DNA demethylation in the *Arabidopsis* genome

Jon Penterman*, Daniel Zilberman†, Jin Hoe Huh*, Tracy Ballinger†‡, Steven Henikoff†‡§, and Robert L. Fischer*§

*Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720; and †Basic Sciences Division and ‡Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

Contributed by Steven Henikoff, February 28, 2007 (sent for review January 30, 2007)

Cytosine DNA methylation is considered to be a stable epigenetic mark, but active demethylation has been observed in both plants and animals. In *Arabidopsis thaliana*, DNA glycosylases of the DEMETER (DME) family remove methylcytosines from DNA. Demethylation by DME is necessary for genomic imprinting, and demethylation by a related protein, REPRESSOR OF SILENCING1, prevents gene silencing in a transgenic background. However, the extent and function of demethylation by DEMETER-LIKE (DML) proteins in WT plants is not known. Using genome-tiling microarrays, we mapped DNA methylation in mutant and WT plants and identified 179 loci actively demethylated by DML enzymes. Mutations in DML genes lead to locus-specific DNA hypermethylation. Reintroducing WT DML genes restores most loci to the normal pattern of methylation, although at some loci, hypermethylated epialleles persist. Of loci demethylated by DML enzymes, >80% are near or overlap genes. Genic demethylation by DML enzymes primarily occurs at the 5′ and 3′ ends, a pattern opposite to the overall distribution of WT DNA methylation. Our results show that demethylation by DML DNA glycosylases edits the patterns of DNA methylation within the *Arabidopsis* genome to protect genes from potentially deleterious methylation.

DNA glycosylase | epigenetics | genome maintenance

Arabidopsis is a eukaryotic model for DNA methylation studies. In *Arabidopsis* cytosine methylation is found in all sequence contexts (CG, CNG, and CNN) and is important for genomic imprinting and genome defense against transposable elements (1). Most DNA methylation is located at transposon-rich heterochromatic regions (2–4). However, genome-wide mapping of methylation has showed that a significant fraction (20–33%) of genes are methylated (3–5). In general, methylation within *Arabidopsis* genes is concentrated in the middle and distributed away from 5′ and 3′ ends, suggesting that 5′ and 3′ methylation is detrimental to gene function (3, 4). The mechanisms that maintain gene ends relatively free of methylation are not known.

DNA methylation is often considered a stable epigenetic mark, but enzymatic DNA demethylation is known to occur. In mammals, the paternal pronucleus is actively demethylated immediately after fertilization (6, 7) but the enzymes responsible are unknown (8). In *Arabidopsis*, DNA demethylation is mediated by the DEMETER (DME) family of bifunctional helix–hairpin–helix DNA glycosylases that have both DNA glycosylase and apurinic/apyrimidinic (AP) lyase activities (9–13). The DNA glycosylase initiates the base excision repair process by specifically excising 5-methylcytosine through cleavage of the *N*-glycosylic bond. AP lyase subsequently nicks the DNA, and an AP endonuclease generates a 3′-hydroxyl to which a DNA repair polymerase adds an unmethylated cytosine. DNA ligase completes the repair process by sealing the nick.

Demethylation by DME is one step in a developmental pathway that establishes genomic imprinting in the *Arabidopsis* endosperm (9, 10, 14). Three targets of DME are *MEDEA* (*MEA*), *FWA*, and *FIS2*. In vegetative tissue, the default state for these genes is methylated (14–16). However, in the reproductive central cell, DME excises 5-methylcytosine at *MEA* and presumably *FWA* and *FIS2*, establishing hypomethylated, transcriptionally active alleles of these genes (9, 10, 14, 15). DNA demethylation by REPRESSOR OF SILENCING1 (ROS1), a DME homolog, occurs in transgenic plants and maintains the transcriptionally active states of a *RD29A*::*LUCERIFERASE* (*RD29A*::*LUC*) reporter gene and the endogenous *RD29A* gene (11–13). In *ros1*; *RD29A*::*LUC* plants, *RD29A*::*LUC* and *RD29A* become heavily methylated and transcriptionally silenced (11). The function of two other members of the DME family (9), DEMETER-LIKE2 (DML2) and DML3, has not been reported.

We used genome-tiling microarrays to identify loci demethylated by ROS1, DML2, and DML3 in WT adult plants. We show that the DML enzymes demethylate ≈180 discrete loci throughout the genome. Of these loci, >80% are within genic regions where DML enzymes primarily demethylate the 5′ and 3′ ends. Our data indicate that demethylation by DML enzymes functions to protect endogenous genes from potentially deleterious methylation. Also, our data strongly imply that DML demethylation defines the genomic methylation profile of *Arabidopsis* genes.

