Duplications and Expression of *RADIALIS*-Like Genes in Dipsacales
Author(s): Geraldine S. Boyden, Michael J. Donoghue, and Dianella G. Howarth
Source: *International Journal of Plant Sciences*, Vol. 173, No. 9 (November/December 2012), pp. 971-983
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/10.1086/667626
Accessed: 29/07/2013 14:56

Int. J. Plant Sci. 173(9):971–983. 2012.

1058-5893/2012/17309-0002$15.00 DOI: 10.1086/667626

DUPLICATIONS AND EXPRESSION OF *RADIALIS*-LIKE GENES IN DIPSACALES

Geraldine S. Boyden,* Michael J. Donoghue,† and Dianella G. Howarth[1,*]

*Department of Biological Sciences, St. John's University, 8000 Utopia Parkway, Queens, New York 11439, U.S.A.; and †Department of Ecology and Evolutionary Biology, Yale University, P.O. Box 208106, New Haven, Connecticut 06520, U.S.A.

Evidence from *Antirrhinum majus* indicates that transcription factors CYCLOIDEA (CYC), DIVARICATA (DIV), and RADIALIS (RAD) play a role in specifying floral symmetry. In bilaterally symmetrical flowers, CYC and RAD are implicated in dorsal identity, while DIV patterns ventral identity. In this study, we examined duplication events within the *RAD*-like gene family in Dipsacales and report the phylogenetic relationships of the *RAD*-like genes across Pentapetaleae. Like *CYC*-like and *DIV*-like genes, we found three Pentapetaleae clades of *RAD*-like genes: RAD1, RAD2, and RAD3, with *AmRAD* in the RAD2 clade. Unlike *CYC*-like and *DIV*-like gene families, only one of these duplications appears to have taken place around the diversification of the Pentapetaleae: RAD1 spans the monocots and eudicots, while RAD2 and RAD3 are Pentapetaleae specific. We found additional duplications within these gene clades in Dipsacales, especially within the Morinaceae, Dipsacaceae, and Valerianaceae, which also contain additional duplications in *CYC*-like and *DIV*-like genes. Using reverse transcription polymerase chain reaction, we show that most *RAD* copies are expressed across floral and leaf tissues in *Lonicera* × *bella*. *DipsRAD2B* (orthologous to *AmRAD*) is expressed in a dorsoventral pattern similar to *DipsCYC2B*, which is similar to the result in *A. majus*. We argue that this favors the hypothesis that a similar interaction between CYC and RAD may occur outside of the Lamiales.

Keywords: *RADIALIS*, floral symmetry, Dipsacales, gene duplication, core eudicots, gene expression, MYB transcription factors.

Online enhancement: appendix table.

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). Such morphological shifts are of special interest in relation to shifts in pollination (see Neal 1998) and, ultimately, perhaps to shifts in rates of diversification (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 in relation 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). This orientation sets up the flower in such a way that the two dorsal petals develop in tandem, the two lateral petals develop in tandem, and the single ventral petal is independent.

Research on the evolution of floral symmetry has been greatly advanced by the discovery of three clades of potential candidate genes, *CYCLOIDEA* (CYC; Luo et al. 1996), *DIVARICATA* (DIV; Galego and Almeida 2002), and *RADIALIS* (RAD; Corley et al. 2005). All three of these gene groups were initially characterized in *Antirrhinum majus* (referred to here as *Antirrhinum*), with function in dorsal (CYC, RAD) or ventral (DIV) portions of the flower. These genes fall into two major transcription factor families: TCP (Cubas et al. 1999*a*) and MYB (classified by the strong conservation of imperfect repeats; Martin and Paz-Ares 1997). A model for these 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*, restricting 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. 1996; Corley et al. 2005), while mutants of *div* (in a *cyc* + *dich* background) result in radially symmetrical, lateralized flowers in *Antirrhinum* (Almeida et al. 1997). Similar phenotypes resulting from CYC loss of function have been corroborated in the closely related *Linaria vulgaris* (Cubas et al. 1999*b*) as well as in two legumes, *Lotus japonica* and *Pisum sativum* (Feng et al. 2006; Wang et al. 2008). Additionally, CYC orthologs in Asteraceae play a role in specifying disk versus ray florets (Broholm 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*, Dipsacales, *Lupinus*, and Malpighiaceae (Cubas et al. 2001; Citerne et al. 2006; Howarth and Donoghue 2006; Busch and

[1] Author for correspondence; e-mail: howarthd@stjohns.edu.

Manuscript received August 2011; revised manuscript received June 2012.

Zachgo 2007; Zhou et al. 2008; Zhang et al. 2010; Howarth et al. 2011). All of these data support parallel recruitment of *CYC*-like genes in specifying floral symmetry and therefore a similar functional system across rosids and asterids (Preston and Hileman 2009). Unlike *CYC*, however, little is known outside of *Antirrhinum* about the expression or function of the MYB genes (*DIV* and *RAD*). Two studies using expression data in other Veronicaceae and in *Bournea* (Gesneriaceae), however, are consistent with the model of *CYC*-like, *DIV*-like, and *RAD*-like gene interaction being conserved at least across Lamiales (Zhou et al. 2008; Preston et al. 2009). Additionally, our previous data on *DIV* expression in *Heptacodium* (Dipsacales) indicated that there is a dorsoventral expression pattern of a *DIV* ortholog in the corolla, supporting the possibility that the pathway is similarly co-opted across asterids (Howarth and Donoghue 2009).

MYB transcription factors, including *RAD* and *DIV*, comprise the largest gene family in *Arabidopsis*, with 198 paralogs (Yanhui et al. 2006), thus contributing to challenges in determining orthology across taxa. MYB genes are highly variable in sequence but are characterized by three α-helices that form a compact bundle with three regularly spaced Trp residues (Stevenson et al. 2006). One subset, the MYB-related family, which contains a total of 64 members in *Arabidopsis*, includes *RAD* and *DIV*. This group is further divided into five subfamilies of genes in which the I-box-like subfamily includes *RAD* and the R-R-type subfamily includes *DIV* (Yanhui et al. 2006). These two subfamilies appear to be sister groups and were previously lumped by Riechmann and Ratcliffe (2000) together under the name I-box-like because of the similarity of the first MYB domain of *DIV* to the single MYB domain of *RAD*. It is possible that I-box-like genes such as RAD were the result of the loss of the second MYB domain of an R-R-type gene such as *DIV* (Stevenson et al. 2006). RAD, which is only 93 aa long, shares significant similarity with the N terminal MYB domain of DIV in the first two helices. The third helix contains only 30% sequence similarity, however, and is notably longer in RAD. Stevenson et al. (2006) suggest that the DNA binding domain of RAD and DIV are in the third helix, and therefore these two proteins may not be binding to the same regions. This clade of I-box-like genes includes six *Arabidopsis* sequences, six *Antirrhinum* sequences, and as many as nine *Oryza* sequences (Yanhui et al. 2006; Baxter et al. 2007). Orthology and evolution of the *RAD*-like gene family have remained unclear because of a lack of taxon sampling across eudicots.

The greatest angiosperm diversity of species and floral forms is found in the Pentapetaleae (core eudicots excluding Gunnerales; see Cantino et al. 2007), which makes up approximately 70% of extant angiosperms. Although lineages of the eudicots are more labile with multiple shifts in merosity and phyllotaxy, flowers of the Pentapetaleae are based on a pentamerous ground plan (as opposed to the dimerous ground plan that characterizes the so-called basal eudicots) and demarcated floral whorls (Magallon et al. 1999; Soltis et al. 2003; Endress 2010, 2011). It has been argued that a large proportion of plant diversity may have arisen following the duplication and subsequent adaptation of preexisting genes (Flagel and Wendel 2009; Van de Peer et al. 2009). Genome comparisons point to a whole genome duplication or even a hexaploidization event that may have occurred before the diversification of the rosids and may be coincident with the origin of the Pentapetaleae (Bowers et al. 2003; De Bodt et al. 2005; Jaillon et al. 2007). This hypothesis has been since supported by other genomic analyses (Tang et al. 2008; Cenci et al. 2010; Chan et al. 2010; Schmutz et al. 2010). Our previous data from other members of the floral symmetry pathway, specifically the ECE clade of *CYC*-like genes (Howarth and Donoghue 2006) and the core RR DIV clade of *DIV*-like genes (Howarth and Donoghue 2009), support this hypothesis, with evidence that two duplications occurred in each group around the divergence of the Pentapetaleae. In this study, we therefore aimed to determine whether *RAD*-like genes also duplicated around the divergence of the Pentapetaleae.

