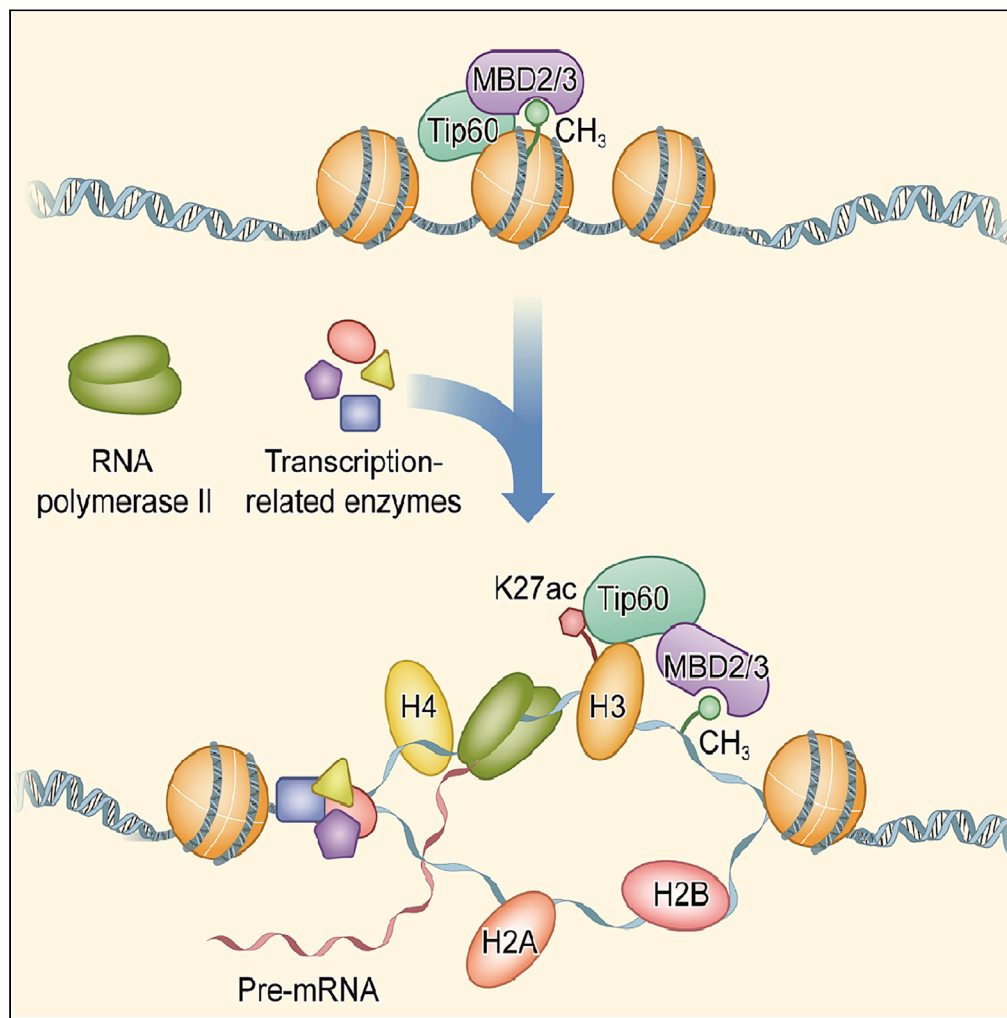# iScience

Article

Intragenic DNA methylation regulates insect gene expression and reproduction through the MBD/Tip60 complex

Guanfeng Xu, Hao Lyu, Yangqin Yi, ..., Xuezhen Peng, Subba Reddy Palli, Sichun Zheng

sczheng@scnu.edu.cn

HIGHLIGHTS

Insect intragenic 5mC enhances gene expression through histone H3K27 acetylation

MBD2/3 binds the intragenic 5mC and recruits Tip60 to promote H3K27 acetylation

Intragenic 5mCs modify protein synthesis-related genes in insect ovaries

The intragenic 5mC plays a role in insect reproduction

Article

Intragenic DNA methylation regulates insect gene expression and reproduction through the MBD/Tip60 complex

Guanfeng Xu,[1,2] Hao Lyu,[1,2] Yangqin Yi,[1,2] Yuling Peng,[1,2] Qili Feng,[1,2] Qisheng Song,[3] Chengcheng Gong,[1,2] Xuezhen Peng,[1,2] Subba Reddy Palli,[4] and Sichun Zheng[1,2,5,*]

SUMMARY

DNA methylation is an important epigenetic modification. However, the regulations and functions of insect intragenic DNA methylation remain unknown. Here, we demonstrate that a regulatory mechanism involving intragenic DNA methylation controls ovarian and embryonic developmental processes in *Bombyx mori*. In *B. mori*, DNA methylation is found near the transcription start site (TSS) of ovarian genes. By promoter activity analysis, we observed that 5′ UTR methylation enhances gene expression. Moreover, methyl-DNA-binding domain protein 2/3 (MBD2/3) binds to the intragenic methyl-CpG fragment and recruits acetyltransferase Tip60 to promote histone H3K27 acetylation and gene expression. Additionally, genome-wide analyses showed that the peak of H3K27 acetylation appears near the TSS of methyl-modified genes, and DNA methylation is enriched in genes involved in protein synthesis in the *B. mori* ovary, with MBD2/3 knockdown resulting in decreased fecundity. These data uncover a mechanism of gene body methylation for regulating insect gene expression and reproduction.

INTRODUCTION

Insects are the largest and most diverse group of organisms on Earth; these animals show fecundity, with large numbers of individuals and wide distributions (Mito et al., 2010). Metamorphosis is an insect-specific developmental process in which the organism undergoes changes in internal organs, external characteristics, and even living habitats from larval to adult stages. In addition to development, metamorphosis improves the adaptability of insects to the environment. Insects are classified into ametabolous, hemimetabolous, and holometabolous groups. During metamorphosis, sequential gene expression promotes ovarian development, and dysfunction in transcription causes dramatic defects in ovarian development and reproduction (Bownes, 1989; Swevers and Iatrou, 2003). For example, knockdown of DNA methyltransferase 1 (*Dnmt1*) in the hemimetabolous species *Oncopeltus fasciatus* resulted in abnormal ovarian development (Bewick et al., 2019). In addition, RNA interference (RNAi) of *Dnmt1* in holometabolous *Bombyx mori* led to a decrease in the egg hatch rate (Xiang et al., 2013). Therefore, it is important to understand how insect DNA methylation regulates gene expression during ovarian development.

Indeed, DNA methylation is an important epigenetic modification in eukaryotes. Dnmt3 and Dnmt1, which function in *de novo* methylation and maintenance of existing methylation, respectively, mediate modifications at the fifth carbon of the pyrimidine ring of cytosines in a gene promoter or body, with 5-methylcytosine (5mC) being the most commonly methylated nucleotide in genomic DNA (Goll and Bestor, 2005; Kim et al., 2009). In vertebrates, studies on DNA methylation regulation are often focused on the CpG islands of gene promoter regions (Bird, 2002; Lister et al., 2009), which inhibits gene expression because methylcytosine-binding proteins (MeCP or MBDs) bind to *cis*-regulatory elements (CREs), blocking the binding of transcription factor(s) (Bird and Wolffe, 1999). MBDs also recruit histone deacetylase (Dobosy and Selker, 2001; Jones et al., 1998; Nan et al., 1998) and histone methyltransferase (Fuks et al., 2003) to silence gene transcription by assembling a repressive nucleosomal array. With regard to gene body methylation in mammals, several reports have demonstrated that gene body 5mC enhances gene expression by promoting

[1]Guangdong Provincial Key Laboratory of Insect Developmental Biology and Applied Technology, Institute of Insect Science and Technology, School of Life Sciences, South China Normal University, Guangzhou 510631, China

[2]Guangzhou Key Laboratory of Insect Development Regulation and Applied Research, Institute of Insect Science and Technology, School of Life Sciences, South China Normal University, Guangzhou 510631, China

[3]Division of Plant Sciences, College of Agriculture, Food and Natural Resources, University of Missouri, Columbia, MO 65211, USA

[4]Department of Entomology, University of Kentucky, Lexington, KY 40546, USA

[5]Lead contact

*Correspondence: sczheng@scnu.edu.cn

https://doi.org/10.1016/j.isci.2021.102040

H3K36m3 (Yang et al., 2014), a mark of transcriptional elongation, or by modulating alternative splicing through MeCP (Maunakea et al., 2013).

In contrast, studies on DNA methylation in insects are not as widespread and in-depth as in mammals. Nonetheless, many insects possess both Dnmt1 and Dnmt3, similar to mammals. However, some insects, such as *B. mori* (silkworm), *Tribolium castaneum* (beetle), and *Schistocerca gregaria* (locust), only have Dnmt1, and some others, such as *Drosophila melanogaster* and *Anopheles gambiae* (mosquito), do not possess Dnmt1 or Dnmt3 (Zemach et al., 2010). Thus, it remains unclear how genes in *D. melanogaster* become methylated without Dnmt1 and Dnmt3 and whether Dnmt1 has dual functions of de novo methylation and maintenance of existing 5mC in *B. mori*. Regulation of gene expression via promoter region DNA methylation was experimentally demonstrated in *D. melanogaster Spok* (Zhang et al., 2018) and in the *B. mori* chitin synthase gene (Xu et al., 2018). However, 5mC DNA methylation in insects occurs mainly in gene bodies (Glastad et al., 2014; Hunt et al., 2013), with higher levels of transcription (Xiang et al., 2010). It is unknown whether intragenic methylation regulates gene expression by modulating alternative splicing, as in mammals. In addition, although *Dmnt1* knockdown in *O. fasciatus* affects ovarian development, it decreases mRNA levels for only 1% of methylated genes, indicating that 5mC might not directly regulate gene expression in this insect (Bewick et al., 2019). Therefore, whether 5mC in gene bodies regulates transcription and how it affects insect reproduction remain unknown.

In mammals, MBD family proteins consist of MeCP2 and MBD1-6 isoforms. MBDs bind to DNA via a highly conserved MBD domain; MBD3 cannot directly bind to DNA because it lacks the MBD domain. MBD2 and MBD3 repress gene transcription as a part of the nucleosome remodeling and histone deacetylation (NuRD) complex; both proteins directly compete with one another for binding to NuRD components through their intrinsically disordered regions and coiled-coil domains (Baubec et al., 2013; Le Guezennec et al., 2006). Sequence comparison in insects has revealed that the orthologs of mammalian MBD2 and MBD3 are the same protein; this MBD protein with MBD and coiled-coil domains is called MBD2/3 (Hendrich and Tweedie, 2003). Nevertheless, it is unknown whether insect MBD2/3 functions in regulating gene expression *in vivo*.

Here, we report that insect-specific MBD2/3 binds to 5mC in the gene body of ovarian development-related genes and enhances gene expression to ensure egg development and hatching. We found that in the holometabolous insect *B. mori*, enrichment of 5mC in ovarian gene bodies promotes transcription via an increase in H3K27ac mediated by the complex of MBD2/3 and acetyltransferase Tip60. We further demonstrate the role of MBD2/3 in the reproduction of holometabolous and hemimetabolous insects. Our study reveals a mechanism for the regulation of insect ovarian development through MBD2/3-mediated intragenic DNA methylation.

RESULTS

5mC is enriched in 5′ untranslated regions of gene bodies in the *B. mori* ovary

To investigate whether and how intragenic methylation regulates ovarian gene expression, 5mC in the ovarian genome was examined. The presence of 5mC in *B. mori* ovaries was confirmed by immunofluorescence staining with anti-5mC antibodies and by electrophoretic mobility shift assays (EMSAs) with an intragenic methylated DNA fragment from the *B. mori* ovary (Figure 1A). Analyses of spatial expression profiles of both mRNA and protein during pupal stages showed that *BmDnmt1*, the only DNA methyltransferase in *B. mori* that introduces 5mC (Figure S1), was highly expressed in the ovary compared with in the other tissues tested (Figure 1B). DNA methylation in the *B. mori* ovary was also analyzed by methylation sequencing. Higher levels of CG DNA methylation, versus at CHG and CHH, were detected. Only 0.52% (303,813 cytosines) of genomic cytosines were methylated, among which 83.61% of mCs were at CG sites and 13.47% and 2.93% were at CHG and CHH sites, respectively (Figure 1C and Table S1). In particular, more 5mC was detected near the transcription start site (TSS) of gene bodies when compared with other regions. Overall, the sequence with the highest frequency of methylation among methylated genes in the genome was TATCGAT (methyl-preferred fragment) (Figure 1D). These results suggest that 5mC mainly occurs at CG sites of 5′ untranslated regions of gene bodies in the *B. mori* ovary.

