Cellular and Molecular Biology
E-ISSN : 1165-158X / P-ISSN : 0145-5680
www.cellmolbiol.org

Original Research

Gene cloning, expression pattern analysis, and subcellular localization of *LIKE HETEROCHROMATIN PROTEIN 1* (*LHP1*) homologs in chrysanthemum (*Chrysanthemum morifolium* Ramat.)

Zhiping Wu, Yike Gao*, Min Fan, Yaohui Gao

Beijing Key Laboratory of Ornamental Plants Germplasm Innovation and Molecular Breeding, National Engineering Research Center for Floriculture, College of Ornamental Horticulture and Landscape Architecture, Beijing Forestry University, Beijing 100083, China

Correspondence to: gaoyk@bjfu.edu.cn

Received August 26, 2018; **Accepted** March 17, 2019; **Published** March 31, 2019

Doi: http://dx.doi.org/10.14715/cmb/2019.65.3.4

Abstract: Flowering is a very important developmental stage in the plant life cycle. *LIKE HETEROCHROMATIN PROTEIN 1* (*LHP1*) has been shown to participate in epigenetic silencing of flowering genes. Here, for the first time, we isolated and characterized six *CmLHP1* homolog genes from the important day-neutral ornamental *Chrysanthemum morifolium* cultivar 'Jin budiao'. These homolog genes were most likely generated by whole-genome duplication. Bioinformatic analysis showed that chrysanthemum *LHP1* homologs present low similarity to other plant *LHP1*-like genes. However, three nuclear localization signals and two domains were highly conserved among them. The secondary structures of the *CmLHP1* homologs mainly include α-helices and random coils, indicating that the proteins are mixed proteins. Phylogenetic tree analysis indicated that the six *CmLHP1* genes constituted a small clade and had the closest relationship with *LsLHP1* (*Lactuca sativa LHP1*). Quantitative RT-PCR analysis showed that the *CmLHP1* homologs were expressed in different tissues during the developmental period of chrysanthemum, but they were highly expressed in the buds, especially during the key S1 stage of the inflorescence. Furthermore, the expression patterns of *CmLHP1* homologs showed divergence under different photoperiods. Both *CmLHP1b* and *CmLHP1e* exhibited photoperiod sensitivity in leaves. Intriguingly, *CmLHP1c* was insensitive to photoperiod in both the shoot apexes and the leaves. Subcellular localization revealed that the six *CmLHP1* proteins were located in the nucleus. These results reveal that *CmLHP1* homolog genes could be strong candidates as important regulators of flowering time in chrysanthemum.

Key words: *Chrysanthemum morifolium*; *LIKE HETEROCHROMATIN PROTEIN1* (*LHP1*); Expression pattern, Subcellular localization; Flowering time.

Introduction

Chrysanthemum (*Chrysanthemum morifolium* Ramat.) is widely cultivated as a popular ornamental plant and is commercially used for potted or cut flowers or for horticulture (1). Flowering time is a very important trait for chrysanthemum. Most chrysanthemums are short-day plants whose flowering period is mainly concentrated in autumn, which limits the annual production of chrysanthemums. However, day-neutral chrysanthemums with longer flowering periods are not limited by the duration of sunshine, and they can flower at a suitable temperature (2). Therefore, identifying key factors regulating the flowering of day-neutral chrysanthemums and using genetic engineering to extend the flowering time of chrysanthemums are currently very important solutions. The molecular regulation pathways for the flowering times of many plants have been discovered, but most detailed studies have concentrated on model plants (3).

Flowering is a key step in the angiosperm life cycle that is strictly regulated by many endogenous and exogenous factors (4-7). In higher eukaryotes, Polycomb group (PcG) proteins play an extremely significant role in the epigenetic regulation of many genes. PcG proteins make up Polycomb Repressive Complex 1 (PRC1) and Polycomb Repressive Complex 2 (PRC2), which both have regulatory functions in epigenetic repression (8-12). *LIKE HETEROCHROMATIN PROTEIN 1* (*LHP1*), a PRC1 core subunit that can recognize trimethylation at lysine 27 of histone H3 (H3K27me3), participates in silencing chromatin genes. *LHP1* gene was originally identified in sieving of inflorescence meristems function (13, 14), and therefore was also known as *TERMINAL FLOWER 2*. Arabidopsis *LHP1* is a single-copy gene that is structurally homologous to the protein HETEROCHROMATIN PROTEIN 1 (HP1) of metazoans (15, 14). *LHP1* interacts with proteins of different cell types to carry out distinct functions. In vivo, *LHP1* is needed to establish complete H3K27me3 levels in protein complexes by direct interaction with *MSI1* (16). In addition, *LHP1* has been shown to affect H3K27me3 levels at thousands of gene loci (17).

Mutations in *LHP1* in Arabidopsis affect flowering time, plant architecture, inflorescence determinacy, leaf and root morphology, temperature and photoperiod sensitivity, and hormone levels (14, 18, 19). DNA microarray analysis has demonstrated that the expression levels of flowering genes such as *FT*, *AGAMOUS* (*AG*), *PISTILLATA* (*PI*), *APETALA3* (*AP3*), and *SEPALLATA3* (*SEP3*) are up-regulated by *LHP1* mutations (20, 14, 21, 22). Molecular studies have revealed that *LHP1* controls flowering time mainly by binding to H3K27me3 and directly interacting with *FLOWERING*

LOCUS T (*FT*) chromatin repression (23). Furthermore, in the autonomous and vernalization pathways, *LHP1* controls flowering time by recognizing and maintaining H3K27me3 in *FLC* (24, 16). To date, many plant *LHP1* homologs have been identified (25, 24). However, little is known about *LHP1* expression, evolution, and function and the regulatory mechanism of flowering time in the chrysanthemum.

