Dev Genes Evol (2008) 218:341–351
DOI 10.1007/s00427-008-0227-y

ORIGINAL ARTICLE

Expression differentiation of *CYC*-like floral symmetry genes correlated with their protein sequence divergence in *Chirita heterotricha* (Gesneriaceae)

Qiu Gao · Ju-Hong Tao · Dan Yan · Yin-Zheng Wang · Zhen-Yu Li

Received: 5 March 2007 / Accepted: 21 May 2008 / Published online: 1 July 2008

Abstract *CYCLOIDIEA* (*CYC*) and its homologues have been studied intensively in the model organism *Antirrhinum majus* and related species regarding their function in controlling floral dorsoventral (adaxial–abaxial) asymmetry, including aborting the adaxial and lateral stamens. This raises the question whether the same mechanism underlies the great morphological diversity of zygomorphy in angiosperms, especially in Lamiales sensu lato, a major clade predominantly with zygomorphic flowers. To address this, we selected a representative in Gesneriaceae, the sister to the remainder of Lamiales s.l., to isolate *CYC* homologues and further investigate their expression patterns using locus-specific semiquantitative reverse transcriptase polymerase chain reaction. Our results showed that four *CYC* homologues in *Chirita heterotricha* differentiated spatially and temporally in expression, in which *ChCYC1D* was only expressed in the adaxial regions, and transcripts of *ChCYC1C* were distributed in both the adaxial and lateral regions, while *ChCYC2A* and *ChCYC2B* transcripts were only detected in the young inflorescences. *ChCYC1C* expression in the lateral regions correlated with abortion of the lateral stamens in *C. heterotricha* hinted at its gain of function, i.e., expanding from the adaxial to the lateral regions in expression. Correlatively, the protein sequences of *ChCYC* genes exhibited remarkable divergences, in which some lineage-specific amino acids between *GCYC1* and *GCYC2* in conserved functional domains and two sublineage-specific motifs between *GCYC1C* and *GCYC1D* in *GCYC1* genes had further been identified. Our results indicated that *ChCYC* genes had probably undergone an expressional differentiation and specialization in establishing the floral dorsoventral asymmetry in *C. heterotricha* responding to different selective pressure after gene duplication.

Keywords Dorsoventral asymmetry · *CYC*-like gene · RT-PCR expression differentiation · Gesneriaceae · *Chirita heterotricha*

Communicated by K. Schneitz

Q. Gao · J.-H. Tao · D. Yan · Y.-Z. Wang (✉) · Z.-Y. Li (✉)
State Key Laboratory of Systematic and Evolutionary Botany,
Institute of Botany, Chinese Academy of Sciences,
Xiangshan,
Beijing 100093, China
email: wangyz@ibcas.ac.cn
email: LiZY@ibcas.ac.cn

Introduction

During the evolution of angiosperms, flower bilateral symmetry, i.e., zygomorphy, is a key innovation associated with important adaptive radiations. Among major clades of angiosperms, zygomorphy is predominant and exhibits great morphological diversity in Asteridae, especially in Lamiales sensu lato that includes a major model organism in developmental biology, i.e., *Antirrhinum majus* (Endress 1998, 1999).

In *A. majus*, *CYCLOIDEA* (*CYC*) and *DICHOTOMA* (*DICH*) are essential for the development of full bilateral symmetry in flowers, in which *CYC* plays a key role in the process (Luo et al. 1996, 1999). *CYC* expression promotes the growth of adaxial petals, while it arrests the growth of the adaxial stamen causing it to become a staminode. Meanwhile, *DICH* activity affects the internal asymmetry of the adaxial petals (Luo et al. 1996, 1999). *CYC* and *DICH* are closely related members of the TCP gene family with the TCP domain related to cell proliferation (Cubas et al. 1999; Kosugi and Ohashi 1997, 2002; Doebley et al.

1995, 1997). A recent study in legumes, distantly related to *A. majus*, shows that a *CYC* homologue *LjCYC*2 also establishes adaxial identity in the legume flowers (Feng et al. 2006). In the basal eudicot family Papaveraceae sensu lato, the duplication and diversification of *CYC*-like TCP genes are accompanied by their divergence in expression patterns, in which one type of them may play a role in flower symmetry (Kölsch and Gleissberg 2006; Damerval et al. 2007). These facts indicate that the basic function of *CYC* and its homologues seems to be conserved in eudicots, i.e., controlling the development of floral asymmetry. Even though the genetic control for the floral dorsoventral asymmetry has been intensively studied in model systems, the evolution of floral symmetry, especially the formation of the great morphological diversity of zygomorphy, is still a vast unexplored field at molecular developmental level.

In *Mohavea confertiflora*, a close relative of *A. majus*, two *CYC* and two *DICH* homologues, i.e., *McCYC1*, *McCYC2*, *McDICH1*, and *McDICH2*, exhibit an expanded expression pattern from the adaxial to the lateral stamens (Hileman et al. 2003). The changes in the expression pattern of *McCYC* and *McDICH* have contributed to the derived flower morphology of *Mohavea*, i.e., the abortion of two lateral stamens besides the adaxial staminode. Even though the expression pattern of *McCYC* and *McDICH* are greatly changed compared to that of *CYC* and *DICH*, there is no noticeable differentiation between copies of *McCYC* and *McDICH* both in expression and sequences of functional domains, respectively. Therefore, Hileman et al. (2003) suggest that changes in the expression domain of an upstream regulator in *CYC/DICH* pathway may be responsible for the expression expansion of *McCYC* and *McDICH* genes from the adaxial to the lateral stamens. However, this suggestion raises the question whether the same mechanism as observed between *Antirrhinum* and *Mohavea* underlies the great morphological diversity of zygomorphy in Lamiales s.l..

The family Gesneriaceae, as the sister group to the remainder of the Lamiales s.l., is characteristic of weak zygomorphy with diverse floral morphologies. In this family, some groups have four fertile stamens with abortion of the adaxial stamen, such as *Oreocharis*, and some have only two abaxial fertile stamens with the adaxial and the laterals aborted, such as *Chirita*. In addition, two lateral fertile stamens plus three staminodes at the adaxial and abaxial positions are characteristic of the genus *Opithandra* (Li and Wang 2004). Given that Gesneriaceae is sister to the remainder of Lamiales s.l. and there are diverse forms of zygomorphy in this family, we considered Gesneriaceae to be the ideal candidate for exploring some possible basic developmental mechanisms in Lamiales s.l. *Chirita heterotricha* is characterized by both adaxial and lateral stamen abortion as in *M. confertiflora* (Fig. 1). Four *CYC* homologues have been isolated from *C. heterotricha* and each belongs to a different lineage of *CYC*-like genes in Gesneriaceae (*GCYC*). Our results further show that the four *CYC*-like genes from *C. heterotricha* differ remarkably in both spatial and temporal expression patterns, which are correlated with their protein sequence divergence.

