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

Studies on DEMETER's Promoter and N-terminal Domain Function in *Arabidopsis thaliana*

애기장대 DEMETER 프로모터와 N-terminal 도메인 기능에 관한 연구

지도교수 최 연 희

이 논문을 이학석사학위논문으로 제출함

2018년 12월

서울대학교 대학원

생명과학부

조브랜던

조브랜던 의 이학석사 학위논문을 인준함

2018년 12월

위 원 장 이 지영 (인)

부 위 원 장 최 연희 (인)

위 원 이 유리 (인)

Studies on DEMETER's Promoter and N-terminal Domain Function in *Arabidopsis thaliana*

A dissertation submitted in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE

to the faculties of

School of Biological Sciences

at

Seoul National University

by

Brandon Cho

Date Approved December 14, 2018

Dissertation Committee

Chair Ji-Young Lee (Seal)

Supervisor Yeonhee Choi (Seal)

Examiner Yuree Lee (Seal)

이학석사 학위논문

애기장대 **DEMETER** 프로모터와 **N-Terminal** 도메인 기능에 관한 연구

Studies on DEMETER's Promoter

and N-terminal Domain Function in *Arabidopsis thaliana*

2018 년 12 월

서울대학교 대학원

생명과학부

조브랜던

ABSTRACT

Studies on *DEMETER's* Promoter and N-Terminal Domain Function in *Arabidopsis thaliana*

Brandon Cho
School of Biological Sciences
Graduate School
Seoul National University

The development of the female gametophyte in plants is tightly regulated by epigenetic mechanisms to ensure the proper development of seeds. DNA methylation is an evolutionarily conserved mechanism that controls numerous biological processes such as gene imprinting, gene expression, and paramutation. In plants, DNA methylation occurs in the CG, CHG, and CHH contexts. The regulation of DNA methylation is done by a family of DNA demethylase genes including *DEMETER (DME), REPRESSOR OF SILENCING 1 (ROS1), DEMETER LIKE 2 (DML2),* and *DEMETER LIKE 3 (DML3)*. Both *DME* and ROS1 are bifunctional DNA glycosylases that initiate active DNA demethylation through the base excision repair pathway. *dme* mutant plants exhibit severe seed abortion phenotypes as maternally expressed genes including *MEDEA (MEA)* cannot be

activated. The structure of DME includes two splice variants, “Spain” and “Berkeley” DME with Spain being 258 amino acids longer on the N-terminal side. DME can further be split up into the N-terminal Domain and the C-terminal domain consisting of the A, G and B domains. Currently, the function of the N-terminal domain is unclear as the C-terminal side was shown to have glycosylase functions *in vitro*. Dominant negative strategies were done to provide a possible explanation to the function of the N-terminal domain. Transcriptional controls of DME show unexpected expression patterns opening the possibility that another factor controls DME expression. A series of differential DME fragments also showed unexpected GFP expression patterns.

Keywords: DEMETER, demethylation, imprinting, ectopic expression, overexpression, domain function, female gametophyte development

Student Number: 2015-22299

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

DME	DEMETER
ROS1	REPRESSOR OF SILENCING
NTD	N terminal Domain
CTD	C terminal Domain
Y2H	Yeast 2 Hybrid
MCC	Mixed Charged Cluster
CC	Central Cell
CCN	Central Cell Nucleus
BER	Base Excision Repair
MEG	Maternally Expressed Genes
cDNA	Complementary DNA
GFP	Green Fluorescent Protein

I. Introduction

1. Overview of Plant Reproductive Development

Plants go through a life cycle known as the alternation of generations. This involves the switch between the haploid generation and the diploid generation in the life cycle of the organism (Figure 1). Sexual reproduction starts with sporogenesis where specialized cells in the sporophyte undergoes meiosis to produce haploid spores. These spores then undergo gametogenesis to develop into multicellular gametophytes which produces the sperm and egg cells. The zygote is formed when the sperm and egg fuse. The development of the embryo leads to the end of this life cycle (Gifford and Foster, 1989).

Angiosperms, known as flowering plants, are heterosporous producing two types of spores that develop into the female megaspore and the male microspore. The female megaspore mother cells undergo meiosis to produce haploid megaspores and then three rounds of syncytial mitotic divisions to eventually haploid female gametophytes. After cellularization, seven cells belonging to four cell types are created: the egg cell, the central cell, two synergid cells, and three antipodal cells (Figure 2). The male microspore mother cells undergo meiosis to produce microspores and eventually male gametophytes. The male gametophyte, or pollen grain, develop within the anther and consists of two sperm cells inside a vegetative cell (Gifford and Foster, 1989).

In sexually reproducing angiosperms, seed formation begins when pollen is transferred from the anther to the carpel's stigma. This process of double fertilization occurs when two sperm cells enter the pollen tube towards

the ovule with one sperm cell fertilizing the egg to produce the diploid embryo and the other sperm cell fusing with the central cell to produce the triploid endosperm (Gifford and Foster, 1989). Despite its reduced size relative to the diploid sporophyte, the development of the FG is tightly regulated as it is essential for successful seed formation.

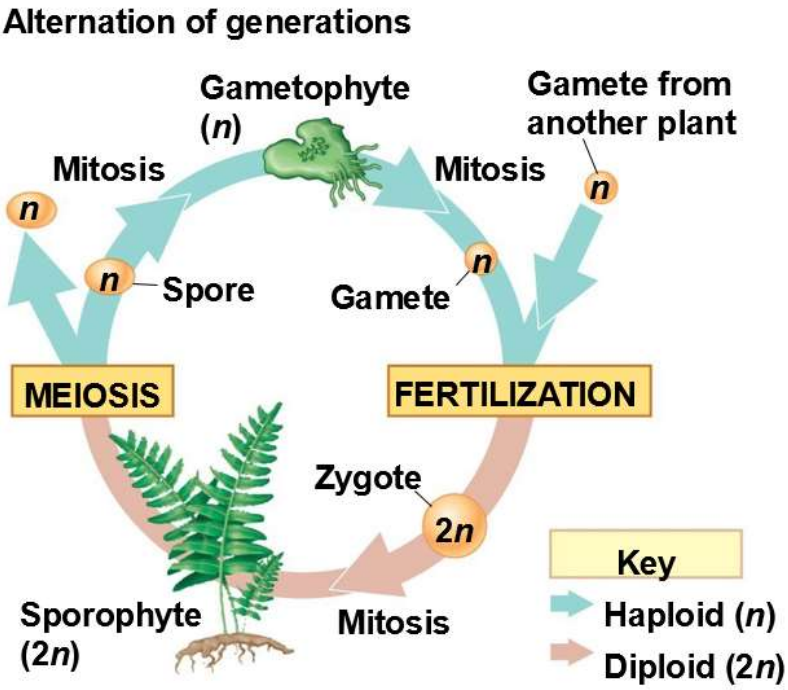

Figure 1. The alternation of generations in plants

The life cycle of the plant starts with the haploid gametophyte to produce gametes. The fusion of these gametes leads to fertilization and starts the diploid generation. The 2n sporophyte undergoes meiosis to produce a haploid gametophyte generation which cycles back to the production of more gametes (Reece et al., 2014).

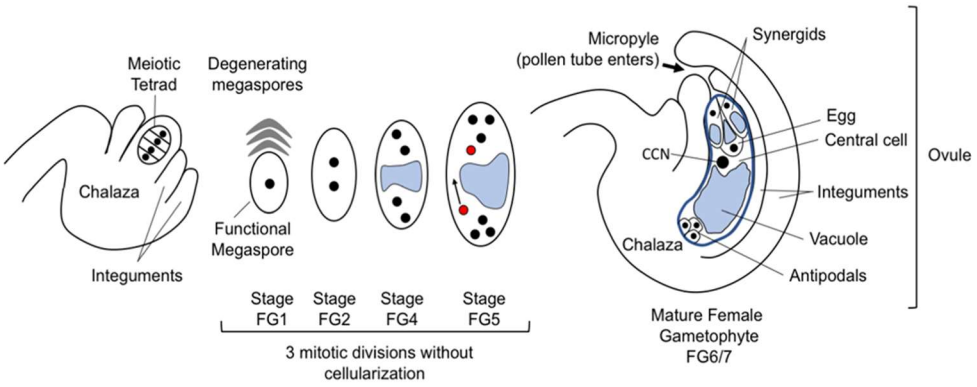

Figure 2. The progression of female gametophyte development

After meiosis, a single haploid cell, usually the basal (chalazal) cell, will enlarge and form the functional megaspore while the remaining products of meiosis degenerate. This haploid megaspore will have three mitotic divisions accompanied by nuclear movement to create a defined pattern at each division. 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, or CCN), and three antipodal cells (Skinner DJ and Sundaresan V, 2018).

2. Epigenetic control for proper seed development in plants

Despite its reduced size relative to the diploid sporophyte, the development of the female gametophyte is tightly regulated by epigenetics to ensure proper development of seeds. Epigenetics refers to stable, heritable changes in gene expression that occur without changes in the DNA sequence (Fujimoto et al., 2012). Three epigenetic mechanisms regulate plant development including 1) DNA methylation, 2) posttranslational histone modifications, and 3) the action of noncoding RNA (siRNAs, microRNAs, miRNAs, long ncRNAs, etc). These mechanisms lead to changes in chromatin structure; the open, active chromatin form allows gene activation while the closed, repressive chromatin form leads to gene silencing (Berger, 2007; Kouzarides, 2007).