Results

DML2 and DML3 Excise 5-Methylcytosine *in Vitro*. To determine whether DML2 and DML3 can excise 5-methylcytosine, we compared the 5-methylcytosine activities of all four DME family members by using an *in vitro* system published in ref. 10. In this assay, the AP lyase activity of DML enzymes is used to monitor their activity on methylated and nonmethylated double-strand oligonucleotides [supporting information (SI) Fig. 5]. DML3 lyase activity was detected in reactions with oligonucleotides bearing 5-methylcytosine in all contexts (CG, CNG, CNN), but no activity was detected with nonmethylated oligonucleotides (SI Fig. 5). DML2 lyase activity was also detected, albeit much weaker, in a reaction with oligonucleotides bearing 5-methylcytosine at a CG site, but not in reactions with nonmethylated oligonucleotides (SI Fig. 5). Thus, DML2 and DML3 are bifunctional DNA glycosylases that can excise 5-methylcytosine *in vitro*.

Genome-Wide Mapping of Methylation Identifies Loci Demethylated by DML Enzymes. To determine the functions of *DML* genes, we isolated loss-of-function T-DNA alleles (Fig. 1*A*). Single, double, and triple F_2 mutants had no overt morphological phenotypes under the growth conditions we used. We examined the expression of *ROS1*, *DML2*, and *DML3* by using RT-PCR and detected expression of all DMLs in the major organ systems of plants: roots, stems, leaves, and inflorescences (SI Fig. 6). Because DML

Author contributions: J.P., D.Z., J.H.H., S.H., and R.L.F. designed research; J.P., D.Z., J.H.H., and T.B. performed research; T.B. contributed new reagents/analytic tools; J.P., D.Z., J.H.H., T.B., S.H., and R.L.F. analyzed data; and J.P., D.Z., S.H., and R.L.F. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Abbreviations: AP, apurinic/apyrimidinic; T-DNA, portion of the Ti (tumor-inducing) plasmid that is transferred to plant cells.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE7336).

§To whom correspondence may be addressed. E-mail: steveh@fhcrc.org or rfischer@berkeley.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0701861104/DC1.

