the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2008) **54**, 899–910 doi: 10.1111/j.1365-313X.2008.03455.x

Arabidopsis PCFS4, a homologue of yeast polyadenylation factor Pcf11p, regulates *FCA* alternative processing and promotes flowering time

Denghui Xing, Hongwei Zhao, Ruqiang Xu[†] and Qingshun Quinn Li[*]
Department of Botany, Miami University, Oxford, OH 45056, USA

Received 23 November 2007; revised 26 January 2008; accepted 8 February 2008; published online 24 April 2008.
[*]For correspondence (fax +1 513 529 4243; e-mail liq@muohio.edu).
[†]Present address: Department of Anatomy and Neurobiology, Virginia Commonwealth University, Richmond, VA 23298, USA.

Summary

The timely transition from vegetative to reproductive growth is vital for reproductive success in plants. It has been suggested that messenger RNA 3′-end processing plays a role in this transition. Specifically, two autonomous factors in the *Arabidopsis thaliana* flowering time control pathway, FY and FCA, are required for the alternative polyadenylation of *FCA* pre-mRNA. In this paper we provide evidence that Pcf11p-similar protein 4 (PCFS4), an Arabidopsis homologue of yeast polyadenylation factor Protein 1 of Cleavage Factor 1 (Pcf11p), regulates *FCA* alternative polyadenylation and promotes flowering as a novel factor in the autonomous pathway. First, the mutants of *PCFS4* show delayed flowering under both long-day and short-day conditions and still respond to vernalization treatment. Next, gene expression analyses indicate that the delayed flowering in *pcfs4* mutants is mediated by *Flowering Locus C* (*FLC*). Moreover, the expression profile of the known *FCA* transcripts, which result from alternative polyadenylation, was altered in the *pcfs4* mutants, suggesting the role of PCFS4 in *FCA* alternative polyadenylation and control of flowering time. In agreement with these observations, using yeast two-hybrid assays and TAP-tagged protein pull-down analyses, we also revealed that PCFS4 forms a complex *in vivo* with FY and other polyadenylation factors. The *PCFS4* promoter activity assay indicated that the transcription of *PCFS4* is temporally and spatially regulated, suggesting its non-essential nature in plant growth and development.

Keywords: Arabidopsis PCFS4, alternative polyadenylation, *FCA*, flowering time, *FLC*, *FY*.

Introduction

Successful reproduction in plants relies on the timely transition from vegetative to reproductive growth. Both environmental cues (such as day length, light quality and temperature) and endogenous factors (such as developmental status and age) regulate this transition (Boss *et al.*, 2004; Simpson and Dean, 2002). Arabidopsis plants perceive and coordinate these environmental and endogenous signals through several signalling pathways, as previously defined by physiological, genetic and molecular analyses (Koornneef *et al.*, 1991, 1998; Pineiro and Coupland, 1998). These pathways include the autonomous pathway, the photoperiod pathway, the vernalization pathway and the GA pathway. It has been further observed that signals from these different pathways converge at only a few defined floral integrators, including *Flowering Locus T* (*FT*), *Suppressor of Overexpression of CONSTANS 1* (*SOC1*) and *Leafy* (*LFY*), which, in turn, control the expression of the floral meristem identity genes (Kardailsky *et al.*, 1999; Kobayashi *et al.*, 1999; Ruiz-Garcia *et al.*, 1997; Samach *et al.*, 2000).

SOC1 and *FT* are direct targets of CONSTANS (CO), a zinc finger transcription factor, and Flowering Locus C (FLC), a MADS box transcription factor (Hepworth *et al.*, 2002; Michaels and Amasino, 1999; Onouchi *et al.*, 2000; Putterill *et al.*, 1995; Samach *et al.*, 2000; Searle *et al.*, 2006; Sheldon *et al.*, 1999). CO relays the signal from the photoperiod and stimulates the expression of *FT* and *SOC1*, while FLC suppresses their expression (Hepworth *et al.*, 2002; Onouchi *et al.*, 2000; Samach *et al.*, 2000). The *FLC* transcription is negatively regulated by both the autonomous and the

vernalization pathways (Michaels and Amasino, 2001). Vernalization and some autonomous pathway factors suppress the expression of *FLC* by means of histone modification of *FLC* chromatin (Bastow *et al.*, 2004; He *et al.*, 2003; Sung and Amasino, 2004). For example, lesions in *FLD* and *FVE* result in hyperacetylation of histone 3, and disruption of *SKB1* causes reduced dimethylation of histone 4 in *FLC* chromatin (He *et al.*, 2003; Wang *et al.*, 2007). The molecular mechanisms of other autonomous factors that regulate *FLC* expression remain elusive. However, the alternative polyadenylation of *FCA*, an autonomous pathway gene encoding an RNA-binding protein, has been associated with *FLC* expression and flowering time, suggesting a role for alternative polyadenylation in the control of *FLC* expression, and hence the regulation of flowering time (Macknight *et al.*, 1997, 2002; Quesada *et al.*, 2003; Simpson *et al.*, 2003).

Apart from regular polyadenylation sites [located at the end of the 3′-untranslated region (UTR)], an alternative polyadenylation site may be located in the coding sequence or an intron where polyadenylation may truncate the messenger. Alternative polyadenylation has been recognized as a powerful means of regulation of gene expression (Meyers *et al.*, 2004; Shen *et al.*, 2008; Tian *et al.*, 2005; Yan and Marr, 2005; Zhao *et al.*, 1999). For example, in animals, the expression of both immunoglobulin genes and the *Drosophila* sex-lethal gene are developmentally regulated by means of alternative polyadenylation (Peterson, 1994; Zhao *et al.*, 1999). The alternative polyadenylation of the calcitonin transcriptional unit is another classical example in which selective use of poly(A) sites resulted in two different hormones being produced (Cote *et al.*, 1992). In plants, there are only a few documented cases of alternative polyadenylation, including the S locus genes in *Brassica*, the lysine-ketoglutarate reductase gene in cotton and *FCA* and *AtCPSF30* in Arabidopsis (Delaney *et al.*, 2006; Giranton *et al.*, 1995; Macknight *et al.*, 1997; Tang *et al.*, 2002; Tantikanjana *et al.*, 1993). Among these, the alternative polyadenylation of *FCA* is the best studied in plants.

It has been demonstrated that FCA physically interacts with FY, a homologue of yeast polyadenylation factor Pfs2p, and that such interaction is essential for the alternative processing of *FCA* pre-mRNA (Simpson *et al.*, 2003). Four transcripts of *FCA* have been detected, and two major forms of them, *FCA*-γ and *FCA*-β, are derived from the use of different poly(A) sites on the *FCA* pre-mRNA (Macknight *et al.*, 1997). *FCA*-γ is the only form encoding the functional FCA (Macknight *et al.*, 2002). Mutants deficient in either FY or FCA produce a decreased amount of *FCA*-β, a non-functional form derived from the use of an alternative poly(A) site at intron 3, with a reciprocal increase in the level of *FCA*-γ (Quesada *et al.*, 2003; Simpson *et al.*, 2003). The significance of alternative polyadenylation of *FCA* pre-mRNA in flowering time is clearly shown by the correlation between alternative polyadenylation and flowering time. Further supporting this notion, the alternative polyadenylation of the *FCA* transcript regulates, both spatially and temporally, the amount of functional FCA protein based on GUS activity analyses in transgenic plants containing a series of *FCA–GUS* fusions (Macknight *et al.*, 2002). The expression of the fused gene containing *FCA* intron 3 is delayed and largely restricted to shoot and root apices, in contrast to the more universal and earlier expression patterns of the intronless fusion genes (Macknight *et al.*, 2002). In addition, FCA is an ABA receptor, and ABA treatment phenocopies the *FCA* mutants leading to delayed flowering. The binding of ABA to FCA disrupts the FY–FCA interaction and reduces the amount of *FCA*-β transcript, as in *fy* and *fca* mutants (Razem *et al.*, 2006). Finally, the *FCA* gene structure and alternative polyadenylation of *FCA* transcripts are conserved across several plant species investigated, including both monocots and dicots (Lee *et al.*, 2005; Macknight *et al.*, 2002), indicative of the evolutionary significance of FCA alternative polyadenylation.

Understanding of the molecular and biochemical mechanisms of polyadenylation in yeast and mammals has greatly advanced over the last decade. Five main protein complexes, cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF), cleavage factors I and II for mammals (CFIm and CFIIm) and poly(A) polymerase (PAP), form the core polyadenylation machinery in mammals (Proudfoot, 2004). This machinery mediates the coupled cleavage and polyadenylation reactions by recognizing and binding to *cis*-elements around the poly(A) site of a pre-mRNA. Plant polyadenylation machinery may possess a similar set of polyadenylation factors, since the homologues of most yeast and mammalian factors have been identified in the Arabidopsis genome based on sequence similarity (Addepalli and Hunt, 2007; Delaney *et al.*, 2006; Forbes *et al.*, 2006; Xu *et al.*, 2006).

