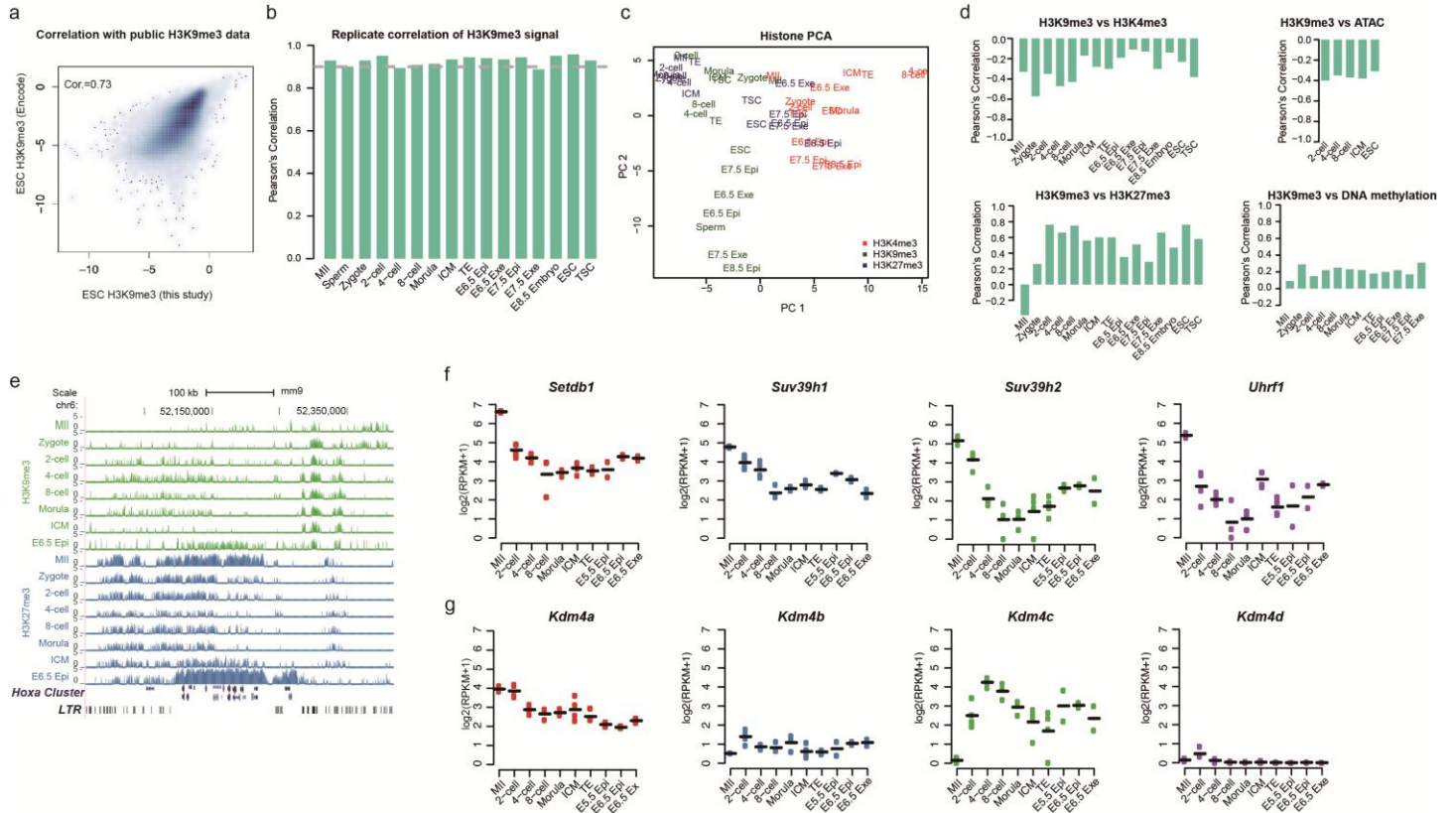


In the format provided by the authors and unedited.

Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development

Chenfei Wang^{1,2}, Xiaoyu Liu^{1,2}, Yawei Gao^{1,2*}, Lei Yang^{1,2}, Chong Li¹, Wenqiang Liu¹, Chuan Chen¹, Xiaochen Kou¹, Yanhong Zhao¹, Jiayu Chen¹, Yixuan Wang¹, Rongrong Le¹, Hong Wang¹, Tao Duan¹, Yong Zhang^{1*} and Shaorong Gao^{1*}

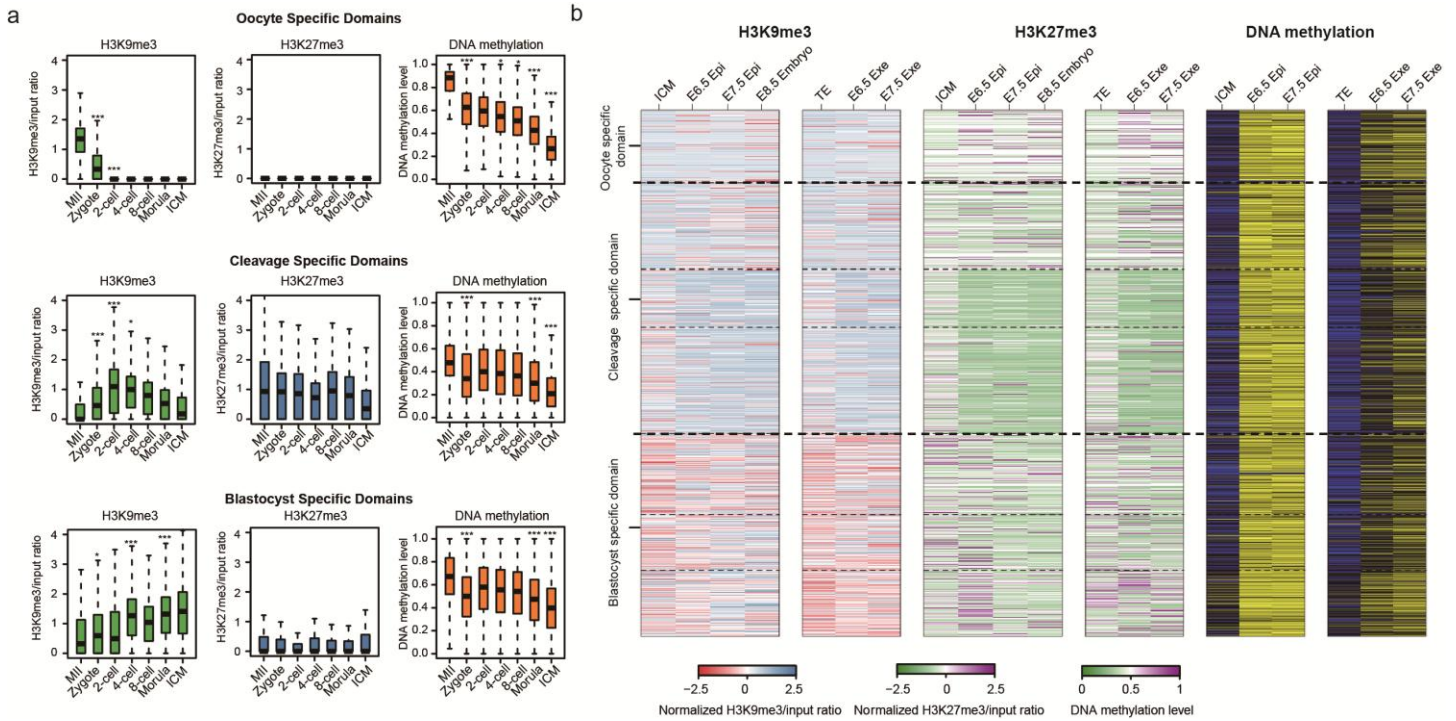
¹Clinical and Translational Research Center of Shanghai First Maternity and Infant Hospital, Shanghai Key Laboratory of Signaling and Disease Research, School of Life Sciences and Technology, Tongji University, Shanghai, China. ²These authors contributed equally: Chenfei Wang, Xiaoyu Liu, Yawei Gao and Lei Yang. *e-mail: gaoyawei@tongji.edu.cn; yzhang@tongji.edu.cn; gaoshaorong@tongji.edu.cn



