Molecular characterization of *CONSTANS-Like (COL)* genes in banana (*Musa acuminata* L. AAA Group, cv. Grand Nain)

Akhilesh Kumar Chaurasia¹ · Hemant Bhagwan Patil¹ · Abdul Azeez¹ ·
Vadakanthara Ramakrishnan Subramaniam¹ · Bal Krishna¹ ·
Aniruddha Prafullachandra Sane² · Prafullachandra Vishnu Sane¹

Received: 30 November 2015 / Accepted: 31 January 2016 / Published online: 11 February 2016
© Prof. H.S. Srivastava Foundation for Science and Society 2016

Abstract The CONSTANS (CO) family is an important regulator of flowering in photoperiod sensitive plants. But information regarding their role in day neutral plants is limited. We report identification of nine Group I type *CONSTANS-like (COL)* genes of banana and their characterization for their age dependent, diurnal and tissue-specific expression. Our studies show that the Group I genes are conserved in structure to members in other plants. Expression of these genes shows a distinct circadian regulation with a peak during light period. Developmental stage specific expression reveals high level transcript accumulation of two genes, *MaCOL3a* and *MaCOL3b*, well before flowering and until the initiation of flowering. A decrease in their transcript levels after initiation of flowering is followed by an increase in transcription of other members that coincides with the continued development of the inflorescence and fruiting. CO binding *cis*-elements are observed in at least three *FT-like* genes in banana suggesting possible CO-FT interactions that might regulate flowering. Distinct tissue specific expression patterns are observed for different family members in mature leaves, apical inflorescence, bracts, fruit skin and fruit pulp suggesting possible roles other than flowering. This is the first exhaustive study of the *COL* genes belonging to Group I of banana.

Keywords CONSTANS · Flowering · Day neutral · Banana · Diurnal · Photoperiod

Introduction

Light is important for plants not just for photosynthesis but also for several other developmental transitions such as germination, seedling establishment and flowering. Of these, flowering in several plants is influenced by both, light quality (blue, red or far red) as well as light duration (photoperiod). Flowering is induced by the FT protein, which is now widely perceived to be the florigen equivalent. It moves from the leaf where it is produced to the apical meristem through the phloem to initiate flowering (Lu et al. 2012). Several studies (Ballerini and Kramer 2011; Bohlenius et al. 2006; Hayama and Coupland 2004) in the past decade have shown that flowering in photoperiod sensitive plants is governed by the so called CONSTANS/FLOWERING LOCUS T (CO/FT) module or regulon. In *Arabidopsis*, expression of the functional *FT* is induced by CO, a central regulator in the photoperiodic pathway of flowering. The *CONSTANS* gene, first isolated by Putterill et al. (Putterill et al. 1993, 1995) expresses in a rhythmic manner and has the ability to coordinate inputs from both, the light pathway and circadian clock (Andres and Coupland 2012; Song et al. 2013). The peaks and troughs in the expression of *CO* mRNA are controlled by light hours and the circadian clock (Suarez-Lopez et al. 2001). Its accumulation is regulated at the transcriptional as well as post-translational level through a number of proteins. At the transcriptional level, *CO* is regulated by GI (GIGANTIA), CDF1 (CYCKLING DOF FACTOR 1) and FKF1 (FLAVIN BINDING, KELCH REPEAT, F-BOX1) that are under the control of clock and different light receptors (Imaizumi et al. 2005; Sawa et al. 2007). Stabilization of the CO protein occurs

✉ Abdul Azeez
dr.azeez.abdul@jains.com

¹ Plant Molecular Biology Lab, Jain R&D Lab, Agri Park, Jain Hills, Shirsoli Road, Jalgaon 425001, India

² Plant Gene Expression Lab, CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow 226001, India

in light while it is degraded in the dark by the E3 ring finger type ubiquitin ligase, CONSTITUTIVE PHOTOMORPHOGENIC1 (COP1) and SPA (SUPPRESSOR OF PHYA-105) protein family (Jang et al. 2008; Laubinger et al. 2006). Blue light receptors like CRY1 and CRY2 stabilize CO through suppression of COP1-SPA1 while red light destabilizes CO transcription through suppression of its transcriptional activator PFT1 and protein levels through the E3 ligase HOS1 (Lazaro et al. 2012; Valverde et al. 2004). Under long day conditions, the CO protein is degraded in the dark and in the early morning by PHYB (red-light receptor) while it is stabilized in the evening by PHYA (far-red receptor) and CRYPTOCHROME 1 (CRY1) and CRY2 (blue-light receptor). In the afternoon CO protein is stabilized through interaction with FKF1 because of the antagonistic effect of PHYA (Song et al. 2013; Wu et al. 2014). The final result of the regulated expression of CO in long day plants such as *Arabidopsis* is its abundance during afternoon and a peak in the evening of a long day season. The CO protein, stabilized under long day conditions, binds to *FT* promoter through a unique *sequence cis*-element that is present in the promoter of *FT* (Tiwari et al. 2010).

In both monocots and dicots the *COL* (*CONSTANS-Like*) family has many members such as 17 in *Arabidopsis*, 16 in rice, 9 in barley (Griffiths et al. 2003), 10 in sugarbeet (Chia et al. 2008), 11 in Medicago (Wong et al. 2014), 26 in soybean (Wu et al. 2014) and 23 in cotton (Zhang et al. 2015). All of these *COL* genes are characterized by the presence of three domains namely the B-box 1, B-box 2 and a CCT (CO, CO-like, and TOC1) domain (Griffiths et al. 2003; Putterill et al. 1995; Strayer et al. 2000). The *COL* (*CONSTANS-like*) members are categorized into three different groups based on the presence/absence of these three domains (Gangappa and Botto 2014). Group I *COL* genes possess all the three domains, and an additional VP motif (valine-proline motif which is involved in the interaction with COP1). Group II contains B-box1, a less conserved or diverged B-box2 and CCT domain but lacks the VP motif while Group III contains B-box 1 and CCT domain only (Gangappa and Botto 2014; Zhang et al. 2015). In a broader way *COL* proteins are a part of the BBX family transcription factors with diverse functions (Almada et al. 2009; Chia et al. 2008; Crocco and Botto 2013; Datta et al. 2006; Huang et al. 2012; Valverde 2011). The characterization and analysis of Group I CO genes has been carried out in many plant species and these cluster with *Arabidopsis* CO and rice Hd1 (Gangappa and Botto 2014; Griffiths et al. 2003; Huang et al. 2012; Zhang et al. 2015). Banana is an important fruit crop that shows day neutral flowering. Our group has embarked upon studies for the identification and characterization of genes that regulate flowering in banana to explore possibilities of either extending or decreasing the life cycle of this crop. Here we characterize the members of *CO* family of banana and their possible role in banana flowering. Our studies on the diurnal expression

patterns, stage dependent expression and tissue-specific expression of Group I *COL* genes besides their sequence similarities with the known *CO* genes suggest that there are at least two *COL* genes that may play an important role in flowering of this day neutral plant.