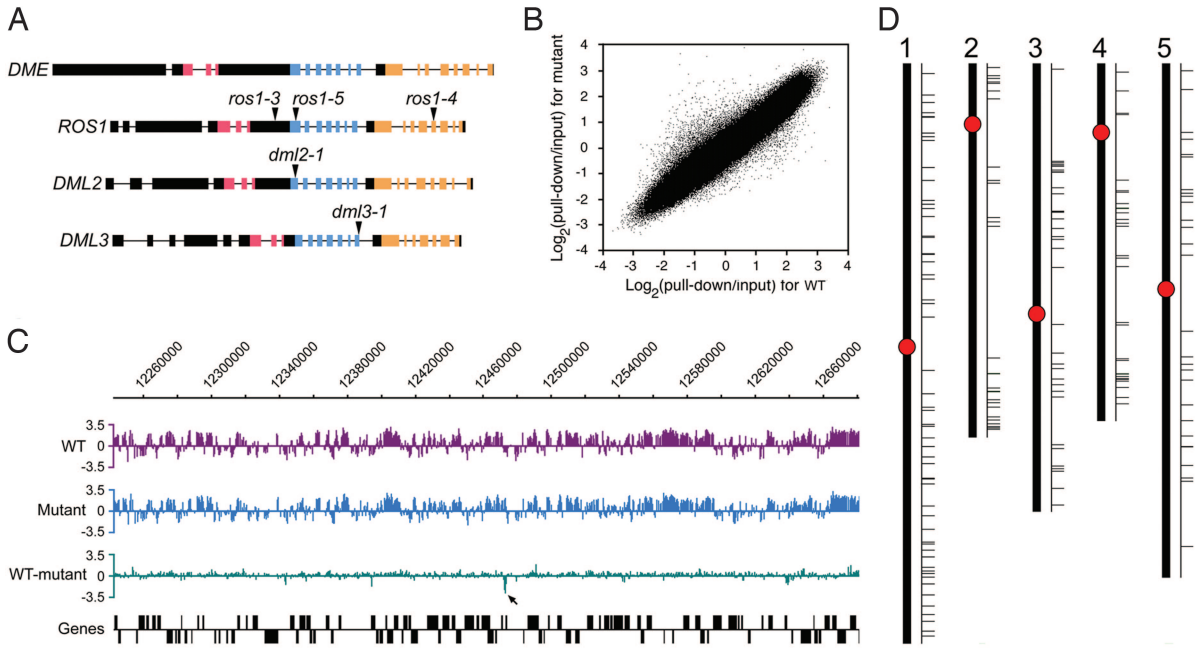

Fig. 1. Mapping genome-wide DNA methylation in WT and *ros1–3*; *dml2–1*; *dml3–1* identifies DML target loci. (*A*) Gene diagrams of the DME family members. Boxed regions are exons, and lines are introns. Blue exons encode the helix–hairpin–helix DNA glycosylase domain, and pink and orange exons encode conserved domains of unknown function (12). Black exons encode amino acids not shared between DML proteins. The position of *ros1–3*, *ros1–4*, *ros1–5*, *dml2–1*, and *dml3–1* T-DNA insertions is marked by a triangle. (*B*) Scatter plot showing the correlation between WT and *ros1–3*; *dml2–1*; *dml3–1* microarray experiments. The correlation coefficient (*r*) of the two data sets is 0.97. (*C*) Example of tiling microarray data. The top scale is the position in base pairs on chromosome 1. Each bar represents a single probe $\log_2$ signal ratio (5′-methylcytosine antibody pull-down/input) for WT and mutant data sets. For the [WT–mutant] data set, each bar represents the subtraction of a mutant $\log_2$ signal ratio from the corresponding WT $\log_2$ signal ratio, and a negative value in the [WT–mutant] data set is indicative of mutant hypermethylation. Genes are represented by black boxes; ones above the line are oriented 5′ to 3′ from left to right, and ones below are oriented 5′ to 3′ from right to left. An arrow indicates a locus hypermethylated in the mutant. Notice how the methylation profile flanking this locus is relatively similar between WT and mutant. (*D*) Genomic location of DML target loci. Shown are chromosomes 1–5 and their centromeres (red circles). To the right of each chromosome are horizontal lines that indicate the positions of each hypermethylated locus in the *ros1–3*; *dml2–1*; *dml3–1* genome.

proteins excise 5-methylcytosine *in vitro*, we hypothesized that, in mutants, loci demethylated by these proteins would become hypermethylated relative to WT. To test this hypothesis, we generated a line with T-DNA mutations in all three DML genes (*ros1–3*; *dml2–1*; *dml3–1*) and a sibling WT line from a self-pollinated triple heterozygote. We then mapped DNA methylation genome-wide in both the WT and triple mutant genomes by using a 5′-methylcytosine antibody and tiling microarrays as described in ref. 3.

We did not observe global DNA methylation changes in *dml* mutant plants: The overall methylation levels of WT and the triple mutant were very similar for the 382,178 isothermal oligonucleotide probes on the array [correlation coefficient (Pearson's r) = 0.97; Fig. 1 *B* and *C*].

We asked whether normal DNA methylation at a subset of genomic loci depends on DML-dependent demethylation. We subtracted the individual triple mutant probe values from WT values to generate the [WT–mutant] data set, in which negative probe values reflect more methylation in the mutant and positive probe values reflect more methylation in WT. We then calculated the average value for four contiguous probes throughout the [WT–mutant] data set (four-probe sliding window) and compared the distribution of [WT–mutant] window values to those derived from a random, normally distributed data set (Table 1). As expected, most window values of the [WT–mutant] data set conformed to a normal distribution. However, at 4 or more SDs from the mean, there were significantly more negative four-probe windows in the [WT–mutant] data set than expected (496 observed vs. 17 expected at 4 SD; 199 observed vs. 0.2 expected at 5 SD) (Table 1). We joined all overlapping negative windows with values greater than 4 SD and identified 179 loci whose methylation levels increased in *ros1–3*; *dml2–1*; *dml3–1* plants (Fig. 1*D* and SI Table 2). The hypermethylated loci were interspersed among the five chromosomes and were solitary features surrounded by equally methylated DNA (Figs. 1 *C* and *D* and 2 *A–C*). There were also more positive four-probe window values in the [WT–mutant] data set than expected at 5 SD from the mean (12 observed vs. 0.2 expected) (Table 1). These positive four-probe windows corresponded to only three loci hypomethylated in mutant plants (SI Table 3). Thus, these data indicate that DML DNA glycosylases demethylate ≈179 discrete loci throughout the genome. These loci represent sites where two opposing pathways, DNA methylation and demethylation, converge.

Bisulfite Sequencing Confirms Microarray Data and Analysis. From the list of hypermethylated loci we chose 17 to verify by using bisulfite sequencing, which is a method to quantitatively measure methylation of each cytosine (SI Tables 4 and 5). We bisulfite sequenced DNA from siblings of the mutant and WT plants used in the microarray experiment. All 17 loci were hypermethylated

Table 1. Distribution of expected and [WT–mutant] four-probe windows whose values are greater than 1, 2, 3, 4, or 5 SDs from the mean

Four-probe windows	1 SD	2 SD	3 SD	4 SD	5 SD
Expected number	51,951	7,966	450	17	0.2
[WT–mutant] negative	52,104	7,428	1,478	496*	199*
[WT–mutant] positive	54,676	6,972	359	27	12*

*Significantly greater than expected (>96.5% confidence level).

