Phylogenetic analysis of the ''ECE'' (CYC/TB1) clade reveals duplications predating the core eudicots

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

Department of Ecology and Evolutionary Biology, Yale University, P.O. Box 208106, New Haven, CT 06520-8106

Contributed by Michael J. Donoghue, April 7, 2006

Flower symmetry is of special interest in understanding angiosperm evolution and ecology. Evidence from the Antirrhineae (snapdragon and relatives) indicates that several TCP gene-family transcription factors, especially *CYCLOIDEA* (*CYC*) and *DICHOTOMA* (*DICH*), play a role in specifying dorsal identity in the corolla and androecium of monosymmetric (bilateral) flowers. Studies of rosid and asterid angiosperms suggest that orthologous TCP genes may be important in dorsal identity, but there has been no broad phylogenetic context to determine copy number or orthology. Here, we compare published data from rosids and asterids with newly collected data from ranunculids, caryophyllids, Saxifragales, and Asterales to ascertain the phylogenetic placement of major duplications in the ''ECE'' (CYC/TB1) clade of TCP transcription factors. Bayesian analyses indicate that there are three major copies of ''CYC'' in the ECE clade, and that duplications leading to these copies predate the core eudicots. CYC1 contains no subsequent duplications and may not be expressed in floral tissue. CYC3 exhibits similar patterns of duplication to CYC2 in several groups. Using RT-PCR, we show that, in flowers of *Lonicera morrowii* (Caprifoliaceae), *DipsCYC2B* is expressed in the four dorsal petals and not in the ventral petal. *DipsCYC3B* is expressed in flower and petal primordia, possibly most strongly in the ventral petal.

CYCLOIDEA | ECE clade | floral symmetry | gene duplication | TCP

Small changes in the molecular mechanisms of floral development can give rise to a vast array of different morphologies that potentially affect reproductive strategies and plant evolution. Increasing knowledge of candidate genes is facilitating research in this area. Studies of MADS-box genes, for instance, have led to a well supported model for specification of floral organ identity (1–3). The symmetry of floral organs, however, is less well understood. Shifts in the symmetry of flowers between polysymmetry (actinomorphy, radial symmetry) and monosymmetry (zygomorphy, bilateral symmetry) have been common within angiosperms (4–6), with much study focused on the sympetalous asterids (7–12). Such morphological shifts are of special interest in relation to pollination (e.g., see ref. 13) and perhaps to rates of speciation (14).

Understanding of the evolution of floral symmetry has been greatly advanced by the study of three transcription factors from two gene families: the TCP family [including *CYCLOIDEA* (*CYC*) (15)] and the MYB family [including *DIVARICATA* (16, 17) and RADIALIS (18, 19)]. *CYC*, the most thoroughly studied to date, has been shown to be involved in specifying dorsal, or adaxial, flower identity (15, 17, 20). Work thus far has shown that both *CYC* (in *Antirrhinum majus*) and its putative ortholog, TCP1 (in *Arabidopsis thaliana*), are expressed dorsally (15, 20, 21).

CYC has been well characterized in *A. majus* (snapdragon). It is a member of the TCP gene family, coined from the conserved basic helix–loop–helix (bHLH) TCP domain found in TEOSINTE BRANCHED1 (*TB1*) in *Zea mays*, CYC in *A. majus*, and the proliferating cell factor (PCF) DNA-binding proteins of *Oryza sativa*. Two closely related copies, *CYC* and *DICH* (*DICHOTOMA*), occur within the Antirrhineae (22), which includes snapdragon. In *A. majus* and *Mohavea* [both nested within *Antirrhinum* (23)], *CYC* and *DICH* have overlapping expression patterns in floral meristems (15, 20, 24), and, at least in *Antirrhinum*, a fully radial and ventralized flower (a peloric form) is produced only in *CYC/DICH* double mutants (15, 17). Although there is partial redundancy in function, they do differ slightly in the timing of expression (20). Additionally, *CYC* and *DICH* both inhibit stamen growth in *A. majus*, with expression in stamen primordia resulting in abortion (15, 20).

The TCP gene family is diverse, with a complement of 24 copies found in *Arabidopsis* (refs. 8 and 25, as well as Fig. 1*A*). This family includes the PCF genes, first described in rice, which control cell growth. The PCF subfamily are easily distinguished from members of the other subfamily, *CYC/TB1*, by differences in the length and sequence of the TCP domain (26). A subset of the *CYC/TB1* subfamily has an additional conserved arginine rich ''R domain'' (26). However, it seems that the R domain originated independently in two separate clades (8). One of these clades, which we call the ''ECE'' clade (Fig. 1), includes both TB1 and *CYC/DICH* (as well as TCP1, -12, and -18 from *Arabidopsis*) and is the clade we focus on in this study. ECE refers to a conserved short motif (glutamic acid–cysteine–glutamic acid) between the TCP and R domains that we have found in most members of this clade (27). The remaining members of the TCP gene family are either suspected to function outside of the flower or do not function in dorsal/ventral patterning (8, 26). Within the ECE clade, only *CYC/DICH* and their apparent orthologs in other species have been assayed for expression, leaving open the possibility that there are other closely related genes that could be important in dorsal/ventral patterning. Little is known about the occurrence or the phylogenetic location of major duplications in the ECE clade of TCP genes, and therefore orthology is difficult to assess, especially in non-model organisms.

