

ISSN 1682-296X (Print)

ISSN 1682-2978 (Online)



# Bio Technology



**ANSI***net*

Asian Network for Scientific Information  
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan



## Research Article

# Characterization and spatial expression pattern of a novel *CONSTANS* gene (*GbCOL-6*) from *Ginkgo biloba*

Mingyue Fu, Jiaping Yan, Lanlan Wang, Xian Zhou, Junping Tan and Feng Xu

College of Horticulture and Gardening, Yangtze University, Jingzhou, 434025, Hubei, China

### Abstract

**Background and Objective:** *Ginkgo biloba* is an important medicinal plant but the transitional period from the sowing to the flowering of *G. biloba* is too long, which severely limits its reproduction and application. The *CONSTANS* (CO), is one of the important transcription factors in the transition from vegetative growth to reproductive growth. The CO plays a central role in photoperiodic flowering control. The aim of this study was to clone and characterize a *CONSTANS-like6* (*GbCOL-6*) gene from *G. biloba*. **Materials and Methods:** A pair of specific primers were designed based on the data of the *G. biloba* transcriptome. The full-length cDNA of the CO gene was amplified from the total RNA of *G. biloba* leaves by RT-PCR, named *GbCOL-6* (GenBank Accession No. MG251395). The expression of *GbCOL-6* gene in different tissues was studied by quantitative RT-PCR. **Results:** Sequence analysis results showed that the cDNA of *GbCOL-6* contained a 1479 bp open reading frame (ORF), which encodes a 492 amino-acid protein. The theoretical molecular weight and pI of the *GbCOL-6* are 54.65 kDa and 5.08, respectively. Similar to other CO proteins, *GbCOL-6* contains two conserved domains (B-box and CCT domain) and the amino acid sequence showed high similarity to other plant CO proteins. Phylogenetic tree analysis revealed that *GbCOL-6* belonged to the second group members of the *CONSTANS* gene family, which has a B-box domain and a CCT conserved domain. These results indicated that *GbCOL-6* is a member of the *CONSTANS* family and belongs to *CONSTANS-LIKE6*. Tissue expression profile analysis showed that *GbCOL-6* was expressed in all tissues, the highest expression level was detected in the female strobili and the lowest in the roots. **Conclusion:** This study found that *GbCOL-6* isolated from *G. biloba* was specifically expressed in the female strobili of *G. biloba*. However, *GbCOL-6* rarely was expressed in roots of *G. biloba*. This laid the foundation for further understanding of the molecular regulation mechanism of the flowering of *G. biloba* and shortening the juvenile phase of woody plant through the transgenic technology and providing genetic resources for promoting early flowering.

**Key words:** *Ginkgo biloba*, *CONSTANS*, flowering time, gene cloning, expression analysis

**Citation:** Mingyue Fu, Jiaping Yan, Lanlan Wang, Xian Zhou, Junping Tan and Feng Xu, 2018. Characterization and spatial expression pattern of a novel *CONSTANS* gene (*GbCOL-6*) from *Ginkgo biloba*. *Biotechnology*, 17: 104-112.

**Corresponding Author:** Feng Xu, College of Horticulture and Gardening, Yangtze University, Jingzhou, 434025, Hubei, China Tel: +86 13177040087

**Copyright:** © 2018 Mingyue Fu *et al.* This is an open access article distributed under the terms of the creative commons attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

**Competing Interest:** The authors have declared that no competing interest exists.

**Data Availability:** All relevant data are within the paper and its supporting information files.

## INTRODUCTION

*Ginkgo biloba* also known as 'gongsun tree'. *Ginkgo biloba* is a relict plant from the Cenozoic Quaternary glacial period, which has a long juvenile phase. *Ginkgo biloba* is a unique peculiar species in China, the general seedling planting 15-20 years to blossom results<sup>1</sup>. *Ginkgo biloba* is a subtropical, temperate deciduous tree species and one of the major economic tree species in China<sup>2</sup>. Its fruits are rich in flavonoids and terpene lactones, which improve cognitive expression, the brain function, antagonize platelet activation factor, scavenging free radicals, inhibition of lipid peroxidation and other effects<sup>3</sup>.

As the juvenile phase of *G. biloba* is particularly long, the selection of fine varieties of *G. biloba* has brought serious obstacles, making the *G. biloba* many economic and social values are subject to a large extent. Although traditional breeding methods such as seedling propagation, grafted seedling propagation, cutting propagation, root tiller and micropropagation<sup>4</sup>, shorten the ginkgo juvenile phase to some extent, it is still relatively longer, therefore, finding new ways to shorten the ginkgo juvenile phase is an urgent problem to be solved.

Flowering is an important signal before plant fruit. The process of plant flowering mainly involves the following regulatory pathways: autonomous pathway, photoperiod pathway, vernalization pathway, gibberellin pathway, carbohydrate induction and flowering inhibition pathway<sup>5</sup>. These regulatory pathways are independent of each other and interact, together constitute plant cell within the precise regulatory network<sup>6</sup>. The *CONSTANS (CO)* gene is one of important regulatory gene in the photoperiod pathway and its expression is affected by the photoperiod pathway Circadian Clock-associated 1 (CCA1), Late Elongated Hypocotyl (LHY), Zeitlure (ZTL), Timing of Cab Expression 1 (TOC1), Gigantea (GI) and F Lavin Binding Kelch-repeat f-box 1 (FKF1)<sup>7</sup> and other upstream genes. The *CO* gene control Flower Locus T (FT) gene synthesis, *FT* as *CO* target gene, together constitute the CO/FT regulatory elements, regulate the flowering time of plants<sup>8</sup>. The results showed that the up-regulation of *CO* gene promoted the expression of FT in the long daylight condition and promoted the flowering<sup>9</sup> and the promoting effect was not obvious under the condition of short daylight<sup>10-12</sup>.

