

Supplementary Information

A stepwise mode of TGF β -SMAD signaling and DNA methylation regulates naïve-to-primed pluripotency and differentiation

Bingnan Zhao^{1,2*}, Xiuwei Yu^{1,2*}, Jintong Shi^{3*}, Shuangyu Ma^{1,2*}, Shizhao Li^{1,2*}, Haitao Shi¹, Shoubing Xia^{1,2}, Youqiong Ye³, Yongchun Zhang^{4#}, Yanhua Du^{3#}, Qiong Wang^{1,2#}

¹ Department of Histoembryology, Genetics and Developmental Biology, Shanghai Key Laboratory of Reproductive Medicine, Key Laboratory of Cell Differentiation and Apoptosis of Chinese Ministry of Education, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China.

² Shanghai Frontiers Science Center of Cellular Homeostasis and Human Diseases, Shanghai 200025, China.

³ Shanghai Institute of Immunology, State Key Laboratory of Oncogenes and Related Genes, Department of Immunology and Microbiology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China.

⁴ State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

* These authors contributed equally to this work.

Correspondence

wangqiong@shsmu.edu.cn

yhdu@shsmu.edu.cn

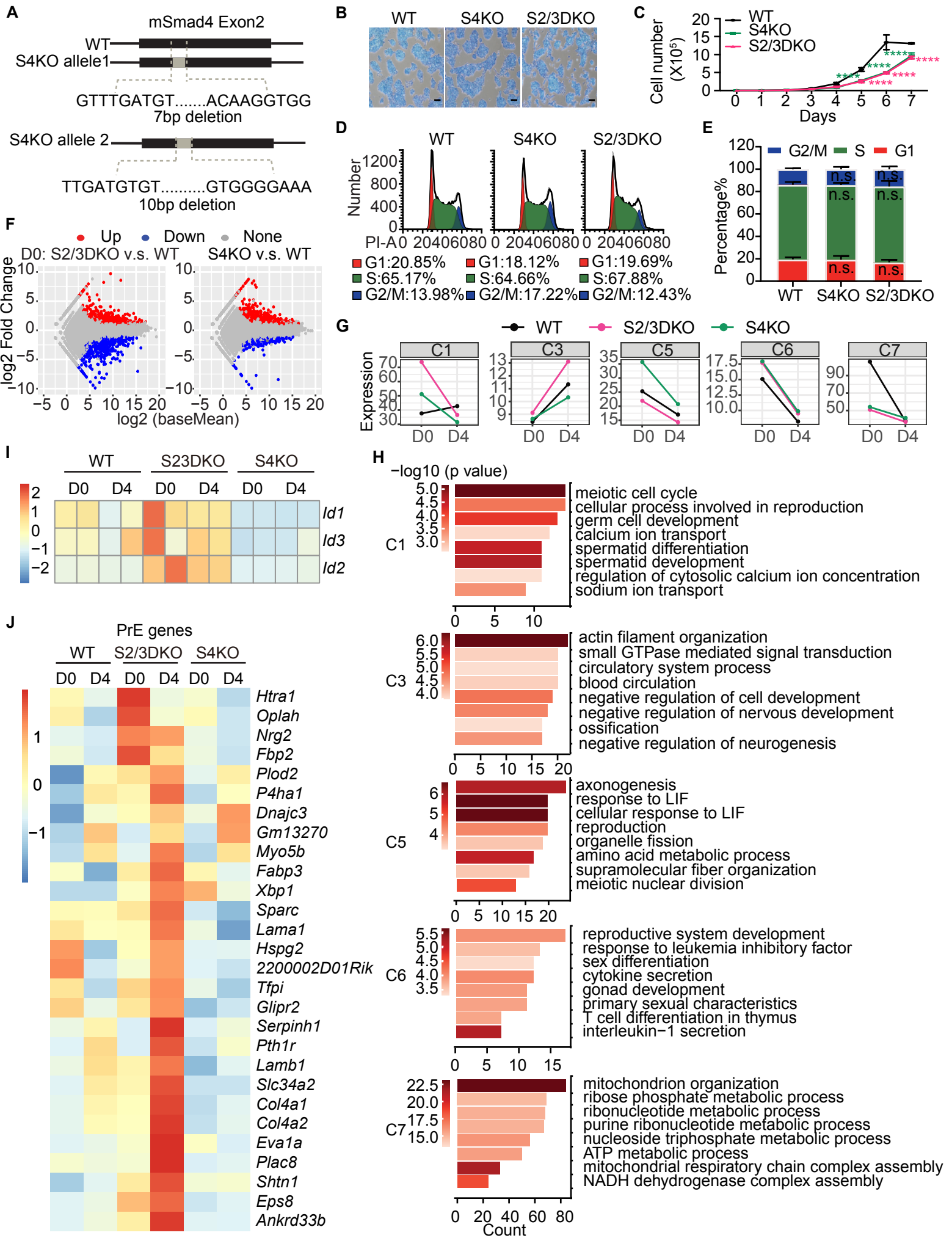
yongchun_zhang@sjtu.edu.cn

The PDF file includes:

Supplementary Figures 1 to 6

Uncropped Western Blots

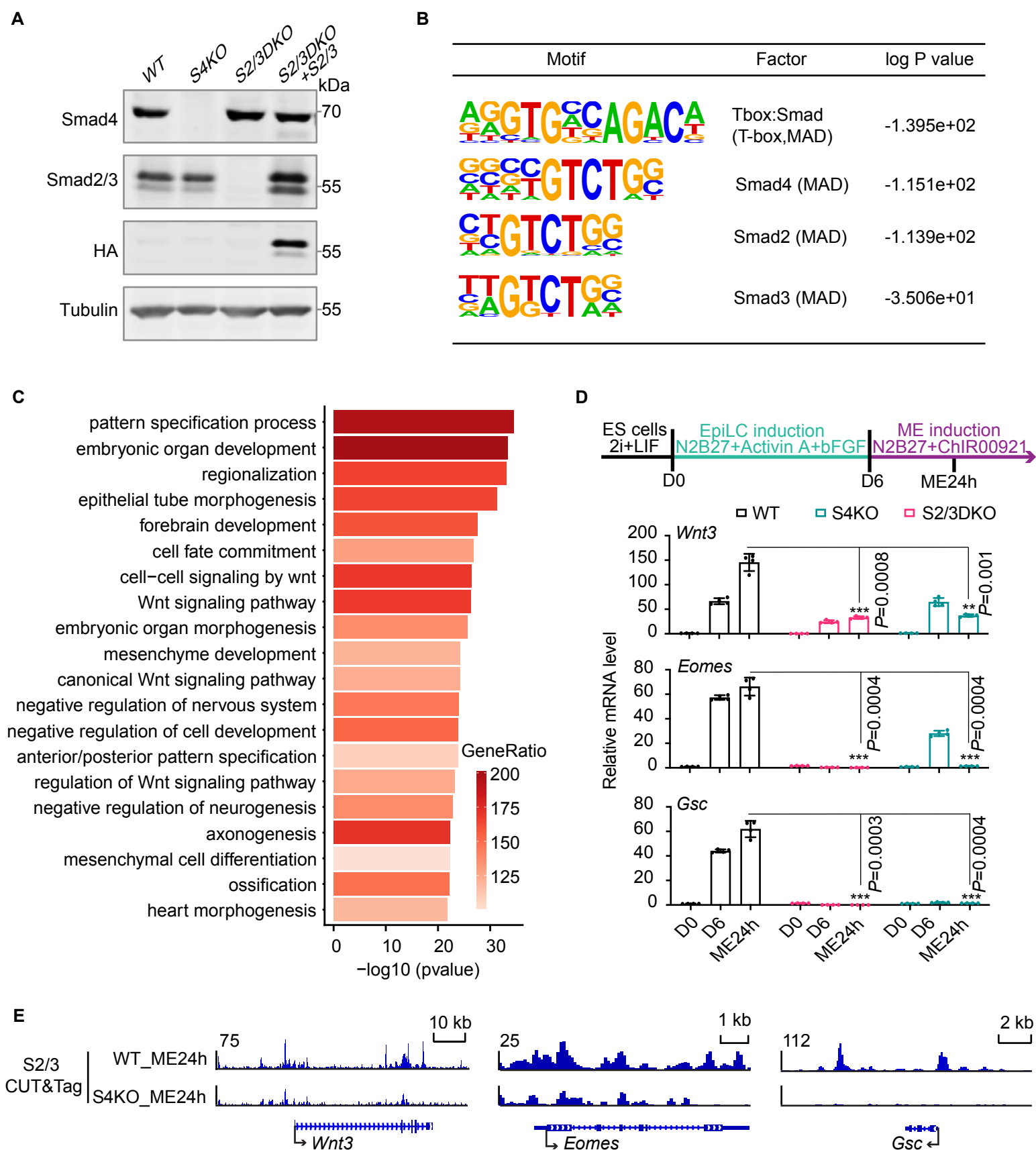
Supplementary Figure 1



Supplementary Figure 1. S2/3DKO displays a more severe phenotype compared to S4KO during EB differentiation.