Duplications in *CYC*-like genes and their apparent orthologs have been common in core eudicots (7, 22, 28–33). Additionally, our recent work on the angiosperm clade Dipsacales (27) demonstrated that three major copies of *CYC*-like genes (*DipsCYC1*, *-2*, and *-3*) were present in the ancestor of that group, and that several additional duplication events then occurred within this clade (Fig. 1*B*). These duplications fell within the ECE clade; however, it could not be determined from studying Dipsacales alone whether they greatly predated the origin of this clade or perhaps took place in the lineage immediately subtending the Dipsacales (Fig. 1*B*). The aim of the present study was to locate the phylogenetic position of these duplication events by using available data and additional sequences we generated from across the eudicots. The resulting ECE gene tree would also provide a framework within which to conduct targeted studies of *CYC* and related genes and thereby obtain a better understanding of the role of these genes in the evolution of flower symmetry.

EVOLUTION

Conflict of interest statement: No conflicts declared.

Abbreviation: PCF, proliferating cell factor.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. DQ462258–DQ462275).

[†]To whom correspondence may be addressed. E-mail: dianella.howarth@yale.edu or michael.donoghue@yale.edu.

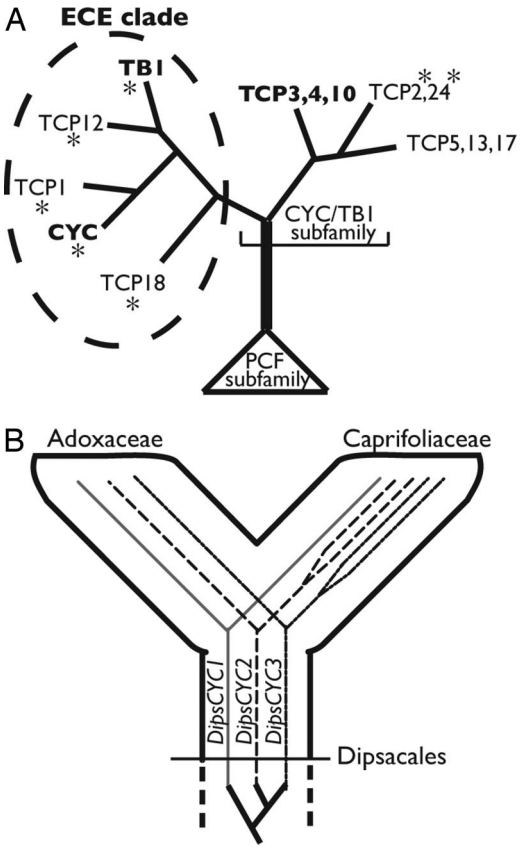

Fig. 1. Published phylogenetic relationships of *Arabidopsis* and Dipsacales TCP genes. (*A*) Phylogeny of *Arabidopsis* TCP genes, modified from ref. 8, showing the major split between the CYC/TB1 subfamily and the PCF subfamily. *, Taxa with R domain. The ECE clade is enclosed in a dashed line. (*B*) Relationships of Dipsacales sequences modified from ref. 27. The major duplications are shown by using lines within a tube, representing the known species tree. Locating the phylogenetic position of the duplications leading to the three major copies (dashed lines) is the aim of this study.

Results

Phylogenetic Analyses. The aligned matrix of the TCP, ECE, and R regions included 225 nucleotides (TCP, 150 nt; ECE, 45 nt; R, 30 nt). The dataset is deposited in TreeBASE (www.treebase.org). By using the sole copy obtained from *Aquilegia* as an outgroup, three major branches within the ECE clade were supported by Bayesian posterior probabilities of CYC1 (99), CYC2 (90), and CYC3 (98) (Fig. 2). These clades corresponded to the three clades of Dipsacales sequences (Fig. 1*B*). Additionally, sequences from all three lineages are also published from *Antirrhinum*, *Arabidopsis*, and *Solanum*. Only two of the copies have been published so far in the other sampled core eudicots.

Additional data that we have generated from outside the core eudicots indicate that the duplications leading to the CYC1, CYC2, and CYC3 clades happened within the eudicot lineage. We found only a single ECE gene copy in monocots (*Zea*, *Pontederia*, and *Smilax*), magnoliids (*Calycanthus* and *Idiospermum*), and ranunculids (*Aquilegia*). Additionally, searches of the *Oryza* database at http://riceblast.dna.affrc.go.jp/using TBLASTN yielded 13 genes in the CYC/TB1 subfamily, of which only *OsTB1* falls within the ECE clade. The two most similar copies in the TCP domain to the ECE clade lack both the R domain and the ECE region (AP003908 and AP005093). Bayesian analyses using only the TCP region of members of the ECE clade (other sequence data were unalignable because of broad and dispersed taxon sampling) show that the single gene copies in monocots, magnoliids, and *Aquilegia* form a grade that subtends the three core eudicot ECE clades: CYC1, CYC2, and CYC3, with CYC1 sister to a clade containing CYC2 and CYC3 (data not shown), although support is weak because of minimal sequence data. Nonetheless, these data support the hypothesis that the duplications leading to these three gene lineages occur after the divergence of ranunculids from the rest of the eudicots (Fig. 2). The analyses we show in this article use data only from eudicots, rooting with the single copy of the ranunculid *Aquilegia*.

