were cloned from all nine taxa within Lamiales using a two-step approach similar to that used for *LFY/FLO*. (1) Two overlapping degenerate primer sets nested within conserved domains of the intronless *UFO*/*FIM* gene were designed from the consensus sequence of *UFO* and *FIM* (first 5′primer 5′-CGAGCTCGATCAGTCTGTAARMGNTGGTA-3′, first 3′primer 5′-CGCTGAACGGGCTGTAGTTCATRCARTARAA-3′; second 5′ primer 5′-CCGCCATTGGTTCCTCTTCTTYAARCARCA-3′, second 3′primer 5′-TCTGCAAACTGATTGTACAGYTGYTGNGGCAT-3′). The PCR was performed on genomic DNA using both primer sets and an identical protocol as for *LFY/FLO.* The PCR fragments were purified, cloned, and at least four clones of each fragment sequenced as before. (2) Sequence of genomic clones of all *UFO/FIM*-like fragments from *UFO/FIM* primer sets 1 and 2 were compared and 3′-RACE primers designed for each unique sequence; 3′-RACE was carried out for *LFY/FLO* and four or more clones of each 3′-RACE product sequenced.

Lamiales homologs of *AP3*/*DEF* were cloned from all nine species directly

TABLE 1. Homologs of *LFY/FLO*, *UFO/FIM*, and *AP3/DEF* cloned previously or in this study. Gene names conform to the cited reference for previously cloned homologs or follow the nomenclature for *Antirrhinum majus* if cloned in this study. For species with more than one homolog, an A or B designation describes the paralogy group within which genes cluster in phylogenetic analyses. Species with multiple A or B copies are given a further numeric designation. An abbreviation incorporating both species and gene names is used to label terminal branches in gene trees.

Species	Gene	Abbreviation	Reference	GenBank accession
Zea mays (Poaceae)	*ZFL1*	*ZeaMaZFL1*	Bomblies et al., 2003	AY179883
	SLK1	*ZeaMaSLK1*	Ambrose et al., 2000	AF181479
Arabidopsis thaliana (Brassicaceae)	*LFY*	*AraThLFY*	Weigel et al., 1992	M91208
	UFO	*AraThUFO*	Ingram et al., 1995	X89224
	AP3	*AraThAP3*	Jack et al., 1992	A42095
Petunia hybrida (Solanaceae)	*ALF*	*PetHyALF*	Souer et al., 1998	AF030171
	TM6	*PetHyTM6*	Kramer et al., 1998	AF230704
	PMADS1	*PetHyPMADS1*	Kush et al., 1993	X69946
Solanum esculentum (Solanaceae)	*FA*	*LycEsFA*	Molinero-Rosales et al., 1999	AF197935
	TM6	*LycEsTM6*	Pnueli et al., 1991	X60759
	AP3	*LycEsAP3*	Kramer et al., 1998	AF052868
Syringa vulgaris (Oleaceae)	*FLO*	*SyrVuFLO*	this study	AY524037
	FIM	*SyrVuFIM*	this study	AY524021
	DEF	*SyrVuDEF*	this study	AY524007
Antirrhinum majus (Veronicaceae)	*FLO*	*AntMaFLO*	Coen et al., 1990	M55525
	FIM	*AntMaFIM*	Simon et al., 1994	S71192
	DEF	*AntMaDEF*	Sommer et al., 1990	S12378
Chelone glabra (Veronicaceae)	*FLO*	*CheGlFLO*	this study	AY524029
	FIM	*CheGlFIM*	this study	AY524022
	DEF	*CheGlDEF*	this study	AY524008
Verbena officinalis (Verbenaceae)	*FLOA*	*VerOfFLOA*	this study	AY524030
	FLOB	*VerOfFLOB*	this study	AY524038
	FIM	*VerOfFIM*	this study	AY524023
	DEFA	*VerOfDEFA*	this study	AY524009
	DEFB	*VerOfDEFB*	this study	AY524014
Salvia coccinia (Lamiaceae)	*FLOA1*	*SalCoFLOA1*	this study	AY524031
	FLOA2	*SalCoFLOA2*	this study	AY524032
	FIM	*SalCoFIM*	this study	AY524024
	DEFB1	*SalCoDEFB1*	this study	AY524015
	DEFB2	*SalCoDEFB2*	this study	AY524016
Pedicularis groenlandica (Orobanchaceae)	*FLOA*	*PedGrFLOA*	this study	AY524033
	FLOB	*PedGrFLOB*	this study	AY524039
	FIM	*PedGrFIM*	this study	AY524025
	DEFA	*PedGrDEFA*	this study	AY524010
	DEFB	*PedGrDEFB*	this study	AY524017
Paulownia tomentosa (Paulowniaceae)	*FLOA*	*PauToFLOA*	this study	AY524034
	FLOB	*PauToFLOB*	this study	AY524040
	FIM	*PauToFIM*	this study	AY524026
	DEFA	*PauToDEFA*	this study	AY524011
	DEFB	*PauToDEFB*	this study	AY524018
Mimulus lewisii (Phyrmaceae)	*FLOA*	*MimLeFLOA*	this study	AY524035
	FLOB	*MimLeFLOB*	this study	AY524041
	FIM	*MimLeFIM*	this study	AY524027
	DEFA	*MimLeDEFA*	this study	AY524013
	DEFB	*MimLeDEFB*	this study	AY524019
Mimulus guttatus (Phyrmaceae)	*FLOA*	*MimGuFLOA*	this study	AY524036
	FLOB	*MimGuFLOB*	this study	AY524042
	FIM	*MimGuFIM*	this study	AY524028
	DEFA	*MimGuDEFA*	this study	AY524012
	DEFB	*MimGuDEFB*	this study	AY524020

from 3′-RACE cDNA pools. A single degenerate primer at the 5′ terminus of the *AP3/DEF*-coding region was designed from the consensus sequence of *AP3/DEF, Petunia hybrida PMADS1,* and *Solanum lycopersicon AP3* (Table 1; 5′-ATGGCTCGTGGGAAGATHCARAT-3′). The 3′-RACE products were cloned, and 12 or more clones of each fragment were sequenced as for *LFY/FLO* and *UFO/FIM.*

Phylogenetic analyses of floral regulatory genes—Coding regions of all *LFY/FLO*- and *AP3/DEF*-like genes we identified were aligned with homologs from *Zea mays*, *A. thaliana*, *Petunia hybrida*, *Solanum lycopersicon*, and published sequences from *A. majus* (Table 1). We used only one of the paralogous *LFY/FLO* loci from *Z. mays* (*ZFL1*) that are believed to be the result of the tetraploid ancestry of *Zea* (Gaut and Doebley, 1997). Coding regions of all *UFO/FIM*-like genes we cloned were aligned with homologs from *A. thaliana* and *A. majus* as well as several other *UFO/FIM*-like homologs. These include F-box genes from *Pisum sativum* (*PisSaSTP*, accession AF004843; Taylor et al., 2001), *Lotus japonica* (*LotJaPFO*, accession AY156687; Zhang et al., 2003), and *Impatiens balsamina* (*ImpBaFIM*, accession AF047392; Pouteau et al., 1998), as well as a *UFO/FIM*-like expressed sequence tag (EST) from a *Solanum lycopersicon* floral bud library (*LycEsEST*, accession BI423409). The *Solanum lycopersicon UFO/FIM*-like EST was the highest scoring hit from BLAST searches of EST databases from *Solonaceae* taxa (Plant Genome

Database, http://plantgdb.org) searched using *A. majus FIM* and all Lamiales *FIM*--like genes we cloned. Because published *A. majus* sequences for all three genes were identical to those we identified in our cloning experiments (see below) and included additional sequence 5′ of ours, we excluded our *A. majus* sequences from further analyses. Nucleotide alignments were carried out initially based on the translated nucleotide (protein) sequence using the ClustalX algorithm implemented in BioEdit (T. Hall, North Carolina State University, Raleigh, North Carolina, USA), followed by visual alignment. Areas with ambiguous alignments were excluded, and the nucleotide alignment was analyzed using likelihood criterion implemented in PAUP* (Swofford, 2002). Likelihood analyses employed the general time reversible model with four rate categories, estimating the gamma shape parameter and the proportion of invariable sites (GTR + 1 + γ). Heuristic search criterion included tree-bisection-reconnection (TBR) branch swapping with 10 random addition replicates. Support for nodes was estimated by 100 bootstrap replicates using the maximum likelihood estimates of substitution parameters and heuristic search criterion described previously.

