Int. J. Plant Sci. 166(3):357–370. 2005.

1058-5893/2005/16603-0001$15.00

DUPLICATIONS IN *CYC*-LIKE GENES FROM DIPSACALES CORRELATE WITH FLORAL FORM

Dianella G. Howarth[1] and Michael J. Donoghue

Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06520-8106, U.S.A.

Flower symmetry is of special interest in understanding the evolution and ecology of angiosperms. Evidence from the Antirrhineae (within the lamiid clade of Asteridae) indicates that several TCP gene family transcription factors, especially *CYCLOIDEA* (*CYC*), play a role in specifying dorsal identity in the corolla and androecium of monosymmetric (bilateral) flowers. We examine the evolution of this gene family in Dipsacales, representing the campanulid clade of Asteridae, in relation to evolutionary shifts in floral symmetry and stamen abortion. We identify three major forms of *CYC*-like genes in Dipsacales. We identify the position of additional gene duplications by comparing each of the three gene trees to a well-supported Dipsacales phylogeny. We infer duplications in two of the major gene lineages along the line leading to the Caprifoliaceae correlated with the origin of monosymmetric flowers. There are no duplications or losses associated with major shifts in stamen number. However, there are several additional duplications within the Caprifoliaceae, especially in the Morinaceae, possibly related to calyx monosymmetry and/or stamen reduction. Within the protein-coding sequences, we identify a new conserved region—the ECE region—that appears to be present across known angiosperm sequences. There are major changes in length and in the presence or absence of the ECE region in the Dipsacales, indicating changes in gene function. These studies expand our understanding of the evolution of *CYC*-like genes in angiosperms and provide a new system for studying the role of this gene family in determining floral form.

Keywords: Dipsacales, *CYCLOIDEA*, TCP, gene duplications, floral symmetry.

Introduction

Shifts in the symmetry of flowers between polysymmetry (actinomorphy, radial symmetry) and monosymmetry (zygomorphy, bilateral symmetry) have been common within angiosperms (Weberling 1989; Endress 1996, 1999). Much of the discussion of such shifts has focused on the sympetalous Asteridae (Donoghue et al. 1998; Ree and Donoghue 1999; Cubas 2002; Gillies et al. 2002; Knapp 2002; Reeves and Olmstead 2003), with ca. 65,000 species, about half of which are monosymmetric and half polysymmetric (Ree and Donoghue 1999). Donoghue et al. (1998) and Ree and Donoghue (1999) inferred that the ancestral asterid was probably polysymmetric and that within this clade there were at least eight major gains of monosymmetry and at least nine reversals to polysymmetry. Such morphological shifts are of special interest in relation to shifts in pollination (Neal 1998) and perhaps, ultimately, in rates of speciation (Sargent 2004). It has also been suggested that the various forms of monosymmetry found among asterids (Donoghue et al. 1998) might reflect underlying developmental constraints in this lineage, especially related to the basic orientation of the flower, that result 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 two potential candidate genes, *CYCLOIDEA* (Luo et al. 1996) and *DIVARICATA* (Galego and Almeida 2002).

CYCLOIDEA (*CYC*) has been well characterized in *Antirrhinum majus* (snapdragon), where it has been shown to be involved in specifying dorsal or adaxial flower identity (Luo et al. 1996, 1999; Almeida et al. 1997). *CYC* is a member of the TCP gene family, named for *TEOSINTE BRANCHED1* (*TB1*) in *Zea mays*, *CYC* in *A. majus*, and the proliferating cell factor (PCF) DNA-binding proteins in rice. All of these transcription factors contain a conserved basic helix-loop-helix (bHLH) TCP domain, and a subset (five in *Arabidopsis*) also includes a conserved arginine-rich R domain (Cubas et al. 1999*a*). *CYCLOIDEA* is found in two closely related copies, *CYC* and *DICHOTOMA* (*DICH*) within the Antirrhineae (Hileman and Baum 2003), which includes snapdragon. In *A. majus* and *Mohavia* (both nested within *Antirrhinum*; Oyama and Baum 2004), *CYC* and *DICH* have overlapping expression patterns in floral meristems (Luo et al. 1996, 1999; Hileman et al. 2003), and a fully radial and ventralized flower (a peloric form) is only produced in *CYC*/*DICH* double mutants. Although there is partial redundancy in function, they do differ slightly in the timing of expression (Luo et al. 1999). In addition, *CYC* and *DICH* both inhibit stamen growth in *A. majus*, with expression in stamen primordia resulting in abortion (Luo et al. 1996, 1999).

Outside of the Antirrhineae, the *CYC* locus has been examined in several other groups of eudicots. Most of this work has

[1] Author for correspondence; telephone 203-432-7168; fax 203-432-3758; e-mail dianella.howarth@yale.edu.

Manuscript received September 2004; revised manuscript received December 2004.

focused on other Lamiales (Cronk and Möller 1997; Citerne et al. 2000; Cronk 2001; Cubas 2002; Hileman and Baum 2003; Reeves and Olmstead 2003; Smith et al. 2004), although a few other groups have been explored, including *Solanum* (Reeves and Olmstead 2003), various legumes (Citerne et al. 2003; Fukuda et al. 2003; Ree et al. 2004), and *Arabidopsis* (Cubas et al. 2001). Multiple copies of *CYC*-like genes have been found in all of these studies. The availability of whole-genome data has also greatly aided our understanding of the gene family, with 24 TCP-containing genes having been identified in *Arabidopsis* (Cubas 2002; Damerval and Manuel 2003) and at least that many in *Oryza* (Goff et al. 2002).

The diversification of this gene family is still very poorly understood because of the limited sample of species studied to date. Within the Asteridae, there is a basic understanding of the number of copies and, in a few cases, of gene function within the lamiid clade (Bremer et al. 2002; the Euasterid 1 clade of APG II 2003), which includes one major origin of bilateral symmetry (Donoghue et al. 1998; Ree and Donoghue 1999; Reeves and Olmstead 2003). The aim of this study was to assess the copy number and establish the location of potentially informative duplications in *CYC*-related genes in the Asteridae outside of the lamiids, in a lineage in which monosymmetry was derived independently (Ree and Donoghue 1999). For this purpose, we focused our attention on the large campanulid clade (Bremer et al. 2002; the Euasterid 2 clade of APG II 2003), which includes three large subclades—Apiales, Asterales, and Dipsacales—and several smaller groups. Specifically, we have examined *CYC*/TCP diversity within the Dipsacales.

Our focus on Dipsacales was motivated by several factors. First, Dipsacales phylogeny is quite well resolved, based on a series of recent 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). This new understanding (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 *CYC* and related genes, including the location of possible gene duplications and losses and their relation to morphological changes.

Second, a variety of floral forms are found within Dipsacales, including polysymmetric, monosymmetric, 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 monosymmetric flowers originated independently within this lineage. The primary split within Dipsacales separates the Adoxaceae from the Caprifoliaceae (fig. 1). Adoxaceae (including *Viburnum*, *Sambucus*, and *Adoxa* and its relatives) have polysymmetric flowers, mostly lacking nectaries, with small calyx lobes, rotate corollas, short styles, and lobed stigmas, whereas Caprifoliaceae (including Diervilleae, Caprifolieae, Linnaeeae, Morinaceae, Dipsacaceae, and Valerianaceae) typically have monosymmetric flowers, nectaries composed of unicellular hairs situated at the base of the corolla, larger calyx lobes, tubular corollas, elongate styles, and capitate stigmas (Fukuoka 1972; Donoghue et al. 2003). Monosymmetry most likely arose along the line leading to the Caprifoliaceae (fig. 1).

Third, there are several forms of monosymmetry present within the Caprifoliaceae. Two-lipped flowers, in which two dorsal petals are differentiated from the two lateral petals and the medial ventral petal (the 2 : 3 form), are widespread within the clade and appear to be ancestral. Within the Caprifolieae, two-lipped flowers, in which the two dorsal and the two lateral petals are differentiated from the medial ventral petal (the 4 : 1 form), are found in *Lonicera* and *Triosteum*, and polysymmetric flowers appear to have reevolved in *Symphoricarpos*. There have also been several shifts in the number of corolla lobes, most notably the reduction from five to four lobes in Dipsacaceae. Stamen number was reduced from five to four in the Linnina clade (fig. 1) through loss of the medial dorsal stamen. Further reductions in stamen number within Linnina occurred within Morinaceae (to two stamens in the *Morina*/*Cryptothladia* clade) and within the Valerianaceae (to three stamens in core Valerianaceae [in *Valeriana*, *Valerianella*, and *Plectritis*]), followed by reduction to two stamens in *Fedia* and independently to just a single stamen in the highly asymmetric flowers of *Centranthus* (Donoghue et al. 2003).

Material and Methods

Plant Material

Twenty-eight individuals were sequenced from 21 genera, representing all major lineages within Dipsacales. We used total genomic DNAs, obtained in many cases from extractions used in previous phylogenetic studies. Table 1 provides a list of included taxa, voucher specimens, and GenBank accession numbers.