Our data show that CYC1 is sister to the clade containing CYC2 and CYC3. Also, CYC1 is the most different in sequence from the other two copies, with roughly 8% more sequence divergence in the TCP domain (comparing either CYC2 or CYC3 with CYC1 vs. comparing CYC2 and CYC3). It is the only clade that does not seem to contain additional duplications in any of the lineages. However, because this clade has the greatest sequence divergence, amplification and sampling were also poorest, and designing specific clade primers was difficult.

CYC2 contains all published *DipsCYC2* sequences and sequences from other asterids (including *CYC* and *DICH* from *Antirrhinum*), rosids (including TCP1 from *Arabidopsis*), caryophyllids, and Saxifragales. CYC2 has the greatest number of within-clade duplications, with multiple independently derived copies in all major groups except *Arabidopsis*, *Plantago*, *Solanum*, and Adoxaceae (each with polysymmetric flowers).

CYC3 contains all published *DipsCYC3* sequences and sequences from other asterids, rosids (including TCP12 from *Arabidopsis*), and caryophyllids. Sampling is poorer in this group than in the heavily studied CYC2 clade; nevertheless, several similar duplication patterns emerged. Specifically, our CYC3 data from Dipsacales and *Scaevola* show the same duplication patterns as those found in CYC2. *Antirrhinum*, however, contains only one published sequence in this clade.

Gene Expression. RNA extracted from whole buds indicates that *DipsCYC* copies from CYC2 and CYC3 are expressed in flower tissue in *Lonicera morrowii* (Fig. 3), whereas *DipsCYC1* (CYC1 clade) is not expressed at this stage of floral development. Therefore, CYC2 (including *DipsCYC2B*) and CYC3 (including *DipsCYC3B*) copies are found late in growth, persisting well beyond the initial patterning of the floral meristem. *DipsCYC2B* is expressed in dorsal petals and not in the ventral petal (Fig. 3). *DipsCYC3B*, alternatively, seems to be expressed in the ventral and the dorsal petals, although it seems to be more strongly expressed in the ventral petal (Fig. 3).

Discussion

The ECE Clade. We have analyzed relationships among known eudicot sequences belonging to the CYC/TB1 clade (Fig. 1*A*). Analysis of our expanded dataset demonstrates that, in contrast to the arrangement shown in Fig. 1*A*, all of the core eudicot sequences form a clade, which is composed of TCP1, -12, and -18 from *Arabidopsis* (8) and their orthologs in other organisms. Because many of these sequences include the amino acid region glutamic acid–cysteine–glutamic acid, we refer to this as the ECE clade (see ref. 27). Within the ECE gene tree, there are three major subclades within core eudicots, which we call CYC1, CYC2, and CYC3. CYC2 has been sampled most thoroughly because it contains *CYCLOIDEA* and its orthologs (Fig. 2).

By comparison to the "known" angiosperm phylogeny (34), the three major ECE clades seem to have arisen in a series of duplication events. Each clade contains asterids and rosids, and several of the better sampled taxa (*Antirrhinum*, *Arabidopsis*, *Solanum*, and most Dipsacales) have gene copies belonging to all three ECE clades. Several lines of evidence indicate that the first duplication led to CYC1 and the ancestor of CYC2 and CYC3. Additional data (data not shown) indicate that there is only a