- A. The genomic sequences of Smad4 exon2 locus in S4KO mutant clones were compared to WT.
- B. Representative alkaline phosphatase (AP) staining of mESC in the absence of Smad2/3 or Smad4. Scale bar: 200 μ m.
- C. The growth curve of mESC. Cells were cultured with LIF and serum for 7 days. $n = 3$ independent experiments for each condition. Data are presented as mean values $\pm$ SD. P value was calculated by two-way ANOVA. ****: $P < 0.0001$.
- D. Cell cycle distributions of mESC were analyzed by ModFit LT. G1 phase (red), S phase (green), and G2/M phase (blue) with percentage were indicated here.
- E. The percentage of each cell-cycle phase was shown in the bar graph. $n = 3$ independent experiments. P value was calculated by the t-test. n.s., no significance.
- F. MA plot of RNA-seq datasets of mESCs (D0) derived from WT, S2/3DKO or S4KO. Genes upregulated (red, $\log_2$ fold-change ≥ 1) or downregulated (blue, $\log_2$ fold-change ≤ -1) in S2/3DKO or S4KO compared to WT. Two biological replicates for each condition were analyzed.
- G. The dot-line plots showed the average gene expression patterns of C1, C3, C5, C6, and C7 in Figure 1E.
- H. Gene Ontology analysis of C1, C3, C5, C6, and C7 in Figure 1E. The top enriched biological processes GO terms with P values were shown here.
- I. Heatmap of *Id1*, *Id2*, and *Id3* mRNA expression levels (RNAseq) in ESCs (D0) and D4 EBs of WT, S2/3DKO and S4KO. The gene expression was z-scaled by row.
- J. Heatmap of PrE genes expression from RNAseq in ESCs (D0) and D4 EBs of WT, S2/3DKO and S4KO. The gene expression was z-scaled by row.

Supplementary Figure 2



Supplementary Figure 2. Both Smad2/3 and Smad4 are required for mesendoderm differentiation

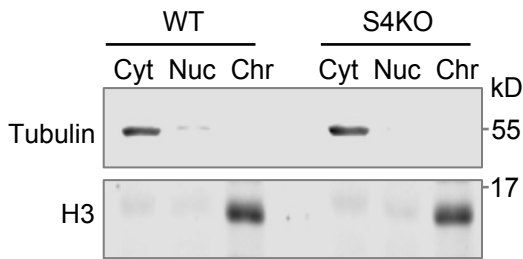
- A. Western blot analysis of Smad2/3, Smad4, HA protein levels in WT, S4KO, S2/3DKO and S2/3DKO cells with overexpression of HA-tagged Smad2/3 (S2/3DKO+S2/3). Tubulin serves as a loading control.
- B. Transcription factor motif analysis revealed enrichment of binding elements for Smad2, Smad3, and Smad4 in these 7,926 Smad4-independent Smad2/3 peaks.
- C. Gene ontology analysis of these 7,926 Smad4-independent Smad2/3 peaks. The top enriched biological processes GO terms with *P* values were shown here.
- D. qRT-PCR showed the mRNA levels of mesendoderm genes (*Wnt3*, *Eomes*, *Gsc*) during the transition from mESC to EpiLC and ME (D0, D6, ME24h). Data are presented as mean values +/- SD. n=4 independent experiments. ***: $P < 0.001$; **: $P < 0.01$.
- E. Gene track of the mesendoderm genes (*Wnt3*, *Eomes*, *Gsc*) loci from Smad2/3 CUT&Tag in ME24h of WT and S4KO.

Supplementary Figure 3

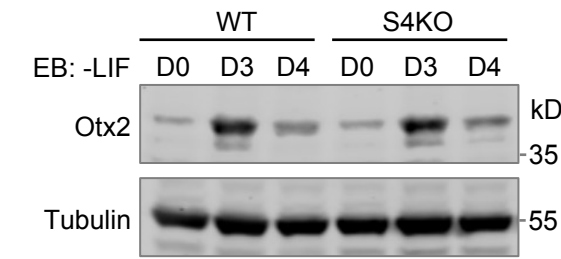
A

Protein	WT	S4KO	S4KO+AC	S4KO+SB
Pomgnt2	0	353603.3438	1265529.669	0
MyI9	0	1288687.75	620234.9174	0
Myo1a	0	1204575.75	519709.3205	0
Slc7a5	0	714238.0625	501610.5552	0
Dnmt3b	0	285801.4375	332614.4778	0
Ccdc61	0	368341.5	127232.194	0
Ppp1cc	0	206684.5156	56176.29524	0
Kctd5	0	464969.75	307307.274	79790.78162
Rraga	0	242838.0313	215242.3745	105434.6949
Skil	0	1000237.5	2167856.504	204988.4954
Tubb6	1170875.86	21817956	13512410.29	7532996.944
Dsp	2703906.985	15410369.25	10227418.56	5572438.055
Hspa1l	305960.5307	3271339	1736724.294	973270.434
Gars	180249.098	553821.5469	801183.1278	184378.0321
Pgk1	136438.4752	1049819	571511.7416	331876.7119
Prpf40a	681057.9579	194369.7031	206584.1167	300419.92

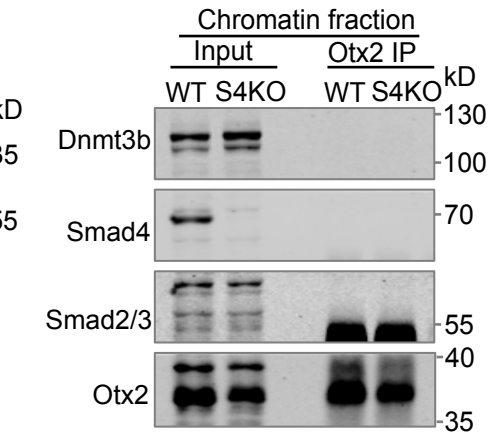
B



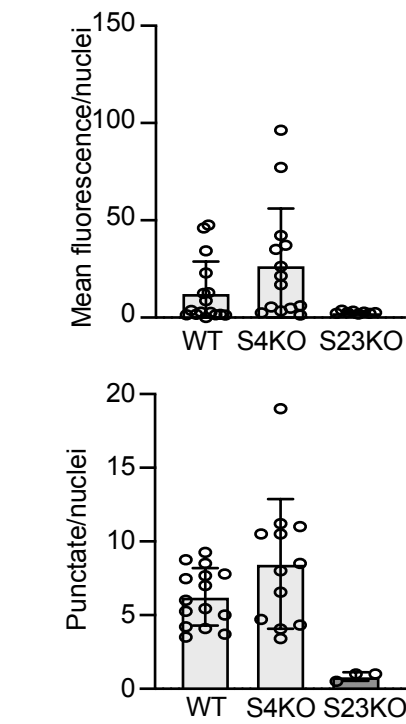
C



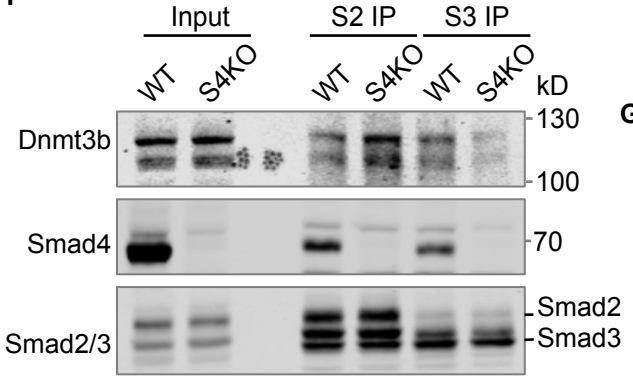
D



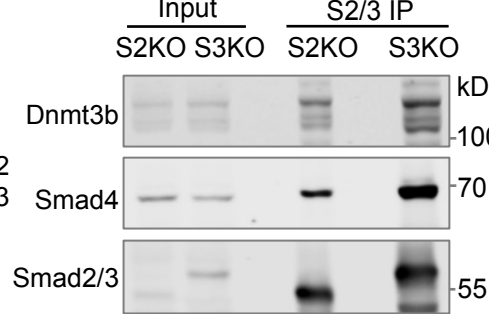
E



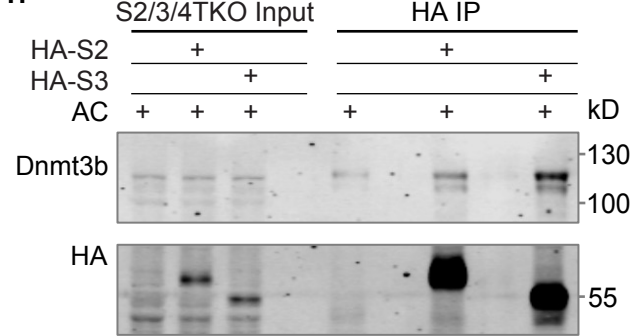
F



G



H

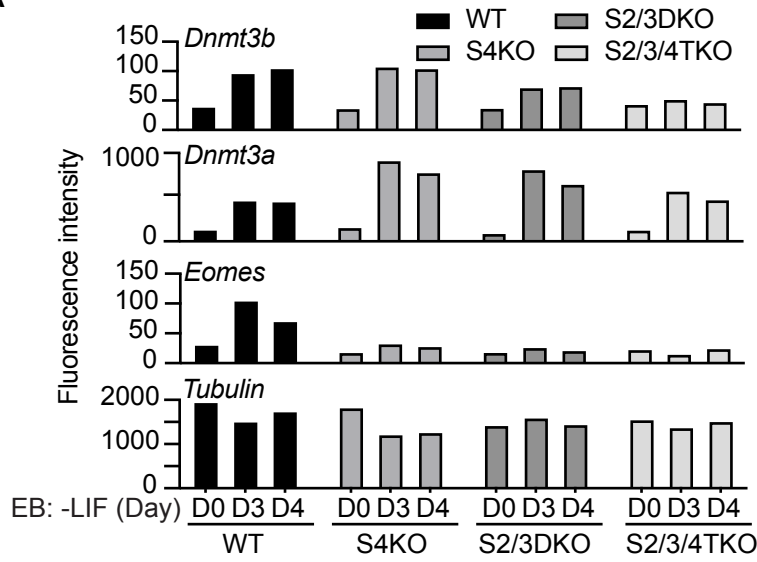


Supplementary Figure 3. Dnmt3b interacts with Smad2/3 in different contexts.

