CHARACTERIZATION OF A PLANT-SPECIFIC PRC1 (POLYCOMB REPRESSIVE COMPLEX 1)-LIKE COMPLEX AND ITS ROLE IN FLOWERING-TIME REGULATION

WANG YIZHONG

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES
NATIONAL UNIVERSITY OF SINGAPORE

2012

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

_____

Wang Yizhong

5/8/2012

ACKNOWLEDGEMENTS

It has been a precious experience for my PhD study journey at NUS. Too many people I have to thank. First, I would like to thank my supervisor, Assoc. Prof. He yuehui. He gave me the chance to study here and work on this project. Thanks for his patient guidance. Thanks our lab member Dr. Gu Xiaofeng for his collaboration and kind help on the ChIP analysis. Thanks Mr. Li Jikun for his proof reading. I also thank other lab members for their help. I am grateful for the financial support from the Singapore Ministry of Education. Most importantly, I would like to thank my parents and my wife Ms. Liu Yi. It is your supports that give me courage.

Yizhong

TABLE OF CONTENTS

SUMMARY

Polycomb group (PcG) proteins are well known for their function of repressing gene expression via repressive chromatin modifications. In *Drosophila* and mammals, PcG-mediated silencing requires Polycomb repressive complex 1 (PRC1). In plants PRC1-like complexes remain elusive. Recent studies suggest that the *Arabidopsis* protein LHP1 (for LIKE HETEROCHROMATIN PROTEIN 1) plays a PRC1-like role to repress gene expression. In this study, using genetic and molecular tools I found that a plant-specific protein, EMBRYONIC FLOWER 1 (EMF1), functions as part of a PRC1-like complex containing LHP1 to silence gene expression. In addition, this EMF1-PRC1 complex also involves other proteins. Furthermore, I found that EMF1 controls the expression of *FLOWERING LOCUS T* (*FT*; encoding florigen) to regulate flowering time in *Arabidopsis*.

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Literature Review

Polycomb group (PcG) proteins and trithorax group (trxG) proteins act antagonistically to regulate gene expression: PcG proteins could repress gene expression while trxG proteins could activate their expression (Schuettengruber et al., 2007). This mechanism seems to be evolutionarily conserved since it has been reported to be present in from animals to plants (Whitcomb et al., 2007; Kohler and Hennig, 2010). PcG proteins were first characterized in *Drosophila melanogaster* for their role in long-term repressing expression of homeotic (*Hox*) genes (Lewis, 1978). With new technology emerging such as genome-wide chromatin immunoprecipitation (ChIP), hundreds of target genes of PcG proteins have been identified (Boyer et al., 2006; Schwartz et al., 2006; Schuettengruber et al., 2009). Besides the *Hox* genes, many of these PcG targets are transcription factors and signaling components involved in many developmental events, which suggest that gene repression by PcG proteins represents a global silencing mechanism.

To perform their function of gene silencing, PcG proteins are normally grouped into multi-protein complexes. Two main complexes have been identified based on their roles in gene silencing: Polycomb Repressive Complex 1 (PRC1) and PRC2. The PRC2 complex could methylate (di- and tri-) the residue of Lys 27 of histone H3 (H3K27me2/3), which is a central feature of PcG targeted chromatin. The PRC1 complex is responsible for the monoubiquitination of Lys 119 of histone H2A (H2AK119ub1) (Simon & Kingston 2009). In addition, some PRC1 complexes may also compact chromatin to repress gene expression (Eskeland, 2010). It is suggested that the regulation of chromatin structure and post-translational modification (PTM) of histone proteins, which surround DNA to form nucleosome structure, are critical to PcG protein-mediated gene silencing (Margueron and Reinberg, 2011). The PRC1 subunit Pc has been

reported to bind to H3K27me3 deposited by PRC2, specifically, indicating that PRC1 may act downstream of PRC2. However, evidence that PRC1 may target genes in the absence of PRC2 challenges this hypothesis (Schoeftner, 2006). But normally, both PRC1 and PRC2 are required to repress gene expression.

1.1 Polycomb Repressive Complex 2 (PRC2)

In *Drosophila*, the core PRC2 complex includes four components: Enhancer of Zeste (E(Z)), Suppressor of Zeste 12 (SUZ12), Extra Sex combs and Extra Sex combs-like (ESC/ESCL), and Nucleosome remodeling factor 55 (p55 and CAF1). E(Z) is a SET domain containing protein serving as the catalytic subunit of PRC2 complex for H3K27 methylation. E(Z) alone does not have activity unless it is assembled at least with SUZ12 and ESC (Pasini et al., 2004; Ketel, 2005). And p55 seems to be related to recognizing and binding to histones, which facilitates PRC2 function (Song et al., 2008)(Table 1).

The core components of PRC2 complex are highly conserved from animals to plants (Table 1). However, there are different variants for PRC2 components providing multiple choices to assemble the PRC2 complex. And these related PRC2 complexes might be adapted for different circumstances. For instance, in mammals, there are two homologues of E(Z): Enhancer of Zeste homologue 1 (EZH1) and EZH2. Homologues of ESC (EED) have several isoforms due to alternative translation start sites. Retinoblastoma-bind protein 48 (RbAp 48) and RbAp 46 are homologues of p55 (Table 1).

Table 1. Core subunits of PRC2 complex in fly, human and *Arabidopsis*

Drosophila	**humans**	***Arabidopsis***	**Protein domains**	**Biochemical function(s)**
Enhancer of Zeste (E(Z))	EZH1 and EZH2	MEA, CLF, and SWN	SET domain	Catalyzes H3K27 methylation
Suppressor of Zeste 12 (Su(Z) 12)	SUZ12	EMF2, FIS2 and VRN2	C2–H2 zinc finger and VEFS domain	Stimulates H3K27 methylation
Extra Sex combs (ESC) and Extra Sex combs-like (ESCL)	EED	FIE	WD repeats	Stimulates H3K27 methylation
Nucleosome remodeling factor 55 (p55 and CAF1)	RbAp48 and RbAp46	MSI1	WD repeats	Histone binding

Derived from the reviews on PRC2 complex of animals and plants (Margueron and Reinberg, 2011; Simon and Kingston, 2009; Hennig and Derkacheva, 2009; Holec and Berger, 2012)

Interestingly, despite of 65% identical sequence between EZH1 and EZH2 and association with almost the same core subunits, the EZH1-PRC2 is quite different from EZH2-PRC2 complex in biochemical functions: EZH1-PRC2 could compact chromatin but has much lower methyltransferase activity than EZH2-PRC2 (Margueron, 2008). As a consequence, EZH1-PRC2 could repress *in vitro* transcription while EZH2-PRC2 could not. These biochemical differences between EZH1-PRC2 and EZH2-PRC2 might account for their different distribution during development: EZH1 is highly expressed in adult tissues and non-dividing cells while EZH2 is the major form in dividing cells (Shen, 2008).

Besides the core components, there are other PcG proteins that can associate with PRC2 complex. *Polycomb*-like protein (PCL) in *Drosophila*, and its homologues PHF1 (PCL1), MTF2 (PCL2), and PHF19 (PCL3) in mammals are found to interact with PRC2 core components (Nekrasov, 2007; Sarma et al., 2008). PCL (PHF1) may enhance PRC2 enzyme activity for H3K27 tri-methylation and facilitate the recruitment of PRC2 to its target genes (Savla et al., 2008). In mammals, two more proteins: AEBP2, a zinc-finger protein, and JARID2, a Jumonji family protein lacking the enzymatic activity of histone demethylation, are also found to interact with PRC2 components. And their colocalization at the target genes with PRC2 suggests that they are also associated with the PRC2 complex (Cao and Zhang, 2004; Jung et al., 2005; Kim et al., 2009). All these non-core components also increase the complexity of PRC2.

To repress gene expression, the PRC2 complex is first recruited to its targets. However, the exact mechanism of PRC2 recruitment is not clear yet. The Polycomb response elements (PREs) of the

target genes are reported to play an important role in PcG proteins recruitment. PREs are *cis*-regulators of PcG targeting genes for the response to PcG proteins. In *Drosophila*, the PREs are identified in *Hox* genes and other PcG targets (Muller and Kassis, 2006; Ringrose and Paro, 2007). However, the exact picture of PREs has not been determined although they have been refined to several hundred base pairs for certain genes (Mihaly et al., 1998; Kozma et al., 2008). It is also noticed that while the elements for PREs function are not uniform for all the PcG target genes, they share some common and key features: they all have multiple binding site for the zinc finger protein Pleiohomeotic (PHO). PHO and other proteins interact with PREs then bridge other PcG proteins to the target genes (Wang et al., 2004). In mammals, the PREs have not been identified yet. It is revealed that most of the PRC2 targets are co-related to CpG islands or similar CG-rich regions which suggests this feature might be a potential sub-element of mammalian PREs (Ku, 2008). Interestingly, the non-coding RNAs (ncRNAs) are also involved in the recruitment of PRC2 complex when studying *Hox* gene clusters silencing (Rinn, 2007), X-chromosome inactivating (Zhao et al., 2008) and paternally imprinting (Pandey, 2008). In plants, ncRNA is also reported to contribute to the recruitment of PRC2 (Heo and Sung, 2011).

After PRC2 is recruited to the target genes, it catalyzes histone methylation on H3K27 by adding up to three methyl groups. Tri-methylated H3K27 (H3K27me3) seems to be the main form for PcG silencing since H3K27me3 is distributed quite restrictively to PcG silencing genes and also attracts PcG complexes (Schwartz, 2006; Schuettengruber et al., 2009). The reciprocal relation between H3K27me3 and PRC2 complex makes it the central feature of PRC2 target genes.

The post-translational modification (PTM) on histone, in general, might function either by regulating chromatin structure directly, affecting the interactions between histones and DNA by changing the electrostatic charge of histones, or by acting as a docking site to facilitate the recruitment of other regulatory complexes. Evidence collected till now indicates that H3K27me3 acts as the latter role to regulate gene expression. The interaction between H3K27me3 and Polycomb (Pc), a subunit of PRC1, and the fact that loss of H3K27me3 also accompanies with loss of affinity of PRC1 to its target genes indicate that H3K27me3 could facilitate the recruitment of PRC1 to its target genes (Min et al., 2003).

In *Arabidopsis thaliana*, the core components of PRC2 complex are well characterized because of high conservation between animals and plants (Hennig and Derkacheva, 2009). There are three homologues of E(Z): MEDEA (MEA), CURLY LEAF (CLF), and SWINGER (SWN). Three homologues of Su(Z)12: EMBRYONIC FLOWER 2 (EMF2), FERTILIZATION INDEPENDENT SEED 2 (FIS2) and VERNALIZATION 2 (VRN2). One homologue of ESC: FERTILIZATION INDEPENDENT ENDOSPERM (FIE). And one homologue of p55: MULTICOPY SUPPRESSOR OF IRA1 (MSI1) (Table 1).

The PRC2 complex functions throughout the *Arabidopsis* life cycle, involved in almost all the key events of *Arabidopsis* development, which indicates a crucial role of PRC2. From a single seed to becoming an adult plant and to eventually getting mature reproductive seeds, the *Arabidopsis* encounters several key transitions in development: seed germination to form a juvenile plant, transition from juvenile to adult stage, floral transition, gametogenesis, fertilization, seeds development (Hennig and Derkacheva, 2009).

Genetic and molecular analyses of PcG proteins, especially the studies of PcG mutants reveal their roles in plants development. In terms of different SUZ12 homologues, there are at least three forms of PRC2 complex in *Arabidopsis* which function at different stages to regulate plant development. The EMF2-PRC2 complex mainly plays role in sporophytic development; the VRN-PRC2 complex is mainly involved in the response of vernalization, a prolonged period of cold exposure like winter at high latitudes; the FIS-PRC2 complex mainly functions during fertilization and seeds development (Holec and Berger, 2012).

Between germination and floral transition, the plant needs a certain period of time for vegetative development. During this period, the shoot apical meristem (SAM) keeps on proliferating to keep itself renewal and to initiate the leaf primordia to form new leaves (Baurle and Dean, 2006). During floral transition, the shoot apical meristem transits to inflorescence meristem and then produces floral meristem to form flowers. It has been reported that the H3K27me3 deposited by PRC2 complexes is distributed in a tissue-specific manner which leads to the spatial-specific expression patterns of target genes. This is especially important for some key regulators whose proper expression in the right place triggers normal development. For instance, the leaf-specific transcription factors such as TCP, CONSTANS-like and GRAS-transcription factors are prevented from meristem with meristem-specific H3K27me3; while HOMEOBOX transcription factors and other meristem-specific gene families are restricted at meristem tissues and get H3K27me3 in a leaf-specific manner (Lafos et al., 2011).

The PRC2 complex is also found to play critical role in meristem determination by regulating the floral identity genes *LEAFY (LFY), AGAMOUS (AG)* and *KNUCKLES (KNU)* expression to allow time for vegetative growth (Sun et al., 2009). The loss of function of PRC2 components leads to the mis-expression of *LFY, AG* and *KNU* and causes a shorter time of vegetative growth comparing to wild type (Kinoshita et al., 2001). The single mutants of *clf* show an early flowering phenotype (Katz et al., 2004); the phenotype of *emf2* mutants is more severe: a very short period for vegetative growth and flowering soon after two cotyledons formation (Yoshida et al., 2001); the *clf/swn* double mutants show much more severe phenotype: the plant no longer maintaining organ identity post-embryonically, producing callus-like tissue, and loss of H3K27me3 at the molecular level (Lafos et al., 2011).

After fertilization, the zygote develops into embryo and the central cell into endosperm. This transition from gametophyte to sporophyte involves the FIS-PRC2 complex which comprises of MEA/SWN, FIS2, FIE and MSI1. Mutants of *fis2* show an over-dividing central cell in the gametophyte without cellularization which indicates the function of PRC2 to maitain normal gametophytic phase (Chaudhury et al., 1997). Mutants of *fie* exhibit defects in germination, enhanced seeds dormancy and unsustainable vegetative development leading to callus-like structure formation at last, which suggests that PRC2 also controls the embryo to seedling phase transition (Ohad et al 1996).

PRC2 complex is also involved in flowering-time control. Mutants of PRC2 components show early-flowering phenotype indicating their roles in preventing flowering. *FLOWERING LOCUS T (FT)*, a flowering-pathway integrator to promote flowering, loses H3K27me3 on its chromatin

and is derepressed in *clf* and *emf2* mutants (Jiang et al., 2008). Another gene, *FLOWERING LOCUS C (FLC)*, which could repress flowering and response to vernalization is also regulated by PRC2 (He, 2012). The mechanism of *FLC* repression by PRC2 is well studied and provides an insight into how PRC2 regulates gene expression in plants.

In *Arabidopsis*, two PRC2 complexes are involved in repressing *FLC* expression. The EMF2-PRC2 complex is considered to be a constitutive *FLC* repressor (He, 2012). Its components such as EMF2 and CLF are found to bind to *FLC* chromatin and defects in these proteins accompany with loss of H3K27me3 on *FLC* locus (Jiang et al., 2008). It is proposed that EMF2-PRC2 coordinates other repressors such as FLOWERING LOCUS D (FLD) and FVE to establish a repressive chromatin environment for *FLC* repression (Gu et al., 2011; Yu et al., 2011).

In winter annuals, *FLC* expression is repressed by vernalization in winter, and remains silenced after the plants are transferred to warm conditions (Kim et al., 2009). Genetic analysis indicates the VRN2-PRC2 complex is required for this process. Cooperated with the PHD domain proteins VERNALIZATION INSENSITIVE 3 (VIN3), VRN5, VERNALIZATION LIKE 1 (VEL1), the VRN2-PRC2 complex including CLF/SWN, VRN2, FIE and MSI1 deposits H3K27me3 on around the first exon region of *FLC* when in cold treatment (Wood et al., 2006; De Lucia et al., 2008), which may resemble the PHD protein PCL mediated PRC2 function in *Drosophila*. After changing to the warm conditions, H3K27me3 spreads to the entire *FLC* chromatin also by VRN2-PRC2 catalyzing. The spread H3K27me3 is recognized and bound by LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) (Sung et al., 2006). LHP1 associated with other repressors leads to the subsequent maintenance of *FLC* repression state.

Recently, it is found that long non coding RNAs are also involved in VRN2-PRC2 mediated *FLC* repression. Vernalization induces transient expression of a non-coding RNA: COLD ASSISTED INTRONIC NONCODING RNA (COLDAIR) which is transcribed from the first intron of *FLC*. COLDAIR is associated with the PRC2 component CLF and assists the recruitment of VRN2-PRC2 to *FLC* chromatin (Heo and Sung, 2011), which suggests that the mechanism of non-coding RNA mediated PRC2 recruitment may be conserved from animals to plants though COLDAIR is the only one found in plants up till now.

1.2 Polycomb Repressive Complex 1 (PRC1)

Another main PcG complex family is PRC1. Unlike the PRC2 complex, PRC1 seems to have much more diversity especially in mammals. There are multiple paralogues of the core subunits, which substitute for various possible combinations. In some cases, some PRC1 proteins may associate with other partners to form complexes which differ from the classic PRC1(Simon and Kingston, 2009).

In *Drosophila*, the core components of PRC1 include four PcG proteins: Polycomb (Pc), Polyhomeotic (Ph), Posterior Sex combs (PSC) and RING (also known as Sex comb extra) (Table 2). In mammals, the homologues of these four proteins are also identified (Francis et al., 2001). However, the mechanisms of PRC1 complex repressing target gene expression are not fully understood yet. It is believed that PRC1 coordinates with PRC2 to maintain the gene

Table 2. Core subunits of PRC1 complex in fly, human and *Arabidopsis*[1]

Drosophila	**humans**	***Arabidopsis***	**Protein domains**	**Biochemical function(s)**
Polycomb (Pc)	CBX2 (PC1), CBX4 (PC2), CBX6, CBX7 and CBX8 (PC3)	LHP1 [2]	Chromodomain and chromo-shadow domain	Binds H3K27me3
Polyhomeotic (Ph)	PH1 and PH2	?	SAM domain and zinc finger	Higher-order interactions
Posterior Sex combs (Psc)	BMI1, MEL18 (PCGF2) and NSPC1 (PCGF1)	AtBMI1a/1b	RING finger	Co-factor for E3 ubiquitin ligase and compacts chromatin
Sex combs extra (RING)	RING1 and RING1B (RNF2)	AtRING1a/1b	RING finger	E3 ubiquitin ligase for H2A

1. Derived from the reviews on PcG proteins in animals and plants (Simon and Kingston, 2009; Whitcomb et al., 2007; Zheng and Chen, 2011; Holec and Berger, 2012).

2. LHP1 is the homologue of HP1 but functions like Pc to recognize and bind to H3K27me3 (Turck et al., 2007).

silencing states. H3K27me3 deposited by PRC2 on target genes are recognized by the subunit of PRC1 Polycomb (Pc) (Fischle, 2003). The recruitment of PRC1 subsequently keeps the gene silenced either by direct repression of transcription, regulation of chromatin structure or compaction of nucleosomes.

Recent studies show that in human, two core subunits of PRC1: RING1B and B lymphoma Mo-MLV insertion region 1 (BMI1) could form a stable heterodimer and catalyze H2A ubiquitination directly (Cao et al., 2005).This suggests that H2A ubiquitination may be the key function of PRC1 (Wang, 2004). However, besides the originally defined PRC1, these two proteins are also found to be present in other complexes with different compositions. In *Drosophila*, a RING-associated factor (dRAF) complex has been reported recently, which contains two PRC1 subunits RING and PSC and a histone lysine demethylase KDM2. KDM2 could demethylate H3K36me2 and enhance the activity of H2A ubiquitination by RING and PSC. Interestingly, unlike KDM2, the addition of Pc and Ph, two other components of PRC1, has no effects on H2A ubiquitination *in vitro* (Lagarou, 2008). A similar complex named as BCL6 corepressor (BCOR) complex is found in mammals, which is composed of the RING homologues RING1 and RING1B, a PSC homologue nervous system Polycomb 1 (NSPC1), BCOR and a KDM2 homologue KDM2B. The BCOR complex could ubiquitylate H2A of its target genes *in vivo* although it is not clear yet whether the association with KDM2B to this complex could stimulate the ubiquitination (Gearhart et al., 2006; Sanchez, 2007). These findings indicate that in both flies and mammals there exists a complex containing the components of PRC1 and a KDM2 subunit which differs from the original PRC1.Considering the fact that the dRAF complex has significantly more ubiquitination activities than the original

PRC1, it is proposed that the original PRC1 regulates chromatin structure, the dRAF complex (BCOR complex in mammals) accounts for H2A ubiquitination, and these two complexes are coordinated for the full function of PRC1. However, the scenarios of gene silencing by these PRC1 complexes are not clear yet.

Possible mechanisms are speculated to interpret PRC1 silencing. It seems that PRC1 proteins do not inhibit transcription initiation by affecting RNA polymerase binding since PRC1 proteins and transcription factors are found to be colocalized at target genes, and RNA polymerases could bind to some PcG silenced targets (Breiling et al., 2001; Bracken et al., 2006; Zhou, 2008). Emerging evidence leads to a hypothesis that PRC1 proteins affect transcription elongation. With the studies on normal and engineered mouse ES cells, it is found that RNA polymerase engaged in PcG targets is present in an inactive phosphorylation state and paused in the middle of target genes. The switch of phosphorylation state from inactive (phosphorylated Ser 5 of C-terminal domain) to active (phosphorylated Ser 2) could release the paused elongation (Stock, 2007).

H2A ubiquitination may play key roles in PRC1 silencing. Ubiquitinated H2A by PRC1 may act as the docking site for repressive transcription complexes or obstruct the recruitment of transcription factors. This is supported by the fact that H2A ubiquitination could inhibit the binding of elongation factor FACT (facilitates chromatin transcription complex) to transcription sites (Saunders, 2003; Zhou, 2008). The genes in mouse ES cells with paused RNA polymerase also possess ubiquitylated H2A (Bernstein, 2006). Removal of H2A ubiquitination in RING1 and RING1B mutants accompanies with the PcG target genes activation (Cao et al., 2005).

Besides, chromatin compaction may also contribute to the inhibition of transcription by PRC1 (Francis et al., 2004). When RNA polymerase transcribes through a gene body, a 'loosen' environment is required where it could get close to DNA sequence. For the nucleosome, the tight interaction between histones and DNA has to be loosened, and this process might be obstructed by compaction and stabilization of the chromatin of gene body (Lorch et al., 1987). The PRC1 subunit PSC itself could regulate chromatin structure by its noncovalent effects. The carboxy-terminal region of PSC (PSC-CTR), which has an intrinsically disordered domain and is poorly conserved in protein sequence, is essential to the biochemical function of PSC. Truncations of PSC-CTR impair compaction chromatin *in vitro* and disrupt gene silencing by PRC1 *in vivo*. Nevertheless, hundreds of genes are targeted by PcG complexes and the mechanisms of silencing may be varied on different target genes (Beh et al., 2012).

