

# Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*

Hume Stroud<sup>1</sup>, Truman Do<sup>1</sup>, Jiamu Du<sup>2</sup>, Xuehua Zhong<sup>1,5</sup>, Suhua Feng<sup>1,3,4</sup>, Lianna Johnson<sup>1</sup>, Dinshaw J Patel<sup>2</sup> & Steven E Jacobsen<sup>1,3,4</sup>

DNA methylation occurs in CG and non-CG sequence contexts. Non-CG methylation is abundant in plants and is mediated by CHROMOMETHYLASE (CMT) and DOMAINS REARRANGED METHYLTRANSFERASE (DRM) proteins; however, its roles remain poorly understood. Here we characterize the roles of non-CG methylation in *Arabidopsis thaliana*. We show that a poorly characterized methyltransferase, CMT2, is a functional methyltransferase *in vitro* and *in vivo*. CMT2 preferentially binds histone H3 Lys9 (H3K9) dimethylation and methylates non-CG cytosines that are regulated by H3K9 methylation. We revealed the contributions and redundancies between each non-CG methyltransferase in DNA methylation patterning and in regulating transcription. We also demonstrate extensive dependencies of small-RNA accumulation and H3K9 methylation patterning on non-CG methylation, suggesting self-reinforcing mechanisms between these epigenetic factors. The results suggest that non-CG methylation patterns are critical in shaping the landscapes of histone modification and small noncoding RNA.

DNA methylation has roles in different biological processes such as gene regulation and imprinting. In *Arabidopsis*, DNA is methylated in three cytosine contexts: CG, CHG and CHH (where H denotes A, T or C)<sup>1</sup>. In mammals, DNA is primarily methylated in CG contexts; however, studies have uncovered the presence of non-CG methylation in certain cell types such as embryonic stem cells and brain cells<sup>2–7</sup>. In *Arabidopsis*, CG methylation is maintained by MET1, the plant homolog of DNMT1. CHG and CHH are site-specifically methylated by CMT3 and DRM2 (refs. 8,9). CMT3 is controlled by H3K9 methylation<sup>10–12</sup>. DRM2 is targeted to certain loci through an RNA-directed DNA methylation (RdDM) pathway involving 24-nt small interfering RNAs (siRNAs)<sup>1</sup>. Heterochromatin in *Arabidopsis* is enriched in both CG and non-CG methylation as well as in H3K9 methylation and 24-nt siRNAs; however, the relationships between each of these marks remain poorly understood.

The abundant non-CG methylation in plants compared to mammals may in part be explained by the presence of plant-specific CMT-encoding genes. In addition to CMT3, the *Arabidopsis* genome contains two other CMT-encoding genes: CMT1 and CMT2. CMT1 is expressed at low levels and is truncated in many *Arabidopsis* ecotypes<sup>13</sup>. CMT2 is expressed and is a putative DNA methyltransferase. A recent study performed whole-genome methylation profiling in *cmt2* mutants and found loss of CHH methylation predominantly at large transposable elements (TEs) that were heterochromatic<sup>9</sup>. Genetic evidence suggested that the chromatin remodeler DDM1 in part allows access for MET1, CMT3 and CMT2 to heterochromatin<sup>9</sup>. However, the mechanism

of CMT2 targeting to heterochromatin, its roles and its relationship with other DNA methyltransferases is not understood.

Here, we set out to characterize the roles of non-CG methylation. We first show that CMT2 is a functional non-CG methyltransferase. CMT2 preferentially methylates unmethylated DNA *in vitro* and methylates both CHG and CHH sites *in vitro* and *in vivo*. We find that CMT2 binds methylated H3K9 *in vitro* and that H3K9 methylation controls non-CG methylation through CMT2. We also uncover that the number of methyl groups on H3K9 may influence targeting of CMT2 and CMT3. Given the identification of CMT2 as a functional methyltransferase, we generated all possible combinations of non-CG methyltransferase mutants and examined the contributions and redundancies between each non-CG methyltransferase in DNA methylation patterning and gene silencing. Although it is clear that 24-nt siRNAs and H3K9 methylation guide non-CG methylation, we reveal extensive dependencies of both 24-nt siRNAs and H3K9 methylation patterning on non-CG methylation. This suggests that non-CG methylation has a critical role in regulating these marks. Furthermore, we find elevated histone acetylation levels throughout sites that lose non-CG methylation. Our results provide insights into targeting of non-CG methylation and will help to guide further studies of the biology of DNA methylation.

## RESULTS

### CMT2 strongly methylates both CHG and CHH sites *in vitro*

To examine whether CMT2 has a role in methylating the genome, we performed whole-genome bisulfite sequencing (BS-seq) in two

<sup>1</sup>Department of Molecular, Cell and Developmental Biology, University of California, Los Angeles, Los Angeles, California, USA. <sup>2</sup>Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. <sup>3</sup>Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, Los Angeles, California, USA. <sup>4</sup>Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles, California, USA. <sup>5</sup>Present address: Wisconsin Institute for Discovery, Laboratory of Genetics, University of Wisconsin–Madison, Madison, Wisconsin, USA. Correspondence should be addressed to S.E.J. (jacobsen@ucla.edu).

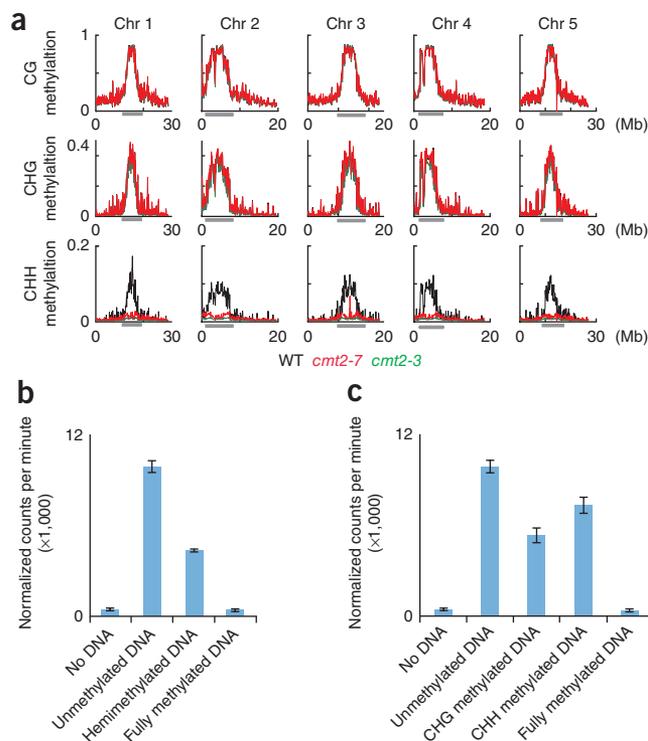
**Figure 1** *In vitro* activity of CMT2. (a) Fractional DNA methylation levels of cytosines in CG, CHG and CHH contexts across chromosomes (chr). Gray bars indicate pericentromeric heterochromatin. WT, wild type. (b) CMT2 *in vitro* methylation activity on DNA of different methylation status. The values for unmethylated and hemimethylated DNA are normalized according to the number of available (i.e., unmethylated) cytosines. Bars represent the range for two technical replicates. (c) CMT2 *in vitro* methylation activity on DNA of different methylation status, assessing sequence specificities of CMT2. Bars represent the range for two technical replicates.

different CMT2 transfer-DNA insertion mutants, *cmt2-7* and *cmt2-3* (ref. 8). We found that global CHH methylation was substantially reduced, whereas CG and CHG methylation were largely undisturbed (Fig. 1a), results consistent with a recent study<sup>9</sup>. For the rest of the study, we focused on *cmt2-7*, which we confirmed to be a null mutant by reverse-transcription PCR (Supplementary Fig. 1a). In contrast to *cmt2* mutants, *cmt3* mutants lost CHG methylation globally, but CHH methylation was lost only at limited sites in the genome<sup>8</sup>. Thus CMT2 and CMT3 appear to have different sequence preferences.

