creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

이학석사학위논문

애기장대에서 DEMETER와 상호작용하는 인자의 선별

Screening for Interaction Partners of DEMETER in *Arabidopsis thaliana*

2019년 8월

서울대학교 대학원

생명과학부

김 효 진

ABSTRACT

Screening for Interaction Partners of DEMETER in *Arabidopsis thaliana*

Hyojin Gim

School of biological science

The Graduate School

Seoul National University

DNA methylation and demethylation plays significant roles in regulating gene expression on variable stress reactions and TEs silencing. A DEMETER (DME), an active DNA demethylase in plants, is expressed in the central cell of the female gametophyte in *Arabidopsis* required for seed development and has glycosylase activity that can actively remove 5-mC that is replaced by cytosine via base excision pathway. DME induces maternal allele expression of the imprinted MEDEA (MEA) polycomb gene and the DNA glycosylase activity of DME leads to the DNA demethylation on its targets. In plants, there is no massive methylation reprogramming in embryo like mammal but instead, the companion cells whose DNA contents are not inherited to the next generation go through global demethylation and the hypomethylated state is maintained in late endosperm stage in which *DME* is no longer expressed.

Despite all these distinct significances, little is known about the interactors that associate with DME. Here, to identify the interacting partners of DME, Ⅰ used Bimolecular Fluorescence Complementation (BiFC). From 83 genes that have been confirmed by the Yeast Two-Hybrid system that interact with DME, I primarily chose 18 candidates to test interaction. While examining these 18 candidates by BiFC, AT5G37930, AT5G60980, AT1G70620, AT5G23090 and the C terminal part of AT1G20960 showed fluorescent signals when it was co-transfected with DME in *Arabidopsis* Col-0 protoplasts. These interactors contain either E3-ubiquitin ligase activity, RNA binding motif, homologous feature of transcription factor which is related to histone acetylation, or are related to RNA splicing.

To further understand its relation with DME in plants one candidate, At1g20960 that is related to RNA splicing, was chosen and its mutant was crossed with *dme* mutant for phenotyping analyses. *dme* mutant allele transmission, segregation and seed abortion ratio shown in *dme* single mutant were not changed in double mutants. Therefore, by doing further experiments, using different candidates, this study would give some specific and clear perspectives and contribute to widen the knowledge by identifying a novel protein involved in the DME demethylation pathway in plants.

Keywords : DEMETER(DME), DNA demethylation pathway, *Arabidopsis*, Epigenetics, BiFC

Student Number : 2017-20959

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

BACKGROUND

1. DNA methylation in Arabidopsis

Epigenetics is a study of heritable but reversible phenotypic changes in organisms caused by modification of gene expression rather than alteration of the genetic code itself and its inheritance. DNA methylation is a major research topic to understand epigenetic phenomena along with histone modifications. DNA methylation is one of the most well-conserved epigenetic makers and has significance in development and stress reaction of plant and animal (Reik et al., 2001; Law and Jacobsen, 2010; Wu and Zhang, 2010; Zhang and Zhu, 2012). Unlike many other popular model organisms, *Arabidopsis* has retained a multi-layered methylation system that contributes to gene and transposon silencing, imprinting, and genome stability and many of the findings are applicable to other eukaryotes (Gehring and Henikoff, 2008). Cytosine can be methylated at the carbon five position, and in plants this can occur on any cytosine regardless of the sequence context (CG; CHG; CHH) (Gehring and Henikoff, 2008). In general, 5-methylcytosine (5mC) is associated with transcriptional silencing.

The *Arabidopsis* genome contains methylation at 24% of CG sites, 6.7% of CHG and 1.7% of CHH (Cokus et al., 2008) and DNA methylation tends to prefer being located at repetitive DNA sequences. The tendency of methylation prefering repetitive DNA sequences suggests that one of methylation's primary functions is to silence the transcription of transposable elements(TEs) (Zilberman and Henikoff, 2004; Gehring and Henikoff, 2007). In *Arabidopsis*, transposons are generally methylated throughout their length at cytosines in all sequence contexts, although distinct patterns do emerge at individual loci (Lippman et al., 2003; Lippman et al., 2004; Zhang et al., 2006; Zilberman et al., 2007). TEs invade genomes and increase in copy number, with strong potential for damaging the host. All organisms have adopted mechanisms to keep TEs silent, including RNA-based chromatin silencing, histone modifications, DNA methylation, or a combination thereof (Slotkin and Martienssen, 2007).

DNA methyltransferase is an enzyme that donates methyl group to the carbon 5 position of cytosine and all known cytosine 5-methyltransferases belong to a single family with several subfamilies (Gehring and Henikoff, 2008). There are three subfamilies of DNA methyltransferases in *Arabidopsis*: CG maintenance methyltransferases that have a role in maintaining 5-mC of CG methylation context, chromomethylases that involved in maintaining 5mC of CHG context, and the *de novo* methyltransferases that have to construct a new 5mC in CHH context since there is no complementary 5-mC to be used as a template for maintaining the methylation. Multiple genes exist for each enzyme class, but only one enzyme of each type appears to be active; the other genes are either expressed at low levels or contain stop codons in various backgrounds, and none have been recovered in mutant screens (Gehring and Henikoff, 2008).

METHYLTRANSFERASE1 (MET1) is the CG maintenance methyltransferase in *Arabidopsis*. This designation is based on sequence similarity to Dnmt1, the orthologous mammalian maintenance methyltransferase, and on the effect mutations in the gene have on DNA methylation (Finnegan and Kovac, 2000). More than half of the regions that are methylated in wild type are lost in *met1*. At some repetitive sequences significant amounts of CHG and CHH methylation are also lost (Gehring and Henikoff, 2008). Many TEs become transcriptionally active in a *met1* mutant (Kato et al., 2003; Lippman et al., 2003; Zhang et al., 2006; Zilberman et al., 2007). Mutations in *MET1* and antisense-directed *MET1* silencing cause various phenotypes from gametogenesis onwards since MET1 is essential to maintain methylation patterns during the haploid gametophyte stage of the plant life cycle (Saze et al., 2003). In the sporophyte, *met1* phenotypes include abnormal embryo patterning, narrow leaves, homeotic transformations of floral organs, altered flowering time, and reduced fertility (Finnegan et al., 1996; Kankel et al., 2003; Saze et al., 2003; Xiao et al., 2006; Mathieu et al., 2007). MET1 might also have *de novo* methyltransferase activity since *de novo* methylation of CG sites is impaired in met1 mutants (Aufsatz et al., 2004).

CMT3 (CHROMOMETHYLASE3) is another methyltransferase unique to plants containing a

chromodomain which makes it possible to bind to methylated lysines in histone tails (Henikoff and Comai, 1998). CMT3 maintains methylation in the CHG sequence context. *cmt3* single mutants do not have any morphological phenotypes, but shows severe defects when combined with a null *met1* allele (Xiao et al., 2006; Zhang and Jacobsen, 2006). The additional loss of CHG methylation in *met1 cmt3* mutants might push the genome over a methylation threshold such that the remaining methylation, in whatever sequence context, is insufficient to accomplish its functions (Gehring and Henikoff, 2008).

DRMs (DOMAINS REARRANGED METHYLTRANSFERASES) were identified as *de novo* methyltransferases based on homology to the mammalian *de novo* methyltransferases Dnmt3a and Dnmt3b (Cao and Jacobsen, 2002). DRM2 appears to be the only functional enzyme in *Arabidopsis*. CHH methylation must be maintained after DNA replication in a *de novo* manner since it is not symmetric between complementary DNA strands. DRM2 is required for establishing methylation at all loci examined and for maintenance, it has locus-specificity on asymmetric methylation (CHH). In in vitro assays, tobacco DRM preferentially methylates CHH and CHG sites, with far less activity at CG sites, and prefers unmethylated DNA over hemimethylated DNA (Wada et al., 2003). Loss of *drm2* and *cmt3* has little overall effect on the distribution of methylation genome-wide (Zhang et al., 2006) since non-CG methylation is always found in the vicinity of CG methylation and CG methylation is the most abundant context for cytosine methylation. Also, *drm2* single mutant does not have significant morphological phenotypes but *drm2 cmt3* mutant plants shows multiple phenotypes including small size, twisted leaves, and reduced fertility (Cao and Jacobsen, 2002). Additionally, *drm2 cmt3* mutants retain some CHG and CHH methylation, particularly in pericentromeric heterochromatin (Cokus et al., 2008). This suggests that another enzyme(s) like MET1 as speculated above *has de novo* methylation activity.

2. DNA demethylation and DEMETER

DNA demethylation can be achieved by two different manners: Passive and active demethylation. Passive demethylation does not require some enzymes to play a role in elimination of methyl group from 5-mC. Instead, DNA methylation gets diluted by going through multiple DNA replications in an absence or reduction of the influence of methyltransferases like MET1 which have critical role in maintaining DNA methylation (Zhu, 2009). In the pollens of flowering plants TEs were found to be unexpectedly reactivated only in the vegetative cell, which does not deliver DNA to the fertilized zygote (Slotkin et al. 2009). In *Arabidopsis*, reduced expression of the RNA-directed DNA methylation (RdDM) pathway components during male gametogenesis results into passive DNA demethylation in the vegetative cell (Slotkin et al. 2009). Similarly, during female gametophyte development passive DNA demethylation may also occur in the central cell which becomes endosperm and does not provide DNA to the next generation (Li et al., 2018). Transcriptional repression of MET1 was found to be associated with genome-wide DNA demethylation in the central cell (Jullien et al., 2008). However, results from a recent study argue against decreased MET1 expression in the central cell (Park et al., 2016), making the involvement of passive DNA demethylation in female gametogenesis controversial.

Unlike a passive DNA demethylation lowering methylated DNA level in a replication-dependent manner, in active DNA demethylation some enzymes do the role to remove 5-mC that is replaced by unmethylated cytosine. In plant, HhH-GPD (helix-hairpin-helix-Gly/Pro/Asp) DNA glycosylases play a role as DNA demethylases using the base excision repair (BER) (Gehring and Henikoff, 2008). In *Arabidopsis*, DEMETER (DME); REPRESSOR OF SILENCING1 (ROS1); DME LIKE2 (DML2); DME LIKE3 (DML3), that have both glycosylase and AP lyase activities can directly cleave the glycosidic bond between the base and sugarphosphate backbone of the 5-mC then the DNA polymerase fill in the abasic site (AP site) with unmethylated cytosine (Fig. 1) (Penterman et al., 2007; Zhu 2009).

DME is expressed primarily in the central cell of the female gametophyte and highly involved in the early female gametophyte development and seed development (Choi et al., 2002). The significance of DME in female gametophyte development and seed development will be discussed later in CHAPTER TWO.

As an active DNA demethylase, DME cannot cleave the unmethylated cytosine but the 5-mC (Morales-Ruiz et al., 2006) in all context but the activity is most efficient in 5meCG whether the DNA strands are both methylated or hemi-methylated (Gehring et al., 2006; Morales-Ruiz et al., 2006). DME has been the most extensively biochemically characterized *in vitro* along with ROS1, although the details from different groups are not always identical. While HhH-GPD DNA glycosylases are the largest class of glycosylases among all organisms, the DME family appears to be unique to the plant lineage (Gehring and Henikoff, 2008). HhH-GPD DNA glycosylases are characterized by a conserved aspartic acid residue and an invariant lysine. When either of these residues is mutated in recombinant DME, 5-mC DNA glycosylase activity is lost (Gehring et al., 2006; Morales-Ruiz et al., 2006).

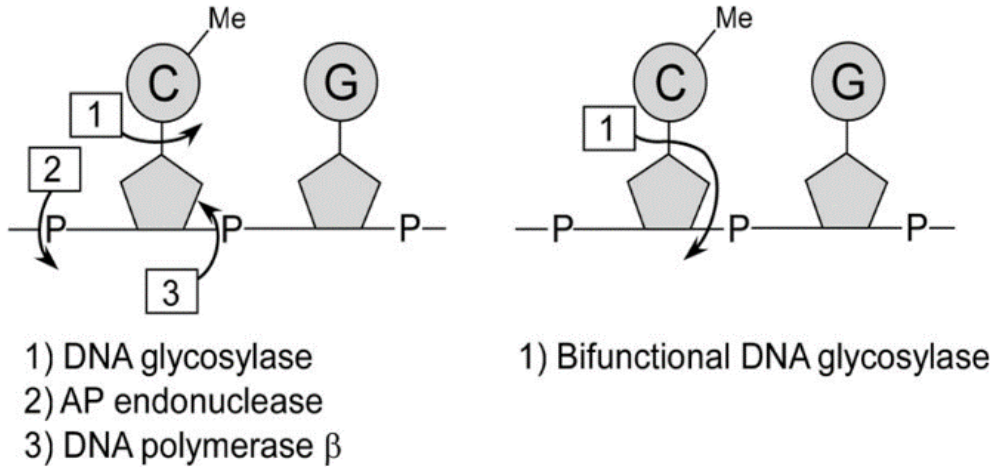


Figure 1. Schematized active CG demethylation by DNA glycosylase (Kress et al., 2006).