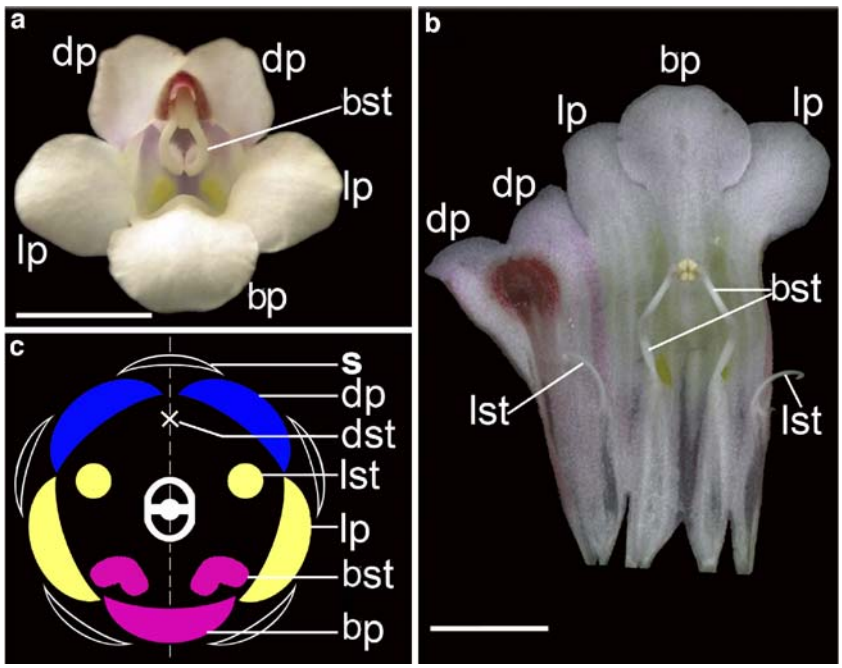

Fig. 1 A flower of *C. heterotricha*. **a** The front view of the flower. **b** The dissected flower. **c** Flower diagram. *s* Sepal, *dp* adaxial petal, *lp* lateral petal, *bp* abaxial petal, *dst* adaxial stamen (undetectable at anthesis), *lst* lateral stamen, *bst* abaxial stamen. *Scale bar*=1 cm

Materials and methods

Plant materials

We used *C. heterotricha* Merr. as the representative of zygomorphic groups in Gesneriaceae for this study. The materials were collected both from plants in field (Baoting county, Hainan province) and from plants that were transplanted from field and cultivated in the greenhouse of the Botanical Garden, Institute of Botany, Chinese Academy of Sciences, Beijing. Flower buds of different stages were collected as follows: Flower buds shorter than 1 cm long were gathered as a whole, while the flower buds that were 1 and 4 cm long, i.e., middle- and late-stage flowers, were dissected into sepals, adaxial/lateral/abaxial corolla plus attached stamens (sepals were removed from the outer whorl and the corolla tube and petals, including the attached adaxial and lateral staminodes and abaxial stamens, were dissected into adaxial, lateral, and abaxial regions). In addition, some late-stage flowers (4 cm long) whose lateral and abaxial stamens could be dissected from the corolla tube were dissected into sepals, adaxial/lateral/abaxial petals, and lateral/abaxial stamens, respectively. Tissues of root, shoot, leaf, and young inflorescence were also collected for reverse transcriptase (RT) polymerase chain reaction (PCR). All materials were frozen in liquid nitrogen immediately after collection for ribonucleic acid (RNA) isolation.

Isolation of *CYC*-like gene in *C. heterotricha*

Deoxyribonucleic acid (DNA) from fresh leaves was extracted following a modified cetyltrimethylammonium bromide procedure of Doyle and Doyle (1987). PCRs were first performed using a forward primer *F1* (5′-ATGTTTG-GAAAGAGCCCATAC-3′) and a reverse primer *GcycR* (Möller et al. 1999) to amplify about 95% of the open reading frame (ORF) of the *CYC*-like genes from *C. heterotricha*.

To get the whole ORF sequence of *ChCYC* genes, we performed rapid amplification of cDNA ends PCR. Total RNA was isolated from young inflorescences and flower buds of *C. heterotricha* using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA), following the manufacturer's instructions. Poly (A) messenger RNA (mRNA) was purified from total RNA using Oligotex mRNA Mini kit (Qiagen, Hilden, Germany). First-strand cDNA was synthesized by SuperScript™ III RT, and amplification was performed with two nested forward primers (5′-TCAACTCTAGCTAAAGAATCAA-3′, 5′-GGCAAGAG CAAGGGCTGGGAA-3′) and polyA primer.

The PCR products were cloned into the pGEM-Teasy vector (Promega, USA) and sequenced. These genes were independently cloned at least twice.

We used contig3.0 (InforMax) to concatenate full-length sequences of *ChCYC* genes.

Phylogenetic analysis of *CYC*-like genes in *C. heterotricha*

We conducted maximum parsimony (MP), maximum likelihood (ML), and neighbor-joining (NJ) analyses with the full-length amino acid sequences of TCP and R domains of four *ChCYC* genes and three TCP class II genes of *Arabidopsis thaliana* (*AtTCP1*, *AtTCP12*, and *AtTCP18*), three TCP genes of *Lotus japonicus* (*LjCYC1*, *LjCYC2*, and *LjCYC3*), *TB1* gene of maize, and *CYC* and *DICH* genes of *A. majus* to identify the position of *ChCYC* genes in the CYC clade of the TCP class II gene family. To clarify the phylogenetic position of *ChCYC* genes in *CYC*-like genes from Old World Gesneriaceae (*GCYC*; subfamily Cyrtandroideae), MP, ML, and NJ analyses were also conducted with nucleotide sequences (MP and NJ) or amino acid sequences (ML) of four *ChCYC* genes and 22 *GCYC* genes from Cyrtandroideae. *CYC* from *A. majus* and a *CYC*-like gene from *Linaria vulgaris* were used as outgroups. The genes used in the phylogenetic analyses are listed in Table 1. Sequences were first aligned using Clustal X (Thompson et al. 1997) and further adjusted manually. PAUP 4.0b10 (Swofford 2002) was used to construct MP and NJ analyses. For MP analysis, the following settings were used: random sequence addition, TBR branching swapping, 1,000 bootstrap replicates. The NJ reconstruction method was carried out on Kimura two-parameter distances, proportion of invariable sites, and parameter for site heterogeneity obtained from MODELTEST (Posada and Crandall 1998) following the Akaike criterion. One thousand bootstrap replicates were done to calculate branch support values. The online version of PHYML (Guindon et al. 2005) was used to construct ML analysis based on amino acid-translated sequences (Guindon and Gascuel 2003) under default settings. The bootstrap values were calculated using 500 replicates. The evolutionary model was a Jones–Taylor–Thornton substitution model, with six substitution rate categories, g-shape parameter, and proportion of invariable sites estimated from the data.

