Novel crosstalk between DNA methylation and H3K27me3 revealed at Arabidopsis transposable elements

Valentin Hure[1], Florence Piron-Prunier[1] and Angélique Déléris[1]

[1] Université Paris-Saclay, Commissariat à l'Energie Atomique (CEA), Centre National de la Recherche Scientifique (CNRS), Institute for Integrative Biology of the Cell (I2BC), 91190, Gif-sur-Yvette, France

ABSTRACT

Transposable elements are primarily silenced by DNA methylation and the associated histone modification H3K9me2 in many multicellular eukaryotes, including plants. However, in the absence of DNA methylation due to mutations in the DNA methylation machinery or in certain developmental contexts, the same TEs can gain Polycomb-associated H3K27me3, another epigenetic silencing mark that is usually linked with the facultative silencing of genes. In this study, we initially aimed to investigate whether DNA methylation and H3K27me3 could compete during the establishment of silencing at TEs in the model plant Arabidopsis. Strikingly, we show that the deposition of the H3K27me3 mark at newly inserted transgenic TE sequences is impaired in plants in which the *de novo* DNA methyltransferase DRM2 is mutated, contrary to the competition model hypothesized. Further profiling of H3K27me3 in *drm2* mutants and in the DNA demethylase mutant *rdd* confirmed this new role of DNA methylation in promoting H3K27me3 deposition at some TEs, in addition to the previously described antagonistic role at others. These findings further reveal a new function of DNA demethylation in modulating H3K27me3 levels in vegetative tissues, which we confirmed via targeted DNA methylation

experiments. Together, our results uncover a novel crosstalk between DNA methylation and Polycomb at TEs and reveal that these two pathways, thought to be specialized and antagonistic, can be interdependent and cooperate more than anticipated to maintain genome and epigenome integrity in eukaryotes.

INTRODUCTION

Transposable elements are repetitive DNA sequences that have the potential to multiply and move around the genome, thus compromising genome integrity[1]. Nevertheless, this threat is largely mitigated by genome defense mechanisms that efficiently target transposable elements and stably repress their expression across cell division and generations through epigenetic marks such as DNA methylation (5mC). These marks can persist even on TEs that have aged and lost mobility; as such, they can play a role as epigenetic regulatory modules that can impact nearby gene regulation[1]. In plants and *Arabidopsis thaliana*, DNA methylation is established and maintained by three classes of DNA methyltransferases: MET1, which maintains CG methylation; CMT3 and CMT2, which maintain CHG and CHH methylation, respectively, and ensure feedback loops with H3K9me2 histone methyltransferases; DRM2, which participates in the maintenance of CHH (at TE loci that are smaller than those targeted by CMT2); and possibly CG methylation[2,3]. Importantly, DRM2 is able to methylate DNA *de novo* in all sequence contexts as part of the RNA-directed DNA methylation (RdDM) pathway[3]. Thus, *drm2* mutation drastically impairs de novo methylation of newly inserted repeats or TE-derived sequences[4,5], a molecular phenotype also observed to various extents in mutants for RNA interference or RdDM factors connected to DRM2 activity[5,6].

The detection of a TE that precedes its *de novo* methylation can be achieved in two manners[7]. The first, referred to as "expression-dependent silencing", is based on TE transcript detection by the RNA interference machinery, which leads to the production of small RNAs[7]. These small RNAs can mediate not only transcript degradation but also RNA-directed DNA methylation by guiding DRM2 to homologous sequences[3]. Alternatively, if the TE is not transcribed upon insertion but displays sequence homology with another DNA-methylated TE copy that accumulates small RNAs in the genome, then the preexisting small RNAs can similarly guide the *de novo* deposition of DNA methylation by DRM2 by virtue of sequence homology with the TE neo-copies ("homology-dependent silencing")[7]. The initial DNA methylation levels on a newly introduced repeat-containing transgene or a TE neo-insertion are usually not as high as those of the corresponding endogenous copy, and several generations are needed to reach these and a stably repressed state[8,9].

While very stable across generations and over evolutionary time, the DNA methylation patterns can nonetheless be constitutively pruned by the four members of the DEMETER (DME) family of 5-methylcytosine DNA glycosylases named DME (DEMETER), ROS1 (REPRESSOR OF SILENCING 1), DML2 and DML3 (DEMETER-LIKE 2 and 3). These DNA demethylases act both redundantly and in a locus-specific manner in vegetative tissues to counteract excessive DNA methylation[10–13]. DME (DEMETER), in particular, demethylates transposable elements in companion cells of male and female gametophytes[14,15] and is required to establish imprints in the endosperm[16]. In contrast to DNA methyltransferases, which have sequence context-dependent activities, DNA glycosylases can demethylate 5mC regardless of the sequence context[11,17,18]. Several functions have been shown for active DNA demethylation, such as preventing the spread of DNA methylation outside of primary targeted sequences[10] or actively participating in gene control, either at specific stages of development[19] or by preventing stress-responsive genes from being locked in a constitutively silent state[20,21].

Interestingly, in recent years, TEs have been shown to be decorated by another epigenetic silencing mark, H3K27me3 (trimethylation of lysine 27 of Histone 3)[22]associated with the evolutionary conserved Polycomb-group (PcG) proteins. This pathway mediates a more dynamic form of transcriptional repression than DNA methylation and has long been thought to specifically target protein-coding genes, particularly development genes for establishing cell identity, or stress-responsive genes[23,24]. H3K27me3 is catalyzed by Polycomb Repressive Complex 2 (PRC2), which is composed of four subunits and is present as different PRC2 cores in Arabidopsis. The catalytic activity is performed by either CLF, SWN or MEA, which are the three plant homologs of animal EZH2[23,24]. PRC2 is associated with PRC1, which is composed of two core subunits and mediates the ubiquitination of H2A (H2Aub) to allow a silent yet responsive transcriptional state of H3K27me3-marked targets[23]. In Arabidopsis, the recruitment of PRC2 to genes has been shown to be mediated via Transcription Factors (TF) that can recognize short recognition sequence motifs called Polycomb responsive elements (PREs), and that were initially described for the recruitment of PRC2 to Drosophila genes[23]. Long noncoding RNAs have also been involved in PRC2 targeting to genes[25,26]. Finally, PRC2 recruitment was also shown at some loci to be dependent on the presence of PRC1 and H2Aub[23].

Interestingly, in recent years, PcG has appeared to be the dominant system for silencing TEs in wild-type unicellular eukaryotes and plants from early lineages, thus illustrating the ancient role of PcG in regulating TE silencing[22,27]. However, in Arabidopsis, PcG can be recruited to three types of TEs[28]. First, approximately one-third of DNA-methylated TEs can gain H3K27me3 in mutants impaired for DNA methylation, such as *met1* or *ddm1*[29–31,] or in specific cell types that are naturally hypomethylated[32]. This finding implies that DNA methylation, particularly CG methylation, can exclude H3K27me3 deposition. In some instances, the two marks can co-localize[33–35,] and this was proposed to be constrained by DNA

methylation density[33,35]. In this context, the two marks can cooperate in restricting TE activation, for example upon biotic stress[35]. Finally, we recently reported that many TE are covered by the H3K27me3 histone mark in the genome of WT Arabidopsis plants, not only short TE relics but also relatively intact copies, which are expected to be targets of DNA methylation[28]. Whether active DNA demethylation can promote H3K27me3 recruitment in some of these cases is unknown.

In this study, we aimed to further investigate the crosstalk between DNA methylation and PcG proteins and reveal novel relationships between these pathways at TEs. Using neo-inserted TE transgenic sequences that are able to recruit both H3K27me3 and DNA methylation *de novo,* we provide evidence that the deposition of H3K27me3 is dependent on DRM2. This finding sheds light on an unsuspected interconnection between the PcG and RdDM pathways and points to a positive crosstalk between them in that context of TE neo-insertion, rather than the antagonism previously proposed. We further identified a subset of endogenous loci that lose H3K27me3 upon the loss of non-CG methylation maintenance in *drm2*, in addition to an anticipated subset the loss of CHH leads to an increase in H3K27me3. These results highlight a dual, locus-specific effect of DNA methylation on PcG recruitment. Accordingly, in the *ros1 dml2 dml3* triple mutant, the gain of 5mC leads to gain of H3K27me3 at some loci but also to loss of H3K27me3 at others, as validated by targeted DNA methylation experiments. These findings reveal a novel function of DNA demethylases in the modulation of H3K27me3 via DNA demethylation. Together, our results uncover new interdependencies between DNA methylation and the Polycomb machinery, particularly in the context of TE neo-insertion. Thus the two major silencing pathways, generally thought to separated, cooperate to shape the epigenomes, a concept that could extend to other multicellular eukaryotes.

RESULTS

DRM2 is involved in the establishment of H3K27me3 at a newly inserted TE sequence

We previously showed that neo-inserted, transgenic sequences of the mobile ATCOPIA21 retrotransposon (*AT5TE65370*) consistently recruit H3K27me3 *de novo* [28]. The endogenous copy is DNA methylated in WT plants and accumulates abundant siRNAs that are likely to target the homologous transgenic sequences (Fig. **S1A**). To verify this, we performed BS-seq on the pools of primary *COPIA21* transformants (TR) used for H3K27me3 analysis [28]. Because of the sequence homology between the endogenous and transgenic *COPIA21* sequences, the DNA methylation signal observed in the transgenic plants is an average of the signals at the endogenous and transgenic copies **(Fig. 1A left**). To circumvent this issue, we calculated, as a negative control, an 'expected' average DNA methylation signal if the transgenes were not methylated (**Fig. 1A right**, grey bars). The real DNA methylation average levels at *COPIA21* in the transgenic pools were greater than those in the negative control : this confirms that the transgenic copies are DNA methylated, yet less than the endogenous copies except in the CHH context (**Fig. 1A**). This higher CHH methylation at transgenes is caused by high *de novo* CHH methylation of the long-terminal repeats (**Fig. S1B**), presumably because small RNAs corresponding to these sequences are particularly abundant in WT plants (**Fig. S1A**).