Thus, six *CmLHP1* homolog genes were isolated and identified from day-neutral chrysanthemum, and their protein sequences, molecular evolution, expression patterns and subcellular localization were studied. The objectives of this study were to elucidate the basic features, evolution, expression patterns and protein functions of *CmLHP1* homologs, with the intention of providing a theoretical basis for the use of *CmLHP1* homolog genes in regulating flowering time in chrysanthemum molecular breeding.

Materials and Methods

Plant materials and growth conditions

The cultivar of day-neutral *Chrysanthemum morifolium* 'Jin budiao' (JBD) was grown in a greenhouse in a modern laboratory at 25/20℃ (day/night) under a nature photoperiod with 60% relative humidity in Xiao Tangshan. The laboratory belonged to Beijing Forestry University, Changping District, Beijing, China.

Isolation of *CmLHP1* homologs

Total RNA was isolated from young leaves of *Chrysanthemum morifolium* 'Jin budiao' using TRIzol reagent (Tiangen, China) according to the manufacturer's instructions. The RNA integrity was assessed by agarose gel electrophoresis and the RNA concentration was measured with a spectrophotometer. Then, a TIANScript RT kit (Tiangen, China) was used to synthesize the first strand of cDNA. Specific primer pairs (*CmLHP1*-F1/R1, Table 1) for amplifying the *CmLHP1* homolog genes were designed based on transcriptome data (NCBI accession number SRP109613) obtained in a previous study. PCR amplification was carried out using a high-fidelity enzyme (TransStart FastPfu DNA Polymerase) (TransGen, China). The PCR amplification conditions were as follows: 95℃ for 20 s, 48℃ for 20 s, and 72℃ for 1 min for 35 cycles. Subsequently, the amplified product was subcloned into a pLB-Simple vector (Tiangen, China) and transformed into *E. coli* DH5α for sequencing. The cloned mRNA coding sequences of these genes were then submitted to GenBank.

Bioinformatic and phylogenetic analysis

Analysis of the deduced protein sequences of the *CmLHP1* homologs was performed using the NCBI BLAST program (https://www.ncbi.nlm.nih.gov/blast/Blast.cgi). The physicochemical properties of the *CmLHP1* homologs were analyzed online with ExPASy analysis software (https://web.expasy.org/protparam/). The secondary structures of the *CmLHP1* proteins were predicted using the SOPMA program (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html). Multiple alignments of *CmLHP1* homolog sequences and *LHP1*-like protein sequences from other plant species were performed using ClustalW (26) with the default parameters. BioEdit software (version 7.0) was used to edit the aligned sequences. Molecular evolutionary and phylogenetic tree analyses were performed with MEGA 5.0 software (27, 28) using the neighbor-joining method with 1000 bootstrap replicates. Conserved motifs of *CmLHP1* homolog genes were predicted using the MEME online tool (http://meme-suite.org/tools/meme).

Real-Time PCR analysis

CmLHP1 homolog expression patterns were analyzed by real-time fluorescence quantitative PCR (qRT-PCR), which was mainly conducted in different tissues (roots, stems, leaves, buds, shoot apexes, and flowers) and in inflorescences at late developmental stages at which the ray florets began to enlarge (stages S1, S2, S3, S5, and S8, Fig. 3b). The late inflorescence development stages of JBD were defined according to the definitions of Gerbera (29). In addition, qRT-PCR was also used to detect the expression pattern of *CmLHP1* homologs in the leaves and shoot apexes of plants that were grown under different photoperiods. The seedlings of JBD were grown in long-day (LD, 16 h light/8 h dark) conditions. Some of the plants continued to grow under LD conditions, and the others were transferred to short-day (SD, 8 h light/16 h dark) conditions for 15 days. All samples were collected from JBD. Once collected, all plant materials were immediately put into liquid nitrogen and stored at −80℃. There were three biological replicates for each sample. Total RNA extraction and cDNA synthesis were completed as described above.

For accurate gene expression analysis, genomic DNA was removed from the total RNA. qRT-PCR was performed on an ABI StepOne system (ABI, USA) with SYBR Premix Ex Taq II (TaKaRa, Japan). All gene-specific primers for qRT-PCR are listed in Table 1. Their specificity and efficiency were examined before performing the qRT-PCR analysis. The qRT-PCR program was 95℃ for 3 min followed by 45 cycles of 95℃ for 7 s, 57℃ for 10 s, and 72℃ for 15 s. To confirm the product specificity of each primer pair, melting curve analysis was performed. The expression levels were calculated

Table 1. Primers used for *CmLHP1* homologs isolation and qRT-PCR in chrysanthemum.

