이학박사 학위논문

TAF15b에 의한 개화 시기의 분자적 조절 기작

Molecular mechanism of TAF15b

on the regulation of flowering time

2018년 8월

서울대학교 대학원

생명과학부

엄현주

Abstract

Molecular mechanism of TAF15b on the regulation of flowering time

Hyunjoo Eom
Department of Biological Sciences
The Graduate School
Seoul National University

TATA-binding protein-associated factors (TAFs) are general transcription factors within the TFIID complex, which recognizes the core promoter of genes. In addition to their biochemical function, it is known that several TAFs are involved in the regulation of developmental processes. In this study, I found that TAF15b regulates flowering time, especially through the autonomous pathway (AP) in *Arabidopsis*. The mutant *taf15b* shows late flowering compared to the wild type during both long and short days, and vernalization accelerates the flowering time of *taf15b*. In addition, *taf15b* shows strong upregulation of *FLOWERING LOCUS C* (*FLC*), a flowering repressor in *Arabidopsis*, and the *flc taf15b* double mutant completely offsets the late

flowering of *taf15b*, indicating that *TAF15b* is a typical AP gene. The *taf15b* mutant also shows increased transcript levels of *COOLAIR*, an antisense transcript of *FLC*. Consistently, ChIP analyses showed that the TAF15b protein is enriched around both sense and antisense transcription start sites of the *FLC* locus. In addition, co-immunoprecipitation showed that TAF15b interacts with RNA polymerase II (Pol II), while ChIP showed increased enrichment of the phosphorylated forms, both serine 2 (Ser2) and Ser5, of the C-terminal domain of Pol II at the *FLC* locus, which is indicative of transcriptional elongation. *taf15b* mutant showed higher enrichment of the active histone marker, H3K4me3, on *FLC* chromatin. To understand the biochemical mechanism of TAF15b, IP-MS assay was performed and components of TAF15b complex were identified. Among them, C-TERMINAL DOMAIN PHOSPHATASE-LIKE 1 (CPL1) was selected as a candidate gene to regulate *FLC* transcription with TAF15b and *FLC* transcripts were indeed increased in the mutant of *cpl1*. Taken together, these results suggest that TAF15b regulates flowering time through transcriptional repression of *FLC* in *Arabidopsis*.

Keywords: *TAF15b, FLC,* RNA polymerase II, transcriptional regulation, autonomous pathway, flowering time, *CPL1*

Student number: 2013-30101

Most of the work in Chapter II, III was published at

Hyunjoo Eom, Su Jung Park, Min Kyung Kim, Hoyeun Kim, Hunseung Kang, Ilha Lee. 2018. TAF15b, involved in the autonomous pathway for flowering, represses transcription of *FLOWERING LOCUS C*. The Plant Journal 93(1):79-91

Chapter V was published at

Hyunjoo Eom and Ilha Lee. 2018. Role of TAF15b in Transcriptional Regulation of Autonomous Pathway for Flowering. Plant Signaling & Behavior (accepted)

Table of contents

List of figures

List of table

List of abbreviations

General abbreviations

3-AT	3-Amino-1,2,4-triazole
3', 5'	3-prime, 5-prime
AP	autonomous flowering pathway
bp	basepair
BiFC	Bimolecular fluorescence complementation
CTD	carboxy terminal domain of RNA polymerase II
CsVp	cassava vein mosaic virus
35Sp	Cauliflower mosaic virus 35S promoter
ChIP	chromatin immunoprecipitation
Col	Columbia
DNA	deoxyribonucleic acid
dsRBD	double-stranded RNA binding domain
GO	gene ontology
GA	gibberellic acid
GFP	green fluorescent protein
GUS	*β-glucuronidase*
h	Hour
IP	immunoprecipitation
kb	kilobase
kDa	kilodalton
LC	low complexity
min	minute
Pol II	RNA polymerase II
PrD	prion-like domain

qRT-PCR	quantitative real-time polymerase chain reaction
RT-PCR	Reverse transcription-polymerase chain reaction
RNA	ribonucleic acid
s	second
SD	standard deviation
TSS	transcription start site
TTS	transcription termination site
UTR	untranslated region
WT	Wild-type
Y2H	yeast two hybrid

Amino acids

Nonpolar side chains

Glycine	Gly	G
Alanine	Ala	A
Valine	Val	V
Leucine	Leu	L
Isoleucine	Ile	I
Methionine	Met	M
Phenylalanine	Phe	F
Tryptophane	Trp	W
Proline	Pro	P

Polar side chains

Serine	Ser	S
Threonine	Thr	T

Cysteine	Cys	C
Tyrosine	Tyr	Y
Asparagine	Asn	N
Glutamine	Gln	Q

Acidic side chains

Aspartic acid	Asp	D
Glutamic acid	Glu	E

Basic side chains

Lysine	Lys	K
Arginine	Arg	R
Histidine	His	H

Abbreviations of gene and protein names

Arabidopsis thaliana

ACT7	ACTIN7	AT5G09810
CDKC;2	CYCLIN-DEPENDENT KINASE C;2	AT5G64960
CPL1	C-TERMINAL DOMAIN PHOSPHATASE-LIKE 1	AT4G21670
CstF64	CLEAVAGE STIMULATION FACTOR 64	AT1G71800
CstF77	CLEAVAGE STIMULATION FACTOR 77	AT1G17760
Distal COOLAIR		GQ352646
FCA	*	AT4G16280
FES1	FRIGIDA-ESSENTIAL 1	AT2G33835
FLC	FLOWERING LOCUS C	AT5G10140
FLD	FLOWERING LOCUS D	AT3G10390

FLK	FLOWERING LOCUS K HOMOLOGY DOMAIN	AT3G04610
FLX	FLC EXPRESSOR	AT2G30120
FPA	*	AT2G43410
FRI	FRIGIDA	AT4G00650
FRL1	FRIGIDA LIKE 1	AT5G16320
FT	FLOWERING LOCUS T	AT1G65480
FVE	*	AT2G19520
FY	*	AT5G13480
GRP7	GLYCINE-RICH RNA BINDING PROTEIN 7	AT2G21660
GRP8	GLYCINE-RICH RNA BINDING PROTEIN 8	AT4G39260
LD	LUMINIDEPENDENS	AT4G02560
LFY	LEAFY	AT5G61850
PCFS4	Pcf11p-SIMILAR PROTEIN 4	AT4G04885
Proximal COOLAIR		GQ342259
PRP39-1	pre-mRNA PROCESSING PROTEIN 39-1	AT1G04080
PRP8	pre-mRNA PROCESSING PROTEIN 8	AT1G80070
SOC1	SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1	AT2G45660
SR45	SERINE/ARGININE-RICH 45	AT1G16610
SUF4	SUPPRESSOR OF FRIGIDA4	AT1G30970
Ta3	TA3/ATCOPIA44 retrotransposon	AT1G37110
TAF14	TATA-BINDING PROTEIN-ASSOCIATED FACTOR 14	AT2G18000
TAF15	TATA-BINDING PROTEIN-ASSOCIATED FACTOR 15	AT1G50300
TAF15b	TATA-BINDING PROTEIN-ASSOCIATED FACTOR 15b	AT5G58470
TUB2	TUBULIN BETA CHAIN 2	AT5G62690

Homo sapiens

EWSR1	EWING'S SARCOMA BREAKPOINT REGION 1	NP_053733
FUS	FUSED IN SARCOMA	NP_004951
TAF15	TATA-BINDING PROTEIN-ASSOCIATED FACTOR 15	NP_631961

CHAPTER Ⅰ

General introduction

1. Flowering-time control in *Arabidopsis thaliana*

Flowering is one of the most important developmental processes for plant survival. Thus, the timing of flowering is finely tuned by both endogenous and environmental signals. In *Arabidopsis thaliana* (L.) Heynh., there are various ecotypes, which can be divided into two groups depending on their flowering times: early- and late-flowering ecotypes. Allelic variations at two genes, *FLOWERING LOCUS C* (*FLC*) and *FRIGIDA* (*FRI*), mainly cause a difference in flowering time (Clarke and Dean, 1994; Duncan et al., 2015; Gazzani et al., 2003; Lempe et al., 2005; Li et al., 2014; Michaels et al., 2003; Shindo et al., 2006). *FLC* encodes a MADS-box transcription factor inhibiting flowering and FRI activates the transcription of *FLC* (Choi et al., 2011; Johanson et al., 2000; Michaels, 1999; Sheldon et al., 1999). The late-flowering ecotypes have functional alleles of both *FRI* and *FLC*, whereas early-flowering ecotypes have null alleles of either *FLC* or *FRI* or both (Gazzani et al., 2003; Johanson et al., 2000; Michaels et al., 2003). Vernalization, a long-term cold exposure which accelerates flowering, causes late-flowering ecotypes to flower early by repressing *FLC* expression (Michaels, 2001).

2. Autonomous flowering pathway

Koornneef *et al*. (1991) isolated a group of mutants that are late-flowering irrespective of day-length and also highly sensitive to vernalization. In addition, double mutations with *flc* completely offset the late-flowering phenotype, indicating that their major downstream target is *FLC* (Michaels, 2001; Rouse et al., 2002). Such mutants have been classified as autonomous pathway (AP) mutants (Bäurle and Dean, 2006; Koornneef et al., 1998; Oh and Lee, 2007; Simpson, 2004).

There are 16 genes involved in the AP and they can be categorized into three groups depending on their putative biochemical functions: RNA processing, transcriptional regulation, and histone modification (Figure 1). In the RNA processing group, *FCA*, *FY*, *FPA, cleavage stimulation factor 64* (*CstF64*), *CstF77*, and *Pcf11p-similar protein 4* (*PCFS4*) are likely to be involved in the cleavage and polyadenylation of mRNA (Macknight *et al.*, 1997; Schomburg *et al.*, 2001; Simpson *et al.*, 2003; Xing *et al.*, 2008; Liu *et al.*, 2010). In addition, some AP genes in the RNA processing group encode proteins involved in RNA splicing, such as glycine-rich RNA binding protein 7 (GRP7), GRP8, pre-mRNA processing protein 8 (PRP8), PRP39-1, and serine/arginine-rich 45 (SR45) (Ali et al., 2007; Marquardt et al., 2014; Streitner et al., 2008; Wang et al., 2007). The gene, *FLOWERING LOCUS KH DOMAIN* (*FLK*), is also a member of the RNA processing group and contains three K homology (KH) domains, a typical arrangement found in poly(C)-binding

ribonucleoproteins (PCBPs) (Lim, 2004; Mockler et al., 2004). All the mutants of these genes show elevated levels of *FLC* transcripts but they do not affect the alternative splicing pattern of *FLC*.

The second group of AP genes encodes putative transcription factors, including *LUMINIDEPENDENS* (*LD*) and *cyclin-dependent kinase C;2* (*CDKC;2*) (Aukerman et al., 1999; Lee et al., 1994; Wang et al., 2014). *LD* encodes a protein with a homeodomain and a domain found in transcription elongation factor S-II (TFIIS) but its transcriptional activity has not been reported. CDKC;2 is a component of transcription elongation factor b (P-TEFb), which affects global RNA polymerase II (Pol II) Ser2 phosphorylation levels (Wang et al., 2014). Although the *cdkc;2* mutant shows increased levels of *FLC*, it is proposed that CDKC;2 does not directly affect *FLC* transcription.

The last group is involved in histone modification and includes *FVE* and *FLOWERING LOCUS D* (*FLD*) (Ausín et al., 2004; He, 2003; Kim et al., 2004; Sanda and Amasino, 1996). FVE is a homologue of human retinoblastoma-associated proteins and is co-immunoprecipitated with histone deacetylase complex (Ausín *et al.*, 2004; Jeon and Kim, 2011). FLD is a homologue of human LYSINE SPECIFIC DEMETHYLASE1 (LSD1) which has H3K4 demethylase activity (He *et al.*, 2003; Liu *et al.*, 2007; Jiang *et al.*, 2009).

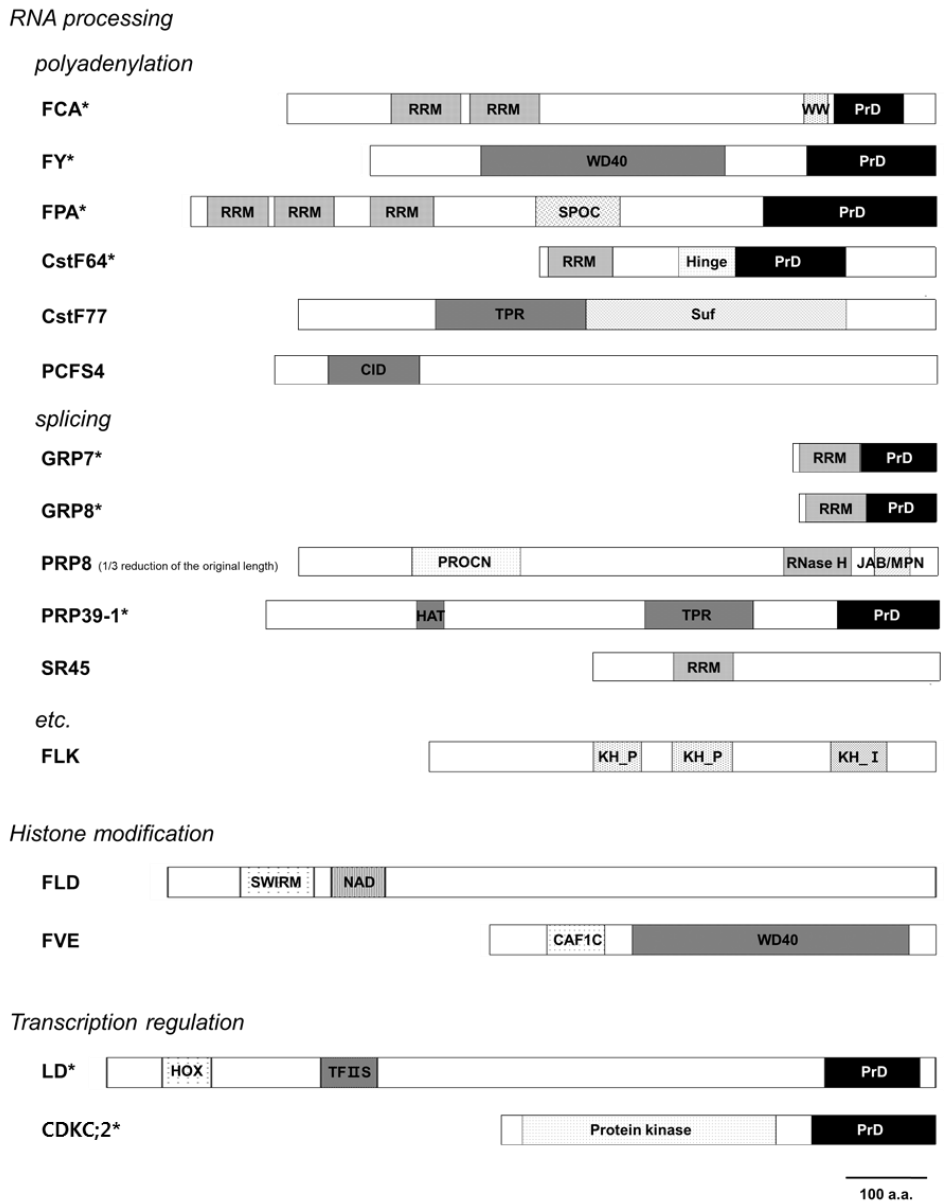

Figure 1. Protein structures of autonomous pathway genes

CAF1C, subunit C of CAF1 complex; CID, CTD-interacting domain; HAT, Half-A-TPR repeats; Hinge, hinge domain of cleavage stimulation factor subunit 2; HOX, Homeodomain;

KH_I, K homology RNA-binding domain_type I; KH_P, K homology RNA-binding domain_PCBP like; NAD, NAD(P)-binding Rossmann-like domain; PrD, predicted prion-like domain; RRM, RNA recognition motif; SPOC, Spen paralogue and orthologue C-terminal domain; Suf, Suppressor of forked protein; SWIRM, SWIRM domain; TFIIS, N-terminal domain of transcription elongation factor S-II; TPR, Tetratricopeptide repeat; WD40, WD40 domain; WW, two conserved tryptophans domain; Zn, zinc finger domain.

3. Aims of the thesis

There is no doubt that autonomous pathway genes repress *FLC* transcription but their biochemical function has not clearly explained yet. From this study, I defined *TAF15b* as a new autonomous pathway gene and tried to understand the molecular mechanism of autonomous pathway. Biological function of TAF15b on the regulation of flowering time was studied and protein localization was observed at both tissue and cellular levels. I suggested the molecular mechanism of TAF15b on the repression of *FLC* transcription and proteomic analysis of TAF15b complex was also performed to figure out the interacting partners of TAF15b.

CHAPTER II.

Biological role of TAF15b on the regulation of flowering time

Construction of *CsVp::TAF15b-GFP* and GUS staining of TAF15bp::GUS were performed by Su Jung Park of Chonnam National University.

1. Introduction

My research about TAF15b started out by professor Hoonseung Kang from Jeonnam national university, who has been investigating about RNA binding proteins and abiotic stress response in plants for a long time. He observed many mutants related to RNA binding and figured out that mutant of *TAF15b* showed late flowering phenotype and he sent us this mutant. TAF15b is one of the TATA-binding protein (TBP)-associated factors (TAFs) and they are usually components of the TFIID complex that directly binds to core promoters to form a pre-initiation complex with RNA polymerase II (Pol II) (Burley and Roeder, 1996; Hampsey and Reinberg, 1997). The TFIID complex is composed of the TATA-box-binding protein (TBP) and 8–12 TBP-associated factors (TAFs) (Albright and Tjian, 2000). In *Arabidopsis*, 18 putative TAFs were identified through BLAST search using TAF protein sequences from diverse organisms and there are two TAF15 proteins (TAF15 and TAF15b), which have RNA-recognition motif (RRM) and Zinc-finger motif in common (Lago et al., 2004).

TAFs are categorized as components of general transcription factors but some *Arabidopsis* TAFs are known to be involved in certain developmental processes, such as reproductive development. For example, TAF1, a histone acetyltransferase, is required for normal flower development, fertility, and the response to DNA damage stress (Waterworth et al., 2015). On the other hand, TAF6

specifically affects pollen tube growth, suggesting that it regulates only a specific subset of genes (Lago et al., 2005). Furthermore, it has been reported that TAF14 interacts with the FRIGIDA complex, which represses flowering by transcriptionally activating *FLC,* a floral repressor in *Arabidopsis* (Choi et al., 2011).

Recently, *TAF15b* has been reported to play a role in plant immunity in *Arabidopsis* (Dong et al., 2016). During the analysis of *MOS11*, a gene involved in plant immunity and mRNA export, TAF15b was identified as a homologue of human FUS, which is a component of a protein complex with CIP29, a human homologue of MOS11. The mutation of *TAF15b* caused an autoimmune phenotype as expected, but the mutations of the other homologues, such as *MOS11* and *DDX39*, did not show such a phenotype. In addition, TAF15b localizes not only to the nucleus but also to processing bodies (p-bodies), cytoplasmic foci involved in mRNA decapping and degradation, suggesting dual functions of TAF15b in both nucleus and cytoplasm (Dong et al., 2016). It is also reported that the *taf15b* mutant has a late flowering phenotype, although a mutant of the close homologue, *taf15*, does not show any flowering phenotype.

In humans, there is only one TAF15, which is more similar to *Arabidopsis* TAF15b, according to domain analysis. Human TAF15 belongs to the FET protein family, which are structurally categorized as heterogeneous nuclear ribonucleoprotein particle (hnRNP) proteins (Schwartz et al., 2015). The FET family

consists of three proteins, FUS (FUSED IN SARCOMA), EWSR1 (EWING'S SARCOMA BREAKPOINT REGION 1), and TAF15 and it is known that mutations in these genes cause neurodegenerative disorders (Kovar, 2011; Schwartz et al., 2015; Svetoni et al., 2016). FET proteins share similar domains, such as an N-terminal low-complexity (LC) domain (also referred to as a QGSY-rich region), a zinc-finger motif, an RRM, and a C-terminal proline-tyrosine nuclear localization signal (PY-NLS) (Svetoni et al., 2016). Because FET proteins have both zinc-finger and RRM, they can bind both RNA and DNA (Bertolotti et al., 1996; Tan and Manley, 2009). Schwartz et al. (2014) nicely presented the molecular functions of FET proteins in his review: they involved in the regulation of transcription, RNA-processing (splicing and polyadenylation) and mRNA transport.

In this study, I report the biological function of TAF15b on flowering time regulation and its protein localization at tissue and cell levels. Mutant of TAF15b flowers late and belongs to the autonomous flowering pathway. *FLC* transcript was reduced in *taf15b* mutant and the *flc taf15b* double mutant completely offsets the late flowering of *taf15b*, indicating that *TAF15b* is epistatic to *FLC* in the control of flowering. TAF15b is expressed in actively dividing tissues and it mainly locates in the nuclei. Similar with human FET proteins, TAF15b protein response to heat stress and forms unidentified nucleoplasmic foci and processing bodies (p-bodies).

2. Results

2.1 *taf15b* mutants show a late-flowering phenotype

Many genes involved in the autonomous pathway for flowering encode a protein with an RRM. Therefore, I searched for additional AP gene candidates using an *Arabidopsis* database for proteins with such domains. Two *taf15b* mutants (SAIL_35_B06, SALK_061974) were obtained from the Arabipdopsis Biological Resource Center (ABRC), with T-DNA insertions in the gene encoding protein with RRM domain (Figure 2A). RT-PCR analysis showed that both mutants do not express detectable levels of *TAF15b*, indicating that they are null mutants (Figure 2B).

In long-day conditions, both alleles flowered later than the wild type (Col-0) and produced twice as many leaves as wild type (WT) before bolting (Figure 3). The SAIL line flowered a little later than the SALK line and, thus, I termed the SAIL line as *taf15b-1* and the SALK line as *taf15b-2*, based on the severity of the flowering phenotype (Figure 3). Introduction of genomic *TAF15b* fused with a C-terminal GFP tag into the *taf15b-1* mutant rescued the late flowering phenotype (Figure 4A). Similarly, overexpression of *TAF15b* driven by the *CsV* promoter from cassava vein mosaic virus caused earlier flowering than the WT (Figure 4B,Verdaguer et al., 1996). These complementation analyses confirmed that the late-flowering phenotype is

caused by loss-of-function of the *TAF15b* gene.

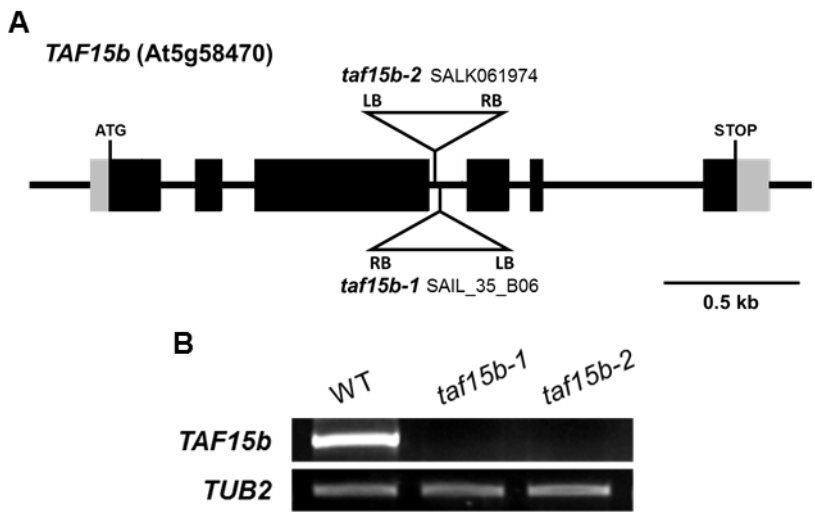

Figure 2. *TAF15b* gene structure and mutant alleles

(A) Schematic illustration of the *TAF15b* locus. The positions of the T-DNA insertion are indicated by triangles. Gray boxes indicate the untranslated region and black boxes indicate exons. Thin lines between exons represent introns.

(B) Transcript levels of *TAF15b* in WT and *taf15b* analyzed by RT-PCR analysis. *TUBULIN2* (*TUB2*) was used as a loading control.

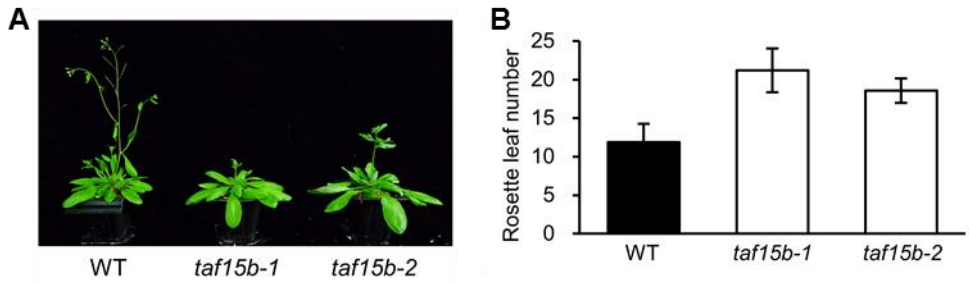

Figure 3. *taf15b* mutants flower late

(A) Late-flowering phenotypes of *taf15b* mutants in long days.

(B) Flowering time measured in terms of rosette leaf number.