Similar to our recent studies in *CYC*-like and *DIV*-like genes, we utilize here the Dipsacales clade as a focal point, along with sequences from published genomes, to determine duplication patterns of *RAD*-like genes across Pentapetaleae. Our focus on Dipsacales was motivated by several factors. The Dipsacales phylogeny is quite well resolved based on a series of molecular and morphological phylogenetic analyses (Judd et al. 1994; Backlund and Donoghue 1996; Bell et al. 2001; Donoghue et al. 2001, 2003; Pyck 2001; Zhang et al. 2003; Jacobs et al. 2011). This backbone lineage phylogeny (fig. 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 the *RAD*-like gene family, including the location of gene duplications.

A variety of floral forms are found within Dipsacales, including radial symmetry, bilateral symmetry, and asymmetric flowers (Donoghue et al. 2003). Recent phylogenetic analyses of campanulid angiosperms support the radially symmetrical Paracryphiaceae as sister to Dipsacales (Winkworth et al. 2008; Tank and Donoghue 2010), which supports an ancestral condition of radial symmetry in the entire clade. The primary split within Dipsacales separates radially symmetrical Adoxaceae (including *Viburnum*, *Sambucus*, and *Adoxa* and its relatives) from the bilaterally symmetrical Caprifoliaceae sensu lato (including Diervilleae, Caprifolieae, Linnaeeae, Morinaceae, Dipsacaceae, and Valerianaceae; Fukuoka 1972; Donoghue et al. 2003). It now appears that Linnaeeae may not be monophyletic, with some more closely related to Morinaceae (Jacobs et al. 2011). In any case, Caprifoliaceae

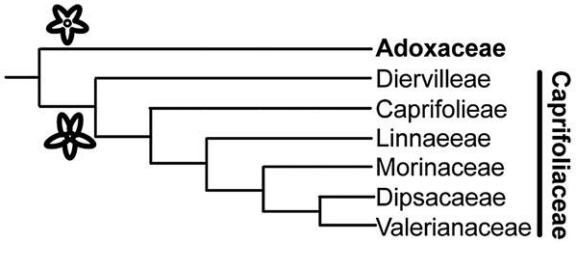

Fig. 1 Simplified Dipsacales phylogeny showing the major lineages. The Dipsacales are divided into two major clades, the radially symmetrical Adoxaceae and the bilaterally symmetrical Caprifoliaceae.

appear to represent an independent derivation of bilateral symmetry within the Asteridae, widely separated from the origin of bilateral flowers in the clade that includes *Antirrhinum*. Within the Caprifoliaceae, there are several forms of bilateral symmetry present. 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. 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 stamen abortion, the number of corolla lobes, a bilaterally symmetrical calyx, and the formation of an epicalyx (Donoghue et al. 2003).

Along with broad duplications leading to three paralogs of *CYC*-like and *DIV*-like genes in Pentapetaleae, several additional duplications have also occurred within Dipsacales. Within the CYC2 and CYC3 clades, for instance, duplications occurred around the divergence of the Caprifoliaceae, coincident with the shift to bilateral symmetry (Howarth and Donoghue 2005). Duplications within Dipsacales also occurred in all three lineages of *DIV*-like genes (DIV1, DIV2, and DIV3), although it remains unclear whether these duplications span the Caprifoliaceae or the entire Dipsacales (Howarth and Donoghue 2009). Additional duplications also occurred in these gene lineages in Morinaceae, Dipsacaceae, and Valerianaceae. Based on these findings and the presumed functional interactions among these genes, we hypothesized that *RAD*-like genes might also have duplicated within the Dipsacales.

In this article, we address the evolution of *RADIALIS*-like genes within the Dipsacales and relate this to general patterns of gene family evolution in eudicots. This study provides the opportunity to examine the coevolution of three independent transcription factors that interact to govern a major floral trait. In addition to highlighting *RAD* duplications in the Dipsacales and across the Pentapetaleae, we present reverse transcriptase PCR (RT-PCR) expression data from one Dipsacales species, *Lonicera* × *bella*, a hybrid cultivar between *L. morrowii* A. Gray and *L. tartarica* L., for each of these copies.

Material and Methods

Dipsacales Sampling

Twenty-one individuals were sequenced representing all major lineages within Dipsacales. We primarily used total genomic DNAs, obtained in many cases from extractions used in previous phylogenetic studies. Some species were sampled using floral cDNA, extracted from whole flower bud tissue. Appendix A provides a list of included taxa, voucher specimens, and GenBank accession numbers for each gene copy. All primers were designed on either end of the single MYB domain. DNA was amplified from each sampled species using combinations of the primers listed in table B1, available in the online edition of the *International Journal of Plant Sciences*. 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 (Promega) in 25 μL, with final concentrations of 2.5 mM $MgCl_2$, 0.5 μM of each primer, 0.8 μM dNTPs. Amplified products were cloned using the StrataClone PCR Cloning Kit (Stratagene). Between 10 and 40 colonies were screened for all potentially different copies or alleles of the *RAD*-like genes obtained. Colonies were selected and used to inoculate 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 30 cycles. Amplification products of the appropriate size were cleaned using a PEG 8000/NaCl precipitation protocol and directly sequenced.

Genome Mining

In our primary analyses *RAD*-like genes were included from four non-Dipsacales taxa, *Antirrhinum majus*, *Arabidopsis thaliana*, *Populus trichocarpa*, and *Vitis vinifera*. *Antirrhinum* and *Arabidopsis* each have six published copies, which were included in our analysis (Yanhui et al. 2006; Baxter et al. 2007). We additionally searched the genomes of *Populus* (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html) and *Vitis* (http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/) to determine their complement of *RAD*-like copies.

In order to determine the rooting of the Pentapetaleae genes, we used published genomic data from the eudicot *Aquilegia coerulea* and the monocot *Oryza sativa* in some analyses. These analyses included *RAD*-like genes obtained from an *Aquilegia* EST database search (http://www.phytozome.net/aquilegia.php) and published *Oryza* sequences (Yanhui et al. 2006).

Alignment and Phylogenetic Analyses

All clones from each DNA extraction (obtained using multiple primer pairs) were compiled into Sequencher 4.2 (Gene Codes, Ann Arbor, MI). All sequenced genes with a MYB domain were included in the matrix. Positive clones were separated into different "clusters" based on shared differences among the clones. These clusters were essentially identical to each other but varied by obvious polymerase error (single base differences in one or two clones out of dozens, with different clones being mutated at different sites). 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 cluster and exported for phylogenetic analysis. Protein sequences for each potential copy/allele obtained were aligned by eye in MacClade 4 (Maddison and Maddison 2003), and the protein-aligned nucleotide matrix was used in subsequent phylogenetic analyses. We hypothesize gene duplication instead of allelic variation when each "copy" is found across multiple species, since gene flow across multiple geographically and phylogenetically divergent species would be un-

likely, although we cannot rule out entirely the possibility of deep coalescence or lineage sorting.

The matrix for each analysis included the entire region between the primer pairs, with only a single amino acid alignment ambiguity within the MYB domain. Parameters for the Bayesian and maximum likelihood (ML) analyses were estimated using jmodeltest 0.1 (Posada and Crandall 1998). The Akaike Information Criterion (AIC; Akaike 1973) was used to recommend the following models for each analysis: TVMef+i+g for the total Pentapetaleae matrix, TIM3+i+g for the clade of Dipsacales copy 1 (156 aligned bases), and K80+G for Dipsacales copy 3 (83 aligned bases). Each analysis was run with the recommended model as well as a general time-reversible (GTR) model with added parameters for invariable sites and a gamma distribution (GTR+i+g) to test the robustness of the data sets. All matrices described here were analyzed using both Garli (ver. 0.951 or 0.96; http://www.bio.utexas.edu/faculty/antisense/garli/Garli.html; Zwickl 2006) and MrBayes 3.12 (Huelsenbeck and Ronquist 2001). At least 10 runs from random starting seeds were performed to check convergence of likelihood scores and topology. ML bootstrap values were estimated on each data set using Garli with 100 random replicates. In MrBayes 3.12, we used the Metropolis-coupled Markov chain Monte Carlo method as implemented in Mr. Bayes (ver. 3.2) to run four chains (three heated). We ran 5 million generations, sampling every 1000 generations, with a burn-in of 1000 trees (1 million generations). Adequate convergence and mixing were assessed by both comparison of the consensus topologies generated from each run and by comparing pairs of individual runs in the program AWTY (Wilgenbusch et al. 2004). Likelihood scores by generation were graphed in Tracer (ver. 1.4.1; http://tree.bio.ed.ac.uk/software/tracer/) to determine the burn-in. A consensus tree was generated in MrBayes, which included branch lengths and posterior probabilities. Because of the difficulties of long branches of possible outgroups, we opted to root these trees with the RAD1 clade, based on where outgroups clustered and on divergent intron location.