Primer Design and Amplification

All primers were designed in the TCP domain (forward primer) and the R domain (reverse primer). Initial primers, designed from published TCP genes, were used for a subset of taxa (*CYC*F1: AAAGAYCGVCACAGCAA; *CYC*R-LH[a]: CTCGCYCTCGCYTTCGCCCTCGACTC). Additional primers were developed through the course of the study as more sequences were obtained. Multiple primer pairs were used for each taxon. Table 2 lists all of the primers used and their success rates.

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 0.5x Q Solution (QIAGEN). Amplified products were cloned using the Invitrogen TOPO TA cloning kit for sequencing (Invitrogen, Carlsbad, CA).

Between 10 and 40 colonies (200–800 bp each) were screened (PCR reactions with the same concentrations as above in a 20-μL volume with standard M13 primers) for all potentially different copies or alleles of the *CYC*-like genes obtained. Selected colonies were then picked and directly mixed 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

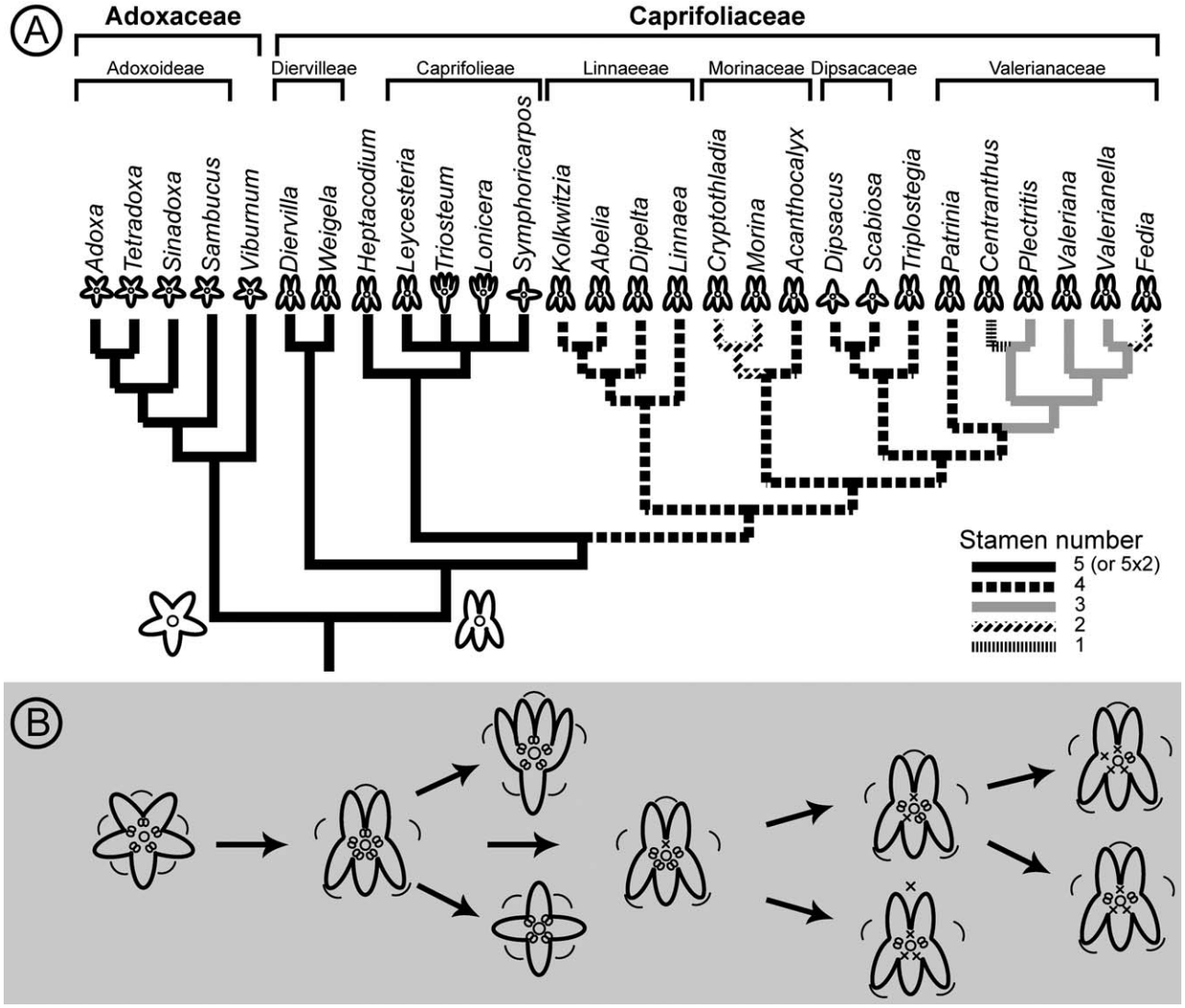

Fig. 1 Evolutionary shifts in floral symmetry and stamen number within Dipsacales. *A*, Summary phylogenetic tree for major lineages within Dipsacales, with the taxonomic names used here based on Donoghue et al. (2001). Floral drawings represent major forms of symmetry (polysymmetric or monosymmetric with 2 : 3 or 4 : 1 corolla orientation). Patterns on the branches represent different stamen numbers (doubled in *Adoxa* and its relatives). *B*, Major shifts in floral form derived from phylogenetic hypothesis for Dipsacales. Outer curved lines represent the calyx lobes. The form of corolla symmetry is shown by placement of the corolla lobes. Stamens are shown with double circles. An "x" represents the loss of a functional structure.

cycles. Amplification products of the appropriate size were cleaned using the QIAGEN PCR cleanup kit (QIAGEN) and were directly sequenced. Cycle sequencing reactions used the BigDye system (Applied Biosystems, Foster City, CA) according to manufacturer's instructions and were electrophoresed on an ABI PRISM 3100 genetic analyzer (Applied Biosystems) or an MJ Research BaseStation51 DNA fragment analyzer (MJ Research, South San Francisco, CA).

Phylogenetic Alignment and Analyses

All clones from each DNA extraction obtained from multiple primer pairs were compiled into Sequencher, version 4.2 (Gene Codes Corporation, Ann Arbor, MI). *CYC*-like genes were determined by the presence of the highly conserved amino acid sequence of the TCP domain. Positive clones were separated into different "types" based on shared differences among the clones. 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 was exported for phylogenetic analysis. Protein sequences for each potential copy/allele obtained were aligned by eye in MacClade, version 4 (Maddison and Maddison 2003), and the entire aligned nucleotide matrix was used in subsequent phylogenetic analyses.

Parameters for the Bayesian and maximum likelihood analyses were estimated using MODELTEST, version 3.06 (Posada and Crandall 1998). The Akaike Information Criterion (AIC; Akaike 1973) recommended a general time reversible (GTR) model with added parameters for invariable sites

Table 1

Dipsacales Species Used, with Collection Locations and Voucher Information

Species	Location	Voucher	GenBank numbers for obtained copies						
			1	2	2A	2B	3	3A	3B
Abelia × *grandiflora*	Cult. Univ. Arizona, Tucson	Donoghue, voucher lacking	165					222, 223	
Centranthus ruber	Cult. San Francisco, CA	Bell 203 (YU)			184	185, 186		224	225
Cryptothladia chinensis	Luhuo, Sichuan, China	Boufford, Donoghue, and Ree 27870 (A)				187, 188		226	227, 228
Cryptothladia kokonorica	Nangqen, Qinghai, China	Boufford et al. 26573 (A)			189	190, 191			229, 230
Diervilla sessilifolia	Cult. Arn. Arb. 960-79	Elsik and Zinman 3286 (A)	166		192	193		231	
Dipelta floribunda	Cult. Arn. Arb. 14514-B	Buckland and Kelly 32 (A)	167		194	195		232	233
Dipsacus pilosus	Nat. Bot. Gard. of Belgium	F. Billiet 1445	168			196, 197			234
Heptacodium miconioides	Cult. Arn. Arb. 1549-80	Koller s.n. 21 Oct. 1984 (A)	169			198		235	236
Kolkwitzia amabilis	Cult. Arn. Arb. 20447-B	Elsik, Michener, and Bailey 844 (A)	170		199	200		237	238
Leycesteria formosa	Yunnan, Lijiang, China	Boufford et al. 29341 (A)	171		201				
Leycesteria sp.	China 6/10/98	Boufford et al. 44597 (A)	172						
Linnaea borealis	Door County, WI	Donoghue 113, voucher lacking	173					239	
Lonicera heteroloba	Cult. Arn. Arb. 838-76-A	Elsik, Dumaine, and Groves 1668 (A)	174		202	203		240, 241	
Lonicera prolifera	Cult. Arn. Arb. 870-74-A	Donoghue 107, voucher lacking			204			242	243
Morina longifolia	Cult. Bergius Bot. Gard., Sweden	Eriksson s.n. 2 Nov. 1999 (SBT)				205, 206			244, 245
Patrinia triloba	Cult. Bergius Bot. Gard., Sweden	Eriksson 807 (SBT)	175		207	208			246
Sambucus adnata	Moxi, Sichuan, China	Boufford, Donoghue, and Ree 27187 (A)					247		
Sambucus canadensis	Cult. Marsh Bot. Gard., Yale U.	Donoghue and Winkworth 37 (A, YU)	176	209			248		
Scabiosa columbaria	Nursery, Marin, CA	Bell 199, voucher lacking	177						
Symphoricarpos occidentalis	Cult. Arn. Arb. 1078-81	Donoghue, voucher lacking	178					249	250
Symphoricarpos orbiculatus	Cult. Arn. Arb. 407-32	Donoghue, voucher lacking	179		210				
Triosteum himalayanum	Kangding, Sichuan, China	Boufford, Donoghue, and Ree 27525 (A)	180		211	212		251	
Triplostegia glandulifera	Moxi, Sichuan, China	Boufford, Donoghue, and Ree 27738 (A)	181			213			252
Valerianella dentata	Cult. Yale Univ.	Bell 106 (YU)	182			214			253
Viburnum lentago	Cult. Arn. Arb. 18021-A	Donoghue and Winkworth 21 (A, YU)		215					
Viburnum prunifolium	Cult. Arn. Arb. 22586-A	Donoghue and Winkworth 13 (A, YU)		216					
Viburnum rhytidophyllum	Cult. Arn. Arb. 133-67-A	Donoghue and Winkworth 8 (A, YU)		217, 218			254		
Viburnum rufidulum	Cult. Arn. Arb. 21418-A	Donoghue and Winkworth 14 (A, YU)		219					
Weigela hortensis	Cult. Arn. Arb. 1897-77-A	Kelly and Buckland 28 (A)	183		220	221		255	

