Plant Science 199–200 (2013) 71–78

Contents lists available at SciVerse ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

The sequence variation responsible for the functional difference between the CONSTANS protein, and the CONSTANS-like (COL) 1 and COL2 proteins, resides mostly in the region encoded by their first exons

Soon-Kap Kim, Hyo-Young Park, Yun Hee Jang, Jeong Hwan Lee, Jeong-Kook Kim*

Plant Signaling Network Research Center, Division of Life Sciences, Korea University, Seoul 136-701, Republic of Korea

ARTICLE INFO

Article history:
Received 30 April 2012
Received in revised form
28 September 2012
Accepted 29 September 2012
Available online 10 November 2012

Keywords:
CO
COL1
COL2
Flowering time
Domain swapping
Evolution

ABSTRACT

Although the protein CONSTANS (CO) and its close relatives CONSTANS-like (COL) 1 and COL2 exhibit high amino acid sequence similarities, only the CO protein regulates floral induction in *Arabidopsis*. To investigate the structural basis for the functional differences between CO, COL1, and COL2 in flowering, we performed domain-swapping between CO, COL1, and COL2, and site-directed mutagenesis on the first exon of *CO*. The results suggest that the lack of flowering promotion activity by COL1 and COL2 is mainly attributed to the differences between CO and the COL1 and COL2 proteins in the amino acid sequence encoded by their first exons.

1. Introduction

Flowering is a very important process that requires proper regulation for successful reproduction. Because various environmental and endogenous cues affect plant growth and development, plants can adjust to continuously changing conditions through phenotypic plasticity and eventually adapt to these changes [1]. Light is one of the major environmental factors influencing the timing of flowering. Molecular genetic analyses in *Arabidopsis thaliana* have revealed that the *Arabidopsis* protein CONSTANS (CO) is the principal regulator of the photoperiod pathway [2]. The *CO* gene encodes a nuclear protein that contains a B-box-type zinc-finger and a CCT (*CO*, *CO*-like, *TOC1*) domain [3]. The photoperiodic regulation of CO protein stability *via* the proteins CONSTITUTIVE PHOTOMORPHOGENESIS 1 (COP1) and SUPPRESSOR OF PHYTOCHROME A (SPA) is a crucial step in this pathway [4,5]. Under long-day (LD) conditions, COP1 interacts with the CCT domain of CO during the night, thereby mediating the degradation of CO. Furthermore, COP1 and SPA cooperate to negatively modulate CO stability under short-day (SD) conditions. The photoperiodic signals mediated by *CO* activate *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*) [6,7] *via FLOWERING LOCUS T* (*FT*) [8,9], thereby promoting flowering [10].

Analysis of the *Arabidopsis* genome revealed the existence of a family of sixteen *CO*-like (*COL*) genes, including *CO* [11]. On the basis of structural variations in the B-box regions, the members of this family are classified into 3 groups, the first group consisting of *CO* and *COL1* to *COL5*, the second group consisting of *COL6* to *COL8*, and the third group consisting of *COL9* to *COL15*. Analysis of various *CO* mutant alleles has shown that the presence of 2 B-boxes and a CCT domain is important for *CO* function in regulating flowering time [11]. Amino acid deletion and substitution in the B-box of the *co-1* and *co-2* alleles, respectively, have been reported to delay flowering time. Amino acid substitution in the CCT domain in 2 *CO* mutant alleles has also been reported to affect *CO* function. Recent studies have shown that the B-boxes and the CCT motif of CO are involved in the recruitment of transcriptional complexes and DNA binding [12–14]. The glutamine-rich region between the B-boxes and the CCT motif has also been identified as a transcriptional activation domain by protoplast transfection assays [14]. Conserved regions, called the M regions (M1 to M4), have been identified between the B-boxes and the CCT motif [15], but their biological functions are not understood yet.

CO function in floral activity does not appear to be conserved in other *COL* genes. *COL3* and *COL9* are known to be floral repressors [16,17]. A mutational lesion in *COL3* and overexpression of *COL9* result in early and late flowering phenotypes, respectively. However, *COL5* acts as a floral activator, similar to *CO*, under SD conditions [18]. Although *COL1* and *COL2*, which are located on

* Corresponding author. Tel.: +82 2 3290 3157; fax: +82 2 927 9028.
E-mail address: jkkim@korea.ac.kr (J.-K. Kim).

0168-9452/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.plantsci.2012.09.019

chromosomes 5 and 3, respectively, are known to have evolved from *CO* through gene duplication [19,20], *35S::COL1* and *35S::COL2* plants do not show altered flowering time [21]. These results suggest that the function of *CO* as a central player in the control of flowering may not have been conserved in the *COL* genes during evolution.

To understand the structural basis of the differences in the functions of CO and the COL1 and COL2 proteins in flowering, we induced the overexpression of various chimeric proteins in wild-type Col-0 plants and a loss-of-function allele, *co-9* mutants [22], in order to identify regions that confer *CO* activity *in vivo*. Collectively, these results suggest that differential activity between CO and the COL1 and COL2 proteins in flowering may result from variations in the amino acid sequences encoded by their first exons.

2. Materials and methods

2.1. Plant materials and growth conditions

Wild-type *Arabidopsis* (Col-0), *co-9* [22], and transgenic *Arabidopsis* plants were grown in Sunshine Mix 5 (Sungro Horticulture, Bellevue, WA, USA) or Murashige and Skoog (MS) medium at 23 °C under LD conditions (16/8-h light/dark) with light supplied at an intensity of 120 μmol m^{-2} s^{-1}.