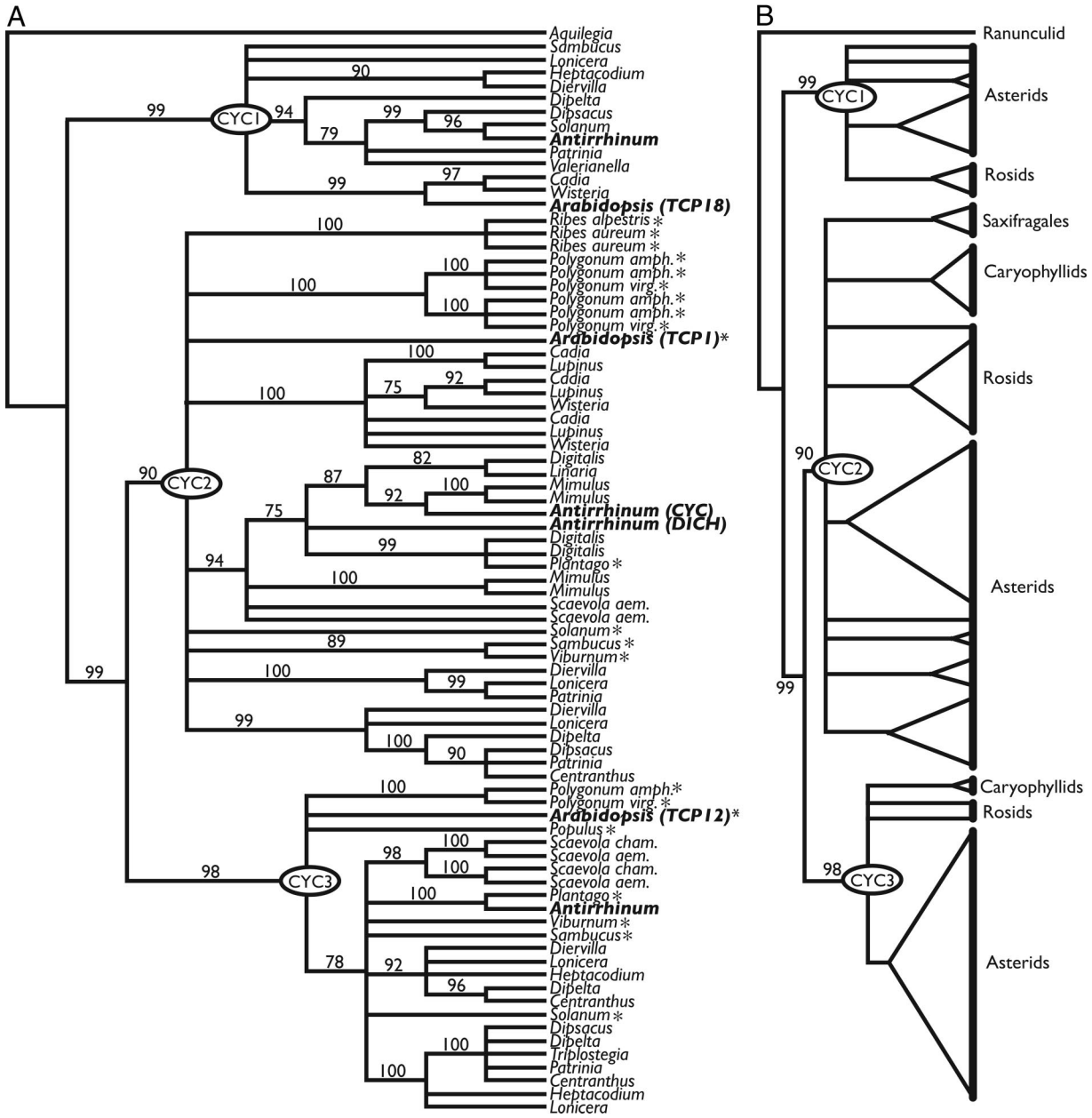

Fig. 2. Phylogeny of *CYC*-like genes using Bayesian analysis. Clades with >75% posterior probability are displayed with numbers above the lines. (*A*) Phylogeny with all taxa sampled. (*B*) Schematic showing just the major groups. The placements of previously named genes from *Arabidopsis* and *Antirrhinum* are labeled; and all included copies from these two model organisms are shown in bold. *, Polysymmetric species in CYC2 and CYC3. Species names are given in Table 1. virg., *virginianum*; amph., *amphibium*; cham., *chamissoniana*; aem., *aemula*.

single ECE gene in monocots, magnoliids, and *Aquilegia*. Our rooting with *Aquilegia* places the first split between CYC1 and CYC2 plus CYC3. Also, members of the CYC1 clade differ the most from the other two copies in sequence divergence, and, unlike CYC2 and CYC3, no duplications have been found within CYC1, here or in previous studies (27, 30, 31, 35).

CYC2 contains *CYCLOIDEA* and *DICHOTOMA* from *Antirrhinum* and has been the target of the most directed study. Expression data in *Antirrhinum* and *Arabidopsis* (TCP1) show that these genes are expressed only dorsally (15, 20, 21). We show here that this pattern is also found in *Lonicera*, with *DipsCYC2B* (CYC2 clade) expressed dorsally. Additionally, double mutants in *Antirrhinum* of *CYC* and *DICH* yield a peloric, ventralized flower (15, 17). Even though *Arabidopsis* flowers seem at maturity to be polysymmetric, TCP1 is expressed only very early in floral buds (21), and the flowers are initially monosymmetric (36).

The CYC2 clade contains the greatest number of additional duplications (7, 22, 28–33). Within the rosid and asterid lineages, the monosymmetric groups have at least two copies of CYC2 genes, whereas the polysymmetric groups (*Arabidopsis*, *Plantago*, *Sambucus/Viburnum*, and *Solanum*) contain only one (Fig. 2). In addition to rosids and asterids, this clade also contains *Ribes* (Saxifragales) and *Polygonum* (Caryophillids). These later groups contain duplications in the CYC2 lineage, despite having polysymmetric flowers. At least in the case of *Polygonum*, however, these gene duplications are likely due to recent polyploidy (37).

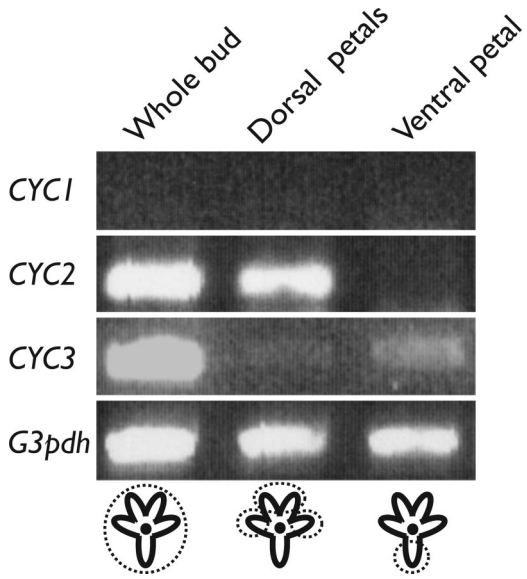

Fig. 3. Image of agarose gel electrophoresis of cDNA from *L. morrowii* flowers. CYC1 indicates expression of *DipsCYC1*, CYC2 indicates expression of *DipsCYC2B*, and CYC3 indicates expression of *DipsCYC3B*. *G3pdh* is included as a control. Bird's-eye view of flower is included showing the differentiation between the four dorsal petals and the single ventral petal. Dotted circles indicate the portions of the flower used in each RNA extraction.