To eliminate 5-mC by using base excision repair, the DNA glycosylase activity that can cleave the bond between base and sugarphosphate and the AP nuclease activity that can remove the sugarphosphate are necessary. In *Arabidopsis* the enzymes that works as DNA demethylases have DNA glycosylase and AP lyase activity at the same time. The AP site that has been emerged by the bi-functional enzyme activity then repaired by DNA polymerase so the 5-mC is replaced with cytosine.

3. Purpose of this study

In *Arabidopsis*, the studies related to DNA methylation and demethylation have been focused on the enzymes and their regulations in molecular and organismal levels or the phenotypic analysis related to their genetic and epigenetic consequences. DNA methylation and demethylation play a key role in epigenetic regulation on gene expressions through their reversibility and can keep the organisms or species from being severely defected by TEs despite going through alternation of generations by silencing them effectively. Especially, in gametogenesis and onward developmental process, there is a possibility that hypomethylation in companion cells like central cell or vegetative cell whose DNA contents are not delivered to the next generation can reinforce the silencing of the egg cell and sperm cells which become a zygote after fertilization and deliver their DNA contents to the next generation. In this way, plants do not have to take a risk of being damaged by having massive global demethylation in egg cell or sperm cells and go through a safer way to achieve TE silencing.

Despite all these significances that DNA methylation and demethylation have had, physical or direct interactors of have not been studied. Therfore, in this study, I will focus on uncovering the physical interactors of DME so that it can contribute to understand more about the active DNA demethylation pathway of DME. To do so, I used BiFC assay to *in vivo* screen the interactors of DME from the candidates selected by yeast two-hybrid (Y2H) screen and observed the interactors' influences on *dme* mutant by using double heterozygous mutants. Doing this research can ultimately give a cue to understand the epigenetic significance and value started from the gametogenesis and seed development in plant more clearly.

CHAPTER ONE:

Interaction partner screening for DEMETER using Bimolecular Fluorescence Complementation(BiFC) Assay

I. Introduction

BiFC assay

BiFC (Bimolecular Fluorescence Complementation) is an effective experimental tool that can visualize protein-protein interactions in living cells. In BiFC, fragmented fluorescent proteins attach to each of the proteins of interest recover their original 3-dimensional conformation and shows fluorescent signal if the two proteins physically interact (Fig.2). In *E. coli,* fluorescence complementation was first detected using fragments of a green fluorescent protein (GFP) variant fused to artificial peptide sequences designed to form an anti-parallel coiled coil (Ghosh, et al., 2000). In tobacco, onion and *Arabidopsis thaliana,* interactions between many different types of proteins have been visualized by introducing expression vectors encoding the fusion proteins using *Agrobacterium* infiltration or particle bombardment (Kerppola, 2006)

The BiFC approach provides several unique advantages for the investigation of molecular complexes in living cells. Examining the interaction in living cells make it possible to observe the interaction in a very similar context of the actual interaction conditions (Kerppola, 2006). Also it gives an easier way to observe the changes followed by varying the given interaction conditions. The BiFC approach is applicable for the visualization of a wide range of molecular interactions proteins from different structural classes in a variety of cell types and species. And as the interaction is visualized, the cellular localization can also be detected. In addition, the intensity of fluorescent signal indicates the intensity of protein-protein interaction itself so the interactions can be compared qualitatively (Kerppola, 2006).

Based on these he advantages, BiFC-based screening is a powerful tool in that the interactions can be detected within the cell, and the effects of stimuli on the interaction can be directly tested. One limitation of BiFC-based screens however is that differences in protein expression levels are likely to influence the partners that can be identified (Keppola, 2006). Nevertheless, BiFC analysis has the potential to identify partners that interact with a protein of interest under specific

cellular conditions. BiFC analysis can also be used to identify synthetic molecules or cellular factors that can modulate protein interactions.

In this study, I used Enhanced yellow fluorescent (EYFP) protein fragments to make fusion proteins. *DME* was fused at the 3' end of the N terminal fragment of EYFP and the genes of interest were fused at the 3' end of the C terminal fragments of EYFP each by using pSAT4 vector system (Fig. 2). To verify the interaction clearly, I used nEYFP-SSRP1N with cEYFP-SPT16C and nEYFP-cDME with cEYFP-SPT16C as positive controls that were previously reported as interactors for transfection efficiency and used confirmed non-interactors of DME which is constructed in a pSAT4-cEYFP vector as negative controls.

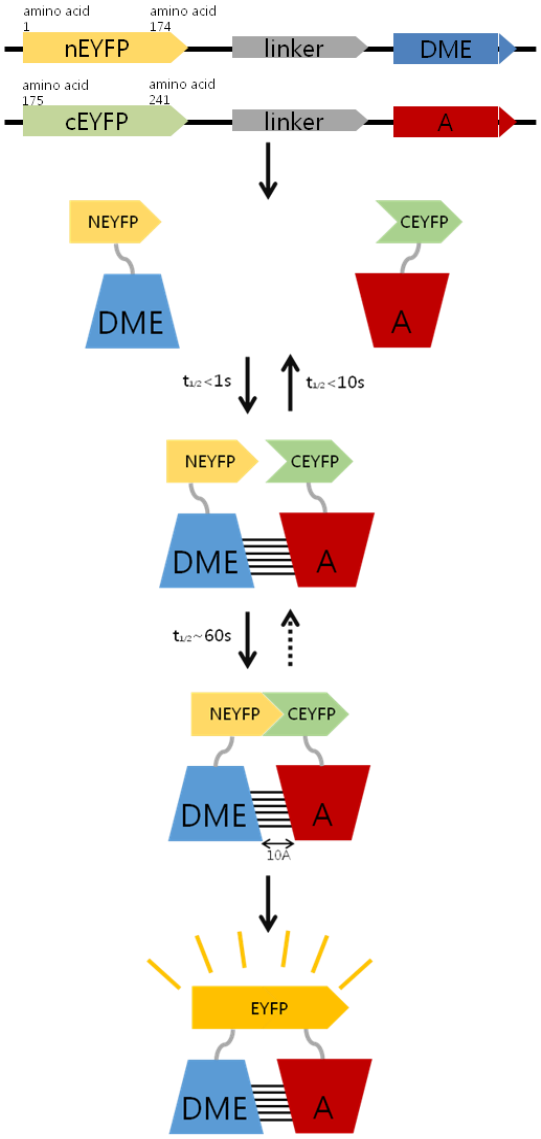


Figure 2. Schematic description of BiFC assay in this study.

The blue boxes indicate either CDS of *DME* of full DME proteins and the red ones indicates the gene of interest or the protein. If the nEYFP-DME fusion protein and cEYFP-*candidat* fusion protein get close enough for the EYFP fragments to resume their original 3D structure because of the physical interaction of DME and the candidate, EYFP signal would be detected by confocal laser scanning microscopy. The entire BiFC reaction to emit the EYFP signal needs at least 6 hours to be expressed.

II. Materials and methods

Plant Material and Growth Conditions

Arabidopsis thaliana Columbia-0 (Col-0) was used as the wild type. Plants were grown in an environmentally-controlled chamber with a long photoperiod (16 hr light and 8 hr dark) at 22℃. The transgenic lines and plasmids were obtained from the Arabidopsis Biological Resource Center (ABRC, Columbus, OH)

Previous yeast two-hybrid data

The yeast two-hybrid (Y2H) library screening using yeast mating (Matchmaker™ Gold Yeast Two-Hybrid System User Manual) was performed to identify physical interactors of partial fragment of DME. A cDNA library which was constructed in an activation domain (AD) vector of GAL4 was made from mRNA extracted from inflorescence meristem of *Arabidopsis* Columbia (col-0) wild type plant and the construct was mated with 667 amino acid-long N terminal fragment of DME(AT5G04560.1) of function is unknown cloned in GAL4 DNA-binding domain (BD). After plating the culture on DDO at 30℃ for 6 to 12 days, the selected colonies tested on QDO plate and the plasmids were extracted and sequenced.

Through these previous experiment, the mating efficiency was about 10% and 239 clones were screened from QDO. After the sequencing 83 genes were found to be physical interactors of DME (Table 1).

Table 1. List of genes showed physical interaction with the N terminal fragment of DME(AT5G04560.1) confirmed with yeast two-hybrid screening

	Gene	repetition	cloned cDNA length (bp)	
1	AT1G07890	15	216	
2	Intergenic (chr2:3337337-3339906)	10	-	
3	AT5G37930	9	982	
4	AT5G40450	7	996	
5	AT5G44510	7	558	
6	AT3G23640	6	1074	
7	AT1G70600	6	708	
8	AT4G00895	4	773	
9	AT3G20015	4	581	
10	AT2G39720	3	1068	
11	AT1G54040	3	477	
12	AT2G35790	3	877	
13	AT1G68875	3	435	
14	AT1G20960	2	553	
15	Intergenic (Chloroplast genome)	2	-	
16	AT5G45775	2	429	
17	AT4G34350	2	617	
18	AT1G01300	2	284	
19	AT5G60980	2	282	
20	AT1G01550	2	1101	
21	AT4G10850	2	804	
22	AT1G70550	2	764	
23	AT5G58680	2	1052	
24	AT5G60670	2	288	
25	AT4G04460	2	681	
26	AT1G20440	1	123	
27	AT1G22920	1	1116	
28	AT4G02760	1	1052	
29	AT3G52150	1	835	from 5' UTR
30	AT3G53110	1	mismatch	from 5' UTR
31	AT3G23150	1	802	
32	AT1G09070	1	1043	out frame
33	AT2G05100	1	250	
34	AT1G64330	1	612	
35	ATCG01180	1	241	
36	AT4G39363	1	-	
37	AT1G52380	1	942	
38	AT1G43670	1	1127	
39	AT4G19985	1	958	
40	AT2G40510	1	247	
41	AT3G16240	1	447	
42	AT4G23990	1	297	
43	AT2G40610	1	935	
44	AT5G46340	1	1075	
45	AT2G02990	1	736	
46	AT2G22780	1	1122	
47	ATCG00080	1	111	
48	ATCG00340	1	-	
49	AT1G12050	1	583	
50	AT2G30200	1	921	
51	AT2G01710	1	905	

52	ATCG00730	1	224	
53	AT3G42800	1	1089	
54	AT3G10912	1	934	
55	AT5G51140	1	176	
56	AT3G58500	1	976	
57	AT3G13510	1	710	
58	AT3G56740	1	1052	
59	AT5G06130	1	1086	
60	AT4G18810	1	1071	
61	AT1G11910	1	703	
62	AT4G20830	1	1015	
63	AT1G07240	1	376	
64	AT1G55560	1	565	
65	AT5G16730	1	1073	
66	AT1G51060	1	637	
67	AT5G37310	1	960	
68	AT5G61900	1	1015	
69	AT1G70620	1	815	
70	Intergenic (chr2:16028595-16031261)	1	-	
71	AT3G20060	1	713	
72	AT1G05780	1	230	
73	AT3G06850	1	1081	
74	AT1G56190	1	950	
75	ATMG00020	1	1046	
76	AT1G15230	1	405	
77	AT4G34320	1	973	
78	AT1G53542	1	185	
79	AT4G20360	1	235	
80	AT5G23090	1	787	
81	AT1G70680	1	871	
82	AT2G05632	1	382	

List-up for BiFC assay

From the 83 genes confirmed as potential interactors for DME in *Arabidopsis* by Y2H assay, I listed them up to 18 candidates as a priority for a further analysis using BiFC (Table 2). These 18 candidate genes were selected by their molecular, structural, physiological validities with DME. Most of them 1) have either RNA or DNA binding motifs, 2) related to the epigenetic modification of DNA or histone proteins or 3) related to protein degradation.

Cloning for BiFC assay

For the BiFC assay, the 18 candidates and DME(AT5G04560.2) were constructed as partial EYFP-fused proteins. pSAT4-nEYFP-C1 vector was used to clone nEYFP-DME construct and pSAT4-cEYFP-C1-B vector was used to clone cEYFP-*candidate gene* constructs (Fig. 3; Fig. 4). 15~42bp-long multiple cloning site(MCS) sequence of the vector between partial EYFP sequence and inserted gene was used as a linker. All genes were amplified as their full CDS forms except for *DME* and *AT1g20960* (Table 3; Table 4) using cDNA synthesized from mRNA extracted from *Arabidopsis thaliana* col-0 inflorescence meristem as a PCR template.

