이학석사학위논문

Study on the Role of *EARLY FLOWERING AND CURLY LEAVES (EFC)* in the FIS2-PRC2 Complex of *Arabidopsis thaliana*

애기장대의 FIS2-PRC2 복합체에서 *EARLY FLOWERING AND CURLY LEAVES (EFC)* 유전자의 역할에 관한 연구

2017 년 2 월

서울대학교 대학원

생명과학부

Nomiun Delgerekh

Abstract

Study on the Role of *EARLY FLOWERING AND CURLY LEAVES* (*EFC*) in the FIS2-PRC2 Complex of *Arabidopsis thaliana*

Nomiun Delgerekh

School of Biological Sciences

The Graduate School

Seoul National University

The Polycomb Repressive Complex 2 (PRC2) is known as one of the master epigenetic regulators in both plants and mammals. It represses transcriptional activity of genes that are not required at specific formative stages by trimethylation of lysine 27 on histone H3 (H3K27me3). In a previous study, a gain-of-function mutant, *early flowering and curly leaves-D* (*efc-D*) was isolated from a population of activation tagging mutants. Overexpression of *EFC* caused increased expression levels of PRC2 target genes, such as *FLOWERING LOCUS C* (*FLC*), *FLOWERING LOCUS T* (*FT*) and *AGAMOUS* (*AG*). Among three PRC2 complexes in *Arabidopsis thaliana*, the FERTILIZATION INDEPENDENT SEED2-PRC2 complex plays essential roles in endosperm development by regulating genes in the central cell and

endosperm. Interestingly, *EFC* was specifically expressed in nuclei of the central cell and endosperm and interacted with MSI1, one of the core members of the PRC2. In this study I characterized a loss-of-function mutant *efc-1*. The *efc-1* mutation caused a seed abortion in *Arabidopsis*. *EFC* showed a partial or incomplete penetrance. The expression of *MEDEA* (*MEA*) and *UPWARD CURLY LEAF1* (*UCL1*), genes that are regulated by FIS2-PRC2, were derepressed in the central cell or endosperm of the *efc-1* mutant. I carried out a complementation experiment of *fis2* with *EMBRYONIC FLOWER2* (*EMF2*), a homolog of *FIS2*, in the seed in absence of *EFC*. Presumably due to the partial penetrance of *EFC*, there was no significant complementation observed in the heterozygous *efc-1*; heterozygous *fis2-11* and hemizygous *FIS2::EMF2* triple mutant plants. Therefore, it needs to be checked in the next generation when I obtain a triple homozygous mutant or double homozygous and *fis2-11* heterozygous mutant. The results so far may suggest that *EFC* might play a role in seed development along with the FIS2-PRC2 complex, however not essential.

Keywords: *EMBRYONIC FLOWER AND CURLY LEAF (EFC),* polycomb repressive complex, seed development, endosperm, loss-of-function, *Arabidopsis*

Student Number: 2014-25243

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

PcG	**Polycomb Group**
PRC2	**Polycomb Repressive Complex 2**
PCR	**Polymerase Chain Reaction**
RT-PCR	**Reverse Transcription-Polymerase Chain Reaction**
GUS	**β-glucuronidase**
H3	**Histone 3**
K27	**Lysine 27**
MS	**Murashige-Skoog**
cDNA	**Complementary DNA**
CAPS	**Cleaved Amplified Polymorphic Sequences**

I. Introduction

The eukaryotic gene expression is regulated meticulously and perpetually in order to maintain its multicellular structure. This intricate control mechanism manipulates the expression of gene at various stages of the classic central dogma. In particular, modifications can be done at genomic, transcriptional, RNA processing, translational or at posttranslational levels through distinctive molecular events.

First found in *Drosophila*, chromatin modifiers formed by Polycomb group proteins (PcG) are one of the well-established transcriptional repression systems in plants, animals and certain unicellular organisms (Mozgova et al., 2015). Multimeric complexes built by PcG proteins are called Polycomb Repressor Complex (PRC). PcG complexes PRC1 and PRC2 are studied very well. In *Arabidopsis*, members that form these complexes are from various protein families (Mozgova et al., 2015).

1. Polycomb Group Protein Complex 2

The biochemical activity of PRC2 is trimethylation of lysine 27 on histone H3 (H2K27me3). In Arabidopsis, a vast number of genes (~4,400) are trimethylated at lysine 27 in histone H3 (Zhang et al., 2007). Four main subunits of PRC2 are histone methyltransferase Enhancer of zeste [E(z)], Supressor of zeste 12 [Su(z)12], Extra sex combs (Esc), and the histone binding nucleosome binding factor 55 kDa (Nurf55 or p55). Homologs of Enhancer of zeste [E(z)] are *MEDEA* (*MEA*), *CURLY LEAF* (*CLF*) and *SWINGER* (*SWN*), homologs of Supressor of zeste 12 [Su(z)12] are

FERTILIZATION INDEPENDENT SEED2 (*FIS2*), *EMBRYONIC FLOWER2* (*EMF2*) and *VERNALIZATION2* (*VRN2*), homologs of Nurf55 are *MULTICOPY SUPPRESSOR OF IRA1-5* (*MSI1-5*), and homolog of Extra sex combs (Esc) is *FERTILIZATION INDEPENDENT ENDOSPERM* (*FIE*) (Grossniklaus et al., 1998; Kiyosue et al., 1999; Luo et al., 1999; Ohad et al., 1999; Kohler et al., 2003). There are three PRC2 complexes, FIS2-PRC2, EMF2-PRC2 and VRN2-PRC2 that exist in *Arabidopsis*, playing essential roles in the different developmental processes (Figure 1).

1.1. FIS2-PRC2

The function of the FIS2-PRC2 complex is crucial for the shift from the female gametophyte to sporophyte because it hinders the central cell proliferation in the absence of fertilization (Figueiredo et al., 2015; Xiao and Wagner, 2015). The FIS2-PRC2 complex consists of *MEA*, *FIS2*, *FIE* and *MSI1*. When a mutation occurs in these genes, endosperm development initiates even when there is no fertilization (Chaudhury et al., 1997; Grossniklaus et al., 1998; Guitton et al., 2004). Endosperm nuclei are overproliferated and cell division is delayed in so called *fis* class mutants that cause the early arrest of embryo development due to a defect of nutrient translocation (Chaudhury et al., 1997; Sorensen et al., 2001; Guitton et al., 2004). It is reported that type I MADS-box protein PHERES1 is strictly regulated by FIS2-PRC2 complex, and deregulation of *PHERES1* (*PHE1*) is extensively responsible for seed-abortion phenotype of mea mutant plants (Kohler et al., 2003). The FIS2 polycomb complex represses the expression of another type I MADS-box transcription factor *AGAMOUS-LIKE 62* (*AGL62*) after certain cycles of nuclei division in the endosperm syncytium in order to trigger cellularization (Kang et al., 2008).

1.2. EMF2-PRC2

The EMF2-PRC2 complex plays a significant role in repression of initiation of the precocious reproductive program in Arabidopsis, allowing normal sporophytic growth of the plant (Yoshida et al., 2001). Member proteins of this complex are *SWN/CLF*, *EMF2, FIE* and *MSI1*. *FLOWERING LOCUS T* (*FT*), *AGAMOUS* (*AG*) and *APETALA3* (*AP3*) are the direct targets of EMF2-PRC2 complex. In the loss-of-function *emf* (*emf1* and *emf2*), development of the floral organ initiates along with the germination (Sung et al., 1992).

1.3. VRN2-PRC2

In angiosperms, the flowering process is induced by exposure to the continuous coldness of winter or artificial cold condition. This process is called vernalization and it allows plants to flower effectively. Similar to the EMF2-PRC2 complex, besides Su(z) 12 homolog *VRN2*, *CLF/SWN*, *FIE* and *MSI1* are the main members of this complex. To maintain the normal development of the plant, the VRN2-PRC2 complex represses expression of *FLOWERING LOCUS C* (*FLC*) in Arabidopsis (Gendall et al., 2001). *FLC* is responsible for floral repression. When a mutation occurs in *VRN2, FLC* level is increased after cold treatment where it is supposed to be downregulated (Gendall et al., 2001).

Figure 1

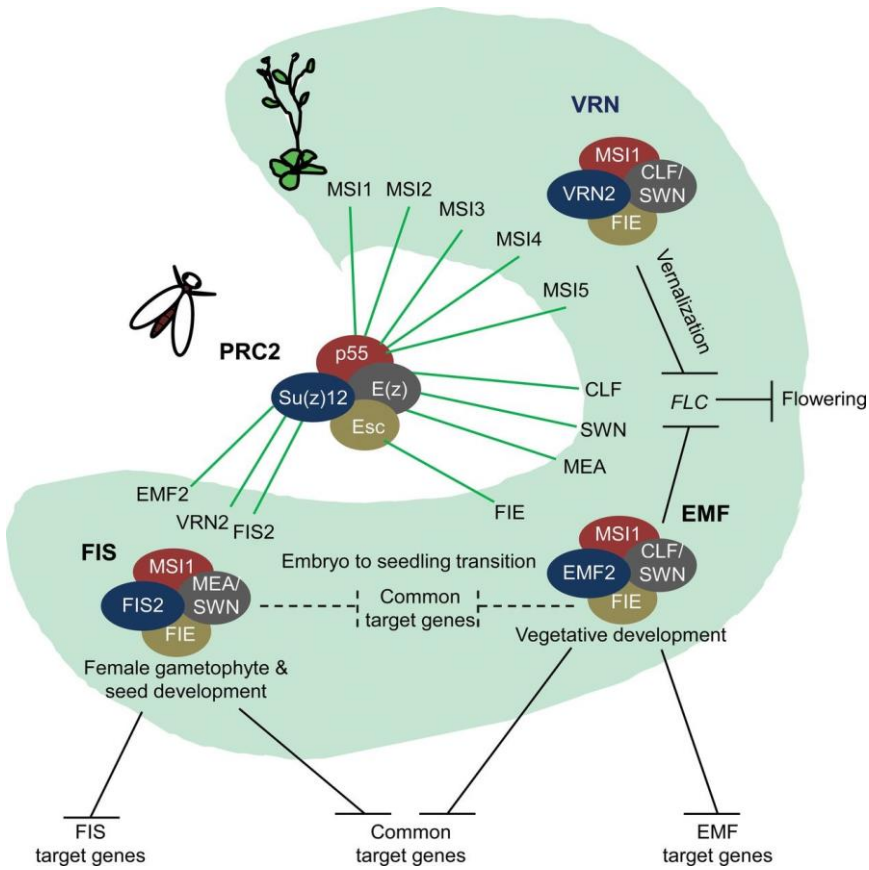


Figure 1. PRC2 complexes in *Arabidopsis thaliana*

Three PRC2 complexes in *Arabidopsis thaliana* are FIS2-PRC2, EMF2-PRC2 and VRN2-PRC2. Note: Figure adapted from Derkacheva, M. and L. Hennig. "Variations On A Theme: Polycomb Group Proteins In Plants". Journal of Experimental Botany 65.10 (2013): 2769-2784. Web.