Tests of alternate topologies—Likelihood-based tests of alternate topologies (Goldman, 1993; Goldman et al., 2000) were used to test different hypotheses regarding the timing of duplication of multiple *LFY/FLO*-like and *AP3/DEF*-like copies found in the Lamiales. Likelihood scores (−lnL) of topologies and maximum likelihood (ML) estimates of substitution parameters (GTR + 1 + γ) were calculated under (1) one constraint topology for *LFY/FLO* and (2) three constraint topologies for *AP3/DEF* using PAUP* (Swofford, 2002). (1) The constraint topology for *LFY/FLO* (*LFY/FLO* $H_0$1) collapsed the nodes within paralogous gene clades with bootstrap support less than 95% and placed *Antirrhinum majus FLO* (*AntMaFLO*) along with *Chelone glabra FLO* (*CheGlFLO*) and *Syringa vulgaris FLO* (*SyrVuFLO*) at an unresolved node sister below two paralogous gene clades. We excluded *Z. mays ZFL1* from the analyses because of problems with long branch attraction (see below). (2) Constraint 1 (*AP3/DEF* $H_0$1) collapsed nodes within the paralogous gene clades of the maximum likelihood tree having bootstrap support less than 95% and moved *Syringa vulgaris DEF* (*SyrVuDEF*) to the arbitrarily defined A clade. Constraint 2 (*AP3/DEF* $H_0$2) is identical to constraint 1 except *SyrVuDEF* is placed within the B clade. Constraint 3 (*AP3/DEF* $H_0$3) collapsed nodes in a similar manner and placed *Chelone glabra* (*CheGlDEF*) and *Antirrhinum majus DEF* (*AntMaDEF*) along with *SyrVuDEF* at an unresolved node sister below the paralogous gene clades (Fig. 3). The test statistic $\delta = (-\ln L_{-ML}) - (-\ln L_{MLH_0})$ was calculated for each constraint, where $(-\ln L_{-ML})$ is the likelihood score maximized over all topologies, and $(-\ln L_{MLH_0})$ is the likelihood score given the topological constraint (Goldman et al., 2000).

Simulated nucleotide sequences for 19 species (N = 400 replicates) for each *LFY/FLO* and *AP3/DEF* constraint were generated using SeqGen (Rambaut and Grassley, 1997). Simulated data used the ML estimates of substitution parameters and topology (including branch lengths) under H_0 (GTR + 1 + γ), and were identical in sequence length to the original alignments used in phylogenetic analyses (*LFY/FLO* = 783 nucleotides, *AP3/DEF* = 666 nucleotides). Simulated data were analyzed according to Goldman et al. (2000) using an approximation under H_A (*posPpud*) in which substitution parameters (GTR + 1 + γ) are provided when maximizing the likelihood over all topologies, and the test statistic $\delta = (-\ln L_{-ML}) - (-\ln L_{MLH_0})$ is calculated for each simulated data set.

RESULTS

Copy number of floral regulatory genes among Lamiales—A single copy of a *LFY/FLO*-like gene was found for both *Syringa vulgaris* and *Antirrhinum majus* (Table 1). Overlapping sequence from 3′-RACE clones corresponding to genomic sequence amplified using first and second intron primer sets were identical. Similarly, a single copy of an *AP3/DEF*-like gene was found for both species based on our copy number assay using 3′-RACE directly. For both genes, the *A. majus* clones we identified were identical to previously published *FLO* (Coen et al., 1990) and *DEF* (Sommer et al., 1990) sequences. In contrast, two copies of both *LFY/FLO*-like and *AP3/DEF*-like genes were found for seven taxa representing five families within Lamiales: *Chelone glabra* (Veronicaceae), *Verbena officinalis* (Verbenaceae), *Salvia coccinia* (Lamiaceae), *Paulownia tomentosa* (Paulowniaceae), *Pedicularis groenlandica* (Orobanchaceae), and *Mimulus lewisii* and *M. guttatus* (Phrymaceae; Table 1). For *LFY/FLO*-like genes, the first intron primer set amplified both copies from most species except *P. groenlandica*, for which we found a single copy using this primer combination. The second intron primer set amplified two copies from *V. officinalis* and *P. tomentosa* but a single copy from all other species. Comparison of sequence from 3′-RACE experiments showed clones corresponding to genomic sequence from first and second intron primer sets were identical for at least one copy except for *P. groenlandica*. Thus first and second intron primer sets together identified two distinct *LFY/FLO*-like genes from *P. groenlandica*.

The single copy of *LFY/FLO*-like and *AP3/DEF*-like genes we found for *S. vulgaris* as well as duplicate copies found for *V. officinalis*, *P. tomentosa*, *P. groenlandica*, *M. lewisii*, and *M. guttatus* appear functional based on correctly spliced cDNAs. Duplicates are well diverged at synonymous codon positions, on average having 44% (*LFY/FLO*-like) and 38% (*AP3/DEF*-like) of silent sites with substitutions. Duplicate copies of *LFY/FLO*-like and *AP3/DEF*-like genes found for *S. coccinia* also appear functional based on cDNAs, but divergence at silent sites is approximately half that seen for other duplicates (19% and 17% for *S. coccinia LFY/FLO*-like and *AP3/DEF*-like genes, respectively). In contrast with other *LFY/FLO*-like duplicates, one copy of both *LFY/FLO*-like and *AP3/DEF*-like genes in *C. glabra* may be nonfunctional as evidenced by no detectable expression in cDNA pools (*LFY/FLO*-like) or premature stop codons found in cDNAs (*AP3/DEF*-like). Duplicates from *C. glabra* that appear nonfunctional (and were sister to the functional copies; see below) were excluded from further analyses.

A single copy of an *UFO/FIM*-like gene was found within all taxa we studied (Table 1). Sequence from genomic clones of both primer sets were identical within the region they overlapped (approximately 500 nucleotides) for all species, and the *A. majus* sequence we identified was identical to previously published *FIM* (Simon et al., 1994). Therefore 3′-RACE using primers designed only from genomic clones of the 5′primer sets were used to obtain cDNA sequences for all taxa, with the exception of *P. groenlandica*. We found no evidence of expression in our floral bud cDNA pools from *P. groenlandica*, and additional 3′- as well as 5′-RACE experiments using multiple nested primer sets showed no evidence of expression in any of three cDNA pools constructed from individual *P. groenlandica* plants collected from separate populations. This suggests the *UFO/FIM*-like genomic sequence from *P. groenlandica* either (1) may not be expressed at detectable levels in the floral tissues we assayed or (2) may correspond to a nonfunctional copy, although there is no evidence of missense or stop codons (see below).

