PNAS
www.pnas.org

# Supplementary Information for

**EHMT2 and SETDB1 protect the maternal pronucleus from 5mC oxidation**

Tie-Bo Zeng[1], Li Han*, Nicholas Pierce[1], Gerd P. Pfeifer[1], and Piroska E. Szabó[1$]

[1]Center for Epigenetics, Van Andel Research Institute, 333 Bostwick Ave. NE Grand Rapids, MI 49503

Corresponding Author: Piroska E. Szabó, Ph.D. Center for Epigenetics, Van Andel Research Institute, 333 Bostwick Ave. NE Grand Rapids, MI 49503
piroska.szabo@vai.org

**This PDF file includes:**





www.pnas.org/cgi/doi/10.1073/pnas.1819946116

**Supplementary Methods:**

Comparing 5hmC enrichment in neurons and H3K9me3 in forebrain at E14.5.

Source of Data: H3K9me3 ChIP-seq data was downloaded from **GSE102487** and **5hmC hMeDIP data** was downloaded from **GSE38118.**

Downloading Raw ChIP-seq Reads:
SRA runs SRR5931465 and SRR5931467 were downloaded from the SRA accession number SRP115381 using the SRA tool kit (version 2.9.1). Briefly, the SRA files were prefetched and the raw-fastq files were extracted using fastq-dump in pair end mode.

Alignment of H3K9me3 ChIP-seq samples to mm9 genome:
Paired-end reads were aligned to the mm9 genome using bowtie2 (version 2.3.4.2) (1) with default parameters. The resultant SAM files were converted into a BAM file, sorted, and indexed using samtools (Version: 1.8) (2). The log2 coverage files (bigwigs) between H3K9me3 and Input were generated using deeptools2.0 bamCompare with parameters --operation log2 --binSize 50.

H3K9me3 Peak Calling using Histone_HMM:
Chromosome information for the mm9 genome was downloaded from ftp://hgdownload.cse.ucsc.edu/goldenPath/mm9/database/. H3K9me3 peaks were called using Histone_HMM (3) with parameters -b 5000 -c mm9_chromInfo.txt. Peaks with an average posterior probability of at least 0.8 were kept for further analysis.

5hmC data:
To convert the tdf file into bedgraph format, igvtools (Version 2.3.98) tdftobedgraph was used. The bedgraph file was sorted and overlapping sites were merged using bedtools (v2.26.0) (4) sort (default parameters) and merge (-c 4 -o mean). Lastly, the bedgraph file was converted into bigwig using the UCSC utility bedGraphToBigWig.

Profile Plots
Profile plots between the 5hmC and H3K9me3 data were generated using the deeptools2.0 computeMatrix and plotProfile modules (5).

Figure S1

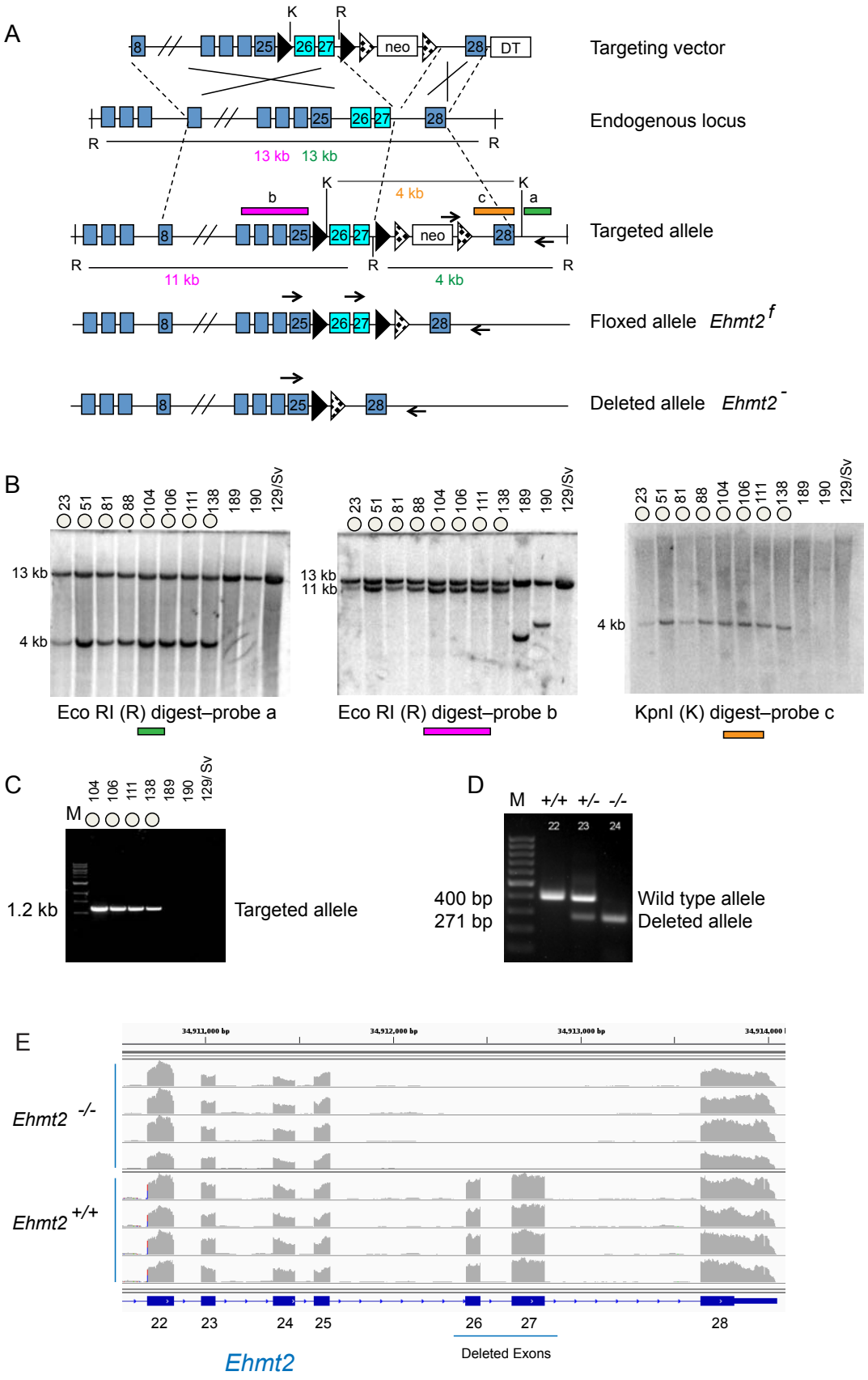
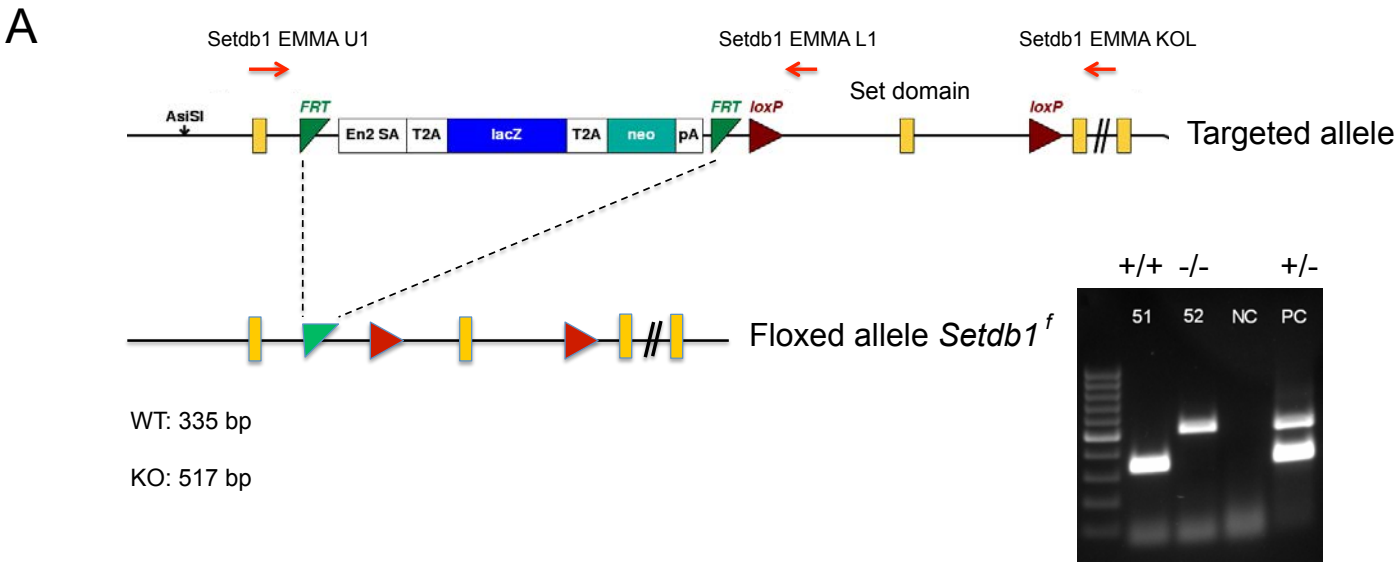

