Control of DEMETER DNA demethylase gene transcription in male and female gamete companion cells in *Arabidopsis thaliana*

Jin-Sup Park[a,1], Jennifer M. Frost[b,1], Kyunghyuk Park[a], Hyonhwa Ohr[a], Guen Tae Park[a], Seohyun Kim[a], Hyunjoo Eom[a], Ilha Lee[a], Janie S. Brooks[c], Robert L. Fischer[b,2], and Yeonhee Choi[a,d,2]

[a]Department of Biological Sciences, Seoul National University, Seoul 151-747, Korea; [b]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720; [c]Department of Science, Seoul Foreign School, Seoul 120-823, Korea; and [d]Plant Genomics and Breeding Institute, Seoul 151-921, Korea

Contributed by Robert L. Fischer, December 22, 2016 (sent for review April 28, 2016; reviewed by Robert G. Franks and Mark A. Johnson)

The DEMETER (DME) DNA glycosylase initiates active DNA demethylation via the base-excision repair pathway and is vital for reproduction in *Arabidopsis thaliana*. DME-mediated DNA demethylation is preferentially targeted to small, AT-rich, and nucleosome-depleted euchromatic transposable elements, influencing expression of adjacent genes and leading to imprinting in the endosperm. In the female gametophyte, *DME* expression and subsequent genome-wide DNA demethylation are confined to the companion cell of the egg, the central cell. Here, we show that, in the male gametophyte, *DME* expression is limited to the companion cell of sperm, the vegetative cell, and to a narrow window of time: immediately after separation of the companion cell lineage from the germline. We define transcriptional regulatory elements of *DME* using reporter genes, showing that a small region, which surprisingly lies within the *DME* gene, controls its expression in male and female companion cells. *DME* expression from this minimal promoter is sufficient to rescue seed abortion and the aberrant DNA methylome associated with the null *dme-2* mutation. Within this minimal promoter, we found short, conserved enhancer sequences necessary for the transcriptional activities of *DME* and combined predicted binding motifs with published transcription factor binding coordinates to produce a list of candidate upstream pathway members in the genetic circuitry controlling DNA demethylation in gamete companion cells. These data show how DNA demethylation is regulated to facilitate endosperm gene imprinting and potential transgenerational epigenetic regulation, without subjecting the germline to potentially deleterious transposable element demethylation.

DNA demethylation | central cell | vegetative cell | cell-specific transcription | DNA enhancer elements

Sexual reproduction is characterized by fertilization of an egg by a sperm cell, generating the embryo. Uniquely in angiosperms, a second sperm cell fertilizes the companion cell of the egg, the central cell, to generate the endosperm, which supports development of the embryo. During reproduction in angiosperm *Arabidopsis thaliana*, the DEMETER (DME) DNA glycosylase exhibits a striking expression pattern. Within the ovule, the female gametophyte is generated by mitosis of the haploid megaspore, forming a mature gametophyte of seven cells. During this process, the egg and central cell lineages are separated, and, at this point, *DME* expression and DNA demethylation is initiated solely in the central cell (1, 2). *DME* expression is switched off after fertilization (2). This precise pattern of expression initiated in the central cell, and not in the egg cell, is responsible for hypomethylation specifically in the maternal endosperm genome and not in the maternal embryo genome (3). *DME* expression in the central cell is essential for plant reproduction and genomic imprinting, whereby its absence results in loss of genomic imprinting, aberrant endosperm development, and early seed abortion (2, 4, 5).

In the male gametophyte, indirect evidence suggests that *DME* is expressed during development of the mature three-cell pollen grain, perhaps originating specifically in the vegetative cell, the companion cell of the two sperm cells (6). During reproduction, the vegetative cell generates a pollen tube that transports two sperm cells to the ovule for double fertilization. Although paternal inheritance of a *DME* mutation is compatible with normal seed development, it does result in decreased pollen viability and germination rates in certain ecotypes (6, 7).

DME is required to demethylate regions of DNA as part of the base-excision repair (BER) pathway. The dual activity helix-hairpin-helix glycosylase family consists of DME, REPRESSOR OF SILENCING1 (ROS1), and DEMETER-LIKE (DML) 2 and 3. Each glycosylase enzyme acts to remove 5-methylcytosine and nick the DNA backbone, followed by repair and replacement with cytosine by downstream enzymes in the BER pathway (4, 8–10). Within the glycosylase family of DNA demethylating enzymes, *DME* is distinguished by its highly restricted pattern of expression in gamete companion cells, as well as its profound effects on plant reproduction. The consequence of silencing the maternal *DME* allele is in the aberrant retention of DNA methylation on the maternal endosperm genome, including the imprinting control regions of imprinted genes (3, 4). Notably, maternal expression of *MEDEA* (*MEA*) and *Fertilization Independent Seed 2* (*FIS2*), which form part of the floral polycomb repressive complex 2 (PRC2), involved

Significance

The *Arabidopsis* DEMETER (DME) DNA demethylase is required for reproduction and endosperm gene imprinting. We investigated mechanisms that restrict DME transcription to the female and male companion cells, the central and vegetative cells, adjacent to egg and sperm cells, respectively. We delineated a region, surprisingly within the *DME* gene, sufficient for proper DME expression, which rescues seed abortion and aberrant endosperm DNA methylation associated with *dme* mutations. We discovered overlapping DNA enhancers promoting vegetative and central cell expression, which can be used to pinpoint candidate transcription factors that regulate DME in companion cells. These results show how reproductive DNA demethylation is restricted to companion cells, thereby protecting egg and sperm cells from deleterious DNA demethylation that would be transmitted to progeny.

Author contributions: J.-S.P., J.M.F., I.L., R.L.F., and Y.C. designed research; J.-S.P., J.M.F., K.P., H.O., G.T.P., S.K., and H.E. performed research; J.-S.P., J.M.F., J.S.B., R.L.F., and Y.C. analyzed data; and J.-S.P., J.M.F., J.S.B., R.L.F., and Y.C. wrote the paper.

