PROF. CLAUDIO PUGLIESI (Orcid ID : 0000-0001-5846-7947)

Ligulate inflorescence of *Helianthus × multiflorus*, cv. Soleil d'Or, correlates with a misregulation of a *CYCLOIDEA* gene characterized by insertion of a transposable element

M. Fambrini[1], M. Bellanca[1], M. Costa Muñoz[1,2], G. Usai[1], A. Cavallini[1] & C. Pugliesi[1]

1 Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, I-56124 Pisa, Italy

2 University of Barcelona, Faculty of Biology, Avenida Diagonal, 643, 08028 Barcelona, Spain

Correspondence

C Pugliesi, Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, I-56124 Pisa, Italy.

email: claudio.pugliesi@unipi.it

ABSTRACT

Members of CYCLOIDEA (CYC)/TEOSINTE BRANCHED1 (TB1) transcription factor family are essential to control flower symmetry and inflorescence architecture. In the *Helianthus annuus* genome, ten *CYC/TB1* genes have been identified. Studies performed on mutants recognized *HaCYC2c* as one of the key players controlling zygomorphism in sunflower.

We identified *CYC2c* genes in the diploid *Helianthus decapetalus* (*HdCYC2c*) and in the interspecific hybrid *Helianthus × multiflorus* (*H×mCYC2cA* and *H×mCYC2cB*), a triploid (2n = 3x = 51), originated from unreduced eggs of *H. decapetalus* fertilized by reduced *H. annuus* male gametes. Phylogenetic analysis showed that *HdCYC2c* and *H×mCYC2c* were placed within a CYC2 subclade together with *HaCYC2c* but distinct from it. The present data showed that in *H. × multiflorus* the allele derived from *H. annuus* is deleted or highly modified.

The *H. × multiflorus* taxon exists in radiate and ligulate inflorescence type. We analyzed the *CYC2c* expression in *H. decapetalus* and in the cultivar "Soleil d'Or" of *H. × multiflorus*, a ligulate inflorescence type with actinomorphic corolla of disk flowers transformed in zygomorphic ray-like corolla.

In *H. decapetalus*, *HdCYC2c* gene showed differential expression between developing flower types, being upregulated in corolla of ray flowers in comparison to disk flower corolla. In *H. × multiflorus*, an insertion of 865 bp, that is part of a CACTA transposable element, was found in the 5'-untraslated region (5'-UTR) of *H×mCYC2cB*. This insertion was likely to cause the ectopic expression of the gene throughout the inflorescence, resulting in the observed loss of actinomorphy and originating a ligulate head.

Keywords

Floral symmetry; *CYCLOIDEA*-like genes; radiate and ligulate inflorescences; TCP transcription factors; transposable elements; *Helianthus × multiflorus.*

INTRODUCTION

The ancestral flower in the Asteraceae family was actinomorphic, and a deeply five-lobed and zygomorphic one is considered derived (Harris 1995). In *Helianthus annuus* inflorescences hermaphrodite flowers with actinomorphic symmetry (disk flowers) are surrounded by a whorl of zygomorphic sterile flowers (ray flowers). The zygomorphic ray flower is made of three petals [0:3 pattern (adaxial:abaxial)] shifted abaxially, forming a short and narrow corolla tube confined to the proximal end (Jeffrey 1977). Actinomorphic disk flowers (tubular flowers) are arrayed in left- and right-turning spiral rows. Each disk flower is subtended by a sharp-pointed chaffy bract, and it consists of an inferior ovary carrying a single ovule, two pappus scales (highly modified sepals), and a five-lobed tubular-like corolla. The five anthers are joined together to form a tube, with separate filaments attached to the base of the corolla tube. Inside the anther tube is the style, terminating in a divided stigma with receptive surfaces in close contact in the bud stage before the flower open (Mizzotti *et al.* 2015).

CYCLOIDEA-like genes are involved in flower symmetry regulation in various plant species (Broholm *et al.* 2014; Fambrini & Pugliesi 2017a; Moyroud & Glover 2017; Spencer & Kim 2017). *CYCLOIDEA* genes encode transcription factors (TFs) of the TCP family, a group of genes that have been associated to the control of growth and development such as cell cycle, axillary shoot outgrowth and leaf development (Luo *et al.* 1996; 1999; Doebley *et al.* 1997; Cubas *et al.* 1999; Theissen 2000; Nath *et al.* 2003; Krizek & Fletcher 2005). The TCP acronym stands for the first three identified members: TEOSINTE BRANCHED1 (TB1) in *Zea mays*, CYCLOIDEA (CYC) in *Antirrhinum majus*, and PROLIFERATING CELL

FACTOR (PCF) 1 and 2 in *Oryza sativa* (Cubas *et al*. 1999). TCP identifies a non-canonical basal helix-loop-helix (bHLH) domain of *circa* 60 residues (Aggarwal *et al*. 2010; Martín-Trillo & Cubas 2010; Uberti Manassero *et al*. 2013; Hileman 2014a,b). Based on the TCP motif, members of this TF family have been classified into PCF (TCP-P or class I) and CYC/TB1 (TCP-C or class II) subfamilies (Cubas *et al*. 1999; Viola *et al*. 2012). Class II TCP TFs have an additional motif, the R domain, which is *circa* 20 residues long and rich in arginine (Cubas *et al*. 1999). Phylogenetic and sequence analysis split class II into two clades, the CYC/TB1-like and the CINCINNATA (CIN)-like clades. The first, called ''ECE'' clade, is further divided into CYC1, 2 and 3 sub-clades, which have evolved due to a series of duplication events (Howarth & Donoghue 2006; Preston & Hileman 2009; 2012; Martín-Trillo & Cubas 2010; Uberti Manassero *et al*. 2013; Hileman 2014b; Specht & Howarth 2015; Zhong & Kellogg 2015). ECE refers to a conserved short motif (glutamic acid-cysteine-glutamic acid) between the TCP and R domains that have been found in many members of this clade (Howarth & Donoghue 2006).

Genes homologous to the CYC2 clade and involved in the evolution of floral symmetry have been amply recognized in many angiosperm families (Coen 1996; Cubas *et al*. 1999; 2001; 2004; Citerne *et al*. 2003; 2010; 2013; Costa *et al*. 2005; Howarth & Donoghue 2005; Feng *et al*. 2006; Damerval *et al*. 2007; Chapman *et al*. 2008; 2012; Wang *et al*. 2008; Fambrini *et al*. 2011; 2014a,b; Tähtiharju *et al*. 2012; Yang *et al*. 2012; 2015; Hileman 2014a; Berger *et al*. 2016; Xu *et al*. 2016). Analogously to *Antirrhinum majus* (Luo *et al*. 1996; 1999), most of the CYC2 clade genes were specifically expressed in dorsal or dorsal plus lateral regions of developing flowers (Costa *et al*. 2005; Damerval *et al*. 2007; Preston & Hileman 2009; Citerne *et al*. 2013); nevertheless, in Asteraceae and in monocots *CYC*-like genes expression in ventral region of floral organs has been observed (Bartlett & Specht 2011; Broholm *et al*. 2014; Juntheikki-Palovaara *et al*. 2014; Berger *et al*. 2016; Garcês *et al*. 2016). Therefore, the origin and evolution of zygomorphism have been correlated to asymmetric *CYC2*-like gene expression.

A complex case of *CYC2* recruitment occurs in members of the Asteraceae and Dipsacaceae families, both characterized by species with radiate inflorescences originated independently (Abbott *et al.* 2003; Broholm *et al.* 2008; Chapman *et al.* 2008; 2012; Kim *et al.* 2008; Busch & Zachgo 2009; Carlson *et al.* 2011; Fambrini *et al.* 2011; Bello *et al.* 2013; Berger *et al.* 2016; Garcês *et al.* 2016). The genome of Asteraceae and Dipsacaceae show the largest numbers of *CYC2* genes (Chapman *et al.* 2008; Carlson *et al.* 2011; Juntheikki-Palovaara *et al.* 2014; Berger *et al.* 2016; Garcês *et al.* 2016) with differential *CYC* transcription linked to changes of both floral symmetry and inflorescence architecture (Broholm *et al.* 2008; Carlson *et al.* 2011; Garcês *et al.* 2016).

In sunflower, ten members of the CYC/TB1 gene family were identified, and phylogenetic analysis showed that these genes occurred in three distinct clades (Chapman *et al.* 2008 Additionally, straightforward evidence of divergence in expression patterns across duplicates within all three clades of sunflower *CYC*-like genes was established (Chapman *et al.* 2008). Noteworthy, prominent differences were detected in the transcription pattern of the *CYC2* lineage, in which three genes were expressed in all floral tissues, one (*HaCYC2d*) was especially transcribed in ray flowers and at a lower level also in disc flowers and one (*HaCYC2c*) was expressed only in ray flowers. Additionally, molecular evolutionary analyses revealed that positive selection had promoted divergence of the *HaCYC2a*, *HaCYC2b* and *HaCYC2c* genes (Chapman *et al.* 2008). Therefore, duplication and functional divergence have played a major role in diversification of the sunflower CYC gene family.

The cultivated ornamental sunflower *Helianthus* × *multiflorus* is known in Europe since the sixteenth century, when it was described by Jacob Theodor Tabernaemontanus (Heiser & Smith 1960). *H.* × *multiflorus* was found to be triploid (2n = 51) and, at meiosis, it generally shows 17 bivalents and 17 univalents in the metaphase stage (Heiser & Smith 1960). The triploid nature explains why the taxon is infertile and only propagates through rhizome fragments. The high degree of similarity observed between the sesquiterpene lactone profile

of *H.* × *multiflorus* and the additive patterns of *H. annuus* and the diploid race of *Helianthus decapetalus* suggested that *H.* × *multiflorus* originated from hybridization between *H. annuus* and the diploid race of *H. decapetalus*, with one parental genome remaining unreduced (Spring & Schilling 1990). Molecular analyses, including comparison of variable regions of the cpDNA, suggested that *H.* × *multiflorus* originated from unreduced eggs of *H. decapetalus* fertilized by reduced *H. annuus* male gametes (Frey & Spring 2015). The taxon exists in both radiate and ligulate inflorescence type and has been known as ornamental, but it was never found in any natural habitat (Frey and Spring 2015).

