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

A Thesis for the Degree of Master of Science

Cellular Response to Ectopic Expression of DNA Demethylase and its Application to Epigenome Editing

DNA 탈메틸 효소의 전위성 발현에 따른 세포 반응과 후성유전체 편집 기술을 위한 응용 연구

FEBRUARY, 2023

JUNWOO YANG

MAJOR IN HORTICULTURAL SCIENCE AND BIOTECHNOLOGY

DEPARTMENT OF AGRICULTURE, FORESTRY AND BIORESOURCES

THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

Cellular Response to Ectopic Expression of DNA Demethylase and its Application to Epigenome Editing

UNDER THE DIRECTION OF DR. JIN HOE HUH
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
SEOUL NATIONAL UNIVERSITY

BY
JUNWOO YANG

MAJOR IN HORTICULTURAL SCIENCE AND BIOTECHNOLOGY
DEPARTMENT OF AGRICULTURE, FORESTRY AND BIORESOURCES
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

FEBRUARY, 2023

APPROVED AS A QUALIFIED THESIS OF JUNWOO YANG
FOR THE DEGREE OF MASTER OF SCIENCE
BY THE COMMITTEE MEMBERS

CHAIRMAN _____
Doil Choi, Ph.D.

VICE-CHAIRMAN _____
Jin Hoe Huh, Ph.D.

MEMBER _____
Cecile Segonzac, Ph.D

Cellular Response to Ectopic Expression of DNA Demethylase and its Application to Epigenome Editing

JUNWOO YANG

DEPARTMENT OF AGRICULTURE, FORESTRY AND BIORESOURCES

THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

ABSTRACT

DNA methylation is a prominent epigenetic alteration that is implicated in genome stability and gene expression in higher eukaryotes. DNA methylation is dynamically regulated to maintain the epigenetic state of the organisms in response to developmental and environmental cues. Although the establishment of DNA methylation is conserved in both mammals and plants, DNA demethylation machineries have distinctively evolved. In *Arabidopsis*, the DEMETER (DME) family genes encode bifunctional 5mC DNA glycosylases that catalyze the excision of 5-methylcytosine (5mC) during DNA demethylation. Among them, DME plays a significant role during female gametogenesis and establishes imprinting that is

crucial for proper seed development. Previous studies verified the effect of plant-specific DNA glycosylase DME activity in human HEK-293T cells revealing that DME-activated transposable elements (TEs) and other repeat elements triggered interferon cascades leading to antiviral responses. In line with the stress response to DNA demethylation in animal cells, the effect of ectopic expression of *DME* was examined in the *Arabidopsis* protoplast cell system. The catalytic activity of DME resulted in the direct 5mC excision on the overall protoplast genome. Expression changes of stress-response genes were conspicuous as the effect of DME accumulates, along with early activations of TEs. In effort to develop epigenome editing tools, targeted DNA demethylation was assessed with the dCas9 editing system fused with DME. The dCas9-SunTag module was manipulated to enhance the efficiency of targeted DNA demethylation. The SunTag-DME fusion protein was able to excise 5mC *in vitro* with the targeting preference in a gRNA-dependent manner. Taken together, this work not only elucidates the impact of DME-mediated DNA demethylation in plant cells but also demonstrates the feasibility of its application to epigenome editing.

Key words: DNA methylation, dCas9-SunTag, DEMETER, epigenome editing, plant protoplasts.

Student number: 2021-28426

CONTENTS

ABSTRACT ····· i

CONTENTS ····· iii

LIST OF TABLES ····· vi

LIST OF FIGURES ····· vii

LIST OF ABBREVIATIONS ····· viii

LITERATURE REVIEWS ····· 1

1. DNA methylation

2. Plant-specific DNA demethylases

3. Epigenome editing

LITERATURE CITED ····· 14

CHAPTER 1. Ectopic expression of DEMETER DNA glycosylase induces changes in the expression of stress-response genes

ABSTRACT ·····22

INTRODUCTION ·····23

MATERIALS AND METHODS ·····27

Plant material and growth conditions

Protoplast isolation and PEG-mediated transfection

In situ 5mC labeling assay

RNA sequencing analysis

RESULTS ·····30

Ectopic expression of DME in plant protoplasts

The effect of DME on *Arabidopsis* protoplast genome

Gene expression profiles of constitutively DME expressing *Arabidopsis* protoplasts

Ectopic expression of DEMETER DNA glycosylase induces changes in the expression of stress-response genes

Transposons are reactivated by the effect of DME in the first 12 h

DICUSSION ·····44

REFERENCES ·····46

CHAPTER 2. Application of 5-methylcytosine DNA glycosylase to DNA-free epigenome editing

ABSTRACT ·····50

INTRODUCTION ·····52

MATERIALS AND METHODS ·····55

Cloning of SunTag-DME module construct

Expression and purification of the proteins

Single-guide RNA (sgRNA) design and *in vitro* transcription

In vitro 5mC excision assay of SunTag-DME

Targeting preference assay of SunTag-DME

RESULTS ·····60

Generation of DNA-free epigenome editing module with SunTag-DME

SunTag-DME is able to excise 5mC *in vitro*

The targeting preference of SunTag-DME is dependent upon a gRNA

DISCUSSION·····69

REFERENCES·····71

ABSTRACT IN KOREAN ·····75

LIST OF TABLES

CHAPTER 2

Table 2-1. Primers used in this study·····59

LIST OF FIGURES

CHAPTER 1

Figure 1-1. Schematic diagrams of the constructs used for the protoplast transfection ····· 32

Figure 1-2. Subcellular localizations of GFP in *Arabidopsis* protoplasts····· 33

Figure 1-3. Distribution of excised 5mC regions in *Arabidopsis* protoplasts ·· 35

Figure 1-4. Relative expression level of *DME* in the transfected protoplasts ·· 38

Figure 1-5. Identification of differentially expressed genes (DEGs) among Mock, DME, and KQ transfected protoplasts ····· 39

Figure 1-6. The gene ontology analysis in DME transfected protoplasts····· 41

Figure 1-7. Distribution of TEs with altered expression level ····· 43

CHAPTER 2

Figure 2-1. Schematic diagrams of SunTag-DME components····· 62

Figure 2-2. SDS-PAGE analysis of purified proteins ····· 63

Figure 2-3. Schematic diagram of A*tFWA* locus and sgRNA ····· 64

Figure 2-4. The *in vitro* 5mC excision activity of SunTag-DME module ···· 66

Figure 2-5. The target preference assay of SunTag-DME module····· 68

LIST OF ABBREVIATIONS

3’-PUA	3’-phosphor-α, β-unsaturated aldehyde
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
ABA	Abscisic acid
AGO	ARGONAUTE
AIW1	ABA-INDUCED WD40-REPEAT 1
AREB1	ABSCIC ACID-RESPONSIVE ELEMENT BINDING PROTEIN 1
BER	Base excision repair
CMT2	CHROMOMETHYLASE 2
CMT3	CHROMOMETHYLASE 3
CRISPR	Clustered regularly interspaced short palindromic repeat
DAPI	4',6-Diamidino-2-Phenylindole, dihydrochloride
DBSH	Double-stranded β-helix
DEG	Differentially expressed genes
DME	DEMETER
DML2	DEMETER-LIKE 2
DML3	DEMETER-LIKE 3
DNMT	DNA methyltransferases
DRM2	DOMAINS REARRANGED METHYLTRANSFERASE 2
FDR	False discovery rate
FIS2	FERTILIZATION INDEPENDENT SEED 2
FLC	FLOWERING LOCUS C
FMR1	Fragile X mental retardation-1
FT	FLOWERING LOCUS T
FWA	FLOWERING WAGENINGEN

GFP	Green fluorescent protein
GO	Gene ontology
GPD	Glycine/proline-rich loop with conserved aspartic acid
HAT	Histone acetyltransferase
HhH	Helix-hairpin-helix
hOGG1	Human 8-oxoguanine DNA glycosylase
KYP	KRYPTONITE
MEA	MEDEA
MET1	METHYLTRANSFERASE 1
NIC3	NICOTINAMIDASE 3
PRC2	Polycomb repressive complex 2
RdDM	RNA-directed DNA methylation
RNP	Ribonucleoprotein
ROS1	REPRESSOR OF SILENCING 1
SAM	S-adenosyl methionine
scFv	Single chain variable fragment
sgRNA	Single-guide RNA
siRNA	Small interfering RNA
SRA	SET and RING associated
SUVH4	SUPPRESSOR OF VARIEGATION 3-9 HOMOLOG 4
TALE	Transcription activator-like effector
TDG	Thymine DNA glycosylase
TE	Transposable element
TET	TEN-ELEVEN TRANSLOCATION
UHRF1	Ubiquitin-like PHD and RING finger domains 1
VIM	VARIANT IN METHYLATION
ZF	Zinc-finger

LITERATURE REVIEWS

1. DNA methylation

DNA methylation is a prevalent epigenetic modification that generally refers to the covalent addition of a methyl group to C5 position of cytosine forming 5-methylcytosine (5mC) (Law and Jacobsen, 2010). 5mC within the genome context is correlated with diverse biological processes including regulation of gene expression, imprinting and X chromosome inactivation (Huh et al., 2008; Law and Jacobsen, 2010; Smith and Meissner, 2013). In mammals, DNA methylation exists predominantly in the symmetric CG dinucleotides contexts. By contrast, 5mC appear in all DNA sequence contexts in plants: the symmetric CG and CHG or the asymmetric CHH, in which H indicates any nucleotide except G (Law and Jacobsen, 2010).

The conversion of cytosine to 5mC is catalyzed by DNA methyltransferases (DNMTs), transferring the methyl group from S-adenosyl methionine (SAM) (Roberts and Cheng, 1998). Both mammals and plants retain two distinct enzyme families to establish and propagate DNA methylation: *de novo* DNA methyltransferases and maintenance DNA methyltransferases (Li and Zhang, 2014). In vertebrates, DNA methyltransferase 3 (DNMT3) family proteins govern *de novo* methylation by binding to unmethylated lysine residue on histone H3 tail (H3K4) through their amino-terminal cysteine-rich domain during early embryogenesis (Ooi et al., 2007). Once established, DNMT1, responsible for maintaining the intact methylation patterns, interacts with ubiquitin-like PHD and RING finger domains 1 (UHRF1) which has a strong preferential for hemimethylated CG sites resulted from

DNA replication due to its SET and RING associated (SRA) domain (Bostick et al., 2007; Law and Jacobsen, 2010).

In plants, *de novo* methylation is mediated by the RNA-directed DNA methylation (RdDM) pathway requiring small interfering RNAs (siRNAs). In the canonical RdDM pathway, siRNA-bound ARGONAUTE (AGO) proteins interact with DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), the plant homolog of DNMT3, which directs DNA methylation in a sequence-independent manner (Law and Jacobsen, 2010; Zhang et al., 2018). Perpetuation of plant DNA methylation is catalyzed by different classes of enzymes. METHYLTRANSFERASE 1 (MET1), a homolog of DNMT1, is in charge of maintaining CG DNA methylation (Zhang et al., 2018). MET1 is recruited by the VARIANT IN METHYLATION (VIM) family of SRA domain proteins which are E3 ubiquitin-protein ligase UHRF1 orthologs in *Arabidopsis*, similar to that of DNMT1 in mammals (Woo et al., 2007). Plant-specific CHH and CHG methylation is mediated by CHROMOMETHYLASE 2 (CMT2) and CMT3, respectively (Lindroth et al., 2001; Zemach et al., 2013). CMT3 binds methylated histone H3 tails, while an SRA domain of SUPPRESSOR OF VARIEGATION 3-9 HOMOLOG 4 (SUVH4) histone methyltransferase recognizes CHG DNA methylation, suggesting that methylated CHG and histone H3 lysine 9 dimethylation (H3K9me2) reinforce one another (Du et al., 2014). The asymmetric CHH methylation is maintained by CMT2 in heterochromatic region, while DRM2 propagates 5mC in CHH contexts through the RdDM pathway at transposons or other repeat sequences (Zemach et al., 2013; Zhang et al., 2018).

DNA demethylation refers to the removal of methylation, achieved by

passive or active mechanisms (Wu and Zhang, 2010). Passive DNA demethylation occurs when DNMT1 or MET1-mediated DNA methylation is diluted during successive rounds of DNA replication (Wu and Zhang, 2010). In contrast, 5mC can be erased in a replication-independent manner, accomplished by certain enzymatic reaction (Wu and Zhang, 2014). Moreover, specific enzymes are committed to reverting 5mC to cytosine in mammals and plants (Wu and Zhang, 2010).

In mammals, TEN-ELEVEN TRANSLOCATION (TET) family proteins, being iron (II)/α-ketoglutarate dioxygenases, catalyze the iterative oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Pastor et al., 2013; Tahiliani et al., 2009). The catalytic core domain of TET proteins includes a double-stranded β-helix (DBSH) domain and a cysteine-rich domain, engaged in bringing 5mC and stabilizing enzyme-DNA interaction, respectively (Hu et al., 2013; Zhang et al., 2010). The 5mC oxidation derivatives, 5fC and 5caC, undergo thymine DNA glycosylase (TDG)-mediated excision coupled with a subsequent base excision repair (BER) pathway, accomplishing TET-dependent active DNA demethylation (He et al., 2011; Weber et al., 2016).

In comparison to mammals, plants utilize *DEMETER* (*DME*) family genes that encode bifunctional DNA glycosylases-lyases which can excise 5mC directly and produce a single strand break during DNA demethylation (Choi et al., 2002; Gehring et al., 2006; Gong et al., 2002). Being bifunctional DNA glycosylases, DME family proteins catalyze 5mC excision by hydrolyzing the glycosylic bond, followed by β- and δ- elimination reactions, giving rise to 3'-phosphor-α, β-unsaturated aldehyde (3'-PUA) and 3'-phosphate, respectively (Agius et al., 2006; Gehring et al.,

2006; Morales-Ruiz et al., 2006). Finally, the gap generated by DNA glycosylase is filled by downstream enzymes in the BER pathway, resulting in the replacement of 5mC with cytosine (Martinez-Macias et al., 2012).

2. Plant-specific DNA demethylases

DNA demethylation is governed by four members of DME family glycosylases in *Arabidopsis*, including REPRESSOR OF SILENCING 1 (ROS1), DEMETER-LIKE 2 (DML2) and DEMETER-LIKE 3 (DML3) (Penterman et al., 2007). They share three common domain structures in C-terminal regions. The glycosylase domain possesses a helix-hairpin-helix (HhH) motif with an iron-sulfur [4Fe-4S] cluster and a glycine/proline-rich loop with conserved aspartic acid (GPD), homologous to the HhH-GPD class DNA glycosylase superfamily which are also found in human 8-oxoguanine DNA glycosylase (hOGG1) or adenine DNA glycosylase (MutY) and endonuclease III (EndoIII) in *Escherichia coli* (Choi *et al.*, 2002; Choi et al., 2004; Guan et al., 1998; Kuo et al., 1992; Mok et al., 2010). Moreover, not only invariant aspartic acid and lysine residues in the HhH motif, but also the conserved cysteine residues to form an iron-sulfur cluster are pivotal to 5mC excision activity of DME glycosylase (Choi et al., 2004; Mok et al., 2010). Recently, the way in which ROS1 recognizes and captures the substrate DNA has been elucidated. The [4Fe-4S] cluster-binding subdomain and a six-helix barrel subdomain both of which are the subset of the glycosylase domain, and the C-terminal domain compose a base-binding pocket which not only recognizes 5mC by hydrogen bonding between asparagine and guanine base but also captures the substrate 5mC (Du et al., 2023). However, two additional domains flanking the glycosylase domain in the center are known to mediate nonspecific DNA binding and promote DME recruitment, but their definite roles remain elusive (Iyer et al., 2011; Mok et al., 2010).