A
K
R
8
25
26
27
neo
28
DT
Targeting vector
Endogenous locus
13 kb
13 kb
b
4 kb
c
a
Targeted allele
11 kb
4 kb
Floxed allele Ehmt2 f
Deleted allele Ehmt2 -
B
23
51
81
88
104
106
111
138
189
190
129/Sv
13 kb
4 kb
13 kb
11 kb
Eco RI (R) digest–probe a
Eco RI (R) digest–probe b
KpnI (K) digest–probe c
C
M
104
106
111
138
189
190
129/Sv
1.2 kb
Targeted allele
D
M
+/+
+/-
-/-
22
23
24
400 bp
271 bp
Wild type allele
Deleted allele
E
34,911,000 bp
34,912,000 bp
34,913,000 bp
34,914,000
Ehmt2 -/-
Ehmt2 +/+
22
23
24
25
26
27
28
Deleted Exons
Ehmt2

**Fig. S1. Generating a conditional allele of the mouse *Ehmt2* gene.**

(A) The targeting construct, the endogenous locus, and the successfully targeted allele is depicted. Gene targeting was done by homologous recombination in ES cells. Exons (blue), targeted exons (turquoise), *loxP* sites (black triangles), the Pgk-neo positive selection cassette (neo), the diphtheria toxin negative selection cassette (DT), and the *frt* sites (patterned triangles) are indicated. Restriction enzymes, EcoRI (R) and KpnI (K), used in Southern blot hybridization are marked in the map. The size of diagnostic restriction fragments is labeled in colored numbers according to the hybridization probes that detect them (colored rectangles a-c). PCR genotyping primers are shown with horizontal arrows. (B) The correctly targeted ES clones were confirmed by Southern blot hybridization using three probes. (C) PCR screening over the short arm identified candidate ES cell clones, marked by gray dots. ES clones were injected to blastocysts to obtain chimeras. F1 individuals were mated with Flp recombinase-transgenic mouse JAX 003946 - 129S4/SvJaeSor-Gt(ROSA)26Sortm1(FLP1)Dym/J to remove the Pgk-neo cassette. The *Ehmt2*$^{-}$ allele was generated by mating to 129S1/Sv-Hprt$^{tm1(CAG-cre)Mann}$/J (6). (D) Genotyping of mice was done by PCR. (E) RNA-seq experiment confirmed that the SET-domain encoding exons of *Ehmt2* were seamlessly excised in the mutant *Ehmt2*$^{-/-}$ embryos at 8.5 dpc.

Figure S2


A
Setdb1 EMMA U1
Setdb1 EMMA L1
Setdb1 EMMA KOL
Set domain
AsiSI
FRT
En2 SA
T2A
lacZ
T2A
neo
pA
FRT
loxP
loxP
Targeted allele
Floxed allele Setdb1 f
WT: 335 bp
KO: 517 bp
+/+
-/-
+/-
51
52
NC
PC

