学位論文

Genome-wide negative feedback drives transgenerational DNA methylation dynamics in Arabidopsis

(シロイヌナズナの継世代的 DNA メチル化動態における

ゲノムワイドの負のフィードバック)

平成26年12月博士 (理学) 申請

東京大学大学院理学系研究科

生物科学専攻

伊藤　佑

Table of contents

Acknowledgements

First, I would like to express my special gratitude to my supervisor, Dr. Tetsuji Kakutani for providing me this precious opportunity as a graduate student in his laboratory. I would like to express my gratitude to Dr. Yoshiaki Tarutani for continuing cooperation in whole genome bisulfite sequencing. I would like to express my gratitude also to Dr. Taiko Kim To for cooperation in analysis of *ddm1* mutants. I would like to thank the other members of Kakutani group. Especially, I really appreciate the technical assistances by Mr. Kazuya Takashima and Ms. Akiko Terui.

I would like to express my appreciation to Dr. Hidetoshi Saze for cooperation in analyses of *ibm1* mutants.

I am very grateful to Mr. Mohamed Kassam, Ms. Evelyne Duvernois-Berthet, Dr. Sandra Cortijo and Dr. Vincent Colot for their cooperation in analyses of epiRILs.

I am very grateful also to Dr. Asao Fujiyama, Dr. Atsuhi Toyoda, and the other members of Fujiyama group for their cooperation in sequencing.

Abbreviations

ChIP	chromatin immunoprecipitation
DNA	deoxyribonucleic acid
DMR	differently methylated region
epiRIL	epigenetic recombinant inbred line
H3K9	lysine 9 of histone H3
LINE	long interspersed nuclear element
LTR	long terminal repeat
MeDIP	methylated DNA immunoprecipitation
mRNA	messenger RNA
PCR	polymerase chain reaction
RdDM	RNA directed DNA Methylation
RNA	ribonucleic acid
SINE	short interspersed nuclear element
TE	transposable element
TEG	transposable element gene
WGBS	whole-genome bisulfite sequencing

Abstract

Epigenetic modifications of chromatin, especially DNA methylation, can be inherited over multiple generations in plants. However, the regulatory mechanisms for the transgenerational dynamics of DNA methylation pattern are still unclear. In order to examine the long-term effect of impaired DNA methylation pattern over multiple generatins, the genome-wide DNA methylation in the mutants of the chromatin remodeler gene *DDM1* (*Decrease in DNA Methylation 1*) in *Arabidopsis thaliana* was analyzed using whole genome bisulfite sequencing. The *ddm1* mutation induces a drastic decrease in DNA methylation of transposable elements (TEs) and repeats in the initial generation, while also inducing ectopic DNA methylation at hundreds of loci. Unexpectedly, this ectopic methylation can only be seen after transgenerational propagation. This ectopic cytosine methylation is found primarily in the non-CG context and starts from the 3' regions of transcription units and spreads upstream. Remarkably, when chromosomes with reduced DNA methylation were introduced from a *ddm1* mutant into a *DDM1* wild-type background through genetic crosses, de novo accumulation of DNA methylation could also be induced in trans on regions of wild type origin. These results suggest that ectopic DNA methylation would be a consequence of redistribution of DNA methylation through genome-wide negative feedback mechanism. This global negative feedback, together with local positive feedback, would ensure robust and balanced differentiation of chromatin states within the genome.

Chapter 1

Preface

Chemical modifications of chromatin, such as methylation of DNA and methylation or acetylation of histone proteins, have important roles in both activation and repression of gene and transposable elements (TEs) in the eukaryotic genome (Saze & Kakutani, 2011). These modifications are called “epigenetic” marks, because the active and inactive chromatin states can be inherited across cell division without any changes in DNA sequences. The genomic pattern of the epigenetic modifications is implicated in maintenance of differentiation during development (Meissner, 2010).

Furthermore, changes of gene expression associated with epigenetic modifications can be inherited through multiple generations. Such transgenerational inheritance of epigenetic variation is found both in animals and plants (Kelly et al., 2014; Heard & Martienssen, 2014). Especially in plants, epigenetic inheritance of DNA methylation is found widely and could have significant impact on the evolution (Kakutani 2002; Richards, 2011; Becker & Weigel, 2012). However, it still remains to be elucidated how DNA methylation is regulated in long-term.

Recently, the long-term dynamics of genome-wide DNA methylation has been analyzed at single base resolution in the flowering plant *Arabidopsis thaliana* using repeatedly self-pollinated wild type plants. Between ancestors and descendants of the self-pollinated line, there were regions that showed heritable gain and loss of DNA

methylation, although the frequencies were generally low (Schmitz et al., 2011; Becker et al., 2011). This observation implies that the background regulatory mechanisms would be dynamic, rather than static. However, it is difficult to understand the mechanisms using only wild-type plants, because the changes of DNA methylation are small and rare. A complementary approach for understanding it would be to analyze mutants of regulatory factors for DNA methylation over multiple generations.

Factors controlling genomic DNA methylation have been studied extensively in Arabidopsis; and many of these factors constitute positive feedback loops to stabilize epigenetic states. Cytosine methylation in the context of dinucleotide CG is maintained by maintenance methyltransferase MET1 (Finnegan et al., 1996; Kankel et al., 2003), while cytosine methylation at non-CG sites is mediated by chromomethylases (CMTs) (Zemach et al., 2013; Stroud et al., 2014). The CMTs are recruited to chromatin by di-methylation of histone H3 lysine 9 (H3K9me2), and the H3K9 methylase KYP/SUH4 is also recruited to chromatin with non-CG methylation, generating a self-reinforcing positive feedback loop (Johnson et al., 2007; Inagaki et al., 2010; Du et al., 2012; Stroud et al., 2014). Both H3K9me2 and non-CG methylation are silent heterochromatin marks normally found in repeats and transposable elements (TEs); and these marks are rarely detectable in transcribed genes. Exclusion of these marks from transcribed genes depends on the H3K9 demethylase IBM1 (Increase in BONSAI Methylation 1) (Saze et al., 2008; Inagaki et al., 2010). IBM1 removes H3K9me2 from transcribed genes, thus generating another positive feedback loop to stabilize active states (Inagaki et al., 2010).

Another positive feedback loop is also found in RNA-directed DNA methylation (RdDM), a de novo DNA methylation process triggered by double-strand RNA; and factors involved in this process have been extensively studied (Mette et al., 2000; Cao et al., 2003; Law & Jacobsen, 2010; Furner & Matzke, 2011; Pikaard et al., 2012). The final step of RdDM is the methylation of both CG and non-CG sites by the de novo DNA methyltransferase DRM2 (Domains Rearranged Methylase 2), with RNAi machinery to generate small interfering RNA (siRNA) functioning as upstream factors. Interestingly, production of siRNA also depends on *DRM2* (Zilberman et al., 2004; Henderson et al., 2010), suggesting another positive feedback between DNA methylation and siRNA production. Genome-wide DNA methylation profiles have been determined in mutants for these and other factors controlling DNA methylation (Zhang et al., 2006; Stroud et al., 2013; Stroud et al., 2014), although information on the transgenerational effects of these mutations is limited.

In this study, I examined the effects of impaired DNA methylation pattern in Arabidopsis mutants over multiple generations to uncover the background mechanisms controlling long-term DNA methylation dynamics.

Chapter 2

Materials and Methods

Plant materials and annotations

Isolation of the *ddm1-1* and *ibm1-4* mutants has been described previously (Vongs et al., 1993; Saze et al., 2008). Self-pollinations of *ddm1* lines were described previously (Kakutani et al., 1996). In order to remove heritable effects of the *ddm1* mutation, the original *ddm1* mutant was backcrossed six times in the heterozygous state. The heterozygous plants were propagated by self-pollination. 1G *ddm1* mutant plants were selected from self-pollinated progeny of the heterozygote. 9G *ddm1* plants were generated by independently self-pollinating different *ddm1* segregants eight times (Figure 1). The annotations of genes and TEs are based on The Arabidopsis Information Resource (http://www.arabidopsis.org/). TAIR8 was used for analyzing ChIP chip data (Figure 4D), TEG (TE gene) data, and epiRILs data. TAIR10 was used for other analyses. The details of the annotation of TEGs were described in a document in TAIR (ftp://ftp.arabidopsis.org/home/tair/Genes/TAIR8_genome_release/Readme-transposons).

DNA methylation analyses

For the 1G and 9G *ddm1* plants and their controls, genomic DNA was isolated from rosette leaves using the Illustra Nucleon Phytopure genomic DNA extraction kit

(GE Healthcare), and genome-wide bisulfite sequencing was performed as described previously (Fu et al., 2013). Raw sequence data were deposited in the DDBJ (DNA Data Bank of Japan) Sequence Read Archive (DRA; accession nos. DRA002545, DRA002546, DRA002548, DRA002549, DRA002551, DRA002554, DRA002555, DRA003018, DRA003019 and DRA003020). The adaptor sequences were clipped out using the FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Reads were trimmed to 90 nucleotide length (45 nucleotide for the data obtained from GEO - GSE39901) and mapped to reference genomes (Release 10 of the Arabidopsis Information Resources) using the Bowtie alignment algorithm (Langmead et al., 2009) with the following parameters, "-X 500 -e 90 -l 20 -n 1". Only uniquely mapped reads were used. Clonal reads were removed except one with the best quality. Any read with three consecutive methylated CHH sites were eliminated. The level of methylation of cytosine in a genomic region was calculated using the ratio of the number of methylated cytosine to that of total cytosine. For the three epiRILs and two parental lines, whole-genome bisulfite sequencing was described previously (Colomé-Tatché et al., 2012) and the data are in GEO (GSE62206).

DMRs (differentially methylated regions) were defined by comparing the methylation level of 100-bp windows throughout the genome between two genotypes. The windows with at least 20 cytosines sequenced were used for the comparison. The level of methylation was calculated using the weighted methylation level of each genotype (Schultz et al., 2012). The windows with difference of methylation 0.5 and 0.3 for CG and CHG, respectively, were selected as DMRs. For defining contiguous DMR

(cDMR), multiple DMRs were merged if they were adjacent to each other or there was only one gap of the 100-bp window. The centroid of methylation of cytosine in cDMR was calculated using the relative position and the methylation level of each cytosine included in the cDMR that was at least 500 bp in length and overlapping with genes. Each contiguous DMR was aligned according to the orientation of the corresponding gene. To plot DNA methylation patterns over genes or TEGs in *ddm1* mutants, #1 samples of each genotype (Figure 3, Figure 5) in 1G *ddm1* and 9G *ddm1* were used. To draw the heatmap of methylation of cytosine, cluster 3.0 (de Hoon et al., 2004) and Java Treeview (Saldanha et al., 2004) were used.

Processing ChIP-seq data

ChIP-seq data of various histone modifications (Luo et al., 2012) in GEO (GSE28398) were used for the analysis. The coordinates were remapped onto TAIR10 annotation using a script in TAIR (Lamesch et al., 2012). Enrichment of histone modification in a DMR was calculated by the density of ChIP-seq reads, and normalized by the mean and the standard deviation of the density of reads in 100,000 windows randomly chosen across the genome.

Processing MeDIP-chip data of epiRILs

The MeDIP-chip data of 123 epigenetic recombinant inbred lines (epiRILs), *ddm1* and WT are in GEO (GSE37284). The regions that were methylated (M) in WT and unmethylated (U) in *ddm1* were selected as targets of *ddm1* mutation using the

values for HMM (hidden Markov model) status (M (methylated) or I (Intermediate) or U (Unmethylated)) (Colomé-Tatché et al., 2012). Global hypo-methylation index of an epiRIL was calculated as the genome-wide average of the values for HMM status of probes on the chip (M=0, I=0.5, U=1) in the target regions of *ddm1* mutation. The data of inference of inherited haplotypes were shown in the previous study (Colomé-Tatché et al., 2012). Following are the names of the lines numbered 1-6 in Figure 22 and Figure 24-29. (Figure 22AB, Figure 24) epiRIL208 epiRIL122 epiRIL98 epiRIL232 epiRIL70 epiRIL114; (Figure 22CD, Figure 25) epiRIL122 epiRIL208 epiRIL114 epiRIL258 epiRIL438 epiRIL508; (Figure 22EF, Figure 26) epiRIL208 epiRIL98 epiRIL438 epiRIL508 epiRIL122 epiRIL114; (Figure 27) epiRIL208 epiRIL73 epiRIL71 epiRIL394 epiRIL98 epiRIL438; (Figure 28) epiRIL508 epiRIL114 epiRIL122 epiRIL438 epiRIL208 epiRIL93; (Figure 29) epiRIL208 epiRIL114 epiRIL556 epiRIL71 epiRIL244 epiRIL98.

Chapter 3

Transgenerational effect of *ddm1* mutation

Introduction

Among the Arabidopsis mutants affecting genomic DNA methylation, *ddm1* (*decrease in DNA methylation 1*) is one of the mutations with the strongest effects. The mutant plants show drastic reduction of DNA methylation at both CG and non-CG sites in repeats and TEs (Vongs et al., 1993; Lippman et al., 2004). The *DDM1* gene encodes a chromatin remodeling factor, which is necessary for DNA methylation in heterochromatic sequences (Jeddeloh et al., 1999; Zemach et al., 2013). Mutation in its mammalian ortholog *Lsh* also induces loss of DNA methylation predominantly at repeat sequences, suggesting conserved functions across the animal and plant kingdoms (Dennis et al., 2001; Tao et al., 2011).

A striking feature of the Arabidopsis *ddm1* mutant is the progressive accumulation of the developmental phenotypes; initial generations of the *ddm1* mutant grow relatively normally, but many types of developmental abnormalities arise after multiple rounds of self-pollinations (Kakutani et al., 1996; Kakutani., 1997). Some of the abnormalities are due to DNA sequence changes, such as insertion mutations of de-repressed endogenous TEs (Miura et al 2001; Singer et al 2001; Tsukahara et al 2009) or rearrangement of repeats (Yi & Richards, 2009), but others are due to

epigenetic changes in gene expression, which correlates with changes in DNA methylation pattern at the affected loci (Soppe et al., 2000; Saze & Kakutani, 2007).

To explore mechanisms controlling long-term DNA methylation dynamics, it is important to compare DNA methylation changes in initial and subsequent generations of *ddm1* mutant plants. Here I analyze the transgenerational effects of the *ddm1* mutation genome-wide, by comparing DNA methylation of the *ddm1* mutants before and after the repeated self-pollinations.

Results

First and subsequent generations of *ddm1* mutants show distinct genomic DNA methylation patterns

I examined DNA methylation in four individuals of *ddm1* homozygous mutants segregated in progeny of a heterozygote (hereafter called 1G for the 1st Generation) and also four lines of *ddm1* plants independently self-pollinated eight times (hereafter called 9G) (Figure 1). In 1G, the *ddm1* mutation already induced reduction of DNA methylation in heterochromatic regions (Vongs et al., 1993; Lippman et al., 2004; Zemach et al., 2013). I separately analyzed normal genes and genes overlapping with annotated transposable elements (hereafter called TE genes, or TEGs; details in Materials and Methods section). Methylation in TEGs (Figure 2D-F) was more severely affected than that in normal genes (Figure 2A-C). The reduction was found for both CG sites (Figure 2D) and non-CG sites. In non-CG sites, both CHG sites (Figure 2E) and CHH sites (Figure 2F) were affected (H can be A, T, or C). These observations in 1G are consistent with previous reports (Vongs et al., 1993; Lippman et al., 2004; Zemach et al., 2013). In 9G, two features were noted: further decrease of CG methylation and an increased methylation at non-CG sites (Figure 2).

Progressive reduction of CG methylation in the self-pollinated *ddm1* lines

Although the *ddm1* mutation immediately induces a drastic loss of DNA methylation in repeats, further reduction of methylation in later generations has been

reported for a few CG sites (Kakutani et al., 1996). Consistent with this, the genome-wide analysis in this study revealed that many loci behave in a similar manner (Figure 3). The progressive reduction of DNA methylation can affect the developmental phenotypes; for example, the promoter of the imprinted gene *FWA* remains methylated in the 1G *ddm1* but the methylation is lost stochastically in 9G *ddm1* (Figure 4A), generating heritable epialleles that cause late-flowering phenotype (Kakutani, 1997; Soppe et al., 2000; Kinoshita et al., 2007). The progressive reduction is seen genome-wide for both genes and TEGs (Figure 2A, Figure 2D).

To compare properties of the regions hypomethylated immediately and gradually, I detected differentially methylated regions (DMRs; details in Materials and Methods). The regions *ddm1* affects immediately (WT-1G DMRs) were enriched in dimethylation of histone H3 lysine 9 (H3K9me2) (Figure 4C left, Figure 4D). H3K9me2 is a mark of silent heterochromatin, and these results are consistent with previous reports (Lippman et al., 2004; Zemach et al., 2013). In marked contrast, however, regions affected later (9G-specific DMRs) have much lower level of H3K9me2 in wild type (Figure 4C middle). *DDM1* gene function is necessary for CG methylation in heterochromatin, but *DDM1* also has significant effects on CG methylation in less heterochromatic regions in the long-term.

Accumulation of non-CG methylation in *ddm1* lines after propagation by self-pollination

More surprisingly, in the self-pollinated *ddm1* lines, non-CG methylation was

increased in a large number of genes and TEs (Figure 5, Figure 7-8). The regions CHG hypermethylated also showed hypermethylation at CHH sites (Figure 6A, Figure 7). In addition, while genic CG methylation tend to decrease progressively from 1G to 9G on average (Figure 2A), non-CG hypermethylated regions show an increase in CG methylation (Figure 6A). The CG and non-CG hypermethylation was found reproducibly at specific loci (Figure 9). The affected loci include *BONSAI* and other sequences that were reported previously (Saze & Kakutani, 2007; Sasaki et al., 2012) but the majority of the affected loci can only be revealed by whole-genome bisulfite sequencing (WGBS), because WGBS enables increased non-CG methylation to be detected with high sensitivity even at loci already CG methylated. In addition to normal genes, a large number of TEGs showed increase in non-CG methylation (Figure 2E-F, Figure 5, Figure 8, Figure 10-13).

A very unexpected feature revealed by WGBS is that non-CG hypermethylation of genes is almost undetectable in the first generation of *ddm1* but is specifically and reproducibly seen in the self-pollinated *ddm1* lines. In Figure 5A and 5B, many black dots can be seen along the vertical axis in the panels for 9G but not for 1G. Hypermethylation is therefore not a simple extension of the effect seen in the first generation. This feature can only be detected in later generations (Figure 6B). In order to further understand the transgenerational dynamics, I examined four independently self-pollinated 2G *ddm1* plants. If the hypermethylation proceeds equally at each self-pollination, the increase from 1G to 2G would be 1/8 of the increase from 1G to 9G. Interestingly, although hypermethylation proceeded in 2G, the difference between 1G

and 2G was much less than 1/8 of that between 1G and 9G, suggesting that the increase is slow initially but accelerated in later generations (Figure 14-16).

Properties of loci hypermethylated in the self-pollinated *ddm1*

Increased non-CG methylation has been reported in mutants of the CG methyltransferase gene *MET1* (Jacobsen et al., 1997; Kishimoto et al., 2001; Mathieu et al., 2007), which results at least in part from a reduction of full-length *IBM1* transcript (Rigal et al., 2012). The *IBM1* gene encodes a demethylase for histone H3K9; and mutation in this gene induces accumulation of H3K9me2 and non-CG methylation in gene bodies. Interestingly, developmental phenotypes of the *ibm1* mutation also become progressively stronger during self-pollinations (Saze et al., 2008). I compared the regions of non-CG hypermethylation in the *ibm1* and self-pollinated *ddm1*. Although an overlap can be detected, the majority of the DMRs in *ddm1* mutants before and after the self-pollinations were distinct from the DMRs of *ibm1* mutants (Figure 17). Just as progressive loss of CG methylation in the *ddm1* mutant, *ibm1* mutant shows progressive accumulation of non-CG methylation in later generations (Figure 18-19). This is consistent with a recent report (Coleman-Derr & Zilberman, 2012) and likely accounts for the progressive developmental defects in the *ibm1* mutant.

In addition, compared to the *ibm1* mutant, the peak in the *ddm1* was shifted toward 3' end, for both gene and TEG transcription units (Figure 19). Furthermore, when hypermethylation levels varied at a given locus among 9G *ddm1* plants, plants with stronger signals tended to show relative centroid positions more upstream than

plants with weaker signals, suggesting that the signal spreads from 3' to 5' (Figure 7, Figure 20).

The bias of the hypermethylation signal toward the 3' region in 9G *ddm1* is especially evident in the hypermethylated TEGs; the peak was often located outside of the transcription unit for both CHG and CHH methylations (Figure 19, bottom half). When different families of TEs are compared, the peak in the downstream region was especially evident in the LTR retrotransposon gypsy (Figure 11-12). Generally, these TEs lost DNA methylation in 1G *ddm1*, but regained methylation during the self-pollinations (Figure 10-13).

Discussion

Using WGBS, I compared DNA methylation patterns of *ddm1* mutants before and after repeated self-pollinations. *ddm1* mutants exhibited loss of DNA methylation immediately when the mutation became homozygous. In later generations, *ddm1* mutation induced non-CG hypermethylation in hundreds of genes and TEs. And unexpectedly, the non-CG hypermethylation was not observed in 1G *ddm1* mutants but only in 9G *ddm1* mutants. The results suggest that there are short and long term effects of *ddm1* mutation.

Spread of H3K9me and non-CG methylation in *ddm1* mutants

These observations raise the question about the mechanisms that trigger the hypermethylation in self-pollinated *ddm1* lines. It has been previously reported that the de novo non-CG methylation in the self-pollinated *ddm1* does not require components of the RdDM machinery, such as RDR2, DCL3, and DRM2 (Sasaki et al., 2012). On the other hand, the non-CG methylase CMT3 and H3K9 methylase KYP are necessary for the de novo methylation, suggesting that the ectopic methylation occurs by mechanisms mediated by the heterochromatin marks H3K9me2 and non-CG methylation. The genome-wide bisulfite analyses (this study) revealed that the genes non-CG hypermethylated in the self-pollinated *ddm1* tend to have low levels of non-CG methylation already in wild type plants (Figure 6A), suggesting that preexisting small heterochromatin domains may function as seed for further heterochromatin formation.