In *H. annuus*, some mutations affect the development of inflorescence and/or florets (Cockerell 1915; Fambrini *et al.* 2003; 2007; 2011; 2014a,b; Berti *et al.* 2005; Chapman *et al.* 2012). Above all, a sunflower mutant named *Chrysanthemoides* (*Chry*), *Florepleno* or *double-flowered* (*dbl*), is characterized by a shift from actinomorphic to zygomorphic-like corolla of disk flowers (Cockerell 1915; Fick 1976; Heiser 1976; Fambrini *et al.* 2003; 2014a; Chapman *et al.* 2012). Although the inflorescence of *Chry* is much larger than that of *H.* × *multiflorus*, the phenotype of both inflorescences is apparently similar (compare Figs. 1B and 1C). In the inflorescences all flowers have zygomorphic-like corollas (Figs. 1A, B, C). The *Chry* mutant of our collection (*Chry2;* Berti *et al.* 2005; Fambrini *et al.* 2014a) as well the *dbl* mutant described by Chapman *et al.* (2012) showed an insertion of truncated versions of CACTA transposable elements (TEs) in the 5' promoter region of the *HaCYC2c* gene (Fambrini *et al.* 2014a). The TE insertion altered *HaCYC2c* expression and is fundamental to generate the *Chry* phenotype (Chapman *et al.* 2012). Nevertheless, Fambrini *et al.* (2014a) demonstrated that much more complex regulatory system stays behind the *Chry2* phenotype.

The goals of this study were: a) identify the gene *CYC2c* in the diploid *H. decapetalus*, one of the parents of *H.* × *multiflorus*; b) identify the *CYC2c* alleles of the interspecific hybrid *H.* × *multiflorus* including the promoter region; c) analyze the transcription of *CYC2c* in flowers with different symmetry to evaluate the putative correlation between the phenotype

of the *Chry* mutant of sunflower and the triploid *H. × multiflorus* in the origin of the ligulate inflorescence in the *Helianthus* genus, in which this inflorescence type is extremely unusual (Funk *et al.* 2009).

MATERIAL AND METHODS

Plant material and growth conditions

Seeds of *H. annuus* (inbred line EF2 from Department of Agriculture, Food and Environment, University of Pisa, Italy) and *H. decapetalus* (Accession PI 468697 from USDA, USA) were germinated at 23 ± 1 °C in Petri dishes in the dark. Rhizomes of *H. × multiflorus* "Soleil d'Or" (a ligulate inflorescence type from L'Erbaio della Gorra, Casalborgone, Torino, Italy) and seedlings of *H. annuus* and *H. decapetalus* were transferred to 30-cm diameter pots containing a mixture of soil and sand. Plants were grown in a growth chamber at 23 ± 1 °C under a 16-h photoperiod (400 μmol photons m^{-2} s^{-1}). Irradiance was provided by a metal halide lamps (Kolorarc Daylight, MBID400/T/H 400 Watts, Betchworth, England).

Morphological analysis of organs in flowers of *H. × multiflorus*

In *H. × multiflorus* "Soleil d'Or", we performed the morphological analysis of floral organs of ray flowers (1st whorl) and "disk-flowers" transformed in ray-like flowers (2nd whorl) taking into consideration the following parameters at the final stage of anthesis: size of corolla, presence of anthers, filaments, style, stigma and size of ovaries. Three-four inflorescences were removed from 3-4 *H. × multiflorus* "Soleil d'Or" random plants. For each flower 20-30 organs were measured. Organ size was measured with an eyepiece micrometer, using a Wild Makroskop M420 inverted microscope (Wild Heerbrugg Ltd., Heerbrugg, Switzerland) as previously described (Fambrini *et al.* 2014b).

DNA extraction and isolation of *CYC* genes from *H. decapetalus* and *H.* × *multiforus*

DNA was extracted from young leaves (2-3 cm long) of *H. decapetalus* and *H.* × *multiflorus* using a modified CTAB procedure (Barghini *et al.* 2015). To isolate the full-length introns/exons region of the *CYC2c* gene the primer combination F73/CYC11 was used (S1). The PCR conditions were: 94 °C for 4 min, 35 cycles (30 s at 94 °C, 30 s at 65 °C, 1 min 50 sec at 72 °C), 72 °C for 7 min. The PCR product was separated by electrophoresis on a 1.5% TAE–agarose gel and visualized with Gel Red TM Nucleic Acid Stain (Biotium) under UV light.

From *H.* × *multiflorus* DNA PCRs were performed also to evaluate the presence of the *CYC2c* allele derived from the genome of the *H. annuus* parent. In particular, the primer combinations F73/CYCREV6, RAG6F/CYC11 and F71/73R were used (S1). The PCR conditions were previously described (Fambrini *et al.* 2011; 2014a,b).

The amplified products were purified using the WizardV R SV Gel and PCR Clean-UP System (Promega Italia, Milano, Italy), ligated into the pGEM-T Vector (Promega), and transformed in *Escherichia coli* JM109 competent cells (Promega). Plasmid DNA was prepared using WizardV R Plus Minipreps DNA Purification Kit (Promega). Several clones were sequenced on both strands.

Database searches, sequence and phylogenetic analyses

Sequence similarity searches were carried out using the BLAST alignment program against the public database of the National Center for Biotechnology Information (NCBI) (Altschul *et al.* 1997); furthermore, PROSITE and PFAM databases were searched to identify conserved domains (Bateman *et al.* 2002; Falquet *et al.* 2002). The TSSPlant program (http://www.softberry.com/berry.phtml) were used to identify the Transcription Start Site (TSS). The amino acid sequences of CYC TFs of several species were aligned by ClustalW

multiple sequence alignment program (Thompson *et al*. 1994) developed by the Kyoto University Bioinformatic Center (http://www.genome.jp/tools/clustalw/) with slow pairwise alignment and subjected to phylogenetic analysis (S2). Phylogenetic analysis was performed using programs from the PHYLIP pachage, PHYLogeny Inference Package, Version 3.64 (Felsenstein 1989). The phylogenetic relationships between species were inferred by a Dayhoff PAM distance matrix (Kosiol & Goldman 2005) generated by PROTDIST program and then subjected to NEIGHBOR program for the UPGMA clustering. Lastly, the strict consensus tree was obtained by CONSENSE program with the Majority rule option (Margush & McMorris 1981). As support for the tree obtained, a bootstrap analysis with 100 replicates was performed by the SEQBOOT program (Felsenstein 1985). The *Oryza sativa* PCF1 (OsPCF1, GenBank accession number XP_015633587) amino acid sequence was used as outgroup.

The databases of sunflower repeated elements (SUNREP, Natali *et al*. 2013) and that of Triticeae repetitive elements (TREP, http://wheat.pw. usda.gov/ITMI/Repeats) were also screened with the BLASTN and BLASTX algorithms. Detailed sequence analysis was performed with the DOTTER graphical dot-plot program (Sonnhammer & Durbin, 1995). Sequence data from this article were deposited in GenBank under the accession numbers MG797677, for *CYC2c* of *H. decapetalus* (*HdCYC2c*), MG797678 and MG797679, for the two alleles of *CYC2c* of *H. × multiflorus* (*HxmCYC2cA* and *HxmCYC2cB*, respectively).

Gene expression analysis by real-time RT-PCR (qPCR)

To analyze *CYC2c* transcript levels, total RNA extractions were carried out from corollas of ray (RF) and disk (DF) flowers of *H. annuus*, *H. decapetalus* and *H. × multiflorus*. In *H. × multiflorus* the DF are substituted by flower with zygomorphic corolla. For this reason, we have introduced the name ray flowers (RF) internal (I) for this type of flowers (RFI); in contrast, the real ray flowers were named external (E) RF (RFE). More precisely, total RNA

was extracted from corollas of RF (0.7-1.0 cm long without ovary, HaRF) and DF (0.5-0.7 cm, without ovary style-stigma and anthers, HaDF) of *H. annuus*; corollas of RF (0.5-0.7 cm long without ovary, HdRF) and corollas of DF (0.4-0.6 cm, without ovary style-stigma and anthers, HdDF) of *H. decapetalus*; corollas of young RFE (1.0-1.2 cm long without ovary, H×mRFE) and corolla of young RFI (0.5-1.0 cm long without ovary, H×mRFI) of *H × multiflorus*. Total RNA was extracted with the TriPure Isolation Reagent, according to the manufacturer's instructions (Roche Diagnostics GmbH, Mannheim, Germany). To exclude DNA contamination, digestion of extracts was performed with DNase I-RNase free (Dasit Sciences S.r.l., Cornaredo, Milan, Italy) as previously described (Sambrook & Russell 2001). Absence of genomic DNA contamination in DNase I-treated samples was tested using a PCR approach with primers PSTOP and INT1R designed to amplify a DNA fragment containing a partial exonic and intronic region of the *CYC2c* gene (S2). To determine the integrity of the RNA and to ensure that equal amounts of RNA were added to each reaction, 1 μg of RNA from each sample was separated via gel electrophoresis in formaldehyde-formamide gels.

Total RNA (1 μg) was used with the iScript™ cDNA synthesis kit (BIO-RAD) to produce the first strand cDNA. Real-time quantitative PCRs were performed using a Real-time Step One (Applied Biosystem, Thermo Fisher Scientific Inc, USA) and gene-specific primers for *CYC2c* and *Ha-18S* mRNA. Quantitative PCR was performed using 20 ng of cDNA and Power SYBR Green RNA-to-Ct 1 Step Kit (Applied Biosystem, Thermo Fisher Scientific Inc. USA, cat. Num. 4389986), according to the manufacturer's instructions. The thermal cycling conditions of RT-PCR were as follows: Reverse transcription: 48 °C-30'; Activation: 95 °C-10'; Cycling: 40 cycles 95 °C-15'/ 59 °C-30'; Melt curve: 95 °C-15"/60° C-15"/95 °C-15". Relative quantification of specific mRNA levels was performed using the comparative $2^{-\Delta\Delta C_T}$ method (Livak & Schmittgen 2001). Briefly, the C_T values of the amplified regions in all samples were normalized with the C_T values of the reference housekeeping

gene (18S ribosomal RNA, *Ha-18S* mRNA, GenBank accession number KF767534.1) to eliminate the variations caused by sample handling. In addition, mRNAs from roots of *H. annuus* (HaR) were used as reference sample. For each gene the normalized C_T values in HaR were subtracted from the corresponding C_T values in the analyzed organs and also in the same HaR. The derived values ($\Delta\Delta C_T$), 0 for the HaR, were inserted in the formula $2^{-\Delta\Delta C_T}$ (according to Livak & Schmittgen 2001) that returns 1 for the reference sample and the final relative quantification for the examined genes in the analyzed organs. Melt-curve analyses were performed after the PCR. A single distinct peak was observed for both the target (*CYC2c*) and control (*Ha-18S*) genes indicating the specific amplification of a single product. The data were the average of three biological replicates (sampled from plants grown at different times), each including three technical replicates. The software Real-time Step One v2.3, provided with the instrument by which we carried out the qPCR, was used. *Ha-18S* was used as the reference gene based on preliminary data that revealed consistent expression levels regardless of this organ type. In particular, the *Ha-18S* was preferred to other putative housekeeping genes [i.e., *actin* (*Ha-ACT*), *phosphoglycerate kinase 2* (*PGK2*) and *SAND*, GenBank accession numbers AF282624.1, HM490307 and GE516373, respectively]. The primers used to amplify *Ha-18S* cDNA were 18F and 18R while the primers used to amplify *CYC2c* cDNA were Q4F and Q4R (S2).