Despite their structural similarity, they have distinct biological roles (Law

and Jacobsen, 2010). DME plays a significant role during gametogenesis on the cellular reprogramming that establishes imprinting (Huh et al., 2008). DME is primarily expressed in the central cell of the female gametophyte before fertilization (Choi et al., 2002). DME-mediated DNA demethylation activates *MEDEA* (*MEA*), *FERTILIZATION INDEPENDENT SEED 2* (*FIS2*) and *FLOWERING WAGENINGEN* (*FWA*) maternal allele expression through antagonizing MET1, while their respective paternal alleles are silenced (Choi et al., 2002; Gehring et al., 2006; Jullien et al., 2006; Kinoshita et al., 2004). Since MEA and FIS2 are the components of the floral polycomb repressive complex 2 (PRC2) necessary for proper seed development, *dme* mutation displayed the loss of seed viability, resembling the *mea* and *fis2* mutants (Choi et al., 2002; Huh et al., 2008). Especially, heterozygous *dme* mutants show a 50% of seed abortion phenotype, disobeying Mendelian rules was due to the maternal inheritance of *MEA* allele (Choi et al., 2002; Gehring et al., 2006).

In contrast to reproductive cell-specific expression of *DME*, other family members, ROS1, DML2 and DML3 facilitate DNA demethylation in somatic cells and function redundantly at the boundaries between euchromatin and heterochromatin in order to restrain the propagation of DNA methylation (Gong et al., 2002; Ortega-Galisteo et al., 2008; Penterman et al., 2007). However, the loss-of-function of these genes does not show discernable phenotypes under optimum growth conditions. Nevertheless, *ros1* mutants showed higher susceptibility to the fungal pathogen, *Fusarium*, due to the down-regulation of defense-related genes, along with increased methylation levels at the transposon residing in their promoters (Schumann et al., 2017). Additionally, *ros1* mutants displayed hypersensitivity to

abscisic acid (ABA) at their early seedling development stages as well as a decrease in the expression level of some ABA-inducible genes. *NICOTINAMIDASE 3* (*NIC3*), among those down-regulated genes, was hypermethylated at the proximal region of the promoter, suggesting that ROS1-dependent DNA demethylation is responsible for both gene activation and ABA responses (Kim et al., 2019). Taken together, it is conceivable that a subset of stress-response genes is under control of ROS1-mediated DNA demethylation (Kim et al., 2019; Zhang et al., 2018)

Furthermore, several studies on *dme ros1 dml2 dml3* (*drdd*) quadruple mutants have been performed in order to elucidate the functions of the DME family proteins in the vegetative tissues. Since *dme* mutants show a seed abortion phenotype whereby reproductive defects need to be overcome, central cell-specific *DME* complementation strategies were adopted (Williams et al., 2022; Zeng et al., 2021). Although DME functions redundantly in somatic tissues along with ROS1, DML2, and DML3, those four DNA demethylases are guided to targets in a tissue-specific manner, with distinctive methylation patterns and gene expression profiles (Williams et al., 2022). *FLOWERING LOCUS C (FLC)*, the repressor of flowering in *Arabidopsis*, was down-regulated in *drdd* quadruple mutants, thereby promoting the transition to flowering compared with wild type, *dme*, and *ros1 dml2 dml3* triple mutants. The early-flowering phenotype in *drdd* mutants was correlated with its hypermethylation at the upstream of transcriptional start site of *FLC*, in which DNA methylation is established via the RdDM pathway. Therefore, four DME family enzymes are responsible for maintaining DNA methylation patterns in *Arabidopsis*, while DME is decisive in achieving full demethylation in the vegetative tissues (Williams et al., 2022). Moreover, *drdd* mutants were more susceptible to pathogens

relative to wild type and *ros1 dml2 dml3* triple mutants. as the expression of a number of defense-related genes was diminished (Zeng et al., 2021).

3. Epigenome editing

Epigenetic alterations generally refer to the structural modification of DNA or chromatin state (Duan et al., 2018; Morikawa et al., 2014). DNA methylation, a prominent epigenetic mark in eukaryotes, is associated with diverse biological processes, including regulation of gene expression, imprinting and X chromosome inactivation (Huh et al., 2008; Law and Jacobsen, 2010; Smith and Meissner, 2013). Histone modifications are diverse post-translational modifications occurring on specific amino acid residues of the histone tails, such as methylation, acetylation, phosphorylation, sumoylation and ubiquitination (Iwasaki and Paszkowski, 2014; Lauria and Rossi, 2011). These heritable and reversible epigenetic reprogramming can be triggered by external environmental factors (Feil and Fraga, 2012).

Genome editing allowing target DNA sequence manipulation has developed with programmable DNA binding modules such as zinc-finger (ZF) proteins, transcription activator-like effector (TALE) proteins tethered with *Fok* I nuclease and the prokaryote-derived clustered regularly interspaced short palindromic repeat (CRISPR)-associated nuclease Cas9 (Yin et al., 2017). Along with genome editing, a control of epigenetic traits via epigenome editing has emerged in order to achieve transcriptional regulation without manipulating DNA sequence with a key feature of aforementioned editing modules, a DNA-binding ability (Kungulovski and Jeltsch, 2016). DNA binding modules such as ZF, TALE, and nuclease-deactivated Cas9 (dCas9) are physically linked to effector proteins (Kim, 2016). The application to manipulate gene regulation by imposed localization was first demonstrated with artificially designed ZF proteins as transcription factors in eukaryotic cells (Beerli et al., 2000). However, while ZF and TALE are expensive and laborious in protein

engineering and assembly, dCas9, containing mutations at both RuvC and HNH nuclease domains, provides an advantageous platform owing to its RNA-guided DNA targeting system (Kim, 2016; Klug, 2010). Therefore, in the CRISPR era, CRISPR/dCas9 platforms become predominant not only in genome editing but also in modifying epigenetic marks (Gilbert et al., 2013; Hilton et al., 2015; Qi et al., 2013; Thakore et al., 2015). Several researches on epigenome editing have been reported in several organisms. TET1, a mammal-specific DNA demethylase, was fused with CRISPR/dCas9 platform to rescue *fragile X mental retardation-1* (*FMR1*) gene expression in the mouse by removing DNA methylation in hypermethylated region causing Fragile X syndrome (Liu et al., 2018). Moreover, selective DNA demethylation at the promoter region of *BRCA1* human cancer gene, leading to breast and cervical cancer when heavily methylated, was achieved by CRISPR/dCas9-TET1 and restored *BRCA1* expression (Choudhury et al., 2016).

In plants, several studies about targeted manipulation of both DNA methylation and demethylation have been reported. The first targeted manipulation of epigenetic traits performed in *Arabidopsis* applied ZF protein tethered with SUVH2 protein. As SUVH2 possesses a SRA domain which preferentially binds to CG methylation site followed by establishing DNA methylation upon by recruiting RNA POLYMERASE V (Pol V) in late RdDM pathway, ZF-SUVH2 fusion protein was sufficient to gain DNA methylation at the unmethylated region in *fwa-4* mutants inducing *FWA* gene silencing and heritable late flowering phenotype (Johnson et al., 2014; Law and Jacobsen, 2010). Aside from adopting RdDM pathway components to establish targeted DNA methylation, bacterial CG-specific DNA methyltransferases were utilized. *Mollicutes* s*piroplasma*, strain MQ1, CG

methyltransferase M.SssI (SssI) was deployed with ZF protein directed against the *FWA* promoter in *Arabidopsis*. The ZF-SssI fusion protein established heritable DNA methylation in the *FWA* promoter and other off-target sites (Liu et al., 2021). In addition, $MQ1^{Q147L}$, a MQ1 variant (MQ1v) with reduced DNA binding affinity, which was introduced to mitigate genome-wide ectopic CG methylation revealed minimal off-target sites when fused with CRISPR/dCas9 (Ghoshal et al., 2021; Lei et al., 2017).

Similar platforms were utilized to target DNA demethylation in plants. The human TET1 catalytic domain (hTET1cd) was tethered with ZF protein to remove DNA methylation at the *FWA* promoter site-specifically. The fusion protein elicited decreased DNA methylation level and led to *FWA* re-activation, followed by heritable late flowering phenotype. Additionally, targeted demethylation at the heterochromatic transposable element – *CACTA1* was performed with identical construct. Contrary to the findings in *FWA*, DNA demethylation at *CACTA1* was insufficient, and erased DNA methylation was re-established once the transgene was segregated out (Gallego-Bartolome et al., 2018).

Various researches were performed to enhance the targeted effect by recruiting multiple effectors into a single locus. dCas9-SunTag is composed of two main modules, dCas9 protein attached to GCN4 epitope tail repeats and a single chain variable fragment (scFv) GCN4 antibody fused to green fluorescent protein (GFP) followed by the effector protein. Harnessing this system, a potent increase in targeted manipulation of epigenetic modification was achieved by recruiting multiple copies of antibody-fusion protein (Tanenbaum et al., 2014). MQ1v and the catalytic domain of *Nicotiana tabacum* DRM methyltransferase (NtDRMcd) were

incorporated into dCas9-SunTag system adopted for targeted methylation to the *FWA* promoter in *Arabidopsis*. Both SunTag system exhibited improved targeting efficiency and in DNA methylation installation compared to straight fusions with dCas9 protein (Ghoshal et al., 2021; Papikian et al., 2019). Moreover, combining more than two distinct guide RNAs simultaneously in a single construct within the *FWA* promoter led to strong DNA methylation over the target locus (Ghoshal et al., 2021; Papikian et al., 2019). In addition, dCas9-SunTag based targeting systems for DNA demethylation elicited enhanced heritable loss of DNA methylation with minor off-target effects (Gallego-Bartolome et al., 2018).

Similarly, various loop structures were adjusted to protrude outside of the dCas9-gRNA ribonucleoprotein (RNP) complex, where MS2 bacteriophage coat proteins bind in order to facilitate the addition of protein-interacting effector protein (Konermann et al., 2015). A few researches were reported in manipulating histone modifications with dCas9-MS2 platforms. The mammalian p300 histone acetyltransferase (HAT) recruited through MS2 to target the flowering regulator *FLOWERING LOCUS T (FT)* locus successfully increased the H3K27 acetylation levels in plants, but revealed marginal effect on gene expression and flowering time (Lee et al., 2019). On the other hand, H3K9 methyltransferase (HMT) G9a and KRYPTONITE (KYP) were utilized for targeted histone methylation at the *FT* locus. Both H3K9 methyltransferase constructs were anticipated to cause late flowering phenotype, and KYP-derived shuttles were able to flower late in T1 populations, while G9a could not affect flowering time. However, dCas9-MS2 linked KYP did not show considerable establishment of H3K9 dimethylation at the targeted *FT* locus (Lee et al., 2019).

In addition to targeting epigenetic alterations, dCas9 platforms were employed in terms of biotic stresses. *Arabidopsis* HAT was fused with dCas9 to target the promoter region of the *ABSCIC ACID-RESPONSIVE ELEMENT BINDING PROTEIN 1* (*AREB1*), which improved drought stress tolerance significantly as well as an up-regulation of *AREB1* and higher chlorophyll content in *Arabidopsis* (de Melo et al., 2020; Roca Paixao et al., 2019).

LITERATURE CITED

Agius, F., Kapoor, A., and Zhu, J.K. (2006). Role of the *Arabidopsis* DNA glycosylase/lyase ROS1 in active DNA demethylation. Proc. Natl. Acad. Sci. U S A 103, 11796-11801.

Beerli, R.R., Dreier, B., and Barbas, C.F., 3rd (2000). Positive and negative regulation of endogenous genes by designed transcription factors. Proc. Natl. Acad. Sci. U S A 97, 1495-1500.

Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 317, 1760-1764.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., and Fischer, R.L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. Cell 110, 33-42.

Choi, Y., Harada, J.J., Goldberg, R.B., and Fischer, R.L. (2004). An invariant aspartic acid in the DNA glycosylase domain of *DEMETER* is necessary for transcriptional activation of the imprinted *MEDEA* gene. Proc. Natl. Acad. Sci. U S A 101, 7481-7486.

Choudhury, S.R., Cui, Y., Lubecka, K., Stefanska, B., and Irudayaraj, J. (2016). CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at *BRCA1* promoter. Oncotarget 7, 46545-46556.

de Melo, B.P., Lourenco-Tessutti, I.T., Paixao, J.F.R., Noriega, D.D., Silva, M.C.M., de Almeida-Engler, J., Fontes, E.P.B., and Grossi-de-Sa, M.F. (2020). Transcriptional modulation of *AREB-1* by CRISPRa improves plant physiological performance under severe water deficit. Sci. Rep. 10, 16231.

Du, J., Johnson, L.M., Groth, M., Feng, S., Hale, C.J., Li, S., Vashisht, A.A., Wohlschlegel, J.A., Patel, D.J., and Jacobsen, S.E. (2014). Mechanism of DNA methylation-directed histone methylation by KRYPTONITE. Mol. Cell 55, 495-504.

Du, X., Yang, Z., Xie, G., Wang, C., Zhang, L., Yan, K., ... & Du, J. (2023). Molecular basis of the plant ROS1-mediated active DNA demethylation. Nat. Plants, 1-9.

Duan, C.G., Zhu, J.K., and Cao, X. (2018). Retrospective and perspective of plant epigenetics in China. J. Genet. Genomics 45, 621-638.

Feil, R., and Fraga, M.F. (2012). Epigenetics and the environment: emerging patterns and implications. Nat. Rev. Genet. 13, 97-109.

Gallego-Bartolome, J., Gardiner, J., Liu, W., Papikian, A., Ghoshal, B., Kuo, H.Y., Zhao, J.M., Segal, D.J., and Jacobsen, S.E. (2018). Targeted DNA demethylation of the *Arabidopsis* genome using the human TET1 catalytic domain. Proc. Natl. Acad. Sci. U S A 115, E2125-E2134.

Gehring, M., Huh, J.H., Hsieh, T.F., Penterman, 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.

Ghoshal, B., Picard, C.L., Vong, B., Feng, S., and Jacobsen, S.E. (2021). CRISPR-based targeting of DNA methylation in *Arabidopsis thaliana* by a bacterial CG-specific DNA methyltransferase. Proc. Natl. Acad. Sci. U S A 118.23.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442-451.

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

Guan, Y., Manuel, R.C., Arvai, A.S., Parikh, S.S., Mol, C.D., Miller, J.H., Lloyd, S., and Tainer, J.A. (1998). MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. Nat. Struct. Biol. 5, 1058-1064.

He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 333, 1303-1307.

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. 33, 510-517.

Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y.G., Zhu, J., Wang, P., and Xu, Y. (2013). Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. Cell 155, 1545-1555.

Huh, J.H., Bauer, M.J., Hsieh, T.-F., and Fischer, R.L. (2008). Cellular Programming of Plant Gene Imprinting. Cell 132, 735-744.

Iwasaki, M., and Paszkowski, J. (2014). Epigenetic memory in plants. EMBO J. 33, 1987-1998.

Iyer, L.M., Abhiman, S., and Aravind, L. (2011). Natural history of eukaryotic DNA methylation systems. Prog. Mol. Biol. Transl. Sci. 101, 25-104.

Johnson, L.M., Du, J., Hale, C.J., Bischof, S., Feng, S., Chodavarapu, R.K., Zhong, X., Marson, G., Pellegrini, M., Segal, D.J., et al. (2014). SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA

methylation. Nature 507, 124-128.