Interestingly, distribution of H3K9me2 around the DMR is asymmetric; it is enriched in 3' of the DMRs (Figure 21). It has been previously shown that the *BONSAI* gene is flanked by insertion of a heterochromatic LINE in the 3' region (Saze & Kakutani, 2007). The *BONSAI* hypermethylation in *ddm1* is induced in a strain with the LINE insertion but not found in a strain without the LINE insertion. The heterochromatin spreads from the 3' LINE to the *BONSAI* region during repeated self-pollination of *ddm1* mutants (Saze & Kakutani, 2007). Spread of non-CG methylation from 3' to 5' regions was also noted in other loci (Figure 7), as well as in genome-wide estimation (Figure 20), suggesting that similar mechanisms may operate in many, even if not all, affected loci.

Figures

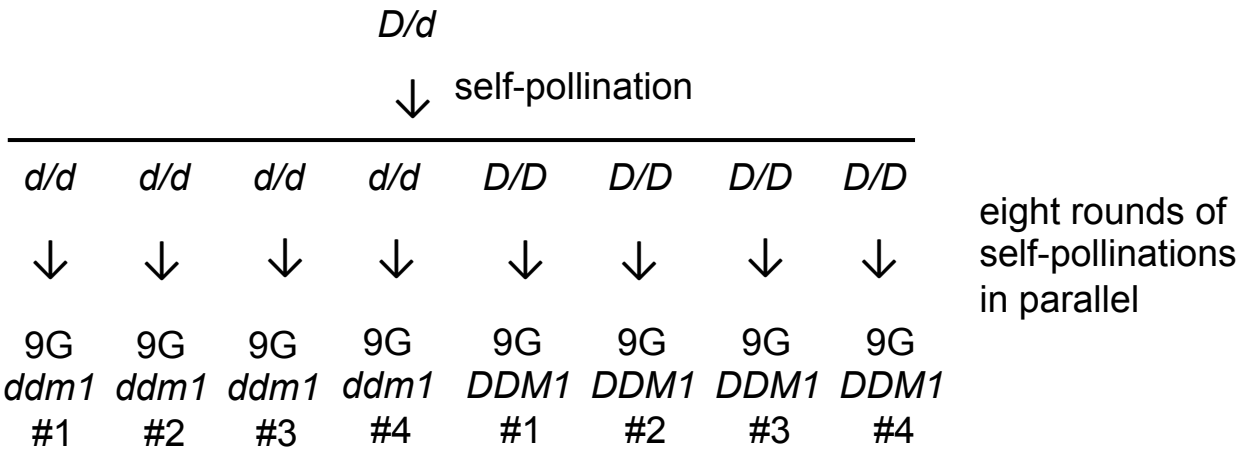


Figure 1. Production of self-pollinated *ddm1* and control *DDM1* lines. Genetic scheme of the production of self-pollinated *ddm1* and control *DDM1* lines. The parental *DDM1/ddm1* (shown as *D/d*) is generated by backcrossing original *ddm1* mutant to wild type six times in the heterozygous state. In the self-pollinated progeny of the heterozygote, multiple *ddm1/ddm1* (*d/d*) and *DDM1/DDM1* (*D/D*) plants were selected and self-pollinated eight times independently to generate 9G *ddm1* and control 9G *DDM1* lines

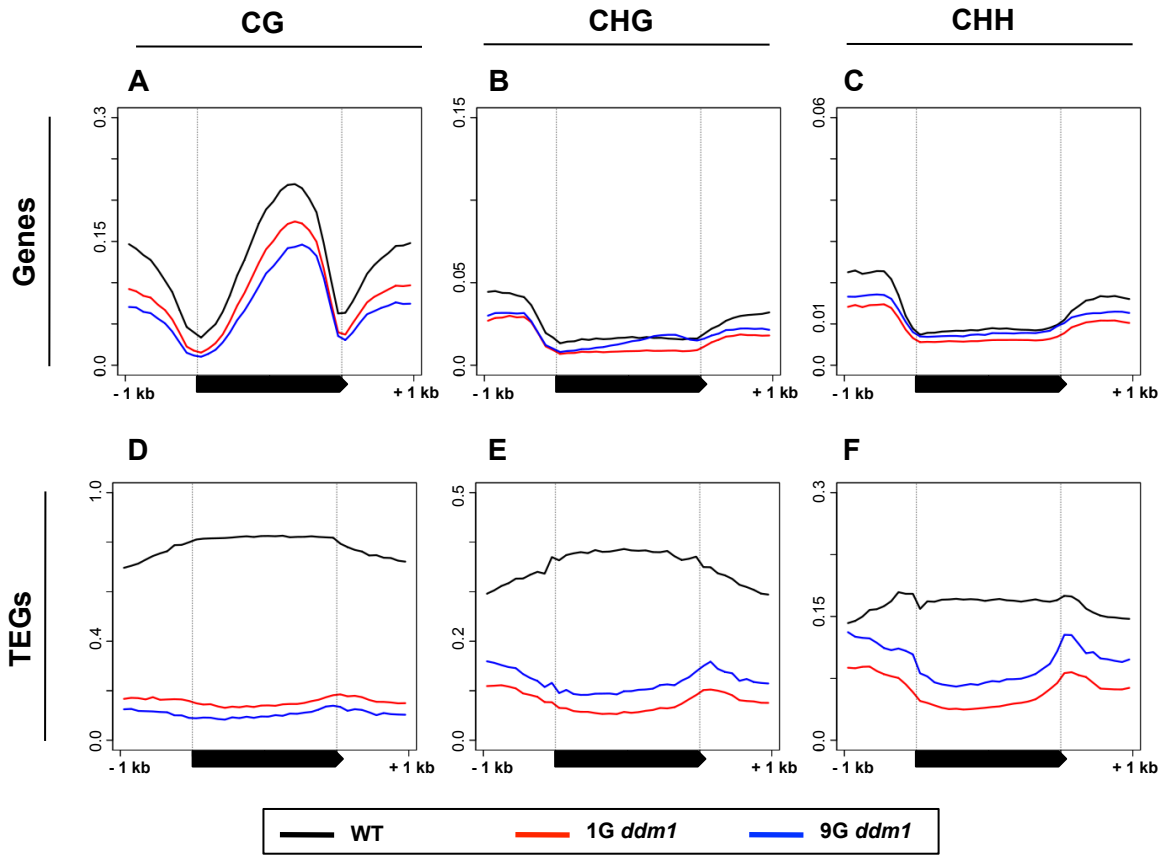


Figure 2. DNA methylation in *ddm1* mutants before and after self-pollination. DNA methylation patterns of WT, 1G *ddm1*, and 9G *ddm1* mutants for normal genes (A-C) or genes associated with transposable elements (transposable element genes, or TEGs, D-F). "WT" is a *DDM1/DDM1* plant segregating as a sibling of the 1G *ddm1/ddm1* plants. The black bars represent transcribed regions. A chromosome-wide view of DNA methylation is also shown in Figure S1.

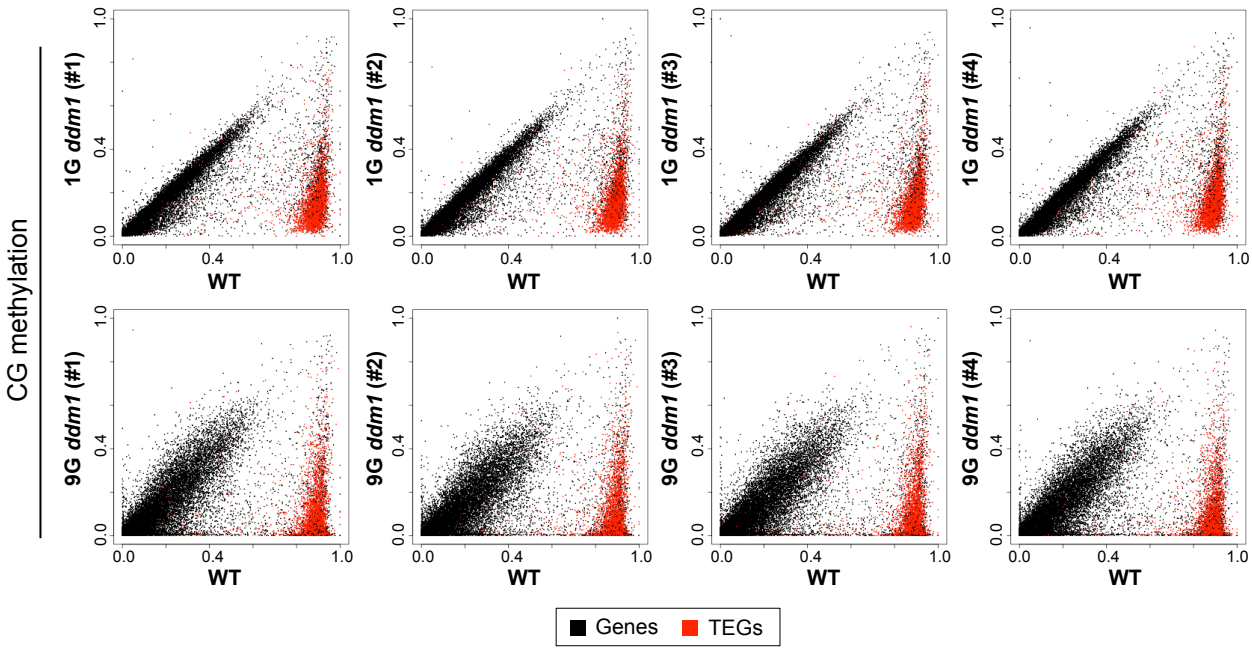


Figure 3. Change of CG methylation during self-pollination of *ddm1* mutants. CG methylation level was compared for each transcription unit. Each dot represents a gene (black dot) or a transposable element gene (TEG, red dot). The top half shows effects in four different 1G *ddm1* plants, while the bottom half shows effects in four different 9G *ddm1* plants. Each of the 9G plants was originated from independent self-pollinations. Comparison of the 9G *ddm1* plants to independently self-pollinated 9G *DDM1* plants (Figure 1) is shown in Figure S2. "WT" is a *DDM1/DDM1* plant segregating as a sibling of the 1G *ddm1/ddm1* plants.

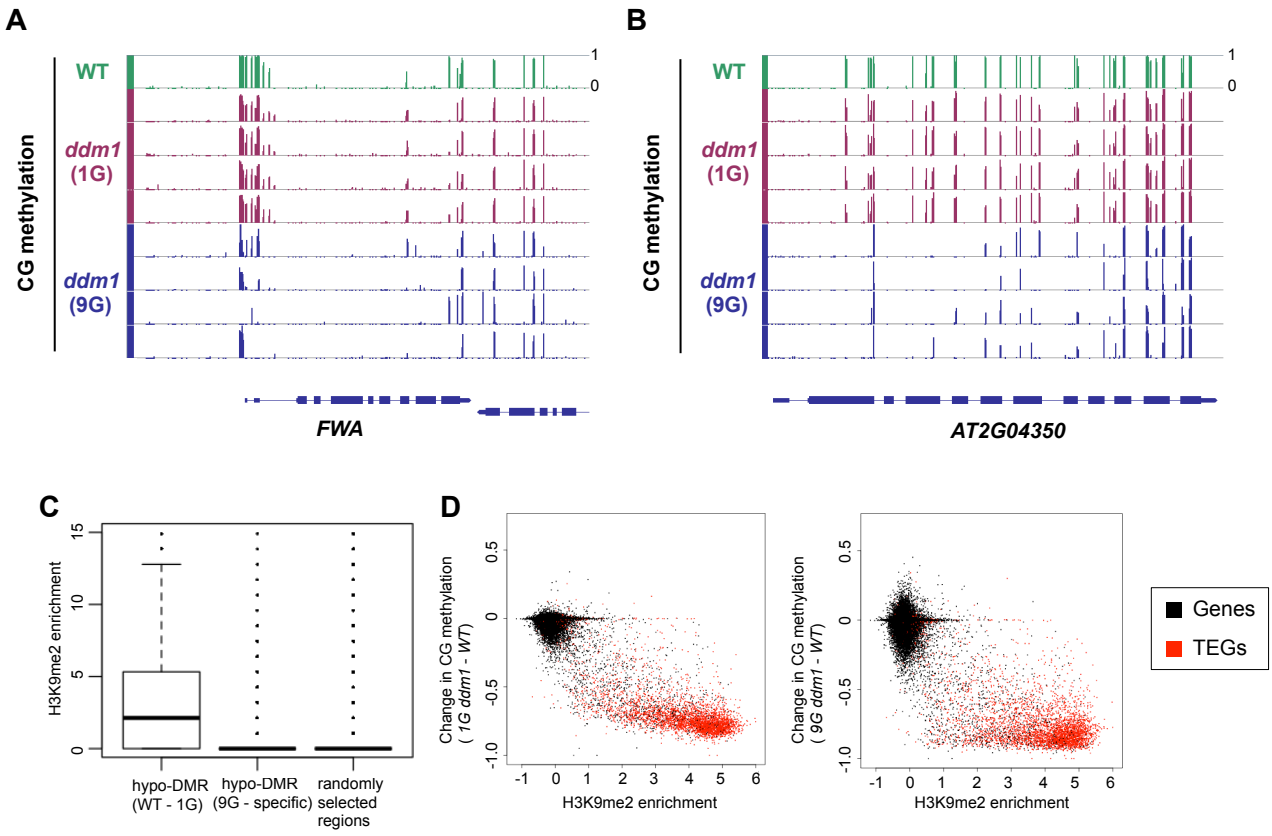


Figure 4. Decrease in CG methylation in 9G *ddm1*.
(A, B) Genome browser views of loci with CG methylation reduced in 9G *ddm1* using the Integrated Genome Browser (Nicol et al., 2009). *FWA* locus (A) and *AT2G04350* locus (B) are shown. The *FWA* gene has dense CG methylation around the 5' end, which is lost during self-pollination of the *ddm1* mutant. (C) H3K9me2 level of differently hypo-methylated regions (hypo-DMRs) in CG context. Left (WT-1G): Distribution of 119,883 DMRs between WT and 1G *ddm1* mutant. Center (9G specific): Distribution of 25,861 DMRs between WT and 9G *ddm1*, excluding DMRs between WT and 1G *ddm1*. Distribution of 100,000 randomly chosen 100 bp regions is also shown as a control (right). H3K9me2 level is shown by reads per million (RPM) in ChIP-seq data obtained from GEO (GSE28398: Luo et al., 2013). (D) Change in CG methylation in 1G *ddm1* (left) and 9G *ddm1* (right) plotted against enrichment of H3K9me2 in WT (data from Inagaki et al., 2010).

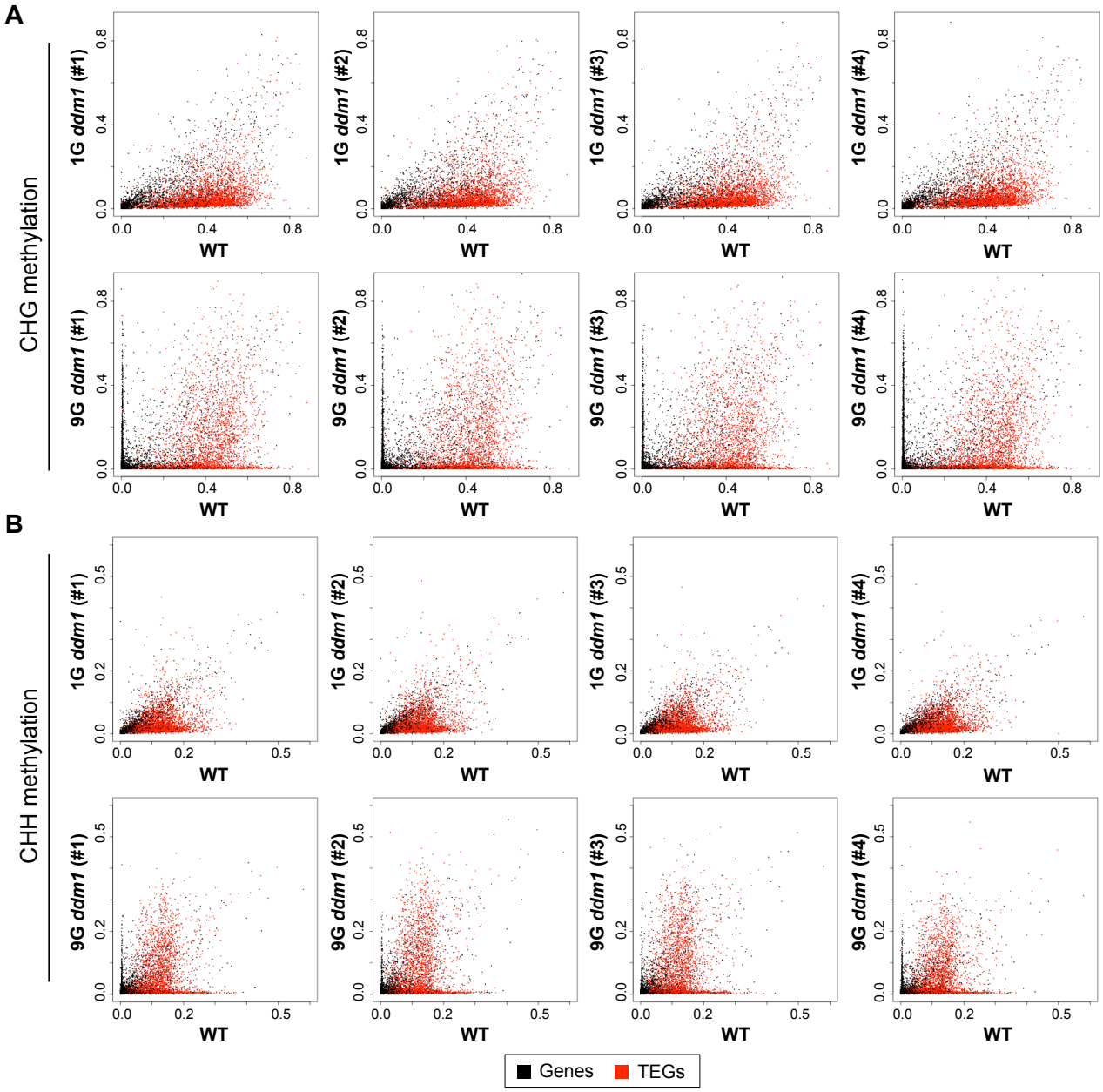


Figure 5. Change of non-CG methylation during self-pollination of *ddm1* mutants. Effects of 1G and 9G *ddm1* mutation on CHG methylation (A) and CHH methylation (B). The format is as shown for CG sites in Figure 3. Comparison of the 9G *ddm1* plants to independently self-pollinated 9G *DDM1* plants is shown in Figure S3.

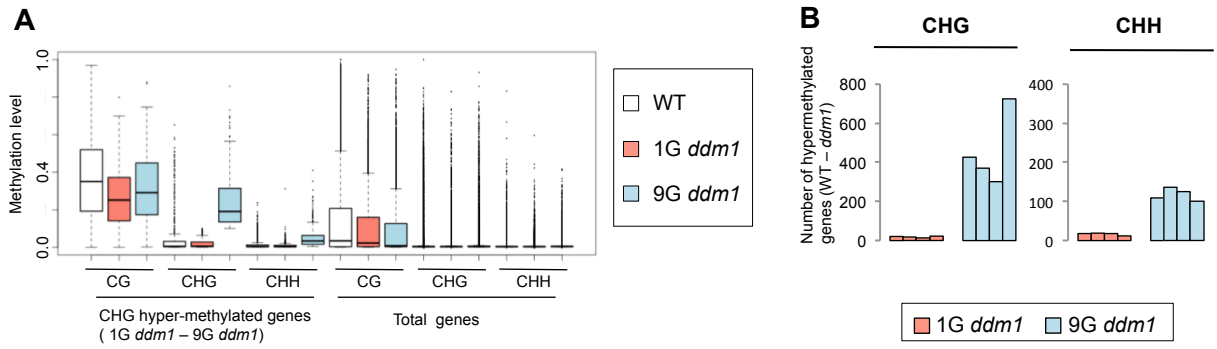


Figure 6. non-CG hypermethylation during self-pollination of *ddm1* mutants. (A) Coordinated hypermethylation of CG, CHG and CHH sites. "CHG-hypermethylated genes" are those with methylation level < 0.1 in 1G *ddm1* and ≥ 0.1 in 9G *ddm1*. DNA methylation levels for three contexts are shown for WT, 1G *ddm1*, and 9G *ddm1*. On the right, total genes are shown as controls. (B) The number of genes that gained non-CG methylation in *ddm1* mutant (methylation level < 0.1 in WT and ≥ 0.1 in *ddm1*). Results for the four 1G and four 9G of *ddm1* mutants are shown for CHG and CHH sites.

