Cell, Volume *126*

Supplemental Data

Resource

Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in *Arabidopsis*

Xiaoyu Zhang, Junshi Yazaki, Ambika Sundaresan, Shawn Cokus, Simon W.-L. Chan, Huaming Chen, Ian R. Henderson, Paul Shinn, Matteo Pellegrini, Steve E. Jacobsen, and Joseph R. Ecker

Supplemental Experimental Procedures

Microarray Design

We designed two (one forward strand and one reverse strand) Affymetrix *Arabidopsis* whole-genome tiling arrays, each contains ~6.4 million 25-nt oligonucleotide features. Each array contains 3.2 million perfect match features (PM) that perfectly match genomic sequence, and every PM feature is accompanied by a single mismatch feature (MM) with the central base (position 13 of 25) substituted by its complement. One array covers ~97% of one strand of all chromosomes; each 35 bp genomic region is represented by a 25-nt feature chosen for optimal hybridization characteristics within its 35 bp window. The complementary strand is covered by the second array. Forward strand arrays were used for DNA methylation analysis while RNA expression analysis was performed using arrays for both strands. The array platform design can be found at the Gene Expression Omnibus (GEO) at NCBI (http://www.ncbi.nlm.nih.gov/geo/, array platform GPL1979 for forward strand design and GPL1980 for reverse strand design).

Of the 3.2 million PM oligonucleotide features, ~90% match a unique position in the genome ("unique features"). Of the ~10% remaining features ("repetitive features"), ~30% match two positions and the remaining ~70% match three or more positions. For expression profiling, only unique features were considered. For DNA methylation analysis, hybridization intensities were mapped to chromosomal positions in two ways. In the first method, only hybridization intensities from unique features were used for subsequent analysis and all repetitive features were ignored. In the second method, the hybridization intensity from a repetitive feature (~10 % of the total features) was mapped to all genomic locations that the feature matched, and these "remapped" intensities were used with those from unique features for subsequent analysis. These two groups of intensities were analyzed independently but with the same procedure to define methylated regions (see below), and the union of the methylated regions from the two methods was presented as the final "methylated regions".

Amplification, Labeling, and Microarray Hybridization for DNA Methylation Analysis

DNA amplification was performed using the GenomePlex® Whole Genome Amplification Kit (Sigma). Six amplification reactions were performed in parallel and the products were pooled to minimize spurious amplification artifacts. 30 µg of amplification product were treated briefly with DNaseI and labeled in a 75 µl mixture containing 5.1 nM biotin-16-ddUTP, 5.1 nM biotin-16-dUTP, 1,360 units of terminal transferase (Roche) in 1X One-Phor-All buffer (Amersham) at 37°C for 1.5 hours. Labeled fragments were purified with MicroSpin G-50 Columns (Amersham). Microarray hybridization, staining, and washing were performed according to the Eukaryotic Target Protocol (Affymetrix Inc, Santa Clara, CA). Scanning was performed at 0.7 µm resolution using a GeneChip Scanner 3000 7G (Affymetrix Inc, Santa Clara, CA).

Whole-Genome Tiling Microarray Analysis of Gene Expression in Wild-Type, *met1*, and *drm1 drm2 cmt3*

RNA Extraction

Total RNA was extracted using the RNeasy Kit (QIAGEN) from the same tissues and at the same developmental stage as those used for DNA methylation analysis and treated with DNase I. The quality and quantity of RNA was examined by both gel electrophoresis and spectrophotometry.

Probe Synthesis and Microarray Hybridization

Five µg of total RNA was used for synthesis of double-stranded cDNA using the GeneChip One-Cycle cDNA Synthesis Kit (Affymetrix). Biotin-labeled cRNA was generated by *in vitro* transcription using the GeneChip IVT Labeling Kit (Affymetrix), fragmented, and approximately 15 µg of cRNA was hybridized to the tiling arrays. Microarray hybridization, staining, and washing were performed according to the Eukaryotic Target Protocol (Affymetrix). Three biological replicates were performed for each genotype on each strand.

Methylation Data Processing and Analysis

Three biological replicates were performed for each genotype with each method, yielding very consistent results (Table S9). Raw tiling microarray data processing was performed in a manner similar to previously described (Bernstein et al., 2005). First, PM intensities were corrected for background using a convolution model-based method (RMA, available in the Bioconductor package of R) (Irizarry et al., 2003). Next, background-corrected values from the three replicates of the same group (where either methylated or unmethylated DNA was used as probes) were subjected to quantile normalization separately. Finally, the two groups were normalized to the same mean. For this final step we also evaluated normalizing two groups according to their mode or median; all analysis procedures yielded very similar results (not shown).

Methylated regions are expected to show a significantly higher hybridization signal when methylated DNA is used as probe compared to unmethylated DNA. We used three different statistical methods to identify these regions and compared the results. In the first method, probe-level intensities from each group were first summarized using a Tukey Biweight estimator and then compared using a Wilcoxon Signed-Rank Test (Hollander and Wolfe, 1999). This was performed using the Tiling Analysis Software (TAS version 1.0.08, Affymetrix) using the "two sample comparison analysis" option with bandwidth 200. In the second method, probe-level *t*-statistics were calculated using an empirical Bayes shrinkage estimator (Ji and Wong, 2005). A two-state hidden Markov model (HMM) was trained to assign probes as either methylated or unmethylated. This was performed using the software package TileMap with the HMM option (Ji and Wong, 2005). The third method was a nonparametric approach that we developed to determine, for a given region, whether the distribution of the logarithm of the ratios between methylated and unmethylated fractions was significantly different from the global distribution. We computed, for each probe, all 9 ratios between the three bound and three unbound fractions, eliminated the single highest and lowest ratio, and computed the average logarithm of the remaining 7 ratios to generate our test statistic. Next, the distribution of our test statistic over each 400 bp (i.e., approximate average sonicated fragment length) window was compared to the global distribution of the test statistic over the entire array. This comparison was performed via a Kolmogorov-Smirnov (KS) test. With F and G being the cumulative distribution functions (cdf's) of the local and global ratios, respectively, the KS-statistic is $K = \max(|F(x) - G(x)|)$. As the (asymptotic) distribution of K is well-known, K is convertible to a probability (p-value) that the ratio distribution is greater in any 400-bp window than in the whole genome. This approach effectively measures the probability that the probe centered in the window has a higher hybridization signal when probed with methylated DNA and also takes into account immediately surrounding probes. This calculation was implemented in MATLAB (MathWorks) using a custom-coded fast KS-test. For each method, "methylated" probes were called by applying a cutoff (posterior probability > 0.5 for TileMap) and "methylated regions" were then derived by combining neighboring methylated probes allowing a maximal gap of 200 bases and requiring a minimal run of 50 bases.

Computational Analyses of DNA Methylation

The *Arabidopsis* genome annotation used in this analysis was based on TAIR/NCBI Arabidopsis Genome Release 6.0 (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=3702). Interspersed repeats were identified with RepeatMasker using two independent repeat libraries, one downloaded from Repbase (http://www.girinst.org/) (Jurka et al., 2005) and the other generated by the *de novo* repeat mining program RECON (Bao and Eddy, 2002) (Z. Bao and S.R. Eddy, unpublished data). Tandem and inverted repeats were identified using the programs Tandem Repeat Finder (Benson, 1999) and Inverted Repeat Finder (Warburton et al., 2004), respectively. The siRNA cluster data was based on a recent large-scale sequencing study and provided by P. Green and B. Meyers (University of Delaware) (Lu et al., 2005).

To correlate tissue-specific gene expression with DNA methylation, we utilized a large public dataset of gene expression measuring different tissues and time points during the development of *Arabidopsis* (http://www.weigelworld.org/resources/microarray/AtGenExpress/) (Schmid et al., 2005). The data set contains 79 samples with three biological replicates each and gene expression values are reported as the logarithm of intensity.

For analysis of the intensity distributions, known and unknown expressed genes were assigned to one of three groups: body-methylated, promoter-methylated, and unmethylated (see text for definition). We then plotted the normalized distribution for each group as seen in Figure 3D.

To examine the tissue distributions of the expressed genes, we computed their entropy. Here, entropy measures the amount of variation in the expression profile of a gene: genes that are uniformly expressed across all tissues have high entropy whereas genes expressed only in a subset of tissues have low entropy. To compute the entropy distributions of genes in each group, we first converted the data from logarithmic scale to linear intensity by taking 2 to the power of the reported value. Since the distribution of intensities for each gene and not their overall values was of interest, we normalized the expression of each gene by dividing the data by the sum over all expression intensities of each condition. We removed the background signal from the expression profile of each gene by subtracting the median value of the profile; although this approach clearly removes part of the signal along with the background, it is essential for estimating the entropy of genes that are only expressed in small numbers of tissues and has only a modest effect on the entropy estimate of globally expressed genes. Finally, using the background-corrected and normalized expression profiles for each gene, we compute the tissue entropy of gene *i* as

$$E_i = \sum_{j=1}^{N} -I_{ij} \log I_{ij}$$

where I_{ij} is the normalized intensity of expression of gene *i* in condition *j* and *N* is the number of conditions. We then plotted the distribution of entropies for each of the three classes of genes as shown in Figure 3E.