Jullien, P.E., Kinoshita, T., Ohad, N., and Berger, F. (2006). Maintenance of DNA methylation during the *Arabidopsis* life cycle is essential for parental imprinting. Plant Cell 18, 1360-1372.

Kim, J.S. (2016). Genome editing comes of age. Nat. Protoc. 11, 1573-1578.

Kim, J.S., Lim, J.Y., Shin, H., Kim, B.G., Yoo, S.D., Kim, W.T., and Huh, J.H. (2019). ROS1-Dependent DNA Demethylation Is Required for ABA-Inducible *NIC3* Expression. Plant Physiol. 179, 1810-1821.

Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, S.E., Fischer, R.L., and Kakutani, T. (2004). One-way control of *FWA* imprinting in *Arabidopsis* endosperm by DNA methylation. Science 303, 521-523.

Klug, A. (2010). The discovery of zinc fingers and their applications in gene regulation and genome manipulation. Annu. Rev. Biochem. 79, 213-231.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517, 583-588.

Kungulovski, G., and Jeltsch, A. (2016). Epigenome Editing: State of the Art, Concepts, and Perspectives. Trends Genet. 32, 101-113.

Kuo, C.F., McRee, D.E., Fisher, C.L., O'Handley, S.F., Cunningham, R.P., and Tainer, J.A. (1992). Atomic structure of the DNA repair [4Fe-4S] enzyme endonuclease III. Science 258, 434-440.

Lauria, M., and Rossi, V. (2011). Epigenetic control of gene regulation in plants. Biochim. Biophys. Acta. 1809, 369-378.

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.

Lee, J.E., Neumann, M., Duro, D.I., and Schmid, M. (2019). CRISPR-based tools for targeted transcriptional and epigenetic regulation in plants. PLoS One 14, e0222778.

Lei, Y., Zhang, X., Su, J., Jeong, M., Gundry, M.C., Huang, Y.H., Zhou, Y., Li, W., and Goodell, M.A. (2017). Targeted DNA methylation *in vivo* using an engineered dCas9-MQ1 fusion protein. Nat. Commun. 8, 16026.

Li, E., and Zhang, Y. (2014). DNA methylation in mammals. Cold Spring Harb. Perspect. Biol. 6, a019133.

Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., and Jacobsen, S.E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science 292, 2077-2080.

Liu, W., Gallego-Bartolome, J., Zhou, Y., Zhong, Z., Wang, M., Wongpalee, S.P., Gardiner, J., Feng, S., Kuo, P.H., and Jacobsen, S.E. (2021). Ectopic targeting of CG DNA methylation in *Arabidopsis* with the bacterial SssI methyltransferase. Nat. Commun. 12, 3130.

Liu, X.S., Wu, H., Krzisch, M., Wu, X., Graef, J., Muffat, J., Hnisz, D., Li, C.H., Yuan, B., Xu, C., et al. (2018). Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the *FMR1* Gene. Cell 172, 979-992 e976.

Martinez-Macias, M.I., Qian, W., Miki, D., Pontes, O., Liu, Y., Tang, K., Liu, R., Morales-Ruiz, T., Ariza, R.R., Roldan-Arjona, T., and Zhu, J.K. (2012). A DNA 3' phosphatase functions in active DNA demethylation in *Arabidopsis*. Mol. Cell 45, 357-370.

Mok, Y.G., Uzawa, R., Lee, J., Weiner, G.M., Eichman, B.F., Fischer, R.L., and Huh, J.H. (2010). Domain structure of the DEMETER 5-methylcytosine DNA glycosylase. Proc. Natl. Acad. Sci. U S A 107, 19225-19230.

Morales-Ruiz, T., Ortega-Galisteo, A.P., Ponferrada-Marin, M.I., Martinez-Macias,

M.I., Ariza, R.R., and Roldan-Arjona, T. (2006). *DEMETER* and *REPRESSOR OF SILENCING 1* encode 5-methylcytosine DNA glycosylases. Proc. Natl. Acad. Sci. U S A 103, 6853-6858.

Morikawa, H., Ohkura, N., Vandenbon, A., Itoh, M., Nagao-Sato, S., Kawaji, H., Lassmann, T., Carninci, P., Hayashizaki, Y., Forrest, A.R., et al. (2014). Differential roles of epigenetic changes and Foxp3 expression in regulatory T cell-specific transcriptional regulation. Proc. Natl. Acad. Sci. U S A 111, 5289-5294.

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. Nature 448, 714-717.

Ortega-Galisteo, A.P., Morales-Ruiz, T., Ariza, R.R., and 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.

Papikian, A., Liu, W., Gallego-Bartolome, J., and Jacobsen, S.E. (2019). Site-specific manipulation of *Arabidopsis* loci using CRISPR-Cas9 SunTag systems. Nat. Commun. 10, 729.

Pastor, W.A., Aravind, L., and Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nat. Rev. Mol. Cell Biol. 14, 341-356.

Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., and Fischer, R.L. (2007). DNA demethylation in the *Arabidopsis* genome. Proc. Natl. Acad. Sci. U S A 104, 6752-6757.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and

Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173-1183.

Roberts, R.J., and Cheng, X. (1998). Base flipping. Annu. Rev. Biochem. 67, 181-198.

Roca Paixao, J.F., Gillet, F.X., Ribeiro, T.P., Bournaud, C., Lourenco-Tessutti, I.T., Noriega, D.D., Melo, B.P., de Almeida-Engler, J., and Grossi-de-Sa, M.F. (2019). Improved drought stress tolerance in *Arabidopsis* by CRISPR/dCas9 fusion with a Histone AcetylTransferase. Sci. Rep. 9, 8080.

Schumann, U., Lee, J., Kazan, K., Ayliffe, M., and Wang, M.B. (2017). DNA-Demethylase Regulated Genes Show Methylation-Independent Spatiotemporal Expression Patterns. Front. Plant Sci. 8, 1449.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat. Rev. Genet. 14, 204-220.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., and Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324, 930-935.

Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell 159, 635-646.

Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat. Methods 12, 1143-1149.

Weber, A.R., Krawczyk, C., Robertson, A.B., Kusnierczyk, A., Vagbo, C.B., Schuermann, D., Klungland, A., and Schar, P. (2016). Biochemical

reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. Nat. Commun. 7, 10806.

Williams, B.P., Bechen, L.L., Pohlmann, D.A., and Gehring, M. (2022). Somatic DNA demethylation generates tissue-specific methylation states and impacts flowering time. Plant Cell 34, 1189-1206.

Woo, H.R., Pontes, O., Pikaard, C.S., and Richards, E.J. (2007). VIM1, a methylcytosine-binding protein required for centromeric heterochroma tinization. Genes Dev. 21, 267-277.

Wu, H., and Zhang, Y. (2014). Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell 156, 45-68.

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

Yin, K., Gao, C., and Qiu, J.L. (2017). Progress and prospects in plant genome editing. Nat. Plants 3, 17107.

Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L., and Zilberman, D. (2013). The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell 153, 193-205.

Zeng, W., Huang, H., Lin, X., Zhu, C., Kosami, K.I., Huang, C., Zhang, H., Duan, C.G., Zhu, J.K., and Miki, D. (2021). Roles of DEMETER in regulating DNA methylation in vegetative tissues and pathogen resistance. J. Integr. Plant Biol. 63, 691-706.

Zhang, H., Lang, Z., and Zhu, J.K. (2018). Dynamics and function of DNA methylation in plants. Nat. Rev. Mol. Cell Biol. 19, 489-506.

Zhang, H., Zhang, X., Clark, E., Mulcahey, M., Huang, S., and Shi, Y.G. (2010). TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. Cell Res. 20, 1390-1393.

CHAPTER I

Ectopic expression of DEMETER DNA glycosylase induces changes in the expression of stress-response genes

ABSTRACT

DNA methylation is crucial for diverse biological processes including gene regulation, imprinting, and X chromosome inactivation. The mode of action in establishing DNA methylation in both plants and mammals is conserved, while DNA demethylation machineries have evolved distinctively across the two kingdoms. DEMETER (DME) DNA glycosylase family proteins specifically excise 5mC and establish imprinting in the endosperm. DME-mediated direct 5mC excision activated endogenous repeat elements, which appeared to form dsRNAs as viral mimics and induced antiviral responses by triggering interferon cascades in mammalian cells. In this study, DME was introduced into the *Arabidopsis* protoplasts to verify the effect of DME in the plant cell system. The erased 5mC by DME was detected globally on the protoplast genome. Also, the impact of ectopically expressing DME induced transcriptional changes in several stress and defense response-related genes. Moreover, the expression changes in TEs were analyzed. Taken together, this study might provide the primary clue of the effect of DME in the plant cell, as well as the overall role of DNA methylation in the genome.

INTRODUCTION

Epigenetic changes are covalent modifications of DNA or histones, influencing chromatin structure and accompanying expression of the underlying genes (Law and Jacobsen, 2010). DNA methylation, a prominent epigenetic modification, referring to the methylation of the 5th position of cytosine, is highly conserved among eukaryotes, including plants, mammals, and fungal models (Law and Jacobsen, 2010). DNA methylation contributes to diverse biological processes and its perturbation can entail developmental abnormalities (Zhang et al., 2018). In mammals, DNA methylation is highly deposited in the CG dinucleotides, while plant genome displays three distinct sequence contexts: symmetric CG and CHG and asymmetric CHH (where H is any nucleotide other than G) (Law and Jacobsen, 2010). The mechanisms in establishing DNA methylation patterns in both mammals and plants are very alike (Li and Zhang, 2014). DNA methyltransferase (DNMT) family governs the addition of methylation marks to genomic DNA in vertebrates (Lyko, 2018). Unmethylated cytosine becomes methylated *de novo* by DNMT3, and DNMT1, the maintenance DNA methyltransferase, restores a symmetrical methylation state after semi-conservative DNA replication (Li and Zhang, 2014). DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), a homolog of the DNMT3 in mammals, governs *de novo* methylation through RNA-directed DNA methylation (RdDM) pathway in plants (Law and Jacobsen, 2010). Moreover, despite their homogeneous roles of adding 5-methylcytosine (5mC), DRM2 has undergone sequence divergence compared to DNMT3, in which DRM2 literally includes rearranged catalytic motifs compared to that of mammals (Cao and

Jacobsen, 2002). However, various enzymes are in charge of maintaining DNA methylation status in plants (Law and Jacobsen, 2010). DNA METHYLTRANFERASE 1 (MET1), CHROMOMETHYLASE 2 (CMT2), and CMT3 are maintaining CG, CHH, and CHG methylation, respectively, while *de novo* methyltransferase DRM2 also catalyzes spreading of 5mC in CHH contexts (Lindroth et al., 2001; Zemach et al., 2013).

Although DNA methylation has been conceived as a stable epigenetic trait, it can be reversibly written or erased (Wu and Zhang, 2010). DNA demethylation can either occur passively or actively, contingent upon the enzymatic activity (Law and Jacobsen, 2010). Passive DNA demethylation takes place in which the maintenance DNA methylation pathway is absent or inhibited during DNA replication in both mammals and plants. However, the mode of action of active DNA demethylation is distinct between two kingdoms. In animals, methylated cytosine is actively reversed into cytosine through TEN-ELEVEN TRANSLOCATION (TET) dioxygenase family-mediated successive oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) (Wu and Zhang, 2017). The 5mC oxidation products, 5fC and 5caC, are cleaved by the thymine-DNA glycosylase (TDG), generating abasic (apurinic or apyrimidic) sites with a single-strand break, followed by downstream repairing processes via base excision repair (BER) pathway (He et al., 2011; Kohli and Zhang, 2013). On the other hand, plant-specific DNA glycosylase is responsible for active DNA demethylation in plants, which can recognize and remove 5mC directly (Zhang et al., 2018). 5mC excision through DNA glycosylase activity accompanies β- and δ- elimination reactions, creating a 3'-phosphor-α, β-unsaturated aldehyde (3'-PUA),

or 3'-phosphate, respectively and processed to 3' OH. Eventually, further BER pathway fills the 3' OH group and terminates the procedure (Lee et al., 2014; Martinez-Macias et al., 2012).

In *Arabidopsis*, there are four members of bifunctional DNA glycosylases: DEMETER (DME), REPRESSOR OF SILENCING 1 (ROS1), DEMETER-LIKE 2 (DML2), and DML3 (Gong et al., 2002; Morales-Ruiz et al., 2006; Ortega-Galisteo et al., 2008). DME is primarily expressed in the central cell and regulates the imprinting of Polycomb Repressor Complex 2 (PRC2), including *MEDEA* (*MEA*) (Choi et al., 2002; Huh et al., 2008). The maternal allele of *MEA* is hypomethylated in the endosperm, and its expression silences paternal *MEA* allele (Gehring et al., 2006). Therefore, DME establishes the imprinting of *MEA* in the endosperm and is prerequisite for proper seed development (Choi et al., 2002). On the contrary, ROS1, DML2 and 3 are expressed in the vegetative tissues (Zhu, 2009). ROS1, as well as DML2 and 3 are crucial for maintaining the dynamics of transposon expression and in shaping genome structure between transposons and genomic regions (Gong et al., 2002; Penterman et al., 2007). Particularly, ROS1 regulates some stress-response genes, the expression of which is controlled by DNA methylation state (Zhu, 2009).

A previous study about introducing plant-specific DNA demethylase DME into human HEK-293T cells revealed that DME expression inhibits cell proliferation rate as well as triggers interferon cascades by TE-derived dsRNAs as viral mimics (Mok et al., 2017). In this study, the effect of ectopically expressing DME in the *Arabidopsis* protoplasts was examined. To confirm the 5mC excising effect of DME on the genome, the removed 5mC was labeled with Cy3 fluorophore. The distribution of expressing DME activity was localized throughout the nucleus, while

the Mock and the catalytic mutant KQ constructs revealed negative. Also, further analysis was conducted to understand the impact of DME in the aspects of gene expression, the transcriptome analysis was performed on the DME transfected, Mock treatment, GFP empty vector, and the catalytic mutant of DME, KQ with two distinct incubated time after transfection. The gene expression level didn't show much differences in the first 12 hours, but showed 183 and 357 differentially expressed genes (DEGs), including several stress-, and defense-response genes, in DME constructs compared to Mock and KQ constructs, respectively. Particularly, *ABA-INDUCED WD40-REPEAT 1* (*AIW1*), and *AIW2* were upregulated in DME constructs, relative to both Mock and KQ transfected cells. Considering ROS1-mediated DNA demethylation is crucial for ABA response, and its downstream gene expression, the transcriptional alteration of *AIW1* and *2* might be also related to the effect of constitutively expressing DME (Kim et al., 2019). The expression of transposable elements (TEs) in DME constructs relative to Mock and KQ in the first 12 h increased, and the majority of whom revealed above 1.3-fold upregulation were overlapped. These findings imply the global effect of DME-mediated DNA demethylation on the expression of both TEs and genes in the protoplasts and insinuate the overall roles of DNA demethylation in plants.

MATERIALS AND METHODS

Plant materials and growth conditions

Arabidopsis thaliana ecotype Columbia-0 (Col-0) was used in this study. The mature seeds were sterilized with 70% ethanol and stratified at 4℃ for 2 day, and sowed on the soil. The germinated seedlings were grown in growth chamber at 23℃ under 16 hours of light. To isolate the mesophyll protoplast, each plant was used when the number of true leaves was 8 or less.