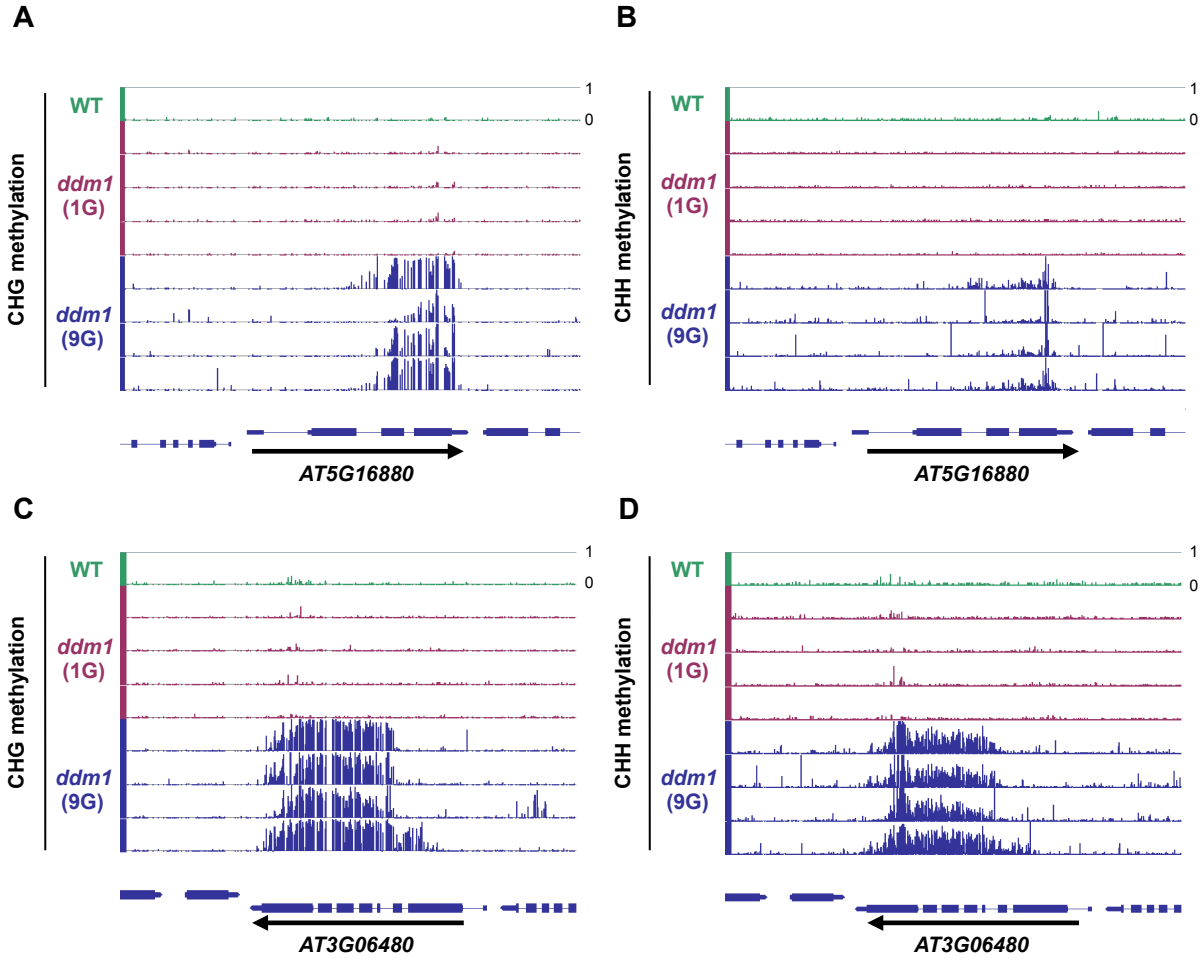


Figure 7. Ectopic non-CG methylation in self-pollinated *ddm1* mutants. (A-D) Genome browser views of loci with non-CG methylation in the 9G *ddm1* plants. *AT5G16880* locus (A-B), *AT3G06480* locus (C-D) are shown for CHG (A, C) and CHH (B, D) contexts.

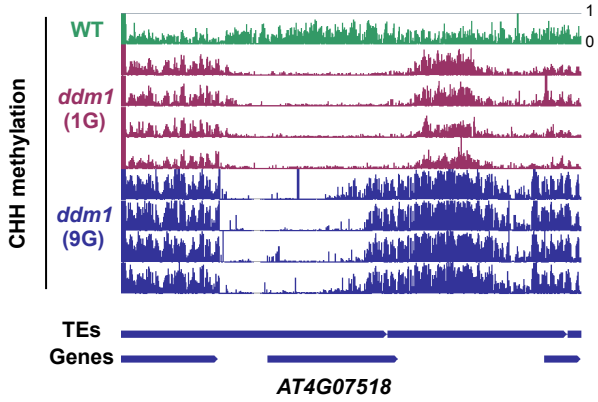


Figure 8. non-CG remethylation of TE in self-pollinated *ddm1* mutants. Genome browser views of *AT4G07518* locus with CHH methylation in the 9G *ddm1* plants.

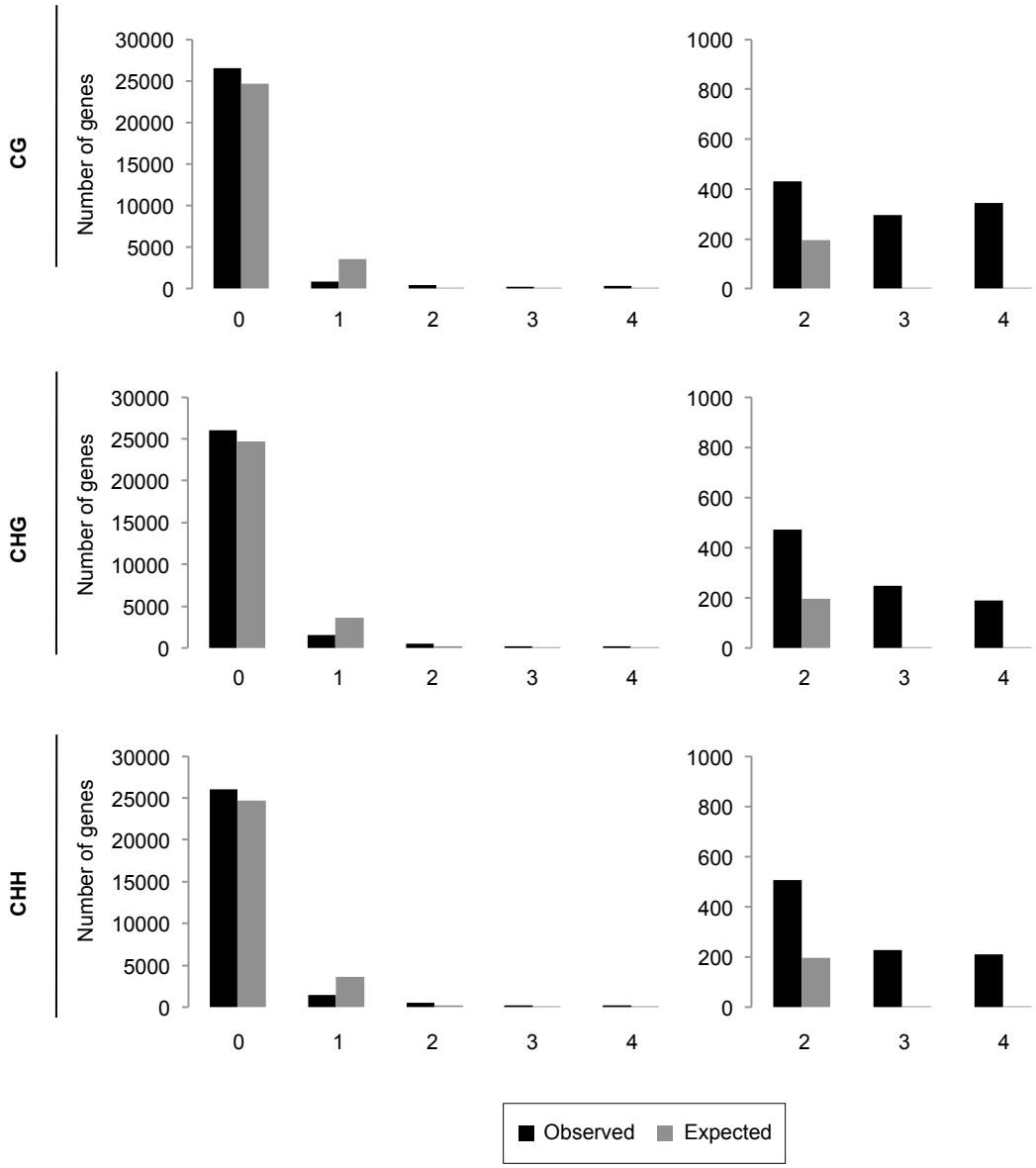


Figure 9. Hypermethylation occurred reproducibly at specific genes during independent repeated self-pollinations of *ddm1* mutants Association of genes hypermethylated in each of the four lines of 9G *ddm1* plants. In each of the four lines, 1000 genes with the largest increase of cytosine methylation were selected. CG, CHG, and CHH contexts are separately shown. "Expected" values were calculated assuming no association (random binominal distribution). Excess of "Observed" values reflects a strong association of the hypermethylated genes in four independently self-pollinated lines. Strong association was found for all three contexts of methylation.

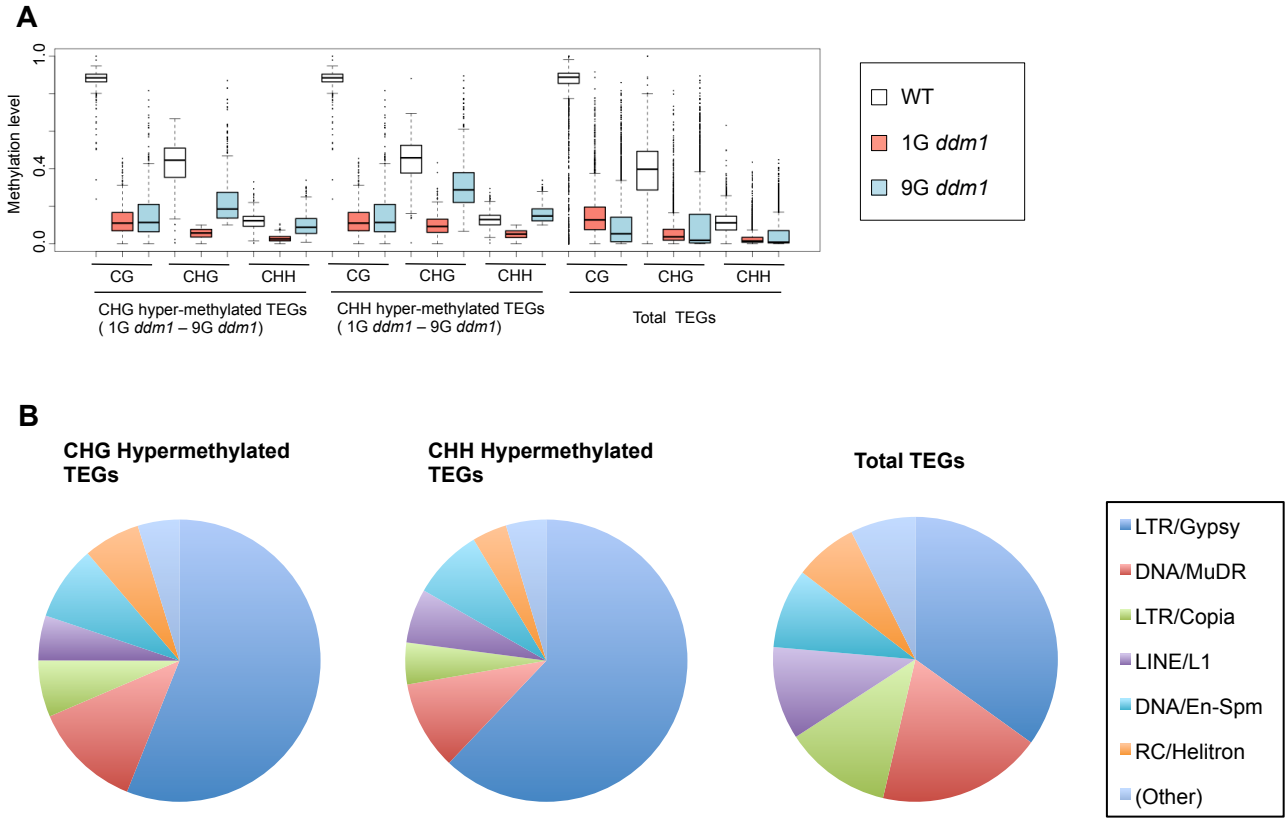


Figure 10. Coordinated remethylation of TEs during self-pollinations of *ddm1*. (A) Change of DNA methylation for CHG-hypermethylated TEGs (1G *ddm1* - 9G *ddm1*) and CHH-hypermethylated TEGs (1G *ddm1* - 9G *ddm1*), compared to all TEGs shown as controls. Three contexts of sites show coordinated hypermethylation in 9G. (B) Pie charts of numbers of non-CG hypermethylated TEGs in each family of TEs shown in (A). TEGs were classified according to the family of the corresponding TE.

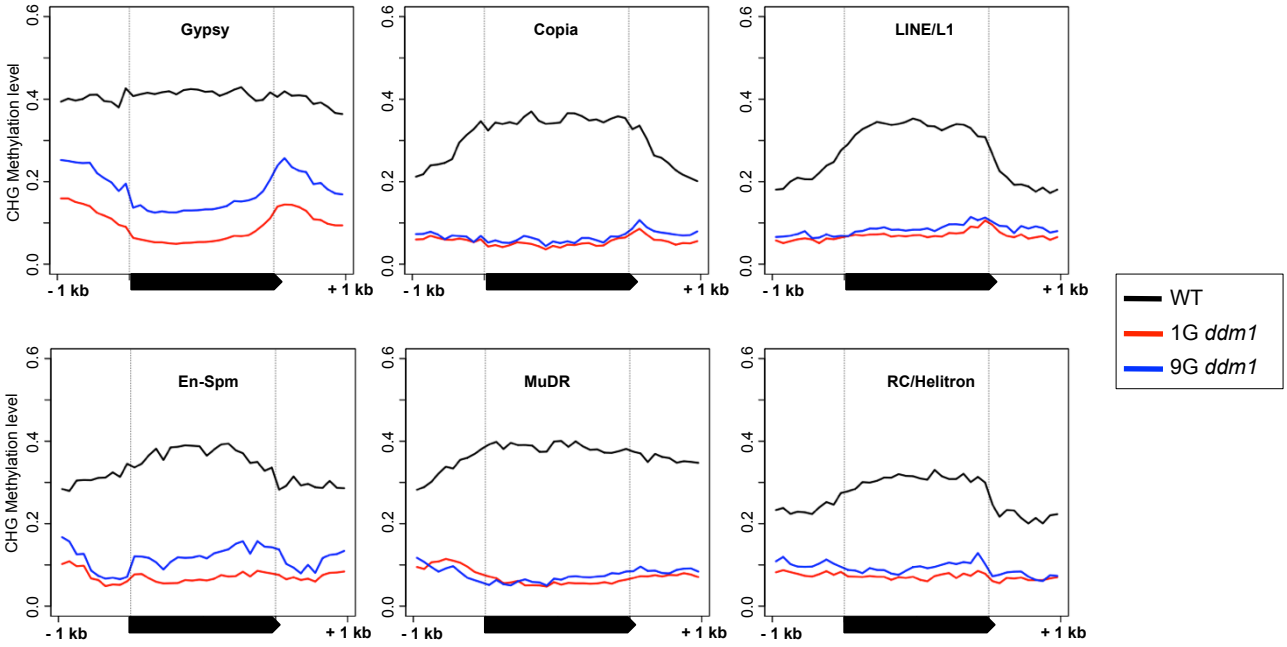


Figure 11. Profiles for CHG remethylation of TEs during self-pollinations of *ddm1*. Pattern of CHG methylation over TEGs are shown for each of TE families. Gypsy show strong peak outside transcription termination site for both CHG and CHH contexts.

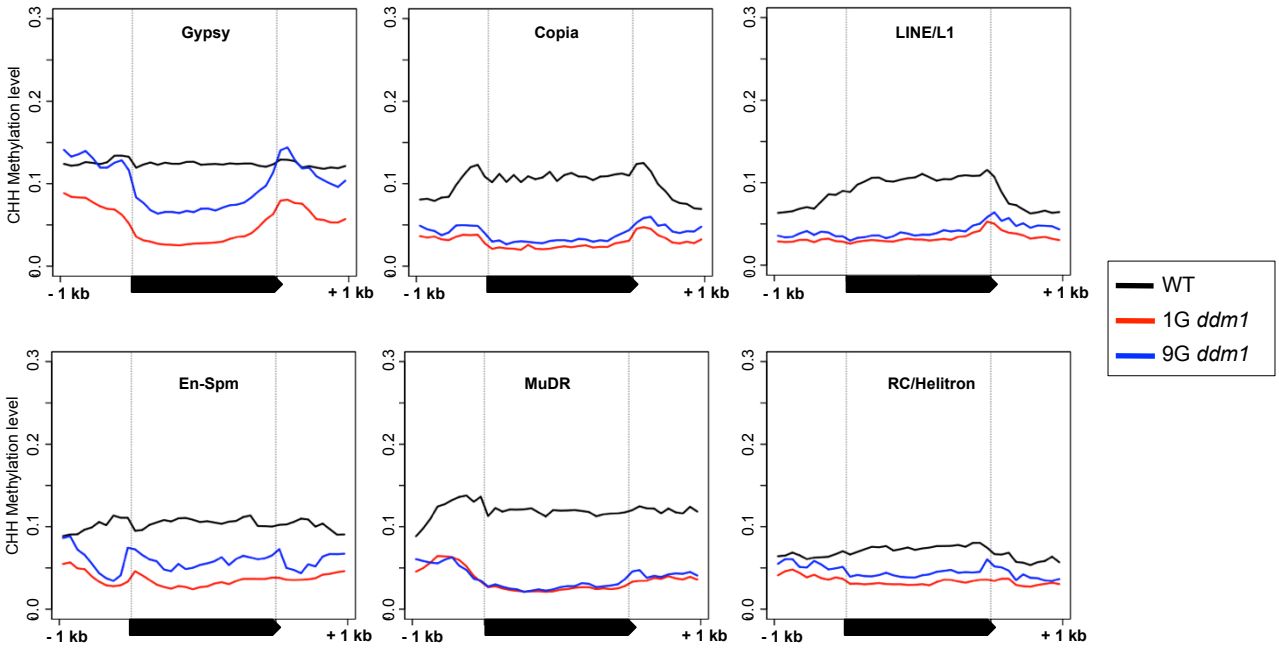


Figure 12. Profiles for CHH remethylation of TEs during self-pollinations of *ddm1*. Pattern of CHH methylation over TEGs are shown for each of TE families. Gypsy show strong peak outside transcription termination site for both CHG and CHH contexts.

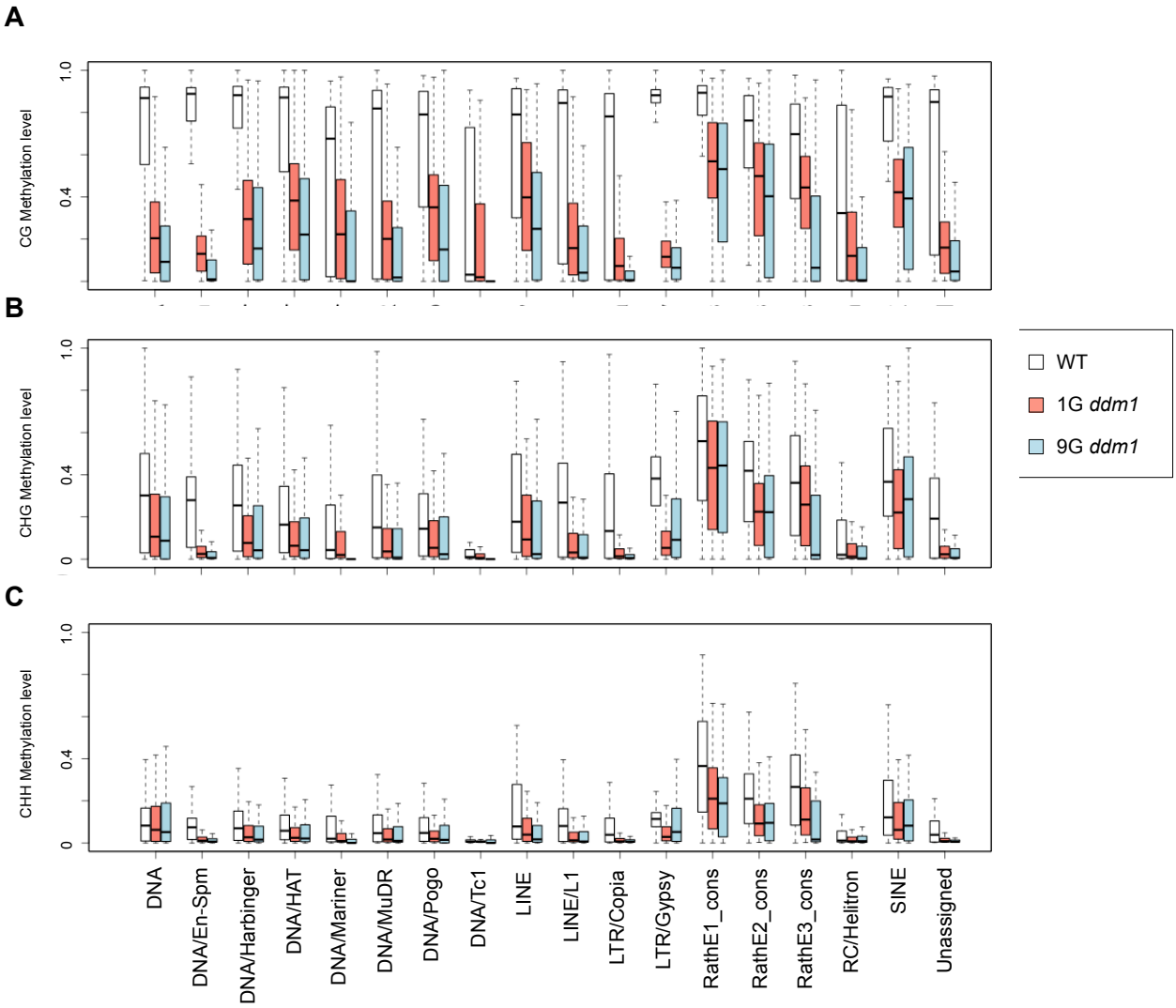


Figure 13. Distribution of the effect of *ddm1* mutation among the TE families. Distribution of methylation change was shown for each of TE families for the three contexts of methylation.

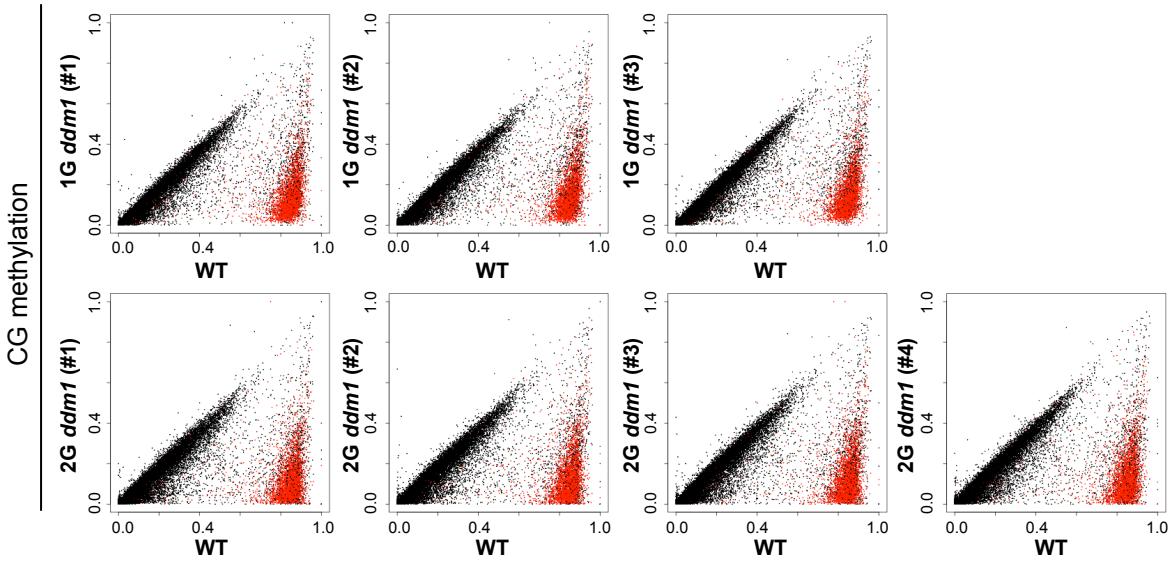


Figure 14. CG methylation level in 2G *ddm1* plants. Methylation level of cytosine was compared for each transcription unit. The top half shows effects in three different 1G *ddm1* plants, while the bottom half shows effects in four different 2G *ddm1* plants. Each of the 2G plants was originated from independent 1G *ddm1* plants. "WT" is a *DDM1*/*DDM1* plant segregating as a sibling of the 1G *ddm1*/*ddm1* plants.