Because of the poor conservation of PRC1 subunits between animals and plants, in the beginning, people doubted whether there were PRC1 or PRC1-like complexes in plants or not. However, it is speculated that there might be plant unique components that function like PRC1 complex. Indeed, the homologues of PRC1 component RING1 and BMI1 are identified recently in *Arabidopsis* and were suggested for the PRC1-like function in plants (Xu and Shen, 2008; Bratzel et al., 2010).

In *Arabidopsis*, AtRING1a and AtRING1b, homologues of mammalian RING1, are characterized for the roles in repressing Class I KNOTTED-like homeobox (*KNOX*) transcription factors expression. *KNOX* genes are expressed specifically in shoot meristems and are key

factors for maintaining meristem function. Loss of function of AtRING1a/1b, leads to the de-repression of *KNOX* genes expression and ectopic meristem formation in the double mutants (Xu and Shen, 2008). AtBMI1a, AtBMI1b and AtBMI1c are homologues of mammalian BMI1. Compared to *AtBMI1a/1b*, the expression level of *AtBMI1c* is very low, suggesting that AtBMI1c might not be involved in the AtBMI1a/1b complex. Like AtRING1a/1b, AtBMI1a/1b are involved in meristem cells determination by repressing the embryonic and stem cell regulators expression. It is also found that AtBMI1a/1b and AtRING1a/1b could mediate H2A monoubiquitination *in vitro* and *in vivo*, just as their homologues in animals. Considering the fact that H3K27me3 levels are not affected in *Atbmi1a/1b* double mutants, these findings suggest that AtBMI1a/1b function downstream of PRC2 and play a PRC1-like role in gene silencing. The *in vitro* pulldown assay hints that AtBMI1a/1b and AtRING1a/1b might be physically associated with LIKE-HETEROCHROMATIN PROTEIN 1 (LHP1) and EMBRYONIC FLOWER 1 (EMF1) (Bratzel et al., 2010).

LHP1 (also known as TERMINAL FLOWER 2 (TFL2)) is the only homologue in *Arabidopsis* of metazoans HETEROCHROMATIN PROTEIN 1 (HP1) (Gaudin et al., 2001). Most of the isoforms of HP1 are enriched in heterochromatic regions to regulate heterochromatin where HP1 proteins bind to histone H3K9me2 and H3K9me3 (Eskeland et al., 2007). However, unlike HP1, LHP1 is distributed predominantly in euchromatin in *Arabidopsis*. *In vitro*, LHP1 could bind to H3K9me2, H3K9me3 and H3K27me3, whereas it recognizes H3K27me3 specifically *in vivo* (Turck et al., 2007). Mutants of *lhp1* show phenotypes of day length independent early flowering, terminal flower, and curled leaves, which resemble mutants of PRC2 components (Gaudin et al., 2001; Kotake et al., 2003). The interactions between LHP1 and AtBMI1a/1b, AtRING1a/1b

hinted in the *in vitro* pulldown assay and the specific association with H3K27me3 *in vivo* suggest that LHP1 functions like the core component of animal PRC1: Pc, which recognizes and binds to H3K27me3 although LHP1 is not the homologue of Pc (Table 2).

As its name suggests, EMBRYONIC FLOWER 1 (EMF1) plays a critical role in floral repression. The mutants of *emf1* skip vegetative development and flowers just after germination which is quite similar to the mutants of a PRC2 component EMF2 (Moon et al., 2003). These observations imply that both genes are needed for the vegetative development of *Arabidopsis*. Later EMF1 is found to coordinate with EMF2-PRC2 complex to repress *AGAMOUS* (a flower organ identity gene) expression (Calonje et al., 2008). Genome wide investigation reveals that EMF1 is involved in the repression of vast of genes expression which are related to key processes of *Arabidopsis* development such as seed, root development, stem cell maintenance, flower organ determination, shoot architecture, and hormone synthesis (Kim et al., 2012). The biochemical function of EMF1 is characterized with the activities of DNA binding, inhibition of *in vitro* transcription and inhibition of chromatin remodeling which resemble the PRC1 component PSC (Calonje et al., 2008; Beh et al., 2012). Like metazoan PSC C-terminal region (PSC-CTR), the structure of EMF1 is largely disordered, and has low contiguous negative charge which might facilitate DNA binding and inhibit chromatin remodeling *in vitro* (Beh et al., 2012).

1.3 Other PcG complexes

Besides these two main families of complexes: PRC1 and PRC2, there might be other PcG-related complexes. Indeed, in *Drosophila*, a PcG protein Pleiohomeotic (Pho) could form a stable

heterodimer with another PcG protein dSfmbt (Scm-related gene containing four mbt domains) which is named as Pho repressive complex (Pho-RC) (Klymenko, 2006). Pho and Pho-like (Phol) are zinc finger DNA binding proteins which could bind to the PREs of *HOX* genes specifically (Brown et al., 1998; Brown et al., 2003). The protein of dSfmbt has four malignant brain tumor (MBT) repeats associated with mono- and di-methylated H3K9 and H4K20 selectively. Loss of function of Pho or dSfmbt result in the de-repression of *HOX* genes, which indicate their function as a repressive complex (Wang et al., 2004). Besides, mutants of *pho* and *phol* lose the association of PRC1 and PRC2 with PREs of *HOX* genes suggesting Pho may act as the recruiter of PRC1 and PRC2. It is proposed that Pho binds to the PREs of target gene, and dSfmbt interacts with the methylated histones of PREs flanking chromatin selectively which directs the recruitment of PRC1 or PRC2, and then mediates target gene repression (Klymenko, 2006).

Another PcG complex named as Polycomb repressive deubiquitinase (PR-DUB) is characterized in *Drosophila* recently which is comprised of two PcG proteins: Calypso and Additional sex combs (ASX) (Scheuermann, 2010). Calypso is the catalytic subunit of PR-DUB while the activity of PR-DUB needs the presence of both subunits. PR-DUB complex specifically deubiquitylates mono-ubiquitinated H2A, which is critical for PR-DUB mediated gene silencing. Disruption of Calypso results in increasing H2A monoubiquitination and de-repression of the target genes (Scheuermann, 2010). This is interesting because H2A monoubiquitination by PRC1 or dRAF complex is also related to gene repression. Thus, it is implied that the balance of H2A monoubiquitination in the target gene chromatin may be crucial in PcG directed repression. In mammals the homologues of PR-DUB components are identified but the function of this complex has not been characterized yet. However, neither Pho-RC like nor PR-DUB like

complexes have been identified to date in plants which may indicate a difference between plants and animals in PcG repression.

1.4 Flowering-time control

Floral transition is a key event in plant life cycle. It turns the plant from vegetative growth to reproductive development. Flowering at the right time is very important for plant to get maximal reproductive success. Coordination with the environmental cues such as temperature, day length, and hormones is crucial for the plant to determine when to start flowering. Floral genes response to the changes of environments and finally induce floral transition. After transition, the shoot apex meristem (SAM) no longer produces new leaf primordia but transits into inflorescence meristem (IM) which produces floral meristem (FM) to develop into flowers finally (Carles and Fletcher, 2003; Irish, 2010). In *Arabidopsis thaliana*, about 180 genes are implicated to be involved in flowering-time control through loss-of-function mutants or transgenic plants analysis (Fornara et al., 2010). These genes form a sophisticated but delicate network to sense the signals of flowering. Interestingly, a small number of these genes, such as *FLOWERING LOCUS C (FLC)* and *FLOWERING LOCUS T (FT)*, act as the integrator to converge signals from other genes(Amasino, 2010). With regards to the signals for flowering, these genes could be roughly grouped into several flowering pathways: photoperiod pathway, vernalization pathway, autonomous pathway, gibberellin pathway, thermosensory pathway and age pathway (Fornara et al., 2010).

Vernalization pathway

As mentioned in the previous section, the winter annual of *Arabidopsis* has to experience the long cold winter and get flowered after returning to the warm condition. *FLC* is the main effector of vernalization, which encodes a MADS-box transcription factor to repress flowering (Sheldon et al., 2000). Vernalization establishes the repressive state of *FLC* which is maintained in the following spring. Repression of *FLC* involves the 'epigenetic' regulation with modifications on histones of *FLC*. VRN2-PRC2 associated with PHD domain proteins deposits H3K27me3 on *FLC* genome leading to the silence of *FLC* (He, 2012). LHP1 recognizes and binds to H3K27me3 and recruits other proteins, which may account for the maintenance of the repressive status of *FLC* (Zheng and Chen, 2011). Long non-coding RNAs are also involved in *FLC* repression. The antisense transcript COLD INDUCED LONG ANTISENSE INTERGENIC RNA (COOLAIR) may mediate *FLC* silencing by a co-transcriptional mechanism (Swiezewski et al., 2009). Another non-coding RNA COLDAIR may act as a recruiter of PRC2 to *FLC* chromatin (Heo and Sung, 2011).

Without vernalization, in winter annuals, *FLC* expression is activated by FRIGIDA (FRI) which acts as a scaffold protein to recruit other regulators (Johanson et al., 2000). FRI interacts with four proteins FRL1, FES1, SUF4 and FLX to form a putative transcription activator complex (FRIc) to activate *FLC* expression (Choi et al., 2011). FRIc may recruit multiple complexes to mediate histone modification on *FLC* chromatin such as H2A.Z deposition (Deal et al., 2007), histone acetylation (Yu et al., 2011), H3K4me3 (Jiang et al., 2009), H2B monoubiquitination(Gu

et al., 2009) and H3K36 di- and tri-methylation (Xu et al., 2008) which lead to the activation of *FLC*.

Autonomous pathway

For the rapid-cycling accessions, unlike the winter annuals, they do not require vernalization to accelerate flowering. Sensing the endogenous signals, the autonomous pathway genes repress *FLC* expression at low level to promote flowering. Mutations of these genes often result in the de-repression of *FLC* expression and delay in floral transition (Amasino, 2010).

Histone modifications on *FLC* also play important roles in autonomous pathway. A corepressor-like complex composed of FLOWERING LOCUS D (FLD), HISTONE DEACETYLASE 6 (HDA6), and FVE or MSI5 is found to repress *FLC* expression (Amasino, 2010; He, 2012). FLD, the homologue of human Lysine-Specific Demethylase1 (LSD1), is a histone H3K4 demethylase. HDA6 could catalyze histone deacetylation (Gu et al., 2011; Yu et al., 2011). FVE or MSI5 is involved in histone binding (Ausin et al., 2004; Jeon and Kim, 2011). Mutations of these proteins result in elevated histone acetylation and H3K4 methylation which lead to *FLC* derepression.

Chromatin immunoprecipitation (ChIP) assay analysis indicates H3K27me3 is also presented in *FLC* chromatin. The EMF2-PRC2 complex is involved in H3K27me3 deposition on *FLC* locus which is also required for *FLC* repression (Jiang et al., 2008).

FCA and FPA, two RNA binding proteins, could promote the processing of 3'-end *FLC* antisense transcripts by the CstF 3'-end processing complex which produce antisense long non-coding RNA to cause FLC repression (Hornyik et al., 2010; Liu et al., 2010). It is also reported that FCA and FPA require FLD for *FLC* downregulation (Liu et al., 2007; Baurle and Dean, 2008). Altogether, these findings suggest that FLD-HDA6-FVE/MSI5 complex cooperates with EMF2-PRC2 and the co-transcription of *FLC* antisense RNA to repress *FLC* expression.

Thermosensory pathway

The surrounding temperature also influences flowering. *Arabidopsis* plants get earlier flowering at high temperature than at low temperature (Balasubramanian and Weigel, 2006). *FT* expression is affected in different temperature. Higher expression of *FT* is observed at high temperature, which leads to earlier flowering. A MADS box transcription factor SHORT VEGETATIVE PHASE (SVP) seems to be critical for this pathway since the *svp* mutants shows uninfluenced flowering time at different temperature. And *FT* expression is repressed by SVP at low temperature thus causes later flowering (Lee et al., 2007). Histone variant H2A.Z deposited by SWR1 complex (SWR1c) on *FT* chromatin is also involved in this pathway. Lower level H2A.Z nucleosomes are detected in high temperature than in low temperature conditions. Impairing SWR1c function causes early flowering and temperature insensitive *FT* activation (March-Diaz et al., 2008; Kumar and Wigge, 2010).

Age pathway

Accompanying with plants aging, transcription factors SQUAMOSA PROMOTER BINDING LIKE (SPLs) are accumulated to promote flowering eventually by activating *LEAFY (LFY), FRUITFUL (FUL) and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)* expression. MicroRNA mir-156 acts as a repressor of SPLs with high expression levels in young plants and decrease progressively during plants aging (Wang et al., 2009).

Gibberellin pathway

The hormone Gibberellin (GA) could promote flowering especially for plants at non inducible conditions such as short day conditions for *Arabidopsis* (Wilson et al., 1992). Any mutants that affect the GA biosynthesis or degradation will affect flowering time. The bioactive GA (GA_4) is detected to be accumulated at SAM before flowering. The petioles appear to be one place for GA biosynthesis where the genes involved in GA synthesis are activated in response to environmental cues. GA could activate the expression of *FT* in leaves and *SOC1, LFY* in SAM which leads to floral induction (Mutasa-Gottgens and Hedden, 2009; Fornara et al., 2010).

Photoperiod pathway

Day length changing especially at high latitudes is another major environmental cues for plant flowering. Some species such as *Arabidopsis thaliana* promote flowering under long days (LDs) of 16 hours of light, whereas some species such as rice (*Oryza sativa*) flower earlier under short

days (SDs) (8 h light). For a long time, it is noticed that plants could sense day length changing in leaves, where plants perceive the signals and transmit to SAM for response (Meacham et al., 1920; Zeevaart, 2006). However, the 'florigen' which is speculated for the signals transmission between leaves and SAM is unclear, until recently FT is identified as the candidate of florigen (Turck et al., 2008).

Lights are perceived by photoreceptors in leaf which give rise to a signal cascade involving GIGANTEA (GI) and CONSTANS (CO). GI with an F-box ubiquitin ligase FLAVIN-BINDING KELCH REPEAT F-BOX 1 (FKF1) negatively regulate *CO* repressors such as *CYCLING DOF FACTOR 1 (CDF1)* expression, which lead to *CO* activation(Imaizumi et al., 2005; Sawa et al., 2007). *CO* expression acts in a diurnal pattern that *CO* mRNA extends its abundance to light in LD conditions, while in SD conditions *CO* is accumulated only in night(Suarez-Lopez et al., 2001). However, in LD conditions, CO proteins are stabilized by light with peaking at the end of light and degraded quickly in the dark (Valverde et al., 2004). The post-transcriptional regulation of CO appears to account for CO function since the presence of *CO* transcripts in SD does not cause flowering. Cryptochromes such as CRYPTOCHROME 2 (CRY2) and PHYTOCHROME A (PhyA) are involved in CO proteins stabilization (Valverde et al., 2004), while PhyB and an E3 ubiquitin ligase CONSTITUTIVE PHOTOMORPHOGENESIS 1 (COP1) mediate CO protein degradation in the dark (Mockler et al., 2003; Wigge et al., 2005; Liu et al., 2008). CO proteins activate *FT* expression that leads to the subsequent promotion of flowering (Wigge et al., 2005).

FT is expressed in the minor veins of leaves, whereas FT proteins are translocated to SAM where they interact with a bZIP transcription factor FLOWERING LOCUS D (FD). FT/FD activates the expression of the floral meristem identity gene *APETALA 1 (AP1)* and another floral promoter gene *SOC1* which in turns promote *LFY* expression, then leads to flower induction (Abe et al., 2005; Wigge et al., 2005; Yoo et al., 2005). This property of *FT* makes it serve as the florigen in *Arabidopsis* (Corbesier et al., 2007). The mechanism of CO and FT regulate flowering in photoperiod pathway seems to be conserved since the functions of their homologues are also identified in rice and trees (Bohlenius et al., 2006; Tamaki et al., 2007).

Epigenetic regulation is also critical for the regulation of *FT* expression. EMF2-PRC2 complex could deposit H3K27me3 on *FT* chromatin leading to *FT* repression (Farrona et al., 2011). Its components CLF and EMF2 are found to bind to *FT* locus and loss of CLF and EMF2 activities results in decrease of H3K27me3 and *FT* derepression (Jiang et al., 2008). REF6, a JmjC-domain H3K27me3 demethylase, could remove H3K27me3 mark and promote *FT* expression, which suggests that H3K27me3 is dynamically regulated on *FT* chromatin (Lu et al., 2011). LHP1, a putative PRC1 component, could bind to *FT* chromatin directly and disruption of LHP1 leads to *FT* derepression and early flowering (Turck et al., 2007).

H3K4me3 is the active mark for gene expression, and the bivalent state between H3K4me3 and H3K27me3 is crucial for *FT* regulation. PKDM7B (also known as JMJ14 or AtJMJ4) catalyzes H3K4 demethylation to repress *FT* expression. Loss of function of PKDM7B accompanies with increasing H3K4me3 and reduction of H3K27me3 which give rise to *FT* derepression thus early flowering (Jeong et al., 2009; Lu et al., 2010; Yang et al., 2010).

1.5 Objectives of this study

For a long time there have been doubts on the existence of PRC1 complex in plants due to poor conservation of PRC1 components between animals and plants. Recent studies revealed that AtBMI1a/1b and AtRING1a/1b, respective homologues of human BMI1 and RING, may function in the context of PRC1-like complex (Bratzel et al., 2010). However, additional components of this complex remain to be identified.

The severe phenotypes of *emf1* mutants indicate the critical roles of *EMF1* in *Arabidopsis* growth and development. Indeed, genome-wide studies imply that *EMF1* is involved in many developmental events including floral repression during the *Arabidopsis* life cycle. It was proposed that EMF1 may function like a PRC1 factor, but whether it acts as part of PRC1 is unclear. The objectives of this study include identifying the plant PRC1 components and understanding the mechanisms for EMF1-mediated flowering-time regulation.

CHAPTER 2

Materials and Methods

2.1 Plant materials

Arabidopsis thaliana ecotypes Columbia (Col) and Wassilewskija (Ws) were used as wild type plants in this study. Mutants *emf1-1, lhp1, co, gi, ft, clf, ebs* (cs906904), *pFT9:GUS* (about 9kb promoter) (Takada and Goto, 2003) and *PKDM7B:HA* (Yang et al., 2010) are in Col background. Mutant *shl* (FLAG546H05) is in Ws backgrounds. To make double mutants of *ebs shl* in a Col-like background, *shl* was crossed to *ebs* and the F_1 generation plant was backcrossed to *ebs* for three times. The final F_1 plants were self-segregated and double mutants of *ebs shl* were genotyped and examined the background by utilizing 24 pairs of Col/Ws molecular markers to make sure all these loci are in Col background. All transgenic plants such as *EMF1-RNAi, SHL:HA, LHP1:HA, EMF1:FLAG* are in Col background.

2.2 Plant growth conditions

For plants grown on plates, surface-sterilized seeds were plated on half Murashige and Skoog medium. For mutants related to *emf1-1* were grown on half Murashige and Skoog agar plates plus vitamins and 1.5% sucrose. For soil grown plants, seeds were directly sowed on soil surface. Seeds were placed at 4 ºC for 2 d before germination and then seedlings were grown under cool white fluorescent light in either long days (LDs) conditions (16 h light/ 8 h dark) or short days (SDs) conditions (8 h light/ 16 h dark) at ~22 ºC.

2.3 Plant transformation

Floral dip method was used to transform plants (Bent, 2006). *Agrobacteria* harboring interested plasmids were subcultured at 28 ºC overnight and collected by centrifuge at 3500 rpm for 15 min. Cell pellets were resuspended in 5% sucrose with 0.05% Silwet L-77 and 0.05% MES, pH 5.7. Flower buds were dipped in the Agro suspension for 40 s and the plants were put at a high humidity environment overnight, and then transferred to normal growth conditions for seeds mature.

2.4 Construction of plasmids

For double stranded RNAi vectors, the SUC2 promoter was amplified from *Arabidopsis* genome and cloned to binary RNAi vector pJawohl8-RNAi (GenBank no. AF408413) at restriction enzyme sites of AscI and XhoI to replace the 35S promoter. Then, the RNAi cassette with the SUC2 promoter and the NOS terminator was cut off with NheI and PmeI and cloned into binary vector pPZP212 (GenBank no. U10462.1) at site XbaI and SmaI.

For *EMF1* RNAi vectors, fragments amplified from target regions by PCR were first cloned into pENTR4 vector (Invitrogen) at site NcoI and XhoI, and were further introduced into the new RNAi binary vector for hairpin RNA production by gateway cloning (Invitrogen).

For *EMF1* overexpressing, the full length *EMF1* coding sequence amplified from Col seedling cDNAs was cloned into pENTR4 vector at site NcoI and NotI, and was subsequently cloned into an overexpression binary vector pLZB2GW7 by gateway cloning (Invitrogen).

For *EMF1:FLAG* construction, *EMF1* coding sequence (CDS) except the stop codon amplified from Col seedling cDNAs was cloned into the modified pENTR4 vector with FLAG (three copies) tag (in frame) at site NcoI and NotI first, and then the *EMF1* promoter (3.8kb upstream translation start site) was inserted just upstream *EMF1* CDS at site NcoI. Then, the *EMF1:EMF1:FLAG* fragment was introduced into binary vector pLZBGW via gateway cloning.

For *EMF1:GUS* vector, the whole genome of EMF1 from 3.8kb promoter upstream translation start site to the stop codon (without stop codon) was cloned into pENTR4 vector at site NcoI and NotI, and subsequently introduced into the upstream of the *GUS* reporter gene in the pMDC162 binary vector via gateway recombination.

For *pFT7:GUS* construction, *FT* native promoter from 6.9kb upstream region to the translation start site was cloned into pENTR4 vector at site NcoI and XhoI, and then inserted into the upstream of the *GUS* reporter gene in the pMDC162 binary vector via gateway cloning.

For the *pFT7:FT:GUS* vector, 6.9kb *FT* native promoter plus the whole *FT* codon region except stop codon was amplified from Col genome and inserted to pENTR4 vector at site NcoI and

XhoI, and then cloned to the upstream of the *GUS* reporter gene in the pMDC162 binary vector via gateway cloning. The genomic coding sequence was in frame with *GUS*.

To construct the *SHL:HA* vector, *SHL* cDNA sequence without stop codon was amplified and fused with a HA tag in pENTR4 vector at site BamHI and XhoI. Then the *SHL:HA* fragment was introduced into downstream of the *35S* promoter in the pMDC32 vector by gateway cloning.