To test for competition between DNA methylation and H3K27me3 in this context of neo-insertion, we transformed the *ATCOPIA21* construct into WT and *drm2* mutant plants impaired in *de novo* DNA methylation. We assessed H3K27me3 levels at the neo-inserted sequences in large pools of primary transformants of the same size for each genetic background. Surprisingly, we observed a drastic decrease in H3K27me3 at *ATCOPIA21* copies in the *drm2* mutant plants compared with the WT plants (**Fig. 1B**). This finding reveals that DRM2 is required for the *de novo* deposition of newly TE inserted sequences.

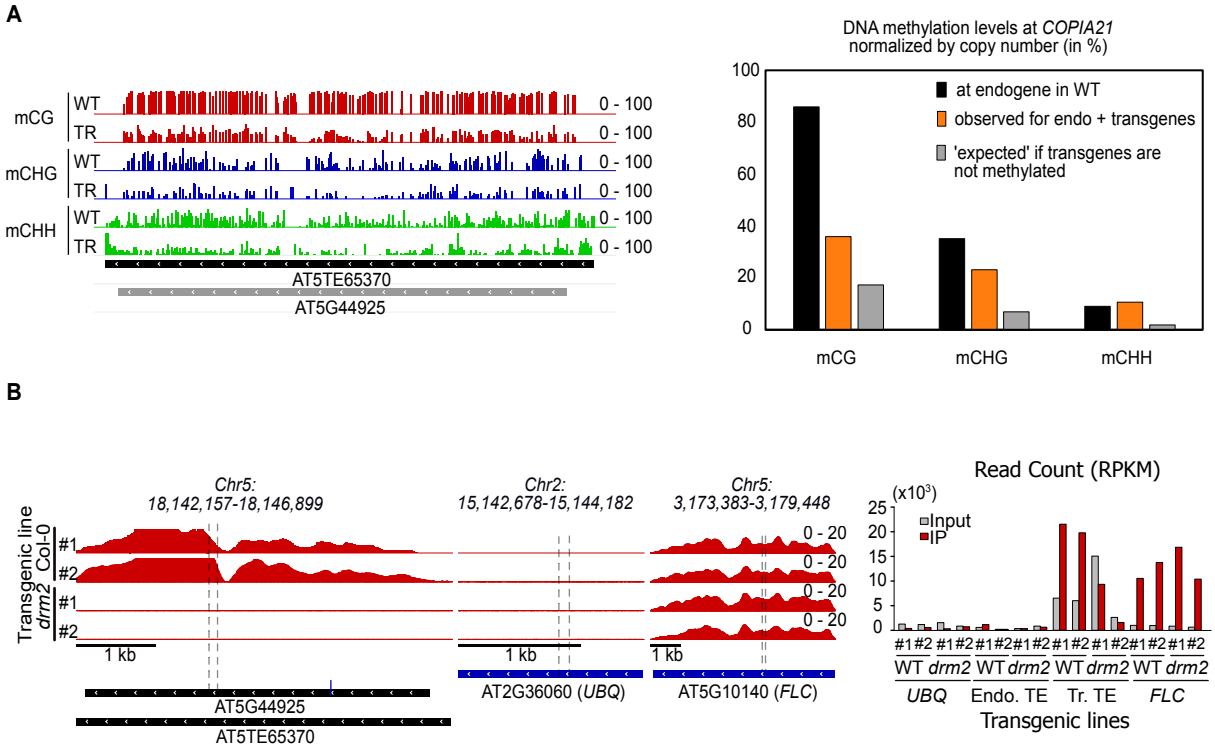

Figure 1. Role of DRM2 in the establishment of *de novo* H3K27me3 at a neo-inserted TE transgenic sequence
(A) Left panel : Representative genome browser view of DNA methylation status (CG : red, CHG : blue, CHH : green) at *COPIA21* in wild-type (WT) and in pools of primary transformants with transgenic *COPIA21* sequence (TR). **Right panel** : Barplot showing the DNA methylation levels in each context at *COPIA21* in WT (black bars), in the transgenic pools (orange bars). Light grey bars represent a prediction of DNA methylation levels in pools of transgenic plants if the transgenic TE sequences were not DNA methylated. **(B) Left panel.** Representative genome browser view of H3K27me3 average profile (IP-INPUT) at *COPIA21* sequence in two independent pools of plants transformed with *COPIA21* transgene (#1 and #2) either in WT "Col-0″ (upper lanes) or in *drm2* mutant (lower lanes). The H3K27me3 signal in these tracks reflects the H3K27me3 status at the endogene + transgene (SNP as blue line). **Middle panel**: *UBQ* and *FLC* are shown are negative and positive controls respectively for H3K27me3 recruitment. **Right panel:** Recruitment on the transgene is done by quantification of endogenous and transgenic *COPIA21* sequences immuno-precipitated with H3K27me3 using CHIP-seq reads at the region where a SNP was introduced to discriminate transgene and endogene and reads overlapping with the SNP at *COPIA21* locus were extracted, counted and normalized by the total read number. Recruitment on the transgene is validated by the enrichment observed in the IP as compared to the Input fraction. At control regions (*UBQ* and *FLC*), reads were extracted at positions *Chr5:3,178,750* (1st intron) and *Chr2:15,143,214* respectively). Copy number variation between transgenic pools can be visualized by the Input DNA fraction (grey bar).

DRM2 is involved in the maintenance of H3K27me3 at endogenous TEs

Next, we profiled H3K27me3 in *drm2* mutants and observed globally unchanged H3K27me3 levels at genes and endogenous TEs (**Fig. S2A**). Thus, DRM2 is neither necessary for general PRC2 activity nor has indirect effects on the PcG machinery. Closer inspection revealed an increase in H3K27me3 in *drm2* plants at a subset of ~900 TEs (**Fig. 2A left panel**): 32% of them overlapped with previously described, differentially methylated regions (DMR[36]) that significantly loose CHH methylation (**Fig. S2B**). This indicates that CHH methylation can antagonize H3K27me3 deposition (**Fig. 2A left, S2B and 2B left**), as previously described for CG methylation[30,31]. Of note, the extent of H3K27me3 gain did not seem proportional to the extent of CHH loss (**Fig. S2B**), possibly because the loss of one or two given cytosines, while not being called DMR, is sufficient to allow H3K27me3 deposition.

On the other hand, we observed a drastic loss of H3K27me3 in *drm2* at another subset of TEs (**Fig. 2A right panel, S2C and 2B right panel**), which is in line with the observations at the *ATCOPIA21* transgene. 36% of these TEs contain previously described CHH-DMRs[36] in *drm2* and showed a more drastic loss of H3K27me3 in *drm2* than the other TEs did (**Fig. S2C).** These results indicate that DRM2-mediated maintenance of CHH methylation at a TE subset can be important for proper H3K27me3 patterning, although a DNA methylation-independent role of DRM2 cannot be excluded.

DRM2-mediated H3K27me3 deposition is mostly specific to TEs

We next aimed to further characterize the contexts in which H3K27me3 deposition is dependent on DRM2 (i.e 7.1% of the H3K27me3 peaks, **Fig. S3A**). First, 17% and 3.6% of the

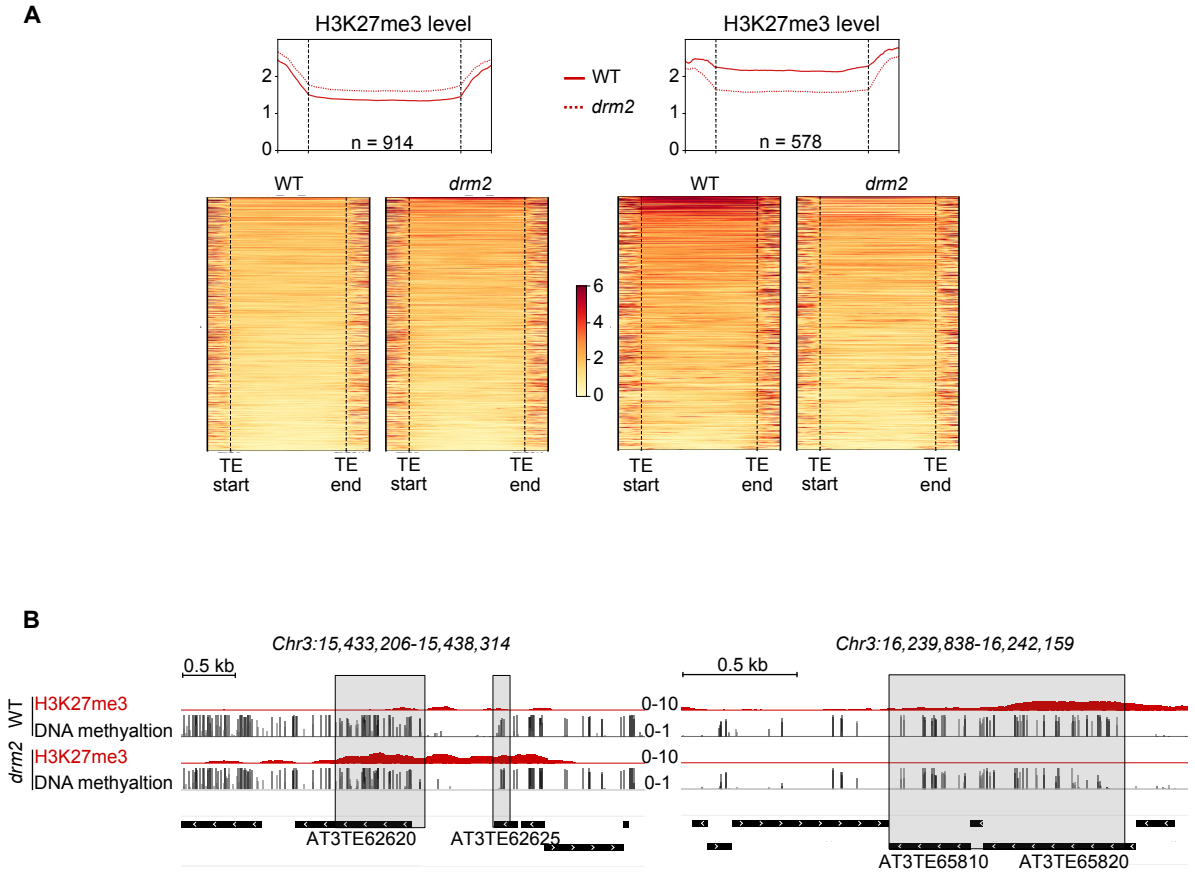

Figure 2. Role of DRM2 in the recruitment of H3K27me3 at TEs
(A**)** Heatmaps and metagenes showing H3K27me3 levels for TEs that gain H3K27me3 (**left panel**) or loose H3K27me3 (**right panel**) in *drm2* compared to WT. **(B)** Representative genome browser view showing H3K27me3 CHIP-seq and DNA methylation BS-seq data in WT and *drm2* mutant. Gain (**left panel**) and loss (**right panel**) of H3K27me3 in *drm2* compared to WT is associated with previously published DMRs (grey boxes (Stroud et al, 2013).