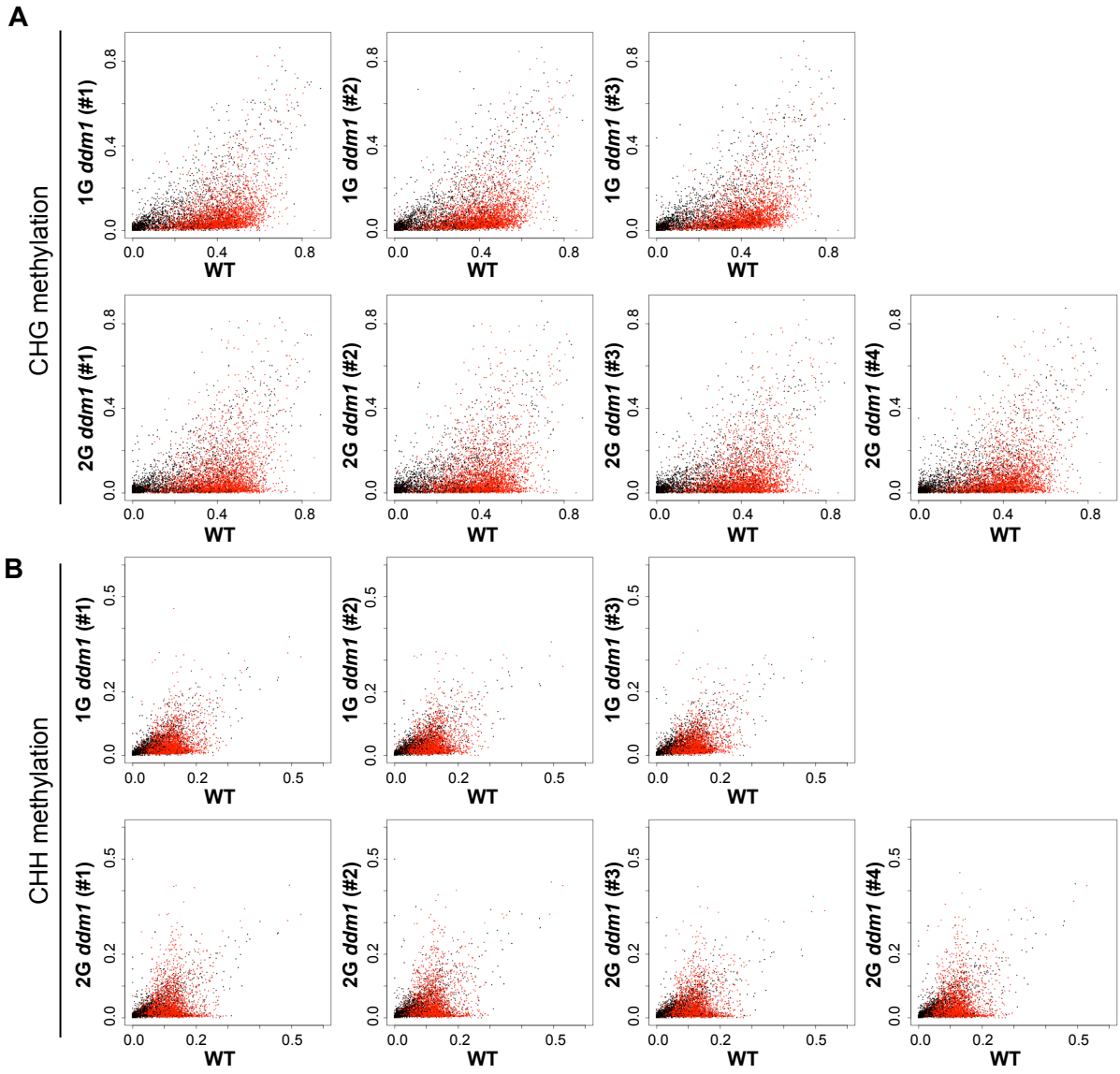


Figure 15. Non-CG methylation level in 2G *ddm1* plants. Methylation level of cytosine was compared for each transcription unit. The top half shows effects in three different 1G *ddm1* plants, while the bottom half shows effects in four different 2G *ddm1* plants. CHG (A), and CHH (B) contexts are separately shown. Each of the 2G plants was originated from independent 1G *ddm1* plants. "WT" is a *DDM1*/*DDM1* plant segregating as a sibling of the 1G *ddm1*/*ddm1* plants.

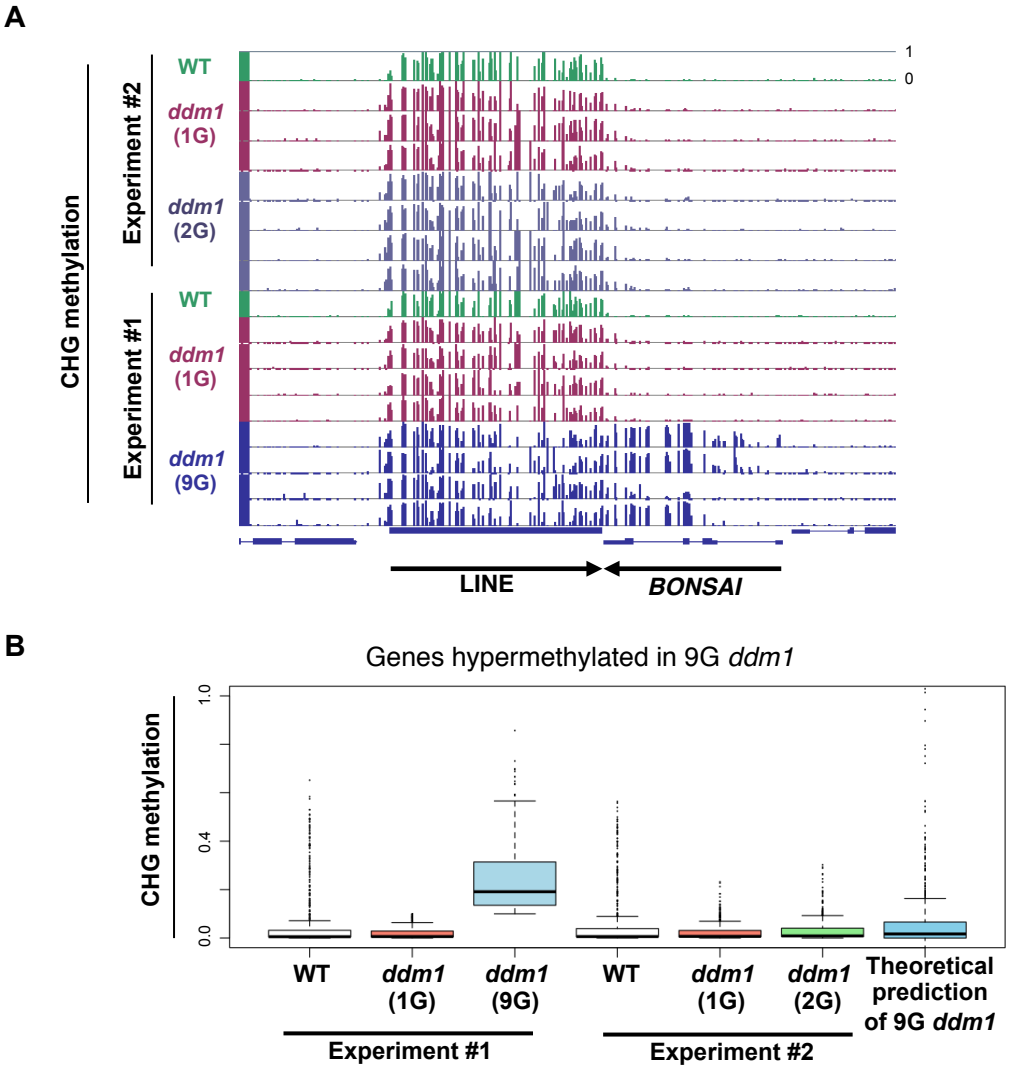
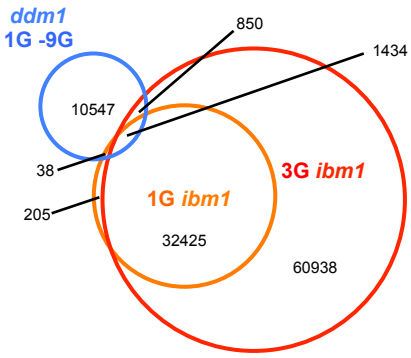


Figure 16. Ectopic non-CG methylation occurring in 2G *ddm1* was slow. (A) Genome browser views of CHG methylation at *AT1G73177* (BONSAI) locus. Spread of CHG methylation from LINE to *BONSAI* gene was still modest in the 2G *ddm1* compared to the 9G *ddm1*. (B) Change of CHG methylation level for genes hypermethylated in 9G *ddm1*. Results are shown for the Experiment #1 (WT, 1G *ddm1*, and 9G *ddm1*) (Figure 5A) and the Experiment #2 (WT, 1G *ddm1*, and 2G *ddm1*) (Figure 15A). The value in the right, "Theoretical prediction of 9G *ddm1*", was calculated by extrapolating signals for 1G and 2G *ddm1* in the experiment #2. In other words, values were calculated by B + (B – A) x 7, where A and B are signals for 1G *ddm1* and 2G *ddm1* in the experiment #2. The value is much less than that in 9G *ddm1* in the experiment #1, suggesting that the ectopic hypermethylation proceed much slower in the initial generations than in later generations.

A



B

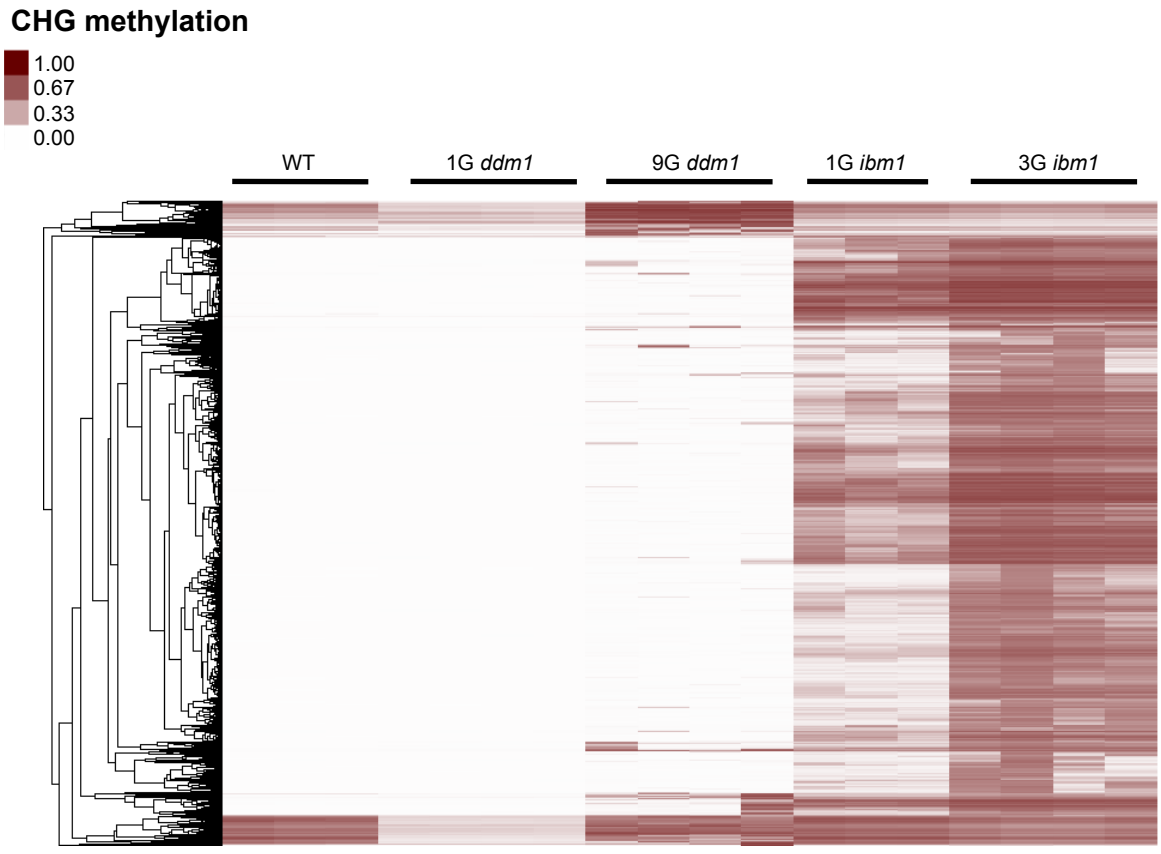


Figure 17. Comparison of CHG methylation level in DMRs.
(A) Comparison of regions CHG hypermethylated in *ibm1* and 9G *ddm1*. DMRs between 1G and 9G *ddm1* (blue), between WT and 1G *ibm1* (orange), and between WT and 3G *ibm1* (red) are shown. (B) Heatmap of CHG methylation for the DMRs shown in A.

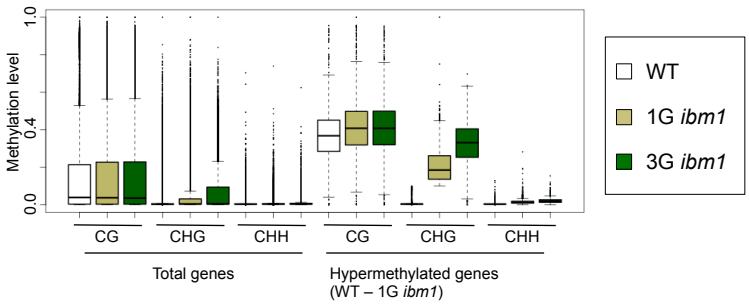


Figure 18. Hypermethylated regions in *ddm1* and *ibm1* mutants.
Increase of CHG methylation in 1G and 3G *ibm1* mutants. Genes hypermethylated in 1G *ibm1* (CHG methylation level < 0.1 in WT and ≥ 0.1 in 1G *ibm1*) are shown (right) with total genes (left). Profiles for multiple 1G and 3G *ibm1* mutant plants are shown in Figure S4.

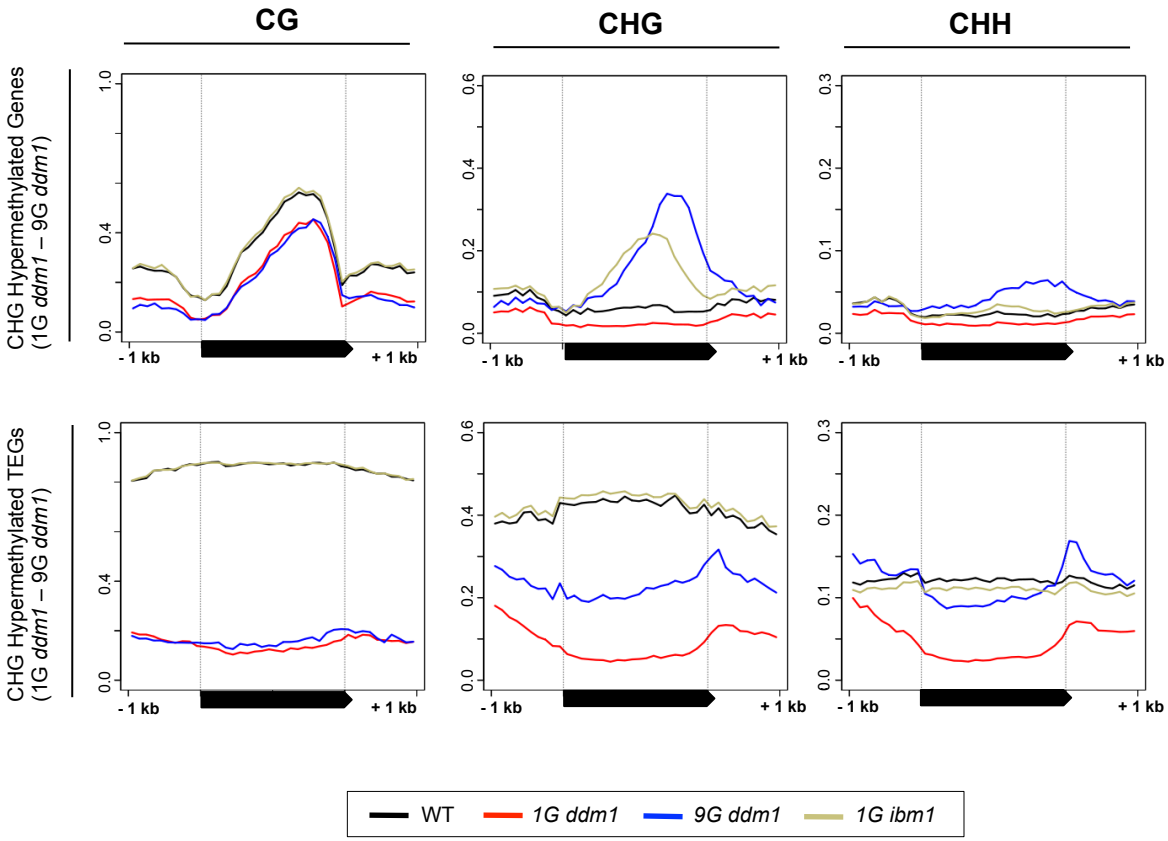


Figure 19. The difference between the effect of self-pollination of *ddm1* mutants and that of *ibm1* mutants. DNA methylation profile for the genes CHG hyper-methylated in 9G *ddm1* (shown in Figure 6A). The top and bottom half represent genes and TEGs, respectively. In these regions, CHH methylation also increased in 9G *ddm1*.

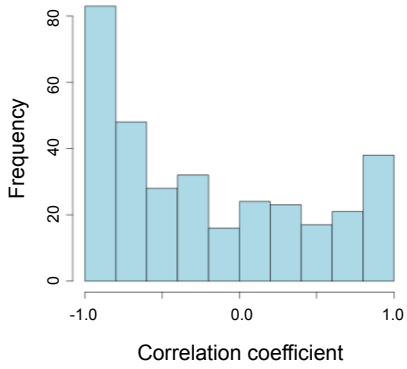


Figure 20. Change of non-CG methylation during self-pollination of *ddm1* mutants. Histogram of correlation coefficient between the level and the relative centroid position of CHG methylation. The coefficient was calculated among the four 9G *ddm1* plants for CHG methylation in contiguous DMRs between 1G and 9G *ddm1* (details in Materials and Methods section) overlapping with genes. The coefficient becomes positive when the centroid moves to the 3' regions as the signal increase. A large proportion of the contiguous DMRs with the coefficient near -1 reflects spread of signals from 3' to 5' regions as the CHG methylation levels increase.

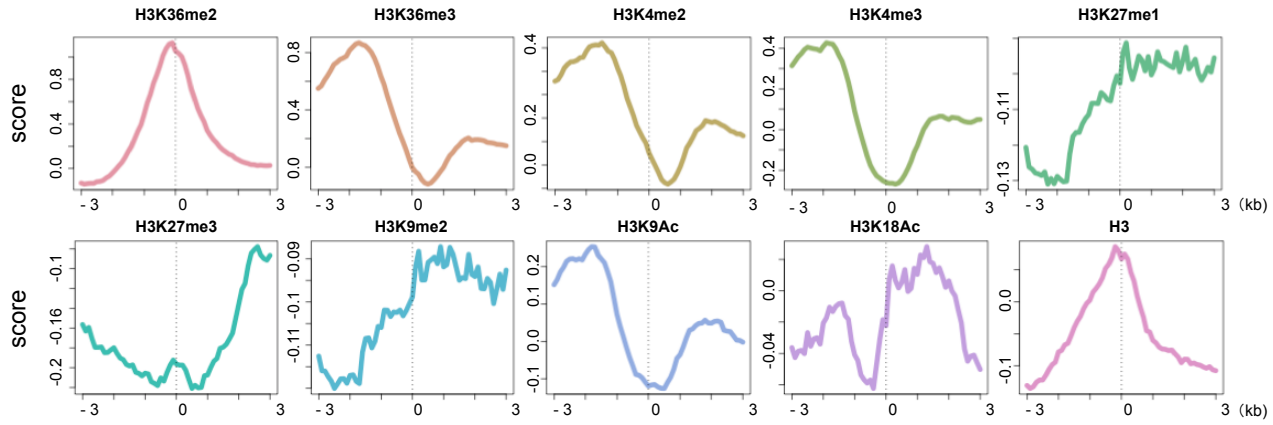


Figure 21. Enrichment of various histone modifications around the DMRs (1G *ddm1* - 9G *ddm1*). Normalized scores were calculated using the 100 thousand regions chosen randomly from the genome. Only DMRs that overlapped with genes were used; Each DMR was aligned according to the orientation of the corresponding gene. ChIP-seq data was obtained from GEO (GSE28398: Luo et al., 2013).