Analysis of Tiling Array Expression Data

Relative Change of Expression Level between Two Genotypes

The relative change of expression level for a given gene between two genotypes (e.g., between *met1* and wild-type) was estimated as follows. The probe-level measurement of expression change was calculated as $\log_{10}(m_1+m_2+m_3) - \log_{10}(c_1+c_2+c_3)$, where (m_1,m_2,m_3) and (c_1,c_2,c_3) are background-corrected and normalized probe intensities from the three replicates of *met1* and wild-type, respectively. The gene-level measurement of fold change was a trimmed geometric average of all probes in the exons and UTRs (on average ~35 probes per gene). The trim was *m* probes with the lowest and *m* probes with the highest ratios, where *m* was the smaller of 5 probes and 10% of probes taken from the gene. This method was also used to measure the relative change of antisense transcription in an entire gene or a specific region (e.g., in a methylated region).

Identification of Differentially Expressed Regions

Although the TileMap package was originally developed for ChIP-on-chip analysis (see above), the same algorithms can also be applied to analyze tiling microarray data measuring transcription activity. The major limitation of this method is that it does not take into account all the probes in a particular gene and thus does not offer a gene-level measurement of transcriptional changes. However, it can be applied to identify regions with significantly up- or down-regulated gene expression, such as a fraction of an exon, intergenic regions, or antisense transcripts. For each strand and each pair of genotypes in comparison, microarray feature intensities for all replicates (e.g., 3 for wild-type and 3 for *met1* for the forward strand) were quantile normalized and subjected to TileMap analysis. Regions with a posterior probability > 0.5 were defined as differentially expressed.

Summary of Expression Analysis

Results presented in Figure 4 and Figure 5 were derived using the "relative change of expression level between two genotypes" method. Differentially expressed intergenic noncoding RNAs (e.g., those shown in Figure 7) were identified using TileMap. A more stringent standard was used to define significantly overexpressed genes (Table S2 and S3): each such gene must contain at least one region identified by TileMap as significantly overexpressed, and the geometric average of all probes within the gene must be at least 1.5-fold higher in the mutant compared to wild-type.

Supplemental References

Ashapkin, V. V., Kutueva, L. I., and Vanyushin, B. F. (2002). The gene for domains rearranged methyltransferase (DRM2) in Arabidopsis thaliana plants is methylated at both cytosine and adenine residues. FEBS Lett 532, 367-372.

Bao, Z., and Eddy, S. R. (2002). Automated de novo identification of repeat sequence families in sequenced genomes. Genome Res *12*, 1269-1276.

Bao, N., Lye, K. W., and Barton, M. K. (2004). MicroRNA binding sites in Arabidopsis class III HD-ZIP mRNAs are required for methylation of the template chromosome. Dev Cell 7, 653-662.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res *27*, 573-580.

Bernstein, B. E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D. K., Huebert, D. J., McMahon, S., Karlsson, E. K., Kulbokas, E. J., 3rd, Gingeras, T. R.*, et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.

Cao, X., and Jacobsen, S. E. (2002a). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc Natl Acad Sci U S A 99 Suppl 4, 16491-16498.

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

Henderson, I. R., Zhang, X., Lu, C., Johnson, L., Meyers, B. C., Green, P. J., and Jacobsen, S. E. (2006). Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nat Genet 38: 721-725

Hollander, M., and Wolfe, D. A. (1999). Nonparametric Statistical Methods, Second Edition (New York: John Wiley and Sons, Inc.).

Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B., and Speed, T. P. (2003). Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31, e15.

Jacobsen, S. E., and Meyerowitz, E. M. (1997). Hypermethylated SUPERMAN epigenetic alleles in arabidopsis. Science 277, 1100-1103.

Jacobsen, S. E., Sakai, H., Finnegan, E. J., Cao, X., and Meyerowitz, E. M. (2000). Ectopic hypermethylation of flower-specific genes in Arabidopsis. Curr Biol 10, 179-186.

Jean Finnegan, E., Kovac, K. A., Jaligot, E., Sheldon, C. C., James Peacock, W., and Dennis, E. S. (2005). The downregulation of FLOWERING LOCUS C (FLC) expression in plants with low levels of DNA methylation and by vernalization occurs by distinct mechanisms. Plant J 44, 420-432.

Ji, H., and Wong, W. H. (2005). TileMap: create chromosomal map of tiling array hybridizations. Bioinformatics 21, 3629-3636.

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110, 462-467.

Lu, C., Tej, S. S., Luo, S., Haudenschild, C. D., Meyers, B. C., and Green, P. J. (2005). Elucidation of the small RNA component of the transcriptome. Science 309, 1567-1569.

Luff, B., Pawlowski, L., and Bender, J. (1999). An inverted repeat triggers cytosine methylation of identical sequences in Arabidopsis. Mol Cell 3, 505-511.

Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H., and Kakutani, T. (2001). Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. Nature 411, 212-214.

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D., and Lohmann, J. U. (2005). A gene expression map of Arabidopsis thaliana development. Nat Genet 37, 501-506.

Stokes, T. L., Kunkel, B. N., and Richards, E. J. (2002). Epigenetic variation in Arabidopsis disease resistance. Genes Dev 16, 171-182.

Warburton, P. E., Giordano, J., Cheung, F., Gelfand, Y., and Benson, G. (2004). Inverted repeat structure of the human genome: the X-chromosome contains a preponderance of large, highly homologous inverted repeats that contain testes genes. Genome Res 14, 1861-1869.

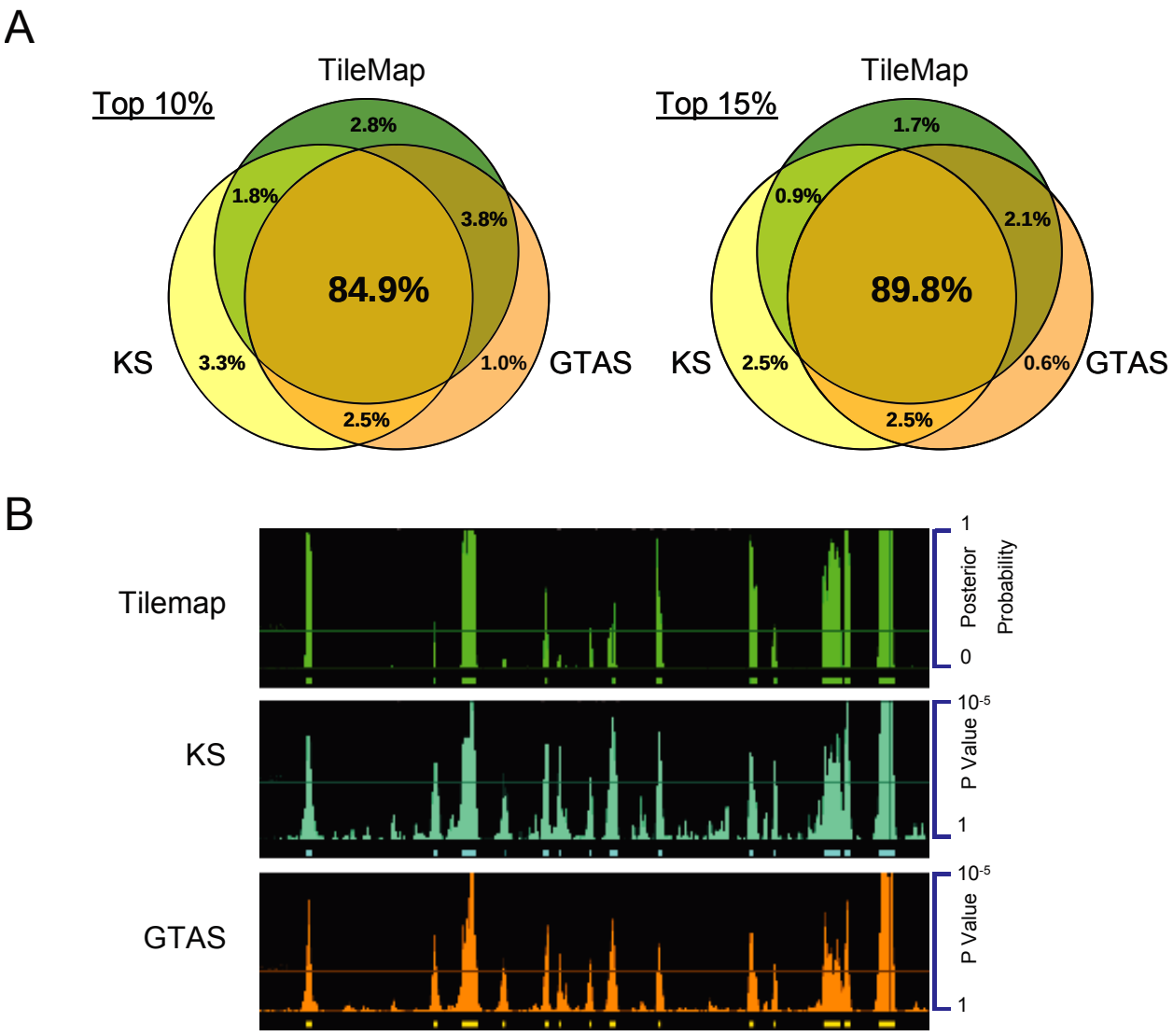

Figure S1. Comparison of Methylated Regions

Comparison of methylated regions detected by a Wilcoxon Signed-Rank Test (GTAS) (Hollander and Wolfe, 1999), a two-state hidden Markov Model (HMM) based on probe-level t-statistics (Tilemap), and a nonparametric Kolmogorov-Smirnov test (KS) (Ji and Wong, 2005).

