**CAPSella rubella** TGA4, A BZIP TRANSCRIPTION FACTOR, CAUSES DELAYED FLOWERING IN ARABIDOPSIS THALIANA

Maofu Li1,2,3, Hua Wang1,2,3, Yuan Yang1,2,3 and Wanmei Jin1,2,3,*

1 Institute of Forestry and Pomology, Beijing Academy of Agriculture and Forestry Science, Beijing, 100093, China
2 Beijing Engineering Research Center for Deciduous Fruit Trees, Beijing, 100093, China
3 Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (North China), Ministry of Agriculture, Beijing, 100093, China

*Corresponding author: jwm0809@163.com

**Received:** April 14, 2015; **Revised:** July 14, 2015; **Accepted:** July 16, 2015; **Published online:** December 22, 2015

**Abstract:** Flowering time is usually regulated by many environmental factors and endogenous signals. TGA family members are bZIP transcription factors that bind to the octopine synthase element, which has been closely linked to defense/stress responses. Most TGA factors interact with non-expressor of PR1 (NPR1) and plant defense responses are strengthened by this interaction. TGA1 and TGA4 factors bind to NPR1 only in salicylic acid (SA)-induced leaves, suggesting that TGA4 has another function during plant development. Here, we isolated a bZIP transcription factor gene, TGA4, from *Capsella rubella*. TGA4 transcripts were detected in most tissues, with high expression in leaves, low expression in stems and flowering buds, and undetectable in siliques. *CruTGA4* was overexpressed in *Arabidopsis thaliana* wild type Col-0 plants. Flowering time and total leaf number in the transgenic plants showed that overexpression of *CruTGA4* could delay flowering in *A. thaliana*. Our findings suggest that TGA4 may act as flowering regulator that controls plant flowering.

**Key words:** A bZIP transcript factor; TGA4 gene; delayed flowering

**Abbreviations:** bZIP – basic leucine zipper; NPR1 – non-expressor of PR1; qRT-PCR – quantitative real-time polymerase chain reaction; FMI – floral meristem identity; SA – salicylic acid; EMSA – electrophoretic mobility shift assays; CO – constants; FT – flowering locus T; ocs – octopine synthase; GST – glutathione S-transferase; CaMV – cauliflower mosaic virus; MS – Murashige and Skoog; ORF – open reading frame; bp – base pair; NFT – neighbor-joining; LD – long day; SOC1 – suppressor of overexpression of CO 1; WT – wild type

**INTRODUCTION**

Floral initiation during the transition from vegetative to reproductive growth is one of the key developmental processes in flowering plants. Many studies have indicated that the integration of multiple environmental factors and endogenous signals regulates this transition. Genetic analysis of the long-day annual plant *Arabidopsis thaliana* has identified five major flowering time regulatory pathways (photoperiod, temperature and vernalization, gibberellin, autonomous and aging pathways) that control floral initiation (Srikanth and Schmid, 2011). More and more floral-pathway integrator genes downstream of the floral pathways are being isolated (Blazquez and Weigel 2000; Hepworth et al., 2002; Kardailsky et al., 1999; Kobayashi et al., 1999; Lee et al., 2000; Moon et al., 2003; Nilsson et al., 1998; Samach et al., 2000). These floral pathway integrator genes cause the floral transition by activating a set of genes known as floral meristem identity (FMI) genes to promote floral development (Blazquez and Weigel 2000; Hepworth et al., 2002; Kardailsky et al., 1999; Lee et al., 2000; Moon et al., 2003; Samach et al., 2000). The photoperiod pathway channels inputs from light, day length, and the circadian clock to promote the floral transition. The B-box transcription factor
CONSTANS (CO) is a central regulator of this pathway and CO can trigger the production of the mobile florigen hormone FLOWERING LOCUS T (FT) that induces flower differentiation (Valverde 2011).

TGA family members, a group of bZIP transcription factors, can bind to the octopine synthase (ocs) element, which is a plant promoter sequence believed to be involved in defense/stress responses (Foley and Singh 2004). Ocs elements were found to regulate the expression of a number of plant pathogen-responsive genes. In Arabidopsis, ocs element-like sequences are important for the expression of specific glutathione S-transferase (GST) and pathogenesis-related genes (Chen and Singh 1999, Lebel et al., 1998). Most TGA factors interact with NPR1 in Arabidopsis (Despres et al., 2003, Kim et al., 2008, Zhang et al., 1999). However, the TGA1 and TGA4 factors do not interact with NPR1 in Arabidopsis, only binding to NPR1 in SA-induced leaves (Despres et al., 2003); this suggests that TGA4 factors have another function. TGA4 can bind to the promoter of the FT gene according to electrophoretic mobility shift assays (EMSA) and can form a protein complex by interacting with CO in the yeast two hybrid system (Song et al., 2008). This suggests that TGA4 may have some function in controlling plant flowering, but there is no evidence of such activity in vivo.