Supplementary Figure 1

Distinct establishment patterns of H3K9me3 on genes and LTRs in mouse early embryos

- a, Scatter plot shows the enrichment of H3K9me3 in union peaks for mESC samples of this study and ENCODE data. Pearson's correlation coefficient is also shown.
- b, Barplot shows the Pearson's correlation coefficients between the replicates of embryo H3K9me3 samples in union peaks.
- c, PCA analysis of H3K9me3, H3K4me3 and H3K27me3 signals in union peaks (n=45 biologically independent samples).
- d, Barplots show the Pearson's correlation coefficients between H3K9me3 and other epigenetic modifications in promoter regions.
- e, The UCSC genome browser view of H3K9me3 and H3K27me3 signal around Hoxa Clusters of mouse gametes and early embryos. Signals represent log₂-transformed ChIP / input ratio.
- f, Graphs show the averaged expression level of H3K9me3 writers, including Setdb1, Suv39h1, Suv39h2 and Uhrf1.
- g, Graphs show the averaged expression level of H3K9me3 erasers, including Kdm4a, Kdm4b, Kdm4c and Kdm4d. In f and g, RNA-seq has been performed 2 times for MII oocyte, morula, E5.5 Epi, E6.5 Epi and E6.5 Exe, 3 times for 8-cell and 4 times for 2-cell, 4-cell, ICM and TE. Results for each experiment are plotted as individual points with the average value indicated by the black line.