Supplementary Figures

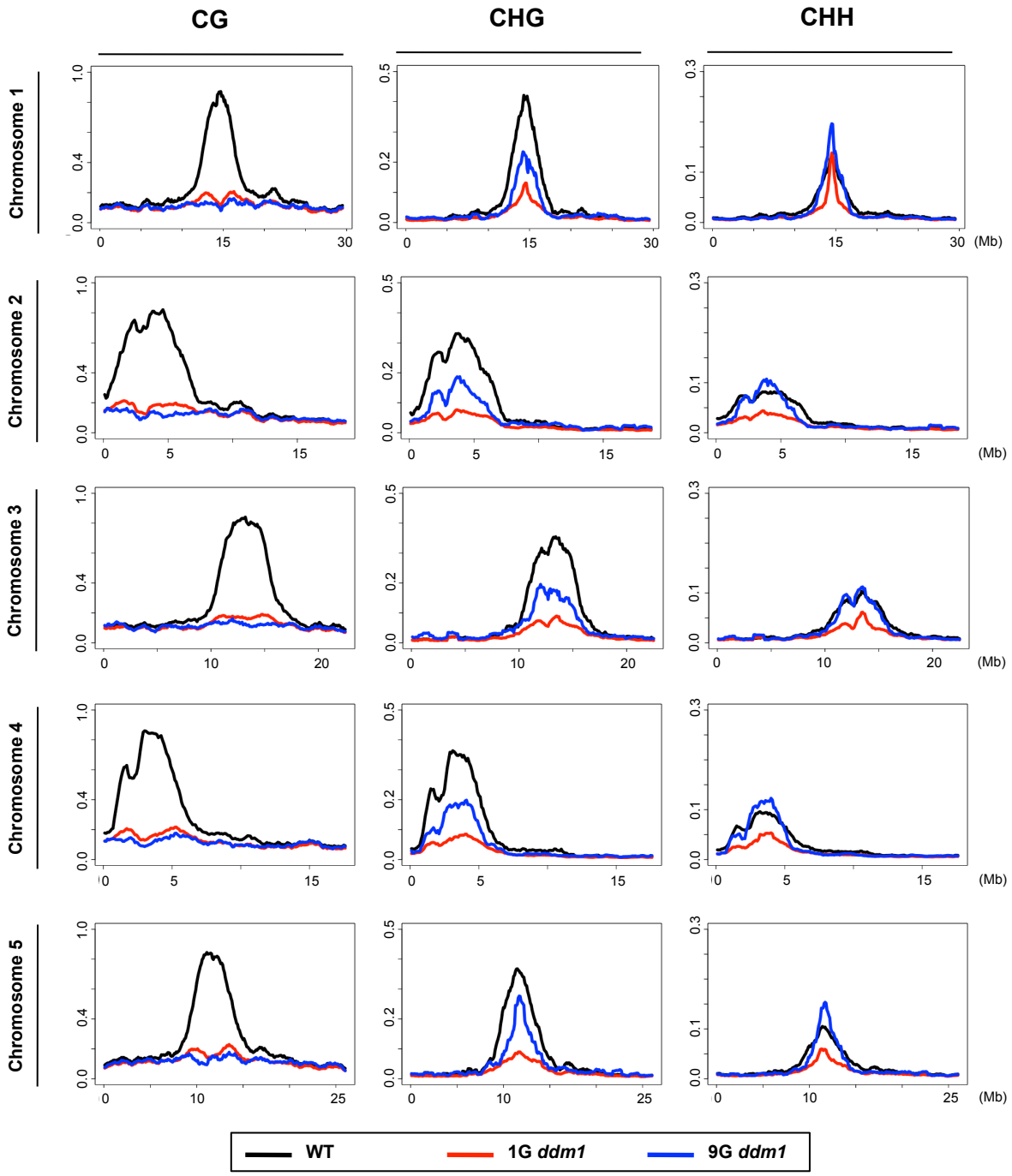


Figure S1. Chromosome-wide view of DNA methylation profiles in *ddm1* mutant lines before and after self-pollinations. Cytosine methylation levels are shown for the three contexts, CG, CHG and CHH, with the sliding windows of 1 Mb.

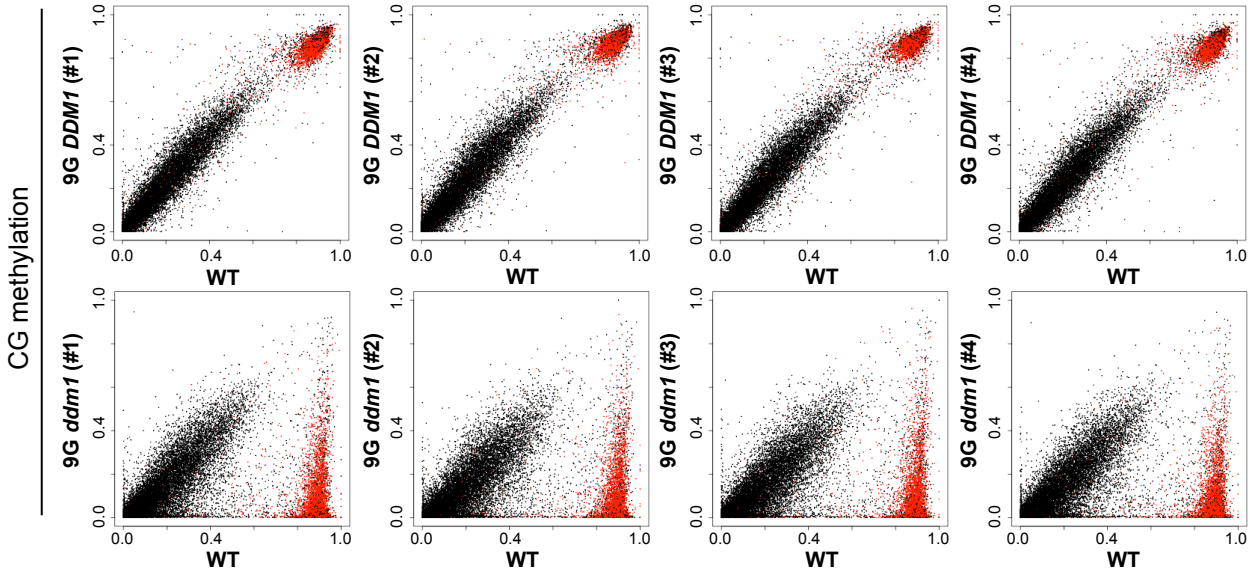


Figure S2. Change of CG methylation in 9G *ddm1* plants compared to control 9G *DDM1* plants. Methylation level of cytosine was compared for each transcription unit between 9G *DDM1/DDM1* plants and WT in the CG context. The format is as shown in Figure 3. Each of the 9G plants was originated from independent self-pollinations (Figure 1). "WT" is a *DDM1/DDM1* plant segregating as a sibling of the 1G *ddm1/ddm1* plants.

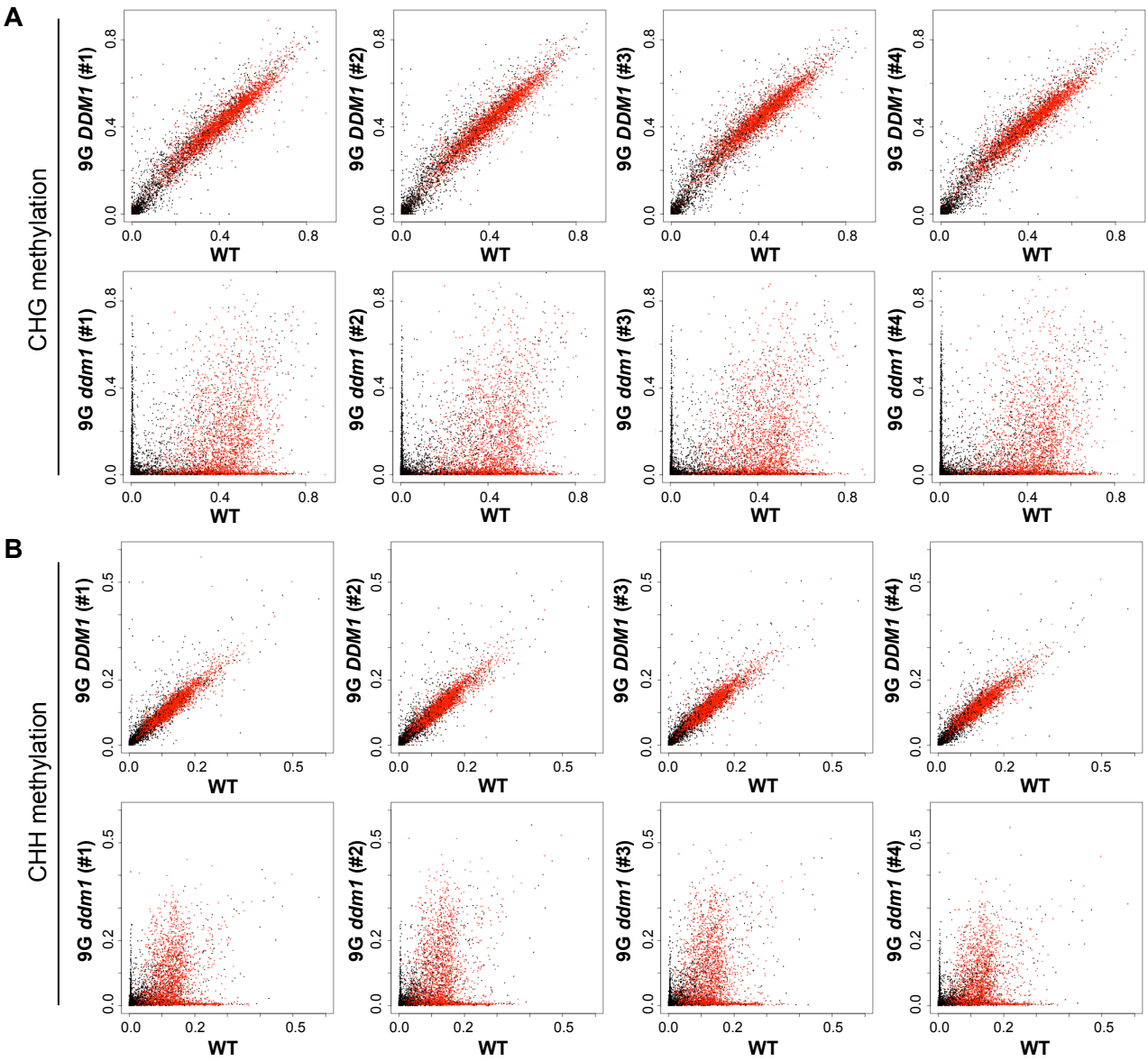


Figure S3. Change of non-CG methylation in 9G *ddm1* plants compared to control 9G *DDM1* plants. Methylation level of cytosine was compared for each transcription unit between 9G *DDM1/DDM1* plants and WT in the CHG (A), and CHH (B) contexts. The format is as shown in Figure 3. Each of the 9G plants was originated from independent self-pollinations (Figure 1). "WT" is a *DDM1/DDM1* plant segregating as a sibling of the 1G *ddm1/ddm1* plants.

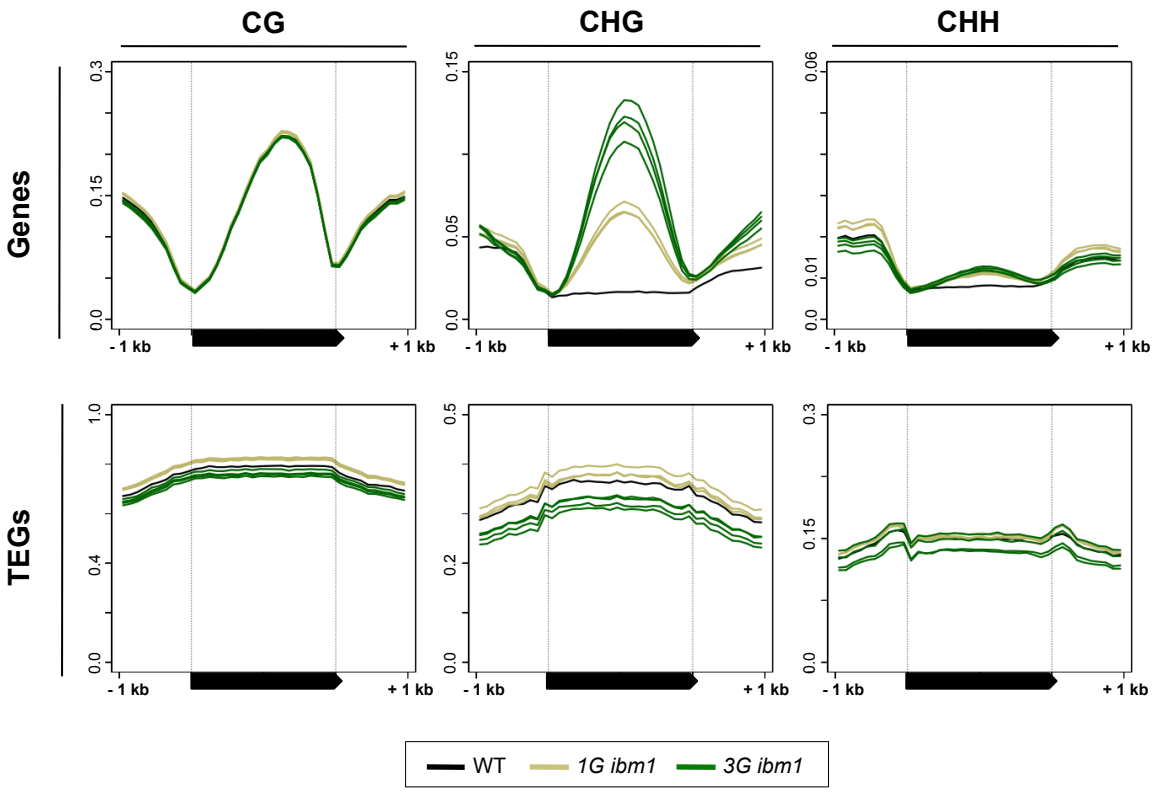


Figure S4. DNA methylation in self-pollinated *ibm1* mutants.
Patterns of DNA methylation over total genes and TEGs are shown for WT, 1G *ibm1*, and 3G *ibm1*. 1G *ibm1* plants are progeny of a *IBM1*/*ibm1* heterozygote. Their *ibm1*/*ibm1* siblings were self-pollinated twice and the progenies were used as 3G *ibm1*.

Chapter 4

Local hypermethylation induced by global hypomethylation

Introduction

In the previous chapter, I have shown that the *ddm1* mutation can induce ectopic DNA methylation at hundreds of genes and TEs. There are two possible explanations for the hypermethylation. One is that the hypermethylation is a direct consequence of impaired *DDM1* function. The other is that the hypermethylation is due to indirect effects of the disrupted heterochromatin in *ddm1* mutants. To test these possibilities, I examined whether the chromosome segments introduced from *ddm1* mutant can induce hypermethylation in wild-type *DDM1* background, using epigenetic recombinant inbred lines (epiRILs) developed by Vincent Colot's laboratory (Johannes et al., 2009).

During the production of the epiRILs, a parental *ddm1* mutant plant was crossed to wild type plant twice, and the F2 plant with *DDM1/DDM1* genotype were selected and self-pollinated seven times (Colomé-Tatché et al., 2012). Thus, each line contains part of chromosome segments derived from *ddm1*, and most of the genomic regions were fixed in *ddm1*-derived haplotype or wild-type-derived haplotype though the process of the repeated self-pollinations (Colomé-Tatché et al., 2012).

In the epiRILs, *ddm1*-derived chromosome segments remain hypomethylated even after introduction into wild type *DDM1* background (Vongs et al., 1993; Kakutani

et al., 1999). Although remethylation can be induced in regions associated with small RNA, hundreds of the hypomethylated regions remain unmethylated in the wild type *DDM1* background (Teixeira et al., 2009; Colomé-Tatché et al., 2012). The haplotypes in the genome of each epiRIL can be inferred using these regions as the markers.

In order to analyze genome-wide DNA methylation in the epiRILs, I utilized public data for epiRILs and examined the effect of *ddm1*-derived hypomethylated chromosome segments in the epiRILs.

Results

The abnormally methylated chromosomes from a *ddm1* mutant could induce hypermethylation in trans even in *DDM1* wild type background

I examined whether the loci exhibiting hypermethylation in the self-pollinated *ddm1* lines also showed hypermethylation in some of the epiRILs using DNA methylation data for the 123 epiRILs, which are based on immunoprecipitation (IP) of genomic DNA by anti-methylcytosine antibody. As methylation of each context cannot be distinguished, I examined seven loci that show increased methylation in 9G *ddm1* but a relatively low level of methylation at CG sites in wild-type. In six out of the seven loci examined, hypermethylation was detected in multiple epiRILs, suggesting that the hypermethylation can be induced or maintained in the *DDM1* background (Figure 22-23). In all of them, the hypermethylation showed positive correlation with the amount of disrupted heterochromatin in each of these lines (Figure 22-23; Table 1), suggesting that the hypermethylation was induced or maintained in the background of disrupted heterochromatin in other genomic regions.

The hypermethylation could be induced de novo or alternatively maintained from the parental *ddm1*. The parental *ddm1* plant originally used for making epiRILs are already self-pollinated three times and that plant also show increased methylation at some loci (Figure 23), which may have the potential to be maintained in *DDM1* background (Saze & Kakutani, 2007). Very importantly, however, the hypermethylation was found even in chromosome segments originated from wild type *DDM1* (Figure

22BDF; Figure 24-29), demonstrating that the hypermethylation could be induced de novo after the initial crosses in the background of functional *DDM1*.

In order to confirm and extend this observation, I used WGBS data for an epiRIL with genome-wide reduction of heterochromatic DNA methylation. The epiRIL98, which contains large amount of chromosomes with reduced DNA methylation, showed CHG hypermethylation in many genes (Figure 30), which include *BONSAI* gene (Figure 31) and genes with body methylation (Figure 32). In the CHG hypermethylated genes, the CHG methylation level was generally much higher than that of the parental 4G *ddm1* plant (Figure 33A), suggesting that the hypermethylation was amplified or induced de novo in the background of functional *DDM1*. A large number of CHG hypermethylated genes were found in chromosome regions of wild type haplotype (Figure 33B, Figure 34), again suggesting that they can be induced de novo. In control epiRILs with much lower levels of disrupted chromatin, the hypermethylation was undetectable, confirming that the disrupted heterochromatin was responsible (Figure 30). Taken together, these results strongly suggest that the hypermethylation can be induced de novo by trans-acting effects of disrupted heterochromatin.

Discussion

Through analysis of DNA methylation data of epiRILs, I showed that the hypermethylation was positively correlated with the amount of disrupted heterochromatin. Importantly, I observed the hypermethylation even in the chromosome segments derived from the wild-type parental plant. These results suggest that the hypermethylation is due to indirect effect of disrupted heterochromatin in the *ddm1* mutants.

Genome-wide negative feedback for heterochromatin and its developmental control

Based on the results, the spread of non-CG methylation observed in the self-pollinated *ddm1* seems to reflect negative feedback of disrupted heterochromatin in other genomic regions, because the hypermethylation could also be induced in *DDM1* wild type background when the genome contains large amount of *ddm1*-derived chromosomal segments with disrupted heterochromatin (Figure 22). How does the negative feedback work? One possible explanation is that disruption of heterochromatin in the *ddm1* mutant results in release of heterochromatin-forming factors such as CMTs and H3K9 smethylases, which then become available in other regions. As these factors are normally recruited to heterochromatin, disruption of a large proportion of heterochromatin in the genome would result in increased level of these factors in released conditions, which would induce spread of heterochromatin into normally

euchromatic regions and its amplification by the self-reinforcing loop of H3K9me2 and non-CG methylation (Figure 35).

This trans-acting negative feedback could also be understood as a hypersensitive reaction to the challenge by active and proliferating TEs. The genome-wide analyses revealed that many of the TEs can be the targets of the negative feedback (Figure 5, Figure 10-13). Active TEs often keep parts of heterochromatin, which can function as seeds of the self-reinforcing heterochromatin formation.

An increase in non-CG methylation is also seen in mutants of the histone demethylase gene *IBM1*. However, targets of IBM1 are generally euchromatic and they do not overlap much with regions hypermethylated in the self-pollinated *ddm1* lines (Figure 17). An increase in non-CG methylation is also found in the maintenance CG methylase gene *MET1* (Jacobsen et al., 1997; Kishimoto et al., 2001; Mathieu et al., 2007). As a mechanism for the *met1*-induced increase in non-CG methylation, loss of *IBM1* function is suggested, as *IBM1* transcripts become truncated in the *met1* mutant (Rigal et al., 2012). On the other hand, Deleris et al. reported that the targets of the *met1*-induced accumulation of H3K9me2 are genes with H3K27me3, another modification for silent chromatin (Deleris et al., 2012). The negative feedback of heterochromatin marks comparable to that seen in the self-pollinated *ddm1* lines may also operate in *met1* mutants. In this study, although spectrums of regions affected by *met1*, *ibm1*, and self-pollinated *ddm1* all differ, significant overlaps are noted (Figure 36). For these mutants, the local triggers for the heterochromatin accumulation appear to be distinct, despite the possible overlap in the downstream mechanisms, including the

self-reinforcing loop of non-CG methylation and H3K9me2.

Other possible mechanisms inferred by observations

I proposed a model that disruption of heterochromatin causes released heterochromatin factors, which induces spread of heterochromatin into normally euchromatic regions and its amplification by the self-reinforcing loop of H3K9me2 and non-CG methylation (Figure 35). Although they are not necessarily mutually exclusive, there are other two possibilities.

(i) Paramutation

Paramutation involves trans-interactions of heritable states between alleles or homologous sequences; when one allele is in silent epigenetic state, the other allele become heritably silent by the interaction between the alleles. The silent state is often associated with DNA methylation and components of RdDM are often necessary for the interaction (Arteaga-Vazquez & Chandler, 2010).

Similar mechanism might operate in the production of epiRILs. In the first progeny of the cross to generate epiRILs, the *ddm1*-derived chromosome and wild-type derived chromosome become heterozygous. Presumably, the former have substantial level of ectopic non-CG methylation, while the latter does not have the ectopic methylation. If paramutation-like interactions occur in this generation, the ectopic hypermethylation may be transmitted from the *ddm1* derived chromosome to the wild-type derived chromosome. That would account for the trans-interaction.

The ectopic hypermethylation in the epiRILs is generally much higher than that of the parental *ddm1* (Figure 33A), suggesting that even if paramutation-like mechanisms are involved, the effect should be much amplified during self-pollinations of epiRILs; and the degree of the amplification correlates with global disruption of heterochromatin (Figure 22-23), which is due to the *ddm1*-derived chromosomes.

(ii) Possible involvement of specific locus

In the model described above, global reduction of heterochromatin induces ectopic non-CG methylation (Figure 35). That would account for the correlation between the global reduction of methylation and ectopic methylation in epiRILs. An alternative mechanism would be that *ddm1* induces change in a specific locus, such as transcriptional de-repression of a specific gene, and the change is inherited in the *DDM1* wild type background and induces the ectopic methylation.