Gene body 5mC enhances gene expression

To investigate whether there is a relationship between gene 5mC levels and gene expression, 5mC and gene expression in the ovarian transcriptome were analyzed in the *B. mori* ovary. The results showed that
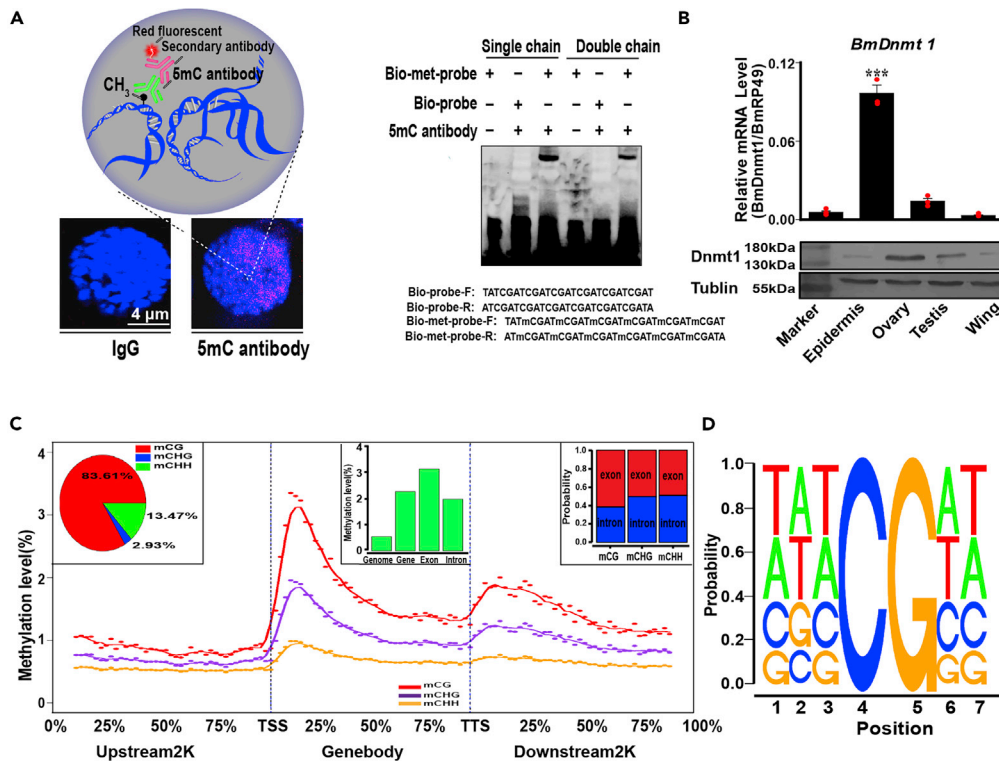
Figure 1. Analyses of DNA methylation in the *Bombyx mori* ovary

(A) 5mC DNA methylation was detected by immunofluorescent staining in the nuclei of *B. mori* ovaries using an anti-5mC antibody (scale bar, 4 μm) (left) (n = 9). An EMSA shows the specific binding of the anti-5mC antibody to mC probes (right).

(B) qRT-PCR (top) and western blot (bottom) analyses of the expression patterns of DNA methyltransferase 1 (*BmDnmt1*) in different tissues of silkworms during the pupal stage (n = 3).

(C) Analyses of whole-genome DNA methylation patterns and distribution in *B. mori* ovaries of 7-day-old pupae (n = 2). Fraction of mCs identified in each sequence context (CG, CHG, and CHH) based on methylated cytosine percentage in the whole genome (left). Methylation profile in the gene region. The gene region consists of 2 kb upstream and downstream of the TSS and TTS regions and the gene body from the TSS to TTS. TSS: transcription start site; TTS: transcription termination site (middle). Methylation level at different regions (exon and intron) in the gene body (right).

(D) Sequence context of CG methylation. CG bases were counted to obtain a base composition of 7 bp near the mCG base in the whole genome, and a Logo map was drawn using seqLogo software. For all measurements, "n" represents the number of biological replicates. Data are presented as the mean ± SEM. ***p < 0.001 by ANOVA.

most genes with methylation rates higher than 1% also had high expression levels (Figures 2A, left panel and Tables S2, S3, S4, and S5), similar to the previous results in insects (Suzuki and Bird, 2008; Zemach et al., 2010; Sarda et al., 2012; Glastad et al., 2014). Indeed, intragenic DNA methylation levels correlated positively with gene expression levels, regardless of the DNA methylation rate in the promoter region (Figure 2A, right panel). To demonstrate whether intragenic DNA methylation influences gene expression, we evaluated mRNA levels in DNA methylation inhibitor-treated ovary-derived Bm12 cells. According to qPCR, the mRNA levels of 14 of 32 genes randomly selected from those with approximately 10% DNA methylation in the ovary (Table S6) were reduced by 5-aza-2'-deoxycytidine (5-aza-dC) (Figure 2B), a DNA methylation inhibitor.

To confirm that upregulation of gene expression is mediated by intragenic DNA methylation, and not through promoter methylation, we determined the transcriptional activity of luciferase reporters in which 0, 2, 4, or 6 repeats of methyl-preferred fragments (TATCGAT) were inserted between the luciferase gene body and actin gene core promoter (Figure S2 and Table S7). Although luciferase activity increased in the presence of four or more repeated methyl-preferred fragments (Figure 2C), the luciferase activity of constructs in which these methyl-preferred fragments were inserted in the promoter region or downstream of the luciferase gene body did not increase (Figure 2D). Additionally, the transcriptional activity of a construct carrying the intragenic methyl-preferred fragment but the mutated methyl-preferred fragment
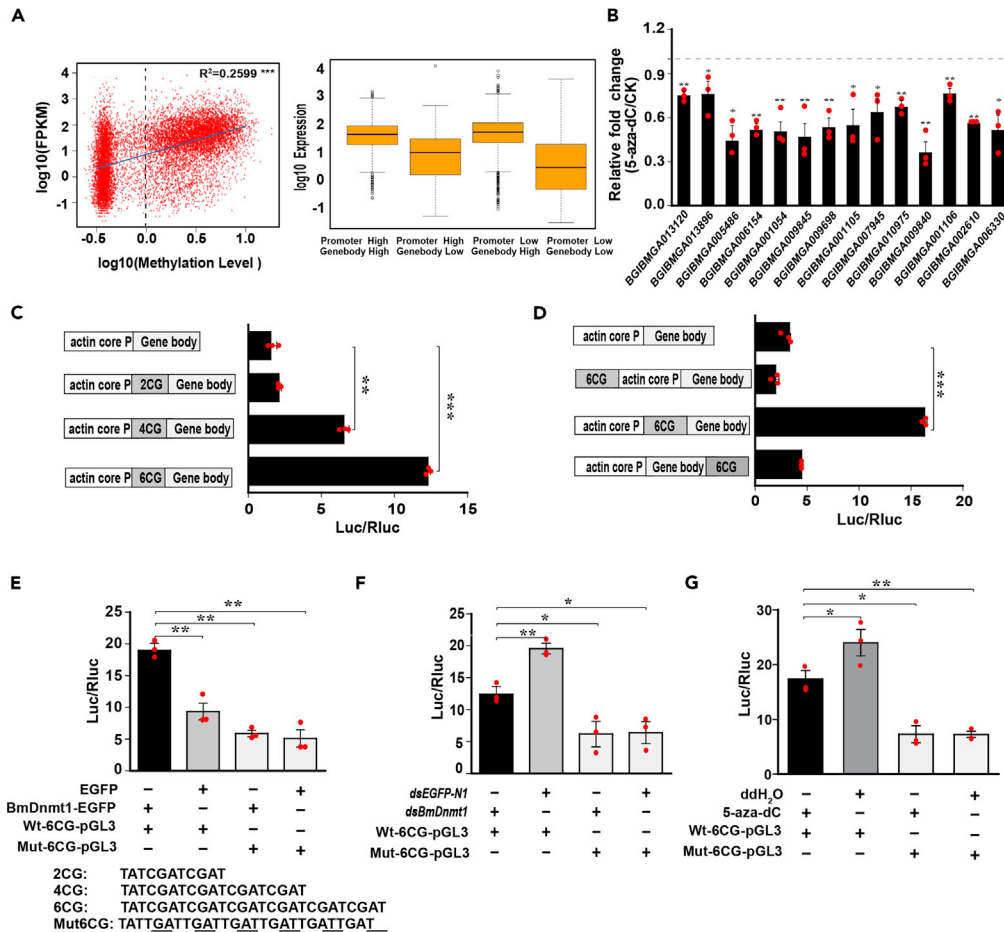
Figure 2. The relationship between intragenic DNA methylation and gene expression in *B. mori* ovarian cells

(A) Analyses of transcriptome and genomic 5mC levels in *B. mori* ovaries reveal a positive correlation between DNA methylation and gene expression levels (left), whereby higher gene expression levels are related to intragenic 5mC but not to promoter 5mC (right) (n = 2). R^2: goodness of fit of the trend lines. *p < 0.05, **p < 0.01, ***p < 0.001 by the F test.

(B) Changes in mRNA levels of 14 genes in *Bm12* cells, a cell line of the *B. mori* ovary, in response to the 5mC inhibitor 5-aza-dC. Relative mRNA levels are presented as the gene mRNA level in 5-aza-dC-treated samples divided by the control, which was normalized to the expression level of the housekeeping gene ribosomal protein 49 (*RP49*) (n = 3).

(C) Effect of the different lengths of methyl-CpG sequences in the 5′-terminus of the intragenic region on luciferase activity. The different repeats of the methyl-CpG sequence, ATCGAT, in which methyl-CpG most frequently appears in this sequence of the ovarian genome, were constructed into the luciferase vector between the actin core promoter and the luciferase ORF and cotransfected with the *BmDnmt1* overexpression vector for the determination of luciferase activity (n = 3).

(D) Effect of the methyl-CpG sequence at positions of the promoter and 5′- or 3′-terminus of the gene body on luciferase activity in *Bm12* cells (n = 3).

(E) Effect of *BmDnmt1-EGFP* overexpression on intragenic 5mC-regulated luciferase activity in *Bm12* cells (n = 3).

(F) Effect of *BmDnmt1* RNAi on intragenic 5mC-regulated luciferase activity in *Bm12* cells (n = 3). The *BmDnmt1-EGFP* vector or *dsBmDnmt1* was cotransfected with the vector including six repeats of the methyl-CpG sequence in the gene body for the determination of luciferase activity. In the mutated CpG sequence, "C" is changed to "T."

(G) Effect of 5-aza-dC (one microgram) on intragenic 5mC-regulated luciferase activity in *Bm12* cells transfected with the construct containing six repeats of methyl-CpG sequence in the gene body; ddH$_2$O was used as a control (n = 3). Data are presented as the mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001 by ANOVA.

was enhanced by an increase in DNA methylation due to *BmDnmt1* overexpression (Figure 2E). Conversely, knockdown of *BmDnmt1* by RNAi (Figure 2F) or treatment with a DNA methylation inhibitor (5-aza-dC) blocked the increase in luciferase activity (Figure 2G). These data suggest that DNA methylation in the gene body enhances gene expression.
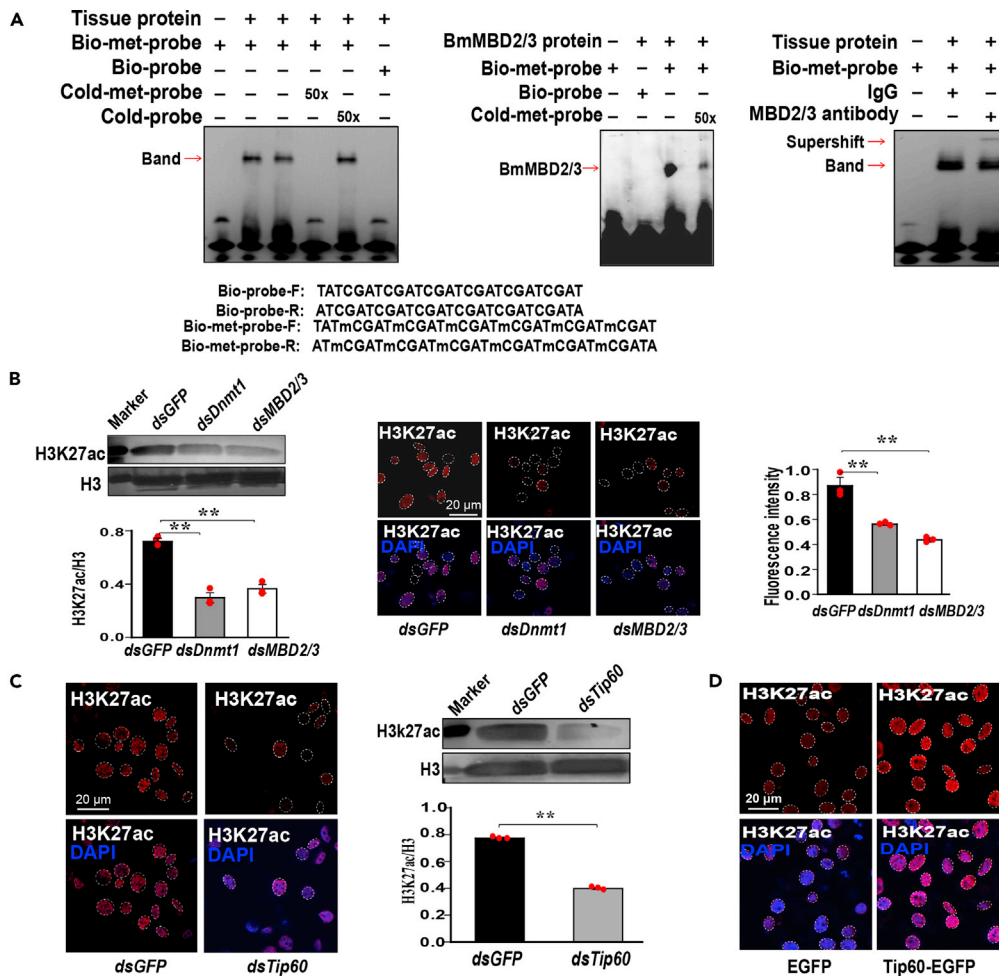
Figure 3. Effects of the methyl-CpG-binding domain protein BmMBD2/3 and transacetylase BmTip60 on H3K27ac

(A) EMSA of the binding of the methylated fragment with specific ovarian nuclear proteins (left) and purified BmMBD2/3 proteins (middle). Binding of BmMBD2/3 in the ovarian nuclear protein extract with the methylated fragment was recognized by the anti-BmMBD2/3 antibody (IgG as a control) (right). The bio-probe: labeled with biotin; the met-probe: methylated; the cold probe: unlabeled.

(B) Western blot and immunofluorescence analyses of H3K27ac levels after *BmDnmt1* and *BmMBD2/3* RNAi in *Bm12* cells (scale bar, 20 μm); quantitative analyses of protein bands (left) (n = 3) and the fluorescence intensity (right) (n = 3) were performed using ImageJ software. Marker: 15 kDa.

(C) Immunofluorescence (left) and western blot (right) analyses of H3K27ac levels after *BmTip60* RNAi in *Bm12* cells (scale bar, 20 μm). Marker: 15 kDa.