Primer name	Forward sequence (5'→3')	Reverse sequence (5'→3')
CmLHP1-F1/R1	AAGAAGGATTTATACAAAATTCACTACT	TGTGTTGAAAAATTTCAAAACTTGGG
CnActin-F1/R1	TTTGAAGTATCCCATTGAGCAC	GCATAAAGAGAAAGCACGGC
CmLHP1b-F1/R1	AAGAAAGCTTAAAATCGGGG	TTGTAAAGGTTCCTCGGTGAT
CmLHP1c-F1/R1	GGAAATGAATACGATGCGATGT	CAGCTTCGATCTCATAAAACCCT
CmLHP1d-F1/R1	AAAGTCAGGGAAAAGGAGATCA	TTGTAAAGGTTCCTCGGTGAT
CmLHP1e-F1/R1	AACAACAACCACAACCACAACC	TACTCCGTCTTCCCCTTTCG
CmLHP1f-F1/R1	CACAACTGTTCCTGTTGGAAAA	CTTTCTGGAGTAGCAACAGTGTCT

Table 2. The features of six chrysanthemum *LHP1*-like cDNAs.

	Formula	Amino acids in length	Estimated molecular weight (kDa)	Theoretical isoelectric point	Instability index	Grand average of hydropathicity
CmLHP1a	$C_{1888}H_{3031}N_{529}O_{643}S_8$	389	43.69	5.11	52.81	-0.926
CmLHP1b	$C_{1888}H_{3033}N_{529}O_{641}S_8$	389	43.66	5.15	52.59	-0.906
CmLHP1c	$C_{1891}H_{3034}N_{530}O_{648}S_8$	391	43.82	5.07	53.28	-0.936
CmLHP1d	$C_{1888}H_{3033}N_{529}O_{643}S_8$	388	43.69	5.12	52.56	-0.936
CmLHP1e	$C_{1908}H_{3063}N_{535}O_{647}S_8$	393	44.11	5.15	54.11	-0.923
CmLHP1f	$C_{1889}H_{3035}N_{529}O_{641}S_8$	389	43.67	5.16	53.30	-0.906

using the $2^{-\Delta\Delta Ct}$ method (30), with *CnActin* (GenBank accession no. KF305683.1) as the internal control (31).

Vector construction and subcellular localization analysis

The transient expression of the six *CmLHP1* homolog genes was analyzed using the expression vector pH7FWG2-35S-GFP (Invitrogen, USA). To construct *35S::CmLHP1a-GFP*, the high-fidelity PCR product of the gene of interest was ligated into the entry vector pENTR (Invitrogen, USA) through TOPO cloning. After transformation, the positive clones were selected and sequenced. The plasmid with the correct insert was used as the entry vector for the LR reaction. The LR reaction included 3 µl of entry vector, 1 µl of the plant expression vector, and 1 µl of LR Clonase enzyme mix for a total reaction volume of 5 µl. After the LR reaction, *E. coli* DH5α competent cells were transformed and incubated at 37°C for 12-15 h. Recombinant positive clones were selected and confirmed by sequencing. Finally, the *35S::CmLHP1a-GFP* vector was obtained. The construction of the GFP fusion vectors of the remaining five genes was performed as described above.

After removing the outer 3-4 layers of scales from a well-grown onion bulb, the remaining inner scales were sterilized with 75% alcohol for 1 min, washed 3 times with sterile water and cut into small pieces of 1 cm^2. The inner epidermis was gently torn off with tweezers for infection. The single-clone bacteria harboring the target plasmid were cultured until they reached an OD_{600} of 0.6 and were then centrifuged at 5000 rpm for 10 min, and the supernatant was discarded. The bacteria were resuspended in 1/2 MS and incubated with the onion inner epidermis for 20 min. Then, the onion epidermis was dried with autoclaved filter paper, transferred to coculture medium (MS medium + AS 15 mg/L) and cultured at 25°C in the dark. After 72 hours, the cocultured onion epidermis was removed, washed with sterile water to remove Agrobacterium, and then surveyed and photographed with laser confocal scanning microscopy (Olympus, Japan), with the *35S::GFP* vector as a reference.

Results

Isolation and characterization of six *CmLHP1* homologs from chrysanthemum

The cDNA sequences of six *CmLHP1* homologs were cloned from chrysanthemum by RT-PCR and named in accordance with the previously published nomenclature as follows: *CmLHP1a*, *CmLHP1b*, *CmLHP1c*, *CmLHP1d*, *CmLHP1e*, and *CmLHP1f* (GenBank accession nos. KX398336-KX398341). Sequencing and bioinformatic analysis showed that the cDNA sequences of the six *CmLHP1* homologs encoded 388-393 amino acid residues with theoretical isoelectric points of 5.07-5.16, estimated molecular weights of 43.66-44.11 kDa, instability index values of 52.56-54.11, and grand average of hydropathicity values of −0.936 to −0.906 (Table. 2). Therefore, all of the encoded proteins are unstable and hydrophilic.

According to the NCBI web server (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), the domain structures of the six *CmLHP1* homolog sequences