For vectors in yeast two-hybrid assay, the full length of coding sequences of *EMF1*, *LHP1*, *SHL*, *EBS, EBSL, PKDM7B, AtBMI1a* or coding sequences of *EMF1*, *LHP1*, *SHL* domains were amplified and cloned into pGADT7 and/or pGBKT7 vectors (Clontech). The coding sequences were in frame with upstream GAL4 activation domain (AD) or DNA binding domain (BD).

2.5 Yeast two-hybrid assay

Yeast experiments were carried out according to the Matchmaker GAL4 Two-Hybrid System User Manual using yeast strain AH109 (Clontech). In short, competent yeast cells were made freshly and transformed with interest plasmids then grown on SD minus Leu and Trp plates. Single colony was picked up to subculture for one day and spotted on selective plates (SD-Trp, -Leu, -His (SD-3) or SD-Trp, -Leu, -His, -Ade (SD-4)) for growing in 30 ºC.

2.6 Histochemical β-glucuronidase (GUS) staining

Plants tissues were first treated with 90% acetone on ice for 30 min for fixing prior to staining. After washed with the staining buffer (5 mM EDTA pH 8.0, 0.05% Triton X-100, 2 mM Potassium ferrocyanide, 2 mM Potassium ferricyanide, 100 mM NaH_2PO_4 and 100 mM Na_2HPO_4), the plants tissues were incubated in the staining buffer containing 0.005 ‰ 5-bromo-4-chloro-3-indolyl-b-d-glucuronic acid (X-Gluc) at 37 ºC till clear signal appeared then stopped and de-stained with 70% ethanol.

2.7 RNA expression analysis

Total RNAs were extracted by using RNeasy mini plus kit (Qiagen) following the manufacturer's instructions and reverse transcribed into cDNAs with oligo (dT) as primer by murine leukemia virus reverse transcriptase (M-MLV) (Promega). For *FT* analysis at one time point, plant seedlings were collected at 8 h after lights on for long day conditions (LDs) and 7 h after lights on for short day conditions (SDs).

For semiquantitative RT-PCR, the cDNAs were first optimized to similar amounts based on *TUBULIN 2 (TUB)* levels. PCR products were viewed on agarose gels after ethidium bromide (EB) staining. *TUB* was used as an internal control.

To perform real time quantitative PCR, PCR amplifications were carried out on an ABI 7900HT Fast Real-Time PCR System with three experimental replicas using SYBR Green PCR master

mix (Fermentas). Following the amplification, a melting curve was done for verifying the specificity of PCR fragments. The relative RNA levels were normalized to at least 2 internal controls as *TUBULIN 2 (TUB)* and *At5g15710* (Vandesompele et al., 2002) by the $2^{-\Delta\Delta Ct}$ method. Primers for RNA expression analysis are shown in table 3.

2.8 Recombinant protein expression in *E.coli*

Plasmids of interest were transformed to the *E.coli* BL21 (DE3) strain for expression. *E.coli* cells with the expression plasmid were cultured at 37 ºC till OD_{600} reached to ~0.4. IPTG was added to the culture to final concentration of 0.5 mM to induce the fusion protein expression. The cells were further cultured for one to two hours at 37 ºC for induction. After induction, the cells were collected by centrifugation and resuspended in extraction buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Triton X-100, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1x Roche proteinase inhibitors). Cells were disrupted by sonication. Cell lysate was centrifuged, and the supernatant was collected for SDS-PAGE or for further analysis.

2.9 Protein pull-down assay

Cell lysate of induced recombinant protein expression cells was centrifuged. The supernatant was collected and mixed with glutathione-linked resins (Sigma). The GST fused proteins were precipitated with the resins and washed for three times with extraction buffer then mixed with plants proteins extract.

About 0.5g *Arabidopsis* seedlings harboring tag fused proteins were harvested and ground to powder in liquid nitrogen. The total proteins were extracted in 1 ml extraction buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Triton X-100, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1x Roche proteinase inhibitors) and mixed with prokaryotic expressed proteins precipitated with GST resins at 4 ºC for 4 h. Subsequently, proteins associated with GST resins were precipitated with the resins, washed with extraction buffer for three times and boiled with 2x SDS PAGE loading buffer.

2.10 Co-immunopreciptation

Co-immunoprecipitation (Co-IP) assay was used for investigating the *in vivo* interactions between EMF1 and its partners. About 0.5g 10-day-old seedlings were harvested and treated with 10 μM MG132 (protease inhibitor, Sigma) first and then grinded to powder in liquid nitrogen. Total proteins were extracted in 1 ml extraction buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.1% NP-40, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1x Roche proteinase inhibitors). Total proteins were mixed and incubated with 50 μl anti-FLAG M2 affinity gel (Sigma) for at 4 ºC for 4 hr. Subsequently, immunoprecipitated proteins were washed for three times with wash buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 10% glycerol, 1 mM PMSF and 1x Roche proteinase inhibitors) then mixed and boiled with 2x SDS-PAGE loading buffer. Western blotting was performed subsequently with anti-FLAG (Sigma) and anti-HA (Sigma) antibodies.

2.11 ChIP assay

For Chromatin immunoprecipitation (ChIP) assay, LD-growing 10-day-old seedlings of *pEMF1:EMF1:FLAG* in *emf1-1* T_3 homozygous at 8 h after light on were harvested and first treated with 10 μM MG132 (protease inhibitor, Sigma) and then ground to powder in liquid nitrogen. Subsequent operations were conducted following the Magana ChIP A kit (Millipore) instructions. Immunoprecipitated DNAs were subsequently quantified by real-time PCR. Of note, the ChIP assay in this thesis study was carried out mainly by my lab mate, Dr. Gu Xiaofeng.

Table 3. Sequences of primers used for RT-PCR

Amplified regions	**Sequences (5' → 3')**
FT	GACCTCAGGAACTTCTATACTTTGGTTATG
	CTGTTTGCCTGCCAAGCTG
EMF1	GCCAAAAGATTGTGGACTGC
	CTCCTAACATTAGAAGCACCCA
EBS	GATGGTATTACCGCCCTGAGGA
	CAACGTTTTCAAGCCTCGTGTAGT
SHL	TTCTGCAAGTGTGAGATGCC
	ACCTGGTCGCTTAGTGTGTTTGTTC
EBSL	GTATAACTCAGATGAACTGATGGAG
	TCATAACTCATAAGAGGCTCTTACATC
TUB	ATCCGTGAAGAGTACCCAGAT
	AAGAACCATGCACTCATCAGC

CHAPTER 3

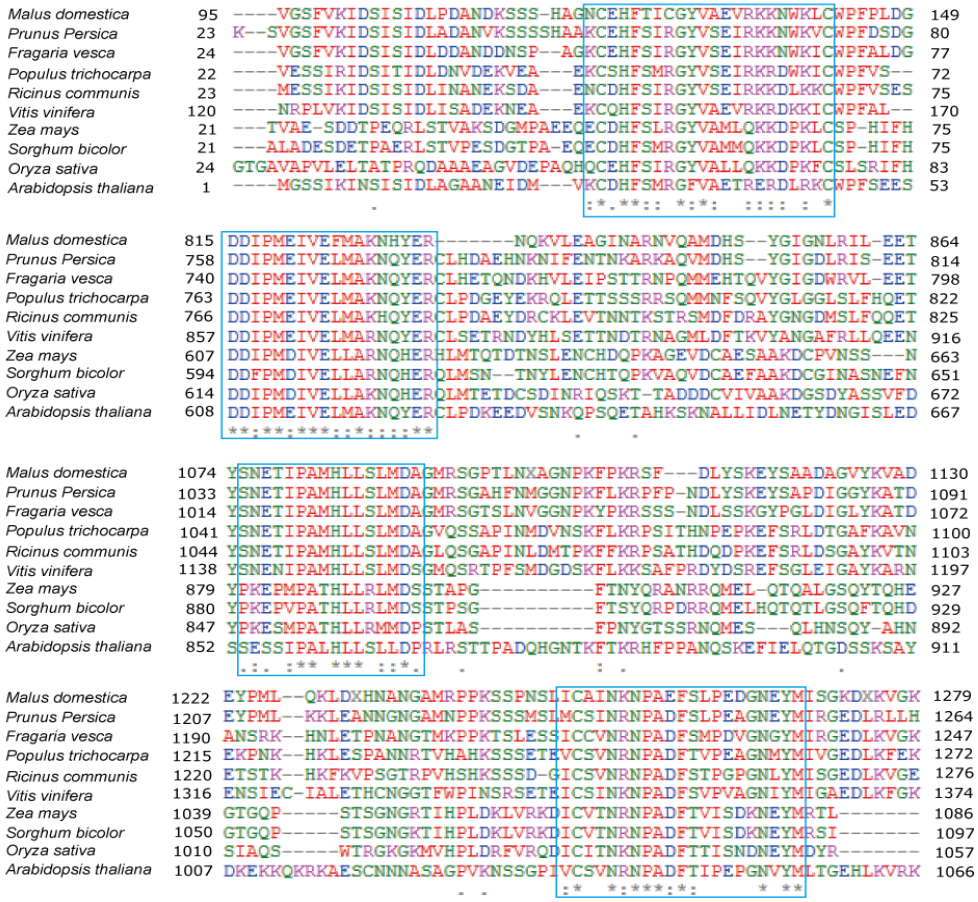
Results

3.1 EMF1 is conserved in higher plants

Previous studies show that EMF1 is involved in the repression of *AG* expression. The embryonic flowering phenotype of *emf1* mutants indicate the crucial role of *EMF1* in flowering repression, which stimulated our interests in how EMF1 is involved in flowering-time controlling (Calonje et al., 2008). EMF1 is a plant-specific protein, and no obvious homologues have been reported other than in plants. Data mining in protein databases showed that EMF1 homologues were only present in angiosperms. Considering the function of EMF1 in *Arabidopsis*, the higher plants might evolve a new path for gene regulation. However, little is known in other species about the EMF1 homologues since they have not been annotated yet.

Sequence comparison and the phylogenetic tree based on multi-sequence alignments indicated that EMF1 in monocots are much closer to *Arabidopsis* than others (Figure 1B). The EMF1 proteins showed high divergence while they still shared some conserved regions. Based on the sequence alignments, there were four conserved regions in EMF1 proteins which are located at the N-terminal (from 25 to 46aa in AtEMF1), the middle domain (608 to 625aa and 853 to 868aa in AtEMF1), and the C-terminal region (1036 to 1056aa in AtEMF1) (Figure 1A). These regions may be crucial for the function of EMF1. Indeed, we found the N-terminal domain and the Middle domain which contain the conserved regions were required for the interaction with other proteins which will be discussed later.

A



B

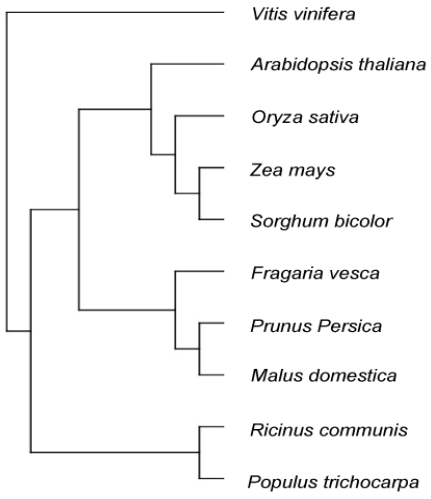


Figure 1. EMF1 is conserved in higher plants. (A) Multi-sequences alignment of the EMF1 homologues in different species by clustalw (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Conserved regions of EMF1 proteins were indicated by blue boxes. (B) Phylogenetic tree of EMF1 homologues generated by the PHYLIP software program.

3.2 EMF1 forms a complex with LHP1

Previous studies indicate that EMF1 participates in vast aspects of *Arabidopsis* development (Kim et al., 2012). We speculate that EMF1 may form a complex with other proteins for its function. First, yeast two-hybrid assay was carried out to screen for EMF1 partners. The interactions between EMF1 and other chromatin modifiers or related proteins such as the PRC2 components, and LHP1 were screened. The full length EMF1 protein was fused with the GAL4 DNA-binding domain (BD) and other proteins were fused with the GAL4 activation domain (AD). The results showed that LHP1 has a strong interaction with EMF1, as indicated with growth of yeast on stringent selection medium (SD-Leu, Trp, His, Ade) (Figure 2A).

Co-immunoprecipitation (Co-IP) assay was further performed to confirm the *in vivo* interaction between EMF1 and LHP1 in *Arabidopsis*. FLAG tagged EMF1 under native *EMF1* promoter (*pEMF1:EMF1:FLAG*) was introduced into *emf1* heterozygous plants by *Agrobacterium*-mediated transformation. Two T_1 plants with single T-DNA insertion and in *emf1* heterozygous backgrounds were selected for segregation to get *emf1* homozygous. The T_2 plants were genotyped and the transgenic plants in *emf1* mutant background were identified. These plants behaved like wild-type plants, which indicated the EMF1:FLAG fusion protein is fully functional. Meanwhile, overexpressing LHP1:HA (*p35S:LHP1:HA*) was introduced into *lhp1* mutants by *Agrobacterium*-mediated transformation. 19 out of 27 T_1 transgenic plants showed rescued phenotype, which indicated the LHP1:HA fusion protein is functional too. The F_1 plants of *pEMF1:EMF1:FLAG* in *emf1* crossing *p35S:LHP1:HA* in *lhp1* were used for Co-IP. Total proteins were extracted from F_1 plant seedling. By using FLAG beads, the LHP1:HA proteins

were pulled down(Figure 2B), which illustrates the interaction between EMF1 and LHP1 in *Arabidopsis* seedlings.

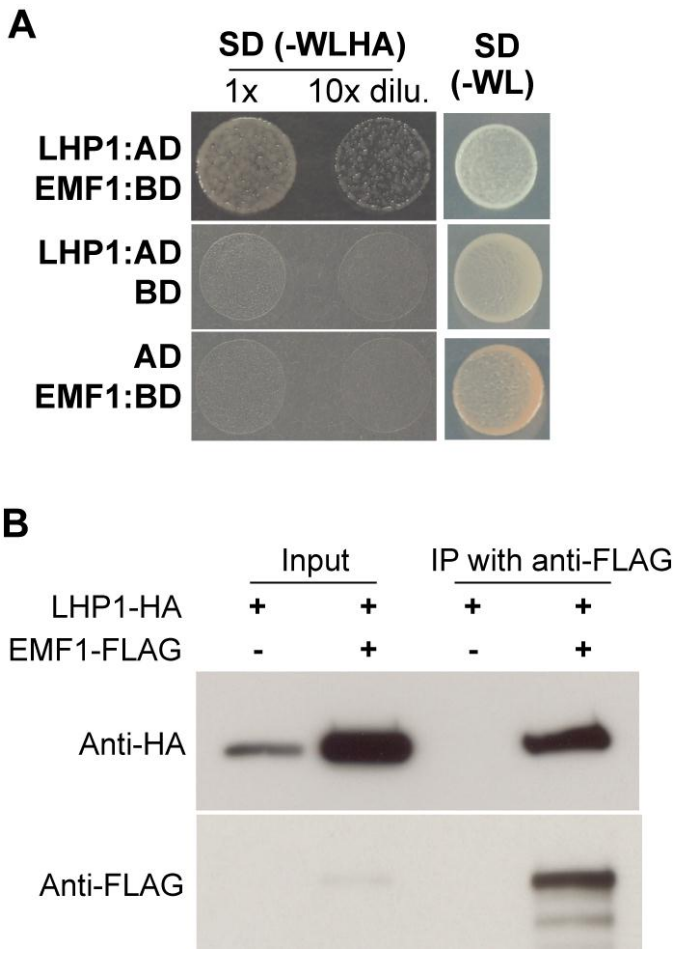


Figure 2. EMF1 and LHP1 form a complex. (A) Interaction of EMF1 with LHP1 in yeast. Full length EMF1 and LHP1 proteins were respectively fused with the GAL4 DNA-binding domain (BD) and activation domain (AD). Yeast cells harboring these fusions as indicated were grown on the highly selective SD media lacking of Trp (W), Leu (L), His (H), Adenine (A) (SD-4). (B) Co-immunoprecipitation analysis of EMF1 with LHP1. Total proteins were extracted from *LHP1:HA* (as a negative control) and hemizygous of *LHP1:HA* and *EMF1:FLAG* seedlings. Protein extracts were immunoprecipitated with anti-FLAG agarose, and the precipitates were analyzed by western blot with anti-HA and anti-FLAG.

No recognizable domains are present in the EMF1 protein except the nuclear localization signals, a putative ATP/GTP binding motif (P-loop) (from 522 to 528aa) and a putative LXXLL motif (from 266 to 270aa) which are present in many transcription factors (Aubert et al., 2001), while the actual functions of these motif for EMF1 are still unclear yet. To identify the interaction regions of EMF1 with LHP1, I broke EMF1 into several fragments based on the sequence alignment results (Figure 1A) and fused them with GAL4 BD domain to check their interactions with LHP1 by yeast two−hybrid assay. The results indicated that LHP1 could interact with both fragments when EMF1 was divided into two parts (N-terminal from 1aa to 608aa, C-terminal from 604aa to the end) (Figure 3A) and with the N-terminal region and the second middle region when EMF1 was broken into four parts (N-terminal from 1aa to 336aa, M_1 region from 337aa to 617aa, M2 region from 618aa to 886aa, and C-terminal from 887aa to the end) (Figure 3B). Both the N-terminal region and M2 region harbor the conserved regions based on the sequence alignment. However, the C-terminal also contained a conserved region without detection of interaction with LHP1, which indicated this region at C-terminal might not be involved in protein interactions. Then, the N-terminal and M2 regions were further divided to refine the interaction region. It was found that as short as 49 amino acids at N-terminal region (N_{49}-EMF1) and the first part of M2 (M-EMF1 from 618 to 745 amino acids) that includes the conserved region, could interact with LHP1 (Figure 3C). These interactions between N_{49}-EMF1, M-EMF1 and LHP1 were further verified by pulldown assays. Prokaryotic expressed proteins N_{49}-EMF1 and M-EMF1 fused with glutathione S-transferase (GST) were extracted and mixed with protein extraction of *p35S:LHP1:HA* in *lhp1* T_3 seedlings. Both GST fused N_{49}-EMF1 and M-EMF1 proteins could pull down HA tagged LHP1, which indicates EMF1 interacts with LHP1 through these two regions (Figure 3D).

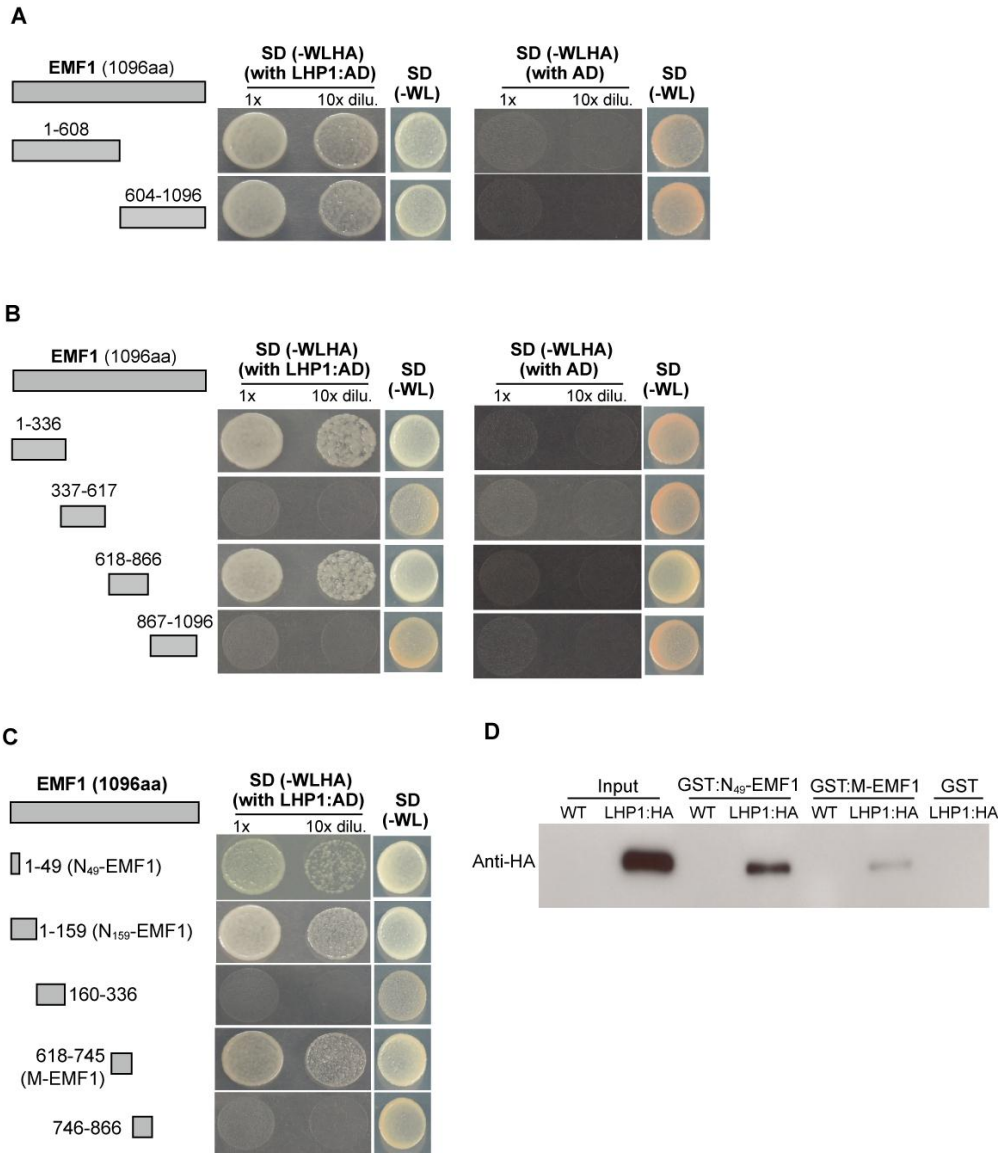


Figure 3. Interactions between EMF1 domains and LHP1. (A) Interactions in yeast between two parts of EMF1 and LHP1. EMF1 was broken to two parts (numbers indicates the amino acids) and fused with the GAL4 BD domain. Yeast cells were grown on SD-4 media. (B) Interactions in yeast between four parts of EMF1 and LHP1. EMF1 was separated to four parts: N-terminal, M1, M2, C-terminal. (C) Interactions of EMF1 N-terminal and M2 regions with LHP1:AD in yeast. EMF1 N-terminal was broken into two parts N_{159}-EMF1 (1-159aa) and the rest (160-336aa) and further refined to N_{49}-EMF1 (1-49aa). M2 region was separated to M-EMF1 (618-745aa) and the rest part. (D) Pulldown analysis of the interactions between the EMF1 domains and LHP1. *E.coli* expressed GST fused N_{49}-EMF1 and M-EMF1 were mixed with protein extracts from seedlings of Col (as control) and *LHP1:HA*. After precipitated with GST-beads, the precipitates were analyzed by western blot with anti-HA.

