Supplemental Data

DEMETER DNA Glycosylase Establishes *MEDEA* Polycomb Gene Self-Imprinting by Allele-Specific Demethylation

Mary Gehring, Jin Hoe Huh, Tzung-Fu Hsieh, Jon Penterman, Yeonhee Choi, John J. Harada, Robert B. Goldberg, and Robert L. Fischer

Supplemental Experimental Procedures

Primers for Bisulfite Sequencing

Primers for the – 4 kb region were MEA3904 (5'-AACTTTATTCATRTAATRRTCRAACACT-3') or MEA3979 and MEA4510. The –3 kb region was amplified with MEA5187BFc (5'-CAAAATACTCTATTCTACATTCCCATCTAT-3') and MEA5810BRc (5-TAAATAAATTAAATGAGTTTGAGTATAAAATG-3'), followed by a nested amplification with MEA5212 and MEA5810BRc. The –500 bp region was amplified with MEA7671 (5'-TAACCATTAAACATTAATTTAAATCTT-3') or MEA7529 and MEA7935. MEA-ISR was amplified from L*er* and Col-*gl* backgrounds using JP1026 and JP1027 (Cao and Jacobsen, 2002). A large deletion and extensive polymorphisms prevented the use of these primers in RLD. Instead, the first repeat was amplified with RLDBi (5'-TAATTTAAAATAATGGTGATGTTGTTAGTTTG-3') and RLDBi4 (5'-AAAAARRTTTTATAAATATTAAATTAATATRA-3'). For *MEA* coding region bisulfite sequencing, Col-*gl* rosette leaf DNA was bisulfite treated as previously (Xiao et al., 2003) and methylation on the bottom strand determined. We sequenced 7 clones from MEA8355F (5'-TTTCACTCCAAACATATATAAATTAAC-3') to MEA8755R (5'-GAYTAATGTATAAYTGTTTATTAGATGTAT-3'), 5 clones each from MEA8646F (5'-CTCTTCTRTATRTTTTTCTRAAAATTAARRA-3') to MEA9066R (5'-TGYATYAATYTTGGYTTTTTTGGYTGAATG-3') and from MEA9294F (5'-CACTTTTRTCRARAATRCAAAACCCACTT-3') to MEA9801R (5'-TAATGYAAAAAYTAAYYATATAAATYGGTY-3'), 8 clones from MEA9810F (5'-CTTRATTATTAATTTRTARTCCATATTTAATAAACTR-3') to MEA10221R (5'-GTGGYTAAATTAAAAAAGAAAGATTYAAAGTTAYYATG-3'), 10 clones from MEA10310F (5'-CCCRARTCTARATCCRTAARCATTAAATC-3') to MEA10650R (5'-GGATYTGAGAYYAYAATYTTGTTTGATATAGAG-3'), 8 clones each from MEA10528F (5'-CTATTCCTTAATTACRTTTATTARTTACTRRT-3') to MEA10905R (5'-GTTTTGTTAAGGTYTAATGAYATAGTAYATTG-3') and MEA10761F (5'-TACTTACACTRTATTCCTTRATTATRC-3') to MEA11285R (5'-TAYAAAYTYATGTTYAAATTAAATYTYATGG-3'), 6 clones from MEA11131F (5'-ATAARCACTACACACCATRCACTTRCAART-3) to MEA11460R (5'-CAAATTCTATAATCAAARTAATTCAAACC), 7 clones from MEA11571F (5'-

CATACAATTCCTCCTTCAAACCAATAA-3') to MEA11987R (5'-GATYATTYAAGGTAAAGAGGTAGGAAGAAYYAA-3'), 8 clones each from MEA11906F (5'-CTRATCACTCATRATRAARCTAATRARCRT-3') to MEA12300R (5'-GAGTTTGAGTTTYTTGGAATATYTTYAATATG-3') and MEA12234F (5'-TCRTRTATCAACTTTACTCRTCRTTRATTRR-3') to MEA12647R (5'-GTTTTGGTTTAGTAAYAYAAAATAGYATTA-3'), and 9 clones from MEA12740F (5'-CAATRTTTATRTTRTTARTTTRCATARACC-3') to MEA13093R (5'-GTTTAGATAYTAAATGTTAGATGYATYAAAT-3'). This covers 91 of the 99 CG sites present from the *MEA* transcription start site to the beginning of the 3' repeats.

