Autocatalytic differentiation of epigenetic modifications within the Arabidopsis genome

Soichi Inagaki^{1,5}, Asuka Miura-Kamio¹,
Yasukazu Nakamura², Falong Lu³, Xia Cui³,
Xiaofeng Cao^{3,*}, Hiroshi Kimura⁴,
Hidetoshi Saze¹ and Tetsuji Kakutani^{1,*}

¹Department of Integrated Genetics, National Institute of Genetics, Mishima, Shizuoka, Japan, ²Center for Information Biology, National Institute of Genetics, Mishima, Shizuoka, Japan, ³State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, PR China and ⁴Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka, Japan

In diverse eukaryotes, constitutively silent sequences, such as transposons and repeats, are marked by methylation at histone H3 lysine 9 (H3K9me). Although selective H3K9me is critical for maintaining genome integrity, mechanisms to exclude H3K9me from active genes remain largely unexplored. Here, we show in Arabidopsis that the exclusion depends on a histone demethylase gene, *IBM1* (increase in *BONSAI* methylation). Loss-of-function *ibm1* mutation results in ectopic H3K9me and non-CG methylation in thousands of genes. The *ibm1*-induced genic H3K9me depends on both histone methylase KYP/SUVH4 and DNA methylase CMT3, suggesting interdependence of two epigenetic marks—H3K9me and non-CG methylation. Notably, IBM1 enhances loss of H3K9me in transcriptionally de-repressed sequences. Furthermore, disruption of transcription in genes induces ectopic non-CG methylation, which mimics the loss of IBM1 function. We propose that active chromatin is stabilized by an autocatalytic loop of transcription and H3K9 demethylation. This process counteracts a similarly autocatalytic accumulation of silent epigenetic marks, H3K9me and non-CG methylation. *The EMBO Journal* (2010) 29, 3496–3506. doi:10.1038/emboj.2010.227; Published online 10 September 2010
Subject Categories: chromatin & transcription; plant biology
Keywords: chromatin; DNA methylation; histone demethylase; IBM1; transposon

Introduction

Chromatin modifications, such as methylation of histone H3 at lysine 9 (H3K9me) and DNA methylation at cytosine

residues (Bender, 2004; Chan *et al*, 2005; Grewal and Elgin, 2007), are important for constitutively silencing transposons (Yoder *et al*, 1997; Walsh *et al*, 1998; Miura *et al*, 2001; Lippman *et al*, 2004; Ding *et al*, 2007; Mirouze *et al*, 2009; Tsukahara *et al*, 2009). On the other hand, these modifications are normally excluded from active genes to ensure their appropriate expression. Mechanisms for the exclusion, however, remain largely unexplored.

In Arabidopsis, methylation in genes is found only at CG sites, although transposons are methylated at all contexts of cytosines, CG, CHG, and CHH (H can be A, C, or T) (Zhang *et al*, 2006; Zilberman *et al*, 2007; Cokus *et al*, 2008; Lister *et al*, 2008). The CHG methylation in transposons is associated with dimethylation of H3K9 (H3K9me2) (Bernatavichute *et al*, 2008). H3K9me2 is catalysed by multiple histone methyltransferases, KRYPTONITE/SUVH4 (hereafter KYP), SUVH5, and SUVH6 (Jackson *et al*, 2002; Malagnac *et al*, 2002; Ebbs *et al*, 2005; Ebbs and Bender, 2006). These histone methyltransferases act upstream of DNA methyltransferase CHROMOMETHYLASE3 (CMT3), which catalyses DNA methylation at CHG sites (Bartee *et al*, 2001; Lindroth *et al*, 2001; Chan *et al*, 2005).

In addition to the positive regulators of DNA methylation, negative regulators have been recently identified (Agius *et al*, 2006; Gehring *et al*, 2006; Saze *et al*, 2008). One of the negative regulators of DNA methylation is IBM1 (INCREASE IN *BONSAI* METHYLATION 1); loss-of-function mutation in the *IBM1* gene induces ectopic DNA methylation at non-CG sites (mainly at CHG sites) in thousands of genes, whereas transposons are unaffected (Saze *et al*, 2008; Miura *et al*, 2009). Interestingly, the *ibm1*-induced genic DNA methylation occurs across each individual transcriptional unit (Miura *et al*, 2009). It is not known how a transcriptional unit can be recognized by the machinery that excludes the transposon-specific modification. The IBM1 protein has a jumoni-C (jmc) domain, and it is a member of the JHDM2/KDM3 family, which contains H3K9 demethylases (Klose *et al*, 2006; Lu *et al*, 2008; Sun and Zhou, 2008; Liu *et al*, 2010).

To understand the mechanisms that establish differential epigenetic modifications between active genes and silent transposons, we examined genome-wide pattern of H3K9me2 using mutants for IBM1, KYP, and CMT3. The *ibm1* mutation induced accumulation of H3K9me2 in thousands of genes, which parallels the increases in DNA methylation. Although IBM1 generally affected H3K9me2 of genes, transposons could also be affected, once they were activated by *kyp* or *cmt3* mutation, suggesting that transcription is an important prerequisite for the IBM1 function. Furthermore, blockage of transcription by T-DNA insertion in the target genes of IBM1 caused ectopic CHG methylation in the body of these genes, which mimics the effect of the loss of IBM1 function. These results lead us to propose that transcription is critical for IBM1 to remove H3K9me2, which mediates stabilization of active chromatin by the autocatalytic loop.