It is believed that the regular polyadenylation machinery is also involved in alternative polyadenylation and that the core polyadenylation machinery is largely the same for both regular and alternative polyadenylation (Zhao *et al.*, 1999). Studies in two classical examples, the alternative polyadenylation of the mammalian calcitonin pre-mRNA and the immunoglobulin pre-mRNA, indicate that two factors, the *cis*-elements and/or the *trans*-acting factors which are specific for alternative polyadenylation, are required to recruit the core polyadenylation machinery to the alternative polyadenylation site (Cote *et al.*, 1991, 1992; Matis *et al.*, 1996; Peterson, 1994). In the FY/FCA-regulated alternative polyadenylation of *FCA* pre-mRNA, both site-specific *cis*-elements and *trans*-acting factors for the alternative polyadenylation at intron 3 remain to be addressed.

When charactering a set of T-DNA insertion mutants of putative Arabidopsis polyadenylation factors, we found that

mutants of an Arabidopsis gene encoding Pcf11p-similar protein (PCFS4), a homologue of yeast polyadenylation factor Protein 1 of Cleavage Factor 1 (Pcf11p), showed a delayed flowering phenotype. Here, we present evidence that PCFS4 is a novel autonomous pathway factor and that PCFS4 forms a complex *in vivo* with FY and other polyadenylation factors which in turn regulates *FCA* alternative polyadenylation. Based on the non-essential nature and the existence of multiple paralogues of Arabidopsis PCFS4, we propose that a specialized polyadenylation machinery, as distinct from the housekeeping one, might have evolved to regulate *FCA* alternative polyadenylation and flowering time.

Results

The Arabidopsis genome contains four homologues of yeast polyadenylation factor Pcf11p

As part of the efforts to characterize plant polyadenylation factors, we identified four *A. thaliana* homologues of yeast polyadenylation factor Pcf11p based on amino acid sequence similarity in the Arabidopsis genome: they are designated as PCF11p-Similar protein 1 or PCFS1, PCFS2, PCFS4 and PCFS5. Yeast Pcf11p is a multi-domain protein required for both transcriptional termination and 3′-end processing of pre-mRNA (Amrani *et al.*, 1997; Birse *et al.*, 1998; Hammell *et al.*, 2002; Licatalosi *et al.*, 2002). Domain mapping revealed that the N-terminal part of the C-terminal interaction domain (CID) of Pcf11p is responsible for the interaction with the C-terminal domain (CTD) of RNA polymerase II large subunit, while its C-terminal part is responsible for interaction with polyadenylation factor Clp1p, Rna14p and Rna15p (Sadowski *et al.*, 2003). The C-terminal part also contains two zinc finger motifs (Sadowski *et al.*, 2003). All four Arabidopsis homologues contain a C-terminal part corresponding to that of yeast Pcf11p with similarity ranging from 34% to 38%, and the two zinc finger motifs in the C-terminal part were also conserved between Arabidopsis and yeast (Figure 1a). Among them, PCFS4 is the only one containing an intact CID domain at its N-terminus with 52% similarity to the CID domain of yeast Pcf11p (Figure 1a and Figure S1). While PCFS2 also contains an N-terminal sequence with 51% similarity to the CID domain of yeast Pcf11p, the conserved sequence is only in the smaller middle part, suggesting a truncated CID domain in PCFS2 (Figure 1a and Figure S1).

In contrast to a single copy of PCFS homologue in other eukaryotic genomes, the multiple copies of PCFS in plant genomes (two in rice, *Oryza sativa*, and four in Arabidopsis) suggest an expansion of *PCFS* genes after the divergence of plants from other eukaryotes (Figure 1b and Figure S1). Phylogenetic analysis of PCFS homologues indicates that the *PCFS* gene was duplicated at least once before the divergence of dicots from monocots since Arabidopsis

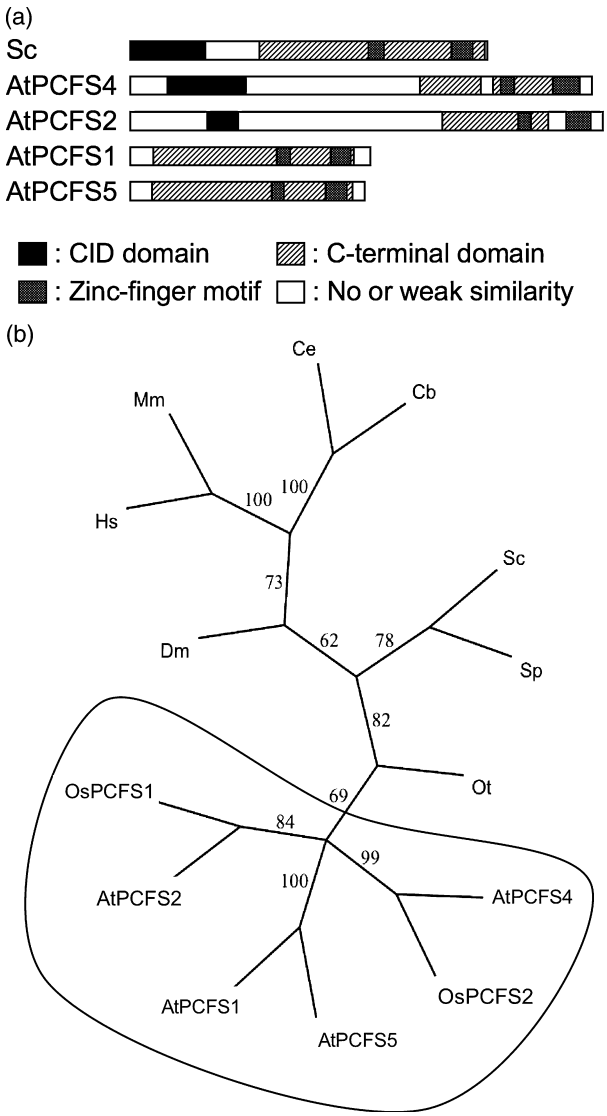

Figure 1. Arabidopsis homologues of yeast polyadenylation factor Protein 1 of Cleavage Factor 1 (Pcf11p).
(a) The domain structure of yeast polyadenylation factor Pcf11p (Sc) and its Arabidopsis homologues. The domain structures of Arabidopsis Pcf11p-similar proteins (PCFSs) were putative based on their sequence similarity to yeast Pcf11p.
(b) Phylogenetic relationship of yeast Pcf11p homologues. The peptide sequences of yeast Pcf11p (Sc, GenBank accession no. NP_010514), homologous from *Homo sapiens* (Hs, EAW75084), *Mus musculus* (Mm, NP_083354), *Caenorhabditis elegans* (Ce, NP_498068), *Caenorhabditis briggsae* (Cb, CAE70080), *Drosophila melanogaster* (Dm, NP_610999), *Schizosaccharomyces pombe* (Sp, NP_593686), *Arabidopsis thaliana* (AtPCFS1, AAW39024; AtPCFS2, NP_565848; AtPCFS4, NP_680598; AtPCFS5, ABF59044), *Oryza sativa* (OsPCFS1, BAD46662; OsPCFS2, NP_001061151) and *Ostreococcus tauri* (Ot, CAL52615), were first aligned with CLUSTALW. The tree was generated from the aligned sequences with parsimony and bootstrap programs within the PAUP 4.0 package. The homologues from plants are circled.

AtPCFS2 and AtPCFS4 are more closely related to rice OsPCFS1 and OsPCFS2, respectively (Figure 1b). Another duplication also occurred after the divergence of monocots

and dicots which gave rise to Arabidopsis AtPCFS1 and AtPCFS5 (Figure 1b). The high homology (90% of identity and 92% of similarity) between AtPCFS1 and AtPCFS5 suggests that the duplication was probably a recent event and that there might be functional redundancy between AtPCFS1 and AtPCFS5 (Figure S1).

PCFS4 mutants cause delayed flowering

To characterize the function of PCFS4, we identified two T-DNA insertion mutants (*pcfs4-1* and *pcfs4-2*) within exon 4 and intron 5, respectively, of the *PCFS4* gene in a Columbia (Col) genetic background from the SALK T-DNA collection (Figure 2a; Alonso *et al.*, 2003). Both mutants flowered later than Col in a long-day growth condition (Figure 2b and Table 1). The flowering delay is associated with an increased number of rosette and cauline leaves (Figure 2b and Table1), suggesting that the late flowering phenotype of the mutants did not result from growth retardation but rather from late floral initiation (Koornneef *et al.*, 1991). Gene expression analysis of *PCFS4* with northern blotting using *PCFS4* full-length cDNA as a probe revealed a shorter transcript in *pcfs4-1* mutant than in Col, suggesting that *PCFS4* transcription was disrupted (Figure 2c). Reverse transcription PCR (RT-PCR) assay of *PCFS4* expression confirmed that the truncated transcript in *pcfs*4-1 was derived from the 5′-end of the *PCFS4* transcript. Reverse transcription-PCR analysis with primer pair p3/p4, which spans the T-DNA insertion site in the *pcfs4-1* mutant, detected a strong fragment in Col, as expected, while no amplification was detected in *pcfs4-1* mutants (Figure 2a,d). On the other hand, an equally strong fragment was amplified in both *pcfs4-1* and Col using primer pair p1/p2, which are located upstream of the T-DNA insertion site of *pcfs4-1* (Figure 2a,d), confirming the existence of shorter transcripts.