(A) Overlap of regions with the highest 10% (left) or 15% (right) scores from the three methods.

(B) A region on chromosome 1 (3,988,000–4,045,000) as an example showing the consistency among results from the three methods. Horizontal thin lines in each track represent the 90th percentile cutoffs. Horizontal thick bars at the bottom of each track denote regions that met the cutoffs.

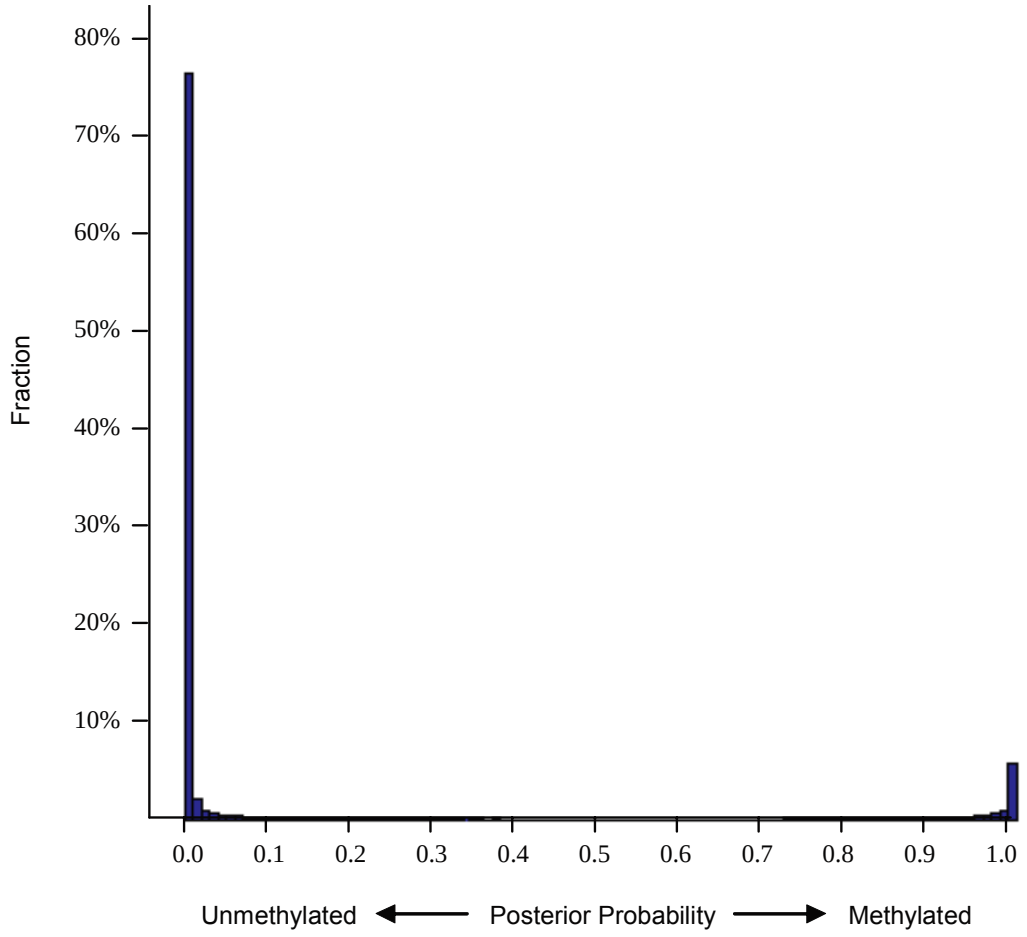

Figure S2. Bimodal Distribution of Posterior Probability Scores Obtained from TileMap with the HMM Option

X-axis: posterior probability scores (0.01 per bin); *y*-axis: fraction of the probes with a given score.

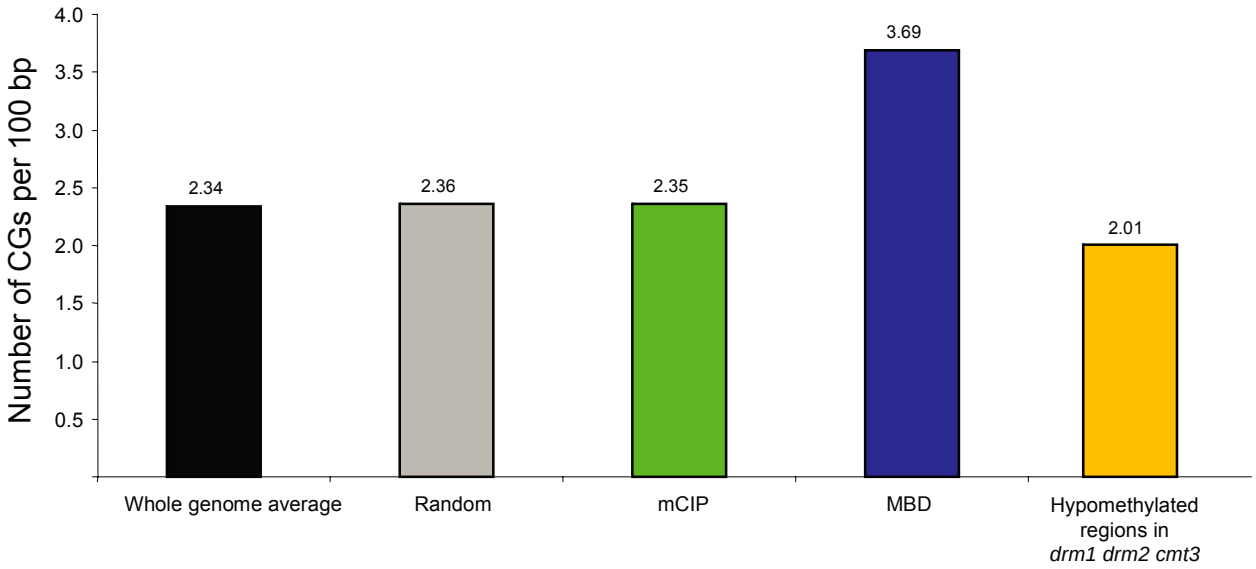

Figure S3. CG Content of Methylated Regions Identified by the mCIP Method (Green) and the HMBD Method (Blue)

The CG content of the entire *Arabidopsis* genome (black) and that of 30,000 random genomic regions (gray) are shown for comparison. Note that the CG content of methylated regions identified by the mCIP method is very close to that of genome average or random regions, whereas the MBD method requires ~1.6 fold higher CG content. The regions that were hypomethylated in *drm1 drm2 cmt3* (yellow) had a lower CG content than average.