*Corresponding authors. X Cao, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China, Tel.: +86 10 64869203; Fax.: +86 10 64873428; E-mail: xcao@genetics.ac.cn or T Kakutani, Department of Integrated Genetics, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan. Tel.: +81 55 981 6801; Fax: +81 55 981 6804; E-mail: tkakutan@lab.nig.ac.jp

⁵Present address: Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan

Received: 6 April 2010; accepted: 13 August 2010; published online: 10 September 2010

Results

IBM1 is necessary for excluding H3K9me2 from genes

To understand how genomic H3K9me2 patterns are generated, we examined the function of the putative H3K9 demethylase, IBM1, as well as the H3K9 methylase KYP and the CHG methylase CMT3. First, we examined demethylase activity of IBM1 *in vivo* by overexpressing FLAG-tagged IBM1 in *Nicotiana benthamiana* leaf. Double immunostaining of nuclei with antibodies against FLAG and methylated histone at different residues shows the histone methylation status at each residue in FLAG-IBM1-overexpressing cells. In the FLAG-stained cells, H3K9me2 was hardly detectable and the H3K9me1 signal was also much reduced compared with that of the cells without FLAG labelling (Figure 1A; Supplementary Figure S1A). The effects on H3K9me2 and H3K9me1 levels were abolished when two amino-acid residues are substituted (H679A and D681A; referred to as FLAG-IBM1m) in the conserved Fe(II) binding site of the IBM1 protein (Figure 1B; Supplementary Figure S1B).

H3K9me3 and the three types of methylation at each of other sites, H3K4, H3K27, and H3K36, were indistinguishable between cells with or without the FLAG-IBM1 signal (Supplementary Figure S1C–F). Among the 12 modifications examined, H3K9me2 and H3K9me1 were the only modifications affected by the FLAG-IBM1. These results support the prediction that IBM1 is a histone demethylase specific for H3K9me2 and H3K9me1.

We then examined H3K9me2 genome-wide in wild-type Columbia-0 (Col-0) and *ibm1* mutant samples by chromatin immunoprecipitation (ChIP) coupled with DNA microarray analysis (ChIP-chip). In order to minimize cross-reacting signals from other modifications, we used a monoclonal antibody specific for H3K9me2 (Kimura *et al*, 2008). In wild type, H3K9me2 level was high in transposons and repeats, but it was very low in annotated expressed genes (Figure 1C and D; Supplementary Figure S2). The *ibm1* mutation caused ectopic H3K9me2 accumulation in transcribed regions of a large number of genes (Figure 1C–E). Essentially the same set of genes was affected in two biological replicates (Supplementary Figures S3). On the other hand, transposons were unaffected by the *ibm1* mutation (Figure 1D and E). The *ibm1*-induced genic H3K9me2 is accompanied by DNA methylation changes (Figure 1F).

As is the case in the *ibm1*-induced DNA methylation in genes, the *ibm1*-induced H3K9me2 showed specific characteristics in relation to the transcription units (Miura *et al*, 2009); long genes and constitutively transcribed genes were affected severely (Supplementary Figure S4A–E), and central region of a transcription unit showed more signal than the 5' and 3' regions (Supplementary Figure S5A–C).

The kyp and cmt3 mutations partially abolish H3K9me2 in transposons

In addition to the negative regulator of H3K9me2, we examined the contribution of positive regulators, KYP and CMT3. Compared with wild-type, *kyp* mutant samples showed reduced H3K9me2 levels in many transposons both in pericentromeric regions and arm regions, although significant levels of H3K9me2 still remained (Figure 2A–C; Supplementary Figure S5D). Other methyltransferase(s) such as SUVH5 and SUVH6 may contribute to the remaining H3K9me2 in

transposons and repeats, as has been suggested previously (Ebbs and Bender, 2006). Consistent with the proposed functional redundancy, the H3K9me2 levels were reduced in the triple mutant for KYP, SUVH5, and SUVH6 (*suvh456*) compared with the single mutant *kyp* (Figure 2D).

The *cmt3* mutant also showed some reduction of H3K9me2 in heterochromatic regions, although the effect was generally less than that seen in *kyp* samples (Figure 2E–H; Supplementary Figure S5D). The dependence of H3K9me2 on CMT3 function suggests a function for DNA methylation in maintaining H3K9me2 in transposons. We then examined the relationships between the levels of DNA methylation at different contexts and the H3K9me2 levels in wild-type, *kyp*, and *cmt3* samples. Consistent with previous report (Bernatavichute *et al*, 2008), the frequency of methylated CHG per length ([mCHG]) was positively correlated with H3K9me2 level in transposons (Supplementary Figure S6A). The *kyp* and *cmt3* mutant samples showed a greater decrease in H3K9me2 at transposons with higher [mCHG] (Supplementary Figure S6C, E, G, and H), suggesting that CHG methylation is indeed required for a high level of H3K9me2 at transposons. Interestingly, the decreases of H3K9me2 in *kyp* and *cmt3* mutants were more pronounced at transposons with lower [mCG] than those with higher [mCG] (Supplementary Figure S6B, D, F, I, and J). The results suggest that an additional pathway, which depends on CG methylation, may also be involved in the H3K9me2 of transposons (Soppe *et al*, 2002; Tariq *et al*, 2003).