On the other hand, we were also interested in the region of LHP1 that interacts with EMF1. LHP1 is the only homologue of animal HP1 protein, but functions differently from HP1 in *Arabidopsis*. In plant, LHP1 is proposed to play the role of Pc which is a core component of animal PRC1 complex. The HP1 protein has two conserved domains: chromodomain and chromo-shadow domain. The chromodomain has been found to mediate the interactions between HP1 and di- or tri-methylated lysine 9 of Histone 3 (H3K9me2 or me3), and the chromo-shadow domain is linked with the recruitment of other proteins (Jacobs and Khorasanizadeh, 2002; Nielsen et al., 2002). Like HP1, LHP1 also contains the chromodomain at the middle region (from 104aa to 167aa) and the chromo-shadow domain at C-terminal. The chromodomain has been reported to be related to the recognizing and binding of H3K27me3. The chromo-shadow domain is speculated to be involved in the recruitment of other repressive proteins (Turck et al., 2007). To identify the interaction region of LHP1 with EMF1, I fused the chromodomain, the chromo-shadow domain, the N-terminal region and the rest part of LHP1 with GAL4 AD domain. Their interactions with EMF1-BD were examined and the results suggested that LHP1 interacts with EMF1 through its chromo-shadow domain. Yeast two-hybrid assay showed that the C-terminal (chromo-shadow domain) of LHP1 could interact with both the whole EMF1 protein (Figure 4A) and the N-terminal (N_{159}-EMF1) and M-EMF1 (Figure 4B).

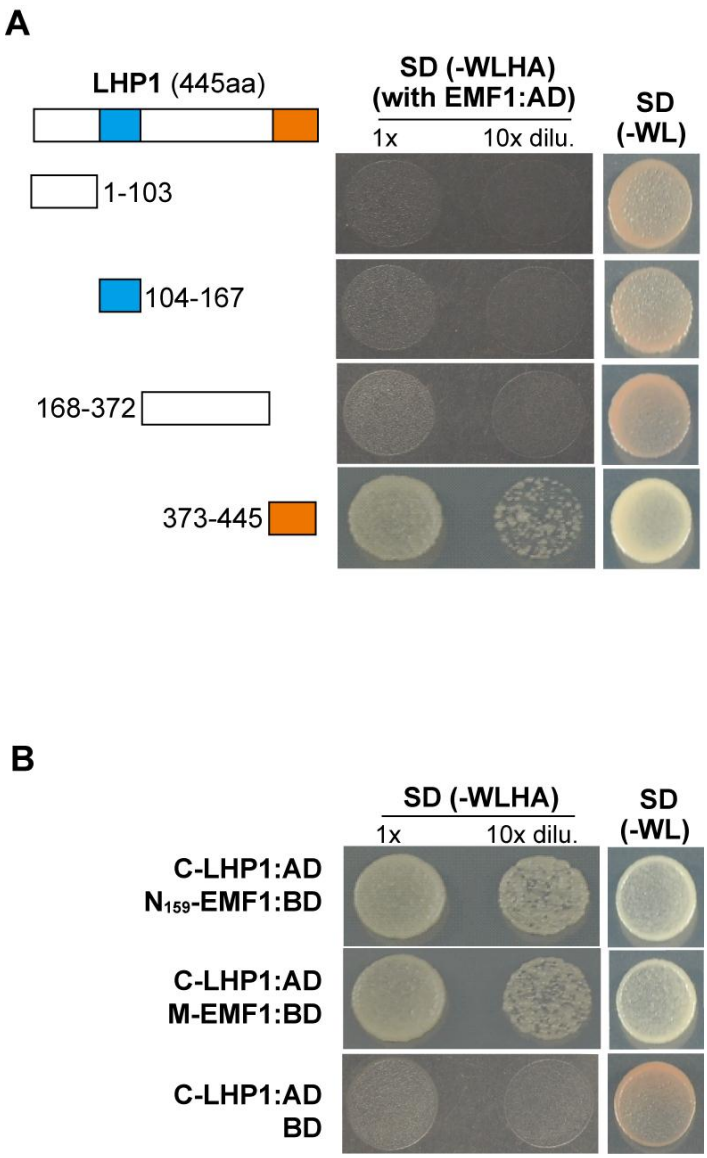


Figure 4. Interactions between EMF1 and LHP1 domains. (A) Interactions of LHP1 domains with full length EMF:AD in yeast. LHP1 was separated into four parts (numbers indicate the amino acids). The blue colored region represents the chromodomain of LHP1. The C-terminal colored with orange indicats the chromo-shadow domain. (B) Interactions between LHP1 C-terminal with EMF1 domains N_{159}-EMF1 and M-EMF1 in yeast.

The genetic interaction between LHP1 and EMF1 was also examined. Because the *emf1* mutants skip vegetative development and are sterile, the *lhp1 -/- emf1 +/-* plants were genotyped to get double mutant plants. The *emf1 lhp1* double mutants behaved identically to *emf1* single mutants. The weak *emf1* mutants (*emf1-1*) could develop to the stage with a short inflorescence which has a terminal flower with pistils, stamens, and sepal structures and is sterile. At 8 DAG (days after germination), comparing with wild type and *lhp1* single mutants with flat, round and long petiole cotyledons (Figure 5A and B), the double mutants (Figure 5D) displayed a phenotype like *emf1* single mutants (Figure 5C) to form oval shaped and petioleless cotyledons, and small short leaf. By 15 DAG, the double mutants and *emf1* single mutants flowered, bypassing rosette growth and forming flower structures (Figure 5F and G) while the WT was still at vegetative stage (Figure 5E). There seemed to be non-additive effects of *lhp1* and *emf1* on *Arabidopsis* development. Together with the molecular interaction between LHP1 and EMF1, our data indicated that LHP1 and EMF1 could form a complex for their functions.

Interestingly, overexpression *EMF1* could partially rescue the *lhp1* mutant phenotypes. The *lhp1* mutants showed phenotypes including dwarf plants, small narrow and curled leaves, in addition to early flowering (Figure 6A)(Gaudin et al., 2001). When *p35S:EMF1* was introduced into the *lhp1* mutants by *Agrobacterium*-mediated transformation, 4 out of 9 transgenic plants showed partially rescued phenotypes. We chose 3 single-copy T-DNA insertion lines as the reference lines and propagated to get homozygous for further examination. The leaves of the transgenic plants exhibited wild type like properties: larger than that of *lhp1* mutants and flat (Figure 6A). The early flowering phenotype was also partially rescued: though still earlier than wild type plants, the *EMF1* overexpression lines in *lhp1* flowered later than *lhp1* mutants (Figure 6A and

C). The extent of the rescue of *lhp1* corresponded to the *EMF1* expression levels: the *p35S:EMF1 in lhp1*-3 line had the lowest *EMF1* expression level and flowered the earliest among these three reference lines, while the other two lines got higher *EMF1* expression level and flowered later (Figure 6A, B and C), which suggests that the rescue is due to *EMF1* overexpression. These data show that increased EMF1 expression can compensate for loss of LHP1 function.

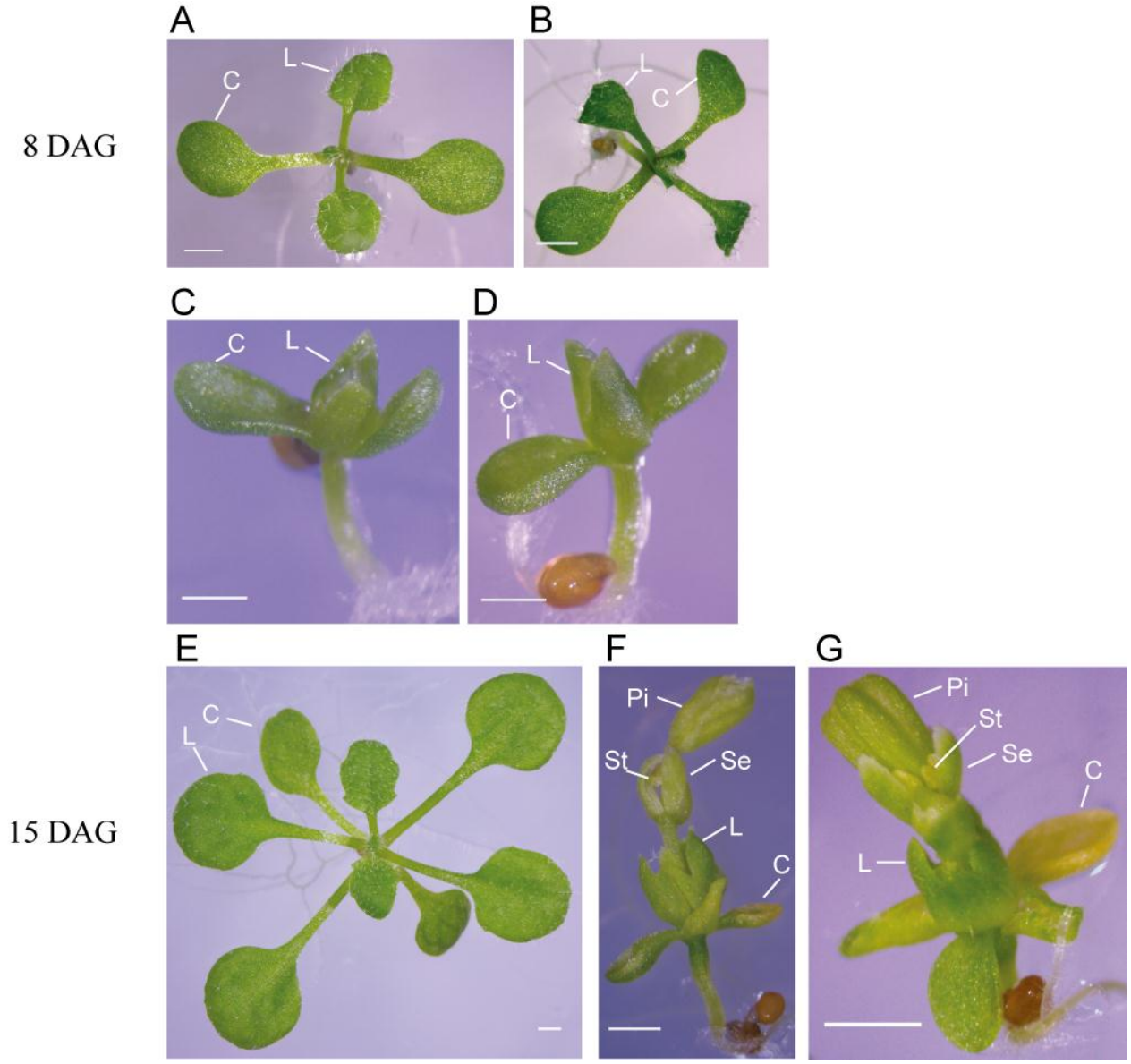
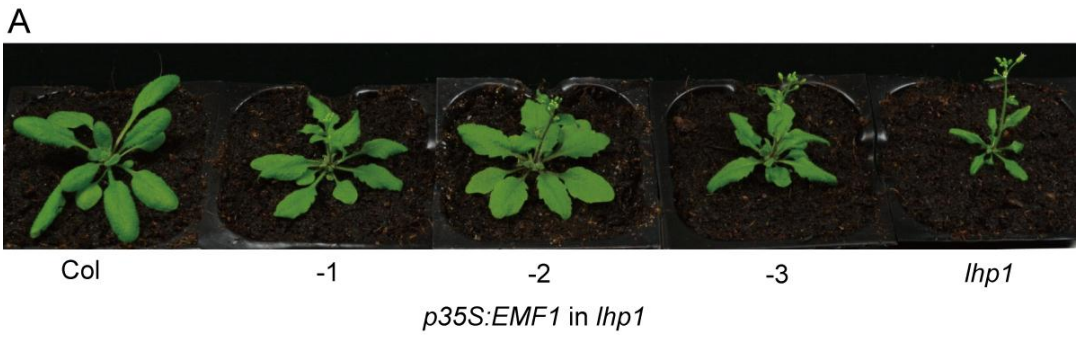
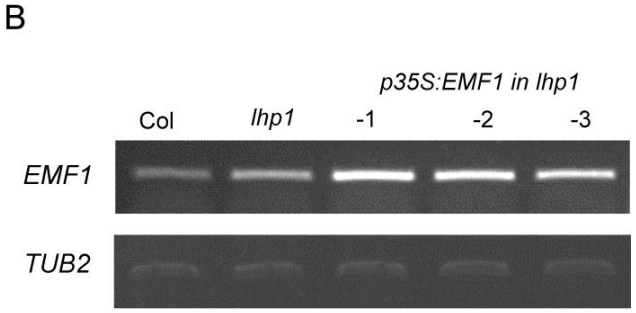


Figure 5. Genetic interactions between *EMF1* and *LHP1*. (A) to (D) 8-day-old seedlings of (A) Col, (B) *lhp1*, (C) *emf1-1*, (D) *lhp1 emf1-1* double. (E) to (G) 15-day-old seedlings of (E) Col, (F) *emf1-1*, (G) *lhp1 emf1-1* double. Scale bar = 1mm. C, cotyledon; L, leaf; Pi, pistil; Se, sepal; St, stamen. DAG, days after germination.





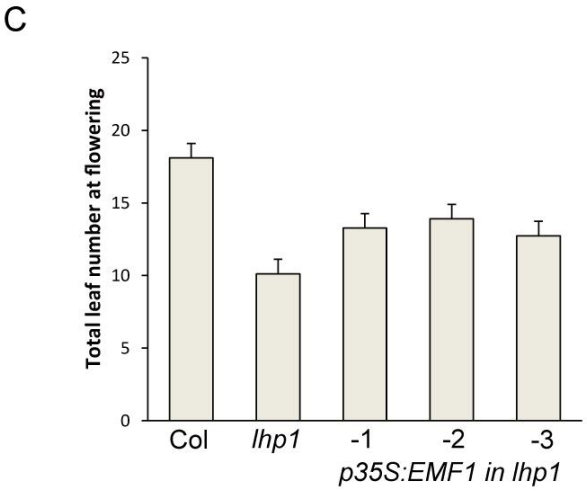


Figure 6. Overexpression of *EMF1* partially rescues *lhp1*. (A) Picture of plants Col, three independent transgenic lines of *p35S:EMF1* in *lhp1* background (T3) and *lhp1*. (B) RT-PCR analysis of *EMF1* expression in seedlings of Col, *lhp1*, and three *p35S:EMF1* in *lhp1* transgenic lines. *TUBULIN 2 (TUB2)* served as an internal control. (C) Flowering time of Col, *lhp1*, and three *p35S:EMF1* in *lhp1* transgenic lines. The flowering time was measured by the total leaf numbers a plant produced prior to flowering. Error bars indicate SD.

3.3 EMF1 associates with the plant-specific SHL family proteins

Like EMF1, SHORT LIFE (SHL) family proteins are plant-specific proteins, which include SHL, EARLY BOLTING IN SHORT DAYS (EBS) and another protein we named as EBS-Like (EBSL) (Figure 7). *EBS* has been reported to be involved in flowering-time control by repressing *FT* expression (Pineiro et al., 2003). *SHL* may affect fertility and influence plant proper development (Mussig et al., 2000; Mussig and Altmann, 2003), while *EBSL* has not been reported to date. Previous studies indicate that EBS may be involved in chromatin modification (Pineiro et al., 2003); hence we explored whether EBS and its homologues could function together with EMF1. Therefore I examined their interactions by yeast two-hybrid analysis and found that the SHL family proteins could also interact with EMF1.

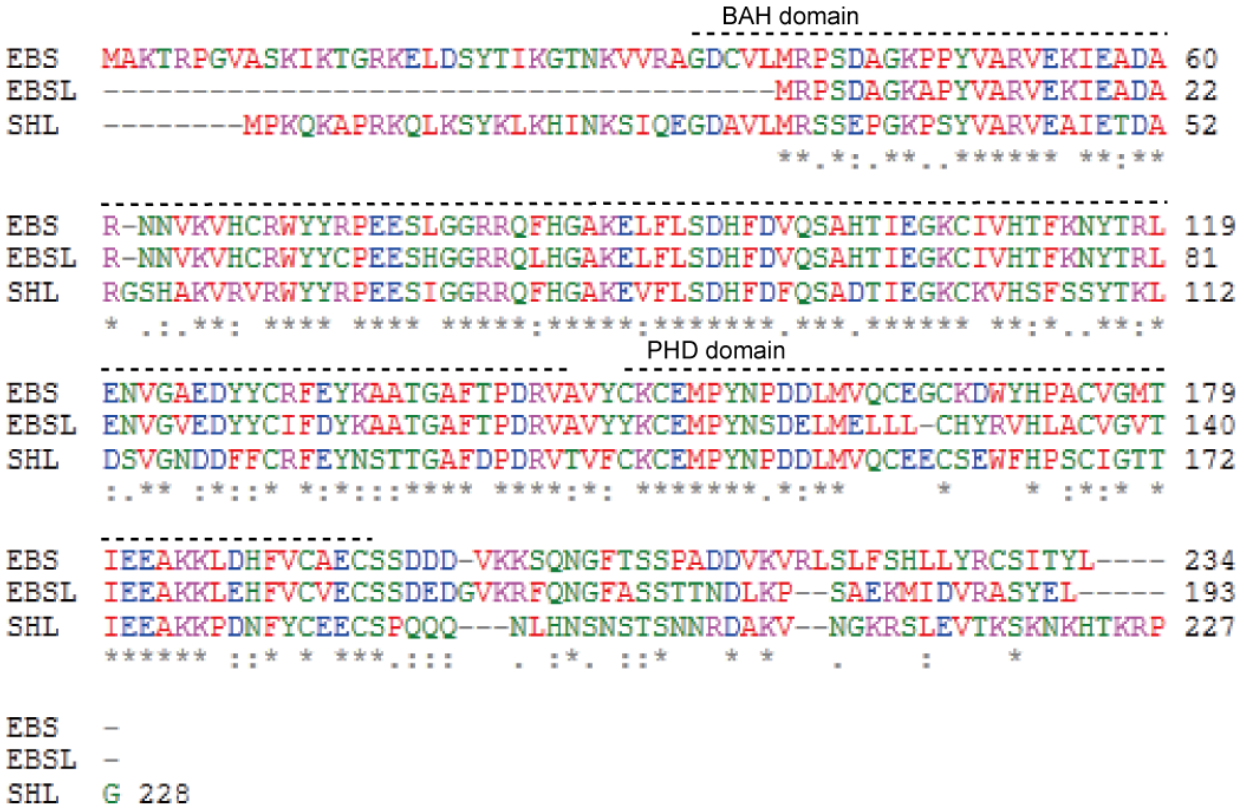


Figure 7. Sequence alignment of SHL family proteins. The BAH domain and PHD domain are indicated above the sequences. Numbers at the right side refer to the amino acid residues.

The results of yeast two-hybrid assay showed that EMF1 strongly interacted with SHL: the yeast cells harboring SHL-AD and EMF1-BD could grow up on the stringent selective medium (SD-Leu, Trp, His, Ade) (Figure 8A). To check the *in vivo* state, the CoIP experiments were further performed. The overexpression of HA fused SHL (*p35S:SHL:HA*) was introduced into Col plants by *Agrobacterium*-mediated transformation. A single copy T-DNA insertion transgenic line was select to get homozygous and further investigation. The F_1 seedlings that co-expressed FLAG tagged EMF1 driven by native *EMF1* promoter and overexpressing SHL:HA were used for analysis. Using FLAG beads, the SHL:HA proteins were pulled down indicating the *in vivo* interaction between EMF1 and SHL (Figure 8B).

The SHL family proteins harbor two conserved motifs: an N-terminal Bromoadiacent Homology (BAH) domain and a C-terminal Plant Homeodomain (PHD) Zn finger domain. The BAH domain is present in many transcriptional regulators related to chromatin remodeling such as POLYBROMO1 (PB) in vertebrates and ORIGIN OF REPLICATION COMPLEX 1 (Orc1) in yeast, which are involved in gene repression (Goodwin and Nicolas, 2001; Noguchi et al., 2006). Like BAH domain, the PHD motif is also reported in chromatin remodeling factors like Polycomb-like (PCL) in *Drosophila*, corepressor KAP1 in human (Schultz et al., 2001). It appears that both BAH domain and PHD domain could mediate protein-protein interactions in protein complexes. Higher conservation was observed in these two domains compared to the rest of regions from the sequence alignments (Figure 7).

To investigate which domain of SHL mediates the interaction of SHL with EMF1, the BAH domain and PHD domain were fused with GAL4 AD domain and the interactions with EMF1-

BD were examined by yeast two-hybrid assay. It was found that the BAH domain, but not the PHD domain, could interact with EMF1, which indicates that the BAH domain mediates the interaction between SHL and EMF1 (Figure 8C).

I further examined the interactions between EMF1 and the other two members of SHL family proteins. The interaction between EBS and EMF1 was detected in the yeast two-hybrid assay (Figure 9A) with the growth of yeast cells harboring both EMF1-AD and EBS-BD on the stringent selective medium (SD-Leu, Trp, His, Ade). And the interaction between EBSL and EMF1 was also detected in our yeast two-hybrid experiments: the yeast cells harboring EBSL-AD and EMF1-BD could grow up on the SD-4 medium (SD-Leu, Trp, His, Ade) (Figure 9B). Together, these results show that EMF1 directly interacts with the SHL family proteins.

