Genetica (2011) 139:1521–1529
DOI 10.1007/s10709-012-9652-y

A transposon-mediate inactivation of a *CYCLOIDEA*-like gene originates polysymmetric and androgynous ray flowers in *Helianthus annuus*

Marco Fambrini · Mariangela Salvini · Claudio Pugliesi

Received: 16 November 2011 / Accepted: 16 April 2012 / Published online: 3 May 2012

Abstract In several eudicots, including members of the Asteraceae family, the *CYCLOIDEA* (*CYC*) genes, which belong to the TCP class of transcription factors, are key players for floral symmetry. The sunflower inflorescence is heterogamous (radiate capitulum) with sterile monosymmetric ray flowers located in the outermost whorl of the inflorescence and hermaphrodite polysymmetric disk flowers. In inflorescence of Heliantheae tribe, flower primordia development initiates from the marginal ray flowers while disk flowers develop later in an acropetal fashion in organized parastichies along a number found to be one of Fibonacci patterns. Mutants for inflorescence morphology can provide information on the role of *CYC*-like genes in radiate capitulum evolution. The *tubular ray flower* (*turf*) mutant of sunflower shows hermaphrodite ray flowers with a nearly polysymmetric tubular-like corolla. Here, we demonstrate that this mutation is caused by the insertion in the TCP motif of a sunflower *CYC*-like gene (*HaCYC2c*) of non-autonomous transposable element (TE), belonging to the CACTA superfamily of transposons. We named this element *Transposable element of turf1* (*Tetu1*). The *Tetu1* insertion changes the reading frame of *turf-HaCYC2c* for the encoded protein and leads to a premature stop codon. Although in *Tetu1* a transposase gene is lacking, our results clearly suggest that it is an active TE. The excision of *Tetu1* restores the wild type phenotype or generates stable mutants. Co-segregation and sequence analysis in progenies of F_2 and self-fertilized plants derived from reversion of *turf* to wild type clearly identify *HaCYC2c* as a key regulator of ray flowers symmetry. Also, *HaCYC2c* loss-of-function promotes the developmental switch from sterile to hermaphrodite flowers, revealing a novel and unexpected role for a *CYC*-like gene in the repression of female organs.

Keywords *CYCLOIDEA*-like gene · Floral morphology · Floral symmetry · Mutant · Sunflower · Transposable element

Electronic supplementary material The online version of this article (doi:10.1007/s10709-012-9652-y) contains supplementary material, which is available to authorized users.

M. Fambrini · C. Pugliesi (✉)
Dipartimento di Biologia delle Piante Agrarie, Sezione di Genetica Agraria, Università di Pisa, Via Matteotti 1B, 56124 Pisa, Italy
e-mail: cpuglies@agr.unipi.it

M. Salvini
Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

Introduction

The evolution in Asteraceae of false but functional flowers (pseudanthium), characterized by combination of different shaped ray and disk florets, in the same inflorescence (radiate capitulum), has a pivotal ecological role (Knowles 1978; Harris 1995). The showy appearance of the sunflower (*Helianthus annuus* L.) inflorescence depends on the outermost whorl of large ray flowers with monosymmetric (zygomorphic) corolla because hermaphrodite disk flowers are smaller and polysymmetric (actinomorphic) (Fig. 1a). Monosymmetric ray flowers have evolved an attraction role as they are sterile because they retain only filamentous remnants of aborted stamens and empty ovaries (Knowles 1978). A variation of ray flower morphology has been described in the *tubular ray flower* (*turf*) mutant where a single nuclear recessive mutation causes the development of atypical capitula characterized by ray flowers with a nearly polysymmetric tubular-like corolla (Fig. 1b, c;

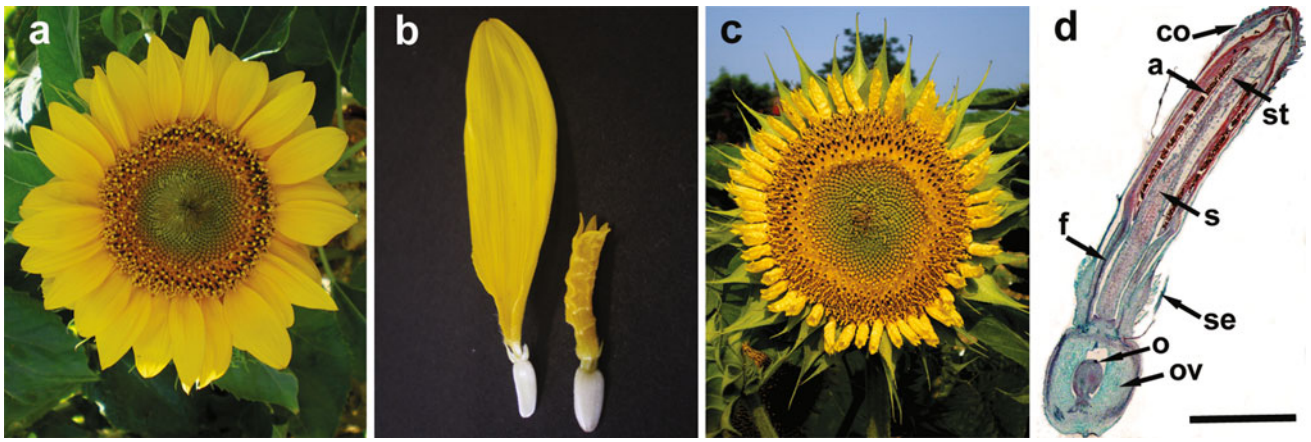

Fig. 1 Morphological characteristics of the *tubular ray flower* (*turf*) mutant of sunflower (*H. annuus* L.). **a** Heterogamous inflorescence of normal (*TURF*) sunflower. **b** A normal ray flower (*left*) and *turf* tubular-like ray flower (*right*). **c** Inflorescence of the *turf* mutant. **d** Longitudinal section of the *turf* tubular-like ray flower 1 week before the anthesis; *a* anther, *co* corolla, *f* filament, *o* ovule, *ov* ovary, *s* style, *st* stigma. *Scale bar* = 1.7 mm