Amplification and Cloning of the *MEA* Allele in *dme-2* Mutant Endosperm

The –500bp region and MEA-ISR were amplified with *Pfu Turbo* DNA polymerase (Stratagene) from the same *dme-2* DNA used for the experiment in Figure 2. PCR products were cloned into the pCR-Blunt II-TOPO vector (Invitrogen) and sequenced. The primers for amplifying the –500 bp region were MEA8323Xba (5'-ATATTCTAGACTTTTTTTCTCGTCTTCTCTGATGTTGGT-3') and UCB3SR12R-sac1 (5'-GGGAGCTCGTTAAGCCTGTGGTTGACAAC-3'). The primers for amplifying the MEA-ISR were B5-7RR (5'- TTAGGTATTAGCTCGTTTGGTTTTA-3') and MEA 3 REP (5'-CTTAAAAGATTTTCAACTCATTTTTTTTAAAAGG-3').

Cloning, Expression, and Purification of DME in *E. coli*

A full-length *DME* cDNA (Choi et al., 2002) was used as template in a PCR reaction with oligonucleotides JH021 (5'-TTAATCTAGAATGCAGAGCATTATGGACTCG-3') and JH017 (5'-CGGTCGACTTAGGTTTTGTTGTTCTTCAATTTGC-3'), which add *Xba*I and *Sal*I restriction sites (underlined), respectively. The 5.2 kb PCR product was digested with *Xba*I and *Sal*I and cloned into the pMAL-c2x vector (NEB) to create c2x-DME. To generate a N-terminal 537 amino acid deletion, c2x-DME was digested with *Xba*I and *Bsu*36I. The 3' overhangs were filled in with T4 DNA polymerase and self-ligated, creating the c2x-DMEΔN537 clone. The construct with an D1304N point mutation was generated using the full-length *DME*(D1304N) cDNA clone (Choi et al., 2004), following the same procedure as above. This fuses DME in frame downstream of maltose-binding protein (MBP). The c2x-DMEΔN537 or c2x-DMEΔN537(D1304N) clones were transformed into *E. coli* Rosetta cells (Novagen). Transformed cells were grown at 28°C in LB supplemented with 0.2% glucose, 100 μg/mL of ampicillin, and 50 μg/mL of chloramphenicol until the OD_{600} reached 0.4. Protein expression was induced with 10 μM of IPTG at 18°C for 1 hr. The culture was centrifuged at 6,500 rpm for 15 min at 4°C and the pellet was resuspended in 30 mL of 4°C column buffer (20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1mM EDTA). Cells were sonicated for 2 min on ice (output power 4; duty cycle 50%; Branson Sonifer 250). The lysate was centrifuged at 9,000 rpm for 25 min at 4°C and the supernatant was collected and subjected to gravity column purification. The MBP-DMEΔN537 and MBP-DMEΔN537(D1304N) fusion proteins were purified following the manufacturer's protocol through amylose resin (New England Biolabs). Eluted protein was dialyzed in the Slide-A-Lyzer dialysis cassette (10,000 MWCO; Pierce) against 50% glycerol at 4 °C overnight. Protein concentration was determined by the Bradford method using the Protein Assay kit (Bio-Rad Laboratories) and stored at –20 °C until use.

Substrate Preparation for DNA Glycosylase Activity Assays

Synthetic oligonucleotides were purchased either from Operon or Midland Certified. All oligonucleotides were 35-nucleotides in length with modifications denoted within parentheses as shown below:

MEA-1.6F, 5'-CTATACCTCCTCAACTCCGGTCACCGTCTCCGGCG
MEA-1.6F18meC, 5'-CTATACCTCCTCAACTC(5-meC)GGTCACCGTCTCCGGCG
MEA-1.6F17meC, 5'-CTATACCTCCTCAACT(5-meC)CGGTCACCGTCTCCGGCG
MEA-1.6F22meC, 5'-CTATACCTCCTCAACTCCGGT(5-meC)ACCGTCTCCGGCG
MEA-1.6F18AP, 5'-CTATACCTCCTCAACTC(abasic)GGTCACCGTCTCCGGCG
MEA-1.6F17AP, 5'-CTATACCTCCTCAACT(abasic)CGGTCACCGTCTCCGGCG
MEA-1.6F15AP, 5'-CTATACCTCCTCAA(abasic)TCCGGTCACCGTCTCCGGCG
MEA-1;6F12AP, 5'-CTATACCTCCT(abasic)AACTCCGGTCACCGTCTCCGGCG
MEA-1.6F18T, 5'-CTATACCTCCTCAACTCTGGTCACCGTCTCCGGCG
MEA-1.6R, 5'-CGCCGGAGACGGTGACCGGAGTTGAGGAGGTATAG
MEA-1.6R17meC, 5'-CGCCGGAGACGGTGAC(5-meC)GGAGTTGAGGAGGTATAG