To understand the difference in sequence specificity between CMT2 and CMT3, we examined CMT2 methyltransferase activity *in vitro*. To test whether CMT2 could methylate DNA *in vitro*, we assayed CMT2 activity toward oligonucleotides of different methylation status. We used oligonucleotides that were unmethylated, methylated in all sequence contexts on only one strand (hemimethylated) or, as a negative control, methylated in all sequence contexts on both strands (fully methylated) (Online Methods)<sup>10</sup>. We found that CMT2 preferentially methylated unmethylated oligonucleotides compared to hemimethylated oligonucleotides *in vitro* (Fig. 1b). This was in contrast to CMT3, which preferentially methylated hemimethylated oligonucleotides<sup>10</sup>. We further assayed sequence specificity of methylation by CMT2 and found that it did not methylate CG sites (Supplementary Fig. 1c). Rather, CMT2 strongly methylated both CHG and CHH sites (Fig. 1c). This was in contrast to CMT3, which substantially preferred to methylate CHG sites compared to CHH sites<sup>10</sup> (Supplementary Fig. 1b). Hence the methyltransferase activity of CMT2 is distinct from that of CMT3, such that it preferentially methylates unmethylated DNA and effectively methylates both CHG and CHH sites *in vitro*. These findings are consistent with our *in vivo* studies (described below) showing that CMT2 not only mediates CHH methylation but also mediates CHG methylation.

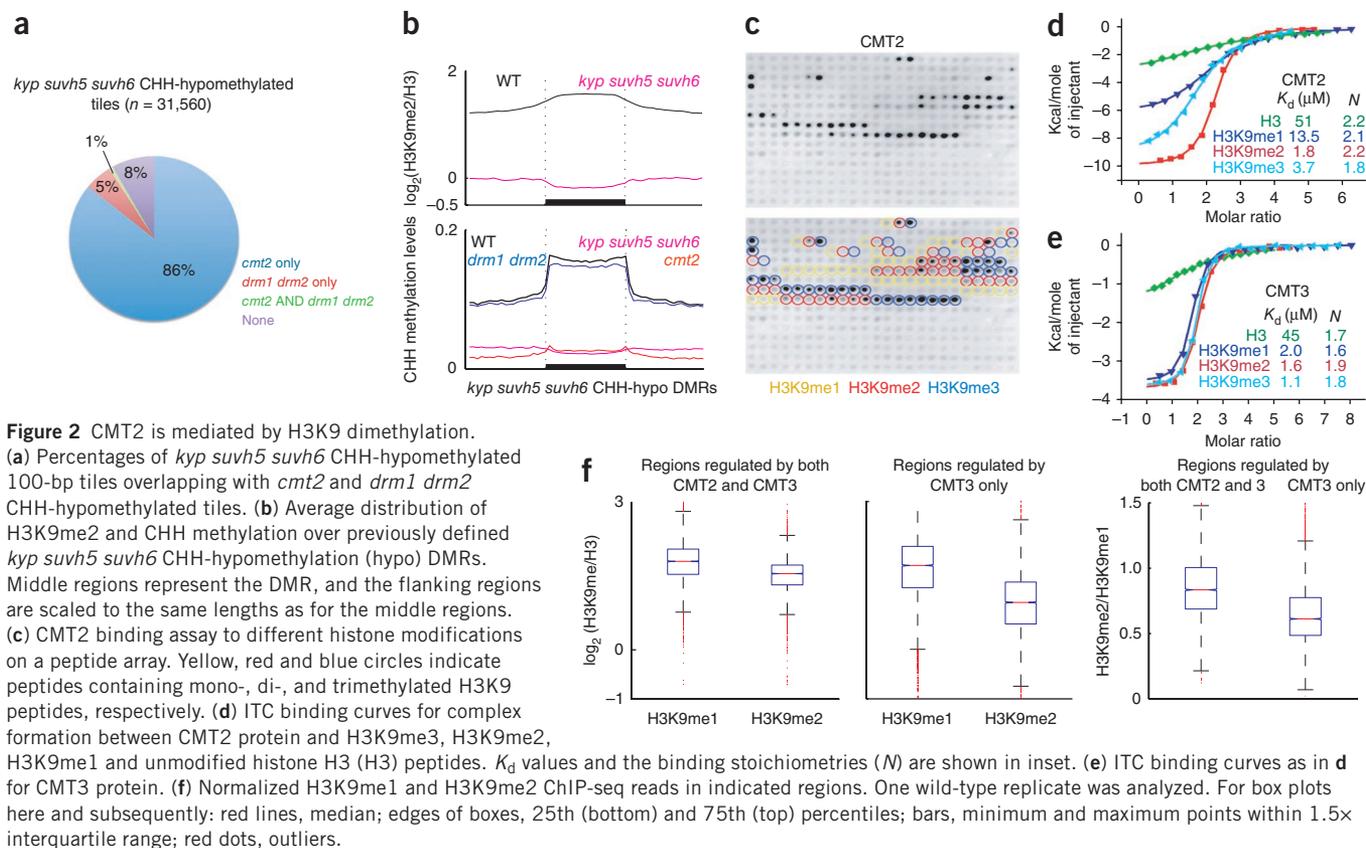
### CMT2 activity is mediated by H3K9 methylation

KRYPTONITE (KYP or SUVH4), SUVH5 and SUVH6 are the major H3K9 methyltransferases in *Arabidopsis*<sup>11,12</sup>. We previously showed that loss of CHG methylation in *kyp* (official symbol *suvh4*) *suvh5* *suvh6* triple mutants mimicked the loss of CHG methylation in *cmt3* mutants genome wide<sup>8</sup>. However, extensive loss of CHH methylation was also present in *kyp* *suvh5* *suvh6* but not in *cmt3*, thus suggesting that there must be another methyltransferase methylating CHH sites<sup>8</sup>. About 86% of *kyp* *suvh5* *suvh6* CHH-hypomethylated sites overlapped with *cmt2* CHH-hypomethylated sites, thus suggesting that H3K9 methylation regulates bulk CHH methylation through CMT2 (Fig. 2a,b). A smaller fraction of CHH sites regulated by KYP, SUVH5 and SUVH6 overlapped with DRM2-target sites (Fig. 2a); this probably is explained by the dependency of RNA polymerase (Pol) IV recruitment on H3K9 methylation through the histone-binding protein SHH1 (refs. 14,15). We performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) on dimethylated H3K9 (H3K9me2) in wild-type and *kyp* *suvh5* *suvh6* mutants and confirmed that loss of CHH methylation in *kyp* *suvh5* *suvh6* was associated with loss of H3K9me2 (Fig. 2b).



Structural and functional work has suggested that the BAH domain and chromodomain of CMT3 bind methylated H3K9 (ref. 10). Because CMT2 and CMT3 proteins have very similar domain configurations (Supplementary Fig. 2a), we hypothesized that CMT2 may also recognize methylated H3K9. To test this, we assayed binding of recombinant CMT2 protein to different histone modifications on a peptide array. Interestingly, we found preferential binding of CMT2 to H3K9 di- and trimethylated peptides (H3K9me2 and H3K9me3) but less binding to monomethylated H3K9 (H3K9me1) peptides (Fig. 2c and Supplementary Fig. 2b), results further confirmed by our isothermal titration calorimetry (ITC) binding data (Fig. 2d). These data were in contrast to CMT3, which bound H3K9me1, H3K9me2 and H3K9me3 equally well (Fig. 2e)<sup>10</sup>. In addition, all the ITC binding curves yielded *N* values around 2, thus indicating that two histone tail peptides bind each CMT molecule and that the dual recognition of methylated H3K9 tails is therefore likely to be a general feature of chromomethylase family of DNA methyltransferases.

The sensitivity of CMT2 to the number of methyl groups on H3K9 *in vitro* led us to investigate whether this property influences the sites to which CMT2 and CMT3 are targeted. To test this, we performed ChIP-seq on H3K9me1 and compared the results to those for H3K9me2. We did not analyze H3K9me3 because this mark is present at extremely low levels<sup>16</sup> and is associated with active genes<sup>17</sup>, which are devoid of non-CG methylation. We compared sites that are regulated by both CMT2 and CMT3 to sites regulated by CMT3 but not CMT2 (Online Methods). At sites regulated by both CMT2 and CMT3, there were higher levels of H3K9me2 compared to those at sites methylated by CMT3 but not CMT2 (Fig. 2f). Hence CMT2 is preferentially associated with H3K9me2, whereas CMT3 does not show such preference. This supports our finding that CMT2 binds H3K9me2 with a substantial preference over H3K9me1, whereas CMT3 can bind both H3K9me1 and H3K9me2 almost equally (Fig. 2c–e)<sup>10</sup>. Our results indicate that the number of methyl groups on H3K9 may influence CMT protein targeting to the genome.