Berti et al. 2005; Fick 1976). These flowers, arranged in the outmost whorl, maintain their positional identity because they are bigger than the polysymmetric disk florets, but also achieve the ability to differentiate both stamens and pistils (Fig. 1d; Berti et al. 2005). The *turf* plants produce tubular-like ray flowers in which stamen filaments, anthers and styles display some developmental defects (Berti et al. 2005), but morphological and genetic analysis demonstrate that both male and female organs are functional (Berti et al. 2005). The *turf* mutation is unstable, and reversions to heads with monosymmetric ray flowers (with about 1 % frequency) have been described (Fambrini et al. 2007).

In several eudicots, the *CYCLOIDEA*-(*CYC*)-like genes, which belong to the TCP class of transcription factors, are key players for floral symmetry (Costa et al. 2005; Donoghue et al. 1998; Martín-Trillo and Cubas 2010; Preston and Hileman 2009). The acronym TCP derives from the first four genes isolated, *TEOSINTE-BRANCHED1* (*Zea mays*), *CYC* (*Antirrhinum majus*) and *PROLIFERATING CELL FACTOR 1* and *2* (*Oryza sativa*) (reviewed in Martín-Trillo and Cubas 2010; Preston and Hileman 2009). All TCP genes contain a 59-amino acid basic helix-loop-helix (bHLH) motif common to all the members that allows DNA binding and protein–protein interactions (Cubas et al. 1999a). This function is consistent with the role for TCP genes as transcriptional regulators of cell division and growth (Martín-Trillo and Cubas 2010). Members of this family are divided into two classes (I and II) based on their structure and TCP domain. In addition to the bHLH motif, *CYC*-like genes, which belong to the TCP class II, have a specific motif, an 18–20 residue arginine-rich motif (the R domain) and a glutamic acid-cysteine-glutamic acid stretch (also known as an ECE motif). Through phylogenetic and sequence analysis, members of class II can be divided into CYC/TB1-like and CINCINNATA-like clades. The first clade is further divided into CYC1, 2 and 3 subclades which have evolved due to a series of duplication events (Howarth and Donoghue 2006; Martín-Trillo and Cubas 2010; Preston and Hileman 2009). Analysis of different angiosperm species shows that most genes of the CYC2 clade were recruited to function in floral symmetry control (Citerne et al. 2000, 2003; Cubas et al. 1999b; Damerval et al. 2007; Luo et al. 1996, 1999). More complicated cases of CYC2 recruitment appear to have occurred in members of the Asteraceae and Dipsacaceae families characterized by radiate inflorescences (Abbott et al. 2003; Broholm et al. 2008; Busch and Zachgo 2009; Carlson et al. 2011; Kim et al. 2008). In particular, the radiate species of Dipsacaceae have undergone more extensive diversification of *CYC*-like genes than discoid species suggesting the hypothesis that these genes could be implicated in the evolution of radiate heads (Carlson et al. 2011). To date, in Asteraceae, the molecular dissection of radiate inflorescences in *Senecio* and *Gerbera* has demonstrated that flower identity within the inflorescence depends on the expression of *CYC*-like genes (Broholm et al. 2008; Kim et al. 2008). In sunflower, ten *CYC*-like genes have been isolated and the diversification of this ample *CYC* gene family depends on duplication events probably followed by functional divergence (Chapman et al. 2008). Recently, a comparative study has been conducted in gerbera and sunflower to explore *CYC/TB1* gene family evolution and diversification in Asteraceae (Tähtiharju et al. 2012). In these two species, marginal florets (ray and trans flowers) showed a strong expression level of the six members of *CYC2* clade in gerbera and the five genes in sunflower. Tähtiharju et al. (2012) hypothesized for these genes early functions in flower type differentiation and late functions in reproductive organ development. Moreover, the null or very low level of expression in disk florets of one gene in gerbera (*GhCYC3*) and two in sunflower (*HaCYC2d* and *HaCYC2c*) could be indicative to their role in controlling ray

flower identity (Chapman et al. 2008; Tähtiharju et al. 2012). Functional specificity for *CYC2* clade members in sunflower has been supposed by flower-type-specific expression and protein–protein interactions (Chapman et al. 2008; Tähtiharju et al. 2012) but the control of ray flower identity is unclear.

Mutants with altered inflorescence morphology can provide information on the roles of *CYC*-like genes in radiate inflorescence evolution (Coen et al. 1995). In this study, we report the molecular characterization of *turf*. Our results identify *HaCYC2c* as a key regulator of ray flower monosymmetry. Also, *HaCYC2c* loss-of-function promotes the developmental switch from sterile to hermaphrodite flowers, revealing a novel and unexpected role for a *CYC*-like gene in the repression of female organs.

