nature cell biology

RESOURCE

https://doi.org/10.1038/s41556-018-0093-4

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 Li[1], Wenqiang Liu[1], Chuan Chen[1], Xiaochen Kou[1], Yanhong Zhao[1], Jiayu Chen[1], Yixuan Wang[1], Rongrong Le[1], Hong Wang[1], Tao Duan[1], Yong Zhang[1*] and Shaorong Gao[1*]

H3K9me3-dependent heterochromatin is a major barrier of cell fate changes that must be reprogrammed after fertilization. However, the molecular details of these events are lacking in early embryos. Here, we map the genome-wide distribution of H3K9me3 modifications in mouse early embryos. We find that H3K9me3 exhibits distinct dynamic features in promoters and long terminal repeats (LTRs). Both parental genomes undergo large-scale H3K9me3 reestablishment after fertilization, and the imbalance in parental H3K9me3 signals lasts until blastocyst. The rebuilding of H3K9me3 on LTRs is involved in silencing their active transcription triggered by DNA demethylation. We identify that *Chaf1a* is essential for the establishment of H3K9me3 on LTRs and subsequent transcriptional repression. Finally, we find that lineage-specific H3K9me3 is established in post-implantation embryos. In summary, our data demonstrate that H3K9me3-dependent heterochromatin undergoes dramatic reprogramming during early embryonic development and provide valuable resources for further exploration of the epigenetic mechanism in early embryos.

Mammals develop from fertilized eggs, which confer distinct epigenetic modifications from both parental genomes. During fertilization, these epigenetic modifications undergo extensive reprogramming for compatibility with embryonic totipotency[1,2]. Specifically, the constitutive heterochromatin marker histone 3 lysine 9 trimethylation (H3K9me3) is rapidly remodelled to create a chromatin environment that is compatible with cellular reprogramming and plasticity[3]. Meanwhile, H3K9me3-dependent heterochromatin is also involved in embryonic gene regulation and imprinted X inactivation during early embryonic development[4–6]. Depletion of the genes encoding H3K9me3-related chromatin factors, such as *Ehmt2* and *Trim28*, causes improper gene expression and embryonic lethality[7,8]. However, previous studies investigating the H3K9me3 dynamics in vivo based on immunofluorescence staining have been limited by the imaging resolution and lack of an overarching regulatory framework[9–11]. Hence, genome-wide analyses of H3K9me3-dependent heterochromatin dynamics in pre-implantation embryos are sorely needed.

The mouse genome is extensively demethylated during early embryonic development[12,13], and a large fraction of long terminal repeats (LTRs) become hypomethylated and are actively transcribed, which are crucial for normal development[14,15]. These LTRs must be properly regulated in later developmental stages to maintain genome stability[16]. Studies in mouse embryonic stem cells (mESCs) have suggested that multiple H3K9me3 modifiers, such as *Suv39h1/h2* and ERG-associated protein with SET domain (ESET) complexes, are responsible for retrovirus silencing by establishing H3K9me3 modifications[17,18]. However, little is known about the involvement of H3K9me3-dependent heterochromatin in LTR silencing during development in vivo.

To address the above questions, we performed comprehensive analyses of H3K9me3-dependent heterochromatin dynamics in pre-implantation mouse embryos as well as in differentiated embryonic tissues after implantation using a recently published ultra-low-input chromatin immunoprecipitation followed by sequencing (ULI-NChIP–seq) method[19,20]. Based on this high-resolution map, we investigated the reprogramming of H3K9me3-dependent heterochromatin during mouse early embryonic development and also explored the molecular mechanism of LTR silencing by H3K9me3-dependent heterochromatin in pre-implantation embryos.

Results

Genome-wide profiling of H3K9me3 in mouse gametes and early embryos. We generated H3K9me3 profiles of mouse metaphase II (MII) oocytes, sperm and embryos from the zygote stage (pronuclear stage 3 (PN3)) to embryonic day 8.5 (E8.5), mESCs and mouse trophoblast stem cells (TSCs). Embryos at the blastocyst stage were separated into the inner cell mass (ICM) and the trophectoderm, and post-implantation embryos were separated into epiblast and extraembryonic tissue (Fig. 1a,b). RNA sequencing (RNA-seq) and whole-genome bisulfite sequencing (WGBS), as well as ChIP–seq of the H3K4me3 and H3K27me3 levels (including some published data[20]) for each corresponding stage, were also performed for comprehensive analysis (Supplementary Table 1). We found that the global H3K9me3 enrichment in ESCs generated from 500 cells was consistent with the ENCODE data from 10–20 million cells (Supplementary Fig. 1a). Moreover, the two replicates of H3K9me3 for each embryonic stage had Pearson's correlation coefficients >0.9, indicating the high reproducibility of the H3K9me3 data (Supplementary Fig. 1b).

