Regulation of DNA methylation turnover at LTR retrotransposons and imprinted loci by the histone methyltransferase Setdb1

Danny Leung, Tingting Du, Ulrich Wagner, Wei Xie, Ah Young Lee, Preeti Goyal, Yujing Li, Keith E. Szulwach, Peng Jin, Matthew C. Lorincz, and Bing Ren

During mammalian development, DNA methylation patterns need to be reset in primordial germ cells (PGCs) and preimplantation embryos. However, many LTR retrotransposons and imprinted genes are impervious to such global epigenetic reprogramming via hitherto undefined mechanisms. Here, we report that a subset of such genomic regions are resistant to widespread erasure of DNA methylation in mouse embryonic stem cells (mESCs) lacking the de novo DNA methyltransferases (Dnmts) Dnmt3a and Dnmt3b. Intriguingly, these loci are enriched for H3K9me3 in mESCs, implicating this mark in DNA methylation homeostasis. Indeed, deletion of the H3K9 methyltransferase SET domain bifurcated 1 (Setdb1) results in reduced H3K9me3 methylation homeostasis. Indeed, deletion of the H3K9 methyltransferase SET domain bifurcated 1 (Setdb1) results in reduced H3K9me3 and DNA methylation levels at specific loci, concomitant with increased 5-hydroxymethylcytosine (5hmC) and ten-eleven translocation 1 binding. Taken together, these data reveal that Setdb1 promotes the persistence of DNA methylation in mESCs, likely reflecting one mechanism by which DNA methylation is maintained at LTR retrotransposons and imprinted genes during developmental stages when DNA methylation is reprogrammed.

Significance

DNA methylation is essential for mammalian development. This modification is nearly completely erased and reestablished in early embryos, but specific classes of DNA elements escape such genome-wide changes via unknown mechanisms. In this study, we identified a likely factor responsible for lack of DNA methylation turnover on a large fraction of such sequences. By focusing on mouse embryonic stem cells depleted of de novo DNA methyltransferases, which exhibit widespread hypomethylation with the exception of particular loci, we show that regions retaining DNA methylation are associated with a specific chromatin state. In cells lacking the enzyme catalyzing this methylation turnover, such regions begin to lose DNA methylation. Our results therefore advance the understanding of how DNA methylation turnover is regulated during development.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE47984).

1 D.L. and U.W. contributed equally to this work.

*To whom correspondence should be addressed. E-mail: biren@usc.edu or mlorincz@mail.ubc.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1322273111/-/DCSupplemental.

6690–6695 | PNAS | May 6, 2014 | vol. 111 | no. 18

www.pnas.org/cgi/doi/10.1073/pnas.1322273111
loci (20, 22–24). Unlike the other K9-specific HMTases, Setdb1 has been shown to repress many ERVs independent of DNA methylation (20, 24). Curiously however, simultaneous depletion of Dnmt1 and Setdb1 results in synergistic reactivation of a subset of ERVs, including IAP elements (20), indicating that both H3K9 and DNA methylation play a role in silencing of such parasitic elements.

Whereas the global level of DNA methylation is dramatically reduced in late passage Dnmt3a/3b double knockout (DKO) ESCs, sequence analysis of specific genomic loci clearly reveals that a subset of regions, including IAP elements, retains relatively high levels of mCG in these cells (7, 23), reminiscent of the pattern observed in postmitigatory PGCs and blastocysts. To ascertain the spectrum of genomic regions that remain methylated in the absence of de novo Dnmt activity, we conducted methylC-seq (methylC-seq) on Dnmt3a/3b DKO mESCs. Thousands of regions with higher than average mCG levels, which we term enriched residual methylation loci (ERML), were identified. We further show that persistence of mCG at a subset of these loci, including ERVs, germ-line genes, and a subgroup of imprinted DMRs, is dependent upon Setdb1. Finally, we provide evidence that Setdb1-mediated H3K9me3 promotes the stability of mCG at ERML at least in part by inhibiting Tcj-dependent oxidation in these regions.