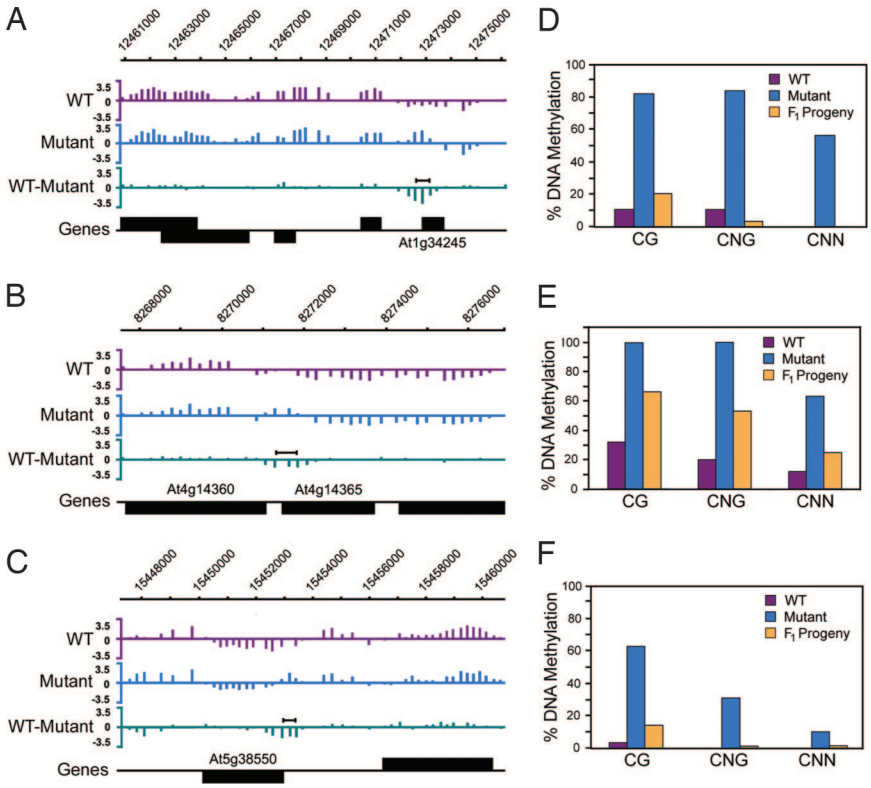

Fig. 2. Bisulfite sequencing confirms methylation profiles of WT and mutant. (*A–C*) Examples of tiling microarray data showing three loci that were confirmed to be hypermethylated in *ros1–3*; *dml2–1*; *dml3–1*. The black bars above the [WT–mutant] data set indicate the positions of bisulfite sequencing. (*D–F*) Bisulfite sequencing data showing percents of CG, CNG, and CNN methylation for loci shown in *A*, *B*, and *C*, respectively, in WT, mutant, and their F_1 progeny. See SI Tables 4 and 5 for more details.

in the mutant background relative to WT, confirming the array and computational analysis (Fig. 2 *D–F* and SI Tables 4 and 5). Hypermethylation was observed in all sequence contexts, consistent with the *in vitro* excision of CG, CNG, and CNN methylation by DML proteins (12, 13) (SI Fig. 5).

Locus-Specific Hypermethylation Is *dml*-Dependent. We tested the specificity of hypermethylation to the *dml* mutations by bisulfite sequencing DNA from an independently generated triple mutant and WT line (SI Table 6) and from mutant and control WT plants that had been inbred for an additional three generations (SI Table 7). In these experiments we discovered that three loci (At3g16000, At4g19720, and At1g53860) became highly methylated in a WT background, making the relationship between their methylation level and DML demethylation less clear; these loci were not analyzed further. Of the remaining loci, hypermethylation at 13/14 cosegregated with the *dml* mutations (SI Table 6) and 13/14 were still hypermethylated in the mutant F_6 plants relative to WT (SI Table 7). Taken together, these data indicate that the hypermethylation at the majority of loci identified by microarray analysis is *dml*-dependent. DML DNA glycosylases regulate the methylation of these loci by excising CG, CNG, and CNN methylation.

Specificity and Redundancy in Demethylation by DMLs. To test whether DML enzymes regulate methylation redundantly or independently, we compared the methylation levels of loci in each single mutant to the levels found in WT and triple mutant by using bisulfite sequencing. For 7 of 14 loci, the methylation levels in every single mutant were much less than in the triple mutant (Fig. 3 and SI Table 6), suggesting that multiple DML enzymes demethylate these loci. We also discovered loci whose methylation was regulated by a single DML. In *ros1–3*, *ros1–4*, and *ros1–5* single mutants, five loci (At1g34245, At5g38550, At5g48280, At1g26400, and At4g14365) were hypermethylated relative to WT (Fig. 3 and SI Tables 6 and 8), indicating that ROS1 specifically demethylates these loci in WT plants. Surprisingly, we found that the *SUPERMAN* gene (At3g23130) and At1g29930 were primarily demethylated by DML2, which showed the least activity on 5-methylcytosine *in vitro* (SI Table 6).