Table 2. List of 18 candidate genes for further analysis using BiFC assay

	Gene	Computational description (*TAIR*)
1	AT5G37930	Protein with RING/U-box and TRAF-like domains; FUNCTIONSIN: ubiquitin-protein ligase activity, zinc ion binding; INVOLVED IN: multicellular organismal development, ubiquitin-dependent protein catabolic process, protein ubiquitination
2	AT2G39720	Encodes aputative RING-H2 finger protein RHC2a
3	AT1G54040	Epithiospecifier protein, interacts with WRKY53. Involved in pathogen resistance and leaf senescence
4	AT2G35790	unknown protein; CONTAINS InterProDOMAIN/s:Protein of unknown function DUF1301. Has 116 Blasthitsto 116 proteins in 49 species: Archae-0; Bacteria-0; Metazoa-53; Fungi-6; Plants-49; Viruses-0; Other Eukaryotes–8
5	AT1G20960	embryo defective1507(emb1507); FUNCTIONS IN: in 6 functions; INVOLVED IN: embryo development ending in seed dormancy
6	AT5G60980	Nuclear transport factor2(NTF2) family protein with RNA binding(RRM-RBD-RNPmotifs)domain; FUNCTIONSIN: RNA binding, nucleotide binding, nucleic acid binding; INVOLVED IN: transport, nucleocytoplasmic transport
7	AT1G70550	Protein of Unknown Function(DUF239); CONTAINS InterProDOMAIN/s: Protein of unknown function DUF239, plant(InterPro:IPR004314); BEST Arabidopsis thaliana protein match is: Protein of Unknown Function(DUF239)
8	AT5G58680	ARM repeat superfamily protein
9	AT3G52150	RNA-binding(RRM/RBD/RNPmotifs)family protein; FUNCTIONS IN: RNA binding, nucleotide binding, nucleic acid binding
10	AT3G53110	Encodes aputative DEAD-Box RNA Helicase and has RNA-dependent ATPase activity. Mutant is Sensitive to chilling stress and heatstress.
11	AT1G52380	NUP50(Nucleoporin50kDa)protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: intracellular transport
12	AT3G58500	PP2A-4, Encodes one of the isoforms of the catalytic subunit of protein phosphatase2A
13	AT3G56740	Ubiquitin-associated(UBA) protein
14	AT1G51060	Encodes HTA10, ahistone H2A protein
15	AT1G70620	cyclin-related; FUNCTIONS IN: molecular_function unknown
16	AT3G20060	UBC19, UBIQUITIN-CONJUGATINGENZYME19. Transcript is always found in dividing cells, but also in other non-dividing cells. Protein is localized to the cytoplasm as well as to the nucleus.
17	AT5G23090	"nuclearfactor Y, subunit B13"(NF-YB13); FUNCTIONS IN: sequence-specific DNA binding transcription factor activity
18	AT4G17245	RING/U-box superfamily protein; FUNCTIONS IN: zinc ion binding

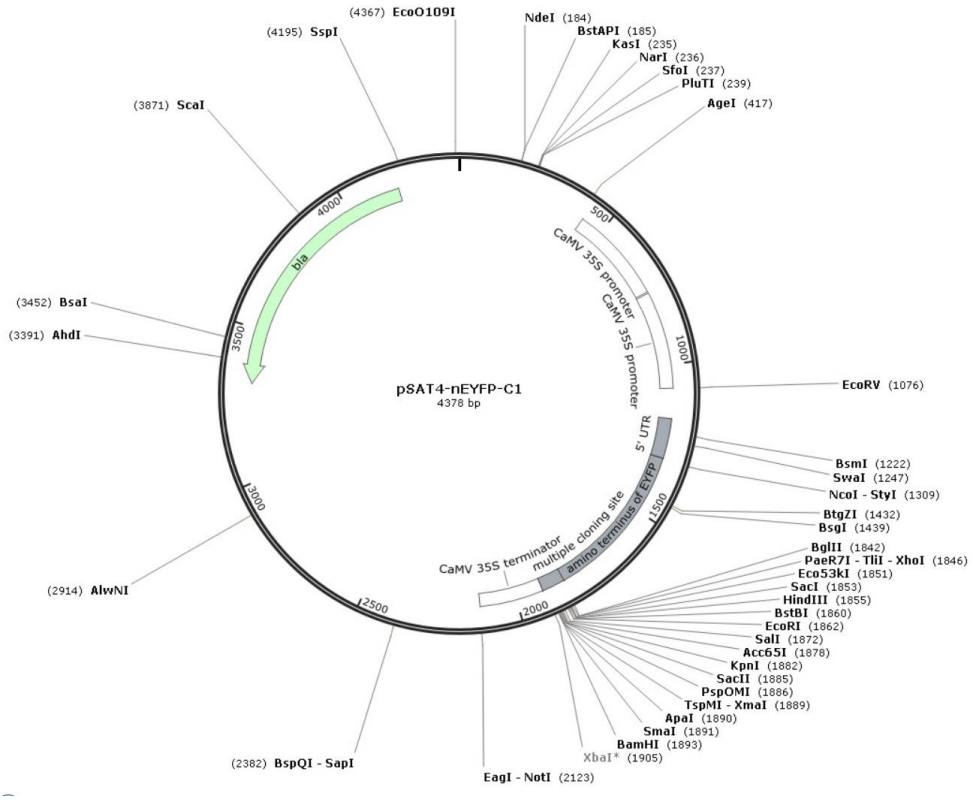


Figure 3. Map of pSAT4-nEYFP C1 vector

pSAT4-nEYFP-C1 vector has two duplicated CaMV 35S promoters and 5’ UTR in front of 174 amino acid-long N terminal fragment of EYFP. MCS lies 3’ end of the nEYFP coding sequence.

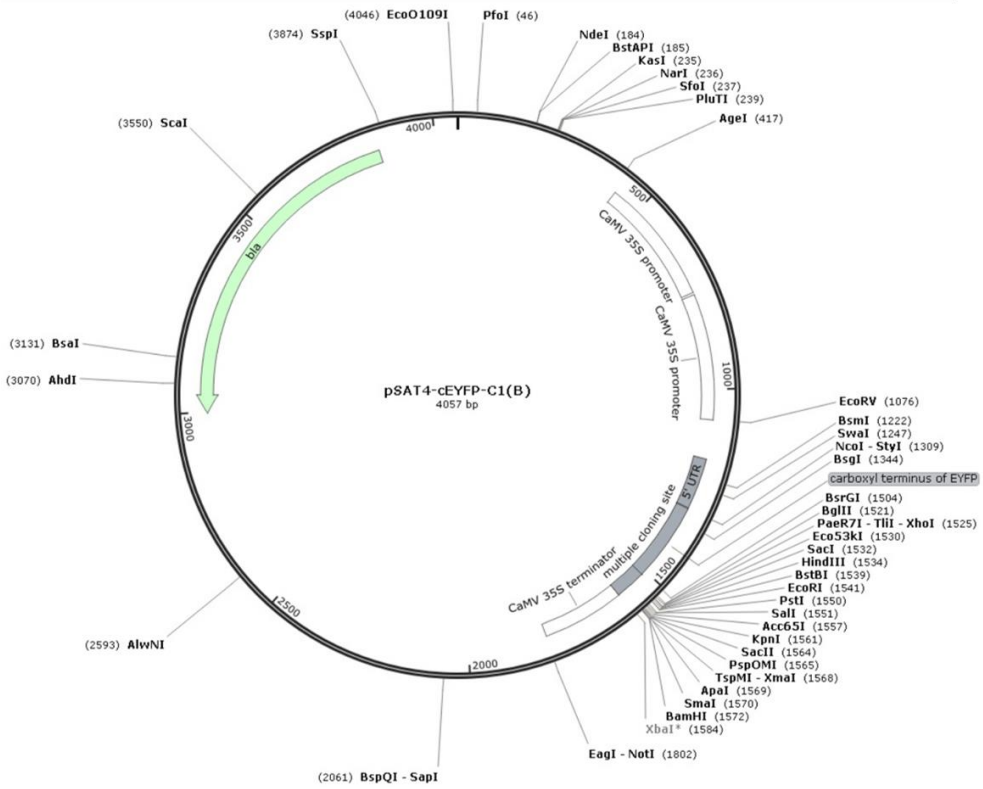


Figure 4. Map of pSAT4-cEYFP-C1-B vector

pSAT4-cEYFP-C1-B vector has two duplicated CaMV 35S promoters and 5' UTR in front of 67 amino acid-long C terminal fragment of EYFP. MCS lies 3' end of the cEYFP coding sequence.

Table 3. List of primers used for PCR amplification in constructing BiFC clones

	Primer label	Sequence (5' to 3')
1	cDME infu F	GAATTCTGCAGTCGACATGAATTCGAGGGCTGATCC
2	cDME infu R	TTTTGCGGACTCTAGATTAGGTTTTGTTGTTCTTCAA
3	berkeley cDME infu F	TTTTGCGGACTCTAGATTAGGTTTTGTTGTTCTTCAA
4	berkeley cDME infu R*	CTGCAGAATTCGAAGCGGTTTTGTTGTTCTTCAATTTGCTCG
5	AT2G35790 infu F	AATTCTGCAGTCGACATGGGAAGATCAGCTCTGAT
6	AT2G35790 infu R	TTTGCGGACTCTAGATTACAAGTTCTTGAAAGCAG
7	AT1G70550 infu F	AATTCTGCAGTCGACATGTGTTTAATAGGTTTGTT
8	AT1G70550 infu R	TTTGCGGACTCTAGATCATGGACATCTAGGGTTTT
9	AT5G37930 infu F	GACTCAGATCTCGAGggATGGCGAGATTCTCAGTT
10	AT5G37930 infu R	TGCAGAATTCTCACGAATGAACAAAGATCC
11	AT2G39720 infu F	AATTCTGCAGTCGACATGGCTTCTGGATCTTACTG
12	AT2G39720 infu R	TTTGCGGACTCTAGATCACGCTAGCCAATTTCGTC
13	AT1G54040 infu F	AATTCTGCAGTCGACATGGCTCCGACTTTGCAAGG
14	AT1G54040 infu R	TTTGCGGACTCTAGATTAAGCTGAATTGACCGCAT
15	AT1G20960C infu F	GGACTCAGATCTCGAGaaGACTTGCAACCTCTCCCAGT
16	AT1G20960 infu R	TCGACTGCAGAATTCTCATTCTTCCATGCGATCTC
17	AT5G60980 infu F	AATTCTGCAGTCGACATGGCACAGCAGGAAGCTAG
18	AT5G60980 infu R	TTTGCGGACTCTAGATCAAGATGAACCACCACCTC
19	AT5G58680 infu F	GACTCAGATCTCGAGggATGGCGAATCACAACAGTTT
20	AT5G58680 infu R	TCGACTGCAGAATTCTTATCTCTCGTTGTCGTTAG
21	AT3G53110 infu F	AATTCTGCAGTCGACATGGCGGATACGGTAGAGAA
22	AT3G53110 infu R	TTTGCGGACTCTAGATCACTCGTCCAGCAGGCCAG
23	AT1G52380 infu F	AATTCTGCAGTCGACATGGGTGACTCGGAAAACGT
24	AT1G52380 infu R	TTTGCGGACTCTAGATCAAGTATCTGTAGCTGTTG
25	AT3G56740 infu F	AATTCTGCAGTCGACATGAACGGCGGTCCCTCCGG
26	AT3G56740 infu R	TTTGCGGACTCTAGATTAGTGGGACTGTGCTTCGA
27	AT1G51060 infu F	AATTCTGCAGTCGACATGGCGGGTCGTGGTAAAAC
28	AT1G51060 infu R	TTTGCGGACTCTAGATCAATCGTCTTCAGCAGATG
29	AT3G20060 infu F	AATTCTGCAGTCGACATGGCGACGGTTAATGGGTA
30	AT3G20060 infu R	TTTGCGGACTCTAGATCATGCGTTTAAAGGCTTGT
31	AT5G23090 infu F	AATTCTGCAGTCGACATGGATCCAATGGATATAGT
32	AT5G23090 infu R	TTTGCGGACTCTAGATTAGCTTTGCGGACTTCTCT
33	AT4G17245 infu F	AATTCTGCAGTCGACATGCGGCTGCTGCTATCATC
34	AT4G17245 infu R	TTTGCGGACTCTAGACTAAGGTGTACTCGTTAGTG
35	AT3G52150 Infu F	GACGGTACCGCGGggATGGCGACTTTCCTAACAAA
36	AT3G52150 infu R	CAGGTGGATCCCGGGCTAAGCCTTATTCACCCGAA
37	AT3G58500 infu F	GACGGTACCGCGGggATGGGCGCGAATTCGCTTCC
38	AT3G58500 infu R	CAGGTGGATCCCGGGTCAAAGGAAATAGTCAGGTG
39	AT1G70620 infu F	GACGGTACCGCGGggATGGACGCGTACCAGCCACC
40	AT1G70620 infu R	CAGGTGGATCCCGGGTCAGACCAGTTTGCTAGAGT
41	AT1G20960C NLS F	GGACTCAGATCTCGAGaaAGGCAGAGTAAGAGGCGGCGTCT AAGGGAAGAGGGAGGATCCGACTTGCAACCTCTCCCAGTG
42	noncatalytic E3 N R	GGACAATCAAGAACATCTGGATCC
43	noncatalytic E3 C F	GATAGGTTATGTTCGTTGCAG