[1]Clinical 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. [2]These 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

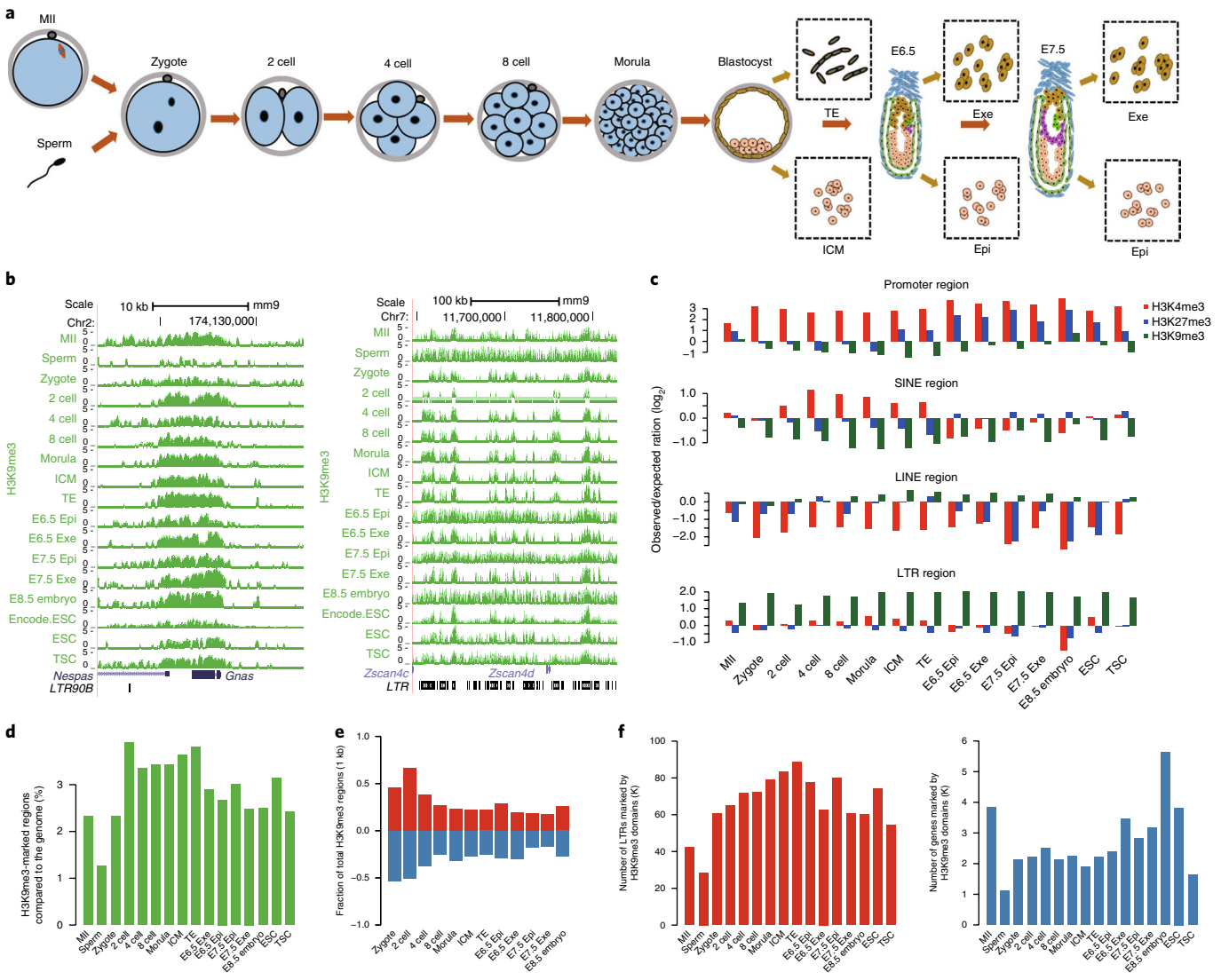

Fig. 1 | Genome-wide profiling of H3K9me3 in mouse gametes and early embryos. a, Schematic showing the preparation of mouse embryos for genome-wide ChIP–seq analysis of H3K9me3. Samples of MII oocytes and sperm; zygotes, 2-cell-, 4-cell-, 8-cell- and morula-stage embryos; the ICM and the trophectoderm (TE) of E3.5 blastocysts; epiblast (Epi) and extraembryonic ectoderm (Exe) of E6.5 and E7.5 of post-implantation embryos are analysed. **b**, The UCSC genome browser view of H3K9me3 signals around *Zscan4d* (right panel) and *Gnas* (left panel) in mouse gametes, early embryos, mESCs (from either this study or from ENCODE) and mouse TSCs. Signals represent the $\log_2$-transformed H3K9me3/input ratio. Chr, chromosome. **c**, Bar plots show the enrichment of H3K9me3, H3K4me3 and H3K27me3 in promoter, SINE, LINE and LTR regions. **d**, Graph showing the percentage of genomic regions covered by H3K9me3 peaks. **e**, Graph showing the fraction of established and disappeared H3K9me3 domains during embryonic development. **f**, Graph showing the number of H3K9me3-marked LTR regions (left panel) and promoter regions (right panel) during embryonic development. In **b**–**f**, ChIP–seq for H3K9me3 was performed two times for each indicated stage, except for TSCs (three times) and E7.5 epiblast tissue (four times), and data shown represent the average values for each stage. In **c**, H3K4me3 and H3K27me3 ChIP–seq data for MII oocyte, sperm, 2-cell-, 4-cell-, 8-cell- and morula-stage embryos, E3.5 ICM, E3.5 trophectoderm and TSCs were from our previous publication (GSE73952)[20] and others were derived from this study: H3K4me3 ChIP–seq was performed two times except for E7.5 epiblast and E7.5 extraembryonic tissues (three times), and H3K27me3 ChIP–seq was performed three times except for E7.5 epiblast and ESCs (two times).