CYC3 contains a separate ECE-type copy from *Antirrhinum* as well as TCP12 from *Arabidopsis*. Until this study, expression data had not been obtained for any members of this clade. Duplications within the CYC3 clade are similar to those within CYC2. For instance, in Dipsacales, it seems that duplications in CYC2 and CYC3 occur in the same places in the phylogeny of Caprifoliaceae, and additional parallel duplications occur within Morinaceae (27). Our data also indicate a duplication in *Scaevola* (Asterales) in both CYC2 and CYC3. Other groups have not yet been sampled as thoroughly, so it is unclear whether such mirrored duplications are common. In addition to rosids and asterids, the CYC3 clade contains sequences from *Polygonum* (Caryophillids). *Polygonum* does not seem to be duplicated in this gene clade; however, we may have failed to amplify all copies in this case.

Importance of Other Copies. A few CYC studies have sequenced outlier sequences, which did not clearly fit in the CYC2 clade (7, 30, 31). Our analyses clarify the position of these outlier (non-CYC2) copies within the ECE gene tree (Fig. 2). This phylogeny should foster more targeted studies of these copies and what role they may play in flower development.

Our data indicate that the single ECE copy found in monocots, magnoliids, and ranunculids duplicated into three separate copies near the divergence of the core eudicots. TB1, the single copy from *Zea*, is expressed across the entire floral meristem, differing from the known CYC2 expression in *Antirrhinum*. Yet, as with CYC2, TB1 expression is correlated with suppression in stamen primordia (38, 39). TB1, however, is most similar to genes in the CYC1 clade, expression of which we have been unable to detect in *Lonicera* flower buds. The CYC3 clade was unrecognized until this study, with the expression patterns and function of CYC3 genes unknown. Our preliminary data from RT-PCR of *DipsCYC3B* (CYC3) in *Lonicera* show that it is also expressed in flowers and, within the petal primordia, is in both the four dorsal petals and the lower single ventral petal (Fig. 3). Differing from the dorsal expression of members of the CYC2 clade, we have seen consistently more amplification of *DipsCYC3B* in the ventral petal as compared with the dorsal petals at this time point in bud development. Given the preliminary expression pattern of *DipsCYC3B* in *Lonicera* and the close sister relationship between CYC3 and CYC2, genes from the CYC3 clade warrant comparative study for floral patterning and inclusion in studies of floral symmetry pathways.

Basic helix–loop–helix (bHLH) transcription factors, such as the TCP genes, usually function as multimers (40), often binding with other bHLH proteins. In *Oryza*, there is evidence that genes in the CYC/TB1 subfamily form homo- and heterodimers with each other (41). Owing to dosage effects, duplications of interacting gene partners often are maintained together (40, 42, 43). These observations may relate to our finding that duplications in CYC2 in various groups are mirrored by duplications in CYC3. It is possible that members of CYC2 and CYC3 interact in areas of coexpression, and members of CYC3 certainly bear close attention from a functional standpoint.

Comparison with MADS-box Genes. Our analyses show that the duplications leading to the three major ECE gene clades that we found previously in Dipsacales (27) took place deep within angiosperm phylogeny, after the origin of the eudicots but before the evolution of the core eudicots. It is possible that these major duplication events played an important role in connection with the major changes in flower form that mark the core eudicot clade. Most importantly, perhaps, flowers of the core eudicots (excluding Gunnerales) are based on a pentamerous ground plan, as opposed to the dimerous ground plan that characterizes the basal eudicots (44, 45). The switch to pentamery entails the establishment of differentiation of dorsal and ventral portions of the flower in relation to the axis on which they are borne.

Our findings on *CYC* evolution provide a remarkable complement to recent studies of MADS-box genes (Fig. 4). It seems that members from each of the major functional categories of the ABC model of floral MADS-box genes [*APETALA1* (*AP1*, A class), *APETALA3* (*AP3*, B class), *AGAMOUS* (*AG*, C class), and *SEPALLATA*] underwent a duplication in a similar location near the base of the core eudicots (46–50). Our data from the ECE clade, a separate major family of transcription factors involved in floral development, show the same pattern (Fig. 4).

Unfortunately, taxon sampling is currently too limited in all of these cases to pinpoint whether these duplications really did occur at the same point in eudicot phylogeny. However, it is worth considering the possibility that these events were tightly correlated with one another and with the major changes in floral organization that occurred during this key period in angiosperm evolution. It is possible that there was a genome-wide duplication at this point, which may have provided the opportunity for functional specialization of the resulting gene copies and therefore for major changes in flower morphology. Based on analyses of gene order (51) and synonymous substitution rate (52), a genome duplication event has been hypothesized before the evolution of the clade including rosids and asterids, but after the split from monocots. Only when we achieve a more complete and comparable sampling will it be possible to evaluate the extent to which these gene duplications actually corresponded in time and fueled a major transition in floral form.