(D) Immunofluorescence analysis of H3K27ac levels after overexpressing *BmTip60* in *Bm12* cells (scale bar, 20 μm). Data are presented as the mean ± SEM. **p < 0.01 by ANOVA.

Gene body 5mC enhances gene expression through MBD2/3 and Tip60-mediated H3K27 acetylation

In mammals, MeCP2 or MBDs link DNA methylation to histone modification by recruiting histone deacetylase (Jones et al., 1998; Nan et al., 1998). Genes encoding two MBDs, BmMBD2/3 and BmMBD4, are found in *the B. mori* genome. To investigate which MBD recruits histone modification enzymes, binding assays of MBD with DNA were performed. The results showed that some proteins from ovarian nuclear extracts (Figure 3A, left panel) or purified BmMBD2/3 (GenBank accession no.: XP_004929675.1) (Figure 3A, middle panel) bound to the methyl-preferred fragment, and this binding was competed by excess unlabeled probe (Figure 3A). No band was detected in control assays with the unmethylated fragment. Moreover, the complex with the proteins from the ovarian nuclear extract was supershifted by a specific

anti-BmMBD2/3 antibody (Figure 3A, right panel). The BmMBD4 protein (GenBank accession no.: XM_021349846.1), which does not contain an MBD domain, did not bind to the methyl-preferred fragment (Figure S3). These EMSA results show that only BmMBD2/3 binds to methyl-CpG.

In addition, the levels of acetylation and methylation of several sites in histone H3 were examined after *BmDnmt1* or *BmMBD2/3* RNAi in *Bm12* cells. The results showed that knockdown of *BmDnmt1* or *BmMBD2/3* reduced H3K27ac (Figure 3B) but had no effect on H3K4me3, H3K9me3, H3K27me3, H3K79me3, H3K36me3 H3K4ac, H3K9ac, H3K23ac, or H3K18ac (Figure S4). Further analysis of the distribution of H3K27ac revealed that similar to the distribution of 5mC, H3K27ac was enriched at the 5′-end of the gene body, near the TSS, suggesting that 5mC is related to H3K27ac (Figure S5). To examine how BmMBD2/3 promotes H3K27ac, we performed an RNAi experiment targeting genes coding for seven histone acetyltransferases (HATs) and eight histone deacetylases (HDACs), which were identified in the *B. mori* genome using the gene sequences of other species (Table S8 and Figure S6). Among these enzymes, BmTip60 (GenBank accession no.: XP_004928297.1), a member of the HAT family, was the only one found to affect H3K27ac; other HATs and HDACs had no effect on H3K27ac (Figure S7). H3K27ac was reduced by *BmTip60* RNAi (Figure 3C) and enhanced by *BmTip60* overexpression (Figure 3D). These data indicate that BmMBD2/3 or BmTip60 enhances the acetylation levels of H3K27.

Interaction between MBD2/3 and Tip60 enhances gene transcription

To determine whether BmMBD2/3 and BmTip60 interact with each other to enhance H3K27ac, far-western blot and protein pull-down analyses were performed. The purified recombinant BmTip60-His protein (approximately 68 kDa) was transferred to a nitrocellulose membrane, incubated with the purified BmMBD2/3-GST protein, and then probed with an anti-GST-tagged antibody. BmTip60-His was able to interact with BmMBD2/3-GST, as indicated by the presence of a band recognized by the anti-GST-tagged antibody (Figure 4A, left panel). Similarly, BmMBD2/3-GST (approximately 54 kDa) was transferred to a membrane and incubated with BmTip60-His, and the anti-His-tag antibody recognized BmTip60-His bound to BmMBD2/3-GST (Figure 4A, second to left panel), indicating interaction between these two proteins. A similar result was obtained in the GST pull-down experiment: BmTip60-EGFP (approximately 87 kDa) as pulled down by BmMBD2/3-GST and recognized by an anti-EGFP antibody (Figure 4A, right panel). Furthermore, immunofluorescence showed that the BmMBD2/3 and BmTip60 proteins colocalized in the nucleus of *Bm12* cells (Figure 4B, left panel). Taken together, these data demonstrate that BmMBD2/3 and BmTip60 interact and form a complex.

To determine whether interaction between BmMBD2/3 and BmTip60 regulates gene transcription, the luciferase activity of a construct containing a fragment of six repeat methyl-preferred sequences was determined. We found that luciferase activity decreased after *BmMBD2/3* or *BmTip60* knockdown by RNAi (Figure 4B, middle panel). Additionally, luciferase activity increased when both BmMBD2/3 and BmTip60 were overexpressed simultaneously in *Bm12* cells (Figure 4B, right panel), suggesting that a complex of the two proteins upregulates gene expression by binding to the methylated gene body region and enhancing H3K27ac.

The aforementioned results demonstrate that *B. mori* MBD2/3 recruits acetyltransferase instead of deacetylase, as observed in mammals. To confirm the function of BmMBD2/3, the MBD and coiled-coil domains of BmMBD2/3 and *Mus musculus* MBD3 (MmMBD3) (GenBank accession no.: NP_038623.1) with approximately 50% amino acid sequence similarity (Figure 4C, top panel) were exchanged to create recombinant Bm:Mm hybrid MBD2/3 proteins (Figure 4C, left and middle panels), after which luciferase activity under control of the actin promoter and the *B. mori* methyl-preferred fragment was determined. Luciferase activity was enhanced by overexpressing BmTip60 and MBD2/3 containing the coiled-coil domain from *B. mori* but not that from MmMBD3 (Figure 4C, middle panel). In addition, GST pull-down assays confirmed that only MBD2/3 containing the *B. mori* coiled-coil domain interacted with BmTip60 (Figure 4C, right panel). These data show that the *B. mori* MBD2/3 coiled-coil domain is essential for recruiting HAT Tip60. Furthermore, to investigate why *B. mori* MBD2/3 recruits acetylase, insect coiled-coil domains, which are involved in protein-protein interactions, were analyzed. According to the results, insect coiled-coil domains are not highly conserved with those of mammals, but they are highly similar among insects (Figure S8), suggesting that the insect-specific coiled-coil domains are responsible for the interaction of MBD and Tip60.

To further demonstrate that the BmMBD/BmTip60 complex binds to mCpG regions in target genes, Chromatin immunoprecipitation assays were performed with four DNA-methylated genes, BGIBGMA001106,
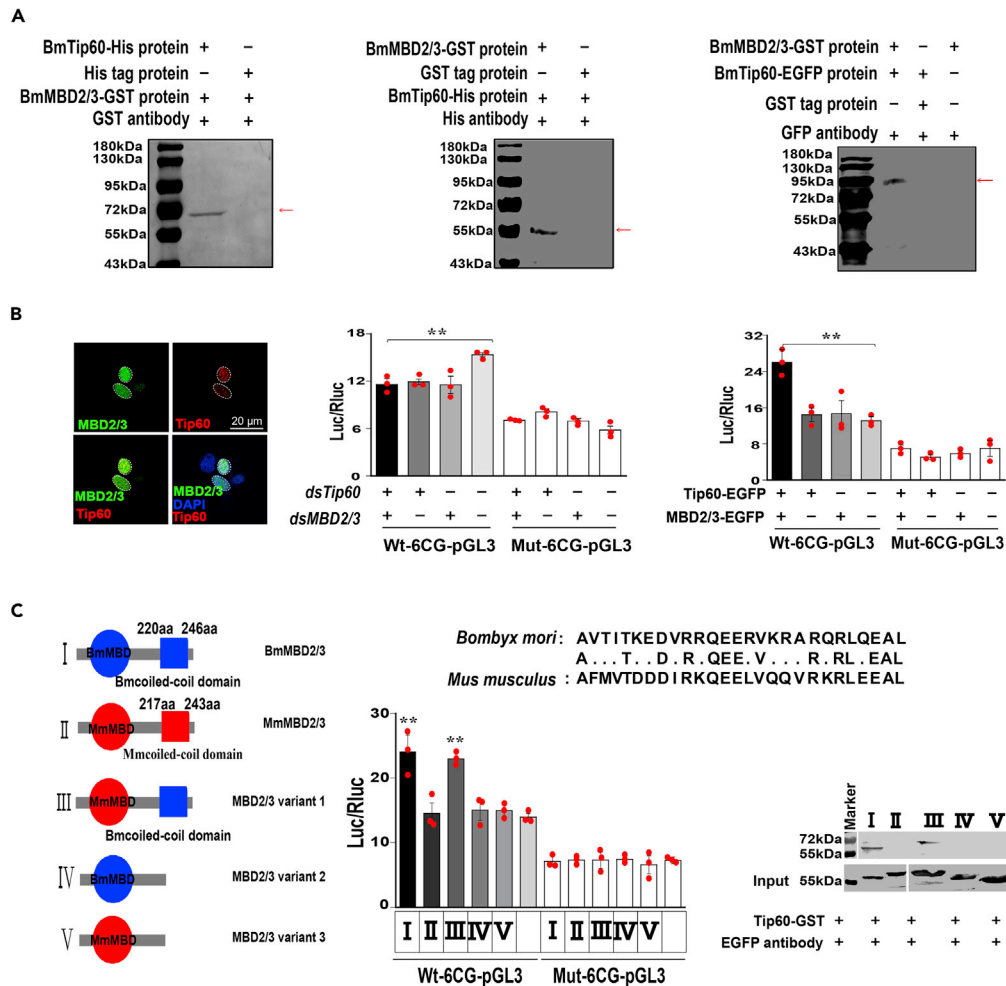
Figure 4. Effect of the interaction of BmMBD2/3 and BmTip60 on gene expression

(A) Far-western blot detection of the complex of BmTip60-His (protein with His tag) and BmMBD2/3-GST (protein with glutathione s-transferase tag) by the anti-GST tag antibody (left panel) or anti-His tag antibody (middle panel). GST pull-down experiments showed that the BmMBD2/3-GST complex and BmTip60-EGFP overexpressed in *Bm12* cells was recognized by the anti-GFP tag antibody (right panel).

(B) Immunofluorescence colocalization of BmMBD2/3 and BmTip60 proteins in *Bm12* cells (scale bar, 20 μm). Blue: DAPI, red: BmTip60, green: BmMBD2/3 (left). Changes in the luciferase activity of reporters containing the wild-type or cytosine-mutated six repeats of the methylated sequence in *Bm12* cells cotransfected with *dsTips60* or *dsMBD2/3* (middle) and cotransfected with Tips60-EGFP or MBD2/3-EGFP constructs (right) (n = 3).

(C) Identification of *B. mori*-specific MBD2/3. The luciferase activities of the reporter constructs with intragenic methyl-CpG fragments were determined under overexpression of MBD2/3s consisting of *B. mori* or *Mus musculus* MBD domains (left and middle) (n=3). The sixth and twelfth columns were controls transfected with *pEGFP*. GST pull-down analyses showed that only MBD2/3 with a *B. mori*-specific coiled-coil domain was able to interact with BmTip60 (right). Data are presented as the mean ± SEM. **p < 0.01 by ANOVA.

BGIBGMA005486, BGIBGMA009840, and BGIBGMA010975, *the mRNA levels of which* were reduced by 5-aza-dC treatment (Figure 2B). *Bm12* cells were transfected with the recombinant plasmid BmMBD2/3-3×FLAG or BmTip60-3×FLAG (EGFP vector as a control), and expression of both proteins in the cells was confirmed by western blot (Figure 5A). In these BmMBD2/3-3×FLAG or BmTip60-3×FLAG-overexpressing cells, anti-FLAG antibodies, but not control IgG, recognized BmMBD2/3-3×FLAG or BmTip60-3×FLAG, precipitating and enriching the mCpG regions *of four test genes* (Figures 5B and 5C). The enriched mCpG region sequences were amplified by PCR and examined by DNA sequencing (Figure 5D). These experiments demonstrated that the BmMBD/BmTip60 complex bound to the intragenic mCpG regions of genes *in vivo*. All these results suggest that BmMBD2/3 binds to intragenic methyl-CpG through its
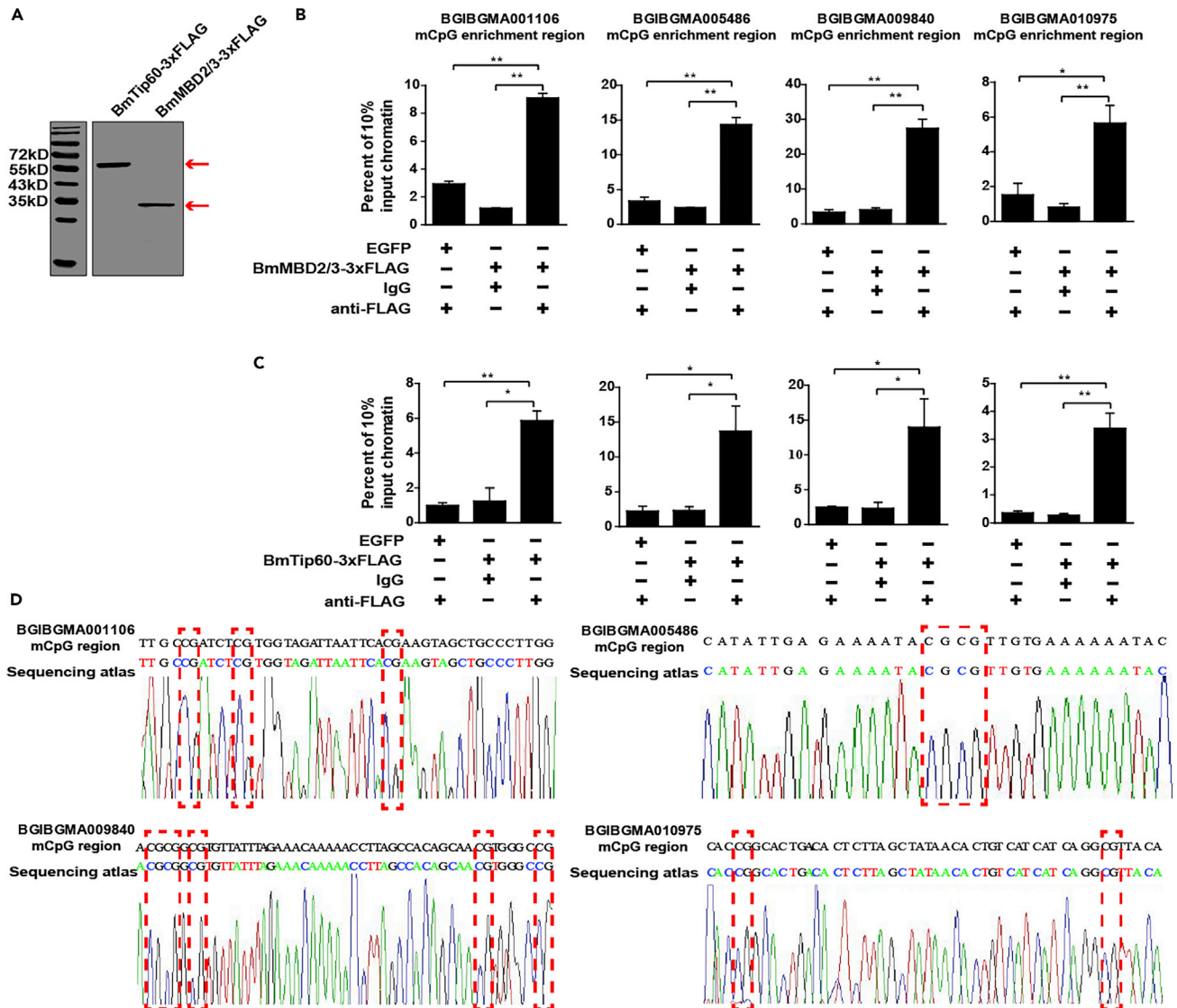
Figure 5. Chromatin immunoprecipitation (ChIP) analysis of the binding of BmMBD2/3 and BmTip60 to the mCpG regions of target genes in Bm12 cells

(A) Western blot analysis showed that BmTip60-3×FLAG and BmMBD2/3-3×FLAG were expressed in Bm12 cells.