Sequence analysis of *ChCYC* genes

We used DAMBE software (Xia and Xie 2001) to translate the nucleotide sequences of *ChCYC*. The amino acid sequences of *ChCYC* genes were compared with CYC by soft DNAMAN 5.29 (Lynnon Biosoft, USA) and then compared with deduced protein sequences of some relatives, *GCYC1* from *Didymocarpus hancei* (data unpublished), *GCYC1C* and *GCYC1D* from *Loxostigma* sp., *GCYC1C* and *GCYC1D* from *Didymocarpus citrinus*, *GCYC1C* and *GCYC1D* from *Opithandra dinghushanensis*

Table 1 Selected species used in the phylogeny reconstruction in this study, with GenBank numbers for the individual genes used

Family	Taxon	Gene name	Reference	GenBank Accession no.
Gramineae	*Zea mays* subsp. *mays*	*TB1*	Hubbard et al. 2002	AF415152
Brassicaceae	*Arabidopsis thaliana*	*AtTCP1*		NM_001084312
		AtTCP12		NM_105554
		AtTCP18		NM_112741
Leguminosae	*Lotus japonicus*	*LjCYC1*	Feng et al. 2006	DQ202475
		LjCYC2	Feng et al. 2006	DQ202476
		LjCYC3	Feng et al. 2006	DQ202477
Plantaginaceae	*Antirrhinum majus* L.	*cycloidea*	Luo et al. 1996	Y16313
		dichotoma	Luo et al. 1999	AF199465
	Linaria vulgaris L.	Lcyc	Cubas et al. 1999	AF161252
Cyrtandroideae	*Bournea leiophylla*	*GCYC1*	Zhou et al. 2008	EF486283
		GCYC2	Zhou et al. 2008	EF486284
	Chrita heterotricha Merr.	*ChCYC1C*	This study	NA
		ChCYC1D	This study	NA
		ChCYC2A	This study	NA
		ChCYC2B	This study	NA
	Conandron ramondioides Siebole & Zucc.	*Gcyc1*	Citerne et al. 2000	AF208321
		Gcyc2	Citerne et al. 2000	AF208316
	Cyrtandra apiculata C.B.Clark	*Gcyc1*	Wang et al. 2004	AY423160
		Gcyc2	Wang et al. 2004	AY423147
	Didymocarpus citrinus Ridl.	*Gcyc1C*	Wang et al. 2004	AY423158
		Gcyc1D	Wang et al. 2004	AY423159
	Haberlea ferdinandi-coburgii Urum.	*Gcyc1*	Citerne et al. 2000	AF208322
		Gcyc2	Citerne et al. 2000	AF208317
	Jankaea heldeichii Boiss.	*Gcyc1*	Möller et al. 1999	AF208332
	Loxostigma sp.	*Gcyc1C*	Wang et al. 2004	AY423161
		Gcyc1D	Wang et al. 2004	AY423162
	Primulina tabacum Hance	*Gcyc1*	Citerne et al. 2000	AF208320
	Oreocharis benthami C.B.Clarke	*GCYC1*	Du and Wang 2008	NA
		GCYC2	Du and Wang 2008	NA
	Ramonda myconi Rchb.	*Gcyc1*	Möller et al. 1999	AF208331
		Gcyc2	Citerne et al. 2000	AF208318
	Saintpaulia ionantha B.L. Burtt.	*Gcyc1A*	Wang et al. 2004	DQ064642
		Gcyc1B	Wang et al. 2004	DQ064644
	Streptocarpus primulifolius Gand.	*Gcyc1A*	Citerne et al. 2000	AF208340
		Gcyc1B	Möller et al. 1999	AF208336

(data unpublished), *GCYC1* and *GCYC2* from *Bournea leiophylla*, *GCYC1* and *GCYC2* from *Oreocharis benthamii*, *GCYC2* from *Conandron ramondioides*, *GCYC2* from *Cyrtandra apiculata*, *GCYC2* from *Ramonda myconi*, and *GCYC2* from *Haberlea ferdinandi-coburgii* (Table 1); thus, their characteristic amino acids and motifs were identified.

Locus-specific semiquantitative RT-PCR expression of *CYC*-like genes in *C. heterotricha*

We performed RT-PCR with flower tissues as described in "Plant materials." The extraction of total RNAs, purification of poly (A) mRNAs, and synthesis of the first-strand cDNAs were performed according to the methods described above. The template quantity was regulated to be uniform using the *ACTIN* gene (Prasad et al. 2001). PCR was performed by using locus-specific primers: *ChCYC1C* (forward 5′-CGCCGTTTATTGAGACTTCAACC-3′, reverse 5′-CTAGAACTCTTCTTTGTATGAG-3′), *ChCYC1D* (forward 5′-TGTCATTTCTTGAGGTTTCAACA-3′, reverse 5′-CTGGAACTTTTCTTTGTATGAA-3′), *ChCYC2A* (forward 5′-ATCAGCATCACCATGACATTCTT-3′, reverse 5′-TCTGTACCAGGTCCTTAATGGCT-3′), and *ChCYC2B* (forward 5′-ACCAGCACCACCATGACATTCTC-3′, reverse 5′-ACTGCACCAGCTCCTTAATTGCC-3′). To make sure each pair of primers was suitable, we first used them to amplify genomic DNA of *C. heteroticha*. The PCR products were then cloned. At least 20 clones of each PCR product were sequenced, and all the primers we used could amplify the specific copy of *ChCYC* genes. The following thermocycling conditions were employed: initial denaturation at 96°C for 3 min, 30 cycles

of 96°C for 30 s, 55–60°C (depending on the Tm value of primer pairs) for 30 s, and 72°C for 1 min, and a final extension at 72°C for 10 min. The amplified products were separated on a 1.5% agarose gel, and the density of ethidium bromide-stained bands was determined using a Bioimaging System (Gene Tools Program, Syngene, UK). We repeated the RT-PCR experiments five times independently with a new RNA extraction each time. In addition, all RT-PCR products were cloned into pGEM-T Easy vector, and at least 20 clones from each product were sequenced to test the locus specificity of RT-PCR. The *ChCYCn/ACTIN* ratio represented the relative level of *ChCYCn* mRNA expression (*n* denoting *1C*, *1D*, *2A*, and *2B*). Data are presented as the mean±SD of independent RT-PCR experiments, and one-way analysis of variance was used to analyze the expression difference of these four transcripts in various tissues from *C. heterotricha*. A *P* value less than 0.05 was taken to indicate statistical significance.

Results

Isolation and phylogenetic analyses of *ChCYC* genes

We isolated four *CYC*-like genes from genomic DNA of *C. heterotricha*. Since BLAST results showed they had high sequence similarity to *GCYC1C*, *GCYC1D*, and *GCYC2*, respectively, we designated them as *ChCYC1C*, *ChCYC1D*, *ChCYC2A*, and *ChCYC2B*. The full-length ORFs of *ChCYC1C*, *ChCYC1D*, *ChCYC2A*, and *ChCYC2B* were 1,020, 1,044, 1,017, and 1,017 bp, respectively, which were 170 bp longer than *CYC* on average. Maximum likelihood analysis of protein sequences of *ChCYC* genes placed all four *ChCYC* genes into the CYC2 subclade including AtTCP1, LjCYC1, LjCYC2, LjCYC3, CYC, and DICH with high support (79% bootstrap; Fig. 2a). In this subclade, the four *ChCYC* genes were clustered into a group (88% bootstrap), sister to *CYC* and *DICH* genes (70% bootstrap) in a monophyletic lineage. Three *LjCYC* genes formed another lineage of CYC2 (90% bootstrap). MP and NJ analyses were congruent with ML analysis (data not shown). These results showed that *ChCYC* genes belonged to the CYC2 subclade of CYC (ECE) clade in the TCP class II gene family, and they were closely related to *CYC* and *DICH* genes. The phylogenetic analyses of four *ChCYC* genes and other *CYC*-like genes from Old World *GCYC* (subfamily Cyrtandroideae) were conducted with their nucleotide sequences or amino acid sequences using the NJ method as well as MP and ML methods, and they had similar topology. The NJ tree (MP and ML trees were not shown) showed that *GCYC* genes in Cyrtandroideae were divided into two major clades, i.e., the *GCYC1* clade (100% bootstrap) and *GCYC2* clade (100% bootstrap; Fig. 2b). Within the *GCYC1* clade, *GCYC1* from three European genera *Haberlea*, *Jankaea*, and *Ramonda* were gathered together (99% bootstrap) and were sister to the other *GCYC1* genes from the groups mainly distributed in