To confirm that the late flowering phenotype of *pcfs4* mutants was indeed caused by the mutation of the *PCFS4* gene, a complementation test was performed in which the *pcfs4-1* mutant was transformed with the genomic fragment spanning the complete *PCFS4* gene (including the 1.4-kb region upstream of translation start site and the 0.7-kb region downstream of translation stop site). We observed that the flowering phenotype of transgenic plants was indeed restored to that of the wild type as demonstrated in both the T_1 (Table S1) and subsequent generations of homozygous transgenic lines (Figure 2b and Table 1), indicating that the mutation of *PCFS4* is responsible for the late flowering phenotype. In addition to the late flowering phenotype, the leaf morphology of *pcfs4* mutants was also altered, with the leaf edge curved towards the abaxial side (Figure 2b), a phenotype that was also complemented by the wild-type *PCFS4* gene in transgenic plants (Figure 2b). Other than the late flowering

Table 1 Flowering time of *pcfs4* mutants

	LD − VRN		LD + VRN	SD
Genotype	FT (day)	No. leaves	No. leaves	No. leaves
Col	28.5 ± 0.34	16.0 ± 0.50	13.5 ± 0.40	56.3 ± 2.12
pcfs4-1	34.1 ± 0.40	26.0 ± 0.67	16.2 ± 0.54	72.3 ± 2.72
pcfs4-2	32.9 ± 0.46	24.4 ± 0.46	16.7 ± 0.34	72.0 ± 2.65
11-2	29.0 ± 0.41	15.8 ± 0.65	n.d.	n.d.
16-6	29.1 ± 0.43	15.7 ± 0.49	n.d.	n.d.
fy-3	33.1 ± 1.08	27.3 ± 0.73	n.d.	n.d.
flc-3	27.1 ± 0.22	13.9 ± 0.30	n.d.	n.d.

Plants were grown in long-day (LD) or short-day (SD) conditions with (+) or without (−) vernalization (VRN). Flowering time (FT) was counted in days from the day of germination to the day of the first flower opening. Total number of leaves (No. leaves) was counted as the sum of rosette leaves and the cauline leaves on the main stem. Genotypes 11-2 and 16-6 are two homozygous complementation lines containing *PCFS4* transgene on the *pcfs4-1* mutant background. Data represent mean ± standard error of samples ($n \geq 12$).
n.d., not determined.

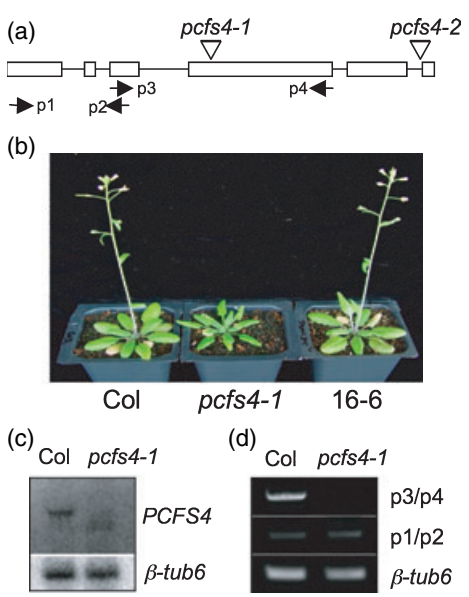

Figure 2. Arabidopsis Pcf11p-similar protein 4 (PCFS4) promotes flowering.
(a) *PCFS4* gene structure. The exons are shown as open boxes, and introns are shown as lines. The locations of the T-DNA insertion in the mutants are indicated by inverted triangles. The arrows represent the primers for analyzing *PCFS4* expression in *pcfs4-1* mutant.
(b) The delayed flowering phenotype of *pcfs4-1* mutant and its genomic complementation line 16-6 that restored wild-type flowering time in T_3 homozygote transgenic lines.
(c) Northern blot analysis of *PCFS4* expression. Ten micrograms of total RNA extracted from 15-day-old seedlings was loaded in each lane. The blot was hybridized first with *PCFS4* probe and subsequently with β-*TUB6* as an equal loading control. In *pcfs4-1* mutant, *PCFS4* transcript was truncated and in reduced expression level.
(d) Reverse transcription-PCR analysis of *PCFS4* expression. The total RNA as in (c) was treated with DNase and used for RT-PCR with the primer pairs p1/p2 and p3/p4 as indicated in (a). Transcription activity was seen for the part in front of the T-DNA insertion site (with primers p1/p2) in *pcfs4-1*.

and the altered leaf morphology, the *pcfs4* mutants grew and developed normally (Figure 2b).

Arabidopsis PCFS4 *is a novel autonomous pathway factor*

Late flowering mutants in Arabidopsis have been classified into three groups based on their responses to photoperiod and vernalization (Koornneef *et al.*, 1991, 1998). One group of mutants, including *fca* and *fy*, which belong to the autonomous pathway, shown delayed flowering in both short-day and long-day growth conditions and after vernalization treatments, while the other two groups are either insensitive to vernalization or to both short-day conditions and vernalization (Koornneef *et al.*, 1991, 1998). To determine which pathway was affected by *pcfs4* mutants, the effects of both short-day conditions and vernalization treatment on *pcfs4* flowering time were tested. The flowering time of *pcfs4-1* and *pcfs4-2*, measured by total leaf number when the first flower was open, was greatly reduced after 22 days of vernalization treatment (Table 1) than without vernalization treatment, suggesting that the vernalization pathway was still intact in *pcfs4* mutants. On the other hand, the leaf number of *pcfs4-1* and *pcfs4-2* grown in a short-day photoperiod was twice that of *pcfs4-1* and *pcfs4-2* grown in long-day conditions (Table 1), suggesting that the photoperiod pathway was not affected in *pcfs4* mutants. In either situation, the leaf numbers of *pcfs4* mutants were still more than that of Col control grown under the same conditions (Table1). As the effect of *pcfs4* mutants on flowering time is independent of photoperiod and vernalization, it can be concluded that Arabidopsis PCFS4 is a novel autonomous pathway factor.

The late flowering of PCFS4 *mutants is mediated by FLC*

The factors in the autonomous pathway promote flowering by suppressing *FLC* expression (Lim *et al.*, 2004; Michaels and Amasino, 2001; Mockler *et al.*, 2004; Simpson *et al.*, 2003; Wang *et al.*, 2007). If the same is true for PCFS4, *FLC* expression should be increased in the *pcfs4* mutants. Compared with wild-type Col, the abundance of *FLC* mRNA was indeed increased in *pcfs4-1* mutant, which was similar to the *FLC* level in *fy-3* mutant, but less abundant than that found in *fy-2* mutant (Figure 3a). Consistent with their similar *FLC* expression levels, *pcfs4-1* and *fy-3* mutants had their flowering delayed to a similar extent (Table 1). More importantly, the *FLC* expression level in the complementation line 16-6 was restored to a level similar to that of the wild type as well (Figure 3a), indicative that the role of PCFS4 on the flowering time is mediated by FLC.

Consistent with the increase of *FLC* expression, the accumulation of *SOC1*, *FT* and *LFY* mRNA (genes whose expressions are negatively regulated by FLC; Boss *et al.*, 2004; Hepworth *et al.*, 2002; Michaels and Amasino, 2001), was decreased in *pcfs4-1* as well as in *fy-2* and *fy-3* mutants (Figure 3b). Again, the transcript abundance of these genes was restored to the level of the wild-type control in the complemented line 16-6. Thus, our results show that Arabidopsis PCFS4 promotes flowering by suppressing FLC-mediated inhibition of flowering.

To test whether *FLC* is the sole target of PCFS4 in regulating flowering time, we created a *pcfs4-1 flc-3* double mutant. In long-day growth conditions, the *pcfs4-1 flc-3* double mutant flowered earlier than the *pcfs4-1* mutant, comparable to the Col control but later than the *flc-3* mutant