Table 4. List of the information used to construct for BiFC clones

	GENE	CDS length	used enzyme site		competent cell	vector
1	AT5G37930	1050bp	EcoRI	XhoI	DH10b	pSAT4-cEYFP-C1-B
2	AT2G39720	1206bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
3	AT1G54040	1026bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
4	AT2G35790	717bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
5	AT1G20960C*	2538bp	EcoRI	XhoI	DH10b	pSAT4-cEYFP-C1-B
6	AT5G60980	1383bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
7	AT1G70550	1398bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
8	AT5G58680	1074bp	EcoRI	XhoI	DH10b	pSAT4-cEYFP-C1-B
9	AT3G52150	762bp	SacII	XmaI	DH10b	pSAT4-cEYFP-C1-B
10	AT3G53110	1491bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
11	AT1G52380	1323bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
12	AT3G58500	942bp	SacII	XmaI	DH10b	pSAT4-cEYFP-C1-B
13	AT3G56740	882bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
14	AT1G51060	399bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
15	AT1G70620	2874bp	SacII	XmaI	DH10b	pSAT4-cEYFP-C1-B
16	AT3G20060	546bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
17	AT5G23090	480bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
18	AT4G17245	501bp	SalI	XbaI	DH10b	pSAT4-cEYFP-C1-B
19	+48 DME.1**	5190bp	HindIII	-	DH10b	pSAT4-nEYFP-C1
20	+111 DME.2***	6075bp	SalI	XbaI	DH10b	pSAT4-nEYFP-C1

*) cEYFP-AT1G20960C construct has 2538bp long-C terminal fragment of original CDS

**) nEYFP-+48 DME.1 construct has no termination codon on its original termination site so extra 15 amino acids of pSAT4-nEYFP-C1 vector MCS sequence would be added until it meets the first termination codon on the vector.

***) nEYFP-+111 DME.2 construct has 2nd intron (111bp) so has longer insert product length(6075bp) then a original CDS length(5964bp). The 2nd intron was expected to be spliced out when the fusion protein is translated.

BiFC Assay

(1) Protoplast preperation

To perform BiFC assay, 15 to 20 leaves (width: 2 cm, length: 5 cm) were collected from 3 to 4-week-old plants grown under optimal light (ca. 150 $\mu E \cdot m^{-2} \cdot s^{-1}$) conditions. Selected leaves were used in a 'Tape-*Arabidopsis* Sandwich' experiment (Wu et. Al., 2009). The upper epidermal surface was stabilized by affixing a strip of Time tape (Time Med, Burr Ridge, IL) while the lower epidermal surface was affixed to a strip of Magic tape (3 M, St. Paul, MN). The Magic tape was then carefully pulled away from the Time tape, peeling away the lower epidermal surface cell layer. The peeled leaves still adhering to the Time tape, were transferred to a Petri dish containing 20 mL of enzyme solution [1% cellulase R-10, 0,2% macerozyme R-10, 20mM KCl, 20mM MES-KOH, 0.4M mannitol, 0.1% BSA, and 10mM $CaCl_2$, pH 5.7]. The leaves were incubated in 28℃ for 1 hour then gently shaken (40 rpm on a platform shaker) for 5 to 10 minutes until the protoplasts were released into the solution. The protoplasts were centrifuged at 850rpm for 5 minutes washed twice with 10 and 5 mL each time of pre-chilled modified W5 solution [154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 5 mM glucose, and 2 mM MES, pH 5.7] and incubated on ice for 30 min in dark. The protoplasts were then centrifuged and resuspended in modified MaMg solution [0.4 M mannitol, 15 mM $MgCl_2$, and 4 mM MES, pH 5.7] to a final concentration of 2 to 5 × 10^5 cells/mL.

(2) DNA-PEG-Calcium transfection

Approximately 5 × 10^4 protoplasts (2 × 10^4 to 1 × 10^5) in 0.2 mL of MaMg solution were mixed with 20 μg each of plasmid DNA (pSAT4-nEYFP-cDME and pSAT4-cEYFP-*candidate gene*) at room temperature. An equal volume of a freshly-prepared solution of 40% (w/v) PEG (MW 4000; Fluka) with 0.1 M $CaCl_2$ and 0.2 M mannitol was added, and the mixture was

incubated at room temperature for 15 minutes. After incubation, 0.8 mL of W5 solution was added slowly, the solution was mixed, and protoplasts were pelleted by centrifugation at 1000rpm for 2 minutes. The protoplasts were resuspended gently in 0.2 mL of W5 and were incubated at 22℃ for 12 to 16 hours in dark.

(3) Confocal laser scanning microscopy

Protoplasts were observed with a Zeiss LSM700 META laser scanning confocal microscope using 20×/0.8 Plan-Apochromat, 40×/1.2 W C-Apochromat or 63×/1.4 Oil Plan-Apochromat in multi-track channel mode. Excitation wavelengths and emission filters were 488 nm/band-pass 505-530 nm for YFP and 488 nm/band-pass 650-710 nm for chloroplast auto-fluorescence. Image processing was performed using Zen lite (Carl Zeiss, blue edition).

III. Results and discussion

Cloning for BiFC assay

As making the constructs for the BiFC assay, there were a few changes to confirm before examine the microscopy results. In pSAT4-nEYFP-DME construct for the general BiFC assay (+111 DME.2) there is 111bp-long 2nd intron of *DME* gDNA but it is expected to be spliced out when the construct is expressed in the protoplast. Also, in pSAT4-nEYFP-+48 DME.1 which is used to compare the interactions related to DME with AT5G37930, the *DME.1* lose its own stop codon, thereby it has 48 more base-pairs. Thus when this construct is expressed in the protoplast, 15 more amino acids extended before it gets to the nearest stop codon of the pSAT4 MCS. Additionally, pSAT4-cEYFP-At1g20960C and pSAT4-cEYFP-+NLS At1g20960C constructs were used to verify its interaction with DME. These constructs do not have full CDSs but a 2538bp-long partial *At1g20960* C terminal fragment and pSAT4-cEYFP-+NLS At1g20960C has NLS that lies on the N terminal part of At1g20960 with a 6 base-pair-long linker sequence. (Fig. 5).

Except for these constructs, all the other constructs that were generated for the initial interaction screening for BiFC were made properly with their main CDS forms (Table 3).

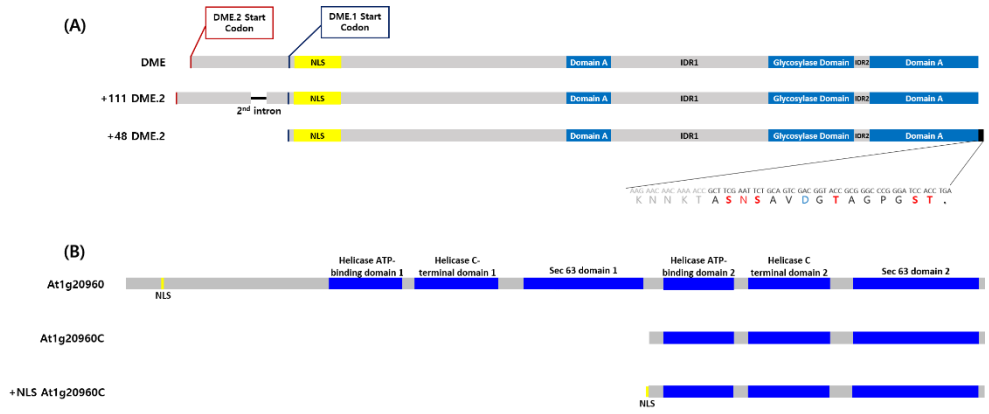


Figure 5. Schematic description of DME and AT1G20960 CDS used for BiFC assay

(A) CDS of DME. First two strands are the original CDS form of DME and the lower two strands refers to be an experimentally used ones. +111 DME.2 has 111 base-pair-long 2nd intron in its N terminal and +48 DME.1 has 48 base-pair-long additional sequence that originates from the MCS of pSAT4-nEYFP vector. The total length is longer than the original form with 111bp and 48bp each.

(B) CDS of At1g20960. The CDS of At1g20960 is 6517bp long but in this study the 2538 base-pair-long C terminal part of At1g20960 had been used. There are two different constructs that have the same C terminal part of At1g20960 but one in the very bottom has NLS on its 5' end. The NLS was from its own NLS sequence located in the N terminal part and was linked with the C terminal part of At1g20960C with 6-base-pair-long linker sequence.

5 genes were identified as interactors of DME through BiFC

Using pSAT4-nEYFP-+111 DME.2, 18 candidate genes (Table 3) were analyzed with BiFC and 5 of them showed positive EYFP signals: AT1G20960; AT5G37930; AT5G60980; AT1G70620; AT5G23090 (Fig. 6; Fig. 7).

AT1G20960C physically interacts with DME regardless of NLS. But with the NLS, DME and AT1G20960C were co-localized in the cell nucleus and the signal intensity was clear and strong (Fig. 6). Also the transfected cell ratio that with positive fluorescent signals were higher in +NLS AT1G20960C and DME interaction (Data not shown). In addition, the EYFP pattern differed when AT1G20960C has NLS. Without NLS, the EYFP signal was faint and blurry but when the NLS is added, the EYFP signals were localized more specifically in the nucleus and had spot-like patterns instead (Fig. 6).

AT1G20960 is U5 small nuclear ribonucleoprotein helicase also known as BRR2a, components of the spliceosome and highly conserved in eukaryotes (TAIR). *Arabidopsis* BRR2a is ubiquitously expressed in all analyzed tissues and involved in the processing of flowering time gene transcripts, mostly FLC (Mahrez et al., 2016). In addition, BRR2a showed physical interaction with yeast PRP6-like splicing factor STA1 which was screened to be a suppressor of *ros1* (Dou et al., 2013). Thereby it is plausible that BRR2a involved in the DNA demethylation and regulating gene expression by physical interaction with DME.

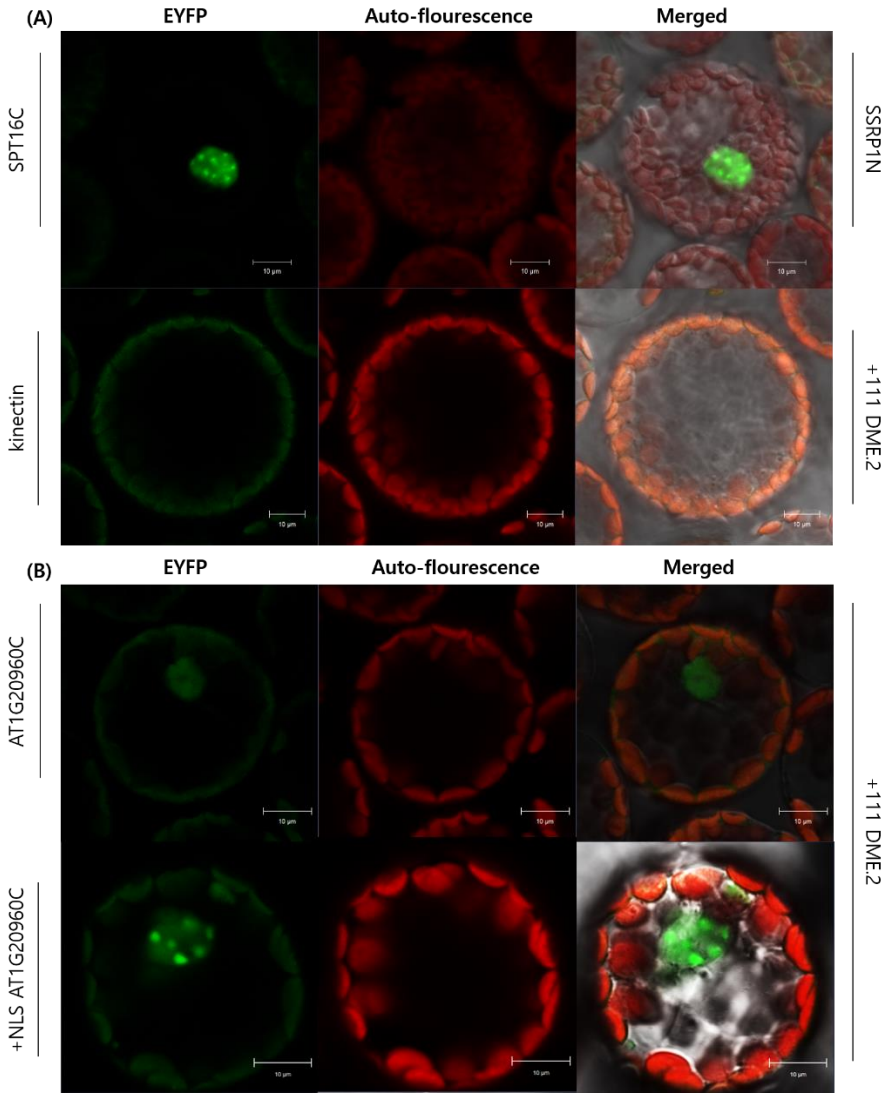


Figure 6. Confocal laser scanning microscopy of BiFC assay shows that DME physically interacts with AT1G20960 C terminal.