2. PRC2 interacting proteins

Intriguing studies unveiled the sequential repression process of flowering repressor *FLC* by VRN2-PRC2 complex associated with PLANT HOMEODOMAIN (PHD) finger proteins. Before cold treatment, VRN2-PRC2 complexes are covering the whole *FLC* locus. And during vernalization, PHD-PRC2 complex is localized on the specific locus of FLC and leads to a decline in histone acetylation. This complex contains VRN2-PRC2 core PcG proteins and *VERNALIZATION INSENSITIVE 3* (*VIN3*), *VERNALIZATION 5* (*VRN5*) and *VIN3-like1* (*VEL1*). By the formation of VRN-PHD, acetylation of histone is reduced and H3k27me3 is increased. When prolonged coldness is gone and the plant is again in the warm condition, expression of *VIN3* terminates and association of VRN5 is increased throughout the *FLC* locus for repression of the gene (De Lucia et al., 2008).

CLF associated protein F-Box protein UPWARD CURLY LEAF1 (UCL1) was recently described. Jeong et al (2015) suggested its function as an assistant of the FIS2-PRC2 complex formation in the endosperm. It is possible that UCL1 prevents CLF to compete with MEA in the endosperm. If CLF is expressed in the endosperm, it harms the seed development by making undesirable PRC2 complex. Specifically, UCL1 leads CLF degradation by interacting with E3 ubiquitin ligase complex (Jeong et al., 2011).

The Enhancer of zeste [E(z)] homolog CLF also interacts with cullin-ring ubiquitin ligase CUL4-DDB1 and MSI4 so that it can repress FLC expression in *Arabidopsis* (Pazhouhandeh et al., 2011). When MSI4 or CUL4 mutated H3K27me3 markers on FLC and FT locus are reduced, allowing misexpression of the genes (Pazhouhandeh et al., 2011).

CUL4-DDB1 also associates with MSI1 for maintenance of H3K27me3 levels at FIS2-PRC2 regulated *MEA* and *PHERES* (Dumbliauskas et al., 2011). Furthermore, MSI1 collaborates with EMBRYONIC FLOWER1 (EMF1) (Calonje et al., 2008) and LIKE HETEROCHROMATIN PROTEIN1(LHP1) (Derkacheva et al., 2013). More recently, it was reported that EARLY IN SHORT DAYS7 (ESD7) interacts with CLF, EMF2 and MSI1 for maintenance of H3K27me3 at FT and SOC1 (Del Olmo et al., 2016).

3. Structural characterization of PRC2

To have a good grasp of the PRC2 complex, it is inevitable to study about the structural biochemistry and the functional domains of the PcG proteins. The SET domain of the Enhancer of zeste [E(z)] protein is responsible for the histone methyltransferase (HMTase) activity. It is interesting that without the association of other members of the PRC2 complex, E(z) does not exhibit HMTase activity *in vivo* (Czermin et al., 2002). The function of Extra Sex Comb (ESC) is to enhance the E(z) enzymatic activity (Nekrasov et al., 2005). Supressor of zeste 12 [Su(z)12] and Nurf55 are important for histone H3 binding (Nekrasov et al., 2005).

4. Purpose of this study

EFC was first identified in the activation tagging mutagenesis study. The dominant *efc-D* mutant plant flowered prior to wild type plants and showed curled leaf phenotype, therefore this mutant was designated as *early flowering and curly leaf* (*efc*) (Kim, 2006). The further study revealed the insertion site of the enhancer element and upregulated adjacent gene that caused this phenotype. It was 16 kDa small unknown protein coding *At4g23110* gene (Kim, 2006). Roh (2013)

reconfirmed that the mutant phenotype was indeed caused by over expression of the *At4g23110* gene because 35CaMV::EFC plants mimicked the phenotype of *efc-D*.

In the study of *EFC* overexpression plants, PRC2 target genes such as *FLOWERING LOCUS T* (*FT*) and floral homeotic genes *AGAMOUS* (*AG*), *SEPALLATA1* (*SEP1*), *SEPALLATA3* (*SEP3*), *PISTILLATA* (*PI*) and *APETALA* (*AP3*) were highly expressed compared to the WT (Roh, 2013). Roh (2013) reported that *EFC* as an upstream regulator of FT in the pathway of flowering timing. Furthermore, genes regulated by PcG such as *AGAMOUS LIKE17* (*AGL17*) and *MEDEA* (*MEA*) were derepressed in 35S::EFC plants because there were significant decrease of H3K27me3 levels on *AG*, *FLC*, *FT* and *MEA* loci according to the ChIP assay (Roh, 2013). The phenotype of *efc-D* and *35S::EFC* was due to misregulation of the genes involved in flowering and determination of leaf morphology (Roh, 2013).

EFC knockdown mutant was obtained through amiRNA methodology. Phenotype of the pFWA::amiR_EFC showed about 30% of seed abortion and endosperm proliferation in the absence of embryonic development. Similar phenotype is seen in the plants with mutations in members of FIS2-PRC2. FIS2-PRC2 complex plays fundamental role in endosperm development by controlling the expression of the genes in the female and male gametophyte and also in the endosperm. Strikingly, subcellular localization of *EFC* in *Arabidopsis* was in the nuclei of the central cell and endosperm (Roh, 2013).

Additionally, co-Immunoprecipitation and glutathione S-transferase assay showed that *EFC* interacts directly with *MSI1*, one of the core members of the PRC2 (Roh, 2013).

In this study I characterized a loss-of-function mutant *early flowering and curly leaf* (*efc-1*). In the *efc-1* mutant, the T-DNA was inserted into the exon of the *EFC* and resulted in null mutant. The *efc-1* mutation exhibited seed abortion phenotype in *Arabidopsis*. *EFC* might have a partial or incomplete penetrance because it showed fluctuation in the seed abortion ratio. The embryonic development was arrested at the globular stage in the *efc-1* mutant seeds. Considering the clues obtained through the previous results, I carried out experiments to check whether EFC plays role in the activity of FIS2-PRC2. The expression of *MEDEA* (*MEA*) and *UPWARD CURLY LEAF1* (*UCL1*), genes that are regulated by FIS2-PRC2, were derepressed in the central cell or endosperm of the *efc-1* mutant. I carried out a complementation experiment of *fis2* with *EMBRYONIC FLOWER2* (*EMF2*), a homolog of *FIS2*, in the seed in absence of *EFC*. Presumably due to the partial penetrance of *EFC*, there was no significant complementation observed in the heterozygous *efc-1*; heterozygous *fis2-11* and hemizygous *FIS2::EMF2* triple mutant plants. Therefore, it needs to be checked in the next generation when I obtain a triple homozygous mutant or double homozygous and *fis2-11* heterozygous mutant. The results so far may suggest that *EFC* might play a role in seed development along with the FIS2-PRC2 complex, however not essential.

I. Materials and Methods

Plant materials and growth conditions

Arabidopsis thaliana ecotype Columbia (Col-0) was used throughout this study as a wild type. *Nicotiana benthamiana* was used in the Co-Immunoprecipitation assay. MEA::GUS, UCL1_4.1k::GUS transgenic plants were applied in GUS expression analysis experiments. In complementation test *fis2-11* mutant plants were utilized. The T-DNA line SALK_018803.52.75.x (*efc-1*) was ordered from Arabidopsis Biological Resource Center (ABRC). Plants were grown in the growth room or growth chamber under long day conditions (16 h of light at 24°C /8 h of darkness at 22°C). Seeds were sown on the surface of soil mixed with vermiculite and moss, then put in 4°C cold chamber and darkness for 2 days before being moved to growth room. For transgenic plant screening, seeds were surface sterilized by treating with 75% ethanol containing 0.08% TritonX-100 (SIGMA) for 15 min twice, followed by washing briefly with 100% ethanol (MERCK). The seeds were dried on Whattman filter paper and plated on solidified MS agar plates containing appropriate antibiotics (25 μg/ml kanamycin). These MS agar plates consisted of 0.5X Murashige and Skoog salts (DUCHEFA), and 1.5% (w/v) plant agar (DUCHEFA), pH between 5.7 and 5.8. MS media were autoclaved in 121°C for 20 min. Before transferring to the growth room, MS plates with seeds were also put in 4°C for cold treatment for 2 days. Per construct, around 20 Colombia-0 (Col-0) wild-type plants and 450 *fis2-11* plants were used for *Agrobacterium*–mediated transformation by the floral dipping method.

Characterization of the *efc-1* allele

To confirm the T-DNA insertion, genotyping PCR was carried out by using genomic DNA as a template. According to the database of SALK institute T-DNA in *efc-1* is inserted 18 nucleotides downstream of the predicted start codon of *EFC*. The left border junction was determined using the T-DNA primer LB1.3 combined with the genomic primer EFC_RP. For amplification of the *EFC* gene in wild-type, gene specific primers EFC_LP and EFC_RP were designed. To test the expression of the *EFC* gene at the transcriptional level, primer sets for RT-PCR were designed. Nucleotide alignment was performed by using the CLUSTAL W.

Recombinant plasmid construction

For complementation test, a pFIS2::cEMF2 construct was generated. The pFIS2::cEMF2 construct includes a 2,247-bp sequences upstream of the predicted translational start codon of FIS2 that was fused with 1,896-bp cDNA sequence of EMF2 including both start and stop codons. The regulatory region of FIS2 was obtained by PCR amplification with a primer set of pFIS2_F and pFIS2_R using wild-type Col-0 genomic DNA as a template, cDNA of EMF2 was amplified by cEMF2_F and cEMF2_R primers using previously cloned pJET-cEMF2 as a template. Fragments were then fused into the pBI101 vector in order by homologous recombination using the infusion technology (TAKARA Clontech). First, the pFIS2 region was cloned into the GUS marker gene containing binary vector pBI101 and later cEMF2 was cloned into the previously generated pBI101-pFIS2 vector.