Twenty pmol of oligonucleotide were end-labeled in a 50 μL reaction using 20 units of T4 polynucleotide kinase in the presence of 30 μCi of (γ-^{32}P)ATP (6000 Ci/mmol, Perkin Elmer Life Sciences) at 37^{o}C for 1 hr. The labeled oligonucleotide was purified using a Qiaquick Nucleotide Removal Kit (Qiagen) as described by the manufacturer.

Labeled oligonucleotides were annealed to the appropriate complementary oligonucleotides in 10 mM Tris-HCl (pH 8.0), 1mM EDTA and 0.1 M NaCl. The mixture was boiled in water for 10 min and then slowly cooled to room temperature overnight. *Msp*I or *Hpa*II restriction endonuclease digestion followed by gel electrophoresis was used to determine the efficiency of annealing. Only substrates that were greater than 90% double-stranded were used in glycosylase activity assays.

$NaBH_4$ Trapping Assays

5'-labeled oligonucleotide substrates (13.3 nM) were incubated with DME protein (250 nM) in a 15 μl reaction with 40 mM HEPES-KOH (pH 8.0), 0.1 M KCl, 0.1 mM EDTA, 0.5 mM dithiothreitol, and 200 μg/mL BSA at 37^{o}. After 1 hr of incubation, 1 M $NaBH_4$ was added to a final concentration of 100 mM and the reaction tubes were placed at 37^{o} for an additional 10 min. An equal volume of 2x SDS-PAGE loading buffer (90 mM Tris-HCl, pH 6.8, 20% glycerol, 2% SDS, 0.02% bromophenol blue, 100 mM dithiothreitol) was added to terminate the trapping reaction. Products were boiled for 10 min before loading onto a 10% SDS-PAGE gel. The wet gel was exposed to Kodak Biomax MS film for 12-18 h at –80^{o}.

Bacterial Cell Toxicity Assays

Bacterial strains AB1157 (F−*thr-1 ara-14 leuB6*(*Am*) *lacY1* (*gpt-proA2*)*62 tsx-33 supE44*(*Am*) *galK2 rac hisG4*(*Oc*) *rfbD1 mgl-51 rpsL31 kdgK51 xyl-5 mtl-1 argE3*(*Oc*) *thi-1*) and its isogeneic AP endonuclease mutant RPC501 (*xth nfo*) were kindly provided by R. P. Cunningham (Cunningham et al., 1986). Strains GM30 (F^{-} *thr-1 ara-14 leuB6 tonA31 lacY1 tsx-78 supE44 galK2 galT22 hisG4 rpsL136 xyl-5 mtl-1 thi-1*) and its isogenic *dcm-6* derivative, GM31, were kindly provided by Martin G. Marinus (Palmer and Marinus, 1994).

The c2x-DMEΔN537 and c2x-DMEΔN537(D1304N) plasmids were individually transformed into the strains above by electroporation and cells were grown on LB/Glu/Amp plates (LB supplemented with 0.2% glucose and 100 μg/mL of ampicillin) at 37^{o} overnight.

Fresh colonies were picked and resuspended in 5 mL of LB/Glu/Amp liquid medium. After 12-14 h incubation at 37^{o}, the culture was diluted 100,000-fold in LB medium and 100 μL was plated on the LB/Glu/Amp plates with 0, 2, 5, 10, 25, 50, and 100 μM of IPTG (isopropyl-β-D-thiogalactopyranoside; Sigma). The plates were incubated at 28^{o} for 20 to 28 hr and the number of colonies was counted.