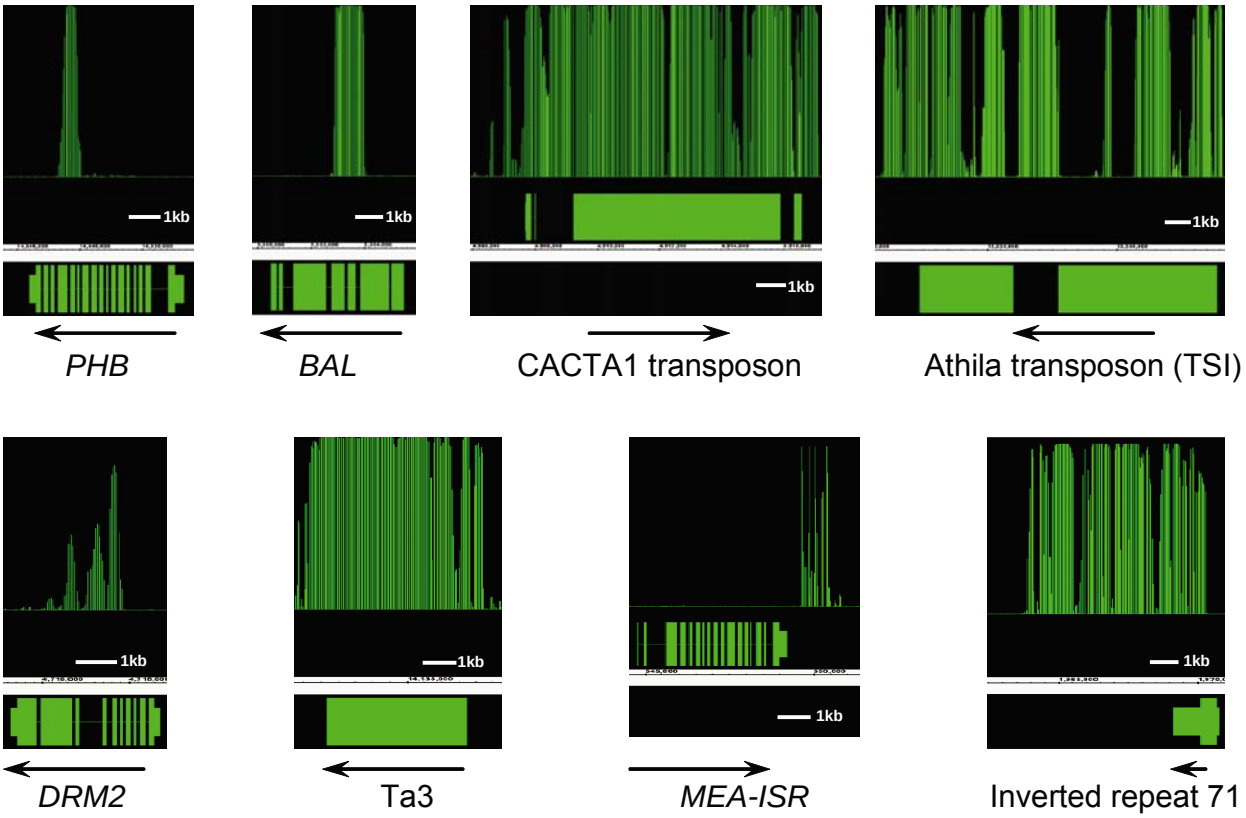

Figure S4. Comparison of mCIP Results and Previously Known Methylated Regions

Each of the methylated regions is detected by the mCIP method. The methylated genes shown are *PHB*, *BAL*, *CACTA1*, *Athila*/*TSI*, *DRM2*, *Ta3*, *MEA-ISR* and Inverted Repeat #71 (Ashapkin et al., 2002; Bao et al., 2004; Cao and Jacobsen, 2002a; Henderson et al., 2006; Miura et al., 2001; Stokes et al., 2002). *PHB* and *DRM2* represent genic DNA methylation (Ashapkin et al., 2002; Bao et al., 2004). *Ta3* is a centromeric retrotransposon (Cao and Jacobsen, 2002b). *MEA-ISR* is a euchromatic tandem repeat locus (Cao and Jacobsen, 2002b). Inverted Repeat #71 is a large inverted repeat (Henderson et al., 2006). Arrows indicate the direction of transcription. Loci from previous studies in the accession Columbia are shown here.

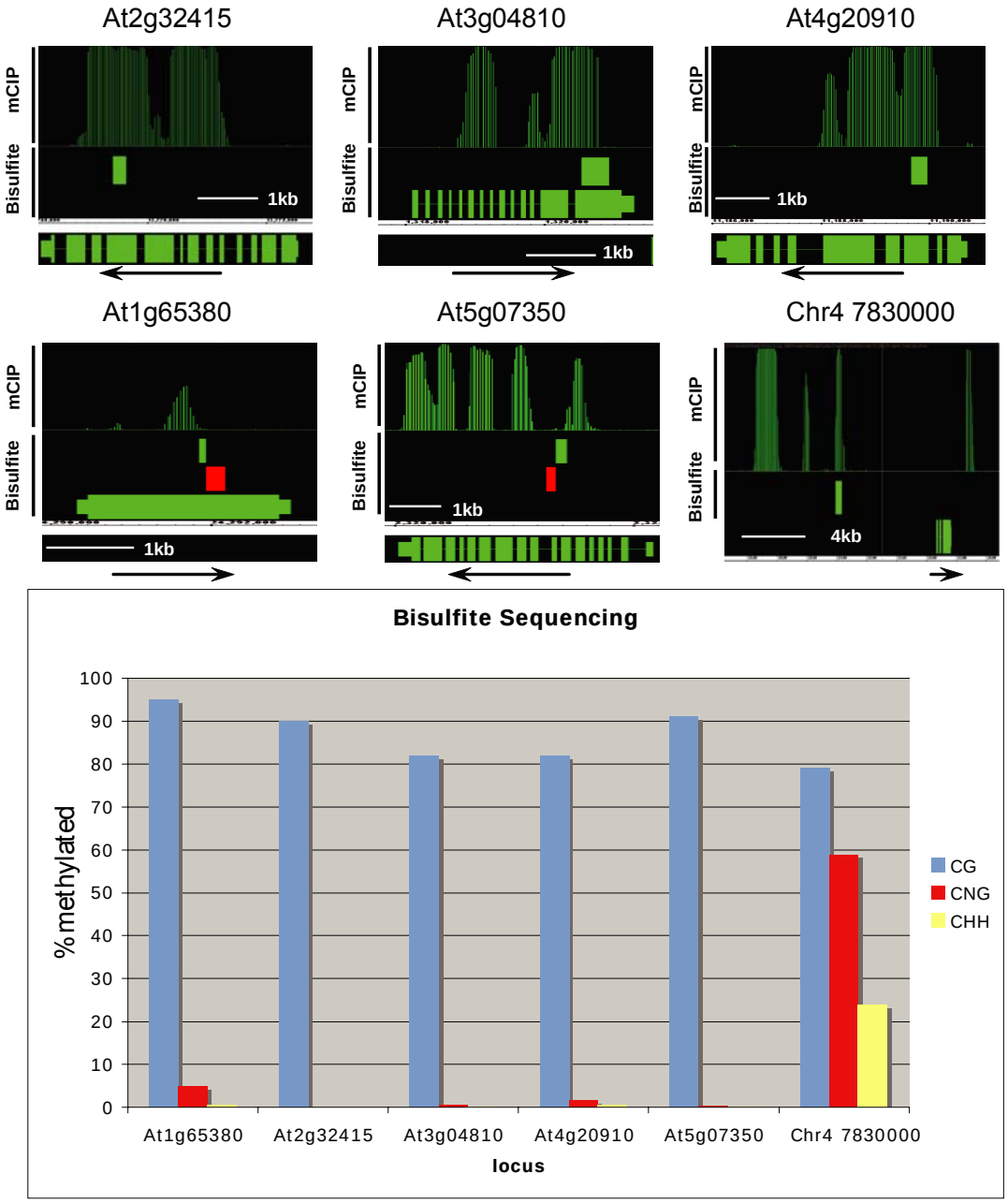

Figure S5A. Bisulfite Genomic Sequencing Validates mCIP Results

We selected five sites of genic methylation (one from each chromosome) and one site with extensive non-CG DNA methylation from chromosome IV. We bisulfite sequenced a 72 bp region from At1g65380, a 206 bp region from At2g32415, a 339 bp region from At3g04810, a 277 bp region from At4g20910, a 227 bp region from At5g07350, and a 236 bp region from chromosome IV at nucleotide 7830000. Regions that were found to be methylated are shown as green boxes and unmethylated as red boxes. The percentage of DNA methylation in CG, CNG and asymmetric (CHH) sequence contexts is shown. Bisulfite sequencing primers are listed in Table S8. Arrows indicate the direction of transcription.