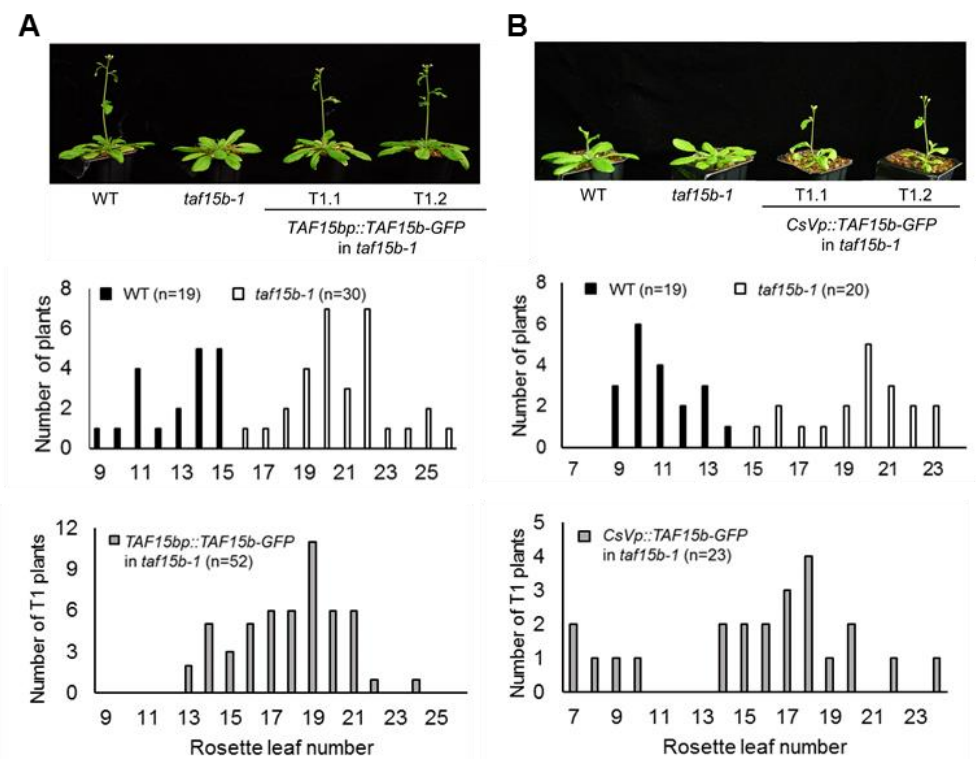

Figure 4. Complementation analysis of *taf15b* mutant with *TAF15b*

(A-B) *taf15b-1* was transformed with genomic *TAF15b* fused with GFP (*TAF15bp::TAF15b-GFP*) (A) or overexpressed *TAF15b* CDS fused with GFP (*CsVp::TAF15b-GFP*) (B). The flowering time of transgenic T1 plants was compared with that of WT and *taf15b-1* in long days.

2.2 Protein structure and localization of TAF15b

In the *Arabidopsis* genome, there are two *TAF15*s, *TAF15* and *TAF15b*, which have RRM and zinc-finger motifs (Lago et al., 2004). When comparing overall protein structures, TAF15b, rather than AtTAF15, seems to be more similar to human TAF15 because they both have an low complexity (LC) domain and a PY motif, which are absent in AtTAF15 (Figure 5). One of the interesting features of TAF15b is that it has an LC domain, which has diverse molecular functions. This domain is composed predominantly of only four amino acids (Q, G, S, and Y) arranged in [S/G]Y[S/G] repeats . Owing to its simple amino acid composition, it is named as 'low complexity' domain (Burke et al., 2015). Because of their self-assembly properties and polar amino acid residues, LC domains are also referred to as prion-like domains (PrDs) (Burke et al., 2015).

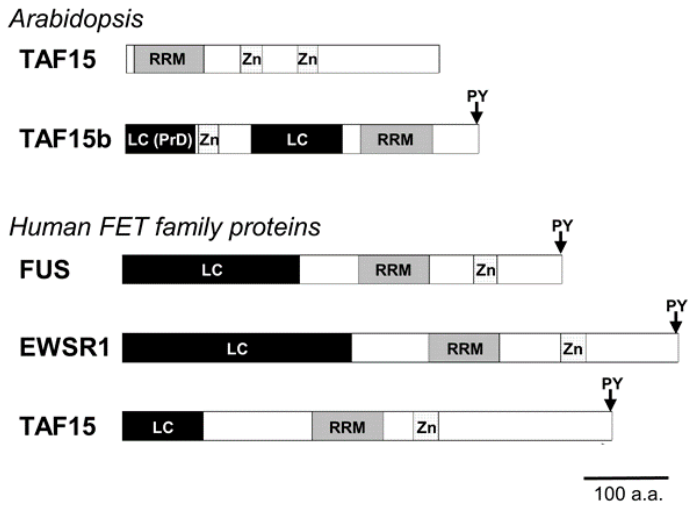

Figure 5. Domain structures in the TAF15b protein

Domain structures in predicted proteins were drawn using the BLAST program from the NCBI. PrD, predicted prion-like domain; LC, low-complexity domain; Zn, zinc finger domain; RRM, RNA recognition motif; PY, proline-tyrosine nuclear localization signal.

To understand the spatial expression pattern of *TAF15b* at the tissue level, I performed histochemical GUS staining using a 1-kb *TAF15b* promoter fused to the *β-glucuronidase* (GUS) reporter gene (Figure 6). In young seedlings, *TAF15b* is broadly expressed in entire leaves and shoot apical meristems, which are actively dividing tissues. In addition, *TAF15bp::GUS* is expressed in the veins of mature leaves and is expressed in anthers and stigmas of flowers.

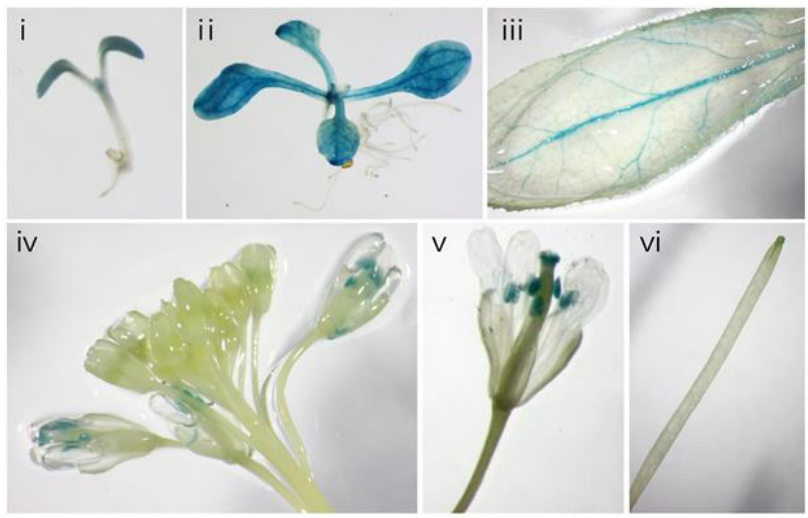

Figure 6. Tissue expression pattern of TAF15b

GUS staining of *TAF15bp::GUS* transgenic plants: i, cotyledons; ii, young seedling; iii, mature rosette leaf; iv, inflorescence; v, flower; vi, silique.

I also analyzed the subcellular distribution of TAF15b in stable transgenic plants expressing *CsVp::TAF15b-GFP*. TAF15b was mainly located in the nuclei and cytoplasm (Figure 7). In the nuclei, TAF15b formed nucleoplasmic foci but it was excluded from the nucleoli, similar to the distribution pattern of human FUS and TAF15 (Marko et al., 2012; Yang et al., 2014). Dong *et al*. (2016) reported that *Arabidopsis* TAF15b is also localized to cytoplasmic ribonucleoprotein granules known as p-bodies. Since it is known that heat stress increases the size of p-bodies in budding yeast (Brengues et al., 2005; Grousl et al., 2009), I checked if TAF15b is also localized to cytoplasmic p-bodies after heat stress. At room temperature, GFP-

fluorescent cytoplasmic p-bodies were undetectable, although green fluorescent nuclei were easily observed in *CsVp::TAF15b-GFP* plants (Figure 7A, B). In contrast, a 42°C heat stress for 1 h caused the appearance of green fluorescent spots in the cytoplasm, which are likely to be p-bodies. In addition, heat stress caused an increase in the number of foci in the nucleus (Figure 7B). Interestingly, a 1 h cold treatment also increased the number of nucleoplasmic foci, although such foci disappeared after vernalization, a 43-day cold treatment. However, cold treatment did not cause the appearance of fluorescent p-bodies. Taken together, these results indicate that subcellular localization of TAF15b is influenced by temperature.

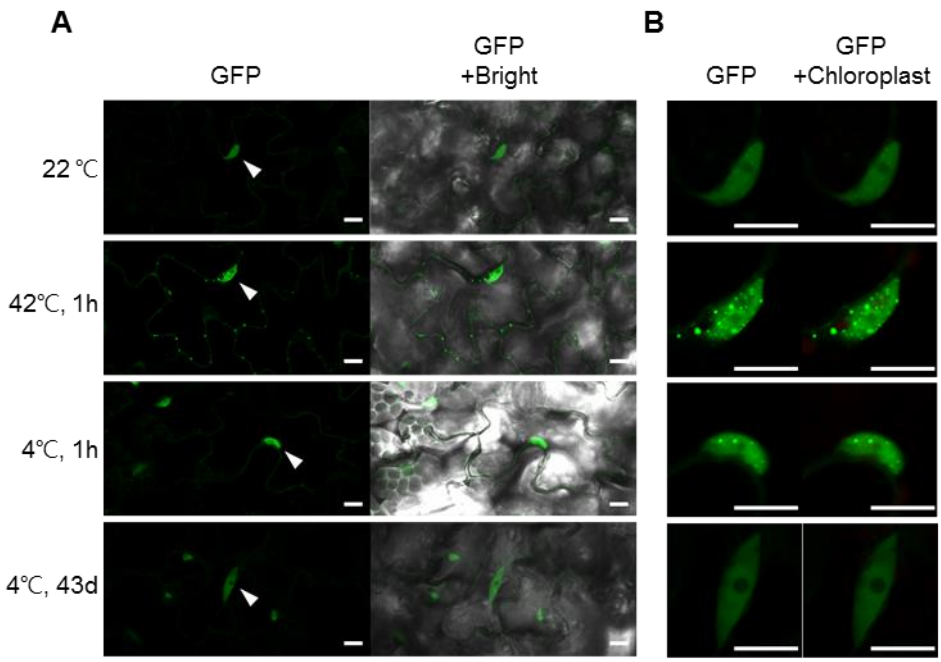

Figure 7. Subcellular localization of TAF15b

(A-B) Subcellular localization of TAF15b-GFP after cold and heat treatments showing entire cells (A) and nuclei (d). *CsVp::TAF15b-GFP* transgenic lines grown in long days (22°C) were treated with heat (42°C) or cold (4°C) for 1 h or 43 days. Then, the upper surfaces of rosette leaves were observed under a confocal microscope. Arrow heads indicate nuclei and the scale bar represents 10 μm.

It is known that LC and prion-like domains induce the formation of membraneless organelles, such as the nucleoli, Cajal bodies, and splicing speckles in the nucleus (March et al., 2016). Therefore, I assumed that LC domain of TAF15b might be responsible for the formation of nucleoplasmic foci. CDS of full length or LC domain deleted TAF15b was transfected into protoplast, which is a leaf cell

without cell wall. Size of nucleoplasmic foci got bigger in protoplasts and the foci disappeared when LC domain deleted TAF15b were transfected (Figure 8). From this result, I could identify that LC domain of TAF15b functions in the formation of nucleoplasmic foci.

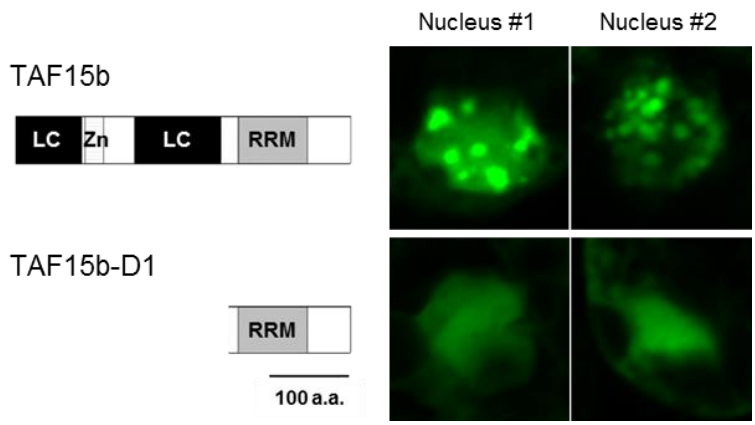

Figure 8. LC domain induces the formation of membraneless organelles in the nucleus *CsVp::TAF15b-GFP* constructs with or without LC domain were transfected into *Arabidopsis* protoplasts and two representative nuclei of each construct were presented. Images were visualized and obtained using confocal microscopy at 19 h after transformation.

2.3 TAF15b is involved in the autonomous pathway for flowering

To address if *TAF15b* is indeed an AP gene, I analyzed the flowering phenotype of *taf15b* under various environmental conditions. The *taf15b* mutants flowered later than the WT in both long and short days and they also showed responses to both

vernalization and GA treatment (Figure 9). Such flowering traits are characteristic of AP mutants.

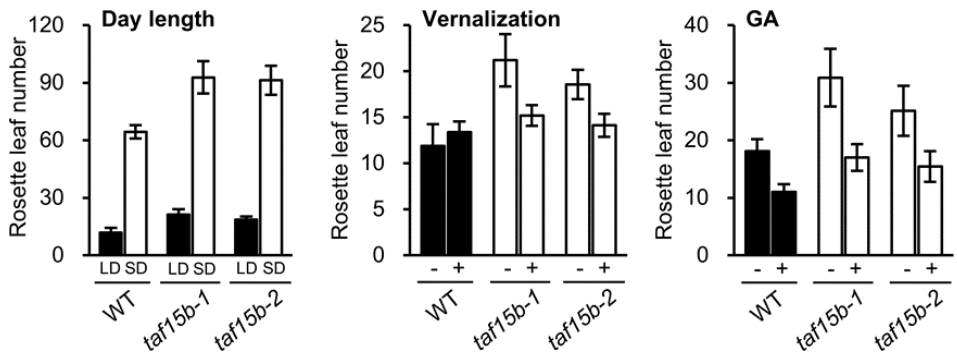

Figure 9. The flowering behavior of *taf15b* mutant in various environmental conditions
Plants were grown in long days (LD) and short days (SD). For the vernalization response, plants were grown in long days with (+) or without (-) 50 days of vernalization treatment. For GA treatment, 10 μM GA_3 was sprayed every three days and for the mock treatment, 0.1% ethanol was sprayed.

In addition, similar to other AP mutants, *taf15b* showed a large increase in *FLC* transcript levels, with the level ten-fold higher in *taf15b-1* and six-fold higher in *taf15b-2* compared to the WT (Figure 10A). *FLC* is a strong repressor of flowering in *Arabidopsis* and suppresses the flowering pathway integrators, *FT* and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*) and *LEAFY* (*LFY*) (Helliwell et al., 2006; Hepworth, 2002; Lee and Lee, 2010; Michaels, 2009). Thus, I checked if expressions of the three integrators were reduced in the *taf15b* mutant.

Indeed, the transcript levels of all of the integrators were decreased in *taf15b* compared to the WT (Figure 10A), which explains why *taf15b* has late flowering. Then, I checked if *taf15b* affects the expression of any other AP genes or vice versa, thus determining if *TAF15b* acts upstream or downstream of AP genes. As shown in Figure 10B, none of the analyzed AP genes showed significant differences in expression between *taf15b* and the WT. Similarly, none of the AP mutants showed significantly different expression of *TAF15b* compared to the WT (Figure 10C). Taken together, my results suggest that *TAF15b* acts neither upstream nor downstream of well-known AP genes.

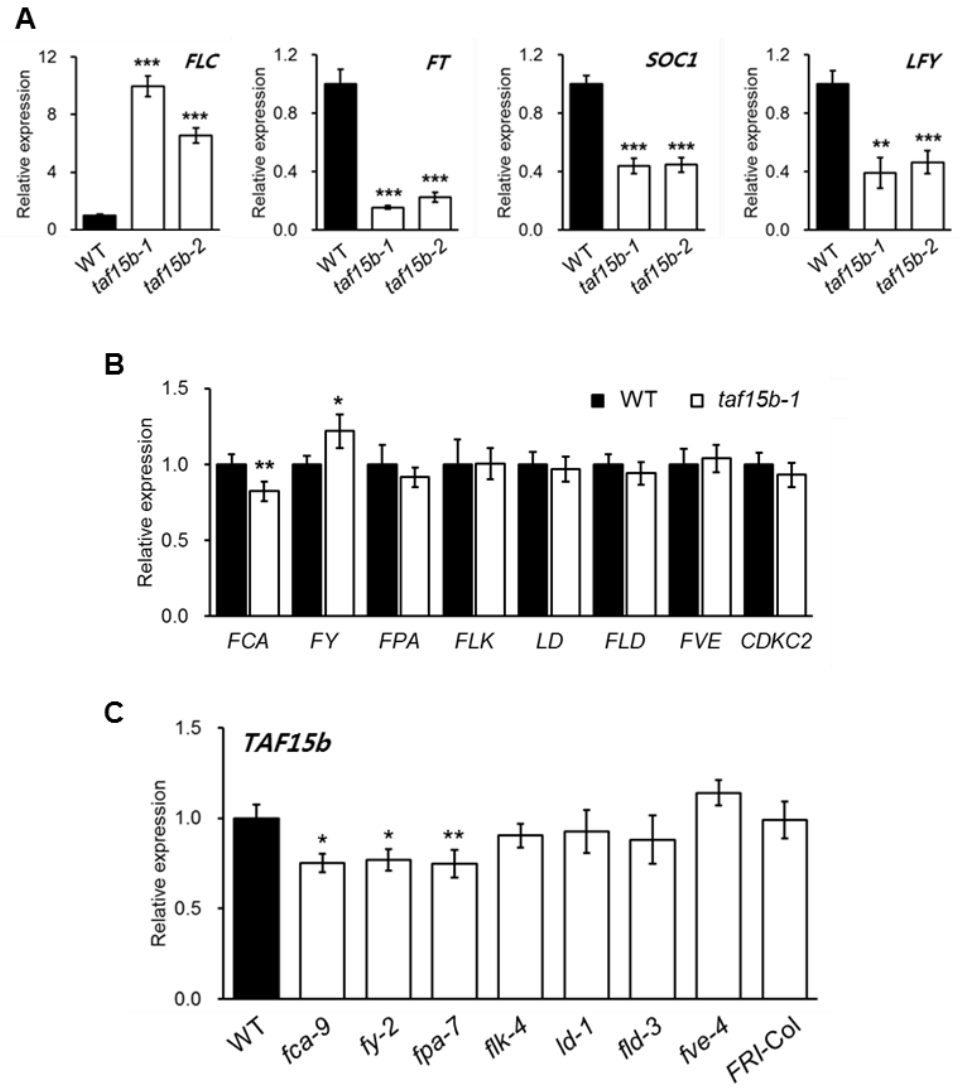

Figure 10. Transcript levels of flowering time and autonomous pathway genes

(A) Transcript levels of flowering time genes in seedlings of WT and *taf15b-1*.

(B) Expression levels of various autonomous pathway genes in WT and *taf15b-1* seedlings.

(C) Expression level of *TAF15b* in AP mutants. Relative transcript levels to *TUB2* are normalized to WT. Means ± SD from three technical replicates are presented. Asterisks indicate statistically significant differences evaluated using Student's *t*-test (* $p<0.05$, ** $p<0.01$, *** $p<0.001$).

Finally, the late-flowering phenotype of *taf15b* was completely offset by the *flc* mutation; thus, the *flc taf15b* double mutant showed early flowering similar to the *flc* single mutant (Figure 11). Altogether, my results suggest that *TAF15b* promotes flowering by repressing *FLC* expression via the autonomous pathway.

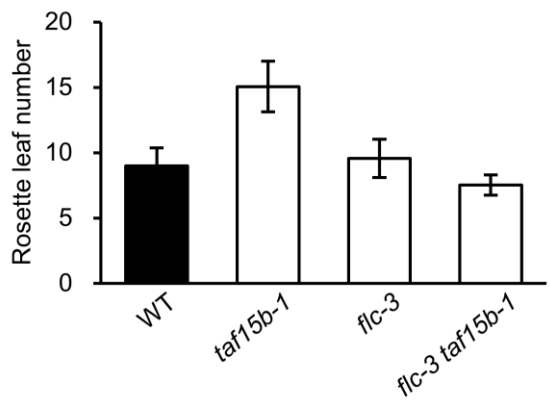

Figure 11. Flowering time of *taf15b flc* double mutant

Plants were grown in long days and at least 16 plants were used to measure flowering time. Error bars represent the SD.

3. Discussion

In this chapter, I discovered *TAF15b* as a new gene involved in the AP, and characterized its protein localization at both the tissue and cellular level. *taf15b* mutants flowered late in both short and long day conditions and this phenomena could be explained by the increase of *FLC* transcript levels in the mutant. Comparing with other AP mutants, *taf15b* showed weak flowering phenotype. TAF15b might have other genes with functional redundancy or simply TAF15b might regulate *FLC* transcription with different mechanisms with AP. This should be confirmed through genetic study between *taf15b* and other AP mutants.

TAF15b showed diverse subcellular localizations in plant cells. It has been figured out that TAF15b is localized to both the nucleus and the processing bodies (p-bodies) in the cytoplasm (Dong et al., 2016). P-bodies are cytoplasmic RNA granules, comprised of mRNAs and components of the mRNA decay machinery, and is involved in the degradation or aggregation of target mRNAs for preventing translation (Anderson and Kedersha, 2009). Stresses or external stimuli that cause translational inhibition, induce the formation of p-bodies and stress granules, which are another type of cytoplasmic RNA granules containing components of the translation initiation machinery instead of components of the mRNA decay machinery (Decker and Parker, 2012). These RNA granules are assembled into membraneless organelles, which form organelle structures that are not lipid-bound.

Examples of membraneless organelles include the nucleoli, Cajal bodies, and splicing speckles in the nucleus (March et al., 2016). Instead of lipids, the LC and prion-like domains of proteins mediate self-aggregation, which induces the formation of membraneless organelles. TAF15b also has an LC domain at its N-terminal region and this domain might be responsible for the formation of p-bodies.

Human FET proteins are also localized in other cytoplasmic assemblies, including stress granules that are generated in response to heat shock and oxidative stress, implying they participate in translational regulation against stress response (Andersson et al., 2008). I therefore speculated whether heat stress could also induce the translocation of TAF15b to the cytoplasmic assemblies in plant. Similar to animal cells, TAF15b is assembled into cytoplasmic granules and unidentified nuclear foci upon heat treatment. Whether such changes in localization are related to the transcription of *FLC* could not be verified; however, it suggests that TAF15b might be involved in response to heat stress and translational repression of other genes.

CHAPTER III

Molecular function of TAF15b on the regulation of *FLC* transcription

1. Introduction

Although many AP genes have been identified and characterized, none of them has been shown to directly regulate either *FLC* transcription or *FLC* RNA processing. The current understanding of how AP genes affect transcript levels of *FLC* is concentrated on the regulation of the noncoding antisense RNA *COOLAIR* at the *FLC* locus (Liu et al., 2007, 2010; Marquardt et al., 2014; Wang et al., 2014). For example, FY, a homologue of a component of a complex required for mRNA cleavage and polyadenylation in yeast interacts with FCA, a plant-specific protein (Simpson et al., 2003). But FCA, together with FY, autoregulates alternative polyadenylation of its own RNA instead of *FLC* RNA (Quesada et al., 2003; Simpson et al., 2003). Interestingly, FCA and FY regulate RNA processing of *COOLAIR*, especially the proximally polyadenylated *COOLAIR* (Liu et al., 2007, 2010). Similarly, FPA, a Spen family protein, also controls alternative cleavage and polyadenylation of its own RNA and regulates the transcript level of distal form of *COOLAIR* (Hornyik et al., 2010). In addition, CstF64 and CstF77 are also required for 3' processing of *COOLAIR* (Liu et al., 2010). Such processing is linked to *FLC* chromatin structure through the genetic interaction of FCA and FLD, an H3K4 demethylase, which causes an inactive form of *FLC* (Liu et al., 2007). Interestingly, CDKC;2 also affects *COOLAIR* transcription, although it does not directly affect *FLC* transcription (Wang et al., 2014). These studies suggest that AP genes affect

FLC transcription through modification of chromatin structure mediated by *COOLAIR*. Therefore, the mechanism of the direct regulation of *FLC* transcription is still unknown.

In the previous chapter, I suggested TAF15b as a new autonomous pathway gene. Because *Arabidopsis* TAF15b has homologous genes in other organisms, the molecular mechanism of *FLC* regulation could be assumed from the earlier research. Human TAF15 belongs to the FET protein family and these proteins are reported to regulate both transcription in the nucleus and RNA processing in cytoplasmic stress granules (Schwartz et al., 2015). For transcriptional regulation, FUS is reported to bind Pol II at the transcription start sites (TSSs) of target genes and inhibit the phosphorylation of Ser2 (Ser2P) of the c-terminal domain (CTD) of Pol II (Schwartz et al., 2012). Because LC domains have self-assembly properties, FUS proteins form polymeric fibers around TSSs and directly bind the CTD of Pol II, which inhibits transcriptional elongation in an RNA-dependent manner (Schwartz et al., 2012, 2013). This leaded me to hypothesize that TAF15b might also regulate *FLC* transcription through the interaction with Pol II.

Here, I report the molecular function of a plant FET protein, TAF15b on the regulation of *FLC* sense and antisense transcript. TAF15b plays a role in decreasing not only *FLC* sense transcription, but also antisense transcription of *COOLAIR*. I show that TAF15b interacts with Pol II and is associated with the TSSs

of both sense and antisense *FLC*. Such results suggest that TAF15b is directly involved in the transcriptional repression of *FLC*.

2. Result

2.1 TAF15b is enriched near the TSSs of *FLC* transcript

The increase of *FLC* transcript levels in the *taf15b* mutant indicates that TAF15b may affect *FLC* at the transcriptional level. Thus, I checked the level of H3K4me3, which is associated with active transcription, by ChIP analysis in *taf15b-1* and compared it to the WT. Similar to previous studies (Wu et al., 2016), H3K4me3 was highly enriched at the *FLC* transcription start site (TSS) in the WT (B region in Figure 12A). In *taf15b*, H3K4me3 enrichment was increased especially at the C region just below the first exon (Figure 12B). This suggests that the *FLC* chromatin of the *taf15b* mutant is in a transcriptionally active state.

As TAF15b has a zinc-finger motif and is a homologue of FET family proteins, which were suggested to affect transcription (Kovar, 2011; Schwartz et al., 2015), I checked if TAF15b directly regulates *FLC* expression. For this, I performed ChIP analysis using *CsVp::TAF15b-GFP* seedlings and identified the enrichment pattern of TAF15b-GFP protein on the *FLC* gene (Figure 13). Interestingly, TAF15b showed enrichment around the TSSs of both *FLC* sense and antisense transcripts (A, B regions, and F region in Figure 13A).

