Duplications and Expression of *DIVARICATA*-Like Genes in Dipsacales

Dianella G. Howarth and Michael J. Donoghue†*

*Department of Biological Sciences, St. John's University, Queens, NY; and †Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT

The genetics underlying flower symmetry shifts between radial and bilateral symmetry has been intensively studied in the model *Antirrhinum majus*. Understanding the conservation or diversification of this genetic pathway in other plants is of special interest in understanding angiosperm evolution and ecology. Evidence from *Antirrhinum* indicates that TCP and MYB transcription factors, especially *CYCLOIDEA* (*CYC*), *DICHOTOMA* (*DICH*), *DIVARICATA* (*DIV*), and *RADIALIS* (*RAD*) play a role in specifying dorsal identity (*CYC*, *DICH*, and *RAD*) and ventral identity (*DIV*) in the corolla and androecium of monosymmetric (bilateral) flowers. Previous data indicate that the ECE clade of TCP genes (including *CYC* and *DICH*) underwent two duplication events around the diversification of the core eudicots. In this study, we examined the duplication events within Dipsacales, which contains both radially and bilaterally symmetrical flowered species. Additionally, we report here the phylogenetic relationships of the *DIV*-like genes across core eudicots. Like TCP genes, we found three core eudicot clades of *DIV*-like genes, with duplications occurring around the diversification of the core eudicots, which we name DIV1, DIV2, and DIV3. The *Antirrhinum* genes, *DIVARICATA* and its sister *DVL1*, fall into the DIV1 clade. We also found additional duplications within these clades in Dipsacales. Specifically, the Caprifoliaceae (bilaterally symmetrical clade) duplicated independently in each of the three core eudicot *DIV* clades. Using reverse transcription polymerase chain reaction (rtPCR), we showed that most of these copies are expressed across floral tissues in the Dipsacales species *Heptacodium miconioides*. One copy, *DipsDIV1A* (orthologous to *DIV* and *DVL1*), was expressed in a dorsal–ventral pattern. *DipsDIV1A* was expressed only in petal tissue, in both dorsal and ventral regions but was lacking from lateral petals. We argue that this suggests that *DipsDIV1A* may be expressed in a similar pattern to DIV in *Antirrhinum*, suggesting a broad conservation of this pathway. Finally, *DIV* contains a large intron near the beginning of the second MYB domain, which shows promise as a highly variable molecular marker for phylogenetic studies.

Introduction

Shifts in the symmetry of flowers between radial symmetry (polysymmetry, actinomorphy) and bilateral symmetry (monosymmetry, zygomorphy) have been common within angiosperms (Weberling 1989; Endress 1996, 1999). Much of the discussion of such shifts has focused on the sympetalous Asteridae (e.g., Donoghue et al. 1998; Ree and Donoghue 1999; Cubas 2002; Gillies et al. 2002; Knapp 2002; Reeves and Olmstead 2003), with approximately 65,000 species, about half of which have radially symmetrical corollas and half have bilaterally symmetrical corollas (Ree and Donoghue 1999). Donoghue et al. (1998) and Ree and Donoghue (1999) inferred that the ancestral asterid was probably radially symmetrical and that within this clade there were at least eight major gains of bilateral symmetry and at least nine reversals to radial symmetry. Such morphological shifts are of special interest in relation to shifts in pollination (e.g., see Neal 1998) and, ultimately, perhaps to shifts in rates of speciation (Sargent 2004). It has also been suggested that the various forms of bilateral symmetry found among asterids (Donoghue et al. 1998) might reflect underlying developmental constraints in this lineage, especially related to the basic orientation of the flower, which results in a medially positioned petal in the abaxial or ventral portion of the flower (Donoghue and Ree 2000).

Ideas on the evolution of floral symmetry have been greatly advanced in recent years by the discovery of three clades of potential candidate genes, *CYCLOIDEA* (Luo et al. 1995), *DIVERICATA* (Galego and Almeida 2002), and *RADIALIS* (Corley et al. 2005). All three of these gene groups were initially characterized in *Antirrhinum*, with function in dorsal (CYC, RAD) or ventral (DIV) portions of the flower. These genes fall into two major transcription factor families: TCP (*Teosinte Branched 1*, *CYCLOIDEA*, and *PCF*: Cubas, Lauter et al. 1999) and MYB (classified by the strong conservation of imperfect repeats: Martin and Paz-Ares 1997). A model for the genes in *Antirrhinum* hypothesizes that the TCP genes *CYC* and *DICH* (a duplicate in Antirrhineae: Hileman and Baum 2003) are expressed in dorsal portions of the corolla and androecium and control the activation of the MYB gene, *RAD*. RAD in turn inhibits another MYB gene, *DIV*, leaving DIV to function only in the ventral portion of the corolla (Corley et al. 2005). Thus, mutants of *cyc* + *dich* and *rad* result in radially symmetrical, ventralized flowers (Luo et al. 1995; Corley et al. 2005), whereas mutants of *div* (in a *cyc* + *dich* mutant background) result in radially symmetrical, dorsalized flowers in *Antirrhinum* (Almeida et al. 1997). Similar CYC function has been corroborated in the closely related *Linaria* (Cubas, Vincent, and Coen 1999), as well as in a legume, *Lotus japonica* (Feng et al. 2006). Additionally, CYC orthologs in Asteraceae play a role in specifying disk versus ray florets (Brohom et al. 2008; Kim et al. 2008). *CYC* has also been shown to be dorsally expressed in an array of other core eudicots: *Arabidopsis*, *Bournea*, *Iberis*, *Lonicera*, and *Lupinus* (Cubas et al. 2001; Citerne et al. 2006; Howarth and Donoghue 2006; Busch and Zachgo 2007; Zhou et al. 2008). An exception is the radially symmetrical *Cadia*, where one *CYC*-like copy (*LegCYC1A*) has a derived, expanded expression across the corolla (Citerne et al. 2006). All of these data suggest that *CYC* expression and function may be similar across rosids and asterids. Unlike *CYC*, however, little is known outside of *Antirrhinum* about the expression or function of the MYB genes (*DIV* and *RAD*). A recent study in *Bournea* (Gesneriaceae),

Key words: *DIVARICATA*, floral symmetry, Dipsacales, gene duplication, core eudicots, gene expression.

E-mail: howarthd@stjohns.edu.

Mol. Biol. Evol. 26(6):1245–1258. 2009
doi:10.1093/molbev/msp051
Advance Access publication March 16, 2009

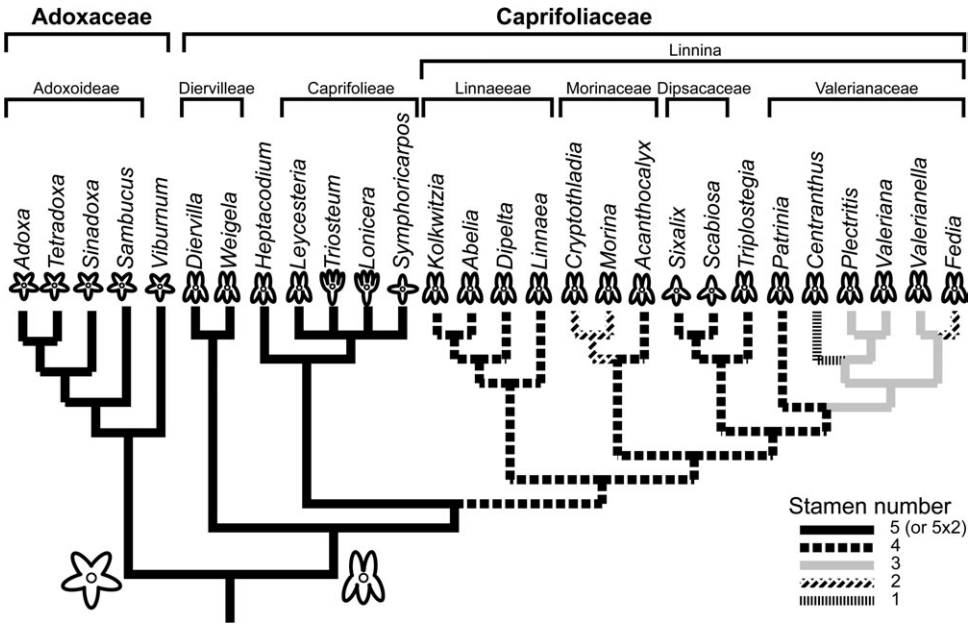

FIG. 1.—Summary phylogenetic tree for major lineages within Dipsacales, with the taxonomic names used here based on Donoghue et al. (2001). Phylogenetic tree represents well-supported relationships of the major groups of Dipsacales with the exception of the placement of *Heptacodium*, which can also be placed as sister to the Linnina clade (Winkworth, Bell, and Donoghue 2008). Floral drawings represent major forms of symmetry (polysymmetric or monosymmetric with 2:3 or 4:1 corolla orientation). Patterns on the branches represent different stamen numbers (doubled in *Adoxa* and its relatives). Modified from Howarth and Donoghue (2005) and Bell and Donoghue (2005).

however, indicates that the model of *CYC*-like, *DIV*-like, and *RAD*-like gene interaction may be conserved at least across Lamiales (Zhou et al. 2008).

Study of these genes in other groups requires the determination of orthologs and paralogs in other species. Analyses of the ECE lineage of *CYC*-like genes uncovered duplications that correlated with shifts in floral form (Howarth and Donoghue 2005). Additionally, a survey of eudicot *CYC* genes uncovered the presence of two duplications leading to three gene clades, also apparently correlated with a shift in floral form, from dimery to pentamery. One of these gene lineages included *CYC* from *Antirrhinum*; its sister clade included genes that may also be important in dorsal–ventral floral symmetry (Howarth and Donoghue 2006). In view of the broad distribution and possibly conserved function of *CYC* genes, we wanted to study another gene in the symmetry pathway outside of *Antirrhinum*. Specifically, we examined gene diversity and expression of *DIVARICATA* in the Dipsacales and across eudicots.