(A) Controls for BiFC assay. For a positive control, nEYFP-SSRP1N and cEYFP-SPT16C had been co-transfected into the protoplast and as a negative control for DME cEYFP-kinectin was introduced into the protoplast.

(B) AT1G20960C interacts with DME whether it has NLS or not. But with the NLS, the co-localization seemed to be more clear and strong in the nucleus and the EYFP pattern differed like spot-like signals.

Another physical interactor of DME that had been screened by BiFC is AT5G37930 which described as E3 ubiquitin-protein ligase SINA-like 10. The EYFP signal patterns of AT5G37930 and DME interaction were variable and scattered (Fig. 7; Fig. 8). More than a half of transfected cells showed irregular patterns of EYFP signals whether cEYFP-AT5G37930 were co-transfected with nEYFP-+111 DME.2 or nEYFP-+48 DME.1(Fig. 8A).

As AT5G37930 has the E3 ligase activity, to verify if the E3 ligase activity may affect the interaction of DME and AT5G37930 itself by protein degradation, non-catalytic form of AT5G37930 which lost its RING-domain catalytic activity thereby cannot catalyze E3 ubiquitin ligase activity was constructed by changing a single histidine of the well-conserved RING-domain into tyrosine. There were two assumptions in setting this BiFC analysis: 1) If the only one form of DME is major in *Arabidopsis*, the other minor form of DME might be degraded by ubiquitin-proteasome pathway or at least, affected by it. 2) If the protein degradation pathway is fast enough, the complemented fluorescent signal may not be emitted even if the two proteins physically binds to each other *in vivo*.

As shown in Figure. 8, ubiquitin ligase activity of AT5G37930 affects the interaction between DME and AT5G37930 itself but not as distinct as all or none. Compared to the control groups, +111 DME.2 showed slightly lower transfection ratio when the AT5G37930 lost its catalytic activity (Fig 8B). But in +48 DME.1, the transfection ratio was elevated as much as the ratio of +111 DME.2 level when the catalytic activity of AT5G37930 is lost (Fig 8B). In other way, the catalytic activity of AT5G37930 seemed to affect the interaction of DME and AT5G37930 especially in the DME.1. If the catalytic activity of AT5G37930 has gone, the transfection ratio which can indicate the degree of some physical interaction got recovered as much as that of DME.2 and AT5G37930.

Previously, DME.2 was reported as a major form of DME in general. Combined with this result, the major form, DME.2 is less affected by the catalytic activity of AT5G37930 so it does not be degraded in *Arabidopsis in vivo.* However, the minor form, DME.1 has higher possibility of going

through an ubiquitin-associated protein degradation pathway.

The other interactors of DME *in vivo* are Nuclear transport factor 2 (NTF2) family protein with RNA binding (RRM-RBD-RNP motifs) domain-containing protein AT5G60980, cyclin-like protein AT1G70620 and nuclear factor Y, subunit B13 (NF-YB13) AT5G23090.

AT5G60980 is a NTF2 family protein which mediates the nuclear import of Ran-GDP (Stewart, 2000). Since NTF2 should import Ran-GDP into the cell nucleus and then have to go back to the cytosol as it is dissociated from Ran-GTP, AT5G60980 does not have any NLS in its CDS. So it makes sense that the BiFC signals of AT5G60980 seems blurry and faint but accumulated quite a few in the nucleus (Fig. 7). Most importantly, AT5G60980 has a RRM (RNA-recognition motif)-RBD (RNA binding domain)-RNP(Ribonucleoprotein) motif so it can recognize and bind to RNA. AtMBD6, a methyl CpG binding domain protein, was previously reported to physically interact with AtNTF2 and the *ntf2* mutant showed decreased DNA methylation at miRNA/siRNA producing loci, pseudogenes and some targets of RdDM like MOP9.1 and SDC (Parida et al., 2017). Therefore, AT5G60980 would be one of the interesting interactor to be further studied.

AT1G70620 shows distinct spot-like EYFP signal patterns specifically located in the cell nucleus (Fig. 7). These spots repeatedly shown in DME interactors' BiFC microscopy (Fig. 6B; Fig. 7) may possibly indicate heterochromatic regions (Lungu et al., 2017). AT1G70620 is described as a cylin-like protein but little is known about this particular gene or protein. However, as the intensity and the transfection ratio of the recovered EYFP signal are the highest amongst, the physical interaction itself is the most distinct *in vivo*.

AT5G23090 is a homolog of transcription factor protein Dr1. Human Dr1 has DNA binding activity and transcriptional corepressor activity, and it is involved in chromatin remodeling and histone H3 acetylation (Vaquerizas et al., 2009). AT5G23090 shows low intense EYFP signal compared to the other interactors of DME that found in this study (Fig.7) but the molecular property is plausible to associate with DME in epigenetic context, so further studies in the plants' phenotypic level are necessary.

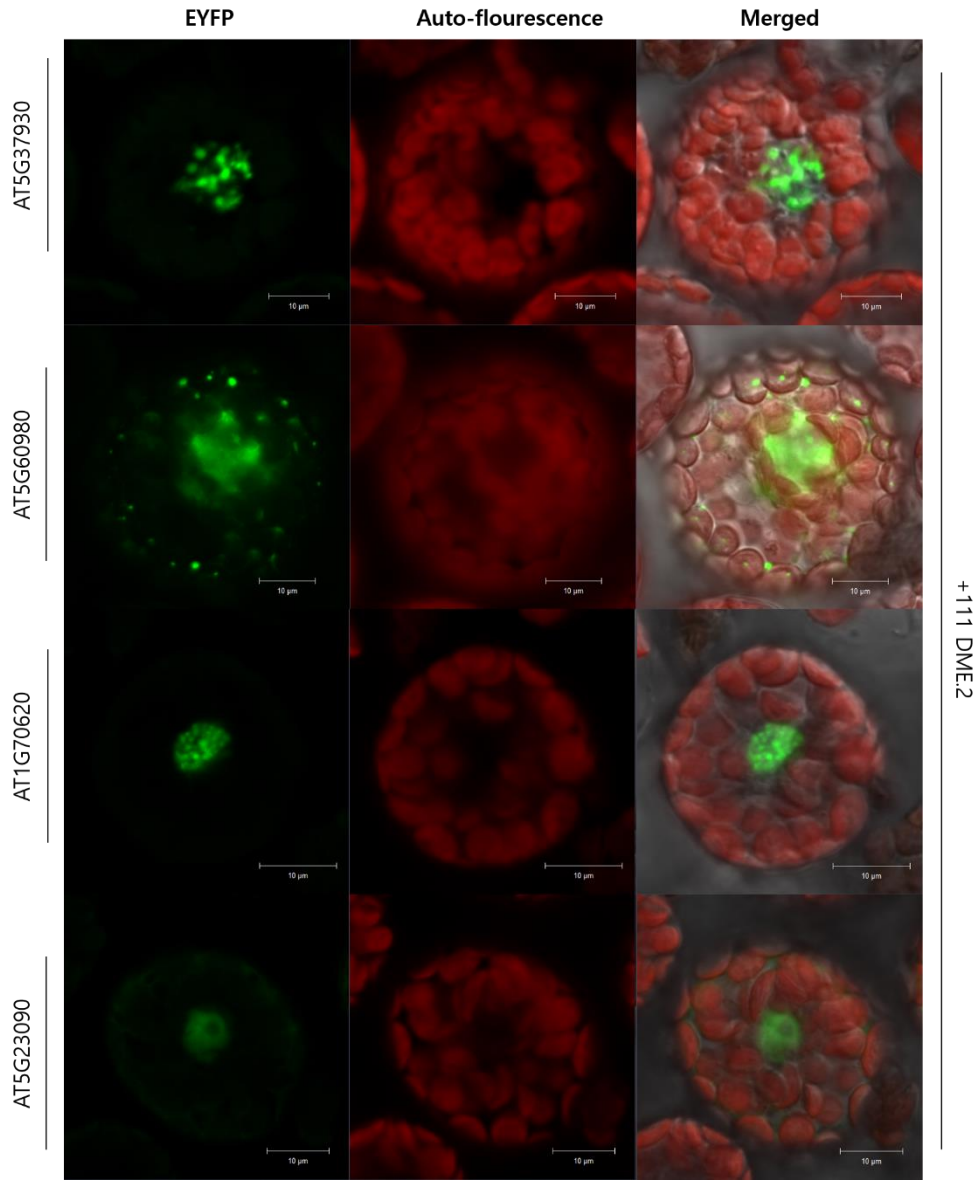
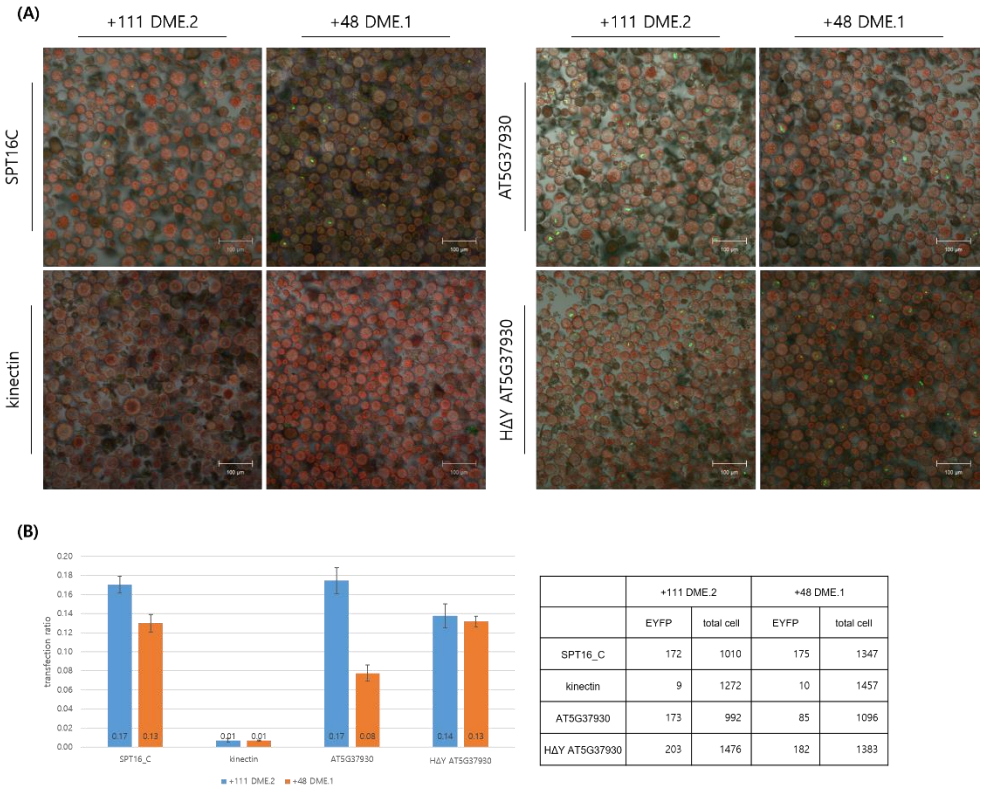


Figure 7. Confocal laser scanning microscopy of BiFC assay shows that DME physically interacts with AT5G37930, AT5G60980, AT1G70620 and AT5G23090.

AT5G37930 showed scattered and irregular patterns of EYFP signals when it was co-transfected with DME. AT5G60980 had lower tendency of being located in a nucleus but showed diffused fluorescent signals in cytosol as well. AT1G70620 showed spot-like signal patterns like +NLS AT1G20960C. AT5G23090 showed weak EYFP signals but those were constantly located in the nucleus.



	+111 DME.2		+48 DME.1	
	EYFP	total cell	EYFP	total cell
SPT16_C	172	1010	175	1347
kinectin	9	1272	10	1457
AT5G37930	173	992	85	1096
HΔY AT5G37930	203	1476	182	1383

Figure 8. DME and AT5G37930 interaction varied whether the E3 ubiquitin ligase activity of AT5G37930(E3 ubiquitin-protein ligase SINA-like 10) is active or not.

(A) Confocal laser scanning microscopy image of +111 DME.2 and +48 DME.1 interacting with AT5G37930 and non-catalytic AT5G37930. The HΔYAT5G37930, a non-catalytic form of AT5G37930 was made by a point mutagenesis that changed 131H of RING-domain into Y.

(B) As compared to the control groups, +111 DME.2 showed slightly lower transfection ratio when the AT5G37930 lost its catalytic activity. But in +48 DME.1, the transfection ratio is elevated about 1.675 times when the catalytic activity of AT5G37930 is lost. In other way, the catalytic activity of AT5G37930 seemed to affect the interaction of DME and AT5G37930.