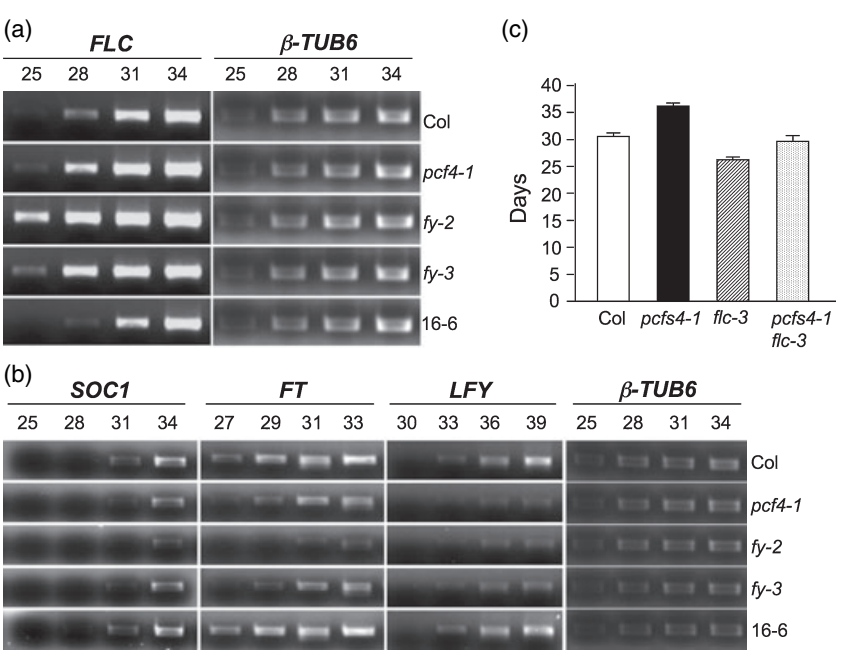

Figure 3. Arabidopsis Pcf11p-similar protein 4 (PCFS4) suppresses *FLC* expression and the phenotype of *pcfs4-1 flc3* double mutant.
Semi-quantitative RT-PCR analysis of *FLC* expression (a) and the expression of floral integrator genes *SOC1*, *FT* and *LFY* (b). Total RNA was extracted from 15-day-old seedlings 8 h after dawn for *FLC* and 16 h after dawn for *SOC1*, *FT* and *LFY*. Total RNA was treated with DNase I and no amplification was detected in any sample without adding reverse transcriptase for each gene investigated (data not shown). β-*TUB6* was used as an internal control. The PCR cycle numbers are listed under the gene names.
(c) The flowering time of the double mutant *pcfs4-1 flc3*. Plants were grown under long-day conditions. Flowering time was counted in days from the day of germination to the day of opening of the first flower. Data represent an average and standard error of samples ($n \geq 16$).

(Figure 3c). On the one hand, this result confirmed the role of *FLC* in mediating the function of PCFS4 in the control of flowering time. On the other hand, it indicated that PCFS4 also regulated flowering time through an *FLC*-independent pathway because it still postponed flowering in the absence of *FLC*.

Arabidopsis PCFS4 *promotes the alternative polyadenylation of FCA pre-mRNA*

Since both PCFS4 and FY are polyadenylation factors and since the characteristics of the *pcfs4* mutants have a remarkable resemblance to those of *fy* and *fca* in both flowering time and molecular phenotype (Figures 2 and 3, Table 1; Macknight *et al.*, 2002; Simpson *et al.*, 2003), we hypothesized that PCFS4 might function together with FY and FCA to regulate *FCA* pre-mRNA processing. To test this hypothesis, we first asked whether the levels of different transcripts of *FCA* were altered in the *PCFS4* mutants. Indeed, the abundance of *FCA*-β was significantly decreased in *pcfs4-1* mutant compared with Col, although to a lesser extent than that observed in the *fy-2* mutant (Figure 4a,b). The reduced *FCA*-β level was accompanied by the increase of *FCA*-γ in both *pcfs4-1* and *fy-2* mutants (Figure 4a,b), as one would expect based on what was observed previously (Simpson *et al.*, 2003). These results demonstrated that, similar to FY and FCA, Arabidopsis PCFS4 stimulates the usage of the alternative poly(A) site within *FCA* intron 3 and thus alters the expression profile of *FCA*.

Arabidopsis PCFS4 *forms a complex* in vivo *with FY and other polyadenylation factors*

In light of the above results, we then ask whether PCFS4 physically interacts with FY and FCA. To make this determination, we conducted yeast two-hybrid (Y2H) assays. The coding sequences of PCFS4 and FY were cloned into

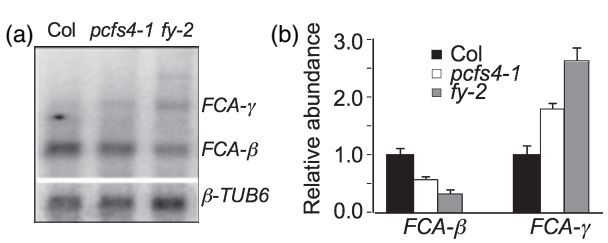

Figure 4. Arabidopsis PCFS4 promotes *FCA* alternative polyadenylation.
(a) Northern blot analysis of *FCA* expression. Total RNA was extracted from 16-day old seedlings. The poly(A) RNA prepared from 400 μg of total RNA was loaded in each lane. The blot was hybridized first with the *FCA* probe (Simpson *et al.*, 2003) and subsequently with β-*TUB6* probe. A representative blot is shown.
(b) Quantification of *FCA* expression based on northern blots. The quantified signal was normalized relative to β-*TUB6* first and then presented as abundance level relative to that of wild type control (Col). The data represent the average and standard error of three independent experimental results.

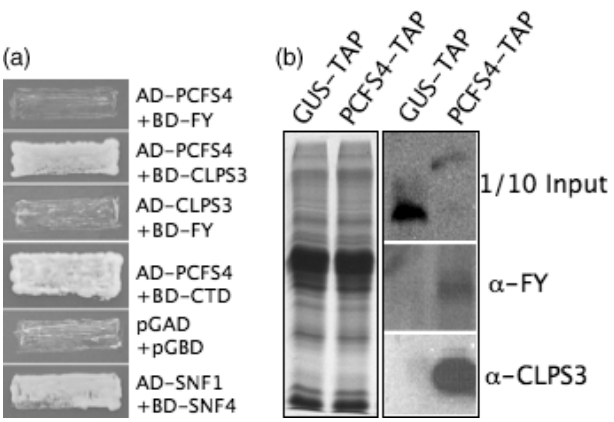

Figure 5. Arabidopsis Pcf11p-similar protein 4 (PCFS4) forms a protein complex with FY and Clp1p-similar protein 3 (CLPS3).
(a) Yeast two-hybrid analyses of pair-wise interactions of PCFS4/FY, PCFS4/CLPS3, CLPS3/FY and PCFS4/CTD. AD, activation domain; BD, binding domain; CTD, RNA polymerase II C-terminal domain domain. The vector plasmids pGAD/pGBD were used as a negative control while AD-SNF1/BD-SNF4 was a positive control (James *et al.*, 1996).
(b) Western blot analyses of total protein and tandem affinity purification (TAP) purified protein from the total protein containing PCFS4–TAP or GUS–TAP (as a control) fusion protein. The affinity-purified protein was analysed by using anti-FY antibody (right middle panel) and anti-CLPS3 antibody (right bottom panel), respectively. One-tenth of the total protein input for the purification was represented by the amount of PCFS4–TAP and GUS–TAP fusion protein detected by peroxidase-conjugated anti-peroxidase IgG (right top panel) and Coomassie blue staining (left panel).

Gateway® compatible pGAD and pGBD plasmids (Xu *et al.*, 2006), resulting in GAD–PCFS4 and GBD–FY fusions, respectively. No interaction was observed between GAD–PCFS4 and GBD–FY (Figure 5a), indicating that PCFS4 did not directly interact with FY. Since both GBD–PCFS4 and GBD–FCA showed self-activation activity for GAL4 expression (data not shown), the interaction between PCFS4 and FCA could not be addressed with the Y2H assay. However, we found that PCFS4 did interact with another putative polyadenylation factor, Clp1p-similar protein 3 (termed CLPS3; GenBank accession no. NP_187119; Figure 5a). CLPS3 is an Arabidopsis homologue of yeast polyadenylation factor cleavage/polyadenylation factor IA subunit (Clp1p) that has been tightly associated with Pcf11p in yeast, both functionally and physically (Minvielle-Sebastia *et al.*, 1997; Noble *et al.*, 2007). This result raised the possibility that the interaction between PCFS4 and FY might be bridged by CLPS3. However, CLPS3 did not directly interact with FY in the Y2H assays either (Figure 3a). These results still do not rule out the possibility that PCFS4 indirectly interacts with FY through other protein factor(s). It is possible that PCFS4 and FY might exist in a multi-subunit complex *in vivo* in a way that cannot be directly detected by the Y2H assay. Therefore, we applied an alternative strategy, tandem affinity purification (TAP; (Rohila *et al.*, 2004), to detect the possible *in vivo* protein–protein associations.

To accomplish this we generated *PCFS4–TAP* and GUS–*TAP* (used as a control) gene fusions driven by the CaMV 35S promoter and transformed the constructs into Arabidopsis suspension culture cells. Following the TAP procedure, immunoblot analyses of the purified proteins were performed with anti-FY antibody (Henderson *et al.*, 2005) and anti-CLPS3 antibody. As shown in Figure 5(b), no signal was detected with either anti-FY or anti-CLPS3 antibodies in the GUS–TAP (negative control) purified sample. In contrast, in the PCFS4–TAP purified sample, a peptide of ~40 kDa and another of ~75 kDa were detected by anti-CLPS3 and anti-FY antibodies, respectively (Figure 5b). The molecular weights of detected peptides are consistent with the predicted sizes for CLPS3 and FY, indicating the presence of these proteins in the PCFS4–TAP purified complex. These results show that Arabidopsis PCFS4 forms a protein complex *in vivo* with FY and CLPS3, and interacts directly with CLPS3 and indirectly with FY potentially through other unknown protein(s).