Protoplast isolation and PEG-mediated transfection

Protoplast was isolated from 4-week-old *Arabidopsis* plant. The fifth, sixth, and the seventh true leaves were chopped and submerged into the enzyme solution (20 mM MES (pH 5.7), containing 1.5% (wt/vol) cellulase R10 (Duchefa, Haarlem, Netherlands), 0.4% (wt/vol) macerozyme R10 (Duchefa, Haarlem, Netherlands), 0.4 M mannitol, and 20 mM KCl) to be digested at room temperature on a shaker at 20 RPM for 3 hours in darkness. Released protoplast from enzyme solution were filtered by 75-μm nylon mesh and collected by centrifugation at 100*g* in a round-bottom 40 mL tube for 2 min. After the centrifugation, the supernatant was removed and the protoplast pellet was washed and resuspended by 5 mL of W5 solution (2 mM MES (pH 5.7) containing 154 mM NaCl, 125 mM $CaCl_2$ and 5 mM KCl) and kept on ice for 30 min under the light. When the protoplasts were settled at the bottom of the tube by gravity, W5 solution was discarded as much as possible. Washed protoplasts were diluted to a density of 3.0 ~ 3.5 x 10^5 protoplasts / mL of MMG solution (4 mM

MES (pH 5.7) containing 0.4 M mannitol and 15 mM $MgCl_2$).

The plasmid DNA was delivered into the protoplasts by PEG-mediated transfection. A mixture of 6.0 ~ 7.0 x 10^4 cells resuspended in 200 µL of MMG solution was gently mixed with 20 µg of plasmid DNA in a 20 µl distilled water in a 2 mL microfuge tube. 220 µL of PEG solution (40% (wt/vol) PEG4000 in distilled water, containing 0.2 M mannitol and 100 mM $CaCl_2$) were added to the tube, and tapped gently to mix completely. Then the tapped tube was incubated for 10 min at room temperature. After incubation, the transfection mixture was by adding 880 µL of W5 solution and inverted gently to stop the transfection process. Transfected protoplasts were pelleted by centrifugation at 200*g* for 2 min at room temperature and the supernatant was removed. Resuspended protoplasts with 1 mL of W5 solution were cultured in each 10% Bovine Serum Albumin (BSA)-coated well of 6-well tissue culture plate and incubated at room temperature under light for either 12 h or 24h. After the incubation of desired period of time, protoplasts were harvested by resuspension and followed centrifugation at 200*g* for 2 min. Supernatant was removed and stored in -80 °C until further use.

In situ 5mC labeling assay

Post-cultured protoplasts transfected with plasmids were pelleted by centrifugation at 200*g* for 2 min at room temperature and the supernatant was removed. The pellets were plated on slides and fixed with 10% formalin and dried at room temperature. After drying, the fixed protoplasts were washed with PBS buffer (10 mM Na_2HPO_4, 143 mM NaCl, pH 7.0) 3 times for 5 min each and digested with 10 µg/ml proteinase K (Invitrogen) for 10 min at room temperature, and rinsed again

with PBS buffer. The cells were labeled with 0.3 nmol Cy3-dCTP (Cytiva) in 100 μL of reaction buffer (NEBuffer 2.0, 15 U Klenow fragment (3'→5' exo-) (NEB), 15 U Endonuclease IV (NEB)) for 90 min at 37 °C. After the incorporation, the slides were washed with PBS buffer 3 times for 5 min each and mounted with 4',6-Diamidino-2-Phenylindole, dihydrochloride (DAPI) to stain the chromosomes and covered with glass coverslip.

RNA sequencing analysis

RNA was extracted from harvested *Arabidopsis* protoplasts with TRIzol reagent (Ambion). For transcriptome analysis, mRNA library construction was achieved in strand-specific manner with TruSeq adapter (Illumina, CA, USA). Sequencing was performed using NovaSeq 6000 platform (Illumina, CA, USA). Raw reads were filtered with Trimmomatic v0.38 with "LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:80" parameters and the filtered reads were mapped on TAIR10 genome TopHat v2.1.1 with default parameters (Bolger et al., 2014; Trapnell et al., 2009). The mapped read counts were calculated using HTSeq 0.6.p1 and EdgeR v3.38.0 was used for the statistical test of differentially expressed genes (DEGs) (Anders et al., 2015; Robinson et al., 2010). Finally, DEGs were defined as genes with a false discovery rate (FDR) < 0.10.

RESULTS

Ectopic expression of DME in plant protoplasts

DME family protein is responsible for the removal of 5mC in plants to catalyze active DNA demethylation (Wu and Zhang, 2010). DME-mediated DNA demethylation is indispensable for proper endosperm development by establishing maternal-specific expression of *MEA* (Choi et al., 2002). Therefore, the loss-of function mutant shows seed abortion phenotype, while the overexpression of DME induces accelerated bud formation in transgenic hybrid poplar (Choi et al., 2002; Conde et al., 2017). Furthermore, it was previously shown that DME-mediated direct 5mC excision induces antiviral responses in HEK-293T cells by activating repeat elements, which results in dsRNA formation followed by triggering interferon cascades (Mok et al., 2017). To verify the effect of ectopic expression of DNA demethylation in the plant cell system, I transfected the construct constitutively expressing DME into *Arabidopsis* protoplasts. In order to obtain the reliable data, the protoplasts with a high level of transfection efficiency, greater than 70% of cells with GFP signals inside the nucleus, were used (Figures 1-1 and 1-2). The GFP signal in the nucleus appeared about 8 hours upon the transfection and the number of cells glowing the signal saturated after 12 hours in the preliminary experiments. However, the longer the incubation time, the more cell aggregation and cell death arose in all transfected cell, but the mock-treated. Therefore, I determined the 12 h and 24 h incubation is suitable to examine the time-course effect of DME in the protoplasts (Figure 1-2). The truncated DME, DMEΔN677ΔIDR1::lnk, comprising the three conserved C-terminal domains, while the N-terminal 677 amino acids and the

interdomain region flanked between A and glycosylase domain are removed were utilized, hereafter named DME (Mok et al., 2010). On the contrary, DMEΔN677ΔIDR1::lnk(K1286Q) (designated as KQ in Figure 1-1), catalytic mutant of DME with reduced DNA glycosylase activity while preserving the structure and stability of the enzyme, was introduced as mutant DME proteins to validate the effect of DME protein itself. Moreover, the mock control, transfected with distilled water, was examined to ascertain the impact of foreign substance delivery.

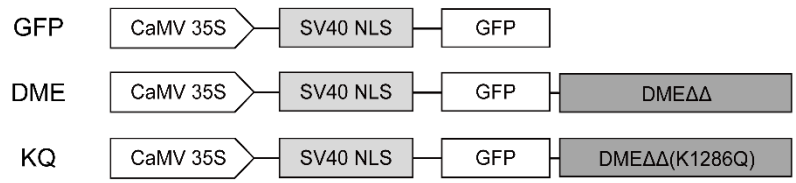

Figure 1-1. Schematic diagrams of the constructs used for the protoplast transfection.

Each construction fused with N-terminal SV40 NLS was driven under CaMV 35S promoter for constitutive expression. GFP was flanked between NLS and the enzyme to determine the transfection efficiency when the constructs were delivered to the protoplasts. SV40-type NLS was added to ensure proper nuclear import in the protoplasts.

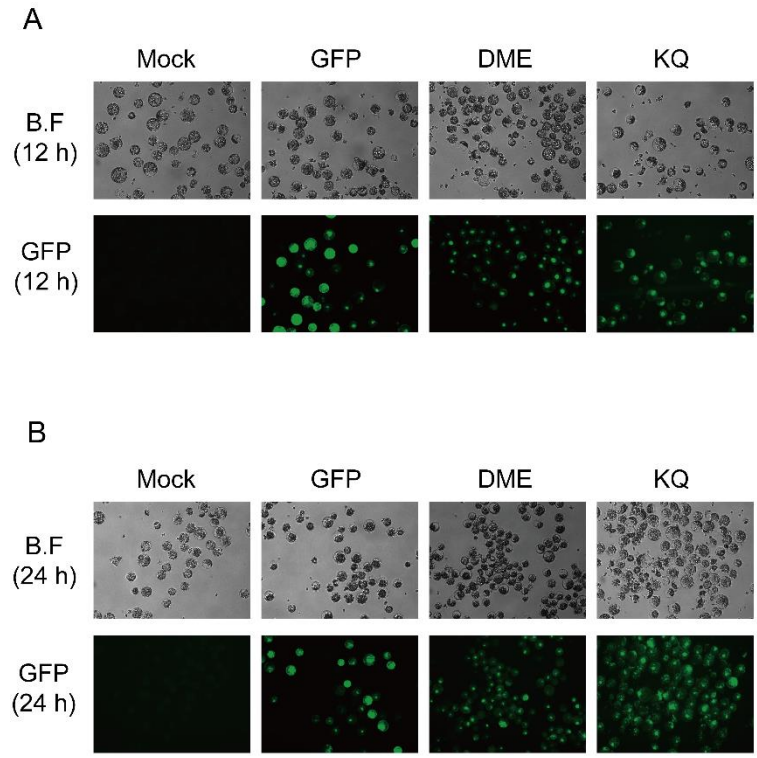

Figure 1-2. Subcellular localizations of GFP in *Arabidopsis* protoplasts.

GFP fluorescence signals (green) indicate the protoplasts expressing GFP-fused constructs under constitutive 35S promoter. (A, B) Images were acquired 12 h (A) and 24 h (B) after transfection, respectively. Bright field indicates the images under bright field microscopy (x200). Mock treatment was transfected with 20 μL of distilled water, while other constructs were transfected with 20 μg of plasmid DNA in 20 μL.

The effect of DME on *Arabidopsis* protoplast genome

Being a bifunctional DNA glycosylase, DME generates nick on the 3' side of the 5mC residue by breaking the phosphodiester bond (Gehring et al., 2006). The DME-induced nick sites were visualized to assess the effect of DME on the protoplast genome. In this study, the replacement of 5mC was subjected to Cy3-dCTP, which had been successfully adopted in the previous research (Lee et al., 2014). However, the abasic site produced by the catalytic activity of DME remains to be processed for succeeding polymerization (Martinez-Macias et al., 2012). Accordingly, Endonuclease IV trimmed the 5mC excision intermediates, the place for which the AP endonuclease is responsible, to provide 3' OH. Cy3-dCTP was sequentially incoporated by Klenow DNA polymerase sealing the gap. As a result, not only the Mock and GFP transfected cells but also the catalytic inactive KQ constructs didn't reveal the Cy3 fluorescence, while DME transfected cells reacted Cy3 positive with the overlapped localization in the nucleus along with the counterstained with DAPI (Figure 1-3). The overlap of DAPI and Cy3 signals in DME-transfected protoplast cells is consistent with the previous study (Mok et al., 2017). Taken together, ectopic-expressed DME protein enabled to excise 5mC and induce single-strand breaks in both plant and mammal cells. It implies that overexpression of DME might change global gene expression in protoplast cells.

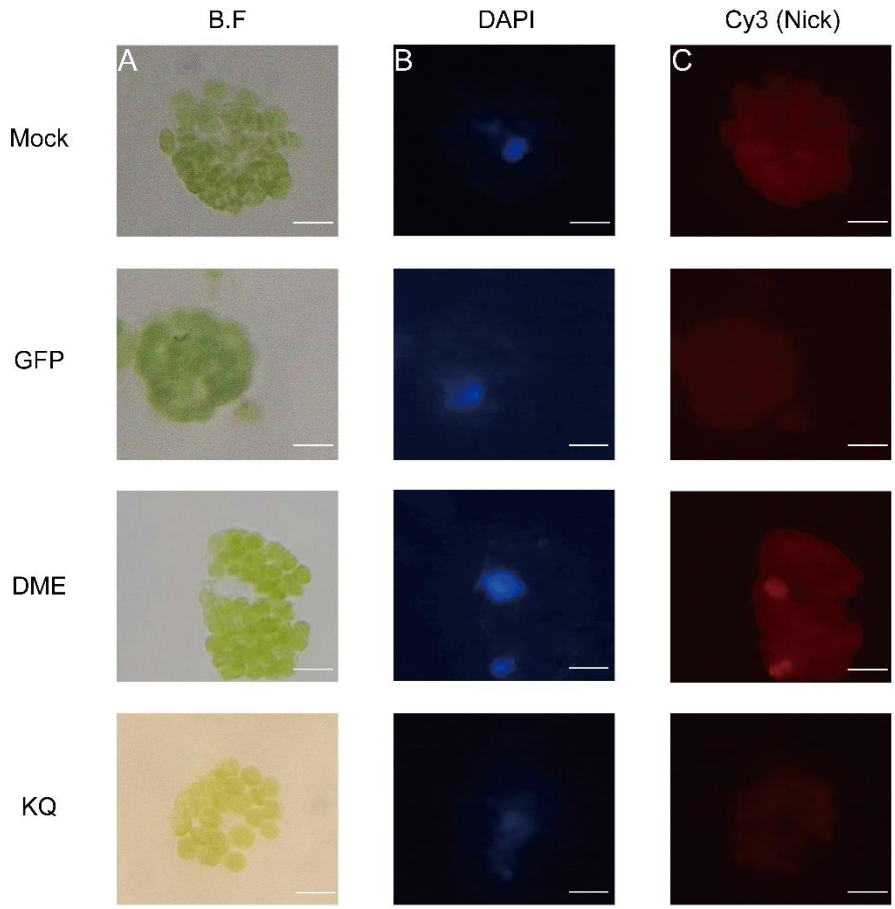

Figure 1-3. Distribution of excised 5mC regions in *Arabidopsis* protoplasts

Arabidopsis protoplasts were mounted on a glass slide 18 h after transfection and labeled with Cy3-dCTP to visualize sublocalization of the excised 5mC regions. (A) Bright field indicates the images under bright field microscopy (x400). (B) DAPI-stained nucleus (indigo). (C) The DME-induced nick sites on DNA are incorporated with Cy3 fluorophore (red). Scale bars = 10 μm.

Gene expression profiles of constitutively DME expressing *Arabidopsis* protoplasts

DNA methylation is related to genome stability in eukaryotes, and the removal of repressive DNA methylation in plants is associated with the activation of both genes and transposons (Ibarra et al., 2012; Zhu, 2009). To understand the effect of DME, the transcriptome analysis was conducted on two replicates for both each construct and incubation time. To satisfy the sufficient amount of total extracted RNA from the transfected protoplasts, every construct was pooled up to the number of 9.0 ~ 10.5 x 10^5 protoplasts per one biological replicate. FPKM values of both annotated genes and TEs were calculated from the RNA-seq data to estimate expression levels. Differentially expressed genes (DEGs) analysis was performed to compare the gene expression between DME and other constructs, Mock, and KQ, respectively. Only 13 genes out of 27,445 protein-coding genes were up-regulated in DME-transfected constructs relative to Mock constructs. DEGs were not observed compared to KQ-transfected cells. On the contrary, as the incubation time rises to 24 h, the number of DEGs in DME constructs increased to 183 and 357 concerning Mock and KQ, respectively (Figure 1-5). To be specific, a total of 101 and 82, 188 and 169 genes were commonly up- and down-regulated, relative to Mock and KQ transfected cells, respectively. Moreover, the number of 46 genes were co-upregulated compared to Mock and KQ, while 55 genes revealed as the commonly down-regulated genes. This data showed that the constitutively expressing DME in the 12 h incubated protoplast cells had a minor effect on the global gene expression changes.

Meanwhile, the enhanced numbers of DEGs corresponding to the incubation

time rising might be due to the accumulated effect of constitutively expressing DME, who shows 756.68 and 42- fold change of expression level in 12 h and 24 h incubation time, compared to GFP constructs (Figure 1-4). Although, even with the high level of DME expression, the first twelve hour might not be enough to elicit transcriptional changes in the cells and resulted in low number of DEGs in 12 h incubated cells.