Because MYB genes make up the largest *Arabidopsis* gene family (Yanhui et al. 2006), determining orthology can be especially tricky. Yanhui et al. (2006) outlined the phylogenetic groupings of these genes in *Arabidopsis* and *Oryza*. *DIV* and *DIV-like1* (a sister gene found in *Antirrhinum* that is not expressed in corolla or androecium tissue) fall into the smaller clade of MYB genes, the R-R-type, which has two MYB domains (the first is similar to the I-box group and the second is similar to the CCA1 group). This clade includes nine *Arabidopsis* and seven *Oryza* genes as defined by Yanhui et al. (2006). Determining the broad gene phylogeny of R-R-type, *DIV*-like genes might uncover other gene regions that could be members of the floral symmetry pathway in *Antirrhinum* or other species. One way to tackle these questions is to generate gene phylogenies for a clade of core eudicots that has a known species phylogeny. We used the asterid clade Dipsacales to examine both the potential role of *DIV* in floral symmetry pathways outside of *Antirrhinum* and the broad phylogenetic context of *DIV*-like genes.

Our focus on Dipsacales was motivated by several factors. First, the Dipsacales phylogeny is quite well resolved based on a series of molecular and morphological phylogenetic analyses (e.g., Judd et al. 1994; Backlund and Donoghue 1996; Bell et al. 2001; Donoghue et al. 2001, 2003; Pyck 2001; Zhang et al. 2003). This species phylogeny, shown in figure 1, provides a solid basis for inferring the location of evolutionary changes in flower characters (Donoghue et al. 2003), as well as a secure framework within which to infer the evolution of *DIV*-like genes, including the location of gene duplications. The species tree reflects previous studies of larger Dipsacales data sets with the exception of *Heptacodium*, which can also be placed as sister to the Linnina clade (Winkworth, Bell, and Donoghue 2008).

Second, a variety of floral forms are found within Dipsacales, including radial symmetry, bilateral symmetry, and asymmetric flowers (Donoghue et al. 2003). Based on character analyses in a broader phylogenetic context (Donoghue et al. 1998; Ree and Donoghue 1999) and on the distribution of floral forms within Dipsacales (Donoghue et al. 2003), it is likely that bilaterally symmetrical flowers originated in this lineage independent of *Antirrhinum* and related plants. The primary split within Dipsacales separates the Adoxaceae from the Caprifoliaceae (fig. 1). Adoxaceae (including *Viburnum*, *Sambucus*, and *Adoxa* and its relatives) have radially symmetrical flowers and rotate corollas, whereas Caprifoliaceae (including Diervilleae, Caprifolieae, Linnaeeae, Morinaceae,

Dipsacaceae, and Valerianaceae) typically have bilaterally symmetrical flowers and tubular corollas (Fukuoka 1972; Donoghue et al. 2003). Dipsacales, in turn, appear to be related to the Paracryphiaceae, which have radially symmetrical flowers (Winkworth, Lundberg, and Donoghue 2008). Therefore, bilateral symmetry most likely arose along the line leading to the Caprifoliaceae (fig. 1).

Third, there are several forms of bilateral symmetry present within the Caprifoliaceae. Two-lipped flowers in which two dorsal petals are differentiated from the two lateral petals plus the medial ventral petals (the 2:3 form) are widespread within the clade and appear to be ancestral. There appears to be some lability within the Caprifolieae. Two-lipped flowers in which the two dorsal petals plus the two lateral are differentiated from the medial ventral petal (the 4:1 form) are found in *Lonicera* and *Triosteum*, and polysymmetric flowers appear to have re-evolved within *Symphoricarpos*. There have also been several shifts within the Caprifoliaceae in other floral traits that could be affected by the floral symmetry pathway, including the number of corolla lobes, stamen abortion, the formation of an epicalyx, and a bilaterally symmetrical calyx (Donoghue et al. 2003).

In this paper, we address the evolution of *DIVARICATA* genes within the Dipsacales and relate this to general patterns for gene family evolution in eudicots. We highlight the major duplications in *DIV* both in the Dipsacales and in core eudicots. Additionally, we present reverse transcription polymerase chain reaction (rtPCR) expression data from one Dipsacales species, *Heptacodium miconioides*, for each of these copies, highlighting dorsal–ventral patterning in an *Antirrhinum DIV* ortholog. Finally, we discuss the utility of an intron in *DIV* as a low-level phylogenetic marker, using *Sixalix atropurpurea* as an example.

Materials and Methods

Dipsacales

Twenty-nine individuals were sequenced from 18 genera, representing all major lineages within Dipsacales. We used total genomic DNAs, obtained in many cases from extractions used in previous phylogenetic studies. We also used floral cDNA from some taxa to confirm intron boundaries and to explore patterns of gene expression. Table 1 provides a list of included taxa, voucher specimens, and GenBank accession numbers.

Table 2 lists the primers used and their success rates. All primers were designed using the two MYB domains, and were developed through the course of the study as more sequences were obtained. *Aquilegia alpina* (Ranunculales) was also mined for any DIV-like copies to be used as outgroups. Multiple primer pairs were used for each taxon.

Amplification utilized the following cycling program: 95 °C for 45 s, 50–56 °C for 1 min, and 72 °C for 1 min 30 s, repeated for 39 cycles. Reactions were performed using *Taq* DNA polymerase (Qiagen, Valencia, CA) in 25 μl, with final concentrations of 2.5 mM $MgCl_2$, 0.5 μM of each primer, 0.8 mM dNTPs, and 0.5× Q Solution (Qiagen). Amplified products were cloned using the Invitrogen TOPO TA Cloning Kit for Sequencing (Carlsbad, CA).

Between 10 and 40 colonies were screened (polymerase chain reaction, PCR, reactions with the same concentrations as above in a 20 μl volume) for all potentially different copies or alleles of the *CYC*-like genes obtained. Selected colonies were then picked into a PCR cocktail. After a 10 min start at 95 °C, amplifications utilized the following cycling program: 95 °C for 30 s, 55 °C for 45 s, and 72 °C for 60 s, repeated for 24 cycles. Amplification products of the appropriate size were cleaned using a PEG/NaCl protocol and directly sequenced.

Genome Mining

In order to determine orthologous genes, *DIVARICATA* and *DIV-like1* genes from *Antirrhinum* were included in the matrix. We also included all known *Arabidopsis* and *Oryza* genes from the clade of R-R-type MYB genes. We searched the genome of *Populus* (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html) for similar genes and included them, and a single gene identified from *Medicago* was included (http://www.tigr.org/tdb/e2k1/mta1/).

Alignment and Phylogenetic Analyses

All clones from each DNA extraction (obtained using multiple primer pairs) were compiled into Sequencher 4.2 (Gene Codes Corporation, Ann Arbor, MI). All sequenced genes with MYB domains were included in the matrix. Positive clones were separated into different "types" based on shared differences among the clones. These types were essentially identical to each other. They only varied by obvious polymerase error (single base differences in one or two clones out of dozens, with different clones being mutated at different sites), with the exception of two copies of *Sixalix* with allelic variation (addressed below). Recombinant PCR sequences (PCR hybrids) were occasionally detected by comparing clones from a single individual (Paabo et al. 1990), and were removed. A consensus sequence was generated for each type and exported for phylogenetic analysis. Protein sequences for each potential copy/allele obtained were aligned by eye in MacClade 4 (Maddison DR and Maddison WP 2003), and the aligned nucleotide matrix was used in subsequent phylogenetic analyses.

Parameters for the Bayesian and maximum likelihood (ML) analyses were estimated using Modeltest 3.06 (Posada and Crandall 1998). The Akaike Information Criterion (Akaike 1973) recommended a General Time Reversible (GTR) model with added parameters for invariable sites and a gamma distribution (GTR + i + g) for Bayesian analyses. The matrix included only the alignable region from both MYB domains. We used the Metropolis-coupled Markov Chain Monte Carlo method as implemented in MrBayes 3.12 (Huelsenbeck and Ronquist 2001) to run four chains (three heated). We ran 5 million generations, sampling every 1,000 generations, with a burn-in of 1,000 trees (1 million generations). A consensus tree was generated in MrBayes, which included branch lengths and posterior probabilities. ML analyses were performed in PAUP* 4.0b10 (Swofford 2001) on the three major clades identified in the Bayesian analyses. These matrices included only coding region (the intron was not included).

Table 1
Taxon List Voucher Information and Genbank Numbers.