Materials and methods

Plant material and growth condition

Sunflower seedlings (*H. annuus* L.) from wild type (*TURF*), mutant *tubular ray flowers* (*turf*), F_1 (*turf* × *TURF*) progenies, F_2 progenies, *turf* plants reverted to *TURF* phenotype and progenies derived from self-fertilization of revertants were grown in field conditions (Experimental fields, University of Pisa, Pisa, Italy). Plant growth conditions were as previously described (Fambrini et al. 2007). Briefly, progenies were grown using 50 cm inter-row spacing with 25–30 cm between plants (about 8–9 plants m^{-2}).

Histological analysis

Ray flowers were fixed as previously described (Fambrini et al. 2006). Paraffin-embedded tissues were sectioned using a rotary microtome (Reichert, Vienna, Austria). Longitudinal sections (10 µm thick) were stained with a mixture of three dyes as described by Graham and Trentham (1998) and examined under a light microscope (DMRB Leitz, Wetzlar, Germany).

Isolation of the *HaCYC2c* gene in wild type (*TURF*) and *turf* mutant

Expression analysis previously reported in sunflower (Chapman et al. 2008), suggested that *HaCYC2c*, a *CYC*-like gene, was the probable candidate for the *TURF* locus. The GenomeWalker DNA walking approach was used to extend the sequence to 5′ untranslated region of the *HaCYC2c* gene (Fambrini et al. 2011; Siebert et al. 1995). DNA was extracted from leaves of 30-days-old plants of both *TURF* and mutant plants with the NucleoSpin® Plant II DNA extraction and purification protocol for plant tissue, according to the manufacturer's instructions (Macherey–Nagel GmbH & Co., Düren, Germany). Genomic DNA digested with the *Dra*I restriction enzyme was ligated to the Genome Walker Adaptor according to the manufacturer's instructions (Clontech Laboratories, Inc., Mountain View, CA, USA). The first PCR was performed with the AP1 primer provided in the kit and the *HaCYC2c*-specific primer RAG5R (reverse, 5′-ACCCAGACCCTTCATGATCAAG CAGCTGGT-3′) (see Supplemental Fig. 1 for primer position in both *TURF-HaCYC2c* and *turf-HaCYC2c* sequences). The nested PCR was performed with the nested adaptor primer (AP2) and the nested *HaCYC2c*-specific primer RAG6R (reverse, 5′-TGGAGGTAGGGTTGATGGTTG TGGAGCA-3′). A PCR product of about 1,750 bp was obtained. All PCRs were performed in conditions recommended by manufacturer (Clontech).

Genomic DNA and first strand cDNA from *TURF H. annuus* were used as templates in PCR reactions to identify the introns/exons regions of *HaCYC2c*. Total RNA was extracted from immature ray flowers (3–4 mm long) with the TRIzol® Reagent, according to the manufacturer's instructions (Invitrogen™ life technologies, Carlsbad, CA, USA). Total RNA (4 µg), was used with the Superscript pre-amplification kit (Invitrogen™), to produce, the first strand cDNA in conditions recommended by the manufacturer. Two gene-specific primers were used: CYC10 (forward, 5′-CAGTGTAAAGCTTAAGAGGGGTGG-3′) placed 67 bp before the start codon and CYC11 (reverse, 5′-CCAAT CAACTAGTGCCGAGAAGTGG-3′) placed 179 bp after the stop codon (Supplemental Fig. 1). PCRs were performed in 20 µl of 1× buffer (Promega Italia, Milano, Italy) containing 0.2 mM dNTPs, 0.5 µM of each primer, 2 mM $MgCl_2$, 0.5 U Go*Taq* DNA polymerase (Promega) and 50 ng of genomic DNA or 1 µl of first strand cDNA. From DNA a PCR product of 1,392 bp was obtained (Supplemental Fig. 2). From cDNA a PCR product of 1,310 bp was obtained. The PCR conditions were: 95 °C for 4 min, 35 cycles (30 s at 94 °C, 1 min 20 s at 60 °C, 30 s at 72 °C), with a final extension of 7 min at 72 °C.

When using mutant DNA as a template, no amplified products were obtained with the primer combination CYC10/CYC11. However, several other primer combinations were able to produce PCR products (Supplemental Fig. 1). In the 5′ region of *HaCYC2c*, a product of 404 bp was obtained with the primer combination CYC10/RAG5R while in the 3′ region, a product of 778 was obtained with the primer combination RAG14F (forward, 5′-ATTGGC TCTTTACCAAGTCCAAGAAGGCGA-3′)/CYC11. The sequences were identical to those of the *TURF* gene (Supplemental Fig. 3a, b). These results suggested that, within the 210 bp-long region between the primers RAG5R and RAG14F of the *turf-HaCYC2c* gene, a chromosomal rearrangement occurred.

The GenomeWalker DNA walking approach was used to find genomic DNA sequences between the RAG5R and RAG14F region of the *turf-HaCYC2c* gene (Supplemental Fig. 1). In the 3′ direction genomic DNA digested with the *Dra*I restriction enzyme was ligated to Genome Walker Adaptor according to the manufacturer's instructions (Clontech). The first PCR was performed with the AP1 primer provided in the kit and the *HaCYC2c*-specific primer CYC6F (forward, 5′-TGCTCCACAACCATCAACCC TACCTCCAGT-3′). The nested PCR was performed with the nested adaptor primer (AP2) and the nested *HaCYC2c*-specific primer CYC5F (forward, 5′-ACCAGCTGCTT GATCATGAAGGGTCTGGGTT-3′). A PCR product of about 1,500 bp was obtained. In the 5′ direction genomic DNA digested with the *Eco*RV restriction enzyme was ligated to Genome Walker Adaptor according to the manufacturer's instructions (Clontech). The first PCR was performed with the AP1 primer provided in the kit and the *HaCYC2c*-specific primer RAG15R (reverse, 5′-CCTG ATCCTCCCTTGATGGTCTCTAGGAA-3′). The nested PCR was performed with the nested adaptor primer (AP2) and the nested *HaCYC2c*-specific primer RAG14R (reverse, 5′-TCGCCTTCTTGGACTTGGTAAAGAGCCA AT-3′). A PCR product of about 5,000 bp was obtained. All PCRs were performed in conditions recommended by manufacturer (Clontech).