Expression of Arabidopsis PCFS4 *is developmentally regulated*

To further probe the function of PCFS4, the temporal and spatial patterns of *PCFS4* expression were analysed using β-glucuronidase (GUS) staining analysis in transgenic plants. The 1.4-kb promoter region upstream of *PCFS4* translation start site, which was proved to be sufficient to activate the *PCFS4* gene in the complementation test described above, was fused to the GUS gene, resulting in $PCFS4_{pro}$–GUS fusion. The fusion construct was transformed into wild-type Col plants. The GUS activity was then analysed for a variety of developmental stages of the transgenic plants (Figure 6a). In embryo development, GUS activity was not detectable until the heart stage (Figure 6a, parts a–d). Starting from the torpedo stage, GUS activity spreads through the whole developing embryo (Figure 6a, parts d–f). Intriguingly, the GUS activity started to disappear in the root tip of the mature embryo (of upturned-U stage), and the lack of GUS activity in the root tip persisted for the rest of the plant's life cycle (Figure 6a, parts f–i). In the seedlings, GUS activity was primarily evident in the shoot apical meristem and young leaves (Figure 6a, parts g–j). In flowers, GUS activity was detectable in sepals, stamens and pistils, but not in petals (Figure 6a, part k). This pattern of GUS activity indicates that *PCFS4* transcription is developmentally and tissue-specifically regulated.

To assess whether the *PCFS4* promoter activity detected by GUS staining indeed reflects the expression of *PCFS4 in vivo*, we performed semi-quantitative RT-PCR assays for its expression in different tissues of 10-day-old wild-type seedlings. As shown in Figure 6(b), the *PCFS4* expression was hardly detectable in root tips and cotyledons, in contrast to the strong signal in shoots and roots (excluding tip). This result is mostly consistent with the promoter activity of seedlings of the same age seen in Figure 6(a, part i), hence validating the GUS staining results.

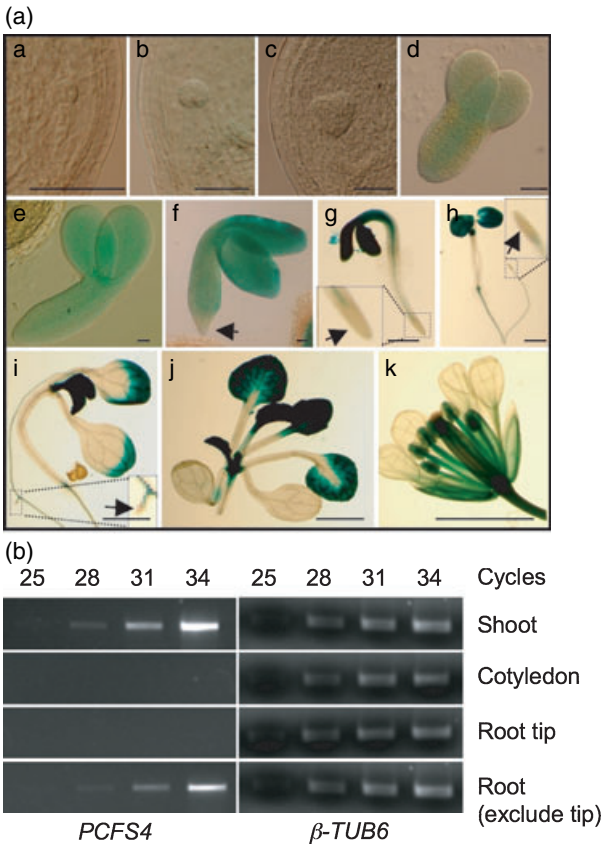

Figure 6. Pcf11p-similar protein 4 (PCFS4) expression pattern.
(a) The expression profiles of Arabidopsis *PCFS4* examined by GUS activity in the transgenic plants containing $PCFS4_{pro}$–GUS fusion in a variety of developmental stages.
(a–c) Early stages (four-cell, globular, heart) and (d–f) late stages (torpedo, early upturned-U, later upturned-U) of embryo development; (g–j) seedlings at 2, 4, 7 and 14 days post-germination, and (k) an open flower. Arrows in (f–i) point to the root tips, noting the disappearance of GUS activity. The boxed inserts (g–i) are enlarged to show root tips. Bar = 25 μm in (a–f) and 1 mm in (g–k).
(b) Semi-quantitative PCR of PCFS4 expression in shoots, cotyledons, root tips and roots (without root tip) from 10-day-old seedlings. β-*TUB6* was used as an internal control.

The subcellular localization of PCFS4 and CLPS3

To determine the subcelluar localization of Arabidopsis PCFS4 and CLPS3, the gene constructs of *PCFS4–GFP* and *CLPS3–GFP* fusions driven by the CaMV 35S promoter were made and transformed into wild-type Col. The GFP fluorescence in the transgenic plants was examined under a confocal microscope. The GFP signals in *PCFS4–GFP* transgenic plants were observed in both cytoplasm and nucleus as in the control plants containing 35S–*GFP* transgene (Figure 7), suggesting that PCFS4 is located in both the cytoplasm and

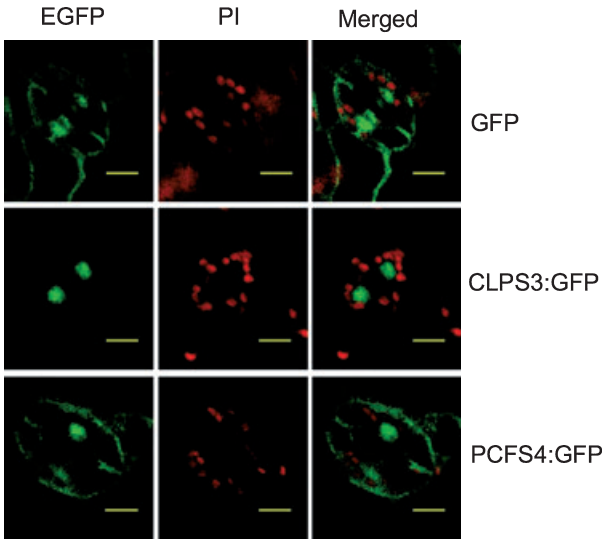

Figure 7. Subcellular localization of Pcf11p-similar protein 4 (PCFS4) and Clp1p-similar protein 3 (CLPS3).
The GFP green fluorescence and the fluorescence from chloroplast were examined with enhanced green fluorescence protein (EGFP) and a PI filter (PI) in the guard cells of wild-type plants (Col), transgenic plants containing *GFP* (GFP), *PCFS4:GFP* (PCFS4:GFP) or *CLPS3:GFP* (CLPS3:GFP) protein fusion, respectively. The green fluorescence and PI pictures were merged to show the accurate localization of GFP fluorescence (Merged). Bar = 10 μm.

the nucleus. In contrast, the GFP signals in *CLPS3–GFP* transgenic plants are solely located in the nucleus (Figure 7), suggesting the nuclear localization of CLPS3. The nuclear localization of both PCFS4 and CLPS3 is consistent with their physical interaction (Figure 5) and with their cellular function in 3′-end processing of mRNA, which occurs in the nucleus. The cytoplasmic localization of PCFS4 suggests that it may have additional cellular function(s).

Discussion

By characterizing the *PCFS4* mutants we provide evidence that Arabidopsis PCFS4 promotes flowering as a novel factor in the autonomous pathway (Figure 2 and Table 1). Similar to other autonomous pathway factors, the control of flowering time by PCFS4 is mediated by FLC and the floral integrators downstream of FLC (Figure 3). However, an FLC-independent pathway(s) was also involved in PCFS4-regulated control of flowering time based on the *pcfs4-1 flc-3* double mutant analysis (Figure 3). The suppression of *FLC* expression by FY and FCA is correlated with the alternative polyadenylation of *FCA* pre-mRNA (Macknight *et al.*, 2002; Simpson *et al.*, 2003). This correlation, along with the evolutionally conserved FCA alternative processing across a variety of species and the developmentally regulated nature of *FCA* expression, indicates that alternative polyadenylation may play a role in the control of *FLC* expression directly or indirectly (Lee *et al.*, 2005; Macknight *et al.*, 2002). Our results demonstrate that, in addition to FY and FCA, PCFS4 is also involved in *FCA* alternative polyadenylation (Figure 4). Furthermore, PCFS4 forms a complex *in vivo* with FY, CLPS3 and other unknown protein(s) (Figure 5). Consistent with this notion, both PCFS4 and CLPS3 are located in the nucleus (Figure 7).