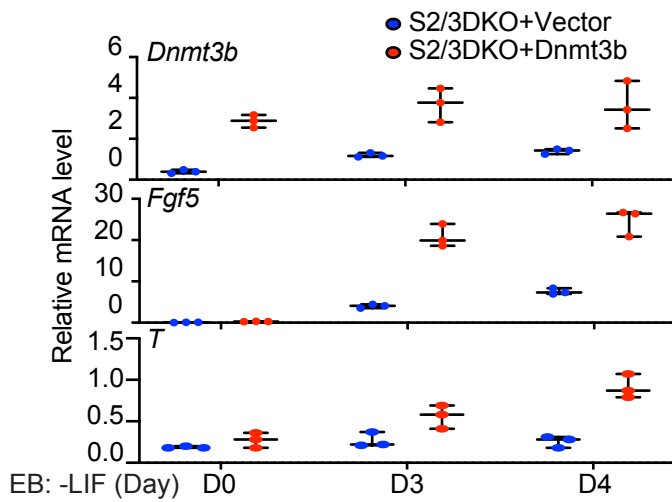
- A. Relative intensity values of 16 proteins from Smad2/3 IP-MS experiment in WT and S4KO EBs, treated with or without SB or AC.
- B. Cytoplasm (Cyt), nucleoplasm (Nuc) and chromatin (Chr) fractions were analyzed by western blot. Tubulin and H3 were assayed as cytoplasm and chromatin markers, respectively.
- C. Western blot displayed the protein levels of Otx2 in WT and S4KO during EB differentiation (D0, D3, D4). Tubulin serves as a loading control.
- D. Western blot analysis of Dnmt3b, Smad2/3, Smad4 and Otx2 in chromatin fraction of WT and S4KO EBs after immunoprecipitation using anti-Otx2 antibody.
- E. Quantitation of PLA signal in Figure 3E. Each dot represents the signal measurement from indicated samples. Mean fluorescence intensities or the number of punctates were calculated using ImageJ and normalized to the number of nuclei.
- F. Dnmt3b, Smad4 and Smad2/3 were analyzed by western blot after Smad2 and Smad3 immunoprecipitation in WT and S4KO EBs.
- G. Smad2/3 immunoprecipitation in S2KO and S3KO D3 EBs under AC addition were analyzed by western blot using Dnmt3b, Smad2/3 and Smad4 antibodies.
- H. Western blot analysis of Dnmt3b and HA expression for the anti-HA immunoprecipitation in Smad2/3/4TKO cells, in which HA tagged Smad 2 or 3 (HA-S2 and HA-S3) were overexpressed with AC stimulation.

Supplementary Figure 4

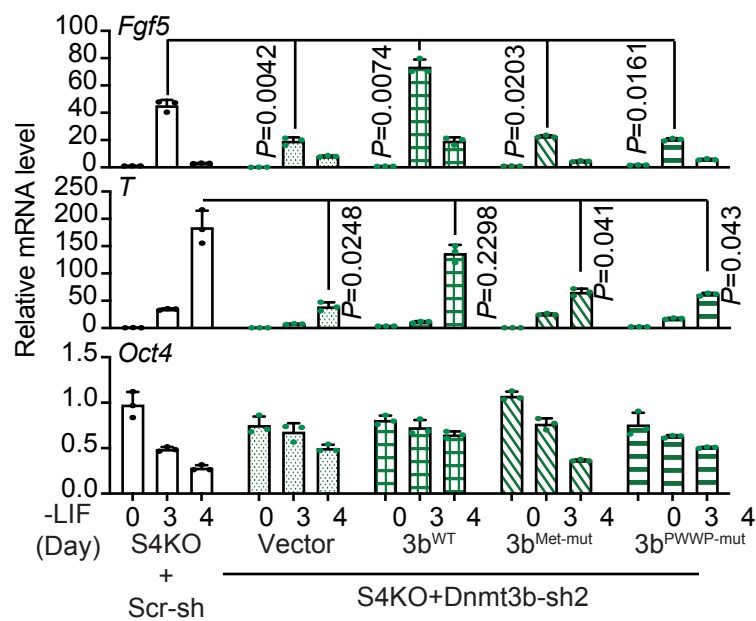
A



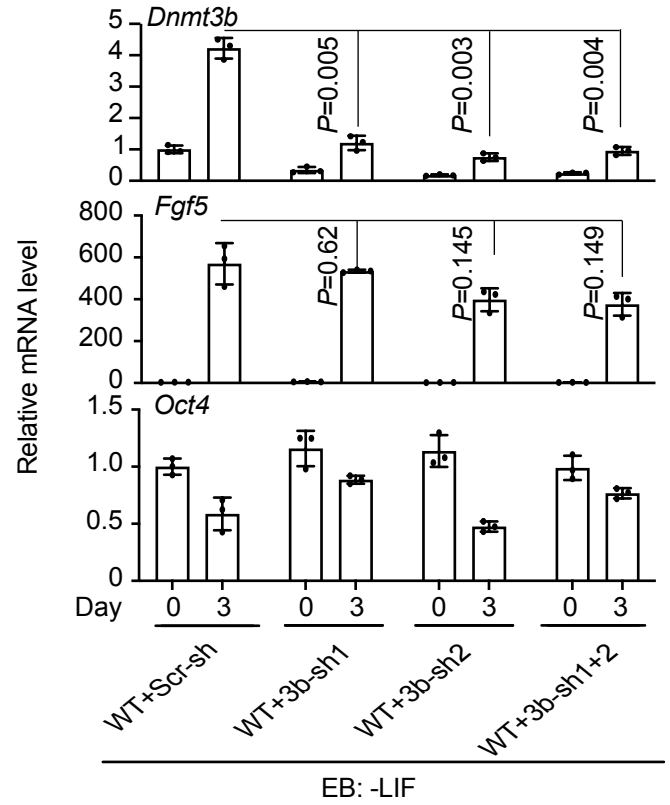
B



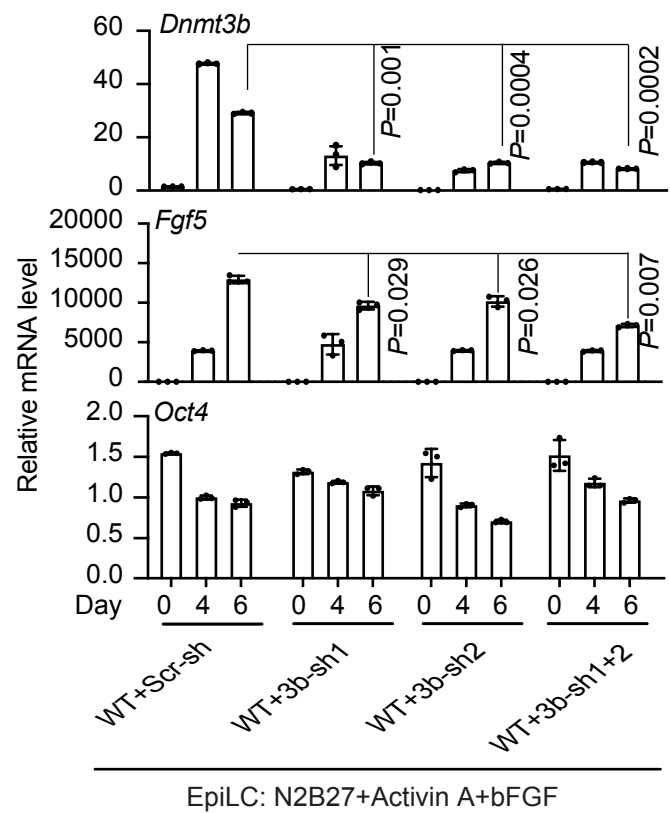
E



C



D



Supplementary Figure 4. Dnmt3b partially rescue epiblast gene expression in S2/3DKO EBs.

A. Plot of fluorescence intensity of Dnmt3b, Dnmt3a, Eomes and Tubulin bands shown in Figure 4A determined by an Odyssey imaging system for each indicated sample.