All PCR products were separated by electrophoresis on a 1.5 % TAE-agarose gel and visualized with ethidium bromide under UV light. All the amplified products were purified using the Wizard® 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 Wizard® *Plus* Minipreps DNA Purification Kit (Promega). Several clones were sequenced on both strands. Sequence data from this article have been deposited in GenBank under the following accession numbers: HE604335 and HE604336.

Database search and sequences analysis

Database searches were carried out using the BLAST program at the National Center for Biotechnology Information (NCBI) (Altschul et al. 1997). PROSITE and PFAM databases were searched to identify conserved domains (Bateman et al. 2002; Falquet et al. 2002). The databases for Triticeae repetitive elements (TREP, http://wheat.pw.usda.gov/ITMI/Repeats) were also screened with the BLASTN and BLASTX algorithms. The chosen sequences were aligned using ClustalW version 1.7 (Thompson et al. 1994), and the EMBOSS pairwais alignment algorithms from EMBL-EBI (European Bioinformatics Institute; http://www.ei.ac.uk/emboss/align/).

Co-segregation analysis

Seedlings from wild type (*TURF*), mutant (*turf*), F_1 (*turf* × *TURF*) progenies, one F_2 progeny, mutant plants reverted to wild type and one progeny derived from self-fertilization of a revertant were grown in field conditions as described above. DNA was extracted from leaves of 20-days-old plants as described above. The primer combination CYC10/CYC11 was used for genotype screening identifying *TURF/TURF* and heterozygous (*TURF/turf*) material. In the *turf* genome (*turf/turf*), the PCR conditions were unable to amplify the ample sequence (7,179 bp) spanning from CYC10 to CYC11 (Supplemental Fig. 1 and Supplemental Fig. 4). The primer combination RAG19 (forward, 5′-GGTTGCATTTGAGTGTAACAAAGG-3′)/ CYC11 amplified a product of 1,259 bp. The sequence of the primer RAG19 is placed in the transposable element and therefore, the PCR product was expected only in homozygous (*turf/turf*), and heterozygous (*TURF/turf*) genotypes (Supplemental Fig. 1 and Supplemental Fig. 4). The PCR conditions for both primer combinations were: 95 °C for 4 min, 30 cycles (30 s at 94 °C, 1 min 20 s at 60 °C, 30 s at 72 °C), with a final extension of 7 min at 72 °C. The PCR products were separated by electrophoresis on a 1.5 % TAE-agarose gel and visualized with ethidium bromide under UV light.

Results and discussion

A transposable element is inserted in the *HaCYC2c* gene of the *turf* mutant

The *HaCYC2c* gene from wild type (*TURF/TURF*) seedlings was isolated using PCR and Genome walking approaches. Sequence analysis showed a 1,146 bp coding sequence (CDS) (Fig. 2a and Supplemental Fig. 1 and Supplemental Fig. 2). The putative *TURF-HaCYC2c* peptide was 381 amino acids long with a calculated molecular mass of 43,570 Da. No nucleotide polymorphisms were detected between the CDS sequences of *TURF-HaCYC2c* and the *HaCYC2c* previously isolated in sunflower (GenBank accession number EU088370; Chapman et al. 2008). However, in the *HaCYC2c* gene isolated from the USDA inbred line Ames 3963 (Chapman et al. 2008), the intron is positioned before the stop codon (see Fig. 2, Chapman et al. 2008). By contrast, the *TURF-HaCYC2c* intron (82 bp) is localized outside the CDS, within the 3′-UTR region, 11 bp after the stop codon (Fig. 2a, Supplemental Fig. 1 and Supplemental Fig. 2). The intron position was deduced from a comparison between *HaCYC2c* cDNA and the genomic DNA sequences. The 3′-UTR position of this intron was confirmed in an assortment of inbred lines of our germplasm

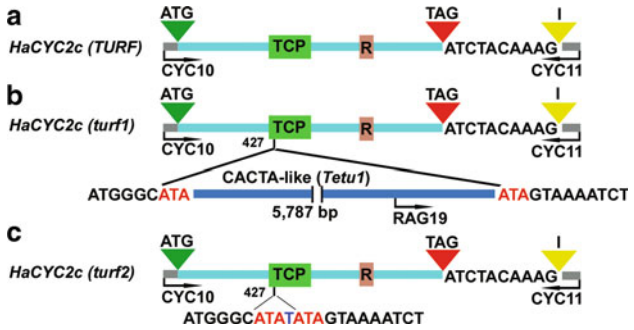

Fig. 2 Schematic representations of *HaCYC2c* gene structure in the wild type (*TURF*, **a**), and mutants (*turf1*, **b** and *turf2*, **c**) of sunflower (*H. annuus* L.). GenBank accession numbers are HE60435 and HE60436. The exons are *boxed* in *light-blue*. The introns are shown as *yellow triangles*. The class II TCP conserved domains, TCP and R, are *boxed* in *green* and *brown*, respectively. The unstable *turf1* allele **b** presents the CACTA-like transposable element, *Tetu1* inserted in the basic motif of the TCP domain (427 bp after the start codon). The stable *turf2* allele **c** shows a footprint (four bp) generated by an imperfect excision of *Tetu1*. The partial 3′- and 5′-UTR are shown in *grey*. The primers used for molecular analysis are indicated. The *arrows* indicate the 5′–3′-primer direction