Statistical analysis

For Table 1 values are means ± SD from 3-4 independent experiments (plants or inflorescences), with 3 replicates (inflorescences or flowers). For each flower 20-30 organs were measured. Data were analyzed using Student's *t*-test. For expression analysis, in each experiment, the values displayed on graphs are means ± SD from 3 independent qPCR analyses with 3 different RNA replicates for each organ or for HaR. All qPCR analyses were performed with three different replicates in each run for all biological samples. Data were

RESULTS AND DISCUSSION

Morphological analysis of floral organs in inflorescences in *H. × multiflorus*

A morphological analysis of ray flowers and "disk flowers" modified in ray-like flowers was performed in inflorescences of *H. × multiflorus* "Soleil d'Or". The results showed slight differences between ray flower placed in the 1st whorl and ray-like flowers placed in the 2nd whorl (Table 1). Significant differences were detected only for corolla width and number of anther/flower. The high number of anther/flower in ray-like flowers might be a reminiscence of the origin of these flowers from hermaphrodite actinomorphic disk flowers (Mizzotti *et al.* 2015). The presence of anther and styles also in true ray flowers (Table 1, Fig. 1D) was surprising. Nevertheless, filaments were aberrant and anthers were undeveloped without pollen grains (Figs. 1D, E) some reproductive organs displayed an abnormal morphology and a homeotic transformation in petaloid structure. Many styles were mono-stigmatic (Fig. 1D). In sunflower, the presence of male organs in ray flowers is not usual. However, when the coding region of the *HaCYC2c* gene is interrupted by the insertion of a defective CACTA-like TE (*Tetu1*) (Fambrini *et al.* 2011) or by the insertion of retrotransposons (Chapman *et al.* 2012), ray flowers switch from zygomorphic to actinomorphic, resembling disc flowers; this trait is peculiar of *tubular ray flower* (*turf*) and *tubular-rayed* (*tub*) mutants (Fambrini *et al.* 2011; Chapman *et al.* 2012). In *turf*, flowers tubular-like ray flowers maintain their positional identity but achieved hermaphrodite features (Fambrini *et al.* 2011; Mizzotti *et al.* 2015). In addition, the incomplete filaments attached at the base of the corolla tube of ray flowers of sunflower are likely a reminiscent of ancestral stamens. The same outcome is displayed by sterile zygomorphic ray flowers of *Gerbera hybrida* (Broholm *et al.* 2008). It is likely that in

Asteraceae the radiate capitulum evolved from a discoid inflorescence. The pseudanthium of sunflower originated after a shift of the more external whorl of actinomorphic flowers in zygomorphic flowers. Nevertheless, these zygomorphic flowers may have maintained a latent potential to develop reproductive organs. In fact, *Tetu1* excision restored the wild type *HaCYC2c* allele, but the excisions also generated footprints and so new *HaCYC2c* alleles (Fambrini *et al.* 2014b). These mutations occurred at the TCP basic motif and caused a change in ray flower phenotype. In particular, ray flowers were often able to differentiate malformed stamens that produced pollen and styles with mono-branched stigma (Fambrini *et al.* 2014b). These data strengthen the assumption of the pleiotropic effects of the *HaCYC2c* gene, which control both ray flower symmetry and reproductive organ development, likely interacting with cell cycle and/or flower organ identity genes (Dezar *et al.* 2003; Fambrini & Pugliesi 2017a). Alternatively, other *CYC*-like gene, controlled by *HaCYC2c,* could be involved in the development of both male and female reproductive organs. It was shown that all genes of CYC2 clade were expressed in both ray flowers and floral reproductive organs (stamen, stigma plus style, and ovary) of sunflower (Chapman *et al.* 2008; Tähtiharju *et al.* 2012).

Sequence and phylogenetic analyses of *CYC2c* genes in *H. decapetalus* and *H. × multiflorus*

The analysis of the sequences isolated from *H. decapetalus* (*HdCYC2c*) and *H. × multiflorus* (*HxmCYC2cA* and *HxmCYC2cB*) are summarized in Table 2. The coding sequence (CDS) of *HdCYC2c* (S3A) encodes for a putative 381 amino acids long peptide (S3B). The CDS of *HxmCYC2cA* from *H. × multiflorus* (Table 2, Fig. 2A, S4A) encodes for a putative 383 amino acids long peptide (S4B). PCR experiments suggested the presence of different alleles in the *H. × multiflorus* genome, as expected because of the hybrid nature of this species. Indeed, we amplified and sequenced a second PCR-product from the *H. × multiflorus*

genome (S5, S6A). This second allele, named *HxmCYC2cB* (S6B), encoded for a putative peptide with the same properties of HxmCYC2cA (Table 2). Notably, alignment of the two sequences reveals an insertion of 865 bp in the 5' untranslated region (5'-UTR) of *HxmCYC2cB*, 53 bp upstream of the start codon and two bp downstream the TSS (Fig. 2B, S6A). In the three genes, the intron (86 bp) is positioned outside the CDS, 11 bp after the stop codon, within the 3'-UTR (Table 2, Fig. 2A, S3A, S4A). The intron position was deduced from a comparison of *CYC2c* cDNAs and the genomic DNA sequences. The three amino acid sequences showed the two highly conserved domains TCP and R (S3B, S4B, S6B). Nevertheless, the ECE residues, present in most proteins of CYC2 clade are absent in these TFs.

Sequence analysis performed in public and custom databases and DOTTER analysis suggests that the insertion in *HxmCYC2cB* is part of a TE belonging to the CACTA superfamily. We named this TE: *CACTA Transposable Element of H. × multiflorus1* (*CTEHM1*) (Fig. 3, S7A, B). The name 'CACTA' refers to the flanking terminal inverted repeats (TIRs), which are 10-28 bp long and terminate in a conserved 5'-CACTA-3' motif (Wicker *et al.* 2003). CACTA TEs include *En/Spm* (Suppressor-mutator/enhancer elements), one of the original maize TEs, first described in the 1940s by Barbara McClintock and Peter Peterson (Pereira *et al.* 1986; Fedoroff *et al.* 1995; Schulman & Wicker 2013). CACTAs are class II TEs that utilize a cut-and-paste mechanism for transposition, which requires a transposase enzyme (Feschotte & Pritham 2007).

The insertion we found in *HxmCYC2B* 5'-UTR near to the TSS displays some clear CACTA signatures (Fig. 3, S7A, B). It is flanked by short TIRs (13 bp, 5'-CACTACAAGAAAC-3') that end with an intact 5'-CACTA-3' motifs known to be recognition sequences for the transposase protein (Lewin 1997). The 865 bp insertion consists of the CACTA element (862 bp) and a perfect 3 bp (TTA) target-site duplication (TSD), which represent a hallmark of this TE superfamily (Fig. 2). It also contains several subterminal repeats (sub-TRs) in direct and inverted orientations that in CACTA elements produce the

characteristic "transposon signature" (Fig. 3, S7A, B) (Wicker *et al.* 2003). The sub-TRs and inverted repeats display a specific pattern when the CACTA sequence is plotted against itself with dot-plot (Fig. 3). The most common motif is a 11 bp sequence (TGTCGCCGCTA) repeated in direct and inverted orientation at the 5'- and 3'-ends (S7A, B). Dot-plot analysis also showed that the TE contains pattern of tandem repeats of variable length. The repeated sequence units range in size from 2 to 22 bp. (S7A, B). The integrity of the TIRs and that of the sub-TRs is essential for the transposition (Schiefelbein *et al.* 1988). The 862 bp insertion in the 5'-UTR of *HxmCYC2cB* is likely a deletion-derivate TE. In fact, it lacks af obvious coding sequences for TNPD that represents the putative transposase required for the excision/integration process during transposition and TNPA a protein with multiple functions, some of them reflecting its ability to bind DNA (Gierl *et al.* 1989; Trentmann *et al.* 1993). We suggest that this insertion is a truncated version of a TE, likely generated from an imperfect excision of an entire transposable element. Because of its small size, the insertion should be classified as a small non autonomous CACTA (SNAC; Wicker *et al.* 2003). Interestingly, the *Chrysanthemoides* (*Chry*) mutant of sunflower showed the insertion in the promoter region of *HaCYC2c* gene of a SNAC, likely occurred in the first hundred years after the sunflower reached Europe from North America (Heiser 1976). The insertion upstream the coding region of *HaCYC2c* causes the ectopic expression of the gene and the shift from actinomorphic to zygomorphic disk flowers (Chapman *et al.* 2012; Fambrini *et al.* 2014a), originating an inflorescence apparently similar to that of *H. x multiflorus*. A comparison of *CTEHM1* from *H x multiflorus* with the SNAC sequences of both *double flowered* (*dbl*) and *Chry2* mutants was performed. *CTEHM1* displayed a relative high identity with the CACTAs of both *dbl* (601/1106, 53.9%) and *Chry2* (615/1142, 53.9%) but confined to the 5' and 3' regions (S8A). By contrast, a high identity (976/1034, 94.4%) between the TEs of *dbl* and *Chry2* (S8B) was detected. It is likely that both SNACs of *dbl* and *Chry2* derived by the same complete TE.