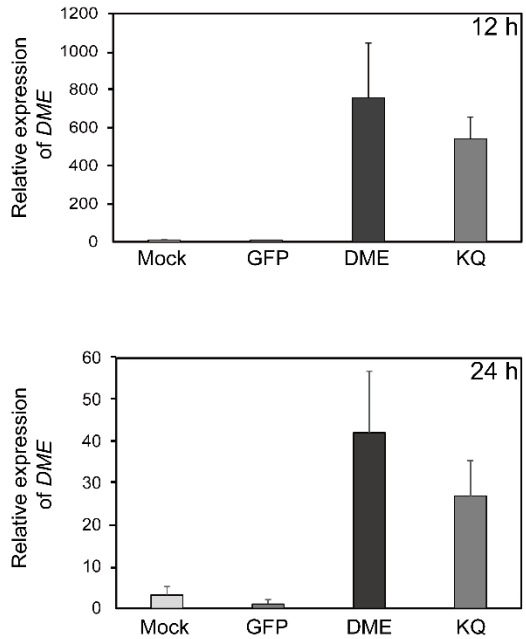

Figure 1-4. Relative expression level of *DME* in the transfected protoplasts.

The expression level of DME was analyzed in every transfected protoplasts with respect to distinct incubation time (A) 12 h and (B) 24 h. Each expression values were normalized with GFP construct as control. Each error bars represent mean ±S.D. of two biological replicates.

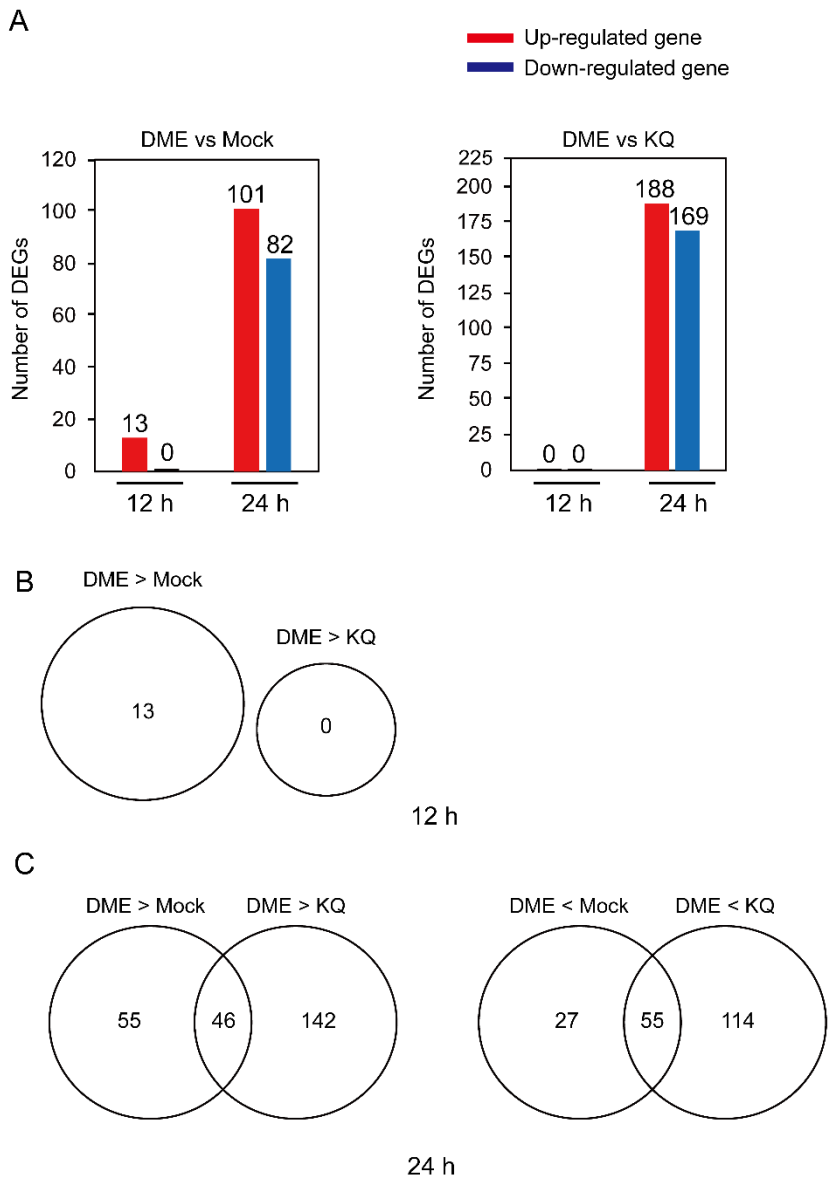

Figure 1-5. Identification of differentially expressed genes (DEGs) among Mock, DME, and KQ transfected protoplasts.

(A) The number of DEGs in DME construct, in comparison to Mock and KQ transfected cells. Each red and blue box indicate the up- and down-regulated genes. (B), (C) The Venn diagram of co-differentially expressed genes of DME construct, compared to Mock and KQ, with respect to incubation time, (B) 12 h, (C) 24 h.

Ectopic expression of DEMETER DNA glycosylase induces changes in the expression of stress-response genes.

DME DNA glycosylase cleaves the phosphodiester bond on the 3' side of the 5mC residue, and preferentially targets the hemimethylated sites. Such DME activity induces single-strand break and is inhibited not to further excise the complementary strand (Gehring et al., 2006). A previous study revealed that DME-dependent 5mC excision triggers the upregulation of interferon-stimulated genes (ISGs) and heat shock proteins (Mok et al., 2017). Gene ontology (GO) analysis was performed to categorize biological functions of both co-differentially expressed genes in DME constructs relative to Mock KQ. Notable among the co-upregulated genes in the DME-transfected cells, a number of stress-related genes were identified such as response to abiotic stimulus, response to abscisic acid, and chemical, etc (Figure 1-6A). Especially, *ABA-INDUCED WD40-REPEAT 1* (*AIW1*), *2* (*AIW2*) were among the up-regulated genes, the expression of which is induced by ABA and showing hypersensitivity to ABA in *aiw1 aiw2* double mutants (Wang et al., 2020). It was previously shown that the down-regulation of *NICOTINAMIDASE 3* (*NIC3*) in *ros1* mutant, an ABA-inducible gene in *Arabidopsis*, was associated with hypermethylation of the promoter region. Therefore, the increased transcription level of both *AIW1*, and *2* might be related to the ectopically expressed DME. Besides, some defense-related genes and response to external biotic factors were identified as commonly down-regulated genes (Figure 1-6B). These findings suggest that the transcriptional changes in the DME-transfected cells may be associated with excessive DNA demethylation induced by DME.

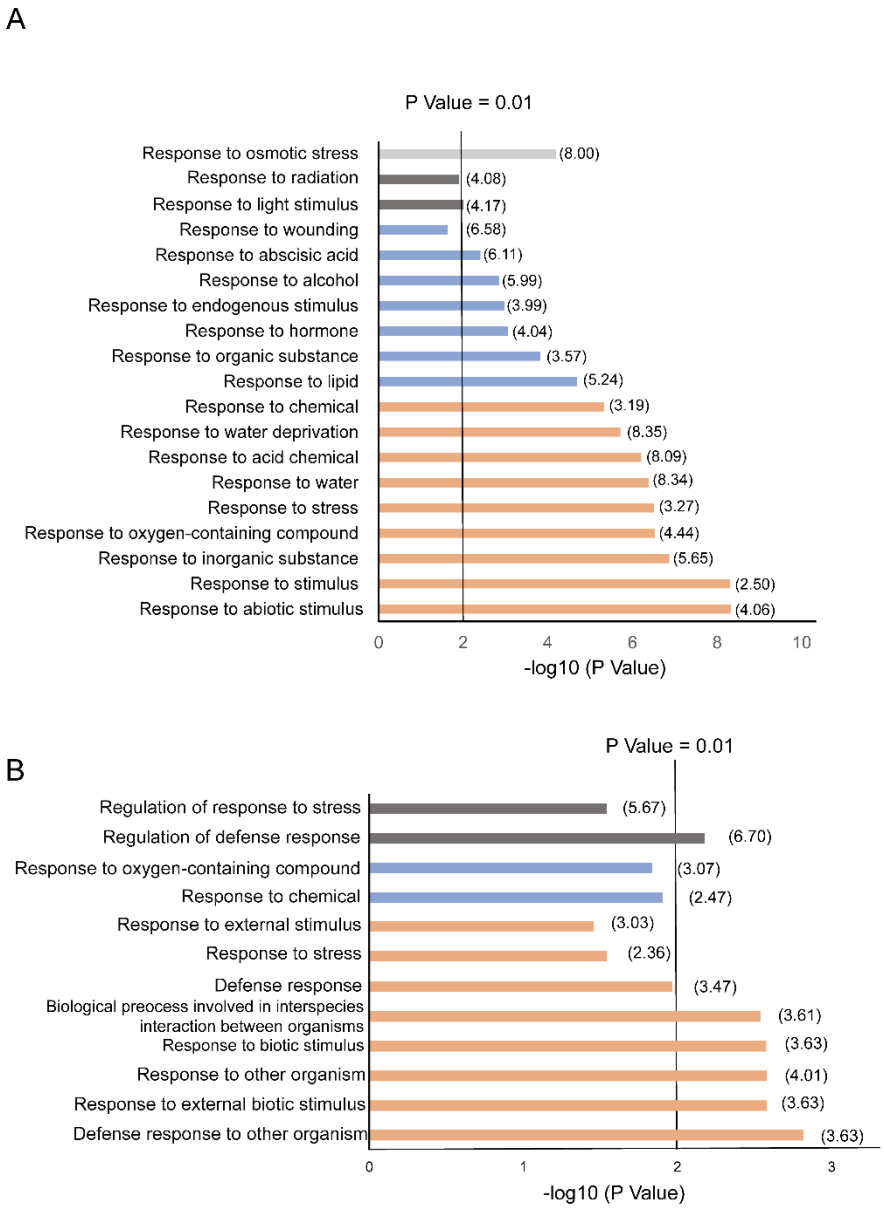

Figure 1-6. The gene ontology analysis in DME transfected protoplasts.

(A, B) GO categories of each differentially expressed genes in DME transfected constructs, relative to both up-regulated (A) and down-regulated (B) in Mock and KQ constructs. Each gene ontology terms filled with identical colors are categorized into an equal gene cluster. Each number indicated in the bracket refers to the fold enrichment.

Transposons are reactivated by the effect of DME in the first 12 h.

Transposable elements (TEs) are residing at the boundaries between euchromatin and heterochromatin and kept silenced to maintain genome stability (Law and Jacobsen, 2010; Zhang et al., 2018). Since TEs are shut down by DNA methylation, their reactivation can be affected by removing it (Ibarra et al., 2012). Therefore, to verify the ectopically expressing DME had an impact on the residing TE families, the expression of which was analyzed in DME construct relative to the Mock and KQ. In 12 h incubated cells, the average expression level change of TEs between DME and KQ, or DME and Mock, was 1.090 and 1.080-folded, respectively, while those revealed as 1.009 and 1.010-folded change in 24 h incubated cells (Figure 1-7A). These imply that ectopic expression of DME slightly activated the expression of TEs at the first 12 h. In particular, a total of 83 and 84 TEs in DME-transfected cells, compared to Mock and KQ, respectively when incubated for 12 h, were activated more than 1.090-fold change. Among those, 76 TEs overlapped, revealed as 91.5% and 90.5%, respectively (Figure 1-7B). However, in terms of 24 h incubated samples, the ratio of overlapped TEs in both DME versus Mock and KQ were reduced to 40.1% and 71.4% respectively, that is 25 co-activated TEs out of 61 and 35. This study implies that TEs which showed an expression change in the primary 12 h were notably up-regulated in the DME construct, compared to Mock and KQ. Furthermore, given the dominant number of overlapped up-regulated TEs in the early 12 h, those activations of TEs might be elicited by the effect of DME-directed DNA demethylation.

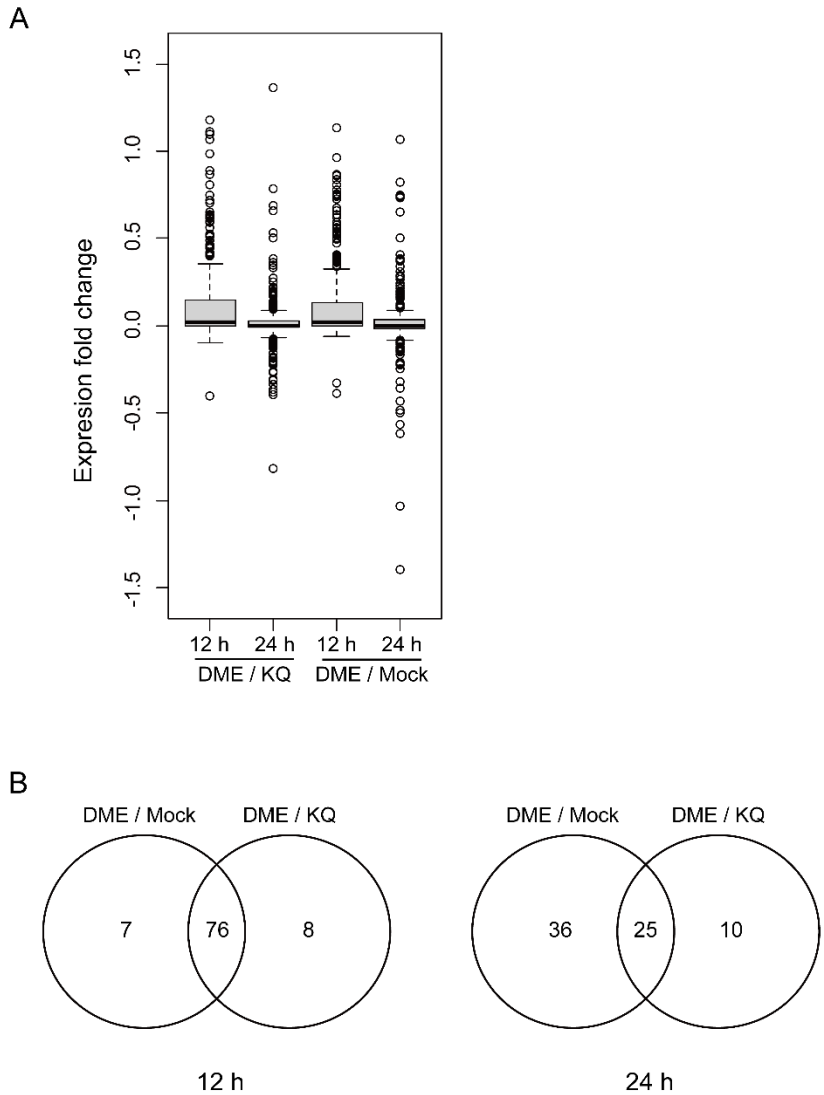

Figure 1-7. Distribution of TEs with altered expression level.

(A) The box plot which shows the distribution of TEs with respect to expression fold change in DME construct relative to other constructs, Mock and KQ along with the time course of incubation time. (B) The two-way Venn diagram of co-activated gene with more than 1.090-fold change in the DME construct against Mock and KQ in both 12 h and 24 h incubation.