Taxon	Voucher	DIV1 A	DIV1 B	DIV2 A	DIV2 B	DIV3 A	DIV3 B
Adoxaceae							
Sambucus caerulea	Donoghue, voucher lacking	FJ805277				FJ805278	
Sambucus canadensis	Donoghue, voucher lacking	FJ805279, FJ805280, FJ805281					
Sambucus nigra	Donoghue, voucher lacking					FJ805282	
Viburnum davidii	Winkworth and Donoghue 42 (A and YU)	FJ805283				FJ805284	
Viburnum ellipticum	Winkworth and Donoghue 31 (A and YU)	FJ805285, FJ805286				FJ805287	
Viburnum jucundum	Winkworth and Donoghue 38 (A and YU)	FJ805288, FJ805289, FJ805290				FJ805291	
Viburnum prunifolium	Winkworth and Donoghue, 6 (A and YU)	FJ805292, FJ805293				FJ805294	
Caprifoliaceae							
Centranthus macrosiphon	Bell, cDNA		FJ805295			FJ805296	
Centranthus ruber	Bell 203 (YU)		FJ805297			FJ805298	FJ805299
Diervilla sessilifolia	Elisk and Zinman 3286 (A)	FJ805300	FJ805301		FJ805302	FJ805303	FJ805304
Dipelta floribunda	Buckland and Kelly 32 (A)		FJ805305	FJ805306	FJ805307		FJ805308
Fedia cornucopiae	Bell, cDNA		FJ805309, FJ805310				
Heptacodium miconioides	Howarth, cDNA	FJ805311	FJ805312	FJ805313	FJ805314		FJ805315
Kolkwitzia amabilis	Elsik, Michener, and Bailey 844 (A)	FJ805316	FJ805317				FJ805318
Leycesteria sp.	Boufford et al. 44597 (A)	FJ805319	FJ805320				FJ805321
Linnaea borealis	Donoghue 113, voucher lacking	FJ805322	FJ805323		FJ805324		FJ805325
Lonicera maackii	Smith 20 (YU)	FJ805326					FJ805327
Lonicera morrowii	Smith 49 (YU)		FJ805328				FJ805329
Lonicera prolifera	Donoghue 107, voucher lacking	FJ805330					FJ805331
Morina longifolia	Eriksson s.n. 2 Nov. 1999 (SBT)	FJ805332, FJ805333	FJ805334, FJ805335, FJ805336			FJ805337	FJ805338, FJ805339
Sixalix atropurpurea	Carlson 144 (YU)	FJ805340, FJ805341, FJ805342, FJ805343					
Sixalix atropurpurea	Carlson 151 (YU)	FJ805344, FJ805345, FJ805346	FJ805347	FJ805348	FJ805349		FJ805350
Symphoricarpos occidentalis	Donoghue, voucher lacking				FJ805351		FJ805352
Symphoricarpos orbiculatus	Donoghue, voucher lacking	FJ805353	FJ805354	FJ805355	FJ805356		FJ805357
Triplostegia glandulifera	Boufford et al. 27738 (A)	FJ805358	FJ805359		FJ805360		FJ805361
Valerianella eriocarpa	Bell, cDNA		FJ805362	FJ805363			
Valerianella locusta	Bell 2006-64 (YU)	FJ805364	FJ805365	FJ805366		FJ805367	FJ805368
Valeriana officinalis	Bell 2006-53 (YU)			FJ805369		FJ805370	FJ805371
Weigela hortensis	Kelly and Buckland 28 (A)	FJ805372	FJ805373		FJ805374		FJ805375
Other							
Antirrhinum majus	Galego and Almeida	AY077453, AY077454					
Aquilegia alpina	Kramer, voucher lacking	FJ805376					
Arabidopsis thaliana	Genome database	AT5G58900, AT2G38090, AT501200		AT5G05790, AT3G11280			
Medicago truncatula	Genome database	TC104048 TC26569 TC33683 TC4550 TC6125 TC90249					
Populus trichocarpa	Genome database	gw1.I.1006.1, estExtfgenesh4pg.CLGIX1149		grail3.0049028801, grail3.0022016501		eugene3.00060843, estExtGenewise1v1CLGXVI3514	

Placement of NonCaprifoliaceae Numbers Does Not Necessarily Reflect Orthology to that Clade.

Table 2
Primers Used in This Study

Initial Primers	
DIVF-1	CCNGGNAARACNGTNKGNGAYGTNAT
DIVR-1	RAARTAYTTYTGNGCRTGNSWNGCNACYTG
DIVF-2	GTGGGGGAYGTGATCAAACAGTAYAG
DIVR-2	CCATACTTRTTWAGCCCSAGCAAAAATTGCCTG
Copy specific	
DIV1F-1	TGGACHARAGARGAGAACAAG
DIV1R-1	GTKSTRATGTCRTGRATRCTWGG
DIV1F-2	GATATWGAARMHGGDNDDRTYCCVVTTCCNGG
DIV1R-2	GAKATRTTYCTCCARTMYCC
DIV2F	GATGTSATYAAACARTAYMRVGMAYTRGA
DIV2R	TGCCTRTGYTCYTCYTCNKTCC
DIV3F-1	TGGACDVYDGHWGAKAACAA
DIV3R-1	TACTTYTGNGCRTGRCTNGCNACYTG
DIV3F-2	GAYGTBAGYDRYATHGARGCHGG
DIV3R-2	GADATRTTYCTCCARTCBCCYTTBCC
Heptacodium rtPCR	
HeptDIV1AF	AGTATTGAAGCTGGGCTAGTTCCAATTCCTGGTTATAATACTAATACTTCTCC
HeptDIV1AR	CCGGGAGATATTTCTCCAGTCCCCCTTTCCGTACTTTTTAAGCCCCAAG
HeptDIV1BF	GTAGTATTGAAGCTGGACTCATTCCAATTCCGGGGTATAGTACCTCTCC
HeotDIV1BR	GTCGGCGTTCTAGTGATTACGAAATTGCGAGAAATATTTCTCCAGTCG
HeptDIV2AF	AAATGGGCTAGTTCCGATTCCCGGGTACATAACCAAATCCTC
HeptDIV2AR	CTGAGTTGGTGTCTTAGTGATCACAAAGTTCCGCGAGATGTTCCTCCAA
HeptDIV2BF	GTTACTGATATAGAAGCTGGGTTGGTTCCAATTCCAGGGTATCTCACCTC
HeptDIV2BR	CCG GGA TAT GTT CCT CCA GTC CCC TCG ACC GTG CTT CTC
HeptDIV3BF	GTGATATTGAGGCCGGTCTAATACCAATACCCGGGTACACTACTAGTGA
HeptDIV3BR	GCGAGATATGTTTCTCCAGTCCCCTTTTCCATACTTTTTAAGCCCTAGC

Heuristic searches were performed with 20 random addition replicates and TBR branch swapping, and 100 ML bootstrap replicates were examined.

A clade of seven potential alleles from two individuals of *S. atropurpurea* was analyzed using parsimony and ML in PAUP* (Swofford 2001). These alleles all fell into two potential sister copies in the *DipsDIV1A* clade. The entire sequenced region was used, including the intron. Bootstrap values were calculated with 1,000 replicates.

rtPCR

Plant tissue was collected from *Heptacodium* growing in the Marsh Botanical Garden of Yale University, dissected, and flash frozen at −80 °C overnight. Corolla lobes were separated from the corolla tubes of multiple flower buds from the first stage that this was possible to nearly mature flower buds. These lobes were separated into three groups: two dorsal petals, two lateral petals, and the single ventral petal. Whole flower buds were also used from 10.5-mm flower buds. Whole calyx, corolla (including the tube), androecium gynoecium (inferior ovary, presumably containing calyx and corolla tissue), and leaf tissues were also separated. Frozen tissues were pulverized with the Bio101 FastPrep system and were subsequently extracted using an RNeasy kit with the optional DNase step (Qiagen). cDNAs was subsequently generated from RNA using Superscript III (Invitrogen).

Full-length *DIV*-like *Heptacodium* sequences were obtained from all of the *Heptacodium* copies using rapid amplification of cDNA ends (RACE) (Smart RACE cDNA Amplification Kit, Clontech, Mountain View, CA). Specific primers were then designed for each individual copy (table 2). The primers spanned the intron between both MYB domains to rule out DNA contamination. PCR was then performed on the cDNA using the same conditions as outlined above. A G3PDH control was also included, using primers from Strand et al. (1997). These results were run on a 1% agarose gel and copies were verified by sequencing.

Results

R-R-Type Genes

Multiple copies of *DIV*-like genes, up to a maximum of eight, were obtained from every Dipsacales species examined. Combining these hypothesized copies resulted in a matrix of 109 taxa, including 94 Dipsacales sequences and 15 sequences from other eudicot species. Additional sequences hypothesized to be alleles were not included in this matrix (discussed below). The final matrix included 192 bases from the first MYB domain and 222 bases from the second MYB domain for a total length of 414 bases. Alignments were unambiguous for this region. The coding region between the two MYBs and an intron were removed from the matrix as they were unalignable. These data can be obtained from TreeBASE (http://www.treebase.org).

Using Bayesian inference and rooting with a single copy obtained from *Aquilegia*, there are three well-supported *DIV*-like genes in Dipsacales (posterior probabilities equal 0.85, 1.0, and 0.88). Additionally, each of the three Dipsacales clades is sister to rosid sequences from *Arabidopsis*, *Populus*, and/or *Medicago*, indicating that these copies characterize a much larger core eudicot clade (posterior probabilities equal 0.85, 1.0, and 0.97). Here we refer to these clades as DIV1 (this includes *DIVARICATA* and *DVL1* from *Antirrhinum*), DIV2, and DIV3 (fig. 2).

We refer to the entire clade, including *Aquilegia* and DIV1, DIV2, and DIV3, as the core RR DIV clade.

Additional R-R-type sequences from *Arabidopsis*, *Populus*, and *Oryza* fall outside of the core RR DIV clade. Due to long branches, analyses of different taxa and alignments provided different placements of the root within the core RR DIV clade. All analyses, however, displayed the three major lineages of core eudicot genes: DIV1, DIV2, and DIV3, with the exception that some outgroups placed the root within the DIV1 clade, causing the rosid sequences (*Arabidopsis* and *Populus*) from DIV1 to fall outside of the core RR DIV clade.

Dipsacales Clades

DIV1

DIV1 includes three *Arabidopsis* sequences, two *Populus* sequences, and *DIVARICATA* and *DVL1* from *Antirrhinum*. Within Dipsacales, this clade appears to include two copies in Adoxaceae and two copies in Caprifoliaceae. The relationships of these copies vary in different analyses. Bayesian analyses indicate that there is a single Adoxaceae clade including both copies, which is sister to the *DipsDIV1A* clade of Caprifoliaceae (fig. 2). ML analyses indicate that the Adoxaceae copies are paraphyletic with respect to the *DipsDIV1A* clade of Caprifoliaceae. Other analyses (not shown) collapsed one Adoxaceae clade into a polytomy, suggesting uncertainty in this relationship. Within Caprifoliaceae, the gene tree of *DipsDIV1* genes generally matches that of the species tree, with a few discrepancies. Within Caprifoliaceae *DipsDIV1A*, *Kolkwitzia* is sister to *Morina* rather than *Linnaea*, and *Valerianella* falls within *Sixalix* (ML) or as sister to the rest of the Linnina clade (Bayesian). *Heptacodium*, which has been especially hard to place in previous studies (see Winkworth, Bell, and Donoghue 2008), is sister to Caprifolieae in *DipsDIV1A* and sister to the Linnina clade in *DipsDIV1B*. DIV1 contains more duplications and potential allelic variation than the other copies. Duplications appear to have occurred in *Morina* and *Sixalix* (or perhaps in the entire Dipsacaceae + Valerianaceae clade) in *DipsDIV1A*, and in *Morina* twice and in the *Valerianella* + *Fedia* clade in *DipsDIV1B*. Additionally, allelic variation (determined by looking at differences in length and base sequence in introns and areas outside of the MYB domains) was found in *Sambucus*, *Viburnum*, *Diervilla*, and *Sixalix* within the *DipsDIV1A* clade.