### Interplay between non-CG methyltransferases in methylation

The finding that CMT2 had an important role in maintaining CHH methylation levels in the genome led us to generate mutants containing all possible combinations of non-CG methyltransferase mutants. We crossed *cmt2* to *cmt3* and to *drm1 drm2* double mutants. (DRM1 is expressed only in female gametes<sup>18</sup>.) We generated single-nucleotide-resolution maps of DNA methylation in the mutants by performing BS-seq. We first looked at non-CG methylation patterns over all TEs and chromosomes. We found that non-CG methylation in the genome was eliminated in *drm1 drm2 cmt2 cmt3* quadruple mutants (Fig. 3a,b and Supplementary Fig. 3a,b). This indicated that DRM1, DRM2, CMT2 and CMT3 are collectively responsible for all non-CG methylation in the *Arabidopsis* genome. This finding enabled us to determine the contributions of each non-CG methyltransferase in DNA methylation patterning. We observed that methylation of both CHG and CHH is redundantly regulated by all non-CG methyltransferases to a certain extent (Fig. 3a–d). This suggests that different pathways cooperate to regulate non-CG methylation patterning.

### CMT2 and CMT3 methylate CHG sites in a redundant manner

CMT3 tends to methylate large TEs and sites distal to genes<sup>8,9</sup>. In *cmt3* mutants, a strong but partial loss of CHG methylation occurs<sup>8,9</sup> (Fig. 3a–d). We found that in *cmt2 cmt3* double mutants there was stronger loss of CHG methylation than in *cmt3* mutants (Fig. 3c,e–g and Supplementary Fig. 3b). These sites were nonoverlapping with DRM2-regulated sites (Fig. 3c). This suggests that although CMT2 preferentially methylates CHH sites, it also methylates CHG sites. This result is consistent with our finding that CMT2 can also methylate CHG sites *in vitro* (Fig. 1c). Hence, although the main role of CMT2 is to methylate CHH sites, CMT2 and CMT3 function partially redundantly to methylate CHG sites.

### DRM2-target sites are methylated by both DRM2 and CMT3

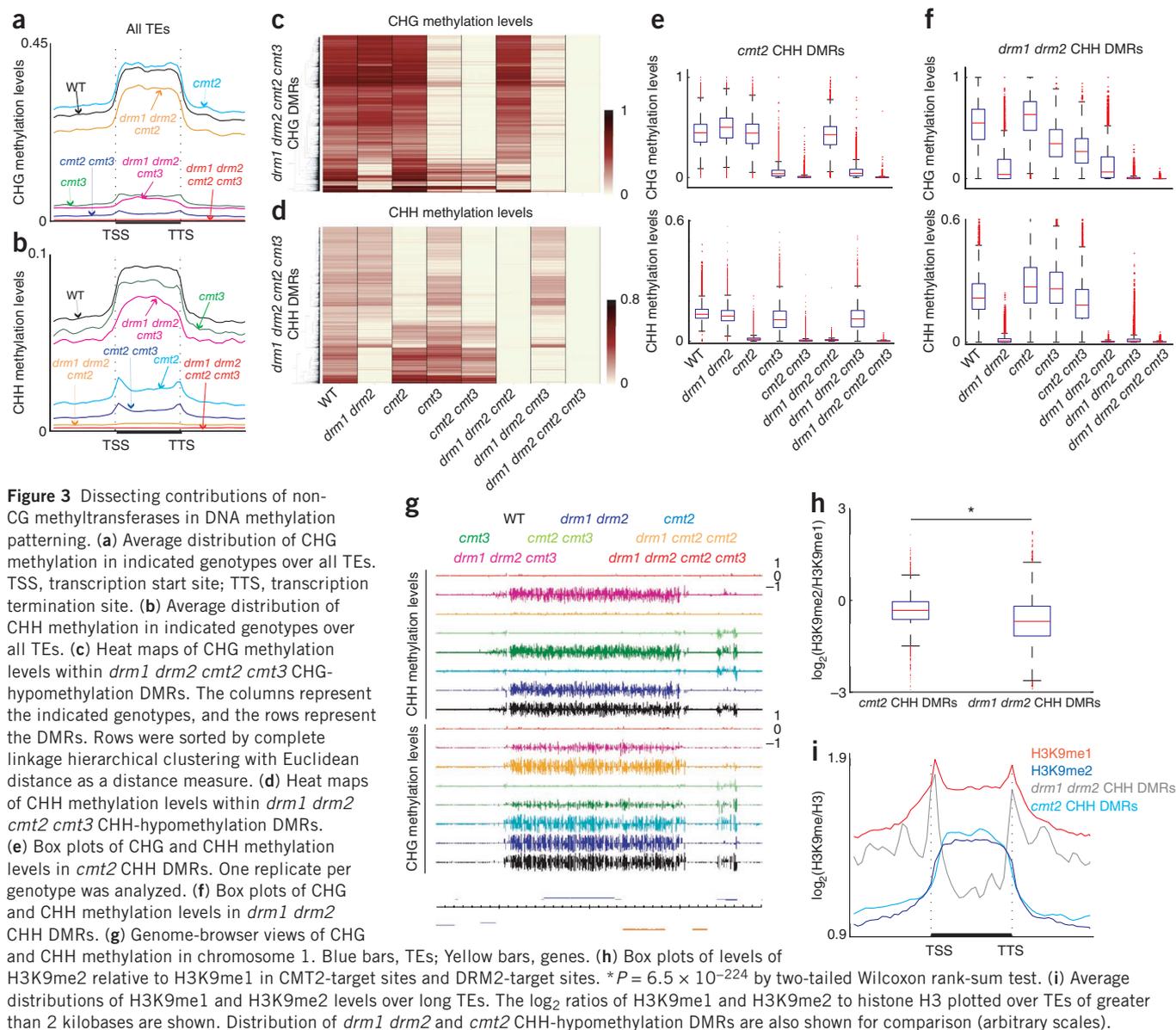
DRM2 tends to methylate the edges of large TEs as well as small TEs that are proximal to genes<sup>8,9</sup>. In *drm1 drm2* mutants, loss of DNA methylation occurs in CHH contexts and to a lesser extent in CHG contexts<sup>8</sup> (Fig. 3f). This suggests that a different methyltransferase is methylating CHG at DRM2-target sites. In *cmt3* mutants, CHG methylation was partially reduced at DRM2-target sites, and in *drm1 drm2 cmt3* triple mutants CHG methylation was nearly completely lost (Fig. 3f). Hence CMT3 also methylates DRM2 sites. There was almost complete loss of non-CG methylation at DRM2 sites in *drm1 drm2 cmt3* mutants in the presence of a functional CMT2 (Fig. 3f). This suggests that CMT2 has a very minor role at DRM2-target sites. Thus, generally at DRM2-target sites, CMT3 and DRM2 methylate cytosines in CHG contexts, and DRM2 methylates cytosines in CHH contexts.

### CMT2 and DRM2 mediate all CHH methylation in the genome

Mutations in CMT2 or DRM2 alone are not sufficient to eliminate CHH methylation in the genome (Fig. 3a–g). However, we found that *drm1 drm2 cmt2* triple mutants essentially eliminated all CHH methylation in the genome (Fig. 3b,d–g). In fact, 99% of *drm1 drm2 cmt2 cmt3* CHH-hypomethylated differentially methylated regions (DMRs) overlapped with *drm1 drm2 cmt2* CHH DMRs (Supplementary Fig. 3c). DRM2 and CMT2 methylate almost completely nonoverlapping sites in the genome (Supplementary Fig. 3d). Hence a large proportion of heterochromatin can be divided into regions that are CMT2 targeted and those that are DRM2 targeted.

### H3K9me1 and H3K9me2 levels at CMT2- and DRM2-target sites

Our finding of CMT2 binding preferentially to H3K9me2 led us to compare H3K9me1 and H3K9me2 levels at target sites of CMT2