DISCUSSION

Active DNA demethylation in plants is catalyzed by DME family DNA glycosylases. Although they share highly conserved domains, their biological roles are distinct. DME is responsible for proper endosperm development during female gametogenesis by directly excising 5mC in the central cell (Choi et al., 2002; Gehring et al., 2006). A previous study revealed that heterogenous *Arabidopsis DME* induced the anti-viral responses in mammalian cells (Mok et al., 2017). In this study, the constitutively expressing DME was delivered into *Arabidopsis* protoplasts to elucidate the time-course effect of ectopic expression of DNA demethylation. The subcellular localization of both GFP and Cy3 represents the 5mC excising activity of DME is concentrated in the nucleus and induces nicks on DNA genome-widely (Figures 1-2 and 1-3). The transcriptome data revealed the increment of differentially expressed genes as the incubation increases (Figure 1-5). It is assumed that the 12 h of incubation wasn't sufficient for cells to react upon the excessive demethylation of DME, there further remains to be addressed by the extended time of incubation. Although, some ABA-inducible genes were up-regulated in the DME construct relative to Mock and KQ-delivered cells at the 24 h incubated condition. In comparison to the aspects of gene expression profiles, the expression analysis on TEs revealed that the effect of DME reactivated TEs, extensively in the first 12 h rather than the last 12 h. These data suggest that the immediate effect of DME transfection was concentrated to TEs rather than genes.

Consequently, it can be hypothesized that the ectopic expression of DME erased the 5mC adjacent to TEs or repeat sequences nearby genes, and elicited their

reactivation upon the transfection. The reactivated TEs, impose deleterious effects on genome stability, and the TEs with repeat sequence might be considered as viral mimics. Therefore, the counteracts to these atypical conditions have resulted in the gene expression changes in stress-response or ABA-inducible genes in the late 12 h. Lastly, the changes in the DNA methylation level need to be investigated to verify whether such expression changes of genes or TEs were directly derived from the DME-mediated DNA demethylation, or the downstream effect of DNA demethylation such as DNA damage is responsible for.

REFERENCES

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166-169.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120.

Cao, X., and Jacobsen, S.E. (2002). Role of the *Arabidopsis* DRM methyltransferases in de novo DNA methylation and gene silencing. Curr. Biol. 12, 1138-1144.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., and Fischer, R.L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. Cell 110, 33-42.

Conde, D., Moreno-Cortes, A., Dervinis, C., Ramos-Sanchez, J.M., Kirst, M., Perales, M., Gonzalez-Melendi, P., and Allona, I. (2017). Overexpression of DEMETER, a DNA demethylase, promotes early apical bud maturation in poplar. Plant Cell Environ. 40, 2806-2819.

Gehring, M., Huh, J.H., Hsieh, T.F., Penterman, 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.

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

He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its

excision by TDG in mammalian DNA. Science 333, 1303-1307.

Huh, J.H., Bauer, M.J., Hsieh, T.-F., and Fischer, R.L. (2008). Cellular Programming of Plant Gene Imprinting. Cell 132, 735-744.

Ibarra, C.A., Feng, X., Schoft, V.K., Hsieh, T.F., Uzawa, R., Rodrigues, J.A., Zemach, A., Chumak, N., Machlicova, A., Nishimura, T., et al. (2012). Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. Science 337, 1360-1364.

Kim, J.S., Lim, J.Y., Shin, H., Kim, B.G., Yoo, S.D., Kim, W.T., and Huh, J.H. (2019). ROS1-Dependent DNA Demethylation Is Required for ABA-Inducible *NIC3* Expression. Plant Physiol. 179, 1810-1821.

Kohli, R.M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. Nature 502, 472-479.

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.

Lee, J., Jang, H., Shin, H., Choi, W.L., Mok, Y.G., and Huh, J.H. (2014). AP endonucleases process 5-methylcytosine excision intermediates during active DNA demethylation in *Arabidopsis*. Nucleic Acids Res. 42, 11408-11418.

Li, E., and Zhang, Y. (2014). DNA methylation in mammals. Cold Spring Harb. Perspect. Biol. 6, a019133.

Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., and Jacobsen, S.E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science 292, 2077-2080.

Lyko, F. (2018). The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. Nat. Rev. Genet. 19, 81-92.

Martinez-Macias, M.I., Qian, W., Miki, D., Pontes, O., Liu, Y., Tang, K., Liu, R.,

Morales-Ruiz, T., Ariza, R.R., Roldan-Arjona, T., and Zhu, J.K. (2012). A DNA 3' phosphatase functions in active DNA demethylation in *Arabidopsis*. Mol. Cell 45, 357-370.

Mok, Y.G., Choi, K.Y., Hong, S.H., and Huh, J.H. (2017). DEMETER plant DNA demethylase induces antiviral response by interferon signalling in animal cells. Sci. Rep. 7, 9160.

Mok, Y.G., Uzawa, R., Lee, J., Weiner, G.M., Eichman, B.F., Fischer, R.L., and Huh, J.H. (2010). Domain structure of the DEMETER 5-methylcytosine DNA glycosylase. Proc. Natl. Acad. Sci. U S A 107, 19225-19230.

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

Ortega-Galisteo, A.P., Morales-Ruiz, T., Ariza, R.R., and 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.

Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., and Fischer, R.L. (2007). DNA demethylation in the *Arabidopsis* genome. Proc. Natl. Acad. Sci. U S A 104, 6752-6757.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105-1111.

Wang, X.T., Wang, W., Wang, Y.T., Zhou, G.H., Liu, S.D., Li, D.Q., Adnan, Hussain,

S., Ahmed, S., Zhang, C., et al. (2020). AIW1 and AIW2, two ABA-induced WD40 repeat-containing transcription repressors function redundantly to regulate ABA and salt responses in *Arabidopsis*. J. Plant Interact. 15, 196-206.

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

Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. Nat. Rev. Genet. 18, 517-534.

Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L., and Zilberman, D. (2013). The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell 153, 193-205.

Zhang, H., Lang, Z., and Zhu, J.K. (2018). Dynamics and function of DNA methylation in plants. Nat. Rev. Mol. Cell Biol. 19, 489-506.

Zhu, J.K. (2009). Active DNA demethylation mediated by DNA glycosylases. Annu. Rev. Genet. 43, 143-166.

CHAPTER II

Application of 5-methylcytosine DNA glycosylase to DNA-free epigenome editing

ABSTRACT

In higher eukaryotes, DNA methylation is one of the prominent epigenetic marks, associated with gene expression. Changes in DNA methylation can be stably inherited and tightly regulated by certain enzymatic activities. Recently, Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR Associated (CRISPR)-based targeted manipulation of epigenetic modifications has been devised as an attractive strategy to regulate gene expression. Here I designed the application of plant-specific DNA glycosylase DEMETER (DME) for targeted DNA demethylation with CRISPR / a catalytically deactivated Cas9 (dCas9)-SunTag system. DME that directly excises 5-methylcytosine from DNA through cleavage of phosphodiester bond was utilized as an effector for the targeted DNA demethylation. Besides, SunTag system in which GCN4 peptides efficiently recruit scFv-fusion proteins was adapted to enhance the efficiency of targeted DNA demethylation. The two components, dCas9-GCN4 and scFv-DMEΔ, were purified and preassembled with single-guide RNA to form ribonucleoprotein (RNP) complex. SunTag-DME system was capable of excising 5mC *in vitro* and showed target preference in a single-guide RNA-dependent manner. These data suggest that SunTag-DME system

might successfully reduce DNA methylation levels at target sites *in vitro* and *in vivo*. This study provides the implication of a DNA-free epigenome editing tool for targeted DNA demethylation which can be applied to the plant system.

INTRODUCTION

Epigenetic modifications encompass a variety of factors that regulate genomic functions apart from DNA sequences (Heard and Martienssen, 2014). Among them, DNA methylation is a stable but reversible epigenetic mark that is pivotal for gene regulation, genome stability, and many other biological processes (Jaenisch and Bird, 2003; Smith and Meissner, 2013). The specific effectors modulate the DNA methylation profile, via establishing, propagating, or removing the 5-methylcytosine (5mC) to coordinate gene expression mitotically or meiotically in a heritable manner (Erdel, 2017; Law and Jacobsen, 2010; Zhu, 2009).

DNA methylation is established in symmetric CG contexts by DNA METHYLTRANSFERASE 3 (DNMT3) and maintained by DNMT1 in mammals. In plants, however, DNA methylation takes place at cytosine bases in several sequence contexts: CG, CHG, and CHH, where H refers to A, C, and T. Accordingly, DNA METHYLTRANSFERASE 1 (MET1), CHROMOMETHYLASE 3 (CMT3), and DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) catalyzes the achievement of DNA methylation in CG, CHG and CHH sequence contexts, respectively (Law and Jacobsen, 2010; Wu and Zhang, 2010). In contrast to the conserved mechanisms in establishing DNA methylation for both mammals and plants, DNA demethylation processes reveal the discrepancies. TEN-ELEVEN TRANSLOCATION (TET) family enzymes govern the removal of DNA methylation in mammals by oxidizing the 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) and coupled with thymine-DNA glycosylase (TDG)-mediated base excision repair (BER) pathway (Wu and

Zhang, 2017). Unlike animals that necessitate the consecutive oxidation for DNA demethylation, plants utilize the *DEMETER* (*DME*) family genes that encode plant-specific DNA glycosylases, which can directly excise 5mC (Choi et al., 2002; Gehring et al., 2006). After the DME-mediated 5mC excision is initiated, a subsequent BER pathway is followed to replace the abasic site (Lee et al., 2014; Martinez-Macias et al., 2012). Alteration of DNA methylation profiles may lead to gene expression changes and entail a change in the corresponding phenotype (Shin et al., 2022). The *FLOWERING WAGENINGEN* (*FWA*) locus is an exemplar of epialleles observed in *Arabidopsis*, the phenotype of which is regulated by DNA methylation level (Soppe et al., 2000).

FWA encoding homeodomain transcription factor, is silenced in wild-type plants with hypermethylation in the upstream region of the coding sequence (Soppe et al., 2000). Besides, the loss of methylation of this locus in *fwa* mutant leads to ectopic expression of *FWA*, accompanied with late-flowering phenotype (Kinoshita et al., 2007). Recently, several pieces of research about targeted manipulation of epimutagenesis on *FWA* locus have been reported, due to its apparent phenotypic difference related to the DNA methylation status (Gallego-Bartolome et al., 2018; Ghoshal et al., 2021; Papikian et al., 2019).

Recent emergence in DNA targeting systems has enabled to target specific sites in the genomes (Gardiner et al., 2022). Notably, clustered regularly interspaced short palindromic repeat (CRISPR)-associated nuclease Cas9 opened a new era with its highly adaptable and accessible manner (Jinek et al., 2012). In addition to using Cas protein to create mutations at target sites, a catalytically deactivated Cas9 protein, dCas9, functions to bind DNA in a guide RNA (gRNA)-targeted fashion (Qi et al.,

2013). Therefore, dCas9 protein was combined with effector domains to apply for the targeted transcriptional regulation or modifying epigenetic states (Nakamura et al., 2021). Beyond the simple direct fusion between dCas9 protein and the effector domain, the SunTag system was incorporated into the dCas9 system. The array of epitope repeats and its antibody fused to an effector domain enabled the SunTag system to recruit multiple copies of the epigenetic modifier at once (Tanenbaum et al., 2014).

Previous studies have reported that dCas9-SunTag system effectively established or removed DNA methylation at target sites (Gallego-Bartolome et al., 2018; Morita et al., 2016; Papikian et al., 2019). Particularly, human DNA demethylase TET1 has been so far utilized in case of targeted removal of DNA methylation (Gallego-Bartolome et al., 2018; Morita et al., 2016). However, plant DNA demethylase DME is not yet applied to epigenome editing. This study introduced the application of DME to targeted DNA demethylation. Compared to TET1, DME catalyzes DNA demethylation via direct 5mC base excision, indicating that DME-based epigenome editing tool is appealing. In addition, dCas9-SunTag system was adopted to improve the targeting efficiency of targeted DNA demethylation. SunTag-DME constructs were able to excise 5mC *in vitro* and shows a preference for the on-target gRNA. These data suggest that SunTag-DME is capable of being utilized as a targeted DNA demethylation system and delivered to the plant for epigenome editing to elicit transcriptional control and further manipulating phenotypes.

MATERIALS AND METHODS

Cloning of SunTag-DME module construct

For the construction of the SunTag-DME module, two separate recombinant plasmids were cloned: pBG102-dCas9-GCN4, and pBG100-scFv-NLS-DME. First, pBG102-dCas9-DMEΔ was utilized to produce pBG102-dCas9-GCN4. pBG102-dCas9-DMEΔ, which DME is flanked between *Spe* I and *Pme* I restriction sites, was digested with these two restriction enzymes to generate pBG102-dCas9. GCN4, serving as peptide epitope tail in SunTag module, was originated from pMOA34-SunTag (Addgene 106435 from Steven Jacobsen). GCN4 and its flanking 2x nuclear localization signal (NLS) were PCR amplified with primers DG4586 and DG4587 (Table 2-1) that harbors *Spe* I and *Pme* I site, and introduced into the pBG102-dCas9 vector, resulting in pBG102-dCas9-GCN4 (Figure 2-1).

The scFv, the antibody of GCN4 peptide epitope, was also derived from pMOA34-SunTag. pMOA34-SunTag plasmid was PCR-amplified with primers DG4588 and DG4589 (Table 2-1) to provide scFv fragment containing *Bam* HI and *Eco* RI restriction sites and subsequently introduced into the corresponding sites of pBG100 vector to produce pBG100-scFv. Thereafter, 2x NLS was adopted in order to ensure proper nuclear import, by which *Eco* RI and *Hind* III restriction sites were added by PCR amplification using DG4591 and DG4592 primer pair (Table 2-1) both at the 5' and 3' end of NLS sequence. Both pBG100-scFv vector and 2x NLS sequence were digested with embracing *Eco* RI and *Hind* III sites, and ligated each other, producing pBG100-scFv-NLS. Finally, DMEΔ fragment was emanated from pEGFP-DMEΔ, which was employed as a template for PCR amplification. The

primer pair DG4593 and DG4594 (Table 2-1) were utilized to clone DMEΔ into the *Hind* III and *Nhe* I restriction sites, located downstream of pBG100-scFv-NLS. DMEΔ, amplified by PCR, was cleaved by *Hind* III and *Nhe* I restriction enzyme, and then cloned into the corresponding sites of the previously manipulated pBG100-scFv-NLS vector (Figure 2-1). Both pBG102-dCas9-GCN4 and pBG100-scFv-NLS-DMEΔ were used for following expression and purification of the proteins for RNP complex assembly.

Expression and purification of the proteins

Recombinant plasmids expressing 6xHis-SUMO-FLAG-dCas9-GCN4 and 6xHis-scFv-DMEΔ were transformed into *E. coli* Rosetta2 (DE3) strain (Novagen), respectively. Cells were grown at 28 °C in 2 L of LB medium containing both 50 μg/mL of kanamycin and chloramphenicol until OD_{600} reached 0.4. Protein expression was induced with 0.1mM IPTG at 16 °C for 16h. Cells were harvested by centrifugation at 7,000 rpm for 20 min at 4 °C and resuspended in 30 mL of lysis buffer (50 mM Tris-HCl, pH 7.4, 100 mM NaCl, 10% glycerol, 0.1 mM DTT, 0.1 mM PMSF). Cells were sonicated for 10 min (0.5x duty cycle; Fisher Scientific 550 Sonic Dismembrator) for lysis on ice. Cell extracts were centrifuged at 12,000 rpm for 30 min at 4 °C. The supernatant was filtered through 0.45 μm nylon membrane (Adventec) and the cleared lysate was sequentially purified by three different types of columns: affinity column (His Trap FF 5 mL, GE Healthcare), ion exchange column (Heparin HP 5 mL, GE Healthcare), and size exclusion (Superdex 200, GE Healthcare). The final eluted fractions were concentrated and aliquoted with 50% glycerol and stored at -80 °C until use (Figure 2-2).