Gene Expression

Plant tissue was collected from *Lonicera* × *bella*, a hybrid cultivar between *L. morrowii* and *L. tartarica*, at the corner of the Grand Central Parkway access road and Midland Parkway in Jamaica, New York. A voucher specimen is housed at YU. Large branches were collected from plants that contained flower buds at multiple stages. These branches were transported to the laboratory and immediately dissected. Dissected material was kept in a freezer at −80°C until RNA extraction could be performed.

The samples utilized for RT-PCR were generated in Howarth et al. (2011). The RNA dissections and extractions contained three separate corolla lobe samples consisting of dorsal (two dorsal lobes), lateral (two lateral lobes), and ventral (single ventral lobe). Corolla lobes were separated from corolla tubes of multiple flower buds from early-stage (first stage possible to dissect) to nearly open flower buds. Through all dissectible bud stages, floral parts were already well developed and only elongating. Bud stages were mixed because of a limited supply of buds from a species with a limited flowering period. Dissections targeted only the lobes, excluding fused corolla tube tissue. Whole-bud and leaf tissues were used for comparison.

Frozen tissues were pulverized with the Bio101 FastPrep system (Qbiogene) and subsequently extracted using an RNeasy kit with the optional DNase step (Qiagen). The samples were subsequently treated with DNase I (NEB), to further eliminate DNA contamination. Standard PCR controls were also run with no reverse transcriptase to test the total RNA for DNA contamination. The cDNA was generated with gene-specific primers using the one-step RT-PCR kit (Invitrogen) according to manufacturer instructions with a 55°C annealing temperature and between 36 and 40 cycles. We determined the linear range by running a whole-bud sample with each primer set and removing 2 mL of the reaction every four cycles from 28–40 cycles. For future reactions, we utilized the number of cycles in which a band first became clear in whole-bud tissue. Reactions amplifying *RAD*-like genes and *DipsCYC2A* and *DipsCYC2B* were in 25-μL reactions (half reactions), while GAPDH controls were in 12.5-μL reactions (quarter reactions).

Gene-specific primers were used to amplify *DipsRAD1A*, *DipsRAD1B*, *DipsRAD2A*, *DipsRAD2B*, and *DipsRAD3* (table B1). Amplification of *DipsCYC2A* and *DipsCYC2B* was performed by Howarth et al. (2011) and is included here for comparison. GAPDH control reactions used primers from Strand et al. (1997). These results were separated on a 1% agarose gel in sodium borate buffer, and copies were verified by sequencing. New sequences were submitted to GenBank (JX123689–JX123751; app. A).

Results

RAD-*Like Genes*

Multiple copies of *RAD*-like genes (up to seven) were obtained from each Dipsacales species examined. A combination of these hypothesized copies resulted in a matrix of 84 taxa, including 63 Dipsacales sequences and 21 sequences from the other eudicot species. The final matrix included 291 bases, although the majority of the taxa were only between 84 and 156 bases in length as a result of primer placement. These data can be obtained from TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S12770).

Using either TVMef+i+g or GTR+i+g models of evolution resulted in the same tree topology for the Pentapetaleae taxa (fig. 2). There were no incongruencies among the trees generated by Bayesian inference and maximum likelihood, although the latter had better resolution. We illustrate the tree from Garli 0.96 using the TVMef+i+g model, with Bayesian posterior probabilities on deep nodes (fig. 2). As anticipated, owing to the very short lengths of individual *RAD* sequences, no nodes had ML bootstrap support values of over 50%.

The phylogeny was divided into three supported clades (posterior probabilities of 0.86, 1.0, and 0.89). Each clade contained copies from *Vitis*, *Populus*, *Arabidopsis*, *Antirrhinum*, and Dipsacales, indicating that each of these gene clades characterize a much larger Pentapetaleae clade. Here we refer to these clades as RAD1, RAD2, and RAD3 (fig. 2). Each clade contained two *Arabidopsis* sequences and a single *Vitis*

Fig. 2 Maximum likelihood phylogram of Pentapetaleae *RAD*-like genes as estimated in Garli (ver. 0.96). Clades labeled 1, 2, and 3 correspond to the three *RAD*-like clades of Pentapetaleae genes, RAD1, RAD2, and RAD3, respectively. Bold branches indicate a Dipsacales clade, with black circles highlighting broad Dipsacales duplications. Numbers above the branches are Bayesian posterior probabilities. Dashed lines indicate Bayesian posterior probabilities less than 0.5. Samples isolated from cDNA are labeled.

sequence along with the sequences outlined below. RAD1 included a single *Antirrhinum* sequence, 21 Dipsacales sequences, and a single *Populus* sequence. *RADIALIS* from *Antirrhinum*, along with its known orthologs in Lamiales, was nested in RAD2 along with 37 Dipsacales sequences and a single sequence from *Populus*. RAD3 included *RAD*-like2 (RL2), RL3, and RL4 from *Antirrhinum*, five Dipsacales sequences, and two sequences from *Populus*. A sixth *Antirrhinum* sequence, RL5, is not displayed in the final tree because of its long branch and uncertain placement; however, our data indicate that it may be another copy within the RAD3 clade.

We also included non-Pentapetaleae copies from *Aquilegia* and *Oryza* to attempt to determine the root of these three Pentapetaleae clades. Through genome mining, we uncovered

three I-box-like *Aquilegia* paralogs (app. A). Yanhui et al. (2006) found nine I-box-like paralogs in *Oryza* genome searches, although several of these do not align well with the rest of the sequences. Unfortunately, because of the short sequence length of RAD and the long branches of these outgroup taxa, the placement of non-Pentapetaleae copies from *Aquilegia* and *Oryza* was not well supported. Variation in taxon sampling, the model used, and alignments led to differences in the placement of the non-Pentapetaleae groups. However, with taxon sampling including all three *Aquilegia* paralogs and only the four *Oryza* paralogs with the closest sequence similarity to *RAD*, multiple models and algorithms converged on a single topology. One copy of *Oryza* and two copies of *Aquilegia* together fall out in the same group as two of the Pentapetaleae clades, RAD2 (which includes *RADIALIS*) and RAD3. The other three *Oryza* copies and a single *Aquilegia* copy fall out with the third Pentapetaleae clade, RAD1. Taken at face value, these data suggest that the duplication separating RAD1 from RAD2 and RAD3 occurred before the diversification of monocots and eudicots. The duplication separating RAD2 from RAD3 appears to have happened around the divergence of the Pentapetaleae. However, our data do not rule out the possibility that the emergence of each of these gene clades predated the diversification of the eudicots (with each *Aquilegia* sequence falling out with each Pentapetaleae clade) or that they occurred around the diversification of the Pentapetaleae (with *Aquilegia* and *Oryza* forming a grade subtending all three Pentapetaleae clades). In any case, all analyses recovered three clades of Pentapetaleae genes: RAD1, RAD2, and RAD3.

Dipsacales Duplications

In the RAD1 clade, which possibly diverged from RAD2 and RAD3 before the diversification of the monocots and eudicots, we uncovered 21 Dipsacales sequences from 15 species (fig. 3). There appears to have been a duplication preceding the divergence of the Dipsacales. RAD1A (posterior probability = 0.69) contains Adoxaceae, Caprifolieae, Linnaeeae, Morinaceae, Dipsacaceae, and Valerianaceae, while RAD1B (posterior probability = 0.69) contains Adoxaceae, Caprifolieae, Diervilleae, Linnaeeae, and Morinaceae. We suspect that missing groups are due to lack of sampling and not gene loss. There is no evidence of duplication in RAD1A, although sampling is minimal. It is possible that there is a duplication within Morinaceae in RAD1B.