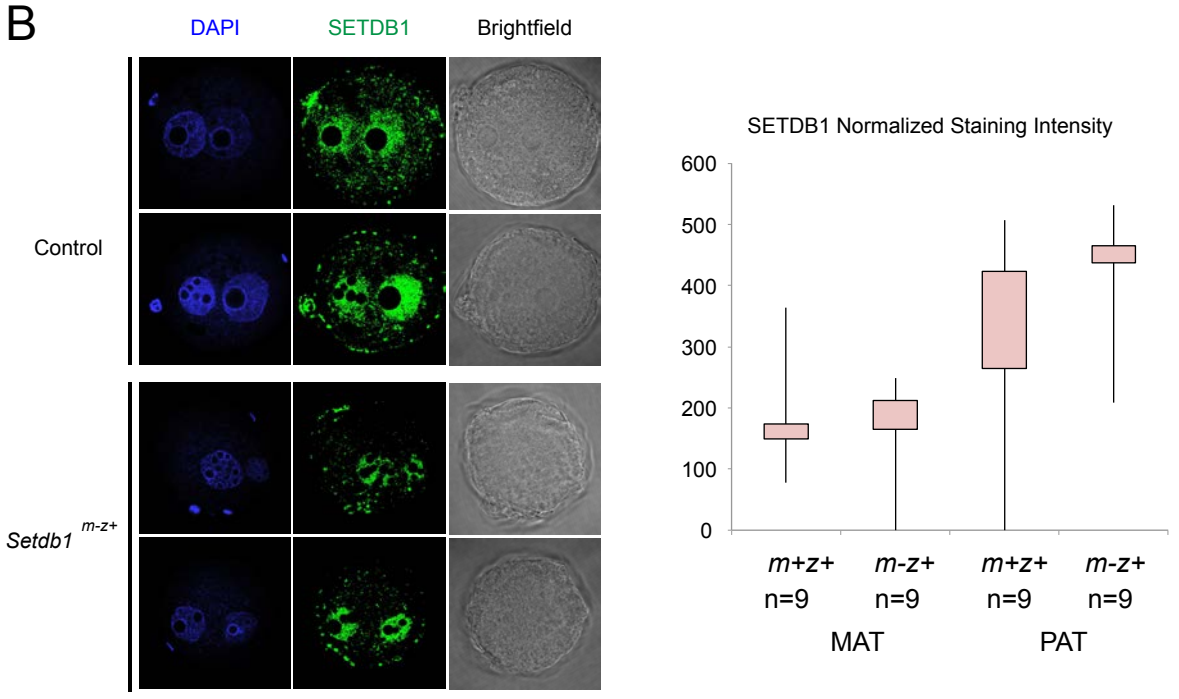

B
DAPI
SETDB1
Brightfield
Control
Setdb1 m-z+
SETDB1 Normalized Staining Intensity
600
500
400
300
200
100
0
m+z+
n=9
m-z+
n=9
m+z+
n=9
m-z+
n=9
MAT
PAT

**Fig. S2. Validation of the *Setdb1*$^{Wtsitm1a(EUCOMM)}$ mutant mouse line.**

(A) The targeting strategy is shown with the genotyping primer locations. The LacZ-neo cassette was removed by crossing the mutant mice with a Flp recombinase-transgenic mouse, JAX 003800 - B6;SJL-Tg(ACTFLPe)9205Dym/J. The inset agarose gel image shows the separation of wild type and KO-specific PCR fragments that identify the genotypes of blastocyst as labeled on top. DNA-free sample was used as negative control (NC) and *Setdb1*$^{+/-}$ mouse tail was used as positive control (PC). (B) Immunostaining of control and *Setdb1*$^{m-z+}$ maternal mutant zygotes (to the left) and quantification results of the SETDB1 image intensities (to the right).

Figure S3

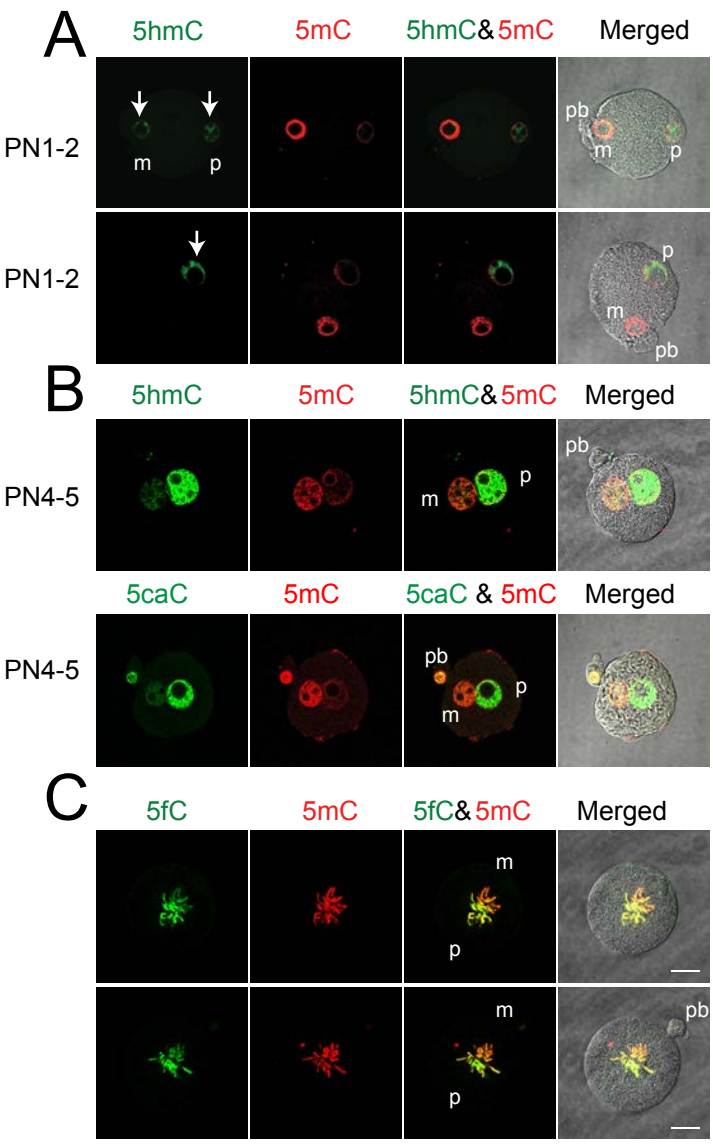
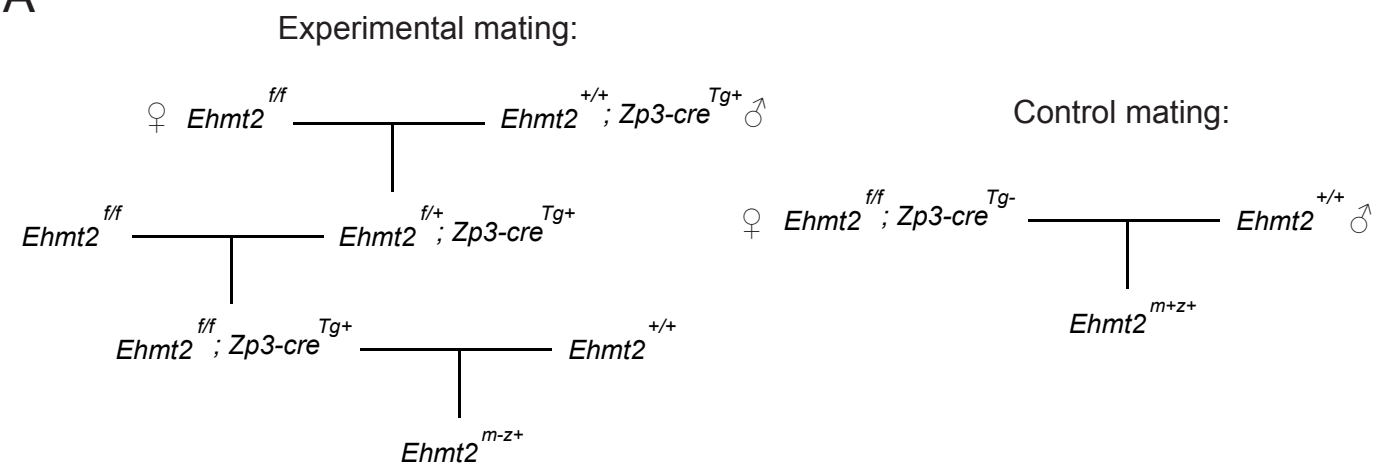