**Results**

**Characterization of Enriched Residual Methylation Loci in Dnmt3a/3b DKO mESCs.** To comprehensively map the genomic regions that remain hypermethylated in the absence of de novo Dnmt3a and/or Dnmt3b, we analyzed the genome-wide distribution of DNA methylation in late passage DKO (passage 33) and WT (passage 39) parental mESCs by methylC-seq. Consistent with previous studies, we found that with the exception of the majority of CpG islands, WT mESCs have high levels of mCG throughout the genome (average methylation level of ~70%) (25). In contrast, mCG levels were dramatically reduced in Dnmt3a/3b DKO cells (average mCG level of 2.7%; Fig. S1 A and B). Non-CG methylation, an epigenetic mark primarily found in embryonic stem cells and neuronal cells, was reduced to background levels, indicating that Dnmt3a and/or Dnmt3b are responsible for deposition of these modifications, as previously reported (26). Although the vast majority of the Dnmt3a/3b DKO genome (>90%) showed a dramatic reduction in methylation, many regions retain relatively high levels of mCG. Using a hidden Markov model (HMM)-based algorithm developed for this study (details in SI Materials and Methods), we identified 6,115 such regions consisting of multiple contiguous bins with greater than 10% mCG, which is considerably higher than the genome average in DKO cells. The average mCG level of all these loci was 38%, with a median size of 5.4 kb. Examples of ERMLs are shown in Fig. 1A and Fig. S1B.

Given our previous observations that specific genes and LTR retrotransposons are regulated by both H3K9me3 and DNA methylation, we analyzed the enrichment of H3K9me3 by ChIP-seq in WT and Dnmt3a/3b DKO cells. Consistent with our previous analysis of Dnmt TKO mESCs (20), we observed that H3K9me3 is generally preserved in DKO mESCs (Fig. 1B).

**Fig. 1.** Genome-wide profile of mCG in wild-type and Dnmt3a/3b double knockout (DKO) mESCs. (A, Upper) University of California Santa Cruz (UCSC) genome browser screen capture showing the DNA methylation profiles for WT and DKO mESCs, mouse E13.5 male primordial germ cells (PGCs) (17), and blastocysts (16) across a region on chromosome 11. ERML are denoted with black bars. H3K9me3 ChIP-seq reads from WT and DKO are also included. H3K9me3-enriched domains in WT are called “ChromaBlocks.” Genes and repetitive elements as annotated by RefSeq and RepeatMasker, respectively, are shown. MethylC-seq reads from WT and DKO are mapped, with 100-bp binned data showing mCG ratios with bar heights between 0 and 1. H3K9me3 profiles are shown as reads per kilobase pair per million reads (RPKM). Mappability of reference genome from the ENCODE project for 50-bp segments is shown with mCG levels (%). (B) Heat map generated by k-means clustering shows the H3K9me3 enrichment (Right) at all identified peaks and mCG (Left) in WT and DKO mESCs. Red boxes outline regions retaining mCG in the DKO cells, which we term ERML. These sites are enriched for strongest H3K9me3 peaks. (C) Pie chart summarizing the overlap between regions enriched for H3K9me3 in both WT and DKO mESCs and ERML. (D) Boxplot showing the difference in mCG levels at ERML and control shuffled regions in PGCs and blastocysts. (E) Barplot of mCG ratios of mappable LTR (red) and non-LTR (blue) repeats in WT (Upper) and DKO (Lower) mESCs. Arrow indicates families that retain higher levels of mCG.
Interestingly, many H3K9me3 peaks showed enrichment of mCG in the DKO cells, suggesting a possible relationship between ERML and this histone mark (Fig. 1B). Indeed, we found that a large proportion of ERML (41.5%, n = 2,538) overlap with H3K9me3-marked regions (Fig. 1B and C). This is considerably higher than the overlap between H3K9me3 peaks and control shuffled regions (5.2%, n = 543) (Fig. S1D), generated by randomizing positions of ERML while maintaining the average region size. For subsequent analyses, we focused on the subset of ERML (n = 2,451) found within the main clusters, generated by k-means clustering, which overlap with H3K9me3 (Fig.1B, red boxes).