DNA methylation, in particular, plays a vital role in *Arabidopsis* development. It is an evolutionarily conserved epigenetic mechanism that controls numerous biological processes including gene imprinting, tissue-specific gene expression, inactivation of transposable elements (TEs), paramutation, and stress responses. In plants, DNA methylation occurs in the CG, CHG, and CHH contexts (H represents A, C, or T). DNA methylation in the promoter leads to gene silencing while methylation in the gene body leads to gene activation. A fine balance between DNA methylation and DNA demethylation maintains proper levels of cytosine methylation and thus proper development. For example, when plants are double or single mutants for DNA methyltransferase, aberrant phenotypes including extreme late flowering, abnormal embryo and endosperm development, reduced leaf size, arrest of vegetative development, and sterility are observed (Cokus et al.,

2008; Law & Jacobsen, 2010; Van Oosten et al., 2014, Matzke and Mosher, 2014; Movahedi et al., 2015).

3. Mechanisms of Active DNA Demethylation in plants

DNA demethylation can be a passive or active process. Passive DNA demethylation occurs when DNA methylation states are lost during DNA replication from the inactivation of enzymes that control DNA methylation (Zhu, 2009). This has been found to occur during both male and female gametophyte development. In the pollen, transposable elements were found to be reactivated only in the vegetative cell but does not provide its DNA to the fertilized zygote. In the female gametophyte, passive DNA demethylation is found in the central cell, the companion cell of the egg cell and fuses with a sperm cell to form the endosperm (Slotkin et al. 2009).

Active DNA demethylation is controlled by a family of DNA demethylase genes: *DEMETER (DME), REPRESSOR OF SILENCING 1 (ROS1), DEMETER LIKE 2 (DML2),* and *DEMETER LIKE 3 (DML3)*. Both *DME* and ROS1 are bifunctional DNA glycosylases that initiate active DNA demethylation through the base excision repair pathway. In this pathway, these enzymes hydrolyze the glycosylic bond between the base and its deoxyribose residue and cleave the DNA backbone at the abasic site. This removal of the 5-methylcytosine (5-meC) base leads to the cleavage of the phosphodiester backbone by either β- or β, δ-elimination and produces a single nucleoside gap that is later filled in with an unmethylated cytosine, leading to transcriptional activation of target genes (Figure 3) (Choi et al., 2002; Gehring et al., 2006, Gong et al., 2002; McCullough et al., 1999; Morales-Ruiz et al., 2006; Ortega-Galisteo et al., 2008).

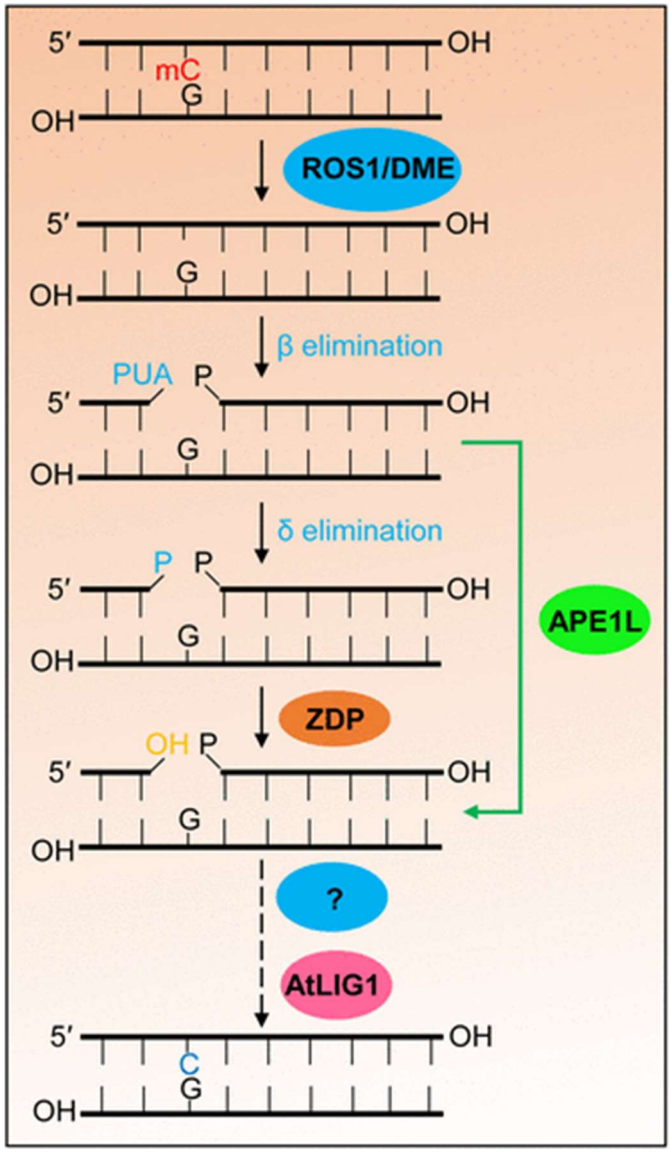

Figure 3. Base excision repair (BER)-mediated active DNA demethylation in plants.

ROS1 and DME are bifunctional DNA glycosylases that remove 5-methylcytosine (mC) and cleave the DNA backbone at abasic site via β or β, δ elimination reaction(s), generating a gap with 3′-PUA or 3′-phosphate terminus. 3′-PUA and 3′-phosphate are converted to 3′-OH by APE1L and ZDP, respectively. The gap is then filled with

a usual (unmethylated) cytosine by the actions of an unknown DNA polymerase and AtLIG1 (Li, 2018).

4. Role of DEMETER, a DNA glycosylase, in plant development

DEMETER (DME) is responsible for the activation of maternally imprinted genes in the central cell which regulates proper seed development (Figure 4). In particular, DME, in the companion cells (central cell and vegetative cell), preferentially targets small, AT-rich, nucleosome-deleted transposable elements (Ibarra et al., 2012).

Maternally expressed genes (MEGs) in the central cell must be tightly regulated to prevent abnormal endosperm and embryo development. *FWA* (Flowering Wageningen), *MEA* (MEDEA), and *FIS2* (Fertilization-Independent Seed 2) are some MEGs that are maintained in a silenced state by DNA methylation and repressive histone modifications. Two genes imprinted in the endosperm, *MEA* and *FWA,* are regulated by DME (Choi et al., 2002; Kinoshita et al., 2004). *MEA* and *FWA* are expressed in the central cell and in the endosperm while *DME* is expressed in the central cell (Kinoshita et al., 1999; Vielle-Calzada et al. 1999; Choi et al., 2002). The maternal alleles of these MEGs are activated by DME-mediated active DNA demethylation. The regulation of MEA is done through promoter demethylation, leading to activation, by DME (Gehring et al., 2006; VanOosten et al., 2014). Although DME is found in the central cell and not in the endosperm, maternal hypomethylation and activation of MEGs are epigenetically maintained in the endosperm. Maternal *DME* mutants result in the failure of MEG activation leading to early seed abortion with phenotypes

of enlarged endosperm and arrested embryo growth. Homozygous *DME*-null mutant plants do not survive whereas heterozygous *DME*/*dme-1* and *DME*/*dme-2* plants exhibit a 50% seed abortion phenotype due to the maternal effect (Figure 5) (Choi et al. 2002).

DEMETER also has similar demethylation patterns in the vegetative cell of the male gametophyte. The suppression of targeted genes caused the reduction of small RNA-directed DNA methylation (RdDM) of transposons in the companion sperm cell indicating movement of small RNAs between the two cell types (Calarco et al., 2012; Ibarra et al.,2012; Schoft et al., 2011). *Arabidopsis* seeds with a paternal *dme* mutant allele still develop normally, suggesting that DME dysfunction does not lead to defective sperm (Choi et al. 2002). However, DME must still function in the male gametophyte because *DME* expression in the vegetative cell was found to be required for DNA demethylation of *MEA*, *FWA*, and the transposon *Mu1a*. A *DME* mutation resulted in impaired vegetative cell germination and pollen tube formation (Schoft et al. 2011). As a result of defective pollen tube formation, self-pollinated *DME*/*dme-1* and *DME*/*dme-2* plants (in the Col-gl background) produce significantly fewer viable heterozygous *DME*/*dme* F1 progenies (approximately 15%) than wild-type progenies (Xiao et al. 2003). Transient DME expression has also been detected in the vegetative cell (Park et al. 2017).