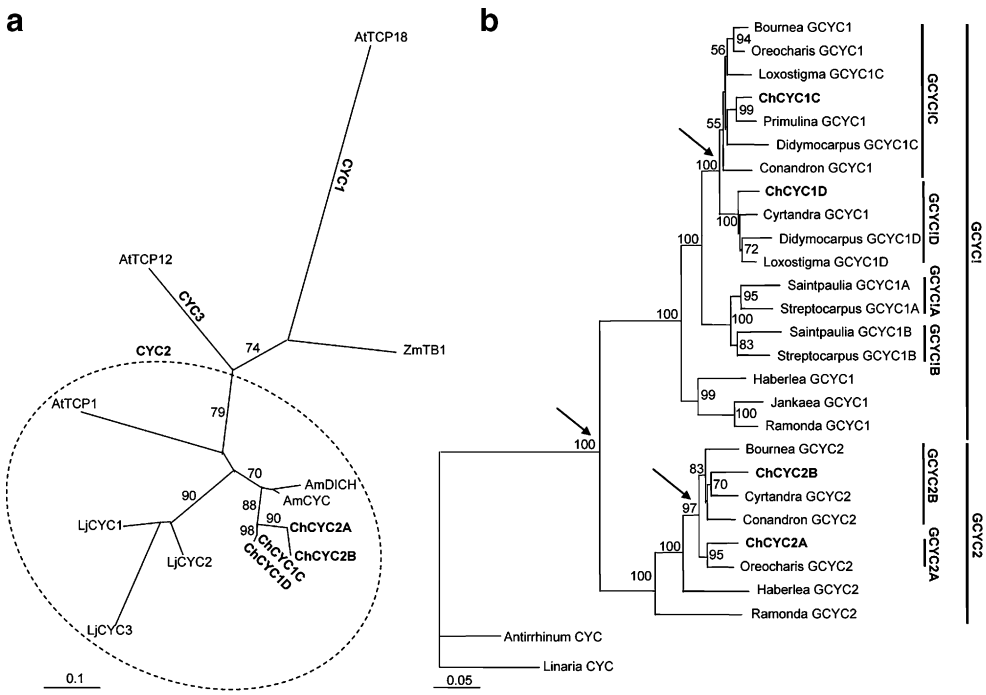

Fig. 2 Phylogenetic reconstructions of *ChCYC* genes with related *CYC*-like genes. **a** Maximum likelihood optimized phylogram of CYC (ECE) clade genes of class II TCP gene family, based on 80 amino acids of TCP and R domains (bootstrap values above 50% are indicated; 500 bootstraps were performed; ChCYC are shown in *bold type*; CYC1/2/3 clades are indicated). **b** Phylogram calculated with the neighbor-joining method based on nucleotide data of *CYC*-like genes from Cyrtandroideae and Plantaginaceae (*ChCYC* genes are shown in *bold type*; numbers indicate bootstrap branch support above 50%; *arrowheads* indicate the duplication nods of *ChCYC*; GCYC1, GCYC2, GCYC1A, GCYC1B, GCYC1C, and GCYC1D clades are indicated)

Africa and Asia, which was further divided into two highly supported subclades, i.e., *GCYC1A*/*GCYC1B* subclade (100% bootstrap) from taxa in Africa and *GCYC1C*/*GCYC1D* subclade (100% bootstrap) from Asiatic groups. These two subclades were both divided into two lineages. The *GCYC1C/GCYC1D* subclade was divided to *GCYC1C* lineage and *GCYC1D* lineage. The single copy of *GCYC1* in *Conandron*, *Bournea*, *Oreocharis*, and *Primulina* were nested within the *GCYC1C* lineage, while *Cytandra GCYC1* was in the *GCYC1D* lineage. All *GCYC2* genes formed a monophyletic clade including the two copies of *GCYC2* from *C. heterotricha*, i.e., *ChCYC2A* and *ChCYC2B*. However, *ChCYC2A* and *ChCYC2B* did not form a clade themselves but were located in different branches of *GCYC2*. *ChCYC2A* was sister to *Oreocharis GCYC2* (95% bootstrap), which was sister to other *GCYC2* from the taxa distributed in Asia (97% bootstrap; Fig. 2b). In these *GCYC2* genes, *ChCYC2B* was sister to *Cyrtandra GCYC2* (70% bootstrap) and together were sister to *Conandron GCYC2*, and inclusively, they were sister to *Bournea GCYC2* (83% bootstrap; Fig. 2b). The *GCYC2* genes in the two European genera were relatively isolated from the two branches of other *GCYC2* genes isolated from Asiatic groups.

RT-PCR expression of *ChCYC* genes

Like *A. majus*, *C. heterotricha* has floral dorsoventral asymmetry in the second and third whorls of floral organs (Fig. 1). To investigate the expression of four *ChCYC* genes, locus-specific semiquantitative RT-PCR was performed on mRNA prepared from dissected sepal and petal plus stamen/staminode tissue of adaxial, lateral, and abaxial organs in middle- (1 cm) and late-stage (4 cm) flowers and dissected sepals, adaxial/lateral/abaxial petals, and lateral/abaxial stamens in late-stage (4 cm) flowers, and root, shoot, leaf, young inflorescence, and flower bud (shorter than 1 cm), respectively. The results showed that the four copies of *CYC*-like genes in *C. heterotricha* were detected only in floral tissues, and they differentiate spatially and temporally in expression in developing flowers. *ChCYC2A* and *ChCYC2B* were detected only in the young inflorescence, while the expression of *ChCYC1C* and *ChCYC1D* was present in the young inflorescence through to the late stage of developing flowers (Fig. 3, 4). In middle stage, transcripts of *ChCYC1C* were detected both in the adaxial and lateral regions of dissected corolla plus staminode tissue, and *ChCYC1D* was mainly expressed in the adaxial petal and staminode. Transcripts of *ChCYC2A* and *ChCYC2B* were undetectable in all floral parts. In the late stage, the mRNA signal of *ChCYC1C* declined earlier in the lateral region than in the adaxial region, and *ChCYC1D* expression was still restricted in the adaxial region but was weak compared with that in middle stage (Fig. 3a). In the lateral floral organs, *ChCYC1C* mainly was expressed in the lateral staminodes in late stage (Fig. 3b). *ChCYC2A* and *ChCYC2B* transcripts were also undetectable in the late stage. The weak mRNA signals in sepals and abaxial petals plus stamens is likely due to the fused parts between them and other floral organs.