Materials and methods

Plant material and sample collection

Banana (*Musa acuminata* L. AAA Group cultivar Grand Nain) tissue culture raised plants were planted in the field and grown in the Jain R&D farms under natural conditions in Jalgaon, Maharashtra, India. For studying the changes associated with different developmental stages the third mature leaf from the top was selected for sample collections. Samples were collected at intervals of 15 days for up to 300 days after planting (DAP). For studying the diurnal patterns 150 day old plants were selected and samples were collected at 4 h intervals for up to 48 h. During the reproductive stage, tissue sampling was performed during different pre and post flowering stages such as mature leaf (ML) at 150 days, bract (BR) at 210 days, apical inflorescence (AI) arrested inside pseudostem after 190 days, flower (FL) including all floral parts at 220 days, young fruit (YF) at 240 days and mature fruit skin/peel (FS) and mature fruit pulp (FP) at 270 days after planting (DAP). All samples were harvested and immediately frozen in liquid nitrogen and stored at -80°C until further use.

Genomic DNA isolation and genome walking library preparation

Genomic DNA was isolated from leaf samples by modified plant DNA mini preparation method (Dellaporta et al. 1983). Genomic DNA purity and quality was checked spectrophotometrically on NanoDrop ND 1000 (Thermo Scientific, USA) and by agarose gel electrophoresis. Genome Walker Universal kit (Clontech Laboratories, Inc. USA) was used for Genome walking library preparation by using four restriction enzymes, namely *DraI*, *PvuII*, *EcoRV* and *StuI* (Fermentas, EU).

RNA isolation and SMART library preparation

Total RNA was isolated by the lithium chloride method from various vegetative and reproductive stage tissues mentioned earlier. Total extracted RNA (10 μg) was treated with RNase-Free DNase (Qiagen) and RNA cleanup was done by RNeasy Mini Kit (Qiagen). First strand cDNA synthesis was performed by iScript cDNA synthesis kit (Bio-Rad, USA) as per manufacturer's instructions by taking 1 μg of total RNA. SMART libraries (5' SMART and 3' SMART) were prepared

using SMART RACE cDNA Amplification Kit (Clontech Laboratories, Inc. USA) as per specified protocol. A pooled mix total RNA of various vegetative and reproductive tissues was used for library preparation.

Isolation of full-length cDNA and genomic sequences of *MaCOL3a* and *MaCOL3b*

One partial genomic sequence (777 bp) of *CO-like* (*COL*) gene of Musa AAB group cultivar ‘Horn plantain’ was searched in NCBI database (<http://www.ncbi.nlm.nih.gov/>) under accession number DQ153049. Two forward MaCON1-F1 (5’AGAGCCGGAGGTTTCGAGAAGACGAT 3’), MaCON1-F2 (5’ AGGTACGCGTCGAGGAAGGCC TAC 3’) and two reverse MaCON1-R1 (5’ AAGGGAAGC AGGGGGATGCGCTCGT), MaCON1-R2 (5’ GCGGAGT GGATGTCGGCGTCACAGTC 3’) were designed to get the complete sequence of this gene. Through genome walking, (performed prior to publication of the banana genome sequence), a 563 bp fragment obtained towards initiation codon and 723 bp towards termination codon were amplified, cloned and sequenced. SMART-RACE PCRs were conducted to obtain the 5’ and 3’ UTR portions. These portions were analyzed and assembled to get the full-length gene sequence designated as *MaCOL3b*. Gene specific primers were used to amplify this gene from cDNA. .

Degenerate primers were designed by iCODEHOP v1.1 (<http://dbmi-iCDe-01.dbmi.pitt.edu/i-COdehop-COtext/>) interactive programme based on an alignment of sequences of *AtCO* (*Arabidopsis*), *OsHd1* (rice), *HvCO1* (barley) and *Zmconz1* (maize). A partial fragment of 758 bp was obtained on genomic DNA by A9 (CGCCTGCCCCATGTA acntgyaargc) and F10 (ACGCATATCGGATTGTC tttycraaytt) primer combinations. Based on this sequence, four gene specific primers MaCON2-F1 (TCCTCGCGTCGGA TGGTGGCTATTTC) and MaCON2-F2 (CAGAATCGGA TCAGTTCGTTATGCCACA) and MaCON2-R1 (GGCCTCGGCTTCTGCTTCATCGT) and MaCON2-R2 (GCAGCGGCTGTGATCAGGGGCTT) were designed from the central portion of the sequence. Using genome walking and SMART-RACE PCRs the complete genomic and cDNA sequences along with UTR portions were obtained. This gene was later named as *MaCOL3a*.

The *MaCOL3a* and *MaCOL3b* sequences were completed prior to the publication of banana genome information (D’Hont et al. 2012).

Search for additional *CONSTANS-like* genes in banana genome database

After availability of the banana genome sequence of DH-Pahang variety (*Musa acuminata*, AA genome) (D’Hont et al. 2012), the Banana Genome Hub ([http://banana-](http://banana-genome.cirad.fr/)

[genome.cirad.fr](http://banana-genome.cirad.fr/)) was searched for more *CONSTANS like* genes using BLASTP tool with protein sequences of *Arabidopsis* AtCO (AT5G15840) and rice OsHd1 (AF490467) rice. Both, the arabidopsis and rice genes are structurally conserved with two B-boxes and one CCT domain. The BLAST query resulted in identification of 25 genes with *CONSTANS-like* features. Of these nine genes belonged to Group I that included *MaCOL3a* and *MaCOL3b*. The details of Group I *CONSTANS-like* genes fetched from banana genome database are given in Table 1.

Quantitative real-time PCR analysis

CFX96 Real-Time PCR Detection System (Bio-Rad, USA) was used to obtain expression patterns of the *COL* genes from different samples. The SsoFast EvaGreen Supermix (Bio-Rad, USA) was used in 96 well low profile PCR plate (Bio-Rad, USA) and programmed for 2 min at 50 °C, 95 °C for 10 min, then 40 cycles for 15 s at 95 °C, 10 s at 57 °C, and 15 s at 72 °C. The *RPS2* (HQ853246) was used as a reference gene (Chen et al. 2011). The final data was analyzed by CFX Manager Software (Bio-Rad). Gene specific primers for each gene were designed by Primer3 (version 0.4.0) software and checked manually too. RT-PCR was used to confirm the corresponding amplicon sizes of each gene by electrophoresis and later confirmed by sequencing, out sourced from SciGenom Labs (India). Primer efficiency of all the primers was checked before performing RT-PCR. The list of primers used for qRT-PCR analyses is given in Table 2.

Bioinformatics analysis

Multiple sequence alignments were performed by ClustalW2.1 programme (<http://www.genome.jp/tools/clustalw/>). BOXSHADE (v3.21) (http://www.ch.embnet.org/software/BOX_form.html) was used for shaded background representation. Phylogenetic trees were constructed by the MEGA6 program using the Neighboring-Joining (NJ) method under default parameters (<http://www.megasoftware.net/>) (Tamura et al. 2013). Amino acid sequence similarity (%) was calculated by ClustalW output file (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). The conserved domains of COL proteins were aligned by weblogo programme using default parameters (<http://weblogo.berkeley.edu/logo.cgi>) (Crooks et al. 2004).