Based on sequence similarity, domain structure and phylogenetic analysis, PCFS4 was found to be one of four Arabidopsis homologues of yeast Pcf11p (Figure 1 and Figure S1). Yeast Pcf11p is essential for cell viability and interacts with multiple proteins *in vivo* (Amrani *et al.*, 1997; Birse *et al.*, 1998; Gross and Moore, 2001; Licatalosi *et al.*, 2002; Minvielle-Sebastia *et al.*, 1997; Sadowski *et al.*, 2003). In yeast, the interaction between the RNA polymerase II CTD and the N-terminal CID domain of Pcf11p is important for transcriptional termination, while the interactions of the CTDs of Pcf11p with Clp1p, RNA14 and RNA15 are required for 3′-end processing of pre-mRNA (Hammell *et al.*, 2002; Licatalosi *et al.*, 2002; Sadowski *et al.*, 2003). Arabidopsis PCFS4, like yeast Pcf11p, also interacts with Arabidopsis RNA polymerase II CTD domain and CLPS3, the Arabidopsis homologue of yeast Clp1p (Figure 5). However, the homozygous mutants of Arabidopsis *PCFS4,* unlike those of yeast Pcf11p, were viable (Figure 2b), suggesting the non-essential nature of *PCFS4*. Although the viability of the *pcfs4* mutants might be explained by the partial expression of the gene in the mutant (Figure 2c,d), the *PCFS4* expression patterns strongly support the notion that Arabidopsis PCFS4 is dispensable for plant viability since *PCFS4* is only expressed in specific tissues and/or at specific developmental stages (Figure 6).

On the other hand, the developmentally regulated nature of *PCFS4* expression supports its specialized biological function(s). The multiple homologues of PCFS in the Arabidopsis genome provide the potential capacity for one of the PCFS homologues, such as PCFS4, to evolve from being a general polyadenylation factor into one with a specialized function, e.g. the alternative polyadenylation of *FCA* pre-mRNA in the context of this study. This specialized function of PCFS4, through interaction with FY and presumably FCA, is further supported by the overlapping expression patterns of *PCFS4*, *FY* and *FCA*. Each of these three genes (*PCFS4*, *FY* and *FCA*), is highly expressed on shoot meristem (Figure 6; Macknight *et al.*, 2002; Henderson *et al.*, 2005) where the floral organ initiates.

In contrast to *FY* expression, which is detectable from the very beginning of embryo development (Henderson *et al.*, 2005), *PCFS4* expression is only detectable from a late stage of embryo development (Figure 6). *FY* expresses universally across different tissues and developmental stages and is indispensable for the growth and development of Arabidopsis (Henderson *et al.*, 2005). The *PCFS4* expression pattern is, however, more similar to that of *FCA*, particularly in the sense that their expressions are developmentally

regulated and not detectable at the early stages of plant development (Macknight *et al.*, 2002; Figure 6). Similar to *FY*, *CLPS3* is constitutively expressed and essential for embryo development (D.X., Q.Q.L., unpublished results).

In summary, our data suggest that PCFS4 and FCA might play a more specific role in the alternative polyadenylation of *FCA* and/or other genes while FY and CLPS3 may be general polyadenylation factors that are recruited into the complex, when needed, for the alternative polyadenylation of *FCA*. As such, alternative polyadenylation of *FCA* becomes one of the mechanisms through which flowering time in Arabidopsis is fine-tuned.

Experimental procedures

Plant materials, growth conditions and flowering time analysis

All *A. thaliana* genotypes used in this study were in Col-0 (CS6000) background. The *pcfs4-1* (SALK_102934), *pcfs4-2* (SALK_063588), *fy-2* and *fy-3* mutants were obtained from the Arabidopsis Biological Resources Center (ABRC, Ohio State University; Alonso *et al.*, 2003). Mutant *flc-3* was a kind gift from Dr Yuehui He. Seeds were sown in plastic pots (50 × 60 mm) with soil (Conrad Fafard, http://www.fafard.com/) stratified at 4°C for 2 days and moved to a growth chamber. The growth conditions were set as the following: temperature 22°C; light intensity 100 μmol m^{-2} sec^{-1}; photoperiods 16-h light/8-h dark for long-day and 8-h light/16-h dark for short-day. Homozygous *pcfs4-1* and *pcfs4-2* were genotyped based on kanamycin resistance and PCR markers using gene-specific primers (p376/p379 for *pcfs4-1* and p401/p422 for *pcfs4-2*) and T-DNA left boarder primer (p341; all primer sequences are listed in Table S2). Genotyping of *fy-2* was carried out as described (Henderson *et al.*, 2005). The double mutant *pcfs4-1 flc-3* was created by crossing *pcfs4-1* and *flc-3* mutants, and the F_2 plants were genotyped with PCR markers as described (Michaels and Amasino, 2001).

Flowering time was analysed as described (Koornneef *et al.*, 1991). In brief, flowering time was scored as the number of days from seed germination to the opening of the first flower. Leaf number was counted as the total number of rosette and cauline leaves on the main stem of each plant.

Plasmid construction and transformation

All gene fusions in this study were constructed by employing the Gateway® cloning technology (Invitrogen, http://www.invitrogen.com/). The DNA fragments of interest were first amplified with PCR and cloned into pTOPO-D vector (Invitrogen), forming entry clones, and then introduced into destination vectors by the LR reaction. All clones were confirmed by DNA sequencing. Gene fusion constructs in appropriate T-DNA plasmids were first introduced into *Agrobacterium tumefaciens* strain GV3101 and then transferred into Col plants (Clough and Bent, 1998) or suspension cells (Menges and Murray, 2004), as described, and screened by appropriate antibiotics or herbicide.

For the complementation test of the wild-type *PCFS4* gene to the *pcfs4-1* mutant allele, a 5.5-kb genomic fragment spanning the *PCFS4* gene (1.5 kb extended upstream of the translation start site and 0.7 kb extended downstream of the translation stop site) was amplified with primers p601 and p602 (Table S2) from a bacterial artificial chromosome (BAC) clone T4B21 (from ABRC) with eLongase Enzyme Mix (Invitrogen), as described by the manufacturer. The PCR product was cloned into pTOPO-D vector, resulting in the $PCFS4_{gene}$–pTOPO entry clone, and it is further cloned into pMDC123 binary vector (Brand *et al.*, 2006), resulting in plasmid $PCFS4_{gene}$–pMDC123, using methods as described by Xu and Li (2008). The *pcfs4-1* mutant plants were transformed, and six independent homozygote lines were identified. All of the transgenic lines were restored to wild type, as shown in Figure 1 and Table 1.

For Y2H assays, the coding sequences (CDSs) of *PCFS4*, *FY*, *CLPS3* and the RNA polymerase II CTD were fused to the Gal4 activation or binding domain, respectively. Briefly, the *PCFS4* CDS was amplified from plasmid RAFL21-50-C04 (purchased from Riken; http://pfgweb.gsc.riken.go.jp) with primer pair p459/p442 and cloned into pTOPO-D, resulting in $PCFS4_{CDS}$-pTOPO(TAG); the *FY* CDS was amplified from plasmid pUNI51 (from ABRC) with primer pairs pFY(ATG)/pFY(TAG), and cloned into pTOPO-D, resulting in plasmid FY_{CDS}–pTOPO; The CDS of the RNA Pol II CTD domain was amplified with primer pair p590/p591 from the CTD–GST construct (a gift from Dr A. Hunt, University of Kentucky, Lexington, KY, USA) and cloned into pTOPO-D vector. LR reactions were then performed to introduce the CDSs of *PCFS4*, *FY* and CTD into pGAD or pGBD vectors, resulting in pGAD–*PCFS4*, pGBD–*FY* and pGBD–CTD constructs, respectively. The *CLPS3* CDS was introduced into pGBD vector from plasmid U17288 (from ABRC), resulting in the pGBD–*CLPS*3 construct. The Y2H assay was performed as described (Xu *et al.*, 2006).

For the TAP purification assay, the *PCFS4* CDS was amplified from RAFL21-50-C04 with primer pair p459/p442 and cloned into pTOPO-D, resulting in $PCFS4_{CDS}$–pTOPO(-TAG). The *PCFS4* CDS was further cloned into cTAPi binary vector (Rohila *et al.*, 2004), resulting in PCFS4–TAP plasmid which was used for Arabidopsis cell culture transformation (Menges and Murray, 2004).

For *PCFS4* promoter activity analysis, $PCFS4_{pro}$–GUS fusion was constructed. A 1.5-kb DNA fragment upstream of the translation start codon of PCFS4 was amplified with primer pair p601 and p584 from BAC clone T4B21. The PCR product was cloned into pTOPO-D vector, resulting in a $PCFS4_{pro}$–pTOPO entry clone and further cloned into pMDC162 (Brand *et al.*, 2006), resulting in plasmid $PCFS4_{pro}$–pMDC162, which was then finally transformed into Col plants.

For generating PCFS4:GFP and CLPS3:GFP fusions, the *PCFS4* CDS was amplified from plasmid RAFL21-50-C04 with primer pair p459/p589 and cloned into pTOPO-D, resulting in $PCFS4_{CDS}$–pTOPO(-TAG). The *PCFS4* CDS was further cloned into pMDC83 by the LR reaction, resulting in PCFS4–pMDC83 plasmid. The CLPS3 CDS was introduced into pMDC83 from plasmid U17288 (from ABRC) by the LR reaction, resulting in plasmid CLPS3–pMDC83.