Previous studies indicate that DNA methylation levels are substantially higher in mESCs than in blastocysts from which they are derived (16). Strikingly, we found that the ERML identified in DKO mESCs also show relatively higher levels of mCG in WT blastocysts (16) (Fig. 1D). Furthermore, we observed a similar retention of mCG in WT male embryonic day 13.5 (E13.5) PGCs at ERML compared with other genomic regions (17) (Fig. 1D). In fact, we observed a high co-occurrence of ERML and hypermethylated regions in both PGCs and blastocysts across large regions of the genome (Fig. 1D and Fig. S1D). Taken together, these data indicate that residual methylation in blastocysts, PGCs, and Dnmt3a/b DKO mESCs may be regulated by a common H3K9me3-dependent pathway that promotes maintenance of DNA methylation and/or protects against DNA demethylation rather than promoting de novo DNA methylation per se.

To determine whether ERML are enriched for specific genomic features, we characterized these regions in greater detail. Interestingly, ERML are enriched for repetitive elements (Table S1). To address whether specific repeat subclasses are enriched, we calculated the average mCG levels of all mappable repeat classes in the Dnmt3a/b DKO versus the WT parental line. Intriguingly, specific subsets of ERVs and their long terminal repeats (LTRs) retained significantly higher mCG levels than other repeat classes (Fig. 1E). We then compared the proportion of repeat subfamilies overlapping with ERML versus the percentage of base pairs of the same subfamilies overlapping with ERML (Fig. S1E and F). This analysis allowed us to discriminate between specific repeat subfamilies showing substantial overlap with ERML and those found by chance to be near or minimally overlapping with ERML due to their high copy numbers in the genome. Repeat subfamilies showing the highest degree of overlap with ERML belong to class I or class II ERVs, including IAP, early mouse transposons (ETn), and Mus musculus ERV using tRNA_Lys type 10 C (MMERVK10C) elements (Fig. S1E and F and Table S1). These ERVs do not show significant overlap with shuffled negative control regions (Fig. S1G). We previously showed that MMERVK10C and ETn subfamilies are induced to a greater extent in Setdb1 KO cells than in Dnmt TKO mESCs, whereas IAP elements are only modestly induced in both KO lines (20, 24). Intriguingly however, simultaneous depletion of Setdb1 and Dnmt1 yielded synergistic activation of IAP ERVs (20), consistent with the hypothesis that Setdb1 and DNA methylation target the same ERV subfamilies.