At present, the homologous genes of *CO* were obtained in the *Nicotiana tabacum*<sup>13</sup>, *Agapanthus africanus Hoffmagg*<sup>14</sup>, *Fragariaananassa duch.*<sup>15</sup>, *Vitis vinifera*<sup>16</sup>, *Malus pumila* Mil<sup>17</sup>, *Festuca elata*<sup>18</sup>, *Pyruspyrifolia nakai*<sup>19</sup>, *Solanum lycopersicum*<sup>20</sup> and *G. biloba*<sup>21</sup>, but their functions were different. The control

route of the CO homologous gene *Vv CO* from *V. vinifera* is similar to that of *Arabidopsis thaliana*. However, there gulation pathway of *MdCOL1* and *MdCOL2* in *M. pumila* Mil is different from that of *A. thaliana*, which shows organ specificity and plays an important role in growth and development of reproductive organs. Based on the data of *G. biloba* transcriptome, the GbCOL-6 cDNA of *G. biloba* was cloned and the bioinformatics analysis was carried out to predict its function. In addition, the expression of *GbCOL-6* gene in *G. biloba* was analyzed.

## MATERIALS AND METHODS

**Materials:** The materials were taken from 31 years old trees of the *G. biloba* cultivar 'Jiafoshou', in the Ginkgo Science and Technology Garden, Yangtze University. The tree was in good condition and normal growth. From February, 2017, the materials of *G. biloba* were collected and the recovered material was stored after liquid nitrogen in a -80°C refrigerator.

PrimeScript™ 1st Strand cDNA Synthesis Kit, Agarose Gel DNA Purification Kit Ver. 4.0, MiniBEST Plant RNA Extraction kit, Taq DNA polymerase RNase, dNTP, pMD18-T vector were purchased from Takara Company (Dalian, China). In this experiment, both the primers synthesis and DNA sequencing were completed by Shanghai Sangon Biotechnology Company, China.

**Cloning of GbCOL-6:** Total RNA was isolated from frozen leaves of *G. biloba* using MiniBEST Plant RNA Extraction kit. The extracted RNA was reverse transcribed into cDNA using PrimeScript™ 1st Strand cDNA Synthesis Kit. The Specific Primers C1 and C2 were designed based on the *COL-6* unigene sequence of *G. biloba* transcriptome data (Table 1). The *GbCOL-6* gene was amplified by reverse transcription polymerase chain reaction (RT-PCR) using cDNA obtained by RNA reverse transcription. The PCR reaction system was 1 µL upstream of the primer (10 µmol L<sup>-1</sup>), Exuq polymerase (5 U µL<sup>-1</sup>) 0.5 µL, 10×PCR buffer 2.5 µL, PCR Enhancer 4 µL, dNTP (2.5 mmol L<sup>-1</sup>) 2 µL, total cDNA (1 µg µL<sup>-1</sup>) 1 µL, ddH<sub>2</sub>O make up 25 µL. The reaction procedure is: Pre-denaturation at 94°C for 3 min, denaturation at 94°C for 30 sec, annealing at 64.5°C for 30 sec, extension at 72°C for 90 sec, a total of 35 cycles, finally at 72°C for 10 min. The amplified products were detected by 1% gel electrophoresis and the PCR products were recovered by Agarose Gel DNA Purification Kit Ver.4.0. The purified product was cloned into the pMD18-T vector, then transformed into *Escherichia coli* DH5α strain. Positive clones were selected and sent to Shanghai Biotechnology Biology Company for sequencing.

Table 1: Primer sequences used in this study

Primers	Sequence (5'-3')
GbCOL-6-up C1	TGAAGAAGATGGAGGGTAGTGAA
GbCOL-6-down C2	TGCGTGAGTTCATTCAAGAGAC
qRT-PCR upstream primer P1	ATGAAGATGTTGTAATGGCCTGGTC
qRT-PCR downstream primer P2	ACTCTTGAAAACCTTCTGTTCCATC
18S upstream primer H1	TTGGTCTCCCGTCTAATGG
18S downstream primer H2	CGAAGCGTCATCCTAAGACAACA

**Bioinformatics analysis:** Using the BLAST tool on the NCBI website to find the sequence with the highest similarity to the nucleotide and deduced amino acid sequences obtained in this experiment. Sequence splicing, protein translation and amino acid sequence multiple comparisons were performed using Vector NTI 11.5 and DNAMAN. The deduced amino acid sequence was analyzed by CLUSTAL X 2 and MEGA6 software. The Neighbor-Joining (NJ) method was used to construct phylogenetic tree and the ExpAsy online server ProtParam predicts the molecular weight and theoretical isoelectric point of the protein.

**Quantitative real-time PCR analysis:** The RNA was extracted from samples including root, stem, leaf, female strobili, male strobili and fruit and cDNA was reverse transcribed. According to the known sequence, real-time quantitative primers P1 and P2 (Table 1) were designed to amplify the *GbCOL-6* gene fragment. Reverse transcribed cDNA was used as a template and 18S was selected as an internal reference gene for qRT-PCR amplification<sup>22</sup>. The upstream and downstream primers of 18S were H1 and H2, respectively (Table 1). Reaction system 25  $\mu$ L, the reaction program with reference TaKaRa real-time quantitative instructions (PrimeScript<sup>TM</sup> RT reagent Kit Perfect Real Time) was carried out. Each sample was set up three times to repeat the ultrapure water as the negative control, 18S to do the internal reference gene. Data processing using the relative quantitative method, concerning  $2^{-\Delta\Delta C_t}$  method for the results of analysis<sup>23</sup>.