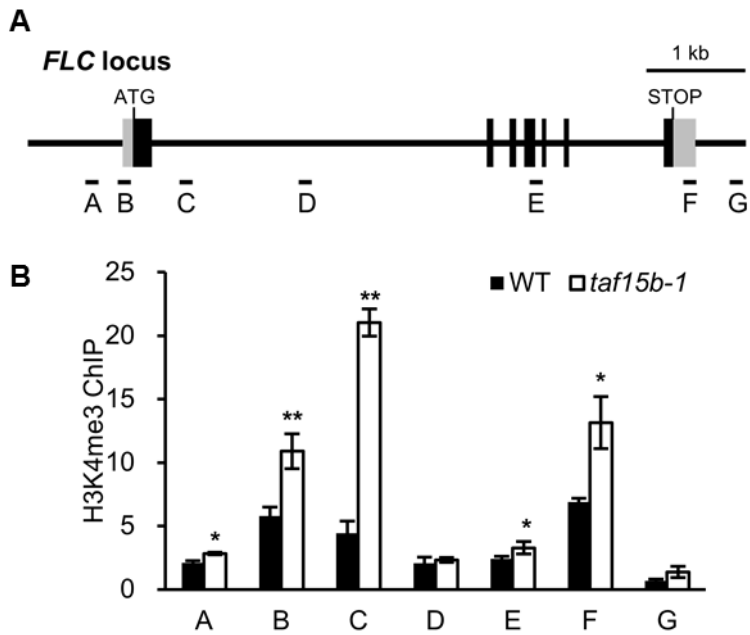

Figure 12. TAF15b affects H3K4me3 level at the *FLC* locus

(A) Schematic diagram of the *FLC* gene indicating the regions (A-G) analyzed by ChIP and real-time qPCR (ChIP–qPCR).

(B) ChIP-qPCR analysis for H3K4me3 levels at the *FLC* locus of WT and *taf15b-1*. Seedlings were grown for 10 days in long days for ChIP analysis. Relative enrichment of the IP/Input was normalized to that of *Ta3*. Means ± SD from three technical replicates are presented. Asterisks indicate statistically significant differences evaluated using Student's *t*-test (* $p<0.05$, ** $p<0.01$).

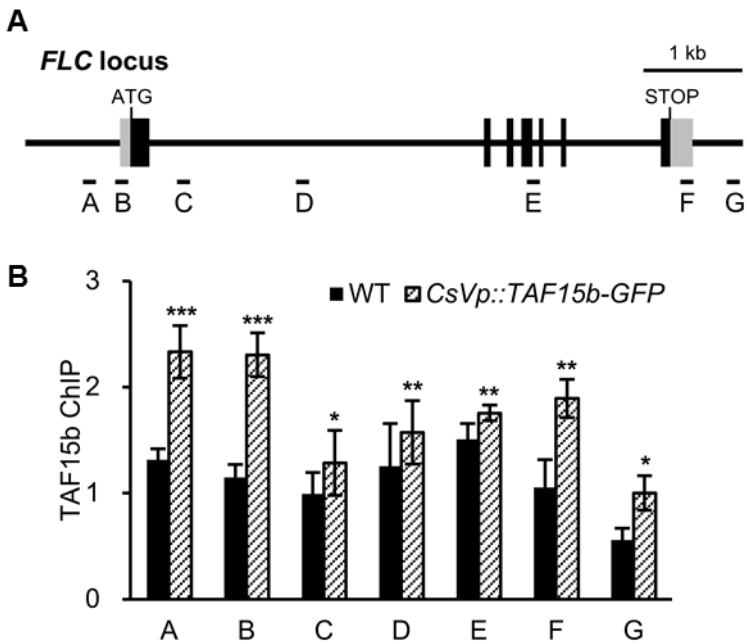

Figure 13. TAF15b binds to the *FLC* locus

(A) Schematic diagram of the *FLC* gene indicating the regions (A-G) analyzed by ChIP–qPCR.

(c) ChIP-qPCR analysis for TAF15b enrichment at the *FLC* locus. WT and epitope-tagged transgenic lines, *CsVp::TAF15b-GFP*, were grown for 10 days in long days for ChIP analysis. Relative enrichment of the IP/Input was normalized to that of *Ta3*. Means ± SD from three technical replicates are presented. Asterisks indicate statistically significant differences evaluated using Student's *t*-test (* $p<0.05$, ** $p<0.01$, *** $p<0.001$).

2.2 TAF15b decreases *FLC* and *COOLAIR* transcripts

The *FLC* locus produces not only a sense transcript but also the antisense transcript *COOLAIR*. Therefore, I analysed the levels of both sense and antisense transcripts in

WT, *taf15b-1*, *fca-9*, which is an intensively studied AP gene, and *FRI*-Col, which shows strong expression of *FLC* (Figure 14). For sense transcripts, I analysed both spliced and unspliced forms of *FLC* because the spliced form represents the level of mRNA, whereas the unspliced form is used as a proxy for transcriptional activity (Swiezewski et al., 2009). In *taf15b-1*, both spliced and unspliced forms of *FLC* were increased compared to the WT (Figure 14B). There are two forms of *FLC* antisense transcripts, the proximal and distal *COOLAIR*s (Liu et al., 2007; Swiezewski et al., 2009). The levels of these two transcripts were also higher in *taf15b* than in the WT, which is similar to those in *fca-9* and *FRI*-Col (Figure 14C). Such results strongly suggest that TAF15b represses the transcription of both sense and antisense *FLC*.

FLC sense and antisense transcripts differentially respond to vernalization: sense transcripts are decreased whereas antisense transcripts are increased, during vernalization (Csorba et al., 2014; Czesnick and Lenhard, 2016; Swiezewski et al., 2009). Because *COOLAIR* is increased to the highest level after 2 weeks of vernalization, I compared the transcript levels at that time (Figure 14B, C). Spliced forms of *FLC* were slightly decreased in *fca-9* and *FRI*-Col, but were slightly increased in WT and *taf15b-1*, which is similar to previous studies showing that *FLC* mRNA is not decreased significantly after 2 weeks of vernalization treatment, probably due to its long half-life (Csorba et al., 2014). However, 6 weeks of vernalization treatment reduced the level of spliced form dramatically in *taf15b*

similar to *fca* and *FRI*-Col, suggesting that *taf15b* responds to vernalization as well as AP mutants (Figure 14B). On the other hand, unspliced *FLC* was significantly decreased by 2 weeks of cold exposure in all genotypes (Figure 14B). In the case of *COOLAIR*, the level of the proximal form was highly increased but the distal form was not much affected by 2 weeks of cold treatment in all genotypes (Figure 14C). Therefore, these results indicate that there is not much differential response to vernalization among *taf15b* and other genotypes, such as *fca* and *FRI*-Col, and the levels of *FLC* sense and *COOLAIR* transcripts were affected oppositely by vernalization.

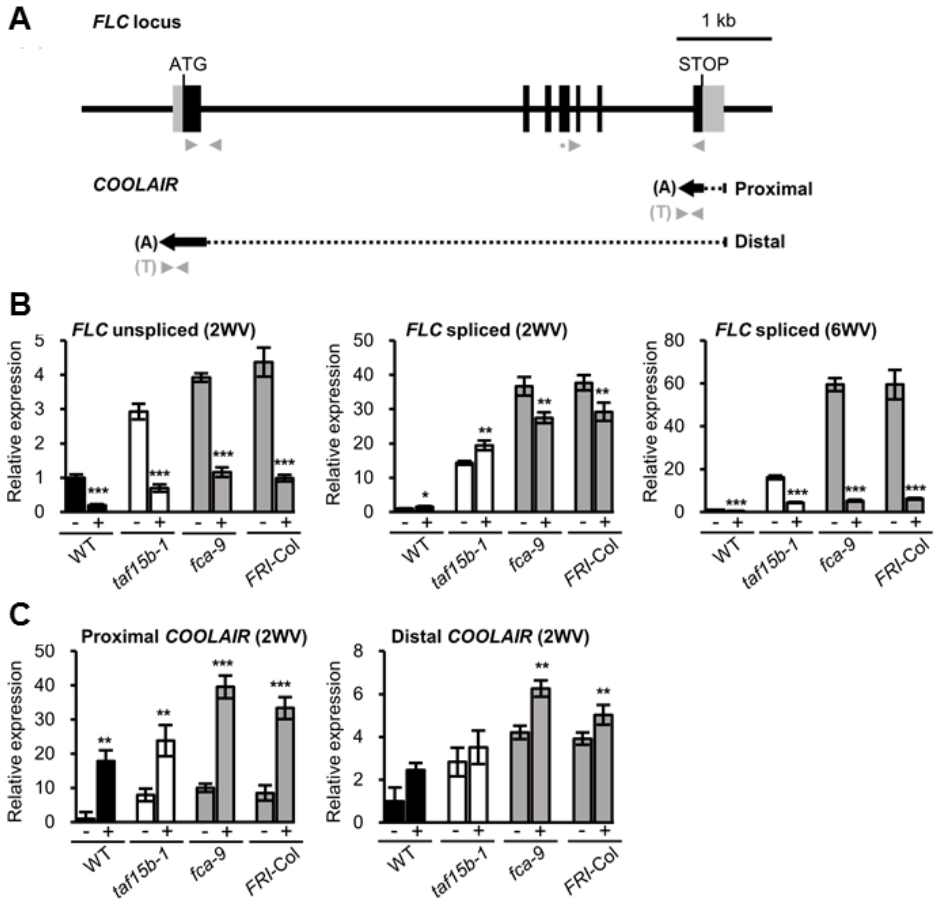

Figure 14. Transcripts of *FLC* and *COOLAIR* are increased in *taf15b*

(A) Schematic diagram of the FLC gene and two forms of the antisense transcript, COOLAIR. Poly A sequences are denoted by (A) and gray arrowheads indicate the primers used in the qRT-PCR analysis. Primers denoted by (T) show poly T sequence primers.

(B, C) qRT-PCR analysis of *FLC* (B) and *COOLAIR* (C) transcripts before vernalization (-), after 2 or 6 weeks of vernalization (+). Non-vernalized samples (-) are seedlings grown for ten days in long days and vernalized samples (+) are seedlings grown for seven days in long days, then vernalized for 2 or 6 weeks. Relative transcript levels to *TUB2* are normalized to non-vernalized WT. Means ± SD from three technical replicates are presented. Asterisks indicate statistically significant differences evaluated using Student's *t*-test (* $p<0.05$, ** $p<0.01$, *** $p<0.001$).

2.3 TAF15b interacts with RNA Polymerase II

TAF15b is enriched at the TSSs of the *FLC* sense and antisense transcripts and is a homologue of a general transcription factor (Lago et al., 2004). It prompted me to check if TAF15b interacts with Pol II by co-immunoprecipitation (co-IP) using *CsVp::TAF15b-GFP* transgenic lines. I used three different antibodies specific for the CTD of Pol II for co-immunoprecipitation (co-IP): an antibody against hypophosphorylated CTD (hypo), an antibody against phosphorylated serine 2 in the CTD (Ser2P), and an antibody against phosphorylated serine 5 in the CTD (Ser5P). The co-IP analysis showed that all types of the CTD were co-immunoprecipitated with TAF15b-GFP, although the Ser2P form showed a stronger signal (Figure 15A). This indicates that TAF15b interacts with Pol II in general. Then, I examined if *taf15b* affects the global phosphorylation status of the Pol II CTD by western blot analysis. The result showed that there were no significant differences in the levels of Pol II and the phosphorylation status of both Ser2P and Ser5P between the WT and the *taf15b-1* mutant (Figure 15B). This result indicates that TAF15b does not affect the global phosphorylation status of the Pol II CTD.

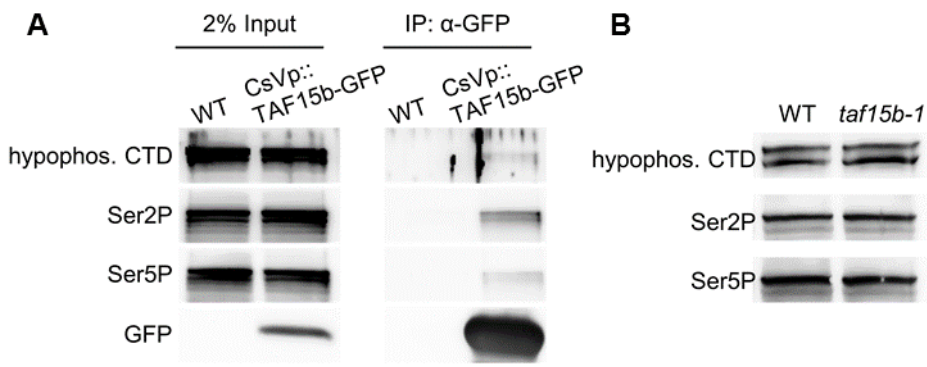

Figure 15. TAF15b interacts with Pol II

(A) TAF15b interacts with Pol II based on phosphorylation status. Total proteins from *CsVp::TAF15b-GFP* seedlings were immunoprecipitated with GFP-trap and detected by Pol II specific antibodies (hypophosphorylated CTD (hypophos), Ser2P, and Ser5P in the CTD). (B) Western blots analyzing proteins from WT and *taf15b*-1 to compare the global levels of Pol II isoforms. Total protein (120 μg) from each line was used for this experiment.

Transcriptional initiation, elongation, and termination are regulated through the posttranslational modification of the CTD of Pol II such that Pol II initiates transcription when Serine 5 of CTD is phosphorylated (Ser5P), elongates transcription when both Ser2 and Ser5 are phosphorylated, and terminates transcription when only Ser2 is phosphorylated (Harlen and Churchman, 2017; Phatnani and Greenleaf, 2006; Zaborowska et al., 2016). Thus, I analyzed if the enrichment of Ser2P or Ser5P of Pol II is affected by the *taf15b* mutation, which would be indicative of direct participation in transcriptional processes. The ChIP

analysis using antibodies specific for Ser2P or Ser5P showed a moderate increase in the enrichment of both Ser2P and Ser5P, indicative of increased transcription elongation, in the *taf15b* mutant compared to the WT throughout the *FLC* locus (Figure 16). In contrast, at the *ACTIN7* locus, whose expression is not influenced by *TAF15b*, there was no difference in the enrichment of Ser2P or Ser5P between *taf15b* and the WT (Figure 17). Taken together, these results suggest that TAF15b interacts with Pol II and decreases transcriptional elongation of *FLC*.

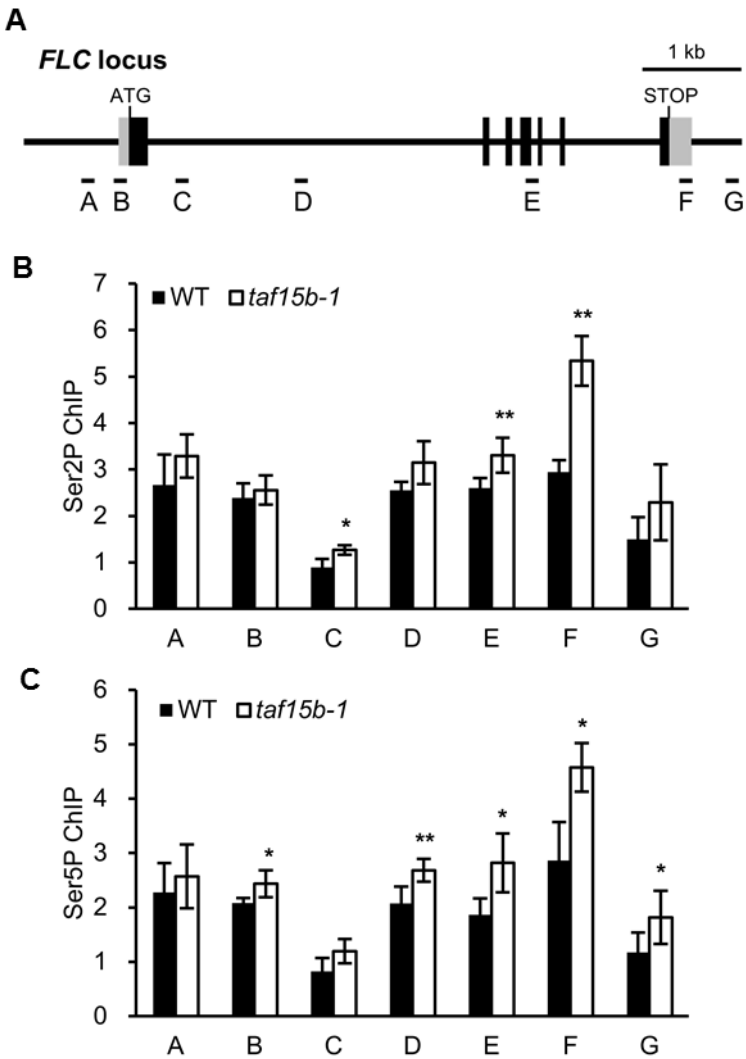

Figure 16. Loss of TAF15b leads to accumulations of Ser2 and Ser5 phosphorylated Pol II

(A) Schematic diagram of the *FLC* gene indicating the regions (A-G) analyzed by ChIP–qPCR.

(B-C) ChIP analysis showing the enrichment of Ser2P (B) and Ser5P (C) of the CTD at the *FLC* locus in WT and *taf15b-1*. Relative enrichment of the IP/Input was normalized to that of *Ta3*. Means ± SD from three technical replicates are presented. Asterisks indicate statistically significant differences evaluated using Student's *t*-test (* $p<0.05$, ** $p<0.01$).

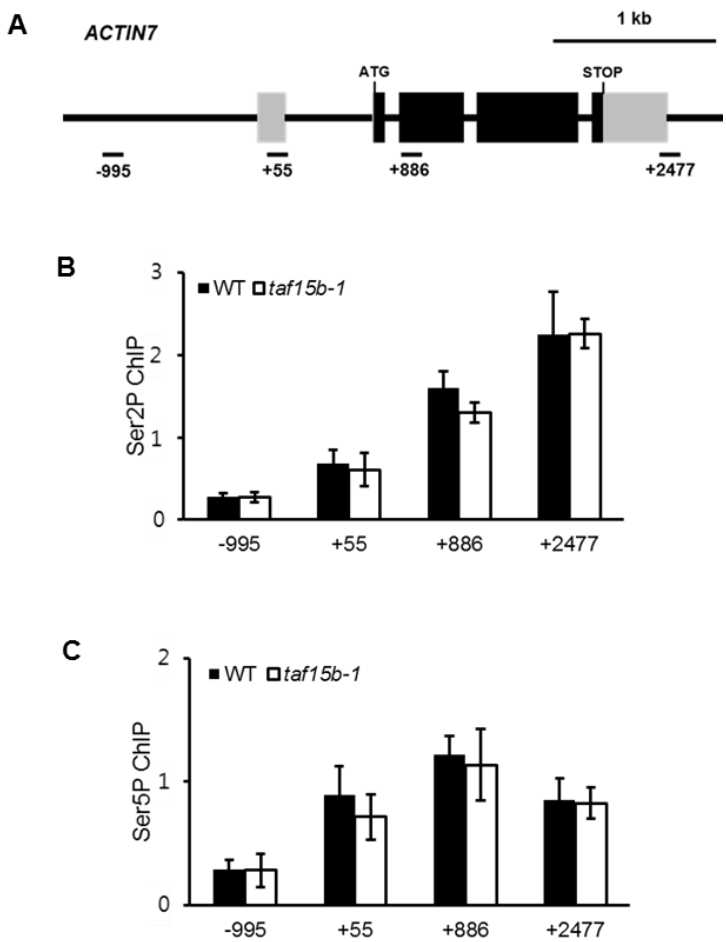

Figure 17. Pol II ChIP analysis at the *ACTIN7* locus

(A) Schematic diagram of the *ACT7* gene indicating the TSS regions analyzed by ChIP-qPCR. Primers used in this assay were those reported by Wu et al. (2016).

(B-C) ChIP assays showing the enrichment of Ser2P (B) and Ser5P (C) of the CTD at the *ACT7* locus in WT and *taf15b-1*. Relative enrichment of the IP/Input was normalized to that of *Ta3*. Means ± SD from three technical replicates are presented. Statistical significance was evaluated using Student's *t*-test but there was no significant difference between the indicated genotypes.

3. Discussion

In general, TAFs are components of the TFIID complex and, thus, are usually considered to function to bring Pol II to a core promoter (Burley and Roeder, 1996; Hampsey and Reinberg, 1997). However, it is known that some TAFs have distinctive roles in developmental processes in various organisms, including *Arabidopsis* (Lago et al., 2005; Goodrich and Tjian, 2010; Choi et al., 2011; Waterworth et al., 2015). Here, I report that TAF15b, which is most likely an orthologue of human TAF15, affects flowering time by repressing transcription of *FLC*. The interaction with Pol II and enrichment at the TSSs of the *FLC* gene strongly support that TAF15b is a transcriptional repressor of *FLC*.

My results reveal that *TAF15b* is a *bona fide* AP gene whose encoded protein is enriched at the promoter region of *FLC* and interacts with Pol II. Although some reports showed that FLD, an AP gene encoding H3K4 demethylase, is enriched at promoter regions (Jiang et al., 2009; Ko et al., 2010), how the epigenetic modifier directly represses transcription of *FLC* has not been resolved. Interestingly, some reports showed that AP proteins are enriched at the *FLC* locus either in the middle or the rear part of the gene such that FLD is enriched in the middle of the first intron (Wu et al., 2016), in contrast to the previous reports, whereas FPA and FCA are enriched at the 3′ end regions of the *FLC* locus (Hornyik et al., 2010; Liu et al., 2007). The locations of FPA and FCA and the newly reported locations of FLD are well

correlated with the hypothesis proposing that *FLC* is suppressed by *COOLAIR*-mediated chromatin inactivation because the enriched regions are near the end of the proximal *COOLAIR* transcript. In contrast, regions enriched with TAF15b at the *FLC* locus correlate well with the hypothesis that TAF15b functions as a transcriptional repressor. Interestingly, TAF15b is enriched at the promoter regions of both sense and antisense transcripts of *FLC*. The suppression of *COOLAIR* by TAF15b may reinforce the transcriptional repression by epigenetic inactivation of *FLC* chromatin mediated by the processing of proximal *COOLAIR*. Otherwise, it may be an indirect effect because the circular structure of genes induced during transcription makes the 5′ and 3′ ends meet (Crevillén et al., 2012). Therefore, the same transcriptional repressor probably suppresses both orientations of transcriptional elongation. Future studies will answer which hypothesis is correct.

Human TAF15 belongs to the FET protein family that is likely to regulate transcription, RNA processing, and nuclear-cytoplasmic shuttling of mRNA (Schwartz et al., 2015). For transcriptional regulation, it was proposed that TAF15 connects the preinitiation and elongation complexes, based on the fact that it can bind both RNA and single-stranded DNA and constitute a functionally different TFIID complex (Bertolotti et al., 1996). However, recent studies showed that FUS binds the CTD of Pol II through the prion-like LC domain and inhibits phosphorylation of Ser2 in the CTD, which prevents transcriptional elongation

(Schwartz et al., 2012, 2013). One of the interesting features of TAF15b is that it has LC domains at the N-terminal and middle regions. Such LC domains are involved in the interaction with the CTD of Pol II in humans (Kwon et al., 2013; Schwartz et al., 2013). It has a simple, low-complexity amino acid composition, mostly glutamine, glycine, serine, and tyrosine (QGSY-rich), which are arranged in [S/G]Y[S/G] repeats (Burke et al., 2015). One of the well-known LC domain-containing peptides is the CTD of Pol II, which is composed of heptad repeats $Y^1S^2P^3T^4S^5P^6S^7$ (34 repeats in *Arabidopsis*) (Koiwa et al., 2004; Schüller and Eick, 2016). Due to the LC domains, FET proteins form higher-order structures and directly bind the CTD of Pol II (Burke et al., 2015; Kwon et al., 2013; Schwartz et al., 2013). As *Arabidopsis* TAF15b interacts with Pol II, it is very likely that the LC domains of TAF15b associate with the CTD of Pol II to inhibit transcriptional elongation, but this remains to be proven in the future.

In *Arabidopsis*, Chakrabortee *et al*. (2016) predicted that approximately 500 proteins have prion-like domains (PrDs) using a computational algorithm. The LC domain before the zinc-finger motif in TAF15b was also predicted to be a PrD (Figure 5). In their report, four out of eight AP genes, *LD, FPA, FCA* and *FY*, were suggested to encode proteins with PrDs, although only LD has prion activity (Chakrabortee et al., 2016). From the list of proteins predicted to contain PrD, I identified five more AP genes, *CstF64, GRP7, GRP8, PRP39-1,* and *CDKC;2,* that

belong to this class. Since PrDs have self-assembly properties and most of AP genes with PrDs belong to the RNA processing group, it is possible that AP proteins interact with each other. The interaction of TAF15b and Pol II and the increase of elongation mark in *taf15b* suggest that it might affect *FLC* transcription through a transcription elongation complex (TEC). A TEC is composed of a growing nascent RNA, Pol II, and accessory proteins, including factors for elongation, RNA processing, and chromatin modification (Perales and Bentley, 2009). In a TEC, the CTD of Pol II provides a platform for direct binding of RNA processing and transcription elongation factors (Phatnani and Greenleaf, 2006). Because the RNA processing group of AP genes encode proteins involved in polyadenylation, splicing, and transcriptional elongation (Rataj and Simpson, 2014; Simpson, 2004), it is possible that AP proteins may function as components of a TEC. Thus, it is of utmost importance to investigate whether any of the AP genes are involved in this process. It would also be interesting to see whether TAF15b has any roles in mRNA processing because the proteins are localized at cytoplasmic p-bodies in certain environments.

CHAPTER IV

Proteomic analysis of TAF15b complex

1. Introduction

Some of AP proteins directly interact each other and form protein complex. FCA, a protein contains two RRM and WW domain interacts directly with RNA 3' end-processing factor FY (Macknight et al., 1997; Simpson et al., 2003). Tandem affinity purification assay also showed that FY stably interacts with CPSF100 and CPSF160, which are RNA processing factors and those proteins forms a protein complex, whose size is around 500-800 kDa (Manzano et al., 2009). In addition to RNA-related AP proteins, FVE, which is involved in histone deacetylase complex, forms 1 Mda huge protein complex (Jeon and Kim, 2011).

Recently, Scott D. Michaels from Indiana university gave interesting seminar about AP protein complex (Michaels, 2016). To figure out the interacting partners with AP proteins, he performed Y2H screening with AP proteins. He found out that and figured out that FLK, FPA, LD, FLD, DCP make a protein complex with Pol II and also identified new interacting proteins. He did not revealed a name but these genes are involved in transcriptional elongation and consist of three redundant genes so only triple mutant shows late flowering phenotype. From these results, a hypothesis was suggested that AP proteins might repress *FLC* transcription through direct interaction with Pol II and regulation of Pol II transcriptional elongation.

FRIGIDA (FRI) is a strong activator of *FLC* transcription and it also forms large protein complex (Choi et al., 2011). FRI-complex consists of FRI, SUF4, FES1,

FLX and FRL1 and these proteins also interact with several chromatin modification factors and general transcription factor. Gel filtration analysis with myc-tagged FRI transgenic plants revealed that the size of FRI-complex is around 400~1000 kDa. TAF14 also interacts with five components of FRI-complex and this implies that the complex also regulate *FLC* transcription with intimate association with Pol II.

