Rapid Molecular Evolution of *CYCLOIDEA*-like Genes in Antirrhinum and Its Relatives

Thomas Gübitz, Ailsa Caldwell, and Andrew Hudson

Institute of Cell and Molecular Biology, University of Edinburgh

The *CYCLOIDEA* (*CYC*) and *DICHOTOMA* (*DICH*) genes encode related TCP transcription factors that control floral asymmetry in *Antirrhinum majus*. Analysis of sequences from relatives of Antirrhinum suggested that *CYC* and *DICH* arose from a gene duplication in an ancestor of the tribe Antirrhineae and have subsequently evolved at similar rates. Coding regions outside the conserved functional TCP and R domains differed by numerous indels, suggesting rapid evolution and low constraint on amino acid sequence. An analysis of variability within the genus Antirrhinum revealed very similar *CYC* alleles in 17 representative species, consistent with most of the species having diverged within the last 1 myr. Whereas substitution mutations appear to have accumulated constantly, one Antirrhinum *CYC* allele provided evidence for sporadic and rapid accumulation of insertion mutations.

Introduction

Antirrhinum majus (tribe Antirrhineae, family Veronicaceae *sensu* Olmstead et al. 2001), produces dorsiventrally asymmetric (zygomorphic) flowers. Floral asymmetry in *A. majus* is dependent on two closely related genes, *CYCLOIDEA* (*CYC*) and *DICHOTOMA* (*DICH*), that are expressed dorsally (adaxially) in floral meristems and developing floral organs (Carpenter and Coen 1990; Luo et al. 1996; Almeida, Rocheta, and Galego 1997; Luo et al. 1999). *CYC* and *DICH* have partially redundant roles in promoting dorsal identity in the floral meristem, corolla, and androecium, and thus zygomorphy of the flower as a whole. Almost all genera within the Antirrhineae and their more distant relatives in the Veronicaceae have zygomorphic flowers. Reduced activity of a *CYC*-like gene in *Linaria vulgaris* (a zygomorphic member of the Antirrhineae) can cause loss of floral asymmetry, indicating conserved function of *CYC*-like genes within this tribe (Cubas, Vincent, and Coen 1999). *CYC*-like genes in the distantly related eudicot, *Arabidopsis thaliana*, showed dorsally restricted expression in floral meristems that develop into radially symmetric (actinomorphic) flowers (Cubas et al. 2001). This suggests that an ancestral *CYC*-like gene was expressed asymmetrically and recruited to control asymmetric flower morphology in Veronicaceae and related families in the Lamiales with asymmetric flowers (Reeves and Olmstead 1998; Cubas, Coen, and Martínez Zapater 2001).

Other functionally characterized *CYC*-related genes include *teosinte branched1* (*tb1*), which represses growth of axillary shoots and floral organs in maize, and the *PCF1* and *PCF2* genes of rice (Doebley, Stec, and Hubbard 1997; Kosugi and Ohashi 1997). All these belong to a larger gene family named TCP (from *tb1*, *CYC*, and *PCF*). This family comprises 25 members in *A. thaliana*, is specific to plants, and encodes noncanonical basic Helix-Loop-Helix (bHLH) transcription factors (Cubas et al. 1999; Riechmann et al. 2000). All *CYC*-like genes share the TCP box, encoding a 55–59 amino acid bHLH motif that is implicated in protein dimerization and DNA binding and an R box encoding a more C-terminal motif of 18 amino acids. Outside these boxes, protein coding sequences (which in *CYC*, *DICH*, and *tb1* are uninterrupted by introns) show low sequence conservation, except in a region encoding a C-terminal domain of unknown function that has been termed the ''end box'' (Cubas et al. 1999; Möller et al. 1999; Citerne, Möller, and Cronk 2000). Because some TCP genes have known roles in floral asymmetry, and because of the potential importance of other members of the family in morphological evolution, they have been the subject of phylogenetic analysis in other genera with zygomorphic flowers (e.g., Gesneriaceae: Möller et al. 1999; Citerne, Möller, and Cronk 2000; and grasses: Lukens and Doebley 2001).

A previous study in Antirrhineae claimed to have identified at least five *CYC*-like loci in Antirrhinum and Misopates species, mainly on the basis of sequence phylogenies (Vieira, Vieira, and Charlesworth 1999). Paralogs and each putative ortholog appeared to be very highly conserved within the genus Antirrhinum, between genera within the Antirrhineae (Antirrhinum, Misopates, Linaria, and Cymbalaria), and with *Digitalis purpurea*, a more distant member of the Veronicaceae (Olmstead et al. 2001). This finding suggested that *CYC*-like sequences have evolved at an unusually slow rate in Veronicaceae.

To investigate further the evolutionary dynamics of *CYC*-like genes, we obtained *CYC* sequences from 17 Antirrhinum species and sequences of *CYC* and *DICH* loci from relatives of Antirrhinum. In contrast to previous studies, we found that *CYC* represents a single locus in Antirrhinum, that *DICH* is its closest paralog, and that the two loci have resulted from a gene duplication in the Veronicaceae in the lineage leading to the tribe Antirrhineae. Moreover we show that both *CYC* and *DICH* genes have evolved quickly within Veronicaceae, mainly as a result of nucleotide insertions and deletions in regions outside the functionally important TCP and R boxes. Antirrhinum species, with one exception, share similar *CYC* alleles, suggesting that the genus has undergone a recent radiation within the last ~5 myr.

Materials and Methods

Genomic DNA was extracted according to the method of Doyle and Doyle (1987). With the exception

Key words: *CYCLOIDEA*, *DICHOTOMA*, TCP genes, Antirrhinum, Antirrhineae.