Results

In *Arabidopsis*, rice, maize, barley and wheat all the functional *CONSTANS* genes that have been demonstrated belong to Group I and have distinct conserved domains designated as B-box1, B-box2 and CCT domains. Hence, we focused our

Table 1 Details of Group I *MaCOL* genes information from the Banana Genome database

Banana genome annotation (Locus)	Function	Chr	Position (bp)		Strand	Length (kb)	Amino acid residues (n)	Name given in this study
			Start	End				
GSMUA_Achr9P22690_001	Putative Zinc finger protein CONSTANS-LIKE 2	chr9	27798729	27799797	+	1.069	291	<i>MaCOL2a</i>
GSMUA_Achr3P28130_001	Putative Zinc finger protein CONSTANS-LIKE 2	chr3	27514010	27515442	–	1.433	302	<i>MaCOL2b</i>
GSMUA_Achr7P16370_001	Putative Zinc finger protein CONSTANS-LIKE 3	chr7	13843660	13845085	–	1.073	324	<i>MaCOL3a</i>
GSMUA_Achr10P23700_001	Putative Zinc finger protein CONSTANS-LIKE 3	chr10	28305407	28306443	+	1.037	319	<i>MaCOL3b</i>
GSMUA_Achr9P19630_001	Putative Zinc finger protein CONSTANS-LIKE 2	chr9	21758449	21759632	–	1.184	362	<i>MaCOL4a</i>
GSMUA_Achr2P04960_001	Putative Zinc finger protein CONSTANS-LIKE 5	chr2	10205177	10206163	+	0.987	269	<i>MaCOL4b</i>
GSMUA_AchrUn_randomP18110_001	Putative Zinc finger protein CONSTANS-LIKE 5	chrUn_random	85817167	85818208	–	1.042	319	<i>MaCOL5a</i>
GSMUA_Achr1P23420_001	Putative Zinc finger protein CONSTANS-LIKE 5	chr1	17763202	17764293	–	1.092	335	<i>MaCOL5b</i>
GSMUA_Achr11P08210_001	Putative Zinc finger protein CONSTANS-LIKE 5	chr11	6362162	6363435	+	1.274	334	<i>MaCOL5c</i>

studies on the Group I *COL* genes of banana for sequence analysis and expression patterns at various developmental stages and in different tissues.

Banana has nine Group I type *COL* genes

In an effort to study the *CONSTANS-like* family in banana, the banana genome database (<http://banana-genome.cirad.fr>) was searched. A total of 25 *COL* genes belonging to Groups I, II and III were obtained. Out of 25 genes, nine genes clustered in Group I. Since the Group I genes play key role in acceleration of flowering under different photoperiodic responses reviewed by Gangappa and Botto (Gangappa and Botto 2014), we focused our research to investigate and characterize the roles of these Group I genes in relation to flowering time. Of these nine, two *COL* genes (*MaCOL3a* and *MaCOL3b*) were obtained using Genome

walking and RACE PCR strategies prior to the release of the banana genome database. The details of the 9 genes belonging to Group I including their location on the chromosome and the size are presented in Table 1. We also provide these genes a uniform nomenclature *MaCOL2a*, *MaCOL2b*, *MaCOL3a*, *MaCOL3b*, *MaCOL4a*, *MaCOL4b*, *MaCOL5a*, *MaCOL5b* and *MaCOL5c*, since *MaCOL1* was already reported, named and characterized (Chen et al. 2012). *MaCOL1* (designated in banana genome database as GSMUA_Achr5P22460_COL-9) in our analysis was clustered with Group II like genes. The genomic organization of *MaCOL* genes belonging to Group I is presented in Fig. 1. Structurally these genes are composed of two exons and one intron except *MaCOL3b* which has two introns (86 and 115 bp). The intron size of these genes ranges in length from 81 to 98 bp except *MaCOL2a* and *2b* which contain introns of 192 and 523 bp respectively. The 5' UTRs of all the genes are small ranging between 20 nt (*MaCOL3a*) to

Table 2 List of primers used in qRT-PCR experiments are displayed in 5' to 3' order and amplicon sizes are given in bp

Gene	Forward primer (5' to 3')	Reverse primer (5' to 3')	Amplicon size (in bp)
<i>MaCOL2a</i>	GATCGGAGTGTGCAATGG	GTTGGATGTATCCGGCACAA	203
<i>MaCOL2b</i>	TCAGCATAGCTTGCAGATGG	TTGAGCTGTGGTTGAACCTTGA	186
<i>MaCOL3a</i>	GAAGTCGGCCGATCTACTC	CATTGAAGAGCTCAATGGAGAG	144
<i>MaCOL3b</i>	ACGCTGATGCTTACCTGGAT	ATGATCGGCTTCTTGTGTTGC	146
<i>MaCOL4a</i>	GGCTCTCATGGCGGACAC	ACTTCCCTTCCAGCTTCGAT	142
<i>MaCOL4b</i>	GAGACGCAATTCCACCAGAT	CTCATGGAGTGGATGGTGTG	161
<i>MaCOL5a</i>	GGTTCGAGAAAACGATCAGG	GTCAAAACGAGGGCACCAC	183
<i>MaCOL5b</i>	CTCCGACGTAGATCCGTACC	GTACCATCGATCGGGAGGA	153
<i>MaCOL5c</i>	GCAACGAGAAGGATGAGGAG	GGCGTCCATCTTAGTGCTGT	173
<i>RPS2</i>	CGGGATTGCCTGATATTGTG	CCCTCACAAATTGCGGATAC	208

Fig 1 The representative structure (drawn to scale) of *MaCOL* genes (Group I). All sequence information related to genomic, CDS and UTR is borrowed from The Banana Genome Hub (<http://banana-genome.cirad.fr>) except for *MaCOL3a* and *MaCOL3b*. The sizes of exons and introns were estimated by comparing genomic and cDNA sequences

143 nt (*MaCOL4b*) while the 3' UTRs are comparatively longer between 90 nt (*MaCOL2b*) and 638 nt (*MaCOL5b*). A closer look at the organization further shows that genes *MaCOL2a* and *MaCOL2b* are devoid of B-box2 although they are grouped with Group I in the phylogenetic analysis (Figs. 2 and 3).