The kyp and cmt3 mutations fully suppress the ibm1-induced H3K9me2

Although the *kyp* mutation induced only a partial reduction of H3K9me2 in transposons (Figure 2A–C), this mutation suppressed the developmental phenotypes induced by the *ibm1* mutation (Saze *et al*, 2008), which are probably due to ectopic H3K9me2 and/or CHG methylation in genic regions. We therefore examined the effects of the *ibm1* mutation in the *kyp* mutant background. Notably, in the *kyp* mutant background, *ibm1*-induced genic H3K9me2 was lost almost completely (Figure 3A and B; Supplementary Figure S7). This result is in striking contrast to the effect of the *kyp* mutation on H3K9me2 in transposons, which is only partial (Figure 2A–C). Taken together, the results suggest that different H3K9 methyltransferases function in genes and transposons: only KYP in genes and KYP and others in transposons.

Surprisingly, the *cmt3* mutation also suppressed the *ibm1*-induced genic H3K9me2 almost completely (Figure 3C; Supplementary Figure S7). This result suggests that CMT3-mediated CHG methylation is required for the increase of H3K9me2 at genes in the *ibm1* mutant. On the other hand, the impact of the *cmt3* mutation on H3K9me2 in transposons was only subtle (Figure 2E–G).

To test whether CHG methylation, rather than an unidentified function of CMT3, mediates its effect on H3K9me2 induction in *ibm1*, we calculated the proportion of CHG sites (total number of CHG sites in transcribed region per length of annotation unit; hereafter [CHG]) for all annotated transcription units and examined the relationship of the [CHG] with the effect of the *ibm1* mutation on DNA methylation and H3K9me2. [CHG] strongly correlated with the degree of response to the *ibm1* mutation not only in DNA methylation, but also in H3K9me2 (Figure 4A and B). Together with

Figure 1 IBM1 downregulates H3K9me2 at genes. (A, B) Overexpression of FLAG-IBM1 in tobacco leaf cells induced reduction in H3K9me2 and H3K9me1. Nuclei transfected with FLAG-IBM1 (A) or a mutated construct, FLAG-IBM1m (B), were mixed with control nuclei without the transfection (see Materials and methods for details). The transfected nuclei and control nuclei could be distinguished by immunostaining with anti-FLAG (shown in red). Histone methylation was visualized by immunostaining with rabbit polyclonal methylation-specific antibodies (green). Nuclei were visualized by DAPI staining (blue). Arrows indicate nuclei transfected with FLAG-IBM1 or FLAG-IBM1m. Quantification of these signals, together with those from other samples, is shown in Supplementary Figure S1A and B. (C, D) Representative examples of H3K9me2 analysed by tiling array for a gene-rich region (C, position 20 829 000–20 934 000 on chromosome 1) and a transposon-rich region (D, positions 16 494 000–16 592 000 on chromosome 1). Profile of DNA methylation (Miura *et al*, 2009) is also shown. Blue and red boxes indicate genes and transposons, respectively. They are oriented 5′–3′ on the top and 3′–5′ on the bottom. Each vertical bar represents the log₂ ratio of the immunoprecipitated DNA to input control for the DNA methylation (DNAm) and H3K9me2, respectively. Each brown bar represents difference of signals for each comparison. (E) H3K9me2 level in each gene (blue dot) and transposon (red dot), compared between wild type and *ibm1*. (F) Comparison between the effect of *ibm1* mutation on DNA methylation (DNAm) and H3K9me2 in each gene (blue dot) and transposon (red dot). The changes in DNA methylation and H3K9me2 levels were not proportional, most likely because each gene differs in the CG methylation level, which is not affected by the *ibm1* mutation (Miura *et al*, 2009).

the result of *ibm1 cmt3* double mutant, these results suggest that CMT3-mediated CHG methylation is required for the induction of genic H3K9me2 in *ibm1*.

Unexpectedly, we also found that [CG] of genes negatively correlated with the induction of H3K9me2 and DNA methylation in the *ibm1* mutant (Figure 4C and D). Despite the

Figure 2 Effects of *kyp* and *cmt3* mutations on H3K9me2 level. (A, B) Global view of H3K9me2 patterns along the chromosome 3, compared between wild type and *kyp*. Average signals in each 100 kb regions (A) or the differences of the signals between wild type and *kyp* (B) were plotted. (C) H3K9me2 levels in genes (blue dots) and transposons (red dots), compared between wild type and *kyp*. (D) H3K9me2 levels in wild type (Col), wild-type sibling of *suvh456* (*SUVH456*), *kyp*, and *suvh456* triple mutant examined with ChIP-PCR. Amplifications from input control and ChIP DNA are shown. (E, F) Global view of H3K9me2 patterns in wild type and *cmt3* (E) or the difference in H3K9me2 between wild type and *cmt3* (F). (G, H) H3K9me2 levels in genes (blue dots) and transposons (red dots), compared between wild type and *cmt3* (G) or *cmt3* and *kyp* (H).

opposite effects, [CHG] and [CG] show a weak positive correlation (Figure 4E), suggesting that [CHG] and [CG] independently affect H3K9me2 and DNA methylation. Genes with higher [CHG] and lower [CG] were most strongly affected by the *ibm1* mutation (Figure 4F).

IBM1 affects transposons de-repressed in *kyp* or *cmt3* mutant backgrounds

In *kyp* or *cmt3* mutant backgrounds, the *ibm1* mutation did not affect H3K9me2 in genes (Figure 3A–C). However, a

significant number of transposons, which were not affected by the *ibm1* single mutation (Figure 3D), responded to the *ibm1* mutation in those backgrounds (Figure 3E and F). These transposons showed a decrease of H3K9me2 in *kyp* and *cmt3* single mutants that was recovered in the *ibm1 kyp* and *ibm1 cmt3* double mutants (Figure 5A–C). Not only H3K9me2, but also CHG methylation was recovered in the *ibm1 kyp* double mutant for these transposons (Figure 5D). These results suggest that a subset of transposons can be targets of IBM1 in a *kyp* or *cmt3* background.

