SEE COMMENTARY

Arabidopsis EDM2 promotes *IBM1* distal polyadenylation and regulates genome DNA methylation patterns

Mingguang Lei[a,1], Honggui La[b,1], Kun Lu[a,c], Pengcheng Wang[a], Daisuke Miki[b], Zhizhong Ren[b], Cheng-Guo Duan[a], Xingang Wang[a], Kai Tang[a], Liang Zeng[b], Lan Yang[b], Heng Zhang[b], Wenfeng Nie[a,d], Pan Liu[b], Jianping Zhou[a,e], Renyi Liu[b], Yingli Zhong[a,f], Dong Liu[g], and Jian-Kang Zhu[a,b,2]

[a]Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN 47906; [b]Shanghai Center for Plant Stress Biology, Shanghai Institute for Biological Science, Chinese Academy of Sciences, Shanghai 200032, China; [c]College of Agronomy and Biotechnology, Southwest University, Chongqing 400715, China; [d]Department of Horticulture, Zhejiang University, Hangzhou 310058, China; [e]School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, Sichuan 610054, China; [f]College of Bioscience and Biotechnology, Hunan Agriculture University, Changsha, Hunan 410128, China; and [g]Ministry of Education Key Laboratory of Bioinformatics, School of Life Sciences, Tsinghua University, Beijing, 100084, China

Contributed by Jian-Kang Zhu, October 28, 2013 (sent for review October 12, 2013)

DNA methylation is important for the silencing of transposons and other repetitive elements in many higher eukaryotes. However, plant and mammalian genomes have evolved to contain repetitive elements near or inside their genes. How these genes are kept from being silenced by DNA methylation is not well understood. A forward genetics screen led to the identification of the putative chromatin regulator Enhanced Downy Mildew 2 (EDM2) as a cellular antisilencing factor and regulator of genome DNA methylation patterns. EDM2 contains a composite Plant Homeo Domain that recognizes both active and repressive histone methylation marks at the intronic repeat elements in genes such as the Histone 3 lysine 9 demethylase gene Increase in *BONSAI* Methylation 1 (*IBM1*) and is necessary for maintaining the expression of these genes by promoting mRNA distal polyadenylation. Because of its role in maintaining *IBM1* expression, EDM2 is required for preventing CHG methylation in the bodies of thousands of genes. Our results thus increase the understanding of antisilencing, genome methylation patterns, and regulation of alternative RNA processing by intronic heterochromatin.

transcriptional gene silencing | gene body methylation | post-transcription regulation | alternative polyadenylation | DNA methylome

DNA methylation is a conserved epigenetic mark in plants and many animals (1–5). Unlike mammals, in which DNA methylation occurs mainly in the CG sequence context, plants have CHG (H is A, T, or C) and CHH, as well as CG methylation (1, 2). In *Arabidopsis*, de novo DNA methylation can be catalyzed by domain rearranged methyltransferase 2 (DRM2) in a pathway known as RNA-directed DNA methylation (RdDM) (2, 6, 7). CG methylation is maintained by DNA methyltransferase 1 (MET1), an ortholog of mammalian DNMT1, and CHG methylation is maintained by the plant-specific chromomethylase CMT3 (8, 9). The asymmetrical CHH methylation cannot be maintained and must occur de novo in every cell cycle through RdDM (10). The RdDM pathway silences thousands of transposable elements (TEs) and some transgenes as well as endogenous genes (11–14).

The genomes of higher eukaryotes harbor numerous transposons and other repeat elements that are subjected to silencing by DNA methylation and repressive histone modifications (15). However, some of the repeat elements are close to or even inside the promoters or other *cis*-elements of genes that need to be expressed during development or in response to environmental cues (16). In addition, some repeat elements and associated heterochromatin have found residence in the introns of genes (17, 18). Presumably, these genes are prone to be silenced because of potential collateral effects of the silencing of the neighboring or genic repeat elements. It is unclear how these genes are kept active when their repeat elements are targeted by DNA methylation and related chromatin silencing pathways. Although Repressor of Silencing 1 (ROS1)-mediated active DNA demethylation is known to be important for preventing the silencing of some of these genes, the mechanisms of keeping genes active and properly expressed appear to be complex. Besides repressing transcription initiation, DNA methylation and heterochromatic histone modifications can also impede transcription elongation and affect alternative splicing or polyadenylation (19–23).

We devised a simple forward genetics screen in *Arabidopsis* to identify antisilencing factors. Our genetic screen led to the identification of Enhanced Downy Mildew 2 (EDM2) as an antisilencing factor. Besides inhibiting promoter DNA methylation to keep transgenes and some endogenous genes active, we discovered that EDM2 also has a role in preventing CHG methylation in thousands of gene bodies, a role that is very similar to

Significance

The histone 3 lysine 9 demethylase Increase in *BONSAI* Methylation 1 (IBM1) is critical for preventing CHG (H=A, T, or C) methylation in plant genes. We found that the putative chromatin regulator Enhanced Downy Mildew 2 (EDM2) controls genome CHG methylation and transgene silencing by regulating alternative polyadenylation of *IBM1*. EDM2 contains a composite plant homeo domain that simultaneously recognizes both repressive and active histone marks, allowing it to associate specifically with the intronic heterochromatin of *IBM1* and other plant genes so that it may affect polyadenylation by methylating the RNA through an N6-adenine methyltransferase-like activity. This work significantly advances our understanding of the regulation of DNA methylome and mechanisms of antisilencing and alternative polyadenylation of intronic heterochromatin-containing genes.

Author contributions: M.L., H.L., D.L., and J.-K.Z. designed research; M.L., H.L., K.L., P.W., D.M., Z.R., C.-G.D., X.W., L.Y., H.Z., W.N., P.L., J.Z., and Y.Z. performed research; M.L., H.L., K.T., L.Z., H.Z., R.L., and J.-K.Z. analyzed data; and M.L., H.L., and J.-K.Z. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The raw data of mRNA-seq of *edm2-4*, WT, and bisulfite sequencing of Col-0, *edm2-4*, and *ibm1-4* have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE51764). The bisulfite sequencing of *ibm1-4* has also been deposited in the GEO database (accession no. GSE48053).