The allele *CYC2c* derived from *H. annuus* genome should be also present in *H. ×*
multiflorus. However, sequencing all PCR-amplified products displayed a higher identity and
a lower number of gaps with *H. decapetalus* than *H. annuus* (Table 3). In addition, the
results of several PCRs conducted in both the 5'- and 3'-region of the *CYC2c* gene
corroborated the possibility that the putative allele *CYC2c* resulting from the sunflower
parent should be deleted or highly modified. For example, in the 5'-region, the sequence
analysis of the amplified PCR-product obtained by the primer combination F71/73R showed
two insertions of 54 and 30 bp in *H. decapetalus* (S9A, B). This should allow an easy
distinction between the two alleles. Indeed, all the amplified PCR-products obtained from *H.*
× multiflorus DNA displayed the same size of *H. decapetalus* samples (S9C). Finally, we
sequenced 19 clones from two regions of the *H×mCYC2c* gene and all sequences showed a
higher similarity with *HdCYC2c* than *HaCYC2c* (S10A, B). Together, these data suggest that
the *CYC2c* allele derived from *H. annuus* is lacking or it has undergone an extensive
genomic rearrangement. Indeed, molecular evidence suggests polyploid genomes display
dynamic and pervasive changes in DNA sequence and gene expression probably as a
response of "genomic shock" (McClintock 1984). The combination of evolutionarily divergent
genomes in polyploids resembles "genomic shock", leading to the activation of quiescent TE
that may induce sequence rearrangements for unequal and illegitimate recombination and
account for a variety of mutations such as deletions, insertions, frameshifts, inversions,
translocations and duplications (Lonnig & Saedler 2002; Chen & Ni 2006).

The phylogenetic tree clearly resolved the three CYC subclades (CYC1, CYC2, and
CYC3) (Fig. 4). HdCYC2c and H×mCYC2c were placed within a CYC2 subclade together
with HaCYC2c but distinct from it (Fig. 4). In addition, the three TFs belonged to a major
clade with GhCYC3, HaCYC2d, SvRAY1 and GhCYC5. *GhCYC3* and *SvRAY1* genes were
indicated to be strong candidates as regulators of ray flower identity in *Gerbera hybrida* and
Senecio vulgaris, respectively (Kim *et al.* 2008; Tähtiharju *et al.* 2012). In sunflower,
HaCYC2d is expressed in ray flowers suggesting a function in zygomorphism (Chapman *et*

al. 2008; Tähtiharju *et al.* 2012). However, its expression was detected at different developmental stages of disk flowers also suggesting its role in the development of organs of this type of flowers (Chapman *et al.* 2008). In *G. hybrida*, *GhCYC5* was expressed in ray and in the intermediate trans flower primordia and at much lower levels in disc flower primordia (Tähtiharju *et al.* 2012). More recently functional analysis showed a main role of *GhCYC5* to regulate the flower density of the gerbera inflorescence (Juntheikki-Palovaara *et al.* 2014). Although most genes playing a role in zygomorphism belong to the CYC2 clade, not all these are orthologs due to the independent duplications in various families (Feng *et al.* 2006; Broholm *et al.* 2008; Kim *et al.* 2008; Wang *et al.* 2008; Jabbour *et al.* 2014). Researches in Poaceae, Zingiberales and Commelinales also suggest a possible role of *Tb1*-like genes in the independent evolution of zygomorphism (Yuan *et al.* 2009; Bartlett & Specht 2011; Preston & Hileman 2012). Additionally, a role of *CIN*-like genes in the complex regulatory network controlling monosymmetry in *Petrocosmea* spp. (Gesneriaceae) (Yang *et al.* 2015) and *Orchis italica* (De Paolo *et al.* 2015) has been shown.

In *H. × multiflorus* the *CYC2c* gene is expressed ectopically in ray-like corollas

HaCYC2c is specifically expressed in ray flowers (HaRF) of sunflower plants and not in disk flowers (HaDF) (Chapman *et al.* 2008; Tähtiharju *et al.* 2012). By contrast, in *Cry* mutant plants the gene is ectopically expressed because of the presence of the TE insertion, resulting in zygomorphic disk flowers (Chapman *et al.* 2012; Fambrini *et al.* 2014a). We analyzed the expression pattern of *CYC2c* genes by real-time RT-PCR (qPCR) in *H. annuus*, *H. decapetalus* and *H. × multiflorus* in order to check whether ligulate capitula of *H. × multiflorus* "Soleil d'Or" are linked to ectopic expression of *CYC2c* genes. As expected, *CYC2c* is specifically expressed in the corolla of ray flowers in *H. annuus* (HaRF) and *H. decapetalus* (HdRF), according with that observed in sunflower for the *HaCYC2c* gene (Chapman *et al.* 2008; Tähtiharju *et al.* 2012). By contrast, *H. × multiflorus* displays an high

transcription level in corollas of both true ray (H×mRFE) and ray-like flowers (H×mRFI) (Fig. 5), similarly to the expression detected in both *dbl* and *Chry2* inflorescences (Chapman *et al.* 2012; Fambrini *et al.* 2014a). Therefore, misexpression of *H×mCYC2c* gene established the shift from actinomorphic to zygomorphic disk flowers originating a ligulate head.

There is evidence that TE insertion in the proximity of genes can influence the regulation of gene expression through a variety of mechanisms (Hirsh & Springer 2017; Dubin *et al.* 2018). In *A. thaliana*, genes lying close to TE sequences were expressed at lower levels compared with the genome-wide distribution of gene expression, suggesting that TE insertions tended to reduce and to a smaller extent to increase the expression of neighbouring genes, with the effect depending of the type of TE and the genic region into which it was transposed (Hollister & Gaut 2009). However, in other cases, TE insertions into the promoter region of a gene determined overexpression of that gene (Yang *et al.* 2005; Chapman *et al.* 2012). In addition, the presence of a TE within the upstream region of a gene encoding a pentatricopeptide repeat (PPR) protein of *A. thaliana* was associated with higher transcript abundance of this gene (Stuart *et al.* 2016). The apparent transcriptional activation, linked with the presence of a TE belonging to the *HELITRON1* family, indicates that this element may carry regulatory information that alters the expression of genes downstream of the TE insertion site (Stuart *et al.* 2016). There are also contrasting evidences about the influence on gene expression of TE insertion in the 5'-UTR or near to TSS (Hirsh & Springer 2017). For example, the insertion of a *Copia*-like retrotransposon in the 5'-UTR region of the *FATTY ACID ELONGATION1* (*FAE1*) gene controlling the content of erucic acid in *Sinapis alba* negatively affects the *FAE1* expression (Zeng & Cheng 2014). By contrast, the insertion of a 1 kb TE in the 5'-UTR in the *HvAACT1* (*HvMATE*), the major gene for aluminium (Al) tolerance in barley, was associated with increased gene expression and Al tolerance (Ma *et al.* 2016). During the study on the role of MORC1 as important repressor of TEs in male germ cells of mouse, Pastor *et al.* (2014) identified three genes (*Nebulin*, *Tmc2* and *Cdkl4*) containing an RLTR10 TE immediately upstream of the TSS. All

three genes showed a statistically significant increase in expression. Indeed, the 5'-UTR contains regulatory sites, and TE insertions may contribute extra regulatory sites leading to increased transcription of the whole gene (Hirsh & Springer 2017). There are also several evidences that TEs may provide novel promoters that would create novel TSSs for nearby genes (Feschotte 2008). This should be the case of the *CTEHM1* TE insertion in the *HxmCYC2cB* allele, which provides a new TSS, internal to the TE (S6). Finally, TEs are known to be a large source of criptic promoters to drive expression of nearby genes, and regulatory regions which can be co-opted by the host genome and contribute to the evolution of regulatory networks (Hirsh & Springer 2017).

CONCLUSIONS

Noteworthy, in a conference devoted to asters and perennial sunflowers held at Chiswick, England, in 1891, DeWar (1893) stated "It ..... seemed probable .... that *H. × multiflorus* is a garden hybrid of *H. annuus* and perhaps *H. decapetalus*, and that it has been produced several times. *H. annuus* is the only known sunflower besides *H. × multiflorus* which assumes double forms (ligulate inflorescence), and its hybrids might inherit this tendency .....". It is likely to hypothesize that the thought of DeWar had proceeded from the knowledge of the mutant *Chry* of *H. annuus,* already present in Europe for about 300 years (Heiser 1976). Here, we demonstrated that in *H. × multiflorus* "Soleil d'Or" the very attractive ligulate inflorescence is correlated to ectopic expression of *CYC2c* genes, probably linked to the insertion of a CACTA defective element. However, other *CYC* genes from the *H. × multiflorus* genome could be also involved in generating unusual inflorescence without actinomorphic disk flowers. To evaluate the role of *CTEHM1* insertion to induce ectopic expression of *CYC2c* genes more in depth, cultivars of *H. × multiflorus* with radiate inflorescence (i.e. "Major" and "Meteor") should be analysed; nevertheless, the new data collected in *H. × multiflorus* provide further information about the involvement of TEs in the

evolution of atypical inflorescence morph in the genus *Helianthus* in which the basic inflorescence architecture is radiate.

CONFLICT OF INTEREST

Marco Fambrini, Matteo Bellanca, M. Costa Muñoz, Gabriele Usai, Andrea Cavallini & Claudio Pugliesi declare that they have no conflict of interest.

REFERENCES

Abbott R.J., James J.K., Milne R.I., Gillies A.C.M. (2003) Plant introduction, hybridization and gene flow. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **358**, 1123-1132.

Aggarwal P., Gupta M.D., Joseph A.P., Chatterjee N., Srinivasan N., Nath U. (2010) Identification of specific DNA binding residues in the TCP family of transcription factors in *Arabidopsis*. *The Plant Cell*, **22**, 1174-1189.

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389-3402.

Barghini E., Mascagni F., Natali L., Giordani T., Cavallini A. (2015) Analysis of the repetitive component and retrotransposon population in the genome of a marine angiosperm, *Posidonia oceanica* (L.) Delile. *Marine Genomics*, **24**, 397-404.

Bartlett M.E., Specht C.D. (2011) Changes in expression pattern of the *TEOSINTE BRANCHED1*-like genes in the Zingiberales provide a mechanism for evolutionary shifts in symmetry across the order. *American Journal of Botany*, **98**, 1-17.

Bateman A., Birney E., Cerruti L., Durbin R., Etwiller L., Eddy S.R., Griffiths-Jones S., Howe K.L., Marshall M., Sonnhammer E.L.L. (2002) The pfam protein families database. *Nucleic Acids Research*, **30**, 276-280.

Bello M.A., Álvarez I., Torices R., Fuertes-Aguilar J. (2013) Floral development and evolution of capitulum structure in *Anacyclus* (Anthemideae, Asteraceae). *Annals of Botany*, **112**, 1597-1612.