Figure 3 The *ibm1*-induced genic H3K9me2 depends on KYP and CMT3. (A) and (D) are from the data shown in Figure 1E. Effect of the *ibm1* mutation on H3K9me2 levels was also examined in the background of *kyp* (B, E) or *cmt3* (C, F) mutant. Genes (A–C) and transposons (D–F) are separately shown.

Figure 4 [CHG] and [CG] correlate with the effect of *ibm1* on H3K9me2. (A, B) Relationships between [CHG] and the *ibm1*-induced changes in H3K9me2 (A) or DNA methylation (B) level in each gene. (C, D) Relationships between [CG] and the *ibm1*-induced changes in H3K9me2 (C) or DNA methylation (D) level in each gene. The effects of [CHG] and [CG] are independent of the effect of gene length (Supplementary Figure S8A–D). (E) Positive correlation between [CHG] and [CG]. (F) The proportion of genes with increased H3K9me2 in *ibm1* (Supplementary Figure S3A), within four gene groups with high and low [CHG] and [CG]. Actual gene numbers (*ibm1* affected/all) in each group are also indicated.

Figure 5 De-repressed transposons can be targets of IBM1. **(A)** Scatter plot showing the relationship between changes in H3K9me2 from wild type to *kyp* (x axis) and from *kyp* to *ibm1 kyp* (y axis). **(B)** Scatter plot showing the relationship between changes in H3K9me2 from wild type to *cmt3* (x axis) and from *cmt3* to *ibm1 cmt3* (y axis). Genes are shown in blue and transposons in red. **(C)** ChIP-chip result of four transposons showing decrease in H3K9me2 in *kyp* and *cmt3*, and restoration in *ibm1 kyp* and *ibm1 cmt3* double mutants. **(D)** DNA methylation at CHG sites analysed by digestion with methylation-sensitive restriction enzymes *MspI* and *BglII* and subsequent PCR (details in Materials and methods). Homozygous plants segregating in self-pollinated progeny of a double heterozygote (*IBM1/ibm1 KYP/kyp*) were used for analyses. Compared with the *kyp IBM1* plants with the hypomethylated transposons, the *kyp ibm1* plants showed increased signal reflecting robust methylation. The sibling wild-type *KYP* plants showed high-level methylation irrespective of the genotypes of the *IBM1* gene. **(E)** RT-PCR analysis for the expression of the four transposons in the wild type, *ibm1*, *kyp*, and *kyp ibm1* double mutant. Total RNA isolated from individual plants was used for each RT-PCR analysis. *GAPC* gene was used to normalize RNA template amounts. Negative control (no RT) lacked reverse transcriptase. We also examined expression pattern of transposons in *kyp* and *cmt3* mutants genome wide (Supplementary Figure S9). Other transposons de-repressed by *kyp* or *cmt3* mutation also showed increased H3K9me2 in the *ibm1* mutant.

One possible mechanism for this observation is that even transposons can be targets of IBM1 when they are transcribed. Indeed, these transposons were transcribed in the *kyp* or *cmt3* single mutant (Figure 5E; Supplementary Figure S9). Even though the transcriptional activation tended to be weak in the absence of IBM1 function (*ibm1 kyp* in Figure 5E), robust activation could be achieved when IBM1 functions (*IBM1 kyp* in Figure 5E). IBM1 apparently stabilizes the transcriptionally active state, most likely by active H3K9 demethylation.

T-DNA insertion induces non-CG methylation within IBM1 target genes

These observations suggest that transcription of the target gene is necessary for IBM1 function. In order to test this

prediction directly, we examined the effect of blocking transcription by T-DNA insertion. DNA methylation was examined in T-DNA insertion alleles for three of the IBM1 target genes (At4g11130, At1g63020, and At4g32700). In all of them, non-CG (mostly CHG) methylation was induced, a situation which is analogous to the effects of the *ibm1* mutation (Figure 6). This observation also supports the prediction that IBM1 function depends on transcription of the target gene.

In addition to the region downstream of the T-DNA insertion (Figure 6), the upstream regions sometimes, but not always, showed increased non-CG methylation (Supplementary Figure S10), suggesting that function of IBM1 might also be affected by the insertion downstream.

Figure 6 Bisulphite sequencing analysis of cytosine methylation at T-DNA-inserted genes. (A–C) Schematic representations of At4g11130 (A), At1g63020 (B), and At4g32700 (C) loci and cytosine methylation level analysed by bisulphite sequencing. The percentage of methylated cytosines is indicated by vertical bar for each site (red, CG; blue, CHG; green, CHH). T-DNA insertion mutants showed decreases of expressions (Supplementary Figure S10E), and increases of non-CG methylation, as in *ibm1* mutant, at downstream regions of T-DNA inserted genes. Mutants for the genes in same pathway (*RDR2* and *NRPD1a*) did not affect the DNA methylation at reciprocal gene. (D) Proportions of methylated cytosines at different contexts in regions shown in (A–C).