Materials and Methods

Plant Material. Sequences from 30 individuals were used in all analyses, consisting of 12 taxa from our previous work in Dipsacales; 11 published taxa from multiple asterid and rosid clades; and sequences obtained for this study (from extracted total genomic DNAs) from 7 taxa representing the ranunculids, Saxifragales, caryophyllids, and Asterales. Multiple copies found in many of these species resulted in a total of 82 separate sequences. (Table 1, which is published as supporting informa-

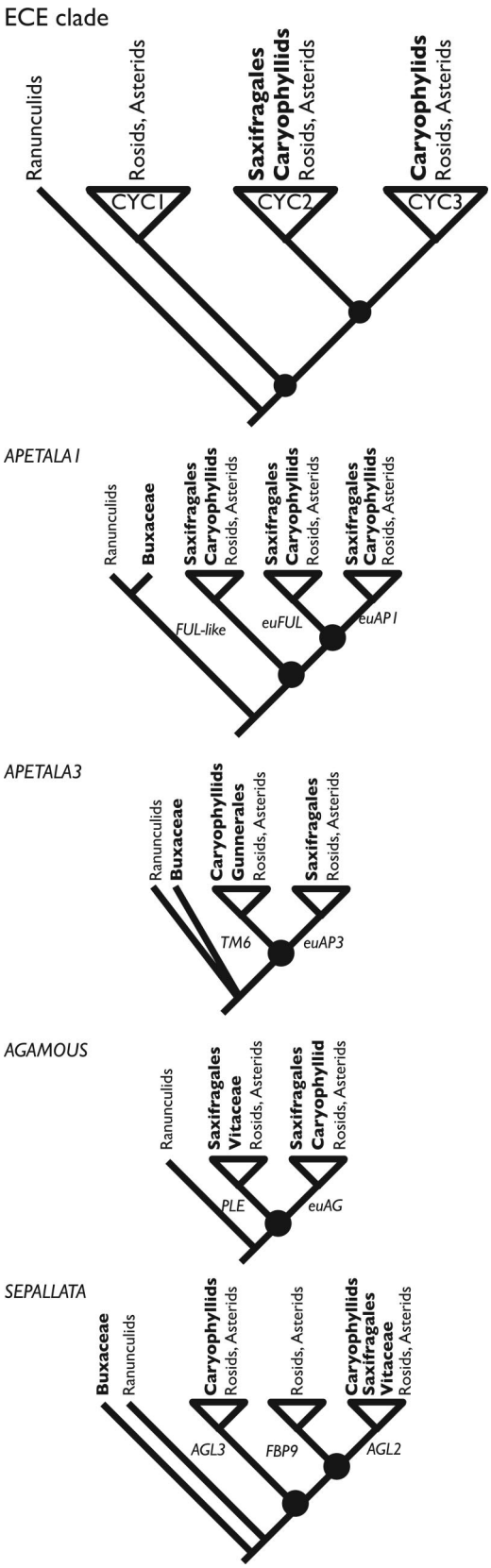

Fig. 4. Phylogenies of floral MADS-box genes and the ECE clade, comparing the phylogenetic location of duplications. Each hypothesized duplication is indicated by a black dot. Non-rosid or -asterid eudicots are shown in bold. Trees of *APETALA1*, *APETALA3*, *AGAMOUS*, and *SEPALLATA* are modified from refs. 47–50, respectively.

tion on the PNAS web site, provides a list of included taxa and GenBank accession numbers.)

Primer Design and Amplification from Additional Taxa. All primers were designed in the TCP domain (forward primer) and the R domain (reverse primer). Primers were designed from published and our amplified TCP genes (see ref. 27 for a full list). Multiple primer pairs were used for each taxon. PCR and cloning were performed as described (27). Between 10 and 40 (depending on cloning success) colonies were screened for all potentially different copies or alleles of *CYC*-like genes. Cloned products of the appropriate size (200–800 bp) were sequenced.

Phylogenetic Alignment and Analyses. All clones from each DNA extraction (obtained from multiple primer pairs) were compiled in SEQUENCHER 4.2 (Gene Codes, 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 from other clones, and a consensus sequence was generated for each type. Protein sequences for each potential copy/allele obtained were aligned by eye in MACCLADE 4 (53) to the other published taxa. Nucleotide sequence data from the three alignable regions TCP, ECE, and R were used in subsequent analyses.

Parameters for the Bayesian analyses were estimated by using MODELTEST 3.06 (54). The Akaike Information Criterion (55) recommended a general time reversible (GTR) model with added parameters for invariable sites and a γ distribution (GTR + I + Γ), for the TCP region alone, the non-TCP region, and the entire sequence. Bayesian analyses were conducted by using this model for the entire matrix. We used the Metropolis-coupled Markov Chain Monte Carlo (MCMCMC) method as implemented in MRBAYES 3.0B4 (56) to run four chains (3 heated). We ran 5 million generations, sampling every 1,000 generations. The trees were analyzed in TRACER 1.0.1 (http://evolve.zoo.ox.ac.uk/software.html?id=tracer) to determine a burn-in of 1 million trees (1,000 sampled trees). A majority rule consensus was calculated from the remaining trees by using PAUP* 4.0B10 (57) to determine posterior probabilities.