2.2. Generation of chimeric constructs

For swapping between the exon sequences of *CO* (GenBank ID: X94937), *COL1* (GenBank ID: Y10555), and *COL2* (GenBank ID: NM_111105), primers were designed to contain the sequences of *CO*, *COL1*, and *COL2*, such that the first half (approximately 12 nucleotides) of each oligonucleotide contained the end sequence of an exon of *CO*, *COL1*, or *COL2*, and the second half contained the starting sequence of the adjacent exon of the respective genes. After amplification of the appropriate fragments of *CO*, *COL1*, and *COL2* open reading frames (ORFs) from their cDNA sequences in the first round of polymerase chain reaction (PCR), gel-purified fragments were mixed and used as templates to obtain the entire chimeric gene in the second round of PCR. Subsequently, sequence-verified chimeras were cloned into the pCHF3 vector. Similar procedures were performed for the generation of constructs with mismatched or substituted mutations. Oligonucleotide sequences are listed in Table S1.

2.3. Generation of transgenic plants and measurement of flowering time

Transgenic plants were generated using the floral dip method with minor modifications [23]. After growth on MS medium for 10 days, kanamycin-resistant transgenic seedlings were transferred to the soil. To show the possible range of phenotypes in the T_1 generation, flowering phenotypes of approximately >30 T_1 seedlings were analyzed for each construct. The flowering phenotypes of some transgenic lines that showed a Mendelian inheritance pattern (3:1 ratio) of kanamycin resistance were also analyzed in the T_2 generation. To measure flowering time in the plants, the total number of primary rosette and cauline leaves formed on the primary inflorescence was counted. To compare flowering times between control (Col-0 and *co-9* plants) and transgenic plants in the T_1 or T_2 generations, we used control plants that were transferred from MS medium to soil to score the flowering time.

2.4. Expression analysis

Total RNA was extracted from whole seedlings using the Trizol reagent (Invitrogen, Carlsbad, CA), and first-strand cDNA was synthesized from 1 to 2 μg of total RNA, in accordance with the manufacturer's instructions (Roche Applied Science, Basel, Switzerland). Gene expression levels were determined *via* quantitative reverse transcription (qRT)-PCR method, as described previously [24].

Oligonucleotide sequences are listed in Table S2. All qRT-PCR experiments were performed in 2 or 3 biological replicates with 3 technical triplicates, each with similar results. The results from a biological triplicate are presented. The procedure used for determining the relative abundance of transcripts has been described in detail elsewhere [24].

3. Results

3.1. Comparison of amino acid sequences between the proteins CO, COL1, and COL2 in Arabidopsis

Overexpression of *COL1* and *COL2* in *Arabidopsis* plants did not affect flowering time [21], although the deduced amino acid sequences of these genes show about 82% similarity to *CO*. However, there are a few genes known to rescue *CO* mutants when they were over-expressed. These include a *Chlamydomonas* CO homolog (CrCO) [25], a sugar beet CO homolog (BvCOL1) [26], a perennial ryegrass CO homolog (LpCO) [27], and a Japanese morning glory CO homolog (PnCO) [28]. The *Arabidopsis* COL1 and COL2 proteins show higher amino acid sequence similarity than any one of these proteins, but still could not affect flowering time [21]. The CrCO protein has 37% amino acid sequence similarity (27% identity) to CO and the zinc-finger motif ($CX_2CX_8CX_7CX_2CX_4HX_8H$) of its B-box2 is not intact. The zinc-finger motif of the B-box2 of the LpCO protein is not intact either. However, the zinc-finger motifs of the B boxes of the COL1 and COL2 proteins are still intact. This suggests that there are some unknown, not conserved, but still essential amino acid residues of the CrCO and LpCO proteins that confer on these proteins the ability to promote flowering in *Arabidopsis*.

We compared the amino acid sequences of CO homologs that promote flowering with those that do not, but it was difficult to find an amino acid residue candidate, variations in which might be responsible for the lack of flowering promotion function in the COL1 and COL2 proteins (Fig. S1). This prompted us to perform a domain-swapping experiment between the CO protein and the COL1 and COL2 proteins to approximately estimate the regions responsible for functional difference between them.

3.2. A domain-swapping between CO and COL1 and COL2 proteins

To map the regions of the CO, COL1, and COL2 proteins responsible for their distinct *in vivo* activities in *Arabidopsis*, we constructed chimeric genes in which individual exons were swapped (Fig. 1). Because each gene has 2 exons, we constructed 4 chimeras (SWA1, SWA2, SWA3, and SWA4) in addition to the wild-type versions of *CO*, *COL1*, and *COL2* (Fig. 1A). All chimeric constructs were expressed from the *35S* promoter in Col-0 plants and *co-9* mutants [22]. To show an unbiased view of the range of possible phenotypes, we first examined the phenotypes in the T_1 generation.

A bimodal distribution of the flowering phenotype was observed in *35S::CO* plants in the T_1 generation, as was previously known [12]. A similar bimodal distribution was also observed in *35S::SWA1* plants and *35S::SWA2* plants, but not in *35S::COL1* plants, *35S::COL2* plants, *35S::SWA3* plants, and *35S::SWA4* plants (Fig. 1B). In order to confirm the inheritance of these phenotypes, we first chose single transgene lines through kanamycin selection and then examined their phenotypes in the T_2 generation (Fig. 1C). The early flowering phenotypes of *35S::SWA1* and *35S::SWA2* plants were more