E-mail: andrew.hudson@ed.ac.uk.

Mol. Biol. Evol. 20(9):1537–1544. 2003
DOI: 10.1093/molbev/msg166
Molecular Biology and Evolution, Vol. 20, No. 9,

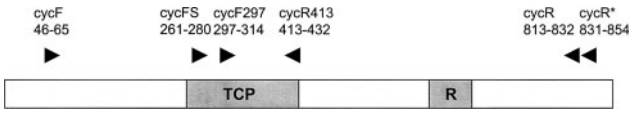

FIG. 1.—Structure of the *CYC* gene. The position and orientation of each oligonucleotide primer used in PCR and sequence analysis is indicated by an arrowhead and positions given in relation to the start codon of the *A. majus CYC* reference sequence (gi:2826867), where the A of the initial ATG(Met) codon is position 1. The TCP and R boxes are situated at positions 259–435 and 599–651 of the *A. majus* sequence, respectively.

of *Digitalis purpurea* and *Cymbalaria muralis*, which were collected from the wild in Edinburgh, all taxa were grown from seed accessions, details of which are available on request. *CYC* sequences were amplified from *D. purpurea*, *C. muralis*, *Chaenorhinum origanifolium*, *Linaria maroccana* with the primers cycFS (5′-atgctaggtttcgacaagcc-3′) and cycR* (5′-atgaatttgtgctgatccaaaatg-3′) and from *Misopates orontium*, *Sairocarpus nuttallianus*, and all Antirrhinum species with primers cycF (5′-tcccttcactctcgcgc-3′) and cycR (5′-tggcgcatagctggttcgac-3′; fig. 1). *DICH* sequences were amplified with primers dichF (5′-acgacgtgatttccgggtac-3′) and dichR (5′-acttcccccacgttctgcttca-3′) or with the primer combination dichFT (5′-ttccaccttcacttaatc-3′) and cycR or cycR*, neither of which is specific for *CYC* orthologs. No candidate *DICH* orthologs were obtained from *Ch. origanifolium* with any of these primer combinations, and further attempts to test whether *DICH* orthologs were present in this species were not pursued. Polymerase chain reaction (PCR) testing was carried out with *Taq* polymerase in buffer containing 2 mM Mg^{2+} at 94°C for 2 min; 30–35 cycles of 94°C for 20 s, annealing for 30 s, and extension at 72°C for 2 min, followed by a final extension at 72°C for 5 min. Different annealing temperatures were used for different primer combinations: 55°C for *trnL + F* primers, 60°C for cycF–cycR48, 48°C for cycFS–cycR*, and 55°C for dichF–dichR. In most cases, PCR products were cloned in pGEM-T Easy vector (Promega), and one clone from each accession was sequenced using T7 and SP6 primers. Sequences were then confirmed by direct sequencing of PCR products using the original amplification primers and the internal sequencing primers, cycF297 (5′-gccaagggacaggagagt-3′) and cycR413 (5′-ctttgatagcggttttcg-3′). The sequences of *CYC* from *M. orontium* and *A. charidemi*, and *DICH* from *A. siculum* were obtained directly from PCR products. *A. majus CYC* (gi:2826867) and *DICH* (gi:6466187) sequences and a *CYC*-like sequence from *L. vulgaris* (gi:5566284) were extracted from GenBank and confirmed by sequencing of the PCR product from the equivalent laboratory lines. Chloroplast sequence-based phylogenies were based on DNA sequences of the *trnL + F* region, consisting of regions of the (UAA) and (GAA) tRNA genes and the intergenic spacer separating them, using the universal primers f and c (Taberlet et al. 1991). *trnL + F* sequences for *Mohavea breviflora* (gi:19224319) and *Mimulus ringens* (gi:19224308) were obtained from GenBank.

Nucleotide sequences were aligned using the *T-Coffee* program (Notredame, Higgins, and Heringa 2000) at www.ch.embnet.org. Phylogenetic analyses on all three data sets (*trnL + F*, *CYC*, and *DICH* homologs from Antirrhineae and *CYC* from Antirrhinum) were performed using three different tree-building methods as implemented in PAUP* 4.b10 (Swofford 1998). The minimum evolution (ME) method (Kumar 1996) was used with Kimura two-parameter (K2P) distances (Kimura 1980). Maximum parsimony (MP) analysis was conducted with unordered and unweighted characters. Maximum likelihood (ML) analysis was performed with the Hasegawa, Kishino, and Yano (HKY) substitution model (Hasegawa, Kishino, and Yano 1985) with empirical base frequencies, ML, estimated transition-transversion ratio, and estimated gamma (among-site rate variation) correction. Heuristic searches were performed with starting trees obtained via random stepwise addition, tree-bisection-reconnection (TBR) branch-swapping, and the ''MulTrees'' option in effect. Support for nodes was assessed using the bootstrap method (Felsenstein 1985) with 1,000 replicates for ME, 500 for MP, or 250 for ML. The number of synonymous (Ks) and nonsynonymous substitution (Ka) rates was calculated with the program K-estimator (Comeron 1999) using the method of Comeron (1995).

For Southern hybridizations, 3 μg of genomic DNA was digested with *Hin*d III, fractionated in 0.7% agarose gels, and blotted onto a nylon membrane. Sequences were detected by hybridization to ^{32}P-labeled probes consisting of either the *CYC* or *DICH* sequences from *A. majus* amplified with primers cycF-cycR or dichF-dichR that had been cloned into pGEM-T. Hybridization was carried out in 2× SSC at 65°C (high stringency), 60°C (medium stringency), or 55°C (low stringency) for 18 hours. Filters were then washed in 0.2× SSC at the equivalent temperature and exposed to X-ray film for 3 days. Between hybridizations, filters were washed in 0.1% SDS at 100°C until no residual signal could be detected with a 3-day exposure to film.