The RAD2 clade, which includes the orthologs of *AmRAD*, also yielded the greatest sampling from Dipsacales, with 37 sequences from 19 species. While this data set was insufficient to generate support for larger clades, a duplication near the divergence of the Dipsacales was clearly evident in RAD2 (fig. 4). Clade RAD2A is weakly supported (posterior probability = 0.52) and contains Adoxaceae sequences and sampling from Diervilleae, Caprifolieae, and Linnaeeae. It also does not conflict with the topology to hypothesize that one group of Morinaceae falls within this group (labeled *A?* in fig. 4). We did not amplify any Dipsacaceae or Valerianaceae from RAD2A, although this is likely due to sampling and not gene loss. The Caprifoliaceae RAD2B clade is relatively well supported (posterior probability = 0.84) and contains clades of all major groups of bilaterally symmetri-

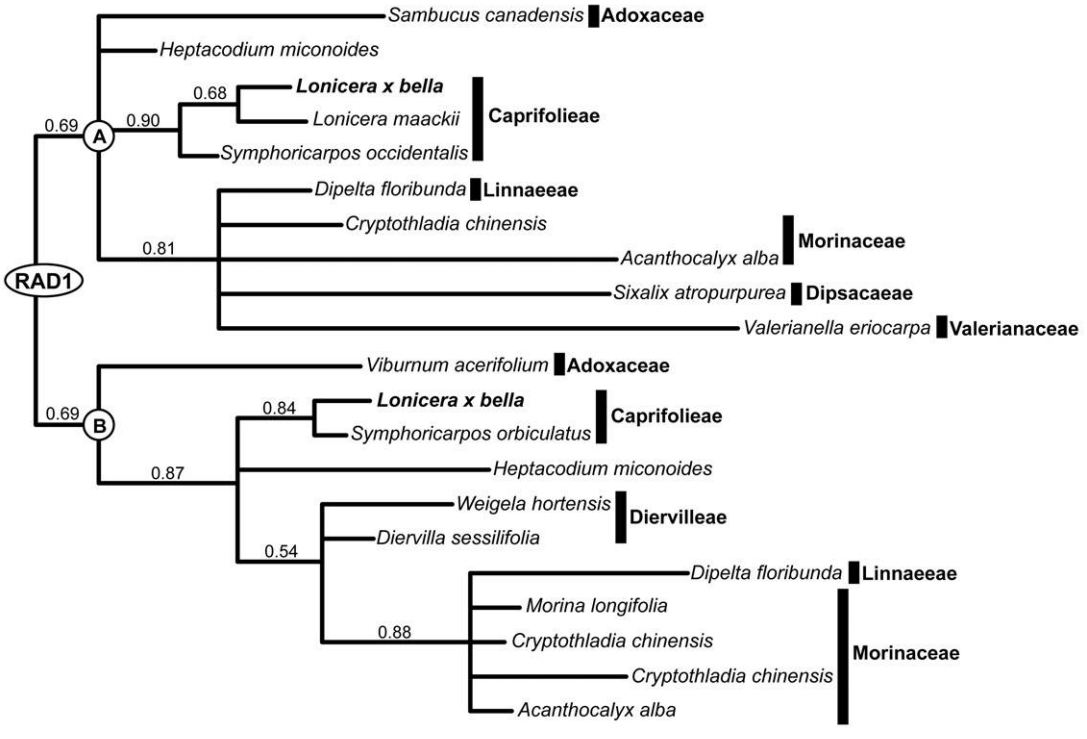

Fig. 3 Phylogram estimated by Bayesian inference of the RAD1 clade of Dipsacales sequences. The two major clades in Dipsacales are labeled *A* and *B*. The numbers above the branches are Bayesian posterior probabilities. Lineages within the Dipsacales are labeled on the right.

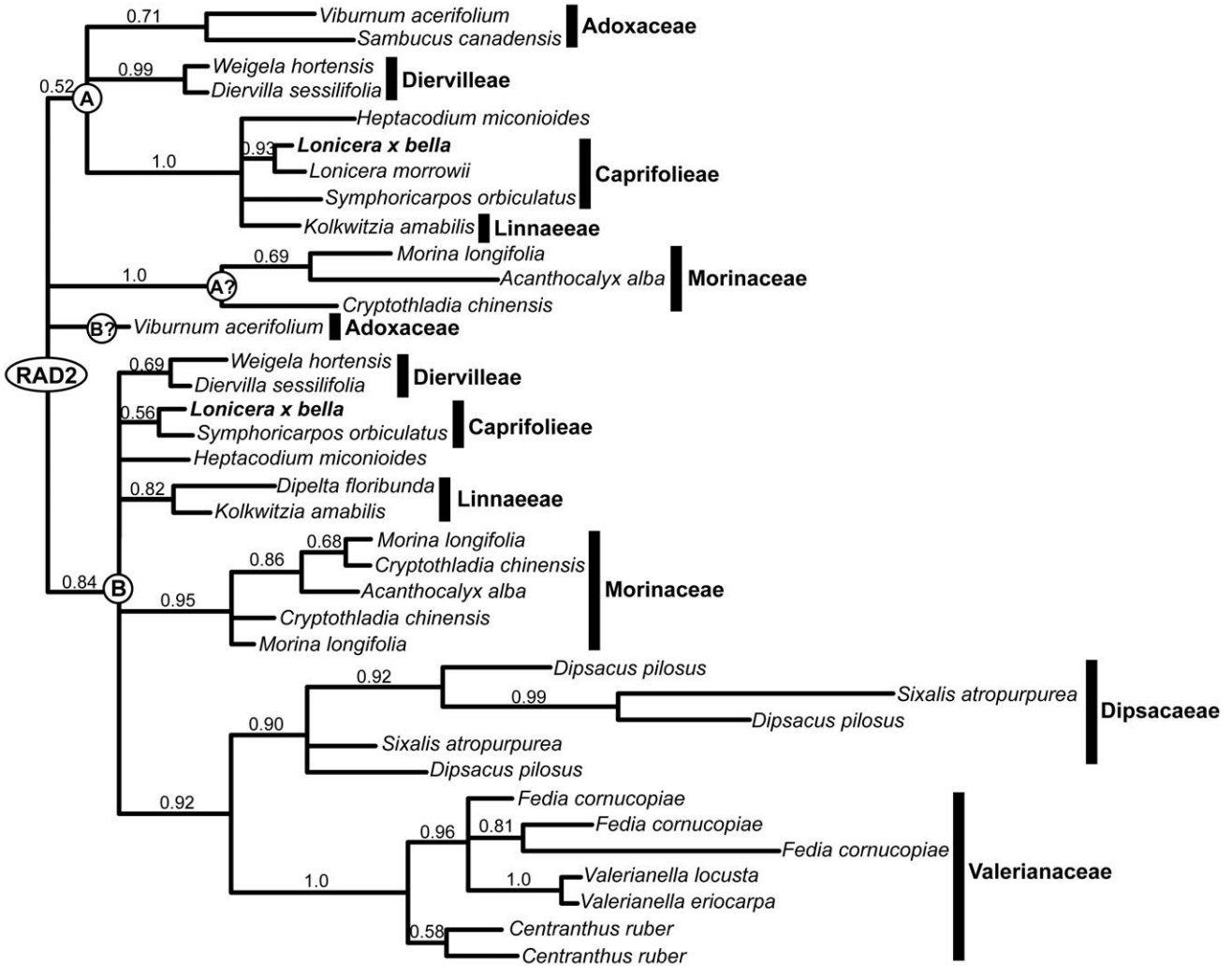

Fig. 4 Phylogram estimated by Bayesian inference of the RAD2 clade of Dipsacales sequences. The two major clades in Dipsacales are labeled *A* and *B*. Two additional clades are hypothesized to potentially be a part of these two clades (labeled *A?* and *B?*). The numbers above the branches are Bayesian posterior probabilities. Lineages within the Dipsacales are labeled on the right.

cal Dipsacales (Caprifoliaceae sensu lato; fig. 1). It is possible that *Viburnum* (Adoxaceae) also falls into this group, given that it is not closely related to the Adoxaceae clade in RAD2A (fig. 4, node labeled *B?*), in which case there was a Dipsacales-wide duplication event in this clade. However, we cannot rule out the possibility that the duplication spans only the bilaterally symmetrical Caprifoliaceae, given the lack of support for the placement of the Adoxaceae sequences. We see no evidence for subsequent duplication in any lineage in RAD2A. RAD2B, on the other hand, contains duplications in the same groups that show duplications in *CYC*-like genes, i.e., independently within Morinaceae, Dipsacaceae, and Valerianaceae (fig. 4).

Dipsacales sequences were identified within the RAD3 clade, including transcripts from Morinaceae and Caprifoliaceae (fig. 2), although we were not able to sample heavily in RAD3. As a result of the short sequence length with only a single conserved MYB domain, successful amplification of any of these gene copies was challenging. We were able to amplify RAD3 from other taxa, but because of primer placement, the sequences were less than 20 bases long and could not be used in the matrix.

RT-PCR

Gene expression was examined by amplifying each gene—*DipsRAD1A*, *DipsRAD1B*, *DipsRAD2A*, *DipsRAD2B*, and *DipsRAD3*—through RT-PCR on dissections of *Lonicera* × *bella* samples (fig. 5). Expression from *DipsCYC2A* and *DipsCYC2B* is provided for comparison along with a GAPDH control. All of the extracted tissues contained similar concentrations of RNA. Most of the *RAD*-like copies were expressed across the petal dissections, in whole buds, and in

leaves (*DipsRAD1A*, *DipsRAD1B*, *DipsRAD2A*, and *DipsRAD3*). *DipsRAD2B*, on the other hand, which is an ortholog of *AmRAD*, was lacking or very weak in the ventral petal dissection. This is similar to the expression of *DipsCYC2B* (fig. 5).