35S::EFC:GFP, 35S::MSI1:T7 and 35S::FIS2-HA constructs in binary vectors were generated for Co-Immunoprecipitation assay. The 444-bp of EFC coding sequence excluding stop codon fused with 720-bp of the EGFP gene was PCR

amplified with the primer set of EFC_infusion_F and EFC_infusion_R using 35S:EFC-GFP non-binary plasmid as a template. This EFC-GFP amplicon was cloned into 35S promoter involving binary vector pBI111-L using the infusion technology (TAKARA Clontech). Similarly, the MSI1 coding sequence in the same frame of T7 tagging protein gene is amplified with a primer set of MSI1_infusion_F and MSI1_infusion_R using non-binary vector 35S::T7::MSI1. Then the amplified fragment was cloned into the 35S promoter involving binary vector pBI111-L using the infusion technology (TAKARA Clontech).

The attB PCR product of EMF2 was gained by PCR amplification using primer set of cEMF2_attB1_F and cEMF2_attB2R. Previously cloned pJET-cEMF2 was used as a template. In order to generate pENTRY-cEMF2, PCR amplicon was inserted into pDONR221 vector using the BP reaction of Gateway technology (Invitrogen). Then pENTRY-cEMF2 and pGWB14 vectors were used in LR reaction of Gateway technology to generate 35S::cFIS2::HAx3.

Agrobacterium tumefaciens transformation and plant transformation using floral dipping

In order to transform Arabidopsis genome, the constructs were introduced into *Agrobacterium* by electroporation. For transformation, 40 µl *Agrobacterium tumefaciens* (GV3101) cell stock in a 1.5 ml tube was thawed on ice for 10 min. 1.5 µl of DNA of interest was added to the competent cell, then microcentrifuge tube was tapped gently for mixing. After transferring the mixture into glass cuvette, it was inserted into the Micro-Pulser™ (BIORAD) in a proper orientation and pursed by 1.8 kV for 5.8 mS. After electroporation 400 µl of LB medium was added to the cuvette and mixture was transferred to a 1.5 ml tube and incubated in shaking incubator at 28°C for 2 h before spread on solidified Luria Broth (LB) plate with

proper antibiotic (50 μg/ml kanamycin,). The plate was incubated also at 28°C for 2 to 3 days.

For *Arabidopsis* transformation, floral dipping method was used. Selected single colonies of transformant *Agrobacterium* were inoculated in 5 ml of LB medium with adequate antibiotics (50 μg/ml kanamycin, 50 μg/ml gentamycin) for 24 h at 28°C shaking incubator. To amplify the bacterium cell, mini-prep cultures were then added into 500 ml of LB medium containing same antibiotics and grown in the same condition for 2 days. Incubated *Agrobacterium* was harvested by centrifuging at 4000 rpm for 20 min at 20°C (SORVALL® RC 6 PLUS with SLC-3000 rotor). After discarding supernatant, the cell pellet was resuspended in infiltration media. 500ml infiltration media contains 1.05 g of MS salt, 25 g of sucrose and 250 μl of Silwet (Vac-In-Stuff, Silwet L-77, LEHLE SEEDS). For successful Arabidopsis transformation, Col-0 and *fis2-11* plants were grown healthily until an adequate number of inflorescences are generated in a long day condition. Right before the transformation siliques and open flowers were removed. Young buds were then dipped into the infiltration media containing *agrobacterium* for 10 sec. After floral dipping, plants were laid on a tray and were left in dark condition for 24 h. Next day, plants were put vertically and grown until their seeds were mature enough to harvest.

Allele-specific expression analysis

To investigate allele-specific expression of *UCL1*, the Derived Cleaved Amplified Polymorphism (dCAPS) assay was used. The experiment design was adopted from Jeong et al (2015). The database of The 1001 Arabidopsis Genome Project was used to detect SNPs among the Col-0, L*er*, RLD ecotypes. Sequences

were aligned using the Clustal W program. In order to analyze allele-specific expression, the PCR products were amplified with a primer set of JCW641/JCW642 from cDNA synthesized with the RNA of the siliques 4 days after crossing. The amplicon was digested with *Eco*RI to detect expressed alleles. RLD produce 222 bp and 54 bp fragments when digested with *Eco*RI, however, Col-0 produces 276 bp uncut DNA fragment. The PCR products and the samples that are digested were analyzed on 4% agarose gels.

Histochemical GUS staining analysis

The expression of GUS was analyzed in the pFIS2::GUS, MEA::GUS and UCL1_4.1k::GUS plants. For analysis of gene expression in female gametophyte, flowers were emasculated and left for 24 h. Whereas, to investigate gene expression in seeds, flowers were emasculated and pollinated, then grown for 12 or 24 h. The tissues were dissected and sampled in the X-Gluc staining solution containing 100 mM sodium phosphate buffer (pH 7.0), 2 mM each of potassium ferricyanide and ferrocyanide, 2 mM X-Gluc and 0.1% (v/v) Triton X-100 for overnight in the dark condition.

Microscopy

The tissues were mounted by the clearing solution (1 ml 70% glycerol, 2.5 g chloral hydrate) on a slide glass. GUS expressing seeds and mutant seeds were observed on a Zeiss Axio Imager A1 light microscope under differential interference contrast optics with 10x, 20x and 40x objectives and photographed by AxioCam HRc camera (Carl Zeiss).

Table 1. List of primer sequences

Oligo name	Sequences
LB1.3	ATTTTGCCGATTTCGGAAC
EFC_LP	GTATTGTAGACGGATTTTTTACCC
EFC_RP	CTATGAACTGTCATCTGCCAA
EFC_qLP	TTGCAAGTTCCAAATGGGGAAA
EFC_qRP	TCTTGTTTTGCTCTTTGTCTGC
pFIS2_F	CGACTCTAGAGGATCAAGCTTGACCTAATCAA AGTCTGTC
pFIS2_R	GACCACCCGGGGATCCCTCTGCTTGATTAATCT ATAAGCTGTCAC
cEMF2_F	TCAAGCAGAGGGATCCATGCCAGGCATTCCTC TTGTTAG
cEMF2_R	GATCGGGGAAATTCGAGCTCTCAAATTTGGAG CTGTTCGAGAAAGG
EFC_infusion_F	CACGGGGGACTCTAGAATGGGGAAAAAATCTA AGCC
EFC_infusion_R	GCTCACTAGTCTCGAGTTACTTGTACAGCTCGT CC
MSI1_infusion_F	CACGGGGGACTCTAGAATGGCTAGCATGACTG GTG
MSI1_infusion_R	GCTCACTAGTCTCGAGCTAAGAAGCTTTTGATG GTTC
cEMF2_attB1_F	GGGGACAAGTTTGTACAAAAAGCAGGCTGATG CCAGGCATTCCTCTTGTTAG

cEMF2_attB2_R	GGGGACCACTTTGTACAAGAAAGCTGGGTGAA TTTGGAGCTGTTCGGAAAGG
JCW641	TCGGAATCGGTAGGGATG
JCW642	CTTTGGGGAGGCGTTTGA
ACT2_F	CCGCTCTTTCTTTCCAAGC
ACT2_R	CCGGTACCATTGTCACACAC

III. Results

1. Characterization of the *efc-1* allele

EFC (At4g23110) encodes a small protein that previously identified during activation tagging experiment in our laboratory (Kim, 2013). When *EFC* was overexpressed, target genes of PRC2 complexes, such as *FLOWERING LOCUS T* (*FT*) and floral homeotic genes *AGAMOUS* (*AG*), *SEPALLATA1* (*SEP1*), *SEPALLATA3* (*SEP3*), *PISTILLATA* (*PI*) and *APETALA* (*AP3*), were upregulated (Roh, 2013). To gain insight into the *EFC* function, I obtained a line containing T-DNA insertion in this gene from the SALK Institute Genomic Analysis Laboratory collection (Alonso et al., 2003) (Figure 2A). I analyzed a T-DNA homozygous knock out line and named *efc-1* (SALK_018803.52.75.x). To identify both left and right flanking sequences of the T-DNA, polymerase chain reaction (PCR) was carried out and sent for sequencing analysis. Molecular evidence showed that the T-DNA was inserted into the exon of *EFC*, 12 nucleotides downstream of the start codon, and is associated with 54-bp deletion (nucleotides -42 to +12) (Figure 3). Additionally, this T-DNA was duplicated to be inserted into the genome, showing two left borders (Figures 2A and 3).

To test the deletion of the *EFC* gene at the transcriptional level, reverse transcription-PCR (RT-PCR) was performed (Figure 2C). *EFC* is known to be specifically expressed in the central cell and endosperm (Roh, 2013). Therefore, RNA was extracted from the siliques 3 DAP. The RT-PCR result showed that *efc-1* is a null allele.

Figure 2

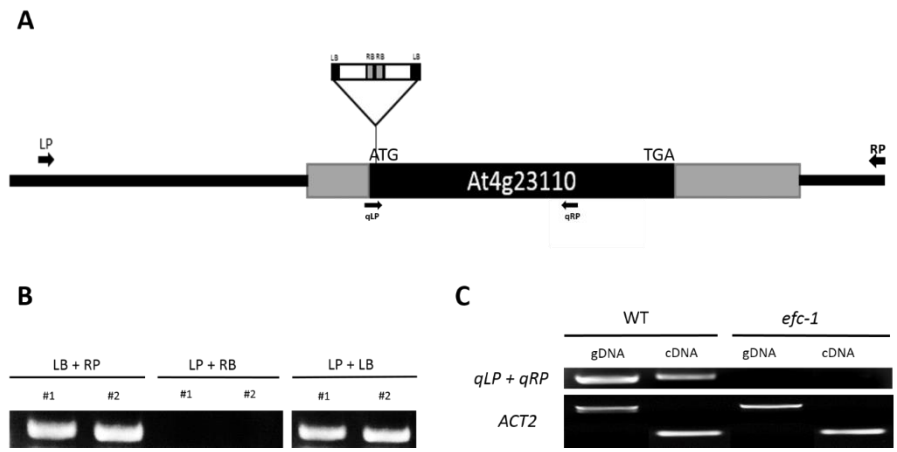


Figure 2. Characterization of the *efc-1* allele

(A) Schematic diagram showing the structure of the *EFC* gene with the T-DNA flanking region of the *efc-1* mutant allele. Black box, translated exon; gray box, untranslated exon. The insertion site of T-DNA is marked by a triangle.

(B) PCR for identification of the T-DNA insertion orientation and structure. The T-DNA contains two LB sequences. #1 and #2 indicates two individual homozygous lines of *efc-1*.

(C) Quantitative RT-PCR analysis was performed to check the deletion of the *EFC* gene at the transcriptional level. Total RNA was extracted from siliques (3DAP). *ACT2* was used as an internal control.