To identify H3K9me3-enriched domains during early embryonic development, we performed peak calling analysis using model-based analysis for ChIP-seq (MACS)[21]. We first examined the distribution of the H3K9me3 domains in early embryos and found that they are distinct from H3K4me3 and H3K27me3 (Supplementary Fig. 1c). Unlike the H3K4me3 and H3K27me3 domains, which are mainly enriched at promoters, the H3K9me3 domain is highly enriched in LTRs (Fig. 1c), indicating that H3K9me3-dependent heterochromatin might be involved in the regulation of LTRs during mouse early embryonic development[17]. H3K9me3 signals are negatively correlated with H3K4me3 and assay for transposase-accessible chromatin using sequencing (ATAC-seq) signals at promoter regions and positively correlated with traditional repressive modifications (Supplementary Fig. 1d,e), reflecting the compatibility between multiple epigenetic modifications in the regulation of embryonic gene expression.

We next focused on the dynamics of the H3K9me3 domains in early embryos. The most dramatic increase in the number of H3K9me3 domains occurs at the 2-cell stage (Fig. 1d,e), suggesting that the global permissive zygotic genome is quickly restrained at the first cell division[22]. Consistent with major changes, the majority of H3K9me3 writers and erasers were highly expressed

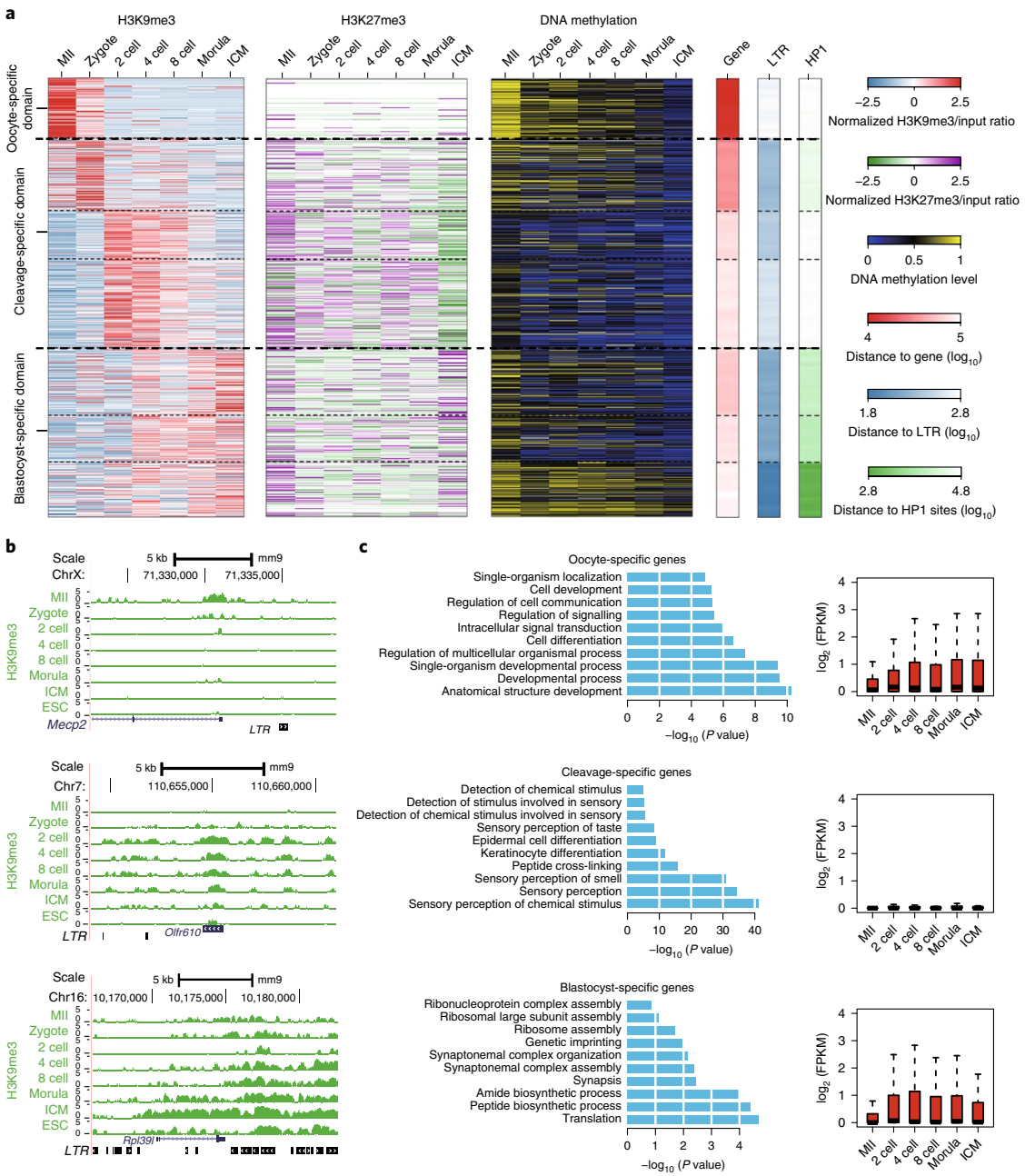

Fig. 2 | Dynamics of H3K9me3-dependent heterochromatin in mouse early embryos. a, Left panel, heatmaps showing the dynamics of H3K9me3 domains during mouse pre-implantation embryo development. The colours represent the $\log_2$-transformed H3K9me3/input ratio scaled by row. Middle and right panels, heatmaps showing the H3K27me3 level and DNA methylation level on H3K9me3 clusters during mouse pre-implantation embryo development. Heatmaps were generated using the same order of H3K9me3 clusters. The colours represent the $\log_2$-transformed H3K27me3/input ratio scaled by row (middle panel) and the absolute DNA methylation levels (right panel). For each cluster, the averaged distance to the TSS, LTR regions and mESC HP1-binding sites are also plotted. The colours represent distance ($\log_{10}$). Data shown represent the averages for two independent H3K9 ChIP–seq experiments. WGBS was performed once for each sample. H3K27me3 ChIP–seq data for zygotes were from experiements performed three times and data for pre-implantation embryos were from our previous publication (GSE73952). **b**, The UCSC genome browser view of representative regions of oocyte-specific (top), cleavage-specific (middle) and blastocyst-specific (bottom) H3K9me3 domains. Signals represent the average $\log_2$-transformed H3K9me3/input ratio from two independent ChIP–seq analyses. **c**, Left, gene ontology analysis of oocyte-specific, cleavage-specific and blastocyst-specific H3K9me3-covered genes. *P* values were determined based on a modified Fisher's exact test. Right, box plot for the expression level of oocyte-specific, cleavage-specific and blastocyst-specific H3K9me3-covered genes during pre-implantation embryo development. Data shown represent the averages for 2–4 independent experiments: RNA-seq was performed 2 times for the MII oocyte and morula, 3 times for the 8-cell stage and 4 times for the 2-cell, 4-cell and ICM stages. The centre of the box plots represents the median value and the lower and upper lines represent the 5% and 95% quantiles, respectively: oocyte specific ($n=1,713$ genes), cleavage specific ($n=1,474$ genes) and blastocyst specific ($n=1,298$ genes).

in oocytes and maintained until the 2-cell stage (Supplementary Fig. 1f,g). Moreover, we observed distinct patterns of H3K9me3 on LTRs and promoters in early embryos (Fig. 1f). The number of LTRs marked by H3K9me3 increased gradually after fertilization and was preserved at a high level in pre-implantation embryos. By contrast, the H3K9me3 domains in the promoters were