See Commentary on page 9.

[1]M.L. and H.L. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: jkzhu@purdue.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1320106110/-/DCSupplemental.

PLANT BIOLOGY

that of Increase in *BONSAI* Methylation 1 (IBM1). We found that EDM2 recognizes methylated Histone 3 lysine 9 (H3K9) and other histone marks in the large, heterochromatin-containing intron of *IBM1* and promotes 3′ distal polyadenylation of *IBM1* transcripts. Other genes with large transposon- or other heterochromatic repeat-containing introns also require EDM2 for 3′ distal polyadenylation and proper expression. This work thus significantly advances our understanding of the regulation of the DNA methylome and antisilencing mechanisms in plants.

Results

EDM2 Is an Antisilencing Factor and Prevents DNA Hypermethylation. To identify cellular factors required for the prevention of transcriptional gene silencing, we developed an efficient transgene-based genetic screen in *Arabidopsis*. When the sucrose transporter 2 gene (*SUC2*) driven by the constitutive cauliflower mosaic virus *35S* promoter (*35S-SUC2*) is expressed in *Arabidopsis* plants, the transgenic plants overaccumulate sucrose, resulting in severe root growth inhibition in sucrose-containing culture media (24). The *35S-SUC2* transgene construct also contains another transgene, *HPTII*, which is driven by the *2x35S* promoter (*2x35S-HPTII*) and which served as a selectable marker for plant transformation (Fig. 1*A*). We mutagenized the *35S-SUC2* transgenic plants (hereafter referred to as the WT) with ethyl methanesulfonate (EMS). M2 seedlings were screened on half-strength Murashige and Skoog (MS) nutrient agar medium containing 2% (wt/vol) sucrose. The screen led to the isolation of the *asi1* mutant, as well as unique alleles of *ros1* and *idm1* (25). *ASI1* and *IDM1* encode an RNA-binding protein and a histone acetyltransferase, respectively (25, 26). In this study, we report three alleles of *EDM2* (Fig. 1 *B* and *C*), which were known to regulate the disease resistance (*R*) gene *RPP7* (27). The three mutants were referred to as *edm2-6*, *edm2-7*, and *edm2-8*. In *edm2-6*, a point mutation (from CAG to TAG) changes Gln-563 to a premature stop, whereas GGA was mutated to GAA, changing Gly-804 to Glu in *edm2-7*, and CCT was mutated to CTT, changing Pro-861 to Leu in *edm2-8* (Fig. 1*C* and Fig. S1*A*). EDM2 was

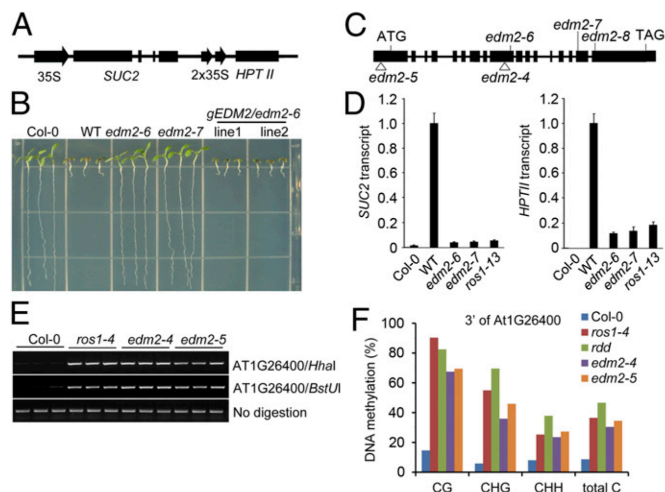

Fig. 1. *edm2* mutant phenotypes. (*A*) Schematic diagram of the construct harboring the *35S-SUC2* and *2x35S-HPTII* cassettes that were transformed in the parental line. Black rectangles represent exons. *SUC2*, sucrose transporter 2 gene; *35S*, cauliflower mosaic virus 35S promoter; *HTPII*, hygromycin phosphotransferase II used for selecting transgenic plants. (*B*) Comparison of root growth among Col-0, WT, two *edm2* mutants, and *edm2-6* transformed with a genomic fragment of *EDM2* (*EDM2* gDNA) driven by its native promoter. (*C*) Diagram showing the mutation sites in *EDM2*. (*D*) Expression of *SUC2* and *HPTII* in *edm2* mutants. Values are means ± SD of three biological replicates where the fold changes are normalized to transcript levels in WT. (*E* and *F*) Analysis of DNA methylation levels at the 3′ region of AT1G26400 in *edm2* mutants by chop PCR (*E*) and by genomic bisulfite sequencing analysis (*F*).

predicted to contain an N-terminal RFD (replication foci domain) module, a composite Plant Homeo domain with three tandem PHD-type zinc finger motifs, and a putative N6-adenine methyltransferase-like domain that was found to be highly conserved among EDM2 orthologs from both monocotyledonous and dicotyledonous species (Fig. S1 *A–C*). In the *edm2* mutant plants, *35S-SUC2* and *2x35S-HPTII* transgenes were silenced (Fig. 1*D*), and the silencing was accompanied by an increased DNA methylation at the *35S* promoter (Fig. S2*A*). Consistent with the notion that increased DNA methylation was responsible for the silencing of the transgenes, expression of the *35S-SUC2* and *2x35S-HPTII* transgenes was recovered in *edm2* mutants after treatment with 20 μM 5-aza-2′-deoxycytidine, a DNA methylation inhibitor (Fig. S2*B*). To further confirm that the *edm2* mutations were responsible for the silencing of *35S-SUC2*, we transformed *EDM2* genomic DNA driven by its native promoter into *edm2-6* and found that the *edm2* mutant phenotypes were rescued (Fig. 1*B*). These results show that EDM2 is required for preventing the TGS of the transgenes.