Discussion

Differentiation of active and inactive chromatin by transcription and histone demethylation

In the absence of IBM1 function, thousands of genes accumulated epigenetic marks of silent chromatin, such as

H3K9me2 and cytosine methylation at non-CG sites. The IBM1 protein is necessary for excluding these marks from transcribed genes. Although transposons were generally not affected by the *ibm1* mutation (Figure 1; Miura *et al*, 2009), de-repressed transposons behaved like genes in regard to their response to *ibm1* (Figure 5; Supplementary

Figure 7 Proposing model for differential epigenetic modifications between genes and transposons. (A) IBM1 demethylates H3K9 of transcribed genes, which in turn ensure robust transcription, generating the autocatalytic loop. In the absence of IBM1 function, H3K9me2 and mCHG accumulate even in transcribed genes by the self-reinforcing mechanism of KYP and CMT3 (grey arrows). (B) Even in wild type, the self-reinforcing loop of KYP and CMT3 operates in transposons. In transposons, additional backup of H3K9me2 machinery exists that depends on SUVH5 and/or SUVH6. When de-repressed, the autocatalytic loop of transcription and H3K9 demethylation runs even in transposons (grey arrows). The genes and transposons differ in that (i) IBM1 functions only in transcribed genes and that (ii) the KYP alone can function in genes for the methylation of H3K9, but additional SUVH proteins operate in transposons.

Figure S9). Conversely, loss of transcription by T-DNA insertion in the target gene of IBM1 caused induction of non-CG methylation (Figure 6). These results lead us to propose that IBM1 demethylates H3K9me2 from transcribed genes. This would stabilize the expressed state of active genes by preventing accumulation of H3K9me2 (Figure 7A).

How could transcription affect IBM1 function? The IBM1 protein may be recruited to transcriptional machinery, as H3K4 and H3K36 methyltransferases are (Li *et al*, 2002; Krogan *et al*, 2003; Ng *et al*, 2003). If so, this recruitment might account for the long-standing enigmatic observations in the mutant of CG methylase gene *MET1*. The *met1* mutation induces not only global loss of CG methylation and de-repression of transposons, but also ectopic methylation at non-CG sites in a subset of genes (Jacobsen and Meyerowitz, 1997; Cokus *et al*, 2008; Lister *et al*, 2008). This might be due to depletion of the IBM1 protein in genes, which is caused by de-repression of transposons and recruitment of IBM1 protein there. Consistent with this interpretation, genes showing ectopic CHG methylation in *ibm1* and *met1* largely overlap (Miura *et al*, 2009).

Interplay between DNA methylation and H3K9me2

The *ibm1*-induced genic H3K9me2 required CMT3, the DNA methylase for CHG sites (Figure 3C). How does CHG methylation affect H3K9me2? One proposed mechanism posits the direct recognition of DNA methylation by the histone methyltransferase; SRA domains within KYP and other SUVH proteins were recently shown to have methylcytosine-binding activity (Unoki *et al*, 2004; Johnson *et al*, 2007; Woo *et al*, 2007). Interestingly, the SRA domain of KYP binds preferentially to methylated CHG and CHH than methylated CG *in vitro* (Johnson *et al*, 2007). If KYP recognizes only CHG and CHH methylation *in vivo*, it could explain why the *cmt3* mutation alone fully suppresses the *ibm1*-induced H3K9me2 despite the presence of CG methylation. All observations are consistent with the model that genic H3K9me2 and CHG methylation are established by the self-reinforcing mechanism (Figure 7).

Unlike the *ibm1*-induced genic H3K9me2, the H3K9me2 in transposons was not much affected by the *cmt3* mutation (Figure 2E–G). Cytosine methylation at context(s) other than

CHG may also function upstream of H3K9me2. Interestingly, *cmt3* mutation affects H3K9me2 at transposons with low CG methylation (Supplementary Figure S6G–J). This observation is consistent with the interpretation that the remaining H3K9me2 in transposons depends on CG methylation. In fact, mutation in the CG methylase gene *MET1* induces a reduction of H3K9me in transposon-rich regions (Soppe *et al*, 2002; Tariq *et al*, 2003). Furthermore, siRNA-guided DNA methylation also occurs in transposon sequences (Matzke *et al*, 2009), which can be excluded from genes by DNA demethylases (Penterman *et al*, 2007; Zhu *et al*, 2007). Taken together, H3K9me of transposons seems to be targeted by multiple pathways, resulting in robust maintenance of heterochromatin marks.

Effect of [CHG] on the *ibm1*-induced H3K9me2 could be understood by the contribution of methylated CHG; a puzzling observation is that [CG] had a strong reverse effect (Figure 4C and D). This could be explained, at least in part, by underrepresentation of CG sites by high mutation rate of methylated cytosine (Tran *et al*, 2005). In fact, [CG] level was lower than that expected from [C] and [G] in genes with high CG methylation rate (Supplementary Figure S8E), and those methylated genes tend to respond more to the *ibm1* mutation (Supplementary Figure S8F). In addition to this indirect effect, it would be interesting if unmethylated CG sites directly inhibit H3K9me2 or CHG methylation in genes (Supplementary Figure S8G–I). If such interaction exists, then that would also affect the differential methylation pattern within the genome.