Discussion

The RAD*-Like Clade*

Here we provide a first attempt to determine the phylogenetic relationships and duplication history of *RAD*-like genes. The full sequence of *AmRAD* is only 93 amino acids long, of which approximately 59 residues make up the MYB domain (Stevenson et al. 2006). The shortness of this data set creates serious challenges for phylogenetic analyses of amplified genes, which include data only within a single conserved domain. Specifically, clade support may not be as strong as in larger data sets. The absence of maximum likelihood bootstraps over 50% clearly reflects the shortness of the sequences because more characters are generally needed to achieve high bootstrap support as compared to high posterior probabilities (Erixon et al. 2003). In a set of experiments in which we successively duplicated our data set, we recovered, as expected, dramatically increased support for the RAD1, RAD2, and RAD3 clades. For their part, posterior probabilities are more sensitive to model misspecification, underparameterization, and slight differences in taxon and character sampling (Douady et al. 2003; Erixon et al. 2003). In simulations, posterior probabilities always assigned support to an equal or greater number of correct clades compared to ML bootstraps at multiple arbitrary cutoffs (including 70% and 95%; Alfaro et al. 2003). It is comforting to note, therefore, that despite the necessarily limited amount of data, we uncovered the same general topology for the clades discussed here under all analyzed models and taxon sampling schemes.

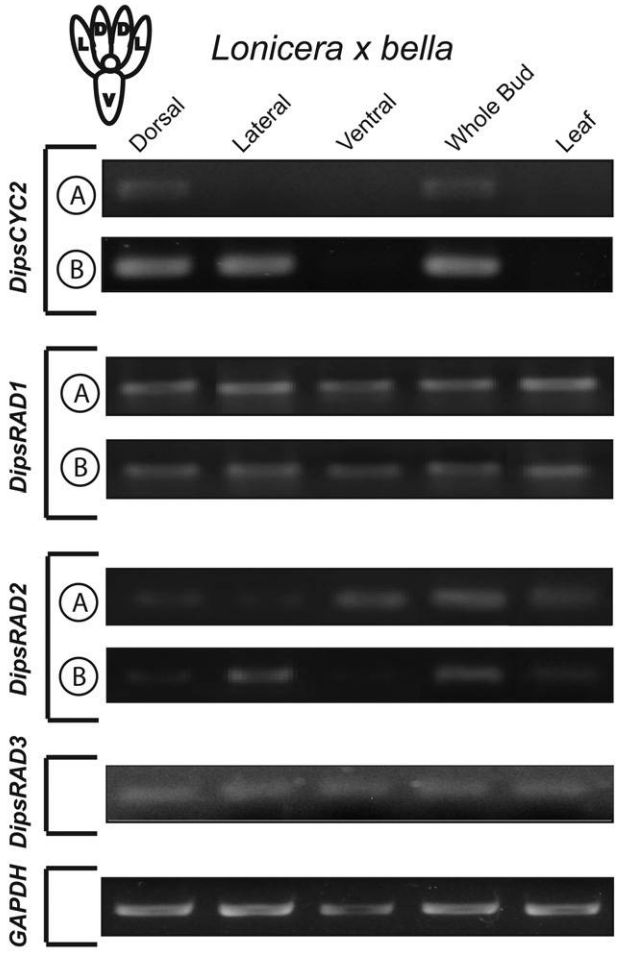

Fig. 5 Expression of *RAD*-like and *DipsCYC2* copies from *Lonicera* × *bella* using RT-PCR. Dorsal, lateral, and ventral corolla lobe dissections are displayed along with whole bud and leaf. Floral diagram indicates petal designations. GAPDH is used as an endogenous control.

Our data provide support for three distinct clades of *RAD*-like genes in the Pentapetaleae (fig. 2). These three clades, which we have named RAD1, RAD2, and RAD3, agree with the three duplets of *Arabidopsis* copies found in smaller data sets (Baxter et al. 2007; Zhou et al. 2008), with each duplet falling, separately, within each of these clades. These data indicate that there were at least two broad duplications of I-box-like genes predating the divergence of the Pentapetaleae.

Because of the short usable sequence region, support for the placement of non-Pentapetaleae taxa, and therefore the rooting of these three clades, was weak. When multiple models and taxon sampling were used, the topology found most frequently showed the general pattern of *Aquilegia* and *Oryza* copies clustering with RAD1 and with RAD2 and RAD3 together, even though posterior probabilities and maximum likelihood bootstrap support values were <0.5 and 50%, respectively. Better sampling will be needed to confidently resolve these broad relationships. The hypothesis that RAD2 and RAD3 are sister clades is also supported by the intron placement in *Arabidopsis* and *Vitis*, where the copies from RAD2 and RAD3 all have an intron in the 3′ end of their coding sequences, differing from the copies from RAD1, where the intron is in the 3′ untranslated region (Baxter et al. 2007). *Antirrhinum* appears to also follow this pattern in RAD2 and RAD3 genes; however, *AmRAD1* (from RAD1) is incomplete, and therefore intron placement is unknown. Intron placement in *Aquilegia* and *Oryza* may be more labile, although several gene sequences are incomplete. Given these data, we hypothesize that the duplication separating RAD1 from RAD2 and RAD3 predated the monocot-eudicot split. The duplication leading to RAD2 and RAD3 appears to have occurred around the divergence of the Pentapetaleae.

Our findings on RAD evolution provide a striking complement to recent studies of floral symmetry and floral organ identity (MADS-box) genes. Proteins within the TCP and MYB families pattern the dorsoventral identity of *Antirrhinum* flowers (Luo et al. 1996; Almeida et al. 1997; Corley et al. 2005; Costa et al. 2005), whereas MADS-box proteins pattern floral organ identity (Coen and Meyerowitz 1991; Meyerowitz et al. 1991). Both the ECE clade of *CYC*-like TCP genes and the core R-R-type clade of *DIV*-like genes apparently duplicated twice before the diversification of the

Pentapetaleae (Howarth and Donoghue 2006, 2009). 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 Pentapetaleae (Litt and Irish 2003; Kim et al. 2004; Kramer et al. 2004; Kramer and Hall 2005; Zahn et al. 2005). More specifically, *AP1* and *SEP* both duplicated twice around the diversification of the Pentapetaleae, while *AP3* and *AG* appear to have duplicated only once (or lost a third copy) around the same time. Additionally, these gene networks may cross since there is evidence that a *RAD* ortholog in *Gossypium* may be controlled in part by AG (Zhang et al. 2011). Our analyses of *RAD*-like genes lead us to hypothesize that these genes duplicated a single time preceding the Pentapetaleae diversification, much like *AP3* and *AG* (Kim et al. 2004; Kramer et al. 2004), as opposed to twice as in the remaining MADS-box gene clades and the other two floral symmetry gene clades, *CYC* and *DIV* (Howarth and Donoghue 2006, 2009). The duplications found among these groups combine to place the potential hexaploidization event hypothesized by Jaillon et al. (2007) between the diversification of the Ranunculales and the rest of the eudicots (Bowers et al. 2003; De Bodt et al. 2005; Jaillon et al. 2007).

Dipsacales Duplications

Owing to the short sequence lengths of the *RAD*-like genes, we did not expect to resolve relationships among groups within the Dipsacales to the same degree that we previously demonstrated with *CYC*-like and *DIV*-like genes (Howarth and Donoghue 2005, 2006, 2009). Nevertheless, we were able to pinpoint major duplication events in the Dipsacales based on clades with >0.5 Bayesian posterior probabilities and the number of copies obtained (figs. 3, 4).

The RAD1 clade, which appears to span monocots + eudicots, had reasonable sampling from Dipsacales and showed clear evidence of a broad duplication event predating the divergence of the Dipsacales (fig. 3). DipsRAD1A contained sampling from Adoxaceae and across Caprifoliaceae sensu lato. DipsCYC1B also contained sampling from Adoxaceae and several groups of Caprifoliaceae but was missing the Dipsacaceae and Valerianaceae. It is likely that these were simply unsampled. The only other possible duplication in Dipsacales is within DipsRAD1B, where there is a possible duplication in Morinaceae (fig. 3), although this could also be allelic variation in *Cryptothladia*.

RAD2, which contains the orthologs of *AmRAD*, was the copy most frequently cloned and sequenced. Based on primer sequence, there is no reason to hypothesize that this copy would be preferentially amplified over copies from RAD1 and RAD3. Nevertheless, we uncovered the greatest number of duplications in RAD2 in Dipsacales (fig. 4). This is similar to other floral symmetry genes where we also found the greatest numbers of duplications within the CYC2 clade (orthologs of *AmCYC*) and the DIV1 clade (orthologs of *AmDIV*). Our interpretation of the RAD2 Dipsacales data set entails a duplication predating the divergence of the Dipsacales (fig. 4). There are two independent groups of radially symmetrical Adoxaceae and bilaterally symmetrical Caprifoliaceae species, and our data do not conflict with these forming two clades.