Berger B.A., Thompson V., Lim A., Ricigliano V., Howarth D.G. (2016) Elaboration of bilateral symmetry across *Knautia macedonica* capitula related to changes in ventral petal expression of *CYCLOIDEA*-like genes. *EvoDevo*, **7**, 8.

Berti F., Fambrini M., Turi M., Bertini D., Pugliesi C. (2005) Mutations of corolla symmetry affect carpel and stamen development in *Helianthus annuus*. *Canadian Journal of Botany*, **83**, 1065-1072.

Broholm S.K., Tähtiharju S., Laitinen R.A.E., Albert V.A., Teeri T.H., Elomaa P. (2008) A TCP domain transcription factor controls flower type specification along the radial axis of the *Gerbera* (Asteraceae) inflorescence. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 9117-9122.

Broholm S.K., Teeri T.H., Elomaa P. (2014) Molecular control of inflorescence development in Asteraceae. *Advances in Botanical Research*, **72**, 297-333.

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

Carlson S.E., Howard D.G., Donoghue M.J. (2011) Diversification of *CYCLOIDEA*-like genes in Dipsacaceae (Dipsacales): implications for the evolution of capitulum inflorescences. *BMC Evolutionary Biology*, **11**, 325.

Chapman M.A., Leebens-Mack J.H., Burke J.M. (2008) Positive selection and expression divergence following gene duplication in the sunflower *CYCLOIDEA* gene family. *Molecular Biology Evolution*, **25**, 1260-1273.

Chapman M.A., Tang S., Draeger D., Nambeesan S., Shaffer H., Barb J.G., Knapp S.J., Burke J.M. (2012) Genetic analysis of floral symmetry in Van Gogh's sunflowers reveals independent recruitment of *CYCLOIDEA* genes in the Asteraceae. *PLoS Genetics*, **8**(3), e1002628.

Chen Z.J., Ni Z. (2006) Mechanism of genomic rearrangements and gene expression changes in plant polyploids. *Bioessays*, **28**, 240-252.

Citerne H.L., Luo D., Pennington R.T., Coen E., Cronk Q.C.B. (2003) A phylogenomic investigation of *CYCLOIDEA*-like *TCP* genes in the Leguminosae. *Plant Physiology*, **131**, 1042-1053.

Citerne H., Jabbour F., Nadot S., Damerval C. (2010) The evolution of floral symmetry. *Advances in Botanical Research*, **54**, 85-137.

Citerne H.L., Le Guilloux M., Sannier J., Nadot S., Damerval C. (2013) Combining phylogenetic and syntenic analyses for understanding the evolution of TCP ECE genes in Eudicots. *PLoS ONE*, **8**, e74803.

Cockerell T.D.A. (1915) Specific and varietal characters in annual sunflowers. *The American Naturalist*, **49**, 609-622.

Coen E.S. (1996) Floral symmetry. *The EMBO Journal*, **15**, 6777-6788.

Costa M.M.R., Fox S., Hana A.I., Baxter C., Coen E. (2005) Evolution of regulatory interactions controlling floral asymmetry. *Development*, **132**, 5093-5101.

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

Cubas P., Coen E., Zapater J.M.M. (2001) Ancient asymmetries in the evolution of flowers. *Current Biology*, **11**, 1050-1052.

Cubas P. (2004) Floral zygomorphy, the recurring evolution of a successful trait. *Bioessays*, **26**, 1175-1174.

Damerval C., Le Guilloux M., Jager M., Charon C. (2007) Diversity and evolution of *CYCLOIDEA*-like *TCP* genes in relation to flower development in Papaveraceae. *Plant Physiology*, **143**, 759-772.

De Paolo S., Gaudio L., Aceto S. (2015) Analysis of the *TCP* genes expressed in the inflorescence of the orchid *Orchis italica. Scientific Reports*, **5**, 16265.

DeWar D. (1893) Perennial sunflowers. *Journal of the Royal Horticultural Society*, **15**, 26-39.

Dezar C.A., Tioni M.F., Gonzalez D.H., Chan R.L. (2003) Identification of three MADS-box genes expressed in sunflower capitulum. *Journal of Experimental Botany*, **54**, 1637-1639.

Doebley J., Stec A., Hubbard L. (1997) The evolution of apical dominance. *Nature*, **386**, 485-488.

Dubin M.J., Mittelsten Scheid O., Becker C. 2018 Transposons: a blessing curse. *Current Opinion in Plant Biology*, **42**, 23-29.

Falquet L., Pagni M., Bucher P., Hulo N., Sigrist C.J., Hofmann K., Bairoch A. (2002) The PROSITE database, its status in 2002. *Nucleic Acids Research*, **30**, 235-238.

Fambrini M., Bertini D., Pugliesi C. (2003) The genetic basis of a mutation that alters the floral symmetry in sunflower. *Annals of Applied Biology*, **143**, 341-347.

Fambrini M., Michelotti V., Pugliesi C. (2007) The unstable *tubular ray flower* allele of sunflower: inheritance of reversion to wild type. *Plant Breeding*, **126**, 548-550.

Fambrini M., Salvini M., Pugliesi C. (2011) A transposon-mediate inactivation of a *CYCLOIDEA*-like gene originates polysymmetric and androgynous ray flowers in *Helianthus annuus*. *Genetica*, **139**, 1521-1529.

Fambrini M., Salvini M., Basile A., Pugliesi C. (2014a) Transposon-dependent induction of Vincent van Gogh's sunflowers: Exceptions revealed. *genesis*, **52**, 315-327.

Fambrini M., Basile A., Salvini M., Pugliesi C. (2014b) Excisions of a defective transposable CACTA element (*Tetu1*) generate new alleles of a *CYCLOIDEA*-like gene of *Helianthus annuus*. *Gene*, **549**, 198-207.

Fambrini M., Pugliesi C. (2017a) CYCLOIDEA 2 clade genes: key players in the of floral symmetry, inflorescence architecture, and reproductive organ development. *Plant Molecular Biology Reporter*, **35**: 20-36.

Fambrini M., Pugliesi C. (2017b) Mobilization of the *Tetu1* transposable element of *Helianthus annuus*: evidnce for excision in different developmental stages. *Biologia Plantarum*, **61**, 55-63.

Fedoroff N., Schläppi M., Raina R. (1995) Epigenetic regulation of the maize Spm transposon. *Bioessays* **17**, 291-297.

Felsenstein J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783-791.

Felsenstein J. (1989) PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics*, **5**, 164-166.

Feng X., Zhao Z., Tian Z., Xu S., Luo Y., Cai Z., Wang Y., Yang J., Wang Z., Weng L., Chen J., Zheng L., Guo X., Luo J., Sato S., Tabata S., Ma W., Cao X., Hu X., Sun C., Luo D. (2006) Control of petal shape and floral zygomorphy in *Lotus japonicus. Proceedings of the National Academy of Sciences of the United States of America*, **103**, 4970-4975.

Feschotte C. (2008) Transposable elements and the evolution of regulatory networks. *Nature Review Genetics*, **9**, 397-405.

Feschotte C., Pritham E.J. (2007) DNA-transposons and the evolution of eukaryotic genomes. *Annual Review* of *Genetics*, **41**, 331-368.

Fick G.N. (1976) Genetics of floral color and morphology in sunflowers. *Journal of Heredity*, **67**, 227-230.

Frey M., Spring O. (2015) Molecular traits to elucidate the ancestry of *Helianthus* × *multiflorus. Biochemical Systematics and Ecology*, **58**, 51-58.

Funk V.A., Susanna A., Stuessy T.F. Robinson H. (2009) Classification of Compositae. In: Funk V.A., Susanna A., Stuessy T.F., Bayer R.J. (Eds), *Systematics, Evolution, and Biogeography of Compositae*. International Association for Plant Taxonomy (IAPT), Vienna, 171-192.

Garcês H.M.P., Spencer V.M.R., Kim M. (2016) Control of floret symmetry by *RAY3*, *SvDIV1B* and *SvRAD* in the capitulum of *Senecio vulgaris. Plant Physiology*, **171**, 2055-2068.

Gierl A., Saedler H., Peterson P.A. (1989) Maize transposable elements. *Annual Review* of *Genetics*, **23**, 71-85.

Harris E.M. (1995) Inflorescence and floral ontogeny in Asteraceae: a synthesis of historical and current concepts. *The Botanical Review*, **61**, 93-278.

Heiser C.B. Jr. (1976) The sunflower. University of Oklahoma Press, Norman, OK.

Heiser C.B. Jr., Smith D.M. (1960) The origin of *Helianthus multiflorus. American Journal of Botany*, **47**, 860-865.

Hileman L.C. (2014a) Bilateral flower symmetry - how, when and why? *Current Opinion in Plant Biology*, **17**, 146-152.

Hileman L.C. (2014b) Trends in flower symmetry evolution revealed through phylogenetic and developmental genetic advances. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **369**, 20130348.

Hirsh C.D., Springer N.M. (2017) Trasposable element influences on gene expression in plants. *Biochimica et Biophysica Acta*, **1860**, 157-165.

Hollister J.D., Gaut B.S. (2009) Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression *Genome Research*, **19**, 1419-1428.

Howarth D.G., Donoghue M.J. (2005) Duplications in *CYC*-like genes from Dipsacales correlate with floral form. *International Journal of Plant Science*, **166**, 357-370.

Howarth D.G., Donoghue M.J. (2006) Phylogenetic analysis of the ''ECE'' (CYC/TB1) clade reveals duplications predating the core eudicots. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 9101-9106.

Jabbour F., Cossard G., Le Guilloux M., Sannir J., Nadot S., Damerval C. (2014) Specific duplication and dorsoventrally asymmetric expression patterns of *Cycloidea*-like genes in zygomorphyc of Ranunculaceae. *PLoS ONE*, **9**(4), e95727.

Jeffrey C. (1977) Corolla forms in Compositae - some evolutionary and taxonomic speculations. In: Heywood V.H., Harborne J.B., Turner B.L. (Eds), *The Biology and Chemistry of the Compositae*, Vol 1. Academic Press, London, UK, 111-118.

Juntheikki-Palovaara I., Tähtiharjiu Lan T., Broholm K., Rijpkena A.S., Ruonala R., Kale L., Albert V.A., Teeri T.H., Elomaa P. (2014) Functional diversification of duplicated CYC2 clade genes in regulation of inflorescence development in *Gerbera hybrida* (Asteraceae). *The Plant Journal*, **79**, 783-796.