Figure 3

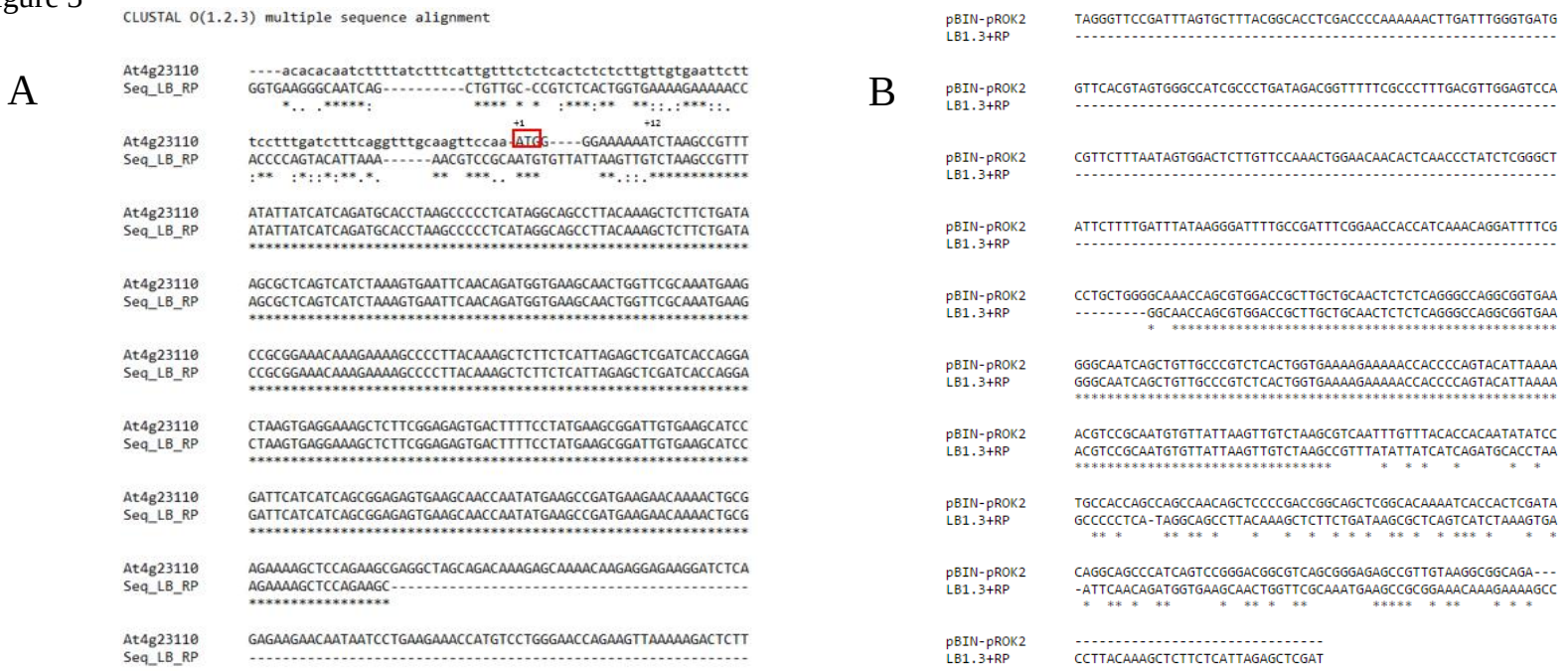


Figure 3. Identification of the T-DNA insertion site and orientation

(A) Comparison of the *At4g23110* genomic DNA sequence and LB+RP PCR product sequence. T-DNA was inserted 12-bp downstream of the *EFC* start codon.

(B) Alignment of LB+RP PCR product sequence and T-DNA left border sequence. There is a 6-bp sequence (TCTAAG) that overlaps with the *At4g23110* sequence.

Figure 3 (continued)

C

D

(C) Comparison of the *At4g23110* genomic DNA sequence and LP+LB PCR product sequence. The 54 nucleotide deletion occurred due to the T-DNA insertion.

(D) Alignment of the LB+RP PCR product sequence and T-DNA left border sequence. There is a 3-bp sequence (TGT) that overlaps with the *At4g23110* sequence.

2. A mutation in *EFC* Affects Seed Development

Since the previous knockdown allele of *EFC* showed 30% seed abortion phenotype, I determined whether *EFC* is required for female gametophyte and/or seed development. I scored the number of seed sets in the siliques of *efc-1* plants. The homozygous *efc-1* mutant plants showed fluctuation in the seed abortion phenotype. Seed abortion ratios ranged from 2% to 40% depending on the plants. The phenotype of the mutant plants were divided into two populations , firstly, ~50% of the individual plants of the same line had around 30% of seed abortion, and the other 50% of the individual plants of the same line had a weaker phenotype that showed ~5% of defect in seed sets in siliques (Table 2). 30% of abnormal seed development is consistent with the phenotype of the *EFC* knockdown mutant (Roh, 2013).

To check whether this phenotypic trait is preserved in the next generation, I chose 5 plants from each population and scored the number of seed sets in the siliques. The seed abortion rate was decreased approximately 6% in progeny of those lines that showed a higher seed abortion rate (in average~27%). On the other hand the abortion rate was increased by around 9% in progeny of the lines that showed 3.2% seed abortion on average (Table 2). Partial penetrance of the gene might be a cause of the inconsistent phenotype. Although *efc-1* doesn't show a high seed abortion rate, presence of abnormal seed suggests that *EFC* might help the normal seed formation.

Table 2

	Lines with high seed abortion ratio (19~41%)	Lines with low seed abortion ratio (3~7%)
Parental	27% (n=1168)	3.2% (n=1278)
Progeny	21.6% (n=9103)	12.6% (n=8640)

Table 2. The abortion rate of *efc-1*

Fluctuation in the seed abortion phenotype of the *efc-1*.

3. Phenotypic analysis of seeds in the *efc-1* mutant

To clarify the phenotype, seeds of the *efc-1* mutants and WT at different developmental stages (1 DAP to 8 DAP) were observed under a uplight microscope. In *efc-1* mutant plants, white seeds were observed (Figure 4A). Then I checked the embryonic development of these mutant seeds. In *efc-1* mutant seeds embryonic development was delayed and arrested at the globular stage (Figure 4B). I could observe arrested embryos at as early as heart stage. At 8 DAP, the white seeds were significantly distinguishable from the normal seeds. White seeds were almost the same size as the normal maturing seeds, however embryonic development was arrested. The phenotype of each seeds in the same silique might vary. Consistent with the previous result, penetration of *EFC* mutation seems to be incomplete.

Figure 4

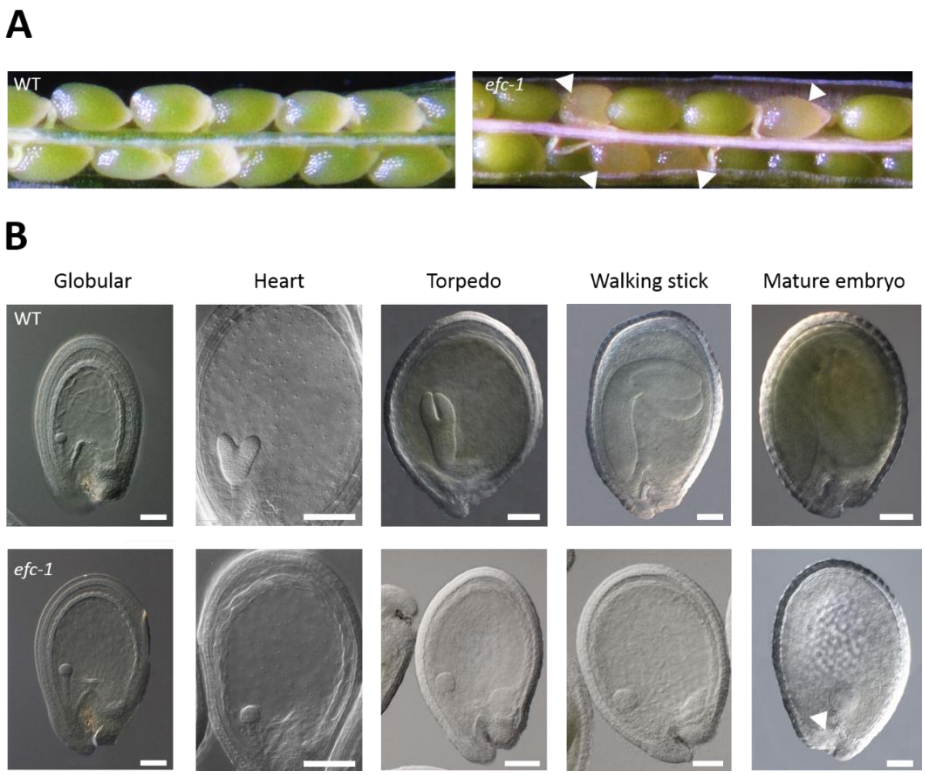


Figure 4. Seed abortion in the *efc-1* knockout line

(A) Seed development in the siliques of WT and *efc-1* mutant plants. Arrowheads indicate aborted seeds.

(B) Normal embryogenesis of *EFC/EFC* plants (upper panel) compared with arrested or delayed embryogenesis of *efc/efc* seeds. (Scale bar: 100μm).

4. The role of *EFC* in the FIS2-PRC2 complex

To identify the role of EFC in the endosperm development I designed experiments using *MEA::GUS* (Jullien et al., 2006) and *UCL1_4.1k::GUS* (Jeong et al., 2015) lines. *MEA* is a maternally expressed imprinted gene. *MEA* is not expressed in the pollen as well as not in the endosperm paternally. By contrast, *UCL1* is a paternally expressed imprinted gene, so *UCL1* expression is absent in ovules before fertilization. I crossed the *efc-1* plant with the *MEA::GUS* plant to see how the *efc-1* mutation affect the imprinting pattern of *MEA* compared to wild type. In the seed, the paternal *MEA* expression was not detected both in WT and *efc-1*$^{+/-}$ plants pollinated with *MEA::GUS* at 24 h after pollination (HAP). In order to detect the full effect of *efc-1*, I created the *efc-1/efc-1*; *MEA::GUS/MEA::GUS* double homozygous plant. Expression of maternal *MEA-GUS* was very strong in the just fertilized seeds and decreased gradually with time. Paternal *MEA-GUS* expression was absent at 12 HAP, 24 HAP in WT seed; however at 24 HAP expression of the gene was derepressed in *efc-1*$^{-/-}$ mutant seed pollinated by *MEA::GUS/MEA::GUS*; *efc-1/efc-1* (Figure 5A). Compared to WT, GUS stained seeds in *efc-1* was increased by 12 to 13% (Figure 5B). This result suggests that *EFC* might play an important role in identity of the FIS2-PRC2 complex.