Chromatin Immunoprecipitation (ChIP) Procedures

LNA nucleotide analogues (Promega) contain a 2'-*O*, 4'-*C* methylene bridge that locks the ribose moiety into a C3'-*endo* conformation (Koshkin et al., 1998; Obika et al., 1998; Singh et al., 1989). Region one (-4 to +440) of *MEA* was amplified with MEA-LNA006 (5'-CACCAACATCAGAGAAGACGAGAAAAG-3') and MEA-LNA004 (5'-GATTATGACTAATGTATAACTGTTTAC-3'). Region 2 (-947 to –547) of *MEA* was amplified with MEA-LNA002 (5'-GGGTCTCAATTTTGTGAACTGGTGTG-3') and MEA-LNA003 (5'-CCGATATTTTTTACTATTTATAACGTTAATTAC-3'). LNA nucleotides are underlined and are complementary to the RLD template sequence but have a mismatch with the L*er* template due to a polymorphism. To demonstrate the specificity of LNA-containing primers, approximately 50 pg of L*er* and RLD genomic DNA were used as a control. To increase the sensitivity of the LNA PCR reaction, 1 μCi of α-dATP-P^{32} was added to each PCR reaction. A polymorphism within region 1 (+60, T in RLD, C in L*er*) was used to check the parental origin of PCR products by sequencing. PCR products from region 1 from wild type (L*er* crossed to RLD) and *mea* (L*er mea/mea* crossed to RLD) were cloned into TOPO TA-cloning vector (Invitrogen, CA). 22 clones each were sequenced to determine the origin of amplification templates. Primer sequences and reaction conditions for *Actin* gene amplification were as described (Johnson et al., 2002).

Supplemental References

Cao, X., and Jacobsen, S. E. (2002). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc Natl Acad Sci USA *99*, 16491-16498.

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.

Cunningham, R. P., Saporito, S. M., Spitzer, S. G., and Weiss, B. (1986). Endonuclease IV (*nfo*) mutant of *Escherichia coli*. *168*, 1120-1127.

Johnson, L. M., Cao, X., and Jacobsen, S. E. (2002). Interplay between Two Epigenetic Marks: DNA Methylation and Histone H3 Lysine 9 Methylation. Curr Biol *12*, 1360-1367.

Koshkin, A. A., Singh, S. K., Nielsen, P., Rajwanshi, V. K., Kumar, R., Meldgaard, M., Olsen, C. E., and Wengel, J. (1998). LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. TETRAHEDRON *54*, 3607-3630.

Obika, S., Nanbu, D., Hari, Y., Andoh, J., Morio, K., Doi, T., and Imanishi, T. (1998). Stability and structural features of the duplexes containing nucleoside analogues with a fixed N-type conformation, 2'-O,4'-C-methyleneribonucleosides. Tetrahedron Lett *39*, 5401-5404.

Palmer, B. R., and Marinus, M. G. (1994). The *dam* and *dcm* strains of *Escherichia coli* - a review. Gene *143*, 1-12.

Singh, H., Clerc, R. G., and Lebowitz, J. H. (1989). Molecular cloning of sequence-specific DNA binding proteins using recognition site probes. Biotechniques *7*, 252-261.

Xiao, W., Gehring, M., Choi, Y., Margossian, L., Pu, H., Harada, J. J., Goldberg, R. B., Pennell, R. I., and Fischer, R. L. (2003). Imprinting of the MEA Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. Developmental Cell *5*, 891-901.

Table S1. Percent Methylation in Rosette Leaves

	-4 kb				-3 kb		-0.5 kb		MEA-ISR			
Accession	N[a]	% CG	% CNG	% CNN	N[a]	% CG	N[a]	% CG	N[a]	% CG	% CNG	% CNN
Col-*gl*	23	91	76	39	5	5	20	79	23	83	48	27
RLD	29	83	68	25	0		17	82	13	78	46	25
L*er*	5	96	80	30	0		15	0	13	68	33	22

a = number of clones sequenced.