Experimental strategies to identify the interacting proteins with AP is limited to *in vitro* methods, such as Y2H screening and affinity purification assay. Results from these methods could be artificial or interactions, which happen in plants, might not be observed in yeast system. DNA and RNA composition of yeast is quite different with plant and yeast could not reflect posttranslational modification of plant. Therefore, in this chapter, I performed immunoprecipitation coupled to mass spectrometry with TAF15b-GFP transgenic plants to identify the TAF15b complex. Identifying the components of TAF15b could help to understand more detailed biochemical mechanism of TAF15b to regulate *FLC* transcription.

2. Result

2.1. TAF15b interacts with FRI-complex and AP proteins

Protein interactions between TAF14, TAF15, TAF15b and FRI-complex genes were analyzed by yeast two-hybrid (Y2H) assay (Figure 18). FES1, FLX, and TAF15 showed strong transcriptional activity of themselves because the combinations of BD-FES1, FLX, TAF15 and AD alone showed cell growth on -LWHA medium. TAF14 interacted with FRI, FRL1, SUF4, which is consistent with the previous results (Choi et al., 2011) and it also interacted with TAF14 itself. Although TAF15 and TAF15b show protein similarity, their interaction patterns with FRI-complex is somewhat different. TAF15b interacted with FRI, SUF4, and TAF14 but TAF15 did not show any interactions with those proteins. This result indicates that TAF15b also interacts with components of FRI-complex.

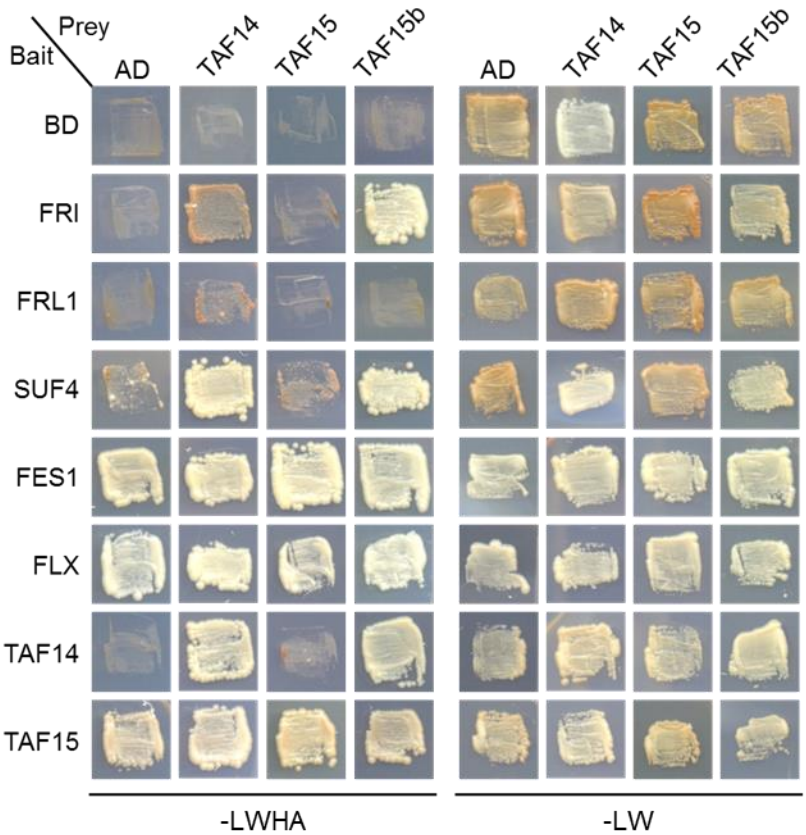

Figure 18. TAF15b interacts with FRI-complex proteins

Interactions of TAF14, TAF15 and TAF15b with FRI-complex genes in yeast cells. Genes were fused in frame to the GAL4 activation domain (AD)–coding and GAL4 DNA binding domain (BD)–coding sequences, respectively. Protein interactions were assayed by cell growth on selective medium. –LWHA indicates Leu, Trp, His, and Ade dropout plates. –LW indicates Leu and Trp dropout plates.

Interactions between TAF14, TAF15, TAF15b and autonomous pathway genes were also analyzed by Y2H assay (Figure 19). BD-FLK had auto-activation so I could not examine the interaction including FLK gene. TAF14 showed strong interaction with FCA and FPA, and weak interaction with FY. Interestingly, TAF15b

interacts with FCA, FY, FPA and LD, which all contain prion-like domains on their c-terminal region. On the other hand, TAF15 did not interact with any autonomous pathway proteins.

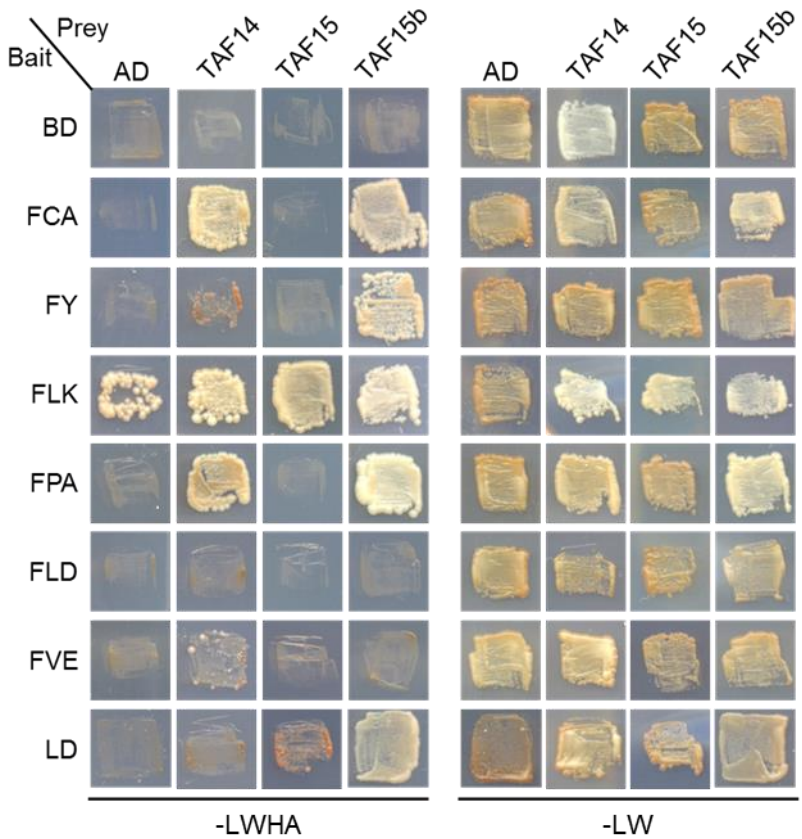

Figure 19. TAF15b interacts with autonomous pathway proteins

Interactions of TAF14, TAF15 and TAF15b with autonomous flowering pathway genes in yeast cells. Genes were fused in frame to the GAL4 activation domain (AD)–coding and GAL4 DNA binding domain (BD)–coding sequences, respectively. Protein interactions were assayed by cell growth on selective medium. –LWHA indicates Leu, Trp, His, and Ade dropout plates. –LW indicates Leu and Trp dropout plates.

I also performed BiFC assay using *Arabidopsis* protoplast to confirm the Y2H result (Figure 20). Unfortunately, false-positive signals were observed from several combinations with empty vector therefore I could not interpret any results from BiFC assay.

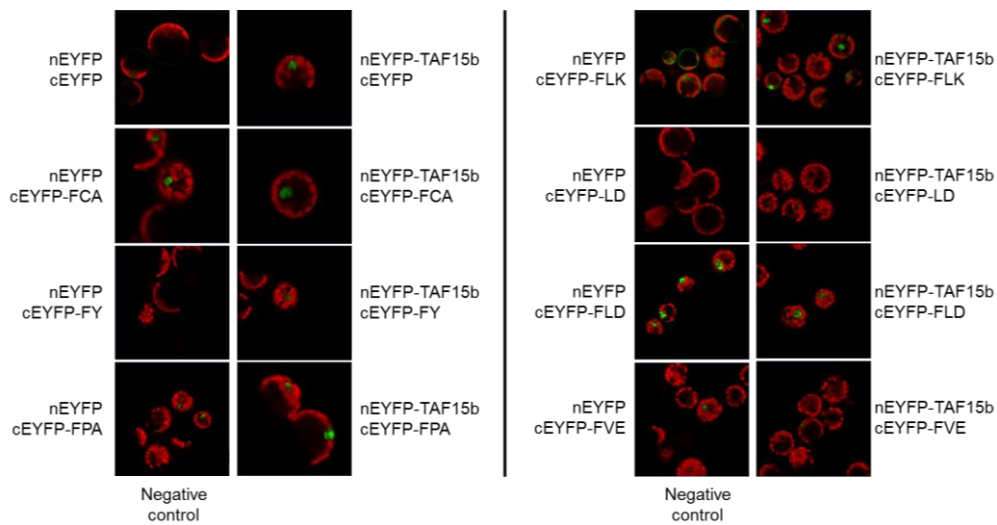

Figure 20. BiFC assay for the interaction of TAF15b with autonomous pathway proteins Different constructs of TAF15b and autonomous pathway genes with the N or C terminus of YFP were cotransformed into *Arabidopsis* protoplasts. Images were visualized and obtained using confocal microscopy at 19 h after transformation. Combination with empty vector (nYFP or cYFP) were used as negative controls.

2.2. Analysis of TAF15b complex

To understand the biochemical function of TAF15b, proteins interacting with TAF15b were identified using immunoprecipitation coupled to mass spectrometry (IP-MS) (Figure 21). Experiments were repeated twice, and ~1,100 and ~600 proteins were purified from the first and second experiments. 386 proteins were purified both from the first and second experiments and I used these proteins for further analysis (Appendix 3).

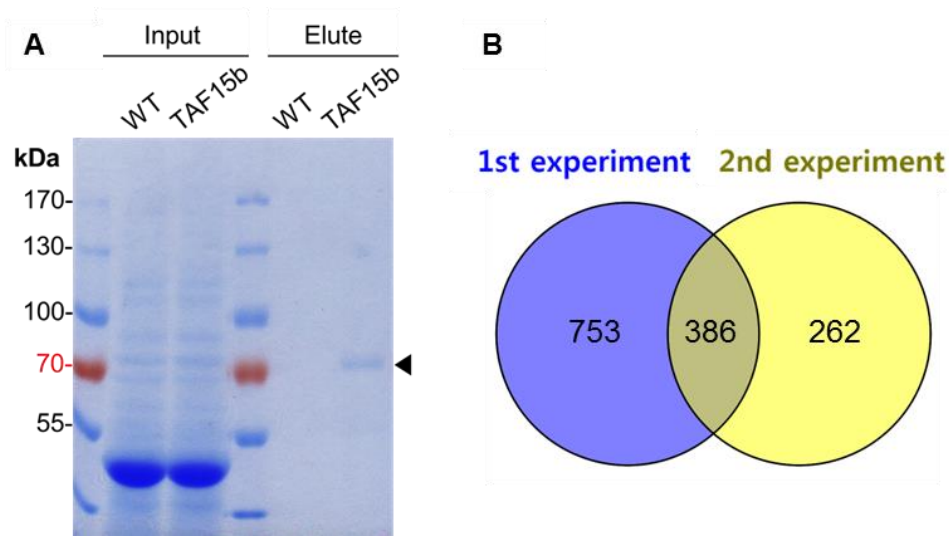

Figure 21. Analysis of TAF15b complex

(A) Co-immunoprecipitated proteins with TAF15b were resolved via 9% SDS-PAGE followed by coomassie blue staining. A control sample from WT was analyzed in parallel. The presence of TAF15b-GFP (~70kDa) is marked with arrowhead (◀).

(B) Venn diagram showing the number of proteins from two independent purifications. The first purification used WT and *CsVp::TAF15b-GFP* seedlings grown for 10 days in long days with GFP-Trap. The second purification used *TAF15bp::TAF15b-GFP* seedlings with binding control and GFP-Trap.

Gene ontology (GO) terms were analyzed against 386 genes and this result suggested unexpected point of views in several ways. In biological process, TAF15b binds to protein related with development and stress response (Figure 22). In cellular component, TAF15b interacts with proteins, which locate at the nucleus, cytoplasm and membrane. This coincides with the previous subcellular distribution patterns of TAF15b, which locates both at the nuclei and cytoplasm.

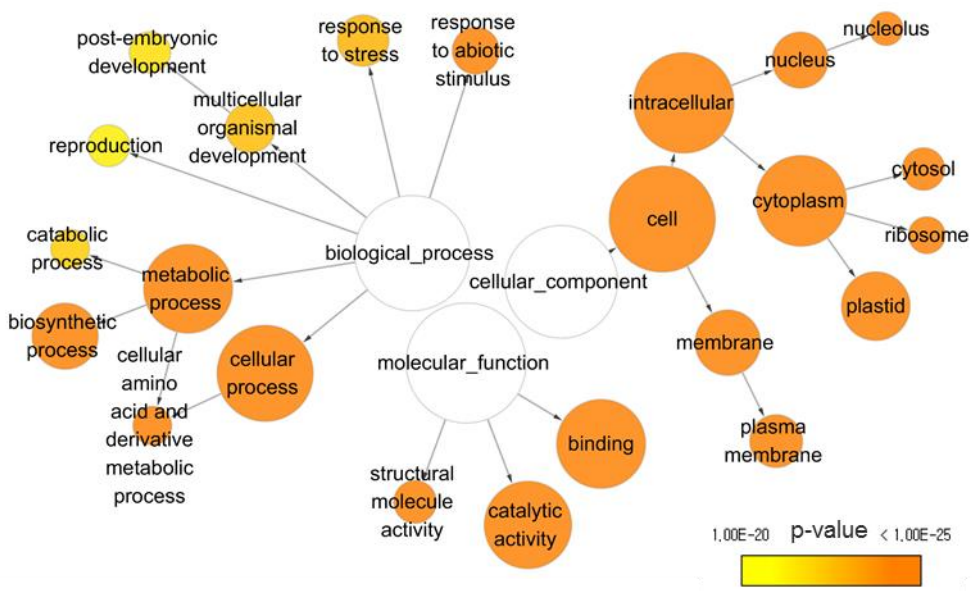

Figure 22. A graphical representation of GO terms associated with TAF15b interacting proteins

GO analysis of terms enriched for 386 genes, which were selected on two IP-MS experiments, was done using BiNGO (Maere et al., 2005). The node size is comparative to the number of proteins denoted by that GO term. Color scale represents the P-value for each enriched GO term.

In the aspect of overall GO terms, top 10 GO terms included transit peptide, RNA-binding, ribosome, and proteasome (Table 1). Transit peptides indicates proteins encoded by nuclear genes and transported to particular organelles, such as chloroplast and mitochondria. In this experiment, most of them were chloroplast enzymes so they might be the target proteins of TAF15b rather than partners regulating transcription. Next, TAF15b interacts with many RNA-binding proteins, which suggests that TAF15b might function together with other RNA-binding proteins. It also bound to macromolecular complexes, such as ribosome and proteasome and this implies that TAF15b could also be involved in the regulation of translation.

Table 1. The top 10 GO terms significantly enriched for TAF15b binding proteins. GO analysis was conducted by DAVID 6.8 and terms listed in descending order of enrichment score. Count indicates the number of genes included in the GO term.

	Term	Enrichment Score	Count
1	Transit peptide	46	132
2	RNA-binding	20	51
3	ATP-binding	14	92
4	Cytosolic ribosome	8	26
5	Proteasome	6	15
6	Biosynthesis of amino acids	6	38
7	Chloroplast thylakoid	6	22
8	Protein biosynthesis	5	21
9	Nucleoid	5	9
10	Microtuble	5	10

Because TAF15b interacts with several FRI-complex and AP proteins in the Y2H assay, I confirmed whether those proteins are also included in the IP-MS result. There are no FRI-complex proteins at all and several AP proteins, FLK, PRP8, FCA, FVE, SR45, were on the list (Table 2). FY, FPA, LD, which showed positive interactions with TAF15b in the Y2H assay, were not detected in IP-MS. Moreover, AP proteins on the IP-MS list did not appear in the second IP experiment and their peptides number is quite small to insist their interaction with TAF15b. Further experiments are needed to confirm the interaction between TAF15b and AP proteins.

Table 2. List of co-immunoprecipitated autonomous pathway proteins with TAF15b. Identified autonomous pathway proteins from two independent purifications listed in descending order of peptides number.

no.	AT number	Descriptive name	Peptides number
1	AT3G04610	RNA-binding KH domain-containing protein (FLK)	8/0
2	AT1G80070	Pre-mRNA-processing-splicing factor (PRP8)	8/0
3	AT4G16280	RNA binding;abscisic acid binding (FCA)	1/0
4	AT2G19520	WD-40 repeat family protein (FVE)	1/0
5	AT1G16610	arginine/serine-rich 45 (SR45, RNPS1)	1/0

2.3. CPL1 is a putative regulator of *FLC* transcription

To find a putative candidate working with TAF15b on the regulation of *FLC* transcription, proteins satisfying the following conditions were selected from the IP-MS result; ⅰ) immunoprecipitated from both two experiments, ⅱ) localize to the nucleus ⅲ) function as transcriptional regulator ⅳ) display late-flowering phenotype. Based on those conditions, I selected *C-TERMINAL DOMAIN PHOSPHATASE-LIKE 1* (*CPL1*) as a putative partners working with TAF15b. CPL1 is a phosphatase, which dephosphorylate Ser5 of the CTD in Pol II and functions as a transcriptional repressor, especially on stress and wound responsive genes (Koiwa et al., 2002, 2004; Matsuda et al., 2009). Also, it localizes to the nucleus and mutant of *cpl1* flowers later than WT (Koiwa et al., 2002, 2004; Manavella et al., 2012).

CPL1 contains Fcp1 homology (FCPH) phosphatase motif at the N-terminal region and double-stranded RNA binding domain (dsRBD) at the C-terminal region (Figure 23A). To confirm the interaction between CPL1 and TAF15b, I performed Y2H assay (Figure 23B). BD-CPL1 had weak auto-activation so 3AT was applied to the selection media to reduce the auto-activity. As expected, CPL1 interacted with TAF15b; however, when LC domains of TAF15b were deleted, the interaction disappeared. From this result, it could be assumed that TAF15b interacts with CPL1 through its LC domains.

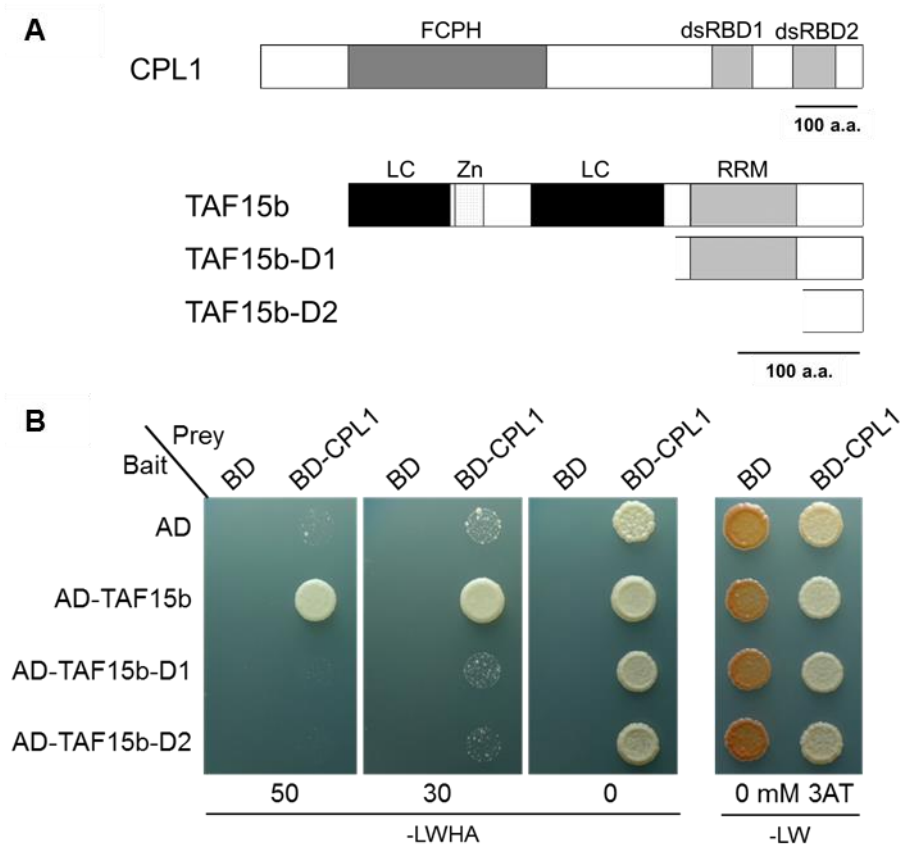

Figure 23. TAF15b interacts with CPL1 through LC domain

(A) Domain structures of CPL1 and TAF15b used for Y2H assay. FCPH, Fcp1 homology domain; dsRBD, double-stranded RNA-binding domain.

(B) Interactions between TAF15b and CPL1. Y2H assay of TAF15b deletion clones with full-length CPL1. Genes were fused in frame to the GAL4 activation domain (AD)–coding and GAL4 DNA binding domain (BD)–coding sequences, respectively. Protein interactions were assayed by cell growth on selective medium. –LWHA indicates Leu, Trp, His, and Ade dropout plates. –LW indicates Leu and Trp dropout plates.

I also performed BiFC assay to confirm the Y2H result (Figure 24). Unfortunately, false-positive signals were observed from several combinations with N terminus of YFP and empty vector. No signal was observed from the combination of C terminus of YFP, therefore I could not interpret any results from BiFC assay.

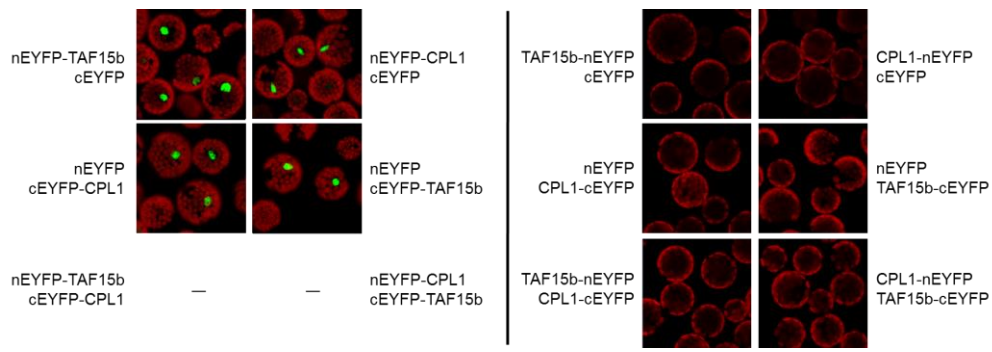

Figure 24. BiFC assay for the interaction of TAF15b with CPL1

Different constructs of TAF15b and CPL1 genes with the N or C terminus of YFP were cotransformed into *Arabidopsis* protoplasts. Images were visualized and obtained using confocal microscopy at 19 h after transformation. Combination with empty vector (nYFP or cYFP) were used as negative controls.

CPL1 also distributed in the nuclei except nucleoli region (Figure 25). The difference with TAF15b in subcellular distribution is that CPL1 did not locate in the cytoplasm. This distribution pattern implies that CPL1 interacts with TAF15b only in the nuclei and functions in the transcriptional regulation.

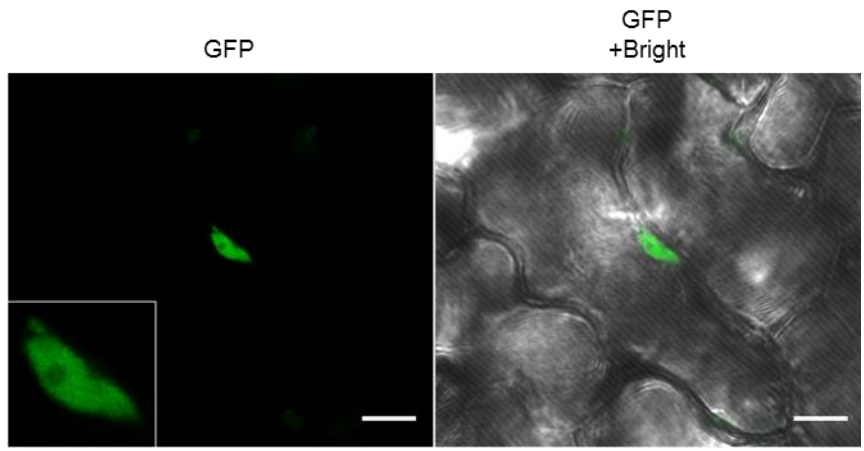

Figure 25. Subcellular localization of CPL1
Upper surfaces of rosette leaves of *35Sp::GFP-CPL1*- transgenic lines were observed under a confocal microscope. The scale bar represents 10 μm.

Finally, *FLC* transcript levels were compared between WT and *cpl1-7* mutant by qRT-PCR analysis. In *cpl1* mutant, both spliced and unspliced forms of *FLC* were increased two-fold compared to the WT (Figure 26). Although the fold change is lower than *taf15b* mutant, such results support the idea that CPL1 represses the transcription of *FLC* with TAF15b.

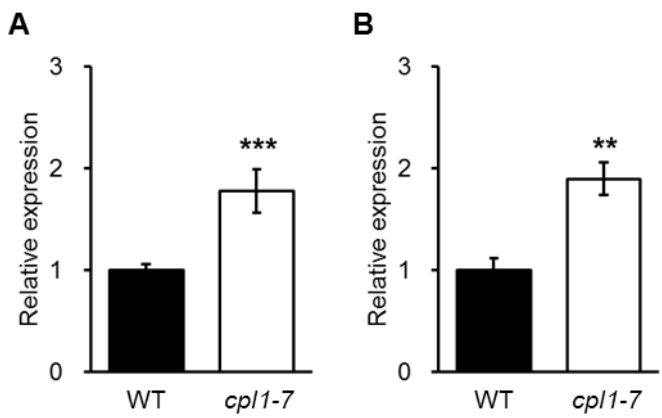

Figure 26. Transcripts of *FLC* increased in *cpl1* mutant

(A-B) qRT-PCR analysis of spliced (A) and unspliced (B) *FLC* transcripts of WT and *cpl1-7* mutant. Seedlings grown for ten days in long days were used for the analysis. Relative transcript levels to *TUB2* are normalized to non-vernalized WT. Means ± SD from three technical replicates are presented. Asterisks indicate statistically significant differences evaluated using Student's *t*-test (** $p<0.01$, *** $p<0.001$).