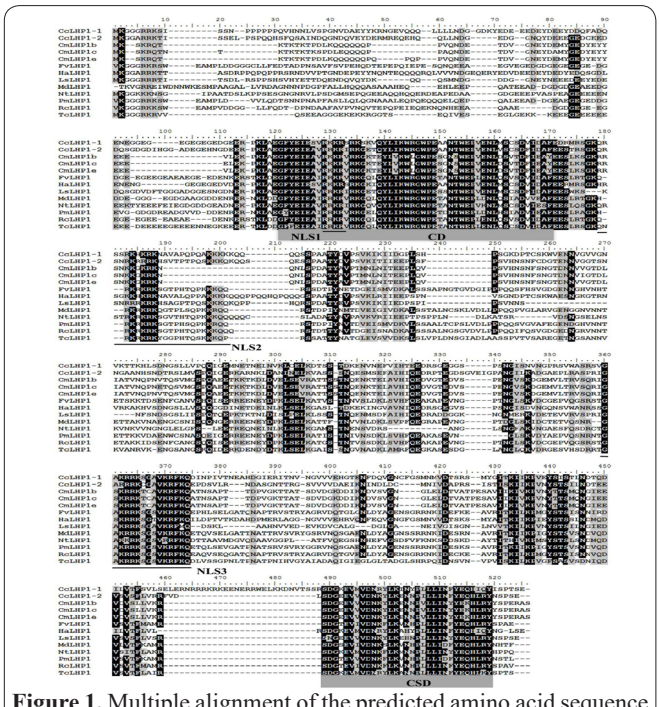

Figure 1. Multiple alignment of the predicted amino acid sequence of *CmLHP1* and selected other homolog proteins in plants. GenBank accession numbers of all listed protein sequences are as follows: *Cynara cardunculus var. scolymus* (*CcLHP1-1*, KVH89325.1), *Cynara cardunculus var. scolymus* (*CcLHP1-2*, KVI00608.1), *Chrysanthemum morifolium* (*CmLHP1b*, KX398337), *Chrysanthemum morifolium* (*CmLHP1c*, KX398338), *Chrysanthemum morifolium* (*CmLHP1e*, KX398340), *Fragaria vesca subsp.vesca* (*FvLHP1*, XP_004307292.1), *Helianthus annuus* (*HaLHP1*, XP_022035403.1), *Lactuca sativa* (*LsLHP1*, XP_023768841.1), *Malus domestica* (*MdLHP1*, BAF75821.1), *Nicotiana tabacum* (*NtLHP1*, NP_001312737.1), *Prunus mume* (*PmLHP1*, XP_008232209.1), *Rosa chinensis* (*RcLHP1*, XP_024192099.1), *Theobroma cacao* (*TcLHP1*, EOY10372.1). Identical amino acids are shown in black (represents amino acid 100% identity), similar in gray (represents amino acid identity > 70%). The conserved regions of *CD*, *CSD* are indicated by gray boxes and underlines for the conserved region of NLS (14, 15, 32).

showed that they were all members of the chromatin organization modifier (CHROMO) and chromo shadow (ChSh) superfamilies.

Sequence analysis of the chrysanthemum *CmLHP1* homologs

An alignment of the six *CmLHP1* homolog proteins and other plant *LHP1* proteins retrieved from NCBI was performed using ClustalW and was further refined manually. Then, we constructed a phylogenetic tree with MEGA 5.0 software using the neighbor-joining method. The phylogenetic analysis revealed that the *CmLHP1* homologs compose a small clade and have the closest relationship with *LsLHP1* (Fig. 2a).

A conserved motif analysis demonstrated that all these proteins included the chromo shadow domain (CSD), chromodomain (CD), hinge region 1 (HR1), hinge region 2 (HR2), and nuclear localization signal 3 (NLS3) motif (Fig. 2b). The CD, CSD and NLS3 were clearly more conserved than HR1 or HR2 (Fig. 2b). The sequence variation among these proteins is primarily located in the nonconserved region at the N-terminus and in the HR with low conservation.

Protein secondary structure prediction of the chrysanthemum *CmLHP1* homologs

The secondary structure of a protein mainly refers to the structure of the main peptide chain under the influence of hydrogen bonding, with regular curling and folding forming a periodic structure in the one-dimensional direction. Analysis of secondary structures aids in the study of the functions of proteins. The results revealed that the secondary structures of the six *CmLHP1* homologs were composed of 21.85%-25.26% α-helices, 62.63%-67.35% random coils, 7.38%-8.51% extended strand structures, and 3.08%-3.61% β-turns. These results indicate that the *CmLHP1* homolog proteins are mixed proteins.

Expression pattern of the chrysanthemum *CmLHP1* homologs

To elucidate the *CmLHP1* homologs expression patterns in chrysanthemum, qRT-PCR experiments were carried out to detect the expression levels of the *CmLHP1* homologs. First, we investigated the expression levels of the *CmLHP1* homologs in different tissues of JBD. The results showed that all the genes were expressed in different tissues and were highly expressed in the buds, followed by the leaves and the shoot apexes (Fig. 3a). The homology with the highest expression in the buds was *CmLHP1e*, the homology with the lowest expression in the buds was *CmLHP1c*, and the homology with moderate expression in the buds was *CmLHP1b* (Fig. 3a). qRT-PCR reactions were also performed to compare the expression patterns of these three genes in the inflorescences of JBD at late developmental stages. Intriguingly, the expression levels of these three genes were downregulated from stage S1 to stage S8, with little expression in stage S8. The degree of the decline in the expression level of *CmLHP1c* was significantly greater than that of *CmLHP1b* and *CmLHP1e* (Fig. 3c). The *CmLHP1* homologs were expressed in different tissues during the developmental period of chrysanthemum, but they were highly expressed in the buds and