Table S2. CG Methylation of *MEA* in the –500 bp and MEA-ISR Regions of Dissected Seeds

		–500 bp		MEA-ISR	
Cross[a]	Allele	N[b]	% CG	N[b]	% CG
RLD x Col-*gl*	Maternal Endosperm	24	24	10	23
	Maternal Embryo	19	77	9	100
	Paternal Endosperm	39	68	0	
	Paternal Embryo	20	85	12	83
L*er* x RLD	Maternal Endosperm	29	22	21	18
	Maternal Embryo	10	18	13	92
	Paternal Endosperm	0		8	100
	Paternal Embryo	4	60	10	97
RLD x L*er*	Maternal Endosperm	16	12	11	11
	Maternal Embryo	12	88	9	100
	Paternal Endosperm	20	8	0	
	Paternal Embryo	26	2	12	82

a = female parent is written on the left.
b = numbers of clones sequenced.

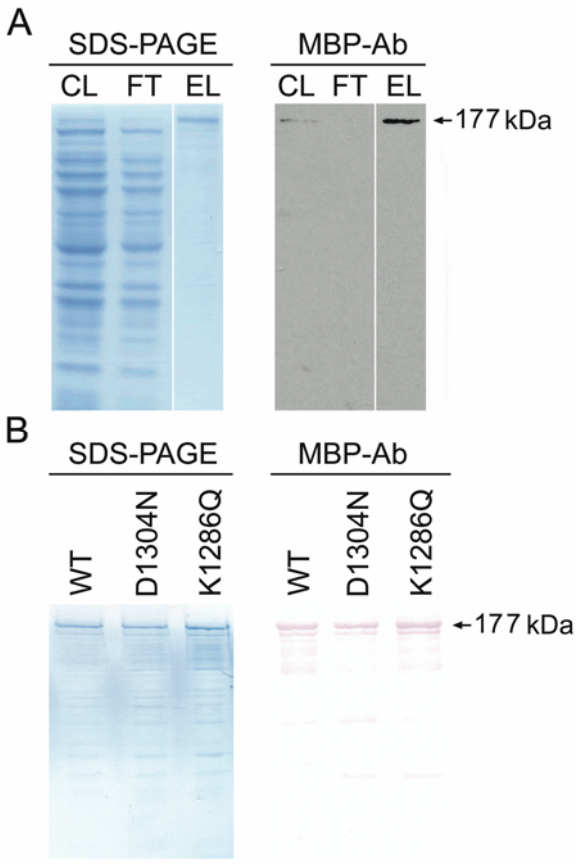

Figure S1. Purification of Wild-Type and Mutant DME Proteins

(A) Wild type (MBP-Δ537DME) DME was produced in bacteria and purified as described in the Supplementary Experimental Procedures Section. Samples at different stages of the purification process (CL, cleared lysate; FT, flow through of the amylose column, EL, eluted fractions from the amylose column) were subjected to electrophoresis on an SDS-polyacrylamide gels. The gels were either stained (left) or blotted (right) and reacted with anti-MBP antibody (MBP-AB) with a chemiluminescent detection system.

(B) Wild type, and mutant DME, D1304N (MBP-Δ537DME(D1304N)) and K1286Q (MBP-Δ537DME(K1286Q)) were purified. Eluted fractions from amylose columns were subjected to electrophoresis on an SDS-polyacrylamide gels. The gels were either stained (left) or blotted (right) and reacted with anti-MBP antibody (MBP-Ab) using a colorimetric detection system.

A

Reference		CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
Clone	1:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	2:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	3:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	4:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	5:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	6:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	7:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	8:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	9:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	10:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	11:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT
	12:	CAATTCGTAAATTTAATTATACGCATTTAAAACTTCTACGTAAATCCGATATTTTTTACTATTTATAACGTTAAT

B

Reference		TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
Clone	1:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	2:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	3:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	4:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	5:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	6:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	7:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	8:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	9:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	10:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	11:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT
	12:	TCCGAAN$_{46}$ATCGCGAACGACN$_{61}$ATCGTTAATGACCACGGTTAAATGGCGGGATCCAAAATCCGGTN$_{20}$TACGT

Figure S2. Sequence of Maternal *MEA* Allele in Endosperm of *dme-2* Col-*gl* Pollinated by RLD

Crosses, seed dissection, and DNA isolation are as described for Figure 2A. Primers and cloning are described in the Supplemental Experimental Procedures section. Reference sequence is Col-*gl* and 'N' is any base. CpG sites in red are hypomethylated in a DME-dependent manner in the endosperm. Twelve sequenced clones of the –500 bp region of the *MEA* promoter (A) and the MEA-ISR (B) had no C→T transition mutations at CpG sites.