However, I could not find the causative locus consistently derived from *ddm1* parent in all of the plants showing the high level of ectopic hypermethylation in the six loci (Figure 24-29). Although it is quite difficult to exclude the possibility that two or more specific loci redundantly mediate the ectopic methylation, a more parsimonious explanation derived from available data would be that the trans-interaction is mediated by global homeostasis.

Table

Table 1. Strong positive correlation between the global hypomethylation from the *ddm1*-derived chromosomes and the local hypermethylation. Pearson correlation coefficients between global hypomethylation and local hypermethylation are shown with p-values using the data of 123 epiRILs. Six out of the seven loci examined in Figure 23 showed significant positive correlation.

Locus ID	*AT1G73177*	*AT5G52480*	*AT5G35510*
r	0.624	0.515	0.577
p-value	1.24E-14	1.07E-09	2.72E-12

Locus ID	*AT4G30975*	*AT2G39540*	*AT1G03660*
r	0.288	0.476	0.273
p-value	0.001226	2.56E-08	0.002223

r: Pearson's correlation coefficient

Figures

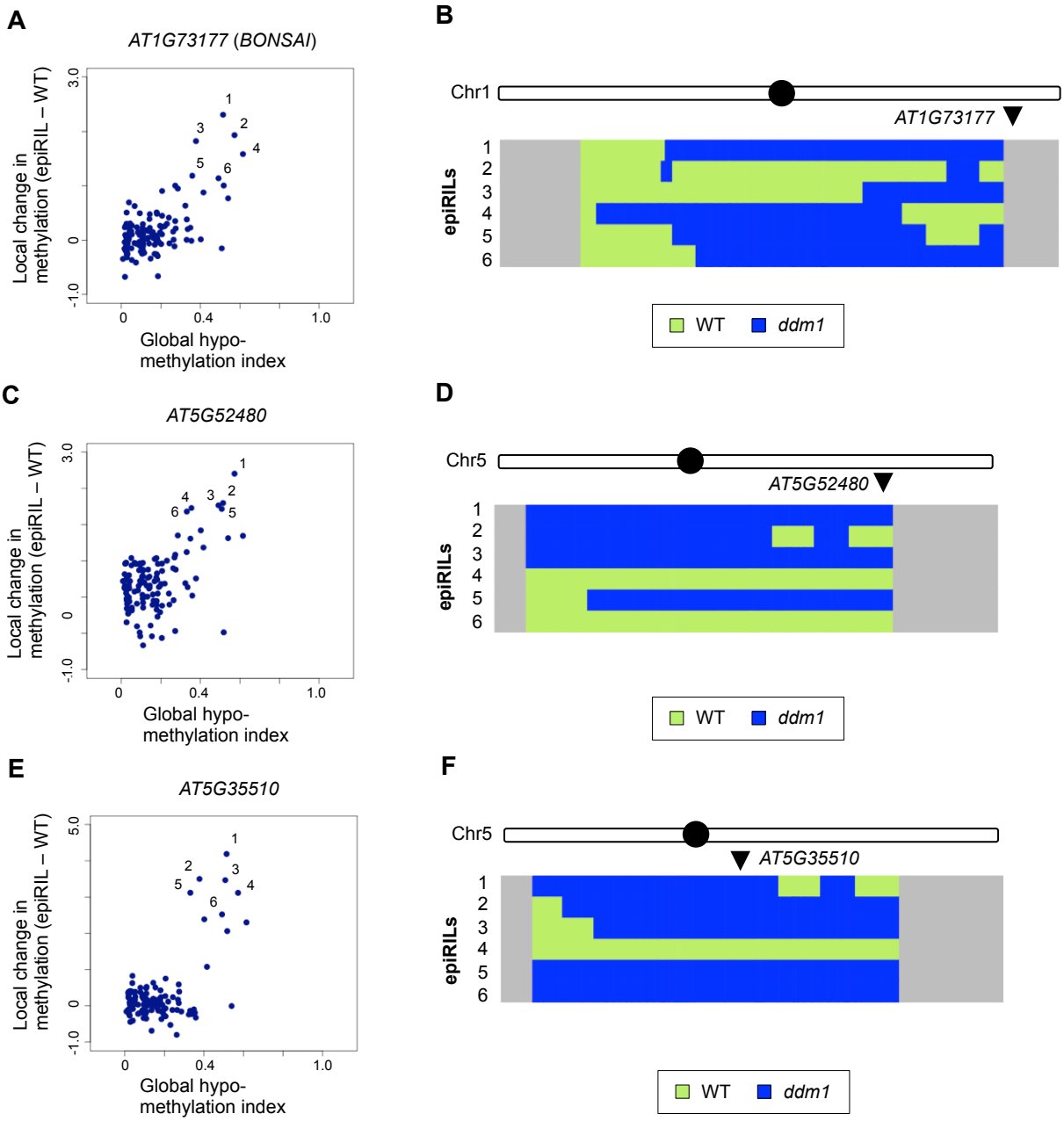


Figure 22. Effects of disrupted heterochromatin in the *DDM1* wild type background on DNA methylation examined by IP. (A, C, E) Changes in local DNA methylation plotted against the global level of DNA hypomethylation in 123 epiRILs. Each dot represents the value for one line. Three loci, *AT1G73177* (*BONSAI*) (A), *AT5G52480* (C), and *AT5G35510* (E) are shown. (B, D, F) WT (light green) / *ddm1* (dark blue) haplotype for epiRILs that showed increase of cytosine methylation for each locus (numbered 1-6 for each locus). In each panel, the chromosome including the target locus (arrowhead) is shown. The filled circles indicate centromere positions. The regions not covered by any markers are indicated in gray. Names of epiRILs numbered 1-6 in each panel are in Materials and Methods. Data of epiRILs were obtained from GEO (GSE37284: Colomé-Tatché et al., 2012).

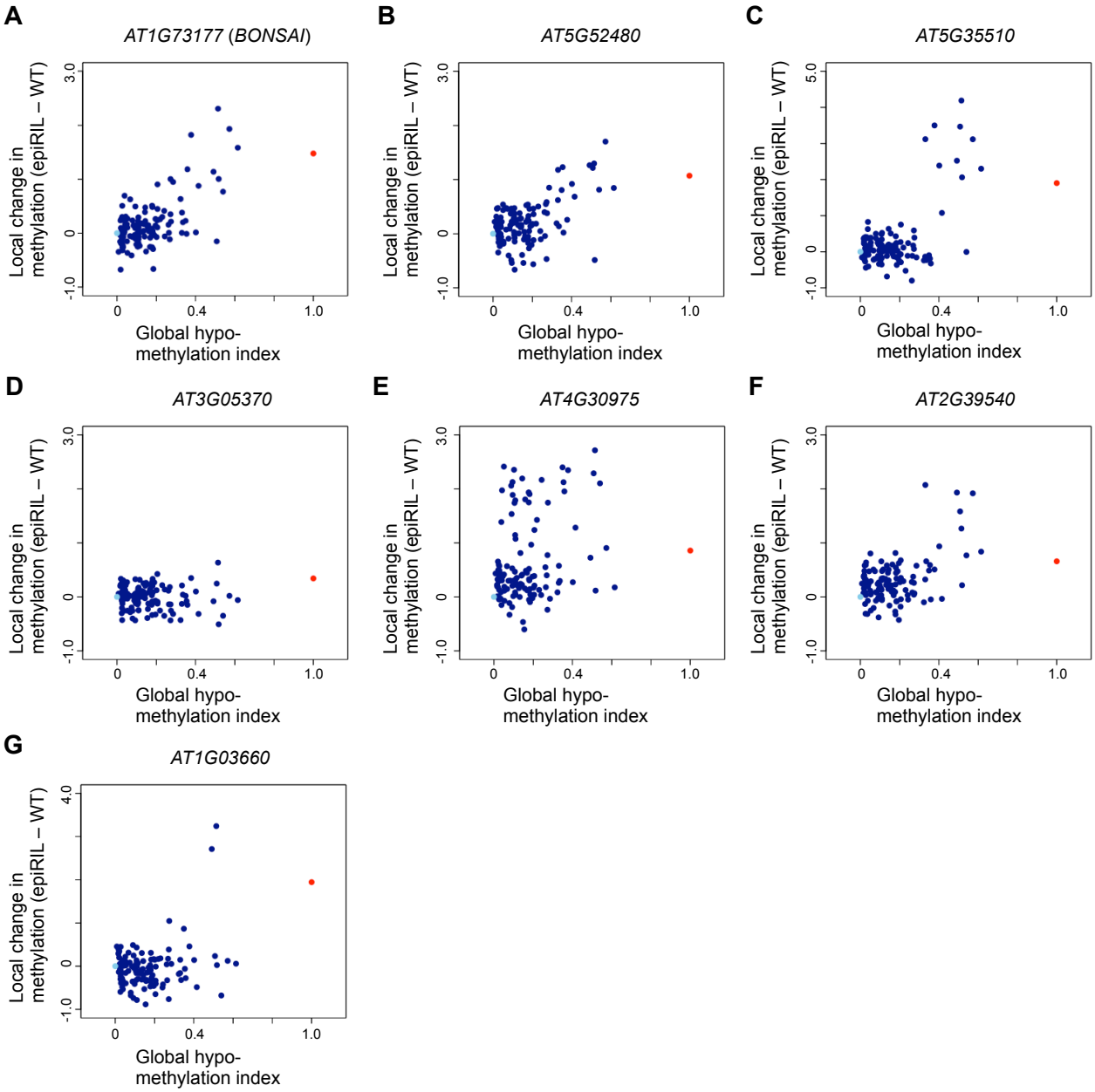


Figure 23. Change of DNA methylation in epiRILs.
For seven loci, changes of local DNA methylation level were plotted against the global hypomethylation as shown in Figure 22. Dots of light blue and red are values for parental *DDM1* and 4G *ddm1* plants, respectively.

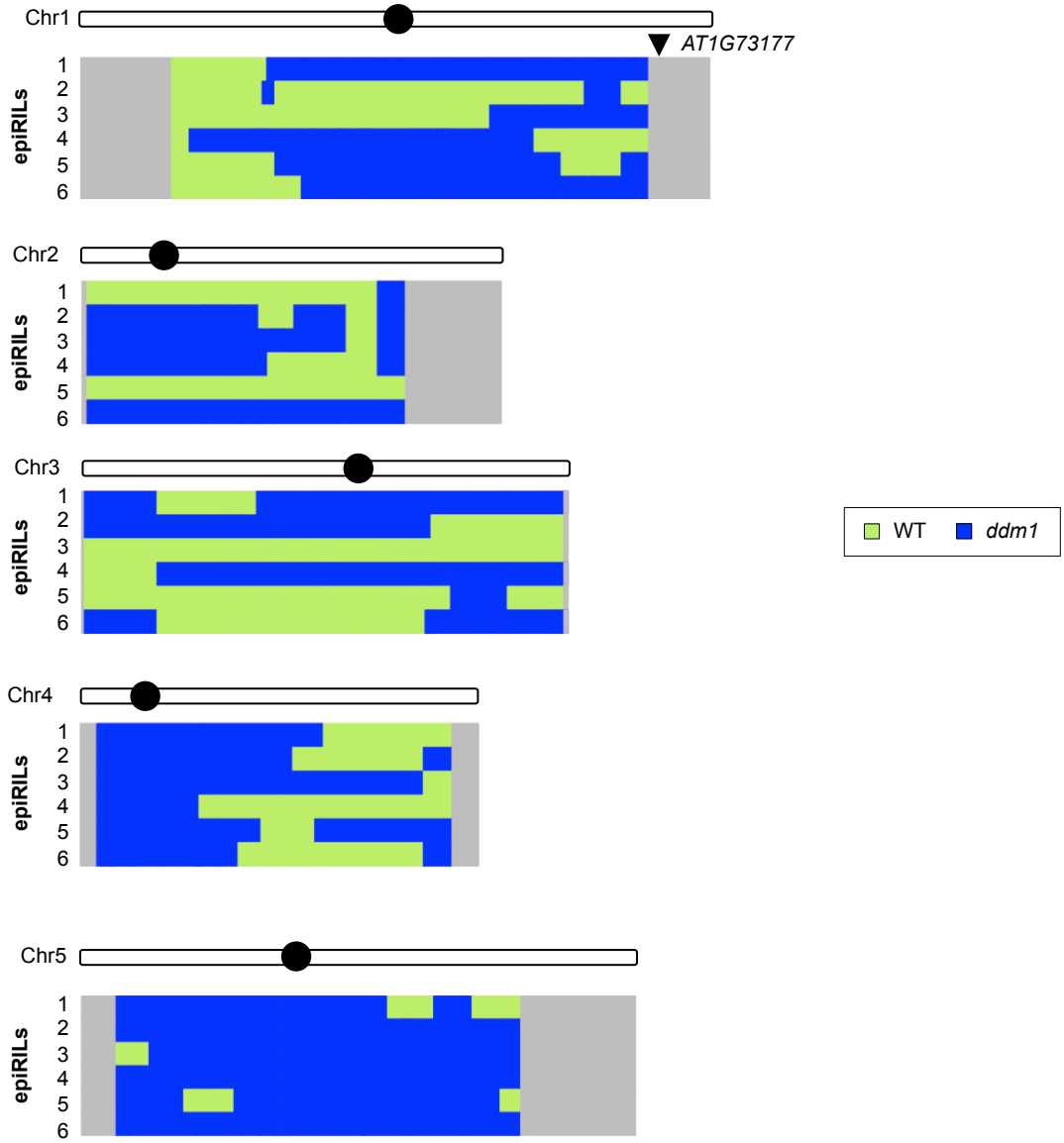


Figure 24. The haplotypes of epiRILs that showed increase of cytosine methylation in *AT1G73177*. Inference of the haplotypes in epiRILs that showed increase of cytosine methylation in *AT1G73177* are shown for all five chromosomes.

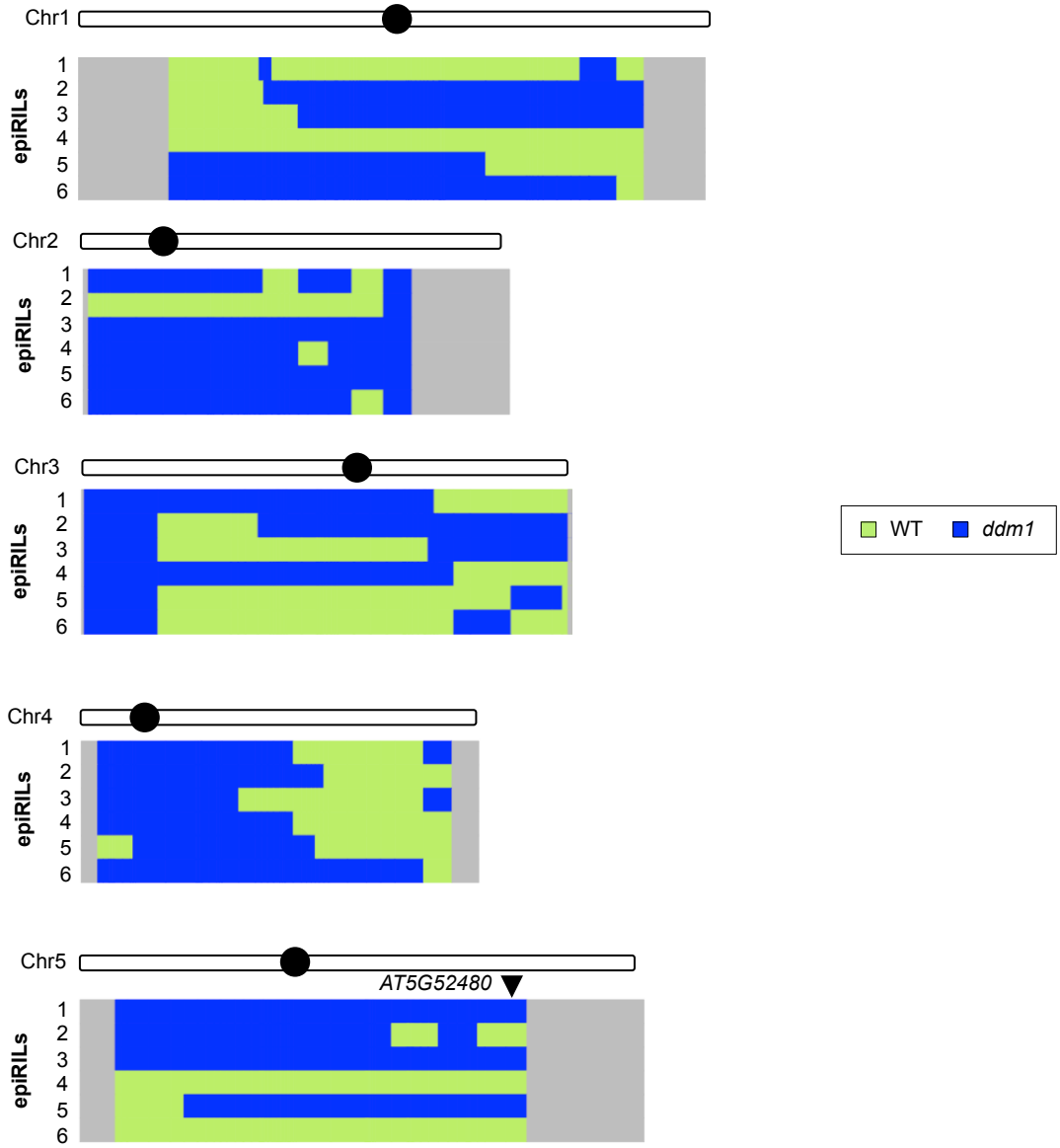


Figure 25. The haplotypes of epiRILs that showed increase of cytosine methylation in *AT5G52480*. Inference of the haplotypes in epiRILs that showed increase of cytosine methylation in *AT5G52480* are shown for all five chromosomes.

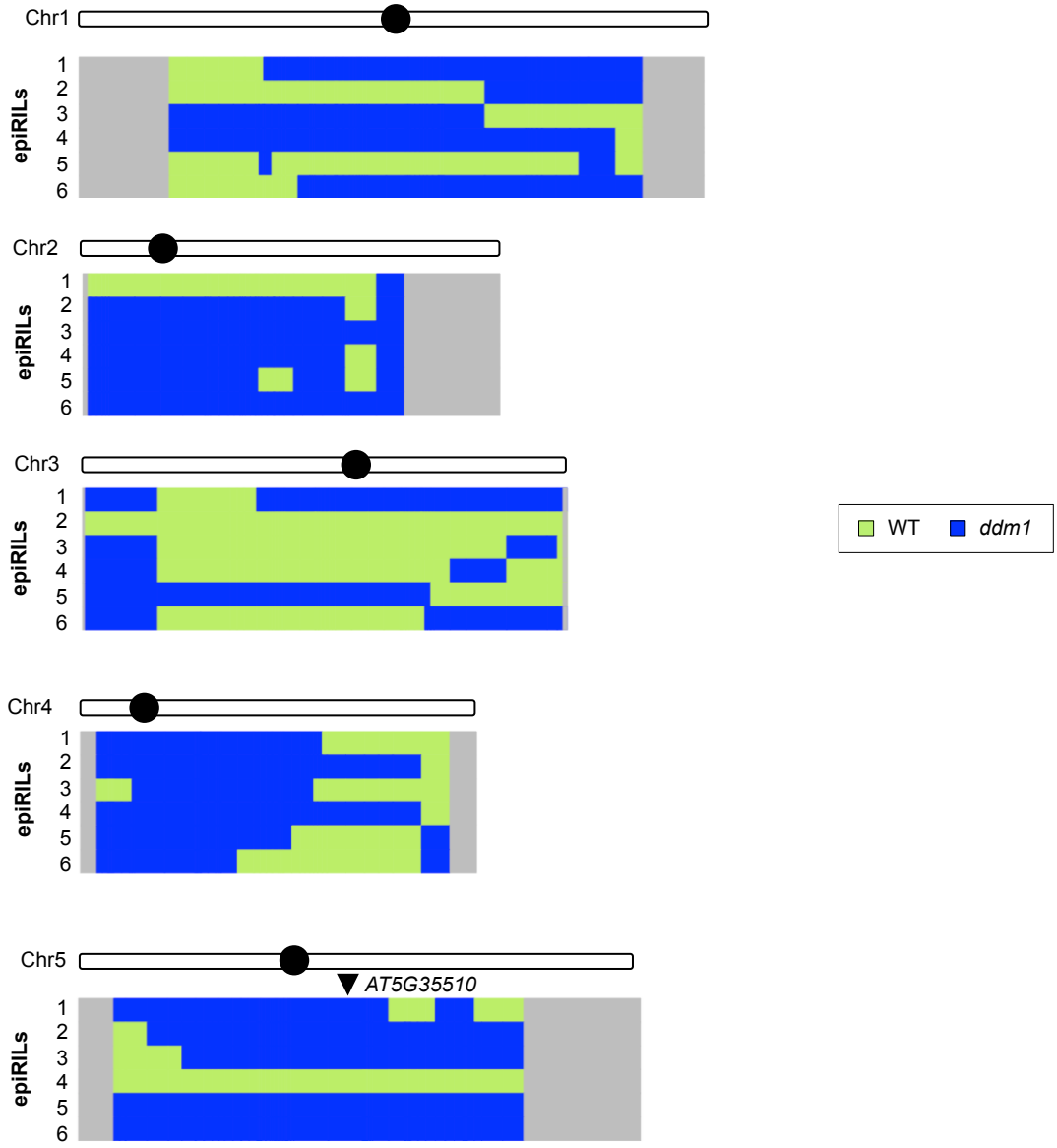


Figure 26. The haplotypes of epiRILs that showed increase of cytosine methylation in *AT5G35510*. Inference of the haplotypes in epiRILs that showed increase of cytosine methylation in *AT5G35510* are shown for all five chromosomes.