The *DipsRAD2A* clade, with posterior probability of 0.69, contains Adoxaceae and three lineages of Caprifoliaceae (Diervilleae, Caprifolieae, and Linnaeeae). We also hypothesize that the separate clade of Morinaceae labeled *A?* in figure 4 may fall into this group, given its distinctness from the other clade of Morinaceae, which falls into *DipsRAD2B*. We assume that the apparent absence from Dipsacaceae and Valerianaceae is simply due to our failure to sample these genes, although it is possible that *DipsRAD2A* has been lost in this group of Dipsacales. There is no evidence for additional duplications within any Dipsacales lineages within *DipsRAD2A*. A second clade of *DipsRAD2* genes, *DipsRAD2B*, is supported by a posterior probability of 0.84 and includes a better sample from across Caprifoliaceae. It does not conflict with the topology for the second copy from *Viburnum acerifolium* (Adoxaceae) to be sister to this DipsRAD2B Caprifoliaceae clade. This would be similar to the patterns seen in all three clades of *DIV*-like genes (Howarth and Donoghue 2009). It is also possible, however, that this copy falls into the DipsRAD2A clade and that the duplication in RAD2 coincided with the shift to bilateral symmetry around the divergence of the Caprifoliaceae, similar to the patterns seen in the DipsCYC2 and DipsCYC3 clades (Howarth and Donoghue 2005). Within the Caprifoliaceae DipsRAD2B clade, we also see clear evidence for a duplication within the Morinaceae and at least one duplication in the Dipsacaceae and Valerianaceae clades. These three groups have also contained duplications in multiple clades of *CYC*-like and *DIV*-like genes (Howarth and Donoghue 2005, 2009).

DipsRAD3 was difficult to amplify from Dipsacales with only sequences from Caprifolieae (*Symphoricarpos*) and Morinaceae obtained. The presence of these two groups indicates, however, that there is a DipsRAD3 clade of Dipsacales genes. Through specific primer design we were able to amplify other gene copies that we hypothesize to be DipsRAD3 copies; however, because of their short sequence (fewer than 20 bases without the flanking primer sequences), they could not be included in the matrix. Future work could include obtaining full-length gene sequence for these copies using randomly amplified cDNA ends, which would allow us to design more specific primers outside of the conserved MYB domain.

Both *Antirrhinum* and *Bournea* appear to have only a single copy in the RAD2 clade, although *Antirrhinum* has at least three RAD3 copies. Duplications may not be as frequent in RAD2 as in CYC2 and DIV1 (Galego and Almeida 2002; Reeves and Olmstead 2003; Howarth and Donoghue 2005, 2006, 2009; Zhou et al. 2008; Jabbour et al. 2009), but within the Dipsacales these duplications are still common. We show evidence in RAD2 for a broad Dipsacales duplication as well as subsequent duplications in Morinaceae, Dipsacaceae, and Valerianaceae.

It is clear that the lineages Morinaceae, Valerianaceae, and Dipsacaceae contain common duplications within each of the *CYC* (Howarth and Donoghue 2005), *DIV*, and *RAD* gene families. These are all groups with divergent floral morphologies including such characteristics in different groups as a bilaterally symmetrical calyx, an epicalyx (sometimes also bilateral), an asymmetric corolla, ventral and asymmetric sta-

men abortion, and multiple flower symmetries within a single head inflorescence. Morinaceae, for instance, are characterized by an unusual bilateral calyx, which appears from developmental studies to have 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 the potentially interacting *RAD*-like, *DIV*-like, and *CYC*-like genes are related to the new function of producing a bilaterally symmetrical calyx.

Expression of DipsRAD

In *Antirrhinum*, the only species in which endogenous RAD function has been studied, Corley et al. (2005) hypothesize that CYC and its immediate sister copy DICH (both in the CYC2 clade; Howarth and Donoghue 2006) together activate *RAD*. RAD, in turn, may downregulate or compete with DIV, given that RAD has a single MYB domain resembling a truncated portion of the 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, RAD, and DIV interact in a regulatory network to pattern floral symmetry in *Antirrhinum*, although the exact nature of the interactions, and even whether this pathway is utilized outside of *Antirrhinum*, remains unknown. Expression data from *Bournea*, *Gratiola*, and *Veronica* suggest that it may be recruited at least across the Lamiales (Zhou et al. 2008; Preston et al. 2009). Costa et al. (2005) 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. In fact, in contrast, expression data from *Arabidopsis* indicate that the *RAD*-like genes may not function in the same way in that species (Costa et al. 2005; Baxter et al. 2007) because there is no expression in the corolla (although the location of expression was not found for one of the RAD2 clade genes, *AtRL1*). A recent study in *Gossypium* (also a radially symmetrical rosid), examining *GbRL1*, which falls into the RAD2 clade (data not shown), indicates that this gene is expressed strongly in corolla and ovule tissues, with potential control of cotton fiber growth, providing evidence for a novel function compared to *Antirrhinum* (Zhang et al. 2011). This pattern could mean that the pathway of interactions between RAD, DIV, and CYC may be specific to Lamiidae or Asteridae. It could also mean, however, that *Arabidopsis*, with four-merous radially symmetrical flowers at maturity, utilizes a derived version of a conserved regulatory network that otherwise does span the Pentapetaleae.

Although RAD function has not yet been corroborated outside *Antirrhinum*, data do indicate that a downstream target of the floral symmetry MYB gene, DIV, has broader function. Perez-Rodriguez et al. (2005) showed that *AmMYBL1*, an *Antirrhinum* gene, is a downstream target of DIV and its ectopic expression in *Nicotiana* affected ventral petal morphology. This suggests the possibility that DIV function may also extend at least across the Lamiidae. Additionally, expression of *DIV* in Gesneriaceae and Dipsacales suggests broader conservation in the pathway (Zhou et al. 2008; Howarth and Donoghue 2009).

Here we present the first expression data supporting the conservation of this floral symmetry regulatory network outside of Lamiidae. We examined all five Dipsacales *RAD*-like genes (*DipsRAD1A*, *DipsRAD1B*, *DipsRAD2A*, *DipsRAD2B*, and *DipsRAD3*) in *Lonicera* × *bella* via RT-PCR. *Lonicera* flowers are strongly bilaterally symmetrical, but, unlike *Antirrhinum*, the two dorsal and the two lateral petals are "up," or dorsally oriented, while the single ventral petal remains "down," or is ventrally oriented (fig. 5). In *Antirrhinum*, evidence indicates that *RAD* is activated by CYC + DICH and therefore expressed in the same zone (Corley et al. 2005). While both *CYC* and *DICH* are expressed in the dorsal petals and staminode, *DICH* is more restricted in expression to the dorsal portion of the dorsal petals (Luo et al. 1996, 1999). *RAD* expression is more similar to the expression of *CYC* than *DICH* (Corley et al. 2005). Therefore, given a hypothesis of similar gene interaction of the floral symmetry pathway in other Pentapetaleae, we would expect a RAD2 clade homologue to express in a pattern similar to that of CYC2 clade genes. In *Lonicera*, *DipsCYC2A* is expressed in only the dorsal two petals, while *DipsCYC2B* has an expanded expression into all four of the petals that are oriented dorsally (fig. 5; Howarth et al. 2011). In *Antirrhinum*, *RAD* is expressed in the broadest zone across which CYC2 members are expressed. Therefore, in *Lonicera*, we would expect to find a *RAD* homologue expressed across the most expanded expression of a CYC2 homologue. In fact, this is what we have discovered in *DipsRAD2B*, which is expressed in the same zone that *DipsCYC2B* is expressed. Although we have examined only crude dissections of whole petals, *DipsRAD2B* appears to be lacking from the ventral petal, which is similar to the expanded zone of expression of *DipsCYC2B* as opposed to the more restricted expression of *DipsCYC2A* (fig. 5; Howarth et al. 2011).

Within the RAD2 clade, expression has been examined only in *Antirrhinum* (and other Veronicaceae), *Bournea*, *Arabidopsis*, and now *Lonicera* (Corley et al. 2005; Baxter et al. 2007; Zhou et al. 2008; Preston et al. 2009). In Veronicaceae, *Bournea*, and *Lonicera*, a RAD2 homologue is expressed in a pattern similar to that of a CYC2 homologue (Corley et al. 2005; Zhou et al. 2008; Preston et al. 2009). In *Arabidopsis*, there are two copies, *AtRL1* and *AtRL2*. Using in situ hybridization, Baxter et al. (2007) did not find *AtRL2* in the corolla but instead in portions of the ovules and embryos, although perhaps in a dorsoventral pattern within the funiculus. They were not able to detect *AtRL1* using in situ hybridization, although it was uncovered from whole-plant RT-PCR, so it is still unknown where it is expressed (Baxter et al. 2007). In *Lonicera*, while *DipsRAD2B* is expressed similarly to *DipsCYC2B* as described above, *DipsRAD2A* is expressed across the corolla and in whole buds and leaves and does not appear to exhibit a dorsoventral expression pattern.