3. Discussion

In this chapter, I presented the protein interactions between TAF15b and FRI-complex, AP proteins and I also identified the components of TAF15b complex. TAF15b bound to the components of FRI-complex and several AP proteins with PrDs in Y2H assay. IP-MS assay with TAF15b-GFP suggested that TAF15b interacted with various proteins of biological functions and cellular components and it might be involved in other molecular mechanisms more than the transcription.

In Y2H assay, TAF15b interacts with several AP proteins such as FCA, FY, FPA and LD and those proteins contain PrDs at the c-terminal regions. For this reason, it could be assumed that TAF15b and AP proteins might interact through PrD. On the other hand, it should not be ignored that those interactions might be artificial because LC and PrD have a tendency to aggregate itself. Therefore, further *in vivo* experiment should be performed to confirm those interactions.

Through IP-MS analysis, CPL1, a phosphatase of Ser5 of the Pol II CTD, was chosen as a candidate for the interacting protein of TAF15b to regulate *FLC* transcription. Indeed, *FLC* transcript level was increased in the *cp1* mutant and CPL1 also localized to the nucleus as TAF15b did; however, I could not verify the biochemical function of how CPL1 regulates *FLC* transcription. A possible role of CPL1 could be assumed from other Ser5 phosphatase, human Ssu72. Ssu72 belongs to Cleavage and Polyadenylation Specificity Factor (CPSF) complex and decreases

the rate of transcriptional elongation (Dichtl et al., 2002; Mayfield et al., 2016; Xiang et al., 2010). It mainly locates at the end of the transcription cycle and is involved in the formation of gene looping, which is for RNA Pol II recycling (Tan-Wong et al., 2012; Zhang et al., 2012). Interestingly, *FLC* locus also has gene loop structure formed by the 5′ and 3′ ends of the gene (Crevillén et al., 2012). Because TAF15b binds Ser2P form of Pol II, which representing the status of transcriptional termination, CPL1 might maintain transcription termination by dephosphorylating of Ser5 of the CTD. It could also have a role in the formation of *FLC* gene loop with the interaction of TAF15b. Further experimentation is required to support this hypothesis.

CHAPTER V

General conclusions

1. General conclusions

The transcription of *FLC* is regulated by various mechanisms, including activation by the FRIGIDA complex and the autonomous pathway (AP), and vernalization-mediated repression (Berry and Dean, 2015). The AP represses the transcription of *FLC* regardless of the day length, and many genes are involved in this pathway, most of which are related to RNA processing and histone modification (Cheng et al., 2017; Rataj and Simpson, 2014). Compared to other mechanisms, little is known about the mechanism by which the genes involved in the AP repress the transcription of *FLC*. In this study, I discovered a new gene involved in the AP, *TAF15b*, and suggested that the TAF15b protein might regulate the transcription of *FLC* by directly interacting with Pol II in the nucleus (Eom et al., 2018). Interestingly, TAF15b is also localized to the cytoplasmic RNA granules in response to heat stress, which suggests that it might be involved in both transcriptional and translational processes.

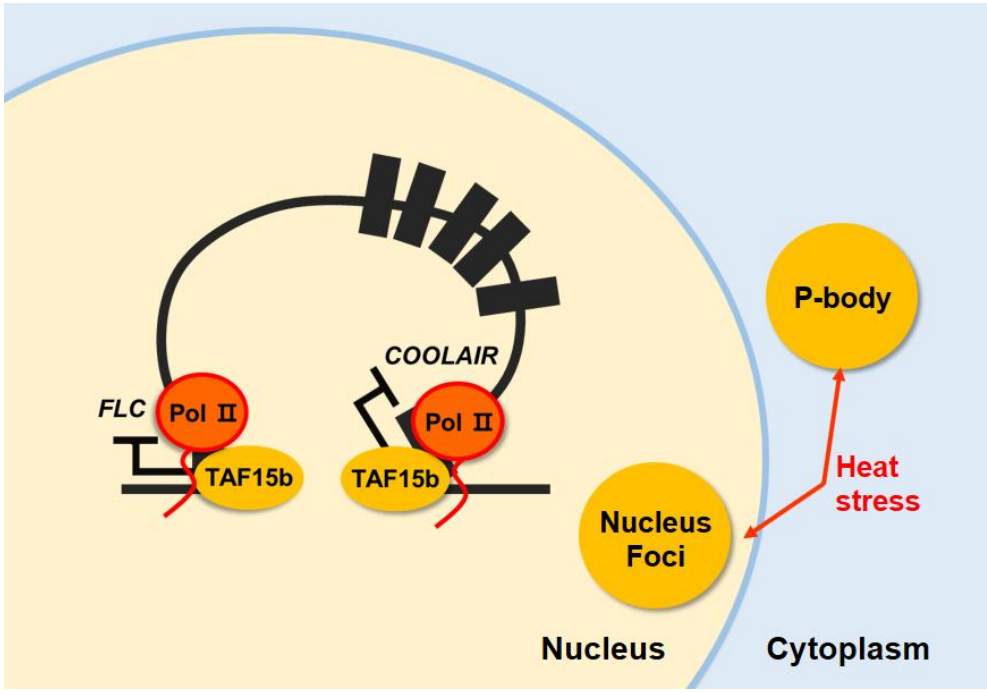

Figure 27. A schematic representing the dual functions of TAF15b in the nucleus and cytoplasm

In the nucleus, TAF15b is localized near the TSSs of *FLC* and *COOLAIR* and represses their transcription. TAF15b interacts with Pol II and affects the phosphorylation status of the CTD of Pol II, on the FLC locus. At normal temperatures, TAF15b is mainly localized in the nucleus, while forming nuclear foci and cytoplasmic p-bodies under heat stress.

The TAF15b protein of *Arabidopsis* is a homolog of the human TAF15 protein, which belongs to the FET protein family. This family consists of three proteins, FUS, EWSR1, and TAF15, which share a similar domain organization, including LC domain, the RRM, and the zinc finger domain, in that order (Schwartz et al., 2015). The molecular functions of FET proteins are related to transcription and

RNA processing, and mutations in their genes cause neurodegenerative diseases (Schwartz et al., 2015). In plants, there is only one FET protein, TAF15b, which is involved in the regulation of flowering and autoimmunity (Dong et al., 2016; Eom et al., 2018). Our study showed that TAF15b promotes flowering by repressing the transcription of *FLC* and binds to the TSSs on both the sense and antisense strands of *FLC*. Because TAF15b binds to Pol II and downregulates the occupancy of Pol II on the *FLC* locus, I assumed that TAF15b might regulate the transcription of *FLC* through direct inhibition of Pol II activity on the TSS.

The exact mechanism as to how TAF15b represses the transcription of *FLC* is unclear; however, the biochemical functions of FET proteins could provide insights into the functions of TAF15b of *Arabidopsis*. FUS regulates transcription by inhibiting the phosphorylation of Ser2 in the CTD of Pol II at the TSSs of the target genes (Schwartz et al., 2012). The target RNA sequences of FUS promote the interactions between the FUS and the CTD and also the formation of fibrous assemblies of FUS protein (Schwartz et al., 2012, 2013). It also represses the activity of histone acetyltransferases by interacting with noncoding RNAs that recruit FUS to the promoter of target genes (Wang et al., 2008). Similar to the human FUS proteins, the TAF15b of *Arabidopsis* also binds to the TSSs of *FLC* and *COOLAIR*, which is an antisense transcript of *FLC*, and represses their transcription (Figure 27). Thus, the transcriptional inhibition caused by TAF15b may result from the regulation

of Pol II activity, as TAF15b binds to Pol II and affects the phosphorylation of the CTD at the *FLC* locus. Otherwise, the enrichment of TAF15b at both the sense and antisense TSSs could be an indirect effect occurring as a consequence of the gene loop structure formed by the 5′ and 3′ ends of the *FLC* gene (Crevillén et al., 2012).

Interaction between TAF15b and Pol II might occur through LC domain of two proteins. The CTD of Pol II also has another well-known LC domain, composed of heptad repeats and acts as a binding platform for RNA splicing and 3' processing factors (Eick and Geyer, 2013; Koiwa et al., 2004). It is known that the FET proteins bind to the CTD of Pol II through their LC domains (Kwon et al., 2013; Schwartz et al., 2013). Interestingly, other proteins in the AP that are associated with RNA processing also have PrDs at their N-termini (Chakrabortee et al., 2016). This suggests a possibility that the proteins involved in the AP may also bind to the CTD of Pol II through their PrDs and regulate the expression of *FLC* cotranscriptionally.

Although several proteins involved in the AP have been identified that contain RNA binding domains, the molecular mechanism by which they repress the transcription of *FLC* is not fully understood. With the identification of the novel AP protein, TAF15b, I suggest that proteins involved in the AP might function along with Pol II to inhibit transcription. Moreover, the translocation of TAF15b to p-bodies in response to heat stress raises a possibility that TAF15b might also function in the translational repression of other mRNAs. Since RNAs are involved in both

transcriptional and translational processes in the central dogma of molecular biology, proteins having RNA binding domains could have dual functions in the regulation of both processes.

2. Hypothesis

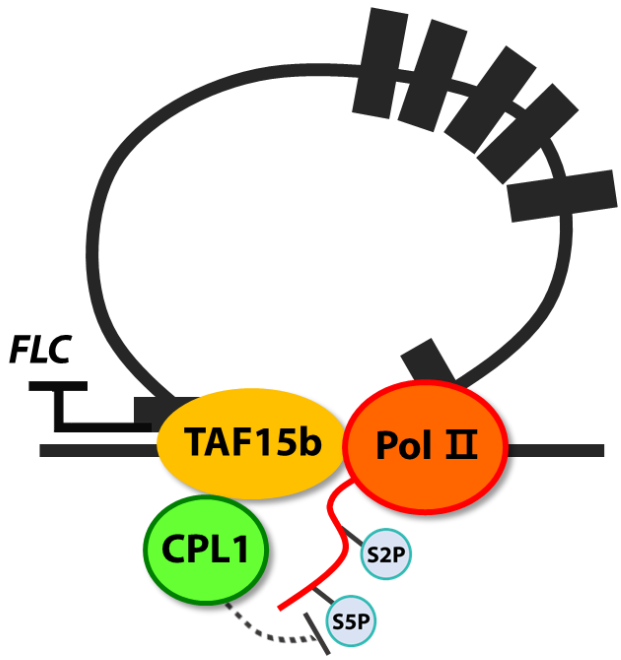

Figure 28. A schematic representing the hypothesis of the *FLC* regulation by TAF15b and CPL1

In this study, I found that TAF15b strongly bound to the Pol II in the transcriptionally terminated status (Ser2P) and also interacted with CPL1, which dephosphorylates Ser5 of the Pol II CTD. It is also well known that *FLC* forms gene loop structure formed by the 5′ and 3′ ends of the gene (Crevillén et al., 2012). This leads me to hypothesize that TAF15b and CPL1 might repress *FLC* transcription by maintain the Pol II in transcriptionally terminated status (Figure 28). TAF15b might

interact with CPL1 and locate at the TSS of *FLC*. After one round of transcription, Pol II would reach to the TTS of *FLC* and at this time, CPL1 might dephosphorylate the Ser5P of Pol II CTD and prevent the recycling of Pol II by keeping it in transcriptionally terminated form.

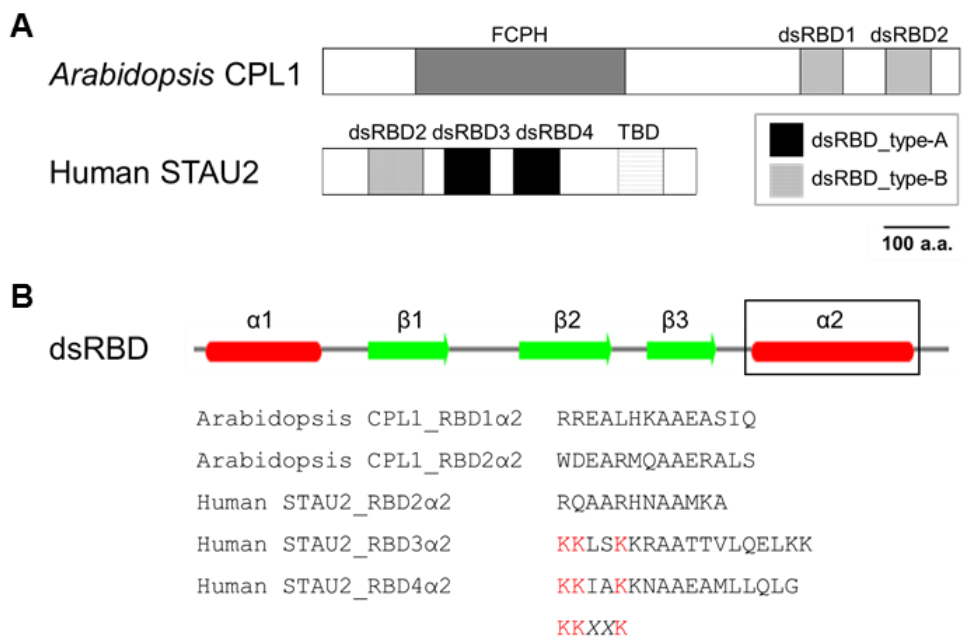

Figure 29. Diagrams of *Arabidopsis* CPL1 and human STAU2

(A) Domain structures of CPL1 and STAU2. FCPH, Fcp1 homology domain; RBD, double-stranded RNA-binding domain. TBD, tubulin-binding domain.

(B) Structural-based alignment of dsRBDs. Amino acid sequences of α2 helix were compared. Only human STAU2 RBD3 and RBD4 contain KK*XX*K motif, which is important dsRNA binding.

Because CPL1 has two dsRBD domains, I assumed that those domains might be important to bind *FLC* RNA (Figure 29). dsRBD consists of an α1-β1-β2-β3-α2 secondary structure and positively charged KK*XX*K motif is a key motif to bind dsRNA (Valente and Nishikura, 2007). I ran BLAST with amino acid sequences of dsRBD1 and dsRBD2 of *Arabidopsis* CPL1 and found out that dsRBD2 of human Staufen 2 (STAU2) shows high similarity with those domains. STAU2 is a dsRNA-binding protein involved in microtubule-dependent transport of RNAs and it contains three dsRBDs, of which dsRBD3 and dsRBD4 (type-A) can bind dsRNA and dsRBD2 (type-B) cannot (Gleghorn and Maquat, 2014). Using a computer program called Jpred 4 (A protein secondary structure prediction server, http://www.compbio.dundee.ac.uk/jpred/), I identified secondary structures of CPL1 and STAU2 dsRBD and compare the amino acids sequences of α2 helices. They were both type-B dsRBD not containing KK*XX*K motif and this means that CPL1 does not have any domains RNA binding capability. type-B dsRBD is responsible for protein-protein interactions rather than dsRNA binding therefore CPL1 dsRBDs might be also involved in protein interaction (Gleghorn and Maquat, 2014).

Based on those facts, I assumed that CPL1 could only regulate the phosphorylation status of Pol II CTD on the *FLC* gene through the protein interaction with TAF15b. To prove this hypothesis, various interactions between protein-protein and protein-nucleotides should be identified.

3. Experimental designs

3.1 CPL1 regulates flowering time and *FLC* expression

> Flowering time of *cpl1* in various environmental conditions (LD/SD, +/- vern, GA treatment)

> Complementation analysis of WT, *cpl1*, 35S::myc-CPL1 in *cpl1*, 35S::myc-CPL1 ΔdsRBD in *cpl1*

> *FLC* and *COOLAIR* expressions in WT, *cpl1*, *taf15b*, *cpl1 taf15b* (cv, mRNA and nascent transcripts)

> *CPL1* expression in *taf15b* and *TAF15b* expression in *cpl1*

> Genetic analysis of WT, *cpl1*, *flc*, *cpl1 flc*, *taf15b*, *cpl1 taf15b*

3.2 Identification target genes of TAF15b and CPL1

> ChIP-seq and RIP-seq with *TAF15bp::TAF15b-GFP* and *35Sp::GFP-CPL1* transgenic seedlings

> DNA and RNA-EMSA; MBP-TAF15b and His-CPL1 with putative target sequences

3.3 TAF15b interacts with CPL1

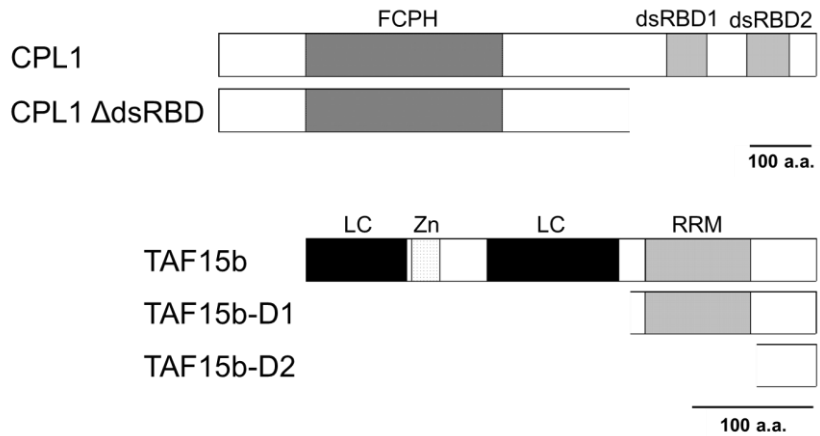

Figure 30. Domain structures of CPL1 and TAF15b that will be used for Y2H assay

> Y2H for deletion series of AD-TAF15b and BD-CPL1 (Figure 30)

> *In vitro* interaction; MBP-TAF15b and His-CPL1

> Co-IP using TAF15bp::TAF15b-GFP 35S::myc-CPL1

3.4 TAF15b mediates the interaction between CPL1 and Pol II CTD

> *In vitro* interaction; MBP-TAF15b and GST-CTD (+/- target RNA sequence)

> *In vitro* interaction; His-CPL1 and GST-CTD (+/- target RNA sequence and MBP-TAF15b)

3.5 *cpl1* effects on the Ser5P of Pol II CTD on *FLC* gene

> Western blot of Ser2P and Ser5P in WT and *cpl1*

> Ser2P and Ser5P ChIP on *FLC* gene (cv, WT, *taf15b*, *cpl1*, *taf15b cpl1*)

> ChIP assay on *FLC* gene (35Sp::myc-CPL1 in *taf15b-1*)

CHAPTER VI

Materials and methods

1. Materials

1.1. Plant materials

Arabidopsis thaliana Columbia-0 ecotype was used as the WT control in this study. *taf15b-1* (SAIL_35_B06) and *taf15b-2* (SALK_061974) were obtained from the Arabidopsis Biological Resource Center. Col:FRI^{SF2} (*FRI*-Col), *fca-9*, and other AP mutants are in the Col-0 background and have been previously described (Feng et al., 2011). *cpl1-7* and 35Sp::eGFP:CPL1 (pPM360) were kindly provided by Dr. Weigel (Manavella et al., 2012).

1.2. Oligonucleotides

Underlined sequences indicate restriction enzyme sites

Cloning of transgenic plants	Sequence (5' to 3')	Enzyme site
TAF15bp::TAF15b-GFP		
EOM126F	CCCGGATCCCATTTCTCCAGAGCTATGGC	BamHI
EOM126R	CCCGGATCCGCATATGGACGAGACCGGTTTC	BamHI
CsVp::TAF15b-GFP		
TAF15b_CDS_F	TCTAGAATGGCTGGGATGTACAATC	XbaI
TAF15b_CDS_R	GGATCCATATGGACGAGACCGGTTTC	BamHI
TAF15bp::GUS		
TAF15b_GUS_F	AAGCTTGTGTCAAAGATCGGTTGGCC	HindIII
TAF15b_GUS_R	GGATCCGACCTAGGGTTTGCGATAAA	BamHI

Cloning for Y2H		Sequence (5' to 3')	Enzyme site
For pGBKT7 vector			
FCA	EOM83F	AAACATATGATGAATGGTCCCCCAGATA	NdeI
	EOM83R	TTTTTGAATTCTCAAGCTTTATTCTTCCACATG	EcoRI
FPA	EOM84F	AAAGGATCCTTATGGCGTTATCTATGAAGCC	BamHI
	EOM84R	TTTGGATCCTCAAGGCCCCTGTCCAG	BamHI
FLK	EOM85F	AAAGGATCCTTATGGCTGAAGCTGAAGATC	BamHI
	EOM85R	TTTGGATCCTCAGTAACCGTAGCCTGAG	BamHI
LD	EOM86F	AAACATATGATGGACGCGTTCAAGGAG	NdeI
	EOM86R	TTTTTGAATTCTTAACGCCATCTTTTGTTTTGG	EcoRI
FY	EOM87F	AAAGGATCCTTATGATGCGGCAGTCGTC	BamHI
	EOM87R	TTTGGATCCCTACTGATGTTGCTGATTGTT	BamHI
TAF15	EOM88F	AAAAGATCTTAATGGCTGGATATCCTACTAATG	BglII
	EOM88R	TTTAGATCTTCAGTTACGGTACCTGCTTC	BglII
CPL1	EOM164F	AAAAGATCTTAATGTATAGTAATAATAGAGTAGAAGT	BglII
	EOM164R	AAAGTCGACTTAAGAGTATCTTCCCGAAGAT	SalI
For pGADT7 vector			
TAF15b	EOM89F	AAAGGATCCTTATGGCTGGGATGTACAATC	BamHI
	EOM89R	TTTGGATCCTCAATATGGACGAGACCGG	BamHI
TAF15b	EOM89F2	AAAGGATCCCAGCTACTGAAAAAGTCAAGCAG	BamHI
-D1	EOM89R	TTTGGATCCTCAATATGGACGAGACCGG	BamHI
TAF15b	EOM89F4	TTTTTGAATTCGAAAAGTCTGCGCCCCGT	EcoRI
-D2	EOM89R	TTTGGATCCTCAATATGGACGAGACCGG	BamHI

Genotyping		Sequence (5' to 3')	Band size
taf15b-1	EOM113F	TCTTTGGTGGCATTGGCC	1,035bp
WT	EOM113R	CTCCTCTACCCCCACCT	
taf15b-1	EOM113R	CTCCTCTACCCCCACCT	~1.2kb
mt	pCSA-LB3	TAGCATCTGAATTTCATAACCAATCTCGATACAC	
taf15b-2	EOM113F	TCTTTGGTGGCATTGGCC	1,035bp
WT	EOM113R	CTCCTCTACCCCCACCT	

taf15b-2	EOM113F	TCTTTGGTGGCATTGGCC	~500bp
mt	agl20-ko3	TTGGGTTCACGTAGTGGGCCATCG	
flc-3	FLC3F	GCTCGTCATGCGGTACACGTG	WT 460bp
	FLC3R	GGCGTACTTATCGCCGGAGG	mt 356bp
fca-9	FCAdCAPS1-1	GCCTCTATCCTTTCCCCC	
	FCAdCAPS2	GTACCCAAGGCATTACCTTG	
	After PCR, Sma Ⅰ cut, WT 80, 20bp mt 100bp		
fy-2 mt	FY4	CTGTTGGAAAGGGTTGTTG	
	LB3	TAGCATCTGAATTTCATAACCAATCTCGATACAC	
fpa-7 WT	EOM119F	GAAAGGGGATTGAAACTAAAC	821bp
	EOM119R	GCAGGTTGAGAAGTTGCAG	
fpa-7 mt	EOM119F	GAAAGGGGATTGAAACTAAAC	400bp
	CSW7	GACGTGAATGTAGACACGTCGAA	
flk-4 WT	FLK_colony_F	CAGGCACAACAGAAAGAGCT	375bp
	FLK_colony_R	CAGGATCATATCACTCAGCC	
flk-4 mt	FLK_colony_F	CAGGCACAACAGAAAGAGCT	747bp
	agl20-ko3	TTGGGTTCACGTAGTGGGCCATCG	
ld-1	SY-ld-1F	GCTGGGTAGCTTTCATCAATGCCA	
	SY-ld-1R	GAATATCTTCCTGTTACGACACG	
	After PCR, Mse Ⅰ cut, WT 381, 36bp mt 333, 48, 36bp		
fld-3 WT	EOM111F	GACGGGTTTATACCAAAAAG	1.4kb
	EOM111R	ATCGATTCTTTTCCTTATGCC	
fld-3 mt	EOM111F	GACGGGTTTATACCAAAAAG	1.2kb
	LBb1	GCGTGGACCGCTTGCTGCAACT	
fve-4	MJH6-1	GTTGATTTCAGGTTCTCATA	100bp
	MJH6-2	AGAAGTGGACATACCAAATC	
	After PCR, sequencing product with MJH6-2 to verify that TTCTCATTTG (WT) changed to TCTCATTTA (*fve-4*)		
FRI	FRI geno F	GAATGAGATTGCCGGTGC	WT 106bp

			mt (Col-0)
	FRI geno R	AGGAACCACCTTTGCAATG	90bp
cpl1-7	EOM177F	ATTCGTGATACAAGTGTTCTTG	158bp
	EOM177R	CCTTCTGGATCAAGGAGC	
	After PCR, sequencing product with EOM177F to verify that ATGGCTGAAA (WT) changed to ATGGTTGAAA (*cpl1-7*)		
RT-PCR			
TAF15b	EOM89F	ATGGCTGGGATGTACAATC	
	EOM89R	TCAATATGGACGAGACCGG	
TUB2	TUB2(qPCR)-F	ATCGATTCCGTTCTCGATGT	
	TUB2-R-1	TACTGCTGAGAACCTCTTGAG	
quantitative PCR for flowering time and autonomous genes			**Reference**
FLC	ONE122	AGCCAAGAAGACCGAACTCA	Swiezewski et
	ONE123	TTTGTCCAGCAGGTGACATC	al. (2009)
FT	FT-SH-F	GCTACAACTGGAACAACCTTTGGCAAT	
	FT-qR	GTTTGCCTGCCAAGCTGTC	
SOC1	SOC1(real)-F	ATGAATTCGCCAGCTCCAAT	
	SOC1(real)-R	GCTTCATATTTCAAATGCTGCA	
LFY	BJN26F	CCGTGAGTTCCTTCTTCAGG	
	BJN26R	GGAGAGCGTAACAGTGAACG	
FCA	EOM141F	GGTCAGAATGCTCAGGATTATG	
	EOM141R	TTCTTGGGGAGATTGACGGCT	
FY	EOM142F	TGGCGGACCCCAAATGTATC	
	EOM142R	GTGGCTGCTGCTGCTGTTG	
FPA	EOM143F	GGGATTGAAACTAAACTGCCTG	
	EOM143R	AACGATCTCACATCCAATGGC	
FLK	FLK_qRT_F	GCTTGTATAGTCAGAGTTCTTGG	
	FLK_qRT_R	CAAGAACTTTCTGAGATGTGACG	
LD	YU291F	GTCTCTCAAATGGAAAGTCAGAG	
	YU291R2	CCTGCGTTCTTTGTTATACGATG	