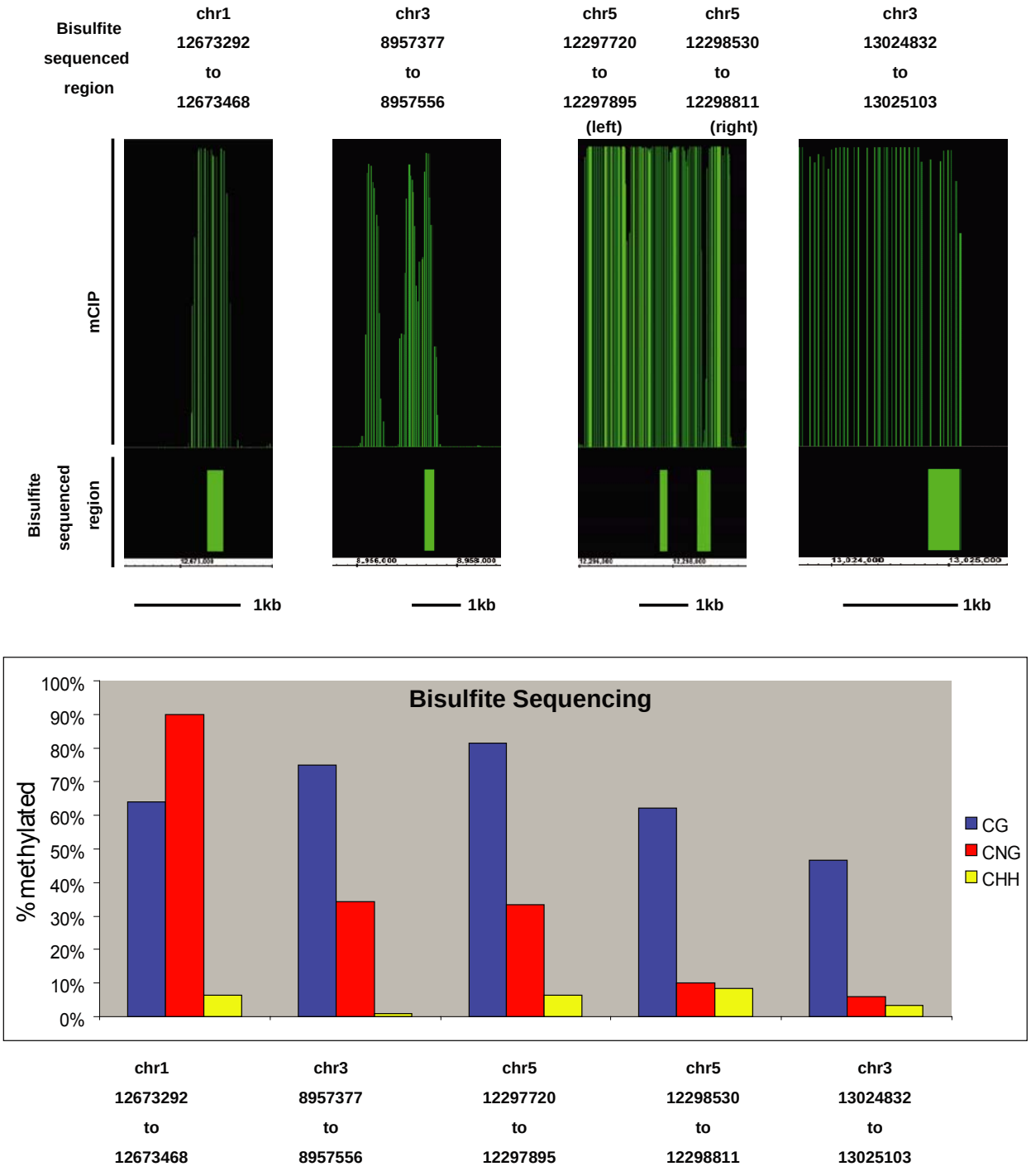

Figure S5B. Bisulfite Genomic Sequencing Validation of the mCIP Results in Intergenic Regions

mCIP results were shown on top. Green boxes represent methylated regions detected by bisulfite sequencing. The results of bisulfite sequencing are shown at the bottom.

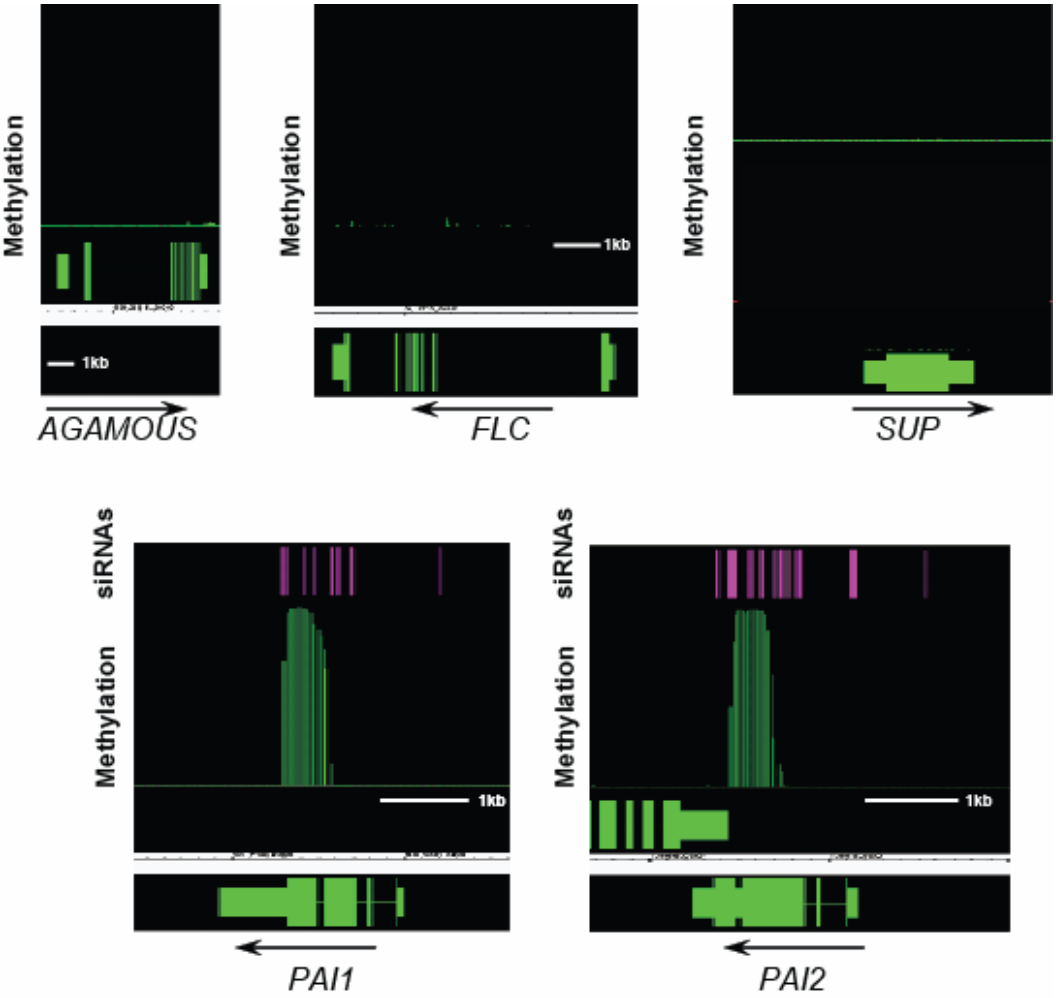

Figure S6. Comparison of mCIP Results to Previously Known Unmethylated Regions in the Columbia Accession

The *AGAMOUS*, *FLC* and *SUP* genes are unmethylated in wild-type plants (Jacobsen and Meyerowitz, 1997; Jacobsen et al., 2000; Jean Finnegan et al., 2005). Note that the promoters of the *PAI1* and *PAI2* genes are unmethylated as previously reported (Luff et al., 1999). However, both genes contain heavy methylation in their transcribed regions, and such methylation is correlated with dense siRNA clusters (purple vertical bars). Arrows indicate direction of transcription.

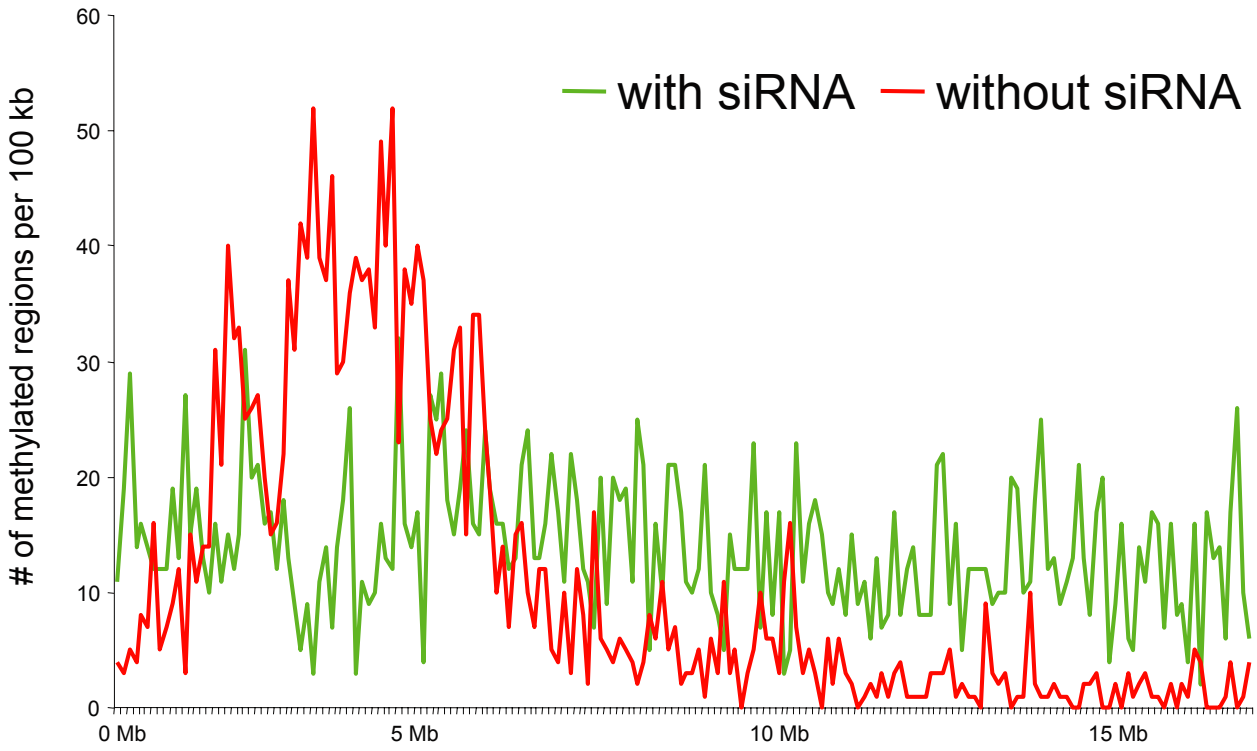

Figure S7. Chromosomal Distribution of DNA Methylation with (red) or without (green) siRNAs

We used an siRNA database compiled by massively parallel signature sequencing (Lu et al., 2005). Chromosome 4 is shown as an example. *Y*-axis: total length of methylated DNA in 100-kb windows. Each vertical bar on the *x*-axis represents a 100-kb region.