## RESULTS

**Analysis of the *GbCOL-6* gene sequence:** Ginkgo cDNA was used as a template for PCR amplification of the *GbCOL-6* gene. The specific primers C1 and C2 (Table 1) were designed and the PCR product was subjected to electrophoresis in 1% agarose gel to obtain a DNA fragment of about 1500 bp (Fig. 1). The product was purified and sent to Shanghai Biotechnology Company for sequencing. The sequencing results showed that the cDNA of *GbCOL-6* was 1594 bp in

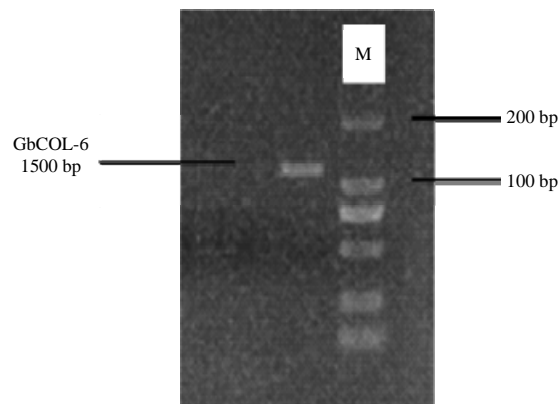


Fig. 1: Electrophoresis of RT-PCR of ORF in *GbCOL-6* gene: M, DL2000 marker

Table 2: Similarity comparison of GbCOL-6-coding protein and other known CO proteins

Species	Accession number	Homology (%)	E value
<i>Populuseuphratica</i>	XP_011016373.1	48	3e-24
<i>Vitis vinifera</i>	CBI27020.3	34	3e-46
<i>Nicotiana tabacum</i>	XP_016463604.1	33	1e-41
<i>Herrania umbratica</i>	XP_021298869.1	33	1e-46
<i>Ricinus communis</i>	EEF30756.1	33	4e-36
<i>Sesamum indicum</i>	XP_020547025.1	33	3e-41
<i>Pyrus x bretschneideri</i>	XP_009356747.1	32	6e-40
<i>Dorcoceras hygrometricum</i>	KZV40687.1	32	4e-40

length. The results of Vector NTI 11.5 showed that *GbCOL-6* contained a 1476 bp ORF encoding 492 amino acids (Fig. 2, GenBank accession number MG251395). The on-line analysis of ExpAsy-ProtParam showed that the theoretical molecular weight and Pi of the *GbCOL-6* protein were 54.65 kDa and 5.08, respectively.

**Analysis of *GbCOL-6* protein:** The amino acid sequence of *GbCOL-6* was compared on NCBI website to obtain protein sequences with high similarity. The homologous alignment results of the deduced amino acid sequence of *GbCOL-6* with BLAST-protein on NCBI site showed that the *GbCOL-6* protein sequence had high homology with the COL protein sequences of other plants and their similarities ranged from 32-48%. The sequence similarities of *GbCOL6* protein and other COL proteins were as follows: *Populus euphratica*, *V. vinifera*, *N. tabacum*, *Herrania umbratica*, *Ricinus communis*, *Sesamum indicum*, *Pyrus x bretschneideri* and *Dorcoceras hygrometricum* were 48, 34, 33, 33, 33, 33, 32 and 32%, respectively (Table 2). In addition, the *GbCOL-6* protein has a conserved B-box structure and a CCT domain of the CO protein family Fig. 3.

```

1      TCGAGAAGCTCAGGGTAGTGAAGATGCAGATGGGTGATGCCGAAACAGCTAAAAATAGTAATAAAAATGACAA
      M E G S E D A D G S M R K T A K N S N K M T
76    TGGCCATAGCGGCAAGAGACTCGTCCATCTGATGCTCCGAGAACAGCTCGCTCGATGGTACTGTGGTCAG
25    M A I G G R R A R P C D V C G E Q L A R W Y C G A
151  ACCAAGCAAACCTATGGGACTCTCCGACGGATCCGTTCACACTGCCAATTCAGTCGGCTGACACACGAAAGAG
50    D Q A N L C D S C D G S V H T A Y S V A C R H E R
226  TCAGACTGACGGCAGCGAGGTTCGATCCGACGAGCATTAAACAAATCATCTTCGGCCCTTCTAACAGCTA
75    V R L T P T Q V R S A G A P L T N H S S A L L T A
301  CTGGAAGCTCCATATGGGCGCCAGCTTCTCGAAACGAAATCGGCCCTTCAGTCGACCGCATCCGGACTACCAGA
100   T R S S I W R H A S R K R S R P S S R P H P D Y Q
378  GACAGACAGCAAAATATAAGCAAGCTCGGTCGACTCAAACTCCATCAAAGTCTTCGACTTCATGTCGACTA
125   R Q T D E I I T K L G R L K P S I K V F D F M S T
451  CTGCAGAAAGTAAATGGGACGACGACTATCCGCTGCACCAAGTCCCAATTTTGGCCGACGAATTCGACTTATGA
150   T A E S X G D D D Y P L H Q V P I L A D E F D F M
528  ATCGGCCGACTGCTACTCTAATCGCTCCATCGTCCGCTCCCTAAATGAGAATTCAGTCGGCCCTATGGGCG
175   N R P T A Y S T L P S S T A P N E N S S A G P M G
601  GCACGGGTAGACCTCATGCGAAGAGCTGCCAGTGGAGTTCTAGAGTCGAAGGCTGTCGATCTGTGTGATTTTG
200   G T G R A S C E E L P V E F L E S K A V D L C D F
878  AAGTGGAGATGAGTTTGGGCTGATGCTTCATTCGCGCGATGCCAATGTGACCGTCGACGACTTGGGAGGCTAG
225   E V E M S L G C S F I G P D A H V T V D D L G G L
751  GCTTGTATAGTTTCAGTAAACCTACGCCGACTTAGATGCTCATGATGCCGTCGGGTTCTTGTATCTCAGAGAC
250   G L Y S F S K P Y A D L D A H D G V G F F D S Q R
828  ACAACTGCGTGTCTCAGACGGATCATTGCCAGGCTCGAGTCCCATGCAATGAAAAGAAATGGGAAATTCAC
275   H N C G C S E T D H C E G S S P M D E K N G E I P
901  TTAGTGGCTTTCGAATCGGTTAAAGTCGAGGAAAGCGAGGACGTGACCATGTGTCGGTAGCAGTGCATCA
300   L S W L C N R V K V E E S E D V D H T F G S S D I
976  AAGTCGAGGAGCAAGAGGATAAGACCTGTGTACATAACTGTGAAAGATGTGGAGGTGAAGATTAAAGCTGG
325   K V E E Q E D I D L C D I T V N E D V E V K I K L
1051 ATTTTGATGATATTTCACAGGAGAGAGAAATATGCTCAGATGGATATGAAGATGTGTTAATGGCT
350   D F D D Y F Q E E E K K I C L R L D Y E D V L I A
1128 GGTCTGATCGAGTTCCTTTGGGCGATAATAAGACTATACAGCCTTTGTGGATGACACCAGCTCCGATGGAA
375   W S D R G S L W A D N K T I Q A F T D D T S S D G
1201 CAGAAGGTTTCAAGAGTATGCAATGTCTCGATTTCAGCTTAGCAATGCAACGGGGGCTCGGGAGATCAAG
400   T E G F Q E Y G I V P D L S L G I A R G A S G D Q
1276 GGGAGCAAGTGCCTGTGATGATGGCGGAGGATCATGAAGGAGGATGGAAGGAGAGAGGCTCGGCTTA
425   G E Q V P V M N G G G D H E G R H G E G R E A R V
1351 TGAGGTACCGTGAGAAAGTCAAGCAGACTTTCTCAAAGAAAGATCAGATACGAAGTTCGCAACTCAACGCAG
450   M R Y R E K R R S R L F S K K I R Y E V R K L N A
1428 AAAAGCGACCGGCATGAAGGTTCGATTGTGAAAAGAAAGCCAGGTCCTCTCTCAGTCAACTCACGCA
475   E K R P R M K G R F V K R T P G L S L E

