Published in Mol.Biol.Evol. 29: 1155-1166 (2012), doi:10.1093/molbev/msr283

Title: Evolution and diversification of the *CYC/TB1* gene family in Asteraceae – a comparative study in gerbera (Mutisieae) and sunflower (Heliantheae)

Submission: Research article

Authors: Sari Tähtiharju[1*], Anneke S. Rijpkema[1*], Adrien Vetterli[2], Victor A. Albert[3], Teemu H. Teeri[1], Paula Elomaa[1]

[1] Department of Agricultural Sciences, P.O. Box 27, FI-00014 University of Helsinki, Finland

[2]Current address: Department of Biosciences, P.O. Box, FI-00014 University of Helsinki, Finland

[3]Department of Biological Sciences, University at Buffalo, Buffalo, NY 14260, United States of America

* These authors contributed equally to this paper

Corresponding author:

Paula Elomaa

Department of Agricultural Sciences

P.O.Box 27

00014 University of Helsinki, Finland

tel: +358-9-19158431

fax: +358-9-19158727

e-mail: paula.elomaa@helsinki.fi

Keywords: Asteraceae, flower development, TCP, *CYCLOIDEA*, *Gerbera*, *Helianthus*

Running head: *CYC/TB1* gene family in gerbera and sunflower

List of nonstandard abbreviations: -

Abstract

Plant-specific TCP domain transcription factors have been shown to regulate morphological novelties during plant evolution, including the complex architecture of the Asteraceae inflorescence that involves different types of flowers. We conducted comparative analysis of the *CYCLOIDEA/TEOSINTE BRANCHED1* (*CYC/TB1*) gene family in *Gerbera hybrida* (gerbera) and *Helianthus annuus* (sunflower), two species that represent distant tribes within Asteraceae. Our data confirms that the *CYC/TB1* gene family has expanded in Asteraceae, a condition that appears to be connected with the increased developmental complexity and evolutionary success of this large plant family. Phylogenetic analysis of the *CYC/TB1* gene family revealed both shared and lineage-specific duplications in gerbera and sunflower, corresponding to the three gene lineages previously identified as specific to core eudicots: CYC1, CYC2, and CYC3. Expression analyses of early stages of flower primordia development indicated that especially within the CYC2 clade, with the greatest number of secondary gene duplications, gene expression patterns are conserved between the species and associated with flower and inflorescence development. All sunflower and gerbera CYC2 clade genes showed differential expression between developing flower types, being upregulated in marginal ray (and trans) flowers. One gene in gerbera (*GhCYC3*) and two in sunflower (*HaCYC2d* and *HaCYC2c*) were indicated to be strong candidates as regulators of ray flower identity, a function that is specific for Asteraceae. Our data further showed that other CYC2 clade genes are likely to have more specialized functions at the level of single flowers, including the late functions in floral reproductive organs that may be more conserved across plant families. The expression patterns of CYC1 and CYC3 clade genes showed more differences between the two species but still pointed to possible conserved functions during vegetative plant development. Pairwise protein-protein interaction assays gave the first

molecular evidence that CYC/TB1-like proteins function in complexes. Compared to sunflower, the gerbera proteins showed higher capacity for dimerization, between as well as within CYC clades. Our data from two distant species within the Asteraceae suggests that the expansion and the apparent conservation of especially the CYC2 clade *CYC/TB1*-like genes is associated with the evolution of the increased complexity of the Asteraceae inflorescence architecture.

Introduction

Asteraceae is one of the largest families of flowering plants. A rapid tribal radiation within the family during the last 40 million years has resulted in nearly 25000 species in about 1600 genera, which are widely distributed all over the world in a variety of ecological habitats (Funk et al. 2005; Barker et al. 2008). The presence of a head-like inflorescence (capitulum) that resembles a large single flower has been proposed as a key innovation behind the evolutionary success of the Asteraceae (Gillies et al. 2002). Typically, an Asteraceae inflorescence consists of tens to hundreds of individual flowers that are often specialized in structure and function (Harris 1995).

Helianthus annuus (sunflower) and *Gerbera hybrida* (gerbera) are well known examples of radiate species, in which inflorescences are composed of both ray and disc flowers. They belong to phylogenetically distant tribes, Heliantheae and Mutisieae, respectively, the latter of which holding a basal position in the family (Bremer 1994). Flowers in gerbera and sunflower develop acropetally on the capitulum; the primordia of the marginal flowers forming first in capitula of only a few millimeters in diameter. In both gerbera and sunflower, mature ray and disc flowers differ in their symmetry. The marginal ray flowers have large attractive petals, whereas the inner bisexual disc flowers tend to be less conspicuous (fig. 1). The species differ in details within their flower types. As is common to Mutisieae, the outer ray flowers of gerbera are female and bilabiate. They have an abaxial ligule made up of three fused petals and two smaller adaxial petals, and staminodia in place of functional stamens (Bremer 1994). Disc flowers are hermaphroditic and develop anthers that congenitally fuse and cover the carpel. Furthermore, gerbera has a third flower type: trans flowers. These are functionally similar to ray flowers, but their ligules are shorter. In contrast to gerbera, sunflower has only

two flower types: sterile ray flowers formed of two abaxial petals and hermaphroditic disc flowers. The individual flowers of gerbera and sunflower do not have leaf-like sepals. Instead, the calyx is composed of many scabrid pappus bristles in gerbera and of two elongate awns in sunflower. In addition, in sunflower each disc flower is supported by a bract, called the palea.

We have demonstrated that ray flower identity in gerbera is affected by the plant-specific CYCLOIDEA/TEOSINTE BRANCHED 1 (CYC/TB1) -like TCP domain transcription factor GhCYC2 (Broholm et al. 2008). A common, conserved function for the *TCP* genes appears to be regulation of cell proliferation and growth (reviewed by Busch and Zachgo 2009; Rosin and Kramer 2009; Martin-Trillo and Cubas 2010). In higher plants *TCP* genes form a gene family that, based on the differences within their DNA-binding basic helix-loop-helix (bHLH) motif, can be divided into classes I and II (Aguilar-Martinez, Poza-Carrion, and Cubas 2007). The *CYC*/*TB1* genes belong to class II. Within the eudicot lineage, predating the core eudicots, the *CYC*/*TB1* subfamily experienced two rounds of gene duplications giving rise to three major clades, CYC1, CYC2 and CYC3 (Howarth and Donoghue 2006). In general, *CYC*/*TB1*-like genes are involved in regulating the development of axillary meristems that form either branches or flowers (Martin-Trillo and Cubas 2010). The CYC1 clade gene *TCP18/BRC1* of *Arabidopsis* and its most similar counterparts in monocots, *TB1* in maize and *OsTB1* in rice, are all involved in the control of shoot branching (Doebley, Stec, and Hubbard 1997; Takeda et al. 2003; Aguilar-Martinez, Poza-Carrion, and Cubas 2007; Finlayson 2007). The *Arabidopsis BRC1* gene as well as the *Arabidopsis* CYC3 clade gene *TCP12/BRC2* are also expressed in developing flower primordia although their potential functions during flower development remain unclear (Aguilar-Martinez, Poza-Carrion, and Cubas 2007; Finlayson 2007). In fact, no functional data for the CYC3 clade genes exists so far whereas those of the CYC2 clade genes are available for some species. *CYC* and its paralog *DICHOTOMA* (*DICH*)

control flower symmetry in *Antirrhinum majus* (Luo et al. 1996; Luo et al. 1999), and many studies across rosids and asterids indicate that CYC2 clade genes have been recruited several times independently to regulate dorsoventral asymmetry of flowers (reviewed by Busch and Zachgo 2009; Preston, Kost, and Hileman 2009). Whereas the CYC2 clade genes in these species are typically expressed in the dorsal parts of the flower (summarized in Preston, Kost, and Hileman 2009), the expression of *GhCYC2* in gerbera is primarily ventral and connected with organ fusion and development of the large ligule (Broholm et al. 2008). Furthermore, *GhCYC2* gene expression follows the radial organization of the capitulum and is specifically localized to the developing marginal, bilaterally symmetrical ray flowers and is completely absent from the centermost, more radially symmetrical disc flowers. Overexpression of *GhCYC2* in transgenic gerbera converted disc flowers into ray-like flowers with enlarged petals and disrupted stamen development. In *Senecio* (Asteraceae: tribe Senecioneae), CYC2 clade genes (*RAY1* and *RAY2*) are similarly associated with the radiate flower head form and ray flower development (Kim et al. 2008). Both the gerbera and *Senecio* data suggest that *CYC2*-like genes have played a key role in the evolution of the complex inflorescence structure of the sunflower family.