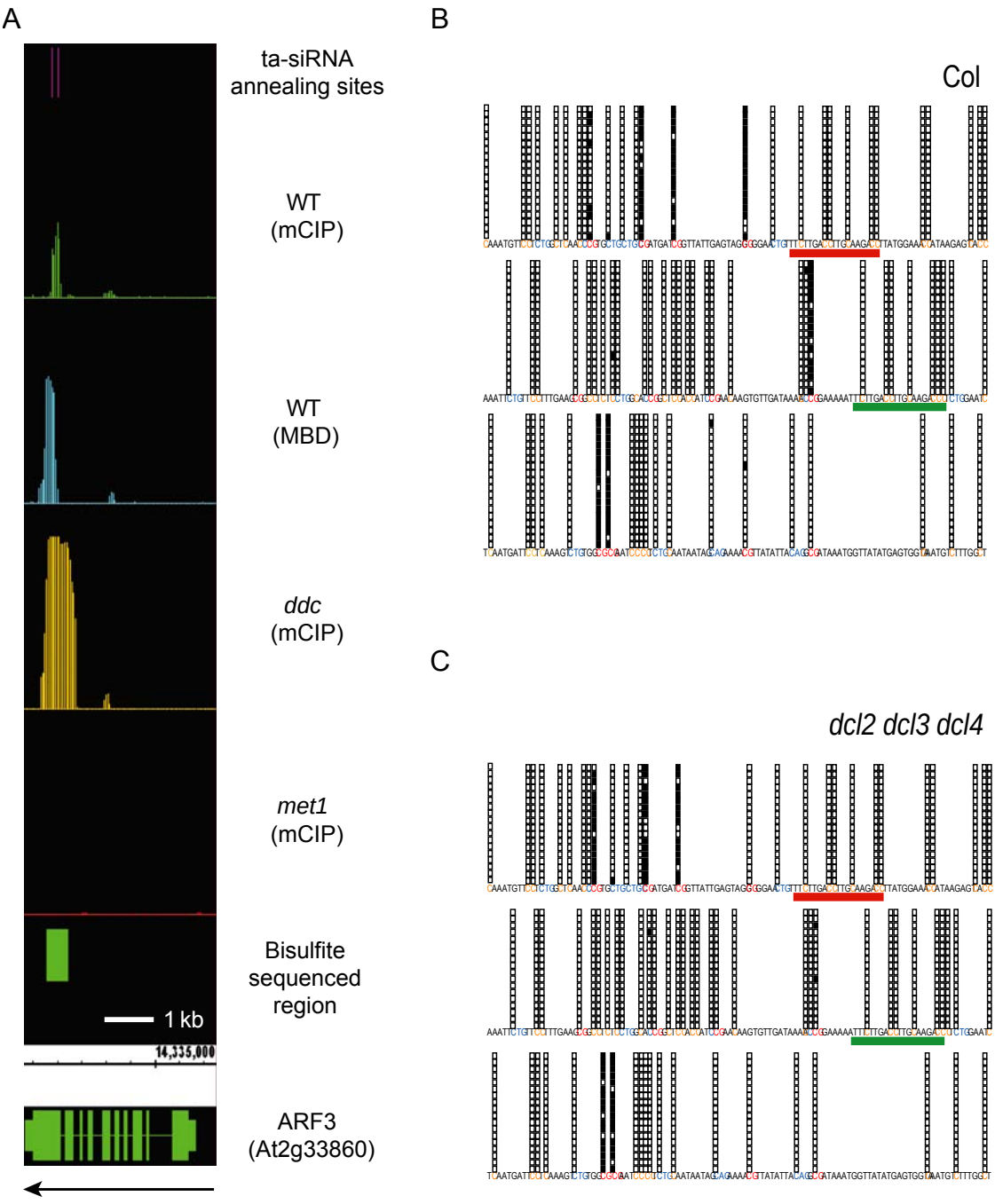

Figure S8. CG Methylation at the tasiRNA Target Locus *ARF3* (At2g33860)

(A) DNA methylation in wild-type (mCIP and MBD methods), *drm1 drm2 cmt3* (mCIP) and *met1* (mCIP). Purple vertical parts indicate the annealing sites of ta-siR2141 (left) and ta-siR2142 (right). Green horizontal bar represent the region subjected to bisulfite sequencing.

(B) Bisulfite sequencing results from wild-type

(C) Bisulfite sequencing results from *dcl2 dcl3 dcl4*. Open boxes represent unmethylated cytosines and filled boxed represent methylated cytosines. The genomic sequence is included (CG shown in red, CNG in blue and CHH in orange). Red and green horizontal bars indicate the annealing sites of ta-siR2141 and ta-siR2142, respectively.

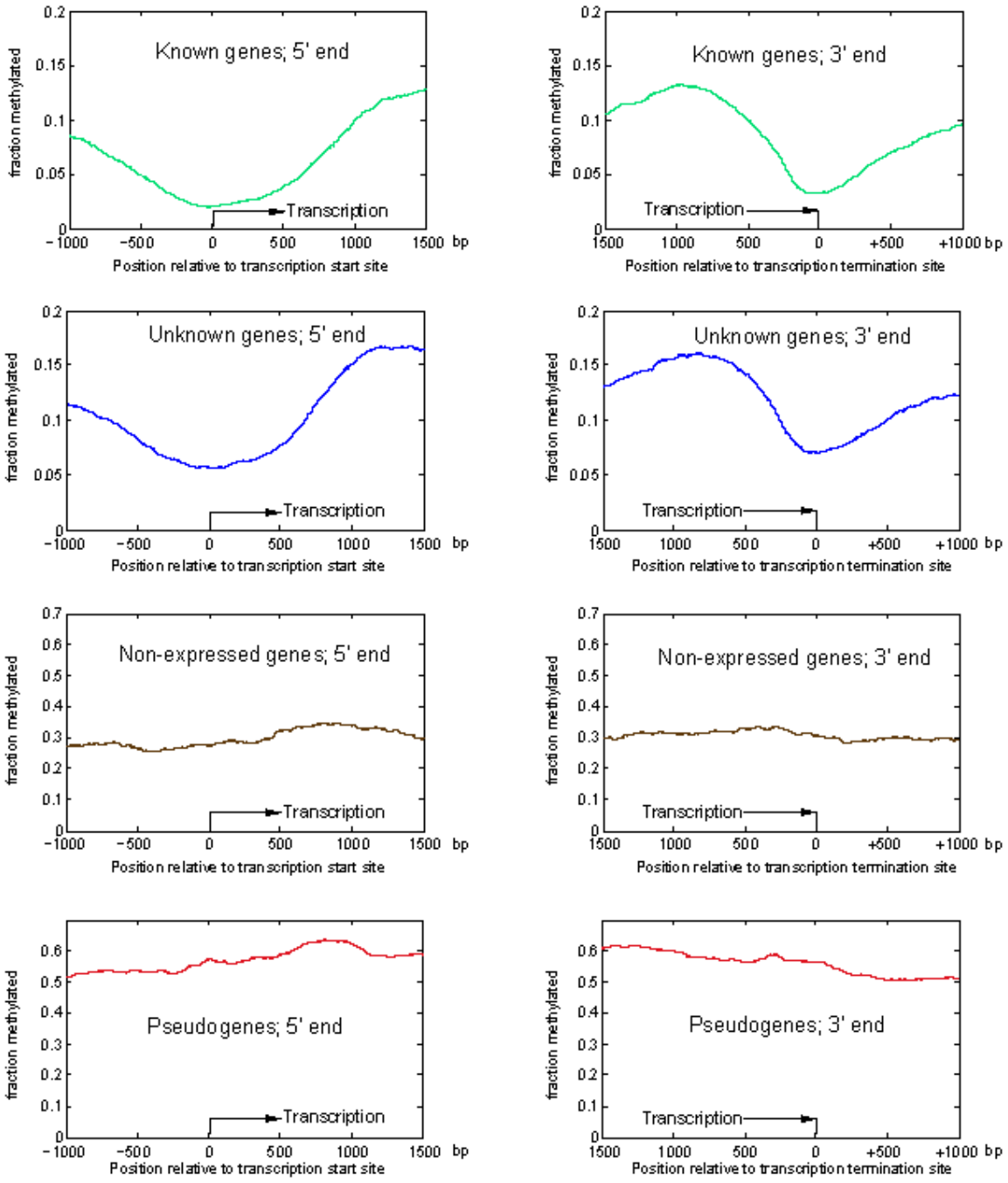

Figure S9. Distribution of DNA Methylation in *Arabidopsis* Genes

For this analysis, the 5' regions of genes including 1 kb upstream and 1.5 kb downstream of transcription start sites (position 0, left panels) were aligned at their transcription start sites, and the 3' regions of genes including 1.5 kb upstream and 1 kb downstream of transcription termination sites (position 0, right panels) were aligned at their transcription termination sites. The distribution of DNA methylation was then determined as the fraction of bases at that position that were in a methylated region (*y*-axis).