H3K27me3 peaks were detected at TEs and genes respectively, showing that this crosstalk is more specific to TEs. Second, H3K27me3 levels are lower at loci where H3K27me3 is dependent on DRM2 than at loci where H3K27me3 is independent of DRM2 (**Fig. 3B)**. In addition, the loci where H3K27me3 marks are DRM2-dependent harbor higher levels of 24 nt small RNAs and DNA methylation (in each context) than the loci where H3K27me3 marks are DRM2-independent and less DNA methylated **(Fig. 3C)**. This indicates a particular state at H3K27me3-marked, DRM2-dependent loci, where both marks with normally separate functions can colocalize and even cooperate in their recruitment; we coin this state 'ambivalent'.

Given that H3K27me3 colocalizes with and can be dependent on H2AK121ub at genes and TEs[23,28], we asked whether H2AK121ub is present at the 'ambivalent loci'. H2Aub peaks were detected at 92% of the loci with DRM2-independent H3K27me3 marks (**Fig. S3B top**) but at only 34% of the loci DRM2-dependent H3K27me3 marks (**Fig. S3B bottom)**. In addition, the length of the overlap between H3K27me3 and H2AK121ub was lower compared to loci where H3K27me3 is independent of DRM2 (mean of 271 bp and 1060 bp, respectively) (**Fig. S3C)**. Thus, 'ambivalent' TEs are less often associated with H2AUb.

Finally, motif enrichment analyses further revealed that loci where H3K27me3 is dependent on DNA methylation are enriched in C-rich motifs (particularly CHH) (**Fig. 3D and S3D)**. Thus, TEs whereby H3K27me3 depends on the presence of DRM2 display both genetic and epigenetic signatures that may underlie this 'ambivalent' chromatin state.

DNA demethylation impacts H3K27me3 patterning at endogenous TEs

To further explore the interconnections between DNA methylation and the PcG pathways, we profiled H3K27me3 in the triple DNA demethylase mutant *rdd (for*

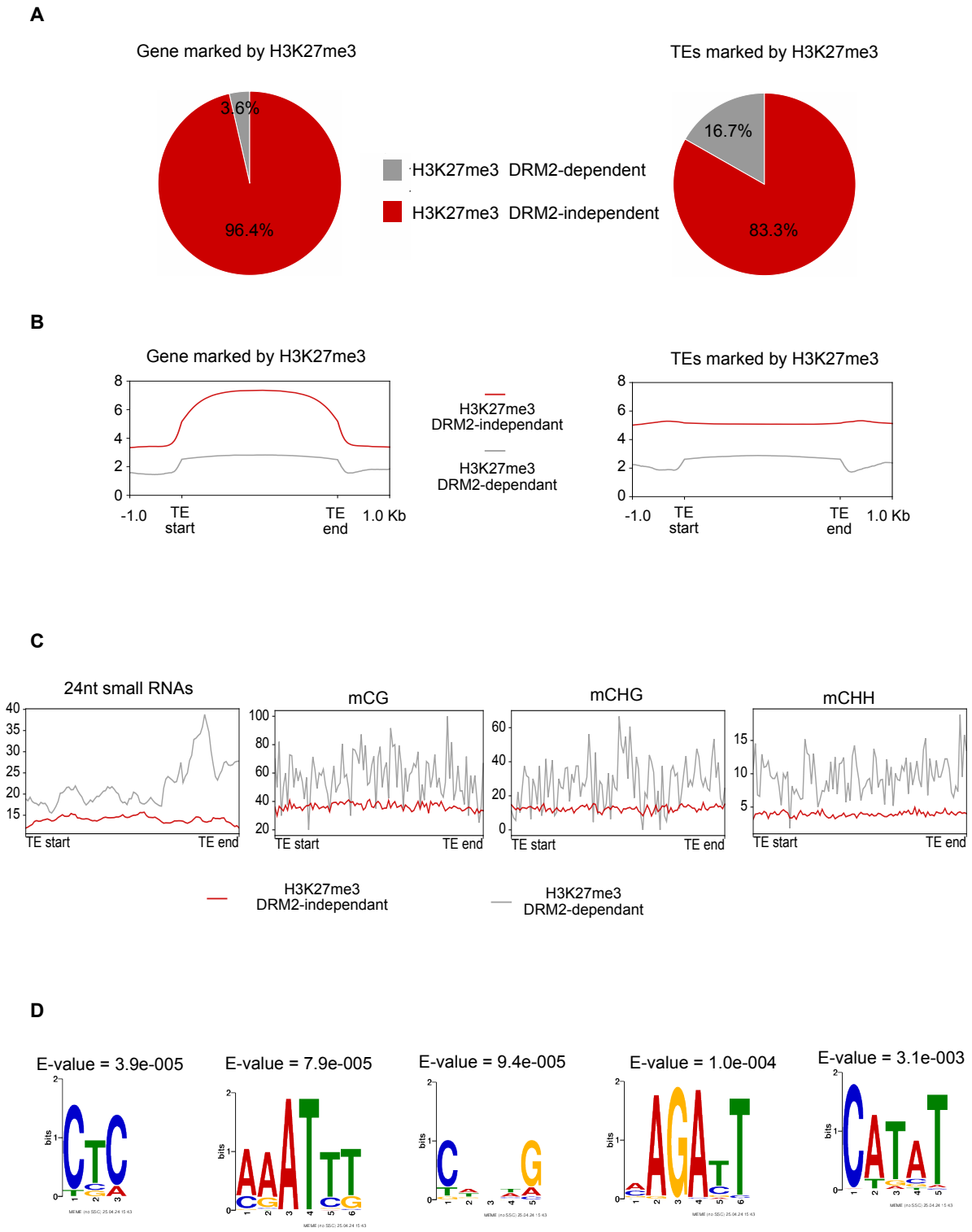

Figure 3. Characterization of loci where H3K27me3 depends on DRM2
(**A**) Pie chart showing the proportion of genes (left) and of TEs (right) marked by H3K27me3 dependently (grey) or independently of DRM2 (red). **(B)** Metaplots showing H3K27me3 profile of genes (left) and TEs (right) marked by H3K27me3 dependently (grey) and independently of DRM2 (red). **(C)** Metaplots showing 24nt small RNAs and DNA methylation (CG, CHG and CHH) profiles (respectively from left to right) for TEs marked by H3K27me3 dependently (grey) or independently of DRM2 (red). **(D)** Motifs enriched in H3K27me3 peaks dependent on DRM2 in comparison to H3K27me3 peaks independent of DRM2, determined by MEME analysis.

ros1/dml1/dml2). In a subset of TEs marked by H3K27me3 in WT and targeted by ROS1/DML1/DML2[37], we observed a loss of H3K27me3 upon DNA hypermethylation (**Fig. 4A-B and S4A left panels**). This finding indicates that H3K27me3 recruitment at some TEs, which we refer to as "Type 1-TEs," could be promoted by active demethylation. This is consistent with the previously described antagonistic effect of DNA methylation on H3K27me3 pathways. Regions with loss of H3K27me3 did not always overlap with the published DMRs[37], which could be explained by the observation that complete loss of H3K27me3 was associated with the gain of only a few cytosines at some loci (**Fig. S4B**). Interestingly, and in line with the loss of H3K27me3 observed at some TEs in *drm2* mutants , we also identified a subset of targets where the gain of DNA methylation in *rdd* leads to a gain of H3K27me3 (**Fig. 4A-B and S4A right panels**). These TEs are not or lowly marked by H3K27me3 in WT, and we refer to them as "Type 2-TE". Again, regions with a gain of H3K27me3 did not always overlap with the published DMRs in *rdd*[37]: this is likely explained by the observation that a gain of H3K27me3 is associated with the gain of only a few cytosines at some loci (**Fig. S4C**).

A targeted DNA methylation approach shows the direct impact of DNA methylation on H3K27me3 deposition in a locus-dependent manner

We next wanted to verify whether the loss of H3K27me3 at the DNA hypermethylated regions was the direct consequence of ectopic DNA methylation. For this purpose, we constructed an inverted repeat transgene to produce small RNAs able to override the dominance of ROS1 activity over RdDM[21] and induce DNA methylation[38] at a specific TE locus (*AT1TE59770*) (**Fig. 5A left panel, Fig. 5B**). In the two independent RNAi lines that we isolated, we observed a clear decrease in H3K27me3 at the endogenous TE compared to the transgenic lines with an unrelated transgene ("Target 1", **Fig. 5C** and **Fig. S5A left panels**).