Although *GhCYC2* plays a major role in differentiating gerbera flower types, our previous study (Broholm et al. 2008) demonstrated that misregulation of *GhCYC2* is not sufficient for complete transformation of one flower type to another. Rather, it suggested involvement of a small family of *CYC/TB1*-like genes in this process in gerbera. Recently, Chapman et al. (2008) identified ten *CYC/TB1*-like genes in sunflower, in stark contrast to *Arabidopsis*, which harbours only three, one for each of the three clades (Howarth and Donoghue 2006; Aguilar-Martinez, Poza-Carrion, and Cubas 2007). Thus, it can be anticipated that the *CYC/TB1*-like gene family has expanded and experienced sub- and/or neofunctionalization

contributing to the evolution of the inflorescence architecture in Asteraceae. In order to unravel the putative functions of the *CYC*/*TB1*-like genes in establishing the unique characteristics of Asteraceae inflorescences, we set out to perform a comparative analysis of *CYC/TB1*-like genes in two species, gerbera and sunflower, representing distant tribes within this plant family (supplementary fig. S1, Supplementary Material online). In this study we isolated six new *CYC/TB1*-like genes from gerbera and examined the phylogenetic relationships between the gerbera and sunflower *CYC*/*TB1*-like genes. We performed a detailed expression analysis of the gerbera and sunflower *CYC*/*TB1*-like genes in different flower types across the capitulum during comparable stages of early flower primordia development, and investigated pairwise protein-protein interactions among the proteins encoded by these genes. Our results indicate that in particular, the CYC2 clade of *CYC/TB1*-like genes in Asteraceae has expanded during evolution and may associate with the increased inflorescence complexity in the family.

Materials and Methods

Plant Material

Gerbera hybrida cultivar Terra Regina (Terra Nigra BV, The Netherlands) and *Helianthus annuus* cultivar Pacino cola (Siemenliike Siren Oy, Finland) were grown under standard greenhouse conditions, as described previously (Ruokolainen et al. 2010).

Scanning Electron Microscopy

Preparation of the gerbera and sunflower capitulum samples for scanning electron microscopy analysis was performed as described in Uimari et al. (2004). Samples were examined using the DSM 962 (Zeiss) scanning electron microscope at the Electron Microscopy Laboratory of the Institute of Biotechnology, or the field emission scanning electron microscope (FESEM, HITACHI S-4800) at the Department of Chemistry, both at the University of Helsinki.

Isolation of Gerbera *CYC/TB1*-like Genes

We have previously identified four full length cDNAs for gerbera *CYC/TB1*-like genes (Broholm et al. 2008). In this study, two new gerbera *CYC/TB1*-like gene fragments were amplified from cDNA prepared from young developing inflorescences, and four additional ones were isolated from genomic DNA using degenerate primers as described in Broholm et al. (2008). The full-length cDNAs were amplified using the SMART RACE cDNA amplification kit (Clontech), and in the case of *GCYC8* the GenomeWalker Universal Kit (Clontech), and were cloned into pCR-Blunt vectors (Invitrogen) for sequencing. The *CYC/TB1*-like coding sequences obtained in this study have been deposited in the GenBank database (accession numbers JN190059-JN190064).

Phylogenetic Analysis

To gain insight into the relationships between the different gerbera and sunflower *CYC/TB1*-like genes, a maximum likelihood phylogenetic tree was constructed using 27 protein sequences (gerbera and sunflower CYC/TB1-like protein sequences and a selection of CYC/TB1 sequences from other eudicots) that were aligned using ClustalW (Thompson, Higgins, and Gibson 1994). All of the gerbera *CYC/TB1*-like genes encoded TCP and R domains. Outside these highly conserved domains, their sequences differed in both content and length, so we used only the extended TCP and R domain portions of the corresponding nucleotide alignments (see supplementary fig. S2, Supplementary Material online), and a single most-optimal tree was computed using the RaxML BlackBox web server (http://phylobench.vital-it.ch/raxmlbb/index.php) running RaxML version 7.2.8 (Stamatakis, Hoover, and Rougemont 2008). Default settings were used with the GTR-gamma model of molecular evolution. Accession numbers of the included sequences are listed in supplementary fig. S2, Supplementary Material online. One hundred bootstrap samples were generated to assess support for the inferred relationships. Local bootstrap values (in percentages) are indicated for branches with >50% support. For parsimony analysis (not shown), TNT software (Goloboff et al. 2008) was used with 100 random replicates of data entry order, saving 10 trees per replicate.

Quantitative Real-time PCR Analyses

For the expression analysis, flower primordia were excised under a stereomicroscope using a scalpel and immediately stored in liquid nitrogen. Samples were taken from gerbera ray, trans and disc flower primordia and sunflower ray and disc flower primordia at different developmental stages (fig. 2). The gerbera stages correspond to those previously described in

Laitinen et al. (2006). For sunflower comparable stages were defined and sampled by comparing the morphology of the emerging gerbera and sunflower flower primordia. For both species, the disc flower sample comprised only the centermost primordia. For later stages of gerbera reproductive organs, samples were collected from ray and disc flowers at different inflorescence development stages (pooled from stages 2, 4, 6, and 8 for ray flowers and from stages 6 and 8 for disc flowers). Again, we defined sunflower inflorescence developmental stages comparable to those defined for gerbera by Helariutta et al. (1993) and used the same method of pooling samples. All sunflower reproductive organ samples were taken from disc flowers only, as the sunflower ray flower lack these organs. In addition, the gerbera stamen sample was pooled from disc flower samples only, and the gerbera style plus stigma and ovary samples were a mix of ray and disc flower organs. Total RNA was isolated using the Trizol reagent (Life Technologies/Gibco-BRL) following the manufacturer's instructions. The RNA was treated with RQ1 RNase-free DNase (Promega) according to the manufacturer's instructions, then extracted with phenol and precipitated with ethanol. After resuspension in diethyl pyrocarbonate-treated water, the RNA concentrations were measured and equalized within a sample set and the RNA integrity was analyzed by gel electrophoresis. The cDNA was synthesized from 1 µg of the treated RNA samples using a AAGCAGTGGTATCAACGCAGACTAC$(T)_{30}$VN primer (SMART RACE primer, Clontech) and Superscript III (Invitrogen) at 50 °C for 1.5 hours, and the polymerase was then inactivated at 70 °C for 10 minutes. The cDNA samples were diluted 20 times by adding 380 µl milliQ water. The real-time RT-PCR was carried out according to the manufacturer's instructions in a LightCycler 480 instrument (Roche Applied Science) and monitored with SYBR-green I dye (Roche Applied Science). Product specificity was evaluated by melting-curve analysis. All qPCR primers used are listed in supplementary table S1, Supplementary Material online. The PCR efficiencies were analyzed for all primer combinations to make sure

that they were close to 100 %. Relative expression levels were calculated using the ΔΔCt method (Pfaffl 2001). All *CYC*/*TB1* gene expression levels were calculated relative to their expression in a root sample of the same species. For all gerbera samples and for sunflower flower primordia samples the expression levels were normalized to *ACTIN* expression levels. Though *ACTIN* was expressed uniformly in the sunflower flower primordia samples, this was not the case for sunflower vegetative and reproductive tissue samples. We tested different candidate reference genes such as three sunflower *ACTIN* genes, *UBIQUITIN* and *GAPDH* but were unable to find a reference with fully uniform expression in all tissues, so the relative expression levels in sunflower plant tissues were normalized against the amount of mRNA in each amplification reaction. For the analysis of *CYC*/*TB1* gene expression levels in sunflower and gerbera flower primordia, the highest expression value encountered in the series of different samples tested was set to 100 for each gene separately, and lower values were normalized to this value. The results are shown as an average of the relative expression values of three biological replicates; the error bars show the standard deviation. The heat map that displays the relative expression levels of the gerbera and sunflower *CYC*/*TB1* genes in different vegetative and reproductive tissues was made using the average relative expression levels of three biological replicate samples for each tissue type.