DML Demethylation Prevents the Formation of Stable Epialleles. We crossed *ros1–3*; *dml2–1*; *dml3–1* to WT and compared the F_1 progeny's and progenitors' methylation levels to see whether

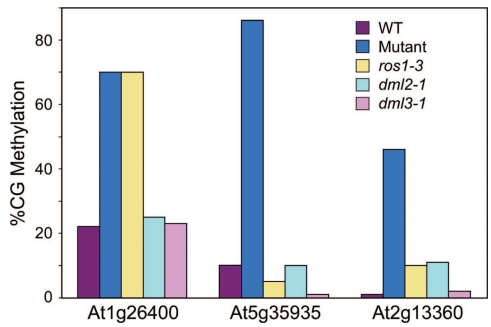

Fig. 3. Bisulfite sequencing of single mutants shows that loci are demethylated exclusively by a single DML or redundantly by multiple DMLs. Graphed are the CG methylation levels of At1g26400, At5g35935, and At2g13360 in WT, triple mutant, *ros1–3*, *dml2–1*, and *dml3–1* backgrounds. For all loci, the behavior of CG methylation is representative of methylation at CNG and CNN sites.

reintroducing WT *DML* alleles to the triple mutant background could restore the methylation to WT levels. In the heterozygous progeny the methylation levels at 11/14 loci were similar to the levels in WT plants (Fig. 2 *D* and *F* and SI Table 4), suggesting that DML enzymes again target these loci for demethylation. The methylation levels of three loci (At5g48280, At4g14365, and At3g45940) in the progeny were intermediate between the levels measured in the WT and triple mutant (Fig. 2*E* and SI Table 4). For these three loci, two differentially methylated alleles could be distinguished and their parent-of-origin could be inferred (SI Fig. 7 and data not shown). The intermediate methylation level in progeny was due to the inheritance of hypermethylated epialleles from the mutant parent. Thus, DML demethylation prevents the formation of stable hypermethylated epialleles at some loci in the *Arabidopsis* genome.

DML Enzymes Demethylate a Gene Set That Is Representative of the Genome. Of loci demethylated by DMLs, >80% were within or near (≤500 bp) genes (SI Table 2). Using the *Arabidopsis* Gene Ontology (GO) resource (17), we compared the functional categories of DML demethylated genes ($n = 146$) to those found in the *Arabidopsis* genome (>26,000). To help us gauge the results of this comparison, we also compared the GO analysis of the genome to genes ($n = 177$) targeted by cloned microRNAs (18), some of which have integral roles in *Arabidopsis* development, growth, physiology, and disease (19–22). For genes demethylated by DMLs, the proportion in all functional categories was very similar to their proportion in the *Arabidopsis* genome (SI Table 9), indicating that DML demethylation is not specific to any particular gene class. By contrast, the GO analysis of microRNA target genes showed a significant enrichment for transcription factors and nucleotide-binding proteins (SI Table 9), which is consistent with their critical roles in *Arabidopsis*. Thus, the functional categories of genes demethylated by DMLs are proportionally representative of the genome.

DML Enzymes Primarily Demethylate the 5′ and 3′ Ends of Genes. We searched for patterns in the location of DML demethylation relative to genic regions (defined here as the transcribed region plus 500 bp of the 5′ and 3′ sequence). We took 85 genes that are not detectably methylated in WT (SI Table 2) and determined the segment of each gene that became hypermethylated in the *dml* triple mutant. We then plotted the $\log_2$ ratio of observed hypermethylated genic segments over the number expected by chance. We found that *ros1–3*; *dml2–1*; *dml3–1* genic hypermethylation is significantly distributed toward the 5′ and 3′ ends and relatively absent from the middle region (Fig. 4*A*), which is opposite to the overall distribution of WT genic DNA methylation (Fig. 4*B*) (3, 4). Furthermore, in our data set we observed genes that normally have methylation in the middle region become hypermethylated in *ros1–3*; *dml2–1*; *dml3–1* at either the 5′ or 3′ end (Fig. 2 *B* and *E* and SI Fig. 8 *A–D*). These observations show that DML enzymes primarily demethylate the 5′ and 3′ ends of these genes, a pattern that is very similar to DME-mediated demethylation of *MEA* (10).

Gene Expression Is Mostly Unaffected by *dml* Mutant Hypermethylation. Transcription of endogenous genes can be repressed by methylation at the 5′ or 3′ end (10, 23, 24). To determine whether *dml* mutant hypermethylation affects gene expression, we used semiquantitative RT-PCR to analyze the expression of 14 genes near or overlapping 12 hypermethylated loci in *ros1–3*; *dml2–1*; *dml3–1* F_3 plants (SI Table 4). For 13 genes, we detected equivalent levels of expression in WT and *dml* mutant plants (SI Fig. 9), indicating that, at this level of detection, expression of these genes was unaffected by the hypermethylation. However, for At5g38550 we detected less expression in *dml* mutants than in WT plants, suggesting that mutant hypermethylation repressed its transcription (SI Fig. 9). Our analysis and bisulfite sequencing data indicated that At5g38550 accumulated heavy methylation just after its transcription start site in both *ros1–3*; *dml2–1*; *dml3–1* and *ros1–3* mutant plants (Fig. 2 *C* and *F*, and data not shown). In addition, a region after the 3′ end of At5g38550 became methylated (Fig. 2*C*), although this was not significant enough to detect in our computational analysis. The absence of a transcriptional effect at most of the hypermethylated genes assayed suggests that WT demethylation is not required for full gene expression.