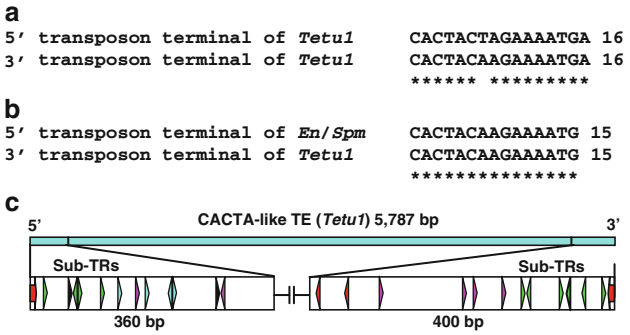

Fig. 3 Characteristics of the CACTA-like transposable element (TE) *Tetu1* inserted in the *HaCYC2c* gene of the *tubular ray flower* (*turf*) mutant of sunflower (*H. annuus* L.). **a** Alignment of terminal inverted repeats (TIRs) of *Tetu1*. **b** Alignment of TIRs comparing *Tetu1* and the CACTA TE *En/Spm*. *Asterisks* indicate identical nucleotides. The 3′ transposon terminal sequence of *Tetu1* is in reverse-complement counterpart. **c** Schematic representation of TIRs and sub-terminal repeat (sub-TRs) regions of *Tetu1*. The *red arrows* indicate the TIRs; the *coloured triangles* indicate the sub-TRs regions in direct and inverted orientations. The TIRs and sub-TRs were deduced by sequence analysis of Supplemental Fig. 4

collection (data not shown). Indeed, intron position is well conserved within the same species; however, we cannot exclude differences related to gene organization between distinct inbred lines. The 404 bp 5′ region as well as the 778 bp 3′ regions amplified from the *turf* genomic DNA were identical with the nucleotide sequence of *TURF-HaCYC2c* (Supplemental Fig. 3a, b). However, PCR experiments suggested a probable chromosomal rearrangement in proximity of the TCP domain of the *turf-HaCYC2c* gene (see also Materials and Methods). Genome walking was used to explore this genomic region. A large insertion (5,787 bp) was found in the basic motif of the TCP domain, 427 nucleotides after the start codon. Sequence analysis performed in public databases and the database for Triticeae repetitive elements (TREP, http://wheat.pw.usda.gov/ITMI/Repeats) suggested that this insertion was a part of a transposable element (TE) belonging to the CACTA superfamily of transposons (*En/Spm*, *Tgm*, and *Tam*) (Fig. 2b, Supplemental Fig. 4 and Supplemental Fig. 5). CACTAs are class II TEs that utilize a cut-and-paste mechanism for transposition, which requires a transposase enzyme (Feschotte and Pritham 2007; Muszynski et al. 1993). The name ‘CACTA’ refers to the flanking terminal inverted repeats (TIRs), which are normally 10–28 bp long and terminate in a conserved 5′-CACTA-3′ motif (Wicker et al. 2003). The TIRs serve as recognition sequences for the transposase enzyme (Lewin 1997). CACTA elements also have characteristic sub-terminal repetitive (sub-TRs) regions of 10–20 bp units that are repeated in direct and inverted orientations (Feschotte and Pritham 2007). In the *HaCYC2c* gene, the insertion of this sequence created a perfect three bp (ATA) target site duplication (TSD), characteristic of the CACTA superfamily (Fig. 2b and Supplemental Fig. 4). Its 5′ and 3′ ends carried imperfect TIRs flanking the conserved CACTA sequence: a 16 bp TIRs contained one mismatch (Fig. 3a), while the outer 15 bp of the 3′ termini of the sunflower CACTA-like element were identical to those of the maize *En/Spm* TE (Fig. 3b). Sub-terminal repeats regions were also identified (Fig. 3c); we named this element *Transposable element of turf1* (*Tetu1*).

The insertion changes the reading frame of *turf-HaCYC2c* for the encoded protein and led to differences in the amino acid sequence as well as to a premature stop codon (Supplemental Fig. 5). Sequence analyses indicated that *Tetu1* does not contain a transposase encoding gene(s). We hypothesized that *Tetu1* is a non-autonomous truncated and/or rearranged version of a CACTA TE, most likely generated by an imperfect excision of a complete element. Alignment of the whole *turf-HaCYC2c* sequence to the *TURF-HaCYC2c* sequence using the EMBOSS pairwise alignment showed that, except for the *Tetu1* insertion, the two allelic sequences were identical (Supplemental Fig. 6).