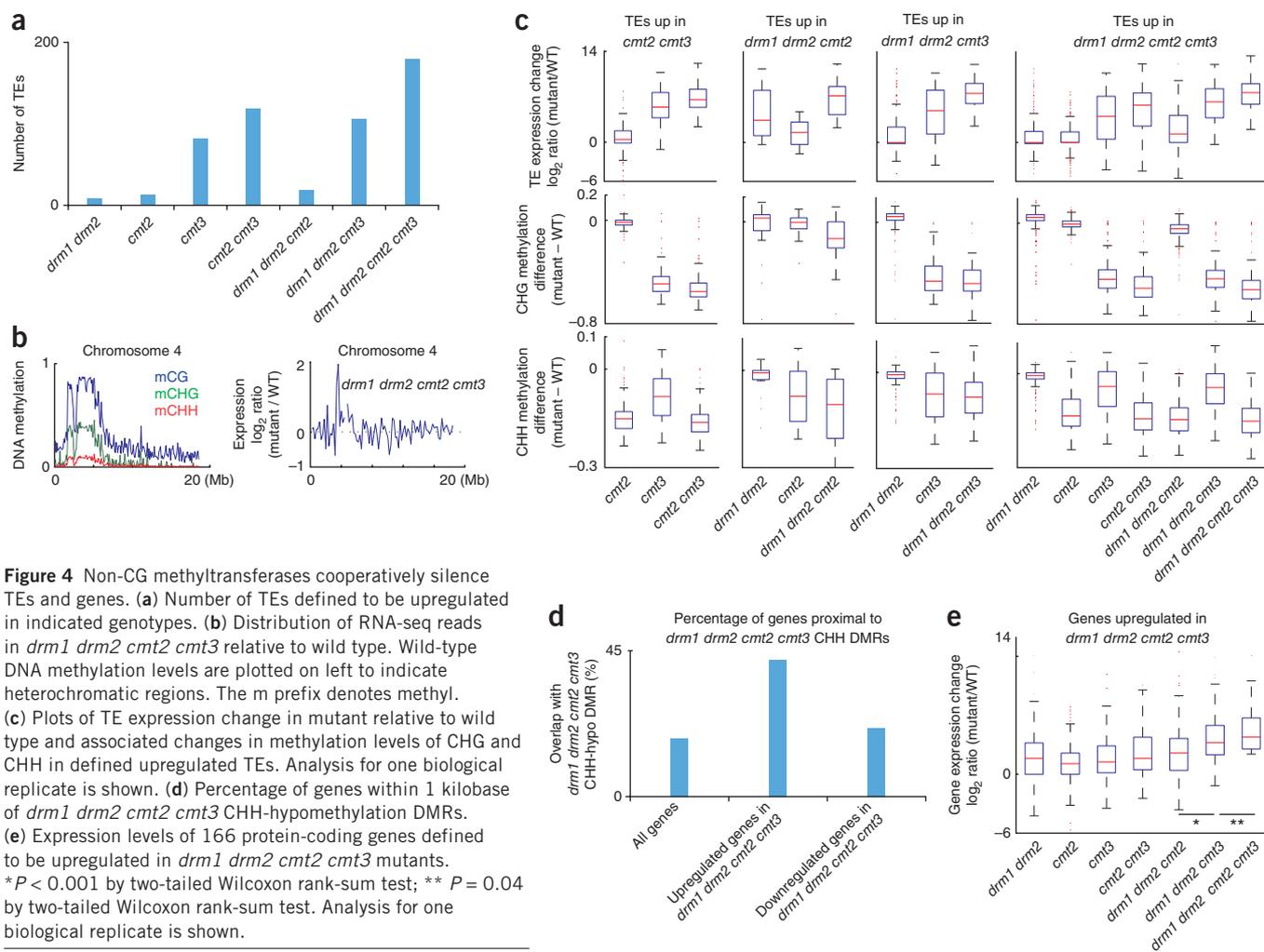


and DRM2. We found that the relative levels of H3K9me2 to H3K9me1 were higher at CMT2-target sites compared to DRM2-target sites (Fig. 3h). Furthermore, because DRM2 targets the edges of TEs<sup>8,9</sup>, we sought to examine the distributions of H3K9me1 and H3K9me2 over TEs. We found that H3K9me1 was especially enriched at boundaries of TEs, whereas H3K9me2 was enriched over the bodies of TEs (Fig. 3i). This distribution of H3K9me1 and H3K9me2 was consistent with the distribution of sites methylated by DRM2 and CMT2 (Fig. 3i). These results are consistent with the fact that SHH1, a factor involved in recruiting Pol IV to promote DRM2 targeting, exhibits similar *in vitro* binding to H3K9me1, H3K9me2 and H3K9me3 as that observed for CMT3 (Fig. 2e)<sup>10,14,15</sup>, whereas CMT2 preferably binds H3K9me2 (Fig. 2c,d). These results further suggest that the number of methyl groups on H3K9 may influence non-CG methyltransferase targeting.

#### CMT2, CMT3 and DRM2 cooperatively regulate TE expression

DNA methylation is implicated in transcriptional regulation. Because for the first time we possessed a mutant with largely normal levels of

CG methylation but a complete lack of non-CG methylation, we were able to test the extent to which non-CG methylation regulates expression of TEs and genes. We performed mRNA sequencing (mRNA-seq) on the different combinations of non-CG methylation mutants (Supplementary Fig. 4a). We defined TE derepression by using stringent cutoffs (Online Methods) and selected only TEs that showed pronounced misregulation in two biological replicates. TE derepression was most prominent in mutants containing *cmt3* mutations, thus suggesting that CMT3 has the strongest role in transcriptional silencing of TEs among non-CG methyltransferases (Fig. 4a). We found relatively minor upregulation of TEs in *cmt2* mutants despite CMT2 methylating a substantial proportion of the genome (Fig. 1a). This, together with the results that *drml1 drml2* or *drml1 drml2 cmt2* mutants showed modest TE derepression defects (Fig. 4a), suggests that CHH methylation itself may not have a major role in TE silencing. However, when combining *cmt3* mutations with *cmt2* or *drml1 drml2* mutations, we observed that an increased number of TEs were upregulated, results suggesting that CHH and CHG methylation redundantly silence TEs (Fig. 4a). Notably, upon loss of all non-CG methylation in *drml1 drml2*



**Figure 4** Non-CG methyltransferases cooperatively silence TEs and genes. **(a)** Number of TEs defined to be upregulated in indicated genotypes. **(b)** Distribution of RNA-seq reads in *drm1 drm2 cmt2 cmt3* relative to wild type. Wild-type DNA methylation levels are plotted on left to indicate heterochromatic regions. The m prefix denotes methyl. **(c)** Plots of TE expression change in mutant relative to wild type and associated changes in methylation levels of CHG and CHH in defined upregulated TEs. Analysis for one biological replicate is shown. **(d)** Percentage of genes within 1 kilobase of *drm1 drm2 cmt2 cmt3* CHH-hypomethylation DMRs. **(e)** Expression levels of 166 protein-coding genes defined to be upregulated in *drm1 drm2 cmt2 cmt3* mutants. \* $P < 0.001$  by two-tailed Wilcoxon rank-sum test; \*\* $P = 0.04$  by two-tailed Wilcoxon rank-sum test. Analysis for one biological replicate is shown.

*cmt2 cmt3* mutants, there was a large increase in the number of TEs upregulated (Fig. 4a and Supplementary Fig. 4a). In fact, there was a global increase in RNA-seq reads in heterochromatic regions in *drm1 drm2 cmt2 cmt3* relative to wild type (Fig. 4b). Although both DNA type and retrotransposons were regulated by non-CG methylation, there was over-representation of DNA Mariner, LINE-1 and long-terminal-repeat copia and gypsy transposons (Supplementary Fig. 4b and Supplementary Table 1). Hence different non-CG methyltransferases cooperate to silence TEs in the genome. We next measured the changes in non-CG methylation levels associated with changes in TE expression. The degree of TE upregulation correlated with the degree of loss of non-CG methylation in the mutants, thus indicating that these TEs are indeed regulated by non-CG methylation (Fig. 4c). Hence non-CG methylation plays important roles in silencing TEs.

#### CMT3 and DRM2, but not CMT2, regulate protein-coding genes

DNA methylation also regulates expression of protein-coding genes. By applying the same stringent cutoffs as we did for TEs, we defined 166 protein-coding genes significantly upregulated and 117 genes downregulated in *drm1 drm2 cmt2 cmt3* mutants. Genes that became upregulated in *drm1 drm2 cmt2 cmt3* mutants were substantially associated with high levels of non-CG methylation in wild type (Supplementary Fig. 4c) as well as non-CG DMRs in *drm1 drm2 cmt2 cmt3* mutants (Fig. 4d), thus indicating that these genes are regulated by non-CG methylation. In contrast, genes downregulated in

*drm1 drm2 cmt2 cmt3* mutants did not show association with non-CG methylation, results suggesting that downregulation of these genes is probably an indirect effect (Fig. 4d and Supplementary Fig. 4c). This result indicates that non-CG methylation primarily acts as a repressor of transcription. Gene ontology analysis of genes upregulated in *drm1 drm2 cmt2 cmt3* mutants indicated some association with response genes (Supplementary Fig. 4d); however, the list contained a variety of genes with different functions (Supplementary Table 2).