Phylogenetic analyses of floral regulatory genes—LFY/FLO—After excluding ambiguous regions, our alignment of *LFY/FLO* homologs from 13 species includes approximately 783 nucleotides from each of 19 homologs (Table 1). Based on this alignment, likelihood criterion found a maximum likelihood tree (−lnL = 6520.4862) placing *Zea mays ZFL1* sister

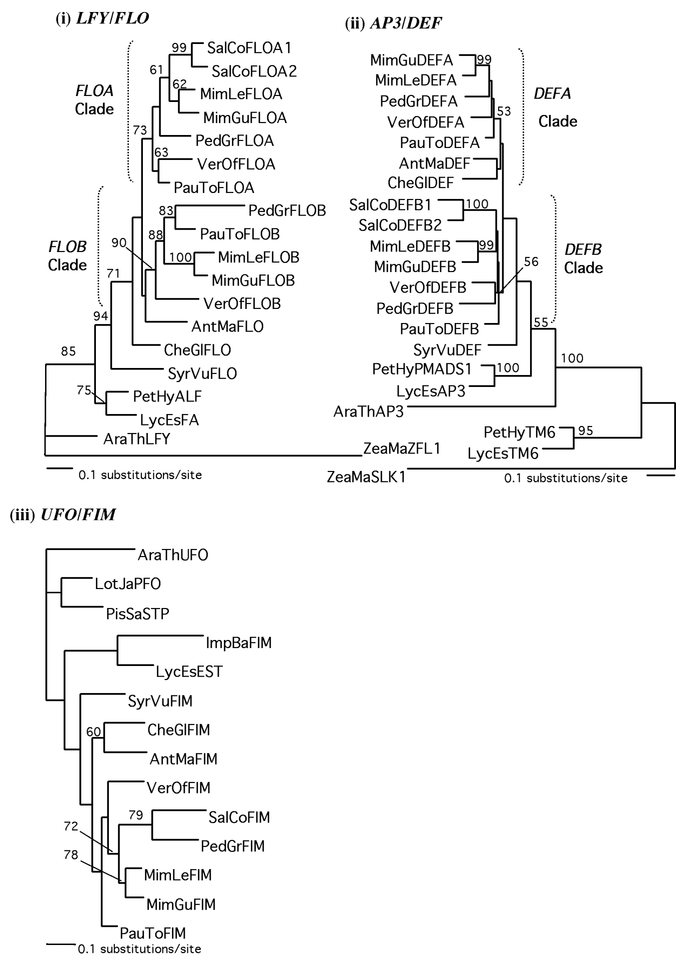

Fig. 2. Likelihood trees for homologs of (i) *LFY/FLO,* (ii) *AP3/DEF,* and (iii) *UFO/FIM* cloned in this study and previously (see Table 1). Paralogy groups of *LFY/FLO* and *AP3/DEF* within the Lamiales are arbitrarily designated *A* or *B.* Likelihood searches were performed under the general time reversible model with four rate categories using PAUP* (Swofford, 2002). All ambiguous regions of alignments were excluded from analyses. Bootstrap support based on 100 bootstrap replicates is shown for nodes with ≥50% support. Because of problems with the long branch leading to *ZeaMaZFL1,* bootstrap support for nodes in the *LFY/FLO* tree was calculated from the maximum likelihood tree that excluded *ZFL1.*

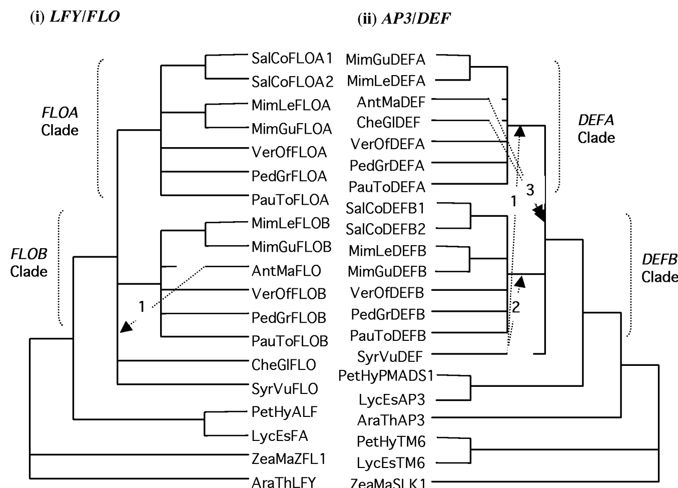

Fig. 3. Constraint trees showing alternative (i) *LFY/FLO* and (ii) *AP3/DEF* topologies tested using the parametric bootstrap approach of Goldman et al. (2000). Topologies are based on the likelihood trees (Fig. 2), with weakly supported nodes within paralogy groups A and B collapsed. (i) 1 = *LFY/FLO* $H_0$1 places *AntMaFLO* along with *CheGlFLO* and *SyrVuFLO* basal to *FLOA* and *FLOB* and implies a duplication of *DEF* after the divergence between Veronicaceae and other families of the Lamiales. (ii) 1 = *AP3/DEF* $H_0$1 places *SyrVuDEF* within the *DEFA* clade; 2 = *AP3/DEF* $H_0$2 places *SyrVuDEF* within the *DEFB* clade; 3 = *AP3/DEF* $H_0$3 places *AntMaDEF* + *CheGlDEF* as sister to *DEFA* + *DEFB* clades. Both *AP3/DEF* $H_0$1 and $H_0$2 imply a duplication of *AP3/DEF* near the base of the Lamiales, while *AP3/DEF* $H_0$3 implies a duplication of *DEF* after the divergence between Veronicaceae and other families of the Lamiales.

to one of two *S. coccinia LFY/FLO* homologs. However, this topology appears to be the effect of the long branch leading to *ZFL1* because an alternate tree of marginally lower score (−lnL = 6523.76183; Fig. 2) was found that has a topology identical to that of the maximum likelihood tree when *ZFL1* is excluded from the analysis. This alternate likelihood tree (Fig. 2) places all *LFY/FLO* homologs from taxa within Lamiales in a well-supported clade (94% bootstrap support) sister to *LFY/FLO* homologs from the Solanales. Within Lamiales, the *S. vulgaris* homolog branches first, sister to a clade comprising the remaining sequences (71%), followed by the single expressed (see above) *C. glabra* homolog. When both *C. glabra LFY/FLO*-like genes are included, they form a well-supported clade (100% bootstrap support; data not shown). Two gene clades belong sister, each including one copy of the *LFY/FLO* homologs from *V. officinalis*, *P. tomentosa*, *P. groenlandica*, *M. lewisii*, and *M. guttatus*. *Antirrhinum majus FLO* is sister to one of these gene clades (*FLOB*), though its placement is weakly supported (49% bootstrap support). The two *LFY/FLO* homologs from *S. coccinia* cluster together as a well-supported sister clade (99%) within the *FLOA* clade.