FLD	EOM145F	AAGCGGAAGATGATCAGAAGG	
	EOM145R	ATTTGCCATTACCTAAAGACTCTG	
FVE	EOM144F	GGTGGAACATTGCAGATATGG	
	EOM144R	TTAAGGCTTGGAGGCACAAGT	
CDKC2	EOM140F	CCATGTGATCCAAAGAGTCTAC	
	EOM140R	GGCTGCTTCCTCATTATGACG	
TAF15b	EOM140F	CCATGTGATCCAAAGAGTCTAC	
	EOM140R	GGCTGCTTCCTCATTATGACG	
TUB2	EOM72F	ACTGTCTCCAAGGGTTCCAG	
	EOM72R	GTGAAGGGAACACCGAGAAG	
quantitative PCR of *FLC* and *COOLAIR*			**Reference**
FLC spliced	spliced exon 4/5	AGCCAAGAAGACCGAACTCA	Swiezewski et al. (2009)
	exon 7	TTTGTCCAGCAGGTGACATC	
FLC unspliced	unspliced exon 1	TTCTCCTCCGGCGATAAGTA	
	ONE125	CCGAGAAACAACAAGAGATCCG	
Proximal pA	LP_FLCin6polyA	TTTTTTTTTTTTTTTACTGCTTCCA	Liu et al. (2010)
	set1_RP	CACACCACCAAATAACAACCA	
Distal pA	LP_FLCprom_polyA	TTTTTTTTTTTTTTTGCGGTACAC	
	set4_RP	GGGGTAAACGAGAGTGATGC	
Chromatin immunoprecipitation of *FLC*			**Reference**
A	FLC B-F Q	TGTAGGCACGACTTTGGTAACACC	
	FLC B-R Q	GCAGAAAGAACCTCCACTCTACATC	
B	FLC C-F Q	TATCTGGCCCGACGAAGAAA	
	FLC C-R Q	TTTGGGTTCAAGTCGCCGGAGATA	
C	FLCp 7QF	TGCATGTCATTCACGATTTG	
	FLCp 7QR	AGATGGCTTGAAACTTCACTCA	
D	FLCp 8QF	AGTAGTTTGGCCATGTTGGTC	
	FLCp 8QR	GTCTCGACAATTCCAAGGCT	
E	FLC body 5 F	AGACTGCCCTCTCCGTGACT	
	FLC body 5 R	AAAGCAGCACGGTTGTTCTC	

F	FLC 3'UTR F	GGGCGAGCGTTTGTATATCTT	
	FLC 3'UTR R	CGATGCAATTCTCACACGAA	
G	FLC 9Q-1_F	GATAACACTTCTTTTTTTCCCACTA	
	FLC 9Q-1_R	GGCGTCTGGAACTGTTGATC	
Ta3	Ta3-F	CTGCGTGGAAGTCTGTCAAA	
	Ta3-R	CTATGCCACAGGGCAGTTTT	
Chromatin immunoprecipitation of *ACT7*			**Reference**
-995	Actin7_-995F	TGGGTCTCATATAGAACACTCACAAAGGT	Wu et al. (2016)
	Actin7_-832R	GACCAAAACCCGAATAGGAGCAAGA	
+55	Actin7_55F	CGTTTCGCTTTCCTTAGTGTTAGCT	
	Actin7_188R	AGCGAACGGATCTAGAGACTCACCTTG	
+886	Actin7_886F	TGCCCCGAGAGCAGTGTTCC	
	Actin7_992R	TGGACTGAGCTTCATCACCAACG	
+2477	Actin7_2477F	GTATCGGGTGACAATGCAGCTATTATGT	
	Actin7_2561R	TGCTGGAGTAAAACATAAGCCACTCAC	

2. Methods

2.1. Plant growth conditions

To measure flowering time, seeds were sterilized using 75% ethanol with 0.05% Triton X-100, rinsed with ethanol, and then dried. Seeds were spread on half-strength Murashige and Skoog (MS) plates with 1% sucrose and 1% plant agar and stratified at 4°C for 3 days. They were grown at 22°C in long days (16 h light/8 h dark) or short days (8 h light/16 h dark) and then seedlings were transplanted to soil at the desired stage. For vernalization treatments, seeds on the MS agar plates were incubated at 4°C for 50 days in short days and transplanted to soil in long days. For gibberellin treatments, I sprayed twice a week with 100 μM GA_3 (Sigma) in 0.1% (v/v) ethanol after 7 long days, with control plants sprayed with only 0.1% (v/v) ethanol. Flowering time was measured by counting the number of rosette leaves when bolting occurred, using at least 16 plants.

2.2. Construction of vectors

For the complementation test, I generated *TAF15bp::TAF15b-GFP* by amplifying the genomic sequence of *TAF15b* with a 1.4 kb promoter fragment as described by Dong et al. (2016) using EOM126F and EOM126R primers. The PCR product was

cloned into the modified binary vector *pCAMBIA1300-NOS* with a C-terminal *GFP* fusion (Cho and Cosgrove, 2002; Lee et al., 2010).

For the overexpression and cellular localization analyses, *TAF15b* cDNA was amplified using TAF15b_CDS_F and TAF15b_CDS_R primers and cloned into the binary vector *CsV-GFP3-PA* between the *CsV* (*Cassava vein mosaic virus*, -368 to +20) and *GFP* sequences.

The *TAF15bp::GUS* construct was generated by amplifying the 1 kb promoter region of *TAF15b* using TAF15b_GUS_F and TAF15b_GUS_R primers and and cloning into the modified binary vector pBI121.

2.3. Generation of transgenic plants

The constructs were transformed into *Agrobacterium* strain GV3101. Transformation of constructs *TAF15bp::TAF15b-GFP* and *CsVp::TAF15b-GFP* into *taf15b-1* and construct *TAF15bp::GUS* into Col-0 was performed using the floral dip method (Clough and Bent, 1998). T1 transgenic plants of the GFP constructs were selected on MS medium with hygromycin.

2.4. Histochemical analysis of GUS activity

GUS expression of *TAF15bp::GUS* transgenic plants in the Col-0 background was detected following the method described previously (Choi et al., 2007).

2.5. Confocal microscopy

CsVp::TAF15b-GFP plants (in the Col-0 background) were grown on soil at 22°C in long days for 12 days and were transferred to a 4°C freezer for cold treatment or a 42°C chamber for heat treatment for 1 h. The upper surfaces of rosette leaves were observed using a Zeiss LSM 700 confocal laser scanning microscope. At least two individual plants with three rosette leaves were used for the observations.

2.6. RNA isolation and reverse transcription

Total RNA was extracted from 10-day-old seedlings grown in long days using the TS^{TM} plant RNA Minikit (Taeshin BioScience, #TS5210) following the manufacturer's instructions. 4 μg of total RNA was treated with 5 U of recombinant DNaseI (TaKaRa, #2270A) at 37°C for 40 min and then incubated at 75°C for 10 min to inactivate DNaseI. cDNA was synthesized using 50 ng of oligo(dT)$_{18}$ primer

with annealing at 65°C for 5 min, and then incubated on ice for another 5 min. Then, 20 mM dNTPs, 5× reaction buffer, and 200 U of RevertAidreverse transcriptase (Fermentas, #EP0441) were added in a total volume of 20 µl and incubated at 42°C for 1 h, and then 72°C for 10 min.

2.7. Quantitative RT-PCR

qRT-PCR analysis was performed using iQSYBR Green SuperMix (Bio-Rad, #170-8882) and monitored by the CFX96 real-time PCR detection system. cDNA (0.4 µl) was used in a 20 µl reaction and the PCR reaction was as follows: 5 min at 95°C, 60 cycles of PCR (30 s at 95°C, 30 s at 60°C, and 30 s at 72°C), and then dissociation from 60°C to 95°C. At least two biological replicates were performed, with three technical replicates for each. Results from one biological replicate are shown.

2.8. Chromatin immunoprecipitation assay

ChIP assay was performed as previously described in Saleh et al. (2008) with some modifications in nuclear lysis, sonication and immunoprecipitation steps. I used 2 g of 10-day-old seedlings grown in long days for each experiment. After cross-linking and nuclear isolation steps, I resuspended the nuclei pellet in 1 ml of nuclei lysis

buffer (50mM HEPES pH 7.5, 150mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% Sodium deoxycholate, 1mM PMSF, cOmplete, EDTA-free protease inhibitor (Roche, #5056489001)) and divided them into three aliquots of 350 µl each in 1.5 ml Bioruptor® Pico Microtubes (#C30010016). Chromatins were sheared by sonicating with Bioruptor® Pico (diagenode) for 15 min (30 sec ON, 30 sec OFF) at 4℃ and centrifuged samples for 10 min at 13,000 rpm at 4℃. After that, I combined all the supernatants and added dilution buffer (20mM Tris-HCl pH 8.0, 150mM NaCl, 2mM EDTA pH 8.0, 1% Triton X-100, 1mM PMSF, protease inhibitor) up to a total volume of 1.2 ml and took 500 µl chromatin for immunoprecipitation, 25 µl for input control (1/20 v/v). Dilution buffer was added to chromatin up to a total volume of 1.4 ml and add appropriate antibody or conjugated beads were applied to the samples. For H3K4me3, Ser2P and Ser5P IP, I added 2 µl of anti-trimethyl-histone H3 (Lys4) (Merk, #07-473), anti-Ser2P (abcam, #ab5095) or anti-Ser5P (abcam, #ab5131) antibody to the samples and incubated overnight at 4℃ with rotation and then applied 25 µl of Protein G PLUS-Agarose beads (Santa Cruz, #sc-2002) with another 4℃ rotation for 2 h. For GFP IP, I added 25 µl of GFP-Trap®_A (ChromoTek, #gta-20) to the samples and incubated only 2 h at 4℃ with rotation. After immunoprecipitation, I centrifuged samples at 3,800 g for 1 min at 4℃, discarded supernatants and added 1ml of dilution buffer to wash beads for 5 min at 4℃ with rotation. I repeated wash twice and added 125µl of elution buffer (4% SDS, 0.1M

um bicarbonate) and incubated the samples at 65℃for 10 min with vortexing. Supernatants were obtained after centrifugation 13,000rpm for 5min and 355 μl of dH_2O were added to dilute the SDS concentration. I also added 125 μl of elution buffer, 330 μl of dH_2O to the input controls and reverse cross-linking, protein digestion and DNA precipitation were proceeded as Saleh et al. (2008) described. DNAs were dissolved with 100 μl of dH_2O and 0.2 μl input control (1% input), 1 μl ChIP DNA were used for qRT-PCR in a final volume of 20μl. ChIP enrichment levels were calculated using Ct values (Lin et al., 2012) and normalized with *TA3* gene.

2.9. Western blot

I used 1 g of 10-day-old Col-0 and *taf15b-1* seedlings grown in long days to compare the amounts of phosphorylated Pol Ⅱ. Samples were grounded in liquid nitrogen and 1 ml of grinding buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 10mM $MgCl_2$, 0.1% Nonidet P-40, 1mM PMSF, protease inhibitor) were added to each samples. After the 15 min incubation at room temperature, I centrifuged samples for 10 min at 13,000 rpm at 4 ℃ to remove cell debris and repeated at least three times. Concentrations of protein samples were measured using protein assay dye (Bio-Rad, #500-0006) and I added 5X SDS loading sample buffer to 120 μg of total proteins and boiled them for 10 min at 95 ℃. Boiled proteins were used for western blot

analysis with anti-CTD (1:2000 diultion, abcam, #ab817), anti-Ser2P (1:2000 diultion, abcam, #ab5095) or anti-Ser5P (1:2000 diultion, abcam, #ab5131).

2.10. Co-immunoprecipitation

Co-IP assays were performed as previously described in Lee and Seo (2016) with few modifications. 2 g of 10-day-old WT and *CsVp::TAF15b-GFP* seedlings grown in long days were grounded in liquid nitrogen and 2ml of grinding buffer were added to each samples. Total proteins were obtained as described above and 25 μl of GFP-Trap®_A (ChromoTek, #gta-20) were added to the samples and incubated for 1 h at 4℃ with rotation. After immunoprecipitation, I centrifuged samples at 3,800 g for 1 min at 4℃, discarded supernatants and added 1ml of grinding buffer to wash beads for 5 min at 4℃ with rotation and repeated wash three times. I added 5X SDS loading sample buffer and dH_2O to the beads and boiled them for 10 min at 95 ℃. Elutes were divided into four and each of them was used for western blot analyses of CTD, Ser2P, Ser5P and GFP as described above. For the western blot of GFP, I used anti-GFP antibody (Clontech, #JL-8) at 1:5000 dilution.

2.11. Yeast two-hybrid (Y2H) assay

For the Y2H analysis, the CDS sequences of genes were cloned into pGBKT7 and pGADT7 vectors then assay was performed the manufacturer's instructions (Clontech, Matchmaker GAL4 Yeast Two-Hybrid System 3). Vectors for FRI-complex and TAF14 were used from the previous study (Choi et al., 2011). AH109 yeast strain were used for the assay and protein interactions were confirmed following the method described previously (Choi et al., 2011).

2.12. Protoplast transient assay

Different constructs of CDS sequences of genes with the N or C terminus of YFP were cotransformed into *Arabidopsis* protoplasts as described (Wu et al., 2009). Images were visualized and obtained using confocal microscopy after 19 h of transformation. Combination with empty vector (nYFP of cYFP) were used as negative controls.

2.13. Immunoprecipitation coupled to mass spectrometry (IP-MS)

Overall procedure of IP-MS is similar with co-IP experiment. The first experiment

used 2g of WT and *CsVp::TAF15b-GFP* seedlings grown for 10 days in long days with GFP-Trap. The second experiment used 18g of *TAF15bp::TAF15b-GFP* seedlings grown for 10 days in long days and protein extract split in half and each half were applied to control bead (ChromoTek, #bab-20) and GFP-Trap. After immunoprecipitation and bead washing, 120 μl of 0.2M glycine (pH 2.5) were added to bead, vortexed for 30 s and centrifuged 13,000rpm for 1 min. Supernatant was transferred to a new tube and 10 μl of 1M Tris-HCl (pH 10.4) was applied to neutralize the solution. Eluted samples were sent to RNA-Proteomics Lab at Seoul national university and analyzed with liquid chromatography–mass spectrometry (LC-MS) after trypsin digestion.

2.14. GO term analysis

GO analysis of terms enriched for 386 genes, which were selected on two IP-MS experiments, was done using DAVID 6.8 and BiNGO (Huang et al., 2009a, 2009b; Maere et al., 2005).

CHAPTER VII

References

Albright, S.R., and Tjian, R. (2000). TAFs revisited: More data reveal new twists and confirm old ideas. Gene *242*, 1–13.

Ali, G.S., Palusa, S.G., Golovkin, M., Prasad, J., Manley, J.L., and Reddy, A.S.N. (2007). Regulation of Plant Developmental Processes by a Novel Splicing Factor. PLoS One *2*, e471.

Anderson, P., and Kedersha, N. (2009). RNA granules: post-transcriptional and epigenetic modulators of gene expression. Nat. Rev. Mol. Cell Biol. *10*, 430–436.

Andersson, M.K., Ståhlberg, A., Arvidsson, Y., Olofsson, A., Semb, H., Stenman, G., Nilsson, O., Aman, P., Stahlberg, A., Arvidsson, Y., et al. (2008). The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. BMC Cell Biol. *9*, 37.

Aukerman, M.J.., Lee, I., Weigel, D., and Amasino, R.M.. (1999). The Arabidopsis flowering-time gene LUMINIDEPENDENS is expressed primarily in regions of cell proliferation and encodes a nuclear protein that regulates LEAFY expression. Plant J. *18*, 195–203.

Ausín, I., Alonso-Blanco, C., Jarillo, J.A., Ruiz-García, L., and Martínez-Zapater, J.M. (2004). Regulation of flowering time by FVE, a retinoblastoma-associated protein. Nat. Genet. *36*, 162–166.

Bäurle, I., and Dean, C. (2006). The Timing of Developmental Transitions in Plants. Cell *125*, 655–664.

Berry, S., and Dean, C. (2015). Environmental perception and epigenetic memory: mechanistic insight through FLC. Plant J. *83*, 133–148.

Bertolotti, A., Lutz, Y., Heard, D.J., Chambon, P., and Tora, L. (1996). hTAF(II)68,

a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. EMBO J. *15*, 5022–5031.

Brengues, M., Teixeira, D., and Parker, R. (2005). Movement of Eukaryotic mRNAs Between Polysomes and Cytoplasmic Processing Bodies. Science (80-.). *310*, 486–489.

Burke, K.A., Janke, A.M., Rhine, C.L., and Fawzi, N.L. (2015). Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. Mol. Cell *60*, 231–241.

Burley, S.K., and Roeder, R.G. (1996). Biochemistry and Structural Biology of Transcription Factor IID (TFIID). Annu. Rev. Biochem. *65*, 769–799.

Chakrabortee, S., Kayatekin, C., Newby, G.A., Mendillo, M.L., Lancaster, A., and Lindquist, S. (2016). Luminidependens (LD) is an Arabidopsis protein with prion behavior. Proc. Natl. Acad. Sci. *113*, 6065–6070.

Cheng, J.-Z., Zhou, Y.-P., Lv, T.-X., Xie, C.-P., and Tian, C.-E. (2017). Research progress on the autonomous flowering time pathway in Arabidopsis. Physiol. Mol. Biol. Plants *23*, 477–485.

Cho, H.-T., and Cosgrove, D.J. (2002). Regulation of root hair initiation and expansin gene expression in Arabidopsis. Plant Cell *14*, 3237–3253.

Choi, K., Park, C., Lee, J., Oh, M., Noh, B., and Lee, I. (2007). Arabidopsis homologs of components of the SWR1 complex regulate flowering and plant development. Development *134*, 1931–1941.

Choi, K., Kim, J., Hwang, H.-J.H.J., Kim, S.S.Y.S.Y., Park, C., Kim, S.S.Y.S.Y., and Lee, I. (2011). The FRIGIDA complex activates transcription of FLC, a strong flowering repressor in Arabidopsis, by recruiting chromatin modification

factors. Plant Cell *23*, 289–303.

Clarke, J.H., and Dean, C. (1994). Mapping FRI, a locus controlling flowering time and vernalization response in Arabidopsis thaliana. Mol. Gen. Genet. *242*, 81–89.

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

Crevillén, P., Sonmez, C., Wu, Z., and Dean, C. (2012). A gene loop containing the floral repressor FLC is disrupted in the early phase of vernalization. EMBO J. *32*, 140–148.

Csorba, T., Questa, J.I., Sun, Q., Dean, C., Goodrich, J., and Ringrose, L. (2014). Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. Proc. Natl. Acad. Sci. *111*, 16160–16165.

Czesnick, H., and Lenhard, M. (2016). Antagonistic control of flowering time by functionally specialized poly(A) polymerases in Arabidopsis thaliana. Plant J. *88*, 570–583.

Decker, C.J., and Parker, R. (2012). P-Bodies and Stress Granules: Possible Roles in the Control of Translation and mRNA Degradation. Cold Spring Harb. Perspect. Biol. *4*, a012286–a012286.

Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H., and Keller, W. (2002). A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. Mol. Cell *10*, 1139–1150.

Dong, O.X., Meteignier, L.-V., Plourde, M.B., Ahmed, B., Wang, M., Jensen, C., Jin, H., Moffett, P., Li, X., and Germain, H. (2016). Arabidopsis TAF15b Localizes to RNA Processing Bodies and Contributes to snc1 -Mediated

Autoimmunity. Mol. Plant-Microbe Interact. *29*, 247–257.

Duncan, S., Holm, S., Questa, J., Irwin, J., Grant, A., and Dean, C. (2015). Seasonal shift in timing of vernalization as an adaptation to extreme winter. Elife *4*, e06620.

Eick, D., and Geyer, M. (2013). The RNA polymerase II carboxy-terminal domain (CTD) code. Chem. Rev. *113*, 8456–8490.

Eom, H., Park, S.J., Kim, M.K., Kim, H., Kang, H., and Lee, I. (2018). TAF15b, involved in the autonomous pathway for flowering, represses transcription of FLOWERING LOCUS C. Plant J. *93*, 79–91.

Feng, W., Jacob, Y., Veley, K.M., Ding, L., Yu, X., Choe, G., and Michaels, S.D. (2011). Hypomorphic alleles reveal FCA-independent roles for FY in the regulation of FLOWERING LOCUS C. Plant Physiol. *155*, 1425–1434.

Gazzani, S., Gendall, A.R., Lister, C., and Dean, C. (2003). Analysis of the Molecular Basis of Flowering Time Variation in Arabidopsis Accessions. Plant Physiol. *132*, 1107–1114.

Gleghorn, M.L., and Maquat, L.E. (2014). "Black sheep" that don't leave the double-stranded RNA-binding domain fold. Trends Biochem. Sci. *39*, 328–340.

Goodrich, J.A., and Tjian, R. (2010). Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. Nat. Rev. Genet. *11*, 549–558.

Grousl, T., Ivanov, P., Frydlova, I., Vasicova, P., Janda, F., Vojtova, J., Malinska, K., Malcova, I., Novakova, L., Janoskova, D., et al. (2009). Robust heat shock induces eIF2 -phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, Saccharomyces cerevisiae. J. Cell Sci. *122*, 2078–2088.

Hampsey, M., and Reinberg, D. (1997). Transcription: Why are TAFs essential? Curr. Biol. *7*, R44–R46.

Harlen, K.M., and Churchman, L.S. (2017). The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat. Rev. Mol. Cell Biol. *18*, 263–273.

He, Y. (2003). Regulation of Flowering Time by Histone Acetylation in Arabidopsis. Science *302*, 1751–1754.

Helliwell, C.A., Wood, C.C., Robertson, M., James Peacock, W., and Dennis, E.S. (2006). The Arabidopsis FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex. Plant J. *46*, 183–192.

Hepworth, S.R. (2002). Antagonistic regulation of flowering-time gene SOC1 by CONSTANS and FLC via separate promoter motifs. EMBO J. *21*, 4327–4337.

Hornyik, C., Terzi, L.C., and Simpson, G.G. (2010). The Spen Family Protein FPA Controls Alternative Cleavage and Polyadenylation of RNA. Dev. Cell *18*, 203–213.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009b). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. *37*, 1–13.

Jeon, J., and Kim, J. (2011). FVE, an Arabidopsis homologue of the retinoblastoma-associated protein that regulates flowering time and cold response, binds to chromatin as a large multiprotein complex. Mol. Cells *32*, 227–234.

Jiang, D., Gu, X., and He, Y. (2009). Establishment of the winter-annual growth habit via FRIGIDA-mediated histone methylation at FLOWERING LOCUS C in Arabidopsis. Plant Cell *21*, 1733–1746.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., and Dean, C. (2000). Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. Science *290*, 344–347.

Kim, H.-J., Hyun, Y., Park, J.-Y., Park, M.-J., Park, M.-K., Kim, M.D., Kim, H.-J., Lee, M.H., Moon, J., Lee, I., et al. (2004). A genetic link between cold responses and flowering time through FVE in Arabidopsis thaliana. Nat. Genet. *36*, 167–171.

Ko, J.-H., Mitina, I., Tamada, Y., Hyun, Y., Choi, Y., Amasino, R.M., Noh, B., and Noh, Y.-S. (2010). Growth habit determination by the balance of histone methylation activities in Arabidopsis. EMBO J. *29*, 3208–3215.

Koiwa, H., Barb, A.W., Xiong, L., Li, F., McCully, M.G., Lee, B.-H., Sokolchik, I., Zhu, J., Gong, Z., Reddy, M., et al. (2002). C-terminal domain phosphatase-like family members (AtCPLs) differentially regulate Arabidopsis thaliana abiotic stress signaling, growth, and development. Proc. Natl. Acad. Sci. U. S. A. *99*, 10893–10898.

Koiwa, H., Hausmann, S., Bang, W.Y., Ueda, A., Kondo, N., Hiraguri, A., Fukuhara, T., Bahk, J.D., Yun, D.-J., Bressan, R.A., et al. (2004). Arabidopsis C-terminal domain phosphatase-like 1 and 2 are essential Ser-5-specific C-terminal domain phosphatases. Proc. Natl. Acad. Sci. *101*, 14539–14544.

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., Peeters, A.J.M., and Soppe, W. (1998).

GENETIC CONTROL OF FLOWERING TIME IN ARABIDOPSIS. Annu. Rev. Plant Physiol. Plant Mol. Biol. *49*, 345–370.

Kovar, H. (2011). Dr. Jekyll and Mr. Hyde: The Two Faces of the FUS/EWS/TAF15 Protein Family. Sarcoma *2011*, 1–13.

Kwon, I., Kato, M., Xiang, S., Wu, L., Theodoropoulos, P., Mirzaei, H., Han, T., Xie, S., Corden, J.L., and McKnight, S.L. (2013). Phosphorylation-Regulated Binding of RNA Polymerase II to Fibrous Polymers of Low-Complexity Domains. Cell *155*, 1049–1060.

Lago, C., Clerici, E., Mizzi, L., Colombo, L., and Kater, M.M. (2004). TBP-associated factors in Arabidopsis. Gene *342*, 231–241.

Lago, C., Clerici, E., Dreni, L., Horlow, C., Caporali, E., Colombo, L., and Kater, M.M. (2005). The Arabidopsis TFIID factor AtTAF6 controls pollen tube growth. Dev. Biol. *285*, 91–100.

Lee, H.G., and Seo, P.J. (2016). The Arabidopsis MIEL1 E3 ligase negatively regulates ABA signalling by promoting protein turnover of MYB96. Nat. Commun. *7*, 12525.

Lee, J., and Lee, I. (2010). Regulation and function of SOC1, a flowering pathway integrator. J. Exp. Bot. *61*, 2247–2254.

Lee, I., Aukerman, M.J., Gore, S.L., Lohman, K.N., Michaels, S.D., Weaver, L.M., John, M.C., Feldmann, K.A., and Amasino, R.M. (1994). lsolation of LUMINIDEPENDENS: A Gene lnvolved in the Control of Flowering Time in Arabidopsis. Plant Cell *6*, 75–83.