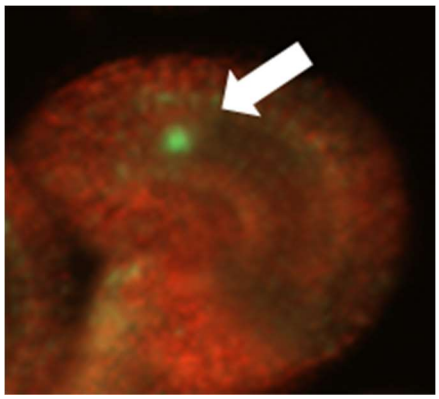

Figure 4. GFP Expression of DME

DME is expressed in the companion cells including the central cell from the female gametophyte

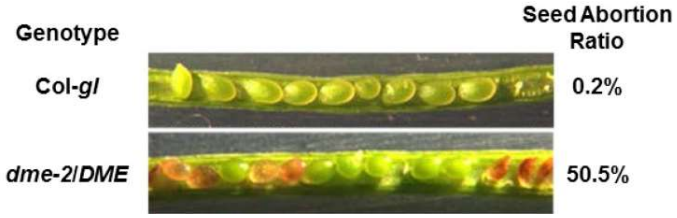

Figure 5. Seed Abortion Ratios for *dme* mutant plants

Col-gl (wild type) has a .2% seed abortion ratio whereas a *dme-2/DME* plant shows almost 50% seed abortion ratio phenotype due to the maternal effect (Park JS, 2017)

5. Structural analysis of DME

Structurally, *DME* has two splice variants: ATG5G04560.1 (Berkeley) and AT5G04560.2 (Spain). The Spain form is the representative model and is 258 residues longer on the N-terminal domain than the Berkeley form. DEMETER can also be split into three regions; the promoter, the N-terminal domain (NTD), and the C-terminal domain (CTD) (Figure 6). The promoter is controlled by positive and negative regulatory elements found between 46 and 473 base pairs downstream the transcription start site and is all that is required for proper DME expression and function (Park et al., 2017). The C-terminal domain contains three regions, the A domain, Glycosylase domain, and B domain. The A domain contains a mixed charge cluster (MCC) responsible for nonspecific DNA binding. The glycosylase domain of DME contains a helix-hairpin-helix (HhH) motif and a glycine/proline-rich loop with a conserved aspartic acid that is flanked by a [4Fe-4S] cluster (Mok et al., 2010).

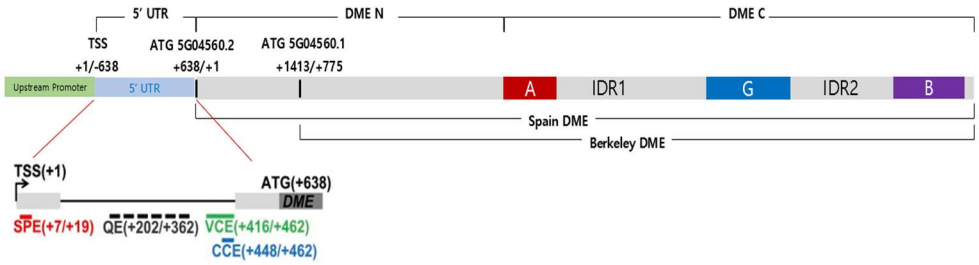

Figure 6. Structure of DEMETER and its Domains

Structure of DEMETER showing both its N-terminal Domain and C-terminal Domain. The N-terminal domain has two start sites: Spain DME starts at ATG.2 while Berkeley DME starts at ATG.1. Spain DME is 774 base pairs/258 amino acids longer on the NTD than Berkeley DME. The 5' UTR is found before ATG.2 and contains various elements that control DME function. The CTD contains the A, G, and B domains with two interdomain regions (IDR) between them.

6. Purpose of this study

Currently, there is no known function of the N-terminal Domain of DEMETER. Previous studies showed that the C-terminal Domain, containing the A, G, and B domains, is enough for DNA glycosylase function in vitro (Mok et al., 2010). The N-terminal Domain contains a stretch with unknown function (residues 1-120), a stretch of basic amino acid-rich direct repeats (residues 291-345), and region of unstructured, low complexity sequences (residues 346-947) (Figure 7).

However, previous studies on another DME family glycosylase, ROS1, gives us insight on some possible functions. When the N-terminal domain of ROS1 was removed, ROS1 activity on long DNA substrates were significantly reduced, while 5-mC excision was not altered in shorter molecules. Domain studies of ROS1 shows the N-terminal to have basic repeats acting as AT-hooks to bind strongly to nonspecific DNA. For short DNA molecules, the N-terminal Domain might not be needed; however, for longer molecules, the N-terminal Domain is used to bind onto DNA, possibly for more stability (Figure 8) (Ponferrada-Marin et al., 2010).

In this report I will try to elucidate the function of DEMETER's N-terminal Domain and also its relationship with its promoter. Because DME is in the DME family of glycosylase, it has a high chance it might act in a similar fashion to ROS1. Finally, this research will also try to find new information on promoter elements that might affect endogenous DME expression patterns.

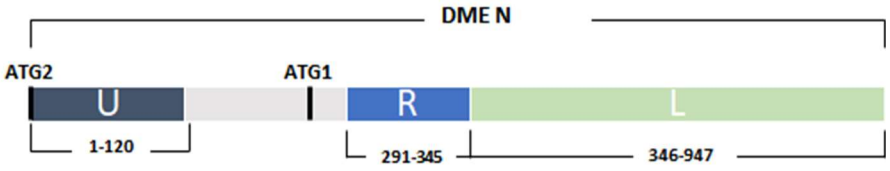

Figure 7. N-terminal Domain structure of DME

The N-terminal Domain contains a stretch of unknown function from amino acids 1-120 (U), a stretch of basic amino acid-rich direct repeats from amino acids 291-345 (R), and a region of unstructured, low complexity sequences from amino acids 346-947 (L).

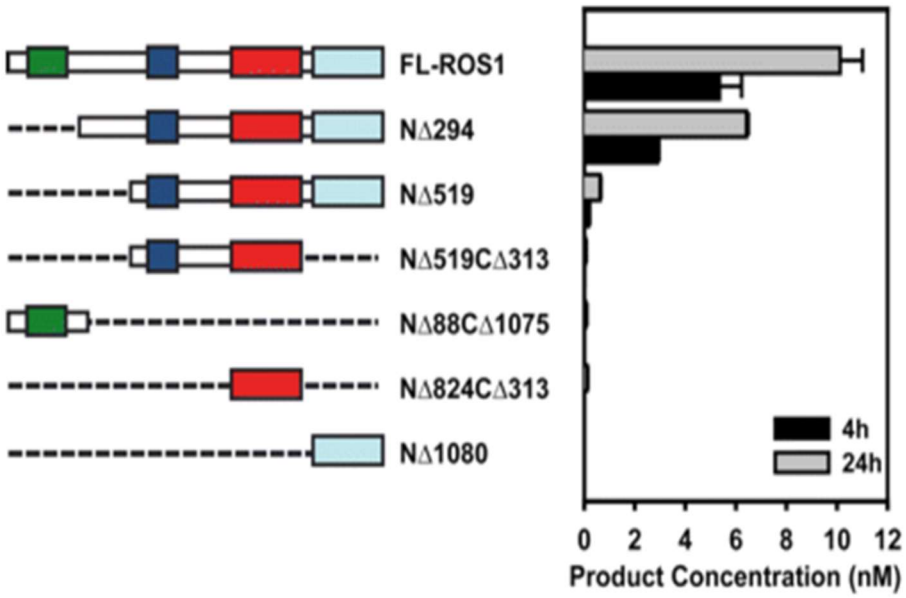

Figure 8. DNA glycosylase activity of truncated ROS1 polypeptides

Truncated ROS1 were incubated with a 51-mer double stranded oligionucleotide for 4h or 24h. Amount of incised oligonucleotide are shown. Green = Lysine rich domain, Red= HhH domain, Blue/Cyan = Unknown Function. (Ponferrada-Marin et al., 2010).

II. MATERIALS AND METHODS

1. Plant materials and growth conditions

Arabidopsis thaliana Columbia-0 (Col-0) was used as the wild type throughout these studies. Plants were grown in a growth room under long day conditions with 16 hours of light and 8 hours of darkness photoperiod at 22°C and 60% relative humidity under cool white fluorescent light (100 μmole/m^2/s). For transgenic plant screening, seeds were sterilized with a seed washing buffer containing 75% ethanol, 250uL TritonX-100 (SIGMA) for 15 minutes, before a 3-minute wash using 100% ethanol (MERCK). The seeds were dried on Whattman filter paper and plated on solidified MS agar plates containing the appropriate antibiotic (25 ug/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 minutes. Before transferring to the growth room, MS plates with seeds were also put in 4°C for cold treatment for 48 hours. Per construct, around 30 Col-0 plants were used for *Agrobacterium* mediated transformation by the floral dipping method.

2. Recombinant plasmid construction

To clone all the constructs used in these experiments, standard Gibson's Assembly (Isothermal Cloning) techniques were used. For the Yeast 2 Hybrid experiments, constructs were put into a GBKT7 vector using a SalI restriction enzyme cut site (NEB) and the proper primer sets. Constructs

inserted into a GADT7 vector were inserted using a BamHI restriction enzyme cut site and proper primer sets. For all other experiments, constructs used was a pBI-GFP vector. The vector was cut using a BamHI restriction enzyme cut site (NEB) and constructs were cloned using proper primer sets (Table 1). Proper primer sets are outlined in the results.

3. *Agrobacterium tumefaciens* transformation and plant transformation by floral dipping

The cloned constructs were introduced into *Agrobacterium* GV3101 by electroporation for transforming the Arabidopsis genome. For transformation, 50uL of Agrobacterium tumefaciens (GV3101) cell stock in a 1.5ml tube was thawed on ice for 10 minutes, and then 2uL of DNA of interest was added into the competent cell mixture by gentle pipetting. After transferring the mixture into a glass cuvette, it was pulsed by 1.8kV for 5.8mS with Micro-Pulser^{TM} (BIORAD). After electroporation, 400 ul of Luria Broth (LB) medium was added to the cuvette and the mixture was transferred into a 1.5mL tube and incubated in a shaking incubator at 250 rpm at 28^{o}C for 2 hours before spreading on solidified LB plate with proper antibiotics (50 ug/ml kanamycin, 50 ug/ml gentamycin). The plates were incubated at 28^{o}C for 48 hours.