```

Fig. 2: Nucleotide sequence and deduced amino acid sequence of GbCOL-6. The start codon and termination codon are indicated by the box, the primers are indicated by the underline

**Phylogenetic analysis of GbCOL-6:** In order to understand the molecular evolution of GbCOL-6 and other plant CO proteins, the GbCOL-6 protein sequence was analyzed by CLUSTAL X 2 and MEGA6 software. The GbCOL-6 protein sequence was homologous analysis with the CO protein of other species on GenBank and the phylogenetic tree was constructed by NJ method (The accession numbers are shown in Table 3). The *CONSTANS* gene is a multi-copy gene family in the plant genome, 17 members in *A. thaliana*<sup>24</sup> and 16 members in *Oryza sativa*<sup>25</sup>. As Fig. 4 shown, the phylogenetic tree is divided into three major branches, one for *CONSTANS* LIKE 6-8 and *CONSTANS* LIKE16, one for *CONSTANS* LIKE 1-5 and *CONSTANS* and one for *CONSTANS* LIKE 9-15. This is consistent with the results of the *CO* gene family classification in *A. thaliana*.

Genetic studies of *A. thaliana* CO show that the CO gene family is divided into three major categories<sup>24</sup>. Class I contains CO and COL1-COL5 (Fig. 4, Group1) and Class I can be further

divided into class Ia (CO, COL1 and COL2) and class Ib (COL3-COL5), Class II contains COL6-COL8 and COL16 (Fig. 4, Group 2), Class III contains COL9-COL15 (Fig. 4, Group 3). Its classification is mainly N-terminal B-box structural features. The N-terminus of Group 1 has two B-boxes, the N-terminus of Group 2 has only one B-box, the N-terminus of Group 3 has a normal B-box and a secondary structure-modified B-box zinc finger structure<sup>24</sup>. The GbCOL-6 has a normal B-box structure and a CCT conserved domain and it shows that GbCOL-6 is a Class II member of the CO family.

**Expression pattern of GbCOL-6 in different tissues:** As shown in Fig. 5, GbCOL-6 was expressed in roots, stems, leaves, male strobilis, female strobilis and fruits, but the expression of GbCOL-6 was various in different tissues. It was the highest expression in female strobilis, followed by stems, leaves, fruits, male strobilis and the lowest level of expression in roots.