In the RAD3 clade, the expression is less known. There are three clear copies of AmRAD2 genes, *AmRL2*, *AmRL3*, and *AmRL4*. Expression data are available only for *AmRL3*, which is in the central pith of inflorescences and the outer layers of the stem, although also possibly with a dorsoventral pattern within bracts (Baxter et al. 2007). In *Arabidopsis*, there are two copies in RAD3, *AtRL5* and *AtRL6*. In situ hybridization did not uncover expression for *AtRL5*, although

whole-plant RT-PCR did, indicating that it is expressed (Baxter et al. 2007). *AtRL6* is not expressed in flowers and appears to be in micropylar endosperm (Boisnard-Lorig et al. 2001). We uncovered only a single copy in Dipsacales from RAD3, which appears to be weakly expressed across flowers and leaves (fig. 5).

The RAD1 clade, which we hypothesize predates the divergence of monocots from eudicots, also has little published expression data. In *Antirrhinum*, *AmRL1* clearly falls in this clade. *AmRL1* appears to be weakly expressed in young floral meristems and around the pedicel. In *Arabidopsis*, there are two copies, *AtRL3* and *AtRL4*, which are expressed just outside vascular bundles and not in flowers (Baxter et al. 2007). In Dipsacales, we show evidence for two independent gene copies within RAD1. In *Lonicera*, both of these copies appear to be expressed across floral and leaf tissue (fig. 5).

There is considerable diversity of expression in the *RAD*-like gene copies, especially when comparing *Lonicera*, *Antirrhinum*, and *Arabidopsis*. *Lonicera* may seem to have broader expression simply because these are RT-PCR results that could pick up much weaker signals than in situ hybridization. Several of the genes do appear to have a dorsoventral pattern of expression in whatever tissue they are expressed in, and in all of these cases they are restricted adaxially, implying that these genes could frequently be co-opted in patterning symmetry. We argue from our circumstantial expression data from *Lonicera* that genes from CYC2, DIV1, and RAD2 possibly form a network of interacting gene partners that are conserved at least across asterids. These interacting gene partners appear to have duplicated in tandem around the diversification of the Pentapetaleae as well as within major clades of Dipsacales. This more complete picture of duplications within each of these three interacting gene lineages sets the stage for an analysis of the coevolution of these genes in the several distantly related lineages in which bilateral flower symmetry originated.

Acknowledgments

We thank the following 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; noneudicot 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. Donoghue laboratory. We thank E. Kramer for helpful discussion on the manuscript. We thank M. Dunn and members of the D. G. Howarth and M. J. Donoghue laboratory for helpful discussions. The work of D. G. Howarth on floral symmetry gene evolution was supported by a National Science Foundation grant (1121301). The work of M. J. Donoghue on angiosperm phylogeny was supported by an Assembling the Tree of Life (ATOL) grant from the National Science Foundation.

Appendix A

Voucher Information

Further duplicated genes in each group are separated by a colon. Taxon; voucher collection and (herbarium); GenBank accession numbers.

Dipsacales; RAD1A, RAD1B, RAD2A, RAD2B, RAD3. *Acanthocalyx alba*; Boufford et al. 28401 (A); JX123730, JX123729, JX123727, JX123728, JX123721. *Centranthus ruber*, Bell 203 (YU); –, –, –, JX123699:JX123692, –. *Cryptothladia chinensis*; Bufford, Donoghue and Ree 27870 (A); JX123736, JX123734:JX123735, JX123732, JX123731:JX123733, JX123720. *Diervilla sessilifolia*; Elisk and Zinman 3286 (A); –, JX123746, JX123745, JX123693, –. *Dipelta floribunda*; Buckland and Kelly 32 (A), JX123743, JX123744, –, JX123742, –. *Dipsacus pilosus*; F. Billiet 1445, Nat. Bot. Gard. of Belgium; –, –, –, JX123714:JX123715:JX123716, –. *Fedia cornucopiae*; Bell, voucher lacking cDNA; –, –, –, JX123689:JX123691:JX123690, –. *Heptacodium miconioides*; Howarth, live specimen Marsh Gardens (YU), cDNA; JX123695, JX123698, JX123697, JX123696, –. *Kolkwitzia amabilis*; Elsik, Michener, and Bailey 844 (A); –, –, JX123707, JX123706, –. *Lonicera* × *bella*; Howarth 1–2010 (YU); JX123719, JX123741, JX123709, JX123694, JX123747. *Lonicera maackii*; Smith 20 (YU); JX123705, –, –, –, –. *Lonicera morrowii*; Smith 49 (YU); –, –, JX123708, –, –. *Morina longifolia*; Eriksson s.n. 2 Nov. 1999 (SBT); –, JX123726, JX123723, JX123722:JX123724, JX123725. *Sambucus canadensis*; Donoghue, voucher lacking; JX123700, –, JX123704, –, –. *Sixalix atropurpurea*; Carlson 144 (YU); JX123712, –, –, JX123713:JX123711, –. *Symphoricarpos occidentalis*; Donoghue, voucher lacking; JX123710, –, –, –, –. *Valerianella eriocarpa*; Bell, voucher lacking, cDNA; JX123701, –, –, JX123703, –. *Valerianella locusta*; Bell 2006–64 (YU); –, –, –, JX123702, –. *Viburnum acerifolium*; Winkworth and Donoghue 27 (A and YU); –, JX123749, JX123748, JX123717, –. *Weigela hortensis*; Kelly and Buckland 28 (A); –, JX123751, JX123750, JX123718, –. **Others;** RAD1, RAD2, RAD3. *Antirrhinum majus*; Baxter et al., 2007; AJ791699 (AmRL1), RADIALIS, DQ375230 (AmRL2):DQ375227 (AmRL3):DQ375228 (AmRL4). *Aquilegia coerulea*; EST database; FTOX81088.g1, RAD2/3-Scaffold 10.3357828.3358083:FTOX54583.b1. *Arabidopsis thaliana*; Baxter et al., 2007; DQ395345 (AtRL4):At4g36570 (AtRL3), At2g21650 (AtRL2):At4g39250 (AtRL1), At1g19510 (AtRL5):At1g75250 (AtRL6). *Bournea leiophylla*; Zhou et al., 2008, –, EF207557, –. *Gratiola officinalis*; Preston et al., 2009; –, FJ649696, –. *Oryza sativa*; Genome database; 0s03g14810:0s03g63890:0s12g33950, RAD2/3-Os07g26150. *Populus trichocarpa*; Genome database; gw1.57.329.1, fgenesh4_pm.C_LG_IV000458, gw1.V.1367.1:eugene3.00020323. *Veronica serpyllifolia*; Preston et al., 2009, –, FJ649695, –. *Vitis vinifera*; Genome database; GSVIVT01018944001, GSVIVT01031975001, GSVIVT01031982001.

Literature Cited

Akaike H 1973 Information theory as an extension of the maximum likelihood principle. *In* BN Petrov, F Csaki, eds. Second International Symposium on Information Theory. Akademiai Kiado, Budapest.

Alfaro ME, S Zoller, F Lutzoni 2003 Bayes or bootstrap? a simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol Biol Evol 20:255–266.

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

Backlund A, MJ Donoghue 1996 Morphology and phylogeny of the order Dipsacales. PhD diss. Department of Systematic Botany, Uppsala University.

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

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

Boisnard-Lorig C, A Colon-Carmona, M Bauch, S Hodge, P Doerner, E Bancharel, C Dumas, J Haseloff, F Berger 2001 Dynamic analyses of the expression of the HISTONE::YFP fusion protein in *Arabidopsis* show that syncytial endosperm is divided in mitotic domains. Plant Cell 13:495–509.

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

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

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

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

Cenci A, M-C Combes, P Lashermes 2010 Comparative sequence analyses indicate that *Coffea* (Asterids) and *Vitis* (Rosids) derive from the same paleo-hexaploid ancestral genome. Mol Genet Genomics 283:493–501.

Chan AP, J Crabtree, Q Zhao, H Lorenzi, J Orvis, D Puiu, A Melake-Berhan, et al 2010 Draft genome sequence of the oilseed species *Ricinus communis*. Nat Biotechnol 28:951–959.

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

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

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

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

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

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

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

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

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

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

Donoghue MJ, RH Ree 2000 Homoplasy, and developmental constraint: a model and an example from plants. Am Zool 40:759–769.

Donoghue MJ, RH Ree, DA Baum 1998 Phylogeny and the evolution of flower symmetry in the Asteridae. Trends Plant Sci 3: 311–317.

Douady CJ, F Delsuc, Y Boucher, WF Doolittle, EJP Douzery 2003 Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. Mol Biol Evol 20:248–254.