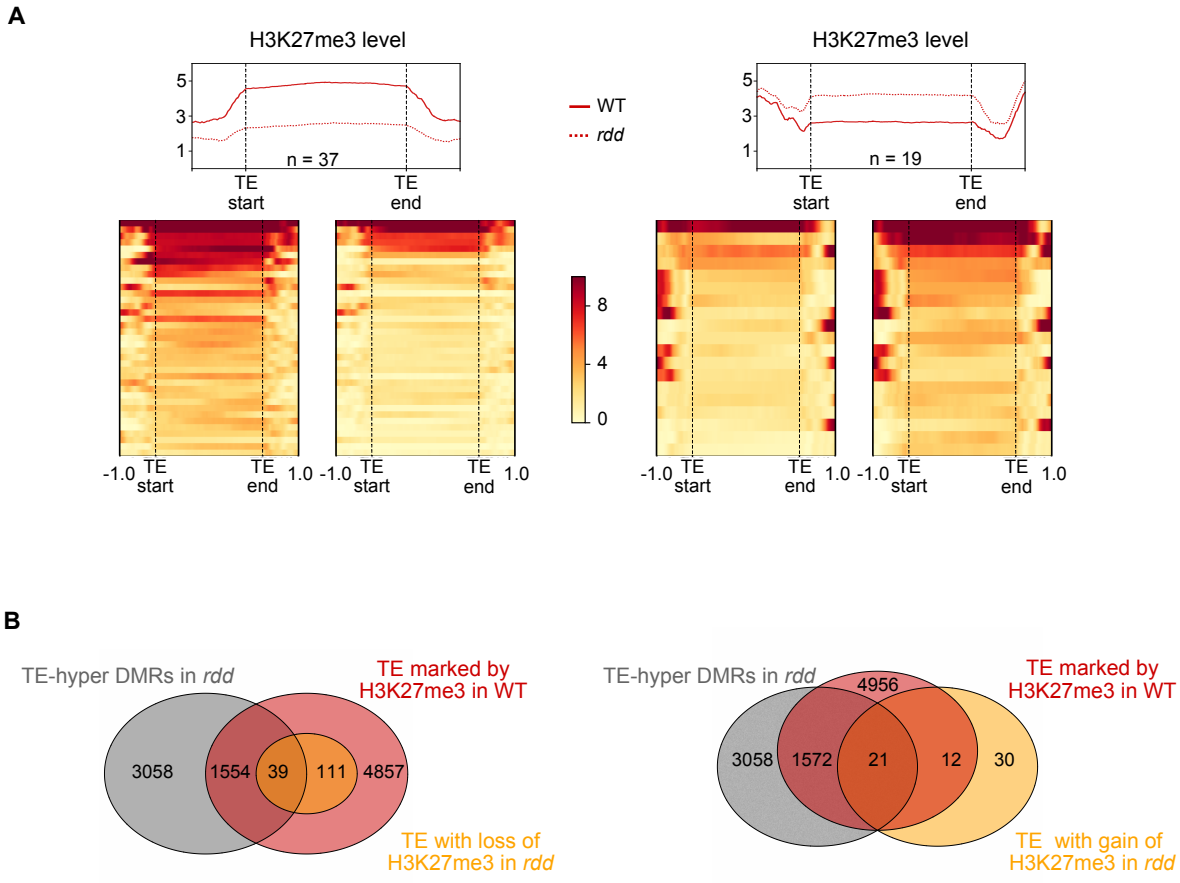

Figure 4. DNA demethylation impacts H3K27me3 patterning at endogenous TEs.
(A) Heatmap and metagene showing H3K27me3 levels for TEs that are DNA hypermethylated in *rdd* compared to WT (*Duan al, 2015*) and loose H3K27me3 (**left panel**) or gain H3K27me3 (**right panel**) in *rdd* compared to WT. **(B)** Venn diagram showing the overlap between DMRs identified in *rdd* compared to WT, TEs marked by H3K27me3 in WT and TEs that loose H3K27me3 marks (**left panel**) or TEs that gain H3K27me3 marks in *rdd* compared to WT (**right panel**).

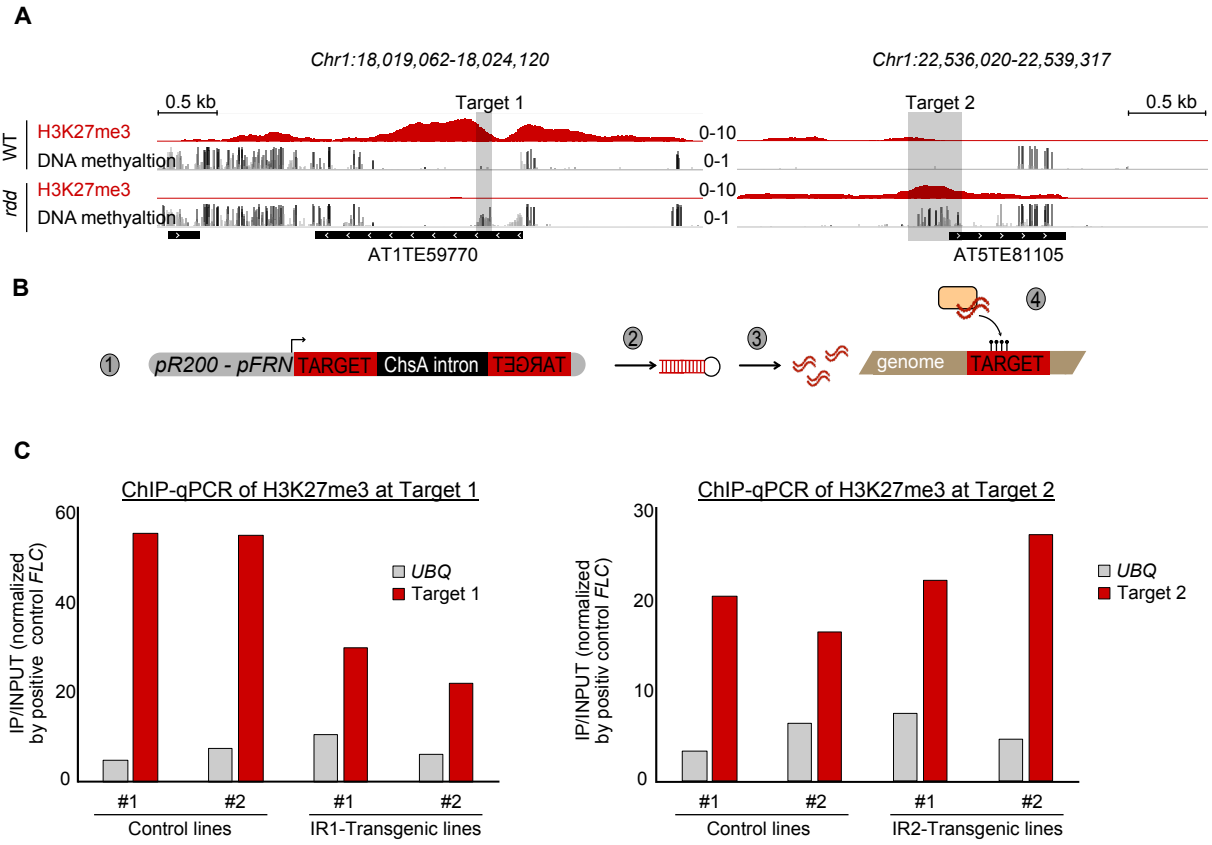

Figure 5. Targeted DNA methylation by inverted-repeats (IR) reveals a dual role of DNA methylation on H3K27me3 patterning at endogenous TEs
(A) Representative genome browser view showing H3K27me3 and DNA methylation in WT and *rdd* mutant. The region highlighted in grey is a DMR and was cloned as an inverted-repeat (IR) for the experiments shown in D and E. **(B)** Experimental design: **1)** Two TE target regions were chosen (1 and 2) ; for each target, the sequence was cloned as an inverted repeat (IR) separated by ChsA intron in the pR200-pFRN vector. **2)** Expression of the IR transgene leads to hairpin formation. **3)** Hairpins are processed by DICER-like proteins into smallRNAs. **4)** Small RNAs induce DNA methylation at the targeted locus. **(C)** ChIP-qPCR analysis of H3K27me3 marks in the control (transgenic lines with an unrelated transgene) and IR-transgenic lines at Target 1(IR1) (**left panel**) or Target 2 (IR2) (**right panel**). Data were normalized by the H3K27me3 ChIP positive control *FLC*; Ubiquitin (*UBQ)* serves as a negative control.

We thus show for the first time that targeting small RNA-directed DNA methylation to a H3K27me3-marked locus can cause a direct loss of H3K27me3.

To verify whether the gain of H3K27me3 at the DNA hypermethylated regions was the direct consequence of ectopic DNA methylation, we constructed another inverted repeat transgene, this time targeting a locus that gains H3K27me3 in *rdd* (*AT5TE81105*) (**Fig. 5A right panel, Fig. 5B**). In two independent RNAi lines, we observed a slight increase in H3K27me3 at the endogenous TE targeted by RNAi ("Target 2", **Fig. 5C** and **Fig. S5A right panels**). This indicates that targeting DNA methylation to a Type 2-TE can increase the recruitment of PRC2. This finding is in accordance with the striking observation that H3K27me3 deposition at a newly inserted TE is dependent on DRM2 (**Fig. 1C**) and provides further evidence that DNA methylation can favour H3K27me3 deposition in a locus-specific manner.

DISCUSSION

We previously showed that H3K27me3 recruitment and patterning at TEs share commonalities with PcG target genes, such as a partial dependency on PRC1, H2AZ variant incorporation, and the activity of JMJ histone demethylases; besides, sequence recognition motifs such as PREs[39] may also be involved in H3K27me3 deposition since H3K27me3 appears to be instructed by the TE sequence itself[28]. Here, by studying the establishment of H3K27me3 on a newly inserted TE sequence, we reveal a novel dependency of PRC2 activity on the DRM2 de novo methyltransferase, which is TE-specific (**Fig. 6 A**). *De novo* deposition of H3K27me3 in this context was impaired in the *drm2* mutant, which has never been reported in any organism before.