Note. Arnold Arboretum (Arn. Arb.) is in Jamaica Plain, MA. A = Arnold Herbarium, Harvard University Herbaria; SBT = Bergius Foundation, Stockholm; YU = Yale University Herbarium. GenBank numbers are given for each copy found. The GenBank numbers are AY851165–AY851255; only the last three digits are shown.

Table 2

Primer Sequences Used in this Study

Name	5′ to 3′	Start base	Success rate
Forward:			
*CYC*F1	AAAGAYCGVCACAGCAA	1	+++
*CYC*F2	AAAGAYCGVCACAGCAARATMA	1	++
*CYC*F3	AAAGAYCGVCACAGCAARATMT	1	++
*CYC*F4	AAAGAYCGVCACAGCAARATATTT	1	−−
*CYC*73aF	GCNCGDAAWTTCTTYRRYCTCCAAG	73	++
*CYC*73aaF	GCNCGDAAWTTHTTYRRYCTCCAVGA	73	+++
*CYC*73bF	GCNCGNARRTTYTTYGATCTDCAAG	73	+++
*CYC*73cF	GCGCGCAAATTCTTTGATCTCCAAG	73	+
*CYCG*1F	AARGAYMGNCAYWSNAARATHWWYACNGCNCANGG	1	−−
*CYCG*34F	GGNBYNMGDGACMGNAGRRTNAGRYTNTC	34	−−
*CYCG*70F	RTNGCNMTNAARTTYTTYGRYYTNCARGA	70	−−
*CYCG*106F	GGNTWYGANAARSCNAGYMRRACNVTNGANTGG	106	−−
Reverse:			
CYCR-LH[a]	CTCGCYCTCGCYTTCGCCCTCGACTC	29	+
CYCR-LH[b]	CTCGCYCTCGCYTTCGCCCTCGWCTC	29	++
*CYCR*1	CTCGCYTTCGCCCTCGWCTC	23	++
*CYCR*2	CTTGCTCTTTCYCTYGCYTTYGCCC	35	+++

Note. The 5′ start base is counted from the first lysine (starts KDRHSK) of the TCP domain for the forward primer and from the first lysine (starts KESRAKAR) of the R domain for the reverse primer. A success rate is indicated for each primer, ranging from use in many of the successful amplifications (+++) to failure to obtain *CYC*-like genes (−−). Many of these primers should be broadly applicable to other eudicot groups (D. G. Howarth and M. J. Donoghue, unpublished data).

[a] L. C. Hileman, unpublished data.

[b] Modified from the L. C. Hileman primer.

and a gamma distribution (GTR + i + g). Bayesian analyses were conducted using this model for the entire matrix. We used the Metropolis-coupled Markov Chain Monte Carlo (MCMCMC) method as implemented in MrBayes, version 3.0b4 (Huelsenbeck and Ronquist 2001), to run four chains (three heated). We ran 5 million generations, sampling every 100 generations. The trees were analyzed in Tracer (Rambaut and Drummond 2003) to determine a burn-in of 1 million trees. A majority-rule consensus was calculated from the remaining trees in PAUP*, version 4.0b10 (Swofford 2001), to determine the posterior probabilities. In addition, 1000 maximum parsimony bootstrap replicates were run to compare parsimony support for major clades. For each replicate, we carried out 25 random addition searches with *maxtrees* set to 1000. Maximum likelihood analyses were performed in PAUP*, version 4.0b10 (Swofford 2001), on the three major clades identified in the Bayesian analyses. The maximum likelihood model was determined using the AIC (Akaike 1973) as calculated in MODELTEST, version 3.06 (Posada and Crandall 1998), resulting in one of the following recommended models for each clade: the K81 model (Kimura 1981), the TrN model (Tamura and Nei 1993), and a GTR model. Heuristic searches were performed with 20 random addition replicates and tree bisection reconnection (TBR) branch swapping, and 100 maximum likelihood bootstrap replicates were examined.

Results

Bayesian Analyses of the Entire Matrix

One to six different copies or alleles were determined for each taxon. The combined matrix therefore contained 91 total sequences. The aligned length was 468 bases, with 72 constant and 374 parsimony informative characters. The alignment was unambiguous in the TCP and R domains across all taxa. Although the intervening region was difficult to align, we chose to use the entire region. Analyses using just the conserved domains yielded the same major clades as those discussed below. These data can be obtained from TreeBASE (http://www.treebase.org).

The majority-rule consensus from the Bayesian analysis strongly supported three distinct clades (fig. 2), with each clade containing members from every major subclade within Dipsacales. This result indicates that there are three major clades of *CYC*-like genes present in the Dipsacales, which we have labeled *DipsCYC1*, *DipsCYC2*, and *DipsCYC3*. The entire gene tree has been rooted in figure 2 with *DipsCYC1*, which is the most diverged from the other two lineages. Each of the three gene subtrees is supported by a posterior probability of 100. Maximum parsimony also recovered these three clades with bootstrap support values of 97, 56, and 77, respectively (fig. 2).

Maximum Likelihood (ML) Analyses of the Major Gene Lineages

Here we describe the results of separate ML analyses of the *DipsCYC1*, *DipsCYC2*, and *DipsCYC3* clades identified in the Bayesian and maximum parsimony analyses and identify additional duplications in some lineages by reference to the species tree for Dipsacales (fig. 1). *DipsCYC* copies missing for particular taxa are here presumed to be from lack of successful amplification rather than absence/loss from those genomes.