A
5hmC
5mC
5hmC&5mC
Merged
PN1-2
m
p
pb
PN1-2
p
m
pb
B
5hmC
5mC
5hmC&5mC
Merged
PN4-5
m
p
pb
5caC
5mC
5caC & 5mC
Merged
PN4-5
pb
p
m
C
5fC
5mC
5fC&5mC
Merged
m
p
m
p
pb

**Fig. S3. The 5mC oxidation process in control zygotes**

Confocal images of zygotes and 1-cell embryos stained with anti-5mC (red) and one of the following: anti-5hmC (A), anti-5hmC or anti-5caC (B), anti-5fC (C) antibodies (green), as indicated. (A) Early pronuclear (PN1-2) stage zygotes already exhibit 5hmC in the paternal and somewhat in the maternal pronucleus (arrows). (B) PN4-5 stage zygotes display high level of 5mC oxidation in the paternal pronucleus (C) Maternal (5mC-rich) and paternal (5fC-rich) chromosomes are being separated during mitosis of the first cell.

A


Experimental mating:
♀ Ehmt2 f/f
Ehmt2 +/+ ; Zp3-cre Tg+ ♂
Ehmt2 f/f
Ehmt2 f/+ ; Zp3-cre Tg+
Ehmt2 f/f ; Zp3-cre Tg+
Ehmt2 +/+
Ehmt2 m-z+
Control mating:
♀ Ehmt2 f/f ; Zp3-cre Tg-
Ehmt2 +/+ ♂
Ehmt2 m+z+

B

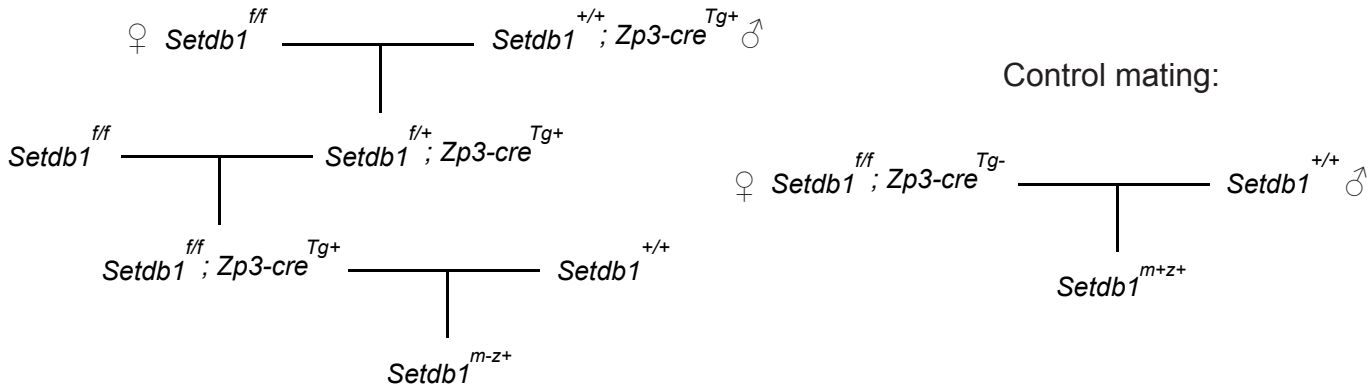

Experimental mating:
♀ Setdb1 f/f
Setdb1 +/+ ; Zp3-cre Tg+ ♂
Setdb1 f/f
Setdb1 f/+ ; Zp3-cre Tg+
Setdb1 f/f ; Zp3-cre Tg+
Setdb1 +/+
Setdb1 m-z+
Control mating:
♀ Setdb1 f/f ; Zp3-cre Tg-
Setdb1 +/+ ♂
Setdb1 m+z+

**Fig. S4. Mating scheme**

The three-step breeding process of generating maternal mutant zygotes is shown. Mothers are to the left and fathers are to the right in each cross. (A) Obtaining *Ehmt2*$^{m\text{-}z+}$ zygotes. The Zp3-cre transgene was crossed into the *Ehmt2*$^{f/f}$ mouse line from the father. This transgenic line was made homozygous for *Ehmt2*$^{f/f}$ by crossing *Ehmt2*$^{f/+}$ male transgene carriers with *Ehmt2*$^{f/f}$ females. The resulting *Ehmt2*$^{f/f}$ homozygous and Zp3-cre transgenic females were crossed to wild type males to obtain the maternal mutant zygotes. Cre-excision of the SET domain in the *Ehmt2*$^{-/-}$ growing oocytes eliminated EHMT2 function in the zygote. The maternal mutant (but zygotic wild type) heterozygous zygote (m-z+) also contains a normal zygotic allele from the sperm, which will switch on later when embryonic genome activation (EGA) takes place in the embryo. (B) Obtaining *Setdb1*$^{m\text{-}z+}$ zygotes.