Lee, O.R., Kim, S.J., Kim, H.J., Hong, J.K., Ryu, S.B., Lee, S.H., Ganguly, A., and Cho, H.-T. (2010). Phospholipase A $_2$ Is Required for PIN-FORMED Protein Trafficking to the Plasma Membrane in the *Arabidopsis* Root. Plant Cell *22*,

1812–1825.

Lempe, J., Balasubramanian, S., Sureshkumar, S., Singh, A., Schmid, M., and Weigel, D. (2005). Diversity of Flowering Responses in Wild Arabidopsis thaliana Strains. PLoS Genet. *1*, e6.

Li, P., Filiault, D., Box, M.S., Kerdaffrec, E., van Oosterhout, C., Wilczek, A.M., Schmitt, J., McMullan, M., Bergelson, J., Nordborg, M., et al. (2014). Multiple FLC haplotypes defined by independent cis -regulatory variation underpin life history diversity in Arabidopsis thaliana. Genes Dev. *28*, 1635–1640.

Lim, M.-H. (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 ONLINE *16*, 731–740.

Lin, X., Tirichine, L., and Bowler, C. (2012). Protocol: Chromatin immunoprecipitation (ChIP) methodology to investigate histone modifications in two model diatom species. Plant Methods *8*, 48.

Liu, F., Quesada, V., Crevillén, P., Bäurle, I., Swiezewski, S., and Dean, C. (2007). The Arabidopsis RNA-Binding Protein FCA Requires a Lysine-Specific Demethylase 1 Homolog to Downregulate FLC. Mol. Cell *28*, 398–407.

Liu, F., Marquardt, S., Lister, C., Swiezewski, S., and Dean, C. (2010). Targeted 3' Processing of Antisense Transcripts Triggers Arabidopsis FLC Chromatin Silencing. Science *327*, 94–97.

Macknight, R., Bancroft, I., Page, T., Lister, C., Schmidt, R., Love, K., Westphal, L., Murphy, G., Sherson, S., Cobbett, C., et al. (1997). FCA, a Gene Controlling Flowering Time in Arabidopsis, Encodes a Protein Containing RNA-Binding Domains. Cell *89*, 737–745.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: A Cytoscape plugin to

assess overrepresentation of Gene Ontology categories in Biological Networks. Bioinformatics *21*, 3448–3449.

Manavella, P.A., Hagmann, J., Ott, F., Laubinger, S., Franz, M., Macek, B., and Weigel, D. (2012). Fast-forward genetics identifies plant CPL phosphatases as regulators of miRNA processing factor HYL1. Cell *151*, 859–870.

Manzano, D., Marquardt, S., Jones, A.M.E., Bäurle, I., Liu, F., Dean, C., Baurle, I., Liu, F., and Dean, C. (2009). Altered interactions within FY/AtCPSF complexes required for Arabidopsis FCA-mediated chromatin silencing. Proc. Natl. Acad. Sci. *106*, 8772–8777.

March, Z.M., King, O.D., and Shorter, J. (2016). Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. Brain Res. *1647*, 9–18.

Marko, M., Vlassis, A., Guialis, A., and Leichter, M. (2012). Domains involved in TAF15 subcellular localisation: Dependence on cell type and ongoing transcription. Gene *506*, 331–338.

Marquardt, S., Raitskin, O., Wu, Z., Liu, F., Sun, Q., and Dean, C. (2014). Functional Consequences of Splicing of the Antisense Transcript COOLAIR on FLC Transcription. Mol. Cell *54*, 156–165.

Matsuda, O., Sakamoto, H., Nakao, Y., Oda, K., and Iba, K. (2009). CTD phosphatases in the attenuation of wound-induced transcription of jasmonic acid biosynthetic genes in Arabidopsis. Plant J. *57*, 96–108.

Mayfield, J.E., Burkholder, N.T., and Zhang, Y.J. (2016). Dephosphorylating eukaryotic RNA polymerase II. Biochim. Biophys. Acta - Proteins Proteomics *1864*, 372–387.

Michaels, S.D. (1999). FLOWERING LOCUS C Encodes a Novel MADS Domain

Protein That Acts as a Repressor of Flowering. PLANT CELL ONLINE *11*, 949–956.

Michaels, S.D. (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 ONLINE *13*, 935–942.

Michaels, S.D. (2009). Flowering time regulation produces much fruit. Curr. Opin. Plant Biol. *12*, 75–80.

Michaels, S.D. (2016). Regulation of flowering time in chromatin structure (Seoul national university).

Michaels, S.D., He, Y., Scortecci, K.C., and Amasino, R.M. (2003). Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. Proc. Natl. Acad. Sci. *100*, 10102–10107.

Mockler, T.C., Yu, X., Shalitin, D., Parikh, D., Michael, T.P., Liou, J., Huang, J., Smith, Z., Alonso, J.M., Ecker, J.R., et al. (2004). Regulation of flowering time in Arabidopsis by K homology domain proteins. Proc. Natl. Acad. Sci. *101*, 12759–12764.

Oh, M., and Lee, U. (2007). Historical perspective on breakthroughs in flowering field. J. Plant Biol. *50*, 249–256.

Perales, R., and Bentley, D. (2009). "Cotranscriptionality": The Transcription Elongation Complex as a Nexus for Nuclear Transactions. Mol. Cell *36*, 178–191.

Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev. *20*, 2922–2936.

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.

Rataj, K., and Simpson, G.G. (2014). Message ends: RNA 3′ processing and flowering time control. J. Exp. Bot. *65*, 353–363.

Rouse, D.T., Sheldon, C.C., Bagnall, D.J., Peacock, W.J., and Dennis, E.S. (2002). FLC, a repressor of flowering, is regulated by genes in different inductive pathways. Plant J. *29*, 183–191.

Saleh, A., Alvarez-Venegas, R., and Avramova, Z. (2008). An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. Nat. Protoc. *3*, 1018–1025.

Sanda, S.L., and Amasino, R.M. (1996). Ecotype-Specific Expression of a Flowering Mutant Phenotype in Arabidopsis thaliana. Plant Physiol. *111*, 641–644.

Schomburg, F.M., Patton, D.A., Meinke, D.W., and Amasino, R.M. (2001). FPA, a gene involved in floral induction in Arabidopsis, encodes a protein containing RNA-recognition motifs. Plant Cell *13*, 1427–1436.

Schüller, R., and Eick, D. (2016). Getting Access to Low-Complexity Domain Modifications. Trends Biochem. Sci. *41*, 894–897.

Schwartz, J.C., Ebmeier, C.C., Podell, E.R., Heimiller, J., Taatjes, D.J., and Cech, T.R. (2012). FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. Genes Dev. *26*, 2690–2695.

Schwartz, J.C., Wang, X., Podell, E.R., and Cech, T.R. (2013). RNA Seeds Higher-Order Assembly of FUS Protein. Cell Rep. *5*, 918–925.

Schwartz, J.C., Cech, T.R., and Parker, R.R. (2015). Biochemical Properties and

Biological Functions of FET Proteins. Annu. Rev. Biochem. *84*, 355–379.

Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J., Dennis, E.S., Feldmann, K.A., and Amasino, R.M. (1999). The FLF MADS Box Gene: A Repressor of Flowering in Arabidopsis Regulated by Vernalization and Methylation. PLANT CELL ONLINE *11*, 445–458.

Shindo, C., Lister, C., Crevillen, P., Nordborg, M., and Dean, C. (2006). Variation in the epigenetic silencing of FLC contributes to natural variation in Arabidopsis vernalization response. Genes Dev. *20*, 3079–3083.

Simpson, G.G. (2004). The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of Arabidopsis flowering time. Curr. Opin. Plant Biol. *7*, 570–574.

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.

Streitner, C., Danisman, S., Wehrle, F., Schöning, J.C., Alfano, J.R., and Staiger, D. (2008). The small glycine-rich RNA binding protein At GRP7 promotes floral transition in Arabidopsis thaliana. Plant J. *56*, 239–250.

Svetoni, F., Frisone, P., and Paronetto, M.P. (2016). Role of FET proteins in neurodegenerative disorders. RNA Biol. *13*, 1089–1102.

Swiezewski, S., Liu, F., Magusin, A., and Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature *462*, 799–802.

Tan, A.Y., and Manley, J.L. (2009). The TET family of proteins: Functions and roles in disease. J. Mol. Cell Biol. *1*, 82–92.

Tan-Wong, S.M., Zaugg, J.B., Camblong, J., Xu, Z., Zhang, D.W., Mischo, H.E., Ansari, A.Z., Luscombe, N.M., Steinmetz, L.M., and Proudfoot, N.J. (2012). Gene Loops Enhance Transcriptional Directionality. Science (80-.). *338*, 671–675.

Valente, L., and Nishikura, K. (2007). RNA Binding-independent Dimerization of Adenosine Deaminases Acting on RNA and Dominant Negative Effects of Nonfunctional Subunits on Dimer Functions. J. Biol. Chem. *282*, 16054–16061.

Verdaguer, B., de Kochko, A., Beachy, R.N., and Fauquet, C. (1996). Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. Plant Mol. Biol. *31*, 1129–1139.

Wang, C., Tian, Q., Hou, Z., Mucha, M., Aukerman, M., and Olsen, O.A. (2007). The Arabidopsis thaliana AT PRP39-1 gene, encoding a tetratricopeptide repeat protein with similarity to the yeast pre-mRNA processing protein PRP39, affects flowering time. Plant Cell Rep. *26*, 1357–1366.

Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature *454*, 126–130.

Wang, Z.-W., Wu, Z., Raitskin, O., Sun, Q., and Dean, C. (2014). Antisense-mediated FLC transcriptional repression requires the P-TEFb transcription elongation factor. Proc. Natl. Acad. Sci. *111*, 7468–7473.

Waterworth, W.M., Drury, G.E., Blundell-Hunter, G., and West, C.E. (2015). Arabidopsis TAF1 is an MRE11-interacting protein required for resistance to genotoxic stress and viability of the male gametophyte. Plant J. *84*, 545–557.

Wu, F.H., Shen, S.C., Lee, L.Y., Lee, S.H., Chan, M.T., and Lin, C.S. (2009). Tape-arabidopsis sandwich - A simpler arabidopsis protoplast isolation method. Plant

Methods *5*, 1–10.

Wu, Z., Ietswaart, R., Liu, F., Yang, H., Howard, M., and Dean, C. (2016). Quantitative regulation of FLC via coordinated transcriptional initiation and elongation. Proc. Natl. Acad. Sci. *113*, 218–223.

Xiang, K., Nagaike, T., Xiang, S., Kilic, T., Beh, M.M., Manley, J.L., and Tong, L. (2010). Crystal structure of the human symplekin-Ssu72-CTD phosphopeptide complex. Nature *467*, 729–733.

Xing, D., Zhao, H., Xu, R., and Li, Q.Q. (2008). Arabidopsis PCFS4, a homologue of yeast polyadenylation factor Pcf11p, regulates FCA alternative processing and promotes flowering time. Plant J. *54*, 899–910.

Yang, L., Gal, J., Chen, J., and Zhu, H. (2014). Self-assembled FUS binds active chromatin and regulates gene transcription. Proc. Natl. Acad. Sci. *111*, 17809–17814.

Zaborowska, J., Egloff, S., and Murphy, S. (2016). The pol II CTD: new twists in the tail. Nat. Struct. Mol. Biol. *23*, 771–777.

Zhang, D.W., Mosley, A.L., Ramisetty, S.R., Rodríguez-Molina, J.B., Washburn, M.P., and Ansari, A.Z. (2012). Ssu72 phosphatase-dependent erasure of phospho-Ser7 marks on the RNA polymerase II C-terminal domain is essential for viability and transcription termination. J. Biol. Chem. *287*, 8541–8551.

Appendix

1. Amino acid sequences of *Arabidopsis* TAF15 and TAF15 b

The underlined amino acid sequences indicate low complexity domain.

>TAF15_ AT1G50300.1
MAGYPTNGSVYVSNLPLGTDENMLADYFGTIGLLKRDKRTGTPKVWLYR
DKETDEPKGDATVTYEDPHAALAAVEWFNNKDFHGNTIGVFMAESKNKN
AGDAVEFVEFDGGAEETNGGAGRGRGQADSSAKPWQQDGDWMCPNTSC
TNVNFAFRGVCNRCGTARPAGASGGSMGAGRGRGRGGGADGGAPGKQPS
GAPTGLFGPNDWACPMCGNVNWAKRLKCNICNTNKPGQNEGGVRGGRG
GGYKELDEQELEETKRRRREAEEDDGEMYDEFGNLKKKYRVKTNQADTR
PAVAAGRAGWEVEELGIDKDGRERSRDRQRDRGRDHHYDKDRRRSRSRER
ERGKERDYDYDHDRDRDRDYGRERGSRYRN

>TAF15b_AT5G58470.1
<u>MAGMYNQDGGGGAPIPSYGGDGYGGGGGYGGGDAGYGGRGASGGGSY
GGRGGYGGGGGRGNRGGGGGGYQGGDRGGRGSGGGGRDG</u>DWRCPNPS
CGNVNFARRVECNKCGALAP<u>SGTSSGANDRGGGGYSRGGGDSDRGGGRG
GRNDSGRSYESSRYDGGSRSGGSYGSGSQRENGSYGQAPPPAAAIPSYDGS
GSYPPPTGYGMEAVPPPTSYSGGPPSYGGPRGGYGSDAPSTGGRGGRSGGY
DGGSAPRRQEASYED</u>AATEKVKQCDADCDDNCDNARIYISNLPPDVTTDE
LKDLFGGIGQVGRIKQKRGYKDQWPYNIKIYTDEKGNYKGDACLAYEDPS
AAHSAGGFFNNYEMRGNKISVTMAEKSAPRAPTFDQRGGGRGGGGGGYG
GGGGDRRRDNYSSGPDRNHHGGNRSRPY

2. Amino acid sequences of human FET proteins

The underlined amino acid sequences indicate low complexity domain.

>FUS_NP004951.1
MASNDYTQQATQSYGAYPTQPGQGYSQQSSQPYGQQSYSGYSQSTDTSGY
GQSSYSSYGQSQNTGYGTQSTPQGYGSTGGYGSSQSSQSSYGQQSSYPGY
GQQPAPSSTSGSYGSSSQSSSYGQPQSGSYSQQPSYGGQQQSYGQQQSYNP
PQGYGQQNQYNSSSGGGGGGGGGGNYGQDQSSMSSGGGSGGGYGNQDQ
SGGGGSGGYGQQDRGGRGRGGSGGGGGGGGGGYNRSSGGYEPRGRGGG
RGGRGGMGGSDRGGFNKFGGPRDQGSRHDSEQDNSDNNTIFVQGLGENV
TIESVADYFKQIGIIKTNKKTGQPMINLYTDRETGKLKGEATVSFDDPPSAK
AAIDWFDGKEFSGNPIKVSFATRRADFNRGGGNGRGGRGRGGPMGRGGY
GGGGSGGGGRGGFPSGGGGGGGQQRAGDWKCPNPTCENMNFSWRNECN
QCKAPKPDGPGGGPGGSHMGGNYGDDRRGGRGGYDRGGYRGRGGDRG
GFRGGRGGGDRGGFGPGKMDSRGEHRQDRRERPY

>EWSR1_NP053733.2
MASTDYSTYSQAAAQQGYSAYTAQPTQGYAQTTQAYGQQSYGTYGQPTD
VSYTQAQTTATYGQTAYATSYGQPPTVEGTSTGYTTPTAPQAYSQPVQGYG
TGAYDTTTATVTTTQASYAAQSAYGTQPAYPAYGQQPAATAPTRPQDGNKP
TETSQPQSSTGGYNQPSLGYGQSNYSYPQVPGSYPMQPVTAPPSYPPTSYSS
TQPTSYDQSSYSQQNTYGQPSSYGQQSSYGQQSSYGQQPPTSYPPQTGSYS
QAPSQYSQQSSSYGQQSSFRQDHPSSMGVYGQESGGFSGPGENRSMSGPD
NRGRGRGGFDRGGMSRGGRGGGRGGMGAGERGGFNKPGGPMDEGPDLD
LGPPVDPDEDSDNSAIYVQGLNDSVTLDDLADFFKQCGVVKMNKRTGQP
MIHIYLDKETGKPKGDATVSYEDPPTAKAAVEWFDGKDFQGSKLKVSLAR
KKPPMNSMRGGLPPREGRGMPPPLRGGPGGPGGPGGPMGRMGGRGGDR
GGFPPRGPRGSRGNPSGGGNVQHRAGDWQCPNPGCGNQNFAWRTECNQC

KAPKPEGFLPPPFPPPGGDRGRGGPGGMRGGRGGLMDRGGPGGMFRGGR
GGDRGGFRGGRGMDRGGFGGGRRGGPGGPPGPLMEQMGGRRGGRGGPG
KMDKGEHRQERRDRPY

>TAF15_NP631961.1
MSDSGSYGQSGGEQQSYSTYGNPGSQGYGQASQSYSGYGQTTDSSYGQN
YSGYSSYGQSQSGYSQSYGGYENQKQSSYSQQPYNNQGQQQNMESSGSQ
GGRAPSYDQPDYGQQDSYDQQSGYDQHQGSYDEQSNYDQQHDSYSQNQ
QSYHSQRENYSHHTQDDRRDVSRYGEDNRGYGGSQGGGRGRGGYDKDG
RGPMTGSSGGDRGGFKNFGGHRDYGPRTDADSESDNSDNNTIFVQGLGEG
VSTDQVGEFFKQIGIIKTNKKTGKPMINLYTDKDTGKPKGEATVSFDDPPSA
KAAIDWFDGKEFHGNIIKVSFATRRPEFMRGGGSGGGRRGRGGYRGRGGF
QGRGGDPKSGDWVCPNPSCGNMNFARRNSCNQCNEPRPEDSRPSGGDFRG
RGYGGERGYRGRGGRGGDRGGYGGDRSGGGYGGDRSSGGGYSGDRSGG
GYGGDRSGGGYGGDRGGGYGGDRGGGYGGDRGGGYGGDRGGYGGDRG
GGYGGDRGGYGGDRGGYGGDRGGYGGDRGGYGGDRSRGGYGGDRGGG
SGYGGDRSGGYGGDRSGGGYGGDRGGGYGGDRGGYGGKMGGRNDYRN
DQRNRPY

3. List of co-immunoprecipitated proteins with TAF15b

386 proteins that were identified from both two independent purifications listed. The first purification used WT and *CsVp::TAF15b-GFP* seedlings grown for 10 days in long days with GFP-Trap. The second purification used *TAF15bp::TAF15b-GFP* seedlings with binding control and GFP-Trap.

no.	AT number	Descriptive name	Peptides number
1	AT5G58470	TBP-associated factor 15B (TAF15b)	968/228
2	AT1G79930	heat shock protein 91 (HSP91)	117/17
3	AT3G11910	ubiquitin-specific protease 13 (UBP13)	75/12
4	AT5G20890	TCP-1/cpn60 chaperonin family protein (CCT2)	61/10
5	AT5G26742	DEAD box RNA helicase (RH3)	40/30
6	AT5G04280	RNA-BINDING GLYCINE-RICH PROTEIN B3 (RBGB3)	40/19
7	AT5G06600	ubiquitin-specific protease 12 (UBP12)	50/5
8	AT3G18190	TCP-1/cpn60 chaperonin family protein	38/12
9	AT5G58290	regulatory particle triple-A ATPase 3 (RPT3)	42/7
10	AT4G02890	Ubiquitin family protein (UBQ14)	25/18
11	AT5G53460	NADH-dependent glutamate synthase 1 (GLT1)	36/4
12	AT2G36530	Enolase (ENO2)	30/8
13	AT5G16070	TCP-1/cpn60 chaperonin family protein (CCT6)	31/7
14	AT4G28520	cruciferin 3 (CRU3)	36/1
15	AT3G03960	TCP-1/cpn60 chaperonin family protein (CCT8)	28/8
16	AT2G20140	AAA-type ATPase family protein	33/3
17	AT1G45000	AAA-type ATPase family protein (RPT4)	25/9
18	AT3G20050	T-complex protein 1 alpha subunit (TCP-1)	26/7
19	AT5G53060	RNA-binding KH domain-containing protein (RCF3)	21/9
20	AT2G46280	TGF-beta receptor interacting protein 1 (TRIP-1)	25/5
21	AT4G11420	eukaryotic translation initiation factor 3A (EIF3A)	26/4
22	AT2G17870	cold shock domain protein 3 (CSP3)	24/6
23	**AT4G21670**	**C-terminal domain phosphatase-like 1 (CPL1)**	**24/5**
24	AT3G58510	DEA(D/H)-box RNA helicase family protein (RH11)	16/9
25	AT1G50480	10-formyltetrahydrofolate synthetase (THFS)	22/3
26	AT1G78900	vacuolar ATP synthase subunit A (VHA-A)	16/8
27	AT3G59770	sacI homology domain-containing protein (SAC9)	23/1
28	AT1G06220	Ribosomal protein S5/Elongation factor G/III/V family protein (MEE5)	19/5
29	AT3G58570	DEAD-box ATP-dependent RNA helicase 52 (RH52)	12/11
30	AT5G55220	trigger factor type chaperone family protein	18/5

31	AT5G44120	RmlC-like cupins superfamily protein (CRA1)	22/1
32	AT4G36020	cold shock domain protein 1 (CSDP1)	18/4
33	AT1G75660	5'-3' exoribonuclease 3 (XRN3)	18/4
34	AT4G25550	Cleavage/polyadenylation specificity factor, 25kDa subunit 2 (CFIS2)	15/7
35	AT2G39800	delta1-pyrroline-5-carboxylate synthase 1 (P5CS1)	15/6
36	AT5G35790	glucose-6-phosphate dehydrogenase 1 (G69D1)	14/7
37	AT5G19990	regulatory particle triple-A ATPase 6A (RPT6A)	17/4
38	ATCG00500	acetyl-CoA carboxylase carboxyl transferase subunit beta (ACCD)	17/3
39	AT4G25630	fibrillarin 2 (FIB2)	15/5
40	AT3G62120	Class II aaRS and biotin synthetases superfamily protein	15/4
41	AT4G13430	isopropyl malate isomerase large subunit 1 (IIL1)	18/1
42	AT1G29900	carbamoyl phosphate synthetase B (CARB)	15/4
43	AT1G09100	26S proteasome AAA-ATPase subunit RPT5B (RPT5B)	12/6
44	AT1G06410	trehalose-phosphatase/synthase 7 (TPS7)	17/1
45	AT2G47940	DEGP protease 2 (DEGP2)	11/7
46	AT1G13440	glyceraldehyde-3-phosphate dehydrogenase C2 (GAPC2)	15/3
47	AT2G27100	C2H2 zinc-finger protein SERRATE (SE)	12/5
48	AT1G23310	glutamate:glyoxylate aminotransferase (GGT1)	12/5
49	AT4G11150	vacuolar ATP synthase subunit E1 (TUF)	13/4
50	AT5G21326	Ca2+regulated serine-threonine protein kinase family protein	15/2
51	AT3G16950	lipoamide dehydrogenase 1 (LPD1)	13/4
52	AT2G33800	Ribosomal protein S5 family protein (RPS5)	14/2
53	AT3G62310	ATP-dependent RNA helicase DEAH2 (DDX15)	14/2
54	AT3G24503	aldehyde dehydrogenase 2C4 (ALDH2C4)	15/1
55	AT3G24430	ATP binding (HCF101)	15/1
56	AT1G64520	regulatory particle non-ATPase 12A (RPN12a)	10/5
57	AT4G20130	plastid transcriptionally active 14 (PTAC14)	12/3
58	AT3G05530	regulatory particle triple-A ATPase 5A (RPT5A)	12/3
59	AT1G08520	ALBINA 1 (ALB1)	13/2
60	AT1G62750	Translation elongation factor EFG/EF2 protein (SCO1)	13/2
61	AT2G04842	threonyl-tRNA synthetase (EMB2761)	14/1
62	AT3G06850	2-oxoacid dehydrogenases acyltransferase family protein (BCE2)	12/3
63	AT4G16143	importin alpha isoform 2 (IMPA-2)	13/2
64	AT3G56150	eukaryotic translation initiation factor 3C (EIF3C)	13/2
65	AT1G32500	non-intrinsic ABC protein 6 (NAP6)	13/2
66	AT1G79920	Heat shock protein 70 (Hsp 70) family protein	14/1
67	AT1G51690	protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform (ATB)	14/1
68	AT5G47010	putative RNA helicase (UPF1)	13/2
69	AT2G14750	APS kinase (APK)	13/1
70	AT1G20110	RING/FYVE/PHD zinc finger superfamily protein (FREE1)	13/1

71	ATCG00180	DNA-directed RNA polymerase family protein (RPOC1)	11/3
72	AT4G27000	RNA-binding (RRM/RBD/RNP motifs) family protein (RBP45C)	12/2
73	AT1G16350	Aldolase-type TIM barrel family protein	13/1
74	AT3G17390	S-adenosylmethionine synthetase family protein (MTO3)	10/3
75	AT1G62640	3-ketoacyl-acyl carrier protein synthase III (KAS III)	10/3
76	AT4G01850	S-adenosylmethionine synthetase 2 (SAM-2)	9/4
77	AT4G21710	DNA-directed RNA polymerase family protein (NRPB2)	10/2
78	AT3G48110	glycine-tRNA ligases (EDD1)	10/2
79	AT1G62020	Coatomer, alpha subunit	11/1
80	ATCG00740	RNA polymerase subunit alpha (RPOA)	9/3
81	AT2G31810	ACT domain-containing small subunit of acetolactate synthase protein	9/3
82	AT3G02360	6-phosphogluconate dehydrogenase family protein	10/2
83	AT1G13320	protein phosphatase 2A subunit A3 (PP2AA3)	11/1
84	AT3G02530	TCP-1/cpn60 chaperonin family protein (CCT6)	10/2
85	AT4G38780	Pre-mRNA-processing-splicing factor (PRP8B)	10/2
86	AT1G09430	ATP-citrate lyase A-3 (ACLA-3)	9/3
87	AT3G27830	ribosomal protein L12-A (RPL12-A)	10/2
88	AT2G18020	Ribosomal protein L2 family (RPL8A)	10/2
89	AT5G54190	protochlorophyllide oxidoreductase A (PORA)	10/2
90	AT2G38270	CAX-interacting protein 2 (CXIP2)	10/2
91	AT2G21060	glycine-rich protein 2B (GRP2B)	11/1
92	AT1G32220	NAD(P)-binding Rossmann-fold superfamily protein	5/7
93	AT4G31490	Coatomer, beta subunit	9/3
94	AT2G34640	plastid transcriptionally active 12 (PTAC12)	10/2
95	AT4G13940	S-adenosyl-L-homocysteine hydrolase (SAHH1)	10/2
96	AT3G48870	Clp ATPase (CLPC)	9/3
97	AT5G20290	Ribosomal protein S8e family protein (RPS8A)	7/4
98	AT5G02870	Ribosomal protein L4/L1 family (RPL4D)	9/2
99	AT4G28470	26S proteasome regulatory subunit S2 1B (RPN1B)	10/1
100	AT5G13650	elongation factor family protein	6/5
101	AT3G57150	homologue of NAP57, Plays a role in ribosomal RNA processing	6/5
102	AT2G36250	Tubulin/FtsZ family protein (FTSZ2-1)	7/4
103	AT5G64040	photosystem I reaction center subunit PSI-N (PSAN)	8/3
104	AT5G14780	formate dehydrogenase (FDH)	8/3
105	AT1G20950	Phosphofructokinase family protein	8/3
106	AT1G65960	glutamate decarboxylase 2 (GAD2)	9/2
107	AT2G20580	26S proteasome regulatory subunit S2 1A (RPN1A)	9/2
108	AT1G74960	fatty acid biosynthesis 1 (FAB1)	10/1
109	AT3G04840	Ribosomal protein S3Ae	10/1
110	AT3G07100	Sec23/Sec24 protein transport family protein (SEC24A)	10/1
111	AT3G48560	chlorsulfuron/imidazolinone resistant 1 (CSR1)	9/2
112	ATCG00900	Ribosomal protein S7p/S5e family protein (RPS7.1)	10/1