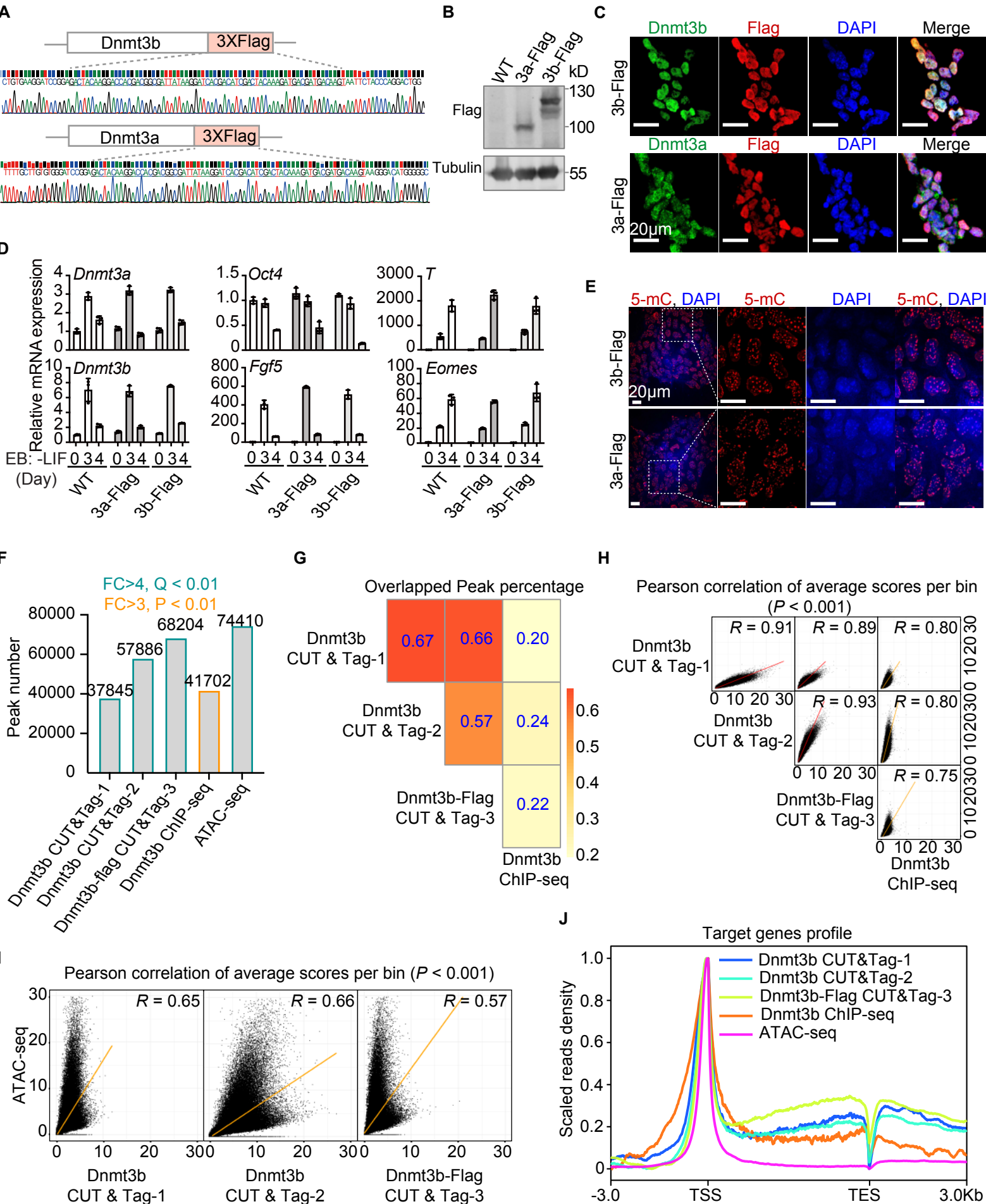
B. qRT-PCR analysis of *Dnmt3b*, *Fgf5* and *T* expression in S2/3DKO expressing Dnmt3b or empty vector during EB differentiation (D0, D3, D4). Data are presented as mean values $\pm$ SD. n=3 independent experiments.

C. The mRNA levels of *Dnmt3b*, *Fgf5* and *Oct4* when Dnmt3b was silenced with shRNA in D0 or D3 WT EBs. Data are presented as mean values $\pm$ SD. n=3 independent experiments. The *P* value of each sample was compared to WT and calculated by paired t-test.

D. Expression of *Dnmt3b*, *Fgf5* and *Oct4* during EpiLC transition (D0, D4, D6) after Dnmt3b knockdown in WT were measured by qRT-PCR. Data are presented as mean values $\pm$ SD. n=3 independent experiments. The *P* value of each sample was compared to WT and calculated by paired t-test.

E. The mRNA levels of *Fgf5*, *T* and *Oct4* during EB differentiation (D0, D3 and D4) were analyzed by qRT-PCR. In S4KO cells expressing Dnmt3b-shRNA2 (S4KO+Dnmt3b-sh2), plasmids were introduced to overexpress Dnmt3b^{WT}, Dnmt3b^{Met-mut}, Dnmt3b^{PWWP-mut} and an empty vector, respectively. S4KO transfected with scramble shRNA was using as a control. Data are presented as mean values $\pm$ SD. n=3 independent experiments. The *P* value of each sample was compared to S4KO and calculated by two-way ANOVA test.

Supplementary Figure 5

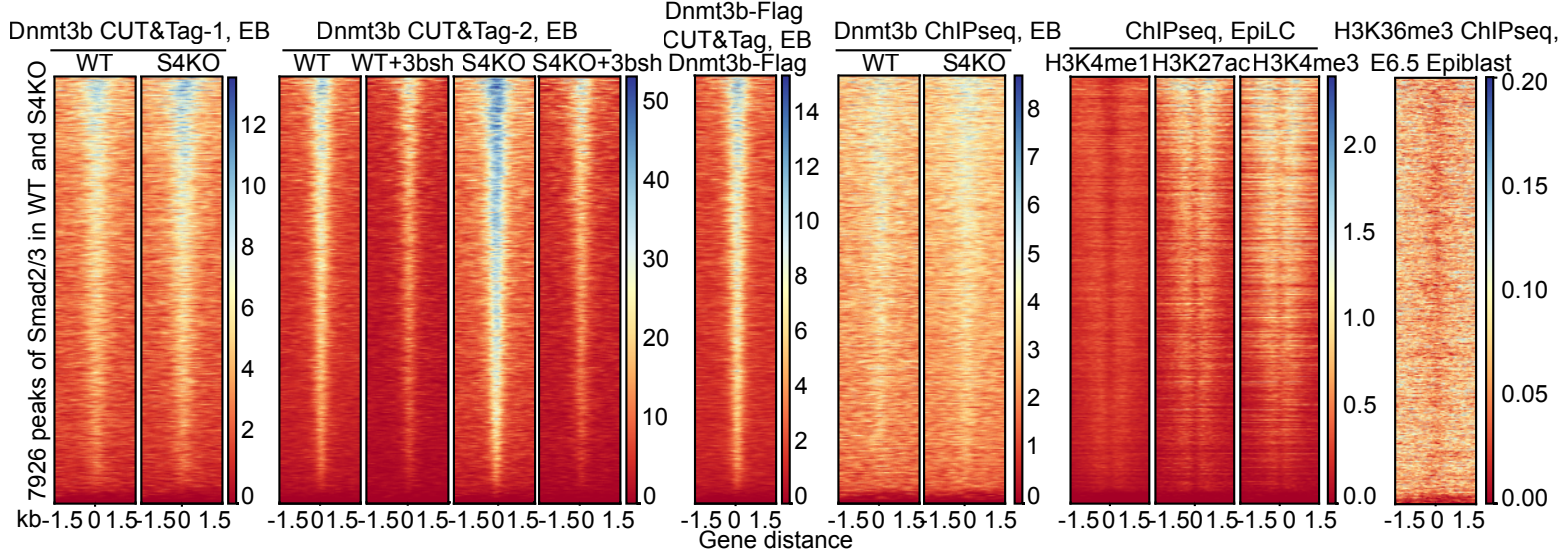


Supplementary Figure 5. The genome-wide binding profiles of Dnmt3b.