Reviewers: R.G.F., North Carolina State University; and M.A.J., Brown University.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]J.-S.P. and J.M.F. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: rfischer@berkeley.edu or yhc@snu.ac.kr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1620592114/-/DCSupplemental.

in chromatin organization and regulation, requires DME action. Without DME-mediated DNA demethylation, the expression of these genes is lost, resulting in a loss of PRC2 and subsequent seed abortion.

DME also has a second function, which potentially impacts plant DNA methylation transgenerationally. DME-mediated DNA demethylation in companion cells is preferentially targeted to small, AT-rich, and nucleosome-depleted euchromatic transposable elements (TEs) (3). Evidence suggests that TE hypomethylation in the companion cells promotes transcription of mobile siRNA at the TEs, mediating RNA-directed DNA methylation (RdDM) in the gametes, so that the same TE sequences become hypermethylated, safeguarding the genomic integrity of the gametes (3, 11, 12). The large overlap between sites demethylated in the central cell, inferred from hypomethylated sites in the maternal endosperm genome (13) and sites demethylated in the vegetative cell, despite their different cell fates, provides evidence toward this common basal function of *DME* expression in gamete companion cells.

Both for the appropriate expression of imprinted genes during seed development, and for the putative role of DME in transgenerational epigenetic regulation, it is vital that *DME* expression is confined to the companion cells of the gametes, and not in the gametes themselves. We therefore sought to delineate the mechanisms affording this important expression profile.

Results

DME Is Expressed Specifically in the Companion Cell of the Male Gametophyte After Separation of the Sperm Cell Lineage. During pollen development, a haploid microspore undergoes an asymmetric mitosis to produce a bicellular pollen with a generative cell engulfed in the vegetative cell. A second mitosis of the generative cell generates two sperm cells (Fig. 1 *A* and *B*). Previously, a low level of *DME* transcripts had been detected in mature pollen grains but not in sperm nuclei whereas DME-mediated DNA demethylation was shown to be restricted to the vegetative cell, implicating the vegetative cell as the site of *DME* expression (6). However, the precise pattern of *DME* expression during male gametophyte development is unknown. To address this issue, we measured β-glucuronidase (GUS) and green fluorescent protein (GFP) reporter expression in pollen from plants bearing the previously described *2.3pDME::GUS/GFP* transgene. The *2.3pDME::GUS/GFP* construct has 2.3 kb of upstream sequence and 2 kb of the *DME* transcriptional unit fused to GUS or GFP and is expressed in the central cell of the female gametophyte (Fig. S1*A*) (2, 14). GUS or GFP reporter expression was detected only in the vegetative cell nucleus of late bicellular pollen: that is, after the first asymmetric mitosis, but not in the generative or sperm cell nuclei, or at any other stages of pollen development (Fig. 1 *A*, *Bottom* and *B*). Real-time quantitative RT-PCR (qRT-PCR) analysis was in accord with these results, showing elevated *DME* RNA expression at the bicellular pollen stage, followed by rapid decreases as pollen matured (Fig. 1*C*). Thus, *DME* expression is not detected until the sperm cell lineage is separated from that of the vegetative cell, at which point DME is active specifically in the vegetative cell.

The DME Promoter Lies Within the DME Transcriptional Unit and Contains both Positive and Negative Regulatory Elements. To identify the elements that promote the striking pattern of *DME* expression in male and female companion cells, we systematically deleted portions of our *2.3pDME::GUS* reference construct (Fig. 2*A*). Deletion of the entire 5′ region, from −2.3 kb to +46 bp downstream of the transcriptional start site (TSS), as defined by 5′ RACE (Fig. 2 *A* and *B* and Fig. S2), had no effect on *DME::GUS* expression in the central and vegetative cells. For each of these deletion constructs, both temporal and spatial *DME::GUS* expression profiles in transgenic plants reflected those of the reference construct (Fig. 2 *A* and *B* and

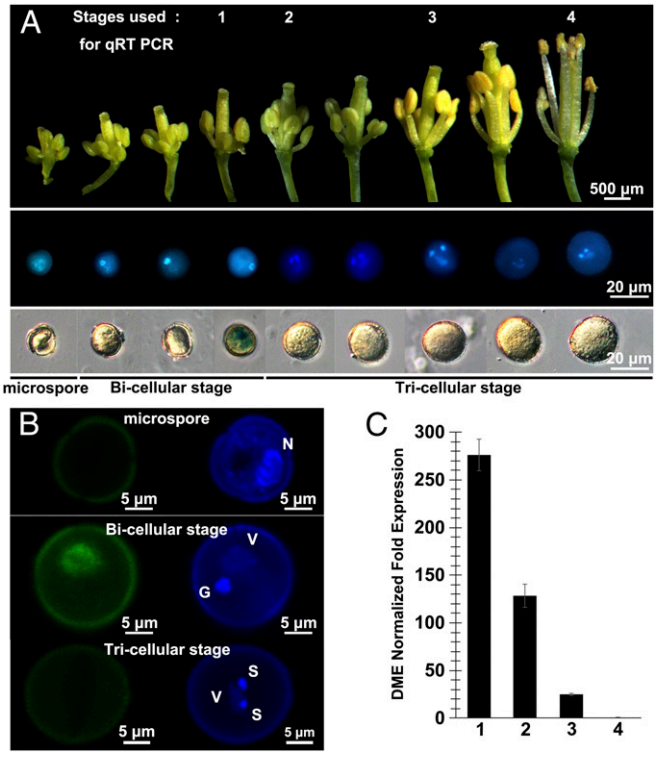

Fig. 1. *DME* is specifically expressed in the vegetative nucleus of late bicellular stage pollen. (*A*) Sequential development of flowers (*Top*) and corresponding pollen development in *2.3 kb DME::GUS* transgenic plants with DAPI (*Middle*) and GUS staining (*Bottom*). (*B*) The *2.3 kb DME::GFP* expression (*Left*) in microspore (*Top*), bicellular (*Middle*), and tricellular (*Bottom*) stage pollen grains stained with DAPI (*Right*). G, generative nucleus; N, microspore nucleus; S, sperm cell nucleus; V, vegetative nucleus. (Scale bars: 5 μm.) (*C*) qRT-PCR analysis of DME expression in WT pollen development after normalization with *ACT1*, *ACT3*, and *ACT12* expression. The four different stages analyzed using qRT-PCR are indicated in *A*. Values are plotted relative to the expression of DME in stage 4 mature pollen, which was set at 1.0, and represent the average of triplicate measurements ± SD.