Single-guide RNA (sgRNA) design and *in vitro* transcription

The *FWA*-sgRNA was designed for target-specific site with higher out-of-frame scores in the upstream of *Arabidopsis FWA* region (Figure 2-3) via CRISPR RGEN tools (Bae et al., 2014). For sgRNA preparation, a 20-bp gRNA sequence was inserted into U6::sgRNA cassette of a pHAtC plasmid (Kim et al., 2016). Then, a template of in vitro RNA transcription that contains 20 bp gRNA sequence and full-length gRNA, followed by T7 promoter was PCR-amplified from the recombinant pHAtC vector using DG4656 and DG4451 (Table 2-1). *In vitro* transcription of sgRNA was undertaken using RNA synthesis kit HiScribe™ T7 High Yield RNA synthesis kit (NEB). Synthesized RNA product was treated with DNase I (NEB) in 1x DNase I reaction buffer and further purified using MEGAclear™ Transcription Clean-Up Kit (Thermo).

***In vitro* 5mC excision assay of SunTag-DME**

The 35-mer oligonucleotide containing 5mC in the middle of the sequence (5'– CTATACCTCCTCAACTC[5mC] GGTCACCGTCTCCGGCC -3') and its complementary strand without 5mC were synthesized (Integrated DNA Technologies) (Figure 2-4A). Forty pmol of oligonucleotide with 5mC was radiolabeled by T4 polynucleotide kinase (TaKaRa) with [γ-^{32}P] ATP (Perkin Elmer). The labeled oligonucleotide was purified by Qiaquick Nucleotide Removal Kit (Qiagen) and annealed with complementary strand oligonucleotide to produce double-stranded DNA substrate. The substrates were boiled in water for 10 min and slowly cooled down for annealing to room temperature for 3h.

For the analysis of 5mC excision activity, 25 nM of radiolabeled

oligonucleotide substrate was incubated with 100 nM DMEΔ or 25 nM SunTag-DME complex in the DNA glycosylase reaction buffer (10 mM Tris-HCl, pH 7.4, 50 mM NaCl, 0.5 mM DTT) at 37 °C for 1h. Reaction was terminated by adding equal volume of stop solution (98% formamide, 10 mM EDTA, 0.2% Xylene cyanol FF, 0.2% bromophenol blue) and heated at 95 °C for denaturation of 10 min. Reaction products were separated on a 15% denaturing polyacrylamide gel containing 7.5 M urea and 1X TBE.

Targeting preference assay of SunTag-DME

To produce the unlabeled cold unlabeled oligo with same concentration of radiolabeled hot oligo substrates, containing a couple of 5mC site in the 14^{th} and 41^{st} of the sequence (5'-GTAAAGCGGTGCT[5mC]GTATGAATGTTGAATGGGAT AAAGAG[5mC]GGCGCAAGATC -3') was purified as described above. For the analysis of substrate competition assay, 25 nM of SunTag-DME complex proteins were premixed with designed sgRNA as a 1:5 M ratio for 15 min at room temperature for the preassembly of RNP complex. 100 nM DMEΔ or aforementioned preassembled SunTag-DME RNP complex was incubated with both 25 nM of radiolabeled oligo and cold competitor substrates in the DNA glycosylase reaction buffer (10 mM Tris-HCl, pH 7.4, 50 mM NaCl, 0.5 mM DTT) at 37 °C for 1h. Reaction was terminated by adding equal volume of stop solution (98% formamide, 10 mM EDTA, 0.2% Xylene cyanol FF, 0.2% bromophenol blue) and heated at 95 °C for denaturation of 10 min. Reaction products were separated on a 15% denaturing polyacrylamide gel containing 7.5 M urea and 1X TBE.

Table 2-1. Primers used in this study

Name	Sequence (5' →3')
DG4586	AATTACTAGTCCTAAGAAAAAGCGGAAAGTGG
DG4587	AATTGTTTAAACTTACCCTGAGCCTGATCCC
DG4588	AATTGGATCCGACATCGTGATGAC
DG4589	AATTGAATTCGCTGCTCACGGTC
DG4591	AATTGAATTCAGCCTGGGCAGC
DG4592	AATTAAGCTTTGATCCGCCACCG
DG4593	AATTAAAGCTTTACAAAGGAGATGGTGC
DG4594	AATTGCTAGCTTAGGTTTTGTTGTTCTTC
DG4656	AATTCTAATACGACTCACTATAGGGGTGCTCGTATGAATGTTGAA
DG4451	GCACCGACTCGGTGCCACTT

RESULTS

Generation of DNA-free epigenome editing module with SunTag-DME

In plants, the DME family genes encode bifunctional 5mC DNA glycosylase/AP lyases that catalyze the excision of 5mC (Gehring et al., 2006; Gong et al., 2002; Penterman et al., 2007). DME family proteins consist of three essential domains, the glycosylase domain and flanking domains A and B. In particular, the engineered DME protein in which N-terminal region was removed and interdomain region I (IDR1) was replaced with a synthetic flexible linker showed 5mC excision activity (Mok et al., 2010). In this study, the engineered DME protein called DMEΔ hereafter was utilized as an effector protein. However, DME binds to DNA independent of sequence contexts and the way of recognizing 5mC remains elusive (Mok *et al.*, 2010). Therefore, tools accompanying DNA-binding ability in sequence specific manner is indispensable to achieve targeted DNA demethylation. Recently, CRISPR/dCas9 system has risen as an advantageous platform for epigenome editing. Unlike other DNA-binding module such as zinc-finger (ZF) protein, or transcription activator-like effector (TALE), CRISPR-dCas proteins retain genomic specificity by their associated gRNA, and are highly adaptable. Therefore, I utilized dCas9-based SunTag system as a programmable DNA binding module which can also bind DNA in a gRNA-dependent manner. Moreover, considering the ability of SunTag to mobilize multiple effector proteins simultaneously, the enhanced effect was anticipated (Tanenbaum et al., 2014). In this study, I designed the two separate components of SunTag system, SunTag-DME, a dCas9 fused to GCN4 epitope repeats and its cognate scFv antibody with DMEΔ (Figure 2-1). Both proteins were

successfully expressed in the *E. coli* system, and purified with high quality by chromatography based on affinity and size exclusion-manner (Figure 2-2). The designed sgRNA imposed SunTag-DME to target the 5'-regulatory upstream regions of *FWA* locus, hypermethylated in wild-type plants (Figure 2-3). Consequently, the two proteins and sgRNA were assembled to generate the RNP complex *in vitro*, give rise to DNA-free epigenome editing tools and further 5mC excision ability or gRNA-directed DNA targeting capability need to be elucidated.

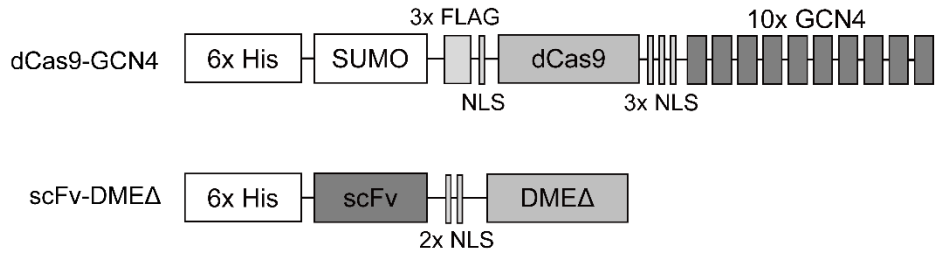

Figure 2-1. Schematic diagrams of SunTag-DME components.

SunTag system requires two different components. dCas9 is followed by ten copies of GCN4 epitope tails, the protein scaffold of which is able to interact with its antibody. Four repeats of NLS were flanked to maximize the nuclear import to target loci specifically in the plant genome. The complementary antibody scFv, interacting with GCN4 repeats to generate SunTag complex, is fused with DMEΔ for direct 5mC excision activity. Histidine-tag was attached to the N-terminus for purification of the proteins.

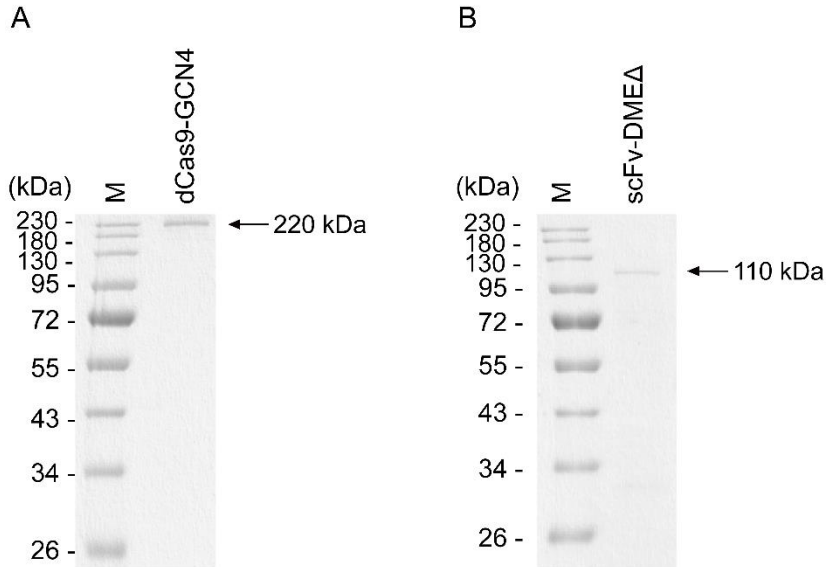

Figure 2-2. SDS-PAGE analysis of purified proteins.

The purified proteins (200 ng) of SunTag-DME components, dCas9-GCN4 (A) and scFv-DMEΔ (B), were electrophoresed on a 10% SDS-PAGE gel. The size of each protein is indicated to the right. M, size maker.

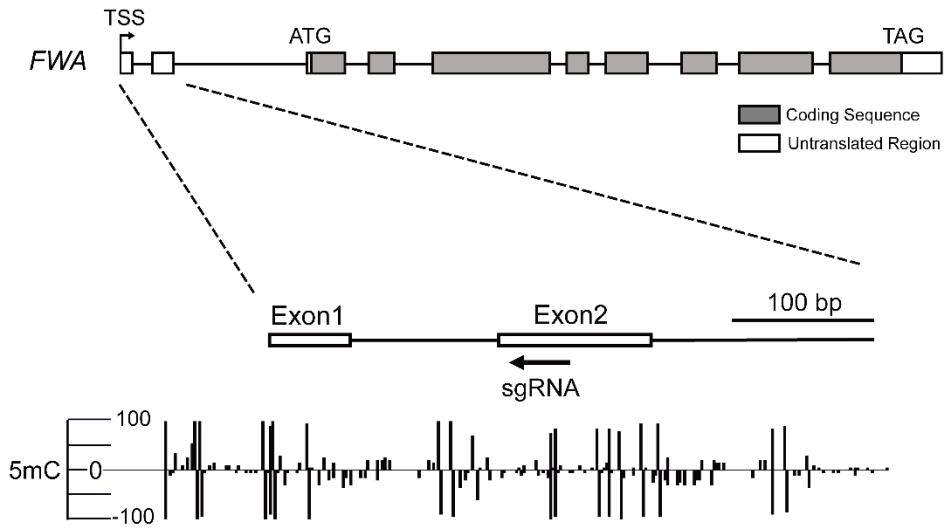

Figure 2-3. Schematic diagram of A*tFWA* locus and sgRNA.

The *Arabidopsis FWA* locus, heavily methylated in wild type plants, is delineated concisely. The methylation level at the *FWA* promoter is indicated by black lines. The designed sgRNA binds at the 5' upstream regions of coding sequence in *FWA* second exon to achieve targeted removal of DNA methylation. TSS refers to transcriptional start site. The coding sequence is indicated by gray boxes.

SunTag-DME is able to excise 5mC *in vitro*.

DME is a bifunctional DNA glycosylase in plants, responsible for DNA demethylation through direct 5mC excision and succeeding BER pathway, resulting in β-, and δ- elimination products (Gehring et al., 2006). In this study, I utilized DMEΔN677ΔIDR1::lnk, referred to as DMEΔ, being an effector protein in SunTag-DME module. Through previous studies that the engineered DME protein fused with Maltose Binding Protein (MBP) or Small Ubiquitin-like Modifier (SUMO) tags successfully remove 5mC *in vitro* (Jang et al., 2014; Mok et al., 2010). I confirmed the 5mC glycosylase activities of both scFv-DMEΔ and SunTag-DME generating β- and δ- elimination product, provided with random 35-bp DNA substrate including 5mC in the 18th position (Figure 2-4). The scFv-DMEΔ retains 5mC excision activity in the way that DMEΔ resulted in β-, and δ- elimination products, despite the scFv fusion at the N-terminal of the DMEΔ (Lane 4 in Figure 2-4C). Particularly, the coexistence of dCas9-GCN4 and scFv-DMEΔ, give rise to SunTag-DME complex, showed a dominant 5mC excision product relative to scFv-DMEΔ solo. Given that the fusion of MBP tag with scFv allowed stable, soluble, and functional fusion protein, the instability and insolubility of scFv-DMEΔ might affect 5mC DNA glycosylase activity (Bach et al., 2001). These results indicate that scFv-DMEΔ is able to bind thoroughly with its complementary dCas9-GCN4, leading to successful 5mC excision activity. Even though, scFv-DMEΔ solo protein showed a hampered DNA glycosylase activity, this might elicit the minor off-target effect of scFv-DMEΔ is expected.

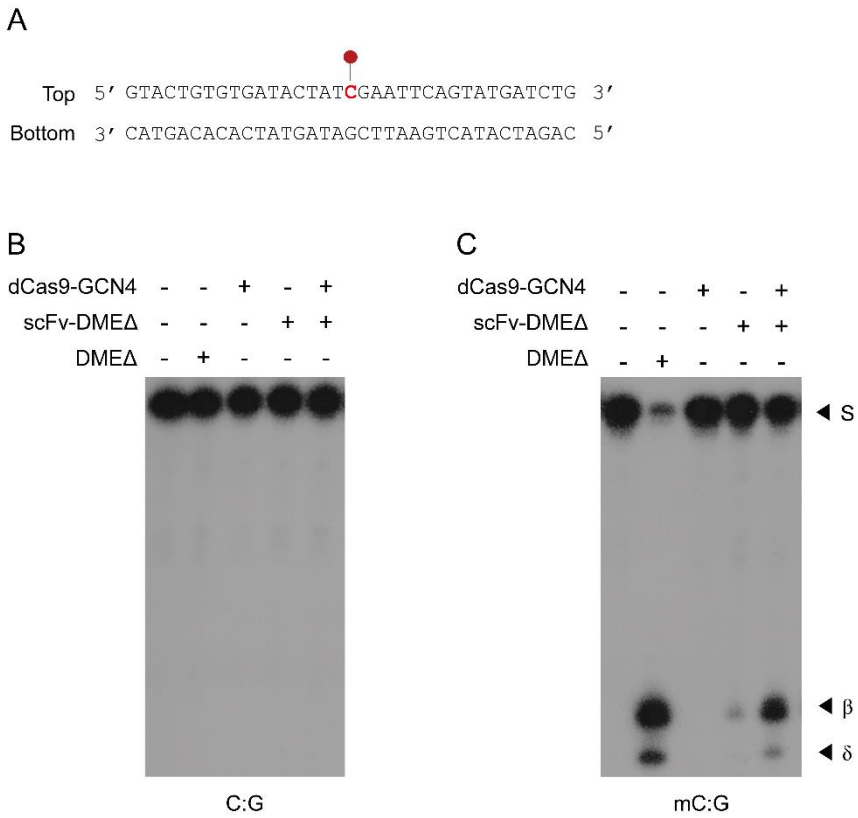

Figure 2-4. The *in vitro* 5mC excision activity of SunTag-DME module.