Comparative analyses of protein sequences of *ChCYC* genes with *CYC* and other *GCYC* genes

The similarities in amino acid and nucleotide sequences among *ChCYC* and *CYC* genes are shown in Table 2. TCP, R, and ECE domains were recognized from the alignment of their amino acid sequences (Fig. 5a). In the TCP domain, there were three amino acid differences between ChCYC and CYC. The first is in the BASIC region with valine (V) replacing tyrosine (Y). The second was between the HELIX I and LOOP regions with glutamic acid (E) replacing aspartic acid (D), and the third was located in the HELIX II region with valine/alanine (V/A) replacing threonine (T). The amino acid sequences of *ChCYC* and *CYC* genes were highly divergent outside TCP and R domains. ChCYC had a long insertion after the R domain compared to CYC, and this is the main reason why ChCYC was 57 amino acid residues longer than CYC. Furthermore, there were two lineage-specific amino acid substitutions in the TCP domain between *ChCYC1* and *ChCYC2* gene lineages. One, located between BASIC and HELIX I, was either isoleucine (I) or methionine (M), and the other was in HELIX II with either valine (V) or alanine (A). One amino acid substitution between ChCYC1 and ChCYC2 in the R domain consisted of either lysine (K) or arginine (R). Additionally, the 17 amino acids around the ECE motif, recognized recently by Howarth and Donoghue (2005, 2006), were lineage-specific for the *GCYC1* and *GCYC2* gene lineages (Fig. 5a).

ChCYC1C and *ChCYC1D* had 84% identity in amino acid sequences (Table 2). Through their alignment with related *GCYC1* protein sequences, two putative sublineage-specific motifs (PSLMs), which diverged between *GCYC1C* and *GCYC1D* sublineages but were conserved inside each sublineage, were further recognized (Fig. 5a,b). One ranged from 40 to 42 amino acids after the start codon of GCYC1, in which there were amino acid substitutions between QQQ and HHH, and the other was within 242–253 amino acids after R domain in GCYC1D with a 12 amino acid deletion in GCYC1C.

We have also underlined the amino acids unique to *ChCYC1C*, *ChCYC1D*, *ChCYC2A*, and *ChCYC2B* genes compared to other *GCYC* genes within the same lineage or sublineage in their amino acid sequences (Fig. 5a; alignment not shown).

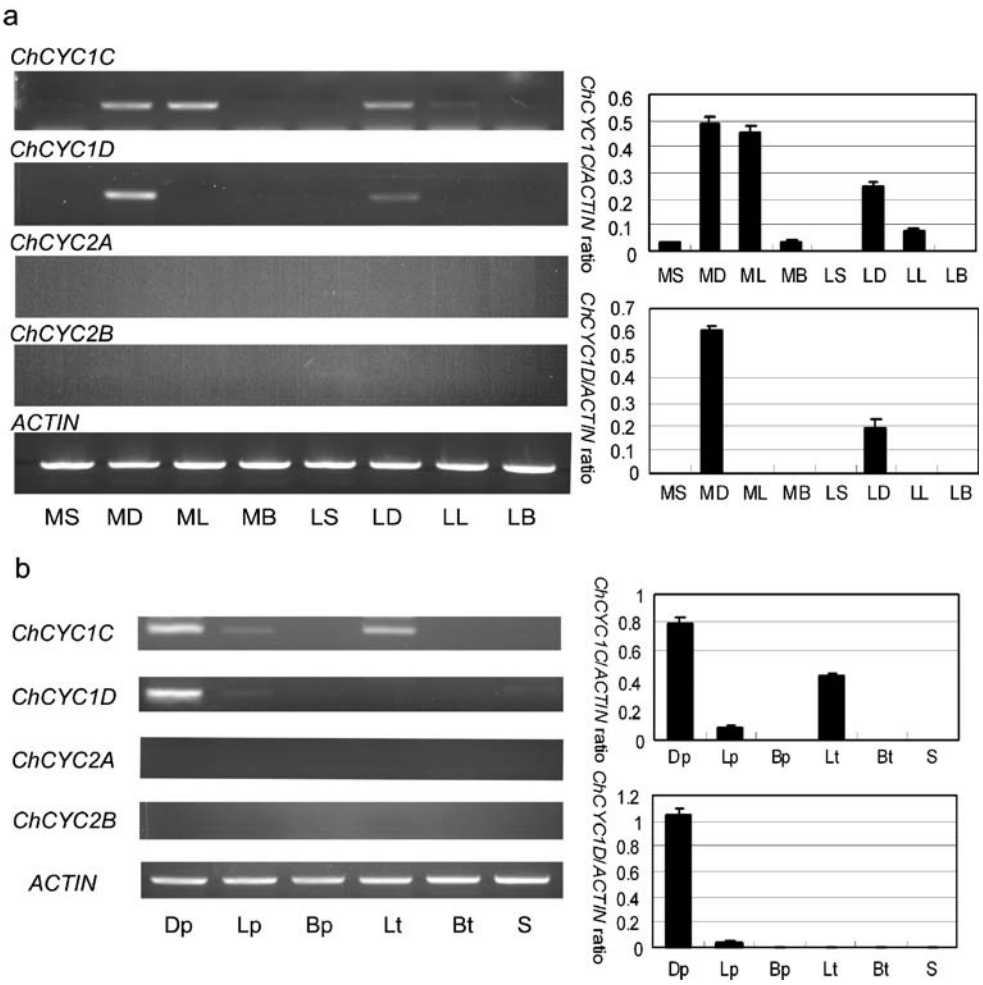

Fig. 3 Locus-specific semi-quantitative RT-PCRs on RNA prepared from dissected *C. heterotricha* flower buds. **a** Sepal (*S*) and adaxial (*D*)/lateral (*L*)/abaxial (*B*) corolla plus attached stamens were dissected from flower buds 1 (*M*) and 4 cm long (*L*). **b** adaxial/lateral/abaxial petals (*Dp*/*Lp*/*Bp*), lateral/abaxial stamens (*Lt*/*Bt*), and sepals (*S*) were dissected from 4-cm-long flower buds. *ACTIN* protein was used as for RT template control. The values (means±SD) shown are determined from five independent experiments

Discussion

Characterization of *CYC*-like genes in *C. heterotricha*

The *CYC*-like genes in Gesneriaceae (*GCYC*) constitute a small gene family that is considered to be derived from gene duplication (Citerne et al. 2000). The *GCYC* genes isolated previously are mostly only 70% of the ORF, in which the important functional domain, the TCP domain, is incomplete. Additionally, only two copies of *CYC*-like genes have been identified in each species investigated to date in Gesneriaceae, i.e., *GCYC1* and *GCYC2* or two copies of *GCYC1*. This is the first time that four copies of full-length ORF *CYC*-like genes from one species have been isolated in Gesneriaceae, i.e., *GCYC1* type, *ChCYC1C* and *ChCYC1D*, and *GCYC2* type, *ChCYC2A* and *ChCYC2B*. Phylogenetic analysis of the four *ChCYC* genes with representatives of the CYC (ECE) clade in class II TCP genes shows that they belong to the CYC2 lineage (Howarth and Donoghue 2006). The CYC2 lineage genes also include the *AtTCP1* gene from *Arabidopsis*, *LjCYC2* gene from *Lotus japonicus*, and *CYC* and *DICH* genes from *A. majus*, which all have roles in controlling the adaxial identity of a dorsoventrally asymmetric flower. Given that there is a closer relationship between *ChCYC* genes and *CYC/DICH*, the four *CYC*-like genes in *C.heterotricha* should be the homologues of *CYC/DICH*. Phylogenetic analysis of the four *ChCYC* genes along with other *GCYC* genes places the new genes described herein among previously published and characterized genes with high