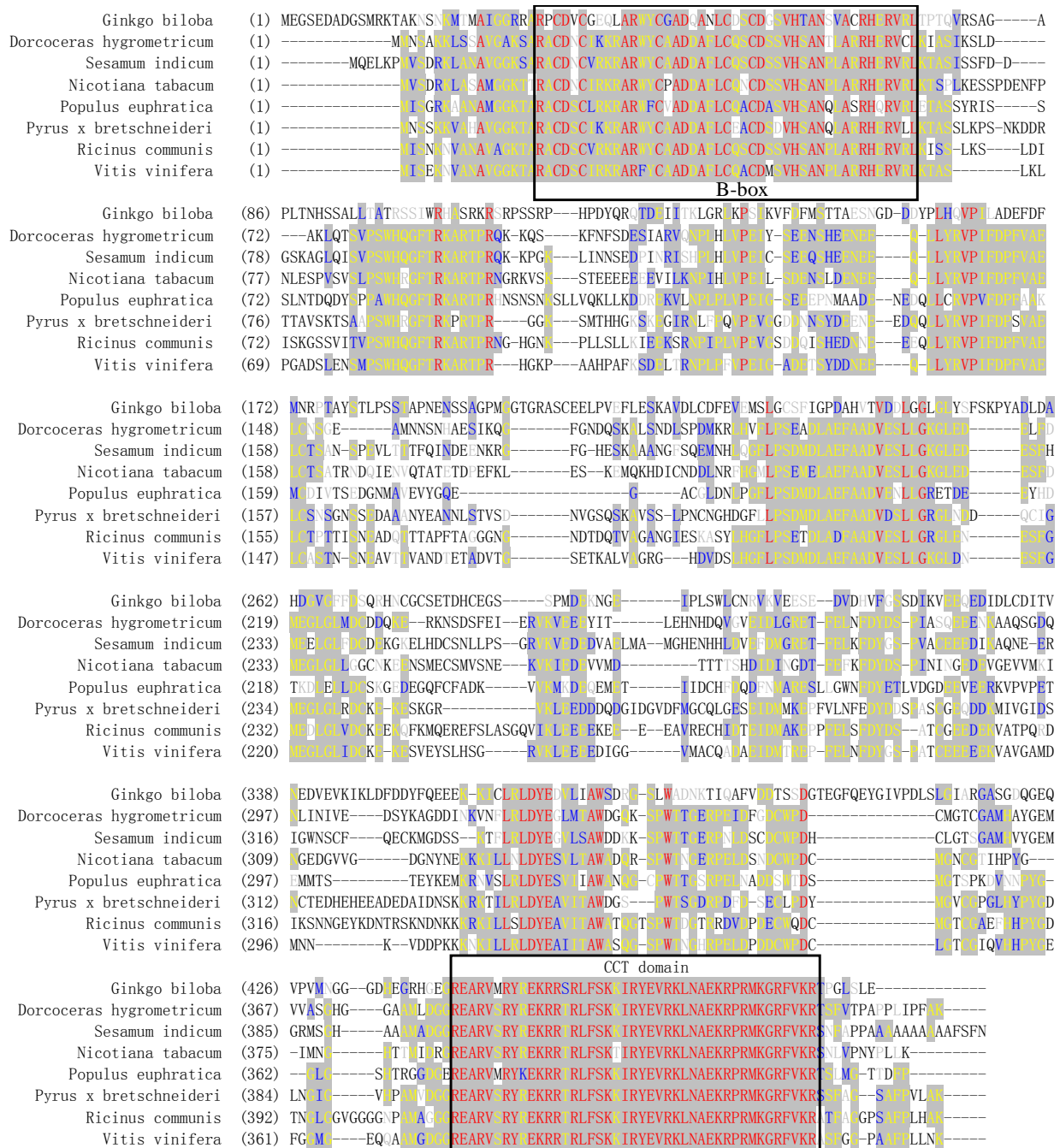


Fig. 3: Sequence multi-alignment of the deduced GbCOL-6 protein with other plant CO proteins. The completely identical amino acids are indicated with red foreground and yellow background, the conservative amino acids are indicated with yellow foreground and cyan background, the block of similar amino acids are indicated with blue foreground and yellow background, the weakly similar amino acids are indicated with gray foreground and white background, the non-similar amino acids are indicated with black foreground and white background

## DISCUSSION

*Ginkgo biloba* is a perennial woody plant with a long juvenile phase. The transformation from vegetative growth to

reproductive growth is regulated by complex gene network. The CO plays a key role in the plant photoperiod pathway and its expression is precisely regulated by the biological clock. In this study, the full-length cDNA of the CO homologous gene

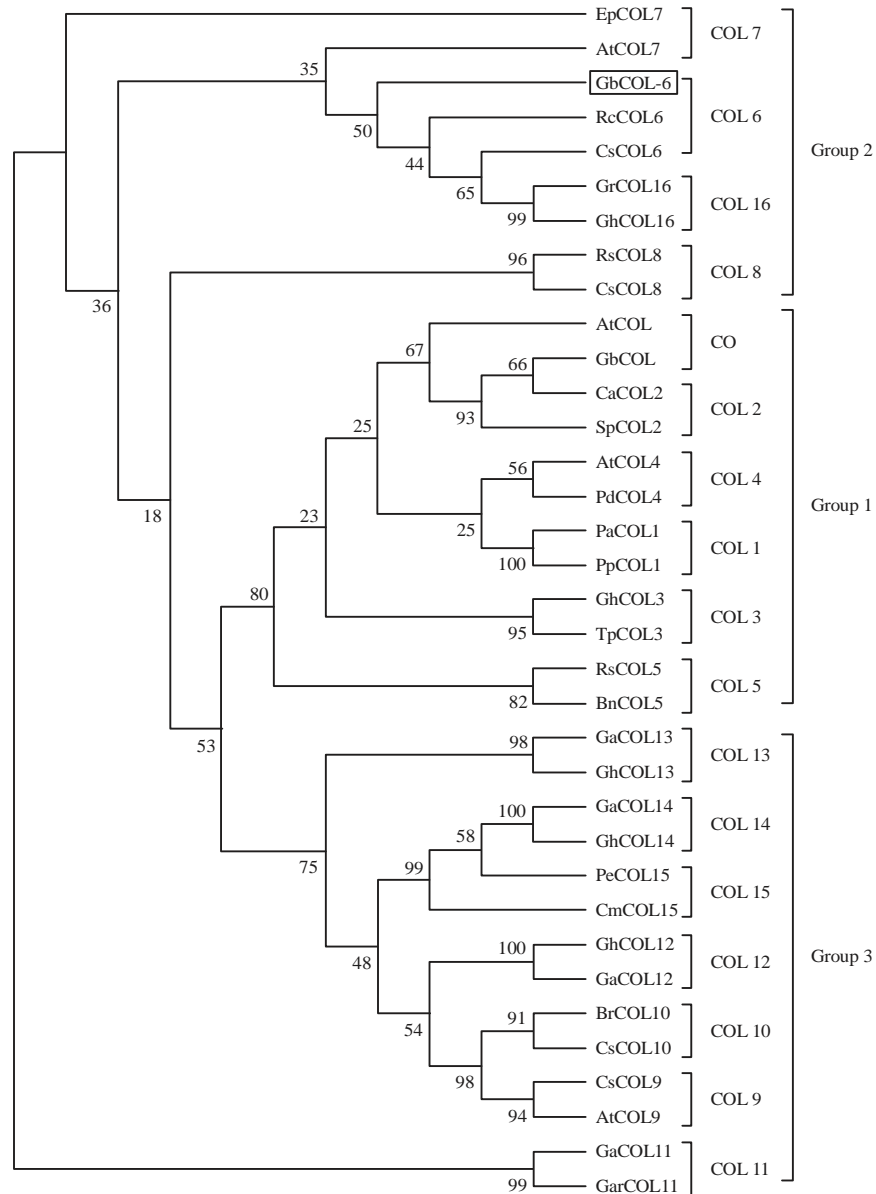


Fig. 4: Phylogenetic tree of CONSTANS from different plants. The numbers of nodes represent the percentage of boot strap value obtained from 1000 sampling