Figure S5

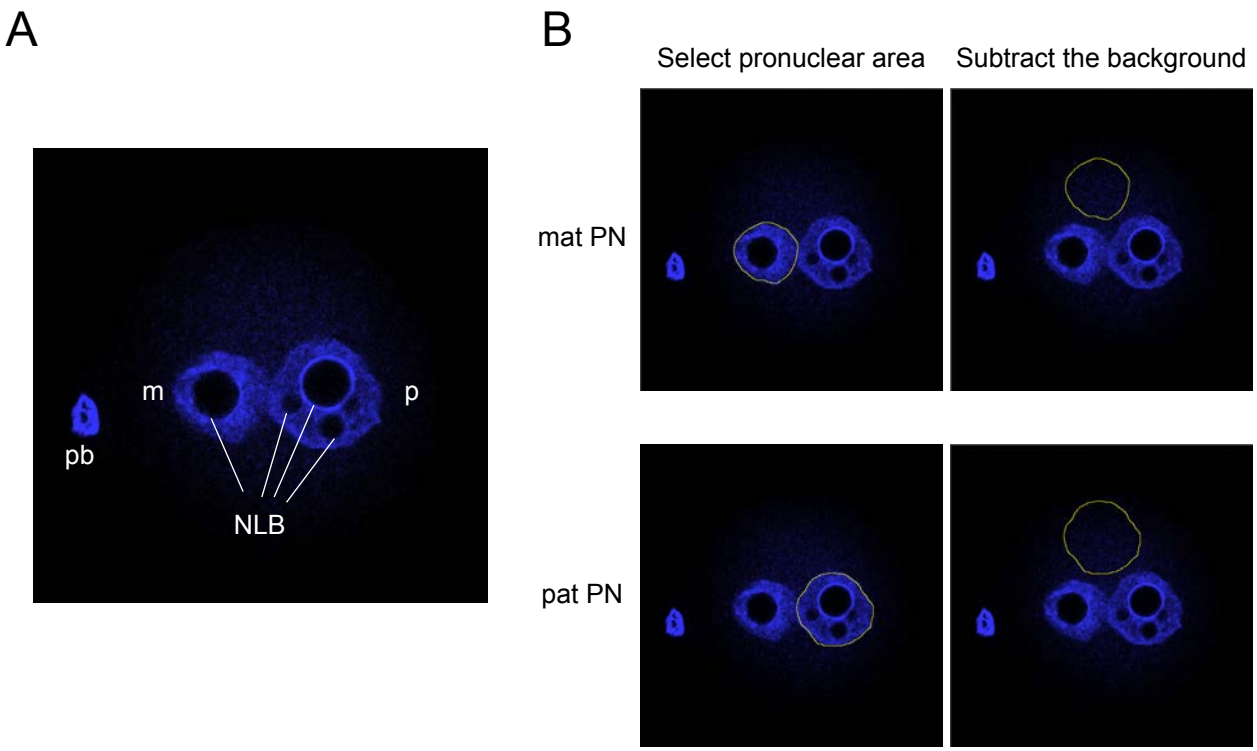
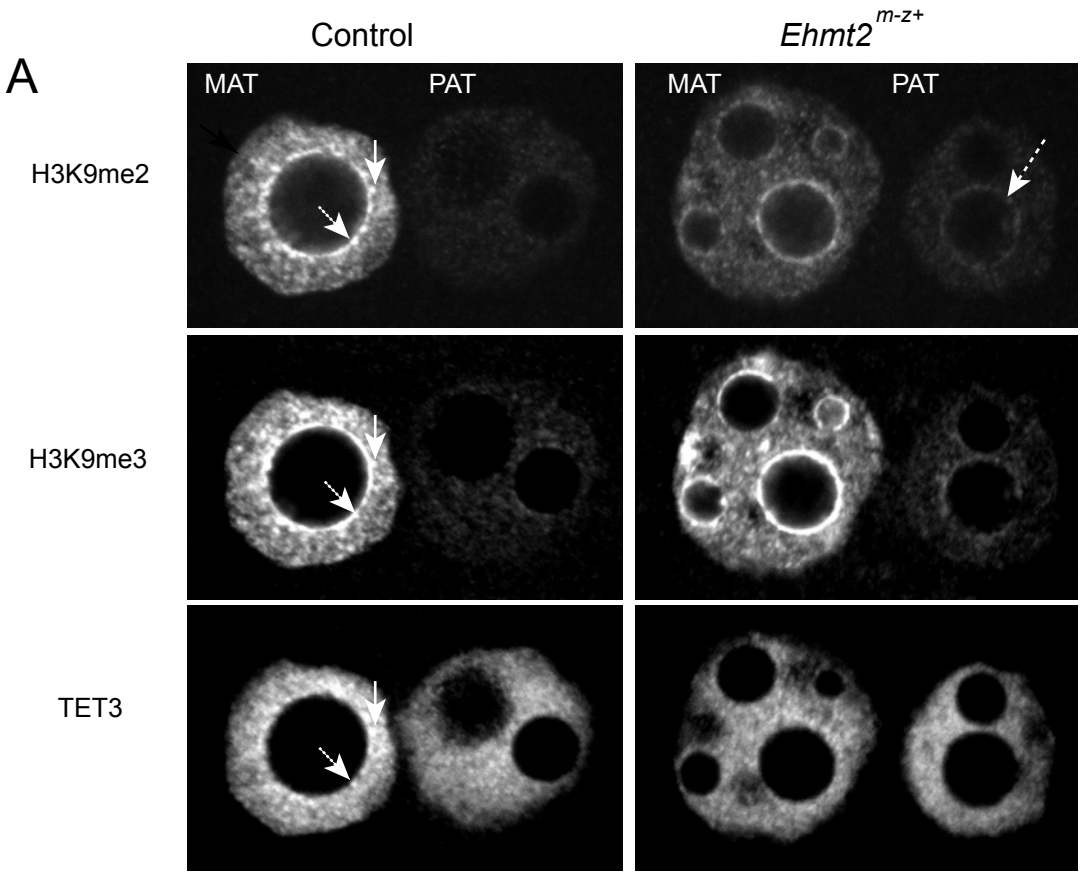

A
m
p
pb
NLB
B
Select pronuclear area
Subtract the background
mat PN
pat PN

**Fig. S5. Quantification of signal intensities in the zygote**

(A) DAPI staining of a control zygote is shown with the structural components, polar body (pb), maternal pronucleus (m), paternal pronucleus (p), and nucleolar-like body (NLB) as marked. (B) The contour of maternal pronucleus is drawn on the image and the intensity is measured inside this line using FIJI software. The pronucleus shape is then copied over to the area in the cytoplasm, where the intensity is measured again. This background intensity is deducted from the intensity measured over the maternal pronucleus. The same process is performed for measuring the staining intensities in the paternal pronucleus. Each pronucleus is measured in its best fitting Z-section. The immunostaining intensity for each specific antibody is measured in the same contour outlined using the DAPI counterstain image.