113	AT5G23010	methylthioalkylmalate synthase 1 (MAM1)	10/1
114	AT3G15000	Encodes RNA-editing factor interacting protein 1 (MORF8)	10/1
115	AT1G10200	GATA type zinc finger transcription factor family protein (WLIM1)	10/1
116	AT5G10450	G-box regulating factor 6 (GRF6)	9/2
117	AT1G01090	pyruvate dehydrogenase E1 alpha (PDH-E1)	7/3
118	AT1G74470	Pyridine nucleotide-disulphide oxidoreductase family protein	7/3
119	AT1G15930	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein	7/3
120	AT1G48600	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein (PMEAMT)	8/2
121	AT2G34160	Alba DNA/RNA-binding protein	7/3
122	AT1G63940	monodehydroascorbate reductase 6 (MDAR6)	7/3
123	AT2G42520	DEAD-box ATP-dependent RNA helicase 37 (RH37)	4/6
124	AT1G69200	fructokinase-like 2 (FLN2)	7/3
125	AT3G13120	Ribosomal protein S10p/S20e family protein	7/3
126	AT1G11650	RNA-binding (RRM/RBD/RNP motifs) family protein (RBP45B)	8/2
127	AT5G12250	beta-6 tubulin (TUB6)	9/1
128	ATCG00160	ribosomal protein S2 (RPS2)	8/2
129	ATCG00800	structural constituent of ribosome	8/2
130	AT1G11720	starch synthase 3 (SS3)	9/1
131	ATCG00750	ribosomal protein S11 (RPS11)	5/4
132	AT3G14790	rhamnose biosynthesis 3 (RHM3)	6/3
133	AT5G11200	DEAD/DEAH box RNA helicase family protein (RH56)	5/4
134	AT5G58040	homolog of yeast FIP1 [V] (ATFIP1[V])	7/2
135	AT5G19550	aspartate aminotransferase 2 (ASP2)	7/2
136	AT3G19760	eukaryotic initiation factor 4A-III (EIF4A-III)	7/2
137	AT5G54160	O-methyltransferase 1 (OMT1)	8/1
138	AT3G61240	DEA(D/H)-box RNA helicase family protein (RH12)	5/4
139	AT4G01370	MAP kinase 4 (MPK4)	7/2
140	AT1G18450	actin-related protein 4 (ARP4)	7/2
141	AT3G54090	fructokinase-like 1 (FLN1)	7/2
142	AT1G30580	GTP binding	7/2
143	AT3G48500	Nucleic acid-binding, OB-fold-like protein (PDE312)	8/1
144	AT5G26710	Glutamyl/glutaminyl-tRNA synthetase, class Ic	8/1
145	AT2G47510	fumarase 1 (FUM1)	8/1
146	AT3G26420	Glycine-rich RNA-binding protein RZ1A (ATRZ-1A)	8/1
147	AT2G36880	methionine adenosyltransferase 3 (MAT3)	5/4
148	AT2G39990	eukaryotic translation initiation factor 2 (EIF2)	6/3
149	AT4G20890	tubulin beta-9 chain (TUB9)	6/3
150	AT5G16715	valine-tRNA ligases (EMB2247)	7/2
151	AT4G15560	Deoxyxylulose-5-phosphate synthase (CLA1)	8/1
152	AT5G45930	magnesium chelatase i2 (CHLI2)	8/1
153	AT2G05520	glycine-rich protein 3 (GRP3)	2/6

154	AT2G36870	xyloglucan endotransglucosylase/hydrolase 32 (XTH32)	5/3
155	AT5G06290	2-cysteine peroxiredoxin B (2CPB)	6/2
156	AT5G15200	Ribosomal protein S4 (RPS9B)	6/2
157	AT2G23350	poly(A) binding protein 4 (PAB4)	7/1
158	AT4G26450	unknown protein	7/1
159	AT1G49760	poly(A) binding protein 8 (PAB8)	7/1
160	AT5G03940	chloroplast signal recognition particle 54 kDa subunit (SRP54CP)	4/4
161	AT4G00660	RNAhelicase-like 8 (RH8)	6/2
162	AT1G10840	translation initiation factor 3 subunit H1 (TIF3H1)	6/2
163	AT4G34110	poly(A) binding protein 2 (PAB2)	7/1
164	AT2G15430	DNA-directed RNA polymerase family protein (RBP36A)	7/1
165	AT5G41670	6-phosphogluconate dehydrogenase family protein	7/1
166	AT1G48030	mitochondrial lipoamide dehydrogenase 1 (mtLPD1)	4/4
167	AT3G52960	Thioredoxin superfamily protein	6/2
168	AT1G64740	alpha-1 tubulin (TUA1)	7/1
169	AT4G27150	seed storage albumin 2 (SESA2)	7/1
170	AT5G54770	thiazole biosynthetic enzyme, chloroplast (THI1)	7/1
171	AT2G32060	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein	7/1
172	AT1G66270	Glycosyl hydrolase superfamily protein (RGLU21)	7/1
173	AT4G16130	arabinose kinase (ARA1)	5/2
174	AT2G19940	oxidoreductases	4/3
175	AT2G16570	GLN phosphoribosyl pyrophosphate amidotransferase 1 (ASE1)	4/3
176	AT1G50200	Alanyl-tRNA synthetase (ALATS)	5/2
177	AT3G55460	SC35-like splicing factor 30 (SCL30)	5/2
178	AT3G19980	flower-specific, phytochrome-associated protein phosphatase 3 (FYPP3)	5/2
179	AT2G32730	26S proteasome regulatory complex	6/1
180	AT2G34590	Transketolase family protein	4/3
181	AT3G46970	alpha-glucan phosphorylase 2 (PHS2)	4/3
182	AT5G51720	2 iron, 2 sulfur cluster binding	5/2
183	AT1G35580	cytosolic invertase 1 (CINV1)	5/2
184	AT4G38630	regulatory particle non-ATPase 10 (RPN10)	6/1
185	AT1G09130	ATP-dependent caseinolytic (Clp) protease/crotonase family protein	6/1
186	AT1G12000	Phosphofructokinase family protein	6/1
187	AT1G70850	MLP-like protein 34 (MLP34)	6/1
188	ATCG00065	ribosomal protein S12A (RPS12A)	5/2
189	AT3G06730	Thioredoxin z (TRX z)	5/2
190	AT5G14320	Ribosomal protein S13/S18 family (EMB3137)	4/2
191	AT3G49010	60S ribosomal protein L13 (RPL13B)	5/1
192	AT5G42970	Proteasome component (PCI) domain protein (COP8)	5/1
193	AT1G02150	Tetratricopeptide repeat (TPR)-like superfamily protein	5/1
194	AT5G25980	glucoside glucohydrolase 2 (TGG2)	5/1

195	AT3G22960	Pyruvate kinase family protein (PKP1)	3/3
196	AT5G43010	regulatory particle triple-A ATPase 4A (RPT4A)	4/2
197	AT4G34740	GLN phosphoribosyl pyrophosphate amidotransferase 2 (ASE2)	4/2
198	AT3G52140	tetratricopeptide repeat (TPR)-containing protein (FMT)	5/1
199	AT2G47400	CP12 domain-containing protein 1 (CP12-1)	5/1
200	ATCG00330	chloroplast ribosomal protein S14 (RPS14)	5/1
201	AT4G34450	coatomer gamma-2 subunit, putative	5/1
202	AT3G46000	actin depolymerizing factor 2 (ADF2)	5/1
203	AT1G07360	CCCH-type zinc fingerfamily protein with RNA-binding domain (MAC5A)	5/1
204	AT3G42170	BED zinc finger ;hAT family dimerisation domain	5/1
205	AT1G80460	Actin-like ATPase superfamily protein (NHO1)	5/1
206	AT5G43780	Pseudouridine synthase/archaeosine transglycosylase-like family protein (APS4)	5/1
207	AT1G76080	chloroplastic drought-induced stress protein of 32 kD (CDSP32)	5/1
208	AT4G26900	HIS HF (HISN4)	5/1
209	AT3G01390	vacuolar membrane ATPase 10 (VMA10)	3/3
210	AT5G51070	Clp ATPase (ERD1)	5/1
211	AT5G10360	Ribosomal protein S6e (RPS6B)	5/1
212	AT3G15660	glutaredoxin 4 (GRX4)	5/1
213	AT3G18060	transducin family protein / WD-40 repeat family protein	2/3
214	AT1G32990	plastid ribosomal protein l11 (PRPL11)	2/3
215	AT1G12410	CLP protease proteolytic subunit 2 (CLPR2)	2/3
216	AT3G48930	Nucleic acid-binding, OB-fold-like protein (RPS11A)	3/2
217	AT5G14060	Aspartate kinase family protein (CARAB-AK-LYS)	4/1
218	AT3G60830	actin-related protein 7 (ARP7)	4/1
219	AT5G24314	plastid transcriptionally active7 (PTAC7)	4/1
220	AT2G37190	60S ribosomal protein L12 (RPL12A)	3/2
221	AT4G33510	3-deoxy-d-arabino-heptulosonate 7-phosphate synthase (DHS2)	3/2
222	AT5G05010	clathrin adaptor complexes medium subunit family protein	3/2
223	AT1G51710	ubiquitin-specific protease 6 (UBP6)	3/2
224	AT2G02160	CCCH-type zinc finger family protein	3/2
225	AT1G29880	glycyl-tRNA synthetase / glycine--tRNA ligase	4/1
226	AT1G79340	metacaspase 4 (MC4)	4/1
227	AT2G43750	O-acetylserine (thiol) lyase B (OASB)	4/1
228	ATCG00380	chloroplast ribosomal protein S4 (RPS4)	4/1
229	AT4G09570	calcium-dependent protein kinase 4 (CPK4)	4/1
230	AT5G45620	Proteasome component (PCI) domain protein	4/1
231	AT4G24820	26S proteasome, regulatory subunit Rpn7	4/1
232	AT3G29360	UDP-glucose 6-dehydrogenase family protein	4/1
233	AT3G29320	Glycosyl transferase, family 35	4/1
234	AT1G47490	RNA-binding protein 47C (RBP47C)	4/1

235	AT2G27600	AAA-type ATPase family protein (SKD1)	4/1
236	AT1G79470	Aldolase-type TIM barrel family protein	3/2
237	AT4G29670	atypical CYS HIS rich thioredoxin 2 (ACHT2)	3/2
238	AT1G74850	plastid transcriptionally active 2 (PTAC2)	4/1
239	AT2G22250	aspartate aminotransferase (MEE17)	4/1
240	AT3G10610	Ribosomal S17 family protein	4/1
241	AT1G65980	thioredoxin-dependent peroxidase 1 (TPX1)	4/1
242	AT5G62790	1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR)	4/1
243	AT3G53580	diaminopimelate epimerase family protein	1/3
244	AT5G48900	Pectin lyase-like superfamily protein	2/2
245	AT5G55230	microtubule-associated proteins 65-1 (MAP65-1)	2/2
246	AT3G06580	Mevalonate/galactokinase family protein (GAL1)	2/2
247	AT4G28080	Tetratricopeptide repeat (TPR)-like superfamily protein	3/1
248	AT1G18080	Transducin/WD40 repeat-like superfamily protein (RACK1A)	3/1
249	AT3G22320	Eukaryotic rpb5 RNA polymerase subunit family protein (RPB5A)	3/1
250	AT2G30110	ubiquitin-activating enzyme 1 (UBA1)	3/1
251	AT1G26910	Ribosomal protein L16p/L10e family protein (RPL10B)	3/1
252	AT4G37910	mitochondrial heat shock protein 70-1 (mtHsc70-1)	3/1
253	AT5G64760	regulatory particle non-ATPase subunit 5B (RPN5B)	3/1
254	AT5G08650	Small GTP-binding protein	3/1
255	AT1G11750	CLP protease proteolytic subunit 6 (CLPP6)	3/1
256	AT3G09680	40S ribosomal protein S23 (RPS23A)	3/1
257	AT1G10670	ATP-citrate lyase A-1 (ACLA-1)	3/1
258	AT1G49970	CLP protease proteolytic subunit 1 (CLPR1)	3/1
259	AT1G53240	Lactate/malate dehydrogenase family protein (Mmdh1)	3/1
260	AT5G52470	fibrillarin 1 (FIB1)	2/2
261	AT5G22800	Alanyl-tRNA synthetase, class IIc (EMB86)	2/2
262	AT5G23060	calcium sensing receptor (CaS)	2/2
263	AT4G10120	Sucrose-phosphate synthase family protein (SPS4F)	3/1
264	AT2G47250	Probable pre-mRNA-splicing factor ATP-dependent RNA helicase DEAH3	3/1
265	AT4G14880	O-acetylserine (thiol) lyase (OAS-TL) isoform A1 (OASA1)	3/1
266	ATCG00770	ribosomal protein S8 (RPS8)	3/1
267	AT1G59830	protein phosphatase 2A-2 (PP2A-1)	3/1
268	AT1G35670	calcium-dependent protein kinase 2 (CPK11)	3/1
269	AT3G50500	SNF1-related protein kinase 2.2 (SPK-2-2)	3/1
270	AT1G62380	ACC oxidase 2 (ACO2)	3/1
271	AT1G17050	solanesyl diphosphate synthase 2 (SPS2)	3/1
272	AT5G60540	pyridoxine biosynthesis 2 (PDX2)	3/1
273	AT3G22200	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein (POP2)	3/1
274	AT5G64260	EXORDIUM like 2 (EXL2)	3/1

275	AT5G37510	NADH-ubiquinone dehydrogenase, mitochondrial, putative (EMB1467)	3/1
276	AT1G63000	nucleotide-rhamnose synthase/epimerase-reductase (NRS/ER)	3/1
277	AT5G02450	Ribosomal protein L36e family protein	3/1
278	AT3G07010	Pectin lyase-like superfamily protein	3/1
279	AT1G17220	Translation initiation factor 2 (FUG1)	3/1
280	AT3G03780	methionine synthase 2 (MS2)	3/1
281	AT1G30380	photosystem I subunit K (PSAK)	3/1
282	AT2G43100	isopropylmalate isomerase 2 (IPMI2)	3/1
283	AT2G29550	tubulin beta-7 chain (TUB7)	3/1
284	AT3G53420	plasma membrane intrinsic protein 2A (PIP2A)	3/1
285	AT1G06680	photosystem II subunit P-1 (PSBP-1)	3/1
286	AT2G35390	Phosphoribosyltransferase family protein	3/1
287	AT3G06480	DEAD box RNA helicase family protein (RH40)	3/1
288	AT3G47810	Calcineurin-like metallo-phosphoesterase superfamily protein (VPS29)	3/1
289	AT1G59900	pyruvate dehydrogenase complex E1 alpha subunit (E1 ALPHA)	1/2
290	AT2G18110	Translation elongation factor EF1B/ribosomal protein S6 family protein	1/2
291	AT2G43030	Ribosomal protein L3 family protein	1/2
292	AT4G09650	ATP synthase delta-subunit gene (ATPD)	1/2
293	AT1G64790	ILITYHIA (ILA)	2/1
294	AT2G40060	Clathrin light chain protein	2/1
295	AT2G44040	Dihydrodipicolinate reductase, bacterial/plant	2/1
296	AT5G05620	gamma-tubulin complex protein 2 (TUBG2)	2/1
297	AT3G20330	PYRIMIDINE B (PYRB)	2/1
298	AT5G43330	Lactate/malate dehydrogenase family protein	2/1
299	AT3G52200	Dihydrolipoamide acetyltransferase, long form protein (LTA3)	2/1
300	AT3G53460	chloroplast RNA-binding protein 29 (CP29)	2/1
301	AT4G33650	dynamin-related protein 3A (DRP3A)	2/1
302	AT4G16340	guanyl-nucleotide exchange factors (SPK1)	2/1
303	AT1G64710	GroES-like zinc-binding dehydrogenase family protein	2/1
304	AT5G03290	isocitrate dehydrogenase V (IDH-V)	2/1
305	AT5G64050	glutamate tRNA synthetase (ERS)	2/1
306	AT5G67030	zeaxanthin epoxidase (ABA1)	2/1
307	AT3G19480	D-3-phosphoglycerate dehydrogenase	2/1
308	AT3G58990	isopropylmalate isomerase 1 (IPMI1)	2/1
309	AT3G53900	uracil phosphoribosyltransferase (UPP)	2/1
310	AT4G29840	Pyridoxal-5'-phosphate-dependent enzyme family protein (MTO2)	2/1
311	AT3G57560	N-acetyl-l-glutamate kinase (NAGK)	2/1
312	AT1G65930	cytosolic NADP+-dependent isocitrate dehydrogenase (cICDH)	2/1
313	AT4G29830	Transducin/WD40 repeat-like superfamily protein (VIP3)	2/1

314	AT3G11710	lysyl-tRNA synthetase 1 (KRS-1)	2/1
315	AT2G05830	NagB/RpiA/CoA transferase-like superfamily protein	2/1
316	AT3G49430	SER/ARG-rich protein 34A (SRp34a)	1/2
317	AT1G07770	ribosomal protein S15A (RPS15A)	1/2
318	AT5G49460	ATP citrate lyase subunit B 2 (ACLB-2)	1/2
319	AT5G14170	SWIB/MDM2 domain superfamily protein (CHC1)	1/2
320	AT2G31570	glutathione peroxidase 2 (GPX2)	2/1
321	AT2G43460	Ribosomal L38e protein family	2/1
322	AT1G05190	50S ribosomal protein L6, chloroplastic (RPL6)	2/1
323	AT1G79230	mercaptopyruvate sulfurtransferase 1 (MST1)	2/1
324	AT2G30970	aspartate aminotransferase 1 (ASP1)	2/1
325	AT4G26720	protein phosphatase X 1	2/1
326	AT3G55750	Ribosomal protein L35Ae family protein	2/1
327	AT3G54470	uridine 5'-monophosphate synthase / UMP synthase (UMPS)	2/1
328	AT5G18550	Zinc finger C-x8-C-x5-C-x3-H type family protein	2/1
329	AT3G16910	acyl-activating enzyme 7 (AAE7)	2/1
330	AT1G04510	MOS4-associated complex 3A (MAC3A)	2/1
331	AT4G21860	methionine sulfoxide reductase B 2 (MSRB2)	2/1
332	AT5G10860	Cystathionine beta-synthase family protein	2/1
333	AT5G17380	Thiamine pyrophosphate dependent pyruvate decarboxylase family protein	2/1
334	AT5G02080	DNA / pantothenate metabolism flavoprotein	2/1
335	AT1G13180	Actin-like ATPase superfamily protein (ARP3)	2/1
336	AT3G13750	beta galactosidase 1 (BGAL1)	2/1
337	AT5G51700	protein binding;zinc ion binding (PBS2)	2/1
338	AT1G78380	glutathione S-transferase TAU 19 (GSTU19)	2/1
339	AT4G14030	selenium-binding protein 1 (SBP1)	1/1
340	AT2G43970	LA RELATED PROTEIN 6B (LARP6B)	1/1
341	AT4G19710	aspartate kinase-homoserine dehydrogenase ii (AK-HSDH II)	1/1
342	AT2G45300	RNA 3'-terminal phosphate cyclase/enolpyruvate transferase, alpha/beta	1/1
343	ATCG00470	ATP synthase epsilon chain (ATPE)	1/1
344	AT1G75780	tubulin beta-1 chain (TUB1)	1/1
345	AT1G43170	60S ribosomal protein L3 (RPL3A)	1/1
346	AT4G14960	Tubulin/FtsZ family protein (TUA6)	1/1
347	AT4G31990	aspartate aminotransferase 5 (ASP5)	1/1
348	AT5G11520	aspartate aminotransferase 3 (ASP3)	1/1
349	AT3G09840	cell division cycle 48 (CDC48)	1/1
350	AT4G01690	Flavin containing amine oxidoreductase family (PPOX)	1/1
351	AT4G09010	ascorbate peroxidase 4 (APX4)	1/1
352	AT3G04880	DNA-damage-repair/toleration protein (DRT102)	1/1
353	AT5G07440	glutamate dehydrogenase 2	1/1

354	AT5G01600	ferretin 1 (FER1)	1/1
355	AT1G03090	methylcrotonyl-CoA carboxylase alpha chain, mitochondrial (MCCA)	1/1
356	AT1G52100	Mannose-binding lectin superfamily protein	1/1
357	AT4G38600	HEAT repeat ;HECT-domain (ubiquitin-transferase) (UPL3)	1/1
358	AT4G21850	methionine sulfoxide reductase B9 (MSRB9)	1/1
359	AT2G28815		1/1
360	AT2G16950	transportin 1 (TRN1)	1/1
361	AT3G01540	DEAD box RNA helicase 1 (RH14)	1/1
362	AT5G47930	Zinc-binding ribosomal protein family protein	1/1
363	AT5G16710	dehydroascorbate reductase 1 (DHAR3)	1/1
364	AT1G64510	30S ribosomal protein S6 alpha, chloroplastic (RPS6)	1/1
365	AT5G12140	cystatin-1 (CYS1)	1/1
366	AT4G27700	Rhodanese/Cell cycle control phosphatase superfamily protein	1/1
367	AT5G20070	nudix hydrolase homolog 19 (NUDT19)	1/1
368	AT5G27410	D-aminoacid aminotransferase-like PLP-dependent enzymes superfamily protein	1/1
369	AT1G32580	plastid developmental protein DAG, putative	1/1
370	AT3G10670	non-intrinsic ABC protein 7 (NAP7)	1/1
371	AT1G34000	one-helix protein 2 (OHP2)	1/1
372	AT5G58430	exocyst subunit exo70 family protein B1 (EXO70B1)	1/1
373	AT5G52920	plastidic pyruvate kinase beta subunit 1 (PKP1)	1/1
374	AT3G27190	uridine kinase-like 2 (UKL2)	1/1
375	AT5G57290	60S acidic ribosomal protein family	1/1
376	AT5G10470	kinesin like protein for actin based chloroplast movement 1 (KCA1)	1/1
377	AT5G03340	ATPase, AAA-type, CDC48 protein	1/1
378	AT3G02230	reversibly glycosylated polypeptide 1 (RGP1)	1/1
379	AT4G13780	methionine--tRNA ligase, putative	1/1
380	AT4G25050	acyl carrier protein 4 (ACP4)	1/1
381	AT1G50370	Calcineurin-like metallo-phosphoesterase superfamily protein	1/1
382	AT4G23850	AMP-dependent synthetase and ligase family protein (LACS4)	1/1
383	AT5G59880	actin depolymerizing factor 3 (ADF3)	1/1
384	AT2G38610	RNA-binding KH domain-containing protein	1/1
385	AT2G29290	NAD(P)-binding Rossmann-fold superfamily protein	1/1
386	AT2G43090	Aconitase/3-isopropylmalate dehydratase protein	1/1

Abstract in Korean 국문초록

TATA-binding protein-associated factors (TAFs)는 유전자의 프로모터를 인식하는 TFIID 복합체를 구성하는 전사인자이다. 이들의 생화학적인 기능과 더불어, 여러 TAF 인자들이 생물의 발달 과정을 조절한다는 사실이 알려져 있다. 본 논문에서는 TAF15b라는 유전자가 autonomous pathway (AP)를 통하여 애기장대의 개화 시기를 조절한다는 사실을 밝혀내었다. *taf15b* 돌연변이체는 장일, 단일 조건에서 모두 야생형에 비해 꽃이 늦게 피었으며 춘화처리에 의하여 개화 시기가 빨라졌다. 또한 *taf15b* 돌연변이체에서 개화억제인자인 *FLOWERING LOCUS C* (*FLC*)의 발현이 상당히 증가하였고 *flc taf15b*의 돌연변이체에서 개화가 빨리 일어나는 현상을 통하여 *TAF15b*가 AP에 속한 유전자라는 것을 확인할 수 있었다. 돌연변이체에서는 *FLC*의 antisense transcript인 *COOLAIR*의 발현 또한 증가하였다. ChIP 분석을 통하여 TAF15b가 *FLC*와 *COOLAIR*의 전사가 시작되는 곳에 위치한다는 것을 밝혀냈으며 co-IP 실험을 통해서는 RNA polymerase II와 붙는다는 사실을 밝혀내었다. *taf15b* 돌연변이체에서 전사 활성화를 나타내는 H3K4me3 마커가 *FLC* 유전자 상에서 증가하였는데, 이 결과를 통하여 TAF15b가 *FLC*의 전사를 억제하다는 것을 알 수 있었다. 생화학적인 작동 기작을 이해하기 위하여 TAF15b로 IP-MS 실험을 수행하여 함

게 붙는 단백질들을 동정하였다. 이들 중 C-TERMINAL DOMAIN PHOSPHATASE-LIKE 1 (CPL1)라는 단백질이 TAF15b와 함께 *FLC*의 전사를 조절할 것으로 예측하였으며, 실제로 *cpl1* 돌연변이체에서는 *FLC*가 야생형에 비하여 더 많이 발현하는 것을 확인할 수 있었다. 이러한 결과들을 통하여 TAF15b가 *FLC*의 전사를 억제하여 애기장대의 개화 시기를 조절한다는 사실을 제시하였다.

주요어: *TAF15b*, *FLC*, RNA 폴리메라제 II, 전사 조절, autonomous pathway, 개화 시기, *CPL1*

학 번: 2013-30101