was cloned from *G. biloba* leaves and named as GbCOL-6. The full length cDNA of GbCOL-6 was 1479 bp and encodes 492 amino acids. Protein analysis showed that GbCOL-6 contained one B-box structure and one CCT motif structure (Fig. 3). Tissue expression profile analysis showed that the *GbCOL-6* gene was expressed in all tissues of *G. biloba*. However, the expression level of GbCOL-6 was different and GbCOL-6 was highest expressed in female strobili.

Phytochrome (PhyA, PhyB, PhyC, PhyD, PhyE) and Cryptochrome (Cry1, Cry2) are the photosynthetic receptor for *CO* gene of plant. This photosynthetic receptor mainly used

to accept the light signal. It affects the stability of protein encoded by *CO* gene to control the expression of *CO* gene, so that the abundance of *CO* protein in the day showed a rhythmic change.

The *CO* protein regulates the expression of downstream genes *FT* and *SOC1* to control flowering time, which reflects the biological clock rhythm of *CO* gene expression<sup>26-28</sup>.

The *CO* homologous genes have been cloned in many species, most of which contain conserved B-box structures and CCT domains<sup>29,30</sup>. B-box was a zinc finger protein that could bind to DNA or protein-protein interaction<sup>31-33</sup>, B-box



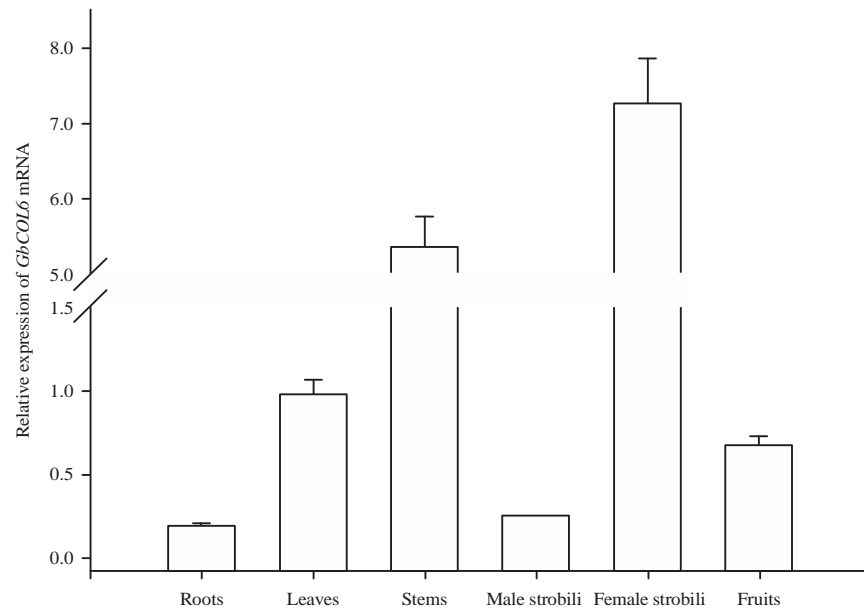


Fig. 5: qRT-PCR relative expression of *GbCOL-6* gene in different tissues

Table 3: Sequence accession numbers of Gene

Gene	Species	Accession number
PaCOL1	<i>Piceaabies</i>	CAK26109.1
PpCOL1	<i>Pinus pinaster</i>	AFV79556.1
CaCOL2	<i>Capsicum annuum</i>	XP_016561206.1
SpCOL2	<i>Solanum pennellii</i>	XP_015065855.1
TpCOL3	<i>Thespesia populneoides</i>	AJR28758.1
GhCOL3	<i>Gossypium hirsutum</i>	AJR28713.1
PdCOL4	<i>Phoenix dactylifera</i>	XP_008789698.1
AtCOL4	<i>Amborella trichopoda</i>	XP_006845571.1
RsCOL5	<i>Raphanus sativus</i>	XP_018452091.1
BnCOL5	<i>Brassica napus</i>	XP_013683363.1
RcCOL6	<i>Ricinus communis</i>	XP_002533412.1
CsCOL6	<i>Camelina sativa</i>	XP_010470801.1
GbCOL-6	<i>Ginkgo biloba</i>	MG251395
AtCOL7	<i>Arabidopsis thaliana</i>	NP_177528.1
GhCOL7	<i>Erycinapusilla</i>	AGI62031.1
CsCOL8	<i>Camelina sativa</i>	XP_010479338.1
RsCOL8	<i>Raphanus sativus</i>	XP_018490515.1
AtCOL9	<i>Arabidopsis thaliana</i>	NP_001118599.1
CsCOL9	<i>Camelina sativa</i>	XP_010486334.1
CsCOL10	<i>Camelina sativa</i>	XP_010493639.1
BrCOL10	<i>Brassica rapa</i>	XP_009151626.1
GaCOL11	<i>Gossypium arboretum</i>	KHG29511.1
GarCOL11	<i>Gossypium arboretum</i>	KHG24248.1
GaCOL12	<i>Gossypium arboretum</i>	KHG03964.1
GhCOL12	<i>Gossypium hirsutum</i>	XP_016666417.1
GaCOL13	<i>Gossypium arboretum</i>	KHG14594.1
GhCOL13	<i>Gossypium hirsutum</i>	XP_016737926.1
GaCOL14	<i>Gossypium arboretum</i>	KHG22096.1
GhCOL14	<i>Gossypium hirsutum</i>	XP_016727140.1
PeCOL15	<i>Populus euphratica</i>	XP_011020834.1
CmCOL15	<i>Cucumis melo</i>	XP_008454034.1
GrCOL16	<i>Gossypium raimondii</i>	XP_012467882.1
GhCOL16	<i>Gossypium hirsutum</i>	XP_016706545.1
AtCOL	<i>Arabidopsis thaliana</i>	CAA71587.1
StCOL	<i>Solanum tuberosum</i>	NP_001274795.1