Figure S6


A
Control
Ehmt2 m-z+
MAT
PAT
MAT
PAT
H3K9me2
H3K9me3
TET3

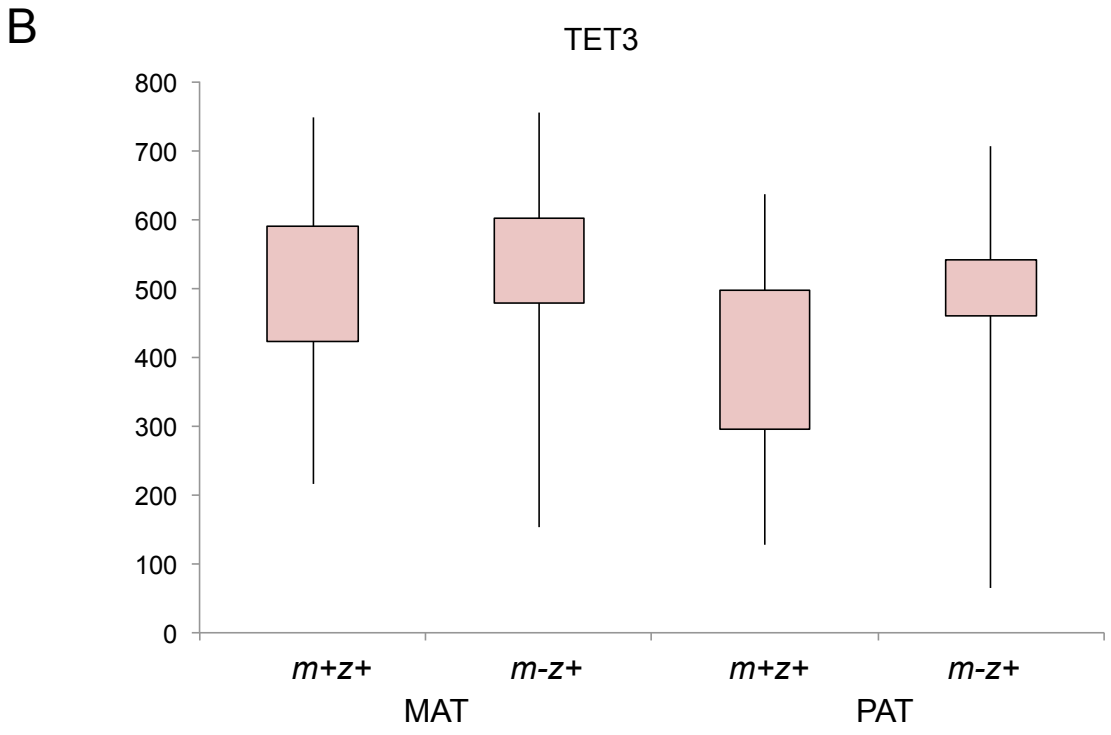

B
TET3
800
700
600
500
400
300
200
100
0
m+z+
m-z+
m+z+
m-z+
MAT
PAT

**Fig. S6. TET3 in *Ehmt2*$^{m-z+}$ maternal mutant zygotes**

(A) Confocal images of pronuclei in control and *Ehmt2*$^{m-/z+}$ zygotes stained with anti-H3K9me2, anti-H3K9me3 and TET3 antibody are displayed. Dotted or dashed arrow points to the NLB in the maternal and paternal pronucleus, respectively. Solid arrow points to an example of a strong dot that is shared between H3K9me2 and TET3. (B) Quantification of TET3 intensities in the pronuclei of wild type (n=13) and *Ehmt2*$^{m-/z+}$ mutant (n=16) zygotes.