Fig. S1*A*). We then deleted a larger block of sequence, up to 395 bp downstream of the TSS, at which point *DME* expression was decreased, and, finally, deletion of *DME* transcriptional unit sequence to 473 bp downstream of the TSS led to the complete loss of *DME::GUS* expression in both central and vegetative cells (Fig. 2 *A* and *B*). These data indicate that the regulatory sequences that are required for the proper expression of DME in the central and vegetative cells lie between 46 and 473 bp downstream of the TSS.

To verify genetically that DNA sequences upstream of the TSS do not regulate DME expression, we obtained two transfer DNA (T-DNA) insertion mutants from the *Arabidopsis* Biological Resource Center (ABRC): *CS857766,* which has a T-DNA insertion 72 bp upstream (−72) of the TSS, and *SALK-036171,* which has a T-DNA insertion 25 bp upstream (−25) of the TSS (Fig. 2*C*). Homozygous mutants of either line were developmentally and morphologically indistinguishable from WT and did not exhibit any defects in fertility or seed viability (Table S1), suggesting that *DME* is appropriately expressed and functions normally in these mutants. *DME* is also expressed in sporophyte tissues (14), and we found the level of *DME* expression in homozygous *CS857766* and *SALK-036171* seedlings to be the same as in WT seedlings (Fig. 2*D*).

In transgenic plants where the sequence downstream from +83 was deleted and the upstream portion fused to GUS directly, "*2.3kb Pro DME::GUS,*" GUS expression was absent from the central and vegetative cells (Fig. 2 *A* and *B*; central cell nucleus within ovule indicated with arrow). However, strong ectopic GUS

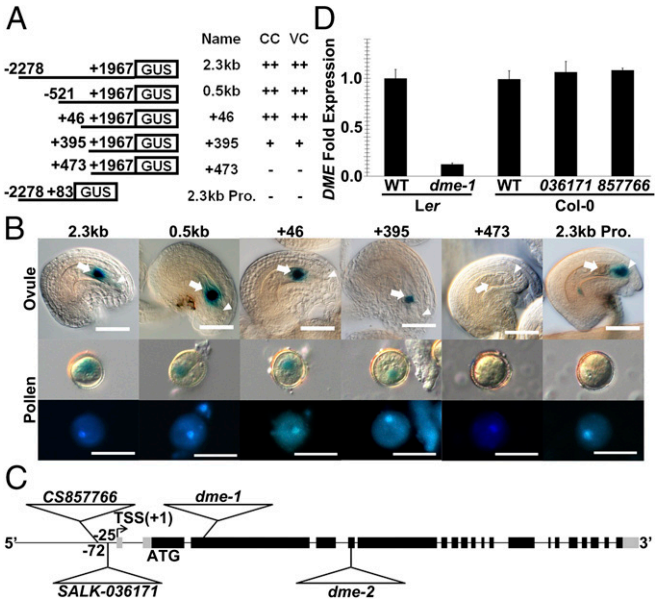

Fig. 2. Diagram of the *DME::GUS* reporter constructs and expression of the T-DNA insertion lines in the *DME* region. (*A*) The name, staining intensity, and the coordinates for each construct are shown. CC, central cells; VC, vegetative cell of pollen; –, none; +, moderate; ++, strong. (*B*) GUS staining is shown in ovules and pollen. DAPI-stained pollen grains are shown in the *Bottom* row. Plants expressing transgenes 2.3kb to +395 displayed GUS expression in the central cell nucleus (arrow) and vegetative cell nucleus. No GUS expression was detected in +473 transgenic plants, and 2.3kb Pro. plants exhibited GUS expression only in the synergid cells (arrowhead). (Scale bars: ovule, 50 μm; pollen, 20 μm.) (*C*) *dme* T-DNA insertion alleles at 72 nt upstream (*CS857766*) and at 25 nt upstream (*SALK-036171*) of the TSS. Black box, translated exon; gray box, untranslated exon; first line, 5′ flanking sequences; other lines, intron. (*D*) qRT-PCR analysis of *DME* expression in homozygous *dme* mutant seedlings after normalization with *ACT1*, *ACT3*, and *ACT12* expression. Values are plotted relative to the expression of *DME* in L*er* WT, which was set at 1.0, and represent the average of triplicate measurements ± SD.

activity was observed in the synergid cells of mature female gametophytes in plants expressing this transgene (Fig. 2 *A* and *B*, arrowhead, and Table S4). Thus, a putative suppressor element that usually represses *DME* expression in synergid cells is present downstream of +83 bp. The lack of a nuclear localization signal (NLS) in this construct resulted in staining of the synergid cells' cytoplasm.