Control of epigenetic marks in the body of transcribed gene

We are proposing that IBM1 negatively regulates accumulation of H3K9me2 in transcribed genes. On the other hand, in the absence of the IBM1 function, constitutively transcribed genes tend to have more H3K9me2 and CHG methylation than genes expressed in a tissue-specific manner or expressed at very low level (Supplementary Figure S4B–E; Miura *et al*, 2009). Exactly the same trend is known for genic methylation at CG sites in wild-type background (Zhang *et al*, 2006; Zilberman *et al*, 2007). The strikingly similar spectra suggest that genic CG methylation might be mechanically connected

to the CHG methylation, detected in the absence of IBM function. However, CHG methylation in the *ibm1* mutant can be found even in genes lacking CG methylation (Miura *et al*, 2009). In addition, mutation in the CG methylase gene *MET1* did not suppress the *ibm1*-induced CHG hypermethylation (SI and TK, unpublished), suggesting that CG methylation is not upstream of the genic CHG methylation. On the other hand, it is tempting to speculate the reverse causative relationship; the component(s) of the self-reinforcing loop of H3K9me2 and CHG methylation may trigger the genic CG methylation. Consistent with that model, the genic methylation level at CG sites positively correlated with [CHG] (proportion of CHG sites, irrespective of methylated or not; data not shown). Once induced, CG methylation can be epigenetically inherited over generations by the maintenance methylation (Saze *et al*, 2003; Takeda and Paszkowski, 2006).

As in Arabidopsis, recent genome-wide analyses of DNA methylation in the human genome revealed that the methylation is found in body of genes with high transcription activity (Lister *et al*, 2009). Furthermore, in human ES cells and iPS cells, the hypermethylation of transcribed genes was also found at non-CG sites, which disappears during differentiation (Lister *et al*, 2009). The correlation with the transcription is much stronger for methylation at non-CG sites than at CG sites (Lister *et al*, 2009). Despite its prevalence in diverse eukaryotes, the underlying mechanisms remain enigmatic for the body methylation of constitutively transcribed genes (Zemach *et al*, 2010). Arabidopsis would further serve as a model organism to understand mechanisms for the interactions between DNA methylation (CG and non-CG sites) and histone modifications/replacements, which can also be affected by transcription (Zilberman *et al*, 2008; Bell and Schübeler, 2009).

Materials and methods

Plant materials

Arabidopsis thaliana strain Col-0 was used as 'wild type'. The isolation of *ibm1* mutants was described previously (Saze *et al*, 2008). We used a presumed null allele *ibm1-4*. *cmt3i1* mutant was originally isolated in Wassilewskija background (Bartee *et al*, 2001), and we crossed that to Col-0 six times before use. The *kyp* mutant (SAIL588F05) is in Col-0 background and has T-DNA insertion in SET domain-coding region; *suvh456* triple mutant was made from *kyp-4* (SALK044606), *suvh5-2* (Ebbs and Bender, 2006), and *suvh6-1* (Ebbs *et al*, 2005), which are all in Col-0 background. T-DNA insertion mutants for At4g11130, At1g63020, and At4g32700 are SALK059661, SALK128428, and SALK035610, respectively. The SALK lines (Alonso *et al*, 2003) were obtained from Arabidopsis Biological Resource Center.

In vivo histone demethylation assay

In vivo histone demethylation activity of IBM1 protein was examined by overexpression and subsequent immunostaining (Seward *et al*, 2007; Lu *et al*, 2010). Full-length IBM1 genomic-coding sequence with stop codon was cloned into pENTR/D-TOPO vector (Invitrogen) using cx3583: 5'-CACCATGGATTCTGTGGAG-GAAGAAG-3' and cx3585: 5'-CTACATCTTCTCCATTCTAAT-3' and then transferred into pEarleyGate202 (pEG202) vector with an N-terminal FLAG tag using LR reaction (Earley *et al*, 2006). IBM1m has two amino-acid substitutions from IBM1, H679A, and D681A. The mutation construct was performed by Quickchange site-directed mutagenesis kit (Stratagene) using the following primers: cx8060: 5'-AGATTCTGTTACTAAGCTCGCCTGCGCCATGTCTGATCGGTTGAAT-3' and cx8061: 5'-ATTCACCGCATCAGACATGGCGCAGCGGAGCTTAGTAACAGAATCT-3'. The constructs were transformed into *Agrobacterium tumefaciens* cells (strain EHA105). Cells were then injected into half of *N. benthamiana* leaves as described

(English *et al*, 1997). FLAG-tagged IBM1 was overexpressed in half of a 4-week-old tobacco leaf, while using the other half without infiltration as a control. After 72 h post-injection, the whole leaves were harvested for nuclei isolation and immunostaining. Nuclei isolation and immunostaining were carried out as previously described (Houben *et al*, 2003; Johnson *et al*, 2002). FLAG M2 antibody from mouse (Sigma) and Histone methylation-specific antibody from rabbit (H3K4me3: Abcam ab8580; H3K4me2: Millipore 07-030; H3K4me1: Millipore 07-436; H3K9me3: Millipore 07-442; H3K9me2: Millipore 07-441; H3K9me1: Millipore 07-450; H3K27me3: gift from Dr Thomas Jenuwein; H3K27me2: Millipore 07-452; H3K27me1: Millipore 07-448; H3K36me3: Abcam ab9050; H3K36me2: Millipore 07-274; H3K36me1: Millipore 07-548) were used for immunoassaying followed by Alexa flour 555-conjugated goat anti-mouse (Invitrogen) and Alexa flour 488-conjugated goat anti-rabbit (Invitrogen). After staining, the slides were mounted by VECTASHIELD mounting medium with DAPI (Vector Laboratory).