domain has two conserved amino acid residues (Cys and His) that bind to the zinc atom to tightly fold the entire zinc finger protein, indicating that GbCOL-6 has similar characteristics as the class II CO gene of the CO family of *A. thaliana* and has a normal B-box structure and a CCT conserved domain, which belongs to the CO homologous gene (Fig. 3). Therefore, it is presumed that the main function of the B-box structure of GbCOL-6 is to maintain the stability of the protein.

The number of members of the CO gene family is also divers in different plants. There are 17 members of *A. thaliana*<sup>24</sup>, 16 members of *O. sativa*<sup>25</sup> and 14 inferred CO members in *V. vinifera*<sup>16,25,34</sup>. According to Yan *et al.*<sup>21</sup>, the CO gene also contains multiple members in *G. biloba*. The CO genes have been cloned from *A. thaliana*<sup>33</sup>, *O. sativa*<sup>34</sup> and *Triticum aestivum*<sup>35</sup> and they have been found to have similar photoperiod regulation and transcription patterns, but the function of plants is different. Therefore, the structure, regulation and function of CO genes in each plant need to be further studied.

This study found that GbCOL-6 isolated from *G. biloba* was specifically expressed in the female strobilis of *G. biloba*. However, GbCOL-6 rarely was expressed in roots of *G. biloba*. In *V. vinifera*, the expression of *VvCO* and *VvCOL1* are also the lowest in roots<sup>16</sup>. According to Ye *et al.*<sup>20</sup> in the study of *S. lycopersicum* *CONSTANS* homologous gene expression in different tissues and organs found that *SLCOL1* in roots, stems and leaves are also expressed, but the highest expression of leaves of *S. lycopersicum*. The reason may be due to the ecological habits and developmental status of different plants.



In this study, the cloning and expression analysis of GbCOL-6 will contribute to the development of flowering gene networks of *G. biloba*. Nevertheless, the flowering regulatory mechanism of GbCOL-6 only stays at the theoretical stage and lacks specific verification. In the following experiment, the specific role of *GbCOL-6* gene can be further verified by transgenic technology.

### CONCLUSION

In conclusion, a CO family gene (*GbCOL-6*) in the photoperiod pathway from *G. biloba* was cloned for the first time. The full length cDNA of GbCOL-6 was 1479 bp and encodes 492 amino acids. The GbCOL-6 protein contains two conserved domains (B-box and CCT domain) and the amino acid sequence showed high similarity to other plant CO proteins. Phylogenetic analysis classified GbCOL-6 as Class II of the CO gene family. The GbCOL-6 is specifically expressed in female strobilis. The flowering regulatory mechanism of GbCOL-6 was elucidated. This finding can provide genetic resources for shortening the juvenile phase of *G. biloba* and other woody plants and it can also lay a theoretical foundation for further study on the molecular mechanism of the flowering pathway of *G. biloba*.

### SIGNIFICANCE STATEMENT

This study discovers a CO family gene (*GbCOL-6*) from *G. biloba* that can be beneficial for shortening the juvenile phase of woody plants. The GbCOL-6 is a key flowering gene in the photoperiod pathway of plants and can be used to regulate the flowering time of *G. biloba* to obtain terpenoids in fruits. This study will help the researcher to uncover the critical areas of key flowering genes in the flowering pathway from *G. biloba* that many researchers were not able to explore. The study may be helpful the extraction of natural terpenoids in *G. biloba*.

### ACKNOWLEDGMENT

This study was supported by National Natural Science Foundation of China (No. 31670608).