Expressing DME Polypeptide in the Central Cell with a Minimal Reproductive Promoter Rescues Seed Abortion and Aberrant DNA Methylation Associated with the *dme-2* Mutation. The +46 *pDME::GUS/GFP* transgene has the shortest sequence that correctly regulates reporter expression in the central cell and vegetative cells, without deleting internal DME coding sequences (Fig. 2 *A* and *B* and Fig. S3*A*). We therefore considered this transgene to contain the minimal reproductive promoter that could be used to drive the correct reproductive expression of a full-length DME polypeptide in a functional assay. We then constructed a +46 *pDME::cDME* transgene (Fig. S1*B*) to determine the functional significance of *DME* expression driven by this minimal reproductive promoter. We transformed *dme-2* heterozygotes with the +46 *pDME::cDME* transgene (Fig. S3*B*). The *dme-2* mutation is a loss-of-function null allele, and, in self-pollinated *dme-2* heterozygous mutant plants, 50% of the F1 progeny seed inherit the maternal *dme-2* mutant allele and abort their development whereas inheritance of the paternal mutant *dme-2* allele has no effect on seed viability (2). To test for +46 *pDME::cDME* transgene function, we analyzed whether it could rescue seed abortion in transgenic lines. In self-pollinated plants that were hemizygous for a single transgene locus, and heterozygous for *dme-2*, 25% of the F1 seed inherited the mutant maternal *dme-2* allele and aborted their development, and 25% inherited both the mutant maternal *dme-2* allele and the transgene. Therefore, full complementation of the mutant maternal *dme-2* allele by the +46 *pDME::cDME* transgene results in 25% seed abortion (2), which we observed (Fig. S3*B* and Table S2). Moreover, self pollination of plants heterozygous for *dme-2* and hemizygous for +46 *pDME::cDME* generated plants homozygous for both the *dme-2* mutation and the +46 *pDME::cDME* transgene, which displayed the same low seed abortion rate (<1%) as both WT plants and homozygous *dme-2* plants expressing the homozygous 2.3kb *pDME::cDME* control transgene (Fig. S3*B* and Table S2), demonstrating the functional activity of the minimal reproductive promoter.

Seed abortion resulting from the *dme-2* mutation is caused, at least in part, by the resultant aberrant expression pattern of imprinted components of the PRC2 in endosperm (2, 4, 15–17). In the absence of DME, PRC2 is defective, and endosperm development is severely compromised, resulting in embryo abortion (18). Because seed abortion is rescued by the +46 *pDME::cDME* transgene, we hypothesized that *DME* expression driven by the minimal reproductive promoter is able to demethylate the central cell genome-wide, including specific PRC2 genes, resulting in a functional endosperm with a distinctive pattern of maternal endosperm genome hypomethylation compared with the paternal endosperm genome. To test this hypothesis, we pollinated *dme-2/dme-2* homozygous Col(*gl*) (Columbia ecotype, homozygous for the *glabrous* mutation) plants that were also homozygous for the +46 *pDME::cDME* transgene, with WT L*er* (Landsberg ecotype homozygous for the *erecta* mutation) pollen. F1 seeds were harvested at 9 days after pollination, endosperm was obtained by manual seed dissection, and genomic DNA was isolated. Maternal and paternal genomes were distinguished by Col versus L*er* single nucleotide polymorphisms, and DNA methylation profiles were obtained by next generation bisulphite sequencing of DNA (3). We analyzed the methylome of F1 endosperm from *dme-2/dme-2* homozygotes that were homozygous for the +46 *pDME::cDME* transgene (*dme-2*; +46 cDME) and compared it with a WT control (Col-0 crossed to L*er*), and with the methylome of seeds inheriting the *dme-2* mutation maternally (3). We found that the maternal allele of F1 *dme-2*; +46 cDME endosperm is normally methylated at maternally (e.g., *FIS2, FWA*) and paternally (e.g., *YUK10, PHE1*) expressed imprinted gene loci and resembles the WT maternal allele whereas these loci are hypermethylated in *dme-2* (Fig. 3*A*). Genome-wide, the hypermethylation phenotype seen in *dme-2* maternal endosperm, demonstrated by the increased density of genomic sites with a fractional methylation level between 0.5 and 1 (Fig. 3*B*, *dme-2* minus WT kernel density trace), is fully complemented in *dme-2*; +46 cDME endosperm and resembles the WT endosperm methylome (Fig. 3*B*, *dme-2*; +46 cDME minus WT trace, and Fig. S4) whereas the paternal allele is unaffected (Fig. 3*C*). Thus, the minimal reproductive promoter promotes functional *DME* expression required for DNA demethylation.

A 357-bp Region of the DME Transcriptional Unit Is both Necessary and Sufficient to Generate the Appropriate *DME* Expression Profile During Female Gametophyte Development. To identify where the precise regulatory elements that control DME expression in the central cell are located, we carried out further deletions within the 2-kb region that we had so far identified to be necessary and sufficient for fully functional DME activity. This gain-of-function (GOF) construct series is denoted "Truncated 5′-UTR" (TU) (Fig. 4*A* and Fig. S5), for which we used increasingly smaller portions of the 748-bp-long –90 to +658 region around the DME TSS to drive GUS expression. The *TU0* reporter construct, containing the full –90 to +658 region, showed the same expression pattern and intensity as the reference *2.3pDME::GUS* construct, except for GUS