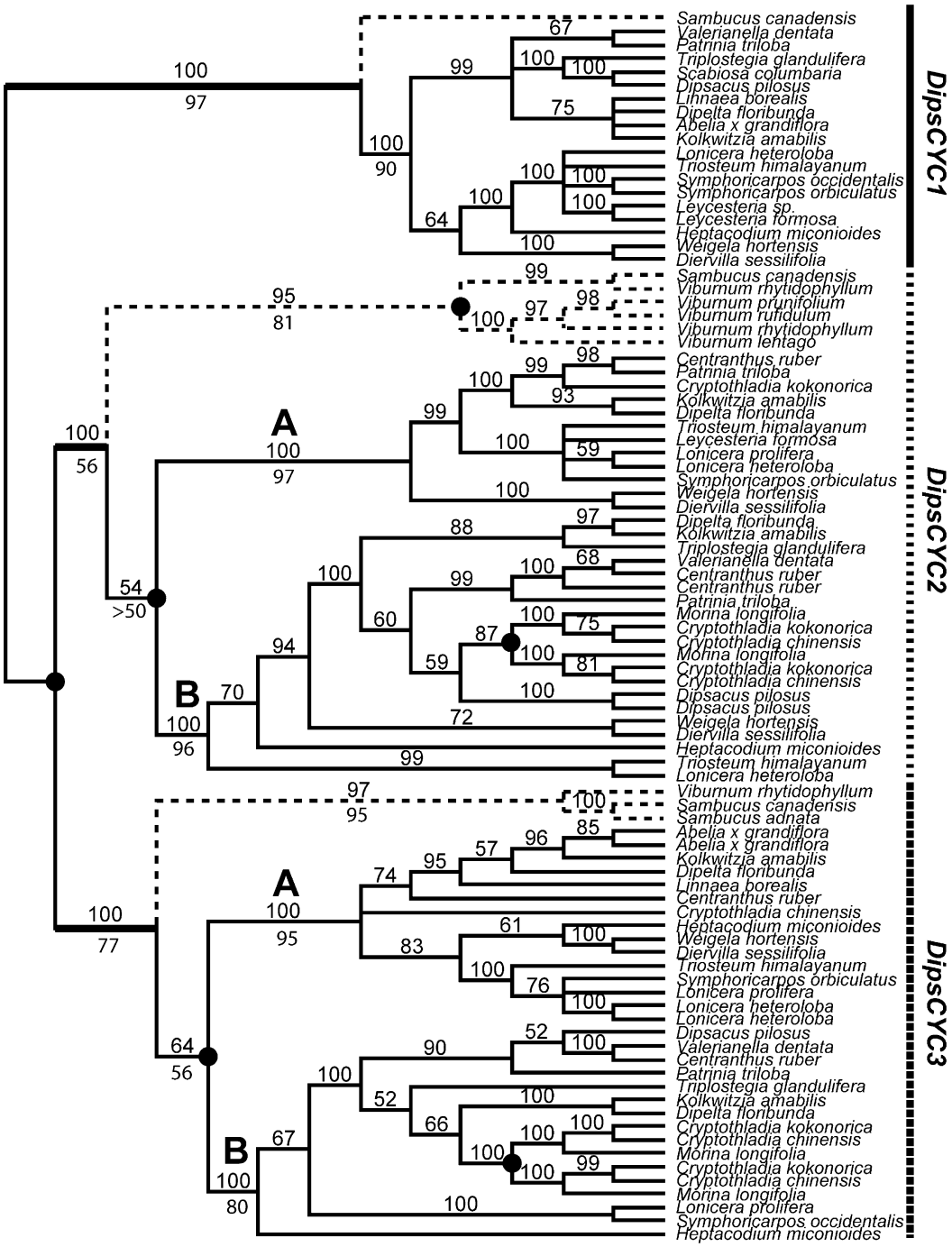

Fig. 2 Gene tree based on Bayesian analyses of Dipsacales *CYCLOIDEA* sequences, showing three major clades of *CYCLOIDEA* (*DipsCYC1*, *DipsCYC2*, and *DipsCYC3*). Each of the three major clades include representatives of Adoxaceae (dashed lines) and Caprifoliaceae (solid lines). Secondary duplications in Caprifoliaceae are labeled "A" and "B." Numbers above the lines are Bayesian posterior probabilities; numbers below the lines are parsimony bootstrap percentages. The nodes at the base of each hypothesized duplication are marked by filled circles.

DipsCYC1. The AIC, as calculated in MODELTEST, version 3.06, recommended the K81 model with a gamma distribution. ML analysis using this model resulted in four equally likely, topologically identical trees (fig. 3). *DipsCYC1* contains *Sambucus* from the Adoxaceae and taxa from all of the major groups within the Caprifoliaceae, except for the Morinaceae. There is no evidence of secondary duplications or allelic diversity within this copy.

DipsCYC2. The GTR model with added parameters for invariable sites and a gamma distribution was selected for *DipsCYC2*. ML analysis resulted in one most likely tree (fig. 4). One major form of *DipsCYC2* is present in Adoxaceae

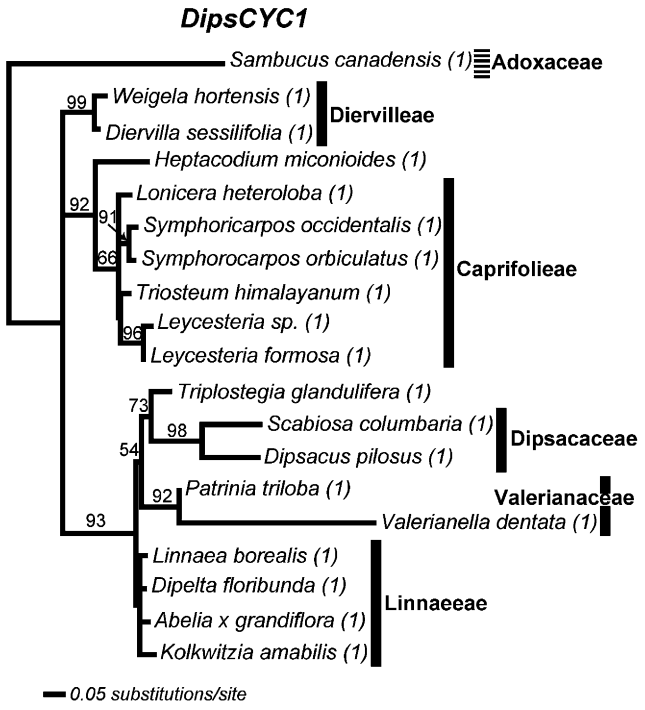

Fig. 3 Maximum likelihood (ML) phylogram of *DipsCYC1*, with ML bootstrap percentages above 50% above the lines. Numbers after the species names indicate copy number.

(represented by *Sambucus* and *Viburnum*), whereas two major forms are found in most of the representatives of the Caprifoliaceae. From this, we infer a duplication event along the line leading to the Caprifoliaceae. What we have labeled "copy 2A" includes species of Diervilleae, Caprifolieae, Linnaeeae, and Valerianaceae, but not Dipsacaceae, while copy 2B contains representatives from all major groups of Caprifoliaceae. In addition, we infer four other, more narrow, duplications. One of these duplications is in *Viburnum* of the Adoxaceae. The other three duplications are in copy 2B in Morinaceae, Dipsacaceae, and *Centranthus*. It is possible that these forms in *Dipsacus* and *Centranthus* represent different alleles, but we hold the view that they represent gene duplications based on sequence divergence. The two types in *Dipsacus* have 17 amino acid length differences and 38 base changes, while those in *Centranthus* have two amino acid length differences and 22 base changes. A larger number of copies (duplication events) occur in *DipsCYC2* than in the other two major forms of the gene.

DipsCYC3. The TrN model, with added parameters for invariable sites and a gamma distribution, was used in the ML analysis of *DipsCYC3*. This resulted in three topologically identical, equally likely trees (fig. 5). Adoxaceae is represented in this phylogeny by both *Sambucus* and *Viburnum*. *DipsCYC3*, like *DipsCYC2B*, appears to undergo a gene duplication at the base of the Caprifoliaceae. Copy 3A contains all of the major subgroups other than Dipsacaceae, while copy 3B contains representatives from all groups except the Diervilleae. In *DipsCYC3B*, there appears to be a secondary duplication in the Morinaceae, mirroring the finding for *DipsCYC2*.

Length Differences and the ECE Region

All the three copies in Dipsacales and the separate subcopies within the Caprifoliaceae appear to differ in sequence length (fig. 6). *DipsCYC1* averages the greatest length, with progressively decreasing lengths in *DipsCYC2* and *DipsCYC3*. However, size differences within the Caprifoliaceae (fig. 6, diamonds) are more extreme than within the Adoxaceae (fig. 6, circles). Within the Caprifoliaceae, each of the two subduplications results in one longer and one shorter copy, both of which are on average shorter than the Adoxaceae sister copy. In *DipsCYC2*, the shorter copy, 2B, has lost much of the alignable intervening sequence with copy 2A and instead contains stretches of asparagines. In length, the only clear overlapping points between copies 2A and 2B are one each (the 2Ba form) of the narrowly duplicated forms in Morinaceae and *Dipsacus* (fig. 4). The other copies (2Bb) from these two groups do not overlap with 2A. In *DipsCYC3*, the Adoxaceae sequences are clearly longer than any sequence from either subcopy from the Caprifoliaceae, and there is no overlap between the two Caprifoliaceae subcopies. Copy 3B is the most reduced, with only ca. 30 amino acids between the TCP and R domains.

Most of the examined *DipsCYC* copies contain a conserved stretch of amino acids within the variable region between the TCP and R domains that we have termed the "ECE region" (fig. 7). This region is comprised of a glutamic acid–cysteine–glutamic acid string that is usually flanked by serines and valines (fig. 7). Both of the short subcopies in the Caprifoliaceae (2B and 3B) lack this ECE region.

Discussion

DipsCYC *Gene Trees in Relation to Dipsacales Phylogeny*

The best fit between the species tree for the major lineages of Dipsacales (fig. 1) and our tree of CYC-like genes implies that crown Dipsacales evolution was preceded, at some unknown point, by two gene duplication events, yielding three major copies (*DipsCYC1*, *DipsCYC2*, and *DipsCYC3*). This comparison further indicates that there were additional duplication events within two of the three gene lineages (*DipsCYC2* and *DipsCYC3*) along the line leading to the Caprifoliaceae. We also infer more narrow duplications in *DipsCYC2* within *Viburnum*, in *DipsCYC2B* and *DipsCYC3B* in Morinaceae, and in *DipsCYC3B* in Dipsacaceae (*Dipsacus pilosus*) and *Centranthus*. These duplications do not appear to be the result of recent polyploidization events, given that they do not occur across all copies. In addition, there is no evidence for genome doubling at the base of the Caprifoliaceae, *Viburnum*, or Dipsacaceae clades. Chromosome counts from *Morina* and *Centranthus* indicate that these groups are polyploids; however, there are currently not enough karyotype data or *DipsCYC* sampling to know whether there is a correlation with gene number (Verlaque 1985; Benko-Iseppon 1992).