CHAPTER TWO:

Phenotypic Study for the DME interactors

I. Introduction

Role of DNA demethylation in gametogenesis and seed development

Formation of the male gametophyte in flowering plants consists of two distinct sequential phases, microsporogenesis and microgametogenesis(Borg et al., 2009). Microsporocytes undergo a meiotic division to produce a tetrad of four haploid microspores. During microgametogenesis, the released microspores undergo a highly asymmetric division, called Pollen Mitosis I (PMI), to produce a bicellular pollen grain with a small germ cell engulfed within the cytoplasm of a large vegetative cell (Fig 9A). While the vegetative cell exits the cell cycle, the germ cell undergoes a further mitotic division at Pollen Mitosis II (PMII) to produce twin sperm cells (Fig 9A).

Female gametogenesis also starts with a meiosis, a single haploid cell, usually the basal (chalazal) cell, enlarging and generating the functional megaspore while the remaining products of meiosis degenerate (Skinner and Sundaresan, 2018). This haploid megaspore will go through three mitotic divisions accompanied by nuclear movement to create a defined pattern at each division (Fig 9B). From stage FG4, the large vacuole (blue) separates the nuclei along the chalazal-micropylar axis. At FG5, the polar nuclei (red) migrate to meet each other and eventually fuse. At FG6/FG7, the mature female gametophyte has seven cells: two synergids, egg cell, central cell with large diploid nucleus (central cell nucleus), and three antipodal cells (Skinner and Sundaresan, 2018).

During the double fertilization of *Arabidopsis*, the fertilization of the central cell(2n) and a sperm cell(n) generates triploid endosperm which does not deliver its DNA contents to the next generation and the fertilization of the egg cell(n) and a sperm cell(n) generates diploid embryo. *DME* is expressed in the vegetative cell nucleus of the male gametophyte (Schoft et al., 2011; Park et al., 2017) and central cell of the female gametophyte during the early gametogenesis and development (Choi et al., 2002).

DME induces hypomehtylation of maternal effect gene *MEDEA* (*MEA*) Polycomb gene (PcG)

in the central cell before fertilization by demethylation so that the *MEA* got transcriptionally activated to produce MEA proteins (Gehring et al., 2006). After fertilization, the FIE-MEA PcG complex activated by maternal DME binds to the paternal *MEA* making it to be silenced (Gehring et al., 2006). So if there are mutations in maternal *dme*, maternal *MEA* cannot be transcriptionally activated so the phenotype of which is identical to *mea* mutant: Endosperm overproliferation; embryo arrest; seed abortion (Choi et al., 2002; Grossniklaus et al., 1998; Kiyosue et al., 1999; Luo et al., 1999).

Meanwhile, in mammal, there reported a genomewide methylation reprogramming by massive demethylation and re-methylation process that take places in preimplantation embryo (Reik et al., 2001). In plants, however, there is no massive methylation reprogramming in embryo but instead, there reported that the companion cells whose DNA contents are not inherited go through global demethylation and the hypomethylated state is maintained in late endosperm stage in which *DME* is no longer expressed (Hsieh et al., 2009; Ibarra et al., 2012 Park et al., 2016; Kim et al., 2019). Based on these researches, the hypothesis that massive TEs demethylation in the central cell makes TE transcriptions highly activated and the small RNAs generated from them may be delivered to the neighboring egg cell or to the embryo so that the TE repressions are reinforced in the embryo by RdDM pathway was developed (Gehring, 2019).

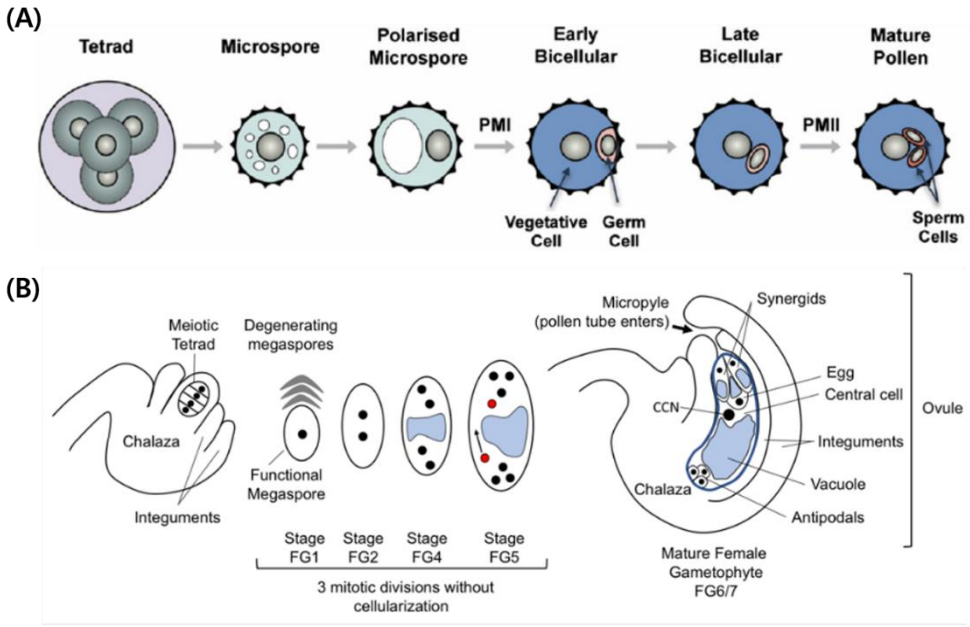


Figure 9. Gametogenesis and development in Arabidopsis

(A) Development of the male gametophyte (Borg et al., 2009). *DME* is expressed in the vegetative cell nucleus. From the mature pollen, two sperm cells participate in the double fertilization and generate embryo and endosperm.

(B) Development of the female gametophyte (Skinner and Sundaresan, 2018). *DME* is expressed in the central cell nucleus (CCN). Mature female gametophyte generates egg cell and central cell. After fertilization, the central cell fertilized with a pollen become an endosperm and the egg cell fertilized with a pollen become an embryo that delivers DNA contents to the next generations.

II. Materials and methods

Plant Material and Growth Conditions

Arabidopsis thaliana Columbia-0 (Col-0) was used as the wild type. Plants were grown in an environmentally-controlled chamber with a long photoperiod (16 hour light and 8 hour dark) at 22℃. The transgenic lines were obtained from the Arabidopsis Biological Resource Center (ABRC, Columbus, OH).

Seed-set analysis and whole-mount clearing

The T-DNA insertion knock-out mutant alleles used in this study are *emb1507-1* (Stock name: CS16090, ABRC) and *dme-2* (Table 5).

Heterozygous *EMB/dme1507-1* plants were first sowed on the MS only plates. To observe seed abortion, siliques (DAP 8 to 10) were dissected on a dissecting microscope (Stemi DV4, Carl Zeiss) and aborted seeds and undeveloped ovules were counted. For whole mount clearing, siliques (DAP 8 to 10) were dissected and mounted in clearing solution [2.5g chloral hydrate; 0.3ml 100% glycerol; 0.7ml distilled water] for 1 hour. Then the samples were observed using an Axio Imager A1 microscope (Carl Zeiss) under DIC optics and were photographed using an AxioCam HRc camera (Carl Zeiss).

After checking the single heterozygotic phenotype of *emb1507-1*, to examine the phenotypic changes in double heterozygous mutant, *EMB/emb1507-1* was crossed with *dme/dme-2* homozygous mutant allele. To obtain the double heterozygote mutant, *EMB/emb1507-1* buds of FG7 stage were dissected and emasculated. The emasculated pistils were placed in 22℃ long day condition growth chamber for 24 hours for maturation then crossed with *dme/dme-2* used as a pollen donor. Fully mature fertilized seeds were harvested and the F1 sowed on MS^{basta} plate for selection. Confirming their genotypes with PCR amplifications (Table 5), check the seed abortion and embryo defects with the same processes used in examining heterozygous *EMB/dme1507-1*

mutant. F2 of these double heterozygote mutants were genotype and in addition, reciprocal cross using Col-0 as a wild type was done to verify the segregation ratio changes.

Table 5. T-DNA insertion alleles and genotyping primers used in this study

Gene	Name	Back ground	insertion	location	primers for genotyping	
AT1G 20960	*emb 1507-1*	col	pCSA 104luc	exon	EMB1507-1 F	GGGATAAATGATG AGGATGC
					EMB1507-1 R	CTCTATCAGCAACA TCTCTCC
					pCSA104 RB3	TAACAATTTCACAC AGGAAACAGCTAT GAC
AT5G 04560	*dme-2*	col-gl	pSKI15	exon	B33F	CACTTGTTCCCTAT GAGAGC
					B33R	CACTGATTGTGATG TTCCAC
					SKI015 LB3	TTGACCATCATACT CATTGCTG

III. Results and Discussion

EMB/emb1507-1 seed phenotype

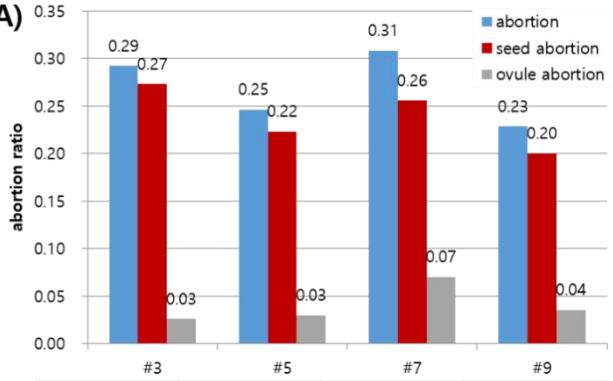
EMB/emb1507-1 showed constant abortion ratio near 25% (Fig. 10A). It seems that *emb/emb1507-1* would arrested in the globular stage of embryo and eventually aborted (Fig. 10B). *DME/dme-2* shows 50% abortion because maternal inheritance of mutant *dme* allele causes seed abort while paternal *dme* did not affect. *EMB/emb1507-1* single heterozygote mutant plants are expected to follow the Mendelian genetics and traditional segregation ratio.

EMB/emb1507-1;DME/dme-2 double heterozygote mutant seed phenotype

To examine the effect of BRR2a on DME, double mutant was generated. Since *DME* is a maternal effect gene, if the *dme-2* allele is derived from maternal allele, the F2 seeds would be all aborted. So using *dme/dme-2* mutant as paternal pollen donor, *EMB/emb1507-1* was crossed to make *EMB/emb1507-1;DME/dme-2* mutant.

EMB/emb1507-1;DME/dme-2 varied seed abortion ratio from 40 to 68% (Fig. 11A). Expected seed abortion ratio is calculated as 62.5%. The lowest seed abortion lines tend to have highest ovule abortion ratio so it is hard to consider that the seed abortion phenotype of *dme-2* was rescued. As examining the arrested seed phenotypes, detectable aborted seeds' ratio is almost 66%, a bit higher than expected. Among the 66% detectable arrested seeds, half of them are arrested in the hear stage which is the seed phenotype of *dme-2* and the other half is arrested in the globular stage which is the seed abortion phenotype of *emb1507-1.* And 5% of seeds were proceeded to torpedo stage.

(A)



Line	seed A	ovule A	normal
#3	41	4	109
#5	29	4	101
#7	29	4	101
#9	44	13	128

(B)

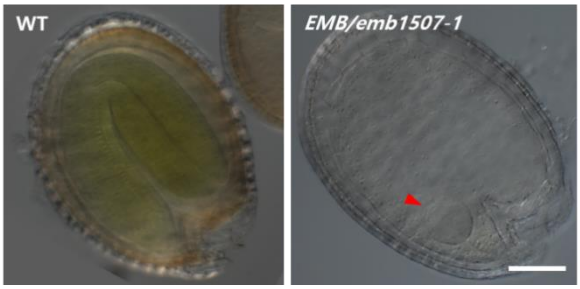


Figure 10. *EMB/enb1507-1* seed abortion phenotypes.

(A) Seed abortion ratio of *EMB/emb1507-1* mutant plants. About 25% of the seeds are aborted (lowest: 20%, highest:27%) while WT siliques showed no seed abortion and slightly lower ovule abortions than *EMB/emb1507-1*. And overall plants, minor ovule abortions were constantly observed.