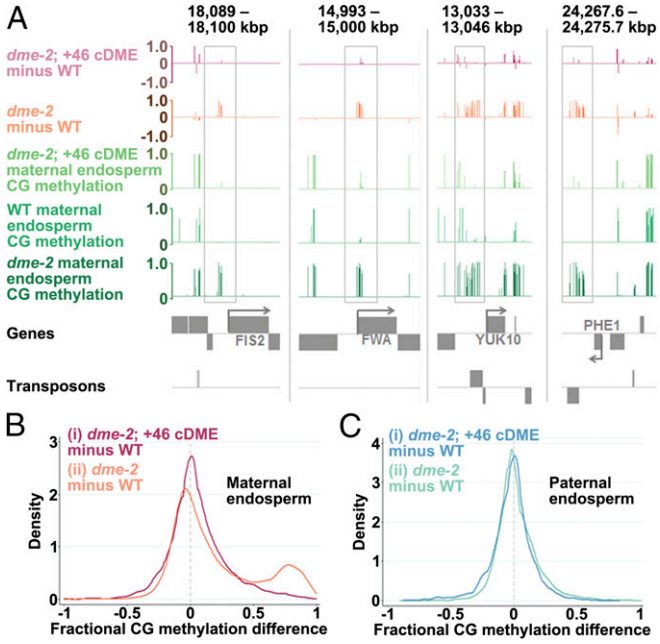

Fig. 3. *DME* expression driven by the +46 transgene can correct the methylation phenotype of homozygous *dme-2* mutant endosperm. (*A*) Snapshots of CG DNA methylation at selected imprinted loci. Each track represents a different genotype: crimson trace, WT subtracted from *dme-2* homozygous endosperm expressing the +46 transgene; orange trace, WT subtracted from *dme-2* heterozygous endosperm; green tracks are raw CG methylation data in the three genotypes compared. Differential methylation at both maternally expressed (*FIS2, FWA*) and paternally expressed (*YUK10, PHE1*) imprinted loci (i.e., maternal hypomethylation of imprinting control regions) is regained in *dme-2* homozygous endosperm when the +46 transgene is expressed. Gray boxes show the imprinting control regions at each locus, and arrows show the direction of gene transcription. (*B*) Kernel density plots of CG methylation differences between the maternal alleles of *dme-2* homozygous endosperm expressing the +46 transgene and WT (i, crimson trace) and *dme-2* heterozygous endosperm and WT (ii, orange trace,). Hypermethylation of the *dme-2* mutant endosperm is evident in the increased density at a fractional methylation difference of between 0.5 and 1 in ii and is corrected by the +46 transgene as seen by the loss of this density increase in i. (*C*) Kernel density plots of CG methylation differences between the paternal alleles of *dme-2* homozygous endosperm expressing the +46 transgene and WT (i, blue trace) and *dme-2* heterozygous endosperm and WT (ii, aquamarine trace). Methylation of the paternal (WT L*er*) alleles is the same in each genotype, showing that the +46 transgene does not affect methylation postfertilization.

expression in the cytoplasm of cells expressing GUS because the endogenous nuclear localization sequence of DME is downstream of 658 bp, and therefore absent from all TU constructs (Fig. 4 and Figs. S5 and S6).

From our GOF TU series, the minimal sequence that we found to be necessary and sufficient to drive *DME* expression in the central cell was 357 bp in length, from +202/+559 (transgene TU34) (Fig. 4). *TU23* (+46/+415) plants did not show any GUS expression, but *TU34* plants displayed GUS activity in the central cell (Fig. 4). Because our previous deletion to 473 bp downstream of the TSS led to the complete loss of DME::GUS expression (Fig. 2*A*), we deduced that the central cell regulatory region lies in a 57-bp fragment between the +416 and +472 positions. We also observed reduced GUS expression in the central cell in *TU45* (+363/+658), indicating that quantitative regulation of central cell expression also involves a region between +202 and +362, denoted the quantitative regulatory element (QE) (Fig. 5*A*).

***DME* Expression in Sporophytic Tissues Is Regulated by Distinct DNA Sequences.** *DME* is expressed in the sporophyte shoot apical meristem (SAM), leaf primordia, and the root apical meristem (Fig. S1*A* and Figs. S5–S8) and is required for floral and vegetative developmental patterning (2, 14). To determine the relationship between the regulation of *DME* in reproductive and sporophytic tissues, we further investigated the regulatory regions of *DME* to elucidate those required for sporophytic *DME* expression. We identified a 349-bp region, from −90 to +259 that is necessary and sufficient for *DME* expression in sporophytic tissues (TU12, Figs. S5 and S6). Next, we generated constructs to narrow this region, identifying 13 bp close to the TSS, between +7 and +19, required for the sporophytic expression of DME, which we designated as a necessary sporophytic enhancer (SPE) (Fig. 5*A* and Figs. S7 and S8). Deletion specifically of the SPE (TU0_*ΔSP*) results in loss of sporophytic, but not reproductive, DME expression (Fig. 5 *B* and *C*).

Overlapping 15- and 47-Base Pair Regions Are Necessary for DME Expression in the Central and Vegetative Cells, Respectively. As stated previously, a 57-bp element necessary for central cell DME expression lies between the +416 and +472 positions. To establish whether this sequence was also sufficient to drive *DME* expression, we generated constructs containing one to four copies of this 57-bp fragment with and without the minimal *CaMV* 35S promoter downstream, but none of these constructs exhibited any GUS expression in any tissue; therefore, we were unable to conclude that this sequence is sufficient for expression (Fig. S9). Nevertheless, to investigate this region further, we generated fine-deletion constructs *TU0_ΔPOL* (Δ+416/+462), *TU0_ΔCC1* (Δ+416/+431), *TU0_ΔCC2* (Δ+432/+447), and *TU0_ΔCC3* (Δ+448/+462) (Fig. 5*B*) to establish the sequence necessary for regulating central cell expression. GUS activity was detected in the central cell in *TU0_ ΔCC1* and *TU0_ΔCC2,* but not in *TU0_ΔCC3* or *TU0_ΔPOL* plants (Fig. 5 *B* and *C*); therefore, the sequence necessary for central cell expression, denoted the "CCE," is ∼15 bp in length and is located between +448 and +462 nt (Fig. 5*A*). Vegetative cell *DME* expression is present in TU0_*ΔSP* but disappears in *TU0_ΔPOL* and in each of *TU0_ΔCC1, TU0_ΔCC2,* and *TU0_ΔCC3* (VC in Fig. 5 *B* and *C*), demonstrating that vegetative cell expression of *DME* specifically requires the 47-bp +416/+462 sequence, denoted the "VCE," which encompasses, but is broader than, the +448 /+462 CCE (Fig. 5*A*).