The topologies of the trees for the three major gene lineages are generally congruent with that of the species tree. Although there are some small differences, support for relationships in the gene tree is generally poor at these nodes.

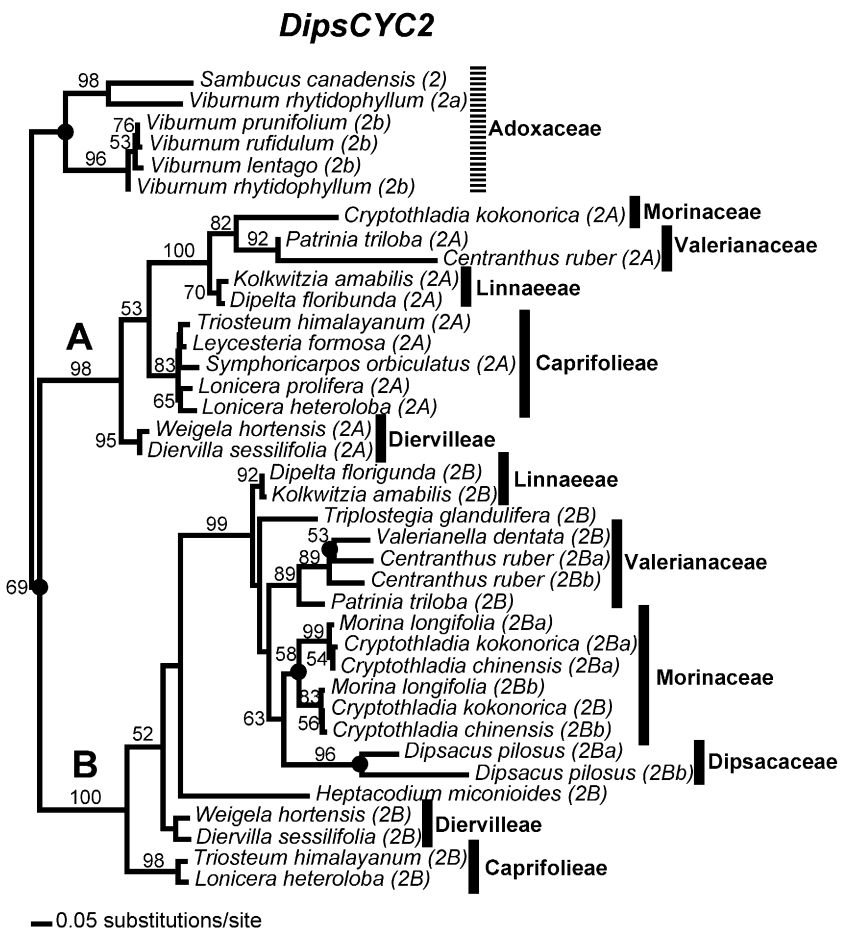

Fig. 4 Maximum likelihood (ML) phylogram of *Dips*CYC2, with ML bootstrap percentages above 50% above the lines. Numbers after the species names indicate copy number. Copies stemming from the major duplication at the base of the Caprifoliaceae are labeled "A" and "B." The nodes at the base of each hypothesized duplication are marked by filled circles.

The *DipsCYC1* tree (with no duplications), insofar as it is resolved, is consistent throughout with the species tree (fig. 3). *Heptacodium* is united with the Caprifolieae, which is concordant with most studies to date (Pyck 2001), and Linnaeeae is united with representatives from the Valerina clade (Valerianaceae, *Triplostegia*, and Dipsacaceae). *DipsCYC1* has not yet been recovered from *Viburnum* or Morinaceae.

Each of the two major copies of *DipsCYC2* is congruent with the Dipsacales phylogeny (fig. 4). In *DipsCYC2A*, Diervilleae is sister to the remaining Caprifoliaceae, whereas the positions of Diervilleae and Caprifolieae are reversed in *DipsCYC2B*. Although Diervilleae is generally preferred as sister to the rest of the Caprifoliaceae in the species tree (Donoghue et al. 2003), this relationship is still somewhat poorly resolved, and the placement of Caprifolieae as sister is only slightly less likely. Within *DipsCYC2A*, we find support for Caprifolieae and for the Linnina clade. So far, *DipsCYC2A* has not been recovered from the *Triplostegia*/Dipsacaceae lineage.

In contrast to the species tree, in *DipsCYC2B Heptacodium* is placed as sister to the Linnina clade, although this has very weak bootstrap support. In maximum parsimony trees for *DipsCYC2B*, *Heptacodium*, Diervilleae, and Caprifolieae all form a clade that is sister to Linnina. Within Linnina, Morinaceae (instead of *Triplostegia*) is linked directly with Dipsacaceae, although again bootstrap values for these nodes are low.

In *DipsCYC3*, as in *DipsCYC2*, we infer a duplication at the base of the Caprifoliaceae, and there is generally good correspondence of the two resulting gene trees with the species tree (fig. 5). In both *DipsCYC3A* and *DipsCYC3B*, *Heptacodium* is sister to Caprifolieae + Linnina rather than to just Caprifolieae, but this is not well supported in either case. Relationships within the Linnina clade in *DipsCYC3* are also not completely congruent with the species tree; again, there is not much support for relationships in this gene tree. Despite targeted searching, we have not yet recovered *DipsCYC3B* copies from Diervilleae. There are no clear duplications within *DipsCYC3A*, but in *DipsCYC3B* a duplication occurs in the Morinaceae, as it does in *DipsCYC2B*.

In summary, these results generally conform to expectation; gene trees match the species tree, with small deviations due either to the limited number of characters in the gene tree or to our failure to date in isolating particular copies

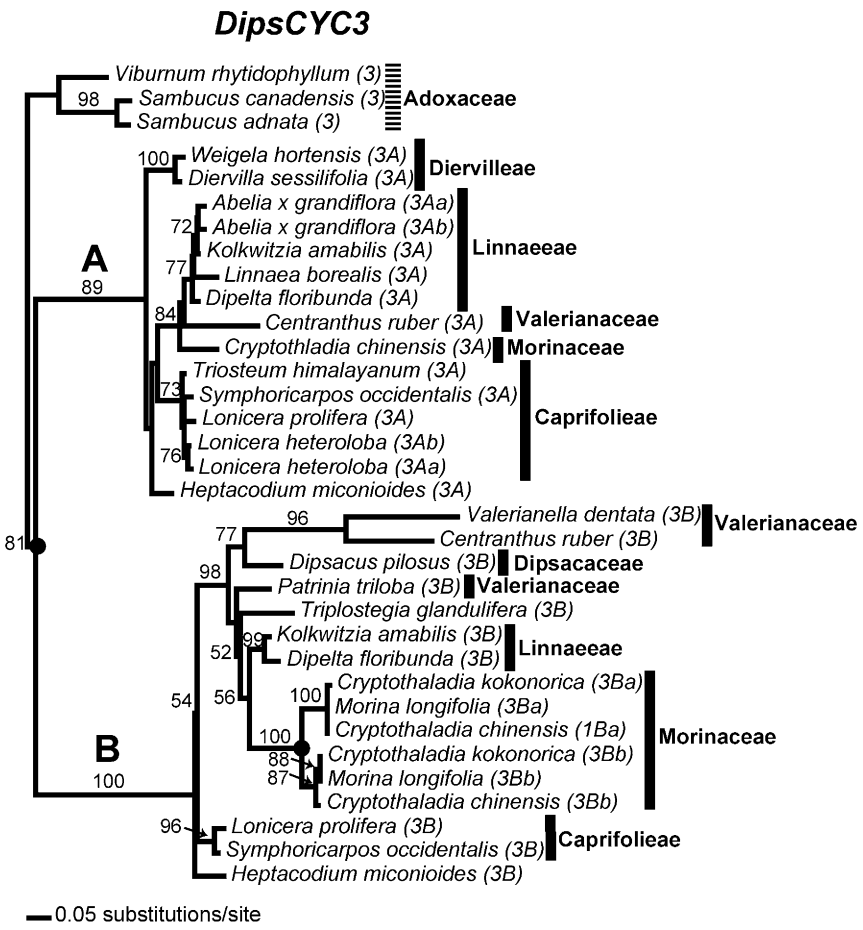

Fig. 5 Maximum likelihood (ML) phylogram of *DipsCYC3*, with ML bootstrap percentages above 50% above the lines. Numbers after the species names indicate copy number. Copies stemming from the major duplication at the base of the Caprifoliaceae are labeled "A" and "B." The nodes at the base of each hypothesized duplication are marked by filled circles.

from some lineages. These results imply that *DipsCYC* genes can be useful for phylogenetic purposes and that the addition of these genes would probably bolster support for the major lineages that have been recognized based on other genes and on morphology. In addition, within the Caprifoliaceae, where relationships have so far resisted resolution based on other data, the *CYC*-like genes may provide some phylogenetically informative characters.