In parallel to the *35S-SUC2* transgene–based screen for antisilencing factors, we also carried out a genetic screen for factors involved in DNA demethylation based on chop PCR analysis of DNA methylation status at the 3′ region of At1g26400, which is known to be hypermethylated in *ros1* and *rdd* (*ros1*/*dml2*/*dml3*) mutants (26). An earlier screen using this chop PCR marker led to the identification of IDM1 (26). In the current study, our further screen of homozygous SALK T-DNA lines of *Arabidopsis* detected two *edm2* T-DNA knockout alleles, *edm2-4* and *edm2-5* (27) (Fig. 1*C* and Fig. S1*A*), that showed hypermethylation at the 3′ region of At1g26400 (Fig. 1*E*). Genomic bisulfite sequencing revealed that DNA methylation levels at the 3′ region of At1g26400 in the two *edm2* mutants were substantially increased, and the increases were comparable to that in *ros1-4* mutant plants (Fig. 1*F*).

EDM2 Prevents Genomewide CHG Methylation in Gene Bodies. To understand the extent of DNA hypermethylation in *edm2* mutant plants, we determined the genomewide DNA methylation profile of *edm2-4* by bisulfite sequencing and compared that profile to those of Col-0 and *rdd* mutant plants (26). There was a large increase in methylated CHG in *edm2-4* (12.5%) compared with Col-0 (6.3%) and *rdd* (7.0%) at the genomewide level. The proportion of CHG methylation in total cytosine methylation was dramatically higher in *edm2-4* (30.5%) than in Col-0 (19.8%) and *rdd* (19.8%) (Fig. S3*A*). Mutations in the histone H3K9 demethylase gene *IBM1* are known to cause very high CHG methylation (28, 29), and whole-genome bisulfite sequencing revealed that the *ibm1-4* mutant also exhibited a very high level (33.4%) of CHG methylation (25).

We identified 6,216, 9,791, and 4,118 hypermethylated regions (hyper DMRs) in *edm2-4*, *ibm1-4*, and *rdd*, respectively, where DNA methylation levels were significantly higher than in the Col-0. The same analysis found only 234, 305, and 426 hypomethylated regions (hypo DMRs) in *edm2-4*, *ibm1-4*, and *rdd*, respectively. Interestingly, more than 95% of the hypermethylated loci in *edm2-4* overlap with genes, which is similar to the case in *ibm1-4*, whereas hypermethylated loci in *rdd* appear to be more evenly distributed on genes, TEs, and intergenic regions (Fig. 2*A*). We analyzed the methylation levels of genes and TEs and regions 2 kb upstream or downstream of the genes or TEs. Like *ibm1*, the *edm2* mutation caused a significant increase in CHG but not CG or CHH methylation in gene bodies (Fig. 2*B*). The *edm2* and *ibm1* mutations did not cause a substantial increase in TE DNA methylation in any sequence context (Fig. 2*C*). We analyzed the effect of *edm2* mutation on genes of different sizes and found that, like *ibm1-4*, *edm2*-induced gene body methylation was preferentially higher on longer genes (Fig. S3*B*). As gene size increased, much more CHG and also slightly more CHH methylation in gene bodies was observed in *edm2-4* and *ibm1-4* mutants than in the Col-0 (Fig. S3*C*). As distinct from genic regions, TEs in *edm2-4* and *ibm1-4* mutants showed reduced CHG and CHH methylation levels regardless of their size (Fig. S3*D*).

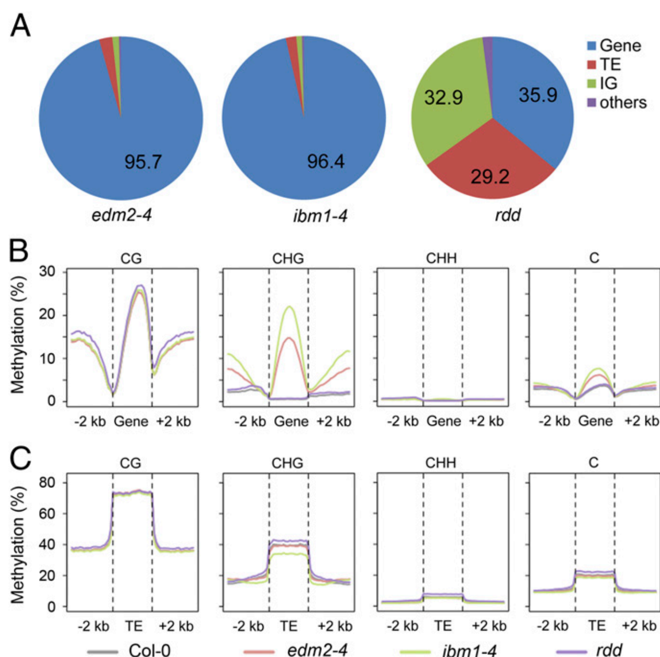

Fig. 2. Analysis of the DNA methylome by whole-genome bisulfite sequencing in Col-0 and in *edm2-4*, *ibm1-4*, and *rdd* mutants. (*A*) Composition of the hypermethylated loci in *edm2-4, ibm1-4*, and *rdd* mutants based on the distribution of hyper DMRs. TE, transposable element; IG, intergenic region. (*B* and *C*) Distribution of DNA methylation along protein-coding genes (*B*) and TEs (*C*).

In contrast, *rdd*-induced hypermethylation was mainly evident on shorter genes and TEs (Fig. S3 *C* and *D*).

There were 5,569 hypermethylated loci overlapping between *edm2-4* and *ibm1-4*, which accounted for 90% of the total hypermethylated loci in *edm2-4* (Fig. 3*A*). As expected, the overlapped hyper DMRs between *edm2-4* and *ibm1-4* showed striking increases in CHG methylation, slight increases in CHH methylation, and no significant increases in CG methylation compared with the Col-0 (Fig. 3*B* and Fig. S4*A*). Even the group of hyper DMRs that were categorized as specifically affected by *ibm1* also showed slight increases in CHG methylation in *edm2-4* (Fig. S4*A*). We developed chop PCR markers for 14 of the hyper DMRs shared by *edm2-4* and *ibm1-4*, including the *BONSAI* locus, to validate the methylome data. The results confirmed that all 14 loci were hypermethylated in the two mutants (Fig. S4*C*). Several of the loci were chosen for further analysis of DNA methylation levels in different sequence contexts, and the data showed that all of the tested loci had substantial increases in CHG but not in CG or CHH methylation (Fig. S4*D*). Chop PCR assays also showed that expression of 3× MYC-tagged *EDM2* genomic DNA under its native promoter complemented the DNA hypermethylation phenotype of the *edm2-4* mutant (Fig. S4*B*).