In order to carry out the second experiment, the *efc-1*$^{-/-}$ plant was crossed with *UCL1_4.1K::GUS*$^{+/+}$ and heterozygous seeds were harvested and grown on soil. I then emasculated the plants to check whether maternal *UCL1* is derepressed in the *efc* heterozygous background. I could observe an abnormal expression of *UCL1_4.1k::GUS* in *efc-1*$^{+/-}$ ovule before fertilization (Figure 6). However, the frequency of the expression was very low. Since *efc-1* has partial penetrance, the

phenotype of this mutant might be even weaker in heterozygous plants. *UCL1::GUS* expression was detected in the ovule before fertilization in *efc*$^{+/-}$, whereas there was no expression detected in the *UCL1_4.1K::GUS* ovules.

Figure 5

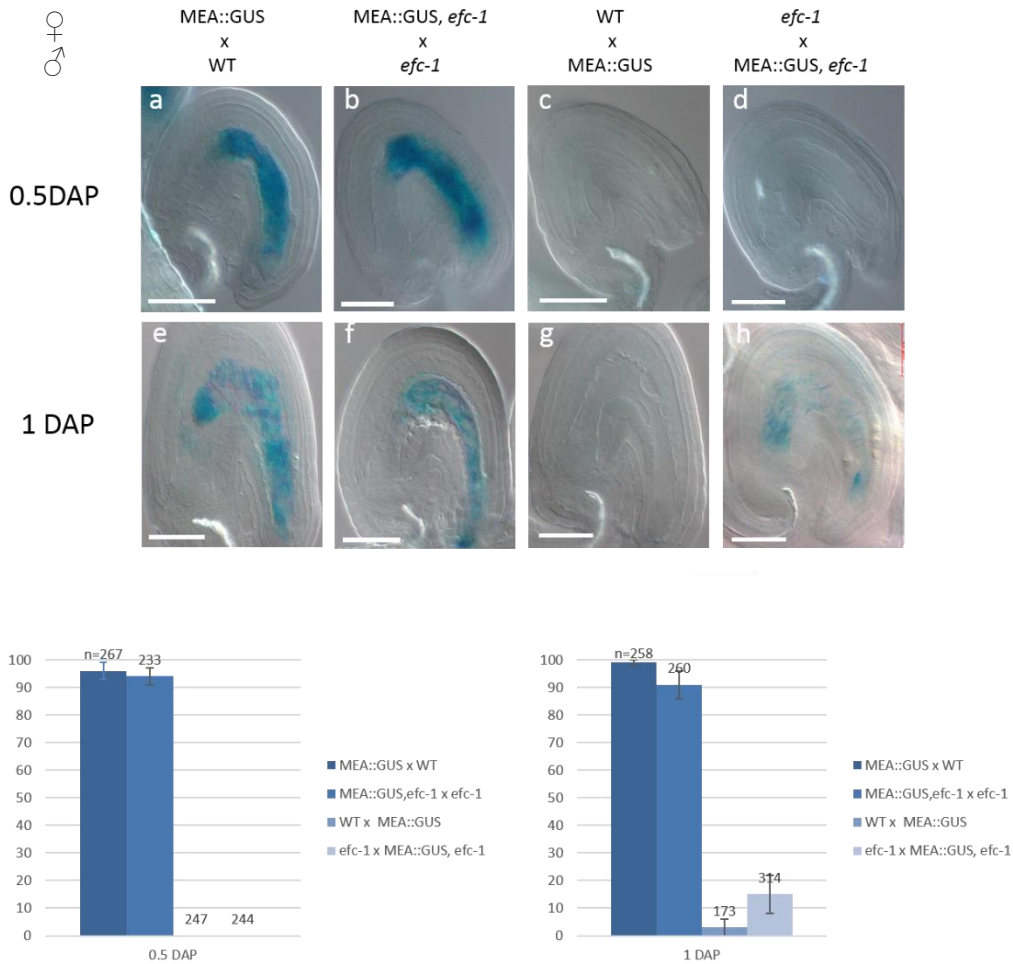


Figure 5. Derepression of paternally derived *MEA* in the *efc-1* mutant

(A) The expression of paternal *MEA* was detected in *efc-1/efc-1* background 24 HAP.

(B) Graphs showing the percentage of the *MEA::GUS* expression.

Figure 6

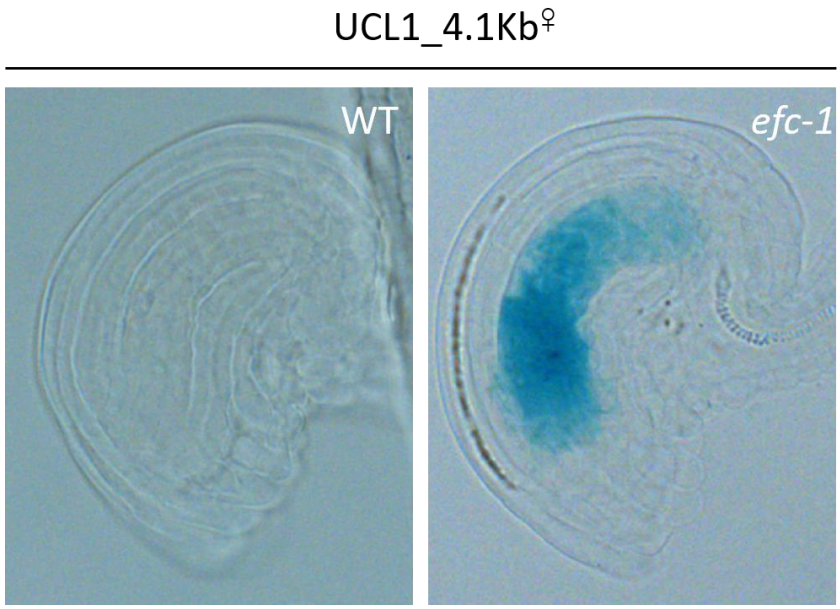


Figure 6. Derepression of *UCL1* in the central cell of the *efc-1* mutant ovule

efc-1$^{-/-}$ was crossed with *UCL1_4.1K::GUS*$^{+/+}$. *UCL1::GUS* expression was detected in the ovule before fertilization in *efc*$^{+/-}$. There was no expression detected in the *UCL1_4.1K::GUS* ovules before fertilization.

5. Allele specific expression of *UCL1* in *efc-1*

UCL1 is silenced in the female gametophyte and expressed in the male gametophyte. It was confirmed that FIS2-PRC2 is responsible for its imprinting, because maternal *UCL1* is derepressed in the *mea-3* homozygous or *fie-1* heterozygous background (Jeong et al., 2015). Thus, to check if *EFC* is crucial for the FIS2-PRC2 complex, I carried out an experiment using RT-PCR and cleaved amplified polymorphic sequence (CAPS) markers of different ecotypes (Jeong et al., 2015) to identify allele specific expression by using different ecotype plants as parents. The database of The 1001 Arabidopsis Genome Project was used to detect SNPs among the Col-0, Ler, RLD, C24, En2 ecotypes. Sequences were aligned using the Clustal W program (Figure 7). Consistent with the finding of Jeong et al (2015), when the *fis2* heterozygous plant was pollinated with the pollen of RLD wild type plants, maternal *UCL1* expression was detected (Figure 8). However, I could not detect maternal expression of *UCL1* in *efc-1* homozygous plant. This result suggests that *EFC* may not be essential for the imprinting of *UCL1*. Thus it needs to be checked whether other target genes of the FIS2-PRC2 complex is regulated by *EFC*.

Figure 7

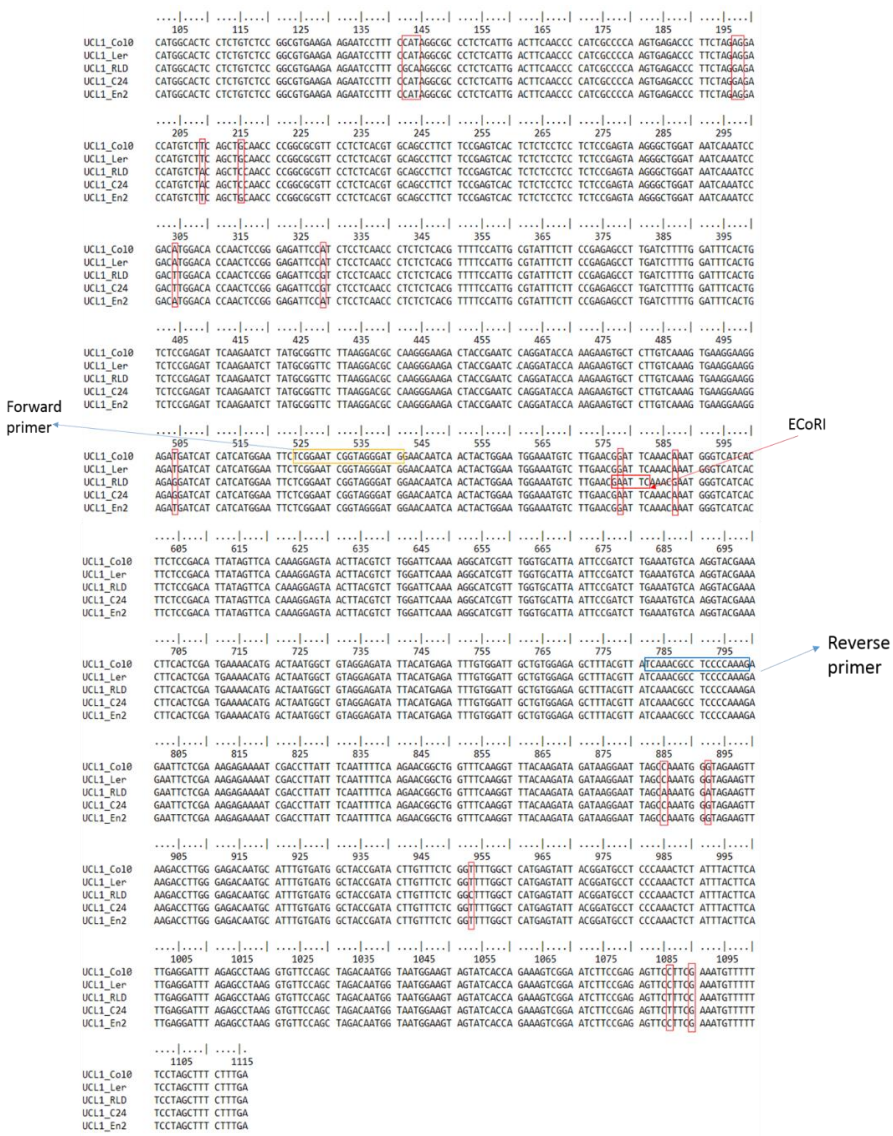


Figure 7. Single Nucleotide Polymorphism of *UCL1* in different ecotypes

A 276-bp RT-PCR product of RLD ecotype contains an *Eco*RI site. When digested with *Eco*RI, the RLD PCR product was cut into 222-bp and 54-bp fragments.