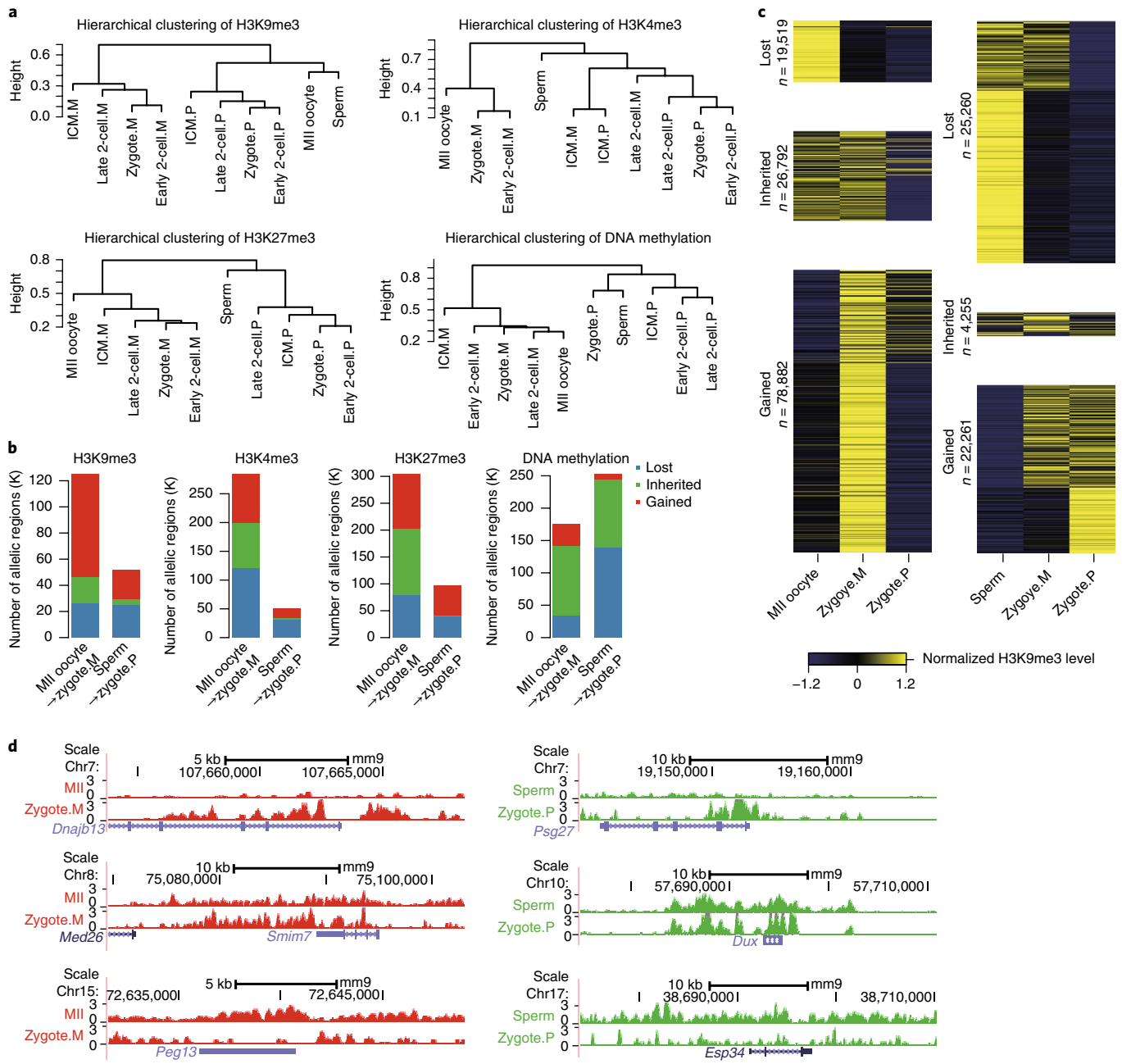

Fig. 3 | Reprogramming of maternal and paternal H3K9me3 during fertilization. a, Hierarchical clustering of H3K9me3, H3K4me3, H3K27me3 and DNA methylation levels in gametes and two alleles of embryos. H3K4me3 and H3K27me3 ChIP-seq data for MII oocyte, sperm, early 2-cell, late 2-cell and E3.5 ICM stages were from public data (GSE71434 and GSE76687)[27,42]. M, maternal; P, paternal. **b**, Graphs showing the number of lost, inherited and gained regions during fertilization for H3K9me3, H3K4me3, H3K27me3 and DNA methylation. **c**, Heatmaps showing the H3K9me3 enrichment in lost, inherited and gained regions of gametes and two alleles of zygote samples. Left panel shows the maternal regions and the right panel shows the paternal regions. The colours represent the normalized SNP-trackable H3K9me3 read counts and the values are scaled by row. **d**, The UCSC genome browser view of the representative lost, inherited and gained regions for maternal (left) and paternal (right) H3K9me3 signal during fertilization. Signals represent ChIP-seq RPM. In **a**–**d**, ChIP-seq for H3K9me3 was performed two times for each indicated stage and the data shown represent the average values.

depleted after fertilization and restored after implantation, suggesting that there are different regulatory mechanisms in these two genomic contexts.

Dynamics of H3K9me3-dependent heterochromatin in mouse early embryos. To investigate the details of the H3K9me3-dependent heterochromatin dynamics, we performed k-means clustering from MII oocytes to the ICM stage and clustered the H3K9me3 domains into three major groups. The first group included 'oocyte-specific domains', which were marked by high levels of H3K9me3 in MII oocytes and were lost quickly after fertilization (Fig. 2a,b). These domains were enriched in the promoters, and functional analysis of the covered genes suggested that most of them are developmental genes that are activated upon zygotic genome activation (Fig. 2c). H3K9me3 signals on 'cleavage-specific domains' are formed temporarily after fertilization and then diminish after the morula stage (Fig. 2a,b), and the covered genes show strong enrichment in sensory-related terms and are seldom