The 15-bp CCE Sequence, Shared by the VCE, Is Required for DME Expression and Is Predicted to Bind Several Key Transcription Factors. DME expression in the vegetative and central cells is thought to have a common function, in regulation of transposon silencing in the germline. As such, the 15-bp common region of the VCE and CCE elements is of particular intrigue. This sequence contains the 9-bp "CATTTATTG" motif, which is

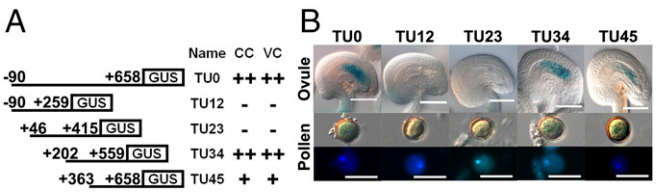

Fig. 4. Diagram of the *DME::GUS* reporter constructs for fine mapping of cis-elements and their expression patterns. The TU (truncated 5′-UTR) series of constructs. (*A*) The name, staining intensity, and coordinates for each construct are shown. CC, central cells; VC, vegetative cell of pollen; –, none; +, moderate; ++, strong. (*B*) GUS staining is shown in ovules and pollen. DAPI-stained pollen grains are shown in the *Bottom* row. TU0, TU34, and TU45 transgenic plants exhibited GUS expression in the central cell and pollen. No GUS expression was detected in TU12 and TU23 plants. (Scale bars: 50 μm.)

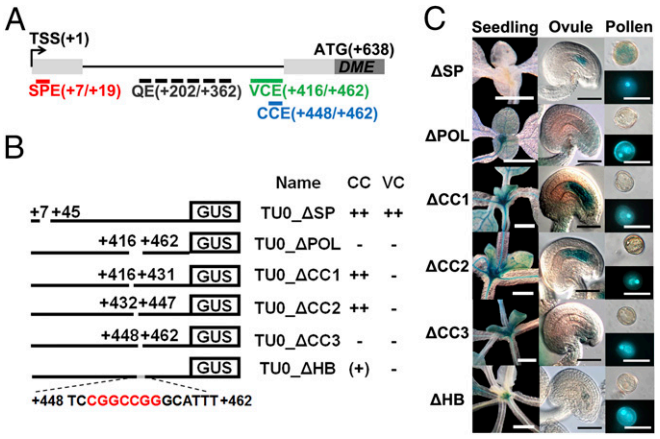

Fig. 5. Internal deletion/substitution of cis-elements. (*A*) Summary of *DME* cis-regulatory elements. Dark gray box, translated exon; light gray box, 5′-UTR; line, first intron; red line, sporophytic element (SPE); blue line, central cell element (CCE); green line, pollen vegetative cell element (VCE) ; dotted line, quantitative regulatory element (QE). (*B*) Diagram of *DME::GUS* internal deletion and substitution constructs of the cis-elements. CC, central cells; VC, vegetative cell of pollen; –, none; (+), weak; ++, strong; Δ, deletions or substitutions. (*C*) GUS staining is shown in ovules and pollen. DAPI-stained pollen grains are shown in the bottom of each pollen. TU0_ΔSP, same GUS expression pattern as TU0; TU0_ΔPOL, central cell and pollen GUS disappeared; TU0_ΔCC1 and TU0_ΔCC2, only the pollen expression disappeared; TU0_ΔCC3, central cell and pollen GUS disappeared. TU0_ΔHB, central cell GUS was significantly reduced and pollen GUS disappeared. (Scale bars: ovule, 50 μm; pollen, 20 μm.)

strikingly similar to the pseudopalindromic targets of the *Arabidopsis* Homeobox HD-ZIP family of plant-specific transcription factors: for example, the recognition sequence "CAAT(T/A)ATTG" of subfamily 1 (19, 20). To examine the role of this AT-rich sequence in the expression of *DME*, 7 bp of an AT-rich sequence in *TU0* was changed to a GC-rich sequence (Fig. 5*B*). This change resulted in a significant reduction of GUS activity in the central cell of *TU0_ΔHB* plants and the complete absence of GUS expression in the vegetative cell of pollen (Fig. 5*C*). Thus, the pseudopalindromic sequence is required for normal central and vegetative cell *DME* expression.

Our identification of precise coordinates for key regulatory elements of DME expression enabled us to carry out preliminary investigations to reveal potential interacting transcription factors. A recent genome-wide analysis to characterize regulatory elements and transcription factor binding sites used a novel high throughput DNA affinity purification sequencing assay (DAP-seq), generating a "cistrome" map for 30% of transcription factors in *Arabidopsis* (21). By correlating our VCE and CCE coordinates with this cistrome dataset, we were able to identify 40 potential candidates that bind these regions in vitro and may therefore be involved in DME regulation in reproductive tissues (Table S3). Among these candidates are 10 HD-ZIP transcription factors, spanning the four subfamilies, which is consistent with our finding functional targets of the HD-ZIP family in the common region of the VCE and CCE elements.

Discussion

Here, we show that the regulation of DME expression is mirrored in both male and female gametophytes, developing simultaneously upon germline differentiation in distinct reproductive organs. *DME* expression is restricted to the vegetative cell nucleus after the first asymmetric mitosis, at the late bicellular stage of pollen development (Fig. 1), which is concurrent with separation of the generative and vegetative cell lineages, so that the demethylation activity of DME is restricted to the vegetative cell whereas the sperm genome remains highly methylated at DME targets. This expression profile is likewise reflected in the female gametophyte. During female gametogenesis, the third mitotic division is followed immediately by cellularization and differentiation, generating antipodal cells at the chalazal pole, and the egg cell, synergids, and two polar nuclei at the micropylar pole (22). It is immediately after this differentiation step that *DME* expression is activated so that expression is confined primarily to the polar nuclei, which fuse to form the central cell, and is absent from the egg (2).