Kim M., Cui M.-L., Cubas P., Gillies A., Lee K., Chapman M.A., Abbott R.J., Coen E. (2008) Regulatory genes control key morphological and ecological trait transferred between species. *Science*, **322**, 1116-1119.

Kosiol C., Goldman N. (2005) Different versions of the Dayhoff rate matrix. *Molecular Biology Evolution*, **22**, 193-199.

Krizek E.M., Fletcher J.C. (2005) Molecular mechanisms of flower development: an armchair guide. *Nature Review of Genetics*, **6**, 688-698.

Lewin B. (1997) Transposons. In: Lewin B. (Ed), *Genes VI*. Oxford University Press, New York, 563-595.

Livak K.J., Schmittgen T.D. (2001) Analysis of relative gene expression data using Real-Time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods*, **25**, 402-408.

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.S. (1999) Control of organ asymmetry in flowers of *Antirrhinum*. *Cell*, **99**, 367-376.

Ma Y., Li C., Ryan P.R., Shabala S., You J., Liu J., Liu C., Zhou M. (2016) A new allele for aluminium tolerance gene in barley (*Hordeum vulgare* L.). *BMC Genomics*, **17**, 186.

Margush T., McMorris F.R. (1981) Consensus n-trees. *Bulletin* of *Mathematical Biology*, **43**, 239-244.

Martín-Trillo M., Cubas P. (2010) TCP genes: a family snapshot ten years later. *Trends in Plant Science*, **15**, 31-39.

McClintock B. (1984) The significance of responses of the genome to challenge. *Science*, **226**, 792-801.

Mizzotti C., Fambrini M., Caporali E., Masiero S., Pugliesi C. (2015) A *CYCLOIDEA*-like gene mutation in sunflower determines an unusual floret type able to produce filled achenes at the periphery of the pseudanthium. *Botany*, **93**, 171-181.

Moyroud E., Glover B.J. (2017) The evolution of diverse floral morphologies. *Current Biology* **27**, R941-R951.

Nath U., Crawford B., Carpenter B., Coen E. (2003) Genetic control of surface curvature. *Science*, **299**, 1404-1407.

Natali L., Cossu R.M., Barghini E., Giordani T., Buti M., Mascagni F., Morgante M., Gill N., Kane N.C., Rieseberg L., Cavallini A. (2013) The repetitive component of the sunflower genome as revealed by different procedures for assembling next generation sequencing reads. *BMC Genomics*, **14**, 686.

Pastor W.A., Stroud H., Nee K., Liu W., Pezic D., Manakov S., Lee S.A., Moissiard G., Zamudio N., Bourc'his D., Aravin A.A., Clark A.T., Jacobsen S.E. (2014) MORC1 represses transposable elements in the mouse male germline. *Nature Communication*, **5**, 5795.

Pereira A., Cuypers H., Gierl A., Sommer Z.S., Saedler H. (1986) Molecular analysis of the *En/Spm* transposable element system of *Zea mays. The EMBO Journal*, **5**, 835-841.

Preston J.C., Hileman L.C. (2009) Developmental genetics of floral symmetry evolution. *Trends in Plant Science*, **14**, 147-154.

Preston J.C., Hileman L.C. (2012) Parallel evolution of TCP and B-class genes in Commelinaceae flower bilateral symmetry. *EvoDevo*, **3**, 6.

Sambrook J., Russell D. (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schiefelbein J.W., Furtek D.B., Dooner H.K., Nelson O.E. Jr. (1988) Two mutations in a maize *bronze-1* allele caused by transposable elements of the *Ac-D*s family alter the quantity and quality of the gene product. *Genetics* **120**, 767-777.

Schulman A.H., Wicker T. (2013) A field guide to transposable elements. In: Fedoroff N.V., (Ed), *Plant Transposons and Genome Dynamics in Evolution*. Wiley-Blackwell, Oxford, UK, 15-40.

Sonnhammer E.L.L., Durbin R. (1995) A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene*, **167**, GC1-GC10.

Specht C.D., Howarth D.G. (2015) Adaptation in flower form: a comparative evodevo approach. *New Phytologist*, **206**, 74-90.

Spencer V., Kim M. (2017) Re"CYC"ling molecular regulators in the evolution and development of flower symmetry. *Seminars in Cell and Developmental Biology*, **17**, 30285-30289.

Spring O., Schilling E.E. (1990). The origin of *Helianthus × multiflorus and H. × laetiflorus* (Asteraceae). *Biochemical Systematics and Ecology*, **18**, 19-23.

Stuart T., Eichten S.R., Cahn J., Karpievitich Y.V., Borevitz J.O., Lister R. (2016) Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation. *eLife*, **5**, 20777.

Tähtiharju S., Rijpkema A.S., Vetterli A., Albert V.A., Teeri T.H., Elomaa P. (2012) Evolution and diversification of the *CYC/TB1* gene family in Asteraceae – a comparative study in gerbera (Mutisieae) and sunflower (Heliantheae). *Molecular Biology Evolution*, **29**, 1155-1166.

Theissen G. (2000) Evolutionary developmental genetics of floral symmetry: the revealing power of Linnaeus' monstrous flower. *Bioessays*, **22**, 209-213.

Thompson J.D., Higgins D.G., Gibson T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673-4680.

Trentmann S.M., Saedler H., Gierl A. (1993) The transposable element En/Spm-ecoded TNPA protein contains a DNA-binding and a dimerization domain. *Molecular Genetics and Genomics*, **238**, 201-208.

Uberti Manassero N.G., Viola I.L., Welchen E., Gonzalez D.H. (2013) TCP transcription factors: architectures of plant form. *Biomolecular Concepts*, **4**, 111-127.

Viola I.L., Reinheimer R., Ripoli R., Uberti Manassero N.G., Gonzalez D.H. (2012) Determinants of the DNA binding specificity of class I and class II TCP transcription factors. *The Journal of Biological Chemistry*, **287**, 347-356.

Wang Z., Luo Y., Li X., Xu S.L., Yang J., Weng L., Sato S., Tabata S., Ambrose M., Rameau C., Feng X.Z., Hu X.H., Luo D. (2008) Genetic control of floral zygomorphy in pea (*Pisum sativum* L.). *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 10414-10419.

Wicker T., Guyot R., Yahiaoui N., Keller B. (2003) CACTA transposons in Triticeae. A diverse family of high-copy repetitive elements. *Plant Physiology*, **132**, 52-63.

Xu Z., Cheng K., Li X., Yang J., Xu S., Cao X., Hu X., Yuan L., Ambrose M., Chen G., Mi H., Luo D. (2016) Trancriptional and post-trancriptional modulation of *SQU* and *KEW* activities in the control of dorsal-ventral asymmetric flower development in *Lotus japonicus. Molecular Plant*, **9**, 722-736.

Yang G., Yeon-Hee L., Jiang Y., Shi X., Kertbundt S., Hall T.C. (2005). A two-edged role for the transposable element *Kiddo* in the *rice ubiquitin2* promoter. *The Plant Cell*, **17**, 1559-1568.

Yang X., Pang H.-B., Liu B.-L., Qiu Z.-J., Gao Q., Wei L., Dong Y., Wang Y.-Z. (2012). Evolution of double positive autoregulatory feedback loops in CYC2 clade genes is associated with the origin of floral zygomorphy. *The Plant Cell*, **24**, 1834-1847.

Yang X., Zhao X.-G., Li C.Q., Liu J., Qiu Z.-J., Dong Y., Wang Y.-Z. (2015) Distinct regulatory changes underlying differential expression of *TEOSINTE BRANCHED1-CYCLOIDEA-PROLIFERATING CELL FACTOR* gene associated with petal variations in zygomorphic flower of *Petrocosmea* spp. of the family Gesneriaceae. *Plant Physiology*, **169**, 2138-2151.

Yuan Z., Gao S., Xue D.W., Luo D., Li L.T., Ding S.Y., Yao X., Wilson Z.A., Qian Q., Zhang D.B. (2009) *RETARDED PALEA1* controls palea development and floral zygomorphy in rice. *Plant Physiology*, **149**, 235-244.

Zeng F., Cheng B. (2014) Transposable element insertion and epigenetic modification cause the multiallelic variation in the expression of *FAE1* in *Sinapis alba*. *The Plant Cell*, **26**, 2648-2659.

Zhong J., Kellogg E.A. (2015) Duplication and expression of *CYC2*-like genes in the origin and maintenance of corolla zygomorphy in Lamiales. *New Phytologist*, **205**, 852-868.

Table 1. Morphological analysis of organs in flowers of *Helianthus × multiflorus* "Soleil d'Or". [*] Values are means ± SD from 3-4 independent experiments (plants or inflorescences), with 3 replicates (inflorescences or flowers). For each inflorescence, 20-30 organs were measured. Values followed by the same letter within a row are not significantly different (p = 0.05, according to Student's *t*-test). [**] Filaments were abnormal and anther were small without pollen grains. Some stamens displayed a partial homeotic transformation in petaloid structures. [***] Many styles were mono-stigmatic.

Characters	Ray flowers 1st whorl	Ray-like flowers of 2nd whorl
Length of the corolla (mm) [*]	32.28 ± 3.83 [a]	33.47 ± 4.49 [a]
Width of the corolla (mm)	14.61 ± 1.12 [a]	11.85 ± 0.50 [b]
No. of petal tips	2.92 ± 0.10 [a]	2.89 ± 0.13 [a]
No. of anthers/flower [**]	3.78 ± 0.20 [a]	4.54 ± 0.40 [b]
Length of the ovary (mm)	5.76 ± 0.07 [a]	5.68 ± 0.24 [a]
Width of the ovary (mm)	1.89 ± 0.35 [a]	1.62 ± 0.59 [a]
No. of style-stigmas/flower [***]	0.11 ± 0.15 [a]	0.23 ± 0.20 [a]

Table 2. Summary of several characteristics of the genes isolated from *Helianthus decapetalus* (*HdCYC2c*) and *H. × multiflorus* (*H×mCYC2cA* and *H×mCYC2cB*).

Gene	Length (bp)	CDS (bp)	Amino acids	Molecular Weight (Da)	Isoelectric Point (pI)	Intron length (bp)
HdCYC2c	1,801	1,146	381	43,652	7.26	86
H×mCYC2cA	1,807	1,152	383	43,679	7.25	86
H×mCYC2cB	2,671	1,152	383	43,679	7.25	86

Table 3. ClustalW (EMBOSS Needle) analysis among the *HaCYC2c* (*Helianthus annuus*), *HdCYC2c* (*H. decapetalus*) and *H×mCYC2cA* (*H. × multiflorus*) nucleotide sequences.