Figure S7

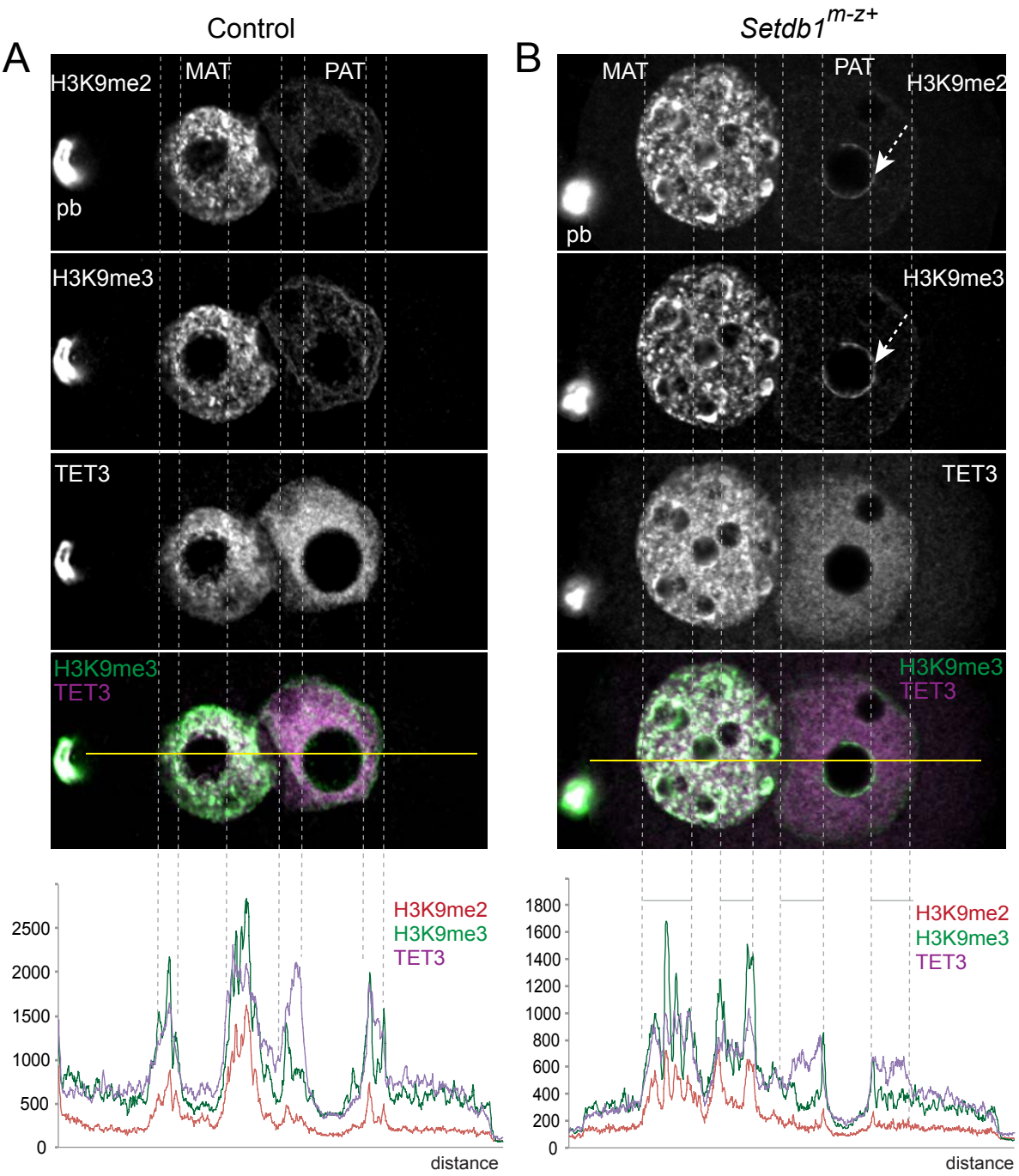
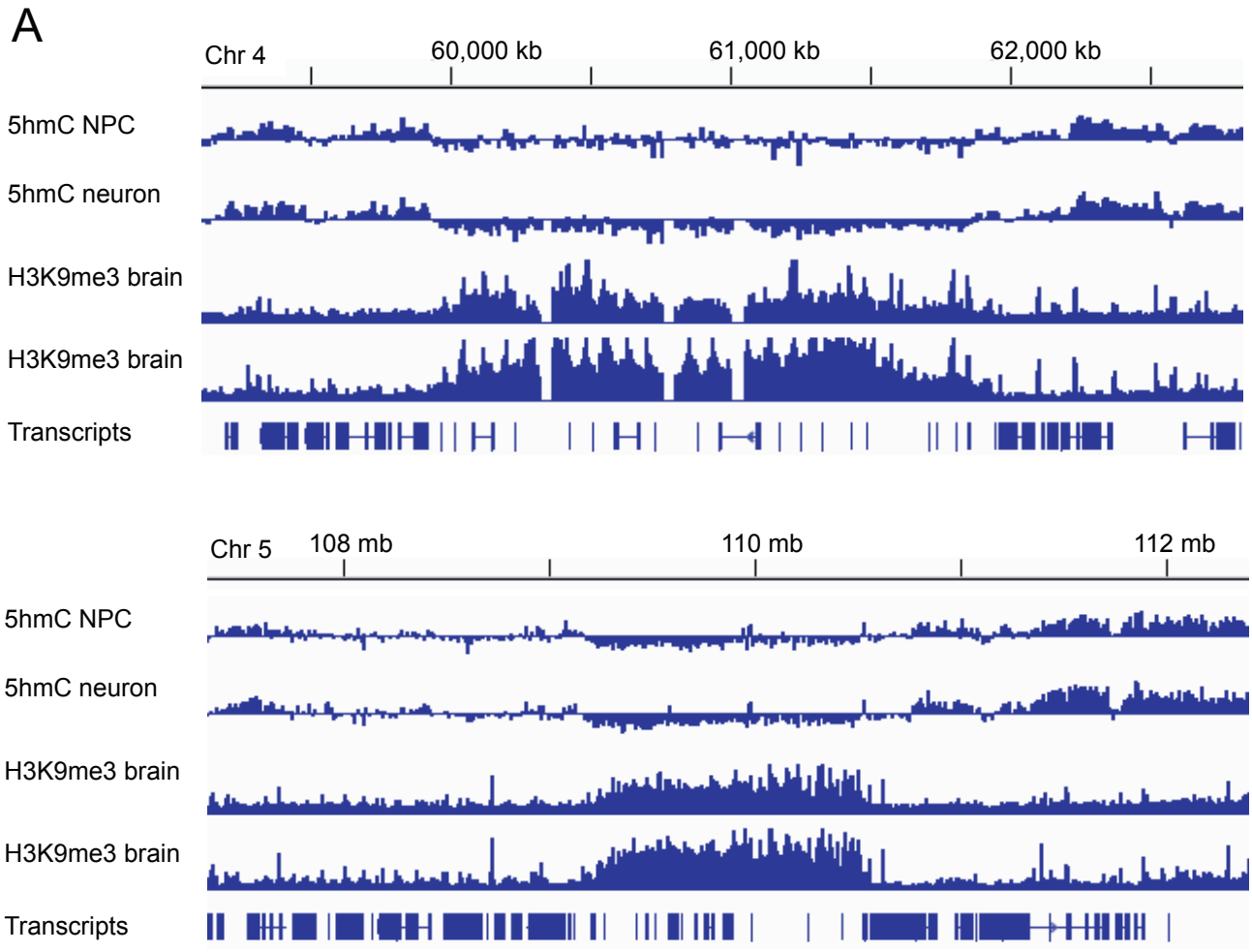

A
Control
H3K9me2
MAT
PAT
pb
H3K9me3
TET3
H3K9me3
TET3
H3K9me2
H3K9me3
TET3
2500
2000
1500
1000
500
0
distance
B
Setdb1m-z+
MAT
PAT
H3K9me2
pb
H3K9me3
TET3
H3K9me3
TET3
H3K9me2
H3K9me3
TET3
1800
1600
1400
1200
1000
800
600
400
200
0
distance

**Fig. S7. The relationship of H3K9me3 and TET3 patterns in the zygote**

Confocal images are shown of a representative control (A) and *Setdb1*$^{m\text{-}z+}$ (B) zygote, stained with anti-H3K9me2, anti-H3K9me3, and TET3 antibodies. These are the neighboring Z-sections of the ones shown in Figure 6 and display the paternal pronucleus of each zygote in focus. A merged image shows the overlap of H3K9me3 and TET3 substructures (in white). Image brightness in this Figure was optimized to aid in observing the substructures. For relative intensities between control and mutant zygotes please see Figure 3. Note an H3K9me3-rich circle at the periphery of the NLB in the mutant but not in the control paternal pronucleus (dashed arrow). Intensity profiles of the above zygotes are depicted under the microscope images. A horizontal line was drawn across each zygote to pass through first the maternal, then the paternal pronucleus and the staining intensity profiles of H3K9me2 and H3K9me3 and TET3 were plotted along this line from point of entry to point of exit of each cell by distance units. Note the similarities in the peak locations. Dashed vertical lines help orientation between microscopic images and profiles.