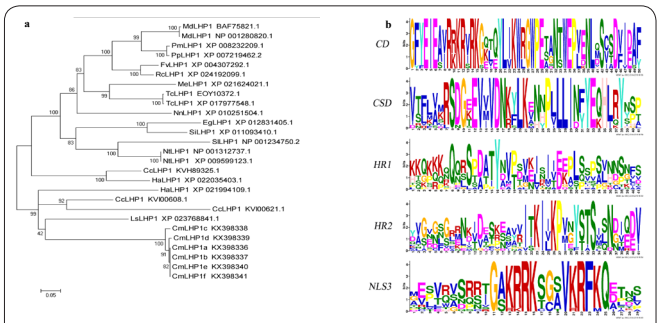

Figure 2. Phylogenetic and conserved motif analysis of *LHP1*-like proteins from chrysanthemum and selected other plant species. **a** Phylogenetic analysis of selected *LHP1*-like proteins using the neighbor-joining method with 1000 bootstrap replicates. GenBank accession numbers of all sequences are displayed following gene names. Species abbreviations are as follows: Cc, *Cynara cardunculus var. scolymus*; Cm, *Chrysanthemum morifolium*; Eg, *Erythranthe guttata*; FV, *Fragaria vesca subsp.vesca*; Ha, *Helianthus annuus*; Ls, *Lactuca sativa*; Md, *Malus domestica*; Me, *Manihot esculenta*; Nn, *Nelumbo nucifera*; Nt, *Nicotiana tabacum*; Nt, *Nicotiana tomentosiformis*; Pm, *Prunus mume*; Pp, *Prunus persica*; Rc, *Rosa chinensis*; Si, *Sesamum indicum*; Sl, *Solanum lycopersicum*; Tc, *Theobroma cacao*. **b** The two conserved CD and CSD domains and three conserved HR1, HR2, and NLS3 motifs were analyzed by the MEME online tool. The total height of the stack demonstrates the amino acid residue conservation at that position.

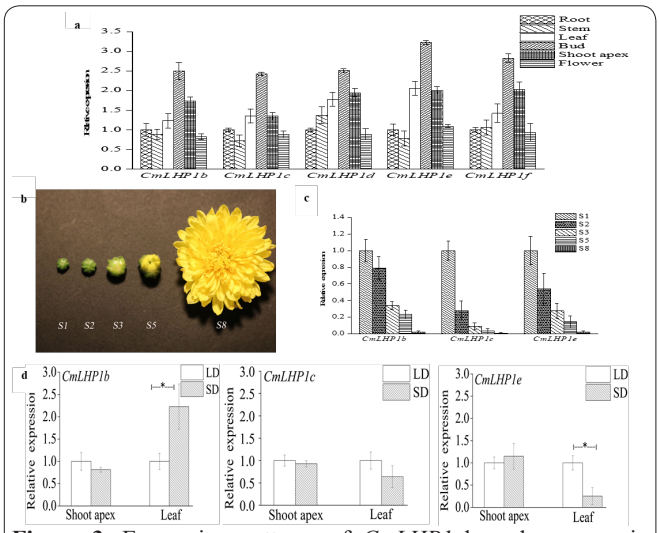

Figure 3. Expression patterns of *CmLHP1* homolog genes in *Chrysanthemum morifolium* 'Jin budiao'. The columns show the mean values of three biological replicates for each sample relative to the expression of the *CnActin* gene as a control. **a** Expression patterns of *CmLHP1* homologs in different tissues. The tissues included roots, stems, leaves, buds, shoot apexes, and flowers. **b** Inflorescences of JBD at late developmental stages at which the ray florets began to enlarge (stages S1, S2, S3, S5, and S8). **c** Expression analysis of *CmLHP1* homologs in inflorescences of JBD during the late development stages. S1-S8 indicate the development stage (stage 1- stage 8) of the inflorescence. **d** Expression patterns of *CmLHP1* homologs in shoot apexes and leaves under different photoperiods including SD (short-day, 8 h light/16 h dark) and LD (long-day, 16 h light/8 h dark) photoperiods. Compared to that of LD condition, the expression level of *CmLHP1b* was significantly increased in leaf, but the expression levels of *CmLHP1e* was significantly decreased in leaf. *p<0.05 (compared with LD condition).

downregulated in the late inflorescence development stages.

As mentioned above, the expression levels of the

CmLHP1 homologs in leaves and shoot apexes were second only to the levels in the buds. To further investigate the expression patterns of the *CmLHP1* homologs under different photoperiods, we used shoot apexes and leaves as samples for analysis. As shown in Figure 3d, under different photoperiods, the expression patterns of *CmLHP1b* in the leaves and shoot apexes were different. Upon the transition of JBD from LD condition to SD condition, *CmLHP1b* was significantly upregulated in the leaves. There was no significant change in the expression of *CmLHP1b* in the shoot apexes. Under the same conditions, the expression patterns of *CmLHP1e* in the leaves and shoot apexes were also different. The expression level of *CmLHP1e* in the leaves was significantly downregulated. There was no significant change in the expression of *CmLHP1e* in the shoot apexes. However, under the same conditions, the expression patterns of *CmLHP1c* in the leaves and shoot apexes were the same; *CmLHP1c* expression was downregulated, but the levels were not markedly different. It can be seen that the expression patterns of *CmLHP1* homologs showed divergence under different photoperiods. Both *CmLHP1b* and *CmLHP1e* exhibited photoperiod sensitivity in leaves. The difference was that the expression trends of *CmLHP1b* and *CmLHP1e* in the leaves were opposite. Interestingly, *CmLHP1c* was insensitive to photoperiod in both the shoot apexes and the leaves.

Overall, these results reveal that *CmLHP1* homologs are likely to play an important role in regulating flowering time in chrysanthemum.