### REFERENCES

1. Chen, L., F. Xu, R. Cai and S.Y. Cheng, 2008. Research advances in regulations of flowering genes and genetic improvement of the juvenile phase of *Ginkgo biloba*. *J. Henan Agric. Sci.*, 37: 13-16.
2. Cheng, S.Y., F. Xu and Y. Wang, 2009. Advances in the study of flavonoids in *Ginkgo biloba* leaves. *J. Med. Plants Res.*, 3: 1248-1252.
3. Smith, J.V. and Y. Luo, 2004. Studies on molecular mechanisms of *Ginkgo biloba* extract. *Applied Microbiol. Biotechnol.*, 64: 465-472.
4. Jia, L., C.D. Li and B.Z. Bai, 2004. Research progress on cultivation and propagation pattern of *Ginkgo biloba*. *J. Agric. Technol.*, 6: 74-76.
5. Jack, T., 2004. Molecular and genetic mechanisms of floral control. *Plant Cell*, 16: S1-S17.
6. Jung, C. and A.E. Muller, 2009. Flowering time control and applications in plant breeding. *Trends Plant Sci.*, 14: 563-573.
7. Jackson, S.D., 2009. Plant responses to photoperiod. *New Phytol.*, 181: 517-531.
8. Chen, F.L., Y.F. Fu and C.T. Lin, 2009. Research progress on CO/FT regulon and its role in adjusting plant flowering. *J. Agric. Sci. Technol.*, 11: 17-22.
9. Wu, L.C., L.L. Chang, X. Chen and L. Ku, 2010. Expression and regulation of plant gene photoperiod CO reaction. *Chin. Agri. Sci. Bull.*, 26: 116-121.
10. Samach, A., H. Onouchi, S.E. Gold, G.S. Ditta, Z. Schwarz-Sommer, M.F. Yanofsky and G. Coupland, 2000. Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. *Science*, 288: 1613-1616.
11. Suarez-Lopez, P., K. Wheatley, F. Robson, H. Onouchi, F. Valverde and G. Coupland, 2001. *CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis*. *Nature*, 410: 1116-1120.
12. Yanovsky, M.J. and S.A. Kay, 2002. Molecular basis of seasonal time measurement in *Arabidopsis*. *Nature*, 419: 308-312.
13. Lu, Y., Y. Liu, Y. Sun, J. Mu, M. Ren, X. Zhang and Z. Wang, 2013. Molecular cloning and analysis of a *CONSTANS* homolog from *Nicotiana tabacum*. *J. Anim. Plant Sci.*, 23: 1360-1365.
14. Shi, Y.B., D. Zhang, X.H. Shen, L. Wang and L.H. Zhuo, 2014. Cloning and expression analysis of *CONSTANS* homologous gene in *Agapanthus africanus* Hoffm. *J. Beijing Forest. Univ.*, 36: 113-120.
15. Zhao, Z.Q., L.J. An, F. Li and T.H. Li, 2010. Cloning and expression analysis of CO gene from *Fragaria ananassa* Duch. *J. China Agric. Univ.*, 15: 19-24.
16. Almada, R., N. Cabrera, J.A. Casaretto, S. Ruiz-Lara and E.G. Villanueva, 2009. *VvCO* and *VvCOL1*, two *CONSTANS* homologous genes, are regulated during flower induction and dormancy in grapevine buds. *Plant Cell Rep.*, 28: 1193-1203.
17. Jeong, D.H., S.K. Sung and G. An, 1999. Molecular cloning and characterization of constans-like cDNA clones of the fuji apple. *J. Plant Biol.*, 42: 23-31.
18. Wang, X.L., X.L. Liu, Y.C. Yang, W. Chen, W.C. Li and J.H. Wu, 2010. Cloning and analysis of transient regulatory gene *CONSTANS* of *Festuca elata*. *Mol. Plant Breed.*, 8: 45-52.

19. Zhang, J.X., D.L. Li, R. Wang, C.L. Liu, Y.B. Yuan and C.H. Ma, 2014. CONSTANS gene cloning from *Pyrus pyrifolia* Nakai and its prokaryotic expression. *J. Plant Genet. Resour.*, 15: 824-830.
20. Ye, X.L., Z.Y. Liu, X.X. Wang, Y.C. Du and Y.M. Guo, 2008. Molecular cloning and expression analysis of a CONSTANS-like gene SLCO1 from *Solanum lycopersicum*. *Acta Hortic. Sin.*, 35: 1607-1612.
21. Yan, J., D. Mao, X. Liu, L. Wang and F. Xu *et al.*, 2017. Isolation and functional characterization of a circadian-regulated *CONSTANS* homolog (*GbCO*) from *Ginkgo biloba*. *Plant Cell Rep.*, 36: 1387-1399.
22. Bustin, S.A., 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.*, 25: 169-193.
23. Schmittgen, T.D. and K.J. Livak, 2008. Analyzing real-time PCR data by the comparative  $C_T$  method. *Nat. Protoc.*, 3: 1101-1108.
24. Robson, F., M.M.R. Costa, S.R. Hepworth, I. Vizir, P.H. Reeves, J. Putterill and G. Coupland, 2001. Functional importance of conserved domains in the flowering time gene *CONSTANS* demonstrated by analysis of mutant alleles and transgenic plants. *Plant J.*, 28: 619-631.
25. Griffiths, S., R.P. Dunford, G. Coupland and D.A. Laurie, 2003. The evolution of *CONSTANS-like* gene families in barley, rice and arabidopsis. *Plant Physiol.*, 131: 1855-1867.
26. Mockler, T., H. Yang, X.H. Yu, D. Parikh, Y.C. Cheng, S. Dolan and C. Lin, 2003. Regulation of photoperiodic flowering by *Arabidopsis* photoreceptors. *Proc. Natl. Acad. Sci. USA.*, 100: 2140-2145.
27. Valverde, F., A. Mouradov, W. Soppe, D. Ravenscroft, A. Samach and G. Coupland, 2004. Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science*, 303: 1003-1006.
28. Klejnot, J. and C. Lin, 2004. A CONSTANS experience brought to light. *Science*, 303: 965-966.
29. Putterill, J., F. Robson, K. Lee, R. Simon and G. Coupland, 1995. The CONSTANS gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell*, 80: 847-857.
30. Yano, M., Y. Katayose, M. Ashikari, U. Yamanouchi and L. Monna *et al.*, 2000. *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell*, 12: 2473-2483.
31. Borden, K.L.B., 1998. RING fingers and B-boxes: Zinc-binding protein-protein interaction domains. *Biochem. Cell Biol.*, 76: 351-358.
32. Torok, M. and L.D. Etkin, 2001. Two B or not two B? Overview of the rapidly expanding B-box family of proteins. *Differentiation*, 67: 63-71.
33. Khanna, R., B. Kronmiller, D.R. Maszle, G. Coupland, M. Holm, T. Mizuno and S.H. Wu, 2009. The *Arabidopsis* B-box zinc finger family. *Plant Cell*, 21: 3416-3420.
34. Ben Naim, O., R. Eshed, A. Parnis, P. Teper Bamnolker and A. Shalit *et al.*, 2006. The CCAAT binding factor can mediate interactions between CONSTANS like proteins and DNA. *Plant J.*, 46: 462-476.
35. Ran, C., H. Shao, J. Yu, C. Kou, Y. Li, L.Q. Li and X.J. Li, 2014. Cloning and analysis of *CO-like* gene *TaCO9* from wheat (*Triticum aestivum* L.). *J. Tritic. Crop*, 34: 1319-1326.