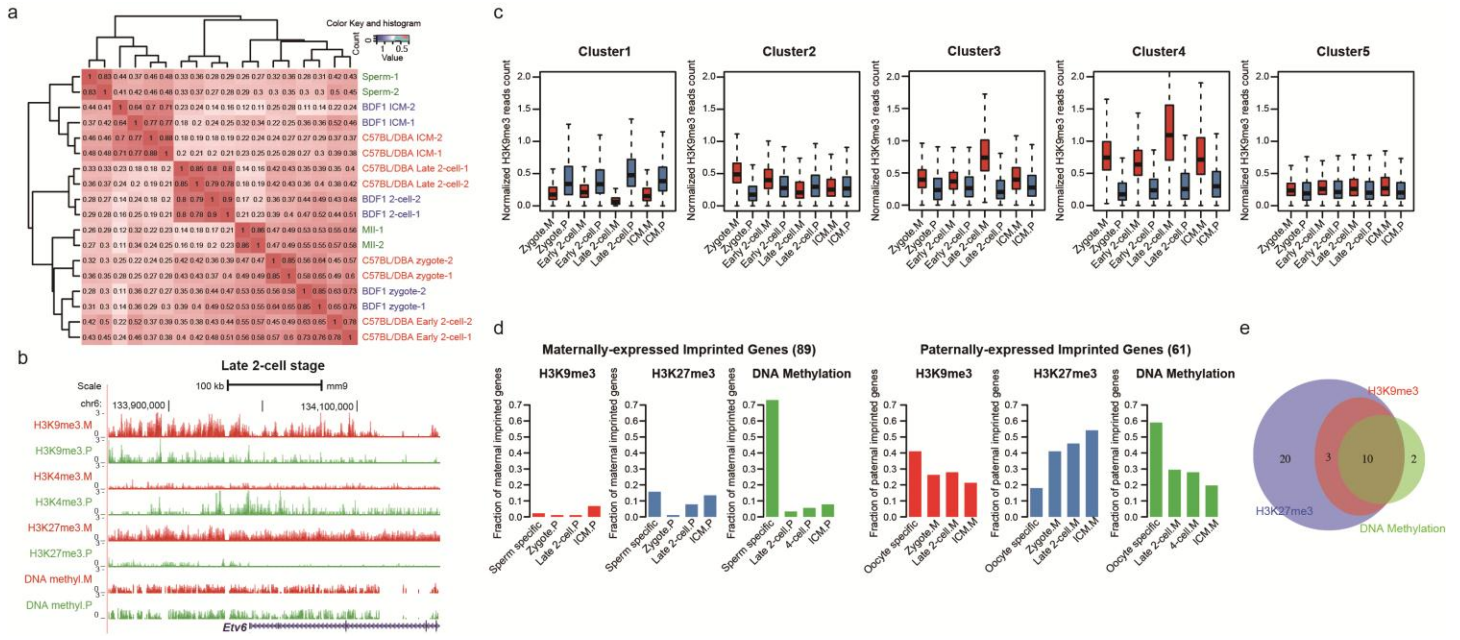


Supplementary Figure 2

Dynamics of H3K9me3 during mouse early embryo development

a, Boxplots show the H3K9me3, H3K27me3 and DNA methylation level of Oocyte specific, Cleavage specific and Blastocyst specific H3K9me3 domains. Statistical significance between stages was evaluated based on one-sided Wilcoxon test, * represent p-value < 0.05, *** represent p-value < 0.001. Oocyte specific (n=38996 H3K9me3 domains), Cleavage specific (n=114033 H3K9me3 domains), Blastocyst specific (n=119837 H3K9me3 domains). The center represents the median value and the upper/lower line represent 5% and 95% quantile for the boxplots.

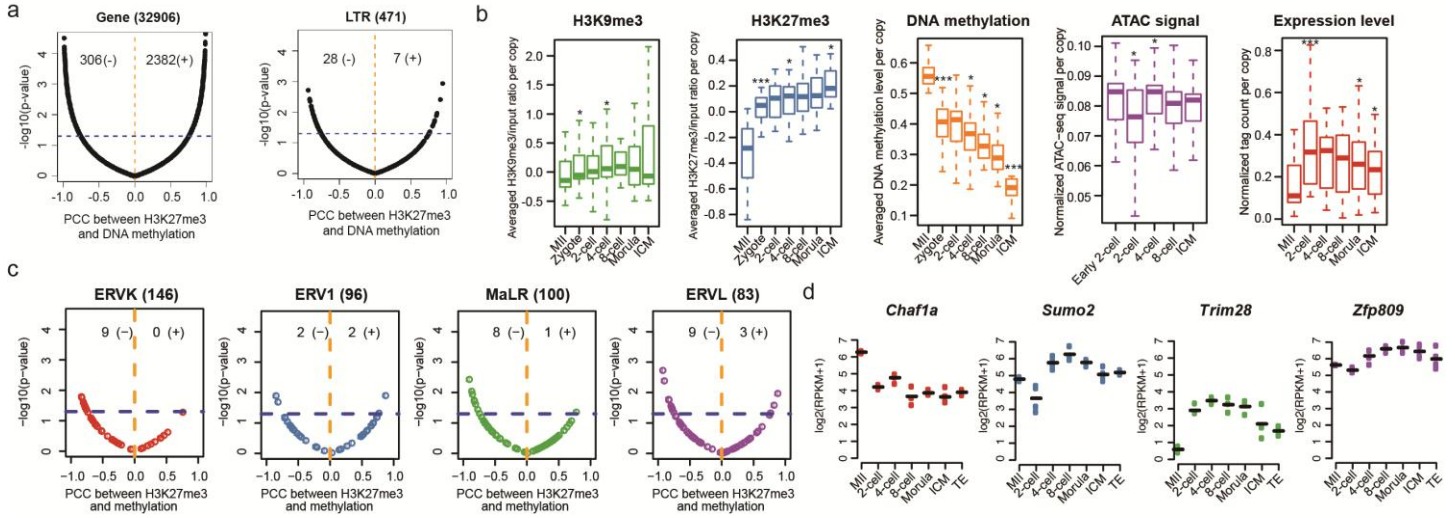
b, Heatmaps show the dynamic of H3K9me3, H3K27me3 and DNA methylation level in blastocysts and post-implanted embryos. The heatmaps were generated using the same order of pre-implantation H3K9me3 heatmap. Colors represent log2-transformed ChIP / input ratio scaled by row, or absolute DNA methylation level. ChIP-seq for H3K9me3 has been performed twice for each sample and data shown represent averages for these independent experiments. H3K27me3 ChIP seq data for zygote was performed three times and data for pre-implantation embryos were from our previous publication (GSE73952). WGBS has been performed once for each sample.