DipsCYC *Diversification and Flower Form*

The number of *CYC*-like genes appears to correlate with flower form in Dipsacales (fig. 8). In the polysymmetric Adoxaceae we find three copies (except for an additional duplication within *DipsCYC2* in *Viburnum*), whereas in monosymmetric Caprifoliaceae we find duplications in two of the major copies (*DipsCYC2* and *DipsCYC3*), resulting in the presence of a base number of five copies in this lineage. Additional duplications nested within Caprifoliaceae, especially additional duplications in two gene lineages (*DipsCYC2B* and *DipsCYC3B*) within the Morinaceae, yield a total of seven copies in some plants. Note that major duplications appear to be situated along the line leading to the whole Caprifoliaceae group, corresponding to the origin of monosymmetric flowers but not along the line to the Linnina clade, where the reduction from five to four stamens occurs (Donoghue et al. 2003). We find no obvious changes in copy number within the Caprifoliaceae associated with the shift from a 2 : 3 to a 4 : 1 flower form in *Lonicera* and *Triosteum* or with the reevolution of polysymmetry within *Symphoricarpos*. This implies that such shifts may result not from changes in copy number but from changes in the expression or function of these genes or from downstream genes. This may also be the case elsewhere, for example, in the reevolution of polysymmetric flowers in *Plantago* (Reeves and Olmstead 2003) and the development of peloric forms in *Linaria*, both within the Plantaginaceae s.l. (Cubas et al. 1999*b*).

A causal relationship between copy number and floral form cannot yet be established because of the lack of gene expression or functional studies. At this stage, we are unsure which copies are active within flowers and where in the flowers they might be expressed. In addition, we assume that these CYC-like genes have been recruited independently for controlling flower symmetry in Dipsacales, since zygomorphy is independently derived in this group. However, the correlations between floral changes and copy number provide viable hypotheses for future research.

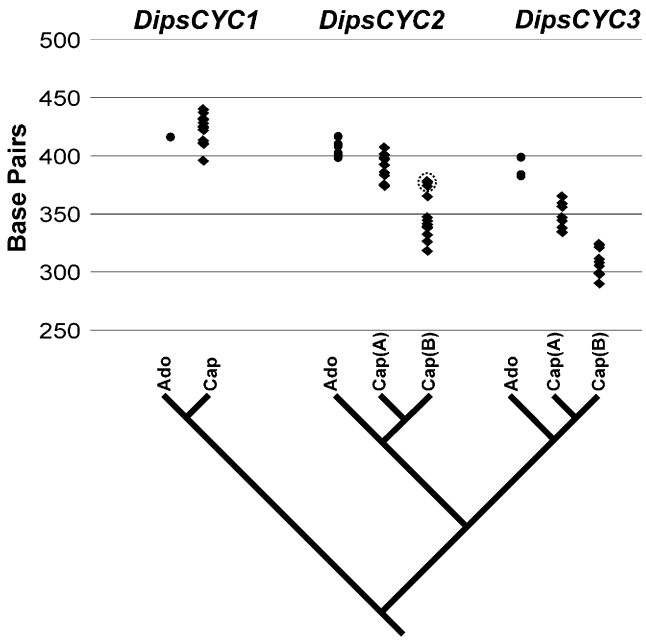

Fig. 6 Base pair length of *CYCLOIDEA* sequences plotted against gene copy. Stylized phylogeny of major gene copies (*Ado* = Adoxaceae, *Cap* = Caprifoliaceae). Lengths are plotted with circles (Adoxaceae) or diamonds (Caprifoliaceae). Circled diamonds in *DipsCYC2B* show one of the two representatives of the secondary duplications within Morinaceae and Dipsacaceae.

The presence of additional duplications/copies in Morinaceae is potentially of great interest because this clade is marked by several major flower changes. First, we see reduction to two functional stamens in the *Morina*/*Cryptothladia* lineage within the Morinaceae. However, stamen number is reduced elsewhere within the Dipsacales (e.g., within Valerianaceae) without a corresponding shift in gene number. In addition, stamen loss in the Dipsacales typically results in the loss of the two ventral stamens, which differs from a simple expansion of *CYC*-like genes from the dorsal into the lateral stamens (Hileman et al. 2003). Second, all Morinaceae are characterized by a highly unusual monosymmetric calyx (figs. 1, 8). 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 this lineage are related to the new function of producing a monosymmetric calyx in addition to the maintenance of a monosymmetric corolla. The fact that *CYCLOIDEA* in *Antirrhinum* controls the abortion of the medial dorsal stamen suggests that *DipsCYC* could play a role in the calyx development of Morinaceae. *CYC* is expressed early in sepal development in *Antirrhinum*, where it initially retards growth (Luo et al. 1996). In addition, there is evidence that the expansion of *CYC* expression can result in the abortion of two additional lateral stamens (Hileman et al. 2003). We therefore postulate that the expansion of expression outward to the calyx could lead to calyx lobe abortion. This would provide an instance of the transference of function through homeoheterotopy (Baum and Donoghue 2002). Unfortunately, it is still unclear how the calyx develops in *Acanthocalyx*, the sister group of the *Morina*/*Cryptothladia* clade (Bell and Donoghue 2003). In addition, we still need to determine precisely where the additional duplications originated. That is, did they occur along the line leading to all Morinaceae (correlated with the origin of the monosymmetric calyx), or along the line leading to the *Morina*/*Cryptothladia* lineage (correlated with the loss of two functional stamens)?

Our analyses provide preliminary evidence for several other patterns in Dipsacales. First, there may be an additional duplication within Dipsacaceae *DipsCYC2B*. The phylogenetic level at which this occurred remains to be determined by better sampling of these organisms. Dipsacaceae, like the Morinaceae, is marked by the presence of an epicalyx surrounding each flower (Donoghue et al. 2003), but the calyx is polysymmetric. Dipsacaceae do, however, undergo a reduction from five to four petals. Second, in *DipsCYC2B*, there could be an additional duplication in *Centranthus* (fig. 4), which would correlate with a shift to producing only one stamen (Donoghue et al. 2003). However, this would be an asymmetric ventral shift, which does not fit other shifts in stamen loss (Hileman and Baum 2003). Finally, we note the presence of relatively long branches in each of the three gene copies in the core Valerianaceae (which excludes *Patrinia*). This correlates with similarly elevated rates of change in other genes (Bell et al. 2001; Donoghue et al. 2001; Bell 2004) and may represent a generalized increase in rate in this lineage, perhaps also correlated with an increased rate of speciation (Bell 2004; Bell and Donoghue 2005).

Gene Sequence Evolution in DipsCYC

There are significant length differences in the region between the TCP and R domains across the different copies and subcopies of *DipsCYC* (fig. 6). Perhaps the most interesting pattern is that size differences among the polysymmetric Adoxaceae are less than those among the monosymmetric Caprifoliaceae. While there appears to be a decrease in length from *DipsCYC1* to *DipsCYC3* in Adoxaceae, there is overlap among the copies. In *DipsCYC1*, which contains no clear duplications in Dipsacales, the lengths in Adoxaceae and Caprifoliaceae are quite similar. However, in *DipsCYC2* and *DipsCYC3*, sequences from Caprifoliaceae are generally shorter than those from Adoxaceae. Furthermore, there is clearly one longer subcopy and one shorter subcopy in the Caprifoliaceae, with minimal to no overlap in length. Although we currently lack information on the expression or function of these genes, this pattern suggests that these genes have likely diverged in function, potentially permitting both novel and redundant functions. Also of note, the longest copies in *DipsCYC2B* and *DipsCYC3B* stem from the extra duplications within Morinaceae and *Dipsacus*. This could relate to the evolution of a monosymmetric calyx and/or stamen abortion in Morinaceae.