For Arabidopsis transformation, the floral dipping method was used. Single colony of transformant Agrobacterium was selected and inoculated into 300mL of LB medium with adequate antibiotics (50 ug/mL kanamycin, 50ug/mL gentamycin) and grown for 48 hours at a 28^{o}C shaking incubator at 250rpm for a total amount of 4.752 x 10^{11} bacterial cells (roughly 330 mL of 1.44x 10^{9} cells/mL). Incubated Agrobacterium cell was harvested by

centrifuge at 4000 rpm for 20minute at 20C (SORVALL® RC 6 PLUS with SLC-3000 rotor). After discarding the supernatant, the cell pellet was resuspended in infiltration media. 300mL of infiltration media contains .63g of MS salt, 15g sucrose, and 150 uL of Silvet (Vac-In-Stuff, Silwet L-77, LEHLE SEEDS). For successful Arabidopsis transformation, Col-0 plants were grown on soil until adequate number of inflorescences were generated in a long day condition (24 plants per transformant). 10 days before transformation, all the plants are clipped to generate bushier plants. Right before transformation, siliques and open flowers were removed. Young buds were then dipped into the filtration media containing Agrobacterium for 10 seconds. After floral dipping, plants were paid on a tray and covered with a black tray to block light for 24 hours. The next day, the plants were uncovered, put vertically and grown until their seeds were mature enough to harvest.

4. Yeast 2 Hybrid Protocol

For DEMETER interaction, various segments of DME NTD were cloned into both pGBKT7 bait vector and pGADT7 prey vector. pGBKT7 bait vector used TRP1 as a selection marker and binding proteins were cloned into a pGADT7 prey vector with Leu as a selection in the Matchmaker Two-Hybrid system in yeast YH109 cells (Clontech). E3 enzyme with locus #AT5G37930 was inserted into a pGBKT7 vector using a SalI restriction enzyme cut site. The DME NTD used were cloned into both pGBKT7 and pGADT7 vectors using a BamHI restriction enzyme cut site from a +46 *DME::*DME:GFP template to create fragments labelled as Spain NTD (residues 1-947), 2) Berkeley NTD (residues 259-947), 3) Bear NTD (residues 407-947) Oski NTD (residues 259-407), 5) 13 NTD (residues 1-406)

and 6) Madrid NTD (residues 1-258). The CTD (residues 948-1987) was cloned into pGADT7. Assay conditions were as described by the manufacturer. From the transformed yeast colonies of each combination, I chose eight independent colonies and examined their growth on -Leu/-Trp/-Ade/-His quadruple dropout media also containing X-alpha-gal to determine interactions. A representative single colony was diluted with water and spotted on new quadruple dropout media to photograph.

5. Microscopy for GFP Analysis

Tissues were mounted using distilled water on glass slides. GFP expressing tissues 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

Label	Oligo name	Sequence
		Cloning
BC1	GBKT7_E3_F	GGCCGAATTCCCGGGGATGGCGAGATTCTCAGTTTGCGG
BC2	GBKT7_E3_R	GGTCGACGGATCCTACGAATGAACAAAGATCCGGCCATGG
BC3	GBKT7_Sp_F	CAGAGGAGGACCTGCATATGATGAATTCGAGGGCTGATCC
BC4	GBKT7_NTD_R	GGCCGCTGCAGGTCGACctaGCTCTCATAGGGAACAAGTGCAC
BC5	GBKT7_Bk_F	CAGAGGAGGACCTGCATATGATGCAGAGCATTATGGACTC
BC6	GBKT7_Bear_F	CAGAGGAGGACCTGCATATGGAAATTGTCCAGAACAGTAG
BC7	GBKT7_Oski_R	GGCCGCTGCAGGTCGACctaAGACTCAGAGTCACCTTGCCTCG
BC8	GBKT7_Madrid_R	GGCCGCTGCAGGTCGACctaACTTTTGTCAGATTTCGGCTCTG
BC9	GBKT7_CTD_F	CCACCCGGGTGGGCATCGATAAGAAGCGAAAACCAAGACCC
BC10	GBKT7_CTD_R	CGATTCATCTGCAGCTCGAGctaGGTTTTGTTGTTCTTCAATTTGC

BC11	**GADT7_Sp_F**	**CCACCCGGGTGGGCATCGATggATGAATTCGAGGGCTGATCC**
BC12	**GADT7_NTD_R**	**CGATTCATCTGCAGCTCGAGctaGCTCTCATAGGGAACAAGTG**
BC13	**GADT7_Bk_F**	**CCACCCGGGTGGGCATCGATggATGCAGAGCATTATGGACTC**
BC14	**GADT7_Bear_F**	**CCACCCGGGTGGGCATCGATggGAAATTGTCCAGAACAGTAG**
BC15	**GADT7_Oski_R**	**CGATTCATCTGCAGCTCGAGctaAGACTCAGAGTCACCTTGCC**
BC16	**GADT7_Madrid_R**	**CGATTCATCTGCAGCTCGAGctaACTTTTGTCAGATTTCGGCT**
BC17	**p1.7_F**	**TGCCTGCAGGTCGACctttaaaaaatcgtattttc**
BC18	**p1.7_R**	**GCTCACCATGGATCCaacacacttgatgaatcac**
BC19	**p1.7_Sp_F**	**gtgattcatcaagtgtgttGATGAATTCGAGGGCTGATC**
BC20	**p1.7_NTD_R**	**GCTCACCATGGATCCGCTCTCATAGGGAACAAGTG**
BC21	**p1.7_Bk_F**	**aagtgtgttGGATCCATGCAGAGCATTATGGACTC**
BC22	**NTD_1013_R**	**CCTTGCTCACCATGGATCGACCAGAACCACCTCCTTGTACCAG**
BC23	**NTD_1055_R**	**CCTTGCTCACCATGGATCGACCAGAACCACCTGGAGGGAATCG**

BC24	**pHTR10_F**	**TGCCTGCAGGTCGACagagagagagataaagtgag**
BC25	**pHTR10_R**	**GCTCACCATGGATCCttcttcgagagaacgatgatg**
BC26	**pHTR12_F**	**GCCTGCAGGTCGACtgtttctgtagacttcagattccacttctc**
BC27	**pHTR12_R**	**CCTTGCTCACCATGGATCCtgatttttactgctggag**
BC28	**pEC_F**	**TGCCTGCAGGTCGACaaacgcctatcatgaattag**
BC29	**pEC_R**	**GCTCACCATGGATCCttctcaacagattgataagg**
BC30	**pZOU_F**	**TGCCTGCAGGTCGACttaattaagggtcataattac**
BC31	**pZOU_R**	**CCTTGCTCACCATGGATCCattgaattgaatgctcatttta ccc**
BC32	**pAT3G10100_ F**	**TGCCTGCAGGTCGACggtttcgtgagagaggactt**
BC33	**pAT3G10100_ R**	**GCTCACCATGGATCCtgccggaaaaatcgattaag**
BC34	**pHTR10_Sp_F**	**catcgttctctcgaagaaGATGAATTCGAGGGCTGATC**
BC35	**pHTR10_DME _R**	**CCTTGCTCACCATGGATCGACCAGAACCACC GGTTTTGTTGTTCTTC**
BC36	**pHTR12_Sp_F**	**ctccagcagtaaaaatcaGATGAATTCGAGGGCTGATC**
BC37	**pHTR12_DME _R**	**CCTTGCTCACCATGGATCGACCAGAACCACC GGTTTTGTTGTTCTTC**
BC38	**pEC_Sp_F**	**cttatcaatctgttgagaaGATGAATTCGAGGGCTGATC**

BC39	**pEC_DME_R**	**CCTTGCTCACCATGGATCGACCAGAACCACC GGTTTTGTTGTTCTTC**
BC40	**pZOU_Sp_F**	**gtaaaatgagcattcaattcaatGATGAATTCGAGGGCTG ATC**
BC41	**pZOU_DME_ R**	**CCTTGCTCACCATGGATCGACCAGAACCACC GGTTTTGTTGTTCTTC**
BC42	**pAT3G10100_ Sp_F**	**cttaatcgatttttccggcaGATGAATTCGAGGGCTGATC**
BC43	**pAT3G10100_ DME_R**	**CCTTGCTCACCATGGATCGACCAGAACCACC GGTTTTGTTGTTCTTC**
BC44	**pHTR10_Bk_F**	**catcgttctctcgaagaaATGCAGAGCATTATGGACTC**
BC45	**pHTR12_Bk_F**	**ctccagcagtaaaaatcaATGCAGAGCATTATGGACTC**
BC46	**pEC_Bk_F**	**cttatcaatctgttgagaaATGCAGAGCATTATGGACTC**
BC47	**pZOU_Bk_F**	**gtaaaatgagcattcaattcaatATGCAGAGCATTATGGA CTC**
BC48	**pAT3G10100_ Bk_F**	**cttaatcgatttttccggcaATGCAGAGCATTATGGACTC**
BC49	**pDME_DME_ R**	**CCTTGCTCACCATGGATCGACCAGAACCACC GGTTTTGTTGTTCTTC**

III. RESULTS

1. Investigations of the possible binding partners of DEMETER's N-Terminal Domain (NTD)

1.1 DEMETER's Interaction with and Ubiquitin-ligase (E3)

Previous analysis showed that DEMETER's NTD had multiple interacting partners. A protein library screening using a yeast-2-hybrid was done using the NTD of both Spain (residues 1-947) and Berkeley cDME (residues 259-947). Both forms were ligated into a GBKT7 binding domain vector (bait) while the protein library was provided in the GADT7 activation domain vector (prey). In this screening, Berkeley cDME NTD was found to have multiple binding partners with a specific ubiquitin-ligase (E3), AT5G37930, being the most common (Park JS, unpublished). However, Spain cDME NTD exhibited autoactivation so it is unclear whether or not Spain cDME NTD has similar binding partners or not. To verify whether or not Berkeley NTD interacts with E3, and to test if Spain NTD also interacts with these enzymes, another yeast-2-hybrid experiment was conducted.