Results

To provide a phylogenetic framework against which to assess the evolution of *CYC*-like genes, sequences were obtained for the *trnL + F* region from the chloroplast genomes of representative taxa. The resulting phylogeny (fig. 2) had essentially the same topology as one obtained using the more slowly evolving *ndhF* chloroplast sequence (Ghebrehiwet, Bremer, and Thulin 2000), except that it resolved Misopates as basal to a clade containing Antirrhinum and the New World genera Sairocarpus and Mohavea.

Oligonucleotide primers complementary to opposite ends of the *CYC* coding sequence from *A. majus* (cycF and cycR; fig. 1) were then used to amplify sequences from Antirrhinum species and their relatives, *S. nuttallianus*, *M. orontium*, and *Chaenorhinum origanifolium* (fig. 3). *DICH*-like genes were amplified from *A. majus*, *S. nuttallianus*, and *M. orontium* using primers (dichF and dichR) complementary to regions of the *A. majus DICH* sequence that were not conserved in *A. majus CYC*. Because these primers were unable to amplify *CYC* and *DICH* loci from Linaria species, *Ch. origanifolium*, *Cymbalaria muralis*, and *Digitalis purpurea*, primers

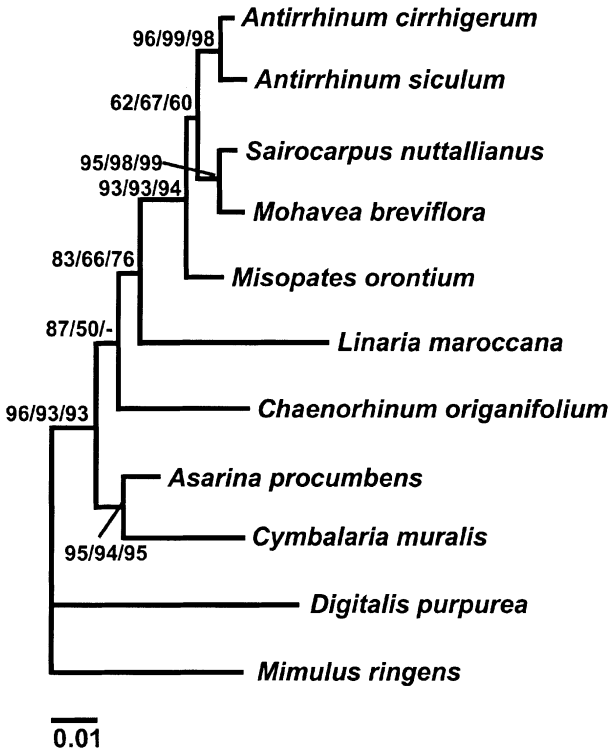

FIG. 2.—Minimum evolution phylogram for chloroplast *trnL + F* sequences of Antirrhineae. Sequences from *Mimulus ringens* and *Digitalis purpurea* are used as outgroups. Numbers at each node give bootstrap support values for ME/MP/ML analyses respectively; values lower than 50% are not shown or are represented by dashes. The scale bar indicates Kimura two-parameter distance.

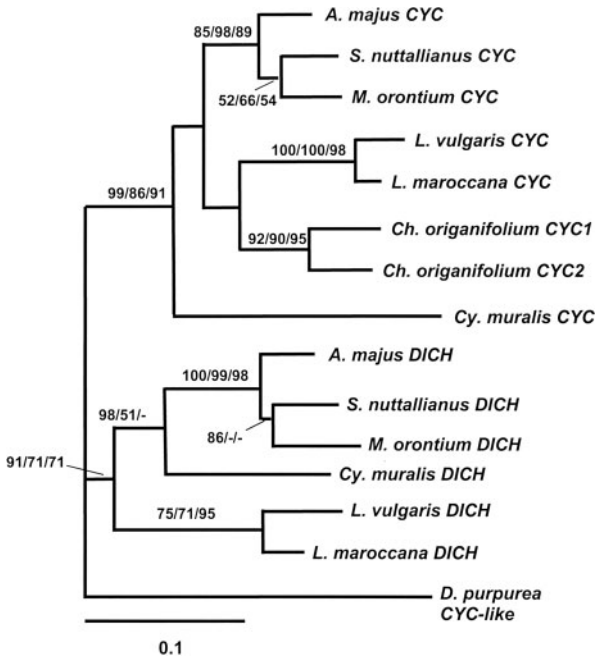

FIG. 3.—Minimum evolution phylogram of *CYC* and *DICH* sequences from Antirrhineae. Bootstrap support values are shown as in figure 1. The scale bar indicates Kimura two-parameter distance.

corresponding to the more highly conserved TCP box and end box (cycFS and cycR*) were used to obtain *CYC* and *DICH* sequences from these more distantly-related taxa. However, no *DICH*-like sequence was obtained from *Ch. origanifolium* with any of these primers. All sequences consisted of open reading frames without frameshifts. *CYC* sequences from the Antirrhineae (i.e., all species except *D. purpurea*) formed a monophyletic clade, as did the *DICH* sequences from the equivalent species with the Digitalis sequence as an outgroup to both (fig. 3).