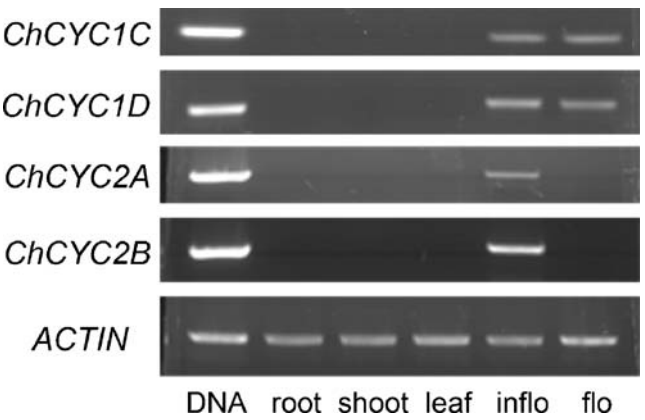

Fig. 4 RT-PCR on RNA prepared from root, shoot, leaf, young inflorescence (*inflo*), and flower bud (*flo*; shorter than 1 cm). *ACTIN* protein and total DNA of *C. heterotricha* were used as controls

Table 2 The similarity among *CYC*-like genes in *C. heterotricha* and *CYCLOIDIEA*

Gene	*ChCYC1C* (%)		*ChCYC1D* (%)		*ChCYC2A* (%)		*ChCYC2B* (%)	
ChCYC1D	88.48	83.81						
ChCYC2A	66.79	54.11	69.73	62.99				
ChCYC2B	64.95	51.56	68.09	58.19	91.26	88.52		
CYCLOIDEA	54.35	45.98	51.04	45.22	51.56	40.00	50.79	41.91

Left data are nucleotide similar identity, and the right data are amino acid similarity

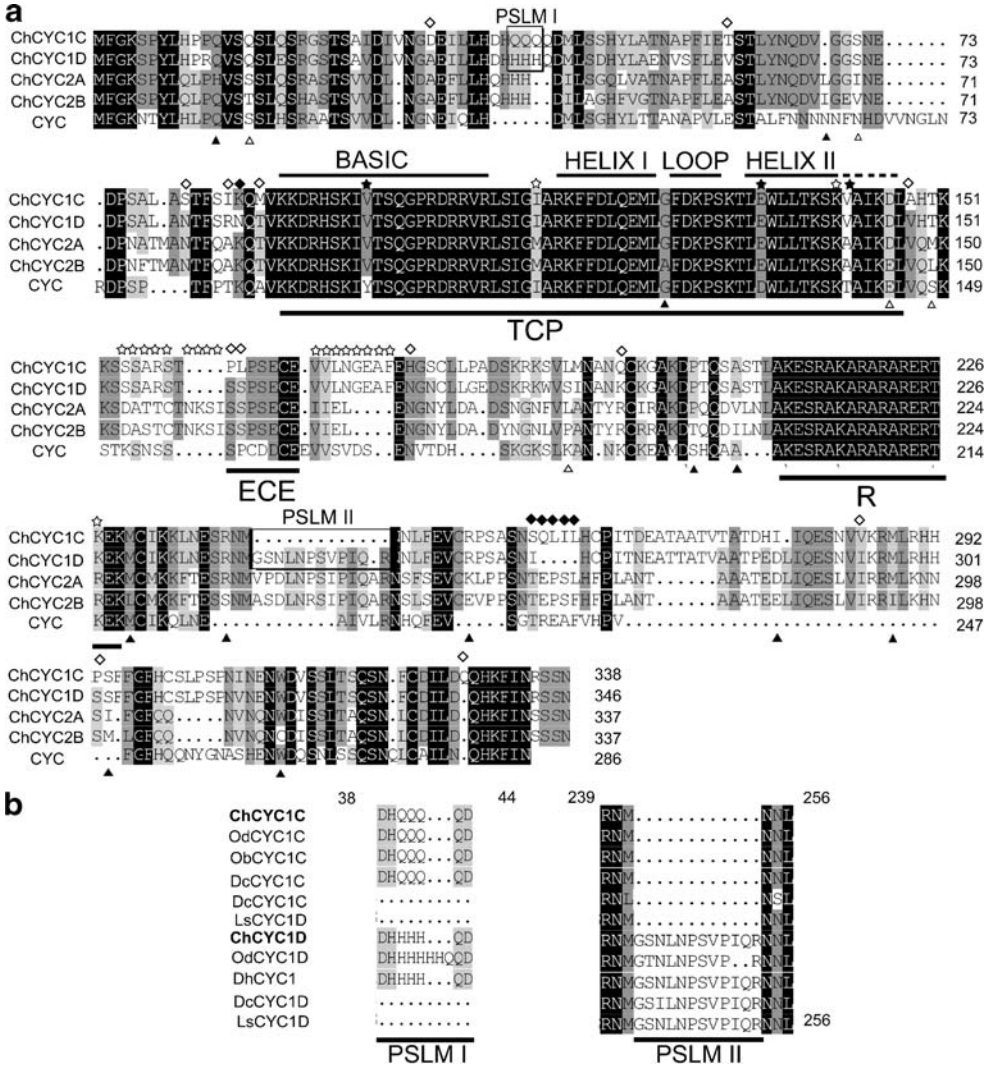

Fig. 5 Alignments of protein sequences of *ChCYC* genes with *CYC* or other *CYC*-like genes from Gesneriaceae (*GCYC*). **a** Alignment of the amino acid sequences of *CYC*-like genes from *C. heterotricha* and *CYC* from *A. majus*. TCP, ECE, R, and two PSLMs are *outlined*, and the identical amino acids are in *black boxes*. TCP domain consists of BASIC, HELIX, LOOP, and HELIX motifs. ◊ and ♦ on the *top* of alignment indicate the amino acids unique to ChCYC1C and ChCYC1D, respectively, compared to other GCYC1C and GCYC1D (data not shown). ▵ and ▲ *below* the alignment indicate the amino acids unique to ChCYC2A and ChCYC2B, respectively, compared to other GCYC2 (data not shown). Along the top of the TCP, R domain, and around the ECE motif, ★ indicate the amino acids divergent between GCYC and CYC, and ☆ indicate the amino acids divergent between GCYC1 and GCYC2. **b** Detail of PSLM I and PSLM II based on alignments of GCYC1C and GCYC1D from *C. heterotricha* (ChCYC1C and ChCYC1D), *Loxostigma* sp. (LsCYC1C and LsCYC1D), *D. citrinus* (DcCYC1C and DcCYC1D), *O. dinghushanensis* (OdCYC1C and OdCYC1D; data unpublished), and GCYC1 from *D. hancei* (data unpublished), *O. benthamii*, and *B. leiophylla* (DhCYC1, ObCYC1, and BlCYC1)

support. The phylogenetic relationships of different clades, e.g., *GCYC1A*/*GCYC1B*, *GCYC1C*/*GCYC1D*, and *GCYC2* clades, conform to previous phylogenetic trees (Möller et al. 1999; Citerne et al. 2000; Wang et al. 2004). Therefore, the phylogenetic analysis strengthens the identification of the four genes based on our BLAST results. In the *GCYC2* clade, however, *ChCYC2A* is sister to *Oreocharis GCYC2*, while *ChCYC2B* is sister to *Cyrtandra GCYC2*, and together they are sister to other *GCYC2* genes from *Conandron* and *Bournea*. These results indicate that the two copies of *GCYC2* genes in *C. heterotricha* are not two alleles from the same locus. In addition, the positions of *ChCYC2A* and *ChCYC2B* located in different branches of *GCYC2*, respectively, imply that the two copies do not come from a recent duplication event, although they are currently only known from *C. heterotricha*. The lack of additional copies of *GCYC2* as well as some *GCYC1* in other genera of Gesneriaceae may be due to incomplete sequence isolation in previous studies (Möller et al. 1999; Citerne et al. 2000; Wang et al. 2004).Thus, the ancestor of *GCYC* might have undergone at least three duplications to result in four copies of *CYC*-like genes in *C. heterotricha*.