A ubiquitin ligase (E3), locus AT5G37930, was cloned into a GBKT7 vector, using primer sets BC1&BC2. Both Spain NTD (residues 1-947), using primer sets BC3&BC4, and Berkeley NTD (residues 259-947), using primer sets BC5&4, were put into both GBKT7 and GADT7 vector as redundancies in the case of autoactivation. However, when DME was put into a GBKT7 vector, auto activation occurred, forcing us to use the DME constructs in the GADT7 vector. Unfortunately, E3 causes autoactivation once in a GBKT7

vector as well preventing us from continuing on with this verification (Figure 9).

AD	Water	Spain NTD	T
BD			
Water	x	x	x
E3	Auto	Auto	Auto
Spain NTD	Auto	Auto	Auto
p53	x	x	o

Figure 9. Result of Yeast 2 Hybrid Results between a E3 ligase and DME
Table showing the results of the Yeast 2 Hybrid experiment. Both E3 and Spain NTD exhibit autoactivation when put into a GBKT7 (bait vector). p53 and T serves as a positive control where water serves as a negative.

1.2 DEMETER's Self-Interaction

The fact that DEMETER was found to have binding partner *in vitro* and having two splice variants, Spain DME and Berkeley DME, there exists a possibility that DEMETER can interact with itself. To verify 1) if DME NTD and DME CTD interacted and 2) if the NTD interacted with among themselves or not, a yeast-2-hybrid analysis was done.

Seven constructs of DEMETER were created and put into both the GBKT7 and GADT7 vectors as backups to in the case any issues of autoactivation arises. The N-Terminal containing constructs include 1) Spain NTD (residues 1-947) using primer sets BC3&BC4, 2) Berkeley NTD (residues 259-947) using primer sets BC5&BC4, 3) Bear NTD (residues 407-947) using primer sets BC6&BC4, Oski NTD (residues 259-407) using primer sets BC5&BC9, 5) 13 NTD (residues 1-406) using primer sets BC3&BC7, and 6) Madrid NTD (residues 1-258) using primer sets BC3&BC8. The C-terminal containing construct labeled as 7) CTD includes residues 948-1987 was put in a GBKT7 vector using primer sets BC9&BC10 (Figure 10) . When GADT7 constructs of Spain NTD, Berkeley NTD, Bear NTD, Oski NTD, 13 NTD, and Madrid NTD were made, primer sets BC11&BC12 / BC13&BC12 / BC14&BC12 / BC14&BC15 / BC11&BC15 / BC11&BC16, respectively, were used for cloning.

Interactions show GBKT7-Spain NTD / -13 NTD / -Madrid NTD having issues of autoactivation, which suggests residues 258 having an autoactivation feature. When GBKT7-Berkeley NTD and GADT7-Spain NTD interacted, there was no positive response. When GBKT7-CTD and GADT7-Spain NTD / - Berkeley NTD were interacted, there was no positive response. To test if there exists a possibility where a fragment of DME will interact with itself, GBKT7-Berkeley NTD / - Bear NTD / - Oski NTD was

tested against GADT7-Berkeley NTD / - Bear NTD / -Oski NTD / -13 NTD / -Madrid NTD. However, all of these interactions between bait and prey combinations showed no positive interaction. This suggests that DEMETER's NTD does not interact with its own NTD and its CTD (Figure 11).

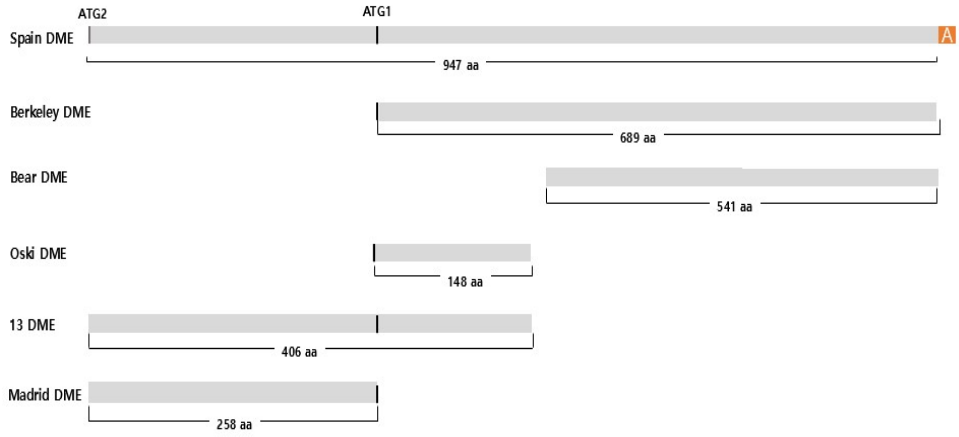

Figure 10. Constructs created for Y2H screening for DME NTD

Six constructs are based off the N-terminal with various segments chosen based off gene analysis. Spain N and Berkeley N are considered to be the full N terminal for each respective form. Bear N, Oski N, 13 N, and Madrid N are segments of the N terminal.

AD									
BD	Empty	Spain-N	Berkeley-N	Bear-N	Oski-N	13-N	Madrid-N	Water	T
Empty	x	x	x	x	x	x	x	x	x
Spain-N	A	A	A	A	A	A	A	A	A
Berkeley-N	x	x	x	x	x	x	x	x	x
Bear-N	x	x	x	x	x	x	x	x	x
Oski-N	x	x	x	x	x	x	x	x	x
13-N	A	A	A	A	A	A	A	A	A
Madrid-N	A	A	A	A	A	A	A	A	A
DME-C	x	x	x	x	x	x	x	x	x
Water	x	x	x	x	x	x	x	x	x
p53	x	x	x	x	x	x	x	x	o

Figure 11. Result of Yeast 2 Hybrid for DME NTD

Table showing the results of the Yeast 2 Hybrid experiment for DME interacting with itself. Spain NTD shows autoactivation when it is in the bait vector. Results show that when Spain NTD is in the prey vector, there is no interaction between the various NTD constructs or the CTD. Berkeley NTD does not interact with the NTD or CTD when put in the bait vector. BD-p53 and AD-T served as a positive control.

1.3 DEMETER's N-terminal cDNA Overexpression

Previous studies show that the C-terminal Domain region of DEMETER exhibits DNA glycosylase function *in vitro* without the presence of the N-terminal Domain (Mok et al., 2010). There is no current known function of the NTD of DEMETER. However, studies on ROS1, a DME family glycosylase, shows that the NTD of ROS1 is necessary for the binding of longer DNA molecules (Ponferrada-Marin et al., 2010).

An overexpression of DEMETER'S NTD was done to test whether DME NTD has a similar function to ROS1 NTD where it used for the binding of longer DNA molecules.

The promoter fusion lines 1.7kb *DME::Spain_NTD:GFP* (residues 1-947), and 1.7kb DME*::Berkeley_NTD:GFP* (residues 259-947) were created in a pBI-GFP vector using primer sets BC19&BC20 and BC21&BC20. A 1.7kb *DME:GFP* was created in a pBI-GFP vector to serve as a transcriptional control using primer sets BC17&BC18. If DME has a binding partner and/or DME NTD is the binding location, then endogenous DME can be blocked from competitive inhibition and expect *dme* mutant phenotypes. If DME does not have a binding partner, and/or if DME NTD is not the binding location of a partner, then normal glycosylase function and demethylation levels, and normal seed development are expected.

Transgenic lines were obtained for each construct. Seed abortion data (Figure 12) showed no significant difference between wild type.

DME::Spain_N:GFP	Normal Seeds	Aborted Seeds	Aborted/Normal %
Line #1	94	5	5.3%
Line #2	122	6	4.9%
Line #3	147	2	1.4%
Line #4	126	4	3.2%
Line #5	119	3	2.5%
Line #6	177	0	0%
Line #7	88	1	1.1%
Line #8	211	5	2.4%
Line #9	108	2	1.9%
Line #10	80	0	0%
Line #11	151	1	0.6%
Line #12	103	2	1.9%
Line #13	108	0	0%
Line #14	126	4	3.2%
Line #15	193	3	1.6%
Col-0 #1	155	3	1.9%
Col-0 #2	168	1	0.6%
Col-0 #3	141	0	0%
Col-0 #4	116	5	4.3%

DME::Berkeley_N:GFP	Normal Seeds	Aborted Seeds	Aborted/Normal %
Line #1	149	2	1.3%
Line #2	174	0	0%
Line #3	163	0	0%
Line #4	139	5	3.6%
Line #5	167	4	2.4%
Line #6	135	8	5.9%
Line #7	153	1	0.6%
Line #8	168	6	3.6%
Line #9	127	0	0%
Line #10	120	5	4.2%
Line #11	181	5	2.7%
Line #12	183	2	1.1%
Line #13	165	1	0.6%
Line #14	124	2	1.6%
Line #15	164	0	0%
Col-0 #1	142	1	0.7%
Col-0 #2	154	0	0%
Col-0 #3	188	1	0.5%
Col-0 #4	176	4	2.2%

Figure 12. Seed abortion data for *DME::DME_N:GFP* constructs

Data table showing seed abortion ratios between 15 transgenic lines of both *DME::Spain_NTD:GFP* and *DME::Berkeley_NTD:GFP*. No line shows significant seed abortions compared to the Col-0 controls.