Comparison between species	Region	Length (bp)	Identity	Gaps
H. annuus vs H × multiflorus	Whole	1798/1807	1759/1808 (97.3%)	11/1808 (0.6%)
H. decapetalus vs H × multiflorus	Whole	1801/1807	1779/1807 (98.5%)	6/1807 (0.3%)
H. annuus vs H × multiflorus	CDS	1146/1152	1118/1152 (97.0%)	6/1152 (0.5%)
H. decapetalus vs H × multiflorus	CDS	1146/1152	1134/1152 (98.4%)	6/1152 (0.5%)
H. annuus vs H × multiflorus	5'-Region	498/497	492/498 (98.8%)	1/498 (0.2%)
H. decapetalus vs H × multiflorus	5'-Region	497/497	490/497 (98.6%)	0/497 (0.0%)
H. annuus vs H × multiflorus	3'-Region	154/158	149/158 (94.3%)	4/158 (2.5%)
H. decapetalus vs H × multiflorus	3'-Region	158/158	155/158 (98.1%)	0/158 (0.0%)

FIGURE CAPTIONS

Fig. 1. The phenotype of *Helianthus × multiflorus* "Soleil d'Or". A: Mature plants of *H. × multiflorus* grown in 30-cm diameter pot. B: Inflorescence of *H. × multiflorus* at complete anthesis; note that all flowers are zygomorphic. C: Inflorescence of *Chrysantemoides2* (*Chry2*), a mutant of *Helianthus annuus*. D: Ray flower; note the monostigmatic style (st) and the extremely small anther (a). E: "Disk flower" transformed in zygomorphic ray-like flower; note the small anther (a) without pollen grains. f: filament. Scale bars: 20 cm in A; 12 mm in B; 12 cm in C; 5 mm in D and E.

Fig. 2. Schematic representation of the two alleles of the *HxmCYC2c* gene in *Helianthus × multiflorus*. The two alleles *HxmCYC2cA* (A) and *HxmCYC2cB* (B) differ for the insertion of a CACTA transposable element (TE) in the 5'-UTR region (53 bp upstream of the ATG and two bp downstream of the TSS) of *HxmCYC2cB*. The CACTA-like TE is 862 bp long. The three nucleotide (TTA), in red letters, represent the duplication in the insertion point of the TE. The TCP and R domains are depicted in green and light-blue color, respectively. The green triangle indicates the "start" (ATG) codon, the red pentagon indicates the "stop" (TAG) codon and the yellow triangle the intron. The schematic representation is drawn in scale.

Fig. 3. Dot-plot analysis of the *CACTA Transposable Element of Helianthus × multiflorus1* (*CTEHM1*) achieved in the 5'-UTR region of the *HxmCYC2cB* gene of *H. × multiflorus* "Soleil d'Or". The sequence of *CTEHM1* is graphically compared with itself. The main diagonal line corresponds to the 100% match when the sequence is plotted against itself. Direct repeats are lines parallel to the diagonal line, and inverted repeats are displayed as lines perpendicular to the diagonal line. The sub-TRs, which produce a very specific pattern ("transposon signature"), can be easily recognized and are also reported in the S7.

Fig. 4. Bootstrap consensus tree based on Dayhoff PAM distance matrix model generated using PROTDIST, followed by UPGMA of clustering by NEIGHBOR program. Bootstrap replicates were 100 (values are given at the nodes). Only bootstrap values higher to 50 are given at the nodes. The species abbreviations are: *Antirrhinum majus* (Am), *Gerbera hybrida* (Gh), *Helianthus annuus* (Ha), *Helianthus decapetalus* (Hd), *Helianthus × multiflorus* (H×m), and *Senecio vulgaris* (Sv). The *Oryza sativa* (Os) PCF1 (OsPCF1) amino acid sequence was used as outgroup.

Fig. 5. Expression of *CYC2c* mRNA. Transcription of *CYC2c* mRNA in corollas of ray (RF) and disk flowers (DF) of *Helianthus annuus* (Ha), *H. decapetalus* (Hd) and *H. × multiflorus* (H×m). Relative transcript values were calculated by qRT-PCR using root mRNA as reference sample and normalized to those of *Ha-18S* ribosomal gene. Details are provided in Materials and Methods. The graph shows the mean (± SD) of three biological replicates (n = 3). Same letters above the bars indicate no significant differences from each other (ANOVA $P \leq 0.05$) according to Tukey's test.

CAPTIONS OF SUPPORTING INFORMATION

(S1) Table S1. List of primers used for amplification and transcription analyses of *CYCLOIDEA* (*CYC*) genes of *Helianthus annuus*, *H. decapetalus* (*Hd*) and *H. × multiflorus* (*H×m*).

(S2) Fig. S2. List of amino acid sequences of CYCLOIDEA (CYC) transcription factors (TFs) used for phylogenetic analysis. The species abbreviations are: *Antirrhinum majus* (Am), *Gerbera hybrida* (Gh), *Helianthus annuus* (Ha), *Helianthus decapetalus* (Hd), *Helianthus × multiflorus* (H×m) and *Senecio vulgaris* (Sv). The GenBank accession numbers are in brackets. The amino acid sequences of CYC TFs were aligned by ClustalW multiple sequence alignment program (Thompson *et al.* 1994) developed by the Kyoto University Bioinformatic Center (http://www.genome.jp/tools/clustalw/) with slow pairwise alignment and subjected to phylogenetic analysis. Phylogenetic analysis was performed using programs from the PHYLIP package, PHYLogeny Inference Package, Version 3.67 (Felsenstein 1985). The Proml program was used that implements the maximum likelihood method for amino acid sequences. The Dayhoff probability model of change between amino acids was chosen. The *Oryza sativa* PCF1 (OsPCF1) amino acid sequence was used as outgroup. As support for the tree, a bootstrap analysis with 100 replicates, was performed by Seqboot program and a consensus tree was obtained by the Consense program using the Majority rule option.

(S3) Fig. S3. Comparison of the *CYC2c* sequences of *Helianthus annuus* (*HaCYC2c*) and *H. decapetalus* (*HdCYC2c*). A: Alignment achieved by ClustalW2 (MUSCLE 3.8) of the nucleotide sequences of *HaCYC2c* and *HdCYC2c*. Highlighted

in green and magenta are the start and stop codons of the coding region, respectively. In yellow are highlighted the introns placed in the 3'-UTR region. Asterisks indicate identical nucleotides. B: Alignment achieved by Clustal Omega of amino acidic sequences HaCYC2c and HdCYC2c. The conserved domains of the CYC2 clade, TCP and R, are highlighted in green and light-blue, respectively. Asterisks indicate identical amino acid residues.

(S4) Fig. S4. Sequence analysis of the *HxmCYC2A* gene of *Helianthus × multiflorus*. A: Nucleotide sequences of *HxmCYC2A*. Highlighted in green and magenta are the start and stop codons of the coding region, respectively. The Transcription Start Site (TSS) is underlined. In yellow is highlighted the intron placed in the 3'-UTR region. B: Amino acidic sequences of HxmCYC2A. The conserved domains of the clade CYC2, TCP and R, are highlighted in green and light-blue, respectively.

(S5) Fig. S5. RT-PCR amplification of two alleles (2,719 and 1,855 bp, respectively) of *Helianthus × multiflorus*. M: marker; 1, 2 and 3: DNA samples.

(S6) Fig. S6. Sequence analysis of two *HxmCYC2c* alleles of *Helianthus × multiflorus*. A: Alignment achieved by ClustalW2 (MUSCLE 3.8) of the nucleotide sequences *HxmCYC2cA* and *HxmCYC2cB*. In *HxmCYC2cB* is inserted the CACTA transposable element (TE). Highlight in green and magenta are the start and stop codons of the coding region, respectively. The Transcription Start Sites (TSSs) are underlined in black for the allele *HxmCYC2cA* and in white for the allele *HxmCYC2cB*. Highlighted in yellow is the intron placed in the 3'-UTR region. In the

5'-UTR, the TE belonging to the CACTA superfamily is highlighted in black with white letters. The TE is flanked by a CACTA sequence at 5'-end and the reverted complementary to the 3'-end (highlighted in blue with white letters). In red, are indicated the three bases duplicated by the TE insertion. Asterisks indicate identical nucleotides. Finally, two single nucleotide polymorphisms (SNPs) between the allele *HxmCYC2cA* and *HxmCYC2cB* are highlighted in gray. B: Alignment achieved by Clustal Omega of amino acidic sequences HxmCYC2cA and HxmCYC2cB. The conserved domains of the clade CYC2, TCP and R, are highlighted in green and light-blue, respectively. Asterisks indicate identical amino acid residues.

(S7) Fig. S7. Some characteristics of the small non autonomous *CACTA Transposable Element of Helianthus × multiflorus1* (*CTEHM1*). A: *CTEHM1* element (862 bp) is inserted 53 bp upstream of the start codon of the *CYCLOIDEA*-like gene *HxmCYC2cB* from *H. × multiflorus* "Soleil d'Or". The short (13 bp) terminal inverted repeats (TIRs), which terminate in an intact 5'-CACTA-3' motif are underlined and in bold characters. Most of sub-terminal repeats (sub-TRs) identified by the DOTTER analysis (Sonnhammer & Durbin 1995) are highlighted with different colors and/or style. A tandem repeat of 22 bp is underlined and highlighted in brown. B: Three tandem sub-TRs (15 bp) in direct orientation next to the 3'-TIR are highlighted in green and underlined. In addition, two tandem sub-TRs (22 bp) in direct orientation are highlighted in grey and white letters and italic style. In both, A and B, further tandem repeats (3-6 bp) are not reported.

(S8) Fig. S8. Sequence analysis, by ClustalW2 (MUSCLE 3.8), of the CACTA of *Helianthus × multiflorus* (*CTEHM1*) and the CACTAs of both the *double flowered* (*dbl*) and *Chrysanthemoides* (*Chry2*) mutants of sunflower (*Helianthus annuus*). A: Alignment of the

nucleotide sequences *dbl* and *Chry2* TEs. B: Alignment of the nucleotide sequences of *dbl*, *Chry2* and *CTEHM1* TEs. The TEs are flanked by a CACTA sequence at 5'-end and a reverted complementary sequence to the 3'-end (highlighted in blue with white letters). Asterisks indicate identical nucleotides.