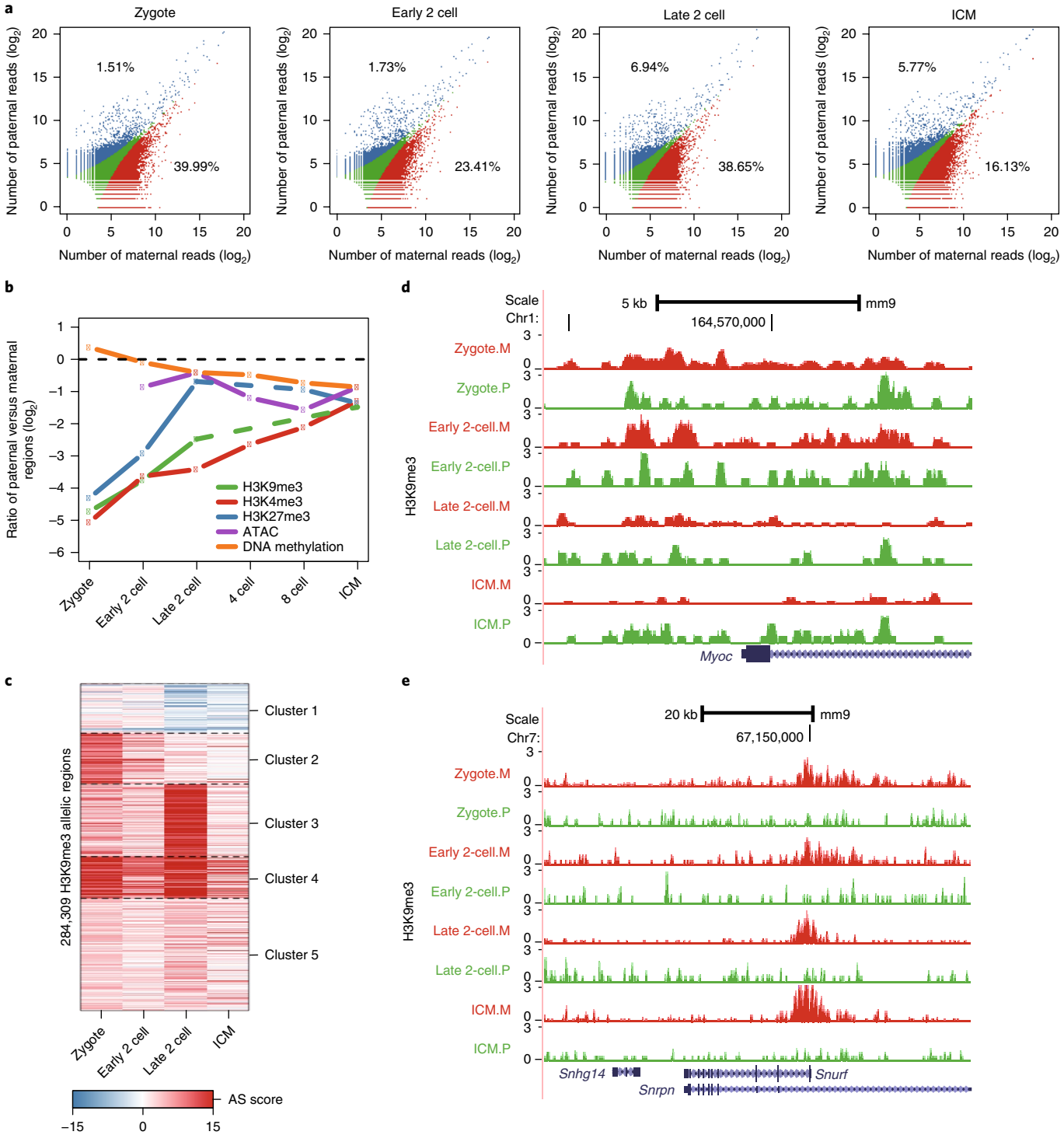

Fig. 4 | Landscape of allelic-specific H3K9me3 during early embryonic development. a, Scatter plots showing the number of H3K9me3 reads assigned to each allele. The red, green and blue colours denote maternal-specific, bi-allelic and paternal-specific H3K9me3 regions. The numbers indicate the percentage of paternal-specific H3K9me3 regions (top left) and maternal-specific H3K9me3 regions (bottom right) for each developmental stage. **b**, Line chart showing the ratio of paternal-specific regions versus maternal-specific regions for epigenetic modifications during mouse pre-implantation development (log_2). The dashed lines represent non-consecutive stages. The black dashed line across the 0 mark indicates equal number of paternal and maternal regions. **c**, Heatmap showing the dynamics of the H3K9me3 AS score during pre-implantation development. The colours represent the H3K9me3 AS score. **d**, The UCSC genome browser view of allelic H3K9me3 signals near the *Myoc* gene (cluster 2 in **c**) during development. Signals represent ChIP-seq RPM. **e**, The UCSC genome browser view of allelic H3K9me3 signals near the *Snrpn* gene (cluster 4 in **c**) during development. Signals represent ChIP-seq RPM. In **a**–**e**, ChIP-seq for H3K9me3 was performed two times for each stage except the late 2-cell stage (three times) and the data shown represent the average values.

expressed in early embryos (Fig. 2c). Most H3K9me3 signals on 'blastocyst-specific domains' are formed after the 4-cell stage and are progressively enhanced during pre-implantation development (Fig. 2a,b). These domains are enriched in LTR retro-transposons and are heterochromatin protein 1 (HP1)binding sites in ESCs, which are the features of classical constitutive heterochromatin[23].