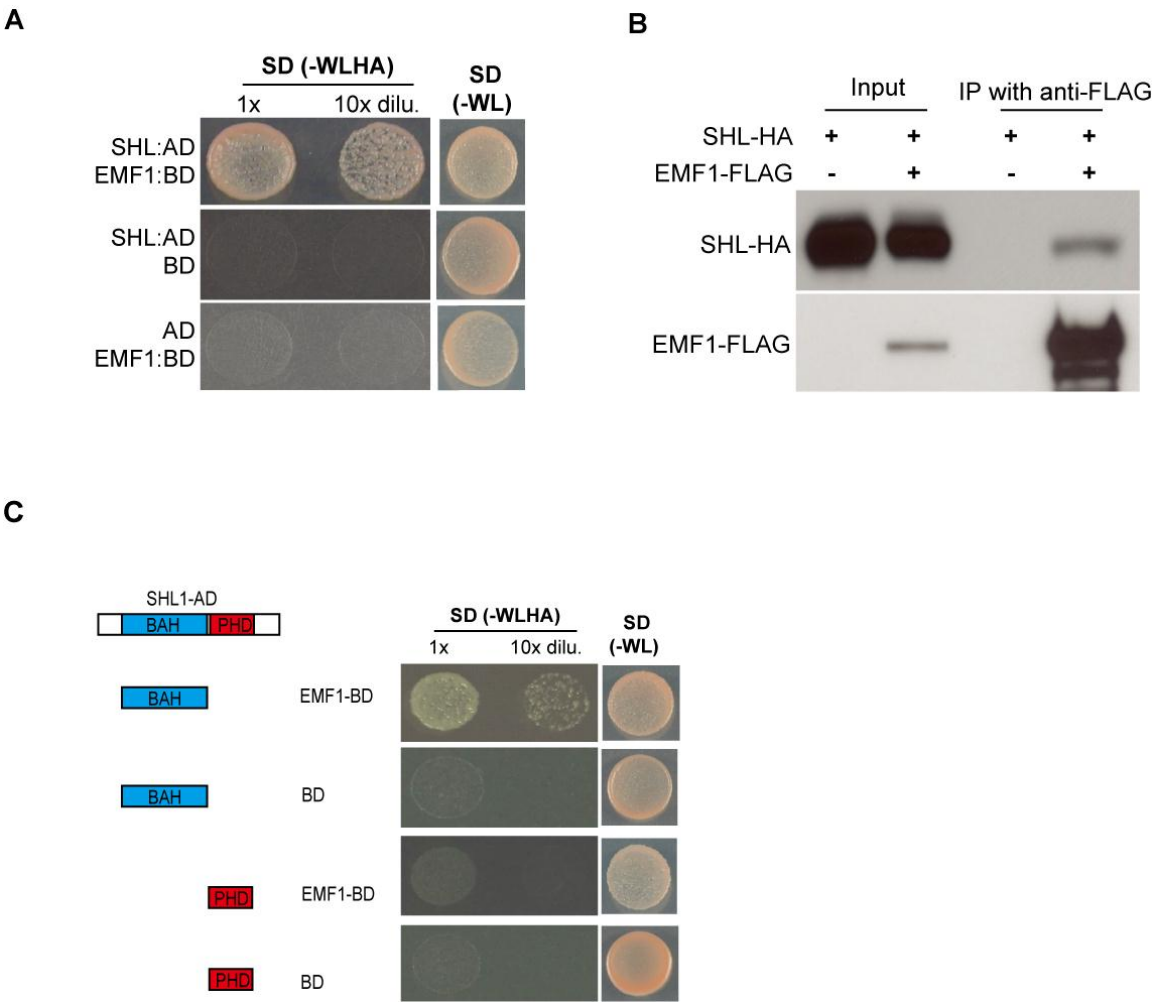


Figure 8. SHL interacts with EMF1. (A) Interaction of EMF1 with SHL in yeast. Full length EMF1 and SHL proteins were respectively fused with the GAL4 BD and AD. Yeast cells harboring these fusions as indicated were grown on the highly selective SD-4 media. (B) Co-immunoprecipitation analysis of the interaction of EMF1 with SHL. Total proteins were extracted from *SHL:HA* (as a negative control) and hemizygous of *SHL:HA* with *EMF1:FLAG* seedlings. Protein extracts were immunoprecipitated with anti-FLAG agarose, and the precipitates were analyzed by western blot with anti-HA and anti-FLAG. (C) Yeast interactions of SHL domains with full length EMF:BD. The BAH domain of SHL was indicated with blue color. The PHD domain was in red.

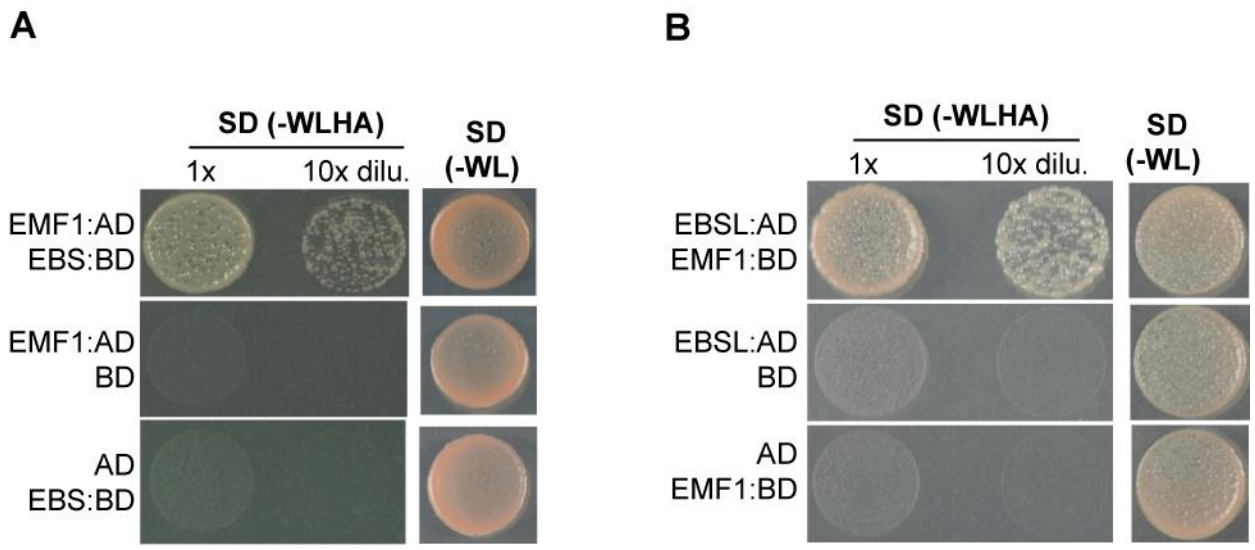


Figure 9. Interactions between EMF1 and other SHL family proteins. (A) EMF1 interacts with EBS in yeast. Full length EMF1 and EBS proteins were fused respectively with the GAL4 AD domain and BD domain. Yeast cells harboring fusion proteins indicated in the left side were grown on the SD-4 media for selection. (B) EMF1 interacts with EBSL in yeast. EBSL was fused with GAL4 AD domain and EMF1 was fused with BD domain. Yeast cells were grown on SD-4 media for selection.

3.4 SHL family proteins function as part of PRC1

Besides the putative PRC1 component EMF1, the interactions between SHL and other components of PRC1 complex were further investigated by yeast two-hybrid assays. It was found that SHL could interact with AtBMI1a in yeast: the yeast cells harboring AtBMI1a-AD and SHL-BD could grow up on the selective medium (SD-Leu, Trp, His) (Figure 10A). And the interaction was further confirmed by the pulldown analysis: *E.coli* expressed GST-fused AtBMI1a was mixed with the protein extraction of *p35S:SHL:HA* T3 seedlings, after immunopreciptation with GST beeds (see Materials and Methods), GST-fused AtBMI1a could pull down SHL:HA proteins from *Arabidopsis* seedlings (Figure 10B).

The *SHL* gene (*At4g39100*) contains five exons and encodes a 228 amino acids protein. The gene *EBS (At4g22140)* encodes a 284aa protein, while *EBSL (At4g04260)* encodes a 211aa protein (Figure 11A). However, when comparing to *EBS* and *SHL* with high expression level in seedlings, we could barely detect *EBSL* expression in our RT-PCR experiments (Figure 11B).

We searched T-DNA insertion mutants of these genes and identified one for both *SHL* and *EBS*, but no insertion lines for *EBSL* available. The T-DNA of FLAG546H05 (named as *shl*) inserted in the third intron of *SHL* was in Ws background (Figure 11A). The insertion of CS906904 (named as *ebs*) happened at the first intron of the *EBS* gene (Figure 11A). Both insertions knocked down respective gene expression. Obvious decrease in *SHL* expression level was detected in *shl* mutants compared to wild type WS (Figure 11D). Although CS906904 was not a

null allele, we could not detect *EBS* expression in CS906904 mutants until up to 40 PCR cycles at which a weak band was detected (Figure 11C).

For the FLAG546H05 line in the Ws background, we attempted to create double mutants of *ebs* and *shl* in a Col-like background, FLAG546H05 was crossed to the *ebs* mutants, and the F1 plant was back crossed to *ebs* for 3 more times and then self-segregated. The *ebs shl* double mutants with the Col background were identified by genotyping with 24 Col/WS molecular markers (see Materials and Methods). The *shl* single mutants (parent line) in Ws behaved the same as wild-type (Figure 11E) which may be due to the redundancy with *EBS* or that sufficient proteins still exist for normal plant development although the transcripts of *SHL* were reduced in the mutants. Unlike the *shl* mutants, the *ebs* single mutants developed normally except slightly earlier flowering than WT (Figure 11E and F) and with reduced fertility (Gomez-Mena et al., 2001; Pineiro et al., 2003). Interestingly, the double mutants of *ebs shl* exhibited much more severe phenotypes than *ebs* single mutants. This suggests that EBS and SHL act redundantly to control *Arabidopsis* development. The *ebs shl* double mutants showed reduced fertility, developed into small and dwarf plants and flowered much earlier, which were similar to the *lhp1* mutants (Figure 11E and F). The genetic evidence, together with the molecular association with EMF1 and AtBMI1a, suggest a LHP1-EMF1-SHL/EBS PRC1 complex in *Arabidopsis*.

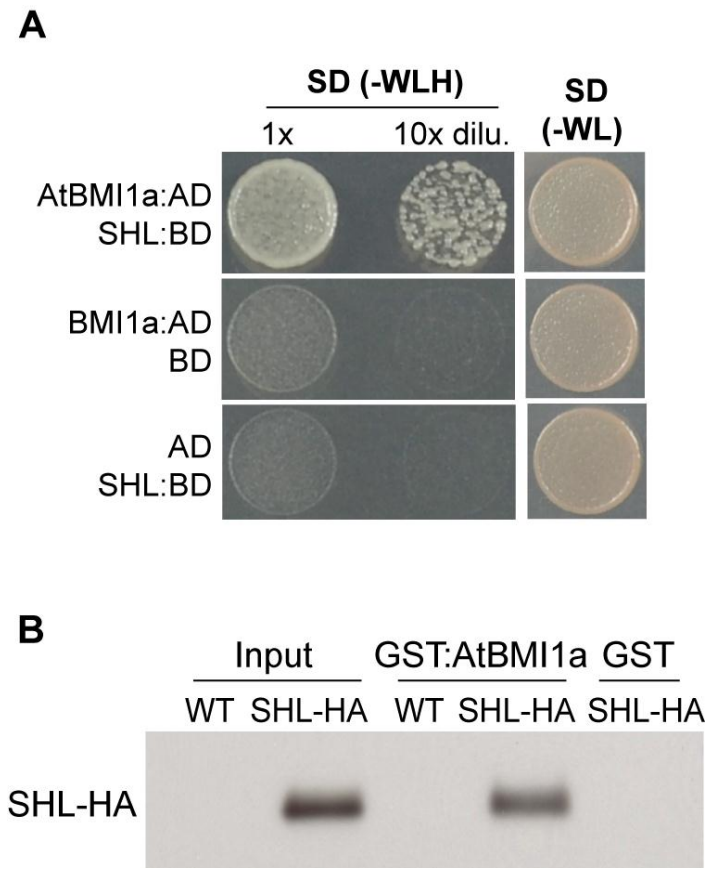


Figure 10. SHL directly interacts with AtBMI1a. (A) Yeast two-hybrid assay on the interaction between SHL and AtBMI1a. Full length SHL and AtBMI1a proteins were fused respectively with BD domain and GAL4 AD domain. Yeast cells harboring fusion proteins indicated in the left side were grown on the SD-3 media (SD-Trp, Leu, His) for selection. (B) Pulldown analysis of the interaction between SHL and AtBMI1a. *E.coli* expressed GST fused AtBMI1a were mixed with protein extracts from plants Col (as control) and *SHL:HA* seedlings. After precipitated with GST-beads, the precipitates were analyzed by western blot with anti-HA.

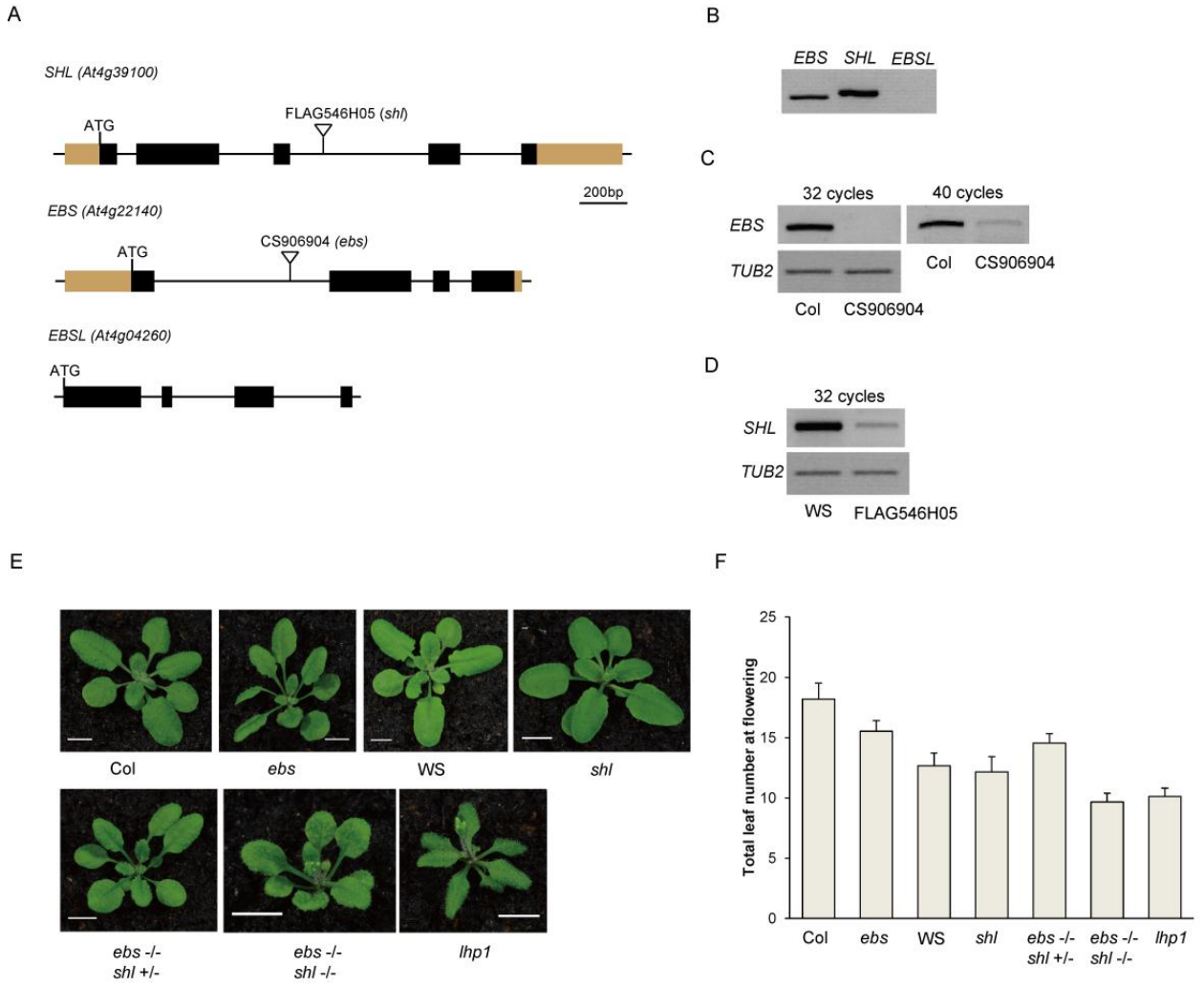


Figure 11. *ebs shl* double mutant causes early flowering. (A) Diagram of gene structures of the SHL family members. Dark boxes represent exons; Orange boxes represent untranslated regions (UTR); lines represent introns; and triangles indicate T-DNA insertions. (B) RT-PCR analysis of *EBS, SHL, EBSL* expression in 10-day-old Col seedlings. (C) RT-PCR analysis of *EBS* expression in Col and *ebs* mutants. *TUBULIN 2 (TUB2)* served as an internal control. Numbers above the gel pictures indicate the PCR cycle numbers. (D) RT-PCR analysis of *SHL* expression in *shl* mutants. (E) 20-day-old seedlings of *ebs shl* single, double and *lhp1* mutants. Col, Ws served as control. The genotypes are indicated at the bottom of picture. The background of *shl* was switched from Ws to Col by backcross of *shl* to *ebs* mutants to get *ebs shl*. Bar = 1cm. (F) Flowering time of *lhp1*, and *ebs shl* mutants. Error bars indicate SD.

3.5 EMF1-PRC1 complex represses *FT* expression to inhibit the floral transition

LHP1 has been reported to play a role in *FT* repression, and LHP1 proteins can directly bind to *FT* chromatin (Adrian et al., 2010). *FT* expression has also been reported to be increased in *emf1* mutants, but it was postulated that EMF1 was not involved in *FT* regulation (Moon et al., 2003). Whether and how EMF1 regulates *FT* expression is hitherto unknown. To understand the mechanism of EMF1 mediated *FT* repression, I knocked down *EMF1* specifically in the vasculature tissues where *FT* is expressed by the RNA interference (RNAi) strategy. To construct the double stranded RNAi (dsRNAi) vectors, the promoter of vasculature-specific *Arabidopsis SUCROSE-PROTON SYMPORTER 2 (SUC2)* (Figure 12D) was used to drive the RNAi cassette. In addition, *EMF1* gene was targeted at different regions: the 5'-UTR (region A), 3'-UTR (region E), the fifth exon (region C), the sixth exon (region D) and the first intron (region B) as a negative control (Figure 12B). The RNAi vectors were introduced into Col plants by the *Agrobacterium*-mediated way and the phenotypes of T_1 transgenic plants were scored in long day conditions (LDs). Except the T1 plants expressing the RNAi targeting the intron region (region B), most of the T_1 plants showed early flowering and curly leaf (Figure 12A). According to the phenotypes, the transgenic plants were roughly grouped into three categories: severe, the plants flowered very early and were very small; moderate, the plants flowered later than the severe plants but still much earlier than wild-types and usually had the first two or three pairs of leaves curved; Col-like, the plants behaved similar to Col plants with both the flowering time and leaf shape (Figure 12A). When the 5'-UTR (n=54) was targeted, half of the T_1 plants showed moderate phenotype and the other half showed Col-like phenotype. When region C (n=50) was targeted, 6% of T_1 plants exhibited severe and 40% plants exhibited moderate phenotype. When

region D (n=55) was targeted, 5.5% T_1 plants behaved severe and 74.5% plant behaved moderate, while region E (n=43) was targeted, 2.3% severe and 58.2% moderate. Comparing to these four regions, when the intron region (n=32) was targeted, the T_1 plants all showed Col-like phenotypes (Figure 12C). The curved leaf phenotype may result from ectopic expressing the flower homeotic gene *AGAMOUS (AG)* (Calonje et al., 2008). Three single-copy inserted T-DNA lines targeting at the last exon of *EMF1* (region D and Figure 12B) with a moderate phenotype were selected as our reference lines, and were named as *EMF1-RNAi-1*, *EMF1-RNAi-2*, *EMF1-RNAi-3*, respectively. The seeds of T_3 homozygous plants were collected and propagated for further analysis.

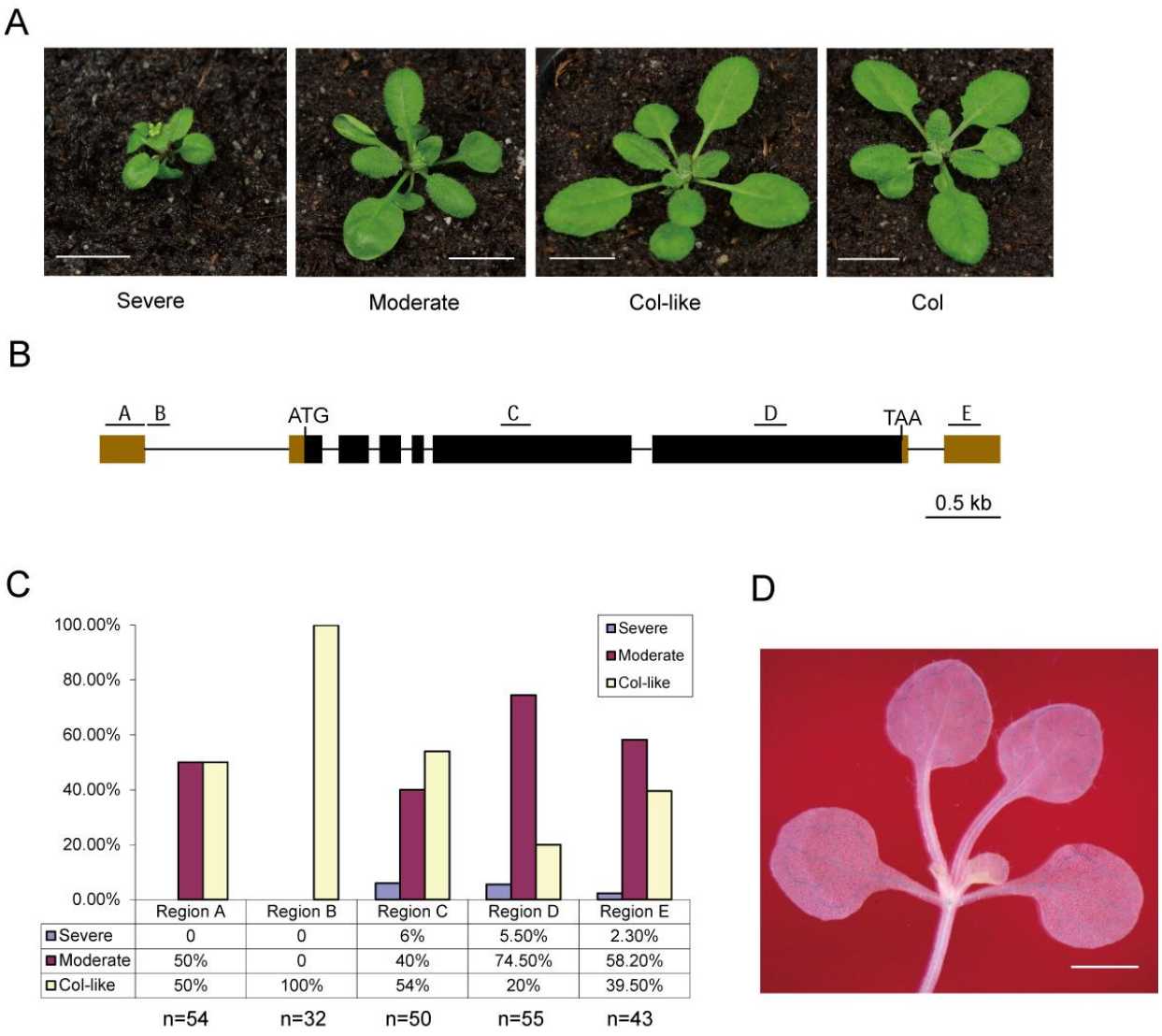


Figure 12. Specific knockdown of *EMF1* expression in the vasculature tissues causes early flowering. (A) Three representative phenotypes of *EMF1-RNAi* transgenic lines. The phenotypes were grouped into three categories: severe, moderate and Col-like as indicated below the plant photos. Bar = 1 cm. (B) *EMF1* structure. Lines above the gene body represent the five *EMF1-RNAi* target regions. Black boxes represent exons, deep yellow boxes represent UTR, and lines in gene body represent introns. (C) Proportions of different *EMF1-RNAi* phenotypes. n indicates the number of T_1 transgenic plants scored. (D) GUS staining of *pSUC2:GUS* 10-day-old seedling. Bar = 2 mm.