Capsella rubella, an Arabidopsis relative, originated through an extreme population bottleneck. This was very important in understanding the common mechanism of adaptation in Capsella rubella compared to the model plant Arabidopsis thaliana. To obtain insight into the possible function of TGA4 in plant development, we isolated a TGA4 gene from Capsella rubella. An overexpression vector, fused to the cauliflower mosaic virus (CaMV) 35S promoter, was constructed and used to transform into Arabidopsis wild-type Columbia-0 (Col-0) by the floral dip method. Flowering time and total leaf number showed that overexpression of CruTGA4 could delay flowering in Arabidopsis. Our findings suggest that TGA4 acts as a flowering regulator.

MATERIALS AND METHODS

Plant growth conditions

Arabidopsis thaliana seeds of ecotype Col-0 wild type and Capsella rubella were surface-sterilized with 75% alcohol for 2 min and washed at least five times with sterile water, dispersed on 1/2 Murashige and Skoog medium supplemented with 0.8% agar and 2% sucrose, stratified at 4°C for 3 days in the dark, and then germinated at 21°C. The plants were grown in plastic pots containing a mixture of substrate and vermiculite (3:1). Controlled environmental conditions were provided in growth chambers (21°C and 50% relative humidity). The plants were illuminated with fluorescent lights (100-150 μmol m⁻² s⁻¹). Long day (LD) conditions consisted of 16 h of light followed by 8 h of darkness. The plants were fertilized with mineral nutrient solution (Estelle and Somerville, 1987).

Genomic DNA extraction, PCR amplification and sequencing

Total genomic DNA was extracted from 250 mg of fresh leaf with an EZ Spin Column Genomic DNA Isolation Kit (Biomega Inc., Foster City, CA, USA). Pairs of primers (Supplemental Table 1) were synthesized by Shanghai Sangon (Shanghai, China). The PCR reaction mix contained 0.5 μM of each primer, 25 ng template including genomic DNA and cDNA (see below), 2.5 mM of each deoxynucleotide triphosphate, 2.5 units of KOD Hot Start polymerase (Toyobo Life Science, Osaka, Japan), 1×PCR buffer and 2.5 mM MgSO₄. The PCR program included an initial step of 94°C for 3 min, 35 cycles of 94°C for 15 s, 55°C for 30 s, 68°C for 1-2 min, and a final extension at 68°C for 8 min. The PCR products were separated on 1% agarose gels, stained with ethidium bromide (EtBr), and visualized with a gel-imaging system. The PCR fragments were purified using an EZ-10 Spin Column DNA gel extraction kit (Bio Basic Inc., Markham, Ontario, Canada). PCR products were ligated into the pEASY-T1 vector (TransGen Biotech, China), cloned into the Escherichia coli strain DH5α and the positive clones were sequenced.
**Sequence alignment and phylogenetic analyses**

Full-length amino acid sequence alignments of TGA4 proteins were performed using ClustalW (Chenna et al., 2003). The genetically mobile structural domains of TGA proteins were determined and annotated using SMART (http://smart.embl-heidelberg.de/). Phylogenetic analysis of the TGA4 genes was based on deduced amino acid sequences. A phylogenetic tree was generated with MEGA version 5.0 using the neighbor-joining (NJ) method with 1000 bootstrap replicates (Tamura et al., 2011).

**Expression pattern analysis by qRT-PCR**

To investigate the spatial and temporal expression patterns of TGA4, qRT-PCR experiments were performed on various organs. Total RNA was extracted from each tissue and first-strand cDNA was synthesized with a RevertAid First Strand cDNA synthesis kit (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions. We performed qRT-PCR with SYBR® Premix Ex Taq™ II (Bio-Rad, Hercules, CA, USA) on a Bio-Rad 480 Real-Time PCR Detection System. Primers used for qRT-PCR are shown in Table S1. Two independent biological replicates and three technical replicates for each sample were performed by qRT-PCR. Data were analyzed using the 2^-ΔΔCT method as outlined by Livak and Schmittgen (2001). The expression levels of specific genes were normalized to tubulin.