EDM2 Also Prevents Hypermethylation in All Sequence Contexts at Some Loci. The *rdd* triple mutant had 306 and 313 hyper DMRs that overlapped with those in *edm2-4* and *ibm1-4*, respectively. The overlapped hyper DMRs between *edm2-4* and *rdd* exhibited increases in methylation at all three sequence contexts, i.e., in CG, CHG, and CHH (Fig. 3*B* and Fig. S4*A*). These overlapped DMRs are mainly enriched in intergenic regions rather than in gene bodies (Fig. 3*B* and Fig. S4*E*). These results suggest that in addition to affecting gene body CHG methylation, the *edm2* mutation may affect 5-methylcytosine DNA glycosylase-mediated active DNA demethylation at some loci. This inference is consistent with EDM2, together with ROS1 and IDM1, playing a critical role in inhibiting the TGS of the *35S-SUC2* transgene and in preventing DNA hypermethylation at the *35S* promoter and 3′ region of At1g26400 (Fig. 1 *E* and *F* and Fig. S2). The small number of overlapped hyper DMRs between *ibm1-4* and *rdd* displayed increases mostly in CG methylation in the two mutants (Fig. S4*A*). The 182 hyper DMRs shared by all three genotypes also showed increases in CG methylation (Fig. 3*A* and Fig. S4*A*).

EDM2 Controls Gene Body CHG Methylation and Plant Development and Prevents Transgene Silencing by Regulating the Expression of *IBM1*. That both *edm2* and *ibm1* caused gene body CHG methylation at thousands of genes prompted us to explore the relationship between EDM2 and IBM1. *IBM1* is known to encode two variant transcripts: *IBM-L* and *IBM-S* (18). Only the longer variant (*IBM-L*) encodes a functional IBM1 protein that contains the jmjC domain (18). Northern blot analysis revealed that the expression level of *IBM-L* but not of *IBM-S* was decreased substantially in *edm2* mutant plants (Fig. 4*A*). We carried out 3′ RACE experiments and found that the WT plants had a high level of long transcript that was polyadenylated at the 3′ distal end of the *IBM1* gene, as well as a substantial amount of short transcript that was polyadenylated and ended inside the large seventh intron (Fig. 4*B* and Table S1). In the *edm2* mutant, the long transcript was almost absent but there was a high level of short transcript that ended inside the large intron (Fig. 4*B*). This result was further confirmed by qRT-PCR using primers specific for the short and long versions of the *IBM1* transcripts (Fig. 4*C*).

A heterochromatic region corresponding to a foreign piece of DNA, a portion of the chloroplast/mitochondrial *YCF1* gene, is located inside the large seventh intron of the *IBM1* gene (30), and it was recently reported that DNA methylation as well as histone H3K9me2 at this intron is required to maintain the expression of *IBM1-L* (18). Our DNA methylome data showed that a region of the large intron of the *IBM1* gene was as heavily methylated in the *edm2-4* and *ibm1-4* mutants as it was in Col-0 (Fig. S5*A*), suggesting that the misregulation of *IBM1* expression in *edm2* was not due to an altered DNA methylation at this intron. We hypothesized that this intronic, methylated, heterochromatic repeat region may be important for EDM2 regulation of *IBM1* expression. To test this hypothesis, we introduced an *IBM1* gene with a 764-bp deletion of the intronic heterochromatin (P_{IBM1}-*gIBM1-ΔIH-3xFLAG*) into *edm2-4* mutant plants. A construct carrying a WT version of the *IBM1* gene (P_{IBM1}-*gIBM1-3xFLAG*) was transformed in parallel into *edm2-4* to serve as a control (Fig. S5*B*). In P_{IBM1}-*gIBM1-ΔIH-3xFLAG* but not P_{IBM1}-*gIBM1-3xFLAG* transformed lines, *IBM1-L* expression (Fig. 4*C*) and *IBM1* 3′ transcript (Fig. S5*C*) was restored. As revealed by chop PCR assays, the deletion version but not the WT version of the *IBM1* gene rescued the DNA hypermethylation of the tested loci including *BNS* (Fig. S5*C*). The results demonstrate that EDM2 controls *IBM1* expression to regulate gene body CHG methylation and that the *IBM1* intronic

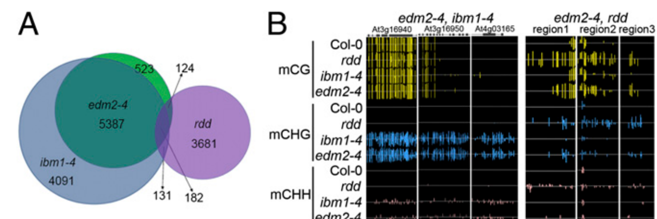

Fig. 3. DNA hypermethylation regions in *edm2-4*, *ibm1-4*, and *rdd* mutants. (*A*) Numbers of DNA hypermethylation regions that overlap among or are unique to the *edm2-4*, *ibm1-4*, and *rdd* mutants. (*B*) Screenshots of DNA hypermethylation regions in *edm2-4* overlapped with *ibm1-4* or with *rdd*. The positions of the *edm2-4* and *rdd* overlapped intergenic region are 1,599,000–1,599,600 bp on chromosome 2 (region 1), 13,361,600–13,362,300 bp on chromosome 5 (region 2), and 12,277,340–12,279,000 bp on chromosome 1 (region 3).