Yeast Two Hybrid Assays

Yeast two-hybrid (Y2H) assays were done using the GAL4-based Proquest^{TM} Two-Hybrid System (Invitrogen) to understand the interaction patterns among the gerbera, sunflower and *Arabidopsis* CYC/TB1-like proteins. *Helianthus HaCYC* sequences were kindly provided by Dr. Mark Chapman. To isolate the corresponding full length cDNAs from *Helianthus annuus*

cv. Pacino cola, polymerase chain reaction (PCR) amplification was performed with gene specific Gateway primers using as template either a mix of cDNA from various vegetative and reproductive tissues, or genomic DNA. Similarly, the full length cDNAs of gerbera *GhCYC* genes were amplified by PCR using the aforementioned cDNA clones as templates. *Arabidopsis TCP12/BRC2* and *TCP18/BRC1* cDNAs were amplified using the plasmids obtained from Dr. Pilar Cubas. A second round of PCR was carried out with adapter primers by using the products of the first round as templates. Subsequent recombination cloning steps, constructs for Y2H experiments, yeast transformations, tests for autoactivation and pairwise interaction studies were performed as described in Broholm et al. (2010). A yeast construct with the full length *Arabidopsis* gene *TCP1* was a gift from Dr. Richard Immink.

All GhCYC proteins except GhCYC5 and GhCYC7 had the ability to autoactivate the expression of *HIS3* reporter genes (supplementary table S2, Supplementary Material online). Likewise, most of the HaCYCs were capable of activating transcription by themselves, as was also seen for all three tested *Arabidopsis* proteins (TCP1, TCP12 and TCP18) (supplementary table S2, Supplementary Material online). In order to use the self activating CYC/TB1-like proteins as baits in the Y2H assay, deletions were made to decrease the transcriptional activity. The N- or C-terminal part of the protein was deleted from those proteins that showed the highest transcriptional activity (growth in the presence of 100 mM 3-amino-1,2,4-triazole, 3-AT). The conserved TCP and R domains were left intact, because they have been suggested to be involved in mediating protein-protein interactions (Pruneda-Paz et al. 2009; Martin-Trillo and Cubas 2010). Depending on the given protein, deletion of either the N- or C-terminal part resulted in decreased transcriptional activity (supplementary table S2, Supplementary Material online). Bait constructs that showed the greatest decrease in the transcriptional activity were used in further interaction studies.

The pairwise Y2H screenings were performed three times with each protein combination in both directions. The matings were also done reciprocally and the mating efficiency was 100%. The empty-vector transformants (pDEST22 and pDEST32) were used as negative controls. A protein combination was scored as a true interaction when it resulted in growth for all selection markers in both reciprocal screenings.

All constructs used in this study were confirmed by sequencing. The Gateway primers used to amplify the full-length and truncated protein coding sequences of GhCYCs and HaCYCs are described in supplementary table S1, Supplementary Material online.

Results

Early Development of Gerbera and Sunflower Ray and Disc Flowers

In earlier work we defined six stages of early flower development for gerbera (cv. Terra Regina) (Laitinen et al. 2006). To compare early flower development in gerbera and sunflower, and to aid the sampling for expression analysis, we defined comparable stages for ray and disc flower development in sunflower (cv. Pacino cola) (fig. 2). At stage 1, flower primordia were small undifferentiated bumps. These began to form ring-shaped petal primordia at stage 2. Whereas in gerbera, ray and disc flower primordia were indistinguishable at these developmental stages, the sunflower ray and disc flower primordia could easily be distinguished by their different shapes: the disc flower primordia being almost

perfectly circular, whereas the ray flower primordia were flatter (fig. 2C and 2D). At stage 3, the individual petals could be clearly distinguished. Gerbera ray and disc flower primordia still appeared identical, and in both flower types the developing pappus bristles (whorl 1) and anthers (whorl 3) were visible (fig. 2A and 2B). In sunflower disc flowers, the anther primordia also started to become recognizable, while in ray flowers only two petal primordia grew (fig. 2C and 2D). At this developmental stage the sunflower disc flowers were completely covered by their outgrowing floral bract (palea) (note that for the SEM-pictures these bracts were removed). At stage 4, petals began to elongate, covering the developing stamen and carpel primordia (if present). At stages 5 and 6 all floral organs grew out further, and on the surface of both gerbera and sunflower disc flower petals, small hair-like structures began to emerge. At stage 5 in gerbera, the stamens lagged in growth and were shorter in the female ray flowers than in the perfect disc flowers (Laitinen et al. 2006). It can be clearly seen that three of the gerbera ray flower petals fuse and grow out further, while the two other petals remain unfused and smaller (fig. 2B). The differences in petal elongation were even more pronounced at stage 6 (fig. 2B). In sunflower, the petals of the sterile ray flowers were clearly fused at stage 5, and had grown out further at stage 6 (fig. 2D). The ray and disc flower primordia looked completely different at this stage.

The differences in early floral development were much greater between sunflower ray and disc flowers than between gerbera ray and disc flowers. In young gerbera ray flowers, the development of all floral primordia initiated as in disc flowers, though some of these organs (stamens and two of the petals) arrested growth later on. On the other hand, in young sunflower ray flowers, only petal primordia appear to develop and grow out. Compared to the development of ray flowers, the early stages of disc flower development are more similar in gerbera and sunflower, except for some species-specific differences such as the presence of

pappus bristles in gerbera, and the presence of a palea that grows over each developing disc flower in sunflower.

CYC/TB1-Like Genes in Asteraceae

Our previous studies already suggested that gerbera may have more than four *CYC/TB1*-like genes (Broholm et al. 2008). The degenerate PCR approach revealed that indeed, like sunflower (Chapman, Leebens-Mack, and Burke 2008), gerbera harbors at least ten gene family members. We performed maximum likelihood phylogenetic analysis to assess the relationships between the different *CYC/TB1*-like genes from gerbera, sunflower and a selection of other species. The ten gerbera *CYC/TB1*-like genes fell, like the sunflower *CYC/TB1*-like genes (Chapman, Leebens-Mack, and Burke 2008), into the three previously described clades: CYC1, CYC2 and CYC3 (Howarth and Donoghue 2005; Howarth and Donoghue 2006; also supported by maximum parsimony analysis [results not shown]). The CYC1 clade contained two gerbera and two sunflower genes (fig. 3), with the gerbera genes grouping together, although with low support. The CYC3 clade was composed of two orthologous gerbera/sunflower gene sets, and the sunflower ortholog of *GhCYC6* was duplicated.

The CYC2 clade seems to have undergone the greatest number of gene duplications within the Asteraceae, resulting in six *CYC2*-like genes in gerbera and five in sunflower (fig. 3; Chapman, Leebens-Mack, and Burke 2008). The most clearly supported orthologs are *GhCYC7* and *HaCYC2a* and the recently duplicated gerbera gene pair *GhCYC4* and *GhCYC9* grouping with *HaCYC2b*. In a previously published phylogenetic analysis by Kim et al.

(2008), *GhCYC4* also grouped with *HaCYC2b*. Although not well supported in our tree, the remaining gerbera and sunflower genes may form two additional orthologous groups, with the exception of *GhCYC3*. Altogether, our phylogenetic analysis of the *CYC/TB1* gene family reveals that although both species have ten CYC/TB1-like genes, they do not represent ten pairs of orthologous loci but rather results of differential gene gain and loss in each species.

Expression Patterns of *CYC/TB1*-Like Genes of Gerbera and Sunflower

Both our own earlier results (Broholm et al. 2008) and those of Kim et al. (2008) indicate that CYC2 clade genes are involved in regulating the complex inflorescence structure of Asteraceae. Retention of such a large number of *CYC/TB1*-like genes in Asteraceae suggests further functional diversification (sub- and/or neofunctionalization) beyond the establishment of the three basic *CYC*-like gene lineages. To look into any flower-type specific roles of the different Asteraceae *CYC*/*TB1* genes, and especially those of the multiplied members of the CYC2 clade, we compared the expression patterns of all gerbera and sunflower *CYC*/*TB1* genes during the early stages of flower primordia development, as defined in fig. 2.