ChIP-chip

Leaves of 4-week-old plants were fixed as described previously (Saze *et al*, 2008). ChIP was performed as described previously (Kimura *et al*, 2008), using antibody against H3K9me2 (CMA307; Kimura *et al*, 2008). Non-immunoprecipitated DNA (input DNA) and ChIP samples were amplified, labelled, and hybridized to microarray according to the manufacturer's instruction (protocols for ChIP and amplification, NimbleGen). Input DNA and ChIP DNA were differentially labelled with Cy3 and Cy5, respectively, and competitively hybridized to a microarray chip. We used NimbleGen 2.1M HD2 array covering entire genome of *A. thaliana*, and 385K array covering entire chromosome 1 and part of chromosome 2. We performed two biological replicate for the H3K9me2 ChIP-chip (two HD2 arrays for wild-type Col-0 and *ibm1*; one HD2 and one 385K arrays for *kyp*, *ibm1 kyp*, *cmt3*, and *ibm1 cmt3*). Each feature on the array has a corresponding scaled log₂ ratio, which is the ratio of the signals for the ChIP and input samples cohybridized to the array. The log₂ ratio is computed and scaled to centre the ratio data around zero. Scaling is performed by subtracting the bi-weight mean for the log₂-ratio values for all features on the array from each log₂-ratio value. Cytosine methylation profilings shown in Figure 1B and C are on 3 × 385K arrays (Miura *et al*, 2009), which show minor difference in the probe sets to the HD2 array. The ChIP-chip data, together with the expression results of *kyp* and *cmt3* mutants, are available as SuperSeries GSE23030 in NCBI (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE23030>).

Analysis of base composition and DNA methylation

Base compositions, such as [C], [G], [CG], and [CHG], were calculated for each transcription unit. Occurrence of each site was counted for both strands and the number was divided by twice of the length (base) of that region. Frequency of methylated sites, [mCG] and [mCHG], was calculated in a similar manner using the wild-type Col results by Lister *et al* (2008).

DNA methylation analysis

DNA methylation was analysed using methylation-sensitive restriction enzyme *MspI* and *BglII*. A 50 ng of genomic DNA was digested with *MspI* and *PstI*, or *BglII* and *EcoRI*, in 20 µl reaction mix. After digestion, PCR was performed using 1 µl of the digested sample as a template. Primers used are shown in Supplementary Table S1.

Genomic bisulphite sequencing was performed using Bisulfast DNA Modification kit (TOYOBO) from 500 ng of genomic DNA. A one-sixth portion of the bisulphite-treated DNA was used as a template for PCR (50 µl reaction). Purified amplified fragment was cloned into pGEM-T easy vector (Promega) and sequenced.

RT-PCR

Total RNA from leaves was isolated using the RNeasy Plant Mini kit (Qiagen) and was treated with DNase I (TAKARA). cDNA was synthesized using the SuperScript III Reverse Transcriptase (Invitrogen) and random hexamer primer. Primers used in RT-PCR are shown in Supplementary Table S1.

Expression microarray

Total RNA from leaves was isolated and treated with DNase I as described above. Double-stranded cDNA was synthesized using the SuperScript Double-Stranded cDNA Synthesis kit (Invitrogen) and oligo-(dT)²⁰ primer. cDNA was labelled with Cy3, and hybridized to

NimbleGen 4 × 72k microarray according to the manufacturer's instruction. Data were normalized by RMA method using Nimble Scan software (NimbleGen).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank Akiko Terui for technical assistance, Judith Bender for *Arabidopsis* seeds, Ryan Lister and Joe Ecker for the cytosine methylation information in the *Arabidopsis* genome, Thomas

Jenuwein for antibody, and Yasushi Hiromi and Eric Richards for critical comments on the paper. This work was supported by grants from Mitsubishi Foundation, Takeda Science Foundation, Japanese Ministry of Education, Culture, Sports, Science and Technology (19207002 and 19060014) (to TK), National Basic Research Program of China (2009CB941500 and 2005CB522400), and National Natural Science Foundation of China (30930048) (to XC).