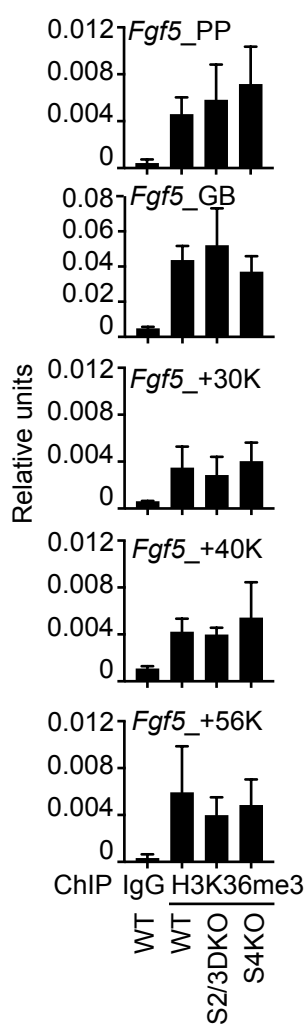
- A. A schematic representation of 3xFlag-tag insertion at the C-terminus of Dnmt3b (3b-Flag) or Dnmt3a (3a-Flag) is shown with the corresponding sequencing data displayed below.
- B. Western blot analyzed Flag-tagged proteins expression in 3a-Flag and 3b-Flag EBs (bottom). Tubulin serves as a loading control.
- C. Representative confocal microscope images of Dnmt3a or 3b (green), Flag (red) and DAPI (blue) staining in 3b-Flag and 3a-Flag EBs. Scale bars, 20 μ M. n=3 independent experiments.
- D. The mRNA levels of *Dnmt3a*, *Dnmt3b*, *Fgf5*, *T*, *Eomes* and *Oct4* during EB differentiation (D0, D3 and D4) in WT, 3a-Flag and 3b-Flag were analyzed by qRT-PCR. Data are presented as mean values $\pm$ SD. n=3 independent experiments.
- E. Representative confocal microscope images of 5-mC (Red) and DAPI (blue) staining in 3b-Flag and 3a-Flag EBs. Scale bars, 20 μ M. n=3 independent experiments.
- F. Bar plots showing the number of Dnmt3b binding peaks identified by MACS2 in three CUT & Tag assays and one ChIP-seq assay, as well as ATAC-seq peaks in D3 EBs. CUT & Tag peaks and ATAC-seq peaks were called with a fold change cutoff of 4 and a Q-value of 0.01(Cyan), while ChIP-seq peaks used a fold change cutoff of 3 and a P-value of 0.01(yellow).
- G. Heatmap displays the overlap percentages of Dnmt3b binding peaks between each pair of experiments as indicated.
- H. Scatterplots show the Pearson correlation (R) of signal intensity for Dnmt3b binding profiles across the whole genome, which were divided into 1 kb bins, between each pair of assays as indicated. An F-test was performed. R values were shown as indicated.
- I. Scatterplots showing the Pearson correlation (R) of signal intensity were compared between CUT&Tag assays for Dnmt3b binding profiles and the ATAC-seq assay across the whole genome, which were divided into 1 kb bins. An F-test was performed. R values were shown as indicated.
- J. Normalized signal density profiles were across 14,461, 16,684, and 17,170 target genes for three Dnmt3b CUT&Tag assays, 14,751 target genes for the Dnmt3b ChIP-seq assay, and 18,120 target genes for the ATAC-seq assay. TSS, transcription start site; TES, transcription end site.

Supplementary Figure 6

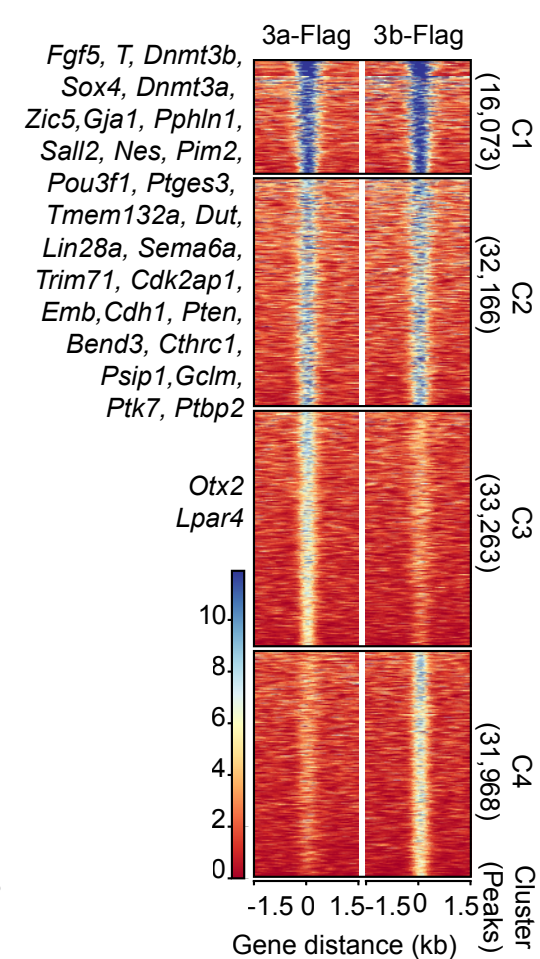
A



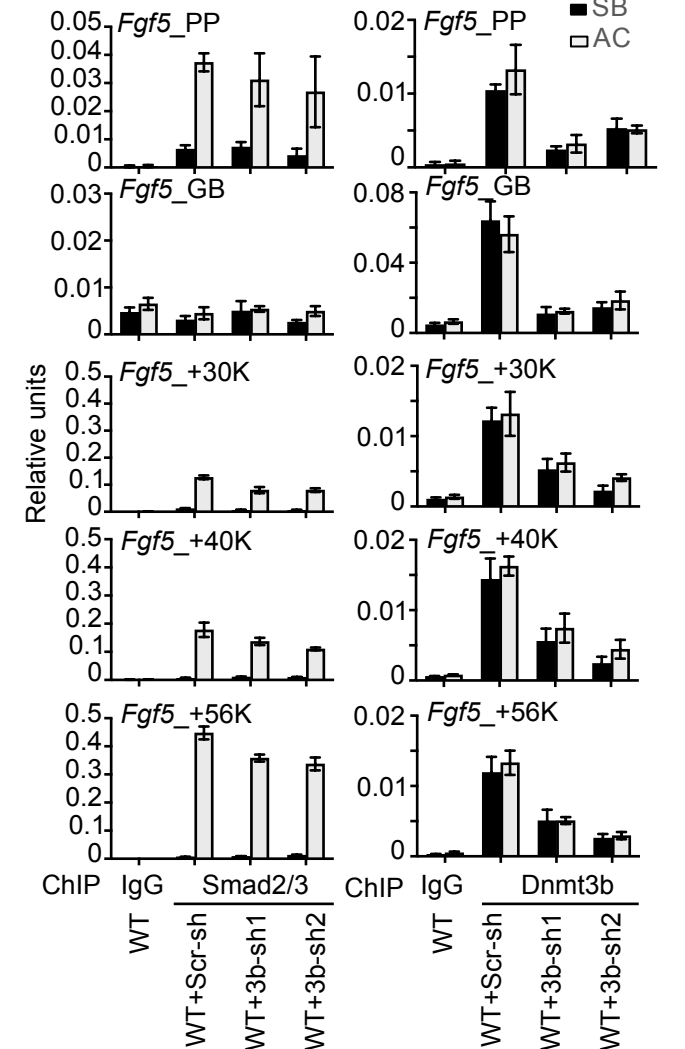
B



C



D



Supplementary Figure 6. Dnmt3b and Smad2/3 converge on epiblast genes across the genome.

- A. Heatmaps of CUT&Tag and ChIP-seq tag densities were within -1.5/+1.5kb genomic regions surrounding the centers of these 7,926 Smad4-independent Smad2/3 binding peaks.
- B. ChIP-qPCR analysis of H3K36me3 binding to *Fgf5*_PP, *Fgf5*_GB, *Fgf5*_+31k, *Fgf5*_+40k, and *Fgf5*_+56k sites in D3 EBs. Non-specific IgG ChIP-qPCR served as a negative control. This experiment was repeated 3 times, and a representative data was shown here. Data are presented as mean values +/- SD.
- C. Heatmap of Flag CUT&Tag densities for 3a-Flag and 3b-Flag in EBs. Representative epiblast maker genes in C1-C2 and C3 were annotated.
- D. ChIP-qPCR analysis of Smad2/3 and Dnmt3b binding to *Fgf5*_PP, *Fgf5*_GB, *Fgf5*_+30k, *Fgf5*_+40k, and *Fgf5*_+56k of D3 EBs treated with SB or AC in WT, WT+ Scr-sh, WT+3b-Sh1 or WT+3b-sh2 cell lines. ChIP with anti-IgG serves as a negative control. This experiment was repeated 3 times, and a representative data was shown here. Data are presented as mean values +/- SD.