**Setdb1 Deletion Leads to Hypomethylation at a Subset of ERVs and ERML.** Given that ERML are enriched for H3K9me3 and Setdb1-regulated ERV subfamilies (20, 24), we next addressed whether this HMTase is required to maintain mCG at ERML. As Setdb1-deficient mESCs are not viable, we used a previously described Setdb1 conditional knockout (CKO) cell line (24). Efficient deletion of Setdb1 was validated by Western blot (Fig. S2A). Analysis of the effects of Setdb1 deletion on H3K9me3 by ChIP-seq in predeletion (hereinafter referred to as WT) and postdeletion (Setdb1 KO) cells revealed a reduction at virtually all mappable H3K9me3 peaks (Fig. S2B) and especially at ERML marked with H3K9me3 (Fig. S2B), consistent with previously published data (20) (Fig. S2C). We then analyzed the methylomes of WT and Setdb1 KO cells by methyl-C-seq at day 6 postinduction with 4-hydroxytamoxifen (Fig. S3A). Global DNA methylation levels were not dramatically reduced upon Setdb1 KO, perhaps due to the limited number of cell divisions before harvest (24). Unfortunately, Setdb1 KO cells survive for only 7 d (24), precluding analysis at later time points. Nevertheless, we sought to identify specific regions where mCG is altered using Fisher’s exact test to calculate the significance of mCG differences across the genome between WT and Setdb1 KO (SI Materials and Methods). We found 2,395 mappable repetitive elements that showed significant hypomethylation. Intriguingly, compared with all mappable ERVs, class I and II ERVs in particular were overrepresented in Setdb1 KO cells, showing a ~2.5-fold (7% versus 18%) and ~2.5-fold increase (29% versus 72%), respectively (Fig. 2A). In contrast, class III ERVs, which are not marked by H3K9me3 (20), are underrepresented (64%)
Setdb1 Antagonizes 5-Hydroxymethylcytosine-Mediated DNA Demethylation. To determine whether Setdb1/H3K9me3 promotes the persistence of DNA methylation at least in part by inhibiting DNA demethylation, we studied the genome-wide distribution of 5hmC, an intermediate product of TET protein-dependent DNA demethylation pathways (9, 10). Using a method that selectively labels 5hmC for affinity capture, we produced genome-wide maps for 5hmC enrichment in WT and Setdb1 KO mESCs at day 6 postdeletion. Intriguingly, the mean 5hmC enrichment level was increased at ERML, showing a significant decrease in methylation in Setdb1 KO mESCs (n = 288) (Fig. 3A). Expanding our analyses to include all ERML that overlap with H3K9me3 (n = 2,451) in WT versus Setdb1 KO mESCs, we found a significantly greater number of loci showing a reduction of H3K9me3 upon Setdb1 deletion, coupled with an increase of 5hmC than at shuffled control regions (Fig. 3B and Fig. S4A). Furthermore, the genomic loci that showed the greatest increase of 5hmC (n = 169) were highly enriched for ERVs, including IAP and ETn elements (Fig. S4B). Analysis of the median levels of 5hmC enrichment revealed a clear increase at mappable ETn (n = 11,207) and IAP elements (n = 22,962). Mappable ERVL elements (n = 9,704), which are not generally marked by Setdb1-deposited K9me3, show a subtle decrease of 5hmC enrichment in Setdb1 KO mESCs (Fig. 3B). Heat maps showing the difference of K9me3 and 5hmC enrichment (RPKM) between Setdb1 KO and WT mESCs at ERML marked by K9me3 (n = 2,451) (Left) and size-matched shuffled control regions (Right). (C) Boxplots illustrate the increase of Tet1 enrichment (RPKM) at ERML marked by K9me3 (n = 2,451) but not at control regions (n = 2,451), in Setdb1 KO mESCs. (*P < 0.05 and ***P < 0.001, respectively, as comparing WT versus Setdb1 KO with the Wilcoxon test.)

**Fig. 3.** ERML gain 5hmC in Setdb1 KO mESCs. (A) Boxplot showing an increase of 5hmC enrichment (RPKM) at ERML, which become hypomethylated upon Setdb1 deletion (n = 288). An increase is also observed at mappable ERVL (n = 9,704) and IAP elements (n = 22,962). Mappable ERVL elements (n = 9,704), which are not generally marked by Setdb1-deposited K9me3, show a subtle decrease of 5hmC enrichment in Setdb1 KO mESCs. (B) Heat maps showing the difference of K9me3 and 5hmC enrichment (RPKM) between Setdb1 KO and WT mESCs at ERML marked by K9me3 (n = 2,451) (Left) and size-matched shuffled control regions (Right). (C) Boxplots illustrate the increase of Tet1 enrichment (RPKM) at ERML marked by K9me3 (n = 2,451) but not at control regions (n = 2,451), in Setdb1 KO mESCs. (*P < 0.05 and ***P < 0.001, respectively, as comparing WT versus Setdb1 KO with the Wilcoxon test.)

Setdb1 Protects Against Demethylation of Germ-Line and Imprinted Genes. We next investigated whether Setdb1/H3K9me3 plays a similar role in preserving DNA methylation at unique gene sequences. We analyzed the methylomes generated for WT and Setdb1 KO mESCs and found that 252 genes showed a significant loss of DNA methylation at their promoters. Gene ontology analysis revealed several classifications associated with meiosis (Fig. 4A). Interestingly, we previously showed that 30 germ-line–specific genes were transcriptionally repressed by both DNA methylation and Setdb1/H3K9me3 (20). Strikingly, there was a significant decrease in the median mCG level at these germ-line gene promoters in Setdb1 depleted cells, whereas no decrease in mCG levels was detected at the promoter regions of all Reference Sequence (RefSeq) annotated genes, which have a significantly lower median mCG methylation level in WT cells (Fig. 4B). This reduction in DNA methylation was clearly apparent at the promoter regions of the germ-line–specific genes Sycp3, Dazl, Mael, and RhoX13 (Fig. 4C and Fig. S6A and B), and hypomethylation of the promoter region of several of these epigenetic pathways, where Setdb1/H3K9me3 protects against Tet1/mCG-associated DNA demethylation at many LTR elements.