Expression of *DME* in companion cells, and the evasion of *DME* expression in gametes, is key for understanding the function of DNA demethylation during plant reproduction. This pattern explains how the maternal endosperm genome is hypomethylated compared with the paternal endosperm genome. Maternally hypomethylated loci are either directly or indirectly (via PRC2 activity) responsible for parent-of-origin gene expression: i.e., gene imprinting, in the endosperm (5, 13). The fact that *DME* is not expressed in the egg or sperm cells is responsible, at least in part, for the similarity of the maternal and paternal embryo methylomes (3) and, therefore, the fact that genes displaying parent-of-origin expression in endosperm do not do so in the embryo (5, 23). Maternal genome hypomethylation is required for seed development, but the demethylation of the vegetative cell does not directly affect seed viability. Instead, demethylation of both the central and vegetative cells at DME targets, such as small, AT-rich, and nucleosome-depleted euchromatic TEs, likely promotes expression of TEs in these cells. Demethylated companion cells do not pass on their genome to the next generation; therefore, the genomic instability resulting from transposon transcription is not deleterious to the species as a whole. Rather, there is evidence to suggest that the RdDM pathway then promotes corresponding TE methylation in the egg and sperm cells, respectively (3, 11, 12). In this way, the companion cell acts sacrificially, reinforcing and protecting the genomic integrity of egg and sperm, which will be inherited by the next generation. The function of *DME* expression in companion cells provides support for the unique importance of double fertilization involving companion cells during evolution.

We explored the regulatory sequences that contribute to this remarkable expression profile by producing a comprehensive array of iteratively deleted reporter transgenes for the regions upstream of the *DME* translational start site. With the exception of a negative regulatory region that suppresses *DME* expression in female gametophyte synergid cells, all other regulatory elements reduced *DME* expression when lost or mutated (Figs. 2 and 4), suggesting that the majority of transcriptional regulation of *DME* is positive. The lack of DME expression in the fertilized endosperm, needed to preserve regions of DNA demethylation that are specific to the maternal endosperm genome (3), is therefore likely caused by a decrease in activity of a positive regulator or regulators, rather than the appearance of a negative regulator.

Using our deletion transgenes, we found that sequences regulating *DME* expression were contained within its transcriptional unit. We designated the +46 transgene, which consists of 592 bp of sequence before the translational start site, as the minimal reproductive promoter and used this +46 minimal promoter in a functional construct to drive expression of *DME* cDNA. The expression of this transgene rescued both the seed abortion and genome-wide DNA methylation phenotypes of *dme-2* heterozygous and homozygous mutants, showing that the expression timing, level, and tissue specificity of *DME* expression in reproductive tissues is recapitulated with a promoter sequence of 592 bp contained within the transcriptional unit. Within this sequence, we identified a 47-bp VCE, overlapping with a 15-bp CCE, necessary for regulation of the vegetative and central cell DME expression patterns, respectively. The CCE and VCE are distinct from the 13-bp SPE close to the TSS that promotes DME expression in the sporophyte. Each of the three enhancer sequences is conserved in closely related *Brassicaea* family members, such as *Arabidopsis*

lyrata, *Capsella rubella*, and *Brassica rapa* (Fig. S10), but they are missing from the *DME* homologs *ROS1, DML2,* and *DML3*, which are expressed much more broadly than *DME* (24) and do not contribute to demethylation in the central cell.

Because the CCE is contained entirely within the VCE, it is possible that the control of *DME* expression in each of the companion cells of the gametes shares a common regulatory pathway. The overlapping VCE/CCE sequence of 15 bp (+448/ +462) is AT-rich, including 9 bp with striking similarity to the pseudopalindromic "CAAT(T/A)ATTG" sequence, which is a target of the HD-ZIP plant-specific homeobox transcription factor family (25–27). Substitution of this motif led to a large reduction in central cell *DME* expression and ablation of vegetative cell *DME* expression, showing that this pseudopalindromic sequence is required for correct DME regulation.

Using the coordinates that we derived for the VCE and CCE and our analyses of the recently published DAP-seq "cistrome" collection (18), we were able to catalog a list of 40 potential transcription factors, including 10 HD-ZIPs, that bind to these elements in vitro (21). MADS-box transcription factor AGL80 is required for *DME* expression in the central cell (28) so it is likely that MADS-box binding domains are present in this regulatory region, and several MADS-box transcription factors were found to bind to the VCE by DAP-seq (Table S3) (21); however, AGL80 was not specifically tested in the DAP-seq screen.

If there is a transcription factor that coregulates *DME* in the central and vegetative cells, it would be expressed in both these tissues, at a similar time to DME itself. However, as we show here, DME expression in the male gametophyte is confined to a short period of the bicellular pollen stage and is often not detected in pollen expression datasets (29–31). Thus, to identify potential transcription factors that may bind the shared sequence of the VCE and CCE and regulate DME, precise establishment of their endogenous expression profile using reporter genes, and their effect on DME expression when mutated or ectopically expressed, will be required in the future.

In summary, we show here that *DME* expression during reproduction is confined to a narrow window of time, and to single companion cells, in female and male gametophytes, which is necessary for its role in seed viability, gene imprinting, and transgenerational transposon silencing. We delineate specific, conserved enhancer sequences required for the precise expression pattern of *DME* and identify candidate transcription factors by their in vitro binding patterns at the VCE and CCE, information that will be valuable in the future to delineate the regulatory pathways that control DME expression.

Materials and Methods

Please see *SI Materials and Methods* full details of methods.

Plant Materials and Growing Conditions. All of the promoter constructs used in this study were transformed into *Arabidopsis* Columbia *gl*. The *dme-1* homozygous mutant allele is in Landsberg *er* (L*er*) background (2). Heterozygous *dme-2* in Col(*gl*) was used for the complementation test.