(B and C) qRT-PCR detection of the enrichment of mCpG regions of BGIBGMA001106, BGIBGMA005486, BGIBGMA009840, and BGIBGMA010975 genes with (B) BmMBD2/3–3×FLAG and (C) BmTip60-3×FLAG (n = 3).

(D) Sequences of the enriched RT-PCR products of the mCpG regions of the BGIBGMA001106, BGIBGMA005486, BGIBGMA009840 and BGIBGMA010975 genes. Data are presented as the mean ± SEM. *p < 0.05, **p < 0.01 by ANOVA.

MBD domain and then recruits HAT BmTip60, resulting in an increase in H3K27ac and target gene transcription.

Gene body DNA methylation affects egg production and hatch rates

The above-mentioned results show MBD2/3-mediated intragenic DNA methylation in insects. To investigate whether *MBD2/3* plays a role in insect reproduction similar to *Dnmt1*, the expression patterns of *MBD2/3* from a set of randomly selected insect species were analyzed. The *B. mori MBD2/3* expression pattern (Figure 6A) was similar to that of *BmDnmt1* (Figure 1B), which was highly expressed in ovaries compared to other tissues. Similar expression patterns of *MBD2/3* were also observed in *Periplaneta americana*, *Apis cerana*, *Spodoptera litura* and *Apis mellifera* (Figure 6A), suggesting that insect MBD2/3 may
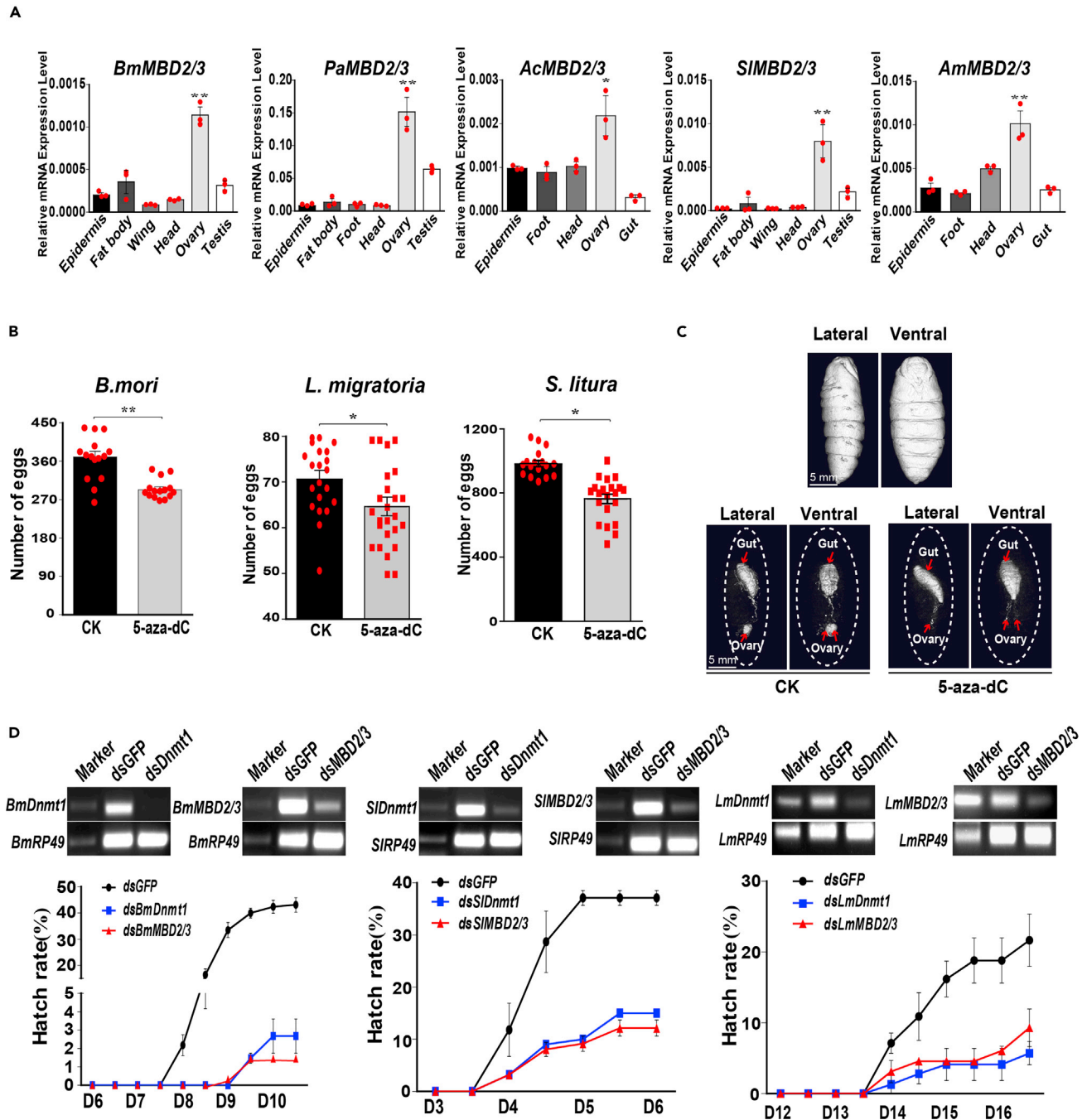
Figure 6. Effect of DNA methylation on insect reproduction

(A) Expression patterns of MBD2/3s in different tissues of different insects, including *B. mori* (Lepidoptera), *P. americana* (Blattaria), *A. cerana* (Hymenoptera), *S. litura* (Lepidoptera), and *A. mellifera* (Hymenoptera) (n = 3).

(B) Changes in the number of eggs (left) after injection with the DNA methylation inhibitor 5-aza-dC in *B. mori*, *S. litura,* and *L. migratoria* (n = 15–25).

(C) Microcomputed tomographic analysis of ovarian morphology in 4-day-old pupae after 5-aza-dC treatment to show the ovary without dissection (scale bar, 5 mm).

(D) Changes in the hatch rate after *Dnmt1* or *MBD2/3* dsRNA was injected into *B. mori* (left, n = 3), *S. litura* (middle, n = 2), and *L. migratoria* (right, n = 2) embryos within 2 h of fertilization. Dn: n day of embryonic development. Marker: 150 bp. Data are presented as the mean ± SEM. *p < 0.05, **p < 0.01 by ANOVA.

have a physiological function in egg production and embryonic development. Next, we examined the holometabolous insects *B. mori* and *S. litura* and the hemimetabolous insect *Locusta migratoria* to study the function of DNA methylation. Injection of the 5mC inhibitor 5-aza-dC into *B. mori*, *S. litura*, and *L. migratoria* decreased the number of eggs oviposited (Figure 6B). Microcomputed tomographic analysis of ovarian morphology in 4-day-old pupae (the middle stage of *B. mori* pupae) after 5-aza-dC treatment revealed that ovarian development was retarded (Figure 6C). These results suggest that DNA methylation enhances oocyte development and egg production and subsequent embryonic development. To determine the involvement of *Dnmt1* and *MBD2/3* in hatching, the dsRNAs targeting *Dnmt1* and *MBD2/3* were injected into eggs on the first day of the embryonic development. Knockdown of these genes significantly reduced hatch rates in *B. mori*, *S. litura*, and *L. migratoria* compared with the control injected with *dsGFP* (Figure 6D). Furthermore, we analyzed the methylated genes in the methylation genome of the *B. mori* ovary and found that the top 20 enriched methylated pathways include protein synthesis-related RNA transport, ribosome, and protein processing in the endoplasmic reticulum (ER) and spliceosome (Figure S9). These results suggest that intragenic DNA methylation may be involved in the upregulation of genes coding for proteins that play key roles in protein synthesis that function in the production of a large number of eggs during insect reproduction.

DISCUSSION

DNA methylation in the promoter of genes has been extensively investigated for its role in inhibiting gene expression (Jones, 2012). Additionally, 5mC enrichment in the gene body was demonstrated to promote gene expression in cancer cells by regulating alternative splicing via H3K36m3 (Yang et al., 2014), and 5mCs enrichment in gene bodies has been detected in insects (Zemach et al., 2010; Glastad et al., 2014; Hunt et al., 2013; Bonasio et al., 2012). However, the role of intragenic DNA methylation in regulating gene expression and the underlying mechanism remain unknown. In the present study, we discovered the function of gene body DNA methylation in enhancing gene transcription and examined the underlying mechanisms. Intragenic 5mCpG near the TSS enhanced the transcriptional activity of genes (Figure 2). Treatment with a 5mC inhibitor resulted in a decrease in the mRNA levels of target genes that are intragenically methylated in *B. mori* ovarian cells (Figure 2). Furthermore, MBD2/3 bound to methyl-CpG in the intragenic regions of genes via its methyl-CpG-binding domain (Figure 3A) and then interacted with the acetyltransferase Tip60 through its coiled-coil domain. The MBD2/3-Tip60 complex increased H3K27ac (Figures 3 and 4). Similar to *BmDnmt1* knockdown, RNAi-mediated knockdown of *BmMBD2/3* decreased the egg number and hatch rates of *B. mori* (Figure 6D). Taken together, these data conclusively show that intragenic DNA methylation upregulates gene expression to ensure a high reproductive rate in insects to enable them to survive challenging environments, raising many intriguing questions for future research.

In insects such as *B. mori*, analyses of the methylated genome and transcriptome have revealed that intragenic DNA methylation is enriched in highly expressed genes (Xiang et al., 2010). However, experimental evidence for the ability of intragenic 5mC to regulate gene expression is lacking. Our results of promoter activity and mRNA levels in 5mC inhibitor-treated cells showed that 5mC in the 5′ UTR region of the gene body enhances gene transcription (Figure 2). In addition, a DNA fragment with more than two CpGs was able to promote gene transcription (Figure 2). These data suggest that in insects, 5′ UTR DNA methylation enhances gene transcription and explains why the insect intragenic DNA methylation rate correlates positively with gene expression levels. Nevertheless, not all methylated genes were affected by 5mC inhibitor treatment (Figure 2). Several reasons may explain this result. First, hypermethylation is more likely to occur in highly expressed housekeeping genes, which may be regulated by multiple pathways or mechanisms, including DNA methylation. Therefore, compared with their high background expression levels, it is difficult to detect small changes induced by 5mC inhibitor treatment on the expression levels of these genes. Second, *Dnmt1* RNAi or 5mC inhibitor treatment may not affect the intragenic methylation complex (Cook et al., 2019), leading to unaltered gene expression. Hence, it will be interesting to investigate differences in intragenic methylation-regulated gene expression by comparing 5mC inhibitor treatment-effective and -ineffective gene expression in future studies.

We discovered that 5′ UTR 5mC enhances the promoter activity of genes via H3K27ac. H3K27ac usually functions as an enhancer of gene expression and is enriched around the TSS (Creyghton et al., 2010; Karlić et al., 2010). We found that DNA methylation affected the level of H3K27ac but not that of H3K4me3 and H3K36me3 in *B. mori*. H3K4me3 is a mark for transcription initiation (Liu et al., 2016), and H3K36me3 is a mark of transcriptional elongation (Yang et al., 2014). We also found that *B. mori* methyl-CpGs and

H3K27ac are enriched in the region near the TSS (Figure 1), suggesting the relationship between H3K27ac and intragenic DNA methylation. Therefore, this work reveals an intragenic methylation regulation, in contrast to intragenic 5mC, which enhances transcription elongation in the methylated area by regulating alternative splicing (Maunakea et al., 2013; Yang et al., 2014). As 5mC has been found to be enriched in the 5′-region of gene bodies in other invertebrates (Zemach et al., 2010; Hunt et al., 2013; Bonasio et al., 2012; Lyko et al., 2010), intragenic DNA methylation might enhance gene transcription in other insects and invertebrates via H3K27ac.