All three reference lines showed a curly leaf phenotype and photoperiod-independent early flowering (Figure 13A, C and table 4) except a small difference: *EMF1-RNAi-1* transgenic plants flowered the quickest and made 11.8 leaves on average in LDs (long days) and 13.0 leaves in SDs (short days) when they bolted, *EMF1-RNAi-2* made 12.1 leaves in LDs and 14.9 in SDs while *EMF1-RNAi-3* 12.0 in LDs and 13.7 in SDs. They made almost the same number of leaves at the time of flowering, while the WT plants made 20.0 leaves on average in LDs and 83.9 leaves in SDs. These data indicate that EMF1 is involved in the photoperiod pathway for flowering-time regulation.

The RNAi effects of *EMF1* dsRNAi were further investigated. It was found that the level of EMF1 proteins was reduced in the RNAi lines. The *pEMF1:EMF:FLAG* in *emf1* T3 line was crossed with three *EMF1* RNAi lines and the F_1 seedlings were used for western bloting. The level of EMF1:FLAG proteins was decreased when natively expressed EMF1:FLAG (*pEMF1:EMF1:FLAG*) was introduced into the three reference lines as revealed by western-blotting analysis (Figure 13B).

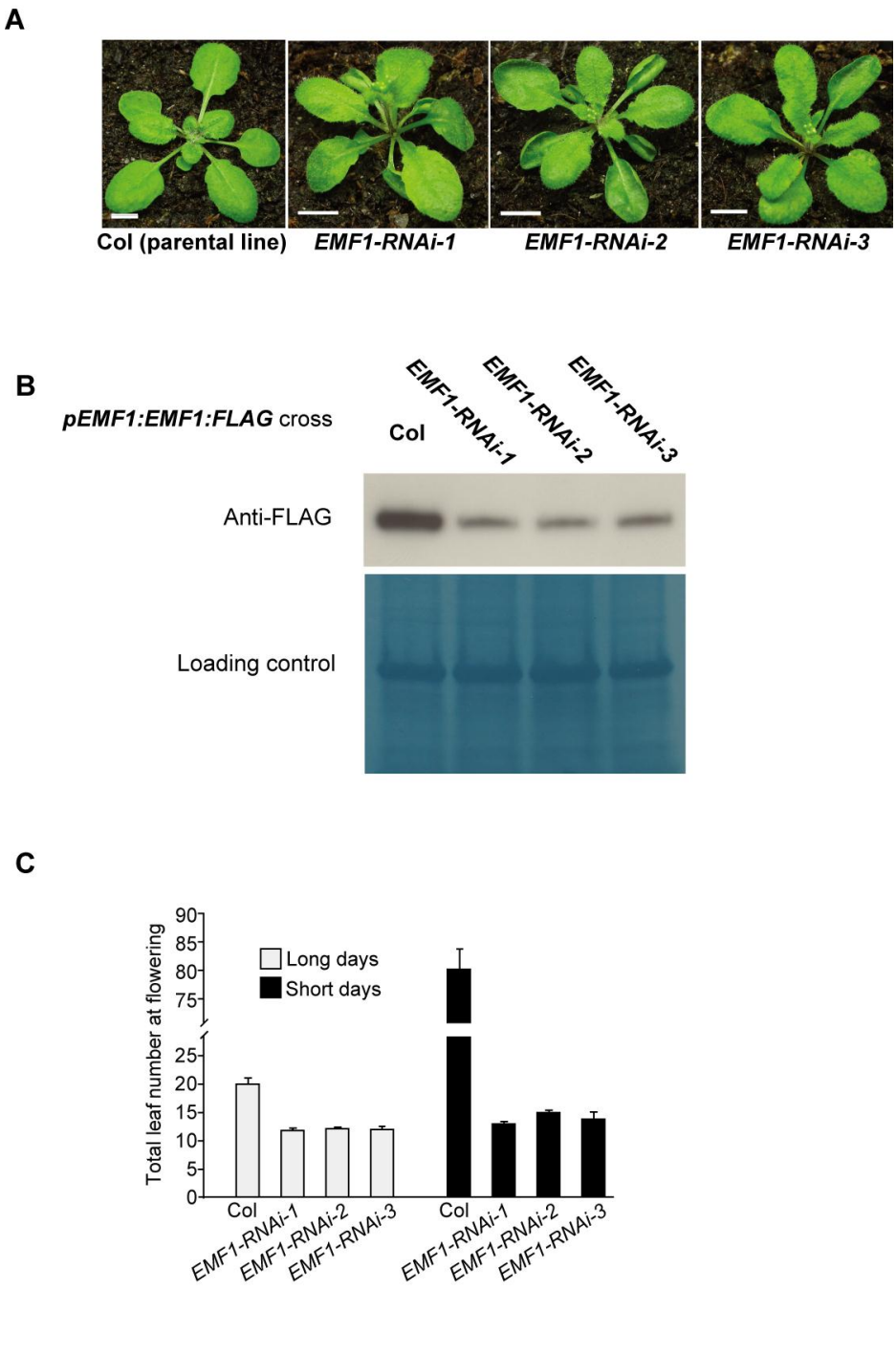


Figure 13. *EMF1-RNAi* in the vascular tissues results in photoperiod independent early flowering. (A) Phenotypes of three reference *EMF1-RNAi* lines. Bar = 1 cm. (B) EMF1 protein levels are decreased in *EMF1-RNAi* lines. Hemizygous of *pEMF:EMF1:FLAG* with Col (as a control) and three *EMF1-RNAi* lines were used for protein analysis. Total proteins were extracted from 10-day-old seedlings and analyzed by western blot with anti-FLAG. The membrane stained with Amido Black served as loading control. (C) Flowering time of *EMF1-RNAi* lines in long days (LD) and short days (SD). Error bars indicate SD.

Table 4. Total leaf number at flowering of Col and *EMF1-RNAi* lines in long days (LDs) and short days (SDs).

Growth condition	Genotypes	Total leaf number (Mean ± SD)	Sample size
LD	Col	19.9 ± 1.1	24
	EMF1-RNAi-1	11.8 ± 0.4	24
	EMF1-RNAi-2	12.1 ± 0.3	24
	EMF1-RNAi-3	12.0 ± 0.6	24
SD	Col	83.9 ± 3.8	15
	EMF1-RNAi-1	13.0 ± 0.4	21
	EMF1-RNAi-2	14.9 ± 0.6	21
	EMF1-RNAi-3	13.7 ± 1.3	21

To identify the targets of EMF1, the cross of *EMF1-RNAi* lines with the mutants of photoperiod pathway including *co* and *ft* were made. The double mutants of *EMF1-RNAi-1* with *co* or *gi* formed almost the same number leaves as *EMF1-RNAi-1* single mutants when flowered, although both the *co* and *gi* single mutants flowered very late in LDs (Figure 14A). However, unlike *co* or *gi*, the double mutants of *EMF1-RNAi-3* with *ft* flowered very late (Figure 14B), even slightly later than *ft* single mutants, which may be resulted from the de-repression of other flower genes such as *FLC* (Kim et al., 2010). These genetic data suggest that *FT* is the target of *EMF1*.

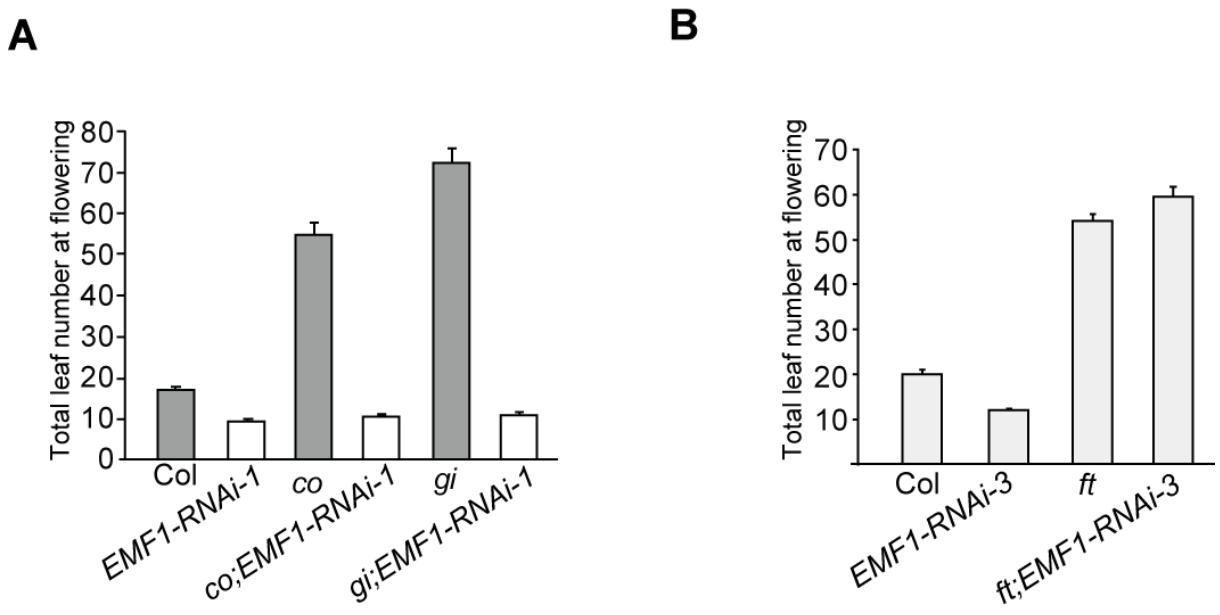


Figure 14 Flowering time of *EMF1-RNAi* in the photoperiod-pathway mutant background. (A) Flowering time of *EMF1-RNAi-1* cross with *co* and *gi*. Genotypes of the mutants were stated below the diagram. Error bars indicate SD. (B) Flowering time of *EMF1-RNAi-3 ft*.

At the molecular level, the expression levels of *FT* were also measured. As expected, *FT* was highly derepressed in *EMF1-RNAi* transgenic lines in both LDs and SDs (Figure 15A). The extent of *FT* derepression in the *EMF1-RNAi* lines was in accordance with their flowering-time phenotype: *EMF1-RNAi-1* flowered the earliest with the highest *FT* expression, while *EMF1-RNAi-2* got flowered the latest with the lowest *FT* expression (Figure 15A). The *GUS* reporter system (*GUS*: beta-glucuronidase) was also introduced for *FT* analysis. I introduced the *pFT9:GUS* (about 9kb *FT* promoter)(Takada and Goto, 2003) into both WT and *EMF1-RNAi-1* backgrounds by crossing *pFT9:GUS* line with both WT and *EMF1-RNAi-1*. The F_1 seedlings were used for GUS staining analysis. As expected, in the *EMF1-RNAi-1* background, the GUS signals were much stronger than WT in the minor veins of leaf vasculature where native *FT* is expressed. Besides, , the GUS signals and were widely spread in the *EMF1-RNAi-1* background indicating the ectopic expression of *FT* resulting from *EMF1* knockdown in the whole vascular tissues (Figure 15C).

We have found that the SHL family proteins appear to function as part of the PRC1-like complex. *SHL* and *EBS* redundantly regulate plant development, and early flowering phenotype was displayed by the *shl ebs* double mutants. Their effects on *FT* regulation were further examined by measuring the *FT* transcript levels in *shl ebs* mutants. *FT* was highly expressed in the *shl ebs* double mutants when compared to WT or single mutants, which indicates that *FT* is also repressed by SHL family proteins (Figure 15B), consistent with that EBS and SHL function as part of the EMF1 complex to repress *FT* expression

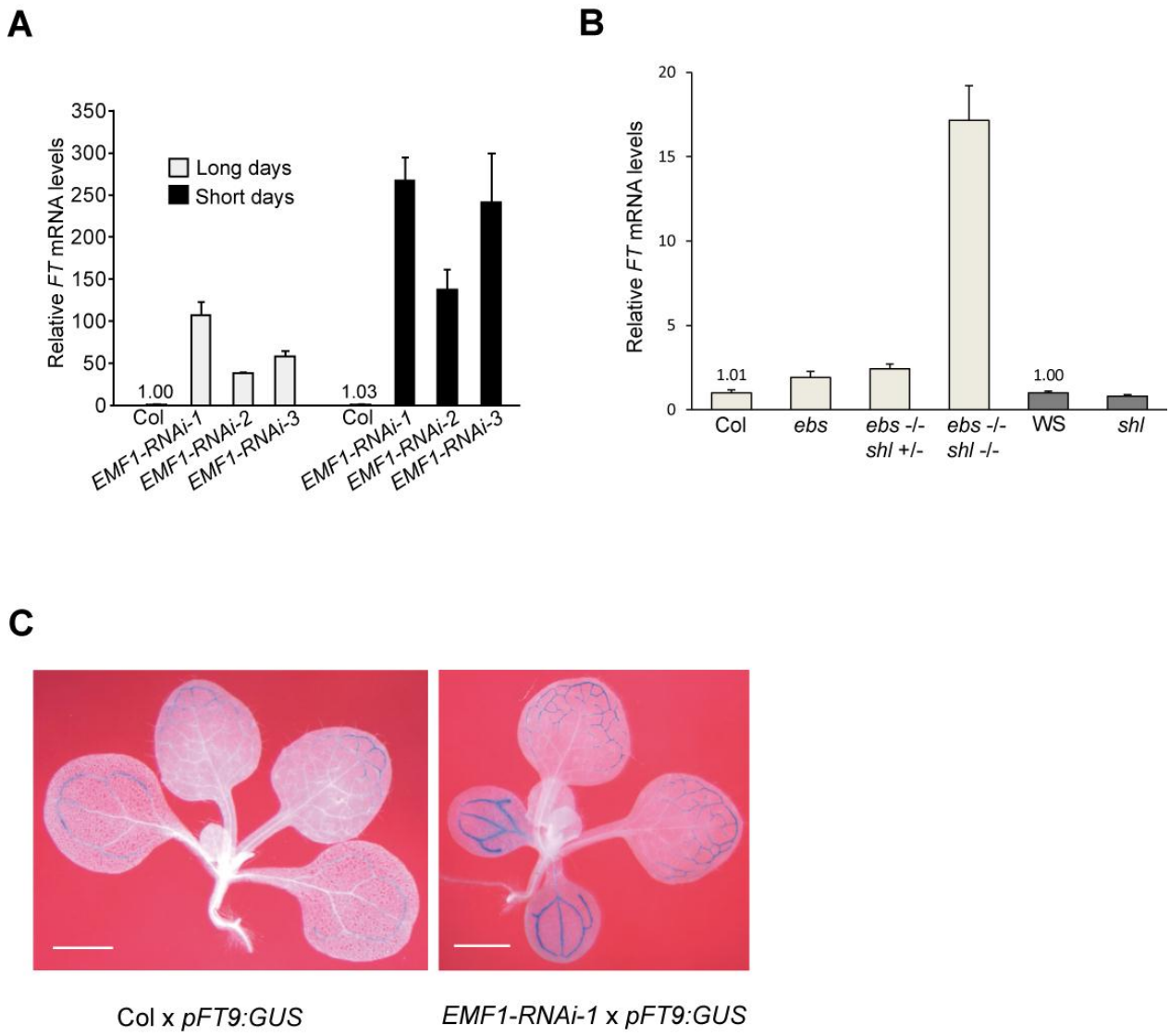


Figure 15 *FT* expression is repressed by EMF1, SHL and EBS. (A) Relative *FT* mRNA levels in seedlings of Col and *EMF-RNAi* lines in both long days (LD) and short days (SD) determined by real-time quantitative PCR. Relative expression to parental Col was presented. Bars indicate SD. (B) Relative *FT* mRNA levels in seedlings of Col, Ws, *ebs* and *shl* mutants. Genotypes of the mutants were stated below the diagram. The *ebs* single mutants and *ebs* hom *shl* het and *ebs shl* double mutants were compared to Col, and *shl* single mutants were compared to Ws. (C) GUS staining of 10-day-old seedlings of hemizygous of *pFT9:GUS* (about 9kb *FT* promoter) with Col and *EMF1-RNAi-1*. Bar = 1 mm.

3.6 *FT* is a direct target of the EMF1-PRC1 complex

The promoter of a gene consists of *cis*-elements for gene regulation such as the binding sites of transcription factors. Previous study implies that a 5.7-kb sequence upstream of the *FT* translation start site is sufficient for *FT* response to day-length change (Adrian et al., 2010). Besides the promoter region, it has also been found that the gene body of *FT* is also involved in *FT* regulation. CLF, the PRC2 component, binds to the introns of *FT* (Jiang et al., 2008). LHP1 is also reported to be directly associated with the introns of *FT* (Adrian et al., 2010). We speculate that the gene body of *FT* has negative effects on *FT* expression. To investigate this hypothesis, I constructed two GUS plasmids which include essential *cis*-elements for *FT* expression (Adrian et al., 2010): one with *GUS* gene directly following 6.9-kb *FT* promoter upstream of the *FT* translation start site (*pFT7:GUS*), the other one with *GUS* following the whole genomic *FT* including 6.9-kb *FT* promoter and the *FT* coding region (*pFT7:FT:GUS*). Both GUS constructions were introduced into Col plants by the *Agrobacterium*-mediated method. The GUS staining of T_1 transgenic seedlings were compared. As expected, the latter GUS construction (*pFT7:FT:GUS*) gave much weaker GUS signals than the *pFT7:GUS*, suggesting that the coding region of *FT* contains negative *cis*-elements for *FT* regulation (Figure 16A).

The GUS staining of p*EMF1:EMF1:GUS* including a 3.8-kb sequence upstream of the translation start site and the *EMF1* coding region, showed that the spatial expression region of *EMF1* overlaps with that of *FT* (Figure 16B). This result implies the possibility of *FT* direct regulation by *EMF1*. The genetic data also indicated that *FT* is the target of EMF1-PRC1 complex (Figure 15A and B). We wonder whether EMF1 could directly regulate *FT* expression. Chromatin immunoprecipitation (ChIP) assay was performed using *pEMF1:EMF1:FLAG* in

emf1-1 transgenic T_3 homozygous plants. 10-day-old seedlings at 8 h after light on in LDs were collected for ChIP analysis (see Materials and Methods). According to the previous studies (Adrian et al., 2010), we quantified the *FT* promoter region and the first intron region by real-time PCR. Results showed that EMF1 is directly associated with both the *FT* promoter and the first intron (Figure 16 C and D).

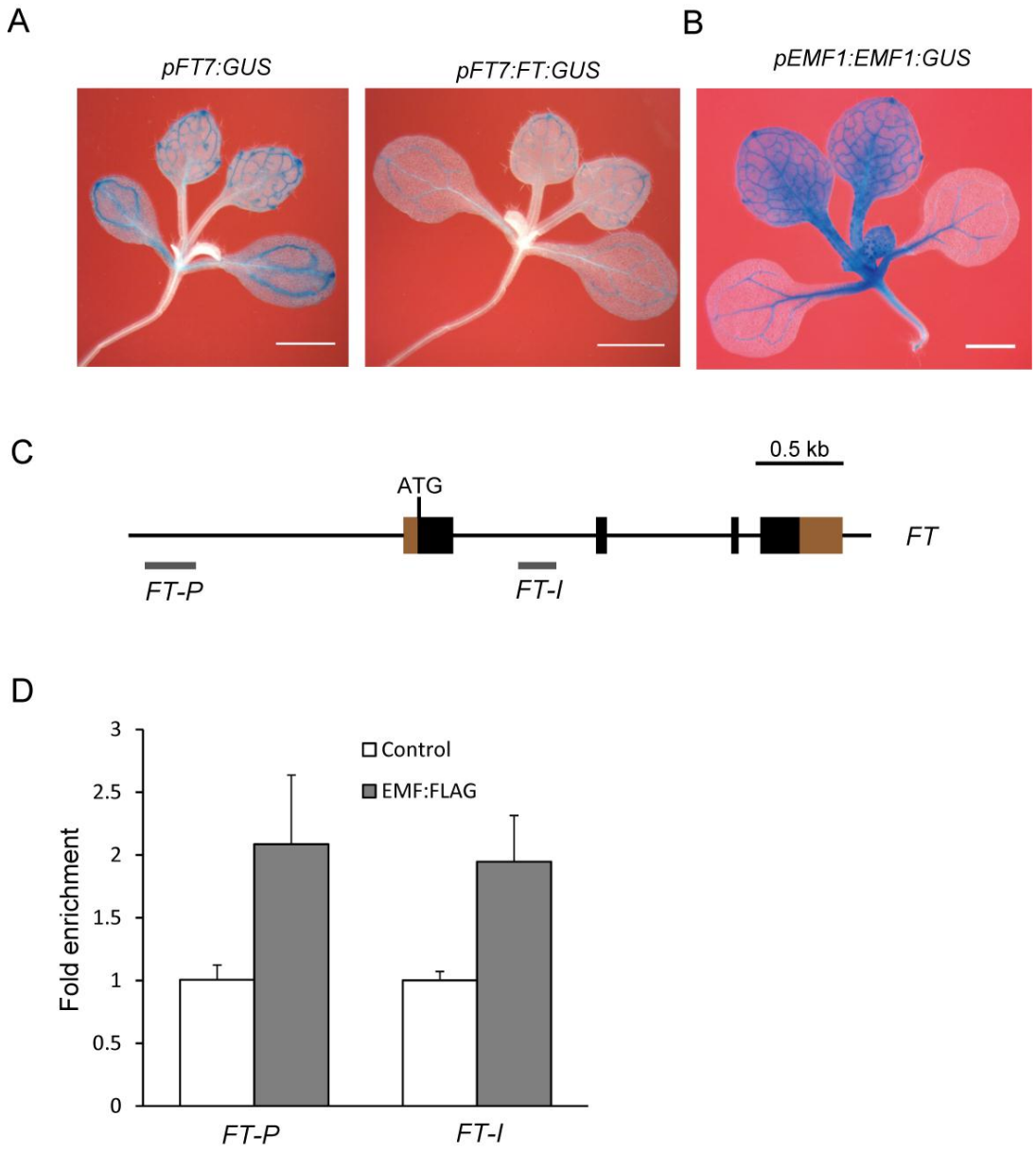


Figure 16 *FT* is a direct target of EMF1. (A) GUS staining of 10-day-old seedlings of *pFT7:GUS* and *pFT7:FT:GUS* (about 7kb *FT* promoter). Bar = 2 mm. (B) GUS staining of 10-day-old seedling of *pEMF:EMF:GUS*. Bar = 1 mm. (C) Diagram of *FT* structure and the regions examined in ChIP. Black boxes represent exons, deep yellow boxes represent UTR, and lines in gene body represent introns. (D) Binding of EMF1 to *FT* chromatin. DNA fragments of *FT-P* and *FT-I*, immunoprecipitated with anti-FLAG from seedlings of a *pEMF1:EMF1:FLAG emf1-1* transgenic line and Col (served as control), were quantified by real-time PCR and subsequently normalized to *TUBLIN 2 (TUB2)*. The fold enrichments of the *pEMF1:EMF1:FLAG emf1-1* line over the control (Col) are shown. Bars indicate SD. Quantified by Dr. Gu Xiaofeng.