Figure 8

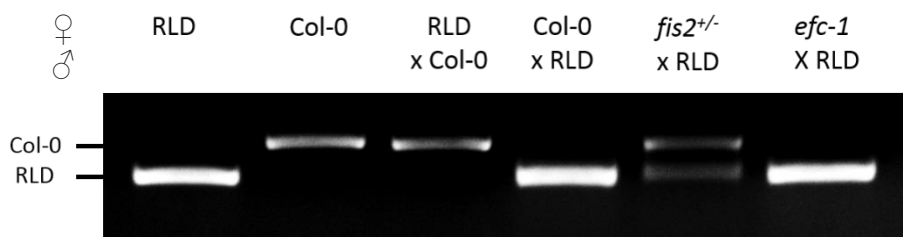


Figure 8. Allele specific expression of *UCL1* using a CAPS marker

RT-PCR analysis was carried out with RNA isolated from siliques (4 DAP) of RLD females crossed with Col-0 males, Col-0 females were crossed with RLD males, and *fis2* or *efc-1* females (Col-0 background) were crossed with RLD males.

6. Complementation of *fis2* with *EMF2* in the *efc-1* mutant

FIS2 and EMF2 are Su(Z) 12 repressor homologs of FIS2-PRC2 and EMF2-PRC2, respectively. It has been previously reported that EMF2 cannot substitute FIS2 (Roszak and Kohler, 2011). However, *EMF2-PRC2* and *FIS2-PRC2* share some target genes (Makarevich et al., 2006) and Su(Z) family members might have roles in identification of target genes, although the mechanism behind is yet to be discovered. To confirm whether *EMF2* indeed does not complement *FIS2* in the FIS2-PRC2 complex, I constructed a *FIS2::EMF2* transgenic plant to ectopically express *EMF2* in endosperm. Firstly, I cloned the promoter region of *FIS2* (*pFIS2*) into GUS marker gene containing pBI101 vector and later I cloned cDNA of *EMF2* (*cEMF2*) into the pBI101-pFIS2 vector (Figure 9A). The *cEMF2* region contains the stop codon, so the GUS is not detectable by histochemical GUS staining after successful transformation. Thus in order to check the accuracy of the promoter region of *FIS2*, the *pBI101-pFIS2-GUS* construct was also used for transformation and later tissues were GUS stained. I confirmed that the pFIS2 fragment was enough for gene expression in the central cell and endosperm (Figure 9B).

To check whether *EFC* is responsible for the activity of the FIS2-PRC2 complex by interrupting *EMF2*, around 450 *fis2-11* homozygous plants (due to low germination ratio) transformed by pFIS2:cEMF2. *fis2-11* has a deletion at 103G, causing a nonsense mutation. *FIS2::EMF2* could not complement seed abortion caused by *fis2-11* mutation (Figure 10). Among 25 transformants, T2 seeds of *fis2-11*; pFIS2::cEMF2, #16 and #17 had germinated and produced more viable seedlings than the other lines. Thus #16 and #17 were used for the further experiment. *efc-1* homozygous plants were crossed with #16 and #17 lines of *fis2-11* containing

FIS2::EMF2. I obtained *fis2-11* heterozygous; *efc-1* heterozygous; *FIS2::EMF2* hemizygous plants and these triple mutants were confirmed by genotyping PCR. I could not observe significant complementation in this generation. So it is necessary to obtain homozygous plants and check for the complementation ratio in the next generation.

I also transformed Col-0 plants with *FIS2::EMF2*. Strikingly, when *EMF2* was overexpressed in WT, it caused ~12.5% seed abortion (24 out of 49). This result suggests that EMF2 might compete with FIS2 in the process of the seed development (Figure 11).

Figure 9

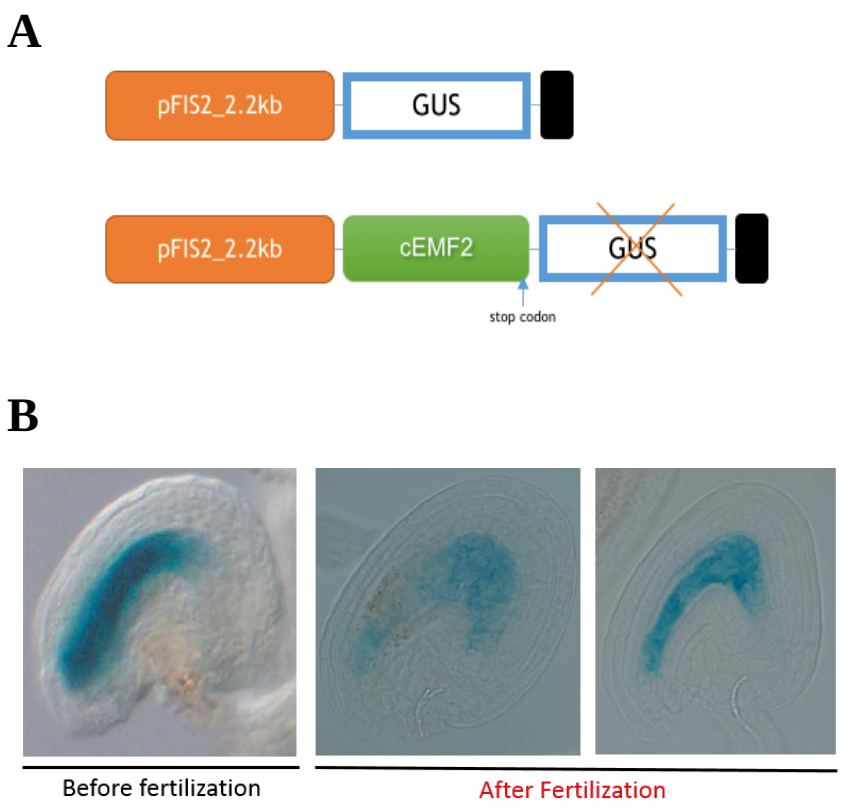


Figure 9. Schematic diagram of *FIS2::GUS* and *FIS2::EMF2* constructs and the expression of *FIS2::GUS*

(A) The 2.2 kb fragment containing the upstream sequence of *FIS2* fused with GUS and cDNA of *EMF2* which contains the stop codon.

(B) Expression of *FIS2::GUS* expression before and after fertilization.

Figure 10

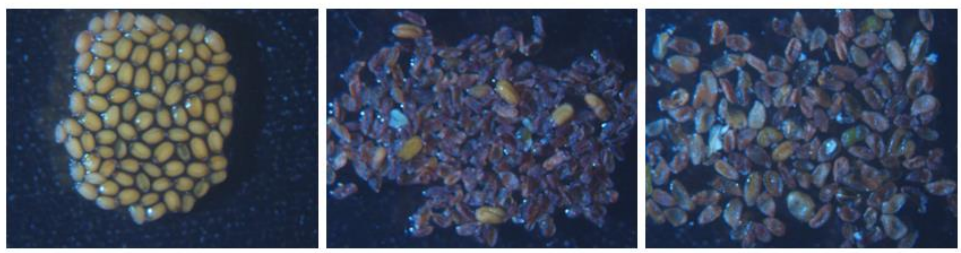


Figure 10. Seeds of Col-0, *fis2-11* mutant and transgenic *FIS2::EMF2* in the *fis2-11* background

Figure 11

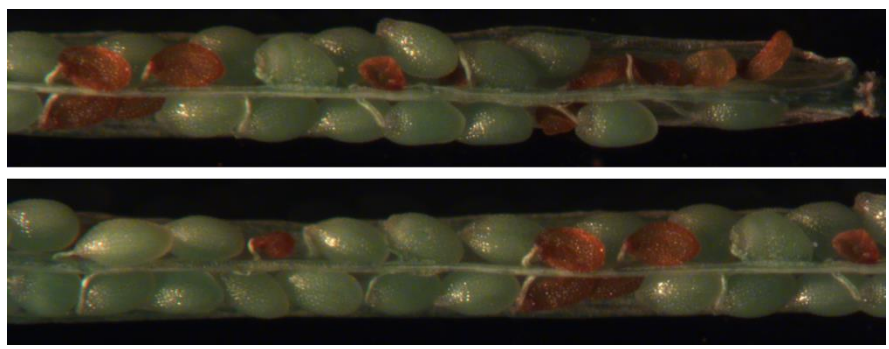
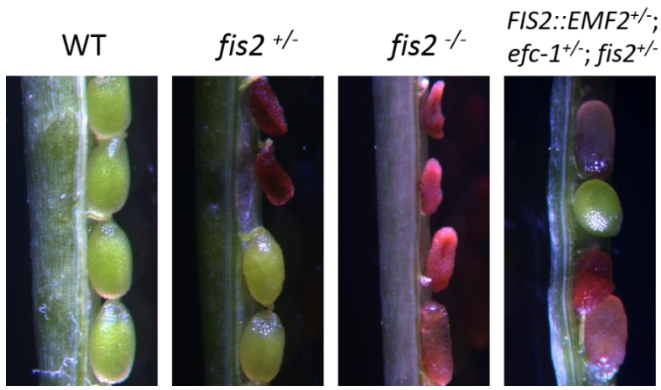


Figure 11. Phenotype of *FIS2::EMF2* transgenic lines

Figure 12

A


WT
fis2 +/-
fis2 -/-
FIS2::EMF2+/-; efc-1+/-; fis2+/-

B

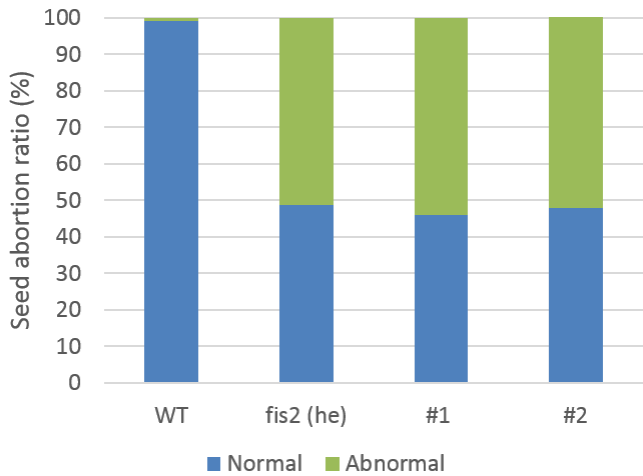

Seed abortion ratio (%)
100
90
80
70
60
50
40
30
20
10
0
WT
fis2 (he)
#1
#2
Normal
Abnormal

Figure 12. Complementation of *fis2-11* with cEMF2 in the absence of EFC

(A) Seed phenotypes of WT, *fis2*$^{+/-}$, *fis2*$^{-/-}$, *FIS2::EMF2* hemizygous; *efc-1*$^{+/-}$*; fis2*$^{+/-}$.