Figure 1A

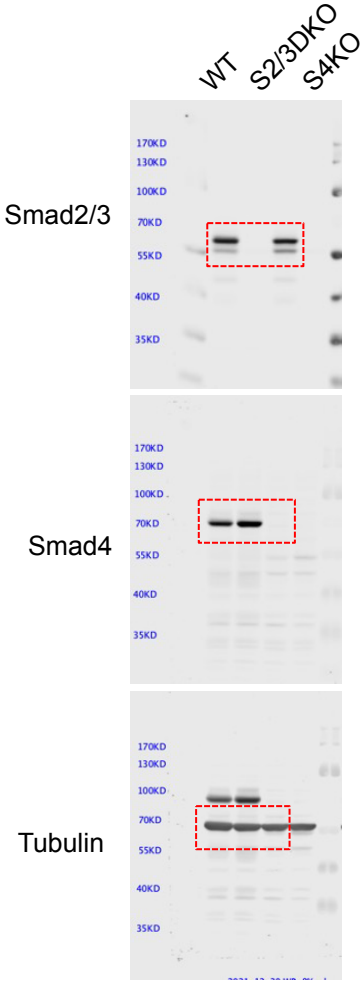


Figure 3C

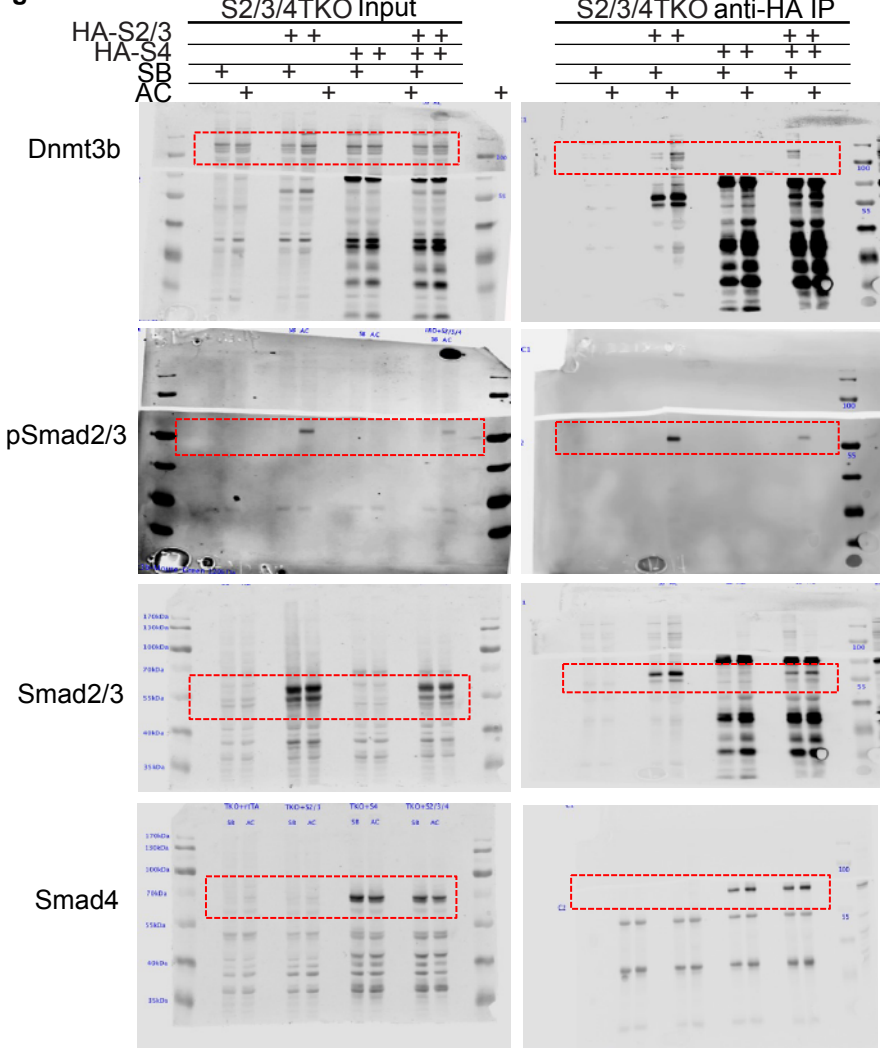


Figure 3B

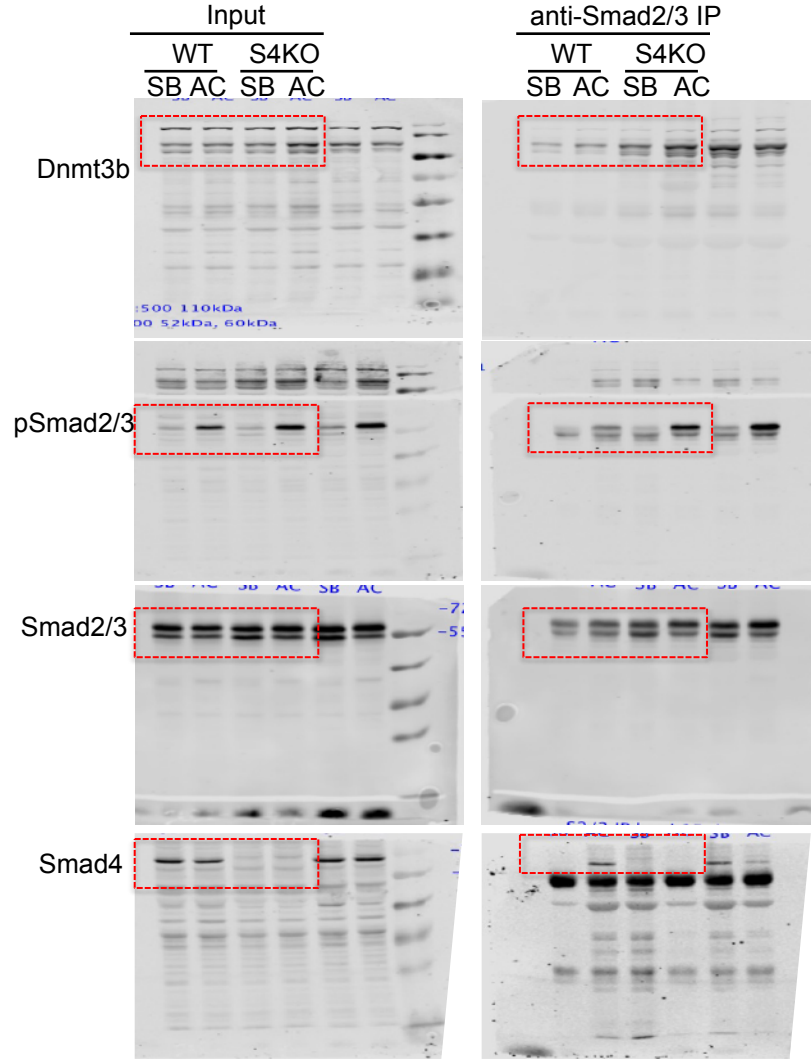


Figure 3D

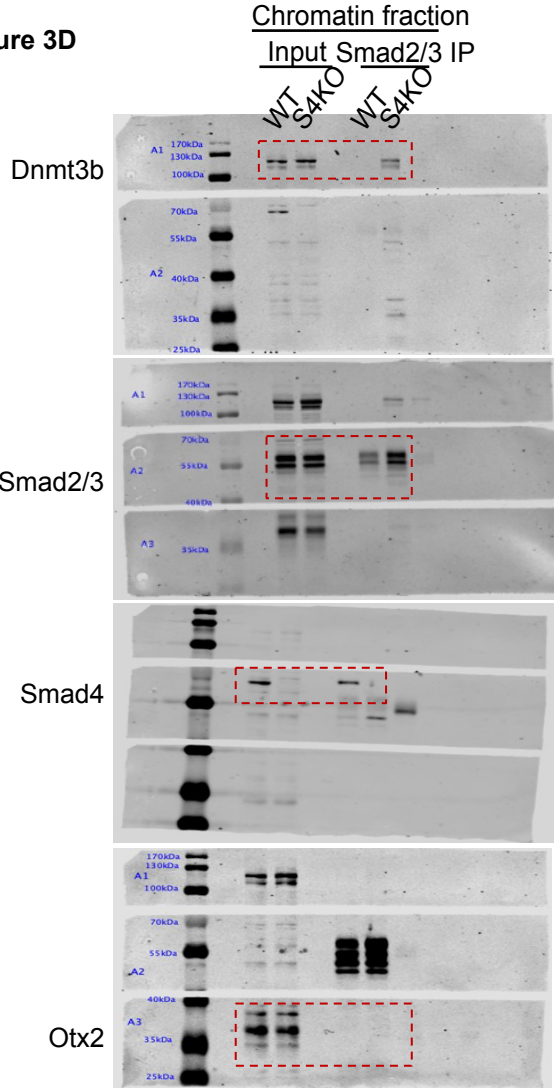


Figure 3F

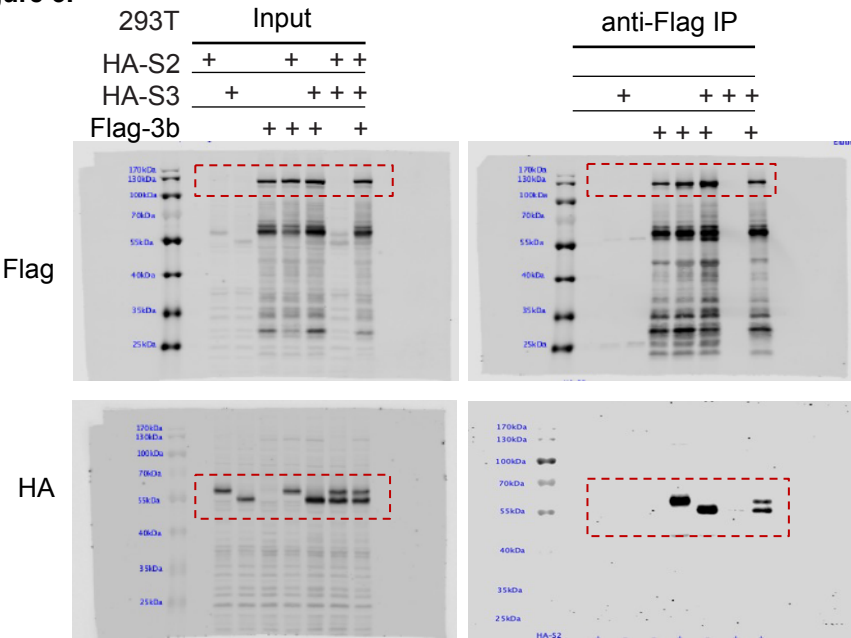


Figure 4A