Next Generation Bisulphite Sequencing. F1 endosperm was hand-microdissected at 8 to 10 days after pollination of homozygous *dme-2*; +46 cDME plants, or WT Col-0, with L*er* pollen, allowing distinction of parental alleles. Bisulphite sequencing libraries were prepared as described previously (3), and in *SI Materials and Methods*.

Recombinant Plasmid Construction. Methods for generating deletion constructs and *TU_GUS* (Table S4) are described in *SI Materials and Methods*. Primers are listed in Table S5.

ACKNOWLEDGMENTS. We thank Jihyun Yu (Seoul National University) for thoughts about the promoter assays and for providing vectors. This study used the Vincent J. Coates Genomics Laboratory at the University of California, Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303. This work was supported by Next-Generation BioGreen 21 Program Grant PJ01101802 (to Y.C.), NRF of Korea Grant 2014R1A2A2A01004887 (to Y.C.), and NIH Grant GM69415 (to R.L.F.). J.-S.P. was supported by the Stadelmann-Lee Scholarship Fund, Seoul National University.

1. Park K, et al. (2016) DNA demethylation is initiated in the central cells of Arabidopsis and rice. *Proc Natl Acad Sci USA* 113(52):15138–15143.
2. Choi Y, et al. (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell* 110(1):33–42.
3. Ibarra CA, et al. (2012) Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* 337(6100):1360–1364.
4. Gehring M, et al. (2006) DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* 124(3):495–506.
5. Hsieh TF, et al. (2011) Regulation of imprinted gene expression in Arabidopsis endosperm. *Proc Natl Acad Sci USA* 108(5):1755–1762.
6. Schoft VK, et al. (2011) Function of the DEMETER DNA glycosylase in the Arabidopsis thaliana male gametophyte. *Proc Natl Acad Sci USA* 108(19):8042–8047.
7. Xiao W, et al. (2003) Imprinting of the MEA Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. *Dev Cell* 5(6):891–901.
8. Morales-Ruiz T, et al. (2006) DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc Natl Acad Sci USA* 103(18):6853–6858.
9. Ortega-Galisteo AP, Morales-Ruiz T, Ariza RR, Roldán-Arjona T (2008) Arabidopsis DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. *Plant Mol Biol* 67(6):671–681.
10. Gong Z, et al. (2002) ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell* 111(6):803–814.
11. Slotkin RK, et al. (2009) Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* 136(3):461–472.
12. Martínez G, Panda K, Köhler C, Slotkin RK (2016) Silencing in sperm cells is directed by RNA movement from the surrounding nurse cell. *Nat Plants* 2:16030.
13. Hsieh TF, et al. (2009) Genome-wide demethylation of Arabidopsis endosperm. *Science* 324(5933):1451–1454.
14. Kim M, et al. (2008) Temporal and spatial downregulation of Arabidopsis MET1 activity results in global DNA hypomethylation and developmental defects. *Mol Cells* 26(6):611–615.
15. Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB (1998) Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. *Science* 280(5362):446–450.
16. Köhler C, et al. (2003) Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *EMBO J* 22(18):4804–4814.
17. Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudhury A (2000) Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds. *Proc Natl Acad Sci USA* 97(19):10637–10642.
18. Hehenberger E, Kradolfer D, Köhler C (2012) Endosperm cellularization defines an important developmental transition for embryo development. *Development* 139(11):2031–2039.
19. Charité J, et al. (1998) Transducing positional information to the Hox genes: Critical interaction of cdx gene products with position-sensitive regulatory elements. *Development* 125(22):4349–4358.
20. Sessa G, Morelli G, Ruberti I (1993) The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. *EMBO J* 12(9):3507–3517.
21. O'Malley RC, et al. (2016) Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* 165(5):1280–1292.
22. Drews GN, Koltunow AM (2011) The female gametophyte. *Arabidopsis Book* 9:e0155.
23. Gehring M, Missirian V, Henikoff S (2011) Genomic analysis of parent-of-origin allelic expression in Arabidopsis thaliana seeds. *PLoS One* 6(8):e23687.
24. Penterman J, et al. (2007) DNA demethylation in the Arabidopsis genome. *Proc Natl Acad Sci USA* 104(16):6752–6757.
25. Chan RL, Gago GM, Palena CM, Gonzalez DH (1998) Homeoboxes in plant development. *Biochim Biophys Acta* 1442(1):1–19.
26. Palena CM, Gonzalez DH, Chan RL (1999) A monomer-dimer equilibrium modulates the interaction of the sunflower homeodomain leucine-zipper protein Hahb-4 with DNA. *Biochem J* 341(Pt 1):81–87.
27. Palena CM, Tron AE, Bertoncini CW, Gonzalez DH, Chan RL (2001) Positively charged residues at the N-terminal arm of the homeodomain are required for efficient DNA binding by homeodomain-leucine zipper proteins. *J Mol Biol* 308(1):39–47.
28. Portereiko MF, et al. (2006) AGL80 is required for central cell and endosperm development in Arabidopsis. *Plant Cell* 18(8):1862–1872.
29. Borges F, et al. (2008) Comparative transcriptomics of Arabidopsis sperm cells. *Plant Physiol* 148(2):1168–1181.
30. Qin Y, et al. (2009) Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. *PLoS Genet* 5(8):e1000621.
31. Boavida LC, Borges F, Becker JD, Feijó JA (2011) Whole genome analysis of gene expression reveals coordinated activation of signaling and metabolic pathways during pollen-pistil interactions in Arabidopsis. *Plant Physiol* 155(4):2066–2080.
32. Hyun Y, et al. (2013) The catalytic subunit of Arabidopsis DNA polymerase α ensures stable maintenance of histone modification. *Development* 140(1):156–166.
33. Yadegari R, et al. (2000) Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms. *Plant Cell* 12(12):2367–2382.