All sunflower and gerbera CYC2 clade genes were found to be expressed at higher levels in ray flower primordia than in disc flower primordia, throughout their development (stages 2-6) (fig. 4). In sunflower the difference in expression level between ray and disc flower primordia was greatest for *HaCYC2c* and *HaCYC2d*. Most gerbera CYC2 clade genes (*GhCYC2*, *GhCYC4*, *GhCYC5*, and *GhCYC9*) were expressed at more or less comparable levels in ray and trans flower primordia, and at much lower levels in disc flower primordia. *GhCYC3*, on the other hand, was expressed at high levels in ray flower primordia only. Unlike the other

gerbera CYC2 clade genes, *GhCYC7* was most highly expressed during the early stages of ray and trans flower development (stage 2), and its expression then decreased (fig. 4). Similar to most CYC2 clade genes, the sunflower CYC3 clade gene *HaCYC3a* was expressed at higher levels in ray flower primordia than disc flower primordia (fig. 4). Thus, of the three sunflower and two gerbera CYC3 clade genes, *HaCYC3a* was the only one that showed a clear differential expression between the different flower types during early flower development. The other exception was the CYC1 clade gene *HaCYC1a*, which was expressed at higher levels in disc flower primordia than ray flower primordia (fig. 4). We observed low expression levels and large variation among the biological replicates for the other CYC3 and CYC1 clade genes, suggesting that they were not expressed during early flower primordia development (data not shown). Alternatively, their expression may be localized to restricted domains which, however, should be verified using e.g. *in situ* hybridization.

To study the expression of the various *CYC*/*TB1*-like genes in a broader context we examined the abundance of their transcripts in vegetative organs and in the reproductive organs of the flower at later stages of development (fig. 5). Many of the *CYC*/*TB1*-like genes (most CYC2 clade genes, both gerbera CYC3 clade genes, and the CYC1 clade genes *GhCYC10* and *HaCYC1a*) were found to be expressed at relatively high levels in ovary tissue (fig. 5), in accordance with the results of Chapman et al. (2008). Furthermore, the CYC2 clade genes of both species, with the exception of *GhCYC7*, were found to be most strongly expressed in the floral reproductive organs (stamen, stigma plus style, and ovary) (fig. 5). *GhCYC7* in contrast was expressed at higher levels in inflorescence stem tissue than in the sampled reproductive organs. Whereas the CYC2 clade genes are typically most strongly expressed in the floral reproductive organs, this is not generally true for the CYC1 and CYC3 clade genes. Both gerbera and sunflower CYC3 clade genes show relatively high expression levels in leaf blade,

verifying earlier results in sunflower (Chapman, Leebens-Mack, and Burke 2008). The expression profiles of the CYC1 clade genes appear to be more divergent, but, like the CYC3 clade genes, these genes are also expressed at relatively high levels outside of the reproductive organs. As previously shown by Chapman et al. (2008), the expression levels of the sunflower *HaCYC1b* gene in the tissues tested were extremely low, so this gene is not displayed in the heat map (fig. 5).

Protein-Protein Interactions among *CYC/TB1*-Like Proteins

Information on specific dimerization patterns can be used, in addition to sequence conservation, phylogenetic relationships and expression patterns, to predict biological functions as well as to evaluate the functional redundancy or diversification of given proteins (Immink, Kaufmann, and Angenent 2010). Given that *CYC/TB1*-like genes of gerbera and sunflower have overlapping expression patterns, and that the TCP domain, through the conserved amphipathic helices, has been shown to mediate protein-protein interaction (Pruneda-Paz et al. 2009), we examined all combinations of pairwise protein-protein interactions for gerbera and sunflower CYC/TB1-like proteins, respectively.

Yeast two-hybrid interaction studies revealed that, in general, the capacity for homo- and heterodimer formation of CYC/TB1-like proteins was greater in gerbera than in sunflower (fig. 6). All except GhCYC1, GhCYC5 and GhCYC8 formed homodimers. Most of the proteins formed heterodimers with other proteins with the exception of GhCYC5 and its ortholog HaCYC2c, which neither interacted with any other CYC/TB1-like proteins nor formed homodimers (fig. 6). However, heterodimerization between certain gerbera proteins

(GhCYC3, GhCYC2 and GhCYC9), as between their orthologs in sunflower, appeared to be conserved, especially within the CYC2 clade. Within the CYC1 clade, the sunflower HaCYC1a and HaCYC1b proteins were the only ones able to form heterodimers. Although the gerbera CYC1 clade proteins did not interact with each other, both formed heterodimers with CYC2 and CYC3 clade proteins. Similar heterodimer formation between the different clades was apparent for several gerbera and sunflower CYC/TB1-like proteins. To examine whether interactions between the CYC clades is conserved in other plant lineages as well, we performed pairwise protein-protein interaction studies for the *Arabidopsis* CYC/TB1-like proteins. All three *Arabidopsis* CYC/TB1-like proteins interacted with each other (supplementary fig. S3, Supplementary Material online), so it is possible that interaction capacity between the proteins belonging to different clades is conserved among rosids and asterid eudicots.

Discussion

Expansion of *CYC/TB1*-gene family in gerbera and sunflower

Our finding that gerbera has ten *CYC/TB1*-like *TCP* genes is consistent with the results of Chapman et al. (2008), who showed that the gene family has experienced a significant expansion in sunflower. As a result, both gerbera and sunflower have more genes in this angiosperm-specific clade of *TCP* genes than any other plant lineages studied thus far (Howarth and Donoghue 2006; Navaud et al. 2007). Phylogenetic analyses by Howarth and Donoghue (2005, 2006) showed that the CYC1, CYC2 and CYC3 clades of *CYC/TB1*-like

genes arose from a single ancestral gene by duplications within the eudicot lineage. These investigators placed CYC1 sister to the clade containing CYC2 and CYC3. Despite extensive analysis of *CYC/TB1*-like genes (Howarth and Donoghue 2005; Howarth and Donoghue 2006) no other duplications within the CYC1 clade similar to those demonstrated in gerbera and sunflower have been found. Taking into account the low support for grouping of gerbera *GhCYC1* and *GhCYC10*, the presence of two genes in both species suggests that they may originate from the shared paleopolyploidization event near the base of Asteraceae 40-45 MYA (Barker et al. 2008). This is supported by divergence estimates for *HaCYC1a* and *HaCYC1b* to the same time period (Chapman, Leebens-Mack, and Burke 2008). Alternatively, the gerbera ortholog of *HaCYC1b* may have been lost, so the *GhCYC1*/*GhCYC10* pair could represent gerbera-specific duplicates.

In contrast to the CYC1 gene clade, the CYC2 and CYC3 gene clades have expanded in plant lineages other than the Asteraceae. In diverse asterid and rosid species, the number of CYC2 copies has been shown to vary from one to four [reviewed by Busch and Zachgo (2009); Howarth et al. (2011)]. Gene isolation from the Dipsacales (asterid) revealed that they have two CYC3 gene copies instead of the one in *Arabidopsis* (rosid) (Howarth and Donoghue 2005). In several species, duplication patterns of CYC3 and CYC2 genes are similar (Howarth and Donoghue 2006). This seems not to be the case in gerbera and sunflower, since they show the highest number of CYC2 paralogs (six for gerbera and five for sunflower) documented so far, but have only three and two CYC3 genes, respectively. The CYC3 genes in gerbera and sunflower form orthologous groups. An additional gene duplication in sunflower, *HaCYC3c/HaCYC3b*, is estimated to have diverged 18.5 MYA, i.e., after the basal Heliantheae whole-genome duplication 26-31 MYA (Chapman, Leebens-Mack, and Burke 2008). The presence of four orthologous groups within the CYC2 clade points to a shared

origin, but species-specific gene duplications not affected by whole genome duplications are also apparent (GhCYC4/GhCYC9 and HaCYC2d/HaCYC2e). Genome and single-gene duplications in plants are known to increase genomic complexity, and have been suggested to produce new genes that allow the evolution of functional novelty (Jiao et al. 2011). Our discovery that especially the CYC2 clade in particular both expanded and maintained duplicates appears to support the idea that members of this clade may have novel and important functions contributing to the complex inflorescence structure in Asteraceae.

Expression analysis indicates both early and late functions for the conserved CYC2 clade genes during flower type differentiation

The CYC2 clade genes that showed the greatest number of secondary duplications in each species also shared strikingly similar expression domains. In both species, all *CYC2* genes were expressed in developing ray flower primordia. In gerbera, however, following its more gradual transition from ray to disc flower structure, most CYC2 clade genes seem to be also involved in defining the development of bilaterally symmetrical trans flowers, and are absent only from the centermost, most strongly actinomorphic disc flowers.