RT-PCR. Total RNA was extracted from middle-stage flower buds (after petal edges had been defined but before bud opening) and dissected corolla lobes of *L. morrowii* Gray [voucher housed in the Yale University Herbarium (YU)], a species where there is differentiation between the four dorsal corolla lobes and a single ventral lobe. Frozen tissue was lysed in a FastPrep Instrument (QBiogene, Irvine, CA) and extracted by using the Qiagen (Valencia, CA) RNeasy kit, using the optional DNase step. cDNA was reverse transcribed with SuperScript III by using the manufacturer's instructions with random hexamers (Invitrogen). Specific primers were used to amplify each of the three copies as follows: *DipsCYC1* [forward (F), 5′-ACCAGAGGCCTYAACTCCAACC-3′; reverse (R), 5′-GCGTTAGCATCRAATGCGATTCTCC-3′]; *DipsCYC2B* (F, 5′-GATGAAAATCAACTGCACTACTGG-3′; R, 5′-AGCATCCCTCTTCTCGTTCCCAAC-3′); and *DipsCYC3B* (F, 5′-TTGAGRGCYAGGAGGATGAGATTAC-3′; R, 5′-ACTCCCTCGCCTTTCCCAATTCTC-3′). Each copy was amplified separately from cDNA generated from whole bud, all four dorsal petal lobes, and the single ventral petal lobe. Additionally, a portion of *G3pdh*, which spans 5 exons, was amplified as a positive control. All of the above reactions included initial stock RNA as a negative control to rule out DNA contamination. Each band was confirmed to be from a specific copy through direct sequencing.

We thank the following for gifts of DNA: E. Edwards (University of California, Santa Barbara; *Pereskia* DNA), E. Kramer (Harvard University, Cambridge, MA; *Aquilegia* DNA), S.-T. Kim (Yale Uni-

versity, New Haven, CT; *Polygonum* DNA), S. Mathews (Arnold Arboretum of Harvard University, Cambridge, MA; non-eudicots DNA), and L. Schultheis (Foothill College, Los Altos Hills, CA; *Ribes* DNA). We thank A. Litt for sharing RNA expertise; L. Hileman (University of Kansas, Lawrence) for kindly providing aliquots of initial primer sequences; P. Soltis and E. Kramer for their gracious help with earlier versions of this manuscript; and M. Dunn, E. Edwards, and members of the M.J.D. laboratory for helpful discussions. D.G.H. is grateful for a Forest B. H. and Elizabeth D. W. Brown Postdoctoral Fellowship from Yale University and for a travel grant from the Molecular and Organismic Research in Plant History (MORPH) Research Coordination Network. The work of M.J.D. on angiosperm phylogeny is supported by a Tree of Life grant from the National Science Foundation.