We tested an alternate topology for *LFY/FLO* homologs using the likelihood-based approach of Goldman et al. (2000; Fig. 3). The alternate topology places *A. majus FLO* sister to the two paralogy groups in Lamiales (*LFY/FLO* $H_0$1). The maximum likelihood tree given this constraint is only slightly lower (−lnL = 5903.04329), although the test statistic comparing the constrained tree with that placing *A. majus FLO* within the *FLO* B clade (δ = 1.8) is near the significance threshold ($P = 0.03$; Fig. 4). This tight distribution apparently reflects the very short branch leading to the node placing *A. majus FLO* sister to the *FLOB* clade (Fig. 2) and the high proportion of invariant sites (36%) in the model used to simulate data.

AP3/DEF—Our analyses examined the phylogenetic placement of tricolpate *AP3/DEF* orthologs relative to the homologous gene from *Z. mays* (*SLK1*) and the *TM6* gene lineage paralogous to *AP3/DEF* (Table 1). Similar to the *LFY/FLO* alignment, all ambiguous regions of the *AP3/DEF* alignment were excluded from phylogenetic analyses (final alignment approximately 666 nucleotides). Likelihood criterion found a maximum likelihood tree placing all Lamiales *AP3/DEF* homologs within a well-supported clade sister to AP3/DEF homologs from the Solanales (Fig. 2). As with the *LFY/FLO* tree, the *S. vulgaris AP3/DEF* homolog we cloned is sister to the rest of the of the clade. Two gene clades each contain one of two *AP3/DEF* homologs from *V. officinalis*, *P. tomentosa*, *P. groenlandica*, *M. lewisii*, and *M. guttatus*. *Antirrhinum majus DEF* and the single expressed (see above) *C. glabra AP3/DEF*

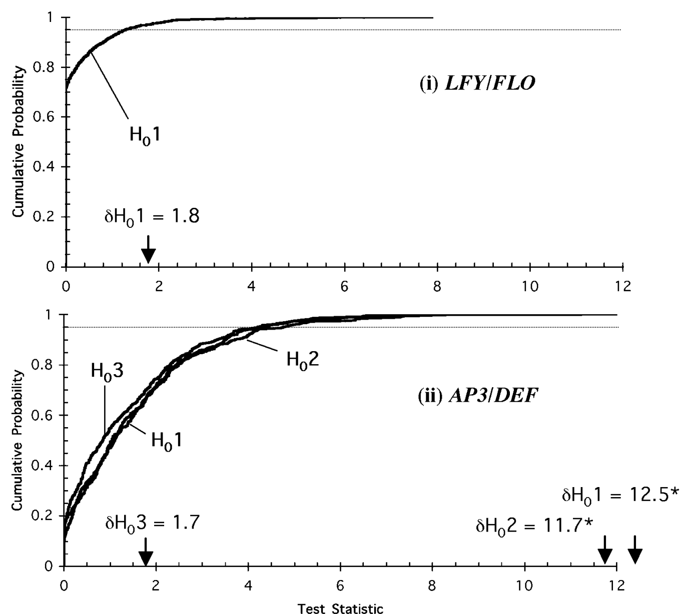

Fig. 4. Distributions of test statistics from parametric bootstrapping for (i) *LFY/FLO* $H_0 1$ and (ii) *AP3/DEF* $H_0 1$, $H_0 2$, and $H_0 3$. Nucleotide substitution parameters from the likelihood search under the specified topological constraints (Fig. 3) were estimated under the general time reversible model using PAUP* (Swofford, 2002) and used to generate 1400 (*LFY/FLO* $H_0 1$) or 400 (*AP3/DEF* $H_0 1$, $H_0 2$, and $H_0 3$) simulated data sets using SEQGEN (Rambaut and Grassley, 1997). Simulated data were analyzed using test *posPpud* approximated under H_A (Goldman et al., 2000). The test statistics for each simulated data set *i* were calculated ($\delta = L_{ML}{}'^{(i)} - L_{constrained}{}^{(i)}$), and the cumulative probability of test statistics plotted. Null hypotheses are rejected for attained δ's that exceed 95% of δ's for simulated data (a one-sided test with 5% significance level; dashed line).

homolog are included in one of the two paralogous clades (DEFA) with weak bootstrap support (53%). When both *C. glabra AP3/DEF*-like genes are included, they form a well-supported clade sister to *A. majus DEF* (100% bootstrap support; data not shown). The *AP3/DEF* homologs from *S. coccinia* also form a well-supported (100%) sister clade nested within the DEFA clade. When we excluded *SLK1* from phylogenetic analyses using likelihood to assess possible long branch effects seen for *LFY/FLO* homologs, a topology identical to that of Fig. 2 was found.

We investigated three alternate topologies for *AP3/DEF* homologs using the likelihood based approach of Goldman et al. (2000; Fig. 3). The first places *S. vulgaris AP3/DEF* within the *DEFA* clade (*AP3/DEF* $H_0 1$) along with *A. majus DEF* and *C. glabra AP3/DEF*. The maximum likelihood score given this constraint ($-\ln L = 6989.50125$) is significantly lower than the unconstrained maximum likelihood tree, and the test statistic ($\delta = 12.5$) lies well beyond 95% of test statistics from data simulated under the constraint (Fig. 4). Similarly, an alternate topology placing *S. vulgaris* within the *DEFB* clade (*AP3/DEF* $H_0 2$) is rejected ($\delta = 11.7$; Fig. 4). However, the third topology placing Veronicaceae homologs (*A. majus DEF* and *C. glabra AP3/DEF*-like) within a clade sister to paralogy groups *DEFA* and *DEFB* cannot be rejected, because the test statistic (= 1.7) lies near the 50th percentile of data simulated under the constraint topology of *AP3/DEF* $H_0 3$ (Fig. 4).

UFO/FIM—Phylogenetic analyses of our alignment (approximately 1113 nucleotides, excluding all ambiguous regions) of *UFO, FIM, UFO/FIM*-like homologs or ESTs, and the eight *UFO/FIM*-like genes we identified among Lamiales taxa show generally weakly supported topologies based on likelihood criterion. Likelihood analyses found a maximum likelihood tree placing *FIM* and all *UFO/FIM*-like genes we cloned from Lamiales taxa in a weakly supported monophyletic clade sister to other Asterid *UFO/FIM*-like genes, including the *UFO/FIM*-like EST from *S. esculentum* (*LycEsEST*; Fig. 2). This is strong evidence the *UFO/FIM*-like genes we cloned from Lamiales taxa are orthologous to *A. majus FIM*. Within the Lamiales clade, the *S. vulgaris UFO/FIM*-like gene branches first. *UFO/FIM*-like homologs from the Veronicaceae including *FIM* form a weakly supported clade sister to the remaining Lamiales *UFO/FIM*-like homologs. Several of the nodes within this clade have moderate bootstrap support including *Mimulus UFO/FIM*-like genes (78%) and a clade with the strongest bootstrap support across the tree (79%) consisting of the *S. coccinia UFO/FIM*-like homolog and the gene from *P. groenlandica*, for which we found no evidence of expression within multiple floral bud cDNAs.