Amino acid similarity and phylogenetic analysis

Identified Group I type banana COL genes were further analyzed for sequence homology through multiple sequence alignment and phylogenetic analysis. The predicted proteins for these genes contain between 269 (*MaCOL4b*) to 362 (*MaCOL4a*) amino acids (Table 3). ClustalW2 analysis revealed that all of these genes contain highly conserved B-boxes with 38 amino acids except *MaCOL5a* and *MaCOL5b* that contain only 35 amino acids. The CCT domain in all is made of 38 residues with a highly variable middle region. The B-boxes are further divided into two types based on their consensus sequences and spacing of zinc-binding residues (Crocco and Botto 2013). In banana, the B-box1 consensus is C-X₂-C-X₈-C-X₂-D-X₃-L-C-X₂-C-D-X₄-H-X₈-H while the B-box2 consensus is C-X₂-C-X₈-C-X₇-C-X₂-C-X₄-H-X₈-H. The exceptions are *MaCOL4b* which has a truncated B-box1 (only 12 amino acids long) (Fig. 2) and *MaCOL2a* and *MaCOL2b* which lack B-box2 completely. Besides these domains, a valine-proline (VP) motif (17 residues) was also present in all the Group I COL genes of banana. In Arabidopsis only COL Group I genes have been

shown to contain this VP motif along with B-boxes and a CCT domain (Gangappa and Botto 2014). In addition to this, a seven amino acid motif (consensus Y/F-G-V-V-P-S-F) was observed at the carboxyl end of these proteins (except *MaCOL5c*) as against a six residue long motif reported in Arabidopsis (Griffiths et al. 2003). The highly conserved domains in the banana *MaCOL* genes of Group I are represented by the logos of the protein alignment in Fig. 3. It shows that in B-box1, 12 amino acids out of 38; in B-box2, 25 out of a total 38; in CCT domain 33 out of 42 while in VP-motif, 2 amino acid residues out of 12 are fully conserved.

The amino acid sequence similarity was investigated at two different levels i.e., overall protein similarity and a concatenation of B-boxes and CCT domain separately (Holefors et al. 2009) (Table 3). Within the Group I *MaCOL* polypeptides, *MaCOL3a* and *MaCOL3b* share highest protein similarity (82.53 %). Functional COL proteins of Arabidopsis (*AtCO*), rice (*OsHd1*), maize, (*Zmconz1*) and barley (*HvCO1*) were also compared to check the similarity of these with banana proteins. The comparison shows that *AtCO* has the highest similarity with *MaCOL3a* (44.96 %) followed by *MaCOL5a* (42.97 %), *MaCOL4a* (41.90 %), *MaCOL3b* (40.46 %), *MaCOL4b* (40.08 %), *MaCOL5b* (39.25 %), *MaCOL2b* (39.10 %), *MaCOL5c* (38.92 %) and *MaCOL2a* (35.56 %). With *OsHd1*, the highest similarity is observed in *MaCOL4a* (46.18 %) followed by *MaCOL3a* (45.32 %), *MaCOL2a* (44.95 %), *MaCOL3b* (43.07 %), *MaCOL4b* (43.03 %), *MaCOL5c* (42.81 %), *MaCOL3b* (42.14 %), *MaCOL5a* (41.23 %) and *MaCOL5b* (41.01 %). The protein similarity

Fig 2 Amino acid ClustalW multiple sequence alignment of 9 MaCOL genes with Group I genes of Arabidopsis. Conserved amino acids are highlighted in *black* and the similar in *grey*. The B-box1, B-box2 and CCT conserved sequences are marked with *horizontal lines*. Small conserved VP motif is also marked. The conserved cysteine residues (*asterisks*) and histidine residues (*filled arrowheads*) are marked in B-box1 and B-box2 domains. In CCT domain, sub-domains are also marked and conserved residues (*open arrowheads*) are also shown

with Zmconz1 ranges between 45.72 and 40.13 % in MaCOL4a and MaCOL5c respectively while it is reduced with HvCO1 (between 42.86 and 34.70 % in MaCOL2a and MaCOL5b respectively).

When the alignment and comparison is performed using concatenated B-boxes and CCT domain residues the maximum similarity is 88.64 % between MaCOL5a and MaCOL5b while the lowest is 70.49 % between MaCOL3a and MaCOL4b. Since MaCOL2a and MaCOL2b lack B-box2, these are not likely to be potential candidates for enhancing FT expression (Table 2). The highest similarity with AtCO is shown by MaCOL4b (69.81 %) followed by MaCOL5c (68.15 %), while MaCOL4b (72.64 %) shows highest similarity with OsHd1 followed by MaCOL5c, and the rest. As with overall protein similarity, a lower similarity with Zmconz1 and HvCO1 proteins (between 70.37 and 63.33 % with Zmconz1 and 60.45 % to 51.11 % with HvCO1) was observed in concatenation of B-boxes and CCT domain.

To study the relationship of these nine COL proteins of banana with all 17 CO/COL proteins of Arabidopsis, we constructed phylogenetic tree by using Neighbor-Joining (NJ) method. In this analysis we also included CONSTANS homologues of other monocots plant species like, rice (OsHd1), maize (Zmconz1) and barley (HvCO1). This analysis revealed that all nine COL proteins clustered with Group I proteins of Arabidopsis. The homologues of other monocots were also clustered in the same group along with banana and *Arabidopsis* (Fig. 4).

Expression of Group I COL genes during plant development

In order to get an insight into the possible function of the COL genes the expression patterns were studied at different developmental stages. COL genes that have been correlated with the initiation of flowering in photoperiod sensitive plants do so by influencing transcription of FT. In banana, under our growth conditions, the transition from the vegetative stage to reproductive stage occurs around 180 days after planting in the field (Fig. 5). It is therefore expected that the functional FT gene expression will coincide with or be just prior to this transition time. Accordingly for any FT that is dependent on CO, the functional CO gene would also have to express

Fig 3 The conserved domains in the MaCOL proteins. Protein alignment logos of the B-box1 (a), B-box-2 (b), CCT-domain (c) and VP-motif (d) are shown. Conserved domains of all nine banana proteins were included in the analysis, except MaCOL4b having incomplete B-box 1 (a) and MaCOL2a and MaCOL2b where B-box 2 domain completely absent (b). The height of each letter designates the conservation of residues across all the proteins. The x-axis and y-axis represents the conserved sequences of the domain and the scale of relative entropy respectively. The relative entropy shows the conservation rate of each amino acid. The conserved domain aligned by weblogo programme using default parameters (<http://weblogo.berkeley.edu/logo.cgi>)

Table 3 Amino acid sequence similarity (%) of the putative CO-like proteins of banana, Arabidopsis (AtCO, AtCOL1, AtCOL2, AtCOL3, AtCOL4) rice (OsHdl), maize (Zmconz1) and barley (HvCO1)

	AtCO	MaCOL5a	MaCOL5b	MaCOL4a	MaCOL4b	MaCOL5c	MaCOL3a	MaCOL3b	OsHdl	Zmconz1	HvCO1	MaCOL2a	MaCOL2b
AtCO	–	62.88	65.15	67.41	69.81	68.15	62.22	67.41	75.56	74.07	61.94	33.33	35.56
MaCOL5a	41.97	–	88.64	82.58	79.61	84.85	73.56	77.27	68.94	68.18	58.02	32.18	34.48
MaCOL5b	39.25	78.53	–	82.58	81.55	84.09	72.41	78.03	71.21	68.94	58.78	31.03	34.48
MaCOL4a	41.90	63.32	63.25	–	79.25	86.67	76.67	84.44	71.11	69.63	59.70	32.22	34.44
MaCOL4b	40.08	60.52	61.63	67.91	–	81.13	70.49	77.36	72.64	70.75	60.00	19.05	22.22
MaCOL5c	38.92	65.02	64.13	68.52	62.75	–	74.44	80.74	71.85	68.89	58.21	31.11	34.44
MaCOL3a	44.96	56.89	60.07	61.87	59.56	58.93	–	86.67	65.56	63.33	51.11	33.33	36.67
MaCOL3b	40.46	56.25	56.00	58.33	54.88	55.96	82.53	–	69.63	70.37	60.45	33.33	36.67
OsHdl	42.94	41.23	41.01	46.18	43.03	42.81	45.32	42.14	–	90.37	73.88	36.67	36.67
Zmconz1	41.57	40.66	39.81	45.72	44.53	40.13	43.93	43.67	74.48	–	74.63	34.44	33.33
HvCO1	35.94	36.57	34.70	38.21	36.48	37.38	39.07	38.38	67.02	64.13	–	34.44	35.56
MaCOL2a	35.56	34.66	35.29	39.75	35.86	35.32	35.56	34.12	44.95	40.77	42.86	–	32.22
MaCOL2b	39.10	36.59	37.05	41.03	39.13	35.95	39.83	37.76	43.07	42.03	41.09	73.20	–