In sunflower, unlike in gerbera, the CYC1 clade gene *HaCYC1a* was the only one that showed upregulation in developing disc flower primordia. In further contrast to the gerbera counterparts, the expression of the sunflower *HaCYC3a* gene suggests that it may also play a role in early ray flower development. Nevertheless, in both species, we can clearly identify genes with specific functions in ray flowers only: *GhCYC3* in gerbera and both *HaCYC2d* and *HaCYC2c* in sunflower. In both species, the restriction of expression to ray flower primordia

suggests that these genes share a general function in defining ray flower identity, while the other CYC2 clade genes are likely to have more localized functions at the level of single ray (and trans) flowers. These functions may include flower symmetry regulation at the dorsoventral axis or suppression of stamen development, as previously shown for several CYC2 clade genes (Luo et al. 1996; Hileman and Baum 2003; Song et al. 2009). Our previous studies showed that suppression of *GhCYC2* expression in transgenic gerbera did not cause alterations in ray flower identity, indicating that some other gene defines ray identity and that *GhCYC2* may function as a modifier gene in this process (Broholm et al. 2008). This is in accordance with classical genetic studies that have indicated that the absence of ray flowers (discoid flower head) in the Asteraceae is determined by one or two major genes (Gillies et al. 2002; Kloos, George, and Sorge 2004). Furthermore, Berti et al. (2005) reported on two sunflower mutants with modified flower type identity. They showed that a single recessive gene controlled the *tubular ray flower* (*turf*) mutant trait that shows development of tubular, disc-like flowers in place of zygomorphic ray flowers, while in the *Chrysanthemoides* (*Chry*) mutant, all flowers were ray-like. Our data suggests that *GhCYC3* in gerbera and *HaCYC2d* and *HaCYC2c* in sunflower are the strongest candidates for conferring ray flower identity, but functional analyses are required to test this hypothesis.

The fact that, in both species, most CYC1 and CYC3 clade genes did not show clear differential expression between different flower types suggests that they, with the possible exception of *HaCYC1a* and *HaCYC3c*, are not involved in the early differentiation of flower types. The relatively high expression levels of these genes in vegetative tissues point to their functions in other aspects of plant development, including leaf and inflorescence stem (scape) development. The *CYC/TB1*-like genes in *Arabidopsis*, in addition to being expressed in flower and shoot primordia (Aguilar-Martinez, Poza-Carrion, and Cubas 2007; Cubas, Coen,

and Zapater 2001), are also expressed in rosette leaves and in inflorescence stems (Aguilar-Martinez, Poza-Carrion, and Cubas 2007; Guo et al. 2010; Koyama, Sato, and Ohme-Takagi 2010). The *Arabidopsis* CYC2 clade gene *TCP1* has been shown to regulate vegetative development by modulating brassinosteroid biosynthesis (Guo et al. 2010), whereas *BRC1/TCP18* and *BRC2/TCP12* were earlier shown to regulate shoot branching (Aguilar-Martinez, Poza-Carrion, and Cubas 2007; Finlayson 2007). The possible functions of these genes in *Arabidopsis* flower development remain to be elucidated. Our results strongly suggest that, in addition to determining ray (and trans) flower specification during the early development of flower types, *CYC2* genes also function in later developmental stages of reproductive and vegetative organs together with *CYC1* and *CYC3* genes.

In conclusion, our results imply that differences, even subtle ones, in patterns of gene expression have contributed to the maintenance of *CYC* gene duplicates within the Asteraceae. Moreover, our study indicates that of the CYC genes, the *CYC2* clade in particular has both early and late functions. Further studies are needed to elucidate the functions of the different gene family members at the organ and tissue level.

Regulatory network involving CYC/TB1 proteins

Pairwise interaction studies showed the complex network of interactions involving CYC/TB1-like proteins. In both species, the capacity of CYC/TB1-like proteins for heterodimerization and the overlap in their expression domains, indicate that these proteins are likely to function in complexes. Our data provides the first experimental proof that the CYC/TB1 proteins can indeed interact *in vitro* and may also form heterodimers among the CYC clades, as initially

proposed by Howarth and Donoghue (2006). However, for some proteins when used as baits, such as the CYC3 clade proteins GhCYC8, HaCYC3a and HaCYC3b, relatively large deletions were necessary to reduce strong autoactivation. Therefore we cannot rule out the possibility that we may have missed some true interactions in our analysis.

In the early stages of flower development, heterodimerization might only occur among CYC2 clade proteins in ray and trans flowers of gerbera, whereas in sunflower, CYC2 proteins might also interact with the CYC3 protein HaCYC3a. Although the gerbera CYC3 clade genes did not show clear expression patterns during early primordia development, their expression in reproductive tissues during later stages of flower development overlapped with that of the CYC2 clade genes and also with the CYC1 clade gene *GhCYC10*. Together with the interaction data, this indicates that the corresponding proteins may function in the same complexes. Similarly, sunflower HaCYC1a interacted with both CYC2 and CYC3 clade proteins and showed overlapping expression especially in ovary and leaf as shown earlier by Chapman et al. (2008). A pair of orthologous proteins in the two species, GhCYC5 and HaCYC2c, did not interact with any other CYC/TB1-like proteins, did not form homodimers, nor show strong transcriptional activity in yeast. Previous studies with the PCF group of TCP proteins have indicated that not all TCPs are able to regulate transcription by themselves and may need other interacting proteins for transcriptional activation (Kosugi and Ohashi 2002). It remains to be elucidated whether this may apply to GhCYC5 and HaCYC2c as well.

The capacity of CYC/TB1-like proteins to form homodimers appears to differ between gerbera and sunflower, whereas that of heterodimer formation was more similar in both species. Heterodimerization seemed to be most highly conserved among those of the CYC2 clade proteins that were most similar to each other, according to our phylogenetic and

expression analyses. In contrast, heterodimerization between the proteins of the three clades show great differences. In general, heterodimerization is known to increase functional specificity since heterodimers combine different DNA-binding domains and can therefore mediate differential gene regulation (Amoutzias et al. 2008). As proposed by Martin-Trillo and Cubas (2010), different heteromeric combinations (as detected in gerbera and sunflower) may, in addition to binding divergent *cis*-regulatory elements, recognize target genes with different affinity, or modulate each others' activity. This could add another level of complexity to the regulatory networks controlled by these factors.

Conclusions

In this study we show for the first time that CYC/TB1-like transcription factors are able to form different homo- and heterodimers, and may therefore form higher-order complexes *in planta*, leading to increased functional specificity. This important regulatory capacity has also been proposed for MADS domain proteins, another major group of transcription factors involved in reproductive development (Immink et al. 2010). In addition to protein complex formation, our data from the Asteraceae show other parallels between the MADS box and *CYC/TB1* genes. Similarly to MADS box genes that function both at the onset of development in determination of flower organ identities and later in regulating organ growth (Dornelas et al. 2011), the expanded Asteraceae CYC2 gene clade appears to show both early functions in flower type differentiation and late functions in reproductive organ development. Interestingly, Howarth and Donoghue (2006) raised the possibility that major duplications of MADS box and *CYC/TB1*-like genes that occurred at the same point in eudicot phylogeny are correlated and may be associated with the evolution of flower form. Our data from two distant species

within the Asteraceae suggests that the expansion and the apparent functional conservation of especially the CYC2 clade *CYC/TB1*-like genes are associated with the evolution of the increased complexity of the inflorescence architecture of the Asteraceae. The CYC2 clade *CYC/TB1*-like genes may therefore have played a major role in establishing this morphological novelty.

Supplementary Material

Supplementary table S1: Oligos cited

Supplementary table S2: Autoactivation of yeast constructs

Supplementary fig. S1: Relationships of Asteraceae tribes

Supplementary fig. S2: Sequence alignment

Supplementary fig. S3: Y2H of *Arabidopsis* CYC/TB1-proteins

Acknowledgements

We thank Dr. Mark Chapman for providing us *Helianthus annuus CYC/TB1*-like gene sequence information, Dr. Richard Immink for the *Arabidopsis* TCP1 yeast construct, and Prof. Pilar Cubas for the *TCP12/BRC2* and *TCP18/BRC1* cDNA clones. Eija Takala, Anu Rokkanen and the greenhouse team are thanked for their skilled technical assistance. Dr. Raili Ruonala is appreciated for valuable discussion and Dr. Fred Stoddard for language revisions. We thank the Electron Microscopy Unit of the Institute of Biotechnology and the Department of Chemistry, both at the University of Helsinki, for their help and for providing laboratory

facilities. This work was supported by the Academy of Finland (grants 115849 and 139092 to P.E.) and the Netherlands Organization for Scientific Research (grant 825.08.037 to A.S.R.).