In addition, DRM2 is also required for proper maintenance of H3K27me3 patterns at a subset of TE loci, as recently reported in rice genes, where H3K27me3 and non-CG methylation

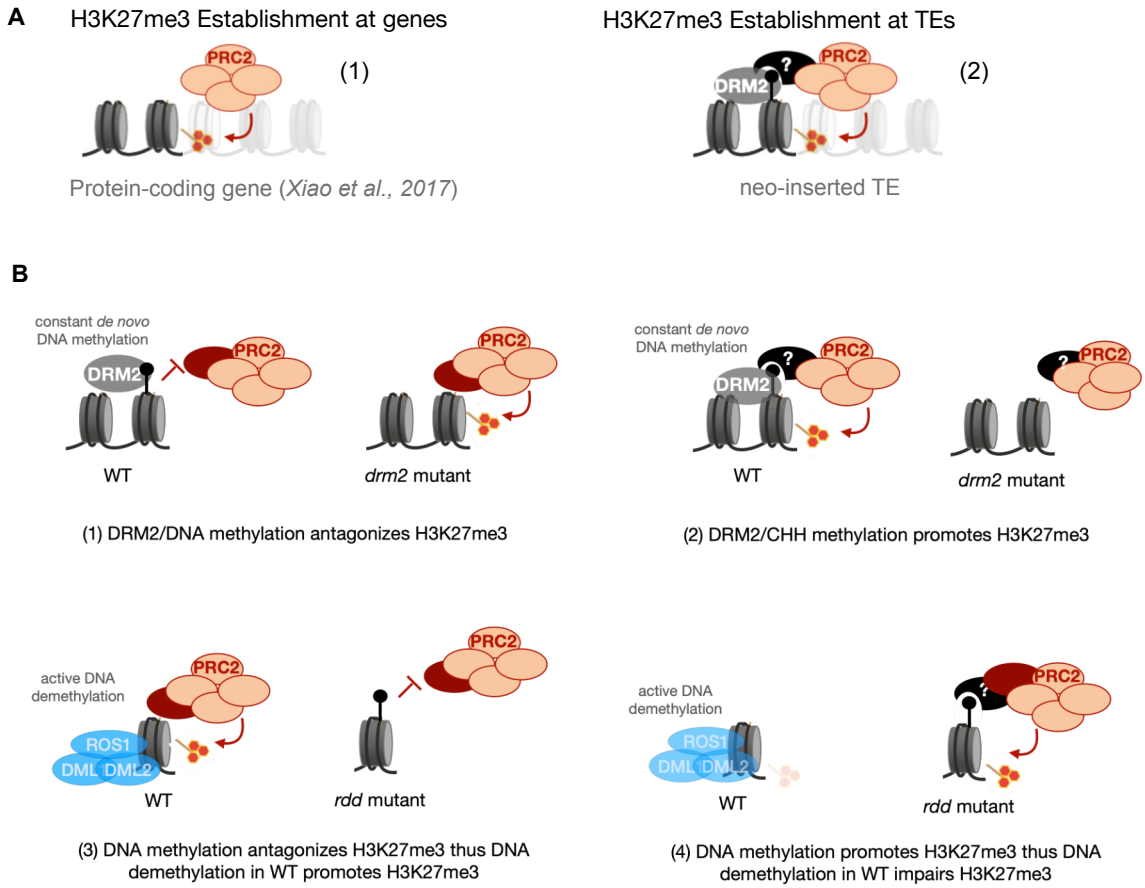

Figure 6. Scenario for differential epigenetic targeting of TEs, in a locus specific manner .
(A) Proposed model for establishment of H3K27me3 at neo-inserted TEs. **1)** PRC2 recruitment on unmethylated protein-coding genes (previously shown to be dependent on specific transcription factors (TF), *Xiao et al., 2017*). **2)** PRC2 recruitment on a neo-inserted TE sequence is DRM2-dependent. Possibly the same TF are involved and may need DNA methylation/DRM2 as a docking site to enhance their recruitment; alternatively unknown PRC2 accessory proteins with affinity to DNA methylation/DRM2 could be involved. **(B)** Model for locus-specific rules of H3K27me3 maintenance at endogenous TEs. **1)** These TEs are DNA methylated in WT and gain H3K27me3 marks in DNA methylation mutants, notably in *drm2*. DNA methylation presumably prevents the binding of a TF or PRC2 co-factor . Depending on the sequence motif recognized by the TF, the context of DNA methylation preventing TF binding may be different, hence the differences seen in *met1* (our previous work) versus *drm2* (this study) in the gain of H3K27me3. **2**) These TEs are both DNA methylated and marked by H3K27me3 in WT, and they loose both CHH and H3K27me3 in *drm2* mutant. DNA methylation (CHH methylation) would be needed to promote PRC2 recruitment via a TF or PRC2 co-factor which is affine to DNA methylation and requires it as a docking site. Presumably, at these TE, H3K27me3 is constantly established *de novo*, similarly as DRM2-mediated CHH methylation. **3**) These TEs (Fig.5 “Type1”) are marked by H3K27me3 only in WT, and loose H3K27me3 upon gain of DNA methylation in *rdd* mutant. PRC2 recruitment is sensitive to DNA methylation, and active demethylation by DML enzymes promotes H3K27me3 recruitment at these TEs. **4**) These TEs (Fig.5 “Type2”) are weakly marked by H3K27me3 and DNA methylation in WT and gain H3K27me3 upon gain of DNA methylation in *rdd*. PRC2 recruitment would be dependent on DNA methylation (as in 2). Together, these TE-specific behaviors point to genetic and epigenetic signatures for PRC2 recruitment and PRC2 antagonism respectively.

can colocalize[40,41]. Accordingly, we show that active DNA demethylation can result in the loss of H3K27me3 at some TEs. In contrast, DNA methylation can antagonize H3K27me3, as indicated by the increase of H3K7me3 in *drm2* and the loss of H3K27me3 in *rdd* in different TE subsets. The antagonistic effect of CG methylation on H3K27me3 deposition was previously shown[30]; however, our results provide two novel insights. First, DRM2-mediated CHH methylation can also antagonize PRC2 recruitment, which is consistent with recent H3K27me3 profiles of rice and maize RdDM mutants[42,43]. Second, we demonstrate a novel role for DNA demethylases in promoting PRC2 recruitment. This result is in line with previous observations that ROS1 targets are enriched in H3K27me3[44] and could provide an explanation as to why metazoan TET enzymes are enriched at the hypomethylated DNA promoters of PcG targets[45]. Taken together, our results point to locus-specific rules for H3K27me3 deposition, which would either require DNA methylation or be antagonized by its presence (**Fig. 6 B**). This suggests a mode of PRC2 recruitment determined by specific genetic and epigenetic signatures and the existence of a set of transcription factor (TF) or PRC2 co-factors that may display different affinities for DNA methylation and need to be identified in the future. This does not exclude direct interactions between PRC2 and DRM2 or the RdDM machinery in a locus-specific manner or in a context-specific manner (in the context of neo-insertion, for example). In that respect, the transgenic system that we have established provides a framework to dissect the complex interactions between H3K27me3 and DNA methylation during establishment and maintenance throughout generations at a single TE copy in future endeavors.

The interconnection that we revealed between *de novo* DNA methylation and deposition of H3K27me3 on newly inserted TEs argues for active cooperation between the two marks at this stage of the TE life cycle. This raises the question as to why such cooperation would take place when a novel copy has just integrated the genome. When they are inserted into the genome, retrotransposons are DNA hypomethylated. The targeting of PcG at this stage could

thus allow rapid silencing of the element while DNA methylation is being established progressively throughout successive generations. One exciting question to address in the future is whether H3K27me3 persists after one round of reproduction or if DNA methylation quickly becomes dense enough to antagonize H3K27me3 in subsequent generations. If this is the case, H3K27me3 could just be a transient silencing form established quickly by the plant to compensate for low *de novo* DNA methylation in the first generation(s) after TE neo-insertion. Our study thus points to a synergy between the two silencing pathways, a concept that also recently emerged at other types of repeats[46,47].

For the endogenous TEs, we identified two additional layers of H3K27me3 regulation linked with DNA methylation. First, we demonstrated that DNA demethylases can target specific TEs to promote H3K27me3 deposition instead of DNA methylation. This could be important for TEs, which, as epigenetic modules, negatively impact nearby gene regulation When marked by H3K27me3 instead of DNA methylation, contribute to a partially repressed state as opposed to a locked silencing state, since H3K27me3 is more plastic and labile in response to developmental or environmental cues. Similarly, the identification of TEs where DNA methylation and H3K27me3 marks not only co-occur but are also interdependent points to what we refer to as an "ambivalent state", which is actively maintained, and where the loss of one mark is linked to the loss of the other. This could similarly be advantageous for nearby gene regulation; for example, a decrease in CHH methylation in response to certain stresses (such as pathogen-induced stress[48]) could lead to a concomitant decrease in the H3K27me3 mark even if the latter is not sensitive per se to that stress. In that sense, these ambivalent TEs could constitute sensitized modules for the dynamic regulation of nearby genes.

By connecting PRC2 to DNA methylation at TEs, our work in Arabidopsis provides important insights into the separation and specialization of the two major silencing pathways throughout eukaryotic evolution. We previously proposed that PcG was an ancestral system of

TE silencing based on H3K27me3 being the dominant mark at TEs in unicellular organisms or ancestral plants. Furthermore, in ciliates, a small RNA-guided enhancer of zeste (the conserved catalytic subunit of PRC2) results in both H3K27me3 and H3K9 methylation[49,50], which either reflects ciliate-specific catalytic activity or suggests that ancestral PRC2 has both activities. We propose that with the evolution of multicellularity and the need for a dynamic system to control developmental transitions, the PcG and H3K9me2/DNA methylation pathways may have specialized for the silencing of genes and TEs, respectively. Our present results nevertheless show that the major silencing pathways in eukaryotes maintain mechanistic connections despite specialization for different functions in higher plants, which is likely to help their functional cooperation in silencing in specific contexts. Such interconnections may exist in other kingdoms, as suggested by the small RNA-driven deposition of H3K27me3 in *C. elegans*[51,52] or the existence of AEBP2, a mammalian PRC2 cofactor that requires DNA methylation for its activity[53]. Conversely, DNA methylation could be dependent on PRC2, as previously suggested[54,55,] and it would be interesting to investigate this possibility in plants. Future work needs to further decipher the connections between PcG and H3K9/DNA methylation and how they evolved from unicellular to multicellular organisms. This should undoubtedly shed light on the evolution of silencing pathways in eukaryotes and how they shape host genome regulation.

MATERIAL AND METHODS

Plant material and growth conditions

All the experiments were conducted on *A. thaliana* on ½ MS plates under short light-day conditions (8-h light/16-h dark photoperiod at 22°C). For the transgenic plants shown in Fig. 1, 4- week-old rosette leaves were pooled and collected for further ChIP-seq and BS-seq analysis.

Mutant lines

We used the *drm1-2 drm2-2* double mutant (Salk-031705, Salk-150863)[4] and the *ros1 dml1 dml2* triple mutant (Col-0 background, derived from Salk_045303 Salk_056440 Salk_131712)[13]

Generation of transgenic lines

The *COPIA21* TE sequence was synthesized and cloned and inserted into pUC57 via Genescript. *COPIA21* TE was subsequently cloned and inserted into pCAMBIA3300. The plants were subsequently transformed via *Agrobacterium tumefaciens* floral dip[56]. For the experiments shown in Fig.1, the transgenic plants were selected on Basta after 2 weeks, transferred in soil and 4 weeks old rosettes leaves were collected in pools of 15-20 plants to perform ChIP-seq and BS-seq (on the same ground tissue).