Upper part of the table (bold) represents concatenated B-box and CCT domain. The accession numbers of these genes are given in Fig. 4

Fig 4 Phylogenetic relationship of nine banana COL genes with all Arabidopsis and functional CO proteins of rice (OsHd1), maize (Zmconz1) and barley (HvCO1). Group I, II and III type genes are separated by vertical lines. Neighbor-Joining (NJ) method was used to construct the tree. The ClustalW alignment was used to align the amino acids and unrooted dendrogram was generated by MEGA5.2 software package. Bootstrap values for 1000 resamplings are shown on each branch in percentage. The average number of substitutions per site is shown in scale below. Banana genes are marked with filled circles. The accession numbers of genes are AtCO (AT5G15840), AtCOL1 (AT5G15850), AtCOL2 (AT3G02380), AtCOL3 (AT2G24790), AtCOL4 (AT5G24930), AtCOL5 (AT5G57660), AtCOL6 (AT1G68520), AtCOL7 (AT1G73870), AtCOL8 (AT1G49130), AtCOL9 (AT3G07650), AtCOL10 (AT5G48250), AtCOL11 (AT4G15250), AtCOL12 (AT3G21880), AtCOL13 (AT2G47890), AtCOL14 (AT2G33500), AtCOL15 (AT1G28050), AtCOL16 (AT1G25440), OsHd1 (AF490467), Zmconz1 (EU098140) and HvCO1 (AF490467)

around the same time. Further since the FT is expressed in leaf (Kardailsky et al. 1999; Wigge 2011) it is expected that the expression of functional CO will also be maximum in the mature leaf at the floral induction stage.

The expression of CO-like genes was studied in the leaf at different stages of plant development. Expression patterns depicted in Fig. 6 show that of all the genes, *MaCOL3a* and *MaCOL3b* had the highest expression levels (relative to *RPS2*). Moreover, the expression of these two genes began to increase well before the induction of flowering in banana. The steady state levels remained high from 90 to 210 days that

Fig 5 Apical meristem section (a) at 150 DAP (days after planting), shows the vegetative growing part and at 195 DAP the flowering transition occurred as vegetative apical dome converted to hand and bract primordia of reproductive phase. Scale bar corresponds to 1 cm

encompassed the period of flowering. In contrast, all other genes showed a comparatively reduced expression with respect to *RPS2* during this period. Moreover, the expression of these genes was low during the period of initiation of flowering but began to increase around the time of fruiting from day 240 onwards. Interestingly the expression of *MaCOL3a* and *MaCOL3b* was at their lowest levels during this period.

Banana Group I COL genes show distinct diurnal expression patterns

Expression of CONSTANS in all plants studied has been shown to be affected by the circadian clock and light leading to diurnal oscillations that govern stability of the CO transcript and protein differently in different photoperiods thereby affecting its interaction with the FT promoter. To check the daily oscillation rhythms of these Group I COL genes in banana the diurnal expression patterns were studied over a 48 h period at 4 h intervals. As shown in Fig. 7 all genes showed a clear diurnal pattern with expression peaks during the day and troughs at night. The only exception was *MaCOL4a* which had peak expression at dusk followed by a decrease in the dark period. Although *MaCOL2a* and *MaCOL2b* had peak expression during the day, the increase began around midnight. The expression peaks of the genes *MaCOL3a*, *MaCOL3b*, *MaCOL5a*, *MaCOL5b* and *MaCOL5c* were in the afternoon while *MaCOL2a* and *MaCOL2b* exhibited their peaks in forenoon. As in the case of developmental expression *MaCOL3a* and *MaCOL3b* had higher expression compared to the other genes.

Tissue specific expression of Group I COL genes

We next studied the expression pattern of the Group I COL genes in different tissues. Samples were taken from various tissues during flowering and up to fruit maturity. Eight of the nine genes expressed in various tissues except *MaCOL4b* for which no expression could be detected. In mature leaf samples (150 days after planting), maximum transcript levels were

Fig 6 Plant developmental stage dependent gene expression analysis of eight *MaCOL* genes by quantitative Real Time PCR. Relative expression levels were checked up to 300 days after planting (DAP) and interval was maintained at 15 days between the two samples. *RPS2* was used as a

reference gene. The dotted vertical line indicates the floral transition stage from vegetative to reproductive phase. Error bars show the $\pm$ SE of minimum two biological and three technical replicates

seen for *MaCOL3a* and *MaCOL3b* (the latter being four times higher than 3a). Reduced expression was seen for the other genes while *MaCOL2b* was barely detectable. Bract, considered as a modified leaf, protects flowers until they are fully opened. The maximum expression in bracts was seen in the two *MaCOL2* genes followed by somewhat reduced expression levels for *MaCOL3a* and *MaCOL3b*. In fact *MaCOL2b* showed an almost bract specific expression. *MaCOL4a* and *MaCOL5c* showed very low expression. Four genes namely *MaCOL4a*, *MaCOL5a-c* were highly expressed in apical inflorescence (AI), arrested inside the pseudostem with *MaCOL4a* being highest. *MaCOL5c* was expressed highly in whole flower tissues, though significant expression was

also observed for *MaCOL2a*, *MaCOL3a* and *MaCOL5b*. When young fruit was examined, none of the genes seemed to be expressed highly, though some expression could be detected for *MaCOL3b* and *MaCOL5b*. Mature fruit peel was found to have the maximum expression for *MaCOL3a* and *MaCOL3b* while the pulp had high expression of *MaCOL3a*, *MaCOL3b*, *MaCOL5b* and *MaCOL5c* (Fig. 8).