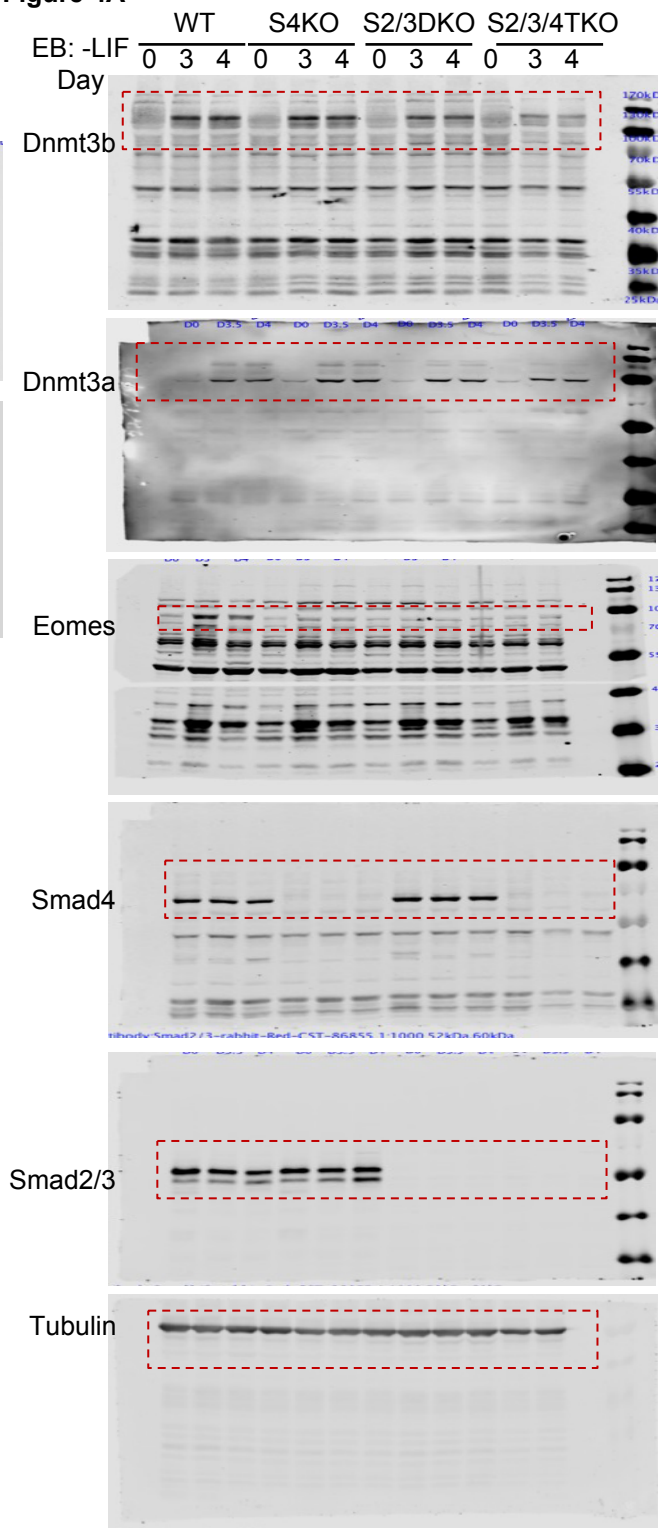


Figure 3G

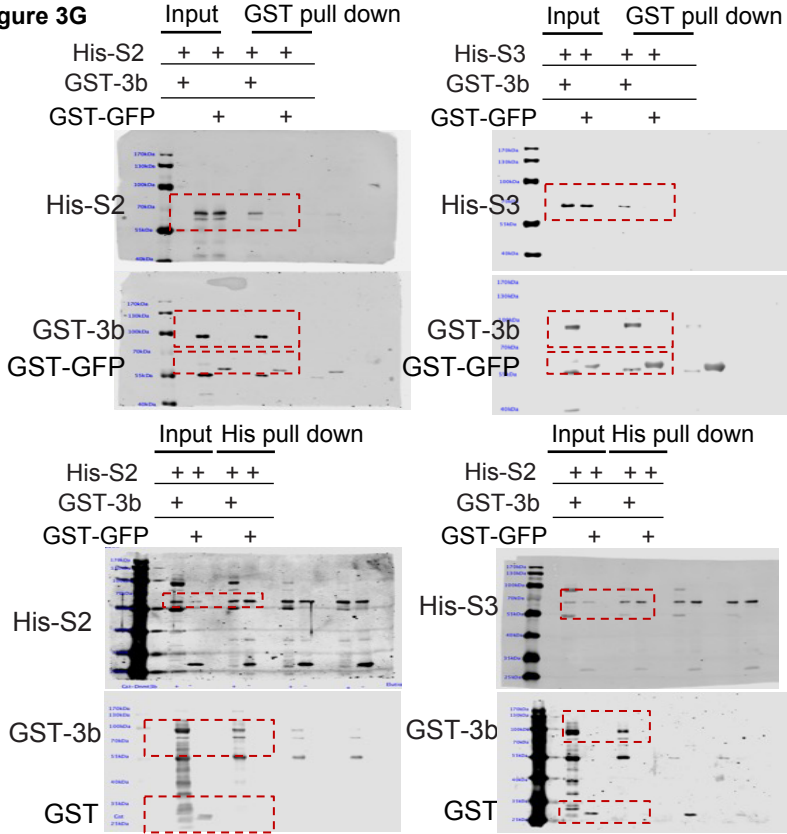


Figure 4B

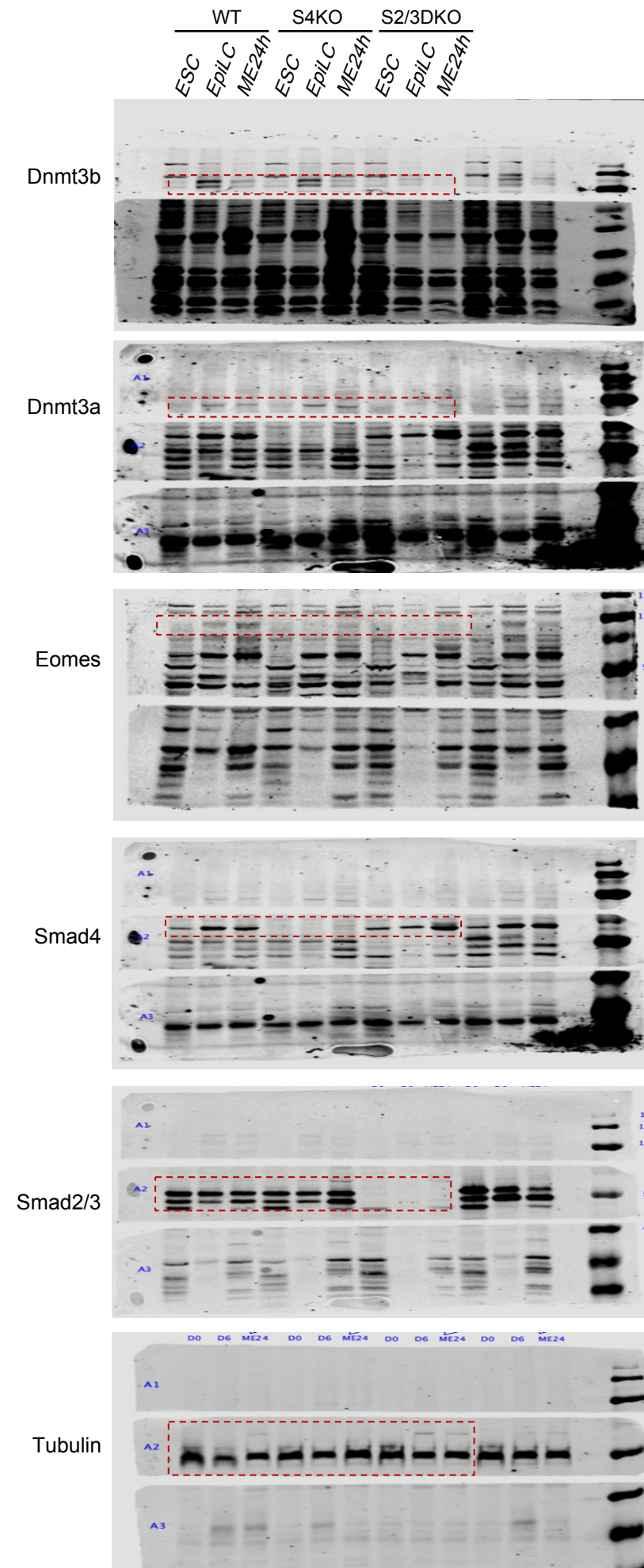


Figure S2A

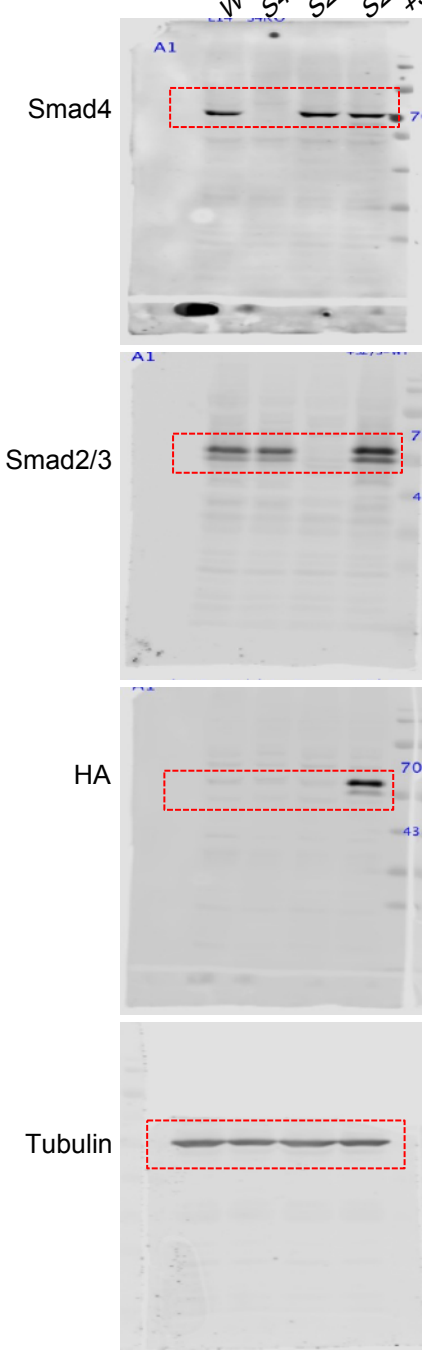


Figure S3B

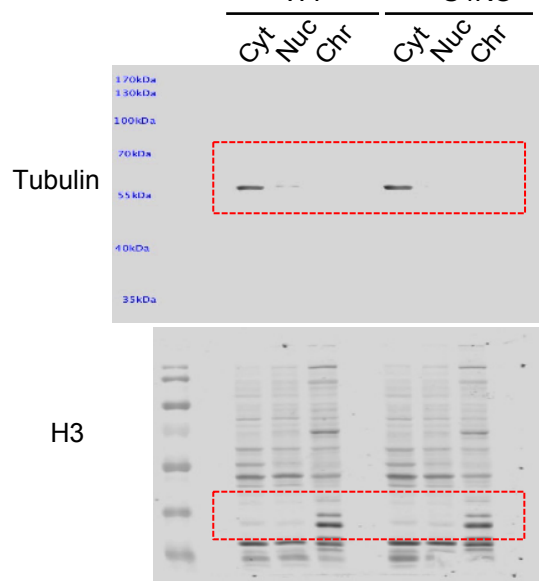


Figure S3C