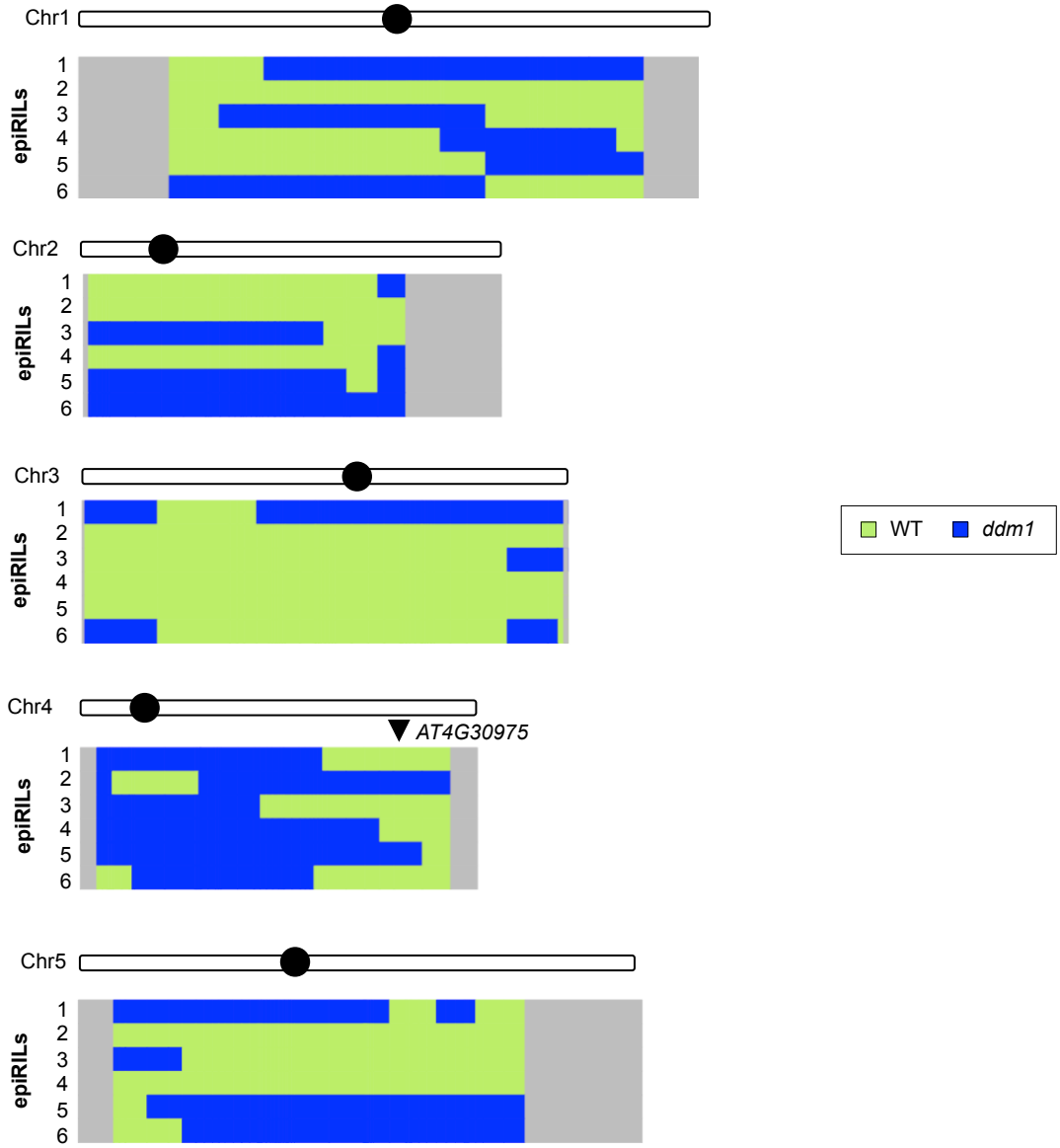


Figure 27. The haplotypes of epiRILs that showed increase of cytosine methylation in *AT4G30975*. Inference of the haplotypes in epiRILs that showed increase of cytosine methylation in *AT4G30975* are shown for all five chromosomes.

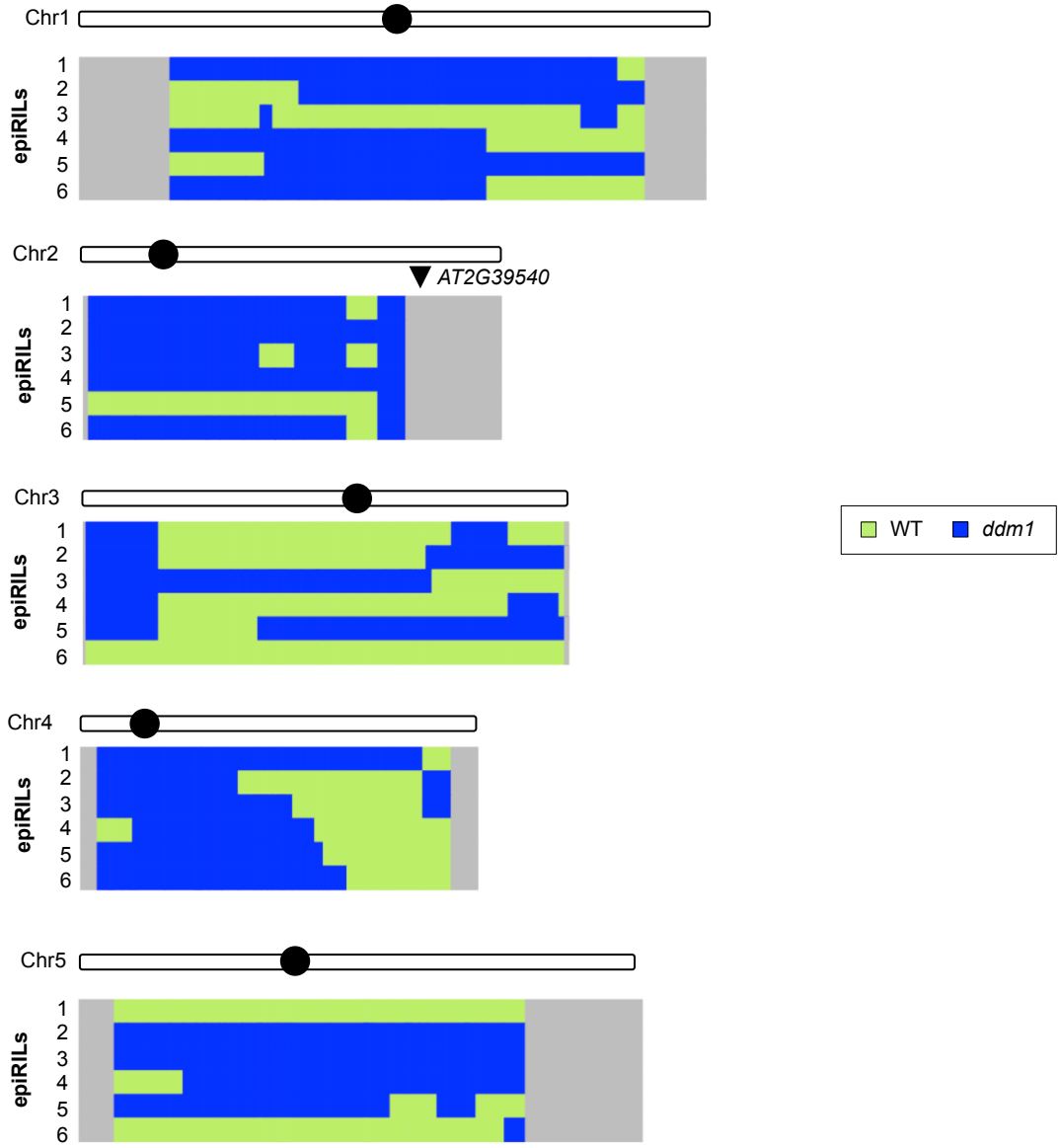


Figure 28. The haplotypes of epiRILs that showed increase of cytosine methylation in *AT2G39540*. Inference of the haplotypes in epiRILs that showed increase of cytosine methylation in *AT2G39540* are shown for all five chromosomes.

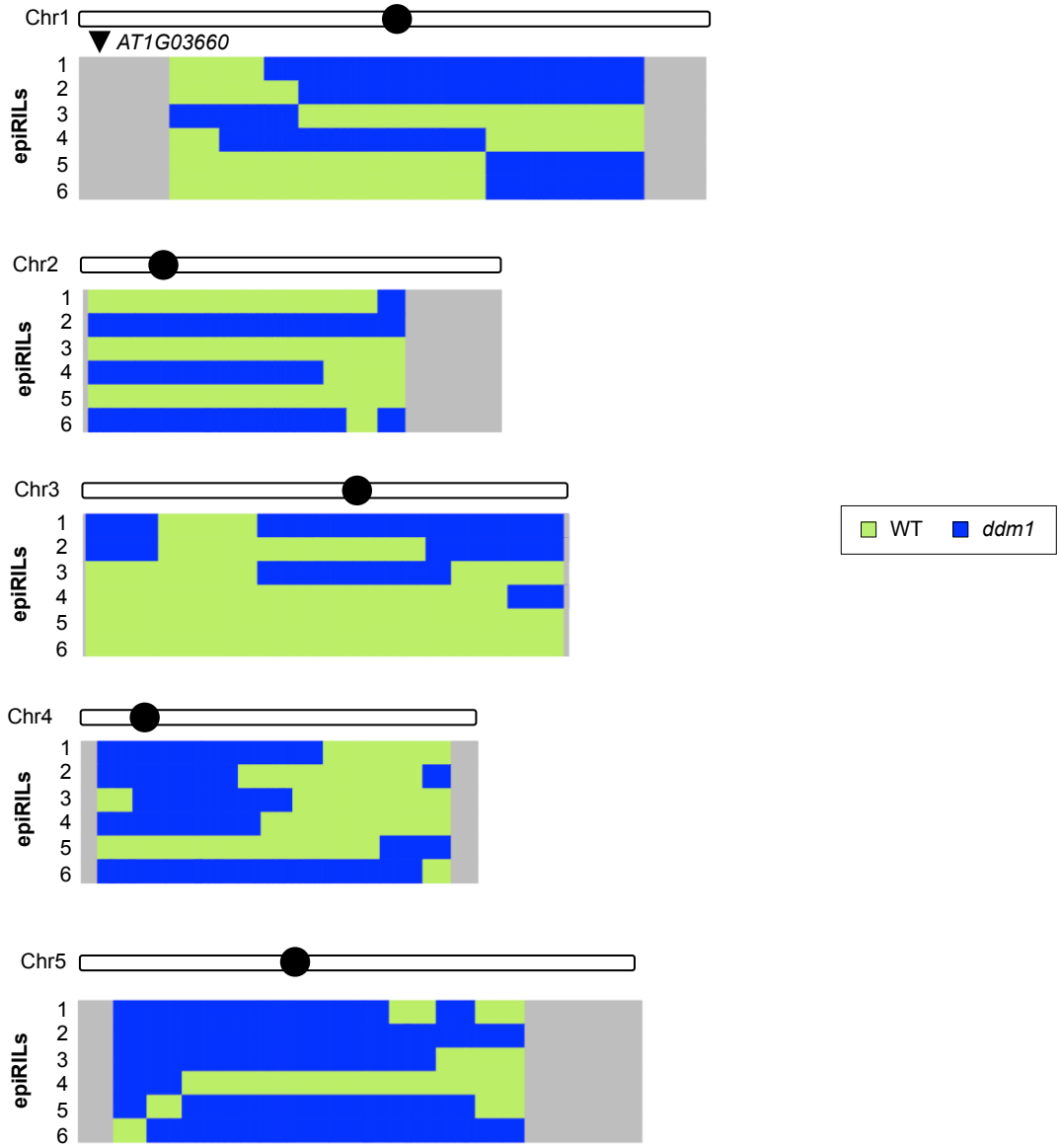


Figure 29. The haplotypes of epiRILs that showed increase of cytosine methylation in *AT1G03660*. Inference of the haplotypes in epiRILs that showed increase of cytosine methylation in *AT1G03660* are shown for all five chromosomes.

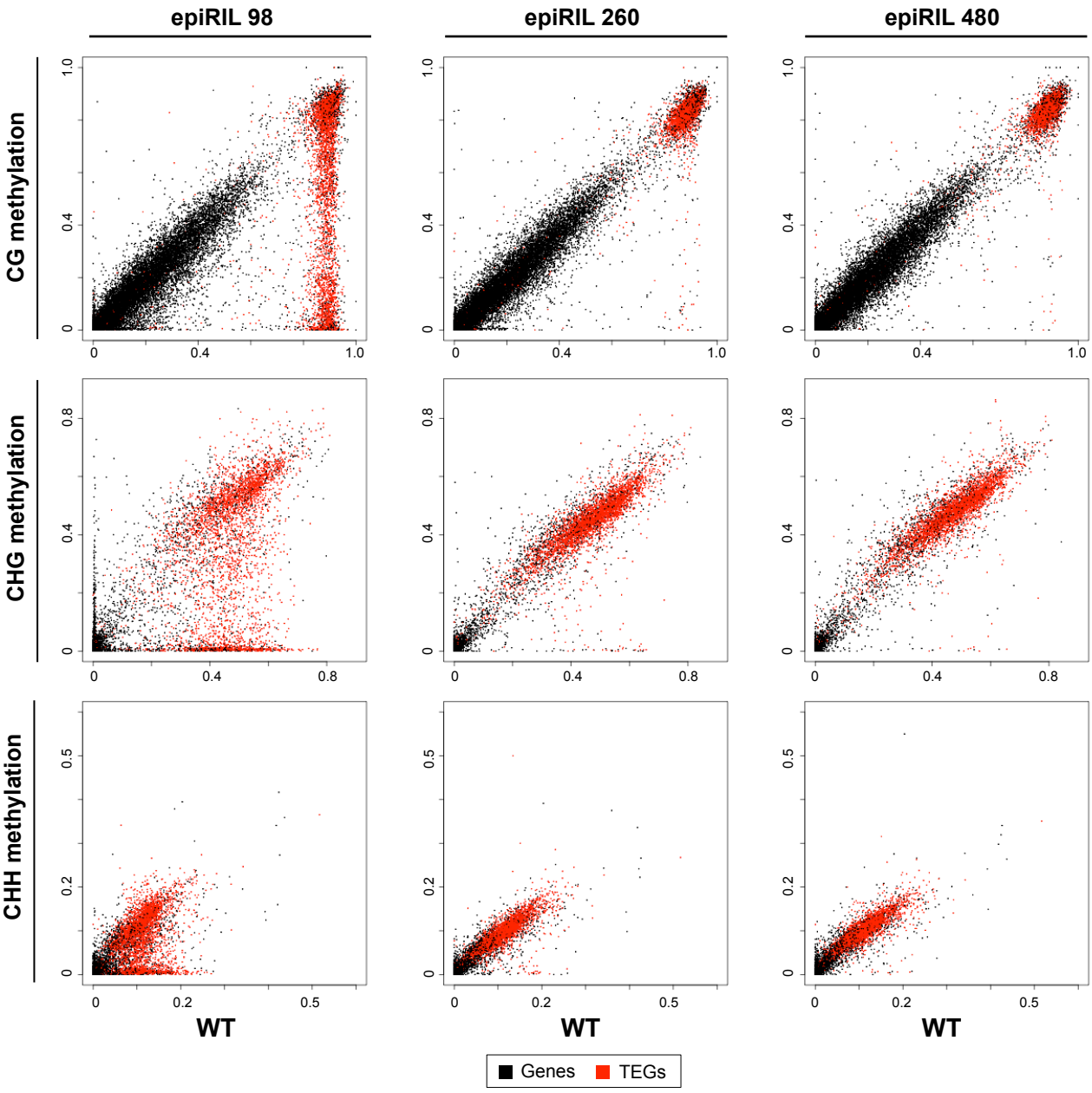


Figure 30. Effects of disrupted heterochromatin in the *DDM1* wild type background examined at single base resolution. Methylation level was compared for each transcription unit in CG, CHG, and CHH contexts. The format is as shown in Figure 3. A globally hypomethylated epiRIL (epiRIL98: plant #3 in Figure 22AB and plant #2 in Figure 22EF) and two epiRILs with lower level of hypomethylation (epiRIL260 and epiRIL480) are shown. Global hypomethylation indexes of epiRIL98, epiRIL260, and epiRIL480 are 0.38, 0.04, and 0.09, respectively. "WT" data are from the parental wild-type Col plant used to generate the epiRILs.

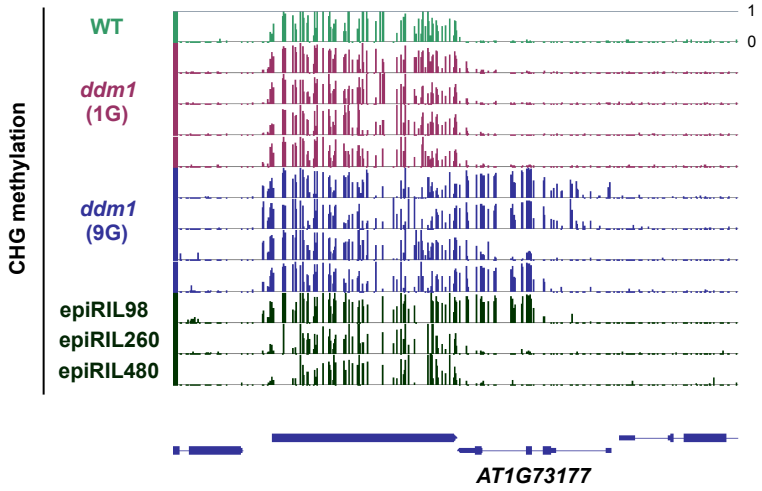


Figure 31. Ectopic non-CG methylation in *BONSAI* locus in epiRIL98.
Genome browser views of CHG methylation in *BONSAI* (*AT1G73177*) locus. CHG methylation increased in the 9G *ddm1* plants and also in epiRIL98.

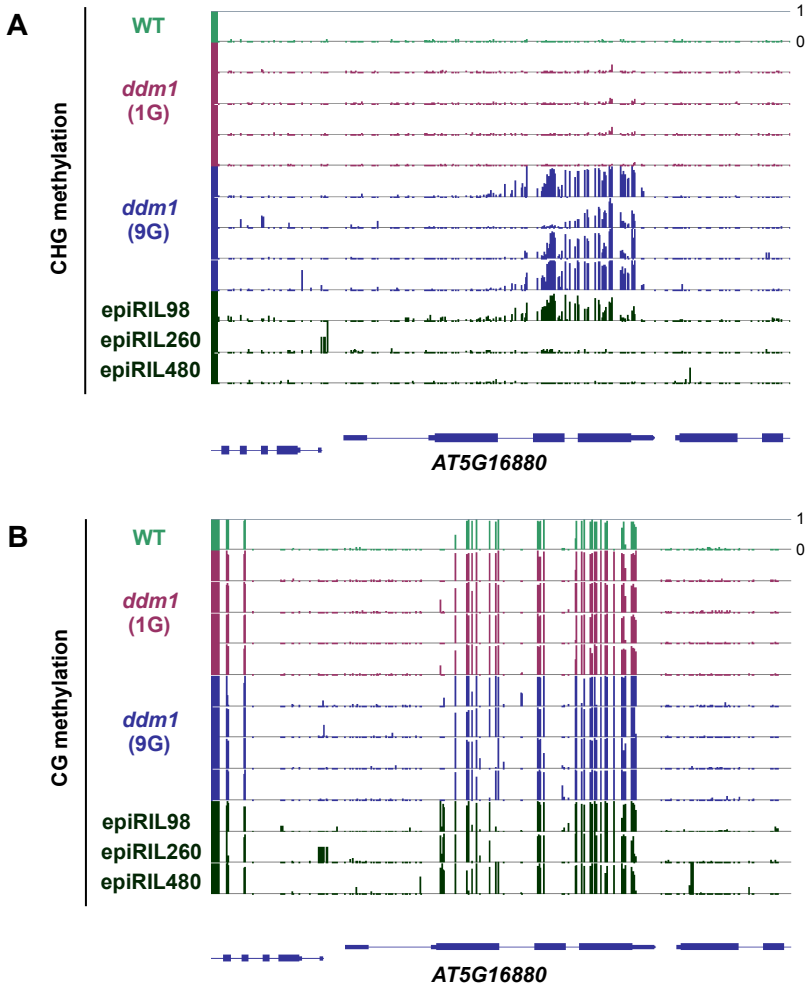


Figure 32. Ectopic non-CG methylation of CG methylated locus was found in epiRIL98.
Genome browser views of DNA methylation in *AT5G16880* locus in CHG (A) and CG (B) contexts. This locus has a high level of CG methylation (B). CHG methylation increased in the 9G *ddm1* plants and also in epiRIL98.

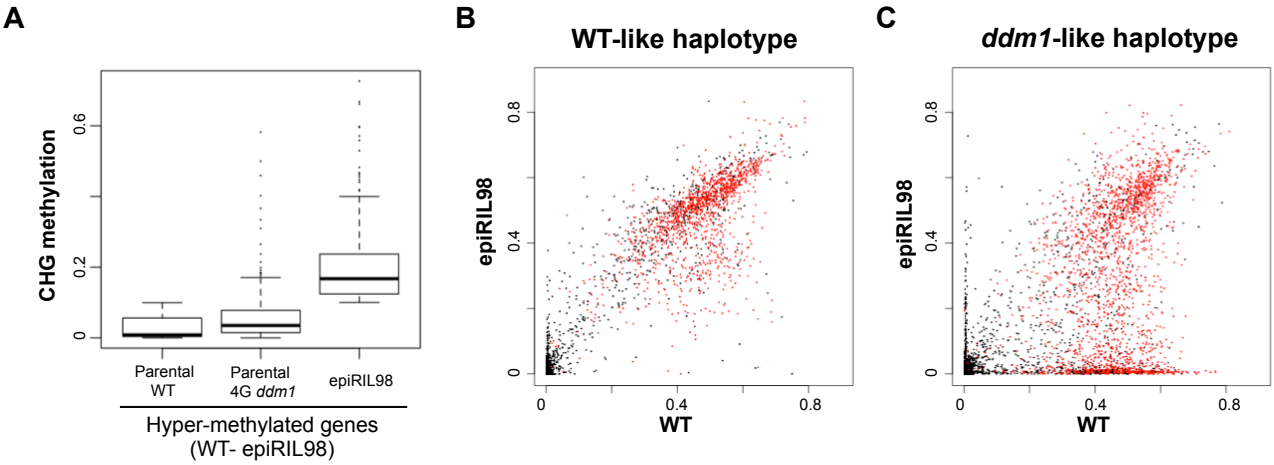


Figure 33. Effects of disrupted heterochromatin in epiRIL98.
(A) CHG methylation levels in the genes that were not methylated in WT but methylated in epiRIL98 (methylation level < 0.1 in WT and ≥ 0.1 in epiRIL98: n = 232). For these transcription units, distributions of the methylation levels were compared among the parental WT, the parental 4G *ddm1* plant, and the epiRIL98. (B-C) CHG methylation level in each transcription unit was compared between epiRIL98 and WT. Each gene was assigned to the inferred haplotypes in epiRIL98: WT-like (B) or *ddm1*-like (C).