DIV2

DIV2 includes the most rosid sequences, with two *Arabidopsis* copies, two *Populus* copies, and the single copy found so far from *Medicago*. Bayesian analyses indicate that together these rosid species are sister to a well-supported Dipsacales clade (fig. 2, ML bootstrap of 83). Within Dipsacales, no DIV2 copies were recovered from the radially symmetrical Adoxaceae. It is possible that they were not successfully identified in our screen; however, at least seven different taxa were sampled with multiple primer pairs with no success, suggesting that this copy may not occur in Adoxaceae. As in DIV1, however, there is a duplication resulting in two separate well-supported Caprifoliaceae clades (ML bootstrap 77 and 86). *DipsDIV2A* (fig. 2) generally supports the species phylogeny with the exception of the reversed positions of *Dipelta* and *Sixalix*. *DipsDIV2B* (fig. 2) also supports the species phylogeny. *Heptacodium* appears as sister to the Linnina clade in *DipsDIV2A*, and in *DipsDIV2B* it is sister to Caprifolieae plus Linnina. There are no other duplications of DIV2 within the Caprifoliaceae.

DIV3

DIV3 includes two rosid sequences, both from *Populus*. Within the known genome of *Arabidopsis*, there appear to be no sequences in this clade. Within Dipsacales, there is a single clade of *Viburnum* and a single clade of *Sambucus* (both Adoxaceae). The relationships of these two clades are unresolved in Bayesian analyses. Within Caprifoliaceae, there was a duplication leading to two well-supported clades (ML bootstrap of 99 and 90). *DipsDIV3A* contains sequences from only five taxa; however, they span the Caprifoliaceae phylogeny and are concordant with the species phylogeny. *DipsDIV3A* is also subtended by a much longer branch than any of the other clades, containing many protein sequence differences within the region between the two MYB domains. *DipsDIV3B* is more thoroughly sampled across Caprifoliaceae and also generally supports the species phylogeny, with the exception of the unresolved placement of *Morina* and the *Sixalix* + *Triplostegia* clade. *Morina* appears to contain a duplication in *DipsDIV3B*.

Floral cDNA Expression

Full-length *Heptacodium* sequences were obtained using RACE for the five *DIV*-like copies found (*DipsDIV1A* and *DipsDIV1B*, *DipsDIV2A* and *DipsDIV2B*, and *DipsDIV3B*). rtPCR results using specific primers for each copy are shown in figure 3. All of the extracted tissues contained similar concentrations of RNA with the exception of calyx tissue, which unfortunately yielded very little RNA, and subsequently little *G3PDH* expression. Most of the copies were expressed in nearly all sampled tissues (*DipsDIV1B*, *DipsDIV2A* and *DipsDIV2B*, and *DipsDIV3A*). *DipsDIV1A*, on the other hand, is found only in extractions that included corolla tissue (although the calyx could not be ruled out due to low yield). Additionally, *DipsDIV1A*, unlike any published pattern for floral symmetry genes, appears to be expressed in dorsal and ventral petals, but not in lateral petals. Its sister copy, *DipsDIV1B*, however, was heavily expressed in all tissues, even in the calyx, which contained less mRNA than the other tissues. *DipsDIV2A*, *DipsDIV2B*, and *DipsDIV3A* were expressed across all tissues (excluding the calyx).

Although only five DIV-like copies were uncovered in *Heptacodium*, there are six copies in the Dipsacales. Genes from all six of these *DIV*-like Dipsacales clades are expressed in flowers as evidenced by the cDNA sequences shown in figure 2. Floral cDNA was used in place of genomic DNA for several species in order to uncover all

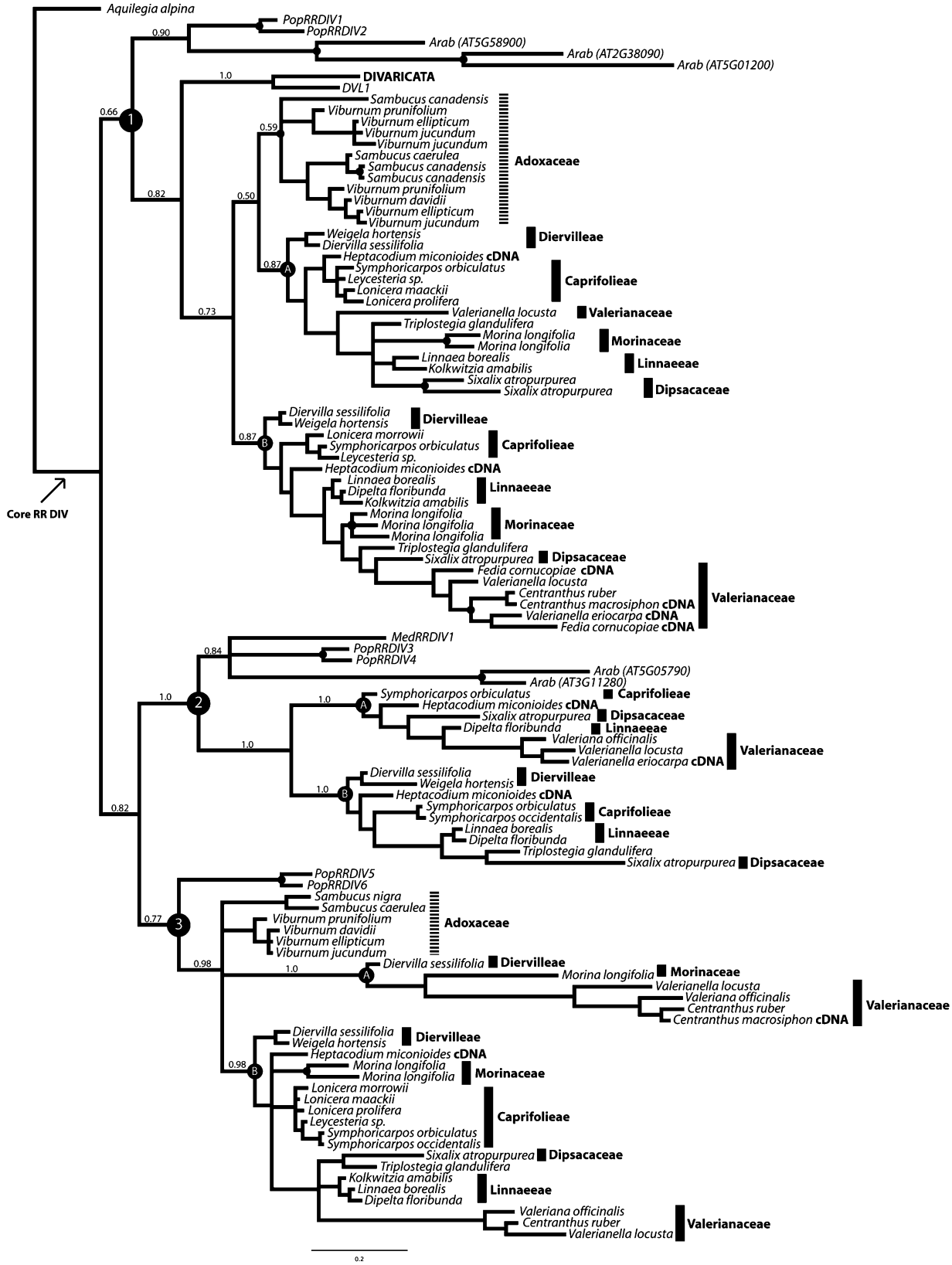

FIG. 2.—Phylogeny of core RR *DIV* genes using Bayesian analysis. Clades with 0.5 posterior probability are displayed with numbers above the lines of the major clades. Phylogeny with all taxa sampled. The placement of *DIVARICATA* and *DVL1* from *Antirrhinum* are bolded. The three major clades are labeled: 1, 2, or 3 and each duplication in Caprifoliaceae is labeled A or B. The Dipsacales lineages are shown on the right side. Arab = *Arabidopsis*, Pop = *Populus*, and Med = *Medicago*.

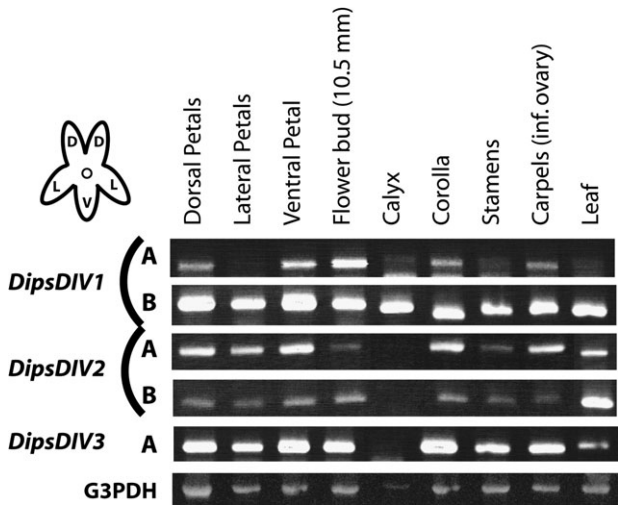

FIG. 3.—Image of agarose gel electrophoresis of cDNA from *Heptacodium miconioides*. Includes expression of five copies of *DIV* found in *H. miconioides*. A G3PDH control is included. Bird's eye view of flower is included showing the differentiation between the dorsal (D), lateral (L), and ventral (V) petal(s).