Gene expression analysis

For semi-quantitative analysis, total RNA, extracted from 16-day-old seedlings, was treated with DNase I (Ambion, http://www.ambion.com/), and then 3 μg of DNA-free total RNA was used for a reverse transcription reaction (20 μl) with $Oligo(dT)_{18}$ primer using Superscipt III reverse transcriptase (Invitrogen) following the manufacturer's instructions. For PCR, 1 μl of reverse transcription product was used as a template in a 20-μl reaction. For northern blot analysis, 10 μg of total RNA, or poly(A) RNA from 400 μg of total RNA, were fractioned on a 1.2% denatured gel and blotted to a nylon membrane. The cDNA probes were labelled with Priming-it II® Random Primer Labeling Kit (Stratagene, http://www.stratagene.com/). Hybridization and washing were as previously described (Xing and Chen, 2006).

Expression of the *PCFS4* promoter was monitored by GUS activity (Jefferson *et al.*, 1987) of the transgenic plants containing $PCFS4_{pro}$–pMDC162. Briefly, seedlings grown on MS medium were harvested at different times. Open flowers and siliques at different developmental stages were collected from plants grown in soil. The siliques were carefully opened (without wounding the developing seeds) with a dissection blade and stained with 5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid, cyclohexylammonium salt (X-Gluc) at 37°C for 16 h. For differential interference contrast (DIC) microscopy examination of GUS activity of the embryos, the siliques were fixed in solution (ethanol:acetic acid = 1:1) for 4–8 h and cleared in Hoyes Medium overnight as described (Stangland and Salehian, 2002). For later-stage (from the torpedo stage) embryos, the embryos were carefully squeezed out from seed coat before observation under a DIC microscope.

Protein complex isolation and western analyses

The TAP assay was performed essentially as described (Rohila *et al.*, 2004). Briefly, the protein extract was cultured with IgG Sepharose beads (GE Healthcare Bioscience; http://www.gelifesciences.com) at 4°C for 4 h then washed four times. The protein complex was released from the beads with AcTEV protease (incubated for 2 h at 16°C; thus the protein A part of the TAP-tag is free from the protein complex). After a brief centrifugation (105, 12 000 ***g***), the supernatant was used for immunoblot analysis. The western blot analysis was performed as described (Xu *et al.*, 2006). The FY antibody was a gift from Caroline Dean (John Innes Centre, UK). The detection of the TAP tag protein A part was done by using the peroxidase-conjugated anti-peroxidase antibody (Sigma, http://www.sigmaaldrich.com/).

The peptide 'ARRRFAKASTSNDPEC' at 111 amino acids downstream of the N-terminal of CLPS3 was synthesized, conjugated to KLH by an Imject maleimide-activated mcKLH kit (Pierce, http://www.piercenet.com/) and used to immunize two New Zealand white rabbits as described (Li *et al.*, 1998). Peptide-specific CLPS3 antibodies were purified from CLPS3 cross-reactive serum using affinity purification techniques (AminoLink; Pierce). The CLPS3 antibody was used at a concentration of 1:2000 according to standard procedures.

PCFS4 and CLPS3 subcellular localization

The transgenic plants containing *PCFS4*–pMDC83 or *CLPS3–pMDC83* plasmid were germinated in MS medium for 3–5 days. The GFP signals in the guard cells of cotyledons were examined under a confocal microscope.

Phylogenetic analysis

The amino acid sequences of homologous PCFS proteins from a variety of eukaryotes were downloaded from NCBI, aligned with CLUSTALW (Felsenstein, 1993) and then analysed using parsimony and bootstrap programs within the PAUP 4.0 package.

Acknowledgements

We thank Caroline Dean for the gift of FY antibody, Yuehui He for the *flc* mutant seeds, Art Hunt for the *FY* and *CTD* clones and Joe Ecker and ABRC for the T-DNA insertion lines. We appreciate Art Hunt, Richard C. Moore and David Martin for critical reading of the manuscript; Chris Wood for assistance in MU's Center for Bioinformatics and Functional Genomics; Richard Edelmann and Matthew Duley in MU's EM facility for microscopy analysis; and assistance from the MU's Instrument Laboratory. Shirisha Yelamanchili contributed in the early stage of the CLPS3 antibody production. This research was funded by US NSF Arabidopsis 2010 Program (MCB 0313472) and, in part, by a NIH grant (1R15GM077192-01A1, both to QQL).

Supplementary Material

The following supplementary material is available for this article online:

Table S1. Flowering time of T_1 transgenic plants containing the *PCFS4* gene in *pcfs4* mutants as measured by leaf number.

Table S2. Oligonucleotide primers used in this study.

Figure S1. Multiple alignment of yeast Pcf11p homologues from variety of eukaryotes as described in Figure 1.

This material is available as part of the online article from http://www.blackwell-synergy.com.

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Addepalli, B. and Hunt, A.G. (2007) A novel endonuclease activity associated with the Arabidopsis ortholog of the 30-kDa subunit of cleavage and polyadenylation specificity factor. *Nucleic Acids Res.* **35**, 4453–4463.

Alonso, J.M., Stepanova, A.N., Leisse, T.J. *et al.* (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, **301**, 653–657.

Amrani, N., Minet, M., Wyers, F., Dufour, M.E., Aggerbeck, L.P. and Lacroute, F. (1997) PCF11 encodes a third protein component of yeast cleavage and polyadenylation factor I. *Mol. Cell. Biol.* **17**, 1102–1109.

Bastow, R., Mylne, J.S., Lister, C., Lippman, Z., Martienssen, R.A. and Dean, C. (2004) Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature*, **427**, 164–167.

Birse, C.E., Minvielle-Sebastia, L., Lee, B.A., Keller, W. and Proudfoot, N.J. (1998) Coupling termination of transcription to messenger RNA maturation in yeast. *Science*, **280**, 298–301.

Boss, P.K., Bastow, R.M., Mylne, J.S. and Dean, C. (2004) Multiple pathways in the decision to flower: enabling, promoting, and resetting. *Plant Cell*, **16**(Suppl.), S18–S31.

Brand, L., Horler, M., Nuesch, E., Vassalli, S., Barrell, P., Yang, W., Jefferson, R.A., Grossniklaus, U. and Curtis, M.D. (2006) A versatile and reliable two-component system for tissue-specific gene induction in Arabidopsis. *Plant Physiol.* **141**, 1194–1204.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.

Cote, G.J., Nguyen, I.N., Lips, C.J., Berget, S.M. and Gagel, R.F. (1991) Validation of an *in vitro* RNA processing system for CT/CGRP precursor mRNA. *Nucleic Acids Res.* **19**, 3601–3606.

Cote, G.J., Stolow, D.T., Peleg, S., Berget, S.M. and Gagel, R.F. (1992) Identification of exon sequences and an exon binding protein involved in alternative RNA splicing of calcitonin/CGRP. *Nucleic Acids Res.* **20**, 2361–2366.

Delaney, K.J., Xu, R., Zhang, J., Li, Q.Q., Yun, K.Y., Falcone, D.L. and Hunt, A.G. (2006) Calmodulin interacts with and regulates the RNA-binding activity of an Arabidopsis polyadenylation factor subunit. *Plant Physiol.* **140**, 1507–1521.

Felsenstein, J. (1993) *PHYLIP (Phylogeny Inference Package) Version 3.5c.* Seattle, WA: Department of Genetics, University of Washington.

Forbes, K.P., Addepalli, B. and Hunt, A.G. (2006) An Arabidopsis Fip1 homolog interacts with RNA and provides conceptual links with a number of other polyadenylation factor subunits. *J. Biol. Chem.* **281**, 176–186.

Giranton, J.L., Ariza, M.J., Dumas, C., Cock, J.M. and Gaude, T. (1995) The S locus receptor kinase gene encodes a soluble glycoprotein corresponding to the SKR extracellular domain in *Brassica oleracea*. *Plant J.* **8**, 827–834.

Gross, S. and Moore, C. (2001) Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I. *Proc. Natl Acad. Sci. USA*, **98**, 6080–6085.

Hammell, C.M., Gross, S., Zenklusen, D., Heath, C.V., Stutz, F., Moore, C. and Cole, C.N. (2002) Coupling of termination, 3′ processing, and mRNA export. *Mol. Cell. Biol.* **22**, 6441–6457.

He, Y., Michaels, S.D. and Amasino, R.M. (2003) Regulation of flowering time by histone acetylation in Arabidopsis. *Science*, **302**, 1751–1754.

Henderson, I.R., Liu, F., Drea, S., Simpson, G.G. and Dean, C. (2005) An allelic series reveals essential roles for FY in plant development in addition to flowering-time control. *Development*, **132**, 3597–3607.

Hepworth, S.R., Valverde, F., Ravenscroft, D., Mouradov, A. and Coupland, G. (2002) Antagonistic regulation of flowering-time gene SOC1 by CONSTANS and FLC via separate promoter motifs. *EMBO J.* **21**, 4327–4337.

James, P., Halladay, J. and Craig, E.A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*, **144**, 1425–1436.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907.

Kardailsky, I., Shukla, V.K., Ahn, J.H., Dagenais, N., Christensen, S.K., Nguyen, J.T., Chory, J., Harrison, M.J. and Weigel, D. (1999) Activation tagging of the floral inducer FT. *Science*, **286**, 1962–1965.