We found that insect intragenic DNA methylation enhances H3K27ac through the methyl-CpG reader MBD2/3 and its partner Tip60. In mammals, MBD2 binds to mCG via the MBD domain and associates with corepressors and other protein complexes through its transcriptional repression domains to inhibit gene expression (Nan et al., 1998; Baubec et al., 2013). MBD3 also recruits deacetylase with its coiled-coil domain to inhibit gene expression (Le Guezennec et al., 2006). These functions of MBD2 and MBD3 appear to be different from those of insect MBD2/3. In fact, MBD2/3 is the only MBD family protein in insects with the same protein structure as mammalian MBD3 (Hendrich and Tweedie, 2003), although the sequence similarity of the coil-coiled domain in insect MBD2/3 and mammalian MBD3 is not highly conserved (Figure S8). Our results demonstrate that *B. mori* MBD2/3 forms a complex with ATCGAT and BmTip60 using its MBD and coiled-coil domains (Figures 3 and 4), respectively. However, the coiled-coil domain of mouse MuMBD3 is not able to recruit BmTip60 (Figure 4). Thus, we identified an insect-specific regulatory mechanism of MBD-mediated DNA methylation. However, several questions about this mechanism need to be addressed in future studies. First, although we demonstrated that the insect MBD coiled-coil domain can recruit HAT Tip60, there is no typical coiled-coil domain in Tip60; thus, further studies are needed to understand how MBD2/3 interacts with Tip60. Second, MBD2/3 is the only MBD protein that binds to 5mC in insects, but it is unknown whether it can recruit deacetylase to inhibit transcription factor binding to the gene promoter in addition to recruiting Tip60 to enhance gene expression. As the NuRD complex is present in *D. melanogaster* (Zhang et al., 2016), it needs to be determined how MBD2/3 can switch binding between deacetylase and acetyltransferase. Third, it remains to be explored how MBD2/3-mediated intragenic DNA methylation regulates housekeeping gene expression to affect insect development. With these questions in mind, the present study, therefore, opens a research perspective for insect epigenetics.

Insects are the largest group of animals on Earth and have high reproduction rates. We found that intragenic DNA methylation regulates insect reproduction. *MBD2/3* RNAi affected the reproduction rates of the holometabolous insects *B. mori* and *S. litura* as well as the hemimetabolous insect *L. migratoria* (Figure 6). The effect of *MBD2/3* RNAi was similar to that of *Dnmt1* RNAi (Figure 6), suggesting that Dnmt1 regulates insect ovarian development, probably through *MBD2/3*. The effect of Dnmt on egg production and hatch rates has been reported for other insects. For instance, *Nlu-Dnmt1* and *Nlu-Dnmt3* knockdown led to fewer offspring in *N. lugens* (Zhang et al., 2015) and *O. fasciatus Dnmt1* RNAi affected the number and development of eggs and damaged the nuclear structure of the follicular epithelium (Bewick et al., 2019). In *Nasonia, Dnmt1* loss resulted in embryonic lethality during the onset of gastrulation (Zwier et al., 2012). MBD2/3 likely regulates reproduction in other insects in addition to those assessed in this study. Thus Dnmt and MBD2/3 may constitute target genes for pest control.

Why does insect intragenic DNA methylation affect reproduction? We analyzed genes that are methylated in the *B. mori* ovary and found that methyl-CpGs mostly appeared in the gene bodies of genes encoding proteins involved in protein synthesis, including genes participating in transcription and translation (Figure S9). Treatment with a 5mC inhibitor at the beginning of ovarian development resulted in retardation of ovary development and a decrease in mature eggs and their hatch rate (Figure 6). These results suggest that intragenic DNA methylation may enhance the protein synthesis required for egg maturation, including chorion synthesis and metabolism, to support high fecundity and hatch rates. By enhancing gene expression through MBD2/3-mediated H3K27ac, intragenic DNA methylation in the ovary and during early embryo development is necessary for a high reproduction rate in insects. Regardless, how intragenic DNA methylation regulates development-related gene expression to affect ovarian and fertilized egg development needs further investigation. Additionally, whether intragenic DNA methylation affects other tissues in addition to the ovary and fertilized egg needs to be clarified in the future. Although higher 5mC levels appear in the gene bodies of ovaries (this study), silk glands (Xiang et al., 2010), and midguts of *B. mori* (Wu et al., 2017), the methylation rates in these tissues are different,

with 0.52% in ovary, 0.6% in midgut and fat body, and 0.11% in silk gland. Of this 83.61%, 81%, and 99.2% of mCs were at CG sites in the ovary, midgut, and silk gland, respectively, whereas only 13.47% and 2.93% were at CHG and CHH sites in the ovary and 19% mC and 0.8% mCs were at both sites in the midgut and the silk gland, respectively. Unlike in the ovary, 5mCs in CHH sites were higher than at CHG sites in the midgut. These data suggest that there may be different functions and regulatory mechanisms of 5mC in different tissues of *B. mori*.

In this study, we show a regulatory mechanism of intragenic DNA methylation in insects. MBD2/3 binds to methyl-CpG in intragenic DNA and recruits Tip60, leading to elevated histone H3K27ac levels and enhanced gene expression. This regulatory mechanism of intragenic DNA methylation facilitates a high rate of reproduction.

Limitations of the study

This study reveals a regulatory mechanism that intragenic DNA methylation enhances gene expression and demonstrates the role of intragenic DNA methylation in insect reproduction. However, several questions about this mechanism and function need to be addressed in future studies. First, the complex of methylated DNA and proteins needs to be explored in depth. It needs to understand how MBD2/3 interacts with Tip60 to enhance H3K27 acetylation. Second, it needs to explain why the expression levels of most methylated genes are not changed after demethylation? Whether there are regulatory pathways to compensate for the loss of DNA methylation and to promote these gene expression needs to be clarified. Third, how intragenic DNA methylation regulates development-related gene expression to affect insect ovarian and fertilized egg development needs further investigation.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Sichun Zheng (sczheng@scnu.edu.cn).

Materials availability

Plasmids generated in this study have been deposited to Guangdong Provincial Key Laboratory of Insect Developmental Biology and Applied Technology, School of Life Sciences, South China Normal University, Guangzhou, China.

The *Bombyx mori* used in this study was obtained from the Research and Development Center of the Sericulture Research Institute of the Academy of Agricultural Sciences of Guangdong Province, China. *Locusta migratoria* used in this study was obtained from the School of Life Sciences, Henan University, Kaifeng, China. *Spodoptera litura* and *Periplaneta americana* used in this study were obtained from School of Life Sciences, South China Normal University, Guangzhou, China. *Apis cerana* and *Apis mellifera* used in this study were obtained from Institute of Biological Resources Application of Guangdong Province, China.

This study did not generate new unique reagents.

All reagents used in this study are available from the Lead Contact with a completed Materials. All unique reagents generated in this study are available from the Lead contact with a completed Materials Transfer Agreement.

Data and code availability

The raw Illumina sequencing data from the BS-seq and RNA-seq generated during this study are deposited in SRA under accession number SRP217273 at NCBI.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2021.102040.

AUTHOR CONTRIBUTIONS

G.X. conducted most of the experiments, participated in the data analyses, and drafted the manuscript. H.L., Y.Y. and Y.P. conducted the gene cloning, protein expression and interaction experiments, cell culture, fertilized egg injection, and sample preparation. C.G. and X.P. assisted with the experiments and insect rearing. Q.F. and Q.S. provided technical and material support, participated in discussions, and helped draft and revise the manuscript. S.R.P. gave suggestions and revised the manuscript. S.Z. conceived the design and oversaw the study, financially supported the research, and drafted and finalized the manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Baubec, T., Ivánek, R., Lienert, F., and Schübeler, D. (2013). Methylation-dependent and -independent genomic targeting principles of the MBD protein family. Cell 153, 480–492.

Bewick, A.J., Sanchez, Z., Mckinney, E.C., Moore, A.J., Moore, P.J., and Schmitz, R.J. (2019). Dnmt1 is essential for egg production and embryo viability in the large milkweed bug, Oncopeltus fasciatus. Epigenetics Chromatin 12, 6.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. 16, 6–21.

Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression—belts, braces, and chromatin. Cell 99, 451–454.

Bonasio, R., Li, Q., Lian, J., Mutti, N.S., Jin, L., Zhao, H., Zhang, P., Wen, P., Xiang, H., Ding, Y., et al. (2012). Genome-wide and caste-specific DNA methylomes of the ants Camponotus floridanus and Harpegnathos saltator. Curr. Biol. 22, 1755–1764.

Bownes, M. (1989). The roles of juvenile hormone, ecdysone and the ovary in the control of Drosophila vitellogenesis. J. Insect Physiol. 35, 409–413.

Cook, N., Parker, D.J., Tauber, E., Pannebakker, B.A., and Shuker, D.M. (2019). Validating the demethylating effects of 5-aza-2′-deoxycytidine in insects requires a whole-genome approach. Am. Nat. 194, 432–438.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hannaet, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. U S A 107, 21931–21936.

Dobosy, J.R., and Selker, E.U. (2001). Emerging connections between DNA methylation and histone acetylation. Cell. Mol. Life Sci. 58, 721–727.

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. J. Biol. Chem. 278, 4035–4040.

Glastad, K.M., Hunt, B.G., and Goodisman, M.A.D. (2014). Evolutionary insights into DNA methylation in insects. Curr. Opin. Insect Sci. 1, 25–30.

Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. Annu. Rev. Biochem. 74, 481–514.

Hunt, B.G., Glastad, K.M., Yi, S.V., and Goodisman, M.A.D. (2013). The function of intragenic DNA methylation: insights from insect epigenomes. Integr. Comp. Biol. 53, 319–328.

Hendrich, B., and Tweedie, S. (2003). The methyl-CpG binding domain and the evolving role of DNA methylation in animals. Trends Genet. 19, 269–277.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. 13, 484–492.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat. Genet. 19, 187–191.

Karlić, R., Chung, H.R., Lasserre, J., Vlahovicek, K., and Vingron, M. (2010). Histone modification levels are predictive for gene expression. Proc. Natl. Acad. Sci. U S A 107, 2926–2931.

Kim, J.K., Samaranayake, M., and Pradhan, S. (2009). Epigenetic mechanisms in mammals. Cell. Mol. Life Sci. 66, 596–612.

Le Guezennec, X., Vermeulen, M., Brinkman, A.B., Hoeijmakers, W.A.M., Cohen, A., Lasonder, E., and Stunnenberg, H.G. (2006). MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. Mol. Cell Biol. 26, 843–851.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462, 315–322.

Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H., et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature 537, 558–562.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., and Maleszka, R. (2010). The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol. 8, e1000506.

Maunakea, A.K., Chepelev, I., Cui, K., and Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. Cell Res. 23, 1256–1269.

Mito, T., Nakamura, T., and Noji, S. (2010). Evolution of insect development: to the hemimetabolous paradigm. Curr. Opin. Genet. Dev. 20, 355–361.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a

histone deacetylase complex. Nature *393*, 386–389.

Sarda, S., Zeng, J., Hunt, B.G., and Yi, S.V. (2012). The evolution of invertebrate gene body methylation. Mol. Biol. Evol. *29*, 1907–1916.

Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat. Rev. Genet. *9*, 465–476.

Swevers, L., and Iatrou, K. (2003). The ecdysone regulatory cascade and ovarian development in lepidopteran insects: insights from the silkmoth paradigm. Insect Biochem. Mol. Biol. *33*, 1285–1297.

Wu, P., Jie, W., Shang, Q., Annan, E., Jiang, X., Hou, C., Chen, T., and Guo, X. (2017). DNA methylation in silkworm genome may provide insights into epigenetic regulation of response to Bombyx mori cypovirus infection. Sci. Rep. *7*, 16013.

Xiang, H., Li, X., Dai, F., Xu, X., Tan, A.J., Chen, L., Zhang, G., Ding, Y., Li, Q., Lian, J., et al. (2013). Comparative methylomics between domesticated and wild silkworms implies

possible epigenetic influences on silkworm domestication. BMC Genomics *14*, 646.

Xiang, H., Zhu, J., Chen, Q., Dai, F., Li, X., Li, M., Zhang, H., Zhang, G., Li, D., Dong, Y., et al. (2010). Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. Nat. Biotechnol. *28*, 516–520.

Xu, G., Zhang, J., Lyu, H., Song, Q., Feng, Q., Xiang, H., and Zheng, S. (2018). DNA methylation mediates BmDeaf1-regulated tissue- and stage-specific expression of BmCHSA-2b in the silkworm, Bombyx mori. Epigenetics Chromatin *11*, 32.

Yang, X., Han, H., De Carvalho, D.D., Lay, F.D., Jones, P.A., and Liang, G. (2014). Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell *26*, 577–590.

Zhang, J., Xing, Y., Li, Y., Yin, C., Ge, C., and Li, F. (2015). DNA methyltransferases have an essential role in female fecundity in brown planthopper, Nilaparvata lugens. Biochem. Biophys. Res. Commun. *464*, 83–88.

Zhang, T., Song, W., Li, Z., Qian, W., Wei, L., Yang, Y., Wang, W., Zhou, X., Meng, M., Peng, J., et al. (2018). Kruppel homolog 1 represses insect ecdysone biosynthesis by directly inhibiting the transcription of steroidogenic enzymes. Proc. Natl. Acad. Sci. U S A *115*, 3960–3965.

Zhang, W., Aubert, A., Gomez de Segura, J.M., Karuppasamy, M., Basu, S., Murthy, A.S., Diamante, A., Drury, T.A., Balmer, J., Cramard, J., et al. (2016). The nucleosome remodeling and deacetylase complex NuRD is built from preformed catalytically active sub-modules. J. Mol. Biol. *428*, 2931–2942.

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science *328*, 916–919.

Zwier, M.V., Verhulst, E.C., Zwahlen, R.D., Beukeboom, L.W., and van de Zande, L. (2012). DNA methylation plays a crucial role during early Nasonia development. Insect Mol. Biol. *21*, 129–138.