All sequences were similar within the TCP and R boxes, with K2P distances ranging between 0.017 and 0.17. Outside these boxes, the sequences differed by numerous insertions and deletions. Indels made sequence alignment problematic, and therefore estimates of divergence distances, and substitution rates were dependent on the alignment chosen. However, alignments based only on the TCP and R boxes produced phylogenies with the same topologies and only slightly lower resolution than more complete sequences (data not shown). The K2P distances for these regions outside the TCP and R boxes were consistently higher (between 0.07 and 0.4) than for the TCP and R boxes. Relative rate tests (Takezaki, Rzhetzky, and Nei 1995) comparing *CYC* and *DICH* sequences from *A. majus*, *S. nuttallianus*, *M. orontium*, *L. vulgaris*, *L. maroccana*, and *Cy. muralis* in relation to the *CYC*-like sequence from *D. purpurea* suggested substitution rates that were not significantly different for the two loci.

To compare the evolutionary rate of *CYC* with another nuclear locus, an *A. majus* EST collection was searched for homologs of *D. purpurea* nuclear loci present in GenBank. Two Antirrhinum ESTs were identified as homologs of the *D. purpurea aldose reductase* gene (gi:13160396). Aligning the Antirrhinum and Digitalis sequences revealed no indels and K2P distances of 0.23 and 0.24 between the Digitalis sequence and each of the Antirrhinum ESTs. In comparison, the K2P distance between Antirrhinum *CYC* and Digitalis *CYC*-like was 0.28 with 15 indels (the number of indels depending on the alignment chosen), and between Antirrhinum *DICH* and Digitalis *CYC*-like the distance was 0.30 with 22 indels. In comparison, the K2P distance between *CYC* and *DICH* loci in Antirrhinum was 0.23 with 19 indels. These results suggest that *CYC* and *DICH* loci have evolved faster than the *aldose reductase* locus, largely as a result of indels outside the TCP and R boxes.

The ratio of nonsynonymous to synonymous substitution rates (K_a/K_s) within the TCP and R boxes ranged from 0.1 to 0.4. Outside these regions values were higher (between 0.1 and 0.6), but comparable to those obtained for TCP genes in grasses (Lukens and Doebley 2001). These K_a/K_s values did not provide evidence for positive selection and suggested that the variation within *CYC* and *DICH* sequences reflected relaxed evolutionary constraints.

The relationships within the *CYC* and *DICH* clades (fig. 3) differed slightly from each other and from the species phylogenies based on chloroplast DNA (fig. 2). Most notably, the clade containing Antirrhinum *DICH* was placed closer to the Cymbalaria sequence than to Linaria, whereas the Antirrhinum *CYC* and *trnL + F* sequences

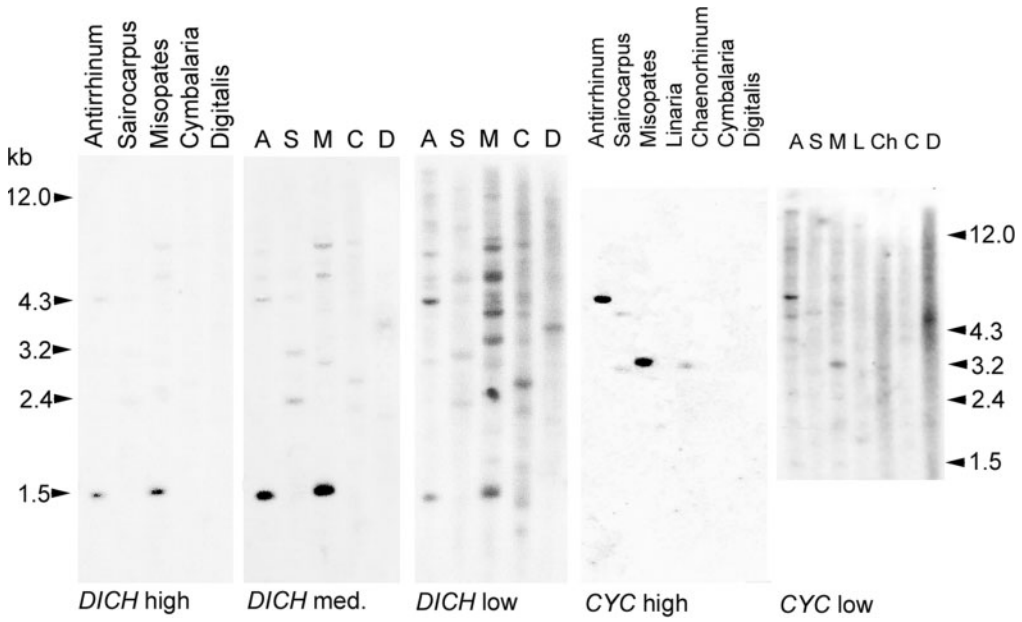

FIG. 4.—Detection of CYC and DICH sequences by Southern hybridization. Genomic DNA (3.5 ng) from *A. majus*, *Sairocarpus nuttallianus*, *Misopates orontium*, *Linaria vulgaris*, *Chaenorhinum origanifolium*, *Cymbalaria muralis*, or *Digitalis purpurea* was digested with *Hind* III. It was then hybridized with either the *A. majus* CYC or DICH sequences at high stringency (65°C), medium stringency (60°C), or low stringency (55°C). Sizes of relevant bands were estimated by interpolation from an internal size standard (BRL 1 kb ladder). All DICH hybridizations were carried out to the same filter, which was different from that used for CYC hybridizations.

appeared closer to Cymbalaria than to Linaria. Possible explanations for these differences is that they resulted from later gene duplications that gave rise to paralogs within each clade and that the sequences did not represent orthologous genes. Evidence of a later duplication was provided by the two *CYC*-like sequences obtained from *Ch. origanifolium*. These were sufficiently different from each other (K2P distance = 0.087) to represent distinct loci rather than alleles. They were, however, more similar to each other than to the sequence from any other taxon, suggesting that they had arisen from a recent duplication in the lineage leading to Chaenorhinum.