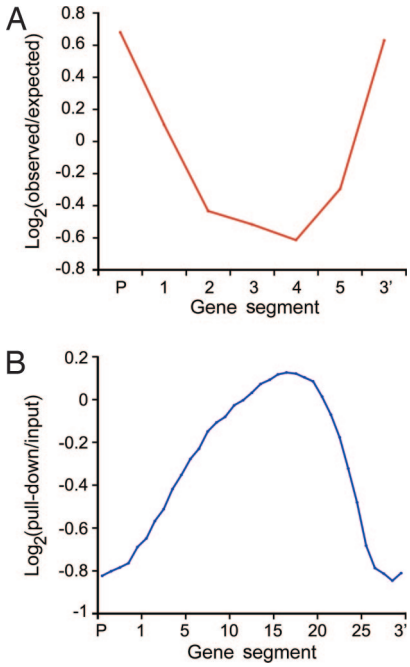

Fig. 4. Distribution of DNA methylation within genes hypermethylated in *ros1–3*; *dml2–1*; *dml3–1* (*A*) and within genes in WT (*B*). (*A*) Eighty-five genes hypermethylated in *ros1–3*; *dml2–1*; *dml3–1* were divided into five equal size segments plus a 500-bp 5′ and 3′ segment. Plotted is the $\log_2$ ratio of observed hypermethylated segments over the number expected through chance alone. The distribution is significantly different than expected ($\chi^2 = 24.2$; df = 6; $P < 0.001$). (*B*) Plotted is the average $\log_2$ signal ratio (pull-down/input) for each segment from 21,583 *Arabidopsis* genes. Genes were divided into 25 equal size segments plus five 100-bp segments on both the 5′ and 3′ ends.

Discussion

In this study, we mapped DNA methylation by using genome tiling microarrays and discovered loci that are actively demethylated by DML enzymes (Fig. 1*C*). We found that two opposing pathways, DNA methylation and demethylation, converge at nearly 180 loci scattered throughout the genome (Fig. 1*D* and SI Table 2). A recent study also found that several transposons and endogenous genes were hypermethylated in *ros1*; *RD29A*::*LUC* plants (25). Most of these loci, which normally contain methylation at CG sites, had increased CNG and CNN methylation levels (25). We identified loci exclusively demethylated by a particular DML and others demethylated by multiple DML enzymes (Fig. 3), and we demonstrated that DMLs prevent the formation of highly methylated epialleles (Fig. 2*E* and SI Fig. 7). At genes, demethylation by DML enzymes primarily occurs at the 5′ and 3′ ends (Fig. 4*A*), and, with one exception, we found that DML demethylation was not required for full gene expression (SI Fig. 9). Our data show that active DNA demethylation is a broadly used process in *Arabidopsis*.

In *Arabidopsis*, there are two known functions for DNA demethylation. Demethylation by DME activates gene expres-

sion during development by establishing a new epigenetic, hypomethylated state (10), whereas ROS1 demethylation prevents transcriptional gene silencing by maintaining a locus relatively free of methylation (11, 25). DMLs likely do not have a major role in plant development, because *ros1–3*; *dml2–1*; *dml3–1* plants grow and develop normally under our conditions. In addition, demethylation by DMLs is likely not required to activate the expression of many genes identified here, because the expression of 13/14 genes was unaffected by the hypermethylation in the triple mutant (SI Fig. 9). Instead, our data indicate that DML demethylation generally prevents the accumulation of methylation at or near genes.

At genes, DML enzymes primarily demethylate the 5′ and 3′ ends. In WT, the 5′ and 3′ ends of genes are less likely to be methylated, whereas the middle segments are more likely to be methylated (Fig. 4*B*) (3, 4). Thus, the genic regions least likely to be methylated in WT correspond to regions most likely demethylated by DML enzymes. These results suggest a connection between DML demethylation and the overall genomic profile of genic methylation. DMLs might be specifically demethylating the 5′ and 3′ ends of genes, while leaving the middle region free to accumulate DNA methylation over evolutionary time (Fig. 4 and SI Fig. 8). Consistent with this idea is the proportional representation of the genome in the genes that DMLs presently demethylate (SI Table 9). These data and observations suggest that DML-dependent 5′ and 3′ demethylation is a factor that defines the methylation profile of genes.

A "housekeeping" function for DML demethylation implies that DML enzymes generally target genes throughout the genome and that DMLs can demethylate many more genes than those identified here. Indeed, the first description of *ROS1* supports this idea. In transgenic plants, the *RD29A*::*LUC* reporter gene causes the methylation of itself and the endogenous *RD29A* gene, and this methylation is removed by ROS1 (11). However, in nontransgenic plants, the endogenous *RD29A* gene lacks DNA methylation and therefore is not actively demethylated by ROS1 (11). DML enzymes are likely capable of demethylating many genes that, like *RD29A*, presently lack DNA methylation.