**Vector construction**

To confirm the in vivo function of the TGA4 gene, we created overexpression constructs by PCR cloning. A genomic DNA fragment from C. rubella containing the entire coding region, the upstream intergenic region and the downstream intergenic region was amplified by PCR using KOD Hot Start polymerase with gene-specific primer sets (Table S1). The PCR condition was the same as for cloning the TGA4 gene. The PCR products were purified with a High Pure PCR purification kit (Bio Basic Inc., Markham, Ontario, Canada). The PCR products and the vector pCAMBIA1300 were digested with Kpn I and Spe I endonucleases. The digested PCR and vector products were ligated with T4 DNA ligase enzyme (Takara Biotechnology, Dalian, China), the positive clones were sequenced and the right clone was used in transgenic tests.

**Transgenic tests and phenotypic data**

The resulting binary vector was introduced into Agrobacterium strain GV3101 and used to transform the Arabidopsis thaliana Col-0 wild-type by the floral dip method (Clough and Bent 1998). Transformed seeds were selected on half-strength MS plates containing 50 mg/L hygromycin. Transgenic plants were grown in plastic pots containing a mixture of substrate and vermiculite (3:1) and identified by PCR using the primer of the hygromycin gene (Table S1).

The total leaf number of transgenic plants and the Col-0 wild-type were scored as the number of leaves in the rosette (excluding cotyledons) plus the number of leaves in the inflorescence at the time of opening of the first flower; flowering time was scored by counting the number of days from sowing to the opening of the first flower. Statistical data were analyzed by SPSS 11.5 (SPSS Inc., Chicago, IL, USA) and Sigmaplot 10.0 (Systat Software Inc., San Jose, CA, USA).

**RESULTS**

**Analysis of TGA4 in Capsella**

TGA4 was amplified from leaf cDNA and the PCR product was cloned and sequenced in pEASY-T1. The full-length sequence of TGA4 was 1095 bp and contained an open reading frame (ORF) of 364 amino acid residues (Fig. S1A). To study the structure of the TGA4 gene, we cloned the genomic sequence of TGA4 from gDNA with the same primers that were used for cDNA cloning (Fig. S1B and Table S1). We compared the gDNA and cDNA sequences of CruTGA4 and another six TGA genes in Capsella. The results revealed that these genes comprised 6-9 exons and 5-8 introns. Exons of 60, 78 and 272 bp in length were much conserved in these genes. Exons of 60 bp were presented in all these genes, exons of 78 bp in all but CruTGA6.
and exons of 272 bp in all except for CruTGA7. However, the introns differed in length. The comparison of exon-intron structures of TGA4 and the other six TGA genes show that these genes were a conserved gene family (Fig. 1A).

**Phylogenetic analysis**

Sequence SMART (http://smart.embl-heidelberg.de/) and plant transcription factor database (http://planttfdb.cbi.pku.edu.cn/) searches revealed that CruTGA4 (Caruv10001361m.g) is located on C. rubella chromosome 6, that it is 1095 bp long, and that it encodes a 364 amino acid protein with a predicted molecular mass of 41.6 kDa. Amino acids 76-136 contain a typical bZIP basic region domain, which may serve as an active domain. All TGA proteins are characterized by an extraordinarily long ZIP domain consisting of multi-heptad repeats (Fig. 1B and Fig. 2A).

To analyze the phylogenetic relationships between members of the TGA4 gene family, we performed a phylogenetic analysis of CruTGA4 and TGA4 genes from other angiosperms. An NJ phylogenetic tree based on deduced amino acid sequences was generated. In the tree, CruTGA4 formed a cluster with CgrTGA4, AthTGA4, BstTGA4, AarTGA4, AlyTGA4, BnaTGA4 and EsaTGA4; together they formed a monophyletic clade in Brassicaceae (Fig. 2B). Therefore, these results further indicate that TGA4 proteins, as a type of bZIP transcription factor, are conserved not only in Brassicaceae but also in other angiosperms.

**Expression patterns of TGA4**

To understand the functions of TGA4 gene in C. rubella development, we studied the organ-specific expression of TGA4 using total RNAs derived from different plant tissues and organs by qRT-PCR analysis. Results showed that TGA4 transcripts were present in most tissues tested, with high expression in leaves, low expression in stems and flowering buds, and undetectable expression in siliques (Fig. 3). This indicated that TGA4 was expressed during plant development, which is consistent with previous reports (Song et al., 2008).
days on average; in plants overexpressing the gene, the flowering time was about 58 days (Fig. 4A). Statistical analysis showed that the transgenic lines had delayed flowering compared with WT Col-0 plants (Fig. 4B).