Subcellular localization of the six *CmLHP1* homolog proteins

To investigate the localization of the chrysanthemum *CmLHP1* homolog proteins in cells, the six *CmLHP1* homolog proteins were first predicted using the online CELLO server (http://cello.life.nctu.edu.tw/). The results showed that all the proteins localized in the nucleus. To further determine the localization of the chrysanthemum *CmLHP1* homologs, six fusion vectors, *35S::CmLHP1a-GFP*, *35S::CmLHP1b-GFP*, *35S::CmLHP1c-GFP*, *35S::CmLHP1d-GFP*, *35S::CmLHP1e-GFP*, and *35S::CmLHP1f-GFP*, were used to transiently transform onion epidermal cells. As shown in the results of laser confocal microscopy, the fluorescence signals of the six chrysanthemum *CmLHP1* proteins were strong in the cell membrane and nucleus, while the fluorescence signal of the empty vector (*35S::GFP*) was strong in the cell membrane (Fig. 4), indicating that the six chrysanthemum *CmLHP1* proteins were located in the nucleus, consistent with the predicted subcellular localization. These results suggested that the six *CmLHP1* homolog proteins were transcription factors.

Discussion

LHP1-like genes have been verified to function as important regulators controlling flowering time in plants (14, 20, 25, 33). However, most studies have focused on Arabidopsis. Until now, there have been no reports on the functions of the *LHP1* gene in chrysanthemum. Therefore, we isolated and characterized six *CmLHP1* genes from chrysanthemum. Multiple alignment of amino acid sequences showed that plant *LHP1* homologs present low similarity, indicating that these genes have rapidly diverged during the evolution of plant species (Fig. 1). The hinge region of *LHP1* is not well conserved, and the HR connects the CD and the CSD to facilitate overall functionality (34). Conserved motif analysis clearly revealed that the HR was the least conserved among all the motifs (Fig. 2b). However, two conserved CD and CSD domains, and three nuclear localization sequences (NLS1, NLS2 and NLS3) of the *LHP1* genes were highly conserved between chrysanthemum and other plant *LHP1* homologs, which is consistent with the results of previous studies (15, 25, 35). It is well established that gene and whole-genome duplications (WGD) have increased genomic complexity and diversity in the evolution of plants (36). WGD have been rampant in the evolution of flowering plants. This is the reason why most flowering plants originate from ancestors with homologous or heterologous polyploidy. Because polyploidy can easily generate functional divergence among homologs, it can enhance the environmental adaptation of plant species (37). The evolution of plant *LHP1* reveals a 'duplication after speciation' topology (38). There are only 2-4 *LHP1* members in other species except for Asteraceae, but there are 6 *LHP1* genes in chrysanthemum, indicating that *CmLHP1* homolog genes are highly likely to have undergone gene duplication events in the evolution of chrysanthemum, which is consistent with the gene duplication event of the *CYC-like* genes in Asteraceae (31, 39). Numerous studies have shown that the genetic background of chrysanthemum is complex and diverse. Variation due to polyploidy and aneuploidy is widespread in Asteraceae. The chrysanthemum JBD

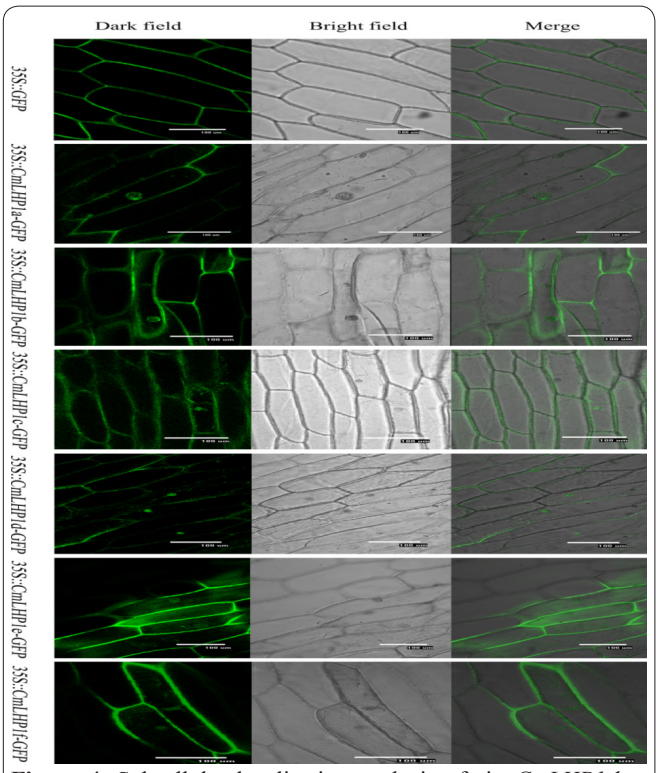

Figure 4. Subcellular localization analysis of six *CmLHP1* homolog proteins in onion epidermal cells. Bars=100um. Images are displayed as dark field, bright field and merged, with *35S::GFP*, *35S:: CmLHP1a-GFP*, *35S::CmLHP1b-GFP*, *35S:: CmLHP1c-GFP*, *35S::CmLHP1d-GFP*, *35S::CmLHP1e-GFP*, and *35S::CmLHP1f-GFP*, respectively.

in this study is a heterologous hexaploid. Intriguingly, in the phylogenetic analysis, the six *CmLHP1* genes first clustered with the homolog genes of lettuce and then clustered with those of sunflowers into a large branch, which distinctly revealed an intimate genetic relationship between *CmLHP1* and *LsLHP1*. The phylogenetic analysis also showed that the relationship between chrysanthemum and lettuce was closer than that between chrysanthemum and sunflower.