(B) Graph shows ratios of seed abortion in each lines. #1 and #2 indicates individual lines of *FIS2::EMF2* hemizygous; *efc-1*$^{+/-}$*; fis2*$^{+/-}$.

IV. Discussion

The previous study on the *EFC* gene suggested that *EFC* might play a positive role on the FIS2-PRC2 complex in the seed development, because the knockdown allele showed a seed abortion phenotype. And when *EFC* was ectopically expressed, target genes of the EMF2-PRC2 complex were misregulated, suggesting that it may have a negative regulator role on the other PRC2 complexes in the vegetative tissue. Since it was confirmed that *EFC* was specifically expressed in the central cell and endosperm, and directly interacted with the main member of the PRC2 complex, MSI1, I assumed that the function of this protein might be related to that of the FIS2-PRC complex.

In this study, I characterized *efc-1*, a loss-of-function mutant. The T-DNA was inserted into the exon of *EFC*. It associated with a 54 bp deletion in the gene and resulted in a null allele. *EFC* may play a role in the normal seed development, because *efc-1* showed the seed abortion phenotype. The embryonic development of the aborted seeds of *efc-1* plants were arrested at the globular stage. However, there was a fluctuation in the seed abortion ratio which suggests that *efc-1* has a partial or incomplete penetrance. Partial penetrance is a phenomenon that development of isogenic organisms being affected by mutations differently (Coote, 1972; Horvitz and Sulston, 1980; Queitsch et al., 2002; Sangster et al., 2008). A well-known example of this phenomenon is polydactyly, when individual is affected by the genetic alteration, he or she will have an extra digit in their hand. In *Arabidopsis*, *turnip* (*tnp*), *sulfurtransferase1* (*str1*), *receptor-like protein kinase1* (*rpk1*) and *resurrection1* (*rst1*) mutant plants were reported to have an incomplete penetrance (Chen et al., 2005; Casson and Lindsey, 2006; Mao et al., 2010; Luichtl et al., 2013).

The *str1-1* mutant exhibited 87.5% seed abortion phenotype while 70% of the *rst1* seeds were highly shrunken and wrinkled (Chen et al., 2005; Mao et al., 2010). According to Luichtl et al (2013) due to functional redundancy only 8.38% of the *rpk1* mutant plants develop monocot seedlings. The reasons behind the incomplete or reduced penetrance of gene are not fully understood. They can be environmental influences, gene interactions that can cover the phenotype or other molecular mechanisms.

According to the experiments that used *MEA::GUS,* I observed the derepression of paternal *MEA* in the seed, in the *efc-1* mutant background. In this experiment, the *MEA::GUS; efc-1* homozygous plant was crossed with pollens of the *efc-1* homozygous plant to check the full influence of *efc-1*. The resulting plant was hemizygous in *MEA::GUS* and homozygous in *efc-1*. There was a ~15% increase in paternal *GUS* expression in the *efc-1* mutant background, compared to no expression in the WT background. This suggests that EFC does play some role in the FIS2-PRC2 complex.

I also observed misregulation of maternal *UCL1* in the *efc-1* background. In order to confirm the derepression of *UCL1* in the natural condition, allele specific gene expression analysis using dCAPS was carried out. However, there was no significant misregulation of *UCL1*. Partially penetrated genes sometimes do not exhibit a phenotype in the heterozygous background. Since *efc-1* and *UCL1* were heterozygous in these experiments, they may have affected the phenotype. Or, this result may suggest that *EFC* does not regulate all the genes that are repressed by FIS2-PRC2.

The current results of the complementation experiment suggests that EMF2 might compete with FIS2 in the endosperm development, because when EMF2 is

ectopically expressed in the WT seed, it caused seed abortion. Because I could not observe no significant complementation in the *FIS2::EMF2; fis2-11; efc-1* heterozygous plants, the complementation experiment of *EMF2* with *fis2-11* in the absence of *EFC* should be done in the *efc-1; fis2-11; FIS2::EMF2* triple homozygous plant or at least, *efc-1* homozygous; *fis2-11* heterozygous; *FIS2::EMF2* homozygous plants. As mentioned above, it maybe because heterozygous *efc-1* might not have any affect in the seed.

EFC encodes a ~16 kDa small protein that has no known domain. So far, by using the online database of genomic sequencing of different genus of plants, I could find only 4 orthologues of *EFC* (At4g23110) in *Arabydopsis lyrata* (ARALYDRAFT_914290), *Capsella rubella* (Carubv10007086m), *Boechera stricta* (Bostr.7867s0100) and *Capsella grandiflora* (Cagra.3145s0009.1) (Figure 13 and 14). It is difficult to point out the exact function of *EFC* due to its incomplete or partial penetrance, however, these results may suggest that *EFC* has evolved recently in order to cooperate with the FIS2-PRC2 complex and help the complex to function more sophisticatedly.

Figure 13

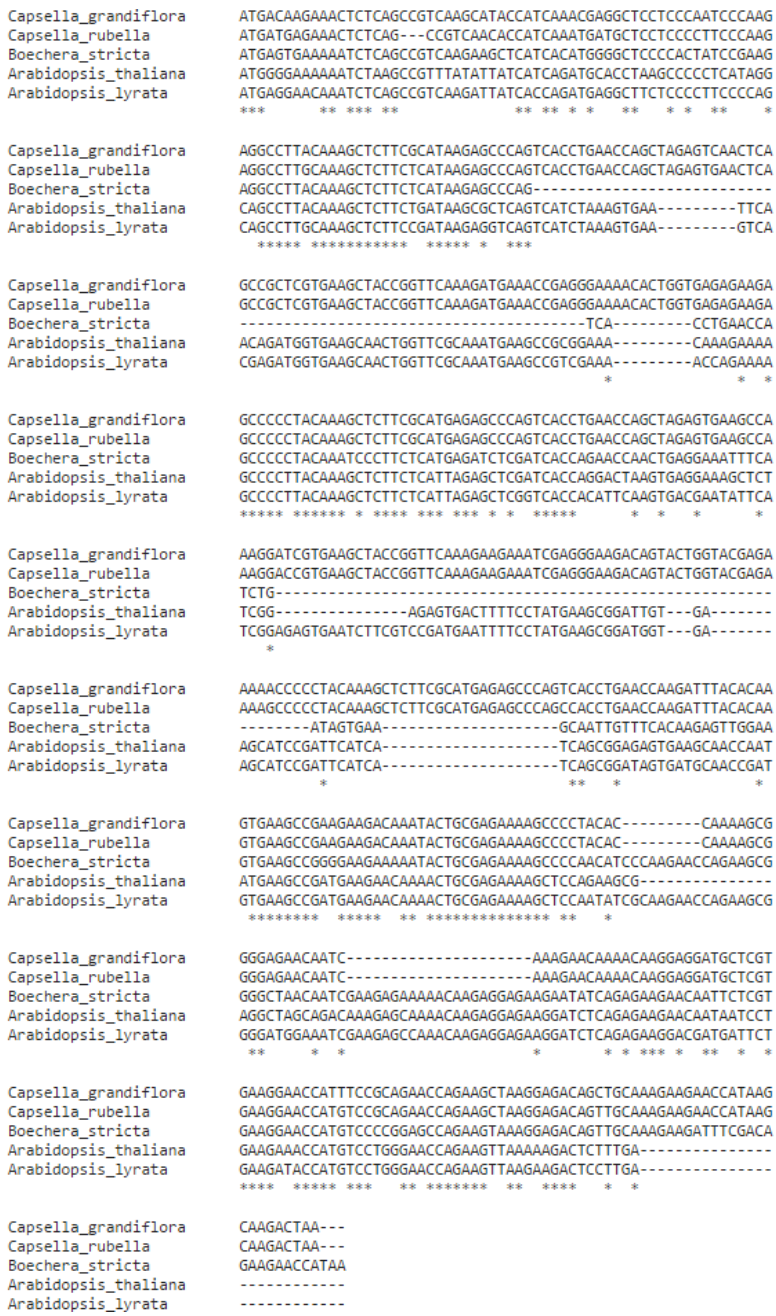


Figure 13. Comparison of genomic sequence homology

Figure 14

Figure 14. Comparison of amino-acid sequence homology

V. References

Calonje M, Sanchez R, Chen L, Sung ZR (2008) EMBRYONIC FLOWER1 participates in polycomb group-mediated AG gene silencing in Arabidopsis. Plant Cell **20:** 1–15

Casson SA, Lindsey K (2006) The turnip Mutant of Arabidopsis Reveals That LEAFY COTYLEDON1 Expression Mediates the Effects of Auxin and Sugars to Promote Embryonic Cell Identity. Plant Physiol **142(2):** 526-541

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

Chen X, Goodwin SM, Liu X, Chen X, Bressan RA, Jenks MA (2005) Mutation of the RESURRECTION1 Locus of Arabidopsis Reveals an Association of Cuticular Wax with Embryo Development. Plant Physiol **139(2):** 909-919.