Supplementary Figure 3

Landscape of allelic specific H3K9me3 during early embryo development

- a, Heatmap shows the Pearson's correlation coefficients of BDF1 H3K9me3 samples and C57BL/DBA H3K9me3 samples.
- b, The UCSC genome browser view of allelic H3K9me3, H3K4me3, H3K27me3 and DNA methylation level near gene *Etv6*. Signals represent ChIP-seq RPM for histone modifications and absolute level for DNA methylation.
- c, Boxplots show the allelic H3K9me3 signal of different clusters in Figure 4c. n=number of H3K9me3 domains between different groups. The center represents the median value and the upper/lower line represent 5% and 95% quantile for the boxplots.
- d, Bar plot show the fraction of maternally-expressed (left) and paternally-expressed (right) imprinted genes that regulated by allelic specific H3K9me3, H3K27me3 and DNA methylation. ChIP-seq for H3K9me3 has been performed 2 time for each stage except late 2-cell (3 times) and data shown represent average values. WGBS for each sample was performed once.
- e, Pie chart show the overlap of maternal specific H3K9me3, H3K27me3 and DNA methylation regulated imprinted genes at ICM stage. H3K27me3 ChIP-seq data for MII oocyte, sperm, late 2cell and E3.5 ICM were from public data (GSE76687).



Supplementary Figure 4

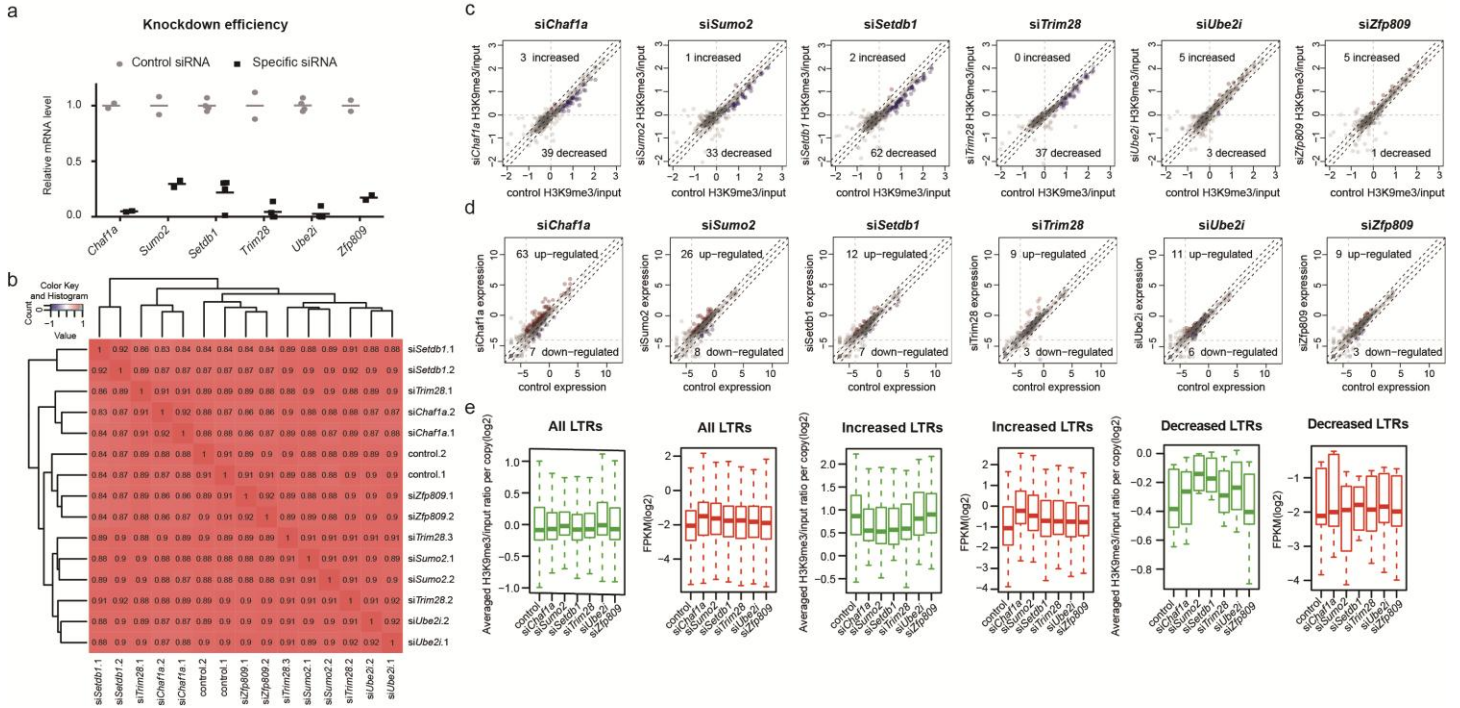
Epigenetic switch of DNA methylation with H3K9me3 on LTR regions