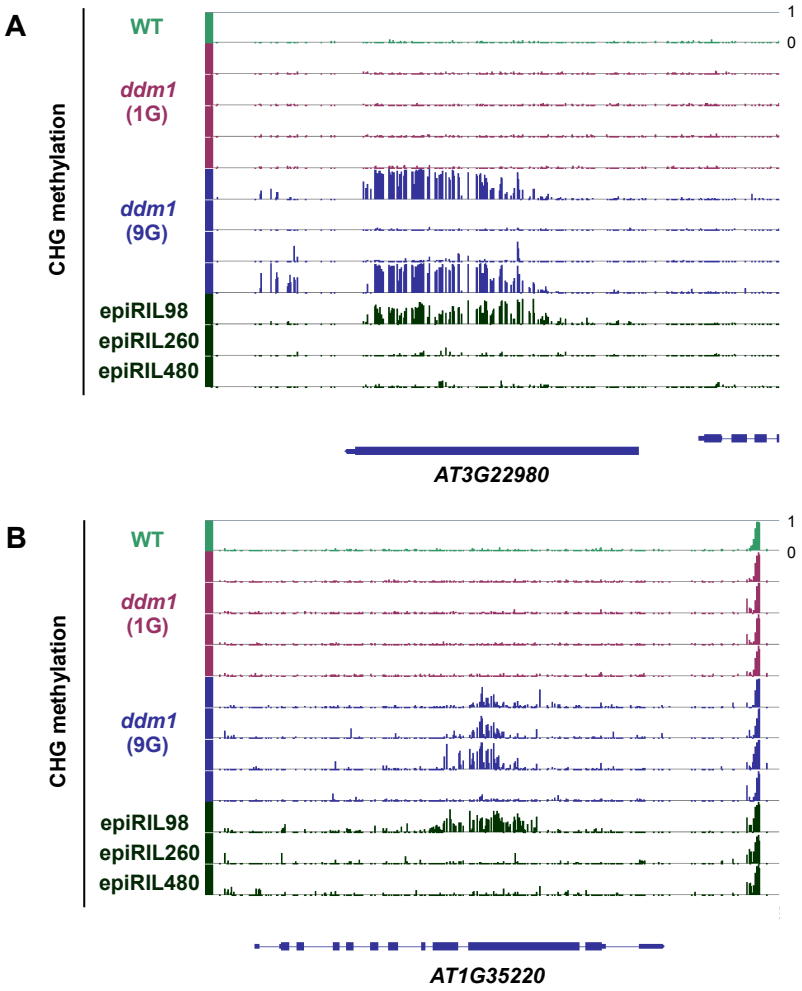


Figure 34. Ectopic non-CG methylation found in WT-like chromosome in epiRIL98. Genome browser views of CHG methylation in *AT3G22980* locus (A) and *AT1G35220* (B) locus. These loci are in the WT-like haplotype in epiRIL98.

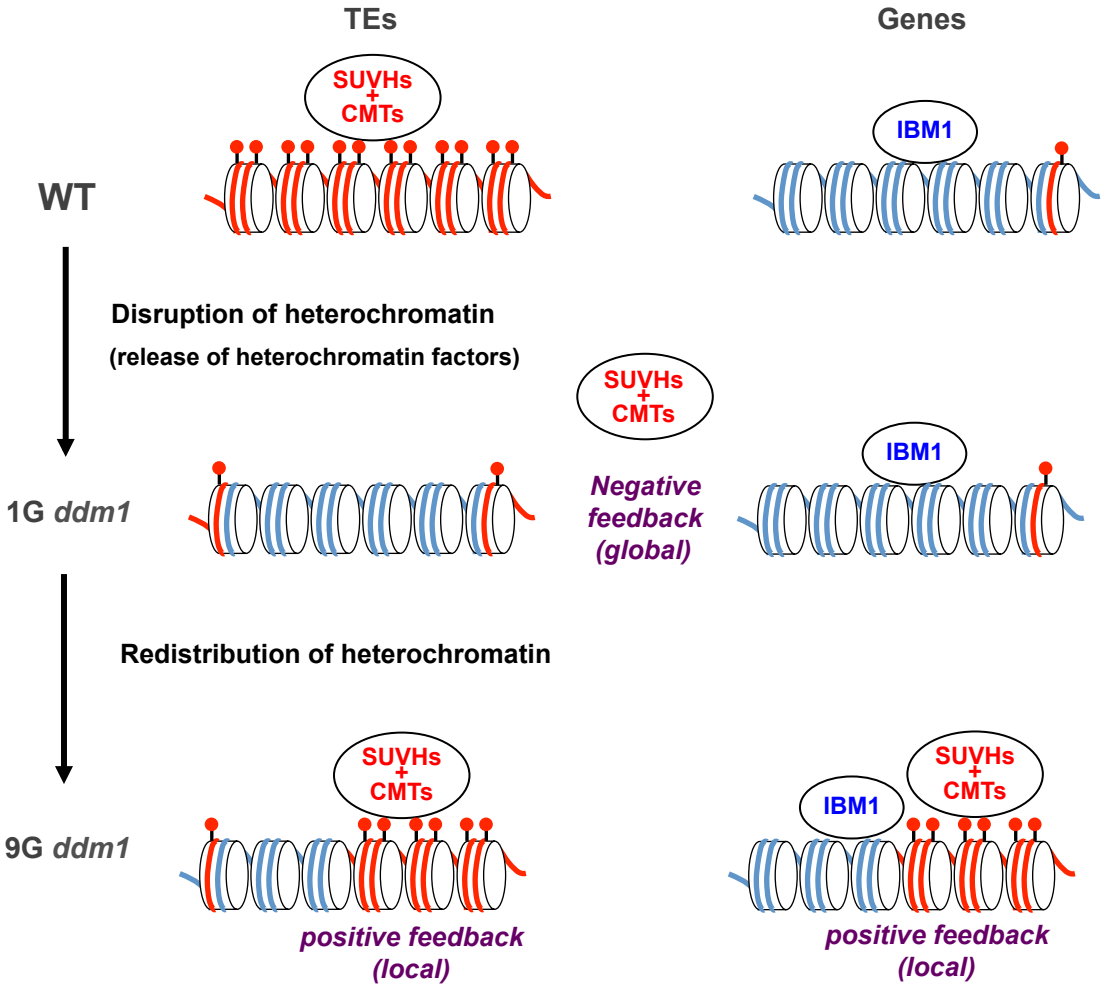


Figure 35. A model for the transgenerational heterochromatin redistribution. The cylinder indicates a nucleosome. Red dots above the nucleosome indicate methylation of H3K9. Red and blue lines indicate DNA with and without non-CG methylation, respectively. The CMTs are non-CG methylases, such as CMT3 and CMT2 (Zemach et al., 2013; Stroud et al., 2014). SUVHs are H3K9 methylases, such as SUVH4/KYP, SUVH5 and SUVH6 (Ebbs & Bender, 2009). In both WT and *ddm1* mutant plants, the histone demethylase IBM1 removes H3K9me from transcribed genes.

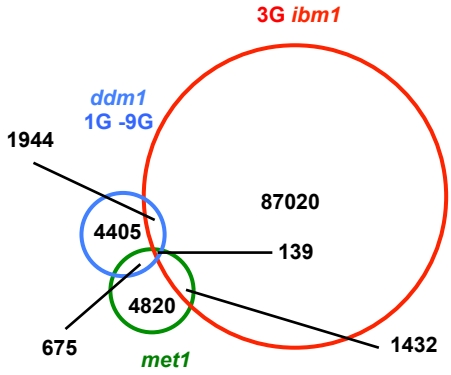


Figure 36. The difference between the effect of self-pollination of *ddm1* mutation and that of *met1* mutation. Overlap of regions CHG hyper-methylated in *met1*, 3G *ibm1* and 9G *ddm1*. DMRs between 1G and 9G *ddm1* (blue), between WT and 1G *met1* (green; Data were obtained from GEO (GSE39901: Stroud et al., 2013), and between WT and 3G *ibm1* (red) are shown.

Chapter 5

General Conclusion and Discussion

I revealed short- and long-term effects of the *ddm1* mutation. *ddm1* mutation induces a drastic loss of DNA methylation in heterochromatic regions in the first generation. In later generations, *ddm1* mutants reproducibly exhibited an increase in DNA methylation in hundreds of genes and TEGs. The increase was evident at both non-CG and CG sites. Importantly and unexpectedly, the *ddm1*-induced increase in DNA methylation was absent in the first generation and specifically seen in the later generations.

Moreover, in the epiRILs, the hypomethylated chromosome segments derived from the parental *ddm1* induced local hypermethylation in wild-type *DDM1* background. This result strongly suggests the existence of a negative feedback mechanism for DNA methylation dynamics. This could be one of the background mechanisms controlling long-term DNA methylation dynamics in Arabidopsis.

Analogous mechanisms may also be operating in other eukaryotes. Mice with a disruption of its *DDM1* homolog *Lsh* show reduced genomic DNA methylation, but interestingly it is also associated with increased DNA methylation at specific regions (Tao et al., 2011). In human cancer, hypomethylation of repeats and TEs are often associated with local hypermethylation of genes, such as tumor suppressor genes (Ehrlich, 2009; Ross et al., 2010). In Drosophila, an increase in the amount of

heterochromatic Y chromosome can results in a release of silencing at multiple loci in trans (Dimitri & Pisano, 1989), suggesting a negative feedback similar to that discussed here. Furthermore, Drosophila modifiers of position effect variegation often function in dosage-dependent manners (Locke et al., 1988; Henikoff, 1996), consistent with the pathway proposed in Figure 35.

Positive feedback loops would stabilize and enhance silent and active states (Zilberman et al., 2004; Johnson et al., 2007; Inagaki et al., 2010; Inagaki & Kakutani, 2013), but they carry the risk of going out of control to excess. A global negative feedback mechanism, together with the local positive feedback, would ensure a robust and balanced chromatin differentiation within the genome, as has been discussed for pattern formation during development (Turing, 1953; Meinhardt & Gierer, 2000).

In the context of evolution in plants, a large variation in the amount of repetitive sequences is often noted between related species or even within a species (Cullis, 2005; Hawkins et al., 2006; Woo & Richards, 2008). On such occasions, fine-tuning of the amounts of the trans-acting heterochromatin factors would be especially important, as an imbalance would not only immediately affect gene expression level but also influence the epigenotype in a transgenerational manner.

References

Arteaga-Vazquez MA, and Chandler VL (2010) Paramutation in maize: RNA mediated trans-generational gene silencing. *Curr Opin Genet Dev* 20: 156-163

Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, and Weigel D (2011) Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* 480: 245-249.

Cao X, Aufsatz W, Zilberman D, Mette MF, Huang MS, Matzke M, and Jacobsen SE (2003) Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr Biol* 13: 2212-2217.

Coleman-Derr D and Zilberman D (2012) Deposition of histone variant H2A.Z within gene bodies regulates responsive genes. *PLoS Genet* 8: e1002988. doi: 10.1371/journal.pgen.1002988.

Colomé-Tatché M, Cortijo S, Wardenaar R, Morgado L, Lahouze B, Sarazin A, Etcheverry M, Martin A, Feng S, Duvernois-Berthet E, Labadie K, Wincker P, Jacobsen SE, Jansen RC, Colot V, and Johannes F (2012) Features of the Arabidopsis recombination landscape resulting from the combined loss of sequence variation and DNA methylation. *Proc Natl Acad Sci U S A* 109:16240-16245.

Cullis CA (2005) Mechanisms and control of rapid genomic changes in flax. *Ann Bot* 95: 201-206.

de Hoon MJ, Imoto S, Nolan J, and Miyano S (2004) Open source clustering software. *Bioinformatics* 20:1453-1454.

Deleris A, Stroud H, Bernatavichute Y, Johnson E, Klein G, Schubert D, and Jacobsen SE (2012) Loss of the DNA methyltransferase MET1 Induces H3K9 hypermethylation at PcG target genes and redistribution of H3K27 trimethylation to transposons in *Arabidopsis thaliana*. *PLoS Genet* 8: e1003062.

Dennis K, Fan T, Geiman T, Yan Q, and Muegge K (2001) Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev* 15: 2940-2944.

Dimitri P and Pisano C (1989) Position effect variegation in Drosophila melanogaster: relationship between suppression effect and the amount of Y chromosome. *Genetics* 122: 793-800.

Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A, Hetzel J, Wohlschlegel JA, Pradhan S, Patel DJ, and Jacobsen SE (2012) Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell* 151:167-180.

Ebbs ML and Bender J (2006) Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. *Plant Cell* 18:1166-1176.

Ehrlich M (2009) DNA hypomethylation in cancer cells. *Epigenomics* 1: 239-259.

Finnegan EJ, Peacock WJ, and Dennis ES (1996) Reduced DNA methylation in

Arabidopsis thaliana results in abnormal plant development. *Proc Natl Acad Sci U S A* 93: 8449-8454.

Fu Y, Kawabe A, Etcheverry M, Ito T, Toyoda A, Fujiyama A, Colot V, Tarutani Y, and Kakutani T (2013) Mobilization of a plant transposon by expression of the transposon-encoded anti-silencing factor. *EMBO J* 32: 2407-2417.

Furner IJ and Matzke M (2011) Methylation and demethylation of the Arabidopsis genome. *Curr Opin Plant Biol* 14: 137-141.

Hawkins JS, Kim H, Nason JD, Wing RA, and Wendel JF (2006) Differential lineage-specific amplification of transposable elements is responsible for genome size variation in Gossypium. *Genome Res* 16: 1252-1261.

Heard E and Martienssen RA (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell* 157: 95-109.

Henderson IR, Deleris A, Wong W, Zhong X, Chin HG, Horwitz GA, Kelly KA, Pradhan S, and Jacobsen SE (2010) The de novo cytosine methyltransferase DRM2 requires intact UBA domains and a catalytically mutated paralog DRM3 during RNA-directed DNA methylation in *Arabidopsis thaliana*. *PLoS Genet* 6: e1001182. doi: 10.1371/journal.pgen.1001182.

Henikoff S (1996) Dosage-dependent modification of position-effect variegation in *Drosophila*. *Bioessays* 18: 401-409.

Inagaki S, Kakutani T (2013) What triggers differential DNA methylation of genes and TEs: contribution of body methylation? *Cold Spring Harb Symp Quant Biol* 2012 77:155-160.

Inagaki S, Miura-Kamio A, Nakamura Y, Lu F, Cui X, Cao X, Kimura H, Saze H, and Kakutani T (2010) Autocatalytic differentiation of epigenetic modifications within the Arabidopsis genome. *EMBO J* 29: 3496-3506.

Jacobsen SE, Meyerowitz E (1997) Hypermethylated *SUPERMAN* Epigenetic Alleles in *Arabidopsis*. *Science* 277: 1100-1103.

Jeddeloh JA, Stokes TL, Richards EJ (1999) Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nat Genet* 22: 94-97.

Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, Bulski A, Albuisson J, Heredia F, Audigier P, Bouchez D, Dillmann C, Guerche P, Hospital F, and Colot V (2009) Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet* 5: e1000530.

Johnson LM, Bostick M, Zhang X, Kraft E, Henderson I, Callis J, Jacobsen SE (2007) The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr Biol* 17: 379-384.

Kakutani T (1997) Genetic characterization of late-flowering traits induced by DNA hypomethylation mutation in *Arabidopsis thaliana*. *Plant J* 12: 1447-1451.

Kakutani T (2002) Epi-alleles in plants: inheritance of epigenetic information over generations. *Plant Cell Physiol* 43: 1106-1111.

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ (1996) Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proc Natl Acad Sci U S A* 22: 12406-12411.

Kakutani T, Munakata K, Richards EJ, Hirochika H (1999) Meiotically and mitotically stable inheritance of DNA hypomethylation induced by ddm1 mutation of *Arabidopsis thaliana*. *Genetics* 151: 831-838.

Kankel MW, Ramsey DE, Stokes TL, Flowers SK, Haag JR, Jeddeloh JA, Riddle NC, Verbsky ML and Richards EJ (2003) Arabidopsis *MET1* cytosine methyltransferase mutants. *Genetics* 163: 1109-1122.

Kelly WG (2014) Transgenerational epigenetics in the germline cycle of *Caenorhabditis elegans*. *Epigenetics Chromatin* 7: 6.

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

Kishimoto N, Sakai H, Jackson J, Jacobsen SE, Meyerowitz EM, Dennis ES and Finnegan EJ (2001) Site specificity of the Arabidopsis METI DNA methyltransferase demonstrated through hypermethylation of the *superman* locus. *Plant Mol Biol* 46: 171-183.

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A and Huala E (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res* 40: D1202-1210.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25.

Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11:204-220.

Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, Carrington JC, Doerge RW, Colot V. and Martienssen R (2004) Role of transposable elements in heterochromatin and epigenetic control. *Nature* 430: 471-476.

Locke J, Kotarski MA, Tartof KD (1988) Dosage-dependent modifiers of position effect variegation in Drosophila and a mass action model that explains their effect. *Genetics* 120: 181-198.

Luo C, Sidote DJ, Zhang Y, Kerstetter RA, Michael TP and Lam E (2013) Integrative analysis of chromatin states in Arabidopsis identified potential regulatory mechanisms for natural antisense transcript production. *Plant J* 73: 77–90.

Mathieu O, Reinders J, Caikovski M, Smathajitt C, Paszkowski J (2007) Transgenerational stability of the *Arabidopsis* epigenome is coordinated by CG

methylation. *Cell* 130: 851-862.

Meinhardt H, Gierer A (2000) Pattern formation by local self-activation and lateral inhibition. *Bioessays* 22: 753-760.

Meissner A. (2010) Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol* 28:1079-1088.

Mette MF, Aufsatz W, van der Winden J, Matzke MA, Matzke AJ. (2000) Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J* 19: 5194-5201.

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.

Nicol JW, Helt GA, Blanchard SG, Raja A, Loraine AE (2009) The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* 25: 2730-2731.

Pikaard CS, Haag JR, Pontes OM, Blevins T, Cocklin R (2012) A transcription fork model for Pol IV and Pol V-dependent RNA-directed DNA methylation. *Cold Spring Harb Symp Quant Biol* 77:205-212.

Richards EJ (2011) Natural epigenetic variation in plant species: a view from the field. *Curr Opin Plant Biol* 14: 204-209.

Rigal M, Kevei Z, Pélissier T, Mathieu O (2012) DNA methylation in an intron of the IBM1 histone demethylase gene stabilizes chromatin modification patterns. *EMBO J* 31: 2981-2993.

Ross JP, Rand KN, Molloy PL (2010) Hypomethylation of repeated DNA sequences in cancer. *Epigenomics* 2: 245-269.

Saldanha AJ (2004) Java Treeview--extensible visualization of microarray data. *Bioinformatics* 20: 3246-3248.

Sasaki T, Kobayashi A, Saze H, Kakutani T (2012) RNAi-independent *de novo* DNA methylation revealed in Arabidopsis mutants of chromatin remodeling gene DDM1. *Plant J* 70: 750-758.

Saze H, Kakutani T (2007) Heritable epigenetic mutation of a transposon-flanked *Arabidopsis* gene due to lack of the chromatin-remodeling factor DDM1. *EMBO J* 26: 3641-3652.

Saze H, Kakutani T. (2011) Differentiation of epigenetic modifications between transposons and genes. *Curr Opin Plant Biol* 14:81-87.

Saze H, Shiraishi A, Miura A, Kakutani T (2008) Control of genic DNA methylation by a jmjC domain-containing protein in *Arabidopsis thaliana*. *Science* 319: 462-465.

Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urich MA, Libiger O, Schork NJ and Ecker JR (2011) Transgenerational epigenetic instability is a source of novel

methylation variants. *Science* 334: 369-373.

Schultz MD, Schmitz RJ, Ecker JR (2012) 'Leveling' the playing field for analyses of single-base resolution DNA methylomes. *Trends Genet* 28: 583-585.

Singer T, Yordan C, Martienssen RA (2001) Robertson's Mutator transposons in A. thaliana are regulated by the chromatin-remodeling gene *Decrease in DNA Methylation* (*DDM1*). *Genes Dev* 15: 591-602.

Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M and Peeters AJ (2000) The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol Cell* 4: 791-802

Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ and Jacobsen SE (2014) Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nat Struct Mol Biol* 21: 64-72.

Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE (2013) Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* 152: 352-364.

Tao Y, Xi S, Shan J, Maunakea A, Che A, Briones V, Lee EY, Geiman T, Huang J, Stephens R, Leighty RM, Zhao K and Muegge K (2011) Lsh, chromatin remodeling family member, modulates genome-wide cytosine methylation patterns at nonrepeat sequences. *Proc Natl Acad Sci U S A* 108: 5626-5631.

Teixeira FK, Heredia F, Sarazin A, Roudier F, Boccara M, et al. (2009) A role for RNAi in the selective correction of DNA methylation defects. *Science* 323: 1600-1604.

Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A and Kakutani T (2009) Bursts of retrotransposition reproduced in *Arabidopsis*. *Nature* 461: 423-426.

Turing AM (1953) The chemical basis of morphogenesis. *Philos Trans R Soc Lond B Biol Sci* 237: 37-72.

Vongs A, Kakutani T, Martienssen RA, Richards EJ (1993) *Arabidopsis thaliana* DNA methylation mutants. *Science* 260: 1926-1928.

Weigel D, Colot V (2012) Epialleles in plant evolution. *Genome Biol* 13: 249.

Woo HR, Richards EJ (2008) Natural variation in DNA methylation in ribosomal RNA genes of Arabidopsis thaliana. *BMC Plant Biol* 8: 92.

Yi H, Richards EJ (2009) Gene duplication and hypermutation of the pathogen Resistance gene *SNC1* in the Arabidopsis *bal* variant. *Genetics* 183: 1227-1234.

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

Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE and Ecker JR (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* 126: 1189-1201.

Zilberman D, Cao X, Johansen LK, Xie Z, Carrington JC and Jacobsen SE (2004) Role of *Arabidopsis ARGONAUTE4* in RNA-directed DNA methylation triggered by inverted repeats. *Curr Biol* 14: 1214-1220.