DISCUSSION

Multiple duplications of LFY/FLO and AP3/DEF in Lamiales—We have identified at least two independent duplications of *LFY/FLO* and *AP3/DEF* homologs within the Lamiales. The first occurred within the Veronicaceae (sensu Olmstead et al., 2001), after the split between the lineage leading to *Antirrhinum majus* and *Chelone glabra*. This duplication is likely the result of a whole genome duplication (polyploidization) as evidenced by the approximate doubling of chromosomes in *C. glabra* ($N = 14$) relative to *A. majus* ($N = 8$; Goldblatt and Krukoff, 1984). In this case, one duplicate of each gene appears to have been silenced based on no detectable expression in RACE cDNA pools (*LFY/FLO*) or premature stop codons (*AP3/DEF*). We also found two copies of both *LFY/FLO*-like and *AP3/DEF*-like genes in representative taxa from Verbenaceae (*Verbena officinalis*), Phrymaceae (sensu Beardsley and Olmstead, 2002; *Mimulus guttatus* and *M. lewisii*), Paulowniaceae (*Paulownia tomentosa*), and Orobanchaceae (*Pedicularis groenlandica*; Fig. 1). Duplicate copies in these taxa are highly diverged at silent sites (44% and 38% for *LFY/FLO*-like and *AP3/DEF*-like, respectively), and phylogenetic analyses consistently place one copy from each species in separate gene clades (Fig. 2). This evidence is consistent with a second ancient duplication(s) in the ancestral lineage of these families, resulting in two paralogs that appear functional based on full-length cDNAs. Finally, two putative copies of *LFY/FLO*-like and *AP3/DEF*-like genes were found in *Salvia coccinia* (Lamiaceae) that cluster together with strong support within the FLOA and DEFB paralogy clades, respectively (Fig. 2). This may represent a third duplication event specific to the Lamiaceae, although there is no evidence of recent polyploidy as for *Chelone glabra* from chromosome numbers. However, we cannot rule out the possibility that duplicate copies in Lamiaceae could be of a hybrid origin (i.e., alleles of a single locus) because divergence at silent sites (19% and 17% for *LFY/FLO*-like and *AP3/DEF*-like, respectively) is near the range found for alleles segregating between interfertile *Mimulus* species (10% and 13%, respectively; J. E. Aagaard, unpublished data; University of Washington).

We examined several alternate topologies for gene trees (Fig. 2) that suggest different points for the ancient duplica-

tion(s) of *LFY/FLO*-like and *AP3/DEF*-like genes relative to divergence among families of Lamiales. First, we tested alternate topologies using a likelihood-based approach (Goldman et al., 2000), suggesting duplication(s) occurred at the base of the Lamiales (*AP3/DEF* $H_0$1, $H_0$2; Fig. 3). In all cases, alternate topologies supporting this hypothesis were rejected (Fig. 4). Second, because the *LFY/FLO* tree places Veronicaceae homologs before (*C. glabra*) and after (*A. majus*) the duplication, we tested alternate topologies constituting a null hypothesis, which implies duplication after the split leading to the Veronicaceae (*LFY/FLO* $H_0$1; *AP3/DEF* $H_0$3; Fig. 3). In this case, we clearly cannot reject the null hypothesis for *AP3/DEF,* though the test result is somewhat ambiguous for *LFY/FLO* (Fig. 4). However, because there is no evidence of duplicate copies from past work in the well-studied model species *A. majus* and *Antirrhinum* and *Chelone* represent a well-supported clade (Olmstead et al., 2001), this argues against rejecting the null hypothesis for *LFY/FLO* as well. In addition, Southern blots probed with the highly conserved C-terminus portion of *FLO* showed no evidence of more than a single copy in taxa from Veronicaceae (data not shown). In sum, our phylogenetic analyses and tests of alternate topologies support ancient duplication(s) of *LFY/FLO*-like and *AP3/DEF*-like genes after the divergence between Veronicaceae and the lineage leading to many of the other families of Lamiales including Verbenaceae, Lamiaceae, Phrymaceae, Paulowniaceae, and Orobanchaceae. We consider this scenario more likely because it is much more parsimonious from the standpoint of number of inferred gene duplication and loss events. Accordingly, we suggest a gene nomenclature for Lamiales *LFY/FLO* and *AP3/DEF* homologs that reflects the precedence of *FLO* (Coen et al., 1990) relative to *LFY* (Weigel et al., 1992), *DEF* (Sommer et al., 1990) relative to *AP3* (Jack et al., 1992), and co-orthology with *FLO* and *DEF.* Below we refer to Lamiales homologs as *FLO* or *DEF,* with co-orthologs further delineated as A or B (see Table 1).

Duplication of both genes at the same relative position in the Lamiales phylogeny suggests an ancient whole genome duplication event in Lamiales. *FLOA* and *DEFB* lie on separate linkage groups in *Mimulus*, as do putative paralogs of another floral regulatory gene (*CYC*; Fishman et al., 2001) that have a similar level of divergence at silent sites (45%) as *FLOA/B* and *DEFA/B* (J. E. Aagaard, unpublished data; University of Washington). Linkage relationships for *FLO* and *DEF* and additional unlinked paralogs are again consistent with whole genome duplication. However, taxa for which we found a single copy of *FLO* and *DEF* varied in their chromosome number (*S. vulgaris*, $n = 22$, *A. majus*, $n = 8$; Goldblatt and Krukoff, 1984) similar to taxa from which two copies of *FLO* and *DEF* were found (*V. officinalis*, $n = 7$; *S. coccinia*, $n = 11$; *M. lewisii*, $n = 8$; *M. guttatus*, $n = 14$; *P. tomentosa*, $n = 20$; *P. groenlandica*, $n = 8$; Goldblatt and Krukoff, 1984). Thus extant chromosome numbers appear to provide little evidence supporting whole genome duplication in the Lamiales phylogeny. As additional gene-based markers are mapped in Lamiales taxa sharing the hypothesized whole-genome duplication event (e.g., *Mimulus*; Bradshaw and Schemske, 1995; Fishman et al., 2001), syntenic groups can be compared with linkage relationships for *A. majus* (Schwarz-Sommer et al., 2003) in order to clarify the source of the ancient *FLO* and *DEF* paralogs we found. Regardless, the clear presence of paralogs for multiple genes among many families of Lamiales suggests future studies in this clade carefully consider orthology/paralogy relationships among homologs.

Patterns of duplication in the AP3/DEF and LFY/FLO gene lineages—The *DEF* duplicates within Lamiales we identified complement a pattern of repeated duplications throughout the *AP3/DEF* lineage reported previously. Duplications have occurred in seed plants (Theissen et al., 2000), within angiosperms at the base of the tricolpates (Kramer et al., 1998), and more recently, among lineages of basal angiosperms (Stellari et al., 2004) and other basal tricolpate families such as Ranunculaceae (Kramer et al., 2003). Interestingly, Irish (2003) suggests the duplication at the base of the tricolpates corresponds with duplication of several other MADS box genes, consistent with an ancient whole genome duplication at the base of the tricolpate lineage. This is similar to what we propose for Lamiales based on the concordant patterns for multiple floral regulatory genes including *FLO* and *DEF.* Duplication of MADS box genes at the base of the tricolpates including the ancestral *AP3/DEF* gene are thought to have contributed to major morphological innovations, including the development of distinct petals and the fixation of whorled phyllotaxy (Kramer and Irish, 1999; Irish, 2003). However, because of the time scale and phylogenetic position of most model species relative to this duplication, it is difficult to reconstruct ancestral gene function. Thus the contribution of earlier duplications to morphological evolution may prove difficult to study. In contrast, we suggest *DEF* duplicates in Lamiales may prove a good system in which to study the mechanisms responsible for duplicate gene preservation and diversification. Specifically, *DEFA* and *DEFB* are present within taxa from multiple Lamiales families separated for more than 30 million years (Wikstrom et al., 2001), providing several independently evolving lineages for comparison with the (single copy outgroup) model species *A. majus*. In addition, because a whole genome duplication is the likely cause of duplication in Lamiales, we believe paralogs of other floral regulatory genes will likely be found, such as those that are known to dimerize with *AP3/DEF,* including *PI/GLO* (Trobner et al., 1992; Goto and Meyerowitz, 1994), or whose expression is regulated by *LFY/FLO,* including *AG/PLENA* (Yanofsky et al., 1990; Bradley et al., 1993).