Like *CO* and *FT*, about 30% of the genes in *Arabidopsis* are regulated by the circadian clock with a periodicity of 24 hours (Fornara et al., 2010). In long days (LDs), the level of *FT* mRNA increases during the day, reaches to the peak at the end of the day and decreases in the following night (Suarez-Lopez et al., 2001)(Figure 17A). To investigate the effects of PRC1 complex on diurnal expression of *FT*, we checked *FT* expression in *EMF1-RNAi-1* and Col over a 24 h time course. As showed by the results, in WT, *FT* mRNA was repressed and kept at low expression level till 8 hours after light on. After that, it got increase in the following 8 hours and reached the peak at 16 h when light off, then gradually decreased in the dark(Figure 17A). In *EMF1-RNAi-1*, for the whole time course, *FT* expression was highly de-repressed comparing to Col. But like in Col, *FT* mRNA in *EMF1-RNAi-1* increased during day time, reached to the highest expression level at the end of the day and decreased in the following dark. Different from Col, *FT* mRNA in *EMF1-RNAi-1* got a rapid increase in the first 8 hours after light on indicating that EMF1 has a much more important role in *FT* repression for this period (Figure 17A). The different expression of *FT* between Col and *EMF1-RNAi-1* suggest that EMF1 be involved in *FT* diurnal expression regulation. To confirm that the EMF1 protein levels were reduced over the 24 h time course, I examined two time points (8 h and 16 h after light on). Indeed, *EMF1* was stably repressed over the 24 h period (Figure 17B). Thus, the different levels of *FT* expression at different time points are not due to the variation in *EMF1* suppression.

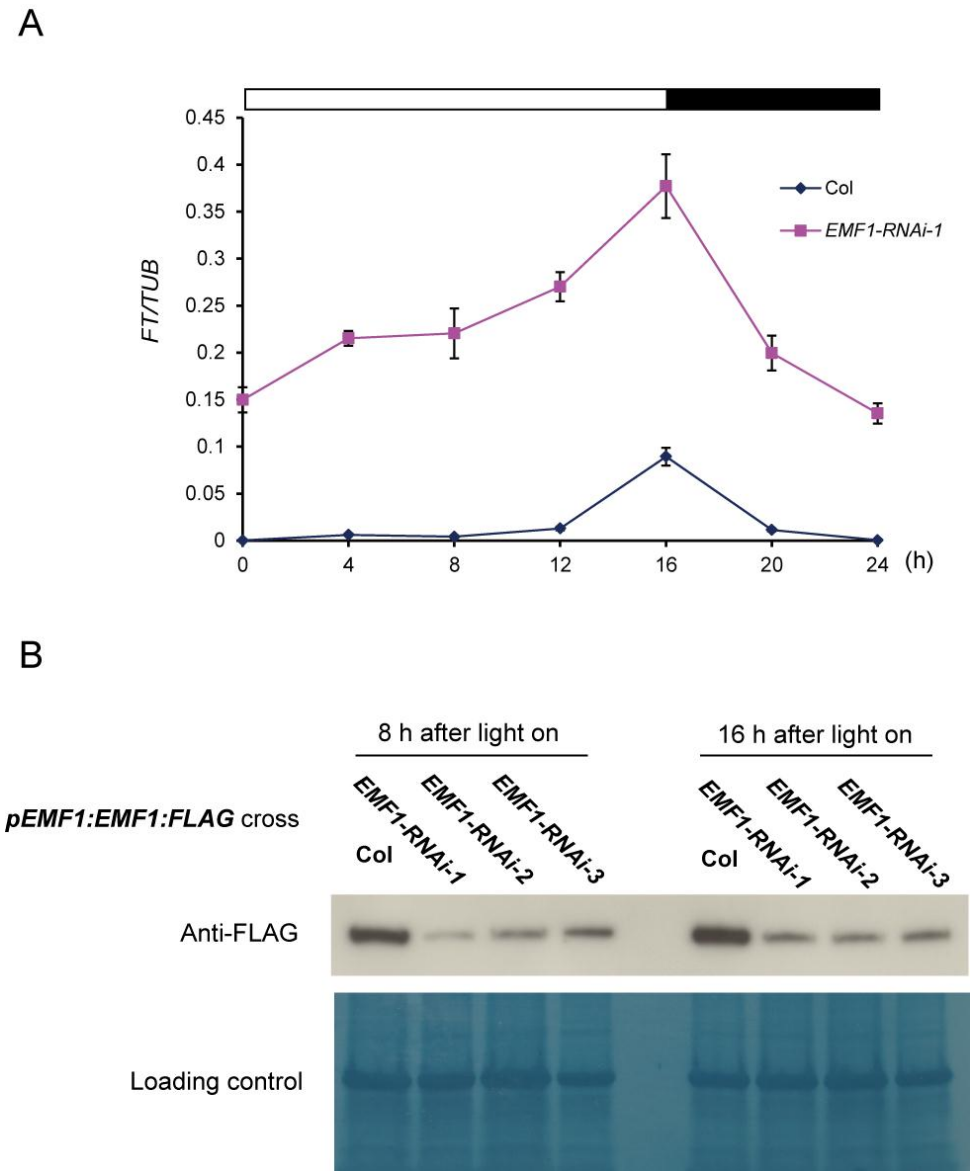


Figure 17. EMF1-PRC1 is involved in the regulation of *FT* diurnal expression. (A) *FT* expression over a 24 h time course in *EMF1-RNAi-1* and Col in LDs. Samples were collected every 4 hours. *FT* expression levels was quantified by real-time PCR and subsequently normalized to *TUBLIN 2 (TUB)*. The values shown are means ±SD. Time "0" represents lights on. White bar above the diagram indicates light period, Black bar indicates dark period. (B) EMF1 protein levels in Col and *EMF1-RNAi* lines at two time points. Seedlings of hemizygous of *pEMF:EMF1:FLAG* with Col (as a control) and three *EMF1-RNAi* lines at 8 h and 16 h after light on were collected for protein analysis. Total proteins were extracted and analyzed by western blot with anti-FLAG. The membrane stained with Amido Black served as loading control.

3.7 EMF1 directly associates with the H3K4 demethylase PKDM7B

In *Drosophila*, the PRC1 complex component PSC is associated with a histone lysine demethylase KDM2 that catalyzes H3K36 demethylation and stimulates H2A ubiquitination by RING and PSC (Lagarou, 2008). PSC, RING and KDM2 form the dRAF complex (BCOR complex in human) to mediate gene repression. The findings from this study suggest that EMF1 plays a PRC1-like role in plants. To test whether there are histone demethylases involved in EMF1-PRC1 in *Arabidopsis*, the interactions between EMF1 and histone demethylases in plants were checked by yeast two-hybrid assay. It was found that PKDM7B, a H3K4 demethylase, can interact with EMF1.

PKDM7B (also known as Jmj14, AtJmj4) is a plant-specific Jumonj C (JmjC) domain-containing protein that demethylates H3K4me3 in *FT* and the *FT* homologue *TWIN SISTER OF FT (TSF)* to repress their expression, and thus, inhibit floral transition (Lu et al., 2010; Yang et al., 2010). In the yeast two-hybrid analysis, I fused PKDM7B with GAL4 AD domain and examined the interactions with full length EMF1-BD or domains of EMF1. The results showed that PKDM7B had weak interactions with the full length of EMF1 on the SD-3 media (SD-Trp, -Leu, -His). However, PKDM7B showed strong interactions with the EMF1-M2 (618aa to 886aa) region (Figure 18A). Furthermore, PKDM7B also showed interaction with M-EMF1 (618aa to 745aa) region on the SD-4 media (SD-Trp, -Leu, -Ade, -His) (Figure 18B). And the interaction was also verified in the pulldown assay, in which *E.coli* expressed GST-fused M-EMF1 could pulldown PKDM7B:HA (Yang et al., 2010) protein from *Arabidopsis* seedlings (Figure 18C). The *in vivo* interaction between PKDM7B and EMF1 was also analyzed by CoIP. 10-day-old F_1 seedlings of *p35S:PKDM7B:HA* crossing with *pEMF1:EMF:FLAG* in *emf1* T3 transgenic line

were used for analysis. The result showed the *in vivo* interaction of PKDM7B with EMF1 since PKDM7B:HA was pulled down by EMF1:FLAG when *pEMF1:EMF1:FLAG* and *p35S:PKDM7B:HA* were coexpressed in plants (Figure 18D).

In *Arabidopsis*, there are more than 20 JmjC-domain containing proteins which act as histone demethylase. We wonder whether other proteins are also involved in the association with EMF1-PRC1 complex. Then the interactions between PKDM7B homologues and EMF1, EMF1 domains were checked by yeast two-hybrid assay. Two close homologues of PKDM7B: JMJ15 (At2g34880) and JMJ18 (At1g30810) which are also H3K4me3 demethylase were fused with GAL4 activation domain (AD). Although they did not show interactions with full length EMF1-BD on SD-3 media (SD-Leu, Trp and His) (data not show), they showed interaction with M-EMF1 (618aa to 745aa) region on the SD-4 media (SD-Trp, -Leu, -Ade, -His) and the interaction of JMJ15 with M-EMF1 was much weaker than that of PKDM7B and JMJ18 (Figure 18B). These data suggested that the association with EMF1 complex may not be restricted to PKDM7B.

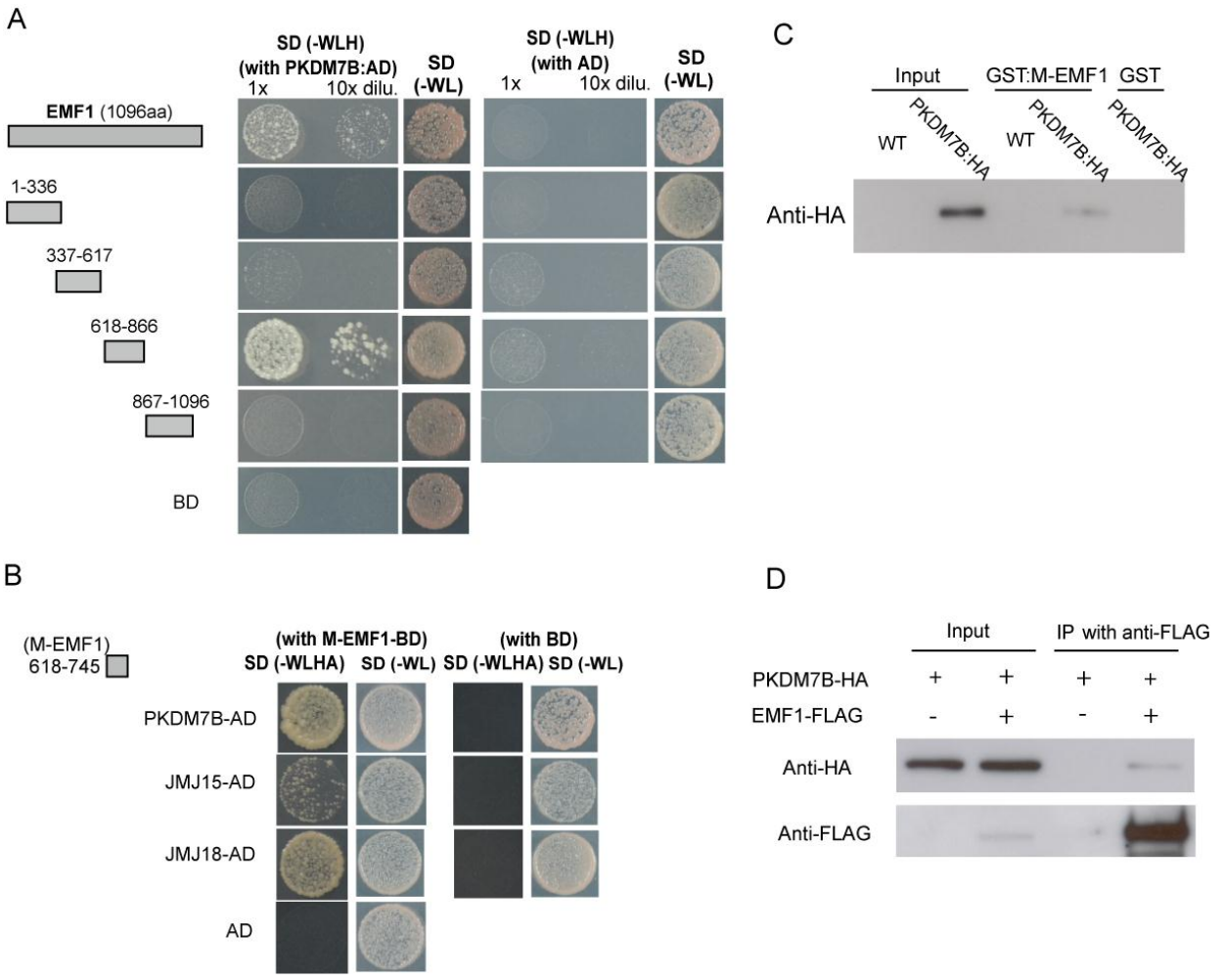


Figure 18. PKDM7B directly associates with EMF1. (A) Yeast interactions between PKDM7B with full length and domains of EMF1. PKDM7B was fused with the GAL4 AD domain and the full length EMF1 and domains were fused to the BD domain. Numbers indicate the amino acid residues of EMF1. Yeast cells harboring these fusion proteins were grown on SD-3 media (SD-Trp, -Leu, -Ade) for selection and SD-2 media (SD-Trp, -Leu) for control. (B) Interactions between M-EMF1 with PKDM7B homologous in yeast. Yeast cells were grown on SD-4 media (SD-Trp, -Leu, -Ade, -His) for selection and SD-2 media for control. (C) Pulldown assay analysis of the interaction between M-EMF1 and PKDM7B. *Ecoli* expressed GST fused M-EMF1 were mixed with protein extracts from plants Col (as control) and *PKDM7B:HA*. After precipitated with GST-beads, the precipitates were analyzed by western blot with anti-HA. (D) Co-immunoprecipitation analysis of EMF1 with PKDM7B. Total proteins were extracted from *PKDM7B:HA* (as a negative control) and hemizygous of *PKDM7B:HA* with *EMF1:FLAG* seedlings. Protein extracts were immunoprecipitated with anti-FLAG agarose, and the precipitates were analyzed by western blot with anti-HA and anti-FLAG.

CHAPTER 4

Discussions & Conclusions

PcG proteins are well known for their function of silencing gene expression. The components of PRC2 in *Arabidopsis* have been well characterized because of the high conservation between animals and plants. However, unlike PRC2, the PRC1 components are poorly conserved. It was speculated that there might be plant-specific proteins that function like PRC1. LHP1 was proposed to act as part of PRC1 in *Arabidopsis* for its role in H3K27me3 recognizing and binding. In this study, we demonstrated that EMF1 interacts with LHP1 both *in vitro* and *in vivo,* suggesting that EMF1 is part of the LHP1-PRC1 complex.

4.1 EMF1 forms a complex with LHP1

EMF1 is a plant-specific protein and conserved in higher plants. We could only find the homologues of EMF1 in angiosperms when we searched the online protein databases. It seems that the higher plants evolved a new mechanism for precise gene regulation. Although we cannot rule out the possible existence of EMF1 homologues in other plants, there might be other components playing the role of EMF1 for gene repression.

Though little was known about the structure of EMF1 protein, we identified the interaction regions of EMF1. For the interaction with LHP1, the regions have been refined to the N-terminal (N_{159}-EMF1) and M domain (618 to 745 aa). Furthermore, the first 49 amino acids of the N-terminal region also showed strong interaction with LHP1 in the yeast experiments. On the other hand, the chromo-shadow domain located at the C-terminal of LHP1 interacts with the two domains of EMF1. Together with the interactions between M-EMF1 and PKDM7B, these data

suggest that the N-terminal and M-EMF1 region mediate the interactions between EMF1 and the LHP1 chromo-shadow domain and other proteins.

The genetic interactions reveal non-additive effects between *emf1* and *lhp1*. The *emf1 lhp1* double mutants behaved identically to *emf1* single mutants and overexpression of *EMF1* could partially rescue *lhp1* phenotype, indicating that they function in the same pathway. Considering the molecular interaction of EMF1 with LHP1, these data suggest that EMF1 and LHP1 are present in the same complex.

4.2 SHL/EBS are part of the EMF1-LHP1 PRC1 complex

Like EMF1, our data have revealed that there are additional components in *Arabidopsis* PRC1 complex. The plant-specific SHL family proteins SHL, EBS and EBSL directly interact with EMF1. It seems that they interact with EMF1 through their BAH domains. Besides, SHL also showed physical interactions with AtBMI1a, a putative component of PRC1. Genetic analysis reveals that SHL and EBS act redundantly for plant development regulation. Both *shl* and *ebs* single mutants exhibited weak phenotypes while the double mutants of *ebs shl* show severe phenotypes resembling to *lhp1*. We could not determine whether *EBSL* also acts redundantly with EBS and SHL due to the lack of *ebsl* mutants. But the gene expression analysis suggests that EBSL is unlikely to overlap with *EBS* and *SHL*. These molecular and genetic data imply that SHL family proteins EBS and SHL function as part of the EMF1-LHP1-PRC1 complex.

4.3 EMF1-PRC1 represses *FT* expression to inhibit floral transition

To investigate the role of EMF1-PRC1 complex on flowering-time control, we specifically knocked down *EMF1* expression in the vasculature tissues. We used RNAi strategy to target the different regions of *EMF1* gene. A majority of the RNAi transgenic plants targeting an EMF1 region, except for the intron, showed moderate phenotypes of early flowering and curly leaf. A small portion exhibited severe phenotypes of much earlier flowering and very small plants, which may due to higher RNAi efficiency or positional effects of T-DNA insertion. The phenotypes of *EMF1-RNAi* transgenic plants and the observation of decrease of EMF1 protein in *EMF1-RNAi* lines suggest our RNAi is effective.

The *EMF1* RNAi transgenic lines exhibited day length independent early flowering, which suggests that EMF1 is involved in the photoperiod pathway for flowering-time control. The genetic interactions between *EMF1-RNAi* lines with the mutants in the photoperiod pathway reveal that *FT* is a target of EMF1. *FT* is highly de-repressed in the *EMF1-RNAi* lines, as revealed by the gene expression analysis and GUS reporter assays. Besides, *FT* transcripts are also increased in the *ebs shl* double mutants and *lhp1* mutants, which indicates that the EMF1-LHP1-SHL/EBS PRC1 complex is involved in floral transition regulation through repressing *FT* expression.

Previous studies indicate the coding region of *FT* may contain *cis*-elements for *FT* regulation (Jiang et al., 2008; Adrian et al., 2010). Introducing the coding region of *FT* into the *pFT:GUS* vector could reduce the GUS signals, suggesting a negative role of *cis*-elements in *FT* coding

region. Consistent with this, we found that EMF1 binds to *FT* chromatin including the promoter and coding region. In addition, LHP1, the EMF1 partner, has been recently shown to bind to *FT* chromatin (Adrian et al., 2010). Taken together, we conclude that the EMF1-LHP1-PRC1 complex directly regulates *FT* expression thus flowering time in *Arabidopsis*.

4.4 PKDM7B is involved in EMF1-mediated *FT* regulation

The active H3K4me3 mark and the repressive H3K27me3 mark co-exist on *FT* chromatin (Jiang et al., 2008). The bivalent of histone H3K4me3 and H3K27me3 may also play critical roles on *FT* regulation (He, 2012). PKDM7B, a H3K4 demethylase, has been reported to regulate flowering by demethylating H3K4me3 at *FT* chromatin (Yang et al., 2010). We detected the physical interactions between PKDM7B and EMF1 both *in vitro* and *in vivo,* suggesting that PKDM7B may also be involved in the EMF1-mediated *FT* regulation. In *Drosophila*, PRC1 components PSC and RING form the dRAF complex with KDM2, an H3K36me2 demethylase. KDM2 stimulates the dRAF activity of ubiquitination. The interaction between EMF1 and PKDM7B may resemble to the association of PSC with KDM2. However, the effects of PKDM7B on EMF1-PRC1 may differ from that of *Drosophila,* and this needs to be further investigated.

Considering all the data obtained from this study, a model on EMF1-PRC1 complex mediated *FT* repression has been proposed. *FT* chromatin is in a bivalent state, simultaneously harboring the active H3K4me3 and the repressive H3K27me3 marks. PRC2 mediates *FT* silencing with H3K27me3 deposition on *FT* chromatin without completely eliminating H3K4me3. LHP1

recognizes and binds to H3K27me3 on *FT* chromatin to anchor PRC1 to the *FT* locus. LHP1 directly interacts with EMF1 to form a PRC1-like complex. EMF1 directly associates with SHL/ EBS, which facilitates AtBMI1a/1b to fit the complex through the interactions with AtBMI1a/1b. AtRING1a/1b are also recruited to the complex to catalyze H2A monoubiquitination. Meanwhile, EMF1 acts as a scaffold protein to recruit PKDM7B for H3K4 demethylation. Decreased H3K4me3 breaks down its balance with H3K27me3. Together, H3K4 demethylation, maintenance of H3K27me3 and H2A monoubquitination form a repressive chromatin environment to silence *FT* expression, which leads to floral inhibition (Figure 19).