In the thoroughly studied *CYC* and *DICH* duplications in the Antirrhineae, the two copies have semiredundant functions (Luo et al. 1996). Both are expressed in the dorsal

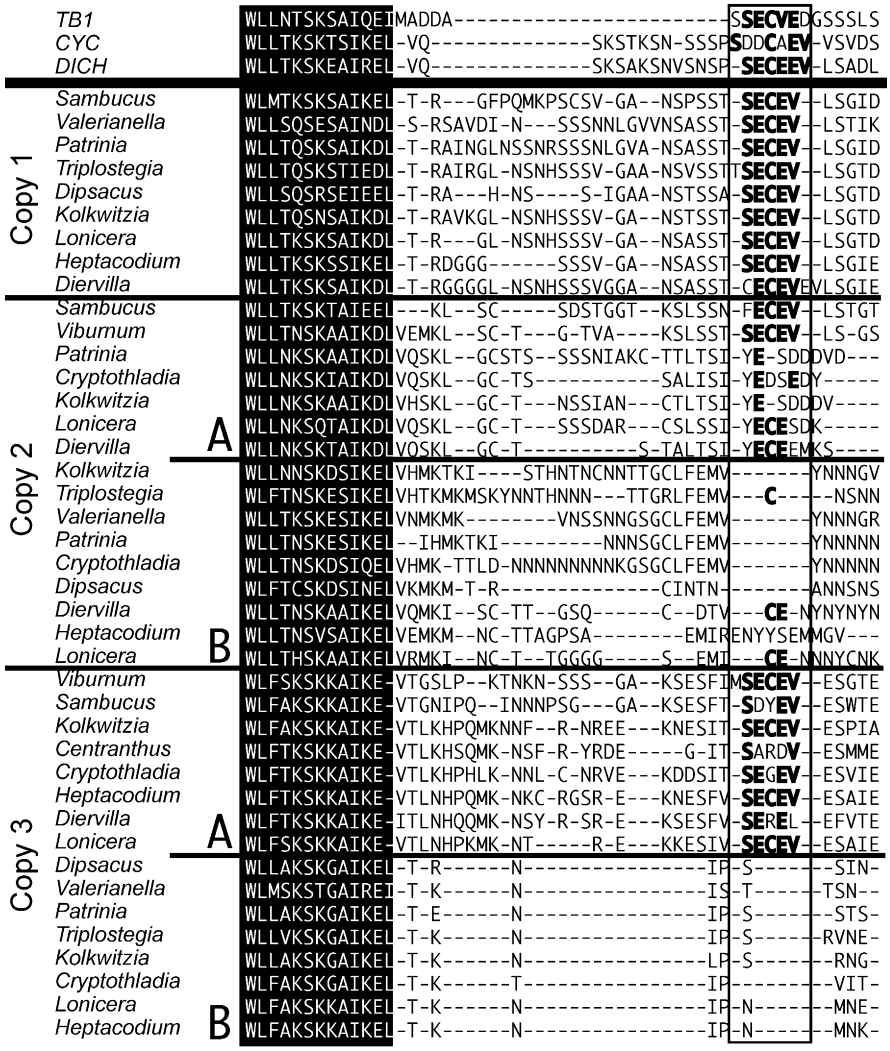

Fig. 7 Protein alignment of the ECE and surrounding regions. The end of the TCP domain is highlighted in black. The boxed residues represent the ECE region with conserved amino acids in boldface. The entire intervening region between the TCP and ECE regions is shown. Alignment includes representatives from the major groups from each copy as well as sequences from *Zea* (*TB1*) and *Antirrhinum* (*CYC* and *DICH*) for comparison.

region of the floral meristem, and both must be knocked out to create a polysymmetric peloric flower. Unlike the duplications we see within the Dipsacales, *CYC* and *DICH* are nearly the same length in *Antirrhinum majus*, ca. 366 bp across the same region (Luo et al. 1996, 1999). Differences in length have, however, been documented in other groups, ranging from 20 to 90 base pair differences in Gesneriaceae and Fabaceae (Citerne et al. 2000, 2003; Fukuda et al. 2003). The three major *CYC*-like copies in *Arabidopsis*—*TCP1*, *TCP12*, and *TCP18*—also differ in length, being 426, 351, and 474 bp, respectively.

Aside from the TCP and R domains, the only other string of amino acids that characterizes all three copies is the ECE region described here, most often being SECEV in sequence (fig. 7). This region occurs clearly in all *DipsCYC1* and in the Adoxaceae in the other two copies. In the Caprifoliaceae, it is present in *DipsCYC2A* and identifiable but varied in *DipsCYC3A* (YECES), but it is lacking in the shorter subcopies (*DipsCYC2B* and *DipsCYC3B*). Although we do not know how this region functions, we hypothesize that it may help stabilize the three-dimensional structure of the protein, because cysteines often form disulfide bonds. We note that this region is also present in *DICH* (*Antirrhinum nuttallianum*, SECEEV; *A. majus*, SECDQE); however, it is slightly diverged in *CYC* (*A. nuttallianum*, SDDCAEV; *A. majus*, CDDCEEV) (Luo et al. 1999; Hileman and Baum 2003). Remarkably, this short region is also recognizable in *TB1* from *Zea mays* (SECVE) (Doebley et al. 1997).

Another protein sequence change is also noteworthy. In all known *TB1*/*CYC* subfamily members, the TCP region contains an RDRRXR motif. However, in *DipsCYC3B*, in the Caprifoliaceae, the aspartic acid (D) has been replaced by an alanine, RARRMR. This region is part of the bipartite nuclear localization signal; however, it is not one of the

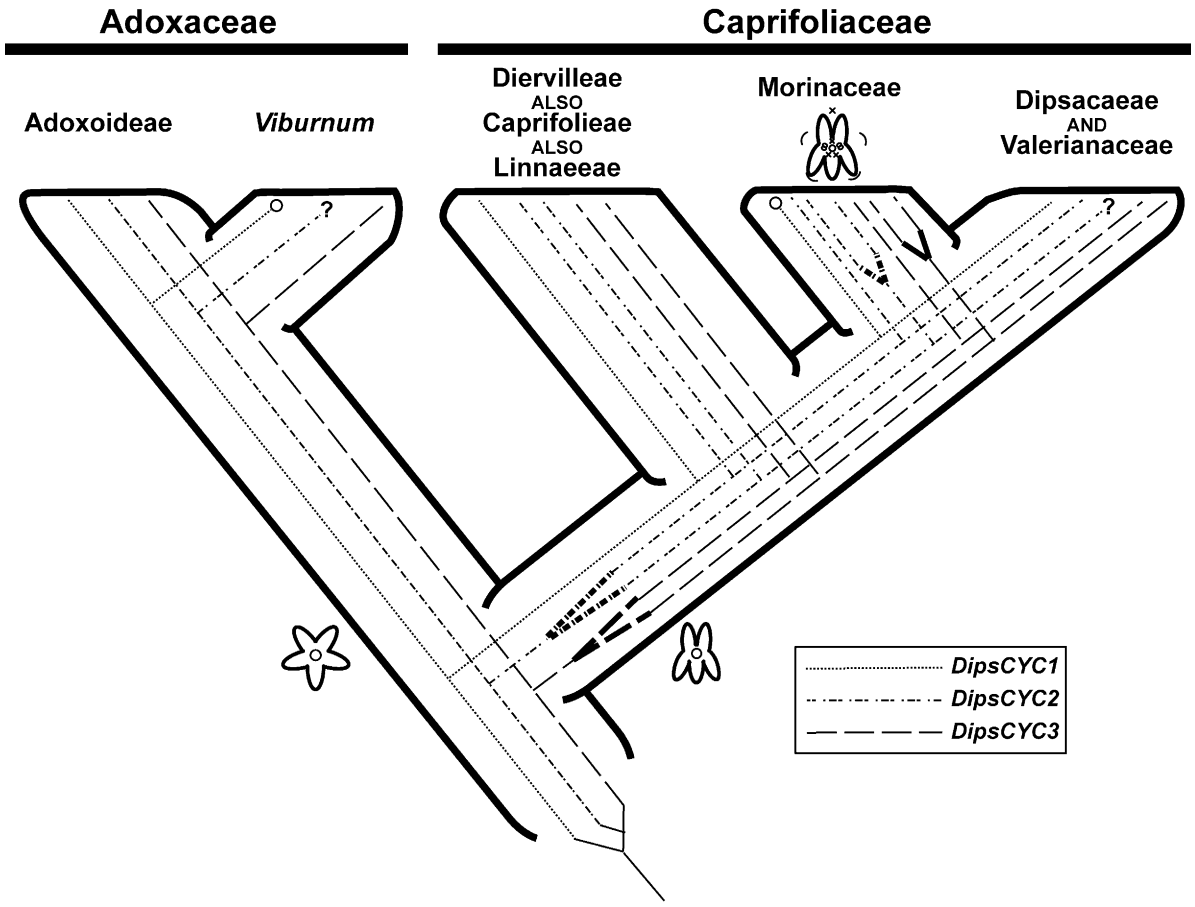

Fig. 8 Summary Dipsacales phylogeny (outer lines) with gene tree plotted within. The species tree is reduced for simplicity; "ALSO" indicates that the same pattern holds in each of the named lineages. Two major duplications are inferred on the line leading to the Caprifoliaceae; two additional duplications are inferred in Morinaceae. Open circles represent copies that we assume to be present, although they have not yet been amplified. Question marks represent possible additional duplications (see text).

required basic residues (Cubas et al. 1999*a*). The nature of this substitution (from a highly reactive aspartic acid to an unobtrusive alanine) and the strict conservation of this sequence in all other sequenced angiosperms indicate a possible change of function in *Dips*CYC3B.