Endress PK 1996 Diversity and evolutionary biology of tropical flowers. Cambridge University Press, Cambridge.

——— 1999 Symmetry in flowers: diversity and evolution. Int J Plant Sci 160(suppl):S3–S23.

——— 2010 Flower structure and trends of evolution in eudicots and their major subclades. Ann Mo Bot Gard 97:541–583.

——— 2011 Evolutionary diversification of the flowers in angiosperms. Am J Bot 98:370–396.

Erixon P, B Svennblad, T Britton, B Oxelman 2003 Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst Biol 52:665–673.

Feng X, Z Zhao, Z Tian, S Xu, Y Luo, Z Cai, Y Wang, et al 2006 Control of petal shape and floral zygomorphy in *Lotus japonicus*. Proc Natl Acad Sci USA 103:4970–4975.

Flagel LE, JF Wendel 2009 Genome duplication and evolutionary novelty in plants. New Phytol 183:557–564.

Fukuoka N 1972 Taxonomic study of Caprifoliaceae. Mem Fac Sci Kyoto Univ Ser Biol 6:15–58.

Galego L, J Almeida 2002 Role of *DIVARICATA* in the control of dorsoventral asymmetry in *Antirrhinum* flowers. Genes Dev 16: 880–891.

Hileman LC, DA Baum 2003 Why do paralogs persist? molecular evolution of *CYCLOIDEA* and related floral symmetry genes in Antirrhineae (Veroniacaceae). Mol Biol Evol 20:591–600.

Hofmann U, J Göttmann 1990 *Morina* L. and *Triplostegia* Wall. ex DC. im Vergleich mit Valerianaceae and Dipsacaceae. Bot Jahrb Syst 111:499–553.

Howarth DG, MJ Donoghue 2005 Duplications in *CYC*-like genes from Dipsacales correlate with floral form. Int J Plant Sci 166:357–370.

——— 2006 Phylogenetic analysis of the "ECE" (CYC/TB1) clade reveals duplications predating the core eudicots. Proc Natl Acad Sci USA 103:9101–9106.

——— 2009 Duplications and expression of *DIVARICATA*-like genes in Dipsacales. Mol Biol Evol 26:1245–1258.

Howarth DG, T Martins, E Chimney, MJ Donoghue 2011 Diversification of *CYCLOIDEA* expression in the evolution of bilateral flower symmetry in Caprifoliaceae and *Lonicera* (Dipsacales). Ann Bot 107:1521–1532.

Huelsenbeck JP, F Ronquist 2001 MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Jabbour F, S Nadot, C Damerval 2009 Evolution of floral symmetry: a state of the art. C R Biol 332:219–231.

Jacobs B, K Geuten, N Pyck, S Huysmans, S Jansen, E Smets 2011 Unraveling the phylogeny of *Heptacodium* and *Zabelia* (Caprifoliaceae): an interdisciplinary approach. Syst Bot 36:231–252.

Jaillon O, J-M Aury, B Noel, A Policriti, C Clepet, A Casagrande, N

Choisne, et al 2007 The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–468.

Judd WS, RW Sanders, MJ Donoghue 1994 Angiosperm family pairs—preliminary phylogenetic analyses. Harv Pap Bot 5:1–51.

Kim M, M-L Cui, P Cubas, A Gillies, K Lee, MA Chapman, RJ Abbott, E Coen 2008 Regulatory genes control a key morphological and ecological trait transferred between species. Science 322: 1116–1119.

Kim S, M-J Yoo, VA Albert, JS Farris, PS Soltis, DE Soltis 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.

Kramer EM, JC Hall 2005 Evolutionary dynamics of genes controlling floral development. Curr Opin Plant Biol 8:13–18.

Kramer EM, MA Jaramillo, VS Di Stilio 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.

Litt A, VF Irish 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, R Carpenter, L Copsey, C Vincent, J Clark, E Coen 1999 Control of organ asymmetry in flowers of *Antirrhinum*. Cell 99:367–376.

Luo D, R Carpenter, C Vincent, L Copsey, E Coen 1996 Origin of floral asymmetry in *Antirrhinum*. Nature 383:794–799.

Maddison DR, WP Maddison 2003 MacClade: analysis of phylogeny and character evolution. Sinauer, Sunderland, MA.

Magallon S, PR Crane, PS Herendeen 1999 Phylogenetic pattern, diversity, and diversification of eudicots. Ann Mo Bot Gard 86:297–372.

Martin C, J Paz-Ares 1997 MYB transcription factors in plants. Trends Genet 13:67–73.

Meyerowitz EM, JL Bowman, LL Brockman, GN Drews, T Jack, LE Sieburth, D Weigel 1991 A genetic and molecular model for floral development in *Arabidopsis thaliana*. Development 113(suppl): 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.

Paabo S, DM Irwin, AC Wilson 1990 DNA damage promotes jumping between templates during enzymatic amplification. J Biol Chem 265:4718–4721.

Perez-Rodriguez M, FW Jaffe, E Butelli, BJ Glover, C Martin 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, KA Crandall 1998 MODELTEST: testing the model of DNA substitutions. Bioinformatics 14:817–818.

Preston JC, LC Hileman 2009 Developmental genetics of floral symmetry evolution. Trends Plant Sci 14:147–154.

Preston JC, MA Kost, LC Hileman 2009 Conservation and diversification of the symmetry developmental program among close relatives of snapdragon with divergent floral morphologies. New Phytol 182:751–762.

Pyck N 2001 Phylogenetic relationships within Dipsacales: a combined molecular and morphological approach. PhD diss. Katholieke University, Leuven, Belgium.

Reeves PA, RG Olmstead 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.

Riechmann JL, OJ Ratcliffe 2000 A genomic perspective on plant transcription factors. Curr Opin Plant Biol 3:423–434.

Sargent RD 2004 Floral symmetry affects speciation rates in angiosperms. Proc R Soc B 271:603–608.

Schmutz J, SB Cannon, J Schlueter, J Ma, T Mitros, W Nelson, DL Hyten, et al 2010 Genome sequence of the palaeopolyploid soybean. Nature 463:178–183.

Soltis DE, AE Senters, MJ Zanis, S Kim, JD Thompson, PS Soltis, LP Ronse De Craene, PK Endress, JS Farris 2003 Gunnerales are sister to other core eudicots: implications for the evolution of pentamery. Am J Bot 90:461–470.

Stevenson CEM, N Burton, MMR Costa, U Nath, RA Dixon, ES Coen, DM Lawson 2006 Crystal structure of the MYB domain of the RAD transcription factor from *Antirrhinum majus*. Proteins Struct Funct Bioinform 65:1041–1045.

Strand AE, J Leebens-Mack, BG Milligan 1997 Nuclear DNA-based markers for plant evolutionary biology. Mol Ecol 6:113–118.

Tang H, X Wang, JE Bowers, R Ming, M Alam, AH Paterson 2008 Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. Genome Res 18:1944–1954.

Tank DC, MJ Donoghue 2010 Phylogeny and phylogenetic nomenclature of the Campanulidae based on an expanded sample of genes and taxa. Syst Bot 35:425–441.

Van de Peer Y, S Maere, A Meyer 2009 The evolutionary significance of ancient genome duplications. Nat Rev Genet 10:724–732.

Wang Z, Y Luo, X Li, L Wang, S Xu, J Yang, L Weng, et al 2008 Genetic control of floral zygomorphy in pea (*Pisum sativum* L.). Proc Natl Acad Sci USA 105:10414–10419.

Weberling F 1989 Morphology of flowers and inflorescences. Cambridge University Press, Cambridge.

Wilgenbusch JC, DL Warren, DL Swofford 2004 AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.

Winkworth RC, J Lundberg, MJ Donoghue 2008 Toward a resolution of Campanulid phylogeny, with special reference to the placement of Dipsacales. Taxon 57:53–65.

Yanhui C, Y Xiaoyuan, H Kun, L Meihua, L Jigang, G Zhaofeng, L Zhiqiang, et al 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, H Kong, JH Leebens-Mack, S Kim, PS Soltis, LL Landherr, DE Soltis, CW dePamphilis, H Ma 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 F, X Liu, K Zuo, J Zhang, X Sun, K Tang 2011 Molecular cloning and characterization of a novel *Gossypium barbadense* L. *RAD*-like gene. Plant Mol Biol Rep 29:324–333.

Zhang W, EM Kramer, CC Davis 2010 Floral symmetry genes and the origin and maintenance of zygomorphy in a plant pollinator mutualism. Proc Natl Acad Sci USA 107:6388–6393.

Zhang WH, JD Chen, J Li, HB Chen, YC Tang 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, Y-Z Wang, JF Smith, R Chen 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.

Zwickl D 2006 Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD diss. University of Texas, Austin.