Literature Cited

Aguilar-Martinez JA, Poza-Carrion C, Cubas P. 2007. *Arabidopsis* BRANCHED1 acts as an integrator of branching signals within axillary buds. Plant Cell 19:458-472.

Amoutzias GD, Robertson DL, de Peer YV, Oliver SG. 2008. Choose your partners: Dimerization in eukaryotic transcription factors. Trends Biochem. Sci. 33:220-229.

Barker MS, Kane NC, Matvienko M, Kozik A, Michelmore RW, Knapp SJ, Rieseberg LH. 2008. Multiple paleopolyploidizations during the evolution of the compositae reveal parallel patterns of duplicate gene retention after millions of years. Mol. Biol. Evol. 25:2445-2455.

Berti F, Fambrini M, Turi M, Bertini D, Puglesi C. 2005. Mutations of corolla symmetry affect carpel and stamen development in *Helianthus annuus*. Can. J. Bot. 83:1065-1072.

Bremer K. 1994. Asteraceae - cladistics & classification. Portland, Oregon, United States: Timber Press Inc.

Broholm SK, Tähtiharju S, Laitinen RA, Albert VA, Teeri TH, Elomaa P. 2008. A TCP domain transcription factor controls flower type specification along the radial axis of the gerbera (Asteraceae) inflorescence. Proc. Natl. Acad. Sci. U. S. A. 105:9117-9122.

Broholm SK, Pöllänen E, Ruokolainen S, Tähtiharju S, Kotilainen M, Albert VA, Elomaa P, Teeri TH. 2010. Functional characterization of B class MADS-box transcription factors in *Gerbera Hybrida*. J. Exp. Bot. 61:75-85.

Busch A, Zachgo S. 2009. Flower symmetry evolution: Towards understanding the abominable mystery of angiosperm radiation. Bioessays 31:1181-1190.

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

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

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

Dornelas MC, Patreze CM, Angenent GC, Immink RGH. 2011. MADS: The missing link between identity and growth? Trends Plant Sci. 16:89-97.

Finlayson SA. 2007. *Arabidopsis* teosinte Branched1-like 1 regulates axillary bud outgrowth and is homologous to monocot teosinte Branched1. Plant Cell Physiol. 48:667-677.

Funk VA, Bayer RJ, Keeley S, Chan R, Watson L, Gemeinholzer B, Schilling E, Panero JL, Baldwin BG, Garcia-Jacas N et al. (12 co-authors). 2005. Everywhere but antarctica: Using a supertree to understand the diversity and distribution of the Compositae. Biol Skr. 55:343-374.

Gillies ACM, Cubas P, Coen E, Abbott RJ. 2002. Making rays in the Asteraceae: Genetics and evolution of radiate versus discoid flower heads. In: Cronk QCB, Bateman RM, Hawkins JA, editors. Developmental Genetics and Plant Evolution. London: Taylor and Francis. p. 233-246.

Goloboff PA, Farris JS, Nixon KC 2008. TNT, a free program for phylogenetic analysis. Cladistics 24: 774-786.

Guo Z, Fujioka S, Blancaflor EB, Miao S, Gou X, Li J. 2010. TCP1 modulates brassinosteroid biosynthesis by regulating the expression of the key biosynthetic gene *DWARF4* in *Arabidopsis thaliana*. Plant Cell 22:1161-1173.

Harris EM. 1995. Inflorescence and floral ontogeny in Asteraceae: A synthesis of historical and current concepts. The Botanical Review 61:93-162.

Helariutta Y, Elomaa P, Kotilainen M, Seppänen P, Teeri TH. 1993. Cloning of cDNA coding for dihydroflavonol-4-reductase (DFR) and characterization of *dfr* expression in the corollas of *Gerbera-hybrida* var Regina (Compositae). Plant Mol. Biol. 22:183-193.

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

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

Howarth DG, Donoghue MJ. 2006. Phylogenetic analysis of the "ECE" (CYC/TB1) clade reveals duplications predating the core eudicots. Proc. Natl. Acad. Sci. U. S. A. 103:9101-9106.

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

Immink RGH, Kaufmann K, Angenent GC. 2010. The 'ABC' of MADS domain protein behaviour and interactions. Semin. Cell Dev. Biol. 21:87-93.

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS et al. (17 co-authors). 2011. Ancestral polyploidy in seed plants and angiosperms. Nature 473:97-U113.

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

Kloos WE, George CG, Sorge LK. 2004. Inheritance of the flower types of *Gerbera hybrida*. J. Am. Soc. Hort. Sci. 129:802-810.

Kosugi S, Ohashi Y. 2002. DNA binding and dimerization specificity and potential targets for the TCP protein family. Plant J. 30:337-348.

Koyama T, Sato F, Ohme-Takagi M. 2010. A role of *TCP1* in the longitudinal elongation of leaves in *Arabidopsis*. Bioscience Biotechnology and Biochemistry 74:2145-2147.

Laitinen RAE, Broholm S, Albert VA, Teeri TH, Elomaa P. 2006. Patterns of MADS-box gene expression mark flower-type development in *Gerbera hybrida* (asteraceae). BMC Plant Biol. 6:11.

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

Luo D, Carpenter R, Copsey L, Vincent C, Clark J, Coen E. 1999. Control of organ asymmetry in flowers of *Antirrhinum*. Cell 99:367-376.

Martin-Trillo M, Cubas P. 2010. TCP genes: A family snapshot ten years later. Trends Plant Sci. 15:31-39.

Navaud O, Dabos P, Carnus E, Tremousaygue D, Herve C. 2007. TCP transcription factors predate the emergence of land plants. J. Mol. Evol. 65:23-33.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29:e45.

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

Pruneda-Paz JL, Breton G, Para A, Kay SA. 2009. A functional genomics approach reveals CHE as a component of the *Arabidopsis* circadian clock. Science 323:1481-1485.

Rosin FM, Kramer EM. 2009. Old dogs, new tricks: Regulatory evolution in conserved genetic modules leads to novel morphologies in plants. Dev. Biol. 332:25-35.

Ruokolainen S, Ng YP, Broholm SK, Albert VA, Elomaa P, Teeri TH. 2010. Characterization of *SQUAMOSA*-like genes in *Gerbera hybrida*, including one involved in reproductive transition. BMC Plant Biology 10:128.

Song CF, Lin QB, Liang RH, Wang YZ. 2009. Expressions of ECE-CYC2 clade genes relating to abortion of both dorsal and ventral stamens in *Opithandra* (Gesneriaceae). BMC Evol. Biol. 9:244.

Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57:758-771.

Takeda T, Suwa Y, Suzuki M, Kitano H, Ueguchi-Tanaka M, Ashikari M, Matsuoka M, Ueguchi C. 2003. The *OsTB1* gene negatively regulates lateral branching in rice. Plant J. 33:513-520.

Thompson JD, Higgins DG, Gibson TJ. 1994. Clustal-W - improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.

Uimari A, Kotilainen M, Elomaa P, Yu DY, Albert VA, Teeri TH. 2004. Integration of reproductive meristem fates by a *SEPALLATA*-like MADS-box gene. Proc. Natl. Acad. Sci. U. S. A. 101:15817-15822.

Figure legends

Fig. 1. Radiate inflorescences of gerbera and sunflower.

Both sunflower (*Helianthus annuus* cv. Pacino cola, left) and gerbera (*Gerbera hybrida* cv. Terra Regina, right) have large solitary capitula composed of hundreds of flowers. The showy marginal ray flowers are sterile in sunflower and female in gerbera. The disc flowers are hermaphroditic in both species. Both gerbera and sunflower ray flowers are bilaterally symmetrical, whereas disc flowers are much more radially symmetrical, especially those of sunflower. In addition to ray and disc flowers, many gerbera varieties have a third flower type: trans flowers that are morphologically similar to ray flowers (bilaterally symmetrical, three fused and two rudimentary petals, rudimentary anthers), but with shorter petals. Pappus bristles forming the calyx in gerbera are visible (right); instead of pappus bristles, sunflower disc flowers have two pappus scales.

Fig. 2. SEM images of early stages of disc and ray flower development in gerbera and sunflower.