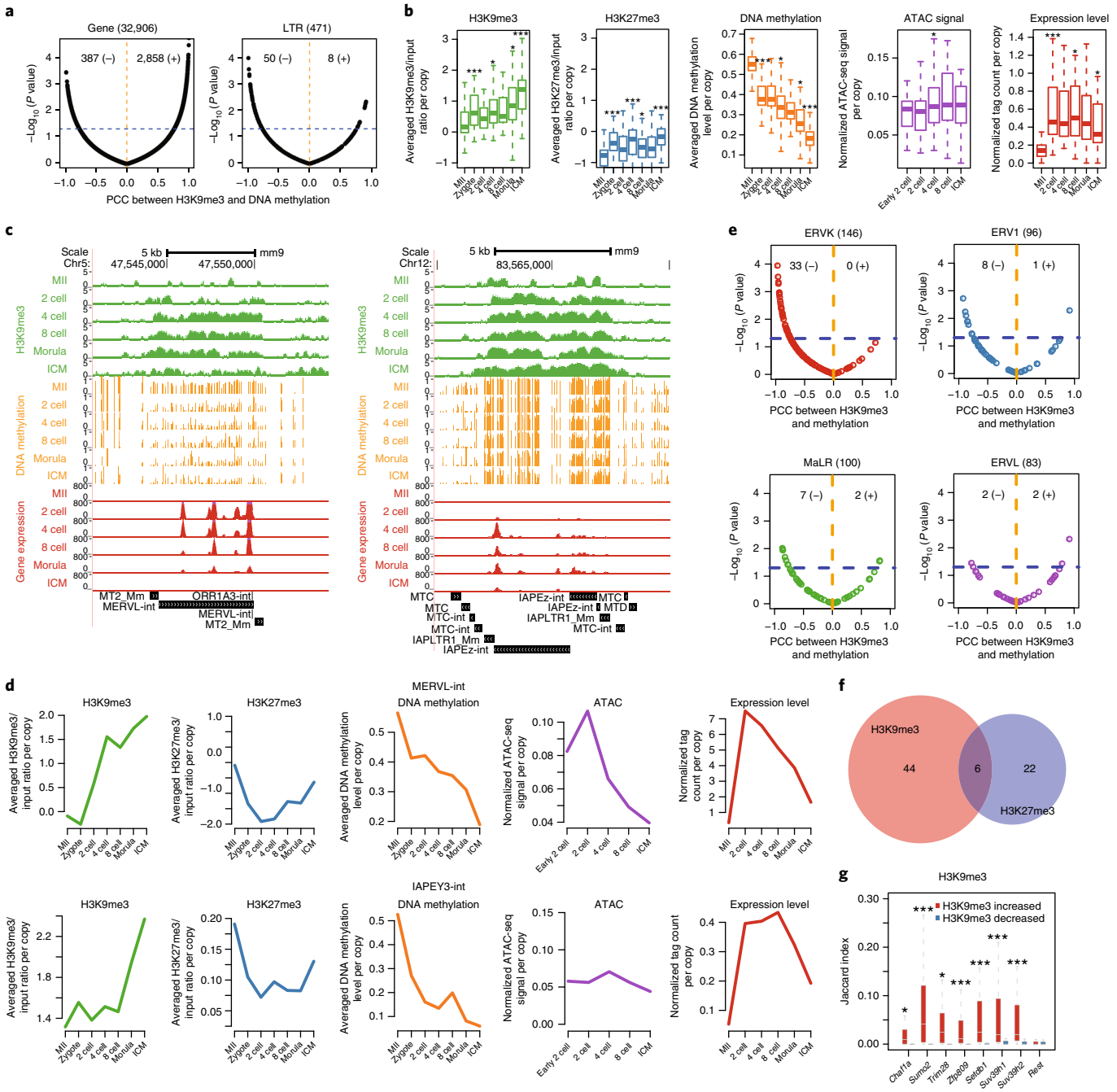

Fig. 5 | Epigenetic switching of DNA methylation with H3K9me3 on LTR regions. a, Scatter plots for the association between the $\log_2$-transformed H3K9me3/input ratio and DNA methylation levels during pre-implantation in promoter and LTR regions. The *x* axis represents the Pearson's correlation coefficients (PCCs) and the *y* axis represents the *P* values of the two-sided association test ($n = 7$ biologically independent samples). The blue horizontal line corresponds to $P = 0.05$. The orange dashed line separates the positive and negative associations. The total number of significant positive (+) and negative (−) correlations ($P < 0.05$) is shown at the top of each plot. **b**, Box plots showing the $\log_2$-transformed H3K9me3/input ratio, the $\log_2$-transformed H3K27me3/input ratio, the DNA methylation, the ATAC signal and the expression levels of the 50 H3K9me3-increased LTRs during pre-implantation. **c**, The UCSC genome browser view of the $\log_2$-transformed H3K9me3/input ratio, the absolute DNA methylation level and the normalized RNA-seq read counts on representative LTRs (the MERVL internal sequences, MERVL-int (left) and the IAPEz internal sequences, IAPEz-int (right)). **d**, Line charts showing the H3K9me3, H3K27me3, DNA methylation, ATAC signal and expression levels of MERVL and IAPEY3 during pre-implantation. **e**, Scatter plots showing the association between the normalized H3K9me3 read counts and the DNA methylation level during pre-implantation in promoter and LTR regions (two-sided association test, $n = 7$ biologically independent samples). The different colours represent the four major LTR families. The total number of significant positive and negative correlations ($P < 0.05$) is shown at the top of each plot. **f**, Venn diagram shows the overlap between the 50 H3K9me3-increased LTRs and the 28 H3K27me3-increased LTRs. **g**, Box plot showing the enrichment of chromatin factor-binding sites on the 50 H3K9me3-increased LTRs and the 8 H3K9me3-decreased LTRs. The *y* axis represents the Jaccard index between ChIP-seq peaks and LTRs. In **a**–**d**, H3K9me3 ChIP-seq was performed twice, with data representing the average values, and WGBS was performed once. In **b** and **d**, H3K27me3 ChIP-seq and ATAC data were from previous publications (GSE73952 and GSE66390, respectively). In **b** and **g**, the centre of the box plots represents the median value and the lower and upper lines represent the 5% and 95% quantile, respectively; significance between the groups was evaluated by one-sided Wilcoxon test ($n = 50$ LTRs), $^*P < 0.05$, $^{***}P < 0.001$.