(A) Schematic diagram of substrate for 5mC excision assay. Double-stranded DNA oligonucleotide substrates were labeled at the 5' end of the top strand. The cytosine base with the red lollipop above indicates 5mC. (B), (C) *In vitro* 5mC glycosylase assay with the SunTag-DME module. The radiolabeled DNA substrate (25 nM) containing cytosine (B) and 5mC (C) was incubated with 100 nM DMEΔ, both 25 nM dCas9-GCN4 and scFv-DMEΔ at 37°C for 1h. Positions of the oligonucleotide substrate (S), and the reaction products (β, δ) were indicated to the right of the panel.

The targeting preference of SunTag-DME is dependent upon a gRNA.

Since DME possesses DNA binding activity itself, the ability of SunTag-DME module to be recruited to the target site designated by sgRNA needed to be elucidated (Mok et al., 2010). To verify the targeting preference of SunTag-DME in a sgRNA-dependent manner, two distinct DNA substrates were used (Figure 2-5A). One was the random 35-bp DNA substrate, identical to the one used in the previous 5mC excision assay. The other is the 52-bp long substrate, the sequence of which is the subset of the FWA promoter region, where designed gRNA can bind. Also, two individual 5mC was contained at the 14th, and the 41st of the sequence. Among those two DNA oligo substrates, the one with the former random 35-bp long substrate was radiolabeled, while the other being unlabeled cold oligo. Lastly, to verify the targeting of SunTag-DME in a gRNA-dependent manner, the gRNA corresponding to 52-bp long unlabeled cold oligo, containing 5mC residing at the 14th position, was used. I hypothesized that SunTag-DME module might be able to excise the 41st position of 5mC predominantly, if it was guided by given gRNA in a proper-manner. The scFv-DMEΔ removing both radiolabeled and the cold oligonucleotides, with reduced amount of 5mC excision product compared to single oligo surroundings, reveals that 5mC excision activity of scFv-DMEΔ occurs regardless of sequence contexts *in vitro* (Figure 2-4C and 2-5B). However, when the scFv-DMEΔ was recruited to dCas9-GCN4, being SunTag-DME, the extent of β-, and δ- elimination products were diminished in the presence of sgRNA (Figure 2-5B, C). These reveal that SunTag-DME was directed toward the unlabeled cold oligonucleotide prevalently by its complementary sgRNA.

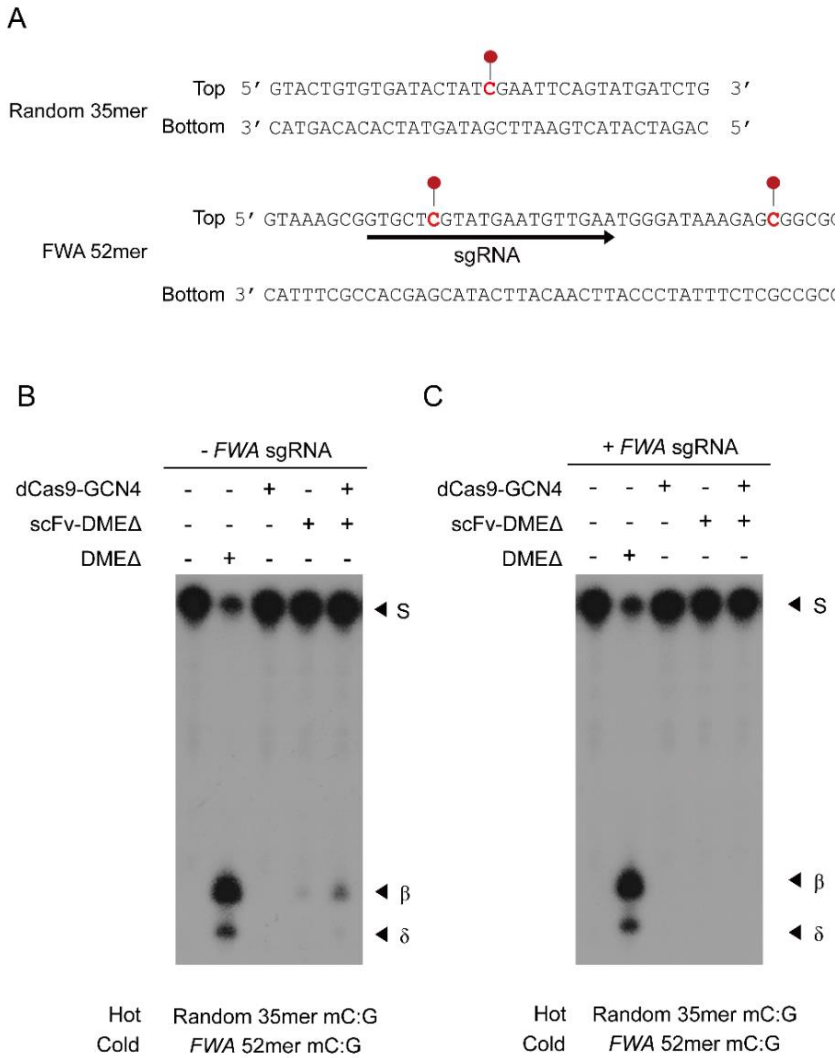

Figure 2-5. The target preference assay of SunTag-DME module.

(A) Schematic diagrams of substrates for target preference assay. Both double-stranded DNA oligonucleotide substrates were labeled at the 5' end of the top strand. The black arrow indicates the position of sgRNA targeting site in the 52 base pair-long substrate, a subset of *FWA* promoter. The lollipops with red indicate 5mC. (B), (C) *In vitro* target preference assay with the SunTag-DME module. The radiolabeled DNA substrate (25 nM) and the cold substrates (25 nM), which is a subset of A*tFWA* promoter region without sgRNA (B), and with sgRNA (125 nM) (C) were incubated with 100 nM DMEΔ, both 25 nM dCas9-GCN4 and scFv-DMEΔ at 37°C for 1h (B). Positions of the oligonucleotide substrate (S), and the reaction products (β, δ) were indicated to the right of the panel.

DISCUSSION

CRISPR/dCas system has been widely employed in epigenome editing with an ability that can bind to DNA sequence-specifically (Gardiner et al., 2022). Notably, TET1 mammalian DNA demethylase was broadly exploited with dCas9 to remove repressive DNA methylation and elicit transcriptional activation (Choudhury et al., 2016; Liu et al., 2018). Recently, the SunTag module successfully accomplished epigenome editing by recruiting multiple effectors to the target loci in plants (Gallego-Bartolome et al., 2018; Ghoshal et al., 2021; Papikian et al., 2019).

In this study, I propose an application of plant-specific DNA glycosylase DME-mediated epigenome editing tools using dCas9/SunTag system. Also, RNP complex, including both dCas9-GCN4 and scFv-DMEΔ proteins and its accessory sgRNA, was assembled to SunTag-DME and verified the feasibility *in vitro*. I showed that SunTag-DME complex is able to excise 5mC (Figure 2-4) and its target preference is determined by sgRNA sequence (Figure 2-5).

The importance of designing a suitable position of sgRNA has been emanated not only to avoid the off-target effect, but also to boost the manipulating efficiency (Uniyal et al., 2019). In addition to sgRNA can designate the targeting position of SunTag-DME, the determination of effective range is required. I hypothesized that the scFv-fused effector protein reacts with DNA in the C-terminus direction of dCas9 due to its configuration of SunTag complex. However, SunTag-DME module catalyzed excision of 5mC at the position of downstream of sgRNA binding site, resulting in dominant β-, and δ- elimination products. To identify the effective range of SunTag-DME, a further experiment using substrate with broader

range including varying DNA methylation sites may be required. Furthermore, given the architecture of SunTag system, which can recruit up to ten effector proteins, the extent of demethylation with the increased amount of scFv-DMEΔ can be elucidated (Morita et al., 2016; Tanenbaum et al., 2014).

REFERENCES

Bach, H., Mazor, Y., Shaky, S., Shoham-Lev, A., Berdichevsky, Y., Gutnick, D.L., and Benhar, I. (2001). Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. J. Mol. Biol. 312, 79-93.

Bae, S., Park, J., and Kim, J.S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30, 1473-1475.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., and Fischer, R.L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. Cell 110, 33-42.

Choudhury, S.R., Cui, Y., Lubecka, K., Stefanska, B., and Irudayaraj, J. (2016). CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at *BRCA1* promoter. Oncotarget 7, 46545-46556.

Erdel, F. (2017). How Communication Between Nucleosomes Enables Spreading and Epigenetic Memory of Histone Modifications. Bioessays 39.

Gallego-Bartolome, J., Gardiner, J., Liu, W., Papikian, A., Ghoshal, B., Kuo, H.Y., Zhao, J.M., Segal, D.J., and Jacobsen, S.E. (2018). Targeted DNA demethylation of the *Arabidopsis* genome using the human TET1 catalytic domain. Proc. Natl. Acad. Sci. U S A 115, E2125-E2134.

Gardiner, J., Ghoshal, B., Wang, M., and Jacobsen, S.E. (2022). CRISPR-Cas-mediated transcriptional control and epi-mutagenesis. Plant Physiol. 188, 1811-1824.

Gehring, M., Huh, J.H., Hsieh, T.F., Penterman, 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.

Ghoshal, B., Picard, C.L., Vong, B., Feng, S., and Jacobsen, S.E. (2021). CRISPR-based targeting of DNA methylation in *Arabidopsis thaliana* by a bacterial CG-specific DNA methyltransferase. Proc. Natl. Acad. Sci. U S A 118.

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

Heard, E., and Martienssen, R.A. (2014). Transgenerational epigenetic inheritance: myths and mechanisms. Cell 157, 95-109.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. 33 Suppl, 245-254.

Jang, H., Shin, H., Eichman, B.F., and Huh, J.H. (2014). Excision of 5-hydroxymethylcytosine by DEMETER family DNA glycosylases. Biochem. Biophys. Res. Commun. 446, 1067-1072.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821.

Kim, H., Kim, S.T., Ryu, J., Choi, M.K., Kweon, J., Kang, B.C., Ahn, H.M., Bae, S., Kim, J., Kim, J.S., and Kim, S.G. (2016). A simple, flexible and high-throughput cloning system for plant genome editing via CRISPR-Cas system. J. Integr. Plant Biol. 58, 705-712.

Kinoshita, Y., Saze, H., Kinoshita, T., Miura, A., Soppe, W.J., Koornneef, M., and Kakutani, T. (2007). Control of *FWA* gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats. Plant J. 49, 38-45.

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.

Lee, J., Jang, H., Shin, H., Choi, W.L., Mok, Y.G., and Huh, J.H. (2014). AP endonucleases process 5-methylcytosine excision intermediates during active DNA demethylation in *Arabidopsis*. Nucleic Acids Res. 42, 11408-11418.

Liu, X.S., Wu, H., Krzisch, M., Wu, X., Graef, J., Muffat, J., Hnisz, D., Li, C.H., Yuan, B., Xu, C., et al. (2018). Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the *FMR1* Gene. Cell 172, 979-992 e976.

Martinez-Macias, M.I., Qian, W., Miki, D., Pontes, O., Liu, Y., Tang, K., Liu, R., Morales-Ruiz, T., Ariza, R.R., Roldan-Arjona, T., and Zhu, J.K. (2012). A DNA 3' phosphatase functions in active DNA demethylation in *Arabidopsis*. Mol. Cell 45, 357-370.

Mok, Y.G., Uzawa, R., Lee, J., Weiner, G.M., Eichman, B.F., Fischer, R.L., and Huh, J.H. (2010). Domain structure of the DEMETER 5-methylcytosine DNA glycosylase. Proc. Natl. Acad. Sci. U S A 107, 19225-19230.

Morita, S., Noguchi, H., Horii, T., Nakabayashi, K., Kimura, M., Okamura, K., Sakai, A., Nakashima, H., Hata, K., Nakashima, K., and Hatada, I. (2016). Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. Nat. Biotechnol. 34, 1060-1065.

Nakamura, M., Gao, Y., Dominguez, A.A., and Qi, L.S. (2021). CRISPR technologies for precise epigenome editing. Nat. Cell Biol. 23, 11-22.

Papikian, A., Liu, W., Gallego-Bartolome, J., and Jacobsen, S.E. (2019). Site-specific manipulation of *Arabidopsis* loci using CRISPR-Cas9 SunTag systems. Nat. Commun. 10, 729.

Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., and Fischer, R.L. (2007). DNA demethylation in the *Arabidopsis* genome. Proc. Natl. Acad.

Sci. U S A 104, 6752-6757.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173-1183.

Shin, H., Choi, W.L., Lim, J.Y., and Huh, J.H. (2022). Epigenome editing: targeted manipulation of epigenetic modifications in plants. Genes Genomics 44, 307-315.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat. Rev. Genet. 14, 204-220.

Soppe, W.J., Jacobsen, S.E., Alonso-Blanco, C., Jackson, J.P., Kakutani, T., Koornneef, M., and Peeters, A.J. (2000). The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. Mol. Cell 6, 791-802.

Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell 159, 635-646.

Uniyal, A.P., Mansotra, K., Yadav, S.K., and Kumar, V. (2019). An overview of designing and selection of sgRNAs for precise genome editing by the CRISPR-Cas9 system in plants. 3 Biotech 9, 223.

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

Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. Nat. Rev. Genet. 18, 517-534.

Zhu, J.K. (2009). Active DNA demethylation mediated by DNA glycosylases. Annu. Rev. Genet. 43, 143-166.

초록

DNA 메틸화는 고등생물에서 지놈의 안정성과 유전자의 발현에 관여하는 주요 후성유전학적 인자이다. DNA 메틸화 양상은 발달과정과 환경적 요인에 따라 특정 효소들에 의해 매개되는 시토신과 5-메틸시토신 간의 가역적인 변화를 의미한다. 이때 동물과 식물의 DNA 메틸화 효소들은 진화생물학적으로 그 기작에 있어 유사성을 보이지만 DNA 탈메틸화 효소의 경우, 각각 특이적으로 진화하였다. 애기장대에서 DNA 탈메틸화를 주관하는 DNA 글라이코실라제 중 DEMETER (DME)는 5-메틸시토신을 직접 제거하며, 식물의 중복수정과 종자발달에 있어 중요한 역할을 담당한다. 제 1 장에서는 애기장대의 원형질체에 DME 단백질을 과발현시킬 수 있는 재조합 플라스미드를 도입하여 시간에 따른 효과를 관찰하였다. 초기 12 시간의 경우, DME 단백질 과발현 조건에서 이동성 유전인자들의 발현량이 대조군인 무처리와 DME 돌연변이 처리구에 비해 10 배 가까이 증가하였다. 추후 시간이 흐름에 따라 24 시간이 경과하였을 때에는 여러 식물스트레스 관련 유전자들의 발현이 변화함을 확인하였다. 제 2 장에서는 DME 단백질의 능동적 탈메틸화능을 응용한 후성유전체 편집모듈인 SunTag-DME 를 제작하였으며, 고효율의 선택적 탈메틸화 편집을 유도하기 위해 고순도의 단백질을 정제할 수 있는 조건을 탐색하였다. 이를 통해 제작한 SunTag-DME 리보핵단백질은 시험관 내에서 5-메틸시토신 제거 활성을 보임과 동시에 single-guide RNA (sgRNA)에

따른 기질선택성을 나타냄을 확인하였다. 따라서 본 연구는 DME 의 식물 원형질체 내에서의 효과 및 기능에 대한 이해를 넓힐 수 있었으며, 단백질의 생화학적 특성에 기반한 선택적 후성유전체 편집의 가능성을 제시하였다.

주요어: DNA 메틸화, DEMETER, 식물 원형질체, 후성유전체 편집, dCas9-SunTag

학번: 2021-28426