Significantly, ours is the first report of duplicate functional copies of *LFY/FLO* homologs outside of tetraploids (e.g., Kelly et al., 1995; Bomblies et al., 2003). Cronk (2001) suggested the conspicuous absence of *LFY/FLO* duplication across angiosperms reported previously (e.g., Frohlich and Parker, 2000) was in sharp contrast to patterns for other loci such as MADS box genes and could be the result of selection against duplicate *LFY/FLO* copies. However, our results suggest duplicate copies of *LFY/FLO* might have simply gone undetected in previous studies due to insufficient sampling of species and genomes. We used two independent primer sets nested within conserved domains of exons based on the *A. majus FLO* sequence (an ingroup) to amplify over introns with known length polymorphism, followed by exhaustive cloning and sequencing of all detectable PCR fragments regardless of signal intensity. Despite our experimental design, in at least one case (*Pedicularis groenlandica*), individual primer sets consistently amplified only one or the other *FLO* paralog. It would be interesting to apply a similar screening approach to assay *LFY/FLO* copy number in other angiosperm lineages, particularly where duplications of other genes regulated by *LFY/FLO* are

known (e.g., *AP3*/*DEF*-like genes in Ranunculaceae; Kramer et al., 2003). Such work will clarify issues regarding the frequency of duplication for *LFY*/*FLO* relative to other floral regulatory genes and may contribute to an understanding of the evolution of regulatory pathways following duplication (see below).

Copy number of UFO/FIM-like genes is inconclusive—Unlike *LFY*/*FLO* and *AP3*/*DEF,* our cloning experiments and phylogenetic analyses of *UFO, FIM,* and *UFO/FIM*-like homologs, including the eight *UFO*/*FIM*-like genes we identified among Lamiales taxa, provide little convincing evidence regarding copy number in Lamiales. Although our cloning experiments only identified a single *UFO*/*FIM*-like gene among all Lamiales taxa from which we cloned, our phylogenetic analyses are suggestive of a pattern of duplication for this gene. For example, it is surprising that support for nodes of the *UFO*/*FIM* phylogeny are weaker than the *LFY*/*FLO* tree since our *UFO*/*FIM* alignment is more than 40% longer than for *LFY*/*FLO* (1113 and 783 nucleotides, respectively), and the *UFO/FIM* alignment has fewer invariable sites than for *LFY/FLO* (20% and 31%, respectively). In addition, the relatively strong support (79%) for the clade containing *S. coccinia* and *P. groenlandica UFO*/*FIM*-like genes to the exclusion of other homologs is surprising, given we do not see this relationship among any other genes we have studied, and this clustering does not reflect the phylogenetic relationships among these taxa in Olmstead et al.'s (2001) study (Fig. 1). Possible explanations include (1) duplicate copies of *UFO*/*FIM* in the Lamiales, which we have incompletely sampled in our PCR-based screen; or (2) duplication of an ancestral *UFO*/*FIM* homolog, followed by repeated loss along the lineages from which we sampled (lineage sorting). These explanations are intriguing, because we can find no evidence of expression for the *P. groenlandica UFO*/*FIM*-like gene we cloned from genomic DNA in any of four cDNA pools constructed from separate individuals (collected from different populations), despite isolating cDNAs for both paralogs of *FLO* and *DEF* from the same RACE pools, and *FIM* should be expressed coincidently (Ingram et al., 1995). This suggests we might have missed a second copy expressed in the *P. groanlandica* floral bud cDNAs in our initial PCR screens from genomic DNA. Further studies, both within Lamiales and among appropriate outgroup taxa (e.g., Solanales), are needed to resolve the issue of copy number and possible duplication of *UFO*/*FIM* within the Lamiales.

Preservation of FLO and DEF paralogs in Lamiales—Paralogs of *FLO* and *DEF* in Lamiales appear to have avoided silencing, the typical fate of most gene duplicates. Lynch and Conery (2000) estimate the average half-life of fully redundant gene duplicates is around 4 million years, with a 95% probability of loss after approximately 17 million years. The *FLO* and *DEF* paralogs have been maintained within at least four independently evolving lineages of Lamiales (Verbenaceae, Phrymaceae, Paulowniaceae, and Orobanchaceae) more than twice as long (estimated divergence times for families range between 32 and 53 million years; Wikstrom et al., 2001). Thus selection must be acting on paralogs in order to maintain them within families over these time scales. Preservation of duplicate genes including floral MADS box genes (e.g., Mena et al., 1996; Lowman and Purugganan, 1999) and orthologs of *LFY/FLO* (Bomblies et al., 2003) have been reported previously and is typically associated with restricted expression of one or both paralogs relative to expression in single-copy lineages. Such parsing of ancestral gene function or subfunctionalization is thought to be the predominant mode of duplicate gene preservation (Force et al., 1999). Consistent with a subfunctionalization hypothesis for both, our studies of the molecular evolution of *FLO* and *DEF* paralogs found no evidence of adaptive divergence acting on duplicate copies but relaxed purifying selection in one (*FLOB*) or both (*DEFA* and *DEFB*) copies following the duplication event (J. E. Aagaard, unpublished data; University of Washington). Future rigorous tests of a subfunctionalization hypothesis for the preservation of *FLO* and *DEF* paralogs will require expression data during floral development as well as information on regulatory elements, which have only been well characterized for *AP3* (Hill et al., 1998).

Our gene trees show an interesting pattern suggesting that preservation of *FLO* and *DEF* paralogs in Lamiales may not have occurred independently. Specifically, all families of Lamiales for which we found two functional co-orthologs of *FLO* also retain two functional co-orthologs of *DEF* (Fig. 2). Such a pattern might be the result of rapid fixation of both *FLO* and *DEF* paralogs independently following duplication through a mechanism such as subfunctionalization (sensu Force et al., 1999), followed by phylogenetic divergence of the Lamiales. Alternately, joint preservation might have occurred non-independently as the result of the cognate relationships between genes. *LFY*/*FLO* regulates the expression of multiple MADS box genes including *AP3*/*DEF* (reviewed in Zhao et al., 2001), in part through binding upstream regulatory sequence (Lamb et al., 2002). Fixation of one set of duplicates (e.g., *DEFA* and *DEFB*) via a mechanism such as subfunctionalization might increase the probability of preserving an interacting gene (*FLOA* and *FLOB*) via a distinctly different mechanism. One possibility is dosage compensation, though this seems unlikely for transcription factors such as *LFY*/*FLO* and *AP3*/*DEF,* which can activate downstream targets when expressed at low levels (*LFY*/*FLO,* Coen et al., 1990; Weigel et al., 1992) or are positively self-regulating (*AP3*/*DEF,* Sommer et al., 1990; Jack et al., 1992). Alternately, a mechanism analogous to subfunctionalization acting at the molecular level could contribute to joint preservation of *FLOA/B* and *DEFA/B.* Joint preservation of duplicate pairs of interacting genes via divergent molecular interactions and/or expression domains is referred to as concerted divergence of gene expression and has been proposed as a mechanism responsible for the preservation of parallel networks of duplicated genes in *Arabidopsis* resulting from ancient polyploidy (Blanc and Wolfe, 2004). In our case, sequence divergence between regulatory elements of subfunctionalized *DEF* paralogs might require *FLO* paralogs with divergent DNA-binding domains. This hypothesis is testable because it predicts increased regulatory activity between pairs of duplicate genes. Interestingly, nonsynonymous substitutions are significantly elevated within the putative DNA binding domain of one of the Lamiales *FLO* paralogs (*FLOB*; J. E. Aagaard, unpublished data; University of Washington). Regardless of the mechanism, it will be interesting in future studies to examine the possibility of a similar pattern of correlated preservation of gene duplicates for *PLE* (Bradley et al., 1993), which is also known to be regulated by *FLO* (McSteen et al., 1998) as well as Lamiales homologs of other genes recently identified as regulated by *LFY* (William et al., 2004).