(B) DIC image of mounted-clearing image of aborted seeds. Most of the aborted seeds of *EMB/emb1507-1* (right) are seemed to be arrested in the globular stage (red arrow) while the WT seeds were almost fully mature. Scale bar=100μm

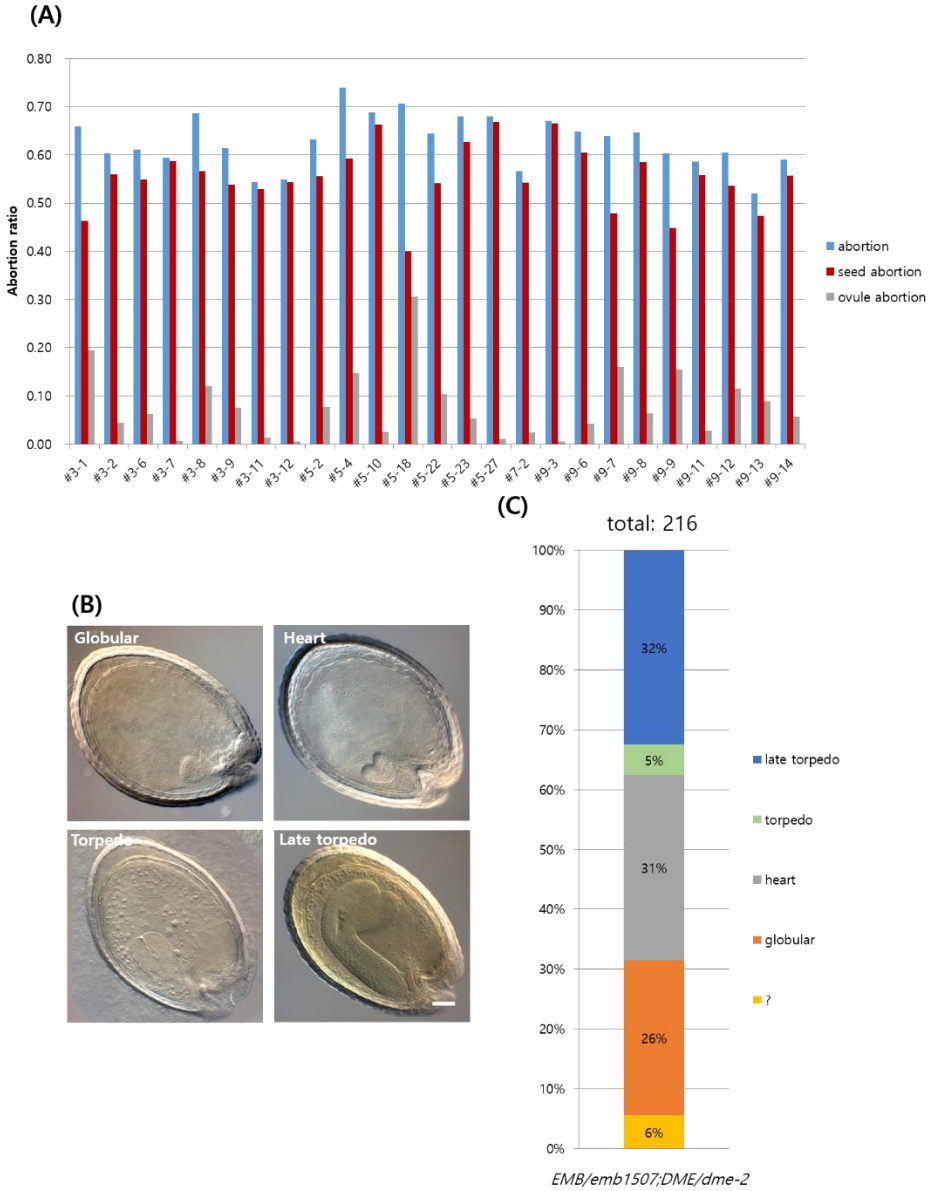


Figure 11. *EMB/emb1507-1;DME/dme-2* seed abortion phenotypes.

(A) Abortion ratio of *EMB/emb1507-1;DME/dme-2.* Expected seed abortion ratio is calculated as 62.5%. The seed abortion ratio varies from 40 to 68%. The lines with lower seed abortion however have significant amount of ovule abortion.

(B) Mounted seed phenotypes of *EMB/emb1507-1;DME/dme-2.* Late torpedo seeds are expected to have WT genotypes. Scale bar=100μm

(C) Distribution of *EMB/emb1507-1;DME/dme-2* seeds that arrested in a specific stages. Color code yellow is undetectable seeds due to its loss of embryos when the seeds were mounted. Among the 66% detectable arrested seeds, half of them are arrested in the hear stage which is the seed phenotype of *dme-2* and the other half is arrested in the globular stage which is the seed abortion phenotype of *emb1507-1.* And 5% of seeds were proceeded to torpedo stage.

Transmission of *emb* and *dme* alleles in *EMB/emb1507-1;DME/dme-2* mutant plants.

In a single heterozygote mutant, if the mutant allele was reciprocally crossed with WT, *emb1507-1* allele seems to have 25% of transmission ratio and as *dme-2* allele is known to have 0% of maternal transmission and 15 to 20% of paternal transmission when it is heterozygous in Col background.

Since *emb1507-1* allele is aborted only when it is homozygous, in a reciprocal cross with WT, seed viability is only affected by *dme-2*. In Figure 12A, the seeds generated from *EMB/emb1507-1;DME/dme-2* reciprocally were viable as expected. Most of the seeds were germinated when the *EMB/emb1507-1;DME/dme-2* was crossed with WT paternally and a bit more than a half of the seeds were germinated when the *EMB/emb1507-1;DME/dme-2* was crossed with WT maternally (Fig. 12A). It seems that the seeds of the maternal *EMB/emb1507-1;DME/dme-2* is a bit more viable compared to the controls but not in a dramatic ratio considering the deviation.

Compared to the control, paternal and maternal transmission of *dme-2* dose not affected whether *dme-2* allele is co-segregated with *emb1507-1*. In case of *emb1507-1* allele, in both cases, transmission ratio of *emb1507-1* got lowered than the expectation when it is with *dme-2* allele (Fig. 12B).

In a heterozygous context, *dme-2* transmission seems not to be affected by *emb1507-1* but the *emb1507-1* allele seems to be affected by *dme-2* a bit. Both in maternal and paternal transmission ratio of *emb1507-1* had been lowered about 7 to 8 % but to verify whether it is truly affected by *dme-2* or not, *EMB/emb1507-1* single mutant must be reciprocally crossed with col-0 and statistically processed.

	total	germination	no germination	%
control	62	60	2	0.97
#6	339	339	0	1.00
#17	331	330	1	1.00
#26	293	293	0	1.00
total	1025	1022	3	0.997
	963	**962**	**1**	**0.999**

	total	germination	no germination	%
control	398	212	186	0.53
#6	179	113	66	0.63
#17	88	61	27	0.69
#26	101	51	50	0.50
total	766	437	329	0.57
	368	**225**	**143**	**0.61**

dme-2	total	+/-	T/R
control	60	28	0.23
#6	319	126	0.20
#17	325	83	0.13
#26	293	128	0.22
total	**937**	**337**	**0.18**
emb1507-1	total	+/-	T/R
#6	327	65	0.10
#17	330	142	0.22
#26	293	107	0.18
total	**950**	**314**	**0.17**

dme-2	total	+/-	T/R
control	212	1	0.002
#6	113	1	0.004
#17	61	1	0.008
#26	51	0	0
total	**225**	**2**	**0.004**
emb1507-1	total	+/-	T/R
#6	113	41	0.18
#17	61	15	0.12
#26	51	25	0.25
total	**225**	**81**	**0.18**

Col-0 X *EMB/emb1507-1;DME/dme-2* *EMB/emb1507-1;DME/dme-2* X Col-0

Figure 12 Seed viability and allele transmission ratio of *EMB/emb1507-1;DME/dme-2*.

(A) Using *DME/dme-2* which was segregated from *EMB/emb1507-1;DME/dme-2* self-pollinated F2 lines as a control, viability of the reciprocally crossed seeds was verified. When *EMB/emb1507-1;DME/dme-2* was crossed with col-0 paternally, almost all the seeds were viable. When *EMB/emb1507-1;DME/dme-2* was crossed with col-0 maternally, control group showed 53% of seed viability similar to expectation but the experimental groups showed about 10% higher seed viability.

(B) Using the same control as (A), paternal and maternal transmission of *dme-2* allele and *emb1507-1 allele* was checked, respectively. Compared to the control, paternal *dme-2* transmission ratio is a bit lower and maternal *dme-2* transmission ratio haven't been changed. In case of *emb1507-1* allele, the expected transmission ratio is 25% whether it is paternal or maternal and in both cases, transmission ratio of *emb1507-1* got lowered when it is with *dme-2* allele.

CONCLUSION

By using yeast two-hybrid (Y2H) mating method, 83 genes were found to be valid interactors of DME N terminal fragment. From these 83 genes, 18 candidate genes were listed-up for further analysis using bimolecular fluorescence complementation (BiFC) assay to verify those interactors actually interact with DME in plant *in vivo*. By using BiFC, 5 genes (AT1G20960; AT5G37930; AT5G60980; AT1G70620; AT5G23090) out of 18 candidates showed positive EYFP signals that indicates interaction between the interactors and DME.

One interactor that was screened from BiFC assay was AT5G37930, as E3 ubiquitin-protein ligase SINA-like 10. The EYFP signal patterns of AT5G37930 and DME interaction were variable and scattered. Ubiquitin ligase activity of AT5G37930 affects the interaction between DME and AT5G37930 especially in the DME.1. When the catalytic activity of AT5G37930 has gone, the transfection indicated by the degree of some physical interaction was recovered as much as that of DME.2 and AT5G37930. This result may indicate that the major form DME.2 which is less affected by the catalytic activity of AT5G37930 seems not to be degraded in *Arabidopsis in vivo* but the minor form, DME.1 has higher possibility of going through an ubiquitin-associated protein degradation pathway.

The other interactors of DME *in vivo* are Nuclear transport factor 2 (NTF2) family protein AT5G60980, cyclin-like protein AT1G70620 and nuclear factor Y, subunit B13 (NF-YB13) AT5G23090.

AT5G60980 is a NTF2 family protein which mediates the nuclear import of Ran-GDP and does not have NLS in its CDS. So it makes sense that the BiFC signals of AT5G60980 seems blurry and faint but accumulated quite a few in the nucleus. Most importantly, AT5G60980 has a RRM-RBP-RNP motif which might be an indication of relevance with RdDM pathway but to verify this, further study is needed.

AT1G70620 shows distinct spot-like EYFP signal patterns specifically located in the cell

nucleus. As mentioned before, these spots possibly indicate heterochromatic regions of DNA contents. AT1G70620 is described as a cylin-like protein but little is known about this particular gene or protein.

AT5G23090 is a homolog of transcription factor protein Dr1. Human Dr1 have DNA binding activity and transcriptional corepressor activity and involved in chromatin remodeling and histone H3 acetylation so the molecular property is plausible enough to do with DME in epigenetic context so further studies in the plants' phenotypic level are necessary.

The other interactor of DME, AT1G20960 is described as U5 small nuclear ribonucleoprotein helicase also known as BRR2a, a component of the spliceosome and highly conserved in eukaryotes. In this study, it was found that AT1G20960C physically interacts with DME whether it has NLS or not. When the NLS is added, the EYFP signals were localized more specifically in the nucleus and had spot-like patterns that might be some heterochromatic regions. In addition in previous report, BRR2a showed physical interaction with yeast PRP6-like splicing factor STA1 which is suppressor of *ros1*. Combined with this study, the probability of BRR2a involved in the DNA demethylation and regulating gene expression by physical interaction with DME seems to be valid so further analysis using T-DNA insertion knock-out mutant plants were planned.

As using *EMB/emb1507-1* heterozygote mutants, *at1g20960* showed some defects in seed development. *EMB/emb1507-1* showed approximately 25% of seed abortion and most of the aborted seeds were arrested in globular stage. *DME/dme-2* heterozygote mutant is known to have 50% of seed abortion ratio and have aborted seeds arrested in heart stage so compare these two mutants, seed abortion in *EMB/emb1507-1* is arrested in earlier stage then *DME/dme-2* mutant.

But examining the transmission ratio and seed abortion phenotype of *EMB/emb1507-1;DME/dme-2* double heterozygote mutant showed no distinct or interesting relevance between AT1G20960 and DME. The abortion ratio of *dme-2* did not either rescued or reinforced so as in transmission of *dme-2* allele. *Emb1507-1* allele transmission was lowered slightly in the

EMB/emb1507-1;DME/dme-2 mutant but there was no significance.

But still, there are some possibilities remains to be addressed. For example, the interactors and DME affect each other but not just be seen as morphological phenotypes. Since the central cell can be successfully isolated using INTACT (Isolation of nuclei tagged in specific cell types) by adopting central cell-specific DD7 promoter (Park et al., 2016), analyzing the central cell methylome in DME interactor mutants and compare them with that of *dme* might be one example that can give some new indications.

REFERENCE

Aufsatz, W., Mette, M.F., Matzke, A.J., and Matzke, M. (2004). The role of MET1 in RNA-directed de novo and maintenance methylation of CG dinucleotides. Plant Mol. Biol. *54*, 793-804.

Borg, M., Brownfield, L., and Twell, D. (2009). Male gametophyte development: A molecular perspective. Journal of experimental botany *60*, 1465-1478.