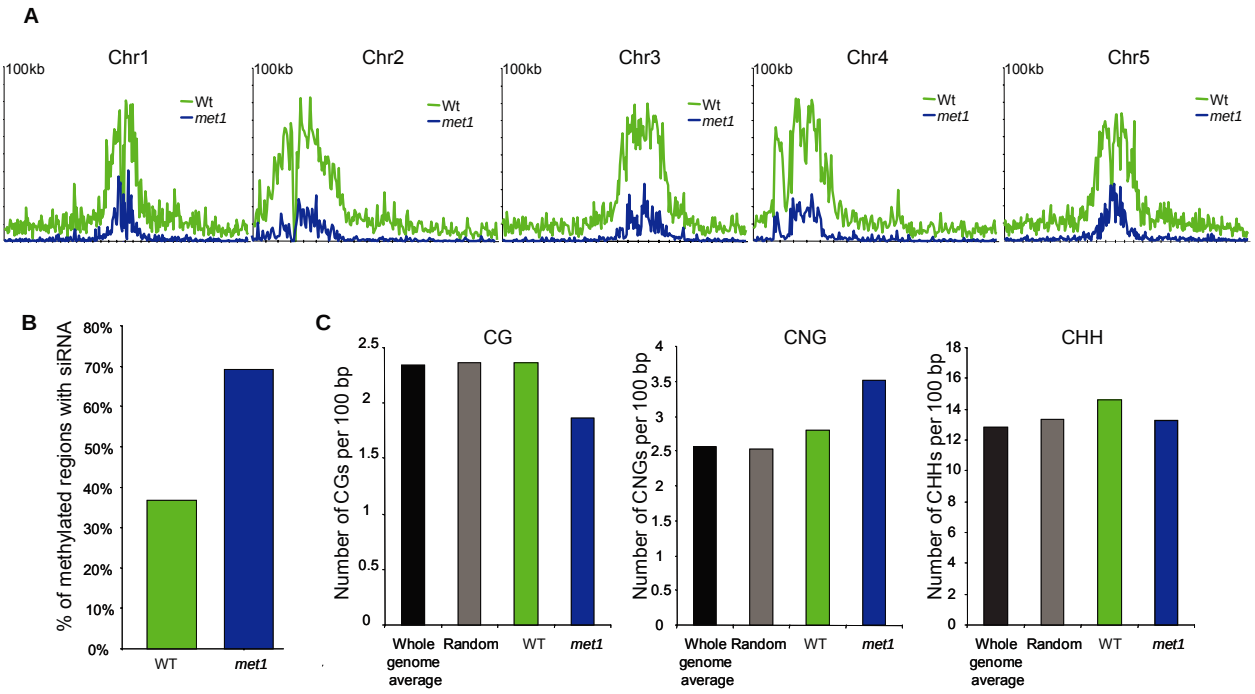

Figure S10. Residual Methylation in *met1*

(A) Chromosomal distribution of residual methylation in *met1* (blue) compared to wild-type (green). Each vertical bar in *x*-axis represents 1Mb. *Y*-axis shows the total length of methylated regions in 100 kb windows.

(B) Percentage of methylated regions associated with siRNA clusters in wild-type (green) or *met1* (blue).

(C) CG, CNG and CHH content in the methylated regions in *met1* (blue) compared to wild-type (green), whole genome average (black) or 30,000 random regions (gray). *Y*-axis: number of CG, CNG or CHH per 100 bp.

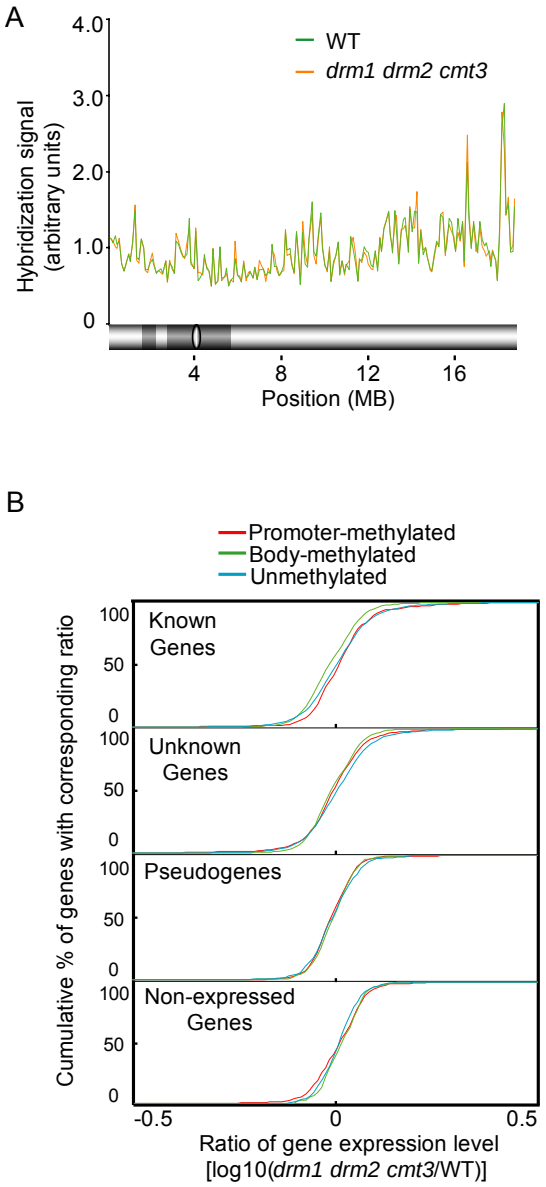

Figure S11. Global Changes of Gene Expression in *drm1 drm2 cmt3* Compared to Wild-Type

(A) Chromosomal-level transcriptional activity in wild-type (green) and *drm1 drm2 cmt3* (orange). Chromosome 4 is shown as an example. *Y*-axis: median probe-level intensity over 100-kb windows. A schematic representation of chromosome 4 is shown at the bottom and labeled as in Figure 1.

(B) Global changes of gene expression in *drm1 drm2 cmt3* compared to wild-type, shown as cumulative distributions. *X*-axis: gene-level expression fold change in *drm1 drm2 cmt3* compared to wild-type (log10 scale); *y*-axis: cumulative percentage of genes in each category with corresponding ratio.

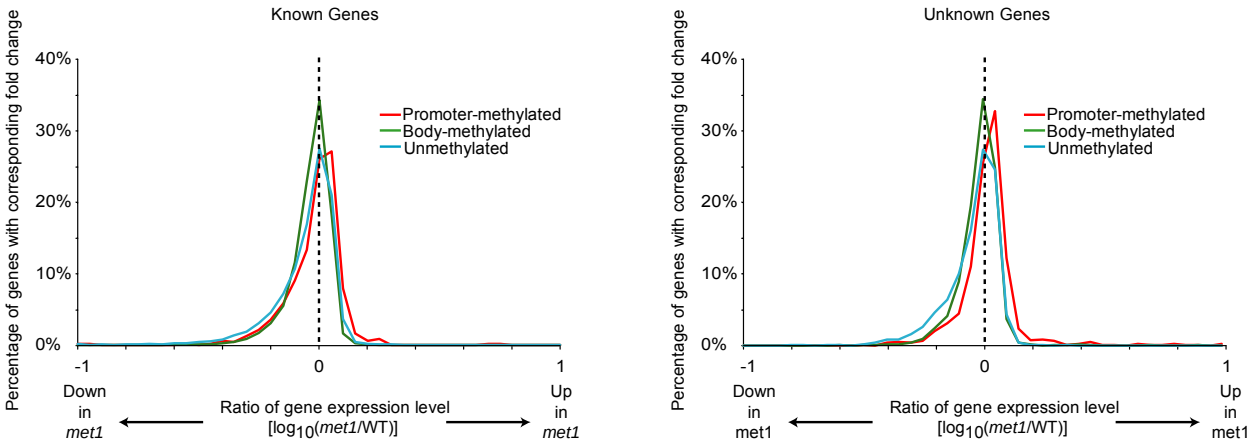

Figure S12. Body-Methylated and Unmethylated Genes Are Not Systematically Upregulated in *met1*

Known genes (left) and unknown expressed genes (right) are shown. *X*-axis: gene-level expression fold change in *met1* compared to wild-type (log10 scale). *Y*-axis: the fractionof genes with corresponding fold change.