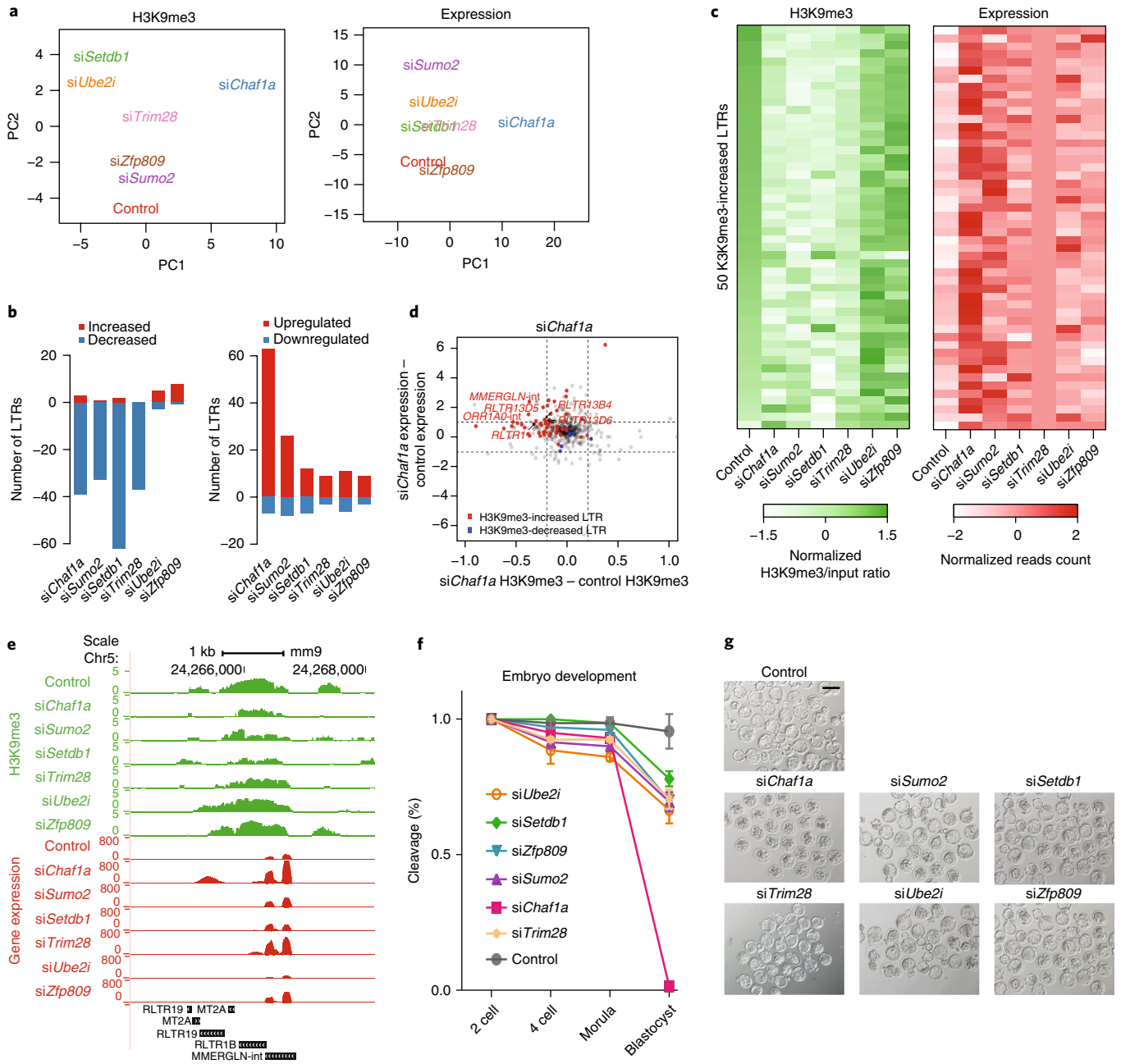

Fig. 6 | *Chaf1a* is involved in LTR silencing by H3K9me3 and is crucial for normal embryonic development. a, Principal component (PC) analysis of H3K9me3 ChIP-seq data (left panel) and RNA-seq data (right panel) for LTRs. Data were obtained at the morula stage in control mouse embryos or embryos with siRNA-mediated KD of the indicated six genes; $n = 7$ biologically independent samples. **b**, Graphs showing the number of LTRs with increased and decreased H3K9me3 signals (left panel) and upregulated and downregulated expression levels (right panel). **c**, Heatmaps showing the H3K9me3 and expression levels of 50 H3K9me3-increased LTRs in control and siRNA-KD samples. The colours represent the $\log_2$-transformed H3K9me3/input ratio and the normalized RNA-seq read counts scaled by row. **d**, Scatter plot showing the changes in H3K9me3 level and expression level of all LTRs in *Chaf1a*-KD embryos. LTRs with both significant H3K9me3 level and expression level changes are labelled on the graph. The grey points indicate other LTRs. **e**, The UCSC genome browser view of H3K9me3 and expression levels on representative LTR (MMERGLN). Signals represent the $\log_2$-transformed H3K9me3/input ratio and the normalized RNA-seq read counts. In **a**–**e**, ChIP-seq for H3K9me3 was performed two times, except for *Trim28*-KD embryos (three times), and RNA-seq was performed two times, except for *Ube2i*-KD embryos (three times); data shown represent the averages for these independent experiments. **f**, Depletion of H3K9me3-related chromatin factors impaired the development of pre