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M. and Araki, T. (1999) A pair of related genes with antagonistic roles in mediating flowering signals. *Science*, **286**, 1960–1962.

Koornneef, M., Hanhart, C.J. and van der Veen, J.H. (1991) A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **229**, 57–66.

Koornneef, M., Alonso-Blanco, C., Blankestijn-de Vries, H., Hanhart, C.J. and Peeters, A.J. (1998) Genetic interactions among late-flowering mutants of Arabidopsis. *Genetics*, **148**, 885–892.

Lee, J.H., Cho, Y.S., Yoon, H.S., Suh, M.C., Moon, J., Lee, I., Weigel, D., Yun, C.H. and Kim, J.K. (2005) Conservation and divergence of FCA function between Arabidopsis and rice. *Plant Mol. Biol.* **58**, 823–838.

Li, Q., Dasgupta, J. and Hunt, A. (1998) Polynucleotide phosphorylase is a component of a novel plant poly(A) polymerase. *J. Biol. Chem.* **273**, 17539–17543.

Licatalosi, D.D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J.B. and Bentley, D.L. (2002) Functional interaction of yeast pre-mRNA 3′ end processing factors with RNA polymerase II. *Mol. Cell*, **9**, 1101–1111.

Lim, M.H., Kim, J., Kim, Y.S., Chung, K.S., Seo, Y.H., Lee, I., Kim, J., Hong, C.B., Kim, H.J. and Park, C.M. (2004) A new *Arabidopsis* gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. *Plant Cell*, **16**, 731–740.

Macknight, R., Bancroft, I., Page, T. *et al.* (1997) FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. *Cell*, **89**, 737–745.

Macknight, R., Duroux, M., Laurie, R., Dijkwel, P., Simpson, G. and Dean, C. (2002) Functional significance of the alternative transcript processing of the Arabidopsis floral promoter FCA. *Plant Cell*, **14**, 877–888.

Matis, S.A., Martincic, K. and Milcarek, C. (1996) B-lineage regulated polyadenylation occurs on weak poly(A) sites regardless of sequence composition at the cleavage and downstream regions. *Nucleic Acids Res.*, **24**, 4684–4692.

Menges, M. and Murray, J.A. (2004) Cryopreservation of transformed and wild-type Arabidopsis and tobacco cell suspension cultures. *Plant J.* **37**, 635–644.

Meyers, B.C., Vu, T.H., Tej, S.S., Ghazal, H., Matvienko, M., Agrawal, V., Ning, J. and Haudenschild, C.D. (2004) Analysis of the transcriptional complexity of *Arabidopsis thaliana* by massively parallel signature sequencing. *Nat. Biotechnol.* **22**, 1006–1011.

Michaels, S.D. and Amasino, R.M. (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*, **11**, 949–956.

Michaels, S.D. and Amasino, R.M. (2001) Loss of FLOWERING LOCUS C activity eliminates the late-flowering phenotype of FRIGIDA and autonomous pathway mutations but not responsiveness to vernalization. *Plant Cell*, **13**, 935–941.

Minvielle-Sebastia, L., Preker, P.J., Wiederkehr, T., Strahm, Y. and Keller, W. (1997) The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3′-end formation. *Proc. Natl Acad. Sci. USA*, **94**, 7897–7902.

Mockler, T.C., Yu, X., Shalitin, D. *et al.* (2004) Regulation of flowering time in Arabidopsis by K homology domain proteins. *Proc. Natl Acad. Sci. USA*, **101**, 12759–12764.

Noble, C.G., Beuth, B. and Taylor, I.A. (2007) Structure of a nucleotide-bound Clp1-Pcf11 polyadenylation factor. *Nucleic Acids Res.* **35**, 87–99.

Onouchi, H., Igeno, M.I., Perilleux, C., Graves, K. and Coupland, G. (2000) Mutagenesis of plants overexpressing CONSTANS demonstrates novel interactions among Arabidopsis flowering-time genes. *Plant Cell*, **12**, 885–900.

Peterson, M.L. (1994) Regulated immunoglobulin (Ig) RNA processing does not require specific cis-acting sequences: non-Ig RNA can be alternatively processed in B cells and plasma cells. *Mol. Cell. Biol.* **14**, 7891–7898.

Pineiro, M. and Coupland, G. (1998) The control of flowering time and floral identity in Arabidopsis. *Plant Physiol.* **117**, 1–8.

Proudfoot, N. (2004) New perspectives on connecting messenger RNA 3′ end formation to transcription. *Curr. Opin. Cell Biol.* **16**, 272–278.

Putterill, J., Robson, F., Lee, K., Simon, R. and Coupland, G. (1995) The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell*, **80**, 847–857.

Quesada, V., Macknight, R., Dean, C. and Simpson, G.G. (2003) Autoregulation of FCA pre-mRNA processing controls Arabidopsis flowering time. *EMBO J.* **22**, 3142–3152.

Razem, F.A., El-Kereamy, A., Abrams, S.R. and Hill, R.D. (2006) The RNA-binding protein FCA is an abscisic acid receptor. *Nature*, **439**, 290–294.

Rohila, J.S., Chen, M., Cerny, R. and Fromm, M.E. (2004) Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J.* **38**, 172–181.

Ruiz-Garcia, L., Madueno, F., Wilkinson, M., Haughn, G., Salinas, J. and Martinez-Zapater, J.M. (1997) Different roles of flowering-

time genes in the activation of floral initiation genes in Arabidopsis. *Plant Cell*, **9**, 1921–1934.

Sadowski, M., Dichtl, B., Hubner, W. and Keller, W. (2003) Independent functions of yeast Pcf11p in pre-mRNA 3′ end processing and in transcription termination. *EMBO J.* **22**, 2167–2177.

Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Z., Yanofsky, M.F. and Coupland, G. (2000) Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. *Science*, **288**, 1613–1616.

Searle, I., He, Y., Turck, F., Vincent, C., Fornara, F., Krober, S., Amasino, R.A. and Coupland, G. (2006) The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. *Genes Dev.* **20**, 898–912.

Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J. and Dennis, E.S. (1999) The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. *Plant Cell*, **11**, 445–458.

Shen, Y., Ji, G., Haas, B.J., Wu, X., Zheng, J., Reese, G.J. and Li, Q.Q. (2008) Genome level analysis of rice mRNA 3′-end processing signals and alternative polyadenylation. *Nucleic Acids Res.* In press.

Simpson, G.G. and Dean, C. (2002) Arabidopsis, the Rosetta stone of flowering time? *Science*, **296**, 285–289.

Simpson, G.G., Dijkwel, P.P., Quesada, V., Henderson, I. and Dean, C. (2003) FY is an RNA 3′ end-processing factor that interacts with FCA to control the Arabidopsis floral transition. *Cell*, **113**, 777–787.

Stangland, B. and Salehian, Z. (2002) An improved clearing method for GUS assay in Arabidopsis endosperm and seeds. *Plant Mol. Biol. Rep.* **20**, 107–114.

Sung, S. and Amasino, R.M. (2004) Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature*, **427**, 159–164.

Tang, G., Zhu, X., Gakiere, B., Levanony, H., Kahana, A. and Galili, G. (2002) The bifunctional LKR/SDH locus of plants also encodes a highly active monofunctional lysine-ketoglutarate reductase using a polyadenylation signal located within an intron. *Plant Physiol.* **130**, 147–154.

Tantikanjana, T., Nasrallah, M.E., Stein, J.C., Chen, C.H. and Nasrallah, J.B. (1993) An alternative transcript of the S locus glycoprotein gene in a class II pollen-recessive self-incompatibility haplotype of *Brassica oleracea* encodes a membrane-anchored protein. *Plant Cell*, **5**, 657–666.

Tian, B., Hu, J., Zhang, H. and Lutz, C.S. (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* **33**, 201–212.

Wang, X., Zhang, Y., Ma, Q., Zhang, Z., Xue, Y., Bao, S. and Chong, K. (2007) SKB1-mediated symmetric dimethylation of histone H4R3 controls flowering time in Arabidopsis. *EMBO J.* **26**, 1934–1941.

Xing, D. and Chen, Z. (2006) Effects of mutations and constitutive overexpression of *EDS1* and *PAD4* on plant resistance to different types of microbial pathogens. *Plant Sci.* **171**, 251–262.

Xu, R. and Li, Q.Q. (2008) Streamline cloning of genes into binary vectors in *Agrobacterium* via the Gateway® TOPO vector system. *Plant Methods*, **4**, 4.

Xu, R., Zhao, H., Dinkins, R.D., Cheng, X., Carberry, G. and Li, Q.Q. (2006) The 73 kD subunit of the cleavage and polyadenylation specificity factor (CPSF) complex affects reproductive development in Arabidopsis. *Plant Mol. Biol.* **61**, 799–815.

Yan, J. and Marr, T.G. (2005) Computational analysis of 3′-ends of ESTs shows four classes of alternative polyadenylation in human, mouse, and rat. *Genome Res.* **15**, 369–375.

Zhao, J., Hyman, L. and Moore, C. (1999) Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63**, 405–445.