a, Scatterplots show the association tests between log2-transformed H3K27me3 / input ratio and DNA methylation level during pre-implantation embryos on specific LTR families. Different colors represent four major family of the LTRs, including ERVK, ERV1, MaLR and ERVL (two-sided association test, n=7 biologically independent samples). The x-axis represents the Pearson's correlation coefficients and the y-axis represents P-value of the tests. The blue horizontal line in each plot correspond p-value of 0.05 significance threshold. The total number of significant positive (+) and negative (-) correlations (P-value < 0.05) for each annotation is shown at the top of each plot.

b, Boxplots show the log2-transformed H3K9me3 / input ratio, log2-transformed H3K27me3 / input ratio, DNA methylation, ATAC signal and expression level of the 28 H3K27me3 increased LTR during pre-implantation development. Statistical significance between stages was evaluated based on one-sided Wilcoxon test, * represent p-value < 0.05, *** represent p-value < 0.001. n=28 different LTRs. The center represents the median value and the upper/lower line represent 5% and 95% quantile for the boxplots.

c, Scatterplots show the association tests between log2-transformed H3K27me3 / input ratio and DNA methylation level during pre-implantation embryos on specific LTR families (two-sided association test, n=7 biologically independent samples). Different colors represent four major family of the LTRs, including ERVK, ERV1, MaLR and ERVL. The total number of significant positive (+) and negative (-) correlations (P-value < 0.05) for each family is shown at the top of each plot.

d, Graphs show the averaged expression level of *Chaf1a*, *Sumo2*, *Trim28* and *Zfp809*. RNA-seq has been performed 2 times for MII oocyte, morula, 3 times for 8-cell and 4 times for 2-cell, 4-cell ICM and TE. Results for each experiment are plotted as individual points with the average value indicated by the black line.



Supplementary Figure 5

Chaf1a is involved in LTR silencing by H3K9me3 and is crucial for normal embryo development

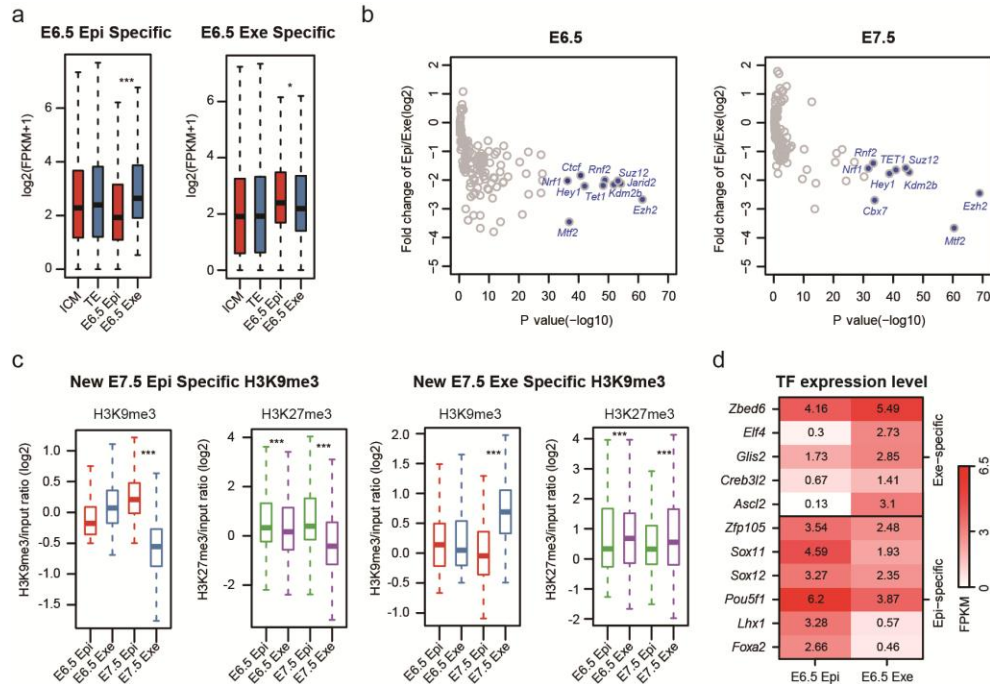
a, Gene expression levels were reduced by corresponding siRNA injection into embryos. RT-qPCR analysis was performed 48 h after injection. The relative expression levels of corresponding genes relative to Gapdh were compared with control embryos. Data shown represent the mean (indicated by the line in the figure) of 2-4 independent experiments: in the order of control siRNA and specific siRNA: *Chaf1a* (n=2, 2), *Sumo2* (n=2, 2), *Setdb1* (n=4, 4), *Trim28* (n=2, 4), *Ube2i* (n=4, 4) *Zfp809* (n=2, 2). "n" represents number of samples repeats in RT-qPCR.