DRM2's targeting of sites proximal to genes suggests that it may function to regulate gene expression<sup>8,9</sup>. These sites are methylated by CMT3 and DRM2 but not CMT2 (Fig. 3f). Consistently with this, gene upregulation was most prominent in *drm1 drm2 cmt3* mutants compared to any other combinations of mutants (Fig. 4e). In fact, *drm1 drm2 cmt2 cmt3* mutants did not show substantial increases in gene expression levels compared to *drm1 drm2 cmt3* mutants (Fig. 4e). This is in contrast to our analysis of TEs (Fig. 4a). *SUPPRESSOR OF drm1 drm2 cmt3* (*SDC*) is a gene redundantly regulated by DRM2 and CMT3 and is responsible for the developmental phenotypes of *drm1 drm2 cmt3* mutants<sup>19</sup>. *SDC* was not more expressed in *drm1 drm2 cmt2 cmt3* compared to *drm1 drm2 cmt3* mutants (Supplementary Fig. 4e), results consistent with the morphological defects that the plants exhibited (Supplementary Fig. 4f). Hence, whereas TEs are cooperatively silenced by DRM2, CMT2 and CMT3, protein-coding genes are largely cooperatively regulated by CMT3 and DRM2 but not CMT2.

**Figure 5** Relationship between non-CG methylation and 24-nt siRNA accumulation. (a) 24-nt siRNA levels in DRM2-target sites. 24-nt siRNA levels are normalized by 21-nt siRNA levels for each genotype. One replicate per genotype was analyzed. (b) 24-nt siRNA levels in CMT2-target sites.

### 24-nt siRNAs and non-CG methylation at DRM2-target sites

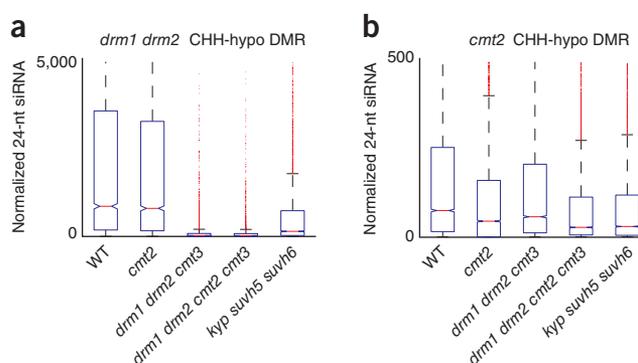
DRM2 is guided by 24-nt siRNAs to target loci<sup>1</sup>. The biogenesis of 24-nt siRNA depends on Pol IV. However, at certain loci siRNA accumulation has also been shown to depend on downstream RdDM factors such as Pol V and DRM2 (refs. 14,20–22). We sought to examine the extent to which siRNA accumulation depends on non-CG methylation by performing small-RNA sequencing. We found that in *drm1 drm2 cmt3* mutants there was strong loss of 24-nt siRNAs (Fig. 5a). This suggests that loss of non-CG methylation at these sites causes loss of 24-nt siRNAs. Loss of 24-nt siRNAs at these sites was not present in *cmt2* mutants, nor was the degree of loss substantially enhanced in *drm1 drm2 cmt2 cmt3* mutants compared to *drm1 drm2 cmt3* mutants (Fig. 5a), results consistent with the finding that CMT2 generally does not act at DRM2-target sites. Our results uncover an almost complete dependency of 24-nt siRNA accumulation on non-CG methylation at DRM2-target sites, suggesting a strong self-reinforcing loop mechanism.

### 24-nt siRNAs and non-CG methylation at CMT2-target sites

Upstream RdDM factors such as Pol IV are responsible for most 24-nt siRNA produced in the genome<sup>23–25</sup>. By analyzing ChIP-seq data on Pol IV<sup>14</sup>, we confirmed that Pol IV protein was physically enriched at CMT2-target sites (Supplementary Fig. 5a). Known upstream RdDM mutants such as *rdm4*, *nprp1a* and *rdr2*, which strongly reduce 24-nt siRNA across the genome<sup>23–26</sup>, did not substantially reduce CHH methylation at CMT2-dependent sites (Supplementary Fig. 5b). In contrast, we observed that both *drm1 drm2 cmt3* triple mutants and *cmt2* single mutants had partial but consistent loss of 24-nt siRNA accumulation at CMT2-target sites (Fig. 5b). There was substantially more loss of 24-nt siRNAs upon loss of all non-CG methylation in *drm1 drm2 cmt2 cmt3* quadruple mutants (Fig. 5b). This suggests that non-CG methylation partially regulates 24-nt siRNAs at these sites. Although these 24-nt siRNAs do not control non-CG methylation in *cis*, one possibility is that they target other elements in *trans*<sup>27</sup>, such as newly inserted TEs<sup>28</sup>. Our results suggest that there is an almost complete dependency of 24-nt siRNA on non-CG methylation at DRM2-target sites and partial dependency of 24-nt siRNA on non-CG methylation at CMT2-target sites. As explored below, a possible mechanism for this dependency may be through H3K9 methylation.

### Non-CG methylation globally controls H3K9 methylation

Most non-CG methylation in the genome is regulated by H3K9 methylation (Fig. 2)<sup>8,10,14,15</sup>. H3K9 methylation has also been suggested to be partially dependent on DNA methylation at certain loci, thus suggesting a self-reinforcing loop between DNA methylation and H3K9 methylation<sup>29–31</sup>. This self-reinforcing loop is probably mediated at least in part by the SRA domains of the H3K9 methyltransferases KYP, SUVH5 and SUVH6, which preferentially bind methylated DNA<sup>29</sup>. However, the extent of this dependency remains poorly understood. We performed ChIP-seq on H3K9me2 in wild type and *drm1 drm2 cmt2 cmt3* mutants and the *kyp suvh5 suvh6* triple H3K9 methyltransferase mutant. Strikingly, by analyzing the distribution of H3K9me2 across chromosomes, we found strong loss of H3K9me2 in *drm1 drm2 cmt2 cmt3* mutants (Fig. 6a). Inspection of the data on the genome browser confirmed loss of H3K9me2 in



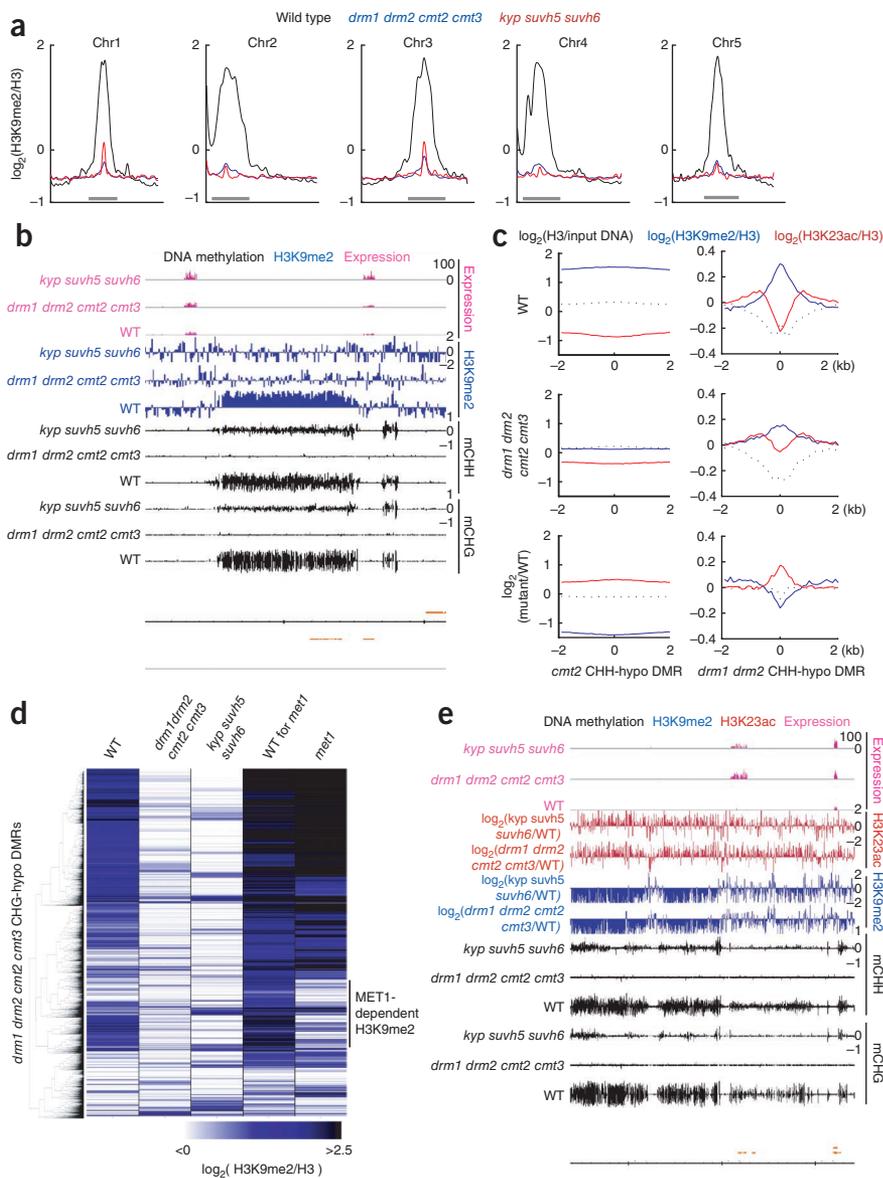
*drm1 drm2 cmt2 cmt3* mutants (Fig. 6b and Supplementary Fig. 6a). In fact, the degree of loss of H3K9me2 in *drm1 drm2 cmt2 cmt3* mutants was as strong as in *kyp suvh5 suvh6* mutants (Fig. 6a,b and Supplementary Fig. 6a). Loss of H3K9me2 in *drm1 drm2 cmt2 cmt3* mutants occurred at both CMT2-targeted sites and DRM2-targeted sites, although the loss appeared stronger at CMT2-dependent sites (Fig. 6c). Strong loss of 24-nt siRNA in *drm1 drm2 cmt2 cmt3* mutants at DRM2-target sites (Fig. 5a) is probably partially explained by loss of H3K9 methylation, because 24-nt siRNA accumulation is dependent on the methyl-H3K9-binding protein SHH1 (refs. 14,15). Our results indicate that non-CG methylation mediates genome-wide H3K9 methylation patterning.