The number of *CYC* and *DICH* loci was tested further by Southern hybridization of genomic DNA. High stringency hybridization with the *A. majus DICH* sequence as a probe detected strongly hybridizing bands (of ~1.5 kb) in *A. majus* and *M. orontium*, but not in their more distant relatives, *Cy. muralis* and *D. purpurea* (fig. 4). This suggested that a single *DICH* locus was present in Antirrhinum and Misopates and that the sequences from these species were more similar to each other than to the *DICH*-like genes of other taxa. This relationship was consistent with the phylogeny of *DICH* sequences obtained from these species, except that *S. nuttallianus* produced a much weaker band than *M. orontium* in Southern hybridization, even though the *DICH* sequences from the two species were equally similar to the *A. majus* probe. One explanation for this discrepancy is that *S. nuttallianus* has a larger genome than *M. orontium*, so fewer *DICH* target sequences were present in equivalent amounts of DNA on Southern blots. Similarly, high stringency hybridization with the *A. majus CYC* probe detected strong bands in *A. majus* and *M. orontium* and weaker bands in *S. nuttallianus* and *Ch. origanifolium*. As for *DICH*, the relative strength of hybridization signals was consistent with the relationship between *CYC* sequences in these taxa, except that *S. nuttallianus* and *L. vulgaris* were weaker than expected. This discrepancy might again reflect relative differences in genome size.

Bands hybridizing weakly with the *DICH* probe at high stringency were also detected in all species except *D. purpurea*. For example the second most strongly hybridizing band in *A. majus* had a size of ~4.3 kb. These bands hybridized more strongly to the *DICH* probe at moderate stringency (fig. 4) and strongly to the *CYC* probe at high stringency. This suggested that they corresponded to the *CYC* locus and therefore that *CYC* and *DICH* are the two most closely related TCP loci in *A. majus*, *S. nuttallianus*, and *M. orontium*. A single band of ~3.7 kb was detected in *D. purpurea* at moderate and low stringency hybridization with the *DICH* probe and at low stringency hybridization with *CYC*, as expected from the presence of a single sequence in Digitalis that is equally similar to the *CYC* and *DICH* loci of *A. majus*.

At low stringency both *CYC* and *DICH* probes detected multiple bands in all species, consistent with the presence of multiple members of the TCP gene family.

To investigate further the variation of *CYC* alleles within the genus Antirrhinum, sequences were obtained from 17 Antirrhinum species accessions representing the full morphological and geographical range of the genus. Phylogenetic analysis gave clear support for monophyly of Antirrhinum *CYC* in relation to the outgroups *M. orontium* and *S. nuttallianus* (fig. 5). The *CYC* alleles of Antirrhinum species showed low substitution divergence (average $P = 0.004$; maximum $P = 0.017$), suggesting that

they had diverged relatively recently. Several groups of species were found to share the same or very closely related sequences including (1) *A. graniticum* and *A. barellieri*; (2) *A. mollissimum* and *A. hispanicum*; (3) *A. cirrigherum*, *A. latifolium*, and *A. australe*; and (4) *A. lopesianum*, *A. meonanthum*, and *A. braun-blanquetii*. However, no bootstrap support was obtained for the distinction of groups (2), (3), or (4) using ML or MP.

As in comparisons of *CYC* and *DICH* sequences from Antirrhineae, *CYC* alleles from Antirrhinum differed by several insertions and deletion polymorphisms. About half the alleles shared an additional trinucleotide (ACT) in their 3′ region. A second trinucleotide repeat (CAA), encoding glutamine residues N-terminal to the TCP domain, was present in five copies in most alleles, four copies in *A. molle*, and two copies in *A. siculum*. The *A. molle CYC* allele differed from other Antirrhinum alleles by four insertions, each of three nucleotides that maintained the reading frame.

Discussion

CYC-like and *DICH*-like sequences were obtained from members of the Antirrhineae and from *D. purpurea*. The sequences formed two well-supported clades within the Antirrhineae—one containing the *A. majus CYC* gene, the other *A. majus DICH*—with the *D. purpurea* sequence basal to both. The basal position of the *D. purpurea* sequence is consistent with the taxonomic relationship between Digitalis and the Antirrhineae (Ghebrehiwet, Bremer, and Thulin 2000; Olmstead et al. 2001). The *D. purpurea* sequence is also more closely related to the Antirrhineae *CYC*- and *DICH*-like genes found in this study than to paralogous TCP genes from *A. majus* (Zsuzsanna Schwarz-Sommer, personal communication), suggesting that it represents a *CYC*/*DICH* homolog.

Our findings differ from those of Vieira, Vieira, and Charlesworth (1999), who proposed the existence of five *CYC*-like loci in Antirrhineae, including *A. majus*. Two of their proposed loci—*cyc1a* and *cyc1b*—correspond to Antirrhinum *CYC* alleles in our study, and their *cyc2*, *cyc3*, and *cyc4* loci correspond to Misopates *CYC*, Misopates *DICH*, and Antirrhinum *DICH*, respectively. The previous study is therefore likely to have classified alleles as paralogous genes. In addition, genera as distantly related to Antirrhinum as Linaria, Cymbalaria, and Digitalis were proposed to carry sequences identical, or nearly identical, to those found in Antirrhinum (Vieira, Vieira, and Charlesworth 1999). We found no evidence for sequences so similar to Antirrhinum *CYC* alleles in more distantly related species, and the sequences that we did obtain from these taxa were not represented in the previous study. Possible explanations for these discrepancies is that the previous PCR primers were able to amplify *CYC* and *DICH* loci only from Antirrhinum and Misopates and that samples from more distantly related taxa had been contaminated with these sequences.