1.4 DEMETER's A-Domain Effect on Expression Patterns

The A domain (residues 948-1055) can be split into roughly the first half (residues 948-1013) and the second half (resides 1014-1055). The first half of the A domain contains stretches of positively and negatively charged residues, labeled the mixed charge cluster, to be necessary for base excision activity through the binding of nonspecific DNA in a methylation independent manner *in vitro* (Mok et al., 2010).

To test the effects of the necessity of the two halves of the A domain *in vivo*, an overexpression was done to compete with binding sites with endogenous DME. If the A domain is necessary for nonspecific DNA binding *in vivo*, then we should expect a varying degree of *dme* mutant phenotypes based off transgene transmission ratios or lower demethylation levels/higher methylation levels from the NTD+A domain interference. Constructs tested were 1.7kb *DME::Spain_NTD(1-1013):GFP* / *-Spain_N (1-1055):GFP* / *-Berkeley_N(259-1013):GFP* / *-Berkeley_N(259-1055):GFP*. Primer sets BC19&BC22 / BC19&BC23 / BC21&BC22 / BC21&BC23 were used respectively to clone into a pBI-GFP vector.

Seed abortion ratio and GFP expression patterns were recorded. There was no observable difference in seed abortion ratios with wild type. Constructs Spain_N (1-1055) and Berkeley_N (259-1055) show GFP pattern in the micropylar end. Spain_N (1-1013) and Berkeley_N (259-1013) shows GFP expression pattern in the central cell. This suggests that the residues 958-1013, the mixed charge cluster, contributes to proper DME localization while the addition of residues 1055 causes interference (Figure 13).

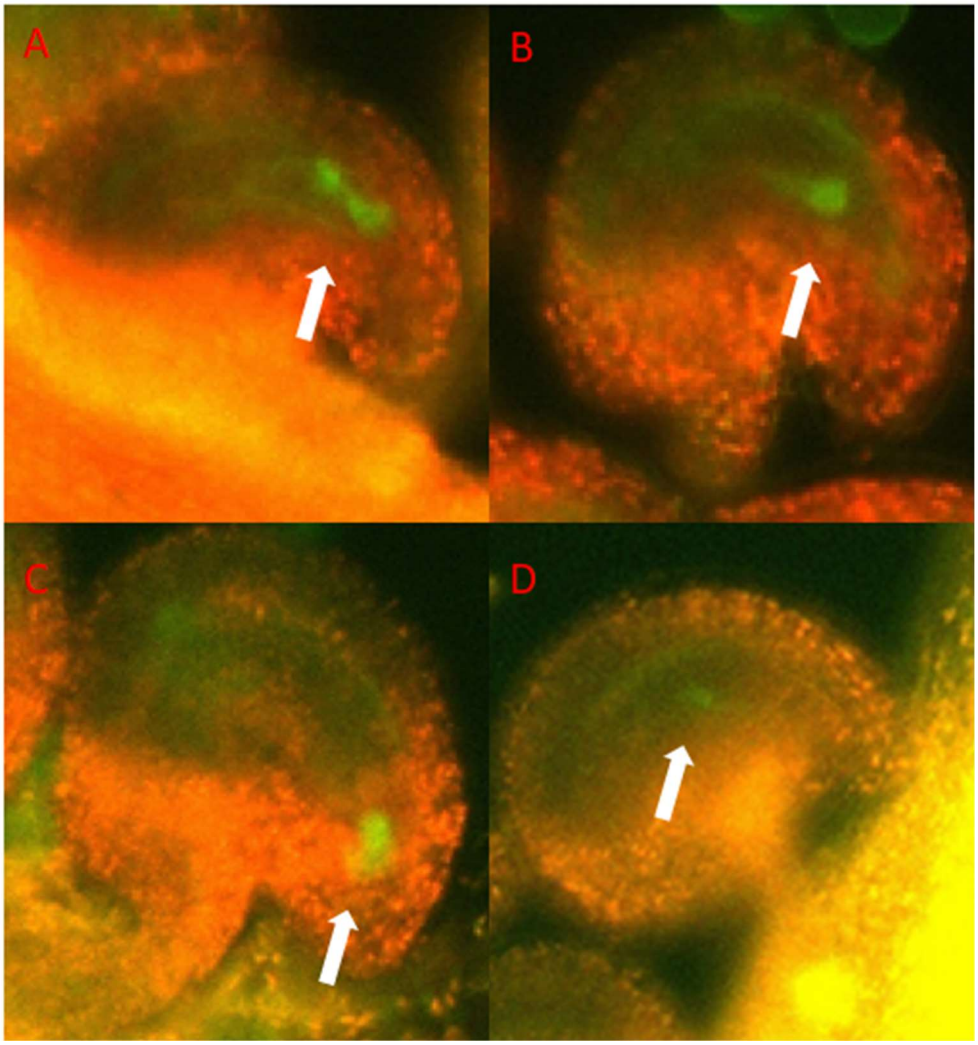

Figure 13. GFP expression of A domain constructs

A) GFP expression in micropyle end for *DME::Berkeley_N(259-1055):GFP*. B) GFP expression in the central cell for *DME::Berkeley_N(259-1013):GFP*. C) GFP expression for in micropyle end for *DME::Spain_N(1-1055):GFP*. D) GFP expression in the central cell for *DME::Spain_N(1-1013):GFP*.

1.5 DEMETER cDNA Ectopic Expression

To test if cDME has ectopic function, an overexpression was done in tissues that are not known to express *DME*. The tissues of interest are those related to *Arabidopsis* reproductive development including the egg cell, sperm cell, sperm cell and early zygote, embryo, and endosperm. From this overexpression, the goal is to see if any phenotypes will arise because of possible interacting partners. Both full length Spain and Berkeley DME cDNA are fused to six promoters including: DME, HTR10 ("Sperm 1"), HTR12 ("Sperm 2"), EC1.1 ("Egg"), ZOU ("Endosperm"), and AT3G10100 ("Embryo") (Figure 14).

Transcriptional controls for pSperm1, pSperm2, pEgg, pEndosperm, pEmbryo were created into a pBI-GFP vector using primer sets BC24&BC25 / BC26&BC27 / BC27&BC29 / BC30&BC31 / BC32&BC33 respectively. Spain DME cDNA was then inserted into the created constructs using primer sets BC34&BC35 / BC36&BC37 / BC38&BC39 / BC40&BC41 / BC42&BC43 respectively. Berkeley DME cDNA constructs were made using primer sets BC44&BC35 / BC45&BC37 / BC46&BC39 / BC47&BC41 / BC48&BC43 respectively. *pDME::Spain_DME:GFP* and *pDME::Berkeley_DME:GFP* were created using primer sets BC19&BC49 / BC21&BC49 respectively.

Transgenic lines were observed for any mutant phenotypes similar to *dme* mutants. There were no GFP expression or abnormal phenotypes when full DME was inserted into our constructs. However, transcriptional controls show promoter activity (Figure 15). This suggests either the transgene did not insert properly or GFP function is blocked due to protein folding.

Promoter	Expression
DME	Bicellular Pollen, Central Cell
HTR10	Sperm cell
HTR12 (CENH3)	Sperm cell, Early zygote
EC1.1	Egg cell
ZOU (RGE1)	Fertilized central cell, early endosperm
AT3G10100	from 2-cell stage embryo

Figure 14. List of Promoters used and Expression Patterns.

Table of promoters and expression patterns used for cDNA overexpression. Includes DME, HTR10, HTR12, EC1.1, ZOU, and AT3G10100 promoters.