### 24-nt siRNA accumulation is mediated by H3K9 methylation

Our finding of extensive self-reinforcing loops between H3K9 methylation and non-CG methylation in part provides an explanation for the self-reinforcing loop between 24-nt siRNA accumulation and non-CG methylation. At DRM2-target sites, non-CG methylation is required for H3K9 methylation (Fig. 6c), which then regulates 24-nt siRNAs through SHH1 binding to methylated H3K9. Consistently with this model, in *kyp suvh5 suvh6* mutants there was a strong loss of 24-nt siRNAs at DRM2-target sites (Fig. 5a). At CMT2-target sites, non-CG methylation is strongly required for H3K9 methylation (Fig. 6c). Consistently with H3K9me2 being lost to a similar extent in *drm1 drm2 cmt2 cmt3* and *kyp suvh5 suvh6* mutants (Fig. 6a), we found similar degrees of loss of 24-nt siRNA in *drm1 drm2 cmt2 cmt3* and *kyp suvh5 suvh6* mutants compared to wild type (Fig. 5b). Hence it is likely that non-CG methylation controls H3K9 methylation, which then regulates the biogenesis of 24-nt siRNA.

### CG methylation and heterochromatic H3K9 methylation

Genome-wide elimination of CG methylation by mutation of the CG methyltransferase MET1 resulted in loss of H3K9me2 at certain sites<sup>32,33</sup>, although the mechanism is not understood. We analyzed H3K9me2 ChIP data in wild type and *met1* mutants<sup>34</sup>. As expected, we observed loss of H3K9me2 at certain sites in *met1* mutants (Fig. 6d). However, we found that these were sites that also lost non-CG methylation in *met1* mutants (Supplementary Fig. 6b). In contrast, we did not observe genome-wide loss of H3K9me2 in *met1* mutants, as we found in *drm1 drm2 cmt2 cmt3* mutants (Fig. 6d and Supplementary Fig. 6c). This suggests that H3K9 methylation is much more dependent on non-CG methylation than on CG methylation. Although we cannot rule out the possibility that loss of H3K9me2 at certain sites in *met1* mutants is directly due to loss of CG methylation, it seems likely that loss of H3K9me2 in *met1* mutants is due to loss of non-CG methylation at these sites. Our results suggest that non-CG methylation has a dominant role in regulating H3K9 methylation patterning throughout the genome.



**Figure 6** Relationship between non-CG methylation and H3K9 methylation. **(a)** Distribution of H3K9me2 relative to histone H3 over chromosomes. The graphs are shifted to align on the euchromatic arms. Gray bars indicate pericentromeric heterochromatin. **(b)** Genome-browser views of DNA methylation, expression levels and H3K9me2 in wild type and *drm1 drm2 cmt2 cmt3* and *kyp suvh5 suvh6* mutants in chromosome 1. Yellow bars, genes. **(c)** Average distribution of H3K9me2 and H3K23ac relative to histone H3 over *cmt2* and *drm1 drm2* CHH-hypomethylation DMRs. **(d)** Heat maps of H3K9me2 levels within *drm1 drm2 cmt2 cmt3* mutant CHG-hypomethylation DMRs. H3K9me2 is normalized to histone H3. Two wild-type H3K9me2 data sets are shown because H3K9me2 data for *met1* have a separate wild-type control<sup>34</sup>. **(e)** Genome-browser views of DNA methylation, expression levels, H3K23ac and H3K9me2 in wild type and *drm1 drm2 cmt2 cmt3* and *kyp suvh5 suvh6* mutants in chromosome 1. Yellow bars, genes.

## DISCUSSION

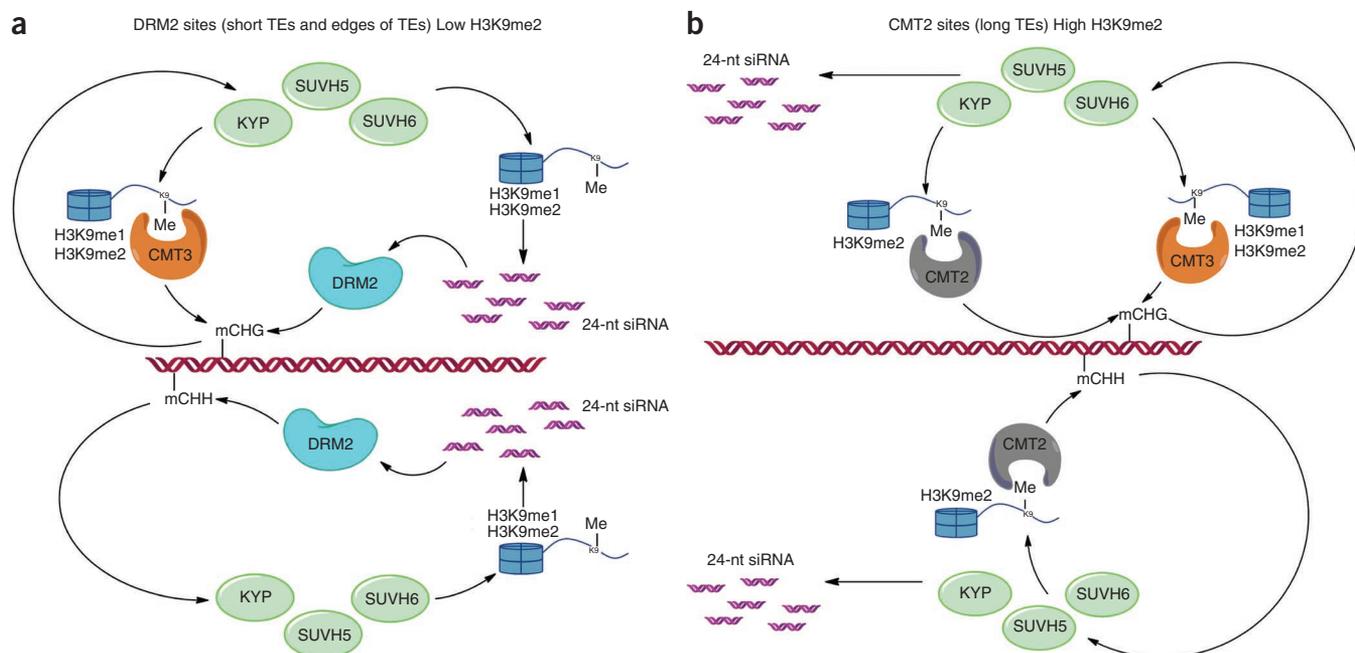
In this study, we characterized a series of mutants affecting non-CG methylation, including the poorly understood methyltransferase CMT2. This analysis has uncovered the roles of each non-CG methyltransferase in DNA methylation patterning and gene silencing. Furthermore, our finding of extensive cross-talk between non-CG methylation and H3K9 methylation provides insights into the mechanisms of cross-talk between different silencing pathways. All data generated in this study can be visualized in a modified UCSC browser (<http://genomes.mcdb.ucla.edu/AthBSseq/>) along with other epigenomic data sets.