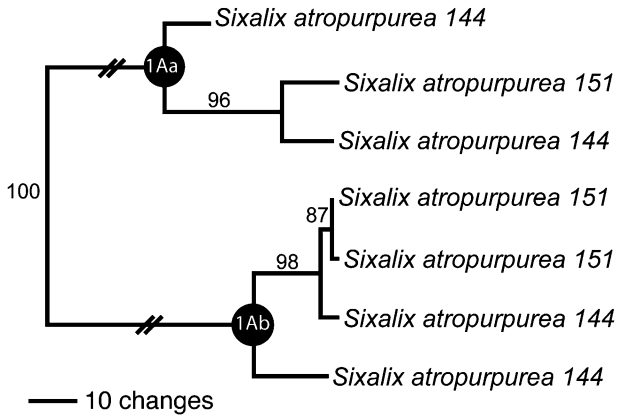

FIG. 4.—Parsimony generated tree of alleles from *DipsDIV1A* from two individuals of *Sixalix atropurpurea*, Carlson 144 and Carlson 151. Two separate putative copies are labeled: 1Aa and 1Ab. Bootstrap support is given above the line. Hatch marks indicate that the branches are much longer than those shown.

potential copies and define the intron/exon boundaries. For instance, copies with long introns may have been too long to amplify and therefore missed in our screens. These data support the expression of each of these copies from multiple Dipsacales species. Most of the copies found from cDNA fall into the *DipsDIV1B* clade, which supports the clear expression of this clade in all tissues in *Heptacodium* (fig. 3).

Intron Utility in Low-Level Phylogenies

DIVARICATA may be a useful marker for phylogeography and species-level phylogenies. The two conserved MYB domains flank a variable intron. In Dipsacales, this intron varies from 76 to 1,280 bases and averages 435 bases. There was no clear distinction in size among DIV1, DIV2, and DIV3 sequences. We show an example of a maximum parsimony tree generated from the *DipsDIV1A* copies of *S. atropurpurea* (fig. 4). The sequences, obtained from two different individuals of *S. atropurpurea*, cluster into two types, indicating an additional duplication event (labeled 1Aa and 1Ab). The intron sequences between these two copies were not alignable. Within each copy, the raw sequence variation varied between 0.026 and 0.039 (copy 1Aa) and 0.01 and 0.043 (copy 1Ab). Both MP and ML analyses produced a single well-supported tree.

Discussion

The Core RR DIV Clade

The entire clade of R-R-type MYB genes includes nine *Arabidopsis* sequences and seven *Oryza* sequences (Yanhui et al. 2006). *DIVARICATA* from *Antirrhinum* and our data from Dipsacales all fall into a crown clade within this group, which also contains five *Arabidopsis* sequences but no *Oryza* sequences. There are additional rosid sequences within this clade as well, obtained from genome analyses of *Populus* and *Medicago*. Therefore, this gene clade appears to have duplicated after the divergence of monocots and eudicots. Using the one *Aquilegia* sequence to root this clade, it would appear that there were two separate duplications that occurred after the split between the ranunculids from the rest of the eudicots (fig. 2). These duplications resulted in three major clades of *DIV*-like sequences in core eudicots (fig. 2). Based on the *Aquilegia* rooting, the first duplication separated DIV1 from the ancestor of DIV2 and DIV3. The DIV1 clade includes *DIV* and *DVL1* from *Antirrhinum*, which form a clade that is sister to our Dipsacales DIV1 sequences.

Unfortunately, the placement of the root of the core RR DIV clade varies depending on outgroup taxon sampling and data partitions. For instance, in analyses using different subsets of data or outgroup sampling, *Aquilegia* is sometimes sister to the entire core RR DIV clade (as shown in fig. 2), sister to the DIV1 clade alone, sister to the DIV2 and DIV3 clades together, or forms a polytomy with all three DIV clades. Aside from this rooting issue, however, there are consistently three broad DIV clades within eudicots, with the single copy from *Aquilegia* falling as sister to one, two, or all three of these clades.

Another potential conflict concerns the rosid sequences in the DIV1 clade. A clade of five rosid sequences (three from *Arabidopsis* and two from *Populus*) is either sister to the other members of DIV1, or forms a polytomy with all three DIV clades, or falls outside of the entire clade including *Aquilegia* and the core RR DIV clade. However, when the three *Arabidopsis* sequences are removed from the analysis, the remaining *Populus* sequences are more likely to form a sister group to the DIV1 clade. This uncertainty is most likely due to minimal taxon sampling across the eudicots. Having only a few sequences from the rosids, especially given the long branches of the *Arabidopsis* copies, is complicating the analysis. Confident rooting of this clade clearly will require data from additional outgroup angiosperms. In the meantime, the tree displayed in figure 2 is our best hypothesis of the relationships of these sequences.

Our findings on *DIV* evolution provide a striking complement to recent studies of TCP and MADS-box genes. Proteins within the TCP and MYB families pattern the dorsal–ventral identity of *Antirrhinum* flowers (Luo et al. 1995; Almeida et al. 1997; Corley et al. 2005; Costa et al. 2005), whereas MADS-box proteins pattern floral organ identities (Coen and Meyerowitz 1991; Meyerowitz et al. 1991). The ECE clade of *CYC*-like TCP genes apparently duplicated twice before the diversification of the core eudicots (Howarth and Donoghue 2006). Likewise, it appears that members of each of the major functional categories of floral MADS-box genes—*APETALA1* (*AP1*, A class), *APETALA3* (*AP3*, B class), *AGAMOUS* (*AG*, C class), and *SEPALLATA*—underwent a duplication in a similar location near the base of the core eudicots (Litt and Irish 2003; Kim et al. 2004; Kramer et al. 2004; Kramer and Hall 2005; Zahn et al. 2005). Our analyses of *DIV*-like genes suggest that the *DIV*-like MYB genes also duplicated at around the time of the diversification of core eudicots.

Bowers et al. (2003) and De Bodt et al. (2005) compared the genomes of *Arabidopsis*, *Populus*, and *Oryza* and hypothesized a whole genome duplication before the diversification of the clade including asterids and rosids but after the split from the monocots. With the completion of the *Vitis* genome, however, Jaillon et al. (2007) were able to more precisely suggest that there were instead two duplication events (or a single hexaploid ancestor) that predate the diversification of the rosids. Our data from the ECE clade of CYC (Howarth and Donoghue 2006) and now the core RR DIV clade, along with data from MADS-box genes, support the hypothesis that there were three independent copies of the gene when the core eudicots diversified. First, there are two or three copies (we assume that in the case of two copies, one was lost) of members of each of these different gene families (MADs, TCP, and MYB). Second, all of these duplications appear to have taken place after the divergence between the ranunculids (represented by *Aquilegia*) and the rest of the eudicots, which places these duplications more precisely than the whole genome comparisons.

It is possible that these major duplication events played an important role in the major changes in flower form that mark the core eudicot clade. Most importantly, perhaps, flowers of the core eudicots (excluding Gunnerales) are based on a pentamerous ground plan, as opposed to the dimerous ground plan that characterizes the so-called basal eudicots (Magallon et al. 1999; Soltis et al. 2003). The switch to pentamery entails the establishment of the differentiation of dorsal and ventral portions of the flower in relation to the axis on which they are borne. It is possible that the shift in floral morphology that marks the core eudicot clade is at least in part the result of the duplication and maintenance of genes from three separate transcription factor gene families that are of known importance in flower development.

Duplications

Gene duplication has been important in the genesis of complicated networks of interacting genes. If a gene duplicate is not quickly lost, it is likely to undergo subfunctionalization in which the two duplicates take on different, complementary, roles (Lynch and Force 2000). Also, owing to dosage effects, duplications of interacting gene partners often are maintained together (Birchler et al. 2001; Papp et al. 2003; Amoutzias et al. 2004). Therefore, uncovering duplications in genes that are potentially interacting partners in a pathway (*CYC* and *DIV* in *Antirrhinum*) are of interest in understanding patterning pathways and their evolution.

The data on the expression and function of *DIV* are confined to *Antirrhinum* in the DIV1 clade only (Almeida et al. 1997). Corley et al. (2005) hypothesize that CYC and its immediate sister copy DICH (both in the CYC2 clade: Howarth and Donoghue 2006) together turn on *RADIALS* (*RAD*, another MYB transcription factor in a separate clade). RAD, in turn, may down regulate or compete with DIV, given that *RAD* has a single MYB domain resembling a truncated *DIV* gene. Stevenson et al. (2006) did not find evidence of dimerization domains in *RAD*, but they hypothesized that *RAD* could mimic either of the MYB domains of *DIV*. *Cyc* + *dich* mutants and *rad* mutants result in dorsalized radially symmetrical flowers, whereas *div* mutants (in a *cyc* + *dich* mutant background) result in ventralized radial flowers (Luo et al. 1995, 1999; Almeida et al. 1997; Corley et al. 2005). Although the exact nature of the interactions between CYC and DIV is unknown, they evidently interact in the same floral pathway in *Antirrhinum*. Although it is unknown whether this pathway is also utilized outside of *Antirrhinum,* data from *Bournea* suggest that it may be conserved at least across the Lamiales (Zhou et al. 2008). Costa et al. (2005) recently showed that *Antirrhinum* CYC cannot turn on *RAD*-like genes in *Arabidopsis*, however, it is unknown whether TCP1 (the endogenous CYC ortholog in *Arabidopsis*) can turn on *Arabidopsis RAD*-like genes.

Although there are no published data confirming similar DIV function outside of *Antirrhinum*, data do indicate that a downstream target of DIV has broader function. *AmMYBML1* is another transcription factor, sister to *MIXTA*, in the R2R3 clade of MYB genes (Perez-Rodriguez et al. 2005). Both AmMYBML1 and MIXTA affect petal morphology in *Antirrhinum*, with the former being specific to the ventral petal. Perez-Rodriguez et al. (2005) showed that AmMYBML1 had a similar effect in *Nicotiana* when it was ectopically expressed and that it is a downstream target of DIV. This suggests that DIV function may extend at least across the asterid clade Lamiidae (sensu Cantino et al. 2007). In contrast, expression data from *Arabidopsis* indicate that the *RAD*-like genes may not function in the same way in that model (Baxter et al. 2007). *RAD*-like genes were not expressed in a dorsal–ventral pattern in *Arabidopsis*, although location of expression was not found for At4g39250. This pattern could mean that the pathway of interactions between RAD, DIV, and CYC may be specific to the entire Asteridae or to the Lamiidae. It could also mean, however, that *Arabidopsis* and its relatives, with radial, four-merous flowers, have lost a pathway that otherwise spans the core eudicots.