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Agius F, Kapoor A, Zhu JK (2006) Role of the *Arabidopsis* DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc Natl Acad Sci USA* **103**: 11796–11801
- Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H *et al* (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**: 653–657
- Bartee L, Malagnac F, Bender J (2001) *Arabidopsis cmt3* chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. *Genes Dev* **15**: 1753–1758
- Bell O, Schübeler D (2009) Chromatin: sub out the replacement. *Curr Biol* **19**: R545–R547
- Bender J (2004) DNA methylation and epigenetics. *Annu Rev Plant Biol* **55**: 41–68
- Bernatavichute YV, Zhang X, Cokus S, Pellegrini M, Jacobsen SE (2008) Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. *PLoS One* **3**: e3156
- Chan SW, Henderson IR, Jacobsen SE (2005) Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat Rev Genet* **6**: 351–360
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE (2008) Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452**: 215–219
- Ding Y, Wang X, Su L, Zhai J, Cao S, Zhang D, Liu C, Bi Y, Qian Q, Cheng Z, Chu C, Cao X (2007) SDG714, a histone H3K9 methyltransferase, is involved in Tos17 DNA methylation and transposition in rice. *Plant Cell* **19**: 9–22
- Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J* **45**: 616–629
- Ebbs ML, Bartee L, Bender J (2005) H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. *Mol Cell Biol* **25**: 10507–10515
- Ebbs ML, Bender J (2006) Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. *Plant Cell* **18**: 1166–1176
- English JJ, Davenport GF, Elmayan T, Vaucheret H, Baulcombe DC (1997) Requirement of sense transcription for homology-dependent virus resistance and *trans*-inactivation. *Plant J* **12**: 597–603
- Gehring M, Huh JH, Hsieh TF, Penterman J, Choi Y, Harada JJ, Goldberg RB, Fischer RL (2006) DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124**: 495–506
- Grewal SI, Elgin SC (2007) Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**: 399–406
- Houben A, Demidov D, Gernand D, Meister A, Leach CR, Schubert I (2003) Methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. *Plant J* **33**: 967–973
- Jackson JP, Lindroth AM, Cao X, Jacobsen SE (2002) Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**: 556–560
- Jacobsen SE, Meyerowitz EM (1997) Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. *Science* **277**: 1100–1103
- Johnson L, Cao X, Jacobsen S (2002) Interplay between two epigenetic marks: DNA methylation and histone H3 lysine 9 methylation. *Curr Biol* **12**: 1360–1367
- Johnson LM, Bostick M, Zhang X, Kraft E, Henderson I, Callis J, Jacobsen SE (2007) The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr Biol* **17**: 379–384
- Kimura H, Hayashi-Takanaka Y, Goto Y, Takizawa N, Nozaki N (2008) The organization of histone H3 modifications as revealed by a panel of specific monoclonal antibodies. *Cell Struct Funct* **33**: 61–73
- Klose RJ, Kallin EM, Zhang Y (2006) JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet* **7**: 715–727
- Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S, Greenblatt J (2003) Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* **23**: 4207–4218
- Li J, Moazed D, Gygi SP (2002) Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *J Biol Chem* **277**: 49383–49388
- Lindroth AM, Cao X, Jackson JP, Zilberman D, McCallum CM, Henikoff S, Jacobsen SE (2001) Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* **292**: 2077–2080
- Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, Carrington JC, Doerge RW, Colot V, Martienssen R (2004) Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**: 471–476
- Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **133**: 523–536
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**: 315–322
- Liu C, Lu F, Cui X, Cao X (2010) Histone methylation in higher plants. *Annu Rev Plant Biol* **61**: 395–420
- Lu F, Cui X, Zhang S, Liu C, Cao X (2010) JMJ14 is an H3K4 demethylase regulating flowering time in *Arabidopsis*. *Cell Res* **20**: 387–390
- Lu F, Li G, Cui X, Liu C, Wang XJ, Cao X (2008) Comparative analysis of JmjC domain-containing proteins reveals the potential histone demethylases in *Arabidopsis* and rice. *J Integr Plant Biol* **50**: 886–896
- Malagnac F, Bartee L, Bender J (2002) An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J* **21**: 6842–6852
- Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ (2009) RNA-mediated chromatin-based silencing in plants. *Curr Opin Cell Biol* **21**: 367–376
- Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, Cao J, Weigel D, Paszkowski J, Mathieu O (2009) Selective epigenetic control of retrotransposition in *Arabidopsis*. *Nature* **461**: 427–430
- Miura A, Nakamura M, Inagaki S, Kobayashi A, Saze H, Kakutani T (2009) An *Arabidopsis* jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. *EMBO J* **28**: 1078–1086

- Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T (2001) Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. *Nature* **411**: 212–214
- Ng HH, Robert F, Young RA, Struhl K (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11**: 709–719
- Penterman J, Uzawa R, Fischer RL (2007) Genetic interactions between DNA demethylation and methylation in *Arabidopsis*. *Plant Physiol* **145**: 1549–1557
- Saze H, Scheid OM, Paszkowski J (2003) Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nat Genet* **34**: 65–69
- Saze H, Shiraishi A, Miura A, Kakutani T (2008) Control of genic DNA methylation by a jmjC domain-containing protein in *Arabidopsis thaliana*. *Science* **319**: 462–465
- Seward DJ, Gabrielle C, Kim S, Schonewald M, Zhang L, Tripet B, Bentley D (2007) Demethylation of trimethylated histone H3 Lys4 *in vivo* by JARID1 JmjC proteins. *Nat Struct Mol Biol* **14**: 240–242
- Soppe WJ, Jasencakova Z, Houben A, Kakutani T, Meister A, Huang MS, Jacobsen SE, Schubert I, Fransz PF (2002) DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis*. *EMBO J* **21**: 6549–6559
- Sun Q, Zhou DX (2008) Rice jmjC domain-containing gene JM1706 encodes H3K9 demethylase required for floral organ development. *Proc Natl Acad Sci USA* **105**: 13679–13684
- Takeda S, Paszkowski J (2006) DNA methylation and epigenetic inheritance during plant gametogenesis. *Chromosoma* **115**: 27–35
- Tariq M, Saze H, Probst AV, Lichota J, Habu Y, Paszkowski J (2003) Erasure of CpG methylation in *Arabidopsis* alters patterns of histone H3 methylation in heterochromatin. *Proc Natl Acad Sci USA* **100**: 8823–8827
- Tran RK, Henikoff JG, Zilberman D, Ditt RF, Jacobsen SE, Henikoff S (2005) DNA methylation profiling identifies CG methylation clusters in *Arabidopsis* genes. *Curr Biol* **15**: 154–159
- Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T (2009) Bursts of retrotransposition reproduced in *Arabidopsis*. *Nature* **461**: 423–426
- Unoki M, Nishidate T, Nakamura Y (2004) ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. *Oncogene* **23**: 7601–7610
- Walsh CP, Chaillet JR, Bestor TH (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet* **20**: 116–117
- Woo HR, Pontes O, Pikaard CS, Richards EJ (2007) VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes Dev* **21**: 267–277
- Yoder JA, Walsh CP, Bestor TH (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* **13**: 335–340
- Zemach A, McDaniel IE, Silva P, Zilberman D (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**: 916–919
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* **126**: 1189–1201
- Zhu J, Kapoor A, Sridhar VV, Agius F, Zhu JK (2007) The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in *Arabidopsis*. *Curr Biol* **17**: 54–59
- Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S (2008) Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* **456**: 125–129
- Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S (2007) Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* **39**: 61–69