Cytosine methylation pathways defend the genome from transposable elements. However, in WT plants and certain genetic backgrounds and conditions, methylation is directed to endogenous genes, causing silencing and sometimes defects in plant development (SI Table 2) (23, 24). DML enzymes specifically remove aberrant 5′ and 3′ methylation from genes where it is most likely to interfere with transcription (3, 4). In agreement, some endogenous genes are transcriptionally repressed by *dml* mutant hypermethylation (SI Fig. 9) (11, 25). Thus, by removing 5′ and 3′ genic methylation, DML enzymes protect endogenous genes from potentially deleterious methylation. This protective function of DML demethylation would have enabled plants to support robust methylation pathways for genome defense.

Materials and Methods

Cloning, Expression, and Purification of ROS1, DML2, and DML3 in *Escherichia coli*. Inflorescence RNA was extracted as described in ref. 9, treated with DNase (Invitrogen, Carlsbad, CA), and reversed transcribed to cDNA by using oligo dT primers (Ambion, Foster City, CA). Partial *ROS1*, *DML2*, and *DML3* cDNAs were PCR-amplified and cloned by using predicted gene annotations. The cDNA ends were cloned by using 5′ and 3′ RACE (Invitrogen). Full-length cDNAs were constructed and sequenced to confirm correctness. The cDNAs encoding $ROS1^{1–1393}$, $DML2^{1–1332}$, and $DML3^{1–1105}$ were inserted in frame and downstream of the maltose binding protein gene in the pMAL-c2x (New England Biolabs, Ipswich, MA). Protein expression and purification were done as described in ref. 10.

Substrate Preparation and Glycosylase Activity Assays. Substrate was prepared and glycosylase activity assays were done as described in ref. 10.

Growth Conditions, Genotyping, and Plant Materials. Growth conditions are as described in ref. 16. PCR was used to genotype plants, and the PCR conditions are in *SI Methods*. The *ros1–3* and *dml2–1* alleles were isolated from Ws-0 heterozygotes (*Arabidopsis* Knockout facility, University of Wisconsin, Madison, WI) and introgressed into the Col-0 background six times (26). The *dml3–1* allele was isolated from a Col-0 heterozygote [Syngenta *Arabidopsis* Insertion Library (SAIL) collection] and backcrossed into Col-0 four times (27). The *ros1–4* allele was isolated from a Col-0 heterozygote (Salk T-DNA collection) (28) and backcrossed into Col-0 twice. A *ros1–5* homozygous T_3 Col-0 plant and its WT counterpart were isolated from the SAIL collection (27).

Triple heterozygous *ros1–3*/*ROS1*; *dml2–1*/*DML2*; *dml3–1*/*DML3* plants were constructed and selfed, generating two independent WT F_2 lines, two independent *ros1–3*, *dml2–1*, *dml3–1* F_2 lines, and all single *ros1–3*, *dml2–1*, and *dml3–1* F_2 lines. The particular WT and *ros1–3*; *dml2–1*; *dml3–1* F_2 lines used in the microarray experiments were also used to generate the WT and *ros1–3*; *dml2–1*; *dml3–1* F_6 plants and the triple heterozygous F_1 progeny. The other independently generated WT and *ros1–3*; *dml2–1*; *dml3–1* F_2 lines were used to test for cosegregation of hypermethylated loci with *dml* mutations. Analyzed DNAs were from 25-day-old F_3 plants (minus roots), unless noted otherwise.

Immunoprecipitation and Microarray Analysis of Methylated DNA. Genomic DNA was immunoprecipitated, amplified, labeled and hybridized to microarrays as described in ref. 3. The array design is described in ref. 3. A single hybridization experiment was performed to map DNA methylation in each genotype. In the mutant microarray data set, parts of chromosome 2 (8,802,496–15,397,296 bp) and chromosome 3 (677,340–5,117,803 bp) were obviously polymorphic (data not shown) and were excluded from further analysis. The polymorphic regions of chromosome 2 and 3 are due to physical linkage of Ws-0 genomic DNA with the *ros1–3* and *dml2–1* alleles, respectively. We also excluded probes spanning 16,316,798–16,318,339 bp on chromosome 4, because this is the location of the *dml3–1* T-DNA insertion.