Within *CYC*-like genes, data from other core eudicots strongly suggest that genes from the CYC2 clade play a role in dorsal–ventral patterning. Specifically, genes from the CYC2 clade are dorsally expressed in *Antirrhinum*, *Arabidopsis*, *Iberis*, *Lonicera*, *Lotus*, *and Lupinus* (Luo et al. 1995; Cubas

et al. 2001; Citerne et al. 2006; Feng et al. 2006; Howarth and Donoghue 2006; Busch and Zachgo 2007), and mutants in the rosid *Lotus* also confer a radial flower (Feng et al. 2006), suggesting conserved function. Preliminary data from CYC3, the sister clade of CYC2, indicates that this clade may also play a role in dorsal–ventral patterning. In *Lonicera*, DipsCYC3B is more clearly expressed in the ventral portion of the corolla (the opposite pattern of CYC2) (Howarth and Donoghue 2006). Therefore, two core eudicot sister copies of *CYC*-like genes may both be important in dorsal–ventral patterning. Additionally, many independent duplication events within different lineages of core eudicots (predominately in bilaterally symmetrical species) occur in both the CYC2 and the CYC3 clades (Howarth and Donoghue 2005, 2006).

We provide the first phylogenetic analyses of *DIV*-like genes, which indicate that the *DIV*-like genes have generally duplicated in parallel with *CYC*-like genes, one of their potential interacting partners. It is striking, for instance, that *DIV* and *CYC* both maintain three copies in core eudicots, and that several further duplications appear to have occurred along the same branches in the species tree for Dipsacales, coincident with certain changes in flower symmetry. Clearly, it will be important to assay the other two *DIV*-like copies (DIV2 and DIV3) to see if either of these plays a role in flower development, especially because it appears that a second CYC-like copy (CYC3) may do so (Howarth and Donoghue 2006). Our preliminary rtPCR data point to *DipsDIV2A* as the most likely candidate given its heavier expression in corolla tissue (fig. 3).

In *DIV*-like genes, there have been multiple duplications in the two major clades of Dipsacales, the radially symmetrical Adoxaceae and the bilaterally symmetrical Caprifoliaceae. It is not clear, however, exactly how the Adoxaceae and Caprifoliaceae clades of *DIV*-like genes are related. DIV1 contains two copies in Adoxaceae and two copies in Caprifoliaceae. The Adoxaceae DIV1 copies tend to link with the *DipsDIV1A* clade of Caprifoliaceae, but it is possible that the Dipsacales sequences within DIV1 have been mis-rooted using the *Antirrhinum* sequences and that a duplication event took place before the diversification of the Dipsacales. DIV2 contains two clades of Caprifoliaceae and no sampled sequences from Adoxaceae. The absence from Adoxaceae could be real, given that we sampled heavily from several lineages to find a copy in this clade. Finally, DIV3 is similar to CYC2 and CYC3 with a single Adoxaceae clade and a duplication inferred along the line leading to the bilaterally symmetrical Caprifoliaceae. Adoxaceae does not appear clearly to be sister to both Caprifoliaceae clades, however, and instead appears as a polytomy or as a paraphyletic grade to *DipsDIV3A*.

Additional duplications in specific lineages within Dipsacales have also been identified, as they were in the *CYC*-like genes (Howarth and Donoghue 2005). As in the *CYC*-like genes, the largest number of copies is found in *Morina,* which contains two copies of *DipsDIV1A*, three copies of *DipsDIV1B*, and two copies of *DipsDIV3B*. *Morina* is the only taxon in which we see evidence for duplication outside of the DIV1 clade. Chromosome counts from *Morina* indicate that it is a polyploid (Verlaque 1985; Benko-Iseppon 1992); however, it appears that not all copies in each clade have been maintained. Other duplications are seen in *Sixalix* of the Dipsacaceae (or perhaps across Dipsacaceae + Valerianaceae) in *DipsDIV1A* and in *Valerianella* + *Fedia* in *DipsDIV1B*; these show no clear evidence of genome doubling (Verlaque 1985; Benko-Iseppon 1992).

The presence of additional copies in these lineages is potentially of great interest, as these clades are marked by several major floral changes. Morinaceae are characterized by an unusual bilateral calyx (in addition to a bilateral corolla). Specifically, it appears from developmental studies that the calyx in *Morina* has only four lobes, with the medial dorsal lobe aborting very early in development (Hofmann and Göttmann 1990). It is possible that the additional duplications in *DIV*-like genes, along with the duplications in the potentially interacting *CYC*-like genes in this lineage, are related to the new function of producing a bilaterally symmetrical calyx. There is evidence that the expansion of *CYC* expression can result in the abortion of two lateral stamens in the *Antirrhinum* relative, *Mohavea* (Hileman et al. 2003), and we have previously postulated that the expansion of expression outward to the calyx could cause calyx lobe abortion (Howarth and Donoghue 2005).

It is clear that the lineages Morinaceae, Valerianaceae, and Dipsacaceae contain common duplications within both the *CYC* (Howarth and Donoghue 2005) and *DIV* gene families. These are all groups with divergent floral morphologies including such characteristics in different groups as a zygomorphic calyx, an epicalyx (sometimes also zygomorphic), an asymmetric corolla, ventral and asymmetric stamens abortion, and multiple flower symmetries in a single inflorescence head. These data, along with the largest number of CYC duplications found to date in the florally complex *Helianthus* (Chapman et al. 2008), imply that copy number in these gene families may be correlated with divergent floral traits. With our current taxon sampling we cannot place these duplications precisely, and thus correlate them with these floral transitions. However, future studies will aim to pin down these duplications within these groups.

Expression Patterns

In *Antirrhinum*, flower symmetry has been shown to involve the interplay between TCP genes (*CYC* and *DICH*) and MYB genes (*RAD* and *DIV*) (Corley et al. 2005). CYC and its duplicate in Antirrhineae, DICH (Hileman and Baum 2003), pattern the dorsal portion of the flower (Luo et al. 1995). DIV, on the other hand, has been shown to confer ventral identity in *Antirrhinum* (Almeida et al. 1997). A model proposed by Corley et al. (2005) suggests that CYC and DICH turn on *RAD* expression and that RAD in turn inhibits DIV function. In *Antirrhinum*, DIV is expressed throughout the flower in early stages and becomes localized to the corolla by stage 9 (Galego and Almeida 2002). In the ventral portion of *Antirrhinum* flowers, *DIV* becomes further localized to the inner epidermis of ventral and lateral petals, whereas it remains diffuse in dorsal regions (Galego and Almeida 2002).

The ECE clade (Howarth and Donoghue 2006) of TCP genes (especially the CYC2 clade) has been the subject of intense interest (see discussion above). Unlike CYC2

genes, however, no functional data are available for *DIV*-like genes outside of *Antirrhinum*. Therefore, it is unknown where or when these genes may be utilized in a similar manner to the model. Expression data from *Bournea*, however, show similar early expression of *DIV*-like genes between it and *Antirrhinum* (Zhou et al. 2008). Here we present the first expression data from *DIV* outside of Lamiales (includes *Antirrhinum* and *Bournea*). Five copies of *Heptacodium* (Dipsacales) were sequenced from this core RR DIV clade, two of which are orthologous to *Antirrhinum DIV* (Dips*DIV*1A and Dips*DIV*1B). Figure 3 shows the rtPCR expression in varying tissues of each of these five copies. We believe that the sixth Dipsacales copy was simply not recovered by cloning, although it is possible that it is was lost in *Heptacodium*.

Within DIV1 (which also includes *Antirrhinum DIV*), expression of the gene copy in the *DipsDIV1A* clade shows the greatest possibility for similar function to *Antirrhinum DIV*, being expressed only in petal tissue. Additionally, there is differing expression along the dorsal–ventral axis. Although expression appears clear in ventral petals, it also occurs unexpectedly in dorsal petals and is lacking in lateral petals. It is possible, however, that this reflects the pattern of expression only in the relatively late buds needed for dissection. In *Antirrhinum*, later stage *DIV* expression remains diffuse in dorsal petals, localizes in lateral and ventral petals, and ultimately wanes from lateral petals (Galego and Almeida 2002). Just the end of this pattern may be what we are seeing here. In any case, these data support that DIV may play a role in dorsal–ventral patterning outside of *Antirrhinum*, although we cannot say how similar it is in expression or function to the model.

Unlike *DipsDIV1A*, *DipsDIV1B* is strongly expressed in all tissues sampled and was also the copy most commonly amplified from floral cDNA across Dipsacales (fig. 2). This is in stark contrast to the clearly paralogous duplicate copy in *Antirrhinum* in this clade (*DVL1*), which is found only weakly in ovules (Galego and Almeida 2002), although there is no reason to expect these copies to be expressed in the same way because they are the result of independent duplications. Like *DipsDIV1A*, *DipsDIV2A* is expressed strongly in corolla containing tissue, although it is also present elsewhere. *DipsDIV2B* and DipsDIV3A are expressed everywhere, including leaves.

In summary, within the bilaterally symmetrical Caprifoliaceae, there are six copies in this core RR DIV clade. Data from cDNA (fig. 2) and from the *Heptacodium* dissections (fig. 3) suggest that all are expressed in flowers. Additionally, *DipsDIV1A* is expressed most clearly in the corolla. *DipsDIV1A*, like *Antirrhinum DIV*, may well play a role in dorsal–ventral patterning and lacks expression in lateral petals.