As mentioned above, NLS1, NLS2 and NLS3 are present in most plant *LHP1* protein sequences. Considering that most plant *LHP1* genes have a KKRK motif corresponding to *SlLHP1* NLS3 (32), suggesting plant *LHP1* homologs maybe target to the nucleus. In eukaryotes, the subcellular localization of proteins is very important for studying protein functions, as it can initially determine where the protein functions. Our data demonstrated that the six *CmLHP1* homolog proteins were localized in the nucleus (Fig. 4). This reveals that the *CmLHP1* homologs are transcription factors, consistent with the findings of previous studies (23).

Previous studies have shown that Arabidopsis *LHP1* is mainly expressed in the lateral roots, shoot apical meristems, young leaves, vascular bundles and flower organs (14, 20). In chrysanthemum, *CmLHP1* was expressed in different tissues, consistent with what has been reported for Arabidopsis (14, 20, 35). Guan *et al.* revealed that the *LHP1* homologs of many plants that are significant in plant evolution also share a conserved expression pattern (35). It is worth mentioning that *OsLHP1* (*Oryza sativa*) is not expressed in shoot apical meristems, indicating that the function of *OsLHP1* may have diverged considerably (35). In this study, the expression of the *CmLHP1* homologs had both redundant and specific patterns, similar to those reported by Huang *et al.* (31). The *CmLHP1* homologs were highly expressed in the buds, especially in the inflorescences at stage S1, when the growth of floral organs begins to transition into the maturation of inflorescences. However, the expression patterns of *CmLHP1* homologs showed divergence under different photoperiods. Both *CmLHP1b* and *CmLHP1e* exhibited photoperiod sensitivity in leaves. Interestingly, *CmLHP1c* expression was insensitive to photoperiod in both the shoot apexes and the leaves, which suggests that *CmLHP1c* may be an important factor in the regulation of flowering in day-neutral chrysanthemum. Together, these evidences indicate that the *CmLHP1* genes are likely to play an important role in regulating flowering time in chrysanthemum. In further research, we will use transgenic technology to elucidate how the *CmLHP1* homologs regulate flowering time in chrysanthemum.

Acknowledgments
This work was supported by the National High Technology Research and Development Program of China (2013AA102706).

Interest conflict
We declare that we have no conflict of interest.

Author's contribution
ZPW and YKG conceived and designed the experiments. MF prepared the plant materials. YHG contributed analysis tools. ZPW performed experiments and wrote the paper. All authors read and approved the manuscript.

References

1. Suvija NV, Suresh J, Kumar RS, Kannan M. Evaluation of chrysanthemum (*Chrysanthemum morifolium* Ramat.) genotypes for loose flower, cut flower and pot mums. Int j innov res adv stud (IJIRAS) 2016; 3(4): 100-104.
2. Anderson NO, Ascher PD. Selection of day-neutral, heat-delay-insensitive *Dendranthema* × *grandiflora* genotypes. J Am Soc Hortic Sci 2001; 126: 710-721.
3. Andrés F, Coupland G. The genetic basis of flowering responses to seasonal cues. Nat Rev Genet 2012; 13: 627-639.
4. Tan FC, Swain SM. Genetics of flower initiation and development in annual and perennial plants. Physiol Plant 2006; 128: 8-17.
5. Wilkie JD, Sedgley M, Olesen T. Regulation of floral initiation in horticultural trees. J Exp Bot 2008; 59: 3215-3228.
6. Guo D, Li C, Dong R, Li X, Xiao X, Huang X. Molecular cloning and functional analysis of the *FLOWERING LOCUS T* (*FT*) homolog *GhFT1* from *Gossypium hirsutum*. J Integr Plant Biol 2015; 57(6): 522-533.
7. Niu L, Fu C, Lin H, Wolabu TW, Wu Y, Wang ZY. Control of floral transition in the bioenergy crop switchgrass. Plant Cell Environ 2016; 39: 2158-2171.
8. Bemer M, Grossniklaus U. Dynamic regulation of Polycomb group activity during plant development. Curr Opin Plant Biol 2012; 15: 523-529.
9. Molitor A, Shen WH. The Polycomb complex PRC1: composition and function in plants. J Genet Genomics 2013; 40: 231-238.
10. Grossniklaus U, Paro R. Transcriptional silencing by polycomb group proteins. Cold Spring Harb Perspect Biol 2014; 6: a019331.
11. Xiao J, Wagner D. Polycomb repression in the regulation of growth and development in Arabidopsis. Curr Opin Plant Biol 2015; 23: 15-24.
12. Mozgova I, Hennig L. The polycomb group protein regulatory network. Annu Rev Plant Biol 2015; 66: 269-296.
13. Larsson AS, Landberg K, Meeks-Wagner DR. The *TERMINAL FLOWER2* (*TFL2*) gene controls the reproductive transition and meristem identity in *Arabidopsis thaliana*. Genetics 1998; 149: 597-605.
14. Kotake T, Takada S, Nakahigashi K, Ohto M, Goto K. Arabidopsis *TERMINAR FLOWER 2* gene encodes a heterochromatin protein 1 homolog and represses both *FLOWER LOCUS T* to regulate flowering time and several floral homeotic genes. Plant Cell Physiol 2003; 44: 555-564.
15. Gaudin V, Libault M, Pouteau S, Juul T, Zhao G, Lefebvre D, et al. Mutations in *LIKE HETEROCHROMATIN PROTEIN 1* affect flowering time and plant architecture in Arabidopsis. Development 2001; 128: 4847-4858.
16. Derkacheva M, Steinbach Y, Wildhaber T, Mozgova I, Mahrez W, Nanni P, et al. Arabidopsis *MSI1* connects *LHP1* to PRC2 complexes. EMBO J 2013; 32: 2073-2085.
17. Wang H, Liu C, Cheng J, Liu J, Zhang L, He C, et al. Arabidopsis flower and embryo developmental genes are repressed in seedlings by different combinations of polycomb group proteins in association with distinct sets of Cis-regulatory elements. PLoS Genet 2016; 12(1): e1005771.
18. Kim JH, Durrett TP, Last RL, Jander G. Characterization of the Arabidopsis *TU8* glucosinolate mutation, an allele of *TERMINAR FLOWER 2*. Plant Mol Biol 2004; 54: 671-682.
19. Cui H, Benfey PN. Interplay between *SCARECROW*, *GA* and *LIKE HETEROCHROMATIN PROTEIN 1* in ground tissue patter-