Discussion

Several studies in both monocots (Miller et al. 2008; Nemoto et al. 2003; Hayama and Coupland 2004) and dicots (Chia

Fig 7 Diurnal expression of 8 *MaCOL*'s genes under natural conditions (day and night time periods are 13 and 11 h respectively). Day and night time periods are based on Sun rise and Sun set time hours. Grey backgrounds present the night periods. Relative gene expression was

studied up to 48 h after each 4 h interval. The x-axis shows the time points and y-axis relative gene expression against *RPS2*. The error bars indicate the $\pm$ SE of two biological and three technical replicates

et al. 2008; Putterill et al. 1995; Suarez-Lopez et al. 2001; Wong et al. 2014; Zhang et al. 2015; Wu et al. 2014) have shown that photoperiod responsive flowering is controlled by the CO/FT module. The control is exerted by a combination of the circadian regulation of the CO gene and differential stability of the transcript and protein in light and dark. Together these allow the CO protein to accumulate in a given photoperiod and control FT expression (Kardailsky et al. 1999; Wigge 2011). In *Arabidopsis*, CO accumulates under inductive long day conditions and induces FT expression and flowering under long day conditions (Putterill et al. 1995) while it does the same job in rice under inductive short day photoperiods by

suppressing FT expression in non-inductive long days (Hayama and Coupland 2004; Izawa 2007). Unlike *Arabidopsis* and rice, banana is a day neutral plant flowering between 160 and 180 days (with 22 to 24 leaves), regardless of the season of planting, suggesting that it has an in-built ability to measure the age via the leaf number. Studies on the role of CO in flowering in day neutral plants are limited and hence it is not yet clear as to whether this is governed by a CO/FT regulon as in photoperiod sensitive plants. Maize CO (*conz1*) is diurnally regulated in temperate (day neutral) and tropical (photoperiod sensitive) genotypes with distinct expression patterns in experimental short and long days

Fig 8 Tissue-specific expression patterns of 8 *MaCOL* genes in mature leaf (ML) at 150 days, bract (BR) at 210 days, apical inflorescence (AI) inside pseudostem after 190 days, flower (FL) including all floral parts at 220 days, young fruit (YF) at 240 days and mature fruit skin/peel (FS) and

mature fruit pulp (FP) at 270 days after planting (DAP). The error bars indicate $\pm$ SE of three biological and technical replicates. Relative expression is normalized against *RPS2*

indicating the ability to sense different photoperiods (Miller et al. 2008).

The mining of the banana genome sequence for *CONSTANS-Like* genes yielded 25 sequences that could be grouped in to 3 groups upon phylogenetic analysis. Of these, nine Group I genes that contained two B-boxes, a CCT domain towards the 3' end and a VP motif (valine-proline repeat) showed similarity to the known *COL* genes involved in photoresponsive plant species such as *Arabidopsis* and rice. We have performed a detailed transcript level characterization of these genes to understand their possible roles. Eight of the nine *COL* genes were expressed and of these two, *MaCOL3a* and *MaCOL3b*, accumulated to high levels well before

flowering and continued to remain high until the initiation of flowering (Fig. 6). What is noteworthy is that both these genes are also the ones that are most highly expressed in mature leaves, the site where perception of light signals and photoperiod occurs. Moreover, expression of both these genes decreases soon after initiation of flowering suggesting a function that is restricted to a period when the plant prepares to flower. It is known that expression of functional *FT* gene(s) occurs just before the transition from vegetative to reproductive stage and the expression of functional *COL* gene(s) coincides with *FT* expression. The expression of *MaCOL3a* and *MaCOL3b* that begins to increase just prior to flowering suggests a possible correlation between induction of flowering

and *MaCOL3a* and *MaCOL3b* expression. Thus an association with flowering initiation can be made with these two genes at least on the basis of their transcription. The other genes show an increase in transcript accumulation after the initiation of flowering during the growth of the inflorescence and development of fruit. Previous studies (Ledger et al. 2001; Zhang et al. 2011) have reported variable expression of *COL* genes in leaves and buds depending upon the stage of development. Our observations in banana also suggest differential expression of different *COL* genes at various stages of plant development in the leaf. In banana there is an extended phase of inflorescence development that continues over a period of about 1 month. It is likely that the late expression of *COL* genes other than *MaCOL3a* and *3b* may be required for the sustained development of flowers in the growing inflorescence. Presuming that at least some of the *MaCOL* genes function as enhancers of banana *FTs*, one would expect these *FTs* to possess a putative motif that is postulated to be responsible for the binding of functional CO. Tiwari et al. (Tiwari et al. 2010) have reported that the functional *FT* (*AtFT*) promoter in *Arabidopsis* possesses a unique *cis*-element TGTG(N2-3)ATG motif that is present in tandem and binds the CO protein. The banana genome data base has a total of 12 *Hd3A/FT-like* genes. The promoter analysis of these 12 genes from the database showed that this unique *cis*-element motif is present in the promoters of 3 genes namely *GSMUA_Achr9P07900_Protein HEADING DATE 3A*, *GSMUA_Achr5P06600_Protein HEADING DATE 3A* and *GSMUA_Achr5P06420_Protein FLOWERING LOCUS T* upstream of initiation codon at –446 and –488 for *GSMUA_Achr9P07900_Protein HEADING DATE 3A*; –341 for *GSMUA_Achr5P06600_Protein HEADING DATE 3A* and –290 for *GSMUA_Achr5P06420_Protein FLOWERING LOCUS T* (Fig. 9). These observations suggest, although do not prove, that there are three *Hd3A/FT-like* genes of banana that have the potential to facilitate the binding of CO protein in their promoter regions. This *in silico* analysis strengthens the

suggestion that in a day neutral banana the CO/FT regulon may be functional.

Another characteristic feature of *CO* genes involved in flowering is the rhythmic diurnal expression governed by the circadian clock and light. Interestingly, even in the day neutral banana plants, all the eight expressed genes showed a clear diurnal rhythm. Importantly, the genes *MaCOL3a* and *MaCOL3b* have just the right rhythmic expression pattern showing a peak late in the afternoon as has been reported for the functional *CO* in *Arabidopsis*. The fact that all the *COL* genes showed a circadian regulation in banana suggest a complex regulation exerted at the transcriptional and post-transcriptional levels even in a day neutral plant.

Since only two of the eight expressed *COL* genes seem to be correlated with initiation of flowering there must be additional functions governed by *COL* genes. Indeed the tissue specific expression patterns and the developmental stage patterns strongly hint towards multiple functions not necessarily related to flowering. Of the different genes, *MaCOL2a* and *2b* express primarily in bracts, *3a* and *3b* in leaves and fruit skin, *4a*, *5a*, *5b* and *5c* in apical inflorescence and *5b* and *5c* in fruit pulp. These suggest specific functions in bract, fruit skin, apical inflorescence and fruit pulp carried out by specific *COL* genes. Chen et al. (Chen et al. 2012) have made a similar observation with another *CONSTANS-like* gene (*MaCOL1*) from banana and shown that the gene was expressed highly in the flower and in the fruit pulp suggesting that *MaCOL1* may be involved in pulp ripening. In *Pharbitis nil* it has been reported that *CO* was expressed in many parts such as in floral buds, floral meristem and floral organs (Kim et al. 2003) while the sugar beet *CO* has been shown to be expressed in floral buds (Chia et al. 2008). It is possible that these *MaCOL* genes may have functions in different plant developmental processes in the vegetative and reproductive stages that need to be studied. Apart from tissue specific regulation, the fact that the *MaCOL* genes also show a distinct circadian regulation and day time peak of transcription indicate that their function in

Fig 9 Schematic representation of promoters of banana *Hd3A/FT-like* genes named *GSMUA_Achr9P07900_Protein HEADING DATE 3A* (a), *GSMUA_Achr5P06600_Protein HEADING DATE 3A* (b) and *GSMUA_Achr5P06420_Protein FLOWERING LOCUS T* (c). The promoter sequences of these genes were taken from banana genome

data base. *CONSTANS* binding unique *cis*-regulatory elements, TGTG(N2-3)ATG depicted with solid boxes in promoter of these genes. The banana *FT-like* proteins annotation taken from banana genome database are given in brackets

these tissues might require inputs from the clock and light as well (although it should be noted that the diurnal regulation of these genes was not examined in tissues other than leaves).