Expression pattern differentiation of *ChCYC1C*, *ChCYC1D*, *ChCYC2A*, and *ChCYC2B*

The locus-specific semiquantitative RT-PCR here shows that the expression of each of the four copies of *CYC*-like genes in *C. heterotricha* (Gesneriaceae) differ spatially and temporally. *ChCYC1C* and *ChCYC1D* are expressed both in the young inflorescence and flowers of different stages as *CYC* in *A. majus*. The expression of *ChCYC1D* is restricted in the adaxial part of flowers in *C. heterotricha* including adaxial petals and staminode, while the expression of *ChCYC1C* ranges from the adaxial to the lateral floral organs. In the lateral floral domain, *ChCYC1C* mainly is expressed in lateral stamens. The comparison of *C. heterotricha* with *A. majus* in floral morphology shows that the specific expression patterns of *ChCYC1C* and *ChCYC1D* are strongly correlated with the differentiation of floral dorsoventral asymmetry in *C. heterotricha*. In its pentamerous flowers, the corolla lobes are characterized by two adaxials, two laterals, and one abaxial, in which the adaxial petals are smaller than the lateral and abaxial ones. In the third whorl of the *C. heterotricha* flower, in contrast to snapdragon with four didynamous stamens plus a staminode at the most adaxial position (Luo et al. 1996), only two abaxial stamens are fertile, while both the adaxial and lateral stamens are aborted (Fig. 1; also see Li and Wang 2004). The adaxial aborted stamen is tiny and barely detectable at anthesis, while the two infertile lateral stamens have short filaments that are one third or one half of the abaxial filaments in length. The remarkable difference between the aborted adaxial and lateral stamens is correlated with the distinctive expression patterns of *ChCYC1C* and *ChCYC1D* relating to different expression levels of *GCYC1* in the two floral regions, i.e., dual expressions of *ChCYC1C/1D* in the adaxials and single expression of *ChCYC1C* in the laterals that declines earlier in floral development. In the model species *A. majus*, the abortion of the adaxial stamen comes from *CYC* and *DICH* activities there (Luo et al. 1996, 1999). In *M. confertiflora*, a close relative of *A. majus*, *McCYC* and *McDICH* expressions expand from the adaxial to the lateral stamens in the third whorl and remains expressed in the adaxial petals in the second whorl, leading to abortion of both the adaxial and lateral stamens. However, no noticeable expressional differentiation takes place among the four genes with respect to abortion of lateral stamens in *Mohavea* (Hileman et al. 2003). Our findings present the first correlation between the abortion of lateral organs and the expression differentiation of *CYC* homologues with respect to a homeotic transformation from the adaxials to the laterals. It seems that the functions of *ChCYC1C* and *ChCYC1D* are largely redundant in the adaxial region, but they differentiate in controlling the lateral organs. Further studies with in situ hybridization and functional analysis are necessary for exploring the localization and functional diversification of *ChCYC* genes' activities in early stages of the flower. As the basal-most group in Lamiales s. l., Gesneriaceae is characteristic of weakly zygomorphic flowers (Endress 1998). The primitive zygomorphic groups in Gesneriaceae often have four didynamous stamens (two abaxials and two laterals) and one staminode (the aborted adaxial stamen). The zygomorphic flowers with three staminodes at the adaxial and lateral positions have been considered to be derived (Burtt 1990; Wang et al. 1992; Weber 2004; Li and Wang 2004). The expression of *ChCYC1C* in the lateral regions (mainly in the lateral stamens) should be interpreted as a gain of function, i.e., expanding from the pre-existing adaxial expression domain to the new lateral domain in the evolution of *GCYC* in Gesneriaceae.

Another interesting finding herein is that the expression of *ChCYC2A* and *ChCYC2B* is only detected in the young inflorescences. Their transcripts are undetectable both in flower buds and dissected petals and stamens. Given that no expression of *ChCYC2* is observed in vegetative organs, their mRNA signal detected in the young inflorescences may be related to a transient expression in the floral meristems as *AtTCP1* in *Arabidopsis* (Cubas et al. 2001). These results further indicate that the duplication of genes is often a major mechanism for the establishment of new genes with novel functions and the generation of evolutionary novelties (Moore and Purugganan 2003). The expressional differentiations of *ChCYC* genes contribute to the elaboration of the floral dorsoventral asymmetry in *C.*

heterotricha, which might be related to the protein sequence changes in their coding regions or some alterations in their upstream or downstream *trans*-regulators or *cis*-elements.

Protein sequence divergence correlated with expression differentiation

Protein sequence changes in the coding region of genes usually play an important role in the functional differentiation of the duplicated genes (Hsia and McGinnis 2003). Transcription factors and their DNA-binding sequences, which lie in the *cis*-regulatory elements of their target genes, are of central importance to the generation of phenotypic variation (Ramsay and Glover 2005). *CYC* and *CYC*-like genes belong to a transcription factor family, i.e., the TCP family, and have two conserved functional domains, i.e., the TCP domain and R domain. The TCP domain contains a conserved basic helix-loop-helix (HLH) motif in which the basic region is important in recognizing special DNA sequences and the HLH may be involved in protein–protein interactions (Cubas et al. 1999). The R domain is predicted to form a coil that may mediate protein–protein interaction. In addition to the amino acid sequences that diverge outside the TCP and R domains between *ChCYC* genes and *CYC*, in the TCP domain, there are three amino acid differences between *ChCYC* genes and *CYC*, which lie in the Basic, loop and Helix II regions, respectively. *ChCYC* genes have an additional long insertion after the R domain compared to *CYC*. These characteristics of *ChCYC* genes might relate to their expressional differentiation from *CYC*.

Within *ChCYC* genes, there are two lineage-specific amino acid alterations in the TCP domain and one amino acid change in the R domain between the two types of *ChCYC* genes, i.e., *ChCYC1* and *ChCYC2*. A motif called ECE was reported recently between the TCP and R domains in *CYC*-like genes, and its function was hypothesized to help stabilize the three-dimensional structure of the protein (Howarth and Donoghue 2005). The 17 amino acids around the ECE motif are lineage specific between *ChCYC1* and *ChCYC2*. The lineage-specific amino acid substitutions in these functional domains, i.e., TCP, R domains, and around ECE motif, are correlated with the expressional divergence between *ChCYC1* and *ChCYC2*.

Comparing the amino acid sequences of *ChCYC1C* and *ChCYC1D* with related *GCYC1* genes and other *CYC*-like genes outside of Gesneriaceae reveals two PSLMs (PSLM I and PSLM II) that are specific to *GCYC1* and are divergent between *GCYC1C* and *GCYC1D* sublineages but conserved inside each sublineage. Since conserved regions are usually functional domains (Cubas et al. 1999), these two submotifs may have special functions in controlling floral symmetry. The changes in protein sequence between *GCYC1C* and *GCYC1D* in the two putative submotifs are also observed between *ChCYC1C* and *ChCYC1D*. In addition to the above protein sequence divergence within domains and motifs, there are also numerous substitutions in the *ChCYC* genes compared to other genes in the same lineage or sublineage of *GCYC*. These may contribute to the floral morphology of *C. heterotricha*.