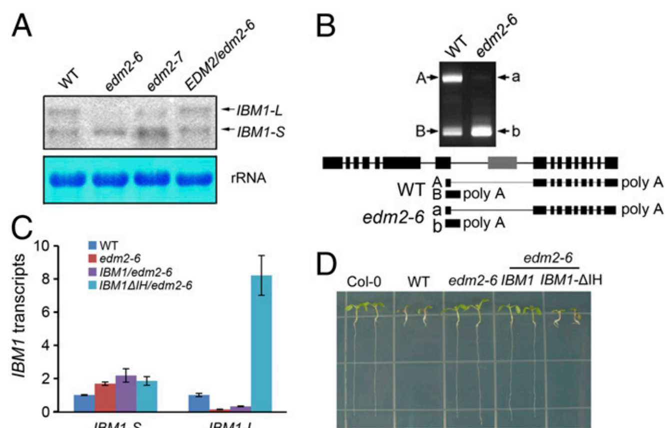

Fig. 4. EDM2 regulation of *IBM1* expression. (*A*) Northern blot analysis of the *IBM1* transcripts. rRNA was used as an RNA loading control. (*B*) 3′ RACE assay of the *IBM1* gene. The RACE PCR products are shown in the gel picture and sequenced clones are shown below the gel picture. (*C*) qRT-PCR analysis of *IBM1* gene expression with *IBM1-S*– and *IBM1-L*–specific primers. Values are means ± SD of three biological replicates where the fold changes are normalized to transcript levels in WT. (*D*) Introduction of the *IBM1-ΔIH* construct rescued *edm2-6* root growth phenotypes.

heterochromatin is necessary for such regulation. Consistent with the *IBM1* expression defect in *edm2*, *edm2* mutant plants exhibited small stature and other developmental abnormalities (31, 32) that were similar to those in *ibm1* mutants (Fig. S5*D*). The deletion version but not the WT version of the *IBM1* gene also rescued the developmental defects of *edm2* mutants (Fig. S5*E*).

In addition to rescuing the CHG methylation and developmental defects of *edm2-4* mutant plants, the deletion version of the *IBM1* construct (P_{IBM1}-*gIBM1-ΔIH*) also rescued expression defect of the *35S-SUC2* transgene (Fig. S5*F*) and suppressed the root growth phenotype of the *edm2-6* mutant (Fig. 4*D*). This result indicates that the antisilencing role of EDM2 is also mediated by IBM1. We introduced the *35S-SUC2* transgene into the *ibm1-1* mutant by crossing the WT plant to *ibm1-1*. In homozygous *ibm1-1* plants, the *35S-SUC2* transgene was silenced and thus did not confer a root phenotype (Fig. S5*G*). The results show that IBM1 has an antisilencing role. This antisilencing role of IBM1 is consistent with the above DNA methylome analysis showing that in addition to regulating gene body CHG methylation, IBM1 also regulates the DNA methylation of some other genomic regions.

EDM2 Regulates the Expression of Genes with Intronic Heterochromatin by Promoting 3′ Distal Polyadenylation. We carried out RNA-seq experiments on Col-0 and *edm2-4* mutant seedlings and identified 59 genes with intronic TEs and with a possible reduced expression of the 3′ regions downstream of the TEs (Table S2). At the top of this list were three genes (At3g05410, At1g58602/*RPP7*, and At1g11270) that contain long (>4 kb) and methylated intronic TEs (Table S2, Fig. 5, and Fig. S6). 3′-RACE assays for the At3g05410 gene revealed a major full-length polyadenylated transcript and a minor short polyadenylated transcript terminating before the intronic, methylated TE in the WT. The *edm2-6* mutant contained very little full-length transcript but an increased quantity of short transcript, as well as a new transcript that terminated inside the intronic, methylated TE (Fig. 5*A* and Table S1). The transcript pattern was similar for the At1g11270 gene, i.e., 3′-RACE results showed that the *edm2* mutant contained substantially less long transcript with polyadenylation after the intronic TE and more short transcripts with polyadenylation before the intronic TE (Fig. 5*B* and Table S1). The intronic TE in At1g11270 is methylated, although the methylation level is not very high (Fig. 5*B*). In At1g58602, DNA methylation appeared to have spread from the TE in the first intron into surrounding sequences (Fig. S6). 3′-RACE experiments detected a variety of polyadenylated At1g58602 transcripts in both the WT and *edm2* mutant plants (Fig. S6 and Table S1). Some of the transcripts ended after the methylated region, whereas some ended before or inside the methylated region. Relative to the WT, the *edm2*-6 mutant had more transcripts that were polyadenylated before or inside the methylated region (Fig. S6).

Because the long intron in *IBM1* lacks a TE but is still heavily methylated, DNA methylation or heterochromatin of sequences other than TEs may also be important for the 3′ expression in the *edm2* mutant. We therefore searched for genes with long (≥500 bp) and methylated (≥20% cytosines in a cluster as meCs) introns that did not contain TEs. This analysis led to the identification of 16 additional genes with a possible reduced 3′ expression in *edm2-4* (Table S3). It is possible that these intronic methylated sequences in *IBM1* and other genes were derived from some non–TE-type invasive repeat elements. Taken together, our results suggest that EDM2 targets genes with methylated intronic TEs or other repeat sequences to regulate their 3′ expression.

It is possible that EDM2 may facilitate RNA polymerase II (Pol II) elongation through intronic heterochromatin. We compared Pol II occupancy at the intronic heterochromatin with that in upstream and downstream regions of the heterochromatic intron in *IBM1*. However, the results revealed no substantial difference in Pol II occupancy at these sites between *edm2-4* and Col-0 (Fig. S7*A*). Similarly, we did not find substantial differences between *edm2-4* and the Col-0 in Pol II occupancy along the At1g58602 gene (Fig. S7*A*). The *edm2* mutation also did not affect Pol II occupancy in two other genes tested (At1g11270 and At3g05410; Fig. S7*A*), although these genes all showed reduced 3′ transcript levels in the *edm2* mutant (Table S2). These data suggest that EDM2 is not required for transcription elongation across the intronic TEs/heterochromatin. Together, our results show that EDM2 promotes 3′ distal polyadenylation rather than influencing transcription elongation of the genes with intronic heterochromatin.

EDM2 Recognizes Both Repressive and Active Histone Marks and Associates with Intronic Heterochromatin. We performed ChIP assays to investigate whether the intronic TEs or other repeat elements are associated with the repressive histone marks H3K9me2 and H3K27me3. H3K9me2 is a well-known mark for heterochromatin at TEs and other repeat elements, whereas H3K27me3 is associated with gene repression during development (33, 34). In the three examined genes (*IBM1*, At1g58602, and AT3g05410), the intronic TEs/repeats were associated with higher levels of H3K9me2 in Col-0 than in the regions upstream or downstream of the TEs/repeats (Fig. S7*B*). In contrast, H3K27me3 was not higher in the intronic TE/repeats than in the surrounding sequences in any of the genes (Fig. S7*C*). There were also higher levels of H3K27me1 and H3K27me2 at the intronic TE in At3g05410 but not in other genes (Fig. S7 *D* and *E*). These results are consistent with the notion that the intronic TEs/repeats are heterochromatic.