Inferring Species Phylogeny

There has been increasing demand in plant systematics for low-level phylogenetic markers, given that many of the commonly sequenced regions are nearly identical at the species level. Nuclear ribosomal internal transcribed spacer (Baldwin et al. 1995) has been the marker of choice, but it is often not variable enough, paralogous copies may be retained, and, owing to concerted evolution, it loses important information for understanding the dynamics of speciation (e.g., hybridization) (Wendel et al. 1995; Buckler et al. 1997; Kim et al. 2008). Recent attention has focused on nuclear introns, due to the potential for high variation, low copy number, the possibility of universal priming sites in flanking exons, and their bipaternal history. Introns from genes involved in flower development, such as *LEAFY*, *PISTILLATA*, *APETALA3/TM6*, and *AGAMOUS*, for instance, have been used to examine closely related species and to uncover hybridization (Barrier et al. 1999; Lee et al. 2002; Nishimoto et al. 2003; Howarth and Baum 2005; Kim et al. 2008).

DIVARICATA contains an intron flanked by two conserved MYB domains, making universal primer design feasible. The location of this intron is conserved, at least in the core RR DIV clade. The intron appears to have considerable variation among and even within species (0.01–0.043 pairwise differences). In the *S. atropurpurea* sequences from *DipsDIV1A*, the intron is AT rich (A = 0.30, T = 0.38, C = 0.14, and G = 0.18), which is similar to other variable introns (Carels and Bernardi 2000; Howarth and Baum 2005). Unlike other AT-rich introns, however, there are no long strings of As or Ts, which necessitate cloning. *DIV* could be directly sequenced in many individuals, although indels among alleles would likely necessitate cloning in a subset of individuals. Although duplications appear to be common, homology assessments should be clear, as introns from even closely related *DIV*-like copies in Dipsacales are unalignable with one another.

Future Directions

The data we present here explore the diversification of a family of genes that seem to be involved with *CYC* and *RAD* in the flower symmetry pathway (Corley et al. 2005). Understanding these gene trees could allow us to work out the codiversification of these presumably interacting genes, and the evolution of the whole system that results in symmetry and evolutionary changes in symmetry. It is encouraging that there does, so far, appear to be a relationship between the diversification of *DIV*-like and *CYC*-like genes (Howarth and Donoghue 2006), two different transcription factors. We have focused on Dipsacales and have found complex duplication patterns within this group in both of these gene families. We now need to extend the sampling further to pin down where critical duplication events occurred in angiosperm phylogeny. This is especially the case for the duplications that seem to have taken place near the origin of the core eudicots in both *CYC* and *DIV* lineages as well as in several separate classes of MADS-box genes (Litt and Irish 2003; Kim et al. 2004; Kramer et al. 2004; Kramer and Hall 2005; Zahn et al. 2005; Howarth and Donoghue 2006). These data, together with genome studies (Bowers et al. 2003; De Bodt et al. 2005; Jaillon et al. 2007), imply that the diversification of the core eudicots may have been marked by a series of whole genome duplications, but it is necessary to improve the comparability of the sampling of key eudicot lineages to test more precisely the coincidence of these duplication events.

Additionally, the expression work that we have started here with rtPCR must be extended. Our data suggest that *DIV*

expression may be similar in Dipsacales and in Lamiales and thus conserved at least across asterid angiosperms. It will be important to determine, however, what parts of these pathways are similar across asterids and across core eudicots, and whether the model really does apply across these groups. Also, it is unknown how many of these duplicated gene copies actually interact with one another and thus potentially play a role in floral symmetry. We are now expanding the expression data with in situ hybridization and are beginning to examine function using virus-induced gene silencing (VIGS) to knockdown these genes in Dipsacales.

Acknowledgments

We thank the following people for gifts of DNA: E. Kramer (Harvard University, Cambridge, MA; *Aquilegia* DNA), A. Litt (New York Botanical Garden, Bronx, NY; Valerianaceae cDNA), and S. Mathews (Arnold Arboretum of Harvard University, Cambridge, MA; noneudicots DNA). We also thank C. Bell, S. Carlson, S. Smith, D. Tank, and R. Winkworth for the use of DNA extractions in the M.J.D. laboratory. We thank M. Dunn and members of the M.J.D. laboratory for helpful discussions. D.G.H. is grateful for a Forest B. H. and Elizabeth D. W. Brown Postdoctoral Fellowship from Yale University and for a travel grant from the Molecular and Organismic Research in Plant History (MORPH) Research Coordination Network. The work of M.J.D. on angiosperm phylogeny is supported by an Assembling the Tree of Life (ATOL) grant from the National Science Foundation.

Literature Cited

Akaike H. 1973. Information theory as an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, editors. Second international symposium on information theory. Budapest: Akademiai Kiado. p. 267–281.

Almeida J, Rocheta M, Galego L. 1997. Genetic control of flower shape in *Antirrhinum majus*. Development. 124: 1387–1392.

Amoutzias GD, Robertson DL, Oliver SG, Bornberg-Bauer E. 2004. Convergent evolution of gene networks by single-gene duplications in higher eukaryotes. EMBO Rep. 5:1–6.

Backlund A, Donoghue MJ. 1996. Morphology and phylogeny of the order Dipsacales. phylogeny of the Dipsacales. Department of Systematic Botany Uppsala University.Backlund A, Ph.D. dissertation.

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann Mo Bot Gard. 82:247–277.

Barrier M, Baldwin BG, Robichaux RH, Purugganan MD. 1999. Interspecific hybrid ancestry of a plant adaptive radiation: allopolyploidy of the Hawaiian silversword alliance (Asteraceae) inferred from floral homeotic gene duplications. Mol Biol Evol. 16:1105–1113.

Baxter CEL, Costa MMR, Coen ES. 2007. Diversification and co-option of RAD-like genes in the evolution of floral asymmetry. Plant J. 52:105–113.

Bell CD, Donoghue MJ. 2005. Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. Org Divers Evol. 5:147–159.

Bell CD, Edwards EJ, Kim ST, Donoghue MJ. 2001. Dipsacales phylogeny based on chloroplast DNA sequences. Harvard pap. Bot. 6:481–499.

Benko-Iseppon AM. 1992. Karyologische Untersuchung der Caprifoliaceae s.l. und möglicher verwandter Familien. Vienna: Naturwiss. Fak. Univ. Wien.

Birchler JA, Bhadra U, Pal Bhadra M, Auger D. 2001. Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantitative traits. Dev Biol. 234:275–288.

Bowers JE, Chapman BA, Rong J, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature. 422:433–438.

Brohom SK, Tähtiharju S, Laitinen RAE, Albert VA, Teeri TH, Elmaa P. 2008. A TCP domain transcription factor controls flower type specification along the radial axis of the Gerbera (Asteraceae) inflorescence. Proc Nat Acad Arts Sci. 105:9117–9122.

Buckler ES, Ippolito A, Holtsford TP. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics. 145:821–832.

Busch A, Zachgo S. 2007. Control of corolla monosymmetry in the Brassicaceae *Iberis amara*. Proc Nat Acad Arts Sci. 104: 16714–16719.

Cantino PD, Doyle JA, Graham SW, Judd WS, Olmstead RG, Soltis PS, Soltis DE, Donoghue MJ. 2007. Towards a phylogenetic nomenclature of *Tracheophyta*. Taxon. 56:822–846.

Carels N, Bernardi G. 2000. Two classes of genes in plants. Genetics. 154:1819–1825.

Chapman MA, Leebens-Mack JH, Burke JM. 2008. Positive selection and expression divergence following gene duplication in the sunflower *CYCLOIDEA* gene family. Mol Biol Evol. 25:1260–1273.

Citerne HL, Pennington RT, Cronk QCB. 2006. An apparent reversal in floral symmetry in the legume *Cadia* is a homeotic transformation. Proc Natl Acad Arts Sci. 103: 12017–12020.

Coen ES, Meyerowitz EM. 1991. The war of the whorls: genetic interactions controlling flower development. Nature. 353:31–37.

Corley SB, Carpenter R, Copsey L, Coen E. 2005. Floral asymmetry involves an interplay between TCP and MYB transcription factors in *Antirrhinum*. Proc Nat Acad Sci. 102:5068–5073.

Costa MMR, Fox S, Hanna AI, Baxter C, Coen E. 2005. Evolution of regulatory interactions controlling floral asymmetry. Development. 132:5093–5101.

Cubas P. 2002. Role of TCP genes in the evolution of morphological characters in angiosperms. In: Cronk QCB, Bateman R, Hawkins J, editors. Developmental genetics and plant evolution. London: Taylor and Francis. p. 247–266.

Cubas P, Coen E, Zapater JMM. 2001. Ancient asymmetries in the evolution of flowers. Curr Biol. 11:1050–1052.

Cubas P, Lauter N, Doebley J, Coen E. 1999. The TCP domain: a motif found in proteins regulating plant growth and development. Plant J. 18:215–222.

Cubas P, Vincent C, Coen E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. Nature. 401:157–161.

De Bodt S, Maere S, Van de Peer Y. 2005. Genome duplication and the origin of angiosperms. Trends Ecol Evol. 20: 591–597.

Donoghue MJ, Bell CD, Winkworth RC. 2003. The evolution of reproductive characters in Dipsacales. Int J Pl Sci. 164(5 Suppl.):S453–S464.

Donoghue MJ, Eriksson T, Reeves PA, Olmstead RG. 2001. Phylogeny and phylogenetic taxonomy of Dipsacales, with special reference to *Sinadoxa* and *Tetradoxa* (Adoxaceae). Harvard Pap Bot. 6:459–479.