The sunflower genome is extremely repetitive with large number of retrotransposons (Kane et al. 2011). Some transposon superfamilies have been identified and their activity is an important component of reorganizations in the genome of interspecific hybrids (Anisimova et al. 2009). However, *turf* is the first spontaneous mutant in sunflower characterized by a CACTA element insertion in a *CYC*-like gene with a distinct role in development of radiate

inflorescence. Nevertheless, transposon-mediate mutations of *CYC* genes affecting flower symmetry were described in the model plant *A. majus* (Carpenter and Coen 1990; Luo et al. 1999). Two independent recessive mutants (*cycloidea-608* and *cycloidea-609*) characterized by a more radially symmetrical flower than the wild type were induced by a transposon-mediated strategy of mutagenesis using CACTA elements (Carpenter and Coen 1990). Moreover, in the spontaneous mutant *backpetals* a transposon inserted in the regulatory region of the *CYC* gene caused ectopic expression in the lateral and ventral regions of the flower (Luo et al. 1999). This mutation is semidominant and genetically unstable; in particular, the transposon excision in revertants demonstrated that the insertion was indispensable for the development of dorsalized petals. Furthermore, in *Antirrhinum*, Roccaro et al. (2005) reported that the binding factor *ROSINA* (*RSI*) can bind the promoter of the MADS-box *DEFICIENS* gene involved in petal and stamen identity. The genomic organization of several *RSI* copies proved that this DNA binding factor is a part of a CACTA TE (Roccaro et al. 2007).

The mutation in the *HaCYC2c* gene co-segregates with the *turf* mutant phenotype

Co-segregation analysis of parental, F_1, F_2 and populations of self-fertilized plants derived from reversion of *turf* to wild type phenotype was performed using a PCR approach with the primer combinations CYC10/CYC11 and RAG19/CYC11 (Fig. 2a, b and Supplemental Fig. 1). We observed full correlation with the segregating phenotypes and inheritance of *turf* alleles (Fig. 4a–c and Supplemental Fig. 7a, b). In the wild type (*TURF/TURF*, *TT*), the only primer combination that gave a PCR product was CYC10/CYC11 (Fig. 4a). By contrast, in the *turf* mutant (*tt*) only the RAG19/CYC11-PCR product was obtained. As expected, from heterozygous (*Tt*) F_1 plants both PCR products were amplified (Fig. 4a). Within one F_2 progeny of 48 plants, 38 individuals showed normal monosymmetric ray flowers, 13 of which displayed a single CYC10/CYC11-PCR product (genotype *TT*) and 25 of which displayed both PCR products (genotype *Tt*). From the 10 individuals with polysymmetric tubular-like flowers, only a RAG19/CYC11-PCR product was obtained (genotype *tt*), in agreement with the expected 1:2:1 genotypic segregation ($\chi 2 = 0.458$; $P = 0.70$–0.80; $df = 2$). A previous genetic analysis demonstrated that reversion of *turf* to wild type affected a single allele (Fambrini et al. 2007). Here, the molecular analysis of one progeny (44 plants) derived from a self-fertilized reverted plant (Fig. 4c), unmistakably validated that the *TURF* locus was the gene *HaCYC2c*. Indeed, in homozygous (*TT*) and heterozygous (*Tt*) individuals, the CYC10/CYC11-PCR products showed identical nucleotide sequences to wild type *HaCYC2c-TURF* gene (Fig. 4c and Supplemental Fig. 7a).

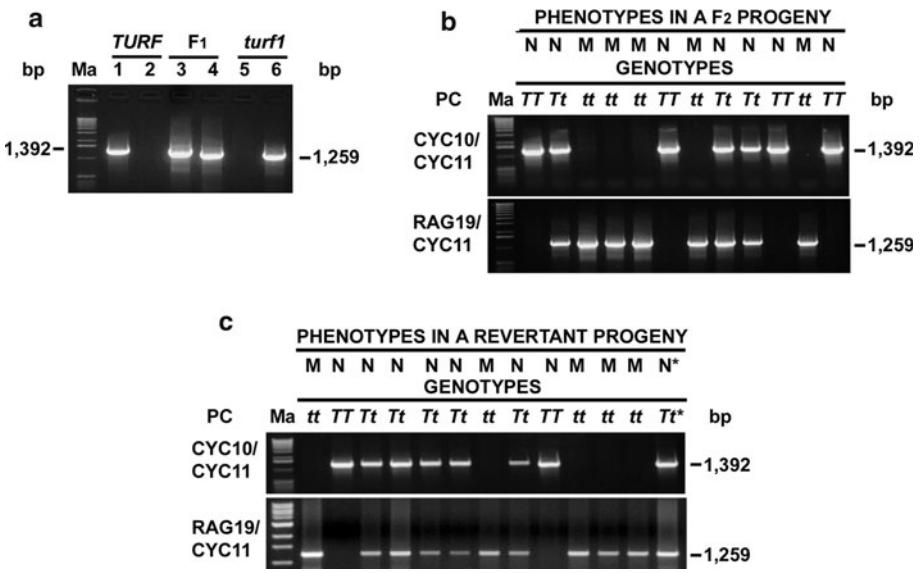

Fig. 4 Co-segregation analysis. **a** PCR screening of genotypes with the primer combination CYC10/CYC11 (*lanes* 1, 3 and 5) and RAG19/CYC11 (*lanes* 2, 4 and 6) in homozygous normal plants (*TURF*), homozygous mutant plants (*turf1*) and F_1 hybrid (*TURF* × *turf1*) (for primer position see also Fig. 2, Supplemental Fig. 1, Supplemental Fig. 2 and Supplemental Fig. 4). **b**, **c** Co-segregation analysis performed as described in **a** in a F_2 progeny (**b**) and in a progeny derived from a self-fertilized reverted plant (**c**). The primer combinations (PC) are indicated on the *left* of gels. M and N indicate the mutant and normal phenotype, respectively. *TT*, *Tt* and *tt* indicate the genotypes for the alleles *TURF* (*T*) and *turf* (*t*). N* and *Tt** indicate the phenotype and the genotype of a reverted plant, respectively. The PCR products were resolved on ethidium bromide-stained 1.5 % (w/v) TAE-agarose gels and visualized under UV light. The size of amplified PCR products is indicated. Ma = 1 kb DNA ladder (BioLabs Inc., New England, Ipswich, MA, USA)