Cao, X., and Jacobsen, S.E. (2002). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc. Natl. Acad. Sci. Suppl *4*, 16491-16498.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature. *452*, 215-219.

Dou, K., Huang, CF., Zhang, CJ., Zhou, JX., Huang, HW., Cai, T., Tang, K., Zhu, JK. And He, X.J,. (2013) The PRP6-like splicing factor STA1 is involved in RNA-directed DNA methylation by facilitationg the production of Pol V-dependent scaffold RNAs, Nucleic Acids Res, *41*(18), 8489-502.

Finnegan, E.J. and Kovac, K. A. (2000). Plant DNA methyltransferases. Plant Mol. Biol. *43*, 189-201.

Finnegan, E.J., Peacock, W.J., and Dennis, E.S. (1996). Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development. Proc. Natl. Acad. Sci. USA *93*, 8449-8454.

Gehring, M. (2019). Epigenetic dynamics during flowering plant reproduction: evidence for reprogramming?, New Phytologist, doi: 10.1111/nph.15856.

Gehring, M., and Henikoff, S. (2008). DNA methylation and demethylation in Arabidopsis. Arabidopsis Book 6: e0102.

Gehring, M., Huh, J.H., Hsieh, T.F., Renterman, J., Choi, Y., Harada, J.J., Goldberg, R.B., and Fischer, R.L. (2006). DEMETER DNA glycosylase establishes MEDEA polycomb gene

self-imprinting by allele-specific demethylation. Cell *124*, 495-506.

Ghosh, I., Hamilton, AD., Regan, L., (2000). Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. J. Am. Chem. Soc *122*, 5658–5659.

Grossniklaus, U., Vielle-Calzada, J.P., Hoeppner, M.A., and Gagliano, W.B. (1998). Maternal control of embryogenesis by MEDEA, a polycomb group gene in *Arabidopsis*. Science *280*, 446-450.

Henikoff, S., and Comai, L. (1998). A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in Arabidopsis. Genetics *149*, 307-318.

Hsieh, T.F., Ibarra, C.A., Dilva, P., Zemach, A., Eshed-Williams, L., Fischer, R.L., and Zilberman, D. (2009). Genome-wide demethylation of *Arabidopsis* Endosperm. Science *324*, 1451-1454.

Ibarra, C.A. et al. (2012). Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. Science *337*, 1360-1364.

Jacobsen, S.E., Sakai, H., Finnegan, E.J., Cao, X., and Meyerowitz, E.M. (2000). Ectopic hypermethylation of flower-specific genes in Arabidopsis. Curr. Biol. *10*, 179-186.

Jullien, PE., Mosquna, A., Ingouff, M., Sakata, T., Ohad, N., and Berger, F. (2008). Retinoblastoma and its binding partner MSI1 control imprinting in *Arabidopsis*. PLoS Biol *6*, e194

Kato, M., Miura, A., Bender, J., Jacobsen, S.E., and Kakutani, T. (2003). Role of CG and non-CG methylation in immobilization of transposons in Arabidopsis. Curr. Biol. *13*, 421-426.

Kankel, M.W., Ramsey, D.E., Stokes, T.L., Flowers, S.K., Haag, J.R., Jeddeloh, J.A., Riddle, N.C., Verbsky, M.L., and Richards, E.J. (2003). Arabidopsis MET1 cytosine methyltransferase mutants. Genetics *163*, 1109-1122.

Kerppola, TK. (2006). Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. Nat. Protoc *1*(3), 1278-1286.

Kim, M.Y., Ono, A., Scholten, S., Kinoshita, T., Zilberman, D., Okamoto, T., and Fischer, RL.

(2019). DNA demethylation by ROS1a in rice vegetative cells promotes methylation in sperm. Proc. Natl. Acad. Sci, *116*(19), 9652-9657.

Kiyosue, T. et al. (1999). Control of fertilization-independent endosperm development by the MEDEA Polycomb gene in *Arabidopsis*. Proc. Natl. Acad. Sci. USA *96*, 4186-4191.

Kress, C., Tomassin, H., and Grange, T., (2006), Active cytosine demethylation triggered by a nuclear receptor involves DNA strand breaks PNAS 103, 11112-11117

Law, J.A., and Jacobsen S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. *11*, 204-220.

Li, Y., Kumar, S., and Qian, W. (2018). Active DNA demethylation: mechanism and role in plant development. *Plant Cell Rep*. *37*(1), 77–85.

Lippman, Z. et al., (2004). Role of transposable elements in heterochromatin and epigenetic control. Nature *430*, 471-476.

Lippman, Z., May, B., Yordan, C., Singer, T., and Martienssen, R. (2003). Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. PLoS Biol. *1*, E67.

Luo, M., Bilodeau, P., Koltunow, A., Dennis, E.S., Peacock, W.J., and Chaudhury, A.M. (1999). Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA *96*, 296-301.

Lungu, C., Pinter, S., Broche, J., Rathert, P., and Jeltsch, A. (2017). Modular fluorescence complementation sensors for live cell detection of epigenetic signals at endogenous genomic sites. Nature communications, *8*(1), 649.

Mahrez, W., Shin, J., Munoz-Viana, R., Figueiredo, DD., Trejo-Arellano, MS., Exner, V., Siretskiy, A., Gruissem, W., Köhler, C., and Hennig, L. (2016). BRR2a affects flowering time via FLC splicing, PLoS Genet, *12*(4), e 1005924.

Mathieu, O., Reinders, J., Caikovski, M., Smathajitt, C., and Paszkowski, J. (2007). Transgenerational stability of the *Arabidopsis* epigenome is coordinated by CG methylation. Cell *130*, 851-862.

Morales-Ruiz, T., Ortega-Galisteo, A.P., Ponferrada-Marin, M.I., Martinez-Macias, M.I., Arixa, R.R., and Roldan-Arjona, T. (2006). DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. Proc. Natl. Acad. Sci. USA *103*, 6853-6858.

Parida, AP., Sharma, A., and Sharma, AK. (2017). AtMBD6, a methyl CpG binding domain protein, maintains gene silencing in Arabidopsis by interacting with RNA binding proteins. J Biosci. *42*(1), 57-68.

Park, K. et al. (2016). DNA demethylation is initiated in the central cells of Arabidopsis and rice. Proc. Natl. Acad. Sci. USA *113*, 15138-15143.

Park, J.S. et al., (2017). Control of DEMETER DNA demethylase gene transcription in male and female gamete companion cells in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA *114*, 2078-2083.

Penterman, J., Zilberman, D., Huh, JH., Ballinger, T., Henikoff, S., and Fischer, RL. (2007). DNA demethylation in the *Arabidopsis* genome. Proc Natl Acad Sci USA *104*, 6752–6757.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089-1093.

Saze, H., Mittelsten Scheid, O., and Paszkowski, J. (2003). Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. Nat Genet *34*, 65-69.

Schoft V.K. et al., (2001). Function of the DEMETER DNA glycosylase in the Arabidopsis thaliana male gametophyte. Proc. Natl. Acad. Sci. USA. *108*, 8042-8047.

Skinner, D.J. and Sundaresan, V. (2018). Recent advances in understanding female gametophyte development. F1000Research *7*, 804.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. Nat. Rev. enet. *8*, 272-285.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, JD., Feijo JA., and Martienssen, RA. (2009), Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell *136*, 461–472.

Stewart, M. (2000). Insights into the molecular mechanism of nuclear trafficking using nuclear

transport factor 2 (NTF2). Cell Struct Funct. *25*(4), 217-25.

Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A. and Luscombe, N.M. (2009). A census of human transcription factors: function, expression and evolution. Nat. Rev. Genet. *10*, 252-263.

Wada, Y., Ohya, H., Yamaguchi, Y., Koizumi, N., and Sano, H. (2003). Preferential de novo methylation of cytosine residues in non-CpG sequences by a domains rearranged DNA methyltransferase from tobacco plants. J. Biol. Chem. *278,* 42386-42393.

Wu, F., Shen, SC., Lee, LY, Lee, SH., Chan, MT., and Lin, CS. (2009). Tape-*Arabidopsis* Sandwitch – a simpler *Arabidopsis* protoplast isolation method, *Plant Methods 5*, 16.

Wu, S.C., and Zhang, Y. (2010). Active DNA demethylation: Many roads lead to Rome. Nat. Rev. Mol. Cell Biol. *11*, 607-620.

Xiao, W., Custard, K.D., Brown, R.C., Lemmon, B.E., Harada, J.J., Goldberg, R.B., and Fischer, R.L. (2006). DNA methylation is critical for Arabidopsis embryogenesis and seed viability. Plant Cell *18*, 805-814.

Zhang, H., and Zhu, J.K., (2012). Seeing the forest for the trees: A wide perspective on RNA-directed DNA methylation. Genes. Dev. *26*, 1769-1773.

Zhang, X., and Jacobsen, S.E. (2006). Genetic analyses of DNA methyltransferases in Arabidopsis thaliana. Cold Spring Harb. Symp. Quant. Biol. *71*, 439-447.

Zhang, X. et al. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. Cell *126*, 1189-1201.

Zhu, J.K., (2009) Active DNA demethylation mediated by DNA glycosylases. Annu Rev Genet. *43*, 143-66.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat. Genet. *39*, 61-69.

Zilberman, D., and Henikoff, S. (2004). Silencing of transposons in plant genomes: kick them when they're down. Genome Biol. *5*, 249.

Zilberman, D., and Henikoff, S. (2007). Genome-wide analysis of DNA methylation patterns. Development *134*, 3959-3965.

국 문 초 록

DNA 의 메틸화와 디메틸화는 유전자의 발현을 조절하여 생물체가 직면하는 변화하는 주변 환경과 다양한 스트레스 상황에 대한 적절한 반응을 가능하게 하며 전이 인자의 발현을 억제하는 등의 역할을 하므로 매우 중요하다. DEMETER(DME)는 식물에서 작용하는 DNA 글리코실화 효소로 애기장대 암배우체의 중심세포에서 발현하며 종자의 발달에 필수적이다. DME 는 염기절제회복방식을 통해 5-mC 를 직접 제거하고 시토신으로 치환하는 방식의 디메틸화를 매개한다. 식물의 배에서는 포유동물과 같은 대규모의 메틸레이션 리프로그래밍이 일어나지 않는 대신 다음 세대로 유전 정보가 전달되지 않는 생식세포나 배의 주변 세포에서 전반적인 디메틸레이션이 일어나고 이러한 하이포메틸레이션 상태가 DME 가 더 이상 발현하지 않는 후기 배주까지 유지된다.

이러한 중요성에도 불구하고 DME 의 디메틸레이션 과정에서 DME 와 상호작용하는 인자들에 대한 연구는 많이 알려진 바가 없다. 본 연구에서는 DME 와 상호작용하는 인자들을 밝히기 위하여 이분자형광상보기법(Bimolecular Fluorescence Complementation)을 이용하였다. 효모이중잡종화(Yeast Two-Hybrid)를 이용한 선행연구를 통해 밝혀진 DME 와 상호작용하는 83 개의 유전자 중 18 개의 후보 유전자를 우선적으로 선정하였고 그 중 AT5G37930, AT5G60980, AT1G70620, AT5G23090, 그리고 AT1G20960 의 C 말단부가 애기장대 야생종(Col-0)의 원형질체에 DME 와 함께 트랜스펙션되었을 때 형광 신호를 보이는 것을 확인할 수 있었다. 이 유전자들은 E3 유비퀴틴 라이게이즈 활성을 가지거나 RNA 결합 모티브를 가지며 히스톤 아세틸레이션과 관계된 전사인자의 상동 유전자를 포함한다.

추가적으로 이분자형광상보기법을 통해 밝혀진 유전자들과 DME 의 관계에 대한 이해를 위해 RNA 스플라이싱에 관여하는 Brr2a 를 암호화하는 유전자인

At1g20960 을 이용하여 돌연변이 연구를 진행하였다. At1g20960 의 T-DNA 삽입 돌연변이체인 *emb1507-1* 식물체를 *dme-2* 돌연변이체와 교배하여 얻은 이중이형접합 돌연변이를 이용하여 종자의 발달단계 및 종자유산 여부와 돌연변이 대립유전자의 분리비, 트랜스미션의 변화를 중심으로 표현형을 분석한 결과 *dme-2* 단일 이형접합돌연변이체와 비교했을 때 해당 표현형들에서의 주목할만한 변화를 관찰할 수는 없었다. 때문에 식물에서 DME 의 디메틸레이션 과정에 관여하는 새로운 인자를 밝혀냄으로써 이에 대한 구체적인 이해를 더하기 위해서는 추가적인 실험을 통해 분자생물학적 관점에서의 표현형 분석이 필요하며 At1g20960 외에 나머지 후보 유전자들에 대한 연구 역시 필요하다.

주요어 : DEMETER(DME), DNA 디메틸레이션, 애기장대, 에피유전학, 이분자형광상보

학 번 : 2017-20959