At DRM2-target sites, there is a self-reinforcing loop between non-CG methylation, H3K9 methylation and 24-nt siRNAs (Fig. 7a). H3K9 methylation is required for CMT3 targeting to methylate CHG sites at a subset of DRM2 sites as well as for DRM2 targeting through binding of SHH1 (ref. 14), which methylates the remaining non-CG sites. SHH1 binding to methylated H3K9 is required for 24-nt siRNA accumulation at a subset of DRM2 sites<sup>14</sup>. The 24-nt siRNAs then direct DRM2 (ref. 35). Our data suggest that non-CG methylation is required for H3K9 methylation, which is largely mediated by KYP, SUVH5 and SUVH6. The H3K9 methylation then directs CMT3 and DRM2 pathways for non-CG methylation.

At CMT2-target sites, there is also a self-reinforcing loop between non-CG methylation, H3K9 methylation and 24-nt siRNAs (Fig. 7b). Our results suggest that both CMT2 and CMT3 mediate CHG methylation, and CMT2 mediates CHH methylation at these sites through binding to methylated H3K9. Non-CG methylation mediated by CMT2 and CMT3 regulates H3K9 methylation mediated by KYP, SUVH5 and SUVH6. H3K9 methylation may then partially regulate 24-nt siRNAs produced at these sites through a similar mechanism of recruitment of H3K9 methylation readers that occurs at DRM2-target sites: Because these 24-nt siRNAs are also dependent on Pol IV<sup>25,26</sup>, there may be H3K9

## Loss of non-CG methylation induces histone hyperacetylation

Histone acetylation is associated with open chromatin and actively transcribed genes. Given the strong loss of the repressive histone mark H3K9me2 in *drm1 drm2 cmt2 cmt3*, we sought to examine the effects on genome-wide histone acetylation patterns. We performed ChIP-seq on acetylated histone H3 K23 (H3K23ac) and histone H3 on wild type and *drm1 drm2 cmt2 cmt3* and *kyp suvh5 suvh6* mutants. As expected, H3K23ac was enriched in promoter regions of active genes in wild type (Supplementary Fig. 6d). We observed genome-wide increases of histone acetylation in *drm1 drm2 cmt2 cmt3* mutants and *kyp suvh5 suvh6* mutants at sites that lost DNA methylation (Fig. 6c,e and Supplementary Fig. 6e). Elevation of histone acetylation levels was not restricted to transcriptionally upregulated TEs and genes (Fig. 6e and Supplementary Fig. 6f), thus suggesting that this phenomenon cannot simply be explained by more transcription in the mutants. Consistently with the elevation in histone acetylation, we found substantial chromocenter decondensation in *drm1 drm2 cmt2 cmt3* mutants (Supplementary Fig. 6g). Hence non-CG methylation is required to keep heterochromatin in a deacetylated and compacted state.



**Figure 7** Non-CG methylation pathways. (a) Non-CG methylation pathways at DRM2-target sites. (b) Non-CG methylation pathways at CMT2-target sites. Descriptions are in main text.

methylation readers other than SHH1 that recruit Pol IV to CMT2 sites. Although these 24-nt siRNAs do not appear to have a major role in guiding DRM2 in *cis*, they might function to silence TEs in *trans*<sup>27,28</sup>.

In summary, our data demonstrate that the CMT2, CMT3 and DRM2 methyltransferases collaborate to control non-CG methylation and participate in self-reinforcing loop mechanisms with H3K9 methylation and small RNAs to control gene silencing throughout the genome.

## METHODS

Methods and any associated references are available in the [online version of the paper](#).

**Accession codes.** All sequencing data have been deposited in GEO with accession [GSE51304](#).

*Note: Any Supplementary Information and Source Data files are available in the [online version of the paper](#).*

## ACKNOWLEDGMENTS

We thank M. Akhavan for Illumina sequencing. Sequencing was performed at the University of California, Los Angeles (UCLA) Broad Stem Cell Research Center BioSequencing Core Facility. We thank W. Yang and M. Pellegrini for help with the UCSC browser. We thank L. Tao and D. Chen for technical help with experiments. We thank F. Berger for helpful comments. This work was supported by the Abby Rockefeller Mauzé Trust, the Maloris and Starr foundations (D.J.P.) and US National Institutes of Health grant GM60398 and National Science Foundation grant 0701745 (S.E.J.). H.S. was supported by a UCLA Dissertation Year Fellowship. X.Z. is supported by a Ruth L. Kirschstein National Research Service Award (F32GM096483-01). S.F. is supported as a Special Fellow of the Leukemia & Lymphoma Society. S.E.J. is supported as an investigator of the Howard Hughes Medical Institute.

## AUTHOR CONTRIBUTIONS

H.S., T.D., J.D., X.Z., S.F. and L.J. performed the experiments. S.E.J. and D.J.P. oversaw the study. H.S. designed the study, analyzed data and wrote the manuscript.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Law, J.A. & Jacobsen, S.E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **11**, 204–220 (2010).
- Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
- Xie, W. *et al.* Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell* **148**, 816–831 (2012).
- Laurent, L. *et al.* Dynamic changes in the human methylome during differentiation. *Genome Res.* **20**, 320–331 (2010).
- Meissner, A. *et al.* Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877 (2005).
- Ramsahoye, B.H. *et al.* Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. USA* **97**, 5237–5242 (2000).
- Varley, K.E. *et al.* Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res.* **23**, 555–567 (2013).
- Stroud, H., Greenberg, M.V., Feng, S., Bernatavichute, Y.V. & Jacobsen, S.E. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell* **152**, 352–364 (2013).
- Zemach, A. *et al.* The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**, 193–205 (2013).
- Du, J. *et al.* Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell* **151**, 167–180 (2012).
- Ebbs, M.L. & Bender, J. Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. *Plant Cell* **18**, 1166–1176 (2006).
- Jackson, J.P., Lindroth, A.M., Cao, X. & Jacobsen, S.E. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**, 556–560 (2002).
- Henikoff, S. & Comai, L. A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in *Arabidopsis*. *Genetics* **149**, 307–318 (1998).
- Law, J.A. *et al.* Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **498**, 385–389 (2013).
- Zhang, H. *et al.* DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. *Proc. Natl. Acad. Sci. USA* **110**, 8290–8295 (2013).
- Johnson, L. *et al.* Mass spectrometry analysis of *Arabidopsis* histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Res.* **32**, 6511–6518 (2004).
- Roudier, F. *et al.* Integrative epigenomic mapping defines four main chromatin states in *Arabidopsis*. *EMBO J.* **30**, 1928–1938 (2011).

18. Jullien, P.E., Susaki, D., Yelagandula, R., Higashiyama, T. & Berger, F. DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. *Curr. Biol.* **22**, 1825–1830 (2012).
19. Henderson, I.R. & Jacobsen, S.E. Tandem repeats upstream of the *Arabidopsis* endogene SDC recruit non-CG DNA methylation and initiate siRNA spreading. *Genes Dev.* **22**, 1597–1606 (2008).
20. Herr, A.J., Jensen, M.B., Dalmay, T. & Baulcombe, D.C. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**, 118–120 (2005).
21. Pontier, D. *et al.* Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev.* **19**, 2030–2040 (2005).
22. Zilberman, D., Cao, X. & Jacobsen, S.E. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299**, 716–719 (2003).
23. Kasschau, K.D. *et al.* Genome-wide profiling and analysis of *Arabidopsis* siRNAs. *PLoS Biol.* **5**, e57 (2007).
24. Mosher, R.A., Schwach, F., Studholme, D. & Baulcombe, D.C. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc. Natl. Acad. Sci. USA* **105**, 3145–3150 (2008).
25. Zhang, X., Henderson, I.R., Lu, C., Green, P.J. & Jacobsen, S.E. Role of RNA polymerase IV in plant small RNA metabolism. *Proc. Natl. Acad. Sci. USA* **104**, 4536–4541 (2007).
26. Wierzbicki, A.T. *et al.* Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the *Arabidopsis* epigenome. *Genes Dev.* **26**, 1825–1836 (2012).
27. McCue, A.D., Nuthikattu, S. & Slotkin, R.K. Genome-wide identification of genes regulated in *trans* by transposable element small interfering RNAs. *RNA Biol.* **10**, 1379–1395 (2013).
28. Zhong, X. *et al.* DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. *Nat. Struct. Mol. Biol.* **19**, 870–875 (2012).
29. Johnson, L.M. *et al.* The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr. Biol.* **17**, 379–384 (2007).
30. Inagaki, S. *et al.* Autocatalytic differentiation of epigenetic modifications within the *Arabidopsis* genome. *EMBO J.* **29**, 3496–3506 (2010).
31. Mathieu, O., Probst, A.V. & Paszkowski, J. Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in *Arabidopsis*. *EMBO J.* **24**, 2783–2791 (2005).
32. Soppe, W.J. *et al.* DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis*. *EMBO J.* **21**, 6549–6559 (2002).
33. Johnson, L., Cao, X. & Jacobsen, S. Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. *Curr. Biol.* **12**, 1360–1367 (2002).
34. Deleris, A. *et al.* 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 (2012).
35. Wierzbicki, A.T., Ream, T.S., Haag, J.R. & Pikaard, C.S. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat. Genet.* **41**, 630–634 (2009).