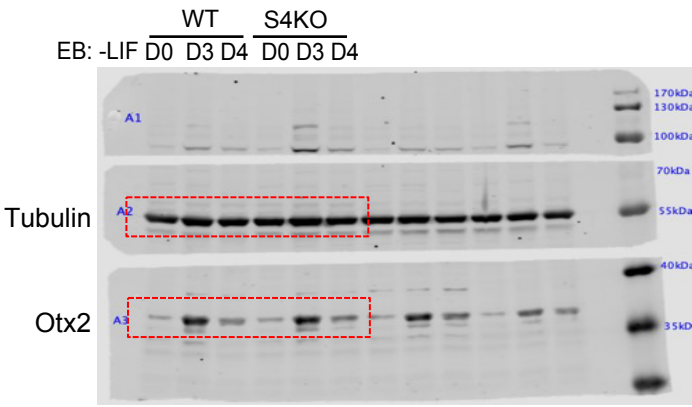


Figure S3D

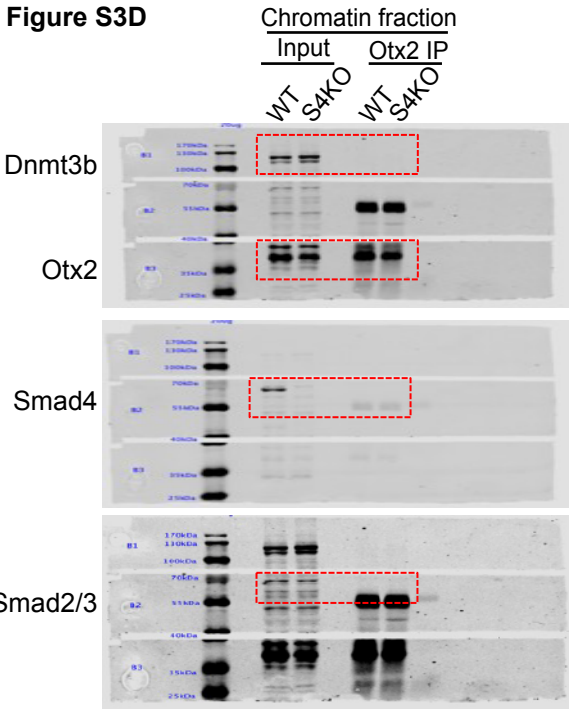


Figure S3H

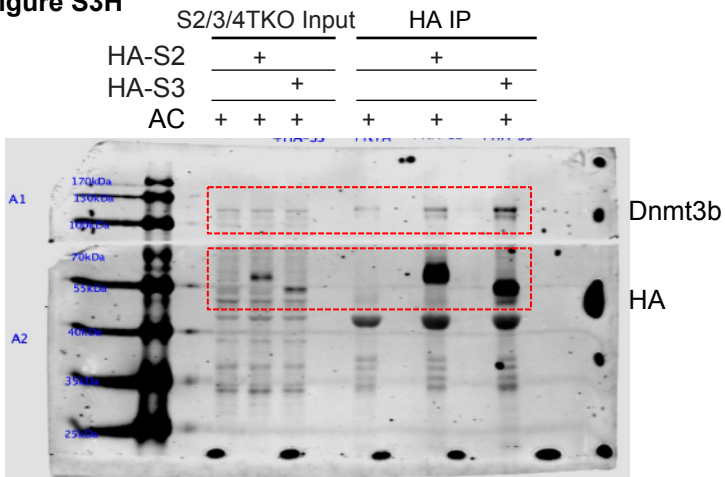


Figure S3F

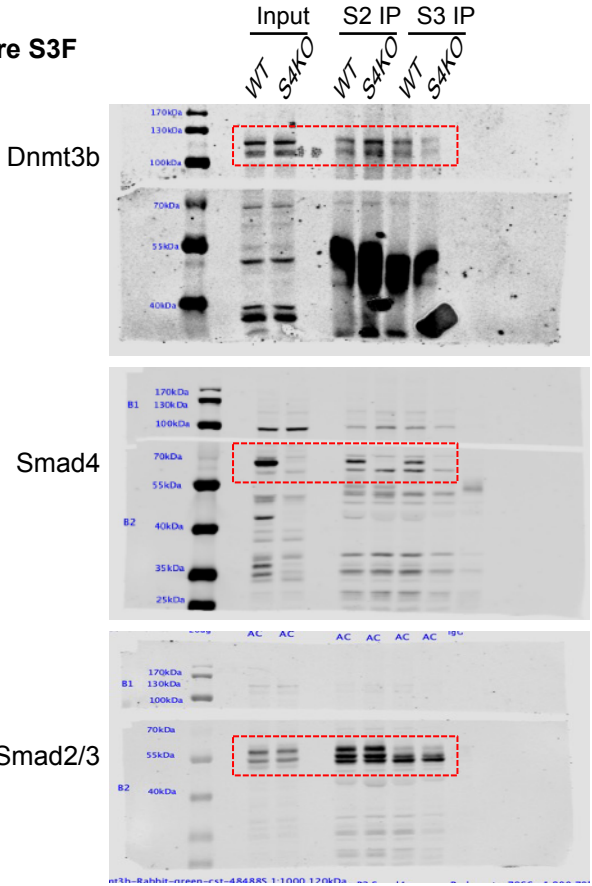


Figure S3G

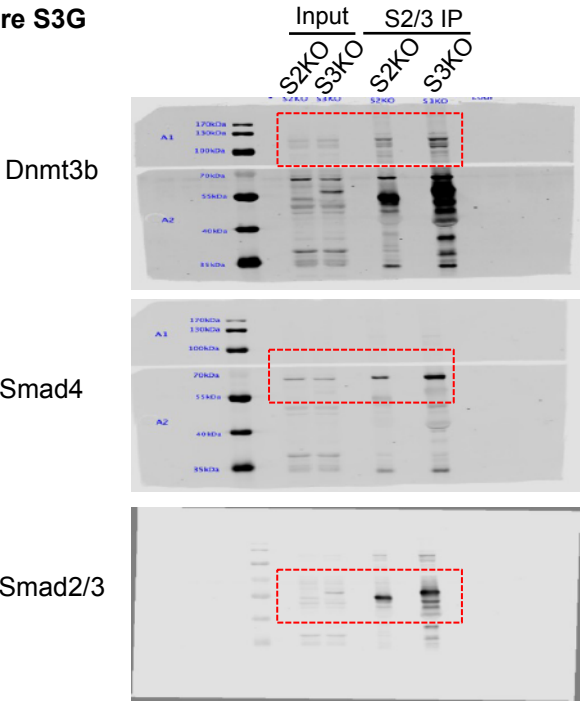


Figure S5B