Evolution of CYC*-like Genes*

The TCP gene family is diverse, with a complement of 24 copies found in *Arabidopsis* (Cubas 2002; Damerval and Manuel 2003). This family includes the proliferating cell factor genes (PCFs), first described in rice, which control cell growth. PCFs are easily distinguished from the other subfamily, *CYC/TB1*, by differences in the length and sequence of the TCP domain. The *CYC/TB1* subfamily appears to play a role in dorsal identity, at least in eudicots (Luo et al. 1996; Cubas et al. 2001). A subset of the *CYC/TB1* subfamily has an additional conserved R domain. Interestingly, however, presence of the R domain is not monophyletic, instead falling into two clades (Cubas 2002). One of these clades, which includes both *TB1* and *CYC/DICH*, is where we focused our search for *CYC*-like genes. Other members of the TCP gene family are presumably present in Dipsacales, but these have not been recovered using our primers. Furthermore, they probably function outside the flower or do not function in dorsal/ventral patterning and are therefore unrelated to flower symmetry (Cubas et al. 1999*a*; Cubas 2002).

Several studies have documented the number of copies in various angiosperm groups. Duplications in *CYC*-like genes have been common in the Asteridae (Reeves and Olmstead 2003); therefore, it is not surprising that we have identified duplications in Dipsacales. In *Antirrhinum* and its relatives, there is one clear duplication at the level of the Antirrhineae, resulting in *CYC* and *DICH* (Gübitz et al. 2003; Hileman and Baum 2003). *Digitalis* (also in Plantaginaceae) lacks this duplication but has three additional copies (Hileman and Baum 2003). In Gesneriaceae, two duplications have been hypothesized in Cyrtandroideae (Citerne et al. 2000) and one more within Gesnerioideae (Smith et al. 2004). Outside of the Asteridae, the same pattern has also been documented in legumes, with three or four copies present in many of the groups that have been examined (Citerne et al. 2003; Fukuda et al. 2003; Ree et al. 2004). However, in Poaceae there do not appear to be duplications within the *TB1*-like genes

(Lukens and Doebley 2001), suggesting that dynamic change in copy number could be specific to eudicots.

Our data on *CYC*-like genes in Dipsacales provide a framework for further studies of this gene family. At the moment, without information on expression or function, we can only point to correlations between changes in gene number and changes in floral morphology (fig. 8). However, these correlations are striking enough to provide clear hypotheses regarding the role of *CYCLOIDEA* in Dipsacales floral evolution. We suggest, for example, that duplications in *DipsCYC2* and *DipsCYC3* along the line leading to the Caprifoliaceae may relate to the independent evolution of monosymmetric flowers in this lineage. Likewise, we hypothesize that duplications in *DipsCYC2B* and *DipsCYC3B* in Morinaceae relate to the abortion of the dorsal calyx lobe and/or the reduction of two stamens. Overall, our studies in the Dipsacales expand our knowledge of *CYC*-like genes in the Asteridae and the angiosperms more broadly and provide a system outside of the Antirrhineae for studying the role of these genes in specifying flower form.

Acknowledgments

We thank L. C. Hileman for sharing her internal *Antirrhinum* primers for initial amplifications. We thank C. D. Bell and R. C. Winkworth for the use of Dipsacales DNA extractions. We greatly appreciate D. A. Baum, M. P. Dunn, E. J. Edwards, A. Litt, A. L. Russell, and two anonymous reviewers, as well as other members of the Donoghue lab, for helpful discussions and insights on the manuscript. For funding, we are grateful for the Forest B. H. and Elizabeth D. W. Brown Postdoctoral Fellowship from Yale University and the MORPH Research Coordination Network grant to attend the 2004 Botanical Society of America meeting.

Literature Cited

Akaike H 1973 Information theory as an extension of the maximum likelihood principle. Pages 267–281 *in* BN Petrov, F Csáki, eds. Second international symposium on information theory. Akadémiai Kiadó, Budapest.

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

APG II (Angiosperm Phylogeny Group) 2003 An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants. Bot J Linn Soc 141:399–436.

Backlund A, MJ Donoghue 1996 Morphology and phylogeny of the order Dipsacales. Paper 4 *in* A Backlund, ed. Phylogeny of the Dipsacales. PhD diss. Uppsala University.

Baum DA, MJ Donoghue 2002 Transference of function, heterotopy, and the evolution of plant development. Pages 52–69 *in* Q Cronk, R Bateman, J Hawkins, eds. Developmental genetics and plant evolution. Taylor & Francis, London.

Bell CD 2004 Preliminary phylogeny of Valerianaceae (Dipsacales) inferred from nuclear and chloroplast DNA sequence data. Mol Phylogenet Evol 31:340–350.

Bell CD, MJ Donoghue 2003 Phylogeny and biogeography of Morinaceae (Dipsacales) based on nuclear and chloroplast DNA sequences. Org Divers Evol 3:227–237.

——— 2005 Dating the Dipsacales: comparing models, genes, and evolutionary implications. Am J Bot 92:284–296.

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

Benko-Iseppon AM 1992 Karyologische Untersuchung der Caprifoliaceae s.l. und möglicher verwandter Familien. PhD diss. Universität Wien.

Bremer B, K Bremer, N Heidari, P Erixon, RG Olmstead, AA Anderberg, M Kallersjo, E Barkhordarian 2002 Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. Mol Phylogenet Evol 24:274–301.

Citerne HL, D Luo, RT Pennington, E Coen, QCB Cronk 2003 A phylogenomic investigation of *CYCLOIDEA*-like TCP genes in the Leguminosae. Plant Physiol 131:1042–1053.

Citerne HL, M Möller, QCB Cronk 2000 Diversity of *cycloidea*-like genes in Gesneriaceae in relation to floral symmetry. Ann Bot 86: 167–176.

Cronk Q, M Möller 1997 Genetics of floral symmetry revealed. Trends Ecol Evol 12:85–86.

Cronk QCB 2001 Plant evolution and development in a post-genomic context. Nat Rev Genet 2:607–619.

Cubas P 2002 Role of TCP genes in the evolution of morphological characters in angiosperms. Pages 247–266 *in* QCB Cronk, R Bateman, J Hawkins, eds. Developmental genetics and plant evolution. Taylor & Francis, London.

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.

Damerval C, M Manuel 2003 Independent evolution of *CYCLOIDEA*-like sequences in several angiosperm taxa. C R Palevol 2:241–250.

Doebley J, A Stec, L Hubbard 1997 The evolution of apical dominance in maize. Nature 386:485–488.

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.

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.

Fukuda T, J Yokoyama, M Maki 2003 Molecular evolution of *Cycloidea*-like genes in Fabaceae. J Mol Evol 57:588–597.

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.

Gillies ACM, P Cubas, ES Coen, RJ Abbott 2002 Making rays in the Asteraceae: genetics and evolution of radiate versus discoid flower heads. Pages 233–246 *in* QCB Cronk, RM Bateman, JA Hawkins,

eds. Developmental genetics and plant evolution. Taylor & Francis, London.

Goff SA, D Ricke, T-H Lan, G Presting, R Wang, M Dunn, J Glazebrook, et al 2002 A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296:92–100.

Gübitz T, A Caldwell, A Hudson 2003 Rapid molecular evolution of *CYCLOIDEA*-like genes in *Antirrhinum* and its relatives. Mol Biol Evol 20:1537–1544.

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

Hileman LC, EM Kramer, DA Baum 2003 Differential regulation of symmetry genes and the evolution of floral morphologies. Proc Natl Acad Sci USA 100:12814–12819.

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

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

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

Kimura M 1981 Estimation of evolutionary distances between homologous nucleotide sequences. Proc Natl Acad Sci USA 78:454–458.

Knapp S 2002 Floral diversity and evolution in the Solanaceae. Pages 267–297 *in* QCB Cronk, RM Bateman, JA Hawkins, eds. Developmental genetics and plant evolution. Taylor & Francis, London.

Lukens L, J Doebley 2001 Molecular evolution of the teosinte branched gene among maize and related grasses. Mol Biol Evol 18:627–638.

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.

Neal PR 1998 Floral symmetry and its role in plant-pollinator systems: terminology, distribution, and hypotheses. Annu Rev Ecol Syst 29:345–373.

Oyama RK, DA Baum 2004 Phylogenetic relationships of North American *Antirrhinum* (Veronicaceae). Am J Bot 91:918–925.

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

Posada D, KA Crandall 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. PhD diss. Katholieke University, Leuven.

Rambaut A, A Drummond 2003 Tracer. Version 1.0.1. University of Oxford, Oxford.

Ree RH, HL Citerne, M Lavin, QCB Cronk 2004 Heterogeneous selection on *LEGCYC* paralogs in relation to flower morphology and the phylogeny of *Lupinus* (Leguminosae). Mol Biol Evol 21: 321–331.

Ree RH, MJ Donoghue 1999 Inferring rates of change in flower symmetry in asterid angiosperms. Syst Biol 48:633–641.

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.

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

Smith JF, LC Hileman, MP Powell, DA Baum 2004 Evolution of *GCYC*, a Gesneriaceae homolog of *CYCLOIDEA*, within Gesnerioideae (Gesneriaceae). Mol Phylogenet Evol 31:765–779.

Swofford DL 2001 PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, MA.

Tamura K, M Nei 1993 Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526.

Verlaque R 1985 Étude biosystématique et phylogénétique des Dipsacaceae. II. Caractères généraux des Dipsacaceae. Rev Cytol Biol Veg Bot 8:117–168.

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

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.