RNAi lines were obtained by cloning approximately 250 bp fragments in an inverted orientation via the pFRN vector. The plants were subsequently transformed via *Agrobacterium tumefaciens* floral dip[56,] and T1 plants were selected in-vitro with Kanamycin resistance.

Chromatin immunoprecipitation (ChIP) and ChIP–qPCR/sequencing analyses

ChIP experiments were conducted in WT or appropriate mutant lines via an anti-H3K27me3 antibody. IP and INPUT DNA were eluted, purified and sequenced (100 bp paired-end; Illumina) by BGI. Reads were mapped via BWA[57] onto TAIR10 *A. thaliana.* Genomic regions significantly marked by H3K27me3 were identified via MACS2[58,] and genes or TEs overlapping these regions were obtained via bedtools[59]. Heatmaps and plotprofiling were generated via bedtools computeMatrix to create a score matrix, and plotHeatmap was used to generate a graphical output of the matrix. For ChIP-seq analyses in *rdd* mutant, regions

inherited from Ws-2 *ros1* and *dml2* mutants (Ws background) after backcrossing into Col-0 were excluded from analysis as previously described[13]

Read count analyses

Reads overlapping with the SNPs (between transgenes and endogenes) at *ATCOPIA21* were extracted, counted and normalized by total read number via SAMtools view. In the control regions (UBQ and FLC), reads were extracted at positions *Chr2:15,143,214* (UBQ) and *Chr5:3,178,750* (1st intron FLC), respectively.

Bisulfite-sequencing analyses

Adapter and low-quality sequences were trimmed via Trimming Galore 0.6.5. Mapping was performed on the TAIR10 genome annotation via Bismark v0.22.2[60] and Bowtie2[61].

ACKNOWLEDGEMENTS

We thank N. Bouché for his help with some bioinformatic analyses. We thank ANRJCJC (ANR-19-CE12-0033-01 to A.D.) for funding and the Genome Biology Department of Institut de Biologie Intégrative de la Cellule (I2BC) for support. We thank the services and platforms of the Institut de Biologie Intégrative de la Cellule (I2BC) for excellent technical support and in particular Véronique Couvreux for excellent plant care.

REFERENCES

1. Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8**, 272–285 (2007).
2. Lyons, D. B. & Zilberman, D. DDM1 and Lsh remodelers allow methylation of DNA wrapped in nucleosomes. *eLife* **6**, e30674 (2017).
3. Zhang, H., Lang, Z. & Zhu, J.-K. Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell Biol.* **19**, 489–506 (2018).
4. Cao, X. & Jacobsen, S. E. Role of the Arabidopsis DRM Methyltransferases in De Novo DNA Methylation and Gene Silencing. *Curr. Biol.* **12**, 1138–1144 (2002).
5. Chan, S. W.-L. *et al.* RNA Silencing Genes Control de Novo DNA Methylation. *Science* **303**, 1336–1336 (2004).
6. Greenberg, M. V. C. *et al.* Identification of genes required for de novo DNA methylation in Arabidopsis. *Epigenetics* **6**, 344–354 (2011).
7. Fultz, D. & Slotkin, R. K. Exogenous Transposable Elements Circumvent Identity-Based Silencing, Permitting the Dissection of Expression-Dependent Silencing. *Plant Cell* **29**, 360–376 (2017).
8. Ausin, I. *et al.* INVOLVED IN DE NOVO 2-containing complex involved in RNA-directed DNA methylation in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 8374–8381 (2012).
9. Teixeira, F. K. & Colot, V. Repeat elements and the Arabidopsis DNA methylation landscape. *Heredity* **105**, 14–23 (2010).
10. Penterman, J. *et al.* DNA demethylation in the Arabidopsis genome. *Proc. Natl. Acad. Sci.* **104**, 6752–6757 (2007).
11. Ortega-Galisteo, A. P., Morales-Ruiz, T., Ariza, R. R. & Roldán-Arjona, T. Arabidopsis DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. *Plant Mol. Biol.* **67**, 671–681 (2008).
12. Zhu, J., Kapoor, A., Sridhar, V. V., Agius, F. & Zhu, J.-K. The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in Arabidopsis. *Curr. Biol. CB* **17**, 54–59 (2007).
13. Williams, B. P., Bechen, L. L., Pohlmann, D. A. & Gehring, M. Somatic DNA demethylation generates tissue-specific methylation states and impacts flowering time. *Plant Cell* **34**, 1189–1206 (2021).
14. Calarco, J. P. *et al.* Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell* **151**, 194–205 (2012).
15. Ibarra, C. A. *et al.* Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* **337**, 1360–1364 (2012).
16. Gehring, M. *et al.* DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124**, 495–506 (2006).
17. Agius, F., Kapoor, A. & Zhu, J.-K. Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11796–11801 (2006).
18. Morales-Ruiz, T. *et al.* DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 6853–6858 (2006).
19. Parrilla-Doblas, J. T., Roldán-Arjona, T., Ariza, R. R. & Córdoba-Cañero, D. Active DNA Demethylation in Plants. *Int. J. Mol. Sci.* **20**, 4683 (2019).
20. Yang, L. *et al.* ROS1-mediated decrease in DNA methylation and increase in expression of defense genes and stress response genes in *Arabidopsis thaliana* due to abiotic

stresses. *BMC Plant Biol.* **22**, 104 (2022).
21. Halter, T. *et al.* The Arabidopsis active demethylase ROS1 cis-regulates defense genes by erasing DNA methylation at promoter-regulatory regions. *eLife* **10**, e62994 (2021).
22. Déléris, A., Berger, F. & Duharcourt, S. Role of Polycomb in the control of transposable elements. *Trends Genet. TIG* **37**, 882–889 (2021).
23. Baile, F., Gómez-Zambrano, Á. & Calonje, M. Roles of Polycomb complexes in regulating gene expression and chromatin structure in plants. *Plant Commun.* **3**, 100267 (2022).
24. Baile, F. & Calonje, M. Dynamics of polycomb group marks in Arabidopsis. *Curr. Opin. Plant Biol.* **80**, 102553 (2024).
25. Wu, H.-W. *et al.* A noncoding RNA transcribed from the *AGAMOUS* (*AGL*) second intron binds to CURLY LEAF and represses expression in leaves. *New Phytol.* **219**, 1480–1491 (2018).
26. Fonouni-Farde, C. *et al.* The Arabidopsis APOLO and human UPAT sequence-unrelated long noncoding RNAs can modulate DNA and histone methylation machineries in plants. *Genome Biol.* **23**, 181 (2022).
27. Hisanaga, T. *et al.* The Polycomb repressive complex 2 deposits H3K27me3 and represses transposable elements in a broad range of eukaryotes. *Curr. Biol. CB* S0960-9822(23)01153–3 (2023) doi:10.1016/j.cub.2023.08.073.
28. Hure, V. *et al.* Alternative silencing states of Transposable Elements in Arabidopsis. 2024.03.16.585326 Preprint at https://doi.org/10.1101/2024.03.16.585326 (2024).
29. 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).
30. 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).
31. Rougée, M. *et al.* Polycomb mutant partially suppresses DNA hypomethylation–associated phenotypes in Arabidopsis. *Life Sci. Alliance* **4**, (2021).
32. Weinhofer, I., Hehenberger, E., Roszak, P., Hennig, L. & Köhler, C. H3K27me3 Profiling of the Endosperm Implies Exclusion of Polycomb Group Protein Targeting by DNA Methylation. *PLoS Genet.* **6**, e1001152 (2010).
33. Brinkman, A. B. *et al.* Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res.* **22**, 1128–1138 (2012).
34. Statham, A. L. *et al.* Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA. *Genome Res.* **22**, 1120–1127 (2012).
35. Zervudacki, J., Agnès, Y., Delase, A., Navarro*, L. & Deleris, A. Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons. *EMBO J.* **37**, e98482 (2018).
36. Stroud, H., Greenberg, M. V. C., 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).
37. Duan, C.-G. *et al.* MET18 Connects the Cytosolic Iron-Sulfur Cluster Assembly Pathway to Active DNA Demethylation in Arabidopsis. *PLOS Genet.* **11**, e1005559 (2015).
38. Sigman, M. J. *et al.* An siRNA-guided ARGONAUTE protein directs RNA polymerase V to initiate DNA methylation. *Nat. Plants* **7**, 1461–1474 (2021).
39. Xiao, J. *et al.* Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. *Nat. Genet.* **49**, 1546–1552 (2017).

40. Zhou, S. *et al.* Cooperation between the H3K27me3 Chromatin Mark and Non-CG Methylation in Epigenetic Regulation1. *Plant Physiol.* **172**, 1131–1141 (2016).
41. Zhang, W. *et al.* Domains Rearranged Methylase 2 maintains DNA methylation at large DNA hypomethylated shores and long-range chromatin interactions in rice. *Plant Biotechnol. J.* (2023) doi:10.1111/pbi.14134.
42. Hari Sundar G, V. *et al.* Plant polymerase IV sensitizes chromatin through histone modifications to preclude spread of silencing into protein-coding domains. *Genome Res.* **33**, 715–728 (2023).
43. Guo, W., Wang, D. & Lisch, D. RNA-directed DNA methylation prevents rapid and heritable reversal of transposon silencing under heat stress in *Zea mays*. *PLoS Genet.* **17**, e1009326 (2021).
44. Tang, K., Lang, Z., Zhang, H. & Zhu, J.-K. The DNA demethylase ROS1 targets genomic regions with distinct chromatin modifications. *Nat. Plants* **2**, 16169 (2016).
45. Verma, N. *et al.* TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat. Genet.* **50**, 83–95 (2018).
46. Eimer, H. *et al.* RNA-Dependent Epigenetic Silencing Directs Transcriptional Downregulation Caused by Intronic Repeat Expansions. *Cell* **174**, 1095-1105.e11 (2018).
47. Sureshkumar, S. *et al.* SUMO protease FUG1, histone reader AL3 and chromodomain protein LHP1 are integral to repeat expansion-induced gene silencing in *Arabidopsis thaliana*. *Nat. Plants* **10**, 749–759 (2024).
48. Deleris, A., Halter, T. & Navarro, L. DNA Methylation and Demethylation in Plant Immunity. *Annu. Rev. Phytopathol.* **54**, 579–603 (2016).
49. Miró-Pina, C. *et al.* Paramecium Polycomb repressive complex 2 physically interacts with the small RNA-binding PIWI protein to repress transposable elements. *Dev. Cell* **57**, 1037-1052.e8 (2022).
50. Balan, T., Lerner, L. K., Holoch, D. & Duharcourt, S. Small-RNA-guided histone modifications and somatic genome elimination in ciliates. *WIREs RNA* **15**, e1848 (2024).
51. Mao, H. *et al.* The Nrde Pathway Mediates Small-RNA-Directed Histone H3 Lysine 27 Trimethylation in *Caenorhabditis elegans*. *Curr. Biol.* **25**, 2398–2403 (2015).
52. Seroussi, U. *et al.* Mechanisms of epigenetic regulation by *C. elegans* nuclear RNA interference pathways. *Semin. Cell Dev. Biol.* **127**, 142–154 (2022).
53. Wang, X. *et al.* Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. *Nat. Struct. Mol. Biol.* **24**, 1028–1038 (2017).
54. Schlesinger, Y. *et al.* Polycomb-mediated methylation on Lys27 of histone H3 premarks genes for de novo methylation in cancer. *Nat. Genet.* **39**, 232–236 (2007).
55. Viré, E. *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871–874 (2006).
56. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J. Cell Mol. Biol.* **16**, 735–743 (1998).
57. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinforma. Oxf. Engl.* **25**, 1754–1760 (2009).
58. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
59. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma. Oxf. Engl.* **26**, 841–842 (2010).
60. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinforma. Oxf. Engl.* **27**, 1571–1572 (2011).
61. 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).