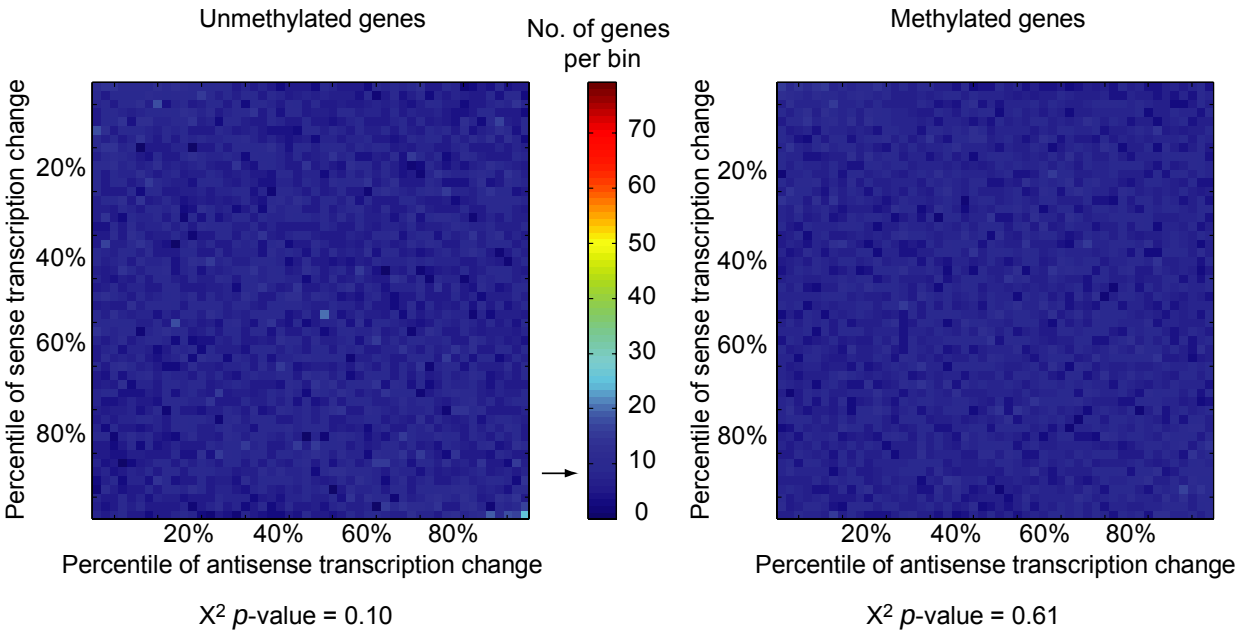

Figure S13. 2-D Histograms Showing that Changes in Sense and Antisense Transcription Are Largely Independent

For this analysis, the ratio of sense transcription level in *met1* compared to wild-type was divided into 50 bins, such that each bin has the same number of corresponding genes (i.e., each bin contains 2% of genes). The ratio of antisense transcription in *met1* compared to wild-type was divided in the same way. Next, each gene was assigned to one of the 2,500 regions defined by the 50 sense transcription ratio bins and the 50 antisense transcription ratio bins. Therefore, if there is a strict positive correlation between sense and antisense transcription changes, all genes should be assigned to the 50 diagonal regions from the lower-left corner to the upper-right corner. If there is a strict negative correlation, all genes should be assigned to the 50 diagonal regions from the upper-left corner to the lower-right corner. Finally, if there is no correlation (i.e., sense and antisense changes are statistically independent), the distribution of genes in the 2,500 regions should be approximately binomial in each square with an average indicated by the arrow on the color scale. The observations are consistent with independence (X2 p-values are shown). These results show that across all levels of gene expression, there is no significant correlation in the changes of sense and antisense transcription in *met1*.

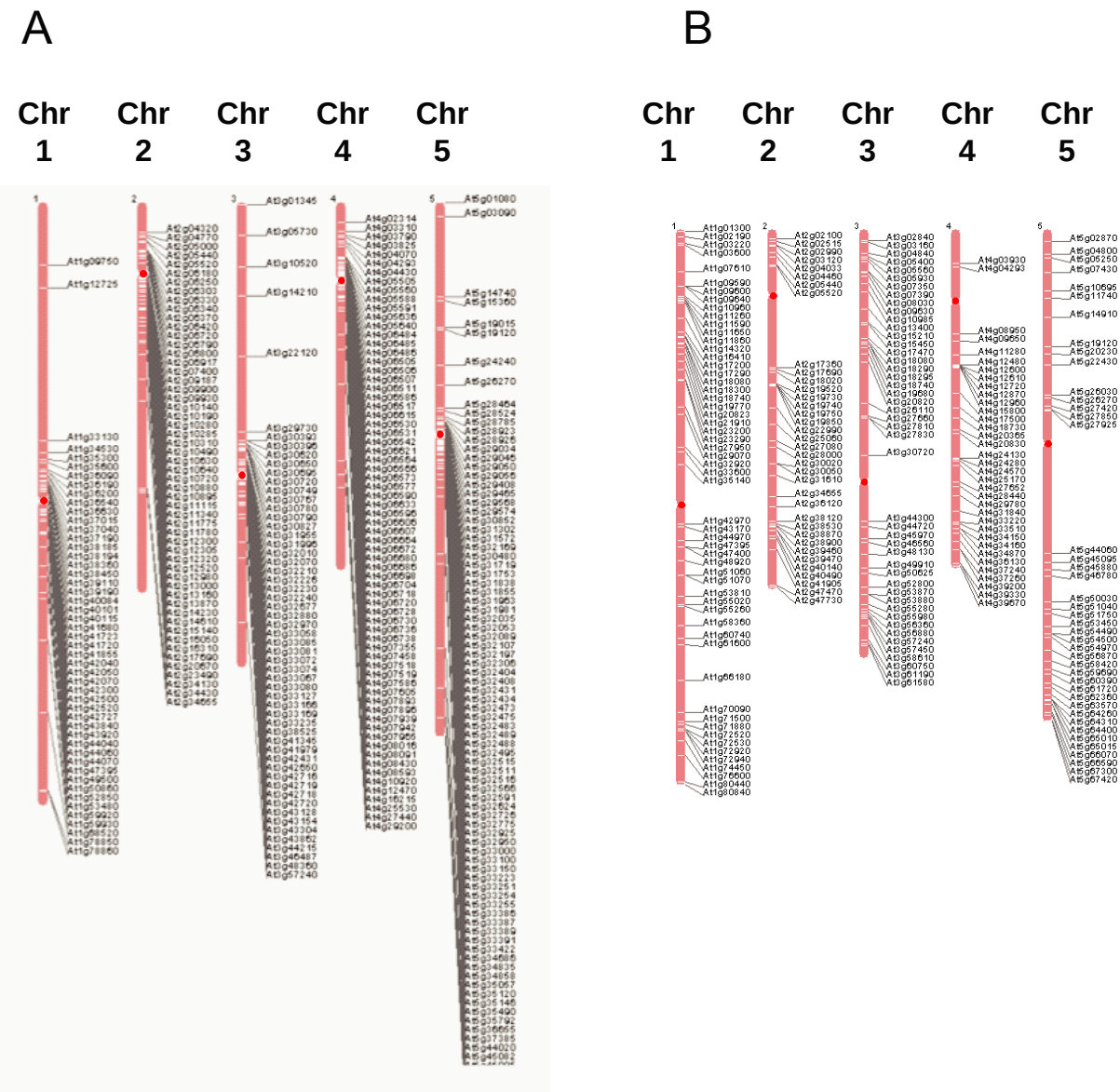

Figure S14. Chromosomal Distribution of Genes that Are Significantly Upregulated in *met1* and *drm1 drm2 cmt3*

See Tables S2 and S3 for lists. Red circles represent centromeres.

(A) *met1*.

(B) *drm1 drm2 cmt3*.

Supplemental Tables

Table S1. Gene Ontology (GO) Analysis of Promoter-Methylated, Body-Methylated, and Unmethylated Genes

Table S2. List of Genes Overexpressed in *met1*

Table S3. List of Genes Overexpressed in *drm1 drm2 cmt3*

Table S4. List of Noncoding RNAs Overexpressed in *met1* or *drm1 drm2 cmt3*

Table S5. List of Noncoding RNAs Whose Expression Is Suppressed in *met1*

Table S6. RT-PCR Primers for Validation of Overexpressed Genes

Table S7. RT-PCR Primers for Validation of Overexpressed Intergenic Noncoding RNAs

Table S8. PCR Primers for Bisulfite Genomic Sequencing Validation of Methylated Regions

Table S9. Reproducibility of the Three Biological Replicates for Each DNA Methylation Experiment