Figure S8


A
Chr 4
60,000 kb
61,000 kb
62,000 kb
5hmC NPC
5hmC neuron
H3K9me3 brain
H3K9me3 brain
Transcripts
Chr 5
108 mb
110 mb
112 mb
5hmC NPC
5hmC neuron
H3K9me3 brain
H3K9me3 brain
Transcripts

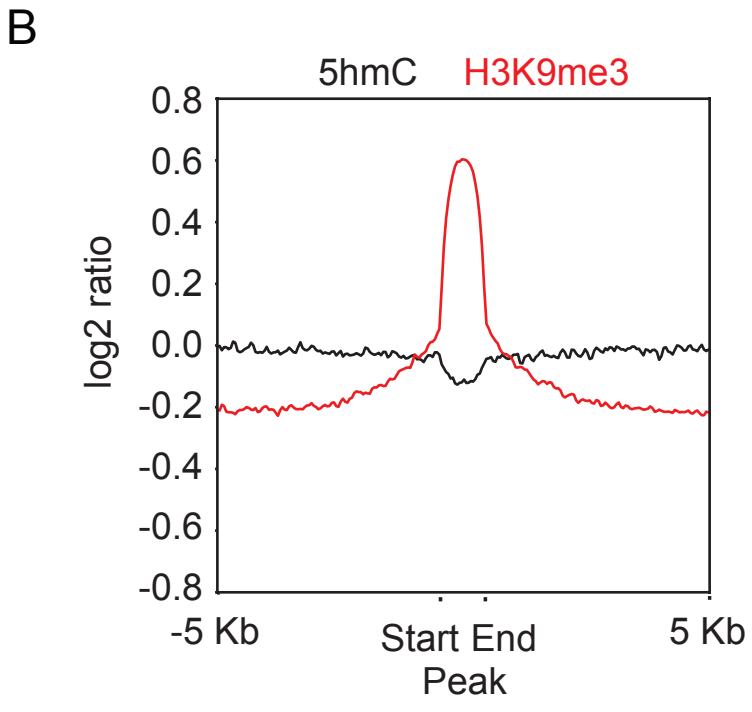

B
5hmC
H3K9me3
log2 ratio
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-5 Kb
Start End
Peak
5 Kb

**Fig. S8. TET3 activity and H3K9me3 are mutually exclusive in mouse embryo neurons**

To gain insight into the relative global patterns of H3K9me3 and 5hmC, we analyzed publically available datasets from mouse embryo neurons, which express TET3 as the major 5mC oxidase (7). We plotted the 5hmC hMeDIP profile (7) along H3K9me3 ChIPseq peaks (8) and found that these marks are mutually exclusive genome-wide, supporting the notion that H3K9me3 inhibits TET3 activity. (A) Genome-wide MeDIP mapping data of 5hmC in mouse embryo neuronal progenitor cells and neurons (7) is aligned with H3K9me3 ChIPseq results from E14.5 mouse embryo brain cells (8). Two genomic regions are shown in the igv genome browser, one along chromosome 4 and another along chromosome 5. Note that H3K9me3 enrichment coincides with 5hmC depleted sequences. (B) The genome-wide enrichment profile of 5hmC was plotted at H3K9me3 peaks and their vicinity.

**Table S1 Oligonucleotides used for mouse genotyping**

| Primers | Sequences (5' – 3') | Tm (°C) | Length |
|---|---|---|---|
| Ehmt2-2lox-U<br>Ehmt2-2lox-L | TACAGAGCCTCCCACCTGCCTGTCT<br>TCACATCAGCCTCGGCATCAGA | 64 | WT: 206 bp<br>*Ehmt2*$^{f}$ : 252 bp |
| Ehmt2-1lox-U1<br>Ehmt2-1lox-L1<br>Ehmt2-1lox-L2 | TCTCCCACCATGCCTGCTCAGG<br>ACTTGCCTGGCTAGTGTTCCTGCA<br>CGCGCCAGATCGAAGTTCCTATTC | 62 | WT: 400 bp<br>*Ehmt2*$^{-}$ : 271 bp |
| Setdb1 EMMA U1<br>Setdb1 EMMA L1<br>Setdb1 EMMA KO-L | ACCTGAGTGGAGTTGACCAGCACG<br>AAGTCAAGGATCAGAGTTCAGACGGAG TAGC<br>GCTCTAACTAGCCTAAAACTCCGTGTA ACCC | 62 | WT: 334 bp<br>*Setdb1*$^{f}$ : 400 bp<br>*Setdb1*$^{-}$ : 500 bp |
| Cre-U<br>Cre-L | TGCTGTTTCACTGGTTGTGCGGCG<br>TGCCTTCTCTACACCTGCGGTGCT | 65 | Cre: 304 bp |

**Table S2 Antibodies used for immunofluorescence staining**

| Host Species | Antibody | Vendor, cat. No. | Dilution | Secondary antibody |
|---|---|---|---|---|
| Rabbit | anti-5hmC | Active motif, 39791 | 1:2000 | Goat anti-Rabbit IgG Alexa Fluor® 488 (Invitrogen, A-11034) |
| | anti-5fC | Active motif, 61227 | 1:2000 | |
| | anti-5caC | Active motif, 61229 | 1:2000 | |
| | anti-H3K9me3 | Millipore, 07-442 | 1:200 | |
| | anti-EHMT2/G9a | Cell Signaling, 68851T | 1:400 | |
| | anti-SETDB1/ESET | Santa Cruz, sc-66884 | 1:60 | |
| Mouse | anti-5mC | Eurogentec, BI-MECY-0100 | 1:2000 | Goat anti-Mouse IgG Alexa Fluor® 568 (Invitrogen, A-11031) |
| | anti-H3K9me2 | Abcam, ab1220 | 1:200 | |
| Rat | anti-TET3 | Active motif, 61743 | 1:200 | Goat anti-Rat IgG Alexa Fluor® 647 (Invitrogen, A-21247) |

**References**


1. Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods* 9(4):357-359.
2. Li H*, et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078-2079.
3. Heinig M*, et al.* (2015) histoneHMM: Differential analysis of histone modifications with broad genomic footprints. *BMC bioinformatics* 16:60.
4. Quinlan AR & Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841-842.
5. Ramirez F*, et al.* (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44(W1):W160-165.
6. Tang SH, Silva FJ, Tsark WM, & Mann JR (2002) A Cre/loxP-deleter transgenic line in mouse strain 129S1/SvImJ. *Genesis* 32(3):199-202.
7. Hahn MA*, et al.* (2013) Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Rep* 3(2):291-300.
8. Kato M, Takemoto K, & Shinkai Y (2018) A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing. *Nature communications* 9(1):1683.