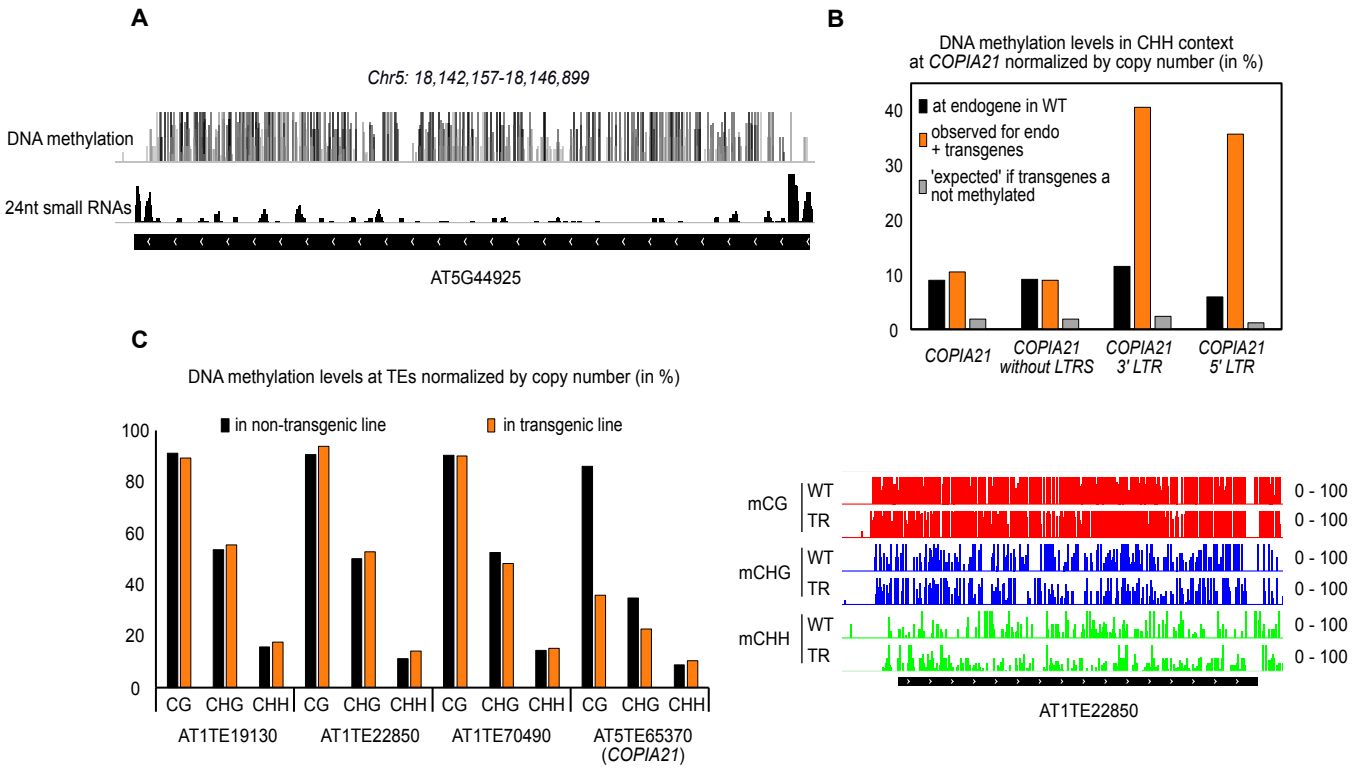

Figure S1. Role of DRM2 in the establishment of *de novo* H3K27me3 at a neo-inserted TE transgenic sequence
(A) Representative genome browser of DNA methylation (Overlay of CG in black, CHG in dark grey and CHH in light grey) and 24nt small RNAs levels at *COPIA21* in wild-type (WT). **(B)** Barplot showing the DNA methylation levels in CHH context at *COPIA21* in WT (black bars), in the transgenic pools (orange bars). Light grey bars represent a prediction of DNA methylation levels in pools of transgenic plants if the transgenic TE sequences were not DNA methylated. **(C) Left panel** : Barplot showing the DNA methylation levels in each context at different TEs in WT (black bars) and in transgenic pools (orange bars). **Right panel** : Representative genome browser of DNA methylation status (CG : red, CHG : blue, CHH : green) at *AT1TE22850* in wild-type (WT) and in transgenic pools (TR).

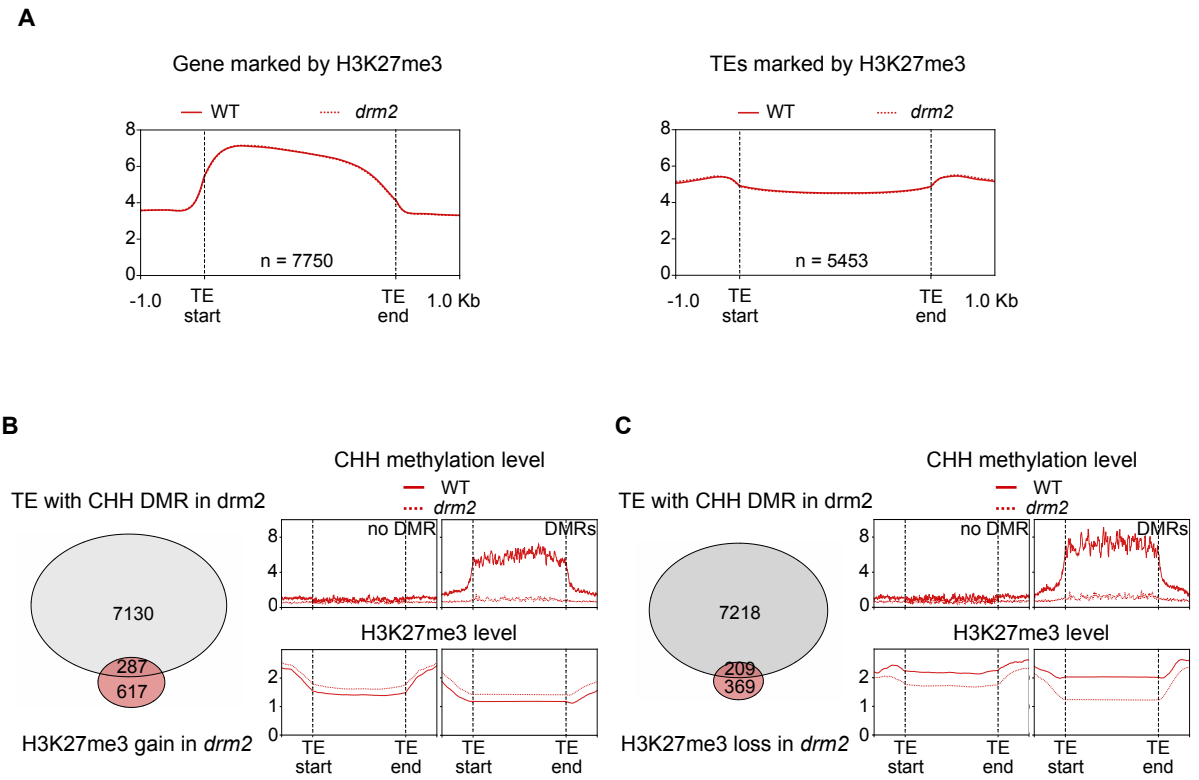

Figure S2. Role of DRM2 in the recruitment of H3K27me3 at TEs
(**A)** Metagenes showing H3K27me3 level for genes (**left panel**) and TEs (**right panel**) marked by H3K27me3, in WT and *drm2* mutant. **(B)** Venn diagram showing the overlap between TEs that gain H3K27me3 in *drm2* compared to WT and TEs which contain DMRs in the CHH context (Stroud al, 2013). TEs with (**right panel**) or without (**left panel**) DMRs and H3K27me3 gain in *drm2* are profiled for CHH methylation and H3K27me3 status. **(C)** Venn diagram showing the overlap between TEs that lose H3K27me3 in *drm2* compared to WT and TEs with DMRs in CHH context. TEs with (**right panel**) or without (**left panel**) DMRs and H3K27me3 loss in *drm2* are profiled for CHH and H3K27me3 status.