(S9) Fig. S9. Comparison of the 5'-region of the *CYC2c* sequences of *Helianthus annuus* (*HaCYC2c*) and *H. decapetalus* (*HdCYC2c*). A: The primer F71 (forward) and 73R (reverse) are highlighted in light blue and brown, respectively. The start codon is highlighted in green. B: Alignment achieved by ClustalW2 (MUSCLE 3.8) of the nucleotide sequences of *HaCYC2c* and *HdCYC2c*. Unambiguous are the insertion of 54 and 30 bp in the *HdCYC2c* sequence indicated in bold red letter. (**C**) PCR conducted with the primer combination F71/73R. M indicates the PhiX 174 DNA *HaeIII* Digest DNA ladder; *Ha*: DNA from *H. annuus*; *Hd*: DNA from *H. decapetalus*; *H×m*: DNA from *H. × multiflorus*; H$_2$O: negative control.

(S10) Fig. S10. Sequence analysis of PCR- amplified products from *Helianthus × multiflorus H×mCYC2cA*. A: Nucleotide sequence of *H. × multiflorus H×mCYC2cA.* Highlighted in green and magenta are the start and stop codons of the coding region, respectively. In yellow is highlighted the intron placed in the 3'-UTR region. Highlighted in brown is the primer F73, highlighted in grey are the reverse primer CYCREV6 and the forward primer RAG6F, highlighted in light-blue is the primer CYC11. B: Sequence identity and gaps resulting from the alignment by EMBOSS Needle of the sequences obtained with the primer combination F73/CYCREV6 and RAG6F/CYC11 (in brackets the number of analyzed clones), with the matching sequences of *H. decapetalus* and *H. annuus*.
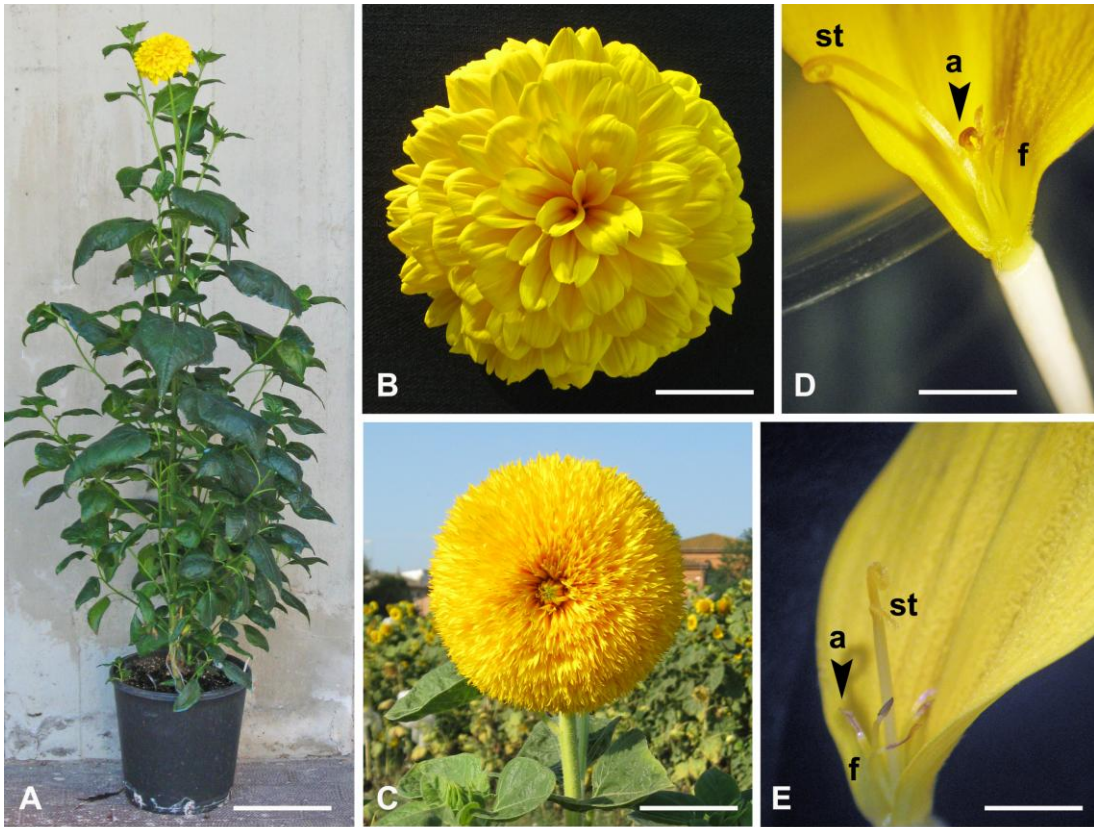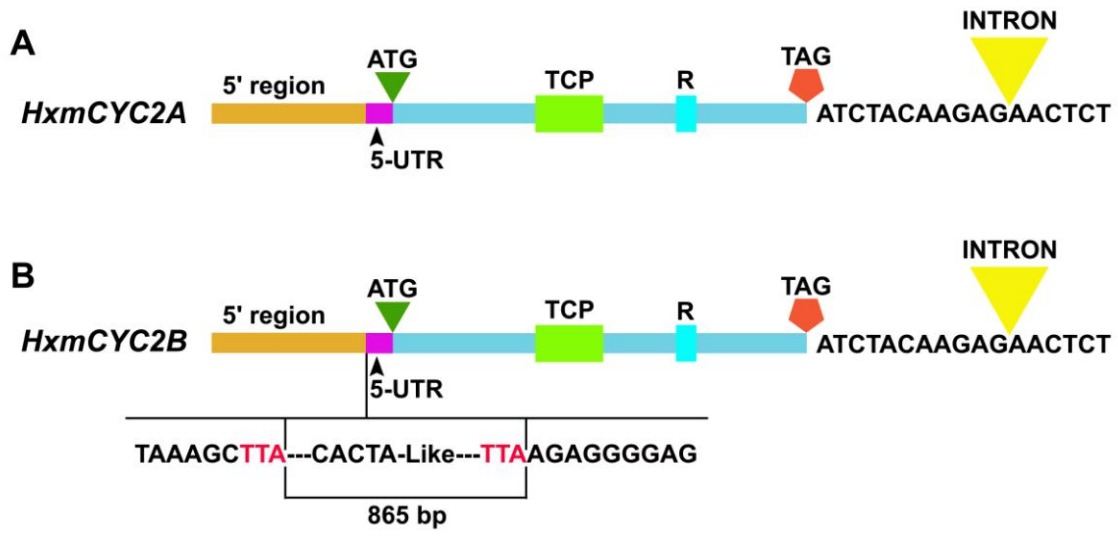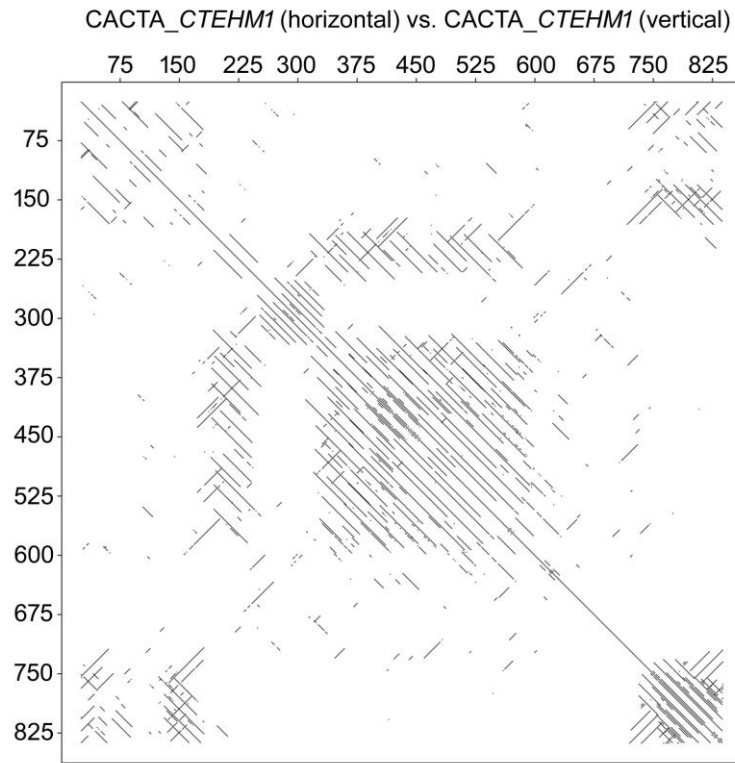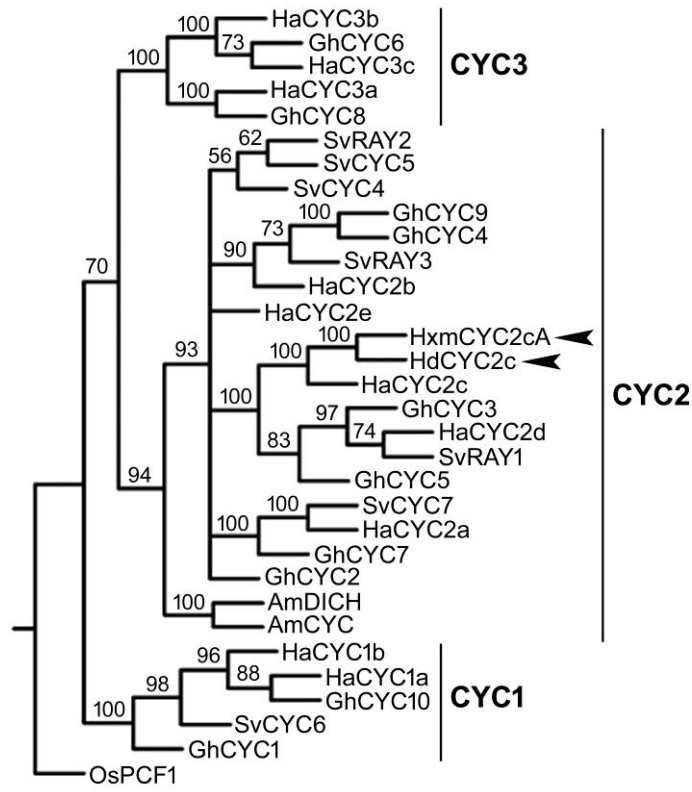
Fig. 1.

Fig. 2.

CACTA_*CTEHM1* (horizontal) vs. CACTA_*CTEHM1* (vertical)

Fig. 3.

Fig. 4.

Fig. 5.