Gene duplication is common in gene evolution and is the primary source of new genes with novel functions (Gu et al. 2002). The fate of a duplicated gene is an interesting question, in which some models have been proposed about the connection between gene duplication and functional diversification, such as pseudogenization, conservation of gene function, neofunctionalization, subfunctionalization, and subneofunctionalization (Ohno 1970; Force et al. 1999; He and Zhang 2005.) A large amount of evidence shows that gene expression and functional divergence generally happen soon after gene duplication because the action of selective pressure on the functional domain of duplicate genes is diverse in gene evolution (Kreitman and Comeron 1999). When specialization or neofunctionalization is complete, duplicate genes are likely to be maintained under different functional constraints and show different substitution patterns (Zhang 2003). Our results show that the four *CYC*-like genes in *C. heterotricha* belong to different lineages and their spatiotemporal differentiations in expression patterns are congruent with their protein sequence divergence, especially in conserved domains. These suggest that *ChCYC* genes might have undergone an expressional differentiation and specialization in establishing the floral dorsoventral asymmetry in *C. heterotricha* responding to different selective pressures. The different selective pressures may further fix their function such that the paralogues with specialized expression patterns derived from gene duplication will persist.

In addition to protein sequence alterations in the coding region, possible changes in the *trans*- or *cis*-regulatory factor or element are also important for the expressional differentiation of *ChCYC* genes. Further functional researches, such as amino acid site-directed mutagenesis, are necessary for revealing the regulation mechanism and evolution of *ChCYC* genes in *C. heterotricha*. More sampling with in situ hybridization and functional investigations on *GCYC* genes including their upstream and downstream regions in Gesneriaceae are necessary to clarify the functional differentiation among different *GCYC* lineages and the evolutionary mechanism underlying gene duplication or loss events in Gesneriaceae, which would shed light on exploring the evolutionary history of *GCYC* genes responsible for the great morphological diversity of floral symmetry in Gesneriaceae.

Acknowledgments We are grateful to James F. Smith for critical comments and improvements on our manuscript, especially in language editing. We also thank Mr. Chen Yan for assistance in the greenhouse. This research was supported by CAS Grant KSCX2-YW-R-135 and National Natural Science Foundation of China Grant, nos. 30770147, 30121003.

References

Burtt BL (1990) Gesneriaceae of the Old World. I. New and little-known species of *Cyrtandra* from Malesia. Edinburgh J Bot 47:201–233

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

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

Cubas P, Coen E, Zapater JMM (2001) Ancient asymmetries in the evolution of flowers. Curr Biol 11:1050–1052

Damerval C, Guilloux ML, 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

Doebley J, Stec A, Gustus C (1995) Teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics 141:333–346

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

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Du ZY, Wang YZ (2008) Significance of RT-PCR expression patterns of *CYC*-like genes in *Oreocharis benthamii* (Gesneriaceae). Journal of Systematics and Evolution 46:23–31

Endress PK (1998) Antirrhinum and Asteridae—evolutionary changes of floral symmetry. Symp Soc Exp Biol 51:133–140

Endress PK (1999) Symmetry in flowers: diversity and evolution. Int J Plant Sci 160:S3–S23

Feng XZ, Zhao Z, Tian ZX, Xu SL, Luo YH, Cai ZG, Wang YM, Yang J, Wang Z, Weng L, Chen JH, Zheng LY, Guo XZ, Luo JH, Sato S, Tabata S, Ma W, Cao XL, Hu XH, Sun CR, Luo D (2006) Control of petal shape and floral zygomorphy in *Lotus japonicus*. Proc Natl Acad Sci USA 103:4970–4975

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151:1531–1545

Gu ZL, Cavalcanti A, Chen FC, Bouman P, Li WH (2002) Extent of gene duplication in the genomes of *Drosophila*, nematode, and yeast. Mol Biol Evol 19:256–262

Guindon S, Gascuel O (2003) A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Syst Biol 52:696–704

Guindon S, Lethiec F, Duroux P, Gascuel O (2005) PHYML Online: a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res 33:557–559

He XL, Zhang JZ (2005) Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. Genetics 169:1157–1164

Hileman LC, Kramer EM, Baum DA (2003) Differential regulation of symmetry genes and the evolution of floral morphologies. Proc Natl Acad Sci USA 100:12814–12819

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

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

Hsia CC, McGinnis W (2003) Evolution of transcription factor function. Curr Opin Genet Dev 13:199–206

Hubbard L, McSteen P, Doebley J, Hake S (2002) Expression patterns and mutant phenotype of *teosinte branched 1* correlate with growth suppression in maize and teosinte. Genetics 162:1927–1935

Kölsch A, Gleissberg S (2006) Diversification of *CYCLOIDEA*-like TCP genes in the basal eudicot families Fumariaceae and Papaveraceae s.str. Plant Biol 8:680–687

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

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

Kreitman M, Comeron JM (1999) Coding sequence evolution. Curr Opin Genet Dev 9:637–641

Li ZY, Wang YZ (2004) Plants of Gesneriaceae in China. Henan Science and Technology publishing house, Zhengzhou, China

Luo D, Carpenter R, Vincent C, Copsey L, Coen ES (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

Möller M, Clokie M, Cubas P, Cronk QCB (1999) Integrating molecular phylogenies and developmental genetics: a Gesneriaceae case study. In: Hollingsworth PM, Bateman RM, Gornall RJ (eds) Molecular systematics and plant evolution. Taylor and Francis, London, pp 375–402

Moore RC, Purugganan MD (2003) The early stages of duplicate gene evolution. Proc Natl Acad Sci USA 100:15682–15687

Ohno S (1970) Euolution by gene duplication. Springer, New York

Prasad K, Sriram P, Kumar SC, Kushalappa K, Vijayraghavan U (2001) Ectopic expression of rice OsMADS1 reveals a role in specifying the lemma and palea, grass floral organs analogous to sepals. Dev Genes Evol 211:281–290

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitutions. Bioinformatics 14:817–818

Ramsay NA, Glover BJ (2005) MYB–bHLH–WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci 10:63–70

Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony software (*and other methods) version 4.0b10. Available at: http://paup.csit.fsu.edu/

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Wang CN, Möller M, Cronk QCB (2004) Phylogenetic position of *Titanotrichum oldhamii* (Gesneriaceae) inferred from four different gene regions. Syst Bot 29:407–418

Wang WT, Pan KY, Li ZY (1992) Keys to the Gesneriaceae of China. Edinburgh J Bot 49:5–74

Weber A (2004) Gesneriaceae. In: Kubitzki K, Kadereit JW (eds) The families and genera of vascular plants. Springer, Berlin, pp 63–158

Xia X, Xie Z (2001) DAMBE: Software package for data analysis in molecular biology and evolution. J Heredity 92:371–373

Zhang JZ (2003) Evolution by gene duplication: an update. Trends Ecol Evol 18:292–298

Zhou XR, Wang YZ, Smith JF, Chen RJ (2008) Altered expression patterns of TCP and MYB genes relating to the floral developmental transition from initial zygomorphy to actinomorphy in *Bournea* (Gesneriaceae). New Phytol 178:532–543