Supplemental Information

Intragenic DNA methylation regulates

insect gene expression and reproduction

through the MBD/Tip60 complex

Guanfeng Xu, Hao Lyu, Yangqin Yi, Yuling Peng, Qili Feng, Qisheng Song, Chengcheng Gong, Xuezhen Peng, Subba Reddy Palli, and Sichun Zheng

Transparent Methods

Insects and inhibitor treatments

The *B. mori* strain P50 was obtained from the Research and Development Center of the Sericulture Research Institute of the Academy of Agricultural Sciences of Guangdong Province, China. Larvae were reared on fresh mulberry leaves at 25°C and a photoperiod of 12 h light:12 h darkness. *S. litura* indivudals were obtained from the Institute of Entomology, Sun Yat-sen University (Guangzhou, China). Larvae were reared on artificial medium at 25°C and 70% humidity under a photoperiod of 12 h light:12 h darkness. Specimens of *L. migratoria* were obtained from the School of Life Sciences, Henan University (Kaifeng, China). Larvae were reared under a 14-h:10-h light:dark cycle and at 30 ± 2°C; larvae and adults were supplied with fresh leaves of *Setaria viridis (L.) Beauv* and wheat bran.

The DNA methylation inhibitor 5-aza-2'-deoxycytidine (5-aza-dC) (Sigma, California, USA) was dissolved in ddH_2O. Two microliters of 5-aza-dC at a concentration of 10 µg/µL was injected into the abdomen of *B. mori* and *S. litura* on pupa day 1 and *L. migratoria* on adult day 1. The same volume of ddH_2O was injected as a control. Three replicates of approximately 30-40 individuals per replicate were carried out.

Library construction and mapping

Genomic DNA was extracted from the ovaries of seven-day-old pupae using the phenol extraction method. After genomic DNA was extracted, DNA concentration and integrity were detected by a NanoDrop spectrophotometer and agarose gel

electrophoresis, respectively. DNA libraries for bisulfite sequencing were prepared by Gene Denovo Biotechnology Co (Guangzhou, China). Briefly, genomic DNA was fragmented into 100-300-bp fragments by sonication (Covaris, Woburn, USA) and purified with a MiniElute PCR Purification Kit (QIAGEN, Dusseldorf, Germany). The fragmented DNAs were end-repaired, and a single "A" nucleotide was added to the 3' end of the blunt fragments. The genomic fragments were ligated to methylated sequencing adapters and fragments with adapters were bisulfite converted using a Methylation-Gold kit (ZYMO, Irvine, USA). Finally, the converted DNA fragments were PCR amplified and sequenced using an Illumina HiSeqTM 2500. Short reads generated by Illumina sequencing were aligned to the *B. mori* reference genome (Wang et al., 2005; Xia et al., 2004).

Methylation level analysis

To obtain high-quality clean reads, raw reads were filtered by removing reads containing more than 10% unknown nucleotides and low-quality reads containing more than 40% low-quality (Q-value $\leqslant$ 20) bases. The obtained clean reads were mapped to the species reference genome using BSMAP software (Xi and Li, 2009) (version: 2.90) by default. Then, a custom Perl script was used to call methylated cytosines, and the methylated cytosines were tested with the correction algorithm described in Lister et al. (Lister et al., 2009). The methylation level was calculated based on the methylated cytosine percentage of each chromosome and in different regions of the genome for each sequence context (CG, CHG, and CHH). To assess

different methylation patterns in different genomic regions, the methylation profile at flanking 2-kb regions and gene bodies (or transposable elements) was plotted based on the average methylation levels for each window.

RNA-seq analyses

Total RNA was extracted from the ovaries of 7-day-old pupae using TRIzol reagent (TaKaRa, Dalian, China) according to the manufacturer's instructions. RNA-seq was performed using the Illumina HiSeq^{TM} 4000 platform by Gene Denovo Biotechnology, and paired reads with an average length of 150 bp were generated. The clean reads that were filtered from the raw reads were used for mapping to reference *B. mori* genomes (Wang et al., 2005; Xia et al., 2004). Using the assembled transcripts generated after mapping, the gene expression level of each gene was normalized using FPKM values (fragments per kilobase of exon per million fragments mapped) with Cufflinks software (Trapnell et al., 2012).

ChIP-seq analyses

H3K27ac ChIP-seq data were obtained from NCBI with the BioProject accession number PRJNA450142 (Cheng et al., 2018). We used Bowtie2 to map paired-end clean reads to the silkworm reference genome. Highly enriched peaks were obtained by MACS2 using standard settings.

Quantitative real-time PCR (qRT-PCR)

Total RNA was extracted using TRIzol (TaKaRa, Dalian, China), and cDNAs were synthesized using First-Strand cDNA Synthesis Kit (TaKaRa, Dalian, China) following the manufacturer's protocol. qPCR was performed using 2×SYBR Premix

EXTaq[TM] Kit (TaKaRa, Dalian, China). The relative mRNA level of gene expression was calculated by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). Data were normalized to the housekeeping gene ribosomal protein 49 (Rp49) and analyzed by the $2^{-\Delta\Delta Ct}$ method. All data consisted of three biological replicates. The sequences of the qRT-PCR primers used are shown in Table S9.

Construction of the reporter luciferase vector

Genomic DNA was extracted from *Bm12* cells. The promoter of the actin gene, including 91 bp of the core promoter and 63 bp of the 5' UTR, was cloned into the pMD-18T vector (TaKaRa, Dalian, China) according to the sequence of the *B. mori* genome SilkDB (http://silkworm.genomics.org.cn). The methyl-CpG fragment was amplified by PCR. The actin promoter and methyl-CpG fragment were cloned into the luciferase reporter pGL3-basic vector (Promega, Madison, USA). The primers used for constructing the plasmids are listed in Table S9.

Cell transfection and transcriptional activity determination

The *B. mori* cell line DZNU-Bm-12 (*Bm12*) originally developed from ovarian tissues (Khurad et al., 2009) was maintained at 28°C on Grace medium (Invitrogen, California, USA) supplemented with 10% fetal bovine serum (FBS) (HyClone, Utah, USA).

Bm12 cells at the logarithmic growth phase were inoculated into Grace Insect medium in 12- or 24-well culture plates (Corning, New York, USA) and cultured for 12 h. Cell transfection was conducted when the cells were at 80% density. The Renilla luciferase vector pRL-SV40 was cotransfected for normalization of firefly luciferase

activity.

For transfection, a mixture of 30 μL containing 1 μg of the pGL3-derived reporter plasmid DNA, 100 ng internal control plasmid (pRL-SV40), and 3 μL Fugene HD transfection reagent (Promega, Wisconsin, USA) in Opti-MEM Reduced Serum Medium (Life Technologies, Massachusetts, USA) was added to the cells, which were cultured for 48 h at 28°C, harvested and used for the transcriptional activity assays or isolation of protein and RNA.

For cotransfection, a mixture of 30 μL containing 0.5 ng of the wild-type or mutated pGL3-derived reporter plasmid DNA, 0.5 μg overexpression plasmid or 1 μg dsRNA, 100 ng internal control plasmid (pRL-SV40 vector) and 4 μL Fugene HD transfection reagent in Opti-MEM Reduced Serum Medium was added to the cells in Grace medium.

Luciferase activity was assessed as follows. Briefly, after 48 h of transfection, the cells were washed twice with filtered PBS and then lysed in 100 μL Passive Lysis Buffer (Promega, Wisconsin, USA). The samples were centrifuged at 800 × g for 5 min at room temperature. The supernatant was used to analyze the luciferase activity with Dual-Luciferase Assay System according to the manufacturer's protocol and a luminometer (IBA7300, Veritas, Turner Biosystems, California, USA). Luciferase activity was normalized to that of Renilla luciferase.

Expression and purification of recombinant proteins

Open reading frames of select genes were amplified using cDNA or synthesized by TSINGKE (Beijing, China). The cDNAs were subcloned into the pPET-28a, pGEX-6P-1 or EGFP vector with a 6×His, GST tag or GFP tag to generate

recombinant expression vectors, respectively. The recombinant proteins were expressed in *Escherichia coli* (BL21) or *Bm12* cells. The specific primers and restriction sites used for each protein expression are shown in Table S9.

For purification of His-tagged protein, the transformed *E. coli* cells were collected by centrifugation and resuspended in binding buffer (0.5 M NaCl, 20 mM Tris-HCl, 5 mM imidazole, pH 7.9 and 1 mM PMSF). The suspension was centrifuged after being lysed by sonication and then purified with Ni-chelating affinity chromatography using His-Bind® 12 Kit according to the manufacturer's protocol (Novagen, Wisconsin, USA).

For purification of GST-tagged protein, the transformed *E. coli* cells were collected by centrifugation and resuspended in PBS. The suspension was centrifuged after being lysed by sonication and then purified using a GST Protein Purification Kit according to the manufacturer's protocol (Beyotime, Shanghai, China).

Nuclear protein preparation and electrophoretic mobility shift assay (EMSA)

Nuclear proteins were extracted from tissues or cells according to the instructions of NE-PER Nuclear and Cytoplasmic Extraction Kit (Thermo Scientific, Massachusetts, USA). EMSA was conducted using LightShift Chemiluminescent EMSA Kit (Thermo Scientific, Massachusetts, USA). Wild-type oligonucleotides labeled with biotin at the 5' end or methylated at cytosine were synthesized by Invitrogen (Shanghai, China). Oligonucleotide probes were heated at 95°C for 10 min in 50 mM Tris-acetate buffer at pH 4.1 and then slowly cooled to room temperature. Binding assays were performed according to the manufacturer's protocol. Briefly,

nuclear protein extracts or purified proteins were incubated with 20 μL of binding buffer containing 50 ng of poly (dI-dC), 2.5% glycerol, 0.05% NP-40, 50 mM potassium chloride, 5 mM magnesium chloride, 4 mM EDTA and 0.1 pmol of a methylated and biotinylated end-labeled double-stranded probe. Cold probes (unlabeled) at different concentrations were added to the binding mixture as competitor. For a mutant probe, 20 fmol of a nonmethylated but biotinylated end-labeled double-stranded probe was used. Two micrograms of MBD2/3 antiserum or 2 μg of normal rabbit IgG (control) was added to detect supershifted bands. After electrophoresis and transfer to nylon membranes, the protein bands were visualized by using LightShift Chemiluminescent EMSA Kit according to the manufacturer's protocol.

GST pull-down assays

Expression of soluble BmMBD2/3-GST protein was induced in DH5α cells. Nuclear proteins of *Bm12* cells overexpressing BmTip60-EGFP proteins were extracted according to the instructions of NE-PER Nuclear and Cytoplasmic Extraction Reagents. BmMBD2/3-GST proteins were mixed with 40 μL of Pierce™ Glutathione Agarose (Thermo Scientific, Massachusetts, USA) and incubated at 4°C with gentle agitation for 1 h. After washing three times with PBS, 100 μg nuclear extract containing BmTip60-EGFP proteins was incubated with Glutathione Agarose-BmMBD2/3 overnight at 4°C in 500 μL PBS with gentle agitation. After incubation, the protein-protein complexes were washed three times with 600 μL PBS. The bound proteins were eluted in 30 μL SDS-PAGE sample buffer (50 mM Tris, 100

mM DTT, 2% SDS, 0.1% bromophenol blue, 10% glycerol) and subjected to 12% SDS-PAGE. Target proteins were identified by Western blot with an anti-GFP antibody (ab290, Abcam, Cambridge, UK).

Chromatin immunoprecipitation (ChIP) assay

ChIP was performed in *Bm12* cells following the instructions of Pierce™ Magnetic ChIP Kit (Thermo Scientific, Massachusetts, USA). Approximately 4 x10^6 cells were crosslinked with 1% formaldehyde for 10 min at room temperature after transfection with overexpression vectors for BmMBD2/3-3×FLAG, BmTip60-3×FLAG or EGFP (as a control) for 48 h and then decrosslinked with glycine. The cells were washed twice with PBS and collected by centrifugation at 3000 × g for 5 min. The cell pellets were broken with extraction buffer containing protease/phosphatase inhibitors. The nuclei were collected and treated with MNase diluted in MNase Digestion Buffer for 15 min at 37°C, and the nuclei were released from the cells by ultrasonic breaking with several pulses and 20 s ice-cold intervals, followed by centrifugation at 9000 × g for 5 min. The protein-DNA complexes were immunoprecipitated by using rabbit anti-FLAG antibodies (#14793, Cell Signaling Technology, MA, USA) or normal rabbit IgG (as a control) (Thermo Fisher Scientific, Massachusetts, USA). The IP reactions were incubated overnight at 4°C with constant mixing. The DNA/protein/antibody complexes were purified by incubation with ChIP Grade Protein A/G Magnetic Beads for 2 h at 4°C with mixing. DNA was crosslinked and purified using the column method according to the manufacturer's instructions

(Thermo Scientific, Massachusetts, USA) and detected by qRT-PCR (Table S9). The enriched sequence in the immunoprecipitated DNA samples was normalized to the DNA present in the 10% input. The PCR products of the enriched promoters were confirmed by sequencing. The antibodies used in this study are listed in Table S10.

RNA interference (RNAi)

For RNAi in the *Bm12* cell line, a 400~600-bp unique fragment of the ORF of target genes was chosen as the template for synthesizing gene-specific dsRNA using T7 RiboMAXTM Express RNAi System (Promega, Wisconsin, USA). dsRNA (4 μg) was used to transfect *Bm12* cells with 6 μL Fugene HD transfection reagent in Opti-MEM Reduced Serum Medium. The cells were collected at 48 h after transfection. For embryo RNAi, dsRNA (20-50 ng) was injected into embryos within 2 h of fertilization, and three replicates of 300~400 embryos per replicate were carried out. To detect the knockdown efficiency, qRT-PCR was performed using the specific primers listed in Table S9.