Conclusions—The *DEFA* and *DEFB* paralogs we cloned in Lamiales represent another example of repeated duplications within the *AP3/DEF* gene lineage (Kramer et al., 1998, 2003). While previous reports have suggested *LFY/FLO* may be constrained as a single copy gene (Cronk, 2001), this is clearly not the case in the Lamiales based on preservation of *FLOA* and *FLOB* paralogs within multiple families. Because our work suggests paralogs of *FLO* and *DEF* may be the result of an ancient whole-genome duplication event in the Lamiales, screens for homologs of other floral regulatory genes may reveal additional paralogous loci. Because duplications appear to have occurred after the split leading to the Veronicaceae, the model species *A. majus* presents a convenient outgroup for phylogenetic and developmental studies of paralogous floral regulatory genes. We believe Lamiales taxa will prove valuable for understanding the mechanisms of duplicate gene preservation and how key regulatory pathways may contribute to morphological diversity in plants.

LITERATURE CITED

AMBROSE, B., D. LERNER, P. CICERI, C. PADILLA, M. YANOFSKY, AND R. SCHMIDT. 2000. Molecular and genetic analyses of the *Silky1* gene reveals conservation in floral organ specification between eudicots and monocots. *Molecular Cell* 5: 569–579.

AOKI, S., K. UEHARA, M. IMAFUKU, M. HASEBE, AND M. ITO. 2004. Phylogeny and divergence of basal angiosperms inferred from APETALA3- and PISTILLATA-like MADS-box genes. *Journal of Plant Research* 117: 229–244.

BEARDSLEY, P., AND R. OLMSTEAD. 2002. Redefining Phrymaceae: the placement of *Mimulus*, tribe Mimuleae, and *Phryma*. *American Journal of Botany* 89: 1093–1102.

BLANC, G., AND K. WOLFE. 2004. Divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* 16: 1679–1691.

BLAZQUEZ, M., AND D. WEIGEL. 2000. Integration of floral inductive signals in *Arabidopsis*. *Nature* 404: 889–892.

BOMBLIES, K., R. WANG, B. AMBROSE, R. SCHMIDT, R. MEELEY, AND J. DOEBLEY. 2003. Duplicate *FLORICAULA/LEAFY* homologs *zfl1* and *zfl2* control inflorescence architecture and flower patterning in maize. *Development* 130: 2385–2395.

BRADLEY, D., R. CARPENTER, H. SOMMER, N. HARTLEY, AND E. COEN. 1993. Complementary floral homeotic phenotypes result from opposite orientations of a transposon at the plena locus of *Antirrhinum*. *Cell* 72: 85–95.

BRADSHAW, H., S. WILBERT, K. OTTO, AND D. SCHEMSKE. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). *Nature* 376: 762–765.

CARROLL, S., J. GRENIER, AND S. WEATHERBEE. 2001. From DNA to diversity. Blackwell Science, Malden, Massachusetts, USA.

COEN, E., J. ROMERO, S. DOYLE, R. ELLIOTT, G. MURPHY, AND R. CARPENTER. 1990. *floricaula*: a homeotic gene required for flower development in *Antirrhinum majus*. *Cell* 63: 1311–1322.

CRONK, Q. 2001. Plant evolution and development in a post-genomic context. *Nature Reviews Genetics* 2: 607–619.

FISHMAN, L., A. KELLY, E. MORGAN, AND J. WILLIS. 2001. A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. *Genetics* 159: 1701–1716.

FORCE, A., M. LYNCH, F. PICKETT, A. AMORES, Y. YAN, AND J. POSTLETHWAIT. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151: 1531–1545.

FROHLICH, M., AND D. PARKER. 2000. The mostly male theory of flower evolutionary origins: from genes to fossils. *Systematic Botany* 25: 155–170.

GAUT, B., AND J. DOEBLEY. 1997. DNA evidence for the segmental allotetraploid origin of maize. *Proceedings of the National Academy of Sciences, USA* 94: 6809–6814.

GOLDBLATT, P., AND B. KRUKOFF. 1984. Index to plant chromosome numbers 1979–1981. *Monographs in Systematic Botany* 4: 1–427.

GOLDMAN, N. 1993. Statistical tests of models of DNA substitution. *Journal of Molecular Evolution* 36: 182–198.

GOLDMAN, N., J. ANDERSON, AND A. RODRIGO. 2000. Likelihood-based tests of topologies in phylogenetics. *Systematic Biology* 49: 652–670.

GOTO, K., AND E. MEYEROWITZ. 1994. Function and regulation of the *Arabidopsis* floral homeotic gene Pistillata. *Genes & Development* 8: 1548–1560.

HILL, T., C. DAY, S. ZONDIO, A. THACKERAY, AND V. IRISH. 1998. Discrete spatial and temporal *cis*-acting elements regulate transcription of the *Arabidopsis* floral homeotic gene APETALA3. *Development* 125: 1711–1721.

HUGHES, A. 1999. Adaptive evolution of genes and genomes. Oxford University Press, Oxford, UK.

INGRAM, G., S. DOYLE, R. CARPENTER, E. SCHULTZ, R. SIMON, AND E. COEN. 1997. Dual role for *fimbriata* in regulating floral homeotic genes and cell division in *Antirrhinum*. *European Molecular Biology Organization Journal* 16: 6521–6534.

INGRAM, G., J. GOODRICH, M. WILKINSON, R. SIMON, G. HAUGHN, AND E. COEN. 1995. Parallels between UNUSUAL FLORAL ORGANS and FIMBRIATA, genes controlling flower development in *Arabidopsis* and *Antirrhinum*. *Plant Cell* 7: 1501–1510.

IRISH, V. 2003. The evolution of floral homeotic gene function. *Bioessays* 25: 637–646.

JACK, T., L. BROCKMAN, AND E. MEYEROWITZ. 1992. The homeotic gene APETALA3 of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* 68: 683–697.

KELLY, A., M. BONNLANDER, AND D. MEEKS-WAGNER. 1995. NFL, the tobacco homolog of FLORICAULA and LEAFY, is transcriptionally expressed in both vegetative and floral meristems. *Plant Cell* 7: 225–234.

KELLY, A., AND J. WILLIS. 1998. Polymorphic microsatellite loci in *Mimulus guttatus* and related species. *Molecular Ecology* 7: 769–774.

KRAMER, E., V. DI STILIO, AND P. SCHLUTER. 2003. Complex patterns of gene duplication in the *AP3* and PI lineages of the Ranunculaceae. *International Journal of Plant Sciences* 164: 1–11.