EDM2 is predicted to have a composite PHD domain consisting of three tandem C4HC3 PHD finger motifs (Fig. S1*B*). PHD fingers of some proteins bind to H3K4me3, an active gene mark, whereas some other PHD fingers recognize the repressive heterochromatic mark H3K9me2 (35–37). To test whether EDM2 can recognize specific histone marks, we probed histone peptide arrays with the recombinant GST-fused PHD domain of EDM2 (221–417 amino acids). Interestingly, the binding specificity of the recombinant protein was unusually broad because it could recognize not only H3K9me1, H3K9me2, and H3K9me3 but also H3K4me1, H3K4me2, and H3K4me3 (Fig. 6*A* and Fig. S7 *F* and *G*). There was little recognition of H3K27me3, H3K4ac, H3S10P, or of unmodified histone H3. In vitro pull-down assays confirmed that recombinant protein of the PHD domain of EDM2 could interact with methylated H3K9 and H3K4 and with H3K14ac but not with H3K27me3 or unmodified H3 (Fig. 6*B*).

We tested the importance of conserved histidine residues in each of the three PHD fingers of EDM2 by site-directed mutagenesis and found that the H246N, H316N, or H378N mutation

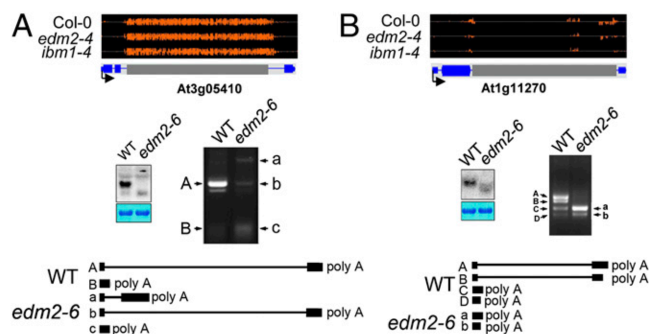

Fig. 5. 3′ RACE assays and Northern blot analysis of transcripts of intronic TE-containing genes. (*A* and *B*) 3′ RACE of At3g05410 (*A*) and At1g11270 (*B*). (*Top*) Screenshots of DNA methylation status in Col-0, *edm2-4*, and *ibm1-4*. The gray boxes inside the genes represent the intronic transposons AT3TE06550, and AT1TE12295, respectively. (*Middle*) Northern blot analysis (*Left*) and 3′RACE assays (*Right*). (*Bottom*) Diagrams of sequencing results of the 3′ RACE products.

in each of the C4HC3 finger motifs abolished or reduced the binding of the PHD domain of EMD2 to the methylated H3K9 and H3K4 peptides and to H3K14ac (Fig. 6*B*). Expression of WT *EDM2* but not mutant forms containing the H246N, H316N, or H378N mutation under its native promoter in *edm2-6* could rescue the root growth phenotype of the mutant plants (Fig. 6*C*). These results suggest that EDM2 can recognize histone marks in the intronic heterochromatin of target genes and that this recognition is important for its function *in planta*.

We carried out ChIP assays using anti-Myc antibody on *edm2-4* mutant plants complemented with native promoter-driven *EDM2-3xMYC*. The results show that there was an enrichment of EDM2 at the intronic heterochromatin region relative to the upstream (exon 1) and downstream (exon 17) regions of *IBM1* (Fig. 6*D*). In At1g58602, EDM2 was enriched to a significantly high level at the intronic Copia element compared with that at upstream (exon 1) and downstream (exon 5) regions (Fig. 6*D*). In contrast, there was no enrichment of EDM2 at the Copia4 element, which is not an intronic TE (Fig. 6*D*).

Discussion

Our results show that EDM2 has roles both in antisilencing and in preventing DNA hypermethylation. The roles of EDM2 in the antisilencing of the *35S-SUC2* transgene and in plant development appear to be mediated by IBM1, because engineered expression of *IBM1* could rescue the antisilencing and developmental defects of *edm2* mutant plants. Consistent with these observations, the *edm2* mutant shared hundreds of hypermethylated regions with *ibm1* and the 5-meC DNA glycosylase triple mutant *rdd* (Fig. 3*A* and Fig. S4*A*). Nevertheless, some ROS1 target sequences like At1g26390 (Fig. S4*E*) and some intergenic regions (Fig. 3*B*, regions 1, 2, and 3) are hypermethylated in *edm2* but not in *ibm1*. Therefore, EDM2 and IBM1 may have distinct and shared roles in antisilencing and in prevention of DNA hypermethylation in repetitive sequences. The antisilencing role of EDM2 is consistent with a recent report (32) of increased DNA methylation and reduced expression of the retrotransposon *COPIA4* in *edm2* mutant plants. Like IDM1 (26), EDM2 and IBM1 may help create favorable chromatin environments that are necessary for active DNA demethylation by ROS1 and related 5-meC DNA glycosylases at some of their target loci. Intriguingly, EDM2 may also have a role in silencing some transposons (32), which is consistent with our observation of DNA hypomethylation in some genomic regions in *edm2* mutant plants, although it is presently unclear how EDM2 may function in silencing.

Whole-genome bisulfite sequencing revealed that in addition to causing DNA hypermethylation in all sequence contexts (CG, CHG, and CHH) at hundreds of TEs, intergenic regions, and repetitive genes or genes containing repeats, the *edm2* mutation also resulted in dramatic increases in CHG methylation in the bodies of thousands of genes, particularly long genes. Our results demonstrated that EDM2 prevents genomewide CHG hypermethylation in gene bodies by controlling *IBM1* expression. EDM2 is required for the proper expression of not only *IBM1* but also of a number of other intronic heterochromatin-containing plant genes. Our results suggest that EDM2 facilitates the 3′ distal polyadenylation or prevents proximal polyadenylation of *IBM1* and other genes with long introns that harbor heterochromatic sequences. Recent studies showed that there is extensive alternative polyadenylation in plants, and more than 2,000 *Arabidopsis* genes have alternative polyadenylation sites in introns, particularly large introns (38). During the preparation of this manuscript, Tsuchiya and Eulgem (39) reported that EDM2 regulates the alternative polyadenylation of RPP7/At1g58602. The findings of Tsuchiya and Eulgem (39) are consistent with our results.