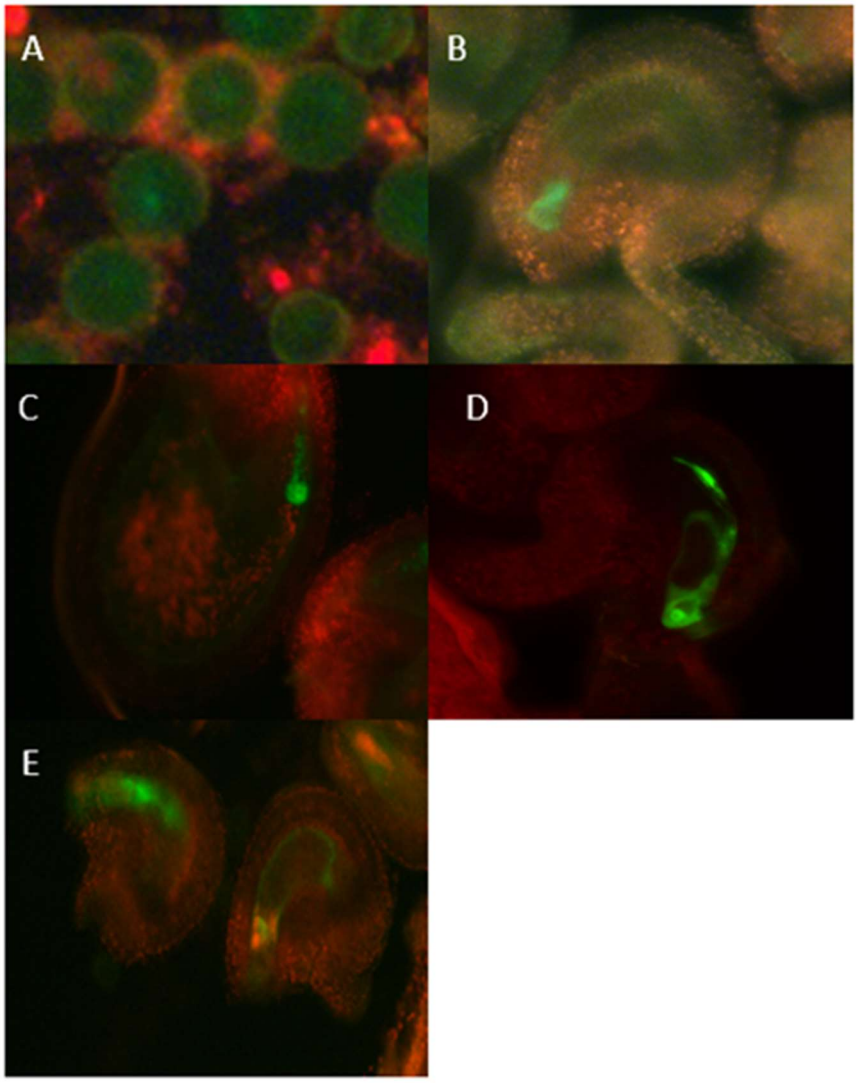

FIGURE 15. GFP photos of transcriptional controls for DME ectopic expression

Expected GFP expression for transcriptional controls. A) Sperm cell for HTR10:GFP. B) Early zygote for HTR12:GFP. C) 4 cell stage embryo for AT3G10100. D) Egg Cell for EC1.1:GFP. E) Endosperm and Late Central Cell for ZOU:GFP.

2. Analysis of the relationship between DEMETER's Promoter and its N-terminal Domain in expression patterns

Currently, published data on DME localization, using a 2.3kb *DME* promoter, is seen in the central cell of the female gametophyte (Choi, 2002). The promoter is controlled by positive and negative regulatory elements in the 5' untranscribed region (UTR) found between 46 and 473 base pairs downstream the transcription start site and is all that is required for proper DME expression and function (Park et al., 2017).

When a 1.7kb *DME* promoter transcriptional control fused to GFP was created, additional expressions in the endosperm and early embryo were observed (Figure 16). However, previous 1.7kb *DME::DME_NTD:GFP* (residues 1-947) constructs created only show expression in the bicellular pollen and the central cell. This suggests the possibility that either the promoter or the NTD influences DME localization. To test if the NTD influences the localization of DME, various segments of DME NTD were created to check for GFP expression patterns. These include 1.7kb *DME::*Oski NTD:GFP (residues 259-407) / - 13 NTD:GFP (residues 1-406)/ -Madrid NTD:GFP (residues 1-258) / -Alpha NTD:GFP (residues 1-104) (Figure 17). Constructs were ligated into a pBI-GFP vector using the primer sets BC21&BC20 / BC19&BC20 / BC19&50 / BC19&BC51 respectively.

GFP expression data was observed and constructs Oski NTD, 13 NTD, and Madrid NTD showed expression patterns similar to 1.7kb *DME::DME_NTD* (residues 1-947)*:GFP* constructs which is expression in the central cell (Figure 18). However, Alpha NTD exhibits GFP in the 4 cell endosperm stage and disappears soon after (Figure 18). This suggests that the

first 104 amino acids might possibly play a role in the interference or degradation of DME in the early endosperm.

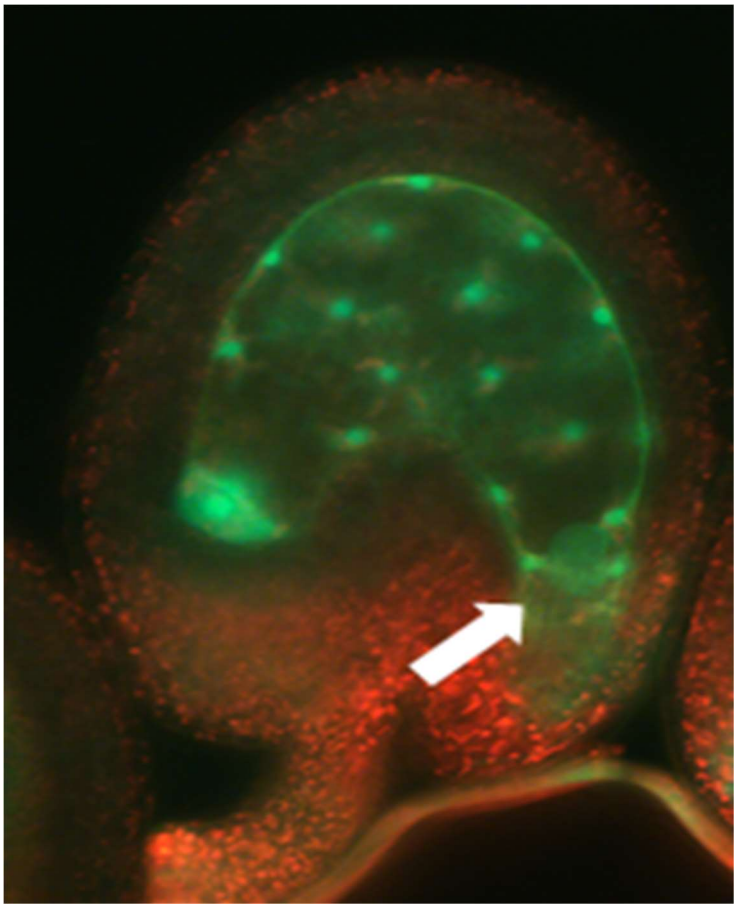

Figure 16. DME promoter expression

DME is being expressed in the early embryo and late endosperm.

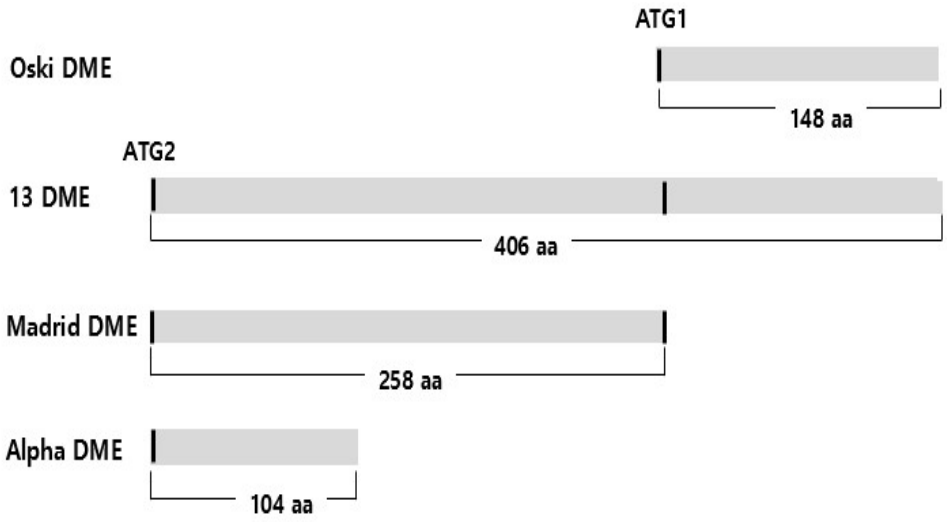

Figure 17. Constructs used to check if N-terminal affects DME expression. Alpha DME is the shortest construct, directly after the promoter and 104 amino acids long. Madrid DME consists of the difference between Spain and Berkeley DME. 13 DME contains Spain-N and the NLS. Oski only contains the NLS and starts at Berkeley ATG.

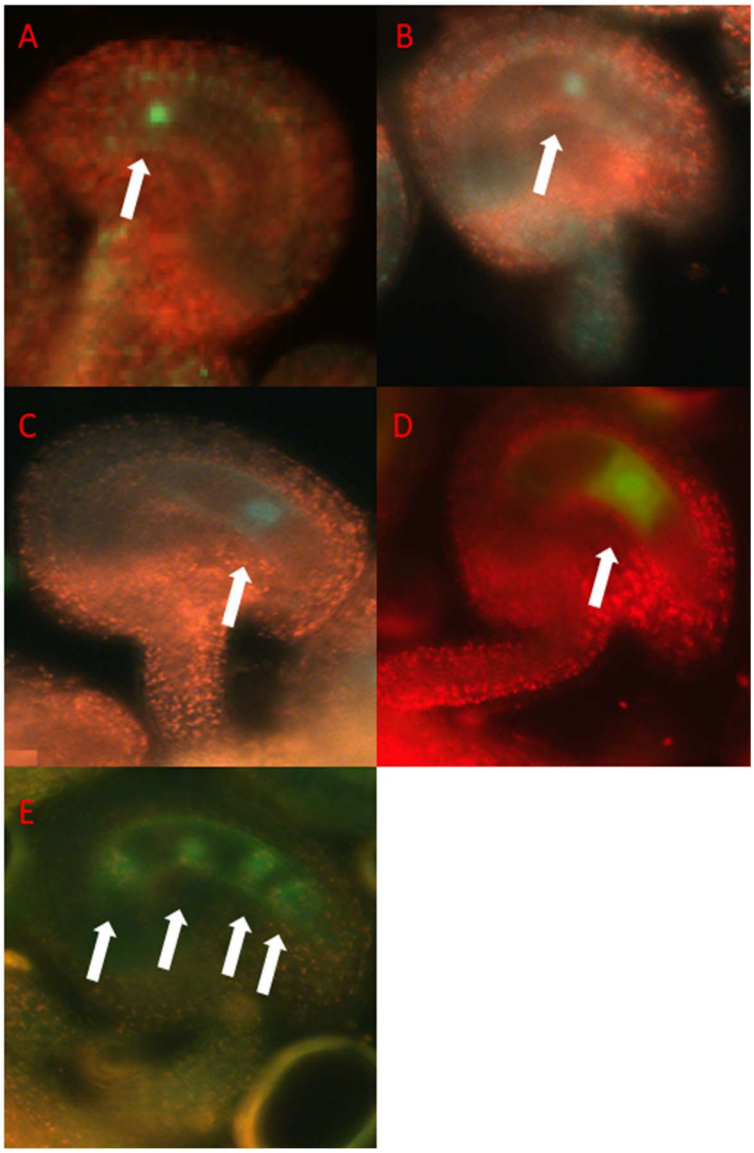

Figure 18. GFP expressions of shortened DME fragments

A) Central cell expression of *DME::Oski:GFP*. B) Central cell expression of *DME::*13:GFP. C) Central cell expression of *DME::Madrid:GFP*. D) Central cell expression of *DME::Alpha:GFP*. E) 4 cell stage endosperm of *DME::Alpha:GFP*.