## ONLINE METHODS

**Plant material.** All mutant lines used in this study were in the Columbia ecotype background. *drm1 drm2 cmt3* and *kyp suvh5 suvh6* mutants were previously described<sup>11,36</sup>. The *cmt2* transfer-DNA allele used in this study was *cmt2-7* (WIGSDSLOX7E02) and *cmt2-3* (SALK\_012874). *cmt2-7* was used for subsequent crosses. Plants were grown under continuous light, and 3-week-old leaves were used for all experiments, except for small-RNA sequencing (described below).

**RT-PCR.** Total RNA was extracted from leaves with TRIzol and treated with DNase I (Roche). cDNA was synthesized with oligo-dTs with Superscript II (Invitrogen). PCR was performed on *CMT2* (JP10697, GAGAAATCCTAAAACGTCCG and JP10698, CAGCCATTTCGTACACGAC) and ACTIN (JP2452, TCGTGGTGGTGTGTTTGTAC and JP2453, CAGCATCATACAAGCATCC).

**Recombinant-protein expression and purification.** The N-terminal fragment of *Arabidopsis* CMT2 (residues 1–503) did not show homology to any known domain, nor did it BLAST to any other plant species, and therefore was not included. The N-terminally truncated CMT2 (residues 504–1,295), including all the functional domains (BAH, chromodomain, and DNA-methyltransferase domains), was cloned into a self-modified vector that fuses an N-terminal hexahistidine plus yeast SUMO tag to the target protein. The recombinant plasmid was transformed into *Escherichia coli* strain BL21(DE3) RIL (Stratagene). The cells were cultured in LB medium at 37 °C until OD<sub>600</sub> reached 0.6. The medium was subsequently cooled to 20 °C, and 0.25 mM IPTG was added to induce the protein expression overnight. The recombinant expressed protein was purified with a HisTrap FF column (GE Healthcare) and then by a Q FF column (GE Healthcare) and a Hiload Superdex G200 16/60 column (GE Healthcare). The purified protein was concentrated to 15 mg/ml and was stocked in –80 °C. The N-terminally truncated *Arabidopsis* CMT3 (residues 46–839), including all the functional domains (BAH, chromodomain, and DNA-methyltransferase domain), was cloned, expressed, and purified with the same protocol as CMT2.

**Isothermal titration calorimetry.** Isothermal titration calorimetry (ITC)-based binding experiments were conducted with a MicroCalorimeter iTC 200 instrument at 4 °C. Purified protein samples were dialyzed overnight against a buffer of 100 mM NaCl, 2 mM β-mercaptoethanol, and 20 mM HEPES, pH 7.5, at 4 °C. Then the protein samples were diluted, and the lyophilized peptides were dissolved with the same buffer. The titration was conducted according to standard protocol, and the data were fitted with Origin 7.0.

**DNA-methyltransferase activity assay.** DNA-methyltransferase assay was performed as previously described<sup>10</sup> except that 2 μg of recombinant CMT2 protein was used. Oligos used for the assays are shown in **Supplementary Table 3**.

**Histone peptide array.** 30 μg of recombinant CMT2 protein was screened on a MODified histone array slide according to the manufacturer's instructions (Active Motif) with antibody to histidine and was developed with Enhanced chemiluminescence (GE Healthcare). All analyses were performed with the manufacturer's software (Active Motif).

**Whole-genome bisulfite sequencing (BS-seq).** 500 ng of genomic DNA was used to generate BS-seq libraries, as previously described<sup>8,37</sup>. One library per genotype was generated. 50-mer sequencing reads were analyzed. Identical reads were collapsed into single reads, and reads were mapped to the TAIR10 genome with BS-seeker by allowing up to two mismatches. Fractional DNA methylation levels were computed by  $\#C/(\#C + \#T)$ . DMRs were defined exactly as previously described<sup>8</sup>. Control datasets used for comparisons (from *cmt2-7*, *cmt3*, *drm1 drm2*, *drm1 drm2 cmt3*, *kyp suvh5 suvh6* mutants) were obtained from ref. 8.

**mRNA sequencing.** RNA was extracted from 0.1 g tissue with TRIzol (Invitrogen). We performed mRNA-seq experiments on two biological replicates for each genotype tested. Libraries were generated and sequenced according to the manufacturer's instructions (Illumina). Data were analyzed as previously described<sup>38</sup>. Reads were mapped to the TAIR10 genome with Bowtie<sup>39</sup> by allowing up to two mismatches and keeping only reads that uniquely mapped to the genome. Genes and TEs were defined as deregulated in a mutant with a four-fold cutoff and a corrected  $P < 0.01$ . Only genes and TEs that showed consistent deregulation in two independent experiments were defined as significantly deregulated. To avoid divisions by zero, elements with zero reads were assigned the lowest nonzero gene or TE expression values within each library.

**smRNA sequencing.** Total RNA was extracted from 0.2 g of flowers with TRIzol (Invitrogen). siRNAs were purified as previously described<sup>40</sup>, with the following modifications. To precipitate high-molecular-weight RNAs, 25% PEG was added to a final concentration of 12.5% instead of 5% PEG. For small-RNA purification from LMW RNA, SYBR Gold was used to stain the gel. The gel was crushed with Gel Breaker Tubes (IST Engineering Inc.), and the debris was filtered with 5-μm filter tubes (IST Engineering Inc.). The final elution of the RNA was done in 5 μL of nuclease-free H<sub>2</sub>O for subsequent generation of libraries for high-throughput sequencing. Libraries were generated and sequenced according to the manufacturer's instructions (Illumina TruSeq Small RNA Sample Preparation Kits). One library per genotype was generated. Adaptor sequences were clipped off before mapping. Reads were mapped to the TAIR10 genome with Bowtie<sup>39</sup> by allowing no mismatches and keeping only reads that uniquely mapped to the genome. For the analyses, the smRNA counts were normalized to the size of each smRNA library by dividing the number of reads by the number of total uniquely mapping reads of 21 bp in size.

**Chromatin-immunoprecipitation (ChIP) sequencing.** 1 g of tissue was ground in liquid nitrogen, and ChIP was performed as previously described<sup>14</sup>, with 1:100 dilutions of antibodies as follows: H3K9me2 (Abcam 1220), H3 (Abcam 1791), H3K9me1 (Upstate 07-450), and H3K23ac (Millipore 07-355). Validations of all antibodies are provided on the manufacturers' websites. Libraries were generated and sequenced according to the manufacturer's instructions (Illumina). Reads were mapped to the TAIR10 genome with Bowtie<sup>39</sup> by allowing up to two mismatches and keeping only reads that uniquely mapped to the genome. Reads mapping to identical locations were collapsed into one read. Two independent ChIP-seq experiments on biological replicates were performed on H3K9me2 and H3K23ac on wild type and *drm1 drm2 cmt2 cmt3* and *kyp suvh5 suvh6* mutants, and all experiments led to similar conclusions.

**Chromocenter compaction assay.** Chromocenter compaction assays were performed as previously described<sup>41</sup> with the following modifications. After post-fix, the slides were washed three times in PBS for 5 min each. The nuclei were then stained and mounted in Vectashield mounting media with DAPI (Vector H-1200). At least 100 nuclei were analyzed for each genotype.

36. Chan, S.W. *et al.* RNAi, DRD1, and histone methylation actively target developmentally important non-CG DNA methylation in *Arabidopsis*. *PLoS Genet.* **2**, e83 (2006).
37. Feng, S., Rubbi, L., Jacobsen, S.E. & Pellegrini, M. Determining DNA methylation profiles using sequencing. *Methods Mol. Biol.* **733**, 223–238 (2011).
38. Stroud, H. *et al.* DNA methyltransferases are required to induce heterochromatic re-replication in *Arabidopsis*. *PLoS Genet.* **8**, e1002808 (2012).
39. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
40. Lu, C., Meyers, B.C. & Green, P.J. Construction of small RNA cDNA libraries for deep sequencing. *Methods* **43**, 110–117 (2007).
41. Moissiard, G. *et al.* MORC family ATPases required for heterochromatin condensation and gene silencing. *Science* **336**, 1448–1451 (2012).