TEs are abundant in many higher eukaryotes, and some plant and mammalian genes contain intronic TEs or other invasive repeat elements (40). It is not surprising that mechanisms have been evolved to cope with intronic heterochromatin elements. EDM2 is a critical part of this mechanism. The intronic repeat sequences are associated with the repressive H3K9me2 mark (Fig. S7*B*) and are also expected to be associated with active H3K4 methylation marks because the intronic sequences are

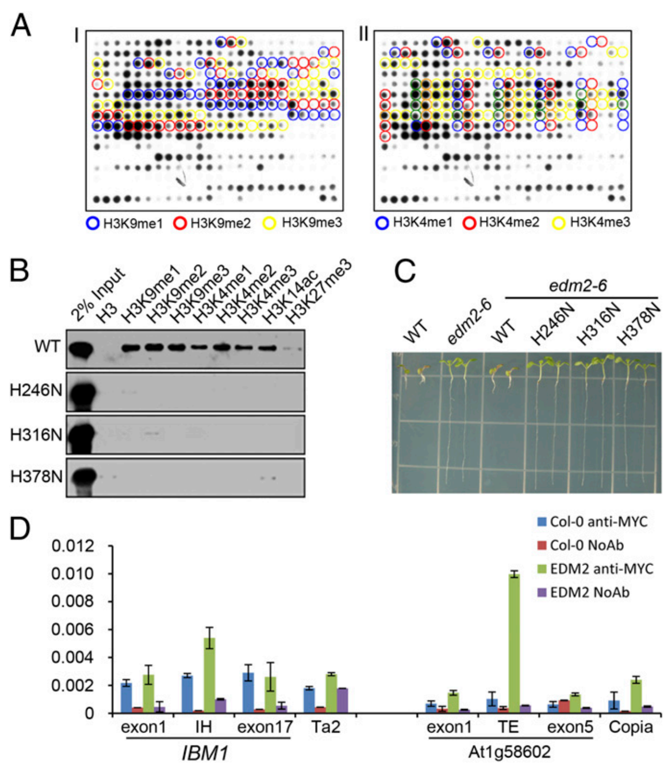

Fig. 6. EDM2 recognition of histone marks at intronic heterochromatin. (*A*) The PHD domain of EMD2 binds to histone H3 with a combination of modifications including H3K9 and H3K4 methylations. Blue, red, and yellow circles in *I* and *II* denote the locations of mono, di-, and trimethylation at lysine 9-containing peptides, and mono, di-, and trimethylation at lysine 4-containing peptides on the arrays, respectively. (*B*) Biotinylated peptide pull-down assays of the in vitro binding of the recombinant EDM2 PHD domain to histone H3 peptides with different modifications. Three mutated forms of the PHD domain (H246N, H316N, and H378N) show abolished or reduced binding to the modified histone peptides. (*C*) Mutations of the PHD domain (shown in *B*) impair the antisilencing function of EDM2. Root growth phenotypes of representative transformed *edm2-6* lines are shown. (*D*) Association of EDM2-3xMyc with intronic heterochromatin in *IBM1* and At1g58602 as revealed by ChIP-qPCR. Col-0 without the EDM2-3xMYC was used as a negative control.

actively transcribed. The composite PHD domain of EDM2 has a broad specificity in binding both repressive and active histone marks (Fig. 6 *A* and *B*) and is thus uniquely suited for recognizing the intronic heterochromatin of EDM2 target genes. Our results thus suggest that EDM2 recognizes the epigenetic marks in the intronic heterochromatin and associates with the intronic heterochromatin to facilitate 3′distal polyadenylation or to prevent proximal polyadenylation. In mammals, intronic CpG methylation is important for the expression of the tumor suppressor gene *EGR2* (41) and of the imprinted genes *H13* and *Herc3* (23, 42). In fact, CG methylation regulates distal polyadenylation of *H13* and *Herc3* transcripts, although the cellular factors involved are unknown (23, 42). Therefore, mechanisms for dealing with intronic heterochromatin may be conserved in higher eukaryotes.

Very recently, we and others have found that the BAH (bromo-adjacent homology) and RNA recognition motif containing protein ASI1/IBM2 regulates the alternative polyadenylation of *IBM1* transcripts and is thus important for preventing CHG methylation in gene bodies (25, 30). ASI1 is an RNA-binding protein, and like EDM2, ASI1 is also critical for antisilencing of the *35S-SUC2* transgene (25). ASI1 associates specifically with the intronic heterochromatin of IBM1 and other genes, possibly through its BAH domain (25). We hypothesize that the methyltransferase domain of EDM2 may catalyze RNA methylation, which may influence the RNA-binding of ASI1, thereby affecting the processing of ASI1-associated pre-mRNAs.

Materials and Methods

All *Arabidopsis* plants used in this study, including mutants and transgenic plants, were in the Columbia-0 (Col-0) genetic background. Details are provided in *SI Materials and Methods*, including plant materials and growth conditions, screening of EMS-mutagenized and T-DNA insertion mutants, plasmid construction and mutant complementation, bisulfite sequencing of individual loci, real-time RT-PCR and Northern analysis, whole-genome bisulfite sequencing and data analysis, mRNA-seq and data analysis, 3′ rapid amplification of cDNA ends, histone peptide array and pull-down assays, and ChIP assay. The primers used in this study are listed in Table S4.

ACKNOWLEDGMENTS. This work was supported by National Institutes of Health Grants R01 GM070795 and R01 GM059138 (to J.-K.Z.), National Natural Science Foundation of China Grant 31071060 (to D.L.), and the Chinese Academy of Sciences.