To identify loci whose methylation differed between WT and *ros1–3*; *dml2–1*; *dml3–1*, mutant probe values were subtracted from their counterparts in the WT data set, generating a [WT–mutant] data set. A positive probe value in [WT–mutant] means that there was less methylation in *ros1–3*; *dml2–1*; *dml3–1* (i.e., mutant hypomethylation), and a negative probe value in [WT–mutant] means that there was more methylation in *ros1–3*; *dml2–1*; *dml3–1* (i.e., mutant hypermethylation). The [WT–mutant] data set was analyzed by using a four-point sliding window to reduce the contribution of individual aberrant probes. The average value of each window was calculated after dropping the highest and lowest probe values. The [WT–mutant] window values have a mean of 0.009 with an SD of 0.25. The distribution of observed [WT–mutant] window values was then compared with the expected values of a normally distributed data set with the same mean and SD (Table 1). Overlapping windows at the 96.5% confidence level were joined to yield discrete loci (SI Tables 2 and 3).

Bisulfite Sequencing. Genomic DNA was digested with Apo I, RsaI, or Sau3A I (New England Biolabs) for 4 h, after which enzymes were inactivated. Bisulfite treatment, PCR amplification, and cloning were done as described in ref. 16. Primers can be found in the *SI Methods*.

RT-PCR Analysis of DML Genes and Genes Demethylated by DMLs. Total mRNA was isolated from roots of 12-day-old seedlings, stems, cauline leaves, inflorescences, and both WT and *ros1–3*; *dml2–1*; *dml3–1* 25-day-old whole plants (minus roots) as de-scribed. Primers and PCR conditions are found in *SI Methods*.

We thank R. Uzawa for technical assistance and M. Gehring for comments on the manuscript. This research was supported by a National Institutes of Health grant and a Ceres, Inc. grant (to R.L.F.), the Howard Hughes Medical Institute (S.H.), a University of California-Berkeley Edge Fellowship (to J.P.), and a Leukemia and Lymphoma Society Postdoctoral Fellowship (to D.Z.).

1. Chan SW, Henderson IR, Jacobsen SE (2005) *Nat Rev Genet* 6:351–360.
2. Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, *et al.* (2004) *Nature* 430:471–476.
3. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S (2007) *Nat Genet* 39:61–69.
4. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, *et al.* (2006) *Cell* 126:1189–1201.
5. Tran RK, Henikoff JG, Zilberman D, Ditt RF, Jacobsen SE, Henikoff S (2005) *Curr Biol* 15:154–159.
6. Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W, Walter J (2000) *Curr Biol* 10:475–478.
7. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T (2000) *Nature* 403:501–502.
8. Morgan HD, Santos F, Green K, Dean W, Reik W (2005) *Hum Mol Genet* 14(Spec No 1):R47–R58.
9. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL (2002) *Cell* 110:33–42.
10. Gehring M, Huh JH, Hsieh TF, Penterman J, Choi Y, Harada JJ, Goldberg RB, Fischer RL (2006) *Cell* 124:495–506.
11. Gong Z, Morales-Ruiz T, Ariza RR, Roldan-Arjona T, David L, Zhu JK (2002) *Cell* 111:803–814.
12. Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marin MI, Martinez-Macias MI, Ariza RR, Roldan-Arjona T (2006) *Proc Natl Acad Sci USA* 103:6853–6858.
13. Agius F, Kapoor A, Zhu JK (2006) *Proc Natl Acad Sci USA* 103:11796–11801.
14. Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, Fischer RL, Kakutani T (2004) *Science* 303:521–523.
15. Jullien PE, Kinoshita T, Ohad N, Berger F (2006) *Plant Cell* 18:1360–1372.
16. Xiao W, Gehring M, Choi Y, Margossian L, Pu H, Harada JJ, Goldberg RB, Pennell RI, Fischer RL (2003) *Dev Cell* 5:891–901.
17. Berardini TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M, Zhang P, Mueller LA, Yoon J, Doyle A, Lander G, *et al.* (2004) *Plant Physiol* 135:745–755.
18. Gustafson AM, Allen E, Givan S, Smith D, Carrington JC, Kasschau KD (2005) *Nucleic Acids Res* 33:D637–40.
19. Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JD (2006) *Science* 312:436–439.
20. Chiou TJ, Aung K, Lin SI, Wu CC, Chiang SF, Su CL (2006) *Plant Cell* 18:412–421.
21. Fujii H, Chiou TJ, Lin SI, Aung K, Zhu JK (2005) *Curr Biol* 15:2038–2043.
22. Chen X (2005) *FEBS Lett* 579:5923–5931.
23. Kishimoto N, Sakai H, Jackson J, Jacobsen SE, Meyerowitz EM, Dennis ES, Finnegan EJ (2001) *Plant Mol Biol* 46:171–183.
24. Jacobsen SE, Sakai H, Finnegan EJ, Cao X, Meyerowitz EM (2000) *Curr Biol* 10:179–186.
25. Zhu J, Kapoor A, Sridhar VV, Agius F, Zhu JK (2007) *Curr Biol* 17:54–59.
26. Sussman MR, Amasino RM, Young JC, Krysan PJ, Austin-Phillips S (2000) *Plant Physiol* 124:1465–1467.
27. Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, *et al.* (2002) *Plant Cell* 14:2985–2994.
28. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, *et al.* (2003) *Science* 301:653–657.