Western blot and far-western blot analyses

Tissues or *Bm12* cells were homogenized in Cell Lysis Buffer (Beyotime, Shanghai, China). For western blot, proteins (30~100 μg) were mixed with 5× loading buffer and then separated by 12% SDS-PAGE, followed by transfer to a nitrocellulose blot membrane. The membrane was blocked with Tris-buffered saline (pH 9.0) with 0.5% Tween-20 (TBST) containing 3% (w/v) BSA, followed by hybridization

overnight in TBST with 1% BSA and the primary antibody. After the membranes were washed in TBST three times, a horseradish peroxidase (HRP)-conjugated antibody (Dingguo Biotechnology, Beijing, China) was applied. An anti-tubulin antibody (Dingguo Biotechnology, Beijing, China) was used to verify equal loading of the protein samples on the gel. The antibodies used in this study are listed in Table S10.

For far-western blot, the purified recombinant BmTip60-His or BmMBD2/3-GST protein (3 mg) was separated by 12% SDS-PAGE and transferred to nitrocellulose membranes. The membranes were washed in TBST to remove SDS, blocked with 3% (wt/vol) BSA in TBST, and then probed with another purified protein (2 mg/mL), BmMBD2/3-GST or BmTip60-His, in TBST containing 0.1% BSA. The blots were incubated overnight; after washing, the membranes were probed with antibodies against the GST or His tag, followed by a horseradish peroxidase (HRP)-conjugated antibody.

Immunofluorescence

Bm12 cells at the logarithmic growth phase were subcultured on glass coverslips (WHB, Shanghai, China) and grown to a proper density. A mixture of 30 μL containing 0.5 μg of BmMBD2/3-EGFP, 0.5 μg of BmTip60-flag-EGFP plasmid DNA, and 3 μL Fugene HD transfection reagent in Opti-MEM Reduced Serum Medium was added, and the cells were cultured for 48 h at 28°C.

The *Bm12* cells on glass coverslips were fixed with 4% paraformaldehyde, blocked in PBS containing 5% BSA and 0.5% Triton-X (PBT) for 1-2 h, and then

incubated with an anti-flag primary antibody (Cell Signaling Technology, D6W5B) at 4°C for 3 h. The cells were washed three times with 0.2% PBT (0.2% Triton-X in PBS) and incubated with Alexa Fluor™594 goat anti-rabbit IgG (Invitrogen, California, USA) for 2 h. DAPI (Beyotime, Shanghai, China) was added to stain the nucleus. Primary and secondary antibodies were diluted 1:200 and 1:400 in PBS containing 5% BSA and 0.2% Triton-X (PBT), respectively. The cells were observed and imaged using an FV3000 confocal microscope (Olympus, Japan).

For 5mC staining, fixed tissues or cells were first incubated with 2 M HCl solution for 20 min and then neutralized with 100 mM Tris-HCl (pH 8.5) for 10 min at room temperature. After blocking with 3% BSA in 0.5% PBT buffer, the tissues or cells were incubated with an anti-5mC antibody (diluted 1:500 in 0.2% PBT buffer, Abcam, Cambridge, UK); Alexa Fluor™ 594 goat anti-rabbit IgG (diluted 1:400 in 0.2% PBT buffer, Invitrogen, California, USA) was used as the secondary antibody. The nuclei were stained with DAPI (Beyotime, Shanghai, China) for 10 min. Wing tissues stained with anti-5mC and DAPI were observed and imaged using an FV3000 confocal microscope (Olympus). The antibodies used in this study are listed in Table S10.

Bioinformatics analysis of protein sequences

Phylogenetic tree analysis was conducted using MEGA7.0. The protein sequences were aligned with ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/), and the protein domains were analyzed with SMART (http://www.smart.embl-heidelberg.de/). The phylogenetic tree was edited with Evolview (http://www.evolgenius.info/).

Egg number and hatch rate statistics

To record egg number, *B. mori* and *S. litura* were injected with 5-aza-dC on the first day of pupation, and *L. migratoria* was injected with 5-aza-dC on the first day of adulthood. Ovaries of *B. mori*, *S. litura* or *L. migratoria* were dissected from adult females on the first, third or tenth day. Eggs were collected from females treated with 5-aza-dC or ddH$_2$O (control) and counted.

To calculate the hatch rate, dsRNA (20-50 ng) was injected into *B. mori, S. litura* and *L. migratoria* embryos within 2 h after fertilization. The larvae hatched from the resulting eggs of *B. mori* and *L. migratoria* were recorded at 9-13 days after dsRNA injection; the hatched larvae of *S. litura* were recorded at 3-5 days after dsRNA injection. The hatch rate was calculated as follows: hatch rate=larval number/egg number.

Statistical analysis

Data are presented as the mean ± SEM. "n" represents the number of biological replicates. P values for the purpose of group comparisons were calculated using ANOVA (*P < 0.05, **P < 0.01, ***P < 0.001).

Supplemental Information
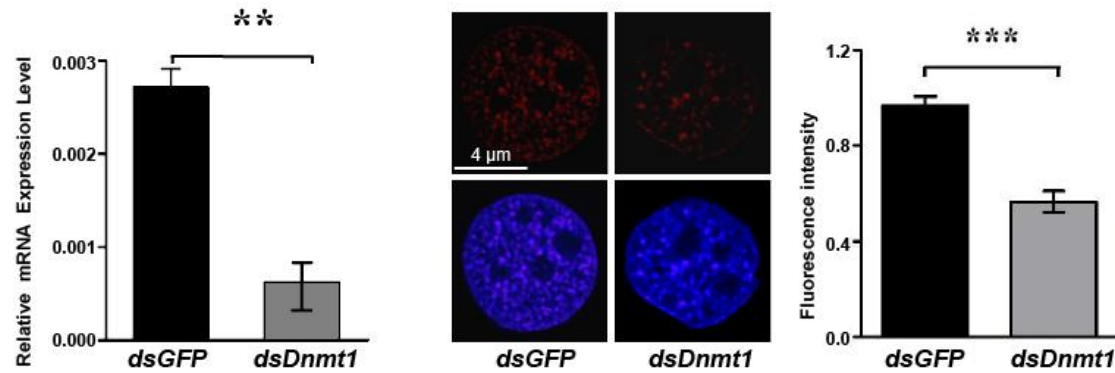
Figure S1. 5mC was detected by immunofluorescent staining in the nuclei of *B. mori* ovarian cell line (Bm12) after *BmDnmt1* RNAi. Related to Figure 1. Relative mRNA levels of *BmDnmt1* after *dsDnmt1* or *dsGFP* (control) was transfected into Bm12 cells. The mRNA levels of *BmDnmt1* was normalized to the expression level of the house-keeping gene, ribosomal protein 49 (*RP49*) (n=3) (left). The immunofluorescent staining (middle) (Scale bar, 4 µm); the fluorescence intensity performed using ImageJ software (right). Data are presented as the mean ± SEM. **$P < 0.01$, ***$P < 0.001$ by ANOVA.
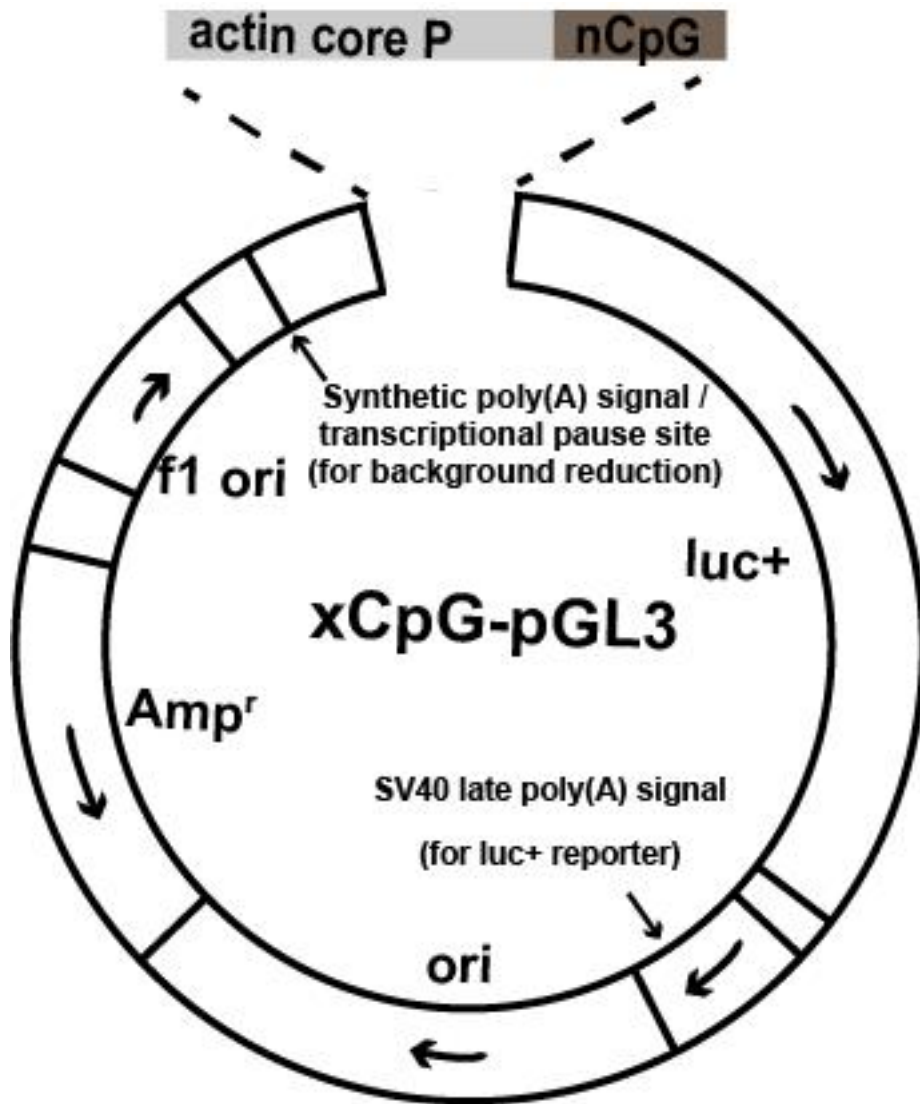
Figure S2. The vector for luciferase activity determination was constructed with 0, 2, 4 and 6 repeats of methyl-CpG fragment (ATCGAT) between the luciferase gene body and actin core promoter. Related to Figure 2. The sequences of the actin core promoter and 5'UTR are shown in Table S2.
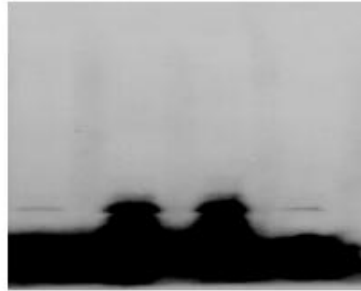
Figure S3. EMSA shows that the purified BmMBD4 did not bind with the methyl-CpG probes. Related to Figure 3. The sequences of the probes in the study are shown in Table S9.
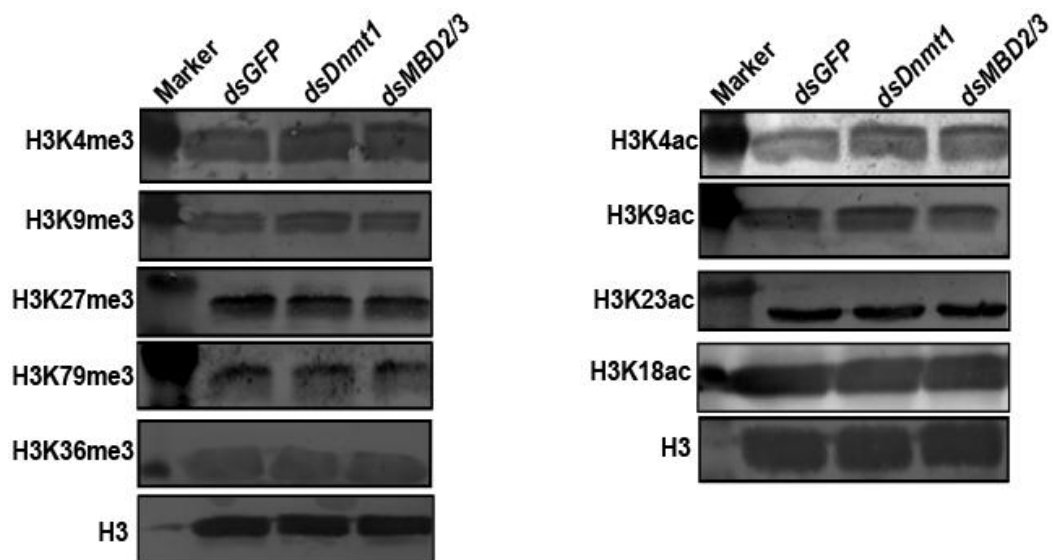
Figure S4. Western blot analyses of histone modification levels after the knockdown of *BmDnmt1* or *BmMBD2/3* in *B. mori* Bm12 cells. Related to Figure 3. Left: H3K4me3, H3k9me3, H3K27me3, H3K79me3, H3K36me3; Right: H3K4ac, H3K9ac, H3K23ac, H3K18ac. Marker: 15 kDa.
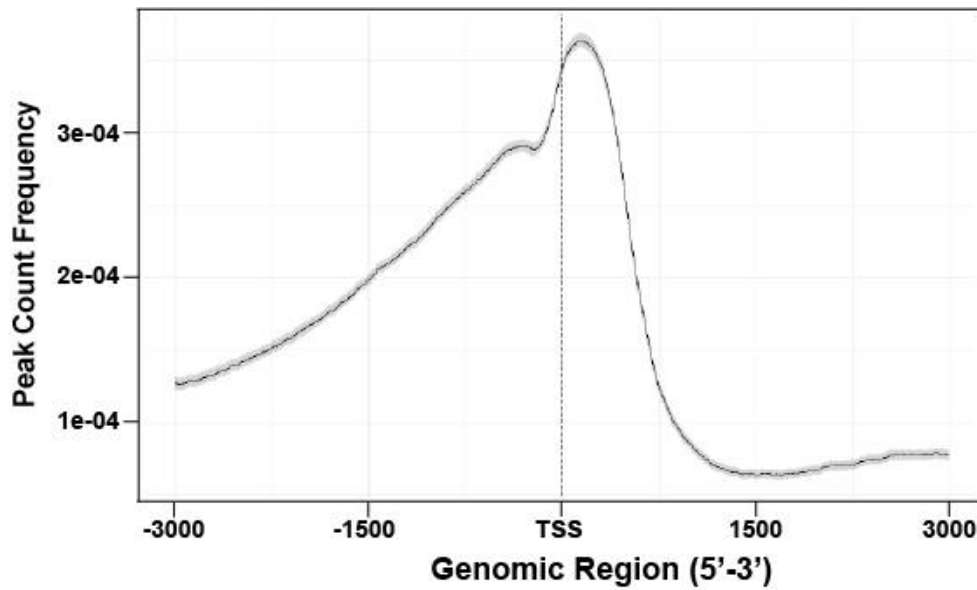
Figure S5. Distribution of H3K27ac in the genome by ChIP-seq using H3K27ac antibody. Related to Figure 3.
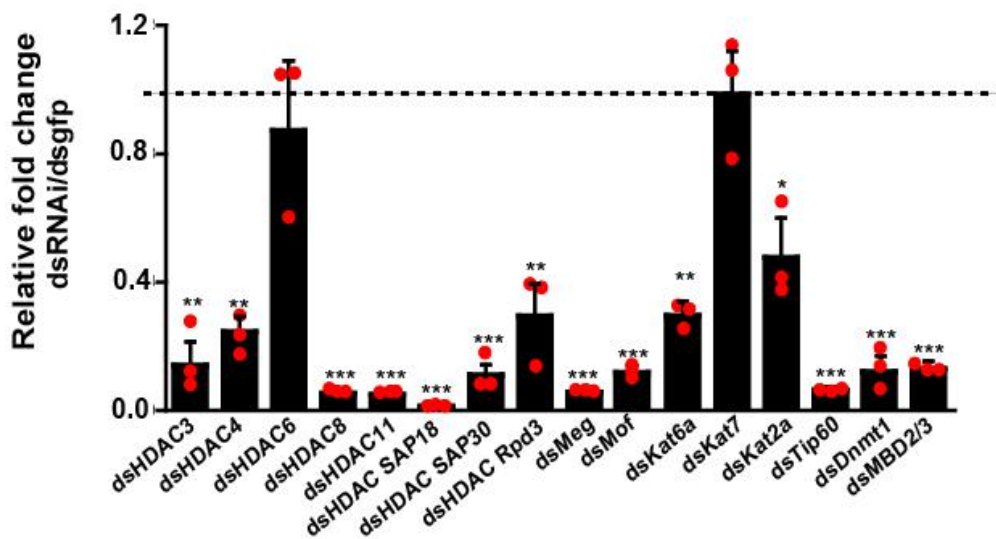
Figure S6. qPCR analyses of RNAi effectiveness in the Bm12 cells. Related to Figure 3. Relative mRNA levels are presented as the mRNA level in dsRNA-transfected samples divided by the control (*dsgfp*), which was normalized to the expression level of the house-keeping gene, ribosomal protein 49 (*RP49*). Genes