b, Heatmap shows the Pearson's correlation coefficients of replicates for control and siRNA knockdown H3K9me3 samples.

c, Scatterplots show the comparison of control and siRNA knockdown H3K9me3 samples. The x-axis stands for log2-transformed H3K9me3 / input ratio on LTRs for control samples (log2), and the y-axis stands for log2-transformed H3K9me3 / input ratio on LTRs for siRNA knockdown experiments. The diagonal dashed line means LTRs with significant H3K9me3 change (log2-transformed absolute fold change > 0.25).

d, Scatterplots show the comparison of control and siRNA knockdown expression samples. The x-axis stands for normalized RNA-seq reads count on LTRs for control samples (log2), and the y-axis stands for normalized RNA-seq reads count on LTRs for siRNA knockdown experiments. The diagonal dashed line means LTRs with significant expression level change (log2-transformed absolute fold change > 1).

e, Boxplots show the H3K9me3 and expression level of all LTRs (left two panel), 50 H3K9me3 increased LTRs (middle two panel) and 8 H3K9me3 decreased LTRs (right two panel). The center represents the median value and the upper/lower line represent 5% and 95% quantile for the boxplots. All LTRs (n=471 LTRs), Increased LTRs (n=50 LTRs), Decreased LTRs (n=8 LTRs).



Supplementary Figure 6

Lineage specific H3K9me3 is established at promoter regions after implantation

a, Boxplots show the expression level of E6.5 Epi specific H3K9me3 marked genes (left panel) and E6.5 Exe specific H3K9me3 marked genes (right panel). Statistical significance between groups was evaluated based on one-sided Wilcoxon test, * represent p-value < 0.05, *** represent p-value < 0.001, n=number of H3K9me3 domains between groups. The center represents the median value and the upper/lower line represent 5% and 95% quantile for the boxplots.

b, Scatterplots showing the preference of epigenetic factor binding in Epi/Exe-specific H3K9me3 domains based on mESC ChIP-seq peaks. The x-axis represents the P-value of Fisher's exact tests, and the y-axis represents the fold-change of binding sites in Epi/Exe-specific H3K9me3 peaks. Factors with significant preferences are labeled on the graph.

c, Boxplots show the H3K9me3 and H3K27me3 signal of E7.5 newly established Epi/Exe specific H3K9me3 marked genes. Statistical significance between groups was evaluated based on one-sided Wilcoxon test, * represent p-value < 0.05, *** represent p-value < 0.001, n=number of H3K9me3 domains between groups. The center represents the median value and the upper/lower line represent 5% and 95% quantile for the boxplots.

d, Graphs show the expression level of identified lineage specific TFs in E6.5 Epi and Exe. Color represent FPKM.

Supplementary Tables:

Supplementary Table 1, Summary of mapping and quality information of ChIP-seq, RNA-seq and WGBS data.

Supplementary Table 2, Detailed information of “H3K9me3 increased LTRs” and “H3K9me3 decreased LTRs”, the $\log_2$ -transformed H3K9me3 / input ratio, DNA methylation level, correlation and p-value of association test were included.

Supplementary Table 3, sequences for PCR primers and siRNAs.