We provided further support for the existence of only one *CYC* and one *DICH* locus in Antirrhineae by direct sequencing and restriction digestion of PCR products. With the exception of Chaenorhinum, each individual

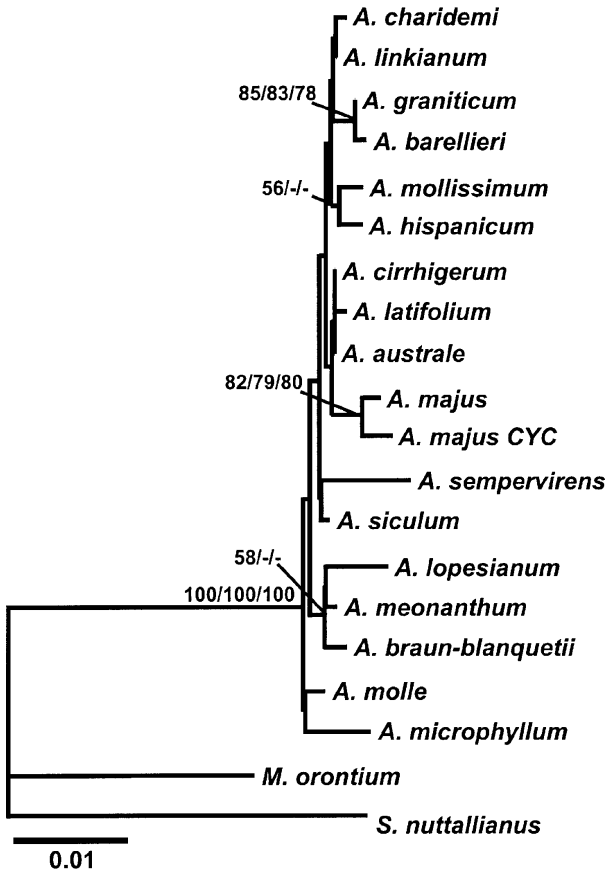

FIG. 5.—Minimum evolution phylogram of *CYC* sequences from Antirrhinum species. *M. orontium* and *S. nuttallianus* are used as outgroups. Bootstrap support values are shown as in figure 1. The GenBank *CYC* reference sequences of *A. majus* is named *A. majus CYC*. The scale bar indicates Kimura two-parameter distance.

carried only one or two *CYC* sequences, as expected of alleles. Moreover, the PCR products from *A. majus* and *A. molle*, distinguished by a restriction site polymorphism, segregated as alleles of the same locus in the F2 of an *A. majus* × *A. molle* hybrid and were used to map the locus to the same position as the *cyc* mutation (Schwarz-Sommer et al. 2003). Southern hybridization also suggested that *DICH* was the locus most similar to *CYC* in Antirrhinum, Sairocarpus, and Misopates, and it failed to provide evidence for sequences identical to Antirrhinum *CYC* in the more distantly related Linaria, Cymbalaria, and Digitalis.

The sequence phylogeny is consistent with the *CYC* and *DICH* loci having arisen from a gene duplication in the Veronicaceae (after the divergence of Digitalis) in an ancestor of the Antirrhineae with asymmetric flowers (Reeves and Olmstead 1998).

Although the relationship between *DICH* or *CYC* sequences from Antirrhinum, Sairocarpus, and Misopates is not well resolved, the phylogenies of both loci place Antirrhinum basal to Sairocarpus and Misopates (fig. 3). This contrasts with the phylogeny of the chloroplast *trnF* + *L* sequences in which Antirrhinum is basal. Such discrepancies between phylogenies of organelle and nuclear sequences are not uncommon in plants at lower taxonomic levels, and they can be explained by reticulation

(e.g., sorting of alleles present in a common ancestral population) or by hybridization (e.g., Schwarzbach and Rieseberg 2002).

At a higher level, the *DICH* sequences from Linaria were placed basal to *DICH* from other Antirrhineae, whereas the Linaria *CYC* and *trnL + F* sequences were nested within the Antirrhineae. This might be explained by an additional duplication of *DICH* followed by loss of different paralogs from Linaria and related taxa. Although the presence of two *CYC* sequences in Chaenorhinum suggested that a more recent duplication had occurred, it appeared likely to have been confined to the lineage leading to Chaenorhinum. Likewise, additional *CYC*-like or *DICH*-like genes in all other taxa could not be detected by PCR amplification or Southern hybridization. An alternative explanation for the anomalous position of Linaria *DICH* sequences is that they had been subjected to different selectional pressures (discussed below).

These comparisons suggested a slightly higher rate of nucleotide substitution in *CYC*-like genes compared to *aldose reductase* genes However, they revealed a much higher turnover of indels in *CYC*-like sequences. Almost all indels are present in regions outside those encoding the functionally important TCP and R boxes. Although the rapid divergence of these regions is likely to represent relaxed evolutionary constraints, adaptive changes cannot be ruled out in the absence of further functional studies.