Coote JG (1972) Sporulation in Bacillus subtilis. Characterization of oligosporogenous mutants and comparison of their phenotypes with those of asporogenous mutants. J Gen Microbiol **71:** 1-15

Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V (2002) Drosophila Enhancer of Zeste/ESC complexes have a histone H3

methyltransferase activity that marks chromosomal polycomb sites. Cell **111:** 185–196

Del Olmo I, López J, Vázquez J, Raynaud C, Piñeiro M, Jarillo J (2016) Arabidopsis DNA polymerase ϵ recruits components of Polycomb repressor complex to mediate epigenetic gene silencing. Nucleic Acids Res **44:** 5597-5614

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 USA **105:** 16831–16836

Derkacheva M, Hennig L (2013) Variations on a theme: Polycomb group proteins in plants. J Exp Bot **65:** 2769-2784

Dumbliauskas E, Lechner E, Jaciubek M, Berr A, Pazhouhandeh M, Alioua M, Cognat V, Brukhin V, Koncz C, Grossniklaus U, Molinier J, Genschik P (2011) The Arabidopsis CUL4–DDB1 complex interacts with MSI1 and is required to maintain MEDEA parental imprinting. EMBO J **30:** 731–743

Figueiredo DD, Batista RA, Roszak PJ, Kohler C (2015) Auxin production couples endosperm development to fertilization. Nat Plants **1:**15184

Gendall AR, Levy YY, Wilson A, Dean C (2001) The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in Arabidopsis. Cell **107:** 525–535

Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB (1998) Maternal control of embryogenesis by MEDEA, a Polycomb group gene in Arabidopsis. Science **280:** 446-450

Guitton AE, Page DR, Chambrier P, Lionnet C, Faure JE, Grossniklaus U, Berger F (2004) Identification of new members of FERTILISATION INDEPENDENT SEED polycomb group pathway involved in the control of seed development in Arabidopsis thaliana. Dev **131:** 2971–2981

Horvitz HR, Sulston JE (1980) Isolation and genetic characterization of cell-lineage mutants of the nematode Caenorhabditis elegans. **Genetics 96:** 435-454

Jeong CW, Roh H, Dang TV, Choi YD, Fischer RL, Lee JS, Choi Y (2011) An E3 ligase complex regulates SET-domain polycomb group protein activity in Arabidopsis thaliana. Proc Natl Acad Sci USA **108:** 8036–8041

Jeong CW, Park GT, Yun H, Hsieh T-F, Choi YD, Choi Y, Lee JS (2015) Control of paternally expressed imprinted UPWARD CURLY LEAF1, a gene encoding an F-Box protein that regulates CURLY LEAF polycomb protein, in the arabidopsis endosperm. PLoS One **10(2):** e0117431. doi:10.1371/journal.pone.0117431

Jullien PE, Katz A, Oliva M, Ohad N, Berger F (2006) Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in Arabidopsis. Curr Biol **16:** 486-492

Kang IH, Steffen JG, Portereiko MF, Lloyd A, Drews GN (2008) The AGL62 MADS domain protein regulates cellularization during endosperm development in Arabidopsis. Plant Cell **20:** 635-647

Kiyosue T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, Fischer RL, Kakutani T (1999) Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. Proc Natl Acad Sci USA **96:** 4186–4191

Kim MD (2006) Characterization of an early flowering mutant of Arabidopsis isolated by activation tagging mutagenesis. Seoul National University, Master's thesis

Köhler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U (2003) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. Genes Dev **17:** 1540–1553

Köhler C, Hennig L, Bouveret R, Gheyselinck J, Grossniklaus U, Gruissem W (2003) Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. EMBO J **22:** 4804–4814

Luichtl M, Fiesselmann B, Matthes M, Yang X, Peis O, Brunner A, Torres-Ruiz R (2013) Mutations in the Arabidopsis RPK1 gene uncouple cotyledon anlagen and primordia by modulating epidermal cell shape and polarity. Biol Open **2:** 1093-1102

Luo M, Bilodeau P, Koltunow A, Dennis ES, Peacock WJ, Chaudhury AM (1999) Genes controlling fertilization-independent seed development in Arabidopsis thaliana. Proc Natl Acad Sci USA **96:** 296–301

Mao G, Wang R, Guan Y, Liu Y, Zhang S (2011) Sulfurtransferases 1 and 2 Play Essential Roles in Embryo and Seed Development in Arabidopsis thaliana. J Biol Chem **286(9):** 7548-7557

Mozgova I, Köhler C, Hennig L (2015) Keeping the gate closed: functions of the polycomb repressive complex PRC2 in development. Plant J **83:** 121-132

Mozgova I, Hennig L (2015) The polycomb group protein regulatory network. Annu Rev Plant Biol **66:** 269-296

Nekrasov M, Wild B, Müller J (2005) Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. EMBO Rep **6(4):** 348--353.

Ohad N, Yadegari R, Margossian L, Hannon M, Michaeli D (1999) Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. Plant Cell **11:** 407–415

Pazhouhandeh M, Molinier J, Berr A, Genschik P (2011) MSI4/FVE interacts with CUL4-DDB1 and a PRC2-like complex to control epigenetic regulation of flowering time in Arabidopsis. Proc Natl Acad Sci USA **108:** 3430–3435

Pien S, Grossniklaus U (2007) Polycomb group and trithorax group proteins in Arabidopsis. Biochim Biophys Acta **1769:** 375-382

Queitsch C, Sangster TA, Lindquist S (2002) Hsp90 as a capacitor of phenotypic variation. Nature **417:** 618-624

Roh HM (2013) Regulation of development and growth mediated by BIA1 and EFC in Arabidopsis, Seoul National University, Doctor's thesis

Roszak P, Köhler C (2011) Polycomb group proteins are required to couple seed coat initiation to fertilization. Proc Natl Acad Sci USA **108:** 20826-31

Sangster TA, Salathia N, Undurraga S, Milo R, Schellenberg K, Lindquist S, Queitsch C (2008) HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. Proc Natl Acad Sci USA **105:** 2963-2968

Sørensen MB, Chaudhury AM, Robert H, Bancharel E, Berger F (2001) Polycomb group genes control pattern formation in plant seed. Curr Biol **11:** 277-281

Sung ZR, Belachew A, Shunong B, Bertrand-Garcia R (1992) EMF an Arabidopsis gene required for vegetative shoot development. Science **258**, 1645–1647

Xiao J, Wagner D (2015) Polycomb repression in the regulation of growth and development in Arabidopsis. Curr Opin Plant Biol **23:** 15–24

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

Zhang X, Germann S, Blus BJ, Khorasanizadeh S, Gaudin V, Jacobsen SE (2007) The Arabidopsis LHP1 protein colocalizes with histone H3 Lys27 trimethylation. Nat Struct Mol Biol **14:** 869–871

국문초록

애기장대의 FIS2-PRC2 복합체에서 *EARLY FLOWERING AND CURLY LEAVES* (*EFC*) 유전자의 역할에 관한 연구

Polycomb Repressive Complex2 (PRC2)는 식물과 동물에서 모두 알려져 있는 주요한 후성 조절인자이다. 이 복합체는 특정한 발달 단계에서 필요하지 않은 유전자의 히스톤 H3 단백질 라이신 27에서 trimethylation (H3K27me3)이 일어나도록 하고, 그 결과로 유전자의 전사 활성이 억제된다.

선행 연구에서는 활성 표지 돌연변이 선별을 통해 기능획득돌연변이인 *early flowering and curly leaves-D* (*efc-D*)가 발견되었다. *EFC*가 과발현 되었을 시 PRC2의 표적 유전자인 *FLOWERING LOCUS C* (*FLC*)와 *FLOWERING LOCUS T* (*FT*) 그리고 *AGAMOUS* (*AG*)의 발현 정도가 증가하였다. 애기장대의 세 가지 PRC2 복합체 중 하나인 FERTILIZATION INDEPENDENT SEED2-PRC2 (FIS2-PRC2 복합체는 중심세포와 배유에서의 유전자 조절을 통해 배유 발달에 필수적인 역할을 수행한다. 흥미롭게도 *EFC*는 중심세포와 배유의 핵에서 특이적으로 발현되며, PRC2의 핵심 구성요소인 MSI1과 상호작용하는 것이 관찰되었다.

본 연구에서는 기능손실돌연변이인 *efc-1*을 획득하였다. 이를 통하여 *efc-1* 돌연변이가 애기장대에서의 종자 낙태를 일으킴을 확인하였으며, 부분적 또는 불완전한 표현율(penetrance)를 보였다. *efc-1* 돌연변이체의 중심세포와 배유를 관찰한 결과 FIS-PRC2에 의해 조절되는 유전자인 *MEDEA* (*MEA*)와 *UPWARD CURLY LEAF1* (*UCL1*)의 발현이 감소됨을 확인할 수 있었다. 또한 *EFC*가 결핍된 종자에서 *fis2*와 *FIS2*의 상동체(homolog)인 *EMBRYONIC FLOWER2* (*EMF2*)의 상보성 실험을 진행하였다. 그 결과 heterozygous *efc-1*; heterozygous *fis2-11* 그리고 hemizygous *FIS2::EMF2* 를 지닌 삼중 돌연변이 식물체에서 유의미한 상보성은 관찰되지 않았으며, 이는 *EFC*의 부분적 표현율(partial penetrance) 때문일 것이라 짐작된다. 따라서 다음 세대에서 삼중 homozygous 돌연변이체 또는 이중 homozygous 돌연변이체와 heterozygous *fis2-11* 지닌 돌연변이체를 획득하여 추가적인 실험을 진행할 예정이다. 현재까지의 진행 결과는 *EFC*가 FIS2-PRC2 복합체와 함께 종자 발달에 역할을 수행하지만 필수적이지는 않다는 것을 암시한다.

주요어: *EMBRONIC FLOWER AND CURLY LEAF* (*EFC*), polycomb repressive complex, 종자 발달, 배유, 기능 손실, 애기장대

학번: 2014-25243

Acknowledgements

First and foremost, I thank my academic advisor, Professor Jong Seob Lee, from the bottom of my heart, for his patient guidance, enthusiastic encouragement and useful critiques of this research work. I would also like to thank my committee members Professor Yeonhee Choi, Professor Ji-young Lee for their encouragement and insightful comments.

I would like to take this opportunity to thank the Government of the Republic of Korea for making my master's degree possible through the scholarship given by NIIED. Also, I was honored to be one of the recipients of the Stadelmann-Lee Scholarship. Thanks to your generous support I could successfully finish my degree.

A very special thanks goes out to Jooyeon Hong and Seoyoung Jeong, the members of plant functional genomics laboratory, for have been a source of friendships as well as good advice and collaboration. I am especially grateful for Geun Tae Park, Younsook Min, Kyunghyuk Park, Seohyun Kim, Jin-sup Park, Goh Choi, Hyojin Kim and Hyunjin Yoo, who were always there to help whenever I was in need.

I am extremely thankful to my friends, Ana Meirelles, Mariana Trifonova, Edwin Jurado, Michael Thompson, Ramziya Bikchentaeva, Khulan Lkhamsuren, Lyudmila Zubkova and Jorge Carpanetti for your endless support and beautiful friendship.

Finally, but by no means least, I would like to thank my father Delgerekh Dorj, my mother Tsetsegmaa Lkhagvadorj and my younger sister Tuvshin Delgerekh for their almost unbelievable support and love. They are the most important people in my world and I dedicate this thesis to them.