Leung et al.
A recent study revealed that H3K9me3 at differentially methylated regions (DMRs) of imprinted genes is dependent upon Kap1, which interacts directly with Setdb1 (29, 30). Consistent with this observation, we also found that H3K9me3 at a subset of DMRs is dependent upon Setdb1 (Fig. S6C). To address whether maintenance of DNA methylation at DMRs is also dependent on depletion in mESCs, we focused on a subset of DMRs marked by H3K9me3, including those for both paternally and maternally imprinted genes. Consistent with previous reports, we found that the methylation level of DMRs in WT cells were not necessarily maintained at the expected 50%, possibly due to culture conditions of mESCs. Regardless, four of the five analyzed DMRs (Meg3, Nespas, Mest, and Pog3) showed a significant decrease of mCG levels, whereas H19 showed no significant change (Fig. 4D and Fig. S4C). Interestingly, four of the five DMRs (Meg3, Nespas, Mest, and H19) also showed a significant increase of 5hmC enrichment (Fig. 4D and Fig. S6C), indicating that Setdb1 may also play a role in protecting imprinted DMRs against 5hmC-dependent demethylation.

**Discussion**

Previous studies of Dnmt3a/3b DKO mESCs identified a few regions in the genome, including IAP elements and several imprinted DMRs, that retain near WT levels of DNA methylation (7, 23), presumably maintained by Dnmt1. Here, we conducted a genome-wide search for hypermethylated regions in these cells. Remarkably, many such ERML correspond to genomic regions that retain high mCG levels in PGCs and blastocysts. A significant fraction of ERML are marked by H3K9me3 in mESCs and are enriched for active class I and II ERVs, parasitic sequences with the potential to deregulate nearby genes and transpose to new genomic sites. Setdb1-deposited H3K9me3 is necessary for silencing of several such ERV subfamilies, including IAP, ETn, and MMERVK10C (20, 24). The overlap between ERML and H3K9me3-marked loci prompted us to investigate the relationship between these epigenetic pathways. Upon depletion of Setdb1 in mESCs, H3K9me3 and DNA methylation are concomitantly reduced at a subset of ERML, which simultaneously gain 5hmC and Tet1 binding. Given that oxidation of 5mC has been shown to act as an intermediate in DNA demethylation (9, 10), these results reveal an antagonistic relationship between Setdb1/K9me3 and 5mC removal. It is noteworthy that Setdb1 KO cells also show DNA hypermethylation at non-LTR retrotransposons and at a subset of single copy genes. However, as these sequences are not generally marked or transcriptionally regulated by H3K9me3, we focused our analyses on the sequences that lose DNA methylation upon Setdb1 depletion.

Reprogramming of DNA methylation patterns in PGCs was recently shown to be a gradual process involving sequential Tet-independent and Tet-dependent pathways. Whereas the reactivation of genes such as Dazl begins concomitantly with the onset of DNA demethylation, the mCG level decreases gradually over the next several days until the Dazl promoter is completely unmethylated (14). This temporal requirement for hydroxymethylation-dependent DNA demethylation turnover likely reflects replication-coupled loss of mCG and may also apply in mESCs. The incomplete demethylation of ERML in Setdb1 null cells could result from a lack of sufficient cell divisions to detect replication-coupled loss of mCG. However, as Setdb1 null ESCs are not viable for more than 7 d postdeletion, it is not possible to determine whether DNA methylation is progressively lost in the absence of H3K9me3.

Interestingly, Uhrl/Np95, which recruits Dnmt1 to replicating DNA and is essential for maintenance DNA methylation, was recently shown to bind cooperatively to hemimethylated CpG and H3K9me2/me3 (31, 32). Moreover, the interaction with methylated H3K9 promotes fidelitous maintenance DNA methylation (31). These observations indicate that in addition to