Rapid sequence evolution can theoretically follow divergence in the expression of duplicate genes, which allows proteins to assume different specialized functions (a case of subfunctionalization; Lynch and Force 2000). Although *CYC* and *DICH* partially overlap in function in the floral meristem of *A. majus*, their expression domains differ, and each also has a unique role, consistent with subfunctionalization following duplication (Carpenter and Coen 1990; Luo et al. 1996, 1999). In *L. vulgaris*, loss of *CYC* activity can have the same phenotypic effect as loss of both *CYC* and *DICH* in *A. majus* (Cubas, Vincent, and Coen 1999). This greater dependence on CYC activity in floral asymmetry is unlikely to reflect absence of the *DICH* locus from Linaria, because we detected *DICH* sequences in both *L. vulgaris* and *L. maroccana*. It might therefore reflect a greater degree of divergence between CYC and DICH functions in Linaria, relative to Antirrhinum. Our finding that the two Linaria *DICH* sequences are placed basally within the *DICH* clade (whereas Linaria *CYC* sequences cluster with Chaenorhinum *CYC* sequences) is consistent with the Linaria *DICH* locus having experienced different constraints following divergence of function. An alternative explanation is that *CYC* and *DICH* have very similar functions in Linaria and that DICH activity was already absent from the *L. vulgaris* material in which *CYC* function was tested experimentally (e.g., because of a recent mutation or epimutation). Floral asymmetry in this case would appear dependent on CYC activity alone.

Rapid sequence evolution has been recorded for several duplicated transcription factor genes in plants, including *CYC*-like genes of Gesneriaceae and *CONSTANS LIKE* genes of Brassicaceae (Citerne, Möller, and Cronk 2000; Lagercrantz and Axelsson 2000). No asymmetry in the rate of molecular evolution was detected for *CYC* and *DICH*, as has been found for other duplicated transcription factors (e.g., Stauber, Jäckle, and Schmidt-Ott 2002; Stauber, Prell, and Schmidt-Ott 2002).

Only a few nucleotide substitutions were found for *CYC* alleles within the genus Antirrhinum. One explanation for the similarity of *CYC* alleles is that speciation in Antirrhinum occurred relatively recently—a view supported by the ability of Antirrhinum species to form fertile hybrids and the low degree of sequence divergence of other nuclear and chloroplast sequences within the genus (our unpublished results). Dating the divergence between the teosinte-branched (*tb1*) gene in maize and its closest homolog in *Oryza sativa* to 65 myr before present (Purugganan 1997), provides a conservative divergence estimate of 3.2 to 4.5×10^{-3} substitutions per million years for *CYC*-like genes. Assuming that the rates of sequence evolution in *CYC* and *DICH* are comparable to their monocot homologs, then the divergence of Antirrhinum species can be estimated to have begun less than 5.3 to 3.7 MYA, and most Antirrhinum species to have arisen within the last million years. These divergence-time estimates for Antirrhinum species are consistent with estimates from other nuclear and chloroplast sequences (our unpublished results).

Several weakly supported clades were detected within *CYC* alleles from different Antirrhinum species. Species groupings largely corresponded to shared morphological characters and similar geographical distributions (Sutton 1988) and to similarities of other DNA sequences (our unpublished results), suggesting that the *CYC* locus provides a weak phylogenetic signal at this taxonomic level. This trend, however, is disrupted by *A. siculum* and *A. charidemi*, which appear geographically and phylogenetically distinct from the other sampled species but have *CYC* alleles that group among them.

CYC is linked to the self-incompatibility (*S*) locus of Antirrhinum by less than 1 cM (Schwarz-Sommer et al. 2003). Antirrhinum species carry functionally equivalent *S* alleles (e.g., Sherman 1939), which are presumed to be ancient and to have been maintained by balancing selection (Vekemans and Slatkin 1994), as in Solanaceae (Ioerger et al. 1990). Where tested, Antirrhinum *S* allele sequences show the expected deep divergences (Xu et al. 1996) and provide little species-phylogenetic signal. Although *CYC* is closely linked to *S*, it appears not to have been subjected to the same selection, suggesting that recombination has been sufficient to uncouple the fate of the *S* and *CYC* loci. However, one exceptional *CYC* sequence, differing by four trinucleotide insertions, was detected in *A. molle*. Although this is placed basally to other Antirrhinum *CYC* sequences, as expected of an old allele, it differs from other Antirrhinum sequences by more trinucleotide insertions than substitutions, and it shares none of its insertions with the outgroup sequences from Misopates or Sairocarpus. The *CYC* allele in *A. molle* is therefore likely to have accumulated its insertions very recently. Because the trinucleotide insertions cause amino acid insertions, rapid evolution of the *CYC* allele in *A. molle* might have been driven by positive selection.

Alternatively, it might reflect an increased rate of insertion mutations or a decreased constraint in the *A. molle* lineage. Further sampling of sequences within *A. molle* populations will be necessary to address this question.

Note Added in Proof

We would like to draw the reader's attention to a study by L. C. Hileman and D. A. Baum, published recently in this journal (*Molecular Biology and Evolution*, 2003, **20**:591–600), which provides similar DNA sequence evidence for the fast evolution of *CYCLOIDEA*-like genes in the Antirrhineae.

Supplementary Material

Sequence alignments are provided online. All novel sequences have been deposited in GenBank under the accession numbers AY316693 to AY316730.

Acknowledgments

Seeds were kindly provided by E. S. Coen, D. Charlesworth, I. Mateu-Andrés, C. Thébaud, and the seedbank at the Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben. We thank J. Golz for *L. maroccana* DNA, M. Möller for primers cycFS, cyc*R, and *trnL* + *F*, and D. Charlesworth and two anonymous reviewers for their constructive comments. This work was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC).