(*A*) Developing gerbera disc flower primordia (stages 1-6). (*B*) Developing gerbera ray flower primordia (stages 1-6). (*C*) Developing sunflower disc flower primordia (stages 1-6). (*D*) Developing sunflower ray flower primordia (stages 1-6). All size bars represent 50 μm.

Fig. 3. Maximum likelihood tree of *CYC/TB1*-like genes from gerbera and sunflower and selected other species.

One hundred bootstrap samples were generated to assess support for the inferred relationships. Local bootstrap probabilities are indicated for branches with >50% support. The alignment on which this tree is based, as well as the accession numbers for the sequences, is given in supplementary fig. S2, Supplementary Material online.

Fig. 4. Relative expression levels of gerbera and sunflower *CYC/TB1*-like genes during early stages of ray and disc flower development.

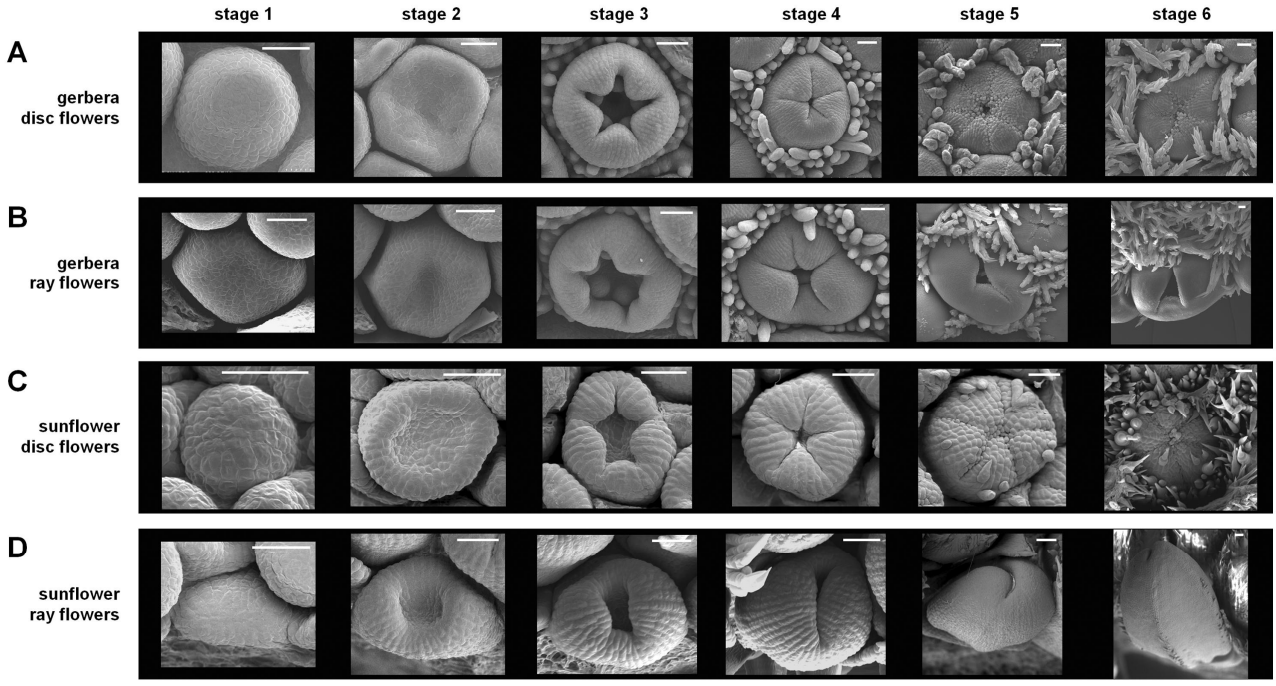
RNA was isolated from gerbera ray, trans and disc flower primordia and sunflower ray and disc flower primordia in developmental stages 2, 3, 4 and 6 (ray2, trans2, disc2 etc.). For comparison, the expression patterns of putative gerbera and sunflower orthologs are presented side by side in the same row. Species-specific duplicates (*GhCYC4/GhCYC9*; *HaCYC3b/HaCYC3c*) correspond to single genes in the other species. Heights of the bars for a given gene indicate relative differences in expression level within the range of tissues tested. The bars show the average of three biological replicates; the error bars show the standard deviation. Gene expression was normalized to a value of 100 for the highest expression value encountered. Expression levels of *ACTIN* were used for normalization.

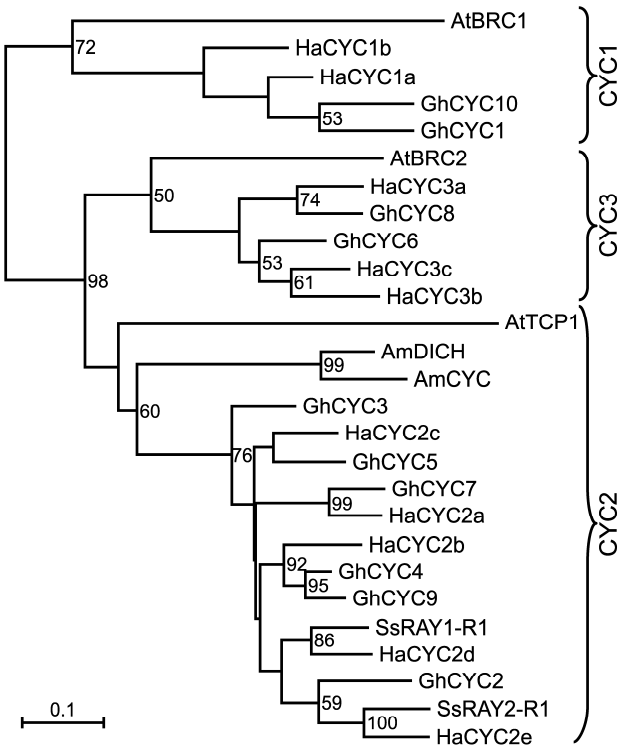
Fig. 5. Heat map displaying the relative expression levels of all gerbera and sunflower *CYC/TB1*-like genes in different vegetative and late developmental stages of reproductive tissues.

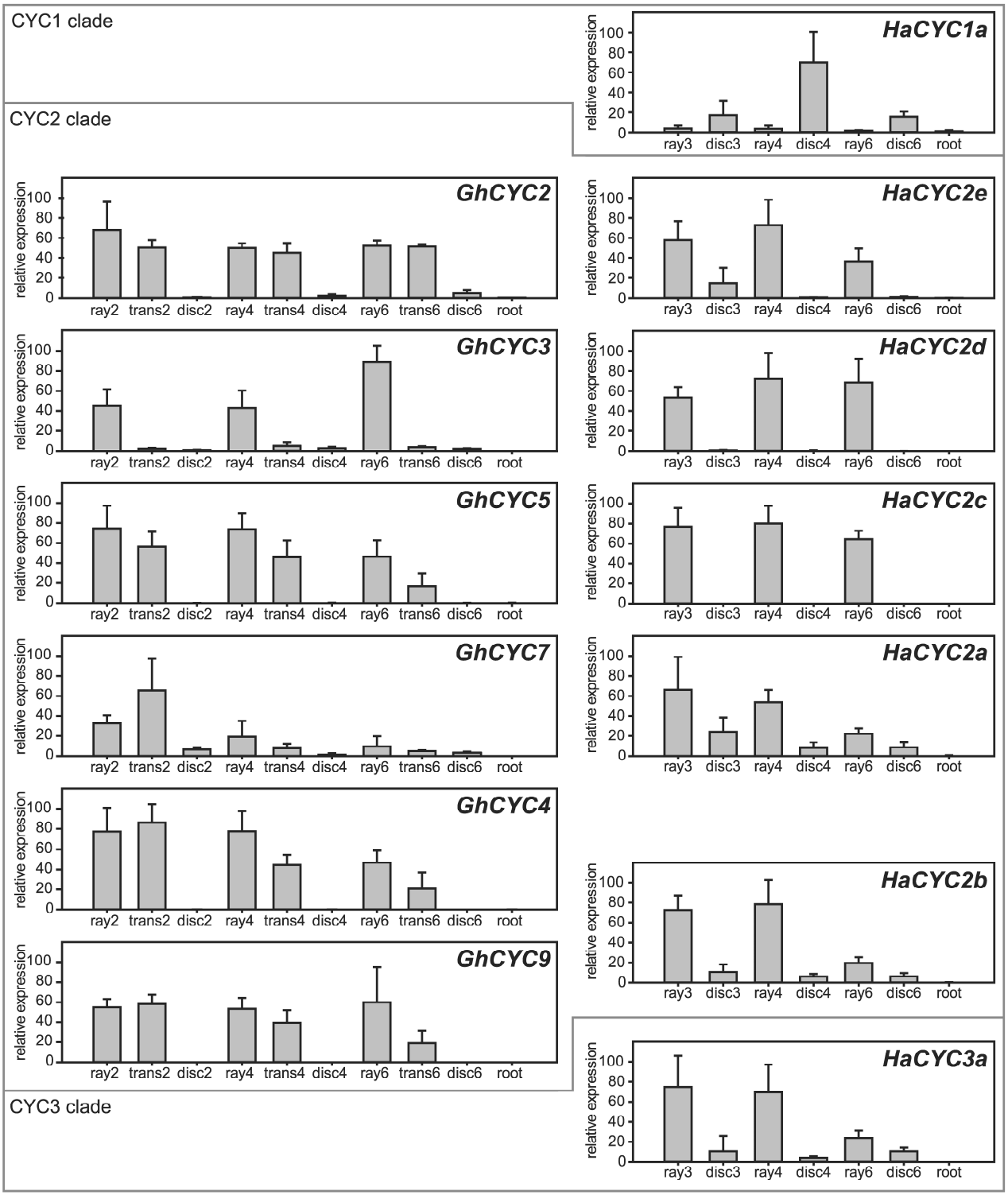
Inflo. stem and invol. bract refer to inflorescence stem and involucral bract, respectively. The darker the color, the higher the expression level of a given gene in a particular tissue relative to the expression of the same gene in the other tissues in this heat map. The bars that connect gerbera and sunflower genes in front of the graph indicate probable orthologous gene combinations. The heat map was made using the average relative expression levels of three biological replicate samples for each tissue type. All expression levels are relative to those of a root sample of the same species. In the case of the gerbera samples, *ACTIN* expression levels were used for normalization.