include *Dnmt1*, *MBD2/3*, *acetyltransferases (HATs)* and *deacetylases* (*HDACs*) (n=3).

"n" represents the number of biological replicates. Data are presented as mean ± SEM.

*P < 0.05, **P < 0.01, ***P < 0.001 by ANOVA analysis.
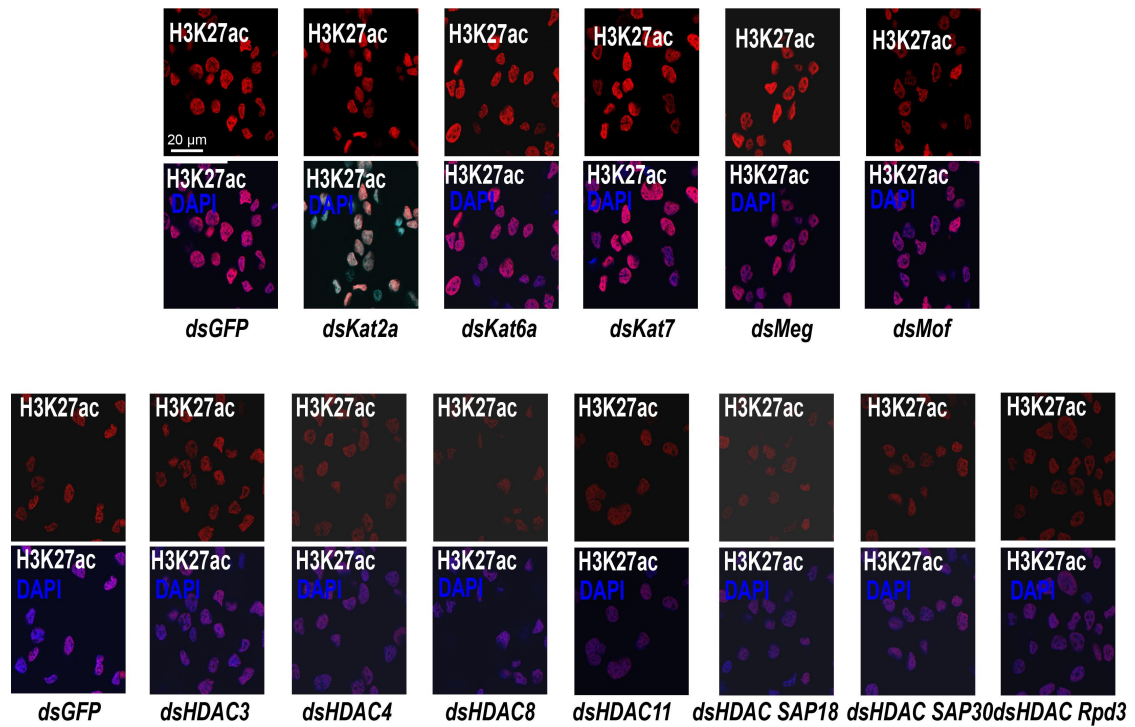
Figure S7. Immunofluorescence analyses of the H3K27ac levels after the knockdown of HATs (a-f) or HDACs (g-n) in the Bm12 cells. Related to Figure 3.

Blue: DAPI, red: H3K27ac (Scale bar, 20 μm).

Figure S8. Conservative analysis of coiled-coil domain of MBD2/3 or MBD2/3 in different species. Related to Figure 4. The MBD2/3 protein contains two domains, methyl-CpG binding domain (MBD, specially binding to mCpGs) and a coiled-coil domain (interacting with other proteins).
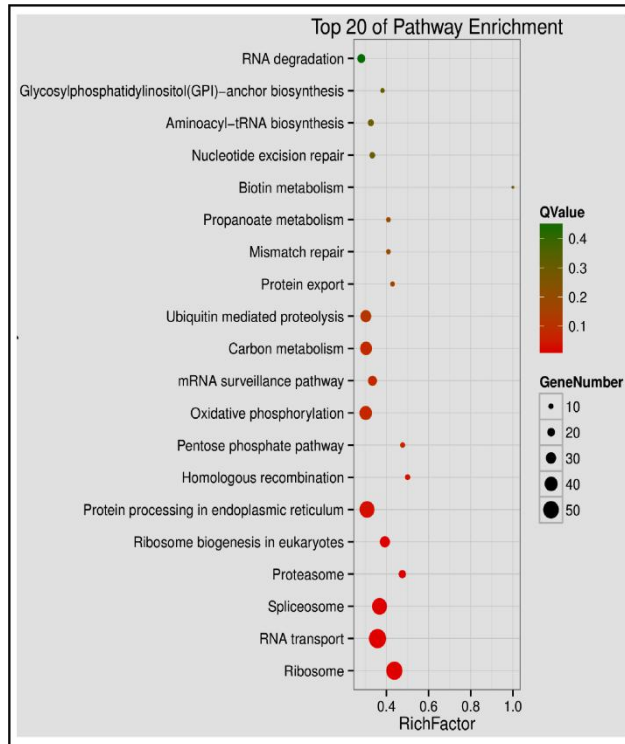
Figure S9. Top 20 pathway enrichment of genes with high DNA methylation rate in the genome of *B. mori* ovary. Related to Figure 6.

Table S1. DNA methylation levels in different regions of the genome. Related to Figure 1. mCs are divided into three categories according to their sequence characteristics: mCG, mCHG and mCHH (H represents A, C or T). The methylation level was calculated as the total number of mCs divided by the total number of Cs, CGs, CHGs and CHHs in the whole genome.

Sample	Region	C(%)	CG(%)	CHG(%)	CHH(%)
	Genome	0.52	0.81	0.42	0.45
	Genebody	0.78	2.24	0.43	0.46
Ovary-1	Exon	1.02	3.24	0.43	0.45
	Intron	0.71	1.9	0.43	0.46
	CDS	1.02	3.24	0.43	0.45
	Genome	0.63	0.96	0.51	0.54
	Genebody	0.93	2.72	0.52	0.55
Ovary-1	Exon	1.23	4	0.52	0.54
	Intron	0.85	2.32	0.52	0.56
	CDS	1.23	4	0.52	0.54

Table S5. Reads and methylated sites number statistics in the whole genome by BS-seq analysis. Related to Figure 1 and 2. mCs are divided into three categories according to their sequence characteristics: mCG, mCHG and mCHH (H represents A, C or T). The methylation proportion was calculated as the number of mCGs, mCHGs and mCHHs divided by the total number of mCs in whole genome.

Sample	Total Reads	Mapped Reads	Mapped Ratio(%)	Sequence Depth
Ovary Bisulfite Sequence-1	227,885,236	196,463,231	86.21	51.08
Ovary Bisulfite Sequence-2	201,184,054	171,451,740	85.22	44.58

Pattern	mCG	mCHG	mCHH
Ovary-1 Number	303,813	10,636	48,939
Ovary-1 Proportion	83.61%	2.93%	13.47%
Ovary-2 Number	333,953	17,738	82,309
Ovary-2 Proportion	76.95%	4.09%	18.97%

Table S6. Identification and description of the hypermethylated genes studied in

Figure 2. Related to Figure 2.

Gene	Chromosome	Methylation level(%)	Description
BGIBMGA013120	nscaf3062	11.95888957	Arp2/3 complex subunit [Bombyx mori]
BGIBMGA013896	nscaf3099	12.1352089	PREDICTED: exosome complex component RRP41 [Bombhyx mori]
BGIBMGA005486	nsca2828	10.07141628	PREDICTED: histone RNA haipinbiding protein [Bombyx mori]
BGIBMGA006154	nscaf2847	10.84927891	mannose-P-dolichol uiliztion defect 1 protein [Bombyx mori]
BGIBMGA001054	nscaf1898	11.84594721	ubiquitin carboxyl-teminal hydrolase 46 [Bombyx mori]
BGIBMGA009845	nsca2970	10.2865735	NADH dehydrogenase isoform 2 [Bombyx mori]
BGIBMGA009698	nscaf2964	10.20651311	nascent polypeptide associated complex protein alpha subunit [Bombyx mori]
BGIBMGA001105	nscaf1898	12.64991839	PREDICTED: nitrie oxide synthase-interacting protein homolog [Bombyx mori]
BGIBMGA007945	nsca2888	10.61053109	PREDICTED: pre-mRNA-splicing factor 38B [Bombyx mori]

BGIBMGA010975	nscaf3013	12.16418691	PREDICTED: protein FRGI homolog [Bombyx mori]
BGIBMGA009840	nscaf2970	11.90055899	PREDICTED: protein lin-37 homolog [Bombyx mori]
BGIBMGA001106	nscaf1898	11.6113326	ribosomal pronein L21 [Bombyx mori]
BGIBMGA002610	nscaf2529	10.09892159	PREDICTED: 39s ribosomal protein L42, mitochondrial [Bombyx mori]
BGIBMGA006330	nscaf2852	11.74673515	PREDICTED: serine/threonine-protein kinase Warts [Bombyx mori]

Table S7. The sequences of the core promoter and 5'UTR of actin gene. The red represents the TATA box and transcription start site (TSS); The blue represents the 5'UTR. Related to Figure 2.

actin core promoter + 5'UTR	aataatacgcgaatgatgataacgtgttacgttacataatcgttgcataactagt gaagtgaaattttttataaaaaaaaacagtttcggaattagtgtaatgccgttgct actaaataagaaataagtttattgaacgcacatttcaaaatg

Table S8. List of histone deacetylases (HDACs) (left) and histone acetyltransferases (HATs) (right) in the silkworm genome. Related to Figure 3.

Histone deacetyase(HDAC)	
Protein introduction	**Accession number**
HDAC3	XM_012697024
HDAC4	XM_021346265
HDAC6	XM_012688445
HDAC8	XM_021349792
HDAC11	XM_004925365
HDAC SAP18	XM_004921692
HDAC SAP30	XM_004927722
HDAC Rpd3	XM_004931383.3
Histone acetyltransferase(HAT)	
Protein introduction	**Accession number**
HAT Tip60	XP_004928297
HAT MOF protein	NP_001093305
CREB-binding protein	XP_021204790
HAT KAT2A	XP_004922629
HAT KAT7	XP_021205742
HAT KAT6A	XP_021202172
Meg protein 3	XP_004928644

Table S10. The list of antibodies used in this study. Related to Figure 1, 3, 4 and 5.

Antibody name	Product code
Anti-5-methylcytosine (5-mC)	ab10805
Anti-histone H3 (tri methyl K27)	ab6002
Anti-histone H3 (tri methyl K36)	ab9050
Anti-histone H3 (tri methyl K79)	ab2621
Anti-histone H3 (acetyl K27)	ab4729
Anti-histone H3 (acetyl K23)	ab177275
Anti-histone H3 (acetyl K4)	ab176799
Anti-histone H3 (acetyl K9)	ab4441
Anti-histone H3 (acetyl K18)	ab1191
Anti-histone H3	ab1791
Anti-EGFP	ab290
Anti-FLAG Antibody	D6W5B