Donoghue MJ, Ree RH. 2000. Homoplasy, and developmental constraint: a model and an example from plants. Am Zool. 40:759–769.
Donoghue MJ, Ree RH, Baum DA. 1998. Phylogeny and the evolution of flower symmetry in the Asteridae. Trends Plant Sci. 3:311–317.
Endress PK. 1996. Diversity and evolutionary biology of tropical flowers. Cambridge: Cambridge University Press.
Endress PK. 1999. Symmetry in flowers—diversity and evolution. Int J Pl Sci. 160:S3–S23.
Feng X, Zhao Z, Tian Z, et al. (21 co-authors). 2006. Control of petal shape and floral zygomorphy in *Lotus japonicus*. Proc Natl Acad Arts Sci. 103:4970–4975.
Fukuoka N. 1972. Taxonomic study of Caprifoliaceae. Mem Fac Sci Kyoto Univ Ser Biol. 6:15–58.
Galego L, Almeida J. 2002. Role of *DIVARICATA* in the control of dorsoventral asymmetry in *Antirrhinum* flowers. Genes Dev. 16:880–891.
Gillies ACM, Cubas P, Coen ES, Abbott RJ. 2002. Making rays in the Asteraceae: genetics and evolution of radiate versus discoid flower heads. In: Cronk QCB, Bateman RM, Hawkins JA, editors. Developmental genetics and plant evolution. London and New York: Taylor & Francis. p. 233–246.
Hileman LC, Baum DA. 2003. Why do paralogs persist? Molecular evolution of *CYCLOIDEA* and related floral symmetry genes in Antirrhineae (Veroniacaceae). Mol Biol Evol. 20:591–600.
Hileman LC, Kramer EM, Baum DA. 2003. Differential regulation of symmetry genes and the evolution of floral morphologies. Proc Natl Acad Sci USA. 100:12814–12819.
Hofmann U, Göttmann J. 1990. *Morina* L. and *Triplostegia* Wall. ex DC. im Vergleich mit Valerianaceae and Dipsacaceae. Bot Jahrb Syst. 111:499–553.
Howarth DG, Baum DA. 2005. Genealogical evidence of homoploid hybrid speciation in an adaptive radiation of *Scaevola* (Goodeniaceae) in the Hawaiian Islands. Evolution. 59:948–961.
Howarth DG, Donoghue MJ. 2005. Duplications in *CYC*-like genes from Dipsacales correlate with floral form. Int J Pl Sci. 166:357–370.
Howarth DG, Donoghue MJ. 2006. Phylogenetic analysis of the ''ECE'' (CYC TB1) clade reveals duplications predating the core eudicots. Proc Natl Acad Arts Sci. 103:9101–9106.
Huelsenbeck JP, Ronquist F. 2001. MrBayes: bayesian inference of phylogenetic trees. Bioinformatics. 17:754–755.
Jaillon O, Aury J-M, Noel B, et al. (56 co-authors). 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 449:463–468.
Judd WS, Sanders RW, Donoghue MJ. 1994. Angiosperm family pairs – preliminary phylogenetic analyses. Harvard Pap Bot. 5:1–51.
Kim M, Cui M-L, Cubas P, Gillies A, Lee K, Chapman MA, Abbott RJ, Coen E. 2008. Regulatory genes control a key morphological and ecological trait transferred between species. Science. 322:1116–1119.
Kim S, Yoo M-J, Albert VA, Farris JS, Soltis PS, Soltis DE. 2004. Phylogeny and diversification of B-function MADS-box genes in angiosperms: evolutionary and functional implications of a 260 million year old duplication. Am J Bot. 91:2102–2118.
Kim S-T, Sultan SE, Donoghue MJ. 2008. Allopolyploid speciation in Persicaria (Polygonaceae): insights from a low-copy nuclear marker. Proc Natl Acad Sci USA.
Knapp S. 2002. Floral diversity and evolution in the Solanaceae. In: Cronk QCB, Bateman RM, Hawkins JA, editors. Developmental genetics and plant evolution. London and New York: Taylor & Francis. p. 267–297.
Kramer EM, Hall JC. 2005. Evolutionary dynamics of genes controlling floral development. Curr Opin Plant Biol. 8: 13–18.
Kramer EM, Jaramillo MA, Di Stilio VS. 2004. Patterns of gene duplication and functional evolution during the diversification of the *AGAMOUS* subfamily of MADS box genes in angiosperms. Genetics. 166:1011–1023.
Lee JY, Mummenhoff K, Bowman JL. 2002. Allopolyploidization and evolution of species with reduced floral structures in *Leidium* L. (Brassicaceae). Proc Natl Acad Arts Sci. 99:16835–16840.
Litt A, Irish VF. 2003. Duplication and diversification in the *APETALA1/FRUITFULL* floral homeotic gene lineage: implications for the evolution of floral development. Genetics. 165:821–833.
Luo D, Carpenter R, Copsey L, Vincent C, Clark J, Coen E. 1999. Control of organ asymmetry in flowers of *Antirrhinum*. Cell. 99:367–376.
Luo D, Carpenter R, Vincent C, Copsey L, Coen E. 1995. Origin of floral asymmetry in *Antirrhinum*. Nature. 383:794–799.
Lynch M, Force A. 2000. The probability of duplicate gene preservation by subfunctionalization. Genetics. 154:459–473.
Maddison DR, Maddison WP. 2003. MacClade: analysis of phylogeny and character evolution. Sunderland (MA): Sinauer Associates.
Magallon S, Crane PR, Herendeen PS. 1999. Phylogenetic pattern, diversity, and diversification of eudicots. Ann Mo Bot Gard. 86:297–372.
Martin C, Paz-Ares J. 1997. MYB transcription factors in plants. Trends Genet. 13:67–73.
Meyerowitz EM, Bowman JL, Brockman LL, Drews GN, Jack T, Sieburth LE, Weigel D. 1991. A genetic and molecular model for floral development in *Arabidopsis thaliana*. Development. Suppl 1:157–167.
Neal PR. 1998. Floral symmetry and its role in plant-pollinator systems: terminology, distribution, and hypotheses. Annu Rev Ecol Syst. 29:345–373.
Nishimoto Y, Ohnishi O, Hasegawa M. 2003. Topological incongruence between nuclear and chloroplast DNA trees suggesting hybridization in the urophyllum group of the genus *Fagopyrum* (Polygonaceae). Genes Genet Syst. 78:139–153.
Paabo S, Irwin DM, Wilson AC. 1990. DNA damage promotes jumping between templates during enzymatic amplification. J Biol Chem. 265:4718–4721.
Papp B, Pál C, Hurst LD. 2003. Dosage sensitivity and the evolution of gene families in yeast. Nature. 424:194–197.
Perez-Rodriguez M, Jaffe FW, Butelli E, Glover BJ, Martin C. 2005. Development of three different cell types is associated with the activity of a specific MYB transcription factor in the ventral petal of *Antirrhinum majus* flowers. Development. 132:359–370.
Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitutions. Bioinformatics. 14:817–818.
Pyck N. 2001. Phylogenetic relationships within Dipsacales: a combined molecular and morphological approach. *Ph.D. dissertation*. Leuven, Belgium: Katholieke University.
Ree RH, Donoghue MJ. 1999. Inferring rates of change in flower symmetry in asterid angiosperms. Syst Biol. 48:633–641.
Reeves PA, Olmstead RG. 2003. Evolution of the TCP gene family in Asteridae: cladistic and network approaches to understanding regulatory gene family diversification and its impact on morphological evolution. Mol Biol Evol. 20: 1997–2009.
Sargent RD. 2004. Floral symmetry affects speciation rates in angiosperms. Proc Roy Soc Lond, B series. 271:603–608.
Soltis DE, Senters AE, Zanis MJ, Kim S, Thompson JD, Soltis PS, Ronse De Craene LP, Endress PK, Farris JS. 2003.

Gunnerales are sister to other core eudicots: implications for the evolution of pentamery. Am J Bot. 90:461–470.
Stevenson CEM, Burton N, Costa MMR, Nath U, Dixon RA, Coen ES, Lawson DM. 2006. Crystal structure of the MYB domain of the RAD transcription factor from *Antirrhinum majus*. Proteins: Struct Funct Bioinf. 65:1041–1045.
Strand AE, Leebens-Mack J, Milligan BG. 1997. Nuclear DNA-based markers for plant evolutionary biology. Mol Ecol. 6:113–118.
Swofford DL. 2001. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sunderland (MA): Sinauer Associates, Inc.
Verlaque R. 1985. Étude biosystématique et phylogénétique des Dipsacaceae. II. Caracteres généraux des Dipsacaceae. Rev Cytol Biol Végét Bot. 8:117–168.
Weberling F. 1989. Morphology of flowers and inflorescences. Cambridge, UK: Cambridge University Press.
Wendel JF, Schnabel A, Seelanan T. 1995. An unusual ribosomal DNA sequence from Gossypium gossypioides reveals ancient, cryptic, intergenomic introgression. Mol Phylogenet Evol. 4:298–313.
Winkworth RC, Bell CD, Donoghue MJ. 2008. Mitochondrial sequence data and Dipsacales phylogeny: mixed models, partitioned Bayesian analyses, and model selection. Mol Phylogenet Evol. 46:830–843.
Winkworth RC, Lundberg J, Donoghue MJ. 2008. Toward a resolution of campanulid phylogeny, with special reference to the placement of Dipsacales. Taxon. 57:1–13.
Yanhui C, Xiaoyuan Y, Kun H, et al. (18 co-authors). 2006. The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol Biol. 60:107–124.
Zahn LM, Kong H, Leebens-Mack JH, Kim S, Soltis PS, Landherr LL, Soltis DE, dePamphilis CW, Ma H. 2005. The evolution of the *SEPALLATA* subfamily of MADS-Box genes: a preangiosperm origin with multiple duplications throughout angiosperm history. Genetics. 169:2209–2223.
Zhang WH, Chen JD, Li J, Chen HB, Tang YC. 2003. Phylogeny of the dipsacales s.l. based on chloroplast *trn*L-F and *ndh*F sequences. Mol Phylogenet Evol. 26:176–189.
Zhou X-R, Wang Y-Z, Smith JF, Chen R. 2008. Altered expression patterns of TCP and MYB genes relating to the floral developmental transition from initial zygomorphy to actinomorphy in *Bournea* (Gesneriaceae). New Phytol. 178:532–543.

Neelima Sinha, Associate Editor

Accepted February 13, 2009