Taken together we suggest that the functional *COLs* in a day neutral banana are likely to be *MaCOL3a* and *MaCOL3b*. It is likely that in day neutral plants the CO function may incorporate not just light and clock signals but also an age dependent signal through as yet unknown mechanisms. The expression of certain miRNAs families such as miRNA156 and miRNA172 has been shown to be required for age dependent manner and control transition from vegetative to reproductive phase in *Arabidopsis* (Huijser and Schmid 2011; Yamaguchi and Abe 2012). These could be potential factors that regulate *COLs* in an age dependent manner. Validating these hypotheses would require further studies through silencing of specific *COLs* that are underway.

To the best of our knowledge this is the first exhaustive report on the possible roles of Group I *MaCOL* genes as enhancers of *FT* in banana.

Acknowledgments Authors thank Jain Irrigation Systems Ltd and its Founder Chairman Dr. B.H. Jain for all financial support and motivation.

Author contributions PVS, BK and APS conceived and designed the research. AKC, HBP and VRS conducted experiments. HBP, BK and VRS contributed new reagents or analytical tools. AKC, AA and PVS analyzed data. AKC, AA, APS and PVS wrote the manuscript. All authors read and approved the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Almada R, Cabrera N, Casaretto JA, Ruiz-Lara S, Gonzalez Villanueva E (2009) VvCO and VvCOL1, two CONSTANS homologous genes, are regulated during flower induction and dormancy in grapevine buds. *Plant Cell Rep* 28(8):1193–1203. doi:10.1007/s00299-009-0720-4
- Andres F, Coupland G (2012) The genetic basis of flowering responses to seasonal cues. *Nat Rev Genet* 13(9):627–639. doi:10.1038/nrg3291
- Ballerini ES, Kramer EM (2011) In the light of evolution: a reevaluation of conservation in the CO-FT regulon and its role in photoperiodic regulation of flowering time. *Front Plant Sci* 2:81. doi:10.3389/fpls.2011.00081
- Bohlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O (2006) CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* 312(5776):1040–1043. doi:10.1126/science.1126038
- Chen L, Zhong HY, Kuang JF, Li JG, Lu WJ, Chen JY (2011) Validation of reference genes for RT-qPCR studies of gene expression in banana fruit under different experimental conditions. *Planta* 234(2):377–390. doi:10.1007/s00425-011-1410-3
- Chen J, Chen JY, Wang JN, Kuang JF, Shan W, Lu WJ (2012) Molecular characterization and expression profiles of MaCOL1, a CONSTANS-like gene in banana fruit. *Gene* 496(2):110–117. doi:10.1016/j.gene.2012.01.008
- Chia TY, Muller A, Jung C, Mutasa-Gottgens ES (2008) Sugar beet contains a large CONSTANS-LIKE gene family including a CO homologue that is independent of the early-bolting (B) gene locus. *J Exp Bot* 59(10):2735–2748. doi:10.1093/jxb/ern129
- Crocco CD, Botto JF (2013) BBX proteins in green plants: insights into their evolution, structure, feature and functional diversification. *Gene* 531(1):44–52. doi:10.1016/j.gene.2013.08.037
- Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14(6):1188–1190. doi:10.1101/gr.849004
- Datta S, Hettiarachchi G, Deng X-W, Holm M (2006) *Arabidopsis* CONSTANS-LIKE3 is a positive regulator of red light signaling and root growth. *Plant Cell* 18(1):70–84. doi:10.1105/tpc.105.038182
- Dellaporta S, Wood J, Hicks J (1983) A plant DNA miniprep: version II. *Plant Mol Biol Report* 1(4):19–21. doi:10.1007/BF02712670
- D'Hont A, Denoeud F, Aury JM, Baurens FC, Carreel F, Garsmeur O, Noel B, Bocs S, Droc G, Rouard M, Da Silva C, Jabbari K, Cardi C, Poulain J, Souquet M, Labadie K, Jourda C, Lengelle J, Rodier-Goud M, Alberti A, Bernard M, Correa M, Ayyampalayam S, McKain MR, Leebens-Mack J, Burgess D, Freeling M, Mbeguie AMD, Chabannes M, Wicker T, Panaud O, Barbosa J, Hribova E, Heslop-Harrison P, Habas R, Rivallan R, Francois P, Poirion C, Kilian A, Burthia D, Jenny C, Bakry F, Brown S, Guignon V, Kema G, Dita M, Waalwijk C, Joseph S, Dievart A, Jaillon O, Leclercq J, Argout X, Lyons E, Almeida A, Jeridi M, Dolezel J, Roux N, Risterucci AM, Weissenbach J, Ruiz M, Glaszmann JC, Quetier F, Yahiaoui N, Wincker P (2012) The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 488(7410):213–217. doi:10.1038/nature11241
- Gangappa SN, Botto JF (2014) The BBX family of plant transcription factors. *Trends Plant Sci* 19(7):460–470. doi:10.1016/j.tplants.2014.01.010
- Griffiths S, Dunford RP, Coupland G, Laurie DA (2003) The evolution of CONSTANS-like gene families in barley, rice, and *Arabidopsis*. *Plant Physiol* 131(4):1855–1867. doi:10.1104/pp.102.016188
- Hayama R, Coupland G (2004) The molecular basis of diversity in the photoperiodic flowering responses of *Arabidopsis* and rice. *Plant Physiol* 135(2):677–684. doi:10.1104/pp.104.042614
- Holefors A, Opseth L, Ree Rosnes AK, Ripel L, Snipen L, Fosslad CG, Olsen JE (2009) Identification of PaCOL1 and PaCOL2, two CONSTANS-like genes showing decreased transcript levels preceding short day induced growth cessation in Norway spruce. *Plant Physiol Biochem* 47(2):105–115. doi:10.1016/j.plaphy.2008.11.003
- Huang J, Zhao X, Weng X, Wang L, Xie W (2012) The rice B-box zinc finger gene family: genomic identification, characterization, expression profiling and diurnal analysis. *PLoS ONE* 7(10):e48242. doi:10.1371/journal.pone.0048242
- Huijser P, Schmid M (2011) The control of developmental phase transitions in plants. *Development* 138(19):4117–4129. doi:10.1242/dev.063511
- Imaizumi T, Schultz TF, Harmon FG, Ho LA, Kay SA (2005) FKF1 F-Box protein mediates cyclic degradation of a repressor of CONSTANS in *Arabidopsis*. *Science* (New York, NY) 309(5732):293–297. doi:10.1126/science.1110586
- Izawa T (2007) Adaptation of flowering-time by natural and artificial selection in *Arabidopsis* and rice. *J Exp Bot* 58(12):3091–3097. doi:10.1093/jxb/ern159
- Jang S, Marchal V, Panigrahi KCS, Wenkel S, Soppe W, Deng X-W, Valverde F, Coupland G (2008) *Arabidopsis* COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic

- flowering response. *EMBO J* 27(8):1277–1288. doi:[10.1038/emboj.2008.68](https://doi.org/10.1038/emboj.2008.68)
- Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Weigel D (1999) Activation tagging of the floral inducer FT. *Science* 286(5446):1962–1965
- Kim SJ, Moon J, Lee I, Maeng J, Kim SR (2003) Molecular cloning and expression analysis of a CONSTANS homologue, PnCOL1, from *Pharbitis nil*. *J Exp Bot* 54(389):1879–1887. doi:[10.1093/jxb/erg217](https://doi.org/10.1093/jxb/erg217)
- Laubinger S, Marchal V, Le Gourrierec J, Wenkel S, Adrian J, Jang S, Kulajta C, Braun H, Coupland G, Hoecker U (2006) Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. *Development* 133(16):3213–3222. doi:[10.1242/dev.02481](https://doi.org/10.1242/dev.02481)
- Lazaro A, Valverde F, Piñeiro M, Jarillo JA (2012) The Arabidopsis E3 ubiquitin ligase HOS1 negatively regulates CONSTANS abundance in the photoperiodic control of flowering. *Plant Cell* 24(3):982–999. doi:[10.1105/tpc.110.081885](https://doi.org/10.1105/tpc.110.081885)
- Ledger S, Strayer C, Ashton F, Kay SA, Putterill J (2001) Analysis of the function of two circadian-regulated CONSTANS-LIKE genes. *Plant J* 26(1):15–22
- Lu KJ, Huang NC, Liu YS, Lu CA, Yu TS (2012) Long-distance movement of Arabidopsis FLOWERING LOCUS T RNA participates in systemic floral regulation. *RNA Biol* 9(5):653–662. doi:[10.4161/rna.19965](https://doi.org/10.4161/rna.19965)
- Miller TA, Muslin EH, Dorweiler JE (2008) A maize CONSTANS-like gene, *conz1*, exhibits distinct diurnal expression patterns in varied photoperiods. *Planta* 227(6):1377–1388. doi:[10.1007/s00425-008-0709-1](https://doi.org/10.1007/s00425-008-0709-1)
- Nemoto Y, Kisaka M, Fuse T, Yano M, Ogihara Y (2003) Characterization and functional analysis of three wheat genes with homology to the CONSTANS flowering time gene in transgenic rice. *Plant J* 36(1):82–93
- Putterill J, Robson F, Lee K, Coupland G (1993) Chromosome walking with YAC clones in Arabidopsis: isolation of 1700 kb of contiguous DNA on chromosome 5, including a 300 kb region containing the flowering-time gene CO. *Mol Gen Genet* 239(1–2):145–157
- Putterill J, Robson F, Lee K, Simon R, Coupland G (1995) The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80(6):847–857
- Sawa M, Nusinow DA, Kay SA, Imaizumi T (2007) FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. *Science (New York, NY)* 318(5848):261–265. doi:[10.1126/science.1146994](https://doi.org/10.1126/science.1146994)
- Song YH, Ito S, Imaizumi T (2013) Flowering time regulation: photoperiod- and temperature-sensing in leaves. *Trends Plant Sci* 18(10):575–583. doi:[10.1016/j.tplants.2013.05.003](https://doi.org/10.1016/j.tplants.2013.05.003)
- Strayer C, Oyama T, Schultz TF, Raman R, Somers DE, Mas P, Panda S, Kreps JA, Kay SA (2000) Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog. *Science* 289(5480):768–771
- Suarez-Lopez P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G (2001) CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature* 410(6832):1116–1120. doi:[10.1038/35074138](https://doi.org/10.1038/35074138)
- Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30(12):2725–2729. doi:[10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197)
- Tiwari SB, Shen Y, Chang HC, Hou Y, Harris A, Ma SF, McPartland M, Hymus GJ, Adam L, Marion C, Belachew A, Repetti PP, Reuber TL, Ratcliffe OJ (2010) The flowering time regulator CONSTANS is recruited to the FLOWERING LOCUS T promoter via a unique cis-element. *New Phytol* 187(1):57–66. doi:[10.1111/j.1469-8137.2010.03251.x](https://doi.org/10.1111/j.1469-8137.2010.03251.x)
- Valverde F (2011) CONSTANS and the evolutionary origin of photoperiodic timing of flowering. *J Exp Bot* 62(8):2453–2463. doi:[10.1093/jxb/erq449](https://doi.org/10.1093/jxb/erq449)
- Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G (2004) Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303(5660):1003–1006. doi:[10.1126/science.1091761](https://doi.org/10.1126/science.1091761)
- Wigge PA (2011) FT, a mobile developmental signal in plants. *Curr Biol* 21(9):R374–R378. doi:[10.1016/j.cub.2011.03.038](https://doi.org/10.1016/j.cub.2011.03.038)
- Wong ACS, Hecht VFG, Picard K, Diwadkar P, Laurie RE, Wen J, Mysore K, Macknight RC, Weller JL (2014) Isolation and functional analysis of CONSTANS-LIKE genes suggests that a central role for CONSTANS in flowering time control is not evolutionarily conserved in *Medicago truncatula*. *Front Plant Sci* 5:486. doi:[10.3389/fpls.2014.00486](https://doi.org/10.3389/fpls.2014.00486)
- Wu F, Price BW, Haider W, Seufferheld G, Nelson R, Hanzawa Y (2014) Functional and evolutionary characterization of the CONSTANS gene family in short-day photoperiodic flowering in soybean. *PLoS ONE* 9(1):e85754. doi:[10.1371/journal.pone.0085754](https://doi.org/10.1371/journal.pone.0085754)
- Yamaguchi A, Abe M (2012) Regulation of reproductive development by non-coding RNA in Arabidopsis: to flower or not to flower. *J Plant Res* 125(6):693–704. doi:[10.1007/s10265-012-0513-7](https://doi.org/10.1007/s10265-012-0513-7)
- Zhang J-X, Wu K-L, Tian L-N, Zeng S-J, Duan J (2011) Cloning and characterization of a novel CONSTANS-like gene from *Phalaenopsis hybrida*. *Acta Physiol Plant* 33(2):409–417. doi:[10.1007/s11738-010-0560-4](https://doi.org/10.1007/s11738-010-0560-4)
- Zhang R, Ding J, Liu C, Cai C, Zhou B, Zhang T, Guo W (2015) Molecular evolution and phylogenetic analysis of eight COL superfamily genes in group I related to photoperiodic regulation of flowering time in wild and domesticated cotton (*Gossypium*) species. *PLoS ONE* 10(2):e0118669. doi:[10.1371/journal.pone.0118669](https://doi.org/10.1371/journal.pone.0118669)