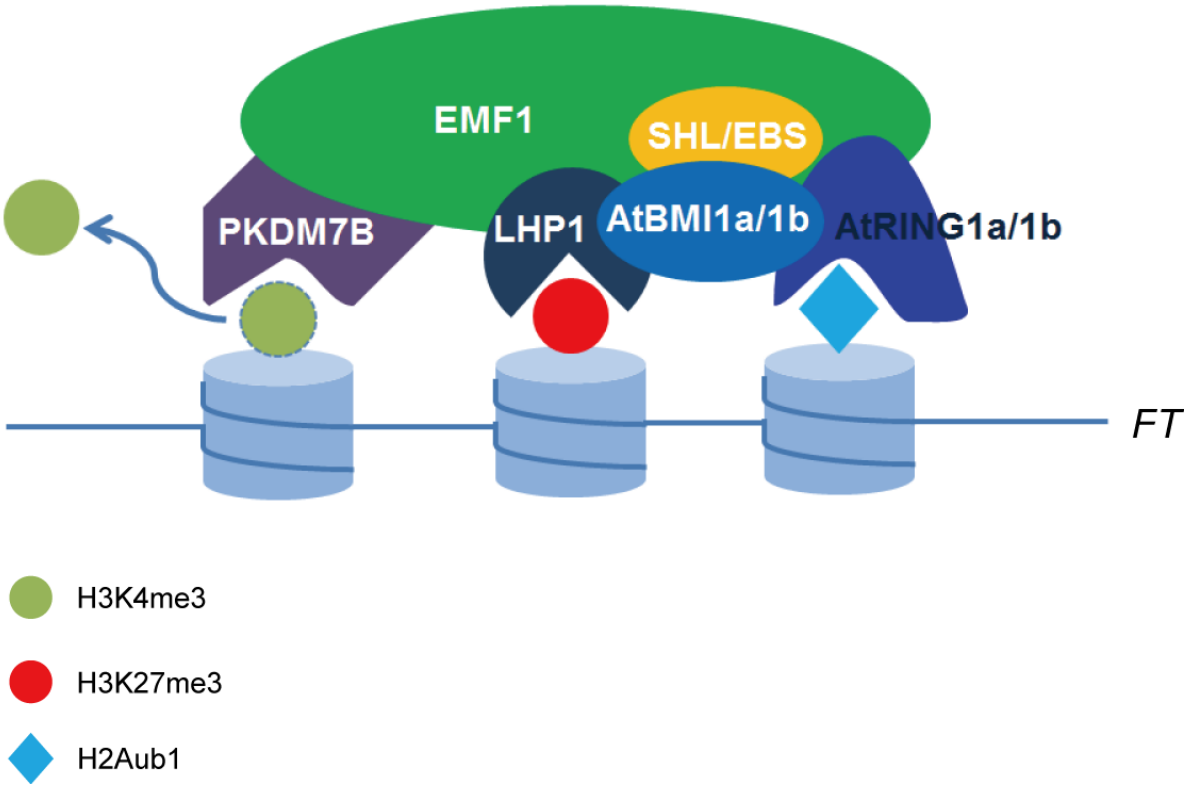


Figure 19. A model for EMF1-PRC1 mediated *FT* silencing. *FT* chromatin is in a bivalent state with both active H3K4me3 and repressive H3K27me3 marks. LHP1 recognizes and binds to H3K27me3 on *FT* chromatin to anchor PRC1 to *FT* locus. LHP1 forms a PRC1-like complex through the interaction with EMF1. EMF1 binds to *FT* chromatin and directly associates with SHL/EBS which facilitates assembling of AtBMI1a/1b & AtRING1a/1b to catalyze H2A monoubquitination on *FT* chromatin. Meanwhile, PKDM7B is recruited by EMF1 to catalyze H3K4 demethylation at the *FT* locus, which leads to the decrease of H3K4me3 levels. Together, H3K4 demethylation, maintenance of H3K27me3 and H2A monoubquitination form a repressive chromatin environment to silence *FT* expression which leads to floral inhibition.

REFERENCES

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T (2005) FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. Science **309:** 1052-1056

Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F (2010) cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in Arabidopsis. Plant Cell **22:** 1425-1440

Amasino R (2010) Seasonal and developmental timing of flowering. Plant J **61:** 1001-1013

Aubert D, Chen L, Moon YH, Martin D, Castle LA, Yang CH, Sung ZR (2001) EMF1, a novel protein involved in the control of shoot architecture and flowering in Arabidopsis. Plant Cell **13:** 1865-1875

Ausin I, Alonso-Blanco C, Jarillo JA, Ruiz-Garcia L, Martinez-Zapater JM (2004) Regulation of flowering time by FVE, a retinoblastoma-associated protein. Nat Genet **36:** 162-166

Balasubramanian S, Weigel D (2006) Temperature Induced Flowering in Arabidopsis thaliana. Plant Signal Behav **1:** 227-228

Baurle I, Dean C (2006) The timing of developmental transitions in plants. Cell **125:** 655-664

Baurle I, Dean C (2008) Differential interactions of the autonomous pathway RRM proteins and chromatin regulators in the silencing of Arabidopsis targets. PLoS One **3:** e2733

Beh LY, Colwell LJ, Francis NJ (2012) A core subunit of Polycomb repressive complex 1 is broadly conserved in function but not primary sequence. Proc Natl Acad Sci U S A **109:** E1063-1071

Bent A (2006) Arabidopsis thaliana floral dip transformation method. Methods Mol Biol **343:** 87-103

Bernstein BE (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell **125:** 315-326

Bohlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O (2006) CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. Science **312:** 1040-1043

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature **441:** 349-353

Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K (2006) Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev. **20:** 1123-1136

Bratzel F, Lopez-Torrejon G, Koch M, Del Pozo JC, Calonje M (2010) Keeping cell identity in Arabidopsis requires PRC1 RING-finger homologs that catalyze H2A monoubiquitination. Curr Biol **20:** 1853-1859

Breiling A, Turner BM, Bianchi ME, Orlando V (2001) General transcription factors bind promoters repressed by Polycomb group proteins. Nature **412:** 651-655

Brown JL, Fritsch C, Mueller J, Kassis JA (2003) The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. Development **130:** 285-294

Brown JL, Mucci D, Whiteley M, Dirksen ML, Kassis JA (1998) The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. Mol. Cell **1:** 1057-1064

Calonje M, Sanchez R, Chen L, Sung ZR (2008) EMBRYONIC FLOWER1 participates in polycomb group-mediated AG gene silencing in Arabidopsis. Plant Cell **20:** 277-291

Cao R, Tsukada Y, Zhang Y (2005) Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol. Cell **20:** 845-854

Cao R, Zhang Y (2004) SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol. Cell **15:** 57-67

Carles CC, Fletcher JC (2003) Shoot apical meristem maintenance: the art of a dynamic balance. Trends Plant Sci **8:** 394-401

Chaudhury AM, Ming L, Miller C, Craig S, Dennis ES, Peacock WJ (1997) Fertilization-independent seed development in Arabidopsis thaliana. Proc Natl Acad Sci U S A **94:** 4223-4228

Choi K, Kim J, Hwang HJ, Kim S, Park C, Kim SY, 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

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, Coupland G (2007) FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. Science **316:** 1030-1033

De Lucia F, Crevillen P, Jones AM, Greb T, Dean C (2008) A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. Proc Natl Acad Sci U S A **105:** 16831-16836

Deal RB, Topp CN, McKinney EC, Meagher RB (2007) Repression of flowering in Arabidopsis requires activation of FLOWERING LOCUS C expression by the histone variant H2A.Z. Plant Cell **19:** 74-83

Eskeland R (2010) Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. Mol. Cell **38:** 452-464

Eskeland R, Eberharter A, Imhof A (2007) HP1 binding to chromatin methylated at H3K9 is enhanced by auxiliary factors. Mol Cell Biol **27:** 453-465

Farrona S, Thorpe FL, Engelhorn J, Adrian J, Dong X, Sarid-Krebs L, Goodrich J, Turck F (2011) Tissue-specific expression of FLOWERING LOCUS T in Arabidopsis is maintained independently of polycomb group protein repression. Plant Cell **23:** 3204-3214

Fischle W (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev. **17:** 1870-1881

Fornara F, de Montaigu A, Coupland G (2010) SnapShot: Control of flowering in Arabidopsis. Cell **141:** 550, 550 e551-552

Francis NJ, Kingston RE, Woodcock CL (2004) Chromatin compaction by a Polycomb group protein complex. Science **306:** 1574-1577

Francis NJ, Saurin AJ, Shao Z, Kingston RE (2001) Reconstitution of a functional core Polycomb repressive complex. Mol. Cell **8:** 545-556

Gaudin V, Libault M, Pouteau S, Juul T, Zhao G, Lefebvre D, Grandjean O (2001) Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis. Development **128:** 4847-4858

Gearhart MD, Corcoran CM, Wamstad JA, Bardwell VJ (2006) Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. Mol. Cell. Biol. **26:** 6880-6889

Gomez-Mena C, Pineiro M, Franco-Zorrilla JM, Salinas J, Coupland G, Martinez-Zapater JM (2001) early bolting in short days: an Arabidopsis mutation that causes early flowering and partially suppresses the floral phenotype of leafy. Plant Cell **13:** 1011-1024

Goodwin GH, Nicolas RH (2001) The BAH domain, polybromo and the RSC chromatin remodelling complex. Gene **268:** 1-7

Gu X, Jiang D, Wang Y, Bachmair A, He Y (2009) Repression of the floral transition via histone H2B monoubiquitination. Plant J **57:** 522-533

Gu X, Jiang D, Yang W, Jacob Y, Michaels SD, He Y (2011) Arabidopsis homologs of retinoblastoma-associated protein 46/48 associate with a histone deacetylase to act redundantly in chromatin silencing. PLoS Genet **7:** e1002366

He Y (2012) Chromatin regulation of flowering. Trends Plant Sci **17:** 556-562

Hennig L, Derkacheva M (2009) Diversity of Polycomb group complexes in plants: same rules, different players? Trends Genet **25:** 414-423

Heo JB, Sung S (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science **331:** 76-79

Holec S, Berger F (2012) Polycomb group complexes mediate developmental transitions in plants. Plant Physiol **158:** 35-43

Hornyik C, Terzi LC, Simpson GG (2010) The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. Dev Cell **18:** 203-213

Imaizumi T, Schultz TF, Harmon FG, Ho LA, Kay SA (2005) FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. Science **309:** 293-297

Irish VF (2010) The flowering of Arabidopsis flower development. Plant J **61:** 1014-1028

Jacobs SA, Khorasanizadeh S (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science **295:** 2080-2083

Jeon J, 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

Jeong JH, Song HR, Ko JH, Jeong YM, Kwon YE, Seol JH, Amasino RM, Noh B, Noh YS (2009) Repression of FLOWERING LOCUS T chromatin by functionally redundant histone H3 lysine 4 demethylases in Arabidopsis. PLoS One **4:** e8033

Jiang D, Gu X, 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

Jiang D, Wang Y, He Y (2008) Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb repressive complex 2 components. PLoS One **3:** e3404

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

Jung J, Mysliwiec MR, Lee Y (2005) Roles of JUMONJI in mouse embryonic development. Dev. Dyn. **232:** 21-32

Katz A, Oliva M, Mosquna A, Hakim O, Ohad N (2004) FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. Plant J **37:** 707-719

Ketel CS (2005) Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. Mol. Cell. Biol. **25:** 6857-6868

Kim DH, Doyle MR, Sung S, Amasino RM (2009) Vernalization: winter and the timing of flowering in plants. Annu Rev Cell Dev Biol **25:** 277-299

Kim H, Kang K, Kim J (2009) AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. Nucleic Acids Res. **37:** 2940-2950

Kim SY, Lee J, Eshed-Williams L, Zilberman D, Sung ZR (2012) EMF1 and PRC2 cooperate to repress key regulators of Arabidopsis development. PLoS Genet **8:** e1002512

Kim SY, Zhu T, Sung ZR (2010) Epigenetic regulation of gene programs by EMF1 and EMF2 in Arabidopsis. Plant Physiol **152:** 516-528

Kinoshita T, Harada JJ, Goldberg RB, Fischer RL (2001) Polycomb repression of flowering during early plant development. Proc Natl Acad Sci U S A **98:** 14156-14161

Klymenko T (2006) A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. Genes Dev. **20:** 1110-1122

Kohler C, Hennig L (2010) Regulation of cell identity by plant Polycomb and trithorax group proteins. Curr Opin Genet Dev **20:** 541-547

Kotake T, Takada S, Nakahigashi K, Ohto M, Goto K (2003) Arabidopsis TERMINAL FLOWER 2 gene encodes a heterochromatin protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes. Plant Cell Physiol **44:** 555-564

Kozma G, Bender W, Sipos L (2008) Replacement of a Drosophila Polycomb response element core, and in situ analysis of its DNA motifs. Mol. Genet. Genomics **279:** 595-603

Ku M (2008) Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet. **4:** e1000242

Kumar SV, Wigge PA (2010) H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. Cell **140:** 136-147

Lafos M, Kroll P, Hohenstatt ML, Thorpe FL, Clarenz O, Schubert D (2011) Dynamic regulation of H3K27 trimethylation during Arabidopsis differentiation. PLoS Genet **7:** e1002040

Lagarou A (2008) dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. Genes Dev. **22:** 2799-2810

Lee JH, Yoo SJ, Park SH, Hwang I, Lee JS, Ahn JH (2007) Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. Genes Dev **21:** 397-402

Lewis EB (1978) A gene complex controlling segmentation in Drosophila. Nature **276:** 565-570

Liu F, Marquardt S, Lister C, Swiezewski S, Dean C (2010) Targeted 3' processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. Science **327:** 94-97

Liu F, Quesada V, Crevillen P, Baurle I, Swiezewski S, 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 LJ, Zhang YC, Li QH, Sang Y, Mao J, Lian HL, Wang L, Yang HQ (2008) COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. Plant Cell **20:** 292-306

Lorch Y, LaPointe JW, Kornberg RD (1987) Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell **49:** 203-210

Lu F, Cui X, Zhang S, Jenuwein T, Cao X (2011) Arabidopsis REF6 is a histone H3 lysine 27 demethylase. Nat Genet **43:** 715-719

Lu F, Cui X, Zhang S, Liu C, Cao X (2010) JMJ14 is an H3K4 demethylase regulating flowering time in Arabidopsis. Cell Res **20:** 387-390

March-Diaz R, Garcia-Dominguez M, Lozano-Juste J, Leon J, Florencio FJ, Reyes JC (2008) Histone H2A.Z and homologues of components of the SWR1 complex are required to control immunity in Arabidopsis. Plant J **53:** 475-487

Margueron R (2008) Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol. Cell **32:** 503-518

Margueron R, Reinberg D (2011) The Polycomb complex PRC2 and its mark in life. Nature **469:** 343-349

Meacham MR, Hopfield JH, Acree SF (1920) The Relative Effect of Phosphate-Acetate and of Phosphate-Phthalate Buffer Mixtures upon the Growth of Endothia parasitica on Malt Extract and Corn Meal Media. J Bacteriol **5:** 305-308

Mihaly J, Mishra RK, Karch F (1998) A conserved sequence motif in Polycomb-response elements. Mol. Cell **1:** 1065-1066

Min J, Zhang Y, Xu RM (2003) Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev. **17:** 1823-1828

Mockler T, Yang H, Yu X, Parikh D, Cheng YC, Dolan S, Lin C (2003) Regulation of photoperiodic flowering by Arabidopsis photoreceptors. Proc Natl Acad Sci U S A **100:** 2140-2145

Moon YH, Chen L, Pan RL, Chang HS, Zhu T, Maffeo DM, Sung ZR (2003) EMF genes maintain vegetative development by repressing the flower program in Arabidopsis. Plant Cell **15:** 681-693

Muller J, Kassis JA (2006) Polycomb response elements and targeting of Polycomb group proteins in Drosophila. Curr. Opin. Genet. Dev. **16:** 476-484

Mussig C, Altmann T (2003) Changes in gene expression in response to altered SHL transcript levels. Plant Mol Biol **53:** 805-820

Mussig C, Kauschmann A, Clouse SD, Altmann T (2000) The Arabidopsis PHD-finger protein SHL is required for proper development and fertility. Mol Gen Genet **264:** 363-370

Mutasa-Gottgens E, Hedden P (2009) Gibberellin as a factor in floral regulatory networks. J Exp Bot **60:** 1979-1989

Nekrasov M (2007) Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. EMBO J. **26:** 4078-4088

Nielsen PR, Nietlispach D, Mott HR, Callaghan J, Bannister A, Kouzarides T, Murzin AG, Murzina NV, Laue ED (2002) Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. Nature **416:** 103-107

Noguchi K, Vassilev A, Ghosh S, Yates JL, DePamphilis ML (2006) The BAH domain facilitates the ability of human Orc1 protein to activate replication origins in vivo. EMBO J **25:** 5372-5382

Pandey RR (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol. Cell **32:** 232-246

Pasini D, Bracken AP, Jensen MR, Lazzerini Denchi E, Helin K (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. **23:** 4061-4071

Pineiro M, Gomez-Mena C, Schaffer R, Martinez-Zapater JM, Coupland G (2003) EARLY BOLTING IN SHORT DAYS is related to chromatin remodeling factors and regulates flowering in Arabidopsis by repressing FT. Plant Cell **15:** 1552-1562

Ringrose L, Paro R (2007) Polycomb/Trithorax response elements and epigenetic memory of cell identity. Development **134:** 223-232

Rinn JL (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell **129:** 1311-1323

Sanchez C (2007) Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor. Mol. Cell Proteomics **6:** 820-834

Sarma K, Margueron R, Ivanov A, Pirrotta V, Reinberg D (2008) Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. Mol. Cell. Biol. **28:** 2718-2731

Saunders A (2003) Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo. Science **301:** 1094-1096

Savla U, Benes J, Zhang J, Jones RS (2008) Recruitment of Drosophila Polycomb-group proteins by Polycomblike, a component of a novel protein complex in larvae. Development **135:** 813-817

Sawa M, Nusinow DA, Kay SA, Imaizumi T (2007) FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. Science **318:** 261-265

Scheuermann JC (2010) Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature **465:** 243-247

Schoeftner S (2006) Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. EMBO J. **25:** 3110-3122

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G (2007) Genome regulation by polycomb and trithorax proteins. Cell **128:** 735-745

Schuettengruber B, Ganapathi M, Leblanc B, Portoso M, Jaschek R, Tolhuis B, van Lohuizen M, Tanay A, Cavalli G (2009) Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. PLoS Biol **7:** e13

Schultz DC, Friedman JR, Rauscher FJ, 3rd (2001) Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. Genes Dev **15:** 428-443

Schwartz YB (2006) Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nature Genet. **38:** 700-705

Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, Pirrotta V (2006) Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat Genet **38:** 700-705

Sheldon CC, Rouse DT, Finnegan EJ, Peacock WJ, Dennis ES (2000) The molecular basis of vernalization: the central role of FLOWERING LOCUS C (FLC). Proc Natl Acad Sci U S A **97:** 3753-3758

Shen X (2008) EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol. Cell **32:** 491-502

Simon JA, Kingston RE (2009) Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol **10:** 697-708

Song JJ, Garlick JD, Kingston RE (2008) Structural basis of histone H4 recognition by p55. Genes Dev. **22:** 1313-1318

Stock JK (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nature Cell Biol. **9:** 1428-1435

Suarez-Lopez P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G (2001) CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature **410:** 1116-1120

Sun B, Xu Y, Ng KH, Ito T (2009) A timing mechanism for stem cell maintenance and differentiation in the Arabidopsis floral meristem. Genes Dev **23:** 1791-1804

Sung S, He Y, Eshoo TW, Tamada Y, Johnson L, Nakahigashi K, Goto K, Jacobsen SE, Amasino RM (2006) Epigenetic maintenance of the vernalized state in Arabidopsis thaliana requires LIKE HETEROCHROMATIN PROTEIN 1. Nat Genet **38:** 706-710

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

Takada S, Goto K (2003) Terminal flower2, an Arabidopsis homolog of heterochromatin protein1, counteracts the activation of flowering locus T by constans in the vascular tissues of leaves to regulate flowering time. Plant Cell **15:** 2856-2865

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K (2007) Hd3a protein is a mobile flowering signal in rice. Science **316:** 1033-1036

Turck F, Fornara F, Coupland G (2008) Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. Annu Rev Plant Biol **59:** 573-594

Turck F, Roudier F, Farrona S, Martin-Magniette ML, Guillaume E, Buisine N, Gagnot S, Martienssen RA, Coupland G, Colot V (2007) Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLoS Genet **3:** e86

Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G (2004) Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science **303:** 1003-1006

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol **3:** RESEARCH0034

Wang H (2004) Role of histone H2A ubiquitination in Polycomb silencing. Nature **431:** 873-878

Wang JW, Czech B, Weigel D (2009) miR156-regulated SPL transcription factors define an endogenous flowering pathway in Arabidopsis thaliana. Cell **138:** 738-749

Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS (2004) Hierarchical recruitment of polycomb group silencing complexes. Mol Cell **14:** 637-646

Whitcomb SJ, Basu A, Allis CD, Bernstein E (2007) Polycomb Group proteins: an evolutionary perspective. Trends Genet **23:** 494-502

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D (2005) Integration of spatial and temporal information during floral induction in Arabidopsis. Science **309:** 1056-1059

Wilson RN, Heckman JW, Somerville CR (1992) Gibberellin Is Required for Flowering in Arabidopsis thaliana under Short Days. Plant Physiol **100:** 403-408

Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA (2006) The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. Proc Natl Acad Sci U S A **103:** 14631-14636

Xu L, Shen WH (2008) Polycomb silencing of KNOX genes confines shoot stem cell niches in Arabidopsis. Curr Biol **18:** 1966-1971

Xu L, Zhao Z, Dong A, Soubigou-Taconnat L, Renou JP, Steinmetz A, Shen WH (2008) Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in Arabidopsis thaliana. Mol Cell Biol **28:** 1348-1360

Yang W, Jiang D, Jiang J, He Y (2010) A plant-specific histone H3 lysine 4 demethylase represses the floral transition in Arabidopsis. Plant J **62:** 663-673

Yoo SK, Chung KS, Kim J, Lee JH, Hong SM, Yoo SJ, Yoo SY, Lee JS, Ahn JH (2005) CONSTANS activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to promote flowering in Arabidopsis. Plant Physiol **139:** 770-778

Yoshida N, Yanai Y, Chen L, Kato Y, Hiratsuka J, Miwa T, Sung ZR, Takahashi S (2001) EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in Arabidopsis. Plant Cell **13:** 2471-2481

Yu CW, Liu X, Luo M, Chen C, Lin X, Tian G, Lu Q, Cui Y, Wu K (2011) HISTONE DEACETYLASE6 interacts with FLOWERING LOCUS D and regulates flowering in Arabidopsis. Plant Physiol **156:** 173-184

Zeevaart JA (2006) Florigen coming of age after 70 years. Plant Cell **18:** 1783-1789

Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science **322:** 750-756

Zheng B, Chen X (2011) Dynamics of histone H3 lysine 27 trimethylation in plant development. Curr Opin Plant Biol **14:** 123-129

Zhou W (2008) Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. Mol. Cell **29:** 69-80