ning in the Arabidopsis root. Plant J 2009; 58: 1016-1027.
20. Takada S, Goto K. Terminal flower 2, an Arabidopsis homolog of heterochromatin protein 1, counteracts the activation of flowering locus T by constans in the vascular tissues of leaves to regulate flowering time. Plant Cell 2003; 15: 2856-2865.
21. Germann S, Juul-Jensen T, Letarnec B, Gaudin V. DamID, a new tool for studying plant chromatin profiling in vivo, and its use to identify putative *LHP1* target loci. Plant J 2006; 48: 153-163.
22. Nakahigashi K, Jasencakova Z, Schubert I, Goto K. The Arabidopsis heterochromatin protein1 homolog (*TERMINAL FLOWER 2*) silences genes within the euchromatic region but not genes positioned in heterochromatin. Plant Cell Physiol 2005; 46: 1747-1756.
23. Feng J, Lu J. *LHP1* Could Act as an Activator and a Repressor of Transcription in Plants. Front Plant Sci 2017; 8: 2041.
24. Exner V, Aichinger E, Shu H, Wildhaber T, Alfarano P, Caflisch A, et al. The chromodomain of *LIKE HETEROCHROMATIN PROTEIN 1* is essential for H3K27me3 binding and function during Arabidopsis development. PLoS One 2009; 4(4): e5335.
25. Mimida N, Kidou S, Kotoda N. Constitutive expression of two apple (*Malus× domestica* Borkh.) homolog genes of *LIKE HETEROCHROMATIN PROTEIN 1* affects flowering time and whole-plant growth in transgenic Arabidopsis. Mol Genet Genom 2007; 278: 295-305.
26. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics 2007; 23: 2947-2948.
27. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol 2011; 28(10): 2731-2739.
28. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 1987; 4: 406-425.
29. Helariutta Y, Elomaa P, Kotilainen M, Seppanen P, Teeri TH. Cloning of cDNA coding for *dihydroflavonol-4-reductase* (*DFR*) and characterization of *dfr* expression in the corollas of *Gerbera hybrida var.* Regina (Compositae). Plant Mol Biol 1993; 22 (2): 183-193.
30. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C_T method. Nat Protoc 2008; 3(6): 1101-1108.
31. Huang D, Li X, Sun M, Zhang T, Pan H, Cheng T, et al. Identification and Characterization of *CYC*-Like Genes in Regulation of Ray Floret Development in *Chrysanthemum morifolium*. Front Plant Sci 2016; 7: 1633.
32. Zemach A, Li Y, Ben-Meir H, Oliva M, Mosquna A, Kiss V, et al. Different domains control the localization and mobility of *LIKE HETEROCHROMATIN PROTEIN 1* in Arabidopsis nuclei. Plant Cell 2006; 18: 133-145.
33. Latrasse D, Germann S, Houba-He´rin N, Dubois E, Bui-Prodhomme D. Control of Flowering and Cell Fate by *LIF2*, an RNA Binding Partner of the Polycomb Complex Component *LHP1*. PLoS One 2011; 6(1): e16592.
34. Lomberk G, Wallrath L, Urrutia R. The heterochromatin protein 1 family. Genome Biol 2006; 7(7): 228.
35. Guan H, Zheng Z, Grey PH, Li Y, Oppenheimer DG. Conservation and divergence of plant *LHP1* protein sequences and expression patterns in angiosperms and gymnosperms. Mol Genet Genom 2011; 285: 357-373.
36. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE. Ancestral polyploidy in seed plants and angiosperms. Nature 2011; 473: 97-100.
37. Adams KL, Wendel JF. Polyploidy and genome evolution in plants. Curr Opin Plant Biol 2005; 8: 135-141.
38. Jacquemin J, Laudie M, Cooke R. A recent duplication revisited: phylogenetic analysis reveals an ancestral duplication highly-conserved throughout the Oryza genus and beyond. BMC Plant Biol 2009; 9: 146.
39. Tahtiharju S, Rijpkema AS, Vetterli A, Albert VA, Teeri TH, Elomaa P. Evolution and diversification of the *CYC/TB1* gene family in Asteraceae–a comparative study in Gerbera (Mutisieae) and sunflower (Heliantheae). Mol Biol Evol 2012; 29: 1155-1166.