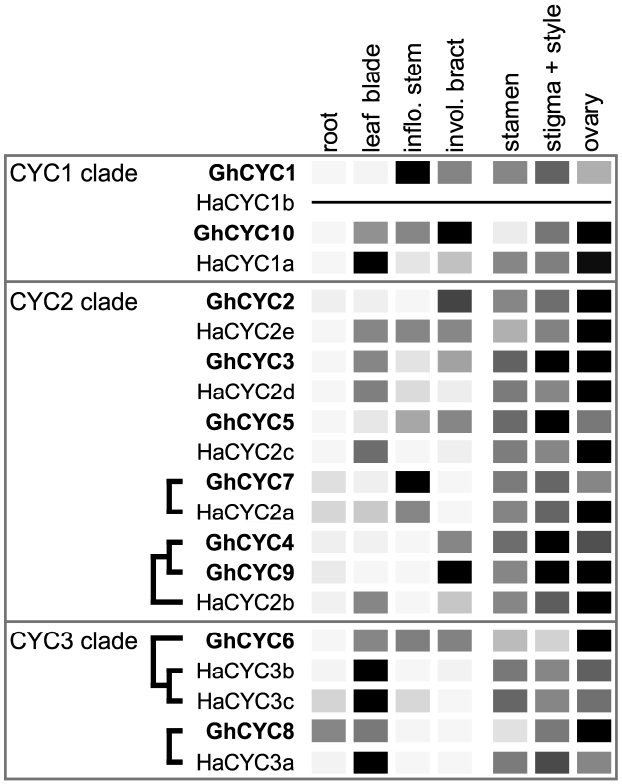
Fig. 6. Pairwise interactions of (A) gerbera and (B) sunflower CYC/TB1-like proteins identified in yeast 2-hybrid assays.

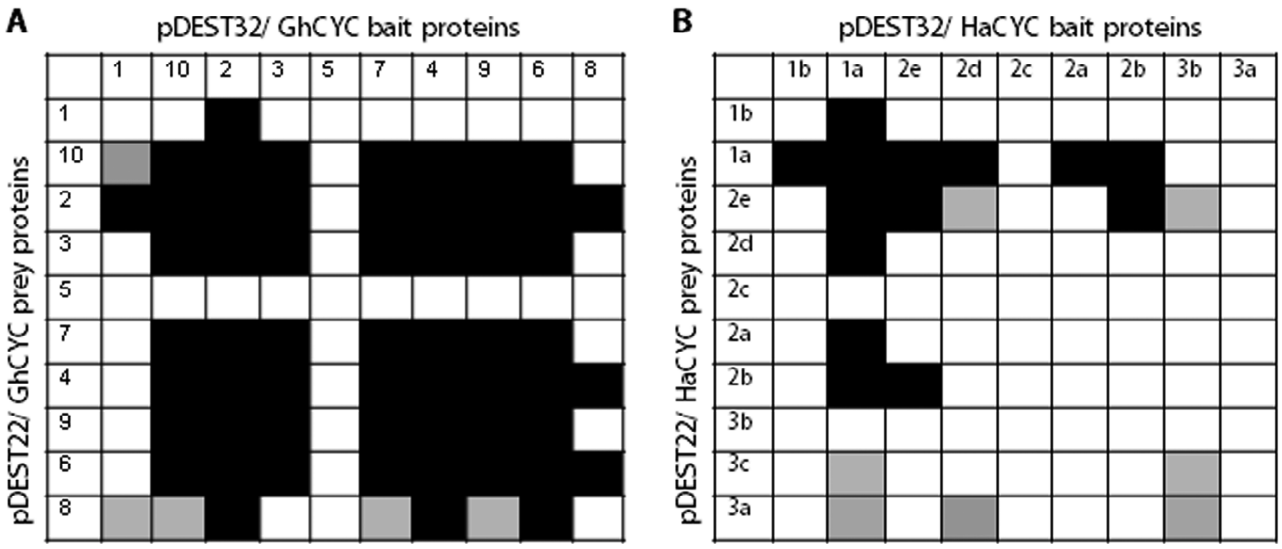
Proteins expressed from pDEST32 (bait proteins) are presented horizontally and proteins expressed from pDEST22 (prey proteins) vertically. Black indicates positive reciprocal interactions and grey positive interactions occurring in only one direction. The HaCYC3c was used only as a prey, due to its high autoactivity as a bait.


stage 1
stage 2
stage 3
stage 4
stage 5
stage 6
A
gerbera
disc flowers
B
gerbera
ray flowers
C
sunflower
disc flowers
D
sunflower
ray flowers


AtBRC1
72
HaCYC1b
HaCYC1a
GhCYC10
53
GhCYC1
CYC1
AtBRC2
50
HaCYC3a
74
GhCYC8
GhCYC6
53
HaCYC3c
61
HaCYC3b
CYC3
98
AtTCP1
AmDICH
99
AmCYC
60
GhCYC3
HaCYC2c
76
GhCYC5
GhCYC7
99
HaCYC2a
HaCYC2b
92
GhCYC4
95
GhCYC9
SsRAY1-R1
86
HaCYC2d
GhCYC2
59
SsRAY2-R1
100
HaCYC2e
CYC2
0.1


CYC1 clade
CYC2 clade
CYC3 clade
HaCYC1a
GhCYC2
HaCYC2e
GhCYC3
HaCYC2d
GhCYC5
HaCYC2c
GhCYC7
HaCYC2a
GhCYC4
HaCYC2b
GhCYC9
HaCYC3a
relative expression
0
20
40
60
80
100
ray2 trans2 disc2 ray4 trans4 disc4 ray6 trans6 disc6 root
ray3 disc3 ray4 disc4 ray6 disc6 root


root
leaf blade
inflo. stem
invol. bract
stamen
stigma + style
ovary
CYC1 clade
GhCYC1
HaCYC1b
GhCYC10
HaCYC1a
CYC2 clade
GhCYC2
HaCYC2e
GhCYC3
HaCYC2d
GhCYC5
HaCYC2c
GhCYC7
HaCYC2a
GhCYC4
GhCYC9
HaCYC2b
CYC3 clade
GhCYC6
HaCYC3b
HaCYC3c
GhCYC8
HaCYC3a


A
pDEST32/ GhCYC bait proteins
1
10
2
3
5
7
4
9
6
8
pDEST22/ GhCYC prey proteins
B
pDEST32/ HaCYC bait proteins
1b
1a
2e
2d
2c
2a
2b
3b
3a
pDEST22/ HaCYC prey proteins