DISCUSSION

In this study, we isolated a TGA4 gene from C. rubella. CruTGA4 had a conserved domain typical of bZIP transcription factors; the bZIP motif plays an important role in determining the function in Arabidopsis. The genomic organization of the TGA family is well conserved in Brassicaceae. In the phylogenetic tree, CruTGA4 was most closely related to TGA4 orthologs from Brassicaceae. Our sequence alignment and phylogenetic tree analysis suggest that CruTGA4 should have a similar function to other TGA4 orthologs. It was reported that TGA family transcription factors, which interact with NPR1, are closely linked to defense/stress responses (Zhou et al., 2000, Zhang et al., 1999). However, TGA4 factors do not interact with NPR1, and bind to NPR1 only in SA-induced leaves (Despres et al., 2003, Kim et al., 2008, Zhang et al., 1999). Thus, TGA4 factors might have multiple roles in plant development and may have other functions besides regulating defense/stress and nitrate responses.

Analysis of TGA4 expression patterns showed that TGA4 transcripts were present in most tissues tested, and were highly expressed in leaves (Fig. 3); this result is in accordance with Song et al. (2008). CO has been demonstrated to promote flowering by activating the expression of FT and SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1) (Kardailsky et al., 1999, Kobayashi et al., 1999, Samach et al., 2000). The diurnal expression patterns of TGA4 were similar to those of CO during the night in both LD and short day conditions (Song et al., 2008). So we infer that TGA4 may interact with CO and bind to the FT promoter, and the

![Fig. 3. Expression analysis of CruTGA4 in different organs of C. rubella. FL – flowering bud; LF – leaf; ST – stem; SL – silique. The relative expression values for CruTGA4 are relative to tubulin and represent means of two biological replicates. SD – standard deviation.](image)

![Fig. 4. Overexpression of CruTGA4 delays flowering in Arabidopsis under long day conditions. A – Representative images of flowering wild type and transgenic overexpression plants grown under LD conditions (16/8 h, light/dark). Photography at 52 days after sowing. B – Flowering time and number of rosette leaves. *** significant at P < 0.001 compared with wild-type levels.](image)
TGA4 mRNA abundance mirrors the diurnal oscillations of CO. Mutations in FT delay flowering in wild-type and in 35S:CO plants (Koornneef et al., 1991; Onouchi et al., 2000). FT mRNA accumulates in the night under LD conditions (Suarez-Lopez et al., 2001), so TGA4 may regulate the expression of FT through its interaction with CO during the dark period, and thereby control plant flowering. Our transgenic plants showed that overexpression of the CruTGA4 gene delayed flowering in A. thaliana in long day conditions. This finding suggests that overexpression of CruTGA4 may repress the expression of FT through its interaction with CO, and then delay flowering in A. thaliana. Although the flowering time between wild-type and transgenic plants demonstrated that CruTGA4 is involved in the regulation of plant flowering, further genetic studies using mutants of co, ft and tga4 genes, are needed for functional complementation experiments and analysis of the relationships among TGA4, CO and FT using transgenic lines.

Acknowledgments: This work was supported by the National Science Foundation for Young Scholars of China (31201606), the Important Crops Molecular Breeding Technology Innovation (KJCX20140202), Special Fund for the Construction of Scientific and Technological Innovation Capability (No. KJCX20140110), and Beijing Municipal Science and Technology Commission (No. D13110000113001) projects.

Authors’ contributions: Maofu Li and Wanmei Jin conceived and designed the experiments; Maofu Li and Hua Wang performed the experiments; Maofu Li and Yuan Yang analyzed the data, Maofu Li and Wanmei Jin wrote the paper.

Conflict of interest disclosure: The authors declare no conflict of interest.

REFERENCES


**Supplementary Material (available online)**

**Supplementary Fig. S1.** The cloned CruTGA4 gene and overexpression vector construct. The CruTGA4 gene cloned from cDNA (A) and gDNA (B). (C) A diagram of the CruTGA4 overexpression vector. (D) Confirmation of the overexpression vector construct by double enzyme digestion, 1-9 indicate the different positive clone, CK – empty vector. **Available at:** http://serbiosoc.org.rs/sup/SupplMat/FIGS1.tif

**Supplementary Fig. S2.** Confirmation of transgenic plants by PCR using the hygromycin gene primer. The numbers 1-15 indicate the 15 transgenic plants, ‘p’ represents the plasmid of the overexpression vector construct, ‘WT’ indicates the non-transgenic plant. **Available at:** http://serbiosoc.org.rs/sup/SupplMat/FIGS2.tif

**Supplementary Table S1.** Primers used in this study. **Available at:** http://serbiosoc.org.rs/sup/SupplMat/SupplTable1.doc

**Supplementary Table S2.** Sequence information of proteins mentioned in Fig. 2B. **Available at:** http://serbiosoc.org.rs/sup/SupplMat/SupplTable2.doc