As expected, identical nucleotide alignment was also detected in RAG19/CYC11-PCR products obtained from homozygous (*tt*) and heterozygous (*Tt*) individuals and the RAG19/CYC11-*turf* sequence (Fig. 4c and Supplemental Fig. 7b). Of the 44 individuals, 12 were homozygous (*TT*) and 24 heterozygous (*Tt*) both with a normal inflorescence, and 8 were homozygous (*tt*) with a *turf* phenotype, in agreement with the expected 1:2:1 ratio ($\chi 2 = 1.091$; $P = 0.50$–0.70; $df = 2$). Therefore, the *turf* mutant co-segregated with the mutation in the *HaCYC2c* gene and the *Tetu1* excision was responsible for the reversion of the mutant to wild type phenotype.

The CACTA excision generates stable *turf* mutant a new wild type phenotype

Excision of the *Tetu1* element generated footprints in the *HaCYC2c* gene resulting in stable mutant phenotypes with frame shift mutations (see for example the sequence of *turf2* generated by a four bp footprint and depicted in Fig. 2c). Notably, an homozygous reverted wild type phenotype displayed an excision of the *Tetu1* that generated a deletion of three bp (GTA) in proximity of the TSD and therefore, located in the basic region of the TCP domain (Supplemental Fig. 8a, b). It is likely that the serine in the bHLH motif was not essential for the activity of the *HaCYC2c* gene. In *HaCYC2c* alleles of other three reverted wild type plants, *Tetu1* excision only generated a canonical three bp (ATA) TSD (data not shown). Indeed, the presence, in some reverted wild type plants, of ray flowers with additional petaloid structures attached at the base of the corolla and filaments with small sterile anthers (Fambrini et al. 2007) could be related to footprints generated by *Tetu1* excision that weakly influenced the HaCYC2c amino acid sequence. Also, the absence of chimerism in inflorescences reverted to a wild type phenotype suggested a preponderance and/or exclusive germinal excision of *Tetu1*, in agreement with the exact mendelian (3:1) segregation detected in progenies of self-fertilized reverted plants (Fambrini et al. 2007).

Although in *Tetu1* a transposase gene is lacking, these results clearly suggest that it is still an active CACTA TE. It is possible that the non-autonomous *Tetu1* element could be mobilized in *trans* by related transposases as long as it contains the sequence motifs recognized by those transposases that are essential for mobilization (e.g. TIRs and sub-TRs).

Conclusion

Our data clearly show that *HaCYC2c* is a key inducer of monosymmetry in ray flowers of sunflower. Also, according to Chapman et al. (2008), the *HaCYC2c* transcript accumulation in sunflower inflorescence is restricted to petals and ovaries of ray flowers. When *HaCYC2c* function is lost, the morphology of *H. annuus* inflorescence is redesigned in a similar manner to transgenic *Senecio* plants overexpressing the *CYC*-like gene *RAY2* in the radiate background (Kim et al. 2008). In *Gerbera hybrida* loss and/or gain of function of *GhCYC2* gene had marked effects on petal length and disc flowers morphology without change of corolla symmetry in ray flowers (Broholm et al. 2008). In particular, constitutive expression of the *GhCYC2* gene resulted in shorter zygomorphic petals of ray flowers while disk flower petals showed a distinct ligular structure that resembled the bilaterally symmetric shape of ray and trans flowers. Moreover in one of the overexpression lines, all petals were fused to give peculiar tubular disk flowers (Broholm et al. 2008). In transgenic lines with reduced *GhCYC2* expression, the length of trans flower petals was shorter than in wild type and occasional splitting of the trans flower ligules was observed (Broholm et al. 2008). Taken together these results suggest that in Asteraceae, species-specific mechanisms are involved in the molecular dissection of inflorescence development. Loss-of-function of *HaCYC2c* affects symmetry and fertility of ray flowers but the floret identity is maintained in *turf* mutant confirming a probable involvement of other genes, e.g. *HaCYC2d* of the CYC2 clade (Tähtiharju et al. 2012). In reproductive organs of mature disk flowers, significant levels of transcription for all genes of the CYC2 clade has recently been reported by Tähtiharju et al. (2012) suggesting late functions; however, the specific role of the *HaCYC2c* gene remains doubtful because morphology and fertility of disk flowers of *turf* are unaffected. Unexpectedly, the transposon-inactivation of the *HaCYC2c* gene was sufficient to promote stamen and pistil differentiation and to recover selfed seeds from outermost whorl of flowers (Fambrini et al. 2007). The relationship between the expression of *CYC*-like genes and stamen abortion has been documented in several species (Cubas et al. 1999a; Luo et al. 1996; reviewed in Preston and Hileman 2009) but their involvement in pistil development is a novel role for a gene of the *CYC2* sub-clade.

This work is the first molecular characterization of a sunflower mutant characterized by an altered floral symmetry. Also, these results open a question about a probable interaction between *TCP* and *MADS* box genes in flower development of sunflower. The regulation of stamen and carpel growth by *HaCYC2c* activity should be mediated by *MADS* box gene with B, C and/or SEPALLATA function (Krizek and Fletcher 2005; Pelaz et al. 2000). On the basis of these data, we retain that further characterization of the *turf* mutant could be very useful to shed light on new aspects about the functional divergence of *HaCYC2c* in *H. annuus*.