1. Coen, E. S. & Meyerowitz, E. M. (1991) *Nature* **353,** 31–37.
2. Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E. & Yanofsky, M. F. (2000) *Nature* **405,** 200–203.
3. Honma, T. & Goto, K. (2001) *Nature* **409,** 525–529.
4. Endress, P. K. (1999) *Int. J. Plant Sci.* **160,** S3–S23.
5. Endress, P. K. (1996) *Diversity and Evolutionary Biology of Tropical Flowers* (Cambridge Univ. Press, Cambridge, U.K.).
6. Weberling, F. (1989) *Morphology of Flowers and Inflorescences* (Cambridge Univ. Press, Cambridge, U.K.).
7. Reeves, P. A. & Olmstead, R. G. (2003) *Mol. Biol. Evol.* **20,** 1997–2009.
8. Cubas, P. (2002) in *Developmental Genetics and Plant Evolution*, eds. Cronk, Q. C. B., Bateman, R. & Hawkins, J. (Taylor and Francis, London), pp. 247–266.
9. Gillies, A. C. M., Cubas, P., Coen, E. S. & Abbott, R. J. (2002) in *Developmental Genetics and Plant Evolution*, eds. Cronk, Q. C. B., Bateman, R. M. & Hawkins, J. A. (Taylor & Francis, London), pp. 233–246.
10. Knapp, S. (2002) in *Developmental Genetics and Plant Evolution*, eds. Cronk, Q. C. B., Bateman, R. M. & Hawkins, J. A. (Taylor & Francis, London), pp. 267–297.
11. Ree, R. H. & Donoghue, M. J. (1999) *Syst. Biol.* **48,** 633–641.
12. Donoghue, M. J., Ree, R. H. & Baum, D. A. (1998) *Trends Plant Sci.* **3,** 311–317.
13. Neal, P. R. (1998) *Annu. Rev. Ecol. Syst.* **29,** 345–373.
14. Sargent, R. D. (2004) *Proc. R. Soc. London Ser. B* **271,** 603–608.
15. Luo, D., Carpenter, R., Vincent, C., Copsey, L. & Coen, E. (1996) *Nature* **383,** 794–799.
16. Galego, L. & Almeida, J. (2002) *Genes Dev.* **16,** 880–891.
17. Almeida, J., Rocheta, M. & Galego, L. (1997) *Development (Cambridge, U.K.)* **124,** 1387–1392.
18. Corley, S. B., Carpenter, R., Copsey, L. & Coen, E. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 5068–5073.
19. Costa, M. M. R., Fox, S., Hanna, A. I., Baxter, C. & Coen, E. (2005) *Development (Cambridge, U.K.)* **132,** 5093–5101.
20. Luo, D., Carpenter, R., Copsey, L., Vincent, C., Clark, J. & Coen, E. (1999) *Cell* **99,** 367–376.
21. Cubas, P., Coen, E. & Zapater, J. M. M. (2001) *Curr. Biol.* **11,** 1050–1052.
22. Hileman, L. C. & Baum, D. A. (2003) *Mol. Biol. Evol.* **20,** 591–600.
23. Oyama, R. K. & Baum, D. A. (2004) *Am. J. Bot.* **91,** 918–925.
24. Hileman, L. C., Kramer, E. M. & Baum, D. A. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 12814–12819.
25. Damerval, C. & Manuel, M. (2003) *C. R. Palevol.* **2,** 241–250.
26. Cubas, P., Lauter, N., Doebley, J. & Coen, E. (1999) *Plant J.* **18,** 215–222.
27. Howarth, D. G. & Donoghue, M. J. (2005) *Int. J. Plant Sci.* **166,** 357–370.
28. Smith, J. F., Hileman, L. C., Powell, M. P. & Baum, D. A. (2004) *Mol. Phylogenet. Evol.* **31,** 765–779.
29. Citerne, H. L., Möller, M. & Cronk, Q. C. B. (2000) *Ann. Bot.* **86,** 167–176.
30. Citerne, H. L., Luo, D., Pennington, R. T., Coen, E. & Cronk, Q. C. B. (2003) *Plant Physiol.* **131,** 1042–1053.
31. Fukuda, T., Yokoyama, J. & Maki, M. (2003) *J. Mol. Evol.* **57,** 588–597.
32. Ree, R. H., Citerne, H. L., Lavin, M. & Cronk, Q. C. B. (2004) *Mol. Biol. Evol.* **21,** 321–331.
33. Gübitz, T., Caldwell, A. & Hudson, A. (2003) *Mol. Biol. Evol.* **20,** 1537–1544.
34. Soltis, D. E., Soltis, P. S., Endress, P. K. & Chase, M. W. (2005) *Phylogeny and Evolution of Angiosperms* (Sinauer, Sunderland, MA).
35. Lukens, L. & Doebley, J. (2001) *Mol. Biol. Evol.* **18,** 627–638.
36. Endress, P. K. (1992) *Int. J. Plant Sci.* **153,** S106–S122.
37. Jaretzky, R. (1928) *Jahrb. Wiss. Bot.* **69,** 357–490.
38. Doebley, J., Stec, A. & Hubbard, L. (1997) *Nature* **386,** 485–488.
39. Hubbard, L., McSteen, P., Doebley, J. & Hake, S. (2002) *Genetics* **162,** 1927–1935.
40. Amoutzias, G. D., Robertson, D. L., Oliver, S. G. & Bornberg-Bauer, E. (2004) *EMBO Reports* **5,** 1–6.
41. Kosugi, S. & Ohashi, Y. (2002) *Plant J.* **30,** 337–348.
42. Papp, B., Pál, C. & Hurst, L. D. (2003) *Nature* **424,** 194–197.
43. Birchler, J. A., Bhadra, U., Pal Bhadra, M. & Auger, D. (2001) *Dev. Biol.* **234,** 275–288.
44. Magallon, S., Crane, P. R. & Herendeen, P. S. (1999) *Ann. Mo. Bot. Gard.* **86,** 297–372.
45. Soltis, D. E., Senters, A. E., Zanis, M. J., Kim, S., Thompson, J. D., Soltis, P. S., Ronse De Craene, L. P., Endress, P. K. & Farris, J. S. (2003) *Am. J. Bot.* **90,** 461–470.
46. Kramer, E. M. & Hall, J. C. (2005) *Curr. Opin. Plant Biol.* **8,** 13–18.
47. Litt, A. & Irish, V. F. (2003) *Genetics* **165,** 821–833.
48. Kim, S., Yoo, M.-J., Albert, V. A., Farris, J. S., Soltis, P. S. & Soltis, D. E. (2004) *Am. J. Bot.* **91,** 2102–2118.
49. Kramer, E. M., Jaramillo, M. A. & Di Stilio, V. S. (2004) *Genetics* **166,** 1011–1023.
50. Zahn, L. M., Kong, H., Leebens-Mack, J. H., Kim, S., Soltis, P. S., Landherr, L. L., Soltis, D. E., dePamphilis, C. W. & Ma, H. (2005) *Genetics* **169,** 2209–2223.
51. Bowers, J. E., Chapman, B. A., Rong, J. & Paterson, A. H. (2003) *Nature* **422,** 433–438.
52. De Bodt, S., Maere, S. & Van de Peer, Y. (2005) *Trends Ecol. Evol.* **20,** 591–597.
53. Maddison, D. R. & Maddison, W. P. (2003) (Sinauer, Sunderland, MA).
54. Posada, D. & Crandall, K. A. (1998) *Bioinformatics* **14,** 817–818.
55. Akaike, H. (1973) in *Second International Symposium on Information Theory*, eds. Petrov, B. N. & Csaki, F. (Akademiai Kiado, Budapest), pp. 267–281.
56. Huelsenbeck, J. P. & Ronquist, F. (2001) *Bioinformatics* **17,** 754–755.
57. Swofford, D. L. (2001) PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods) (Sinauer, Sunderland, MA), Version 4.0B10.