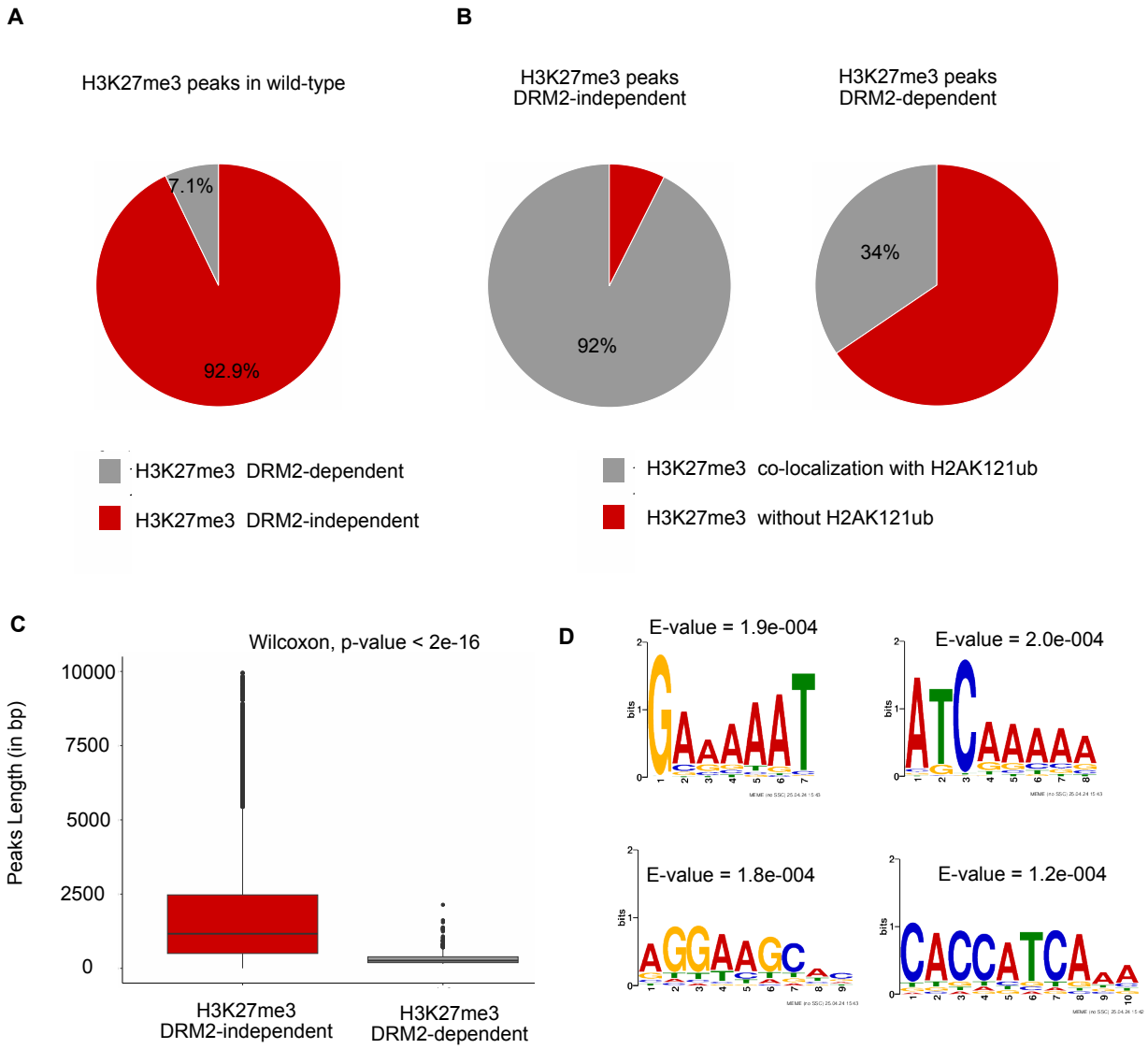

Figure S3. Characterization of H3K27me3 loci which are dependent on DRM2
(A) Pie chart showing proportion of peaks marked by H3K27me3 dependently (grey) and independently of DRM2 (red). **(B)** Pie charts showing proportion of peaks marked by H3K27me3 that co-localize with H2AK121ub (grey) or not (red). **(C)** Boxplots showing the length of overlap between H3K27me3 and H2AK121ub peaks, either for loci where H3K27me3 are DRM2-independent (red) or DRM2-dependent (grey). Statistical analysis was performed using Wilcoxon-Mann-Whitney U-test. **(D)** Motifs enriched in peaks dependent on DRM2 in comparison to DRM2-independent peaks and determined by MEME analysis with minimal width at 7.

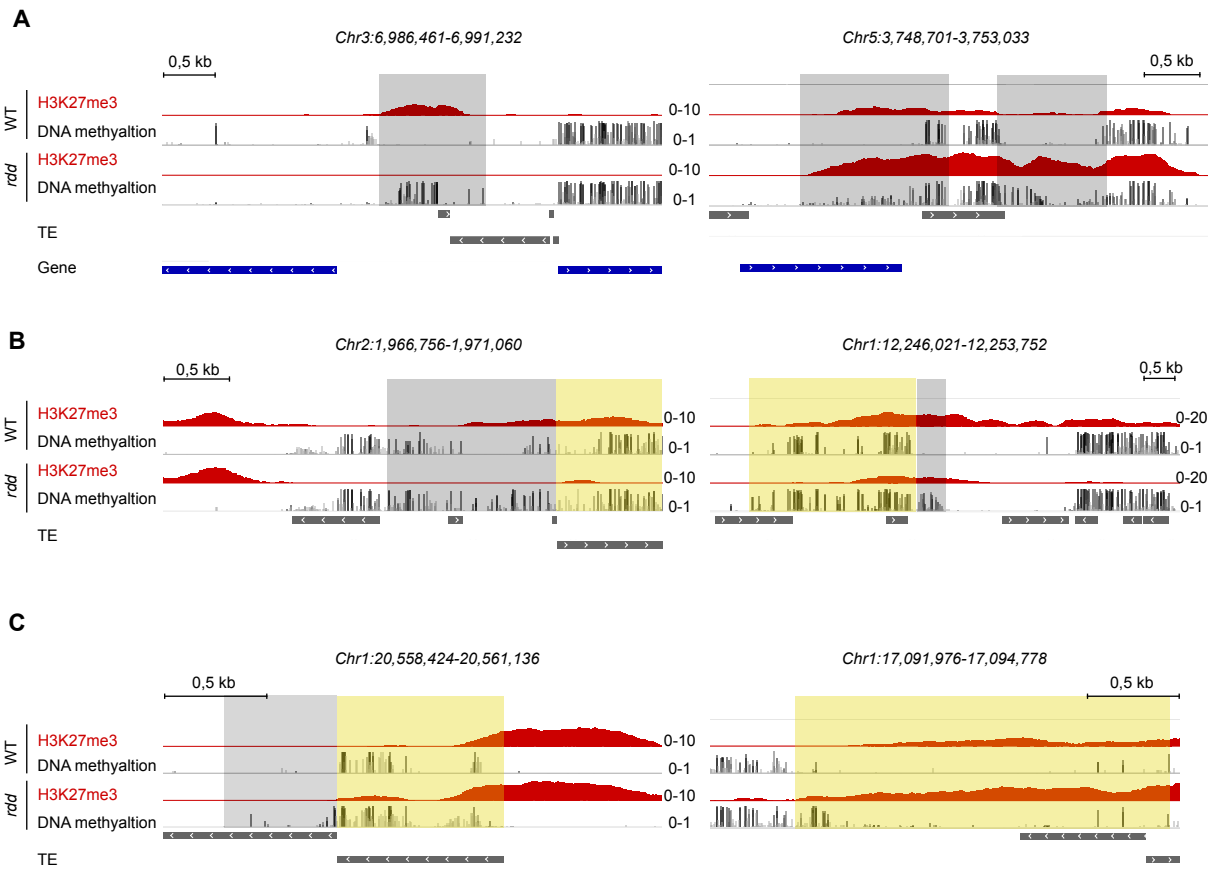

Figure S4. DNA demethylation impacts H3K27me3 patterning at endogenous TEs
Representative genome browser views showing H3K27me3 and DNA methylation in WT and *rdd* mutant. Regions highlighted in grey are DMRs previously identified (*Duan et al,. 2015*) whereas regions highlighted in yellow are regions that were not previously identified as DMRs despite some changes in cytosine methylation. We prepose that these regions could explain why there are some H3K27me3 changes in *rdd* which are in appearance not overlapping with DMRs. **(A)** and **(B)** panels show loss of H3K27me3 in *rdd* and **(C)** panels show gain of H3K27me3 in *rdd*.

A

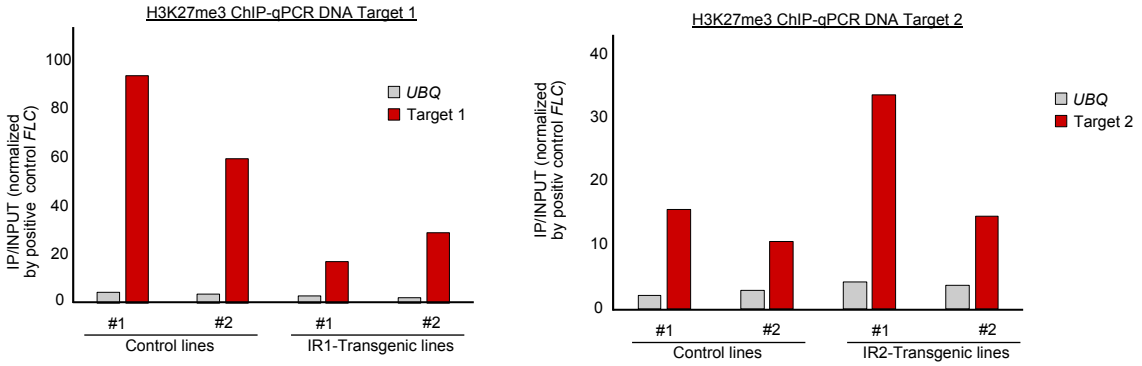

Figure S5. Targeted DNA methylation by inverted-repeats (IR) reveals a dual role of DNA methylation on H3K27me3 patterning at endogenous TEs

(A) ChIP-qPCR analysis of H3K27me3 marks in control and IR-transgenic lines at Target 1 (IR1) (**left panel**) or Target 2 (IR2) (**right panel**). A biological replicate of the experiment Fig.5E is shown. Data were normalized by the H3K27me3 ChIP positive control *FLC*; Ubiquitin (*UBQ)* serves as a negative control. Data were not averaged and biological replicates are shown separately because of the differences in ChIP efficiencies between the two batches of experiments.