IV. DISCUSSION

The N-terminal Domain of DEMETER serves an unknown function, if there is one. In a study by Mok et. al (2010), it was shown that DEMETER retains its glycosylase function *in vitro* with only its C-terminal Domain. However, the C-terminal domain can be split leaving only the A, G, and B domains and still retain glycosylase function making the interdomain regions between these 3 domains unnecessary. The N-terminal domain can also be split into three regions as well; a stretch with unknown function (residues 1-120), a stretch of basic amino acid-rich direct repeats (residues 291-345), and region of unstructured, low complexity sequences (residues 346-947).

ROS1, another glycosylase in the DME family, has a similar N-terminal domain structure as DME. When the N-terminal domain of ROS1 was removed, ROS1 activity on long DNA substrates were significantly reduced, while 5-mC excision was not altered in shorter molecules. Domain studies of ROS1 shows the N-terminal to have basic repeats acting as AT-hooks to bind strongly to nonspecific DNA. For short DNA molecules, the N-terminal Domain might not be needed; however, for longer molecules, the N-terminal Domain is used to bind onto DNA, possibly for more stability (Ponferrada-Marin et al., 2010).

To understand if DEMETER's acid-rich direct repeats (resudies 291-345) acts as a binding site, the N-terminal domain was overexpressed. From previous yeast 2 hybrid experiments, Berkeley NTD was found to interact with an E3 ligases among other binding partners. Thus, it is possible that the NTD serves as a binding site or acts similar to ROS1's AT hook where it is serves as stability for target DNA molecules. If the NTD serves as a binding site for either putative binding partners or for binding stability, then we should

expect endogenous DME not functioning properly leading to varying levels of *dme* like phenotypes. However, all transgenic lines did not show any *dme* mutant phenotypes. These results suggest the NTD does not interact with E3 ligases nor is it needed for stability. However, the possibility that other, unknown mechanisms influencing our results can exist. These two conclusions must be tested separately. To test either the NTD interacts with our E3 ligase, a BiFC experiment can be done. To test if the NTD is needed for DNA binding, constructs containing the NTD+CTD and CTD only can be tested in a electrophoretic mobility shift assay or testing for 5mC excision activity.

When the A-domain was introduced to the NTD and overexpressed, confounding results arose. When the NTD and the first half of the A domain, containing the mixed charged cluster which is thought to be important for nonspecific DNA binding, was overexpressed, normal DME localization and expression was found. However, if the full length A-domain is added to the NTD, then localization patterns shifts to the micropylar end representing the egg cell or early zygote stage. The MCC does not seem to block endogenous DME from functioning, whereas the full A domain causes some unexpected localization patterns. It is possible that the full A domain, or the 2nd half is needed for proper DNA binding for glycosylase function. In order to confirm this, another overexpression involving cutting the MCC out while keeping the 2nd half of the A domain must be done.

Currently, DME is known to be expressed only in the central cell for the female gametophyte. However, a DME transcriptional fusion construct shows DME to be expressed in the central cell, endosperm as well as early embryo. Interestingly, when the first 104 amino acids directly after the promoter was attached, expression up to the 4 cell endosperm was observed which quickly faded away. When any other lengths of DME were attached,

normal expression patterns was observed again. It is clear that the promoter has additional expressions, however, it is unclear whether the loss of these additional expressions are caused by the addition of a protein or from another mechanism such as being degraded when proteins have to be made. One way to test if the expression is caused by an addition of a protein is fusing a the DME promoter to a random protein and seeing its expression pattern. It is also possible that another element in the promoter controls the expression. To test this, cuts of DMEpro must be done and observed.

After all of these experiments, the function and relationship between DEMETER's NTD and promoter activity is unclear. However, more insights were obtained and give us a better lead on future experiments for possible functions.

V. REFERENCES

1. Berger SL. The complex language of chromatin regulation during transcription. Nature. 2007;447:407–12.

2. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB et al (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. Cell 110:33–42

3. Calarco JP, Borges F, Donoghue MT, Van Ex F, Jullien PE, Lopes T et al (2012) Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. Cell 151:194–205

4. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD *et al*. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature; 452: 215–219.

5. Gehring M, et al. (2006) DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. Cell 124(3):495–506.

6. Gifford E.M., Foster A.S. Morphology and Evolution of Vascular Plants. New York: W.H. Freeman and Company; 1989.

7. Gong Z, Morales-Ruiz T, Ariza RR, Roldan-Arjona T, David L, Zhu JK (2002) ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. Cell 111:803–814

8. Ibarra CA, Feng X, Schoft VK, Hsieh TF, Uzawa R, Rodrigues JA et al (2012) Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. Science 337:1360–1364

9. Kinoshita T, Yadegari R, Harada JJ, Goldberg RB, Fischer RL. (1999) Imprinting of the MEDEA polycomb gene in the Arabidopsis endosperm. Plant Cell, 11, pp. 1945-1952.

10. Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, et al. (2004) One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation. Science 303: 521–523.

11. Kouzarides, T. (2007) Chromatin Modifications and Their Function. Cell, 128, 693-705

12. Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 11:204–220

13. Li, Y., Kumar, S. & Qian, W. Plant Cell Rep (2018) 37: 77.

14. Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet 15:394–408

15. McCullough AK, Dodson ML, Lloyd RS (1999) Initiation of base excision repair: glycosylase mechanisms and structures. Annu Rev Biochem 68:255–285

16. Mok YG, Uzawa R, Lee J, Weiner GM, Eichman BF, Fischer RL et al (2010) Domain structure of the DEMETER 5-methylcytosine DNA glycosylase. Proc Natl Acad Sci USA 107:19225–19230

17. Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marin MI, Martinez-Macias MI, Ariza RR, Roldan-Arjona T (2006) DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. Proc Natl Acad Sci USA 103:6853–6858

18. Movahedi A, Sun W, Zhang J, Wu X, Mousavi M, Mohammadi K et al (2015) RNA-directed DNA methylation in plants. Plant Cell Rep 34:1857–1862

19. Ortega-Galisteo AP, Morales-Ruiz T, Ariza RR, Roldan-Arjona T (2008) *Arabidopsis*DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. Plant Mol Biol 67:671–681

20. Ponferrada-Marin MI, Martinez-Macias MI, Morales-Ruiz T, Roldan-Arjona T, Ariza RR (2010) Methylation-independent DNA

binding modulates specificity of repressor of silencing 1 (ROS1) and facilitates demethylation in long substrates. J Biol Chem 285:23030–23037

21. Park JS, Frost JM, Park K, Ohr H, Park GT, Kim 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

22. Reece, J. B., Urry, L. A., Cain, M. L. 1., Wasserman, S. A., Minorsky, P. V., Jackson, R., & Campbell, N. A. (2014). *Campbell biology* (Tenth edition.). Boston: Pearson.

23. Schoft VK, Chumak N, Choi Y, Hannon M, Garcia-Aguilar M, Machlicova A et al (2011) Function of the DEMETER DNA glycosylase in the *Arabidopsis thaliana* male gametophyte. Proc Natl Acad Sci USA 108:8042–8047

24. Skinner DJ and Sundaresan V. Recent advances in understanding female gametophyte development [version 1; referees: 2 approved]. *F1000Research* 2018, 7(F1000 Faculty Rev):804 (doi: 10.12688/f1000research.14508.1)

25. Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, Feijo JA et al (2009) Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell 136:461–472

26. Van Oosten MJ, Bressan RA, Zhu J-K, Bohnert HJ, Chinnusamy V (2014) The Role of the Epigenome in Gene Expression Control and the Epimark Changes in Response to the Environment. Critical Reviews in Plant Sciences; 33(1):64–87.

27. Vielle-Calzada JP, Thomas J, Spillane C, Coluccio A, Hoeppner MA, Grossniklaus U. (1999) Maintenance of genomic imprinting at the Arabidopsis medea locus requires zygotic DDM1 activity. Genes Dev, 13 pp. 2971-2982.

28. Xiao W, Gehring M, Choi Y, Margossian L, Pu H, Harada JJ et al (2003) Imprinting of the MEA Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. Dev Cell 5:891–901

29. Zhu JK (2009) Active DNA demethylation mediated by DNA glycosylases. Annu Rev Genet 43:143–166