1. He XJ, Chen TP, Zhu JK (2011) Regulation and function of DNA methylation in plants and animals. *Cell Res* 21(3):442–465.
2. Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11(3):204–220.
3. Zemach A, Zilberman D (2010) Evolution of eukaryotic DNA methylation and the pursuit of safer sex. *Curr Biol* 20(17):R780–R785.
4. Martienssen RA, Richards EJ (1995) DNA methylation in eukaryotes. *Curr Opin Genet Dev* 5(2):234–242.
5. Tariq M, Paszkowski J (2004) DNA and histone methylation in plants. *Trends Genet* 20(6):244–251.
6. Matzke MA, Birchler JA (2005) RNAi-mediated pathways in the nucleus. *Nat Rev Genet* 6(1):24–35.
7. Zhang HM, Zhu JK (2011) RNA-directed DNA methylation. *Curr Opin Plant Biol* 14(2): 142–147.
8. Saze H, Mittelsten Scheid O, Paszkowski J (2003) Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nat Genet* 34(1):65–69.
9. Lindroth AM, et al. (2001) Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* 292(5524):2077–2080.
10. Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJM (2009) RNA-mediated chromatin-based silencing in plants. *Curr Opin Cell Biol* 21(3):367–376.
11. Wierzbicki AT, et al. (2012) Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the Arabidopsis epigenome. *Genes Dev* 26(16):1825–1836.
12. Zhong X, et al. (2012) DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. *Nat Struct Mol Biol* 19(9):870–875.
13. Huettel B, et al. (2006) Endogenous targets of RNA-directed DNA methylation and Pol IV in Arabidopsis. *EMBO J* 25(12):2828–2836.
14. Li X, et al. (2012) Antisilencing role of the RNA-directed DNA methylation pathway and a histone acetyltransferase in *Arabidopsis*. *Proc Natl Acad Sci USA* 109(28):11425–11430.
15. Dooner HK, Weil CF (2007) Give-and-take: Interactions between DNA transposons and their host plant genomes. *Curr Opin Genet Dev* 17(6):486–492.
16. Soppe WJ, et al. (2000) The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol Cell* 6(4):791–802.
17. Liu J, He Y, Amasino R, Chen X (2004) siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. *Genes Dev* 18(23):2873–2878.
18. Rigal M, Kevei Z, Pélissier T, Mathieu O (2012) DNA methylation in an intron of the IBM1 histone demethylase gene stabilizes chromatin modification patterns. *EMBO J* 31(13):2981–2993.
19. Crevillén P, Dean C (2011) Regulation of the floral repressor gene FLC: The complexity of transcription in a chromatin context. *Curr Opin Plant Biol* 14(1):38–44.
20. Lorincz MC, Dickerson DR, Schmitt M, Groudine M (2004) Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol* 11(11):1068–1075.
21. Saint-André V, Batsché E, Rachez C, Muchardt C (2011) Histone H3 lysine 9 trimethylation and HP1γ favor inclusion of alternative exons. *Nat Struct Mol Biol* 18(3): 337–344.
22. Spies N, Nielsen CB, Padgett RA, Burge CB (2009) Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* 36(2):245–254.
23. Wood AJ, et al. (2008) Regulation of alternative polyadenylation by genomic imprinting. *Genes Dev* 22(9):1141–1146.
24. Lei MG, et al. (2011) Genetic and genomic evidence that sucrose is a global regulator of plant responses to phosphate starvation in Arabidopsis. *Plant Physiol* 156(3): 1116–1130.
25. Wang X, et al. (2013) RNA-binding protein regulates plant DNA methylation by controlling mRNA processing at the intronic heterochromatin-containing gene IBM1. *Proc Natl Acad Sci USA* 110(38):15467–15472.
26. Qian W, et al. (2012) A histone acetyltransferase regulates active DNA demethylation in Arabidopsis. *Science* 336(6087):1445–1448.
27. Eulgem T, et al. (2007) EDM2 is required for RPP7-dependent disease resistance in Arabidopsis and affects RPP7 transcript levels. *Plant J* 49(5):829–839.
28. 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(5862): 462–465.
29. Inagaki S, et al. (2010) Autocatalytic differentiation of epigenetic modifications within the Arabidopsis genome. *EMBO J* 29(20):3496–3506.
30. Saze H, et al. (2013) Mechanism for full-length RNA processing of Arabidopsis genes containing intragenic heterochromatin. *Nat Commun* 4:2301.
31. Tsuchiya T, Eulgem T (2010) The Arabidopsis defense component EDM2 affects the floral transition in an FLC-dependent manner. *Plant J* 62(3):518–528.
32. Tsuchiya T, Eulgem T (2013) Mutations in EDM2 selectively affect silencing states of transposons and induce plant developmental plasticity. *Sci Rep* 3:1701.
33. Probst AV, Fransz PF, Paszkowski J, Mittelsten Scheid O (2003) Two means of transcriptional reactivation within heterochromatin. *Plant J* 33(4):743–749.
34. Zhang XY, et al. (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. *PLoS Biol* 5(5):e129.
35. Iwase S, et al. (2007) The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell* 128(6):1077–1088.
36. Li H, et al. (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* 442(7098):91–95.
37. Peña PV, et al. (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* 442(7098):100–103.
38. Wu X, et al. (2011) Genome-wide landscape of polyadenylation in Arabidopsis provides evidence for extensive alternative polyadenylation. *Proc Natl Acad Sci USA* 108(30):12533–12538.
39. Tsuchiya T, Eulgem T (2013) An alternative polyadenylation mechanism coopted to the Arabidopsis RPP7 gene through intronic retrotransposon domestication. *Proc Natl Acad Sci USA* 110(37):E3535–E3543.
40. Wei FS, et al. (2009) Detailed analysis of a contiguous 22-Mb region of the maize genome. *PLoS Genet* 5(11):e1000728.
41. Unoki M, Nakamura Y (2003) Methylation at CpG islands in intron 1 of EGR2 confers enhancer-like activity. *FEBS Lett* 554(1-2):67–72.
42. Cowley M, Wood AJ, Böhm S, Schulz R, Oakey RJ (2012) Epigenetic control of alternative mRNA processing at the imprinted Herc3/Nap1l5 locus. *Nucleic Acids Res* 40(18):8917–8926.