KRAMER, E., R. DORIT, AND V. IRISH. 1998. Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the APETALA3 and PISTILLATA MADS-box gene lineages. *Genetics* 149: 765–783.

KRAMER, E., AND V. IRISH. 1999. Evolution of genetic mechanisms controlling petal development. *Nature* 399: 144–148.

KURODA, H., N. TAKAHASHI, H. SHIMADA, M. SEKI, K. SHINOZAKI, AND M. MATSUI. 2002. Classification and expression analysis of *Arabidopsis* F-box-containing protein genes. *Plant Cell Physiology* 43: 1073–1085.

KUSH, A., A. BRUNELLE, D. SHEVELL, AND N. CHUA. 1993. The cDNA sequence of two MADS box proteins in petunia. *Plant Physiology* 102: 1051–1052.

LAMB, R., T. HILL, Q. TAN, AND V. IRISH. 2002. Regulation of *APETALA3* floral homeotic gene expression by meristem identity genes. *Development* 129: 2079–2086.

LOWMAN, A., AND M. PURUGGANAN. 1999. Duplication of the *Brassica oleracea* APETALA1 floral homeotic gene and the evolution of domesticated cauliflower. *Journal of Heredity* 90: 514–520.

LYNCH, M., AND J. CONERY. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290: 1151–1155.

MCSTEEN, P., C. VINCENT, S. DOYLE, R. CARPENTER, AND E. COEN. 1998. Control of floral homeotic gene expression and organ morphogenesis in *Antirrhinum*. *Development* 125: 2359–2369.

MENA, M., B. AMBROSE, R. MEELEY, S. BRIGGS, M. YANOFSKY, AND R. SCHMIDT. 1996. Diversification of C-function activity in maize flower development. *Science* 274: 1537–1540.

MOLINERO-ROSALES, N., M. JAMILENA, S. ZURITA, P. GOMEZ, J. CAPEL, AND R. LOZANO. 1999. FALSIFLORA, the tomato orthologue of FLORICAULA and LEAFY, controls flowering time and floral meristem identity. *Plant Journal* 20: 685–693.

OLMSTEAD, R., C. DEPAMPHILIS, A. WOLFE, N. YOUNG, W. ELISONS, AND P. REEVES. 2001. Disintegration of the Scrophulariaceae. *American Journal of Botany* 88: 348–361.

PARCY, F, O. NILSSON, M. BUSCH, I. LEE, AND D. WEIGEL. 1998. A genetic framework for floral patterning. *Nature* 395: 561–566.

PNUELI, L., M. ABU-ABEID, D. ZAMIR, W. NACKEN, Z. SCHWARZ-SOMMER, AND E. LIFSCHITZ. 1991. The MADS box gene family in tomato: temporal expression during floral development, conserved secondary structures and homology with homeotic genes from *Antirrhinum* and *Arabidopsis*. *Plant Journal* 1: 255–266.

POUTEAU, S., D. NICHOLLS, F. TOOKE, E. COEN, AND N. BATTEY. 1998.

Transcription pattern of a *FIM* homologue in *Impatiens* during floral development and reversion. *Plant Journal* 14: 235–246.

RAFF, R. 1996. The shape of life. The University of Chicago Press, Chicago, Illinois, USA.

RAMBAUT, A., AND N. GRASSLEY. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Computer Applications in the Biosciences* 13: 235–238.

SCHWARZ-SOMMER, Z., E. SILVA, R. BERNDTGEN, W. LONNIG, A. MULLER, I. NINDL, K. STUBER, J. WUNDER, H. SAEDLER, T. GUBITZ, A. BORKING, J. GOLZ, E. RITTER, AND A. HUDSON. 2003. A linkage map of an F_2 hybrid population of *Antirrhinum majus* and *A. molle*. *Genetics* 163: 699–710.

SIMON, R., R. CARPENTER, S. DOYLE, AND E. COEN. 1994. Fimbriata controls flower development by mediating between meristem and organ-identity genes. *Cell* 78: 99–107.

SOMMER, H., J. BELTRAN, P. HUIJSER, H. PAPE, W. LONNIG, H. SAEDLER, AND Z. SCHWARZ-SOMMER. 1990. Deficiens, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. *European Molecular Biology Organization Journal* 9: 605–613.

SOUER, E., A. VAN DER KROL, D. KLOOS, C. SPELT, M. BLIEK, J. MOL, AND R. KOES. 1998. Genetic control of branching pattern and floral identity during Petunia inflorescence development. *Development* 125: 733–742.

STELLARI, G., M. JARAMILLO, AND E. KRAMER. 2004. Evolution of the APETALA3 and PISTILLATA lineages of MADS-box-containing genes in the basal angiosperms. *Molecular Biology and Evolution* 21: 506–519.

SUNDSTROM, J., A. CARLSBECKER, M. SVENSON, M. E. SVENSSON, AND P. ENGSTROM. 1999. MADS-box genes active in developing pollen cones of Norway spruce are homologous to the B-class floral homeotic genes in angiosperms. *Developmental Genetics* 25: 253–266.

SWOFFORD, D. 2002. PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4.0 beta for Macintosh. Sinauer, Sunderland, Massachusetts, USA.

TAYLOR, S., J. HOFER, AND I. MURFET. 2001. *Stamina pistilloida*, the pea ortholog of *FIM* and *UFO,* is required for normal development of flowers, inflorescences, and leaves. *Plant Cell* 13: 31–46.

THEISSEN, G., A. BECKER, A. DI ROSA, A. KANNO, J. KIM, T. MUNSTER, K. WINTER, AND H. SAEDLER. 2000. A short history of MADS-box genes in plants. *Plant Molecular Biology* 42: 115–149.

TROBNER, W., L. RAMIREZ, P. MOTTE, I. HUE, P. HUIJSER, W. LONNIG, H. SAEDLER, H. SOMMER, AND Z. SCHWARZ-SOMMER. 1992. *GLOBOSA*: a homeotic gene which interacts with *DEFICIENS* in the control of *Antirrhinum* floral organogenesis. *European Molecular Biology Organization Journal* 11: 4693–4704.

WEIGEL, D., J. ALVAREZ, D. SMYTH, M. YANOFSKY, AND E. MEYEROWITZ. 1992. LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* 69: 843–859.

WIKSTROM, N., V. SAVOLAINEN, AND M. CHASE. 2001. Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society of London, B, Biological Sciences* 268: 2211–2220.

WILLIAM, D., Y. SU, M. SMITH, M. LU, D. BALDWIN, AND D. WAGNER. 2004. Genomic identification of direct target genes of LEAFY. *Proceedings of the National Academy of Sciences, USA* 101: 1775–1780.

YANOFSKY, M., H. MA, J. BOWMAN, G. DREWS, K. FELDMAN, AND E. MEYEROWITZ. 1990. The protein encoded by the *Arabidopsis* homeotic gene agamous resembles transcription factors. *Nature* 346: 35–39.

ZHANG, S., N. SANDA, P. POLOWICK, J. STILLER, J. STOUGAARD, AND P. ROBERT. 2003. *Proliferating floral organs* (*Pfo*), a *Lotus japonicus* gene required for specifying floral meristem determinacy and organ identity, encodes an F-box protein. *Plant Journal* 33: 607–619.

ZHAO, D., Q. YU, M. CHEN, AND H. MA. 2001. The ASK1 gene regulated B function gene expression in cooperation with *UFO* and LEAFY in *Arabidopsis*. *Development* 128: 2735–2746.