References

Abbott RJ, James JK, Milne RI, Gillies ACM (2003) Plant introduction, hybridization and gene flow. Philos Trans R Soc B Biol Sci 358:1123–1132

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Anisimova IN, Tumanova LG, Gavrilova VA, Dyagileva AV, Pasha LI, Mitin VA, Timofeyeva GI (2009) Genomic instability in sunflower interspecific hybrids. Russ J Genet 45:934–943

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer ELL (2002) The pfam protein families database. Nucleic Acids Res 30:276–280

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

Broholm SK, Tähtiharju S, Laitinen RAE, 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 USA 105:9117–9122

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

Carlson SE, Howard DG, Donoghue MJ (2011) Diversification of *CYCLOIDEA*-like genes in Dipsacaceae (Dipsacales): implications for the evolution of capitulum inflorescences. BMC Evol Biol 11:325. doi:10.1186/1471-2148-11-325

Carpenter R, Coen ES (1990) Floral Homeotic mutations produced by transposon-mutagenesis in *Antirrhinum majus*. Genes Dev 4:1483–1493

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

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

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

Coen ES, Nugent JM, Luo D, Bradley D, Cubas P, Chadwick M, Copsey L, Carpenter R (1995) Evolution of floral symmetry. Philos Trans R Soc B Biol Sci 350:35–38

Costa MMR, Fox S, Hana AI, Baxter C, Coen E (2005) Evolution of regulatory interactions controlling floral asymmetry. Development 132:5093–5101

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

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

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 Physiol 143:759–772

Donoghue MJ, Ree RH, Baum DA (1998) Phylogeny and the evolution of flower symmetry in the Asteridae. Trends Plant Sci 3:311–317

Falquet L, Pagni M, Bucher P, Hulo N, Sigrist CJ, Hofmann K, Bairoch A (2002) The PROSITE database, its status in 2002. Nucleic Acids Res 30:235–238

Fambrini M, Bonsignori E, Rapparini F, Cionini G, Michelotti V, Bertini D, Baraldi R, Pugliesi C (2006) *stem fasciated*, a recessive mutation in sunflower (*Helianthus annuus*), alters plant morphology and auxin level. Ann Bot 98:715–730

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

Fambrini M, Mariotti L, Parlanti S, Picciarelli P, Salvini M, Ceccarelli N, Pugliesi C (2011) The extreme dwarf phenotype of the GA-sensitive mutant of sunflower, *dwarf2*, is generated by a deletion in the *ent-kaurenoic acid oxidase1* (*HaKAO1*) gene sequence. Plant Mol Biol 75:431–450

Feschotte C, Pritham EJ (2007) DNA-transposons and the evolution of eukaryotic genomes. Annu Rev Genet 41:331–368

Fick GN (1976) Genetics of floral color and morphology in sunflowers. J Hered 67:227–230

Graham ET, Trentham R (1998) Staining paraffin extracted, alcohol rinsed and air dried plant tissue with aqueous mixture of three dyes. Biotech Histochem 73:178–185

Harris EM (1995) Inflorescence and floral ontogeny in Asteraceae: a synthesis of historical and current concepts. Bot Rev 61:93–278

Howarth DG, Donoghue MJ (2006) Phylogenetic analisys of the "ECE" (CYC/TB1) clade reveals duplications predating the core eudicots. Proc Natl Acad Sci USA 103:9101–9106

Kane NC, Gill N, King MG, Bowers JE, Berges H, Gouzy J, Bachlava E, Langlade NB, Lai Z, Stewart M, Burke JM, Vincourt P, Knapp SJ, Rieseberg LH (2011) Progress towards a reference genome for sunflower. Botany 89:429–437

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

Knowles PF (1978) Morphology and anatomy. In: Carter JF (ed) Sunflower science and technology. ASA, CSSA, SSSA, Inc. Publishers, Madison, pp 55–87

Krizek EM, Fletcher JC (2005) Molecular mechanisms of flower development: an armchair guide. Nat Rev Genet 6:688–698

Lewin B (1997) Transposons. In: Lewin B (ed) Genes VI. Oxford University Press, New York, pp 563–595

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 ES (1999) Control of organ asymmetry in flowers of *Antirrhinum*. Cell 99:367–376

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

Muszynski MG, Gierl A, Peterson PA (1993) Genetic and molecular analysis of a three-component transposable-element system in maize. Mol Gen Genet 237:105–112

Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF (2000) B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature 405:200–203

Preston JC, Hileman LC (2009) Developmental genetics of floral symmetry evolution. Trends Plant Sci 14:147–154

Roccaro M, Li Y, Masiero S, Saedler H, Sommer H (2005) ROSINA (RSI), a novel protein with DNA-binding capacity, acts during floral organ development in *Antirrhinum majus*. Plant J 43:238–250

Roccaro M, Li Y, Sommer H (2007) ROSINA (RSI) is a part of a CACTA transposable element, *TamRSI*, and links flower development to transposon activity. Mol Genet Genomics 278:243–254

Siebert PD, Chenchik A, Kellog DE, Lukyanov KA, Lukyanov SA (1995) An improved method for walking in uncloned genomic DNA. Nucleic Acids Res 23:1087–1088

Tähtiharju S, Rijpkema AS, Vetterli A, Albert VA, Teeri TH, Elomaa P (2012) Evolution and diversification of the CYC/TB1 gene family in Asteraceae—a comparative study in gerbera (Mutisieae) and sunflower (Heliantheae). Mol Biol Evol 29:1155–1166

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

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