Literature Cited

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

Carpenter, R., and E. S. Coen. 1990. Floral homeotic mutation produced by transposon-mutagenesis in *Antirrhinum majus*. Genes Dev. **4**:1483–1493.

Citerne, H. L., M. Möller, and Q. C. B. Cronk. 2000. Diversity of *cycloidea*-like genes in Gesneriaceae in relation to floral symmetry. Ann. Bot. **86**:167–176.

Comeron, J. M. 1995. A method for estimating the numbers of synonymous and nonsynonymous substitutions per site. J. Mol. Evol. **41**:1152–1159.

———. 1999. K-Estimator: calculation of the number of nucleotide substitutions per site and the confidence intervals. Bioinformatics **15**:763–764.

Cubas, P., E. Coen, and J. M. Martínez Zapater. 2001. Ancient asymmetries in the evolution of flowers. Current Biol. **11**: 1050–1052.

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

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

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

Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolated procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**:11–15.

Felsenstein, J. 1985. Confidence-limits of phylogenies—an approach using the bootstrap. Evolution **39**:783–791.

Ghebrehiwet, M., B. Bremer, and M. Thulin. 2000. Phylogeny of the tribe Antirrhineae (Scrophulariaceae) based on morphological and *ndhF* sequence data. Plant Syst. Evol. **220**:223–239.

Hasegawa, M., H. Kishino, and T. A. Yano. 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **22**:160–174.

Ioerger, T. R., A. E. Clarke, and T.-H. Kao. 1990. Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. Proc. Natl. Acad. Sci. USA **87**:9732–9735.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.

Kosugi, S., and Y Ohashi. 1997. PCF1 and PCF2 specifically bind to *cis* elements in the rice proliferating cell nuclear antigen gene. Plant Cell **9**:1607–1619.

Kumar, S. 1996. Stepwise algorithm for finding minimum evolution trees. Mol. Biol. Evol. **13**:584–593.

Lagercrantz, U., and T. Axelsson. 2000. Rapid evolution of the family of *CONSTANS LIKE* genes in plants. Mol. Biol. Evol. **17**:1499–1507.

Lukens, L., and 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, and E. S. Coen. 1999. Control of organ asymmetry in flowers of *Antirrhinum*. Cell **99**:367–376.

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

Lynch, M., and A. Force. 2000. The probability of duplicate gene preservation by subfunctionalization. Genetics **154**:459–473.

Möller, M., M. Clokie, P. Cubas, and Q. C. B. Cronk. 1999. Integrating molecular phylogenies and developmental genetics: a Gesneriaceae case study. Pp. 375–402 in P. M. Hollingsworth, R. M. Bateman, and R. J. Gornall, eds. Molecular systematics and plant evolution. Taylor & Francis, London.

Notredame, C., D. G. Higgins, and J. Heringa. 2000. T-coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. **302**:205–217.

Olmstead, R. G., C. W. de Pamphilis, A. D. Wolfe, N. D. Young, W. J. Elisens, and P. A. Reeves. 2001. Disintegration of the Scrophulariaceae. Am. J. Bot. **88**:348–361.

Purugganan, M. D. 1997. The MADS-box floral homeotic gene lineages predate the origin of seed plants: phylogenetic and molecular clock estimates. J. Mol. Evol. **45**:392–396.

Reeves, P.A., and R. G. Olmstead. 1998. Evolution of novel morphological and reproductive traits in a clade containing *Antirrhinum majus* (Scrophulariaceae). Am. J. Bot. **85**:1047–1056.

Riechmann, J. L., J. Heard, G. Martin et al. (14 co-authors). 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science **290**:2105–2110.

Schwarz-Sommer, Z., E. de Andrade Silva, R. Berndtgen et al. (11 co-authors). 2003. A linkage map of an F2 hybrid population of *Antirrhinum majus* and *A. molle*. Genetics **163**:699–710.

Schwarzbach, A. E., and Rieseberg L. H. 2002. Likely multiple origins of diploid hybrid sunflower species. Mol. Ecol. **11**:1703–1715.

Sherman, M. 1939. The inheritance of self-sterility in various species and varieties of Antirrhinum. Z. f. indukt. Abst. Vererbungsl. **77**:3–17.

Stauber, M., H. Jäckle, and U. Schmidt-Ott. 1999. The anterior determinant bicoid of Drosophila is a derived Hox class 3 gene. Proc. Natl. Acad. Sci. USA **96**:3786–3789.

Stauber, M., A. Prell, and U. Schmidt-Ott. 2002. A single Hox3 gene with composite bicoid and zerknüllt expression characteristics in non-Cyclorrhaphan flies. Proc. Natl. Acad. Sci. USA **99**:274–279.

Sutton, D. A. 1988. A revision of the tribe Antirrhineae. Oxford University Press, London.
Swofford, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Mass.
Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three coding regions of chloroplast DNA. Plant Mol. Biol. **17**:1105–1109.
Takezaki, N., A. Rzhetzky, and M. Nei. 1995. Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. **12**:823–833.
Vekemans, X., and M. Slatkin 1994. Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics **137**: 1157–1165.
Vieira, C. P., J. Vieira, and D. Charlesworth. 1999. Evolution of the *cycloidea* gene family in Antirrhinum and Misopates. Mol. Biol. Evol. **16**:1474–1483.
Xue, Y., R. Carpenter, H. G. Dickinson, and E. S. Coen. 1996. Origin of allelic diversity in Antirrhinum S locus RNases. Plant Cell **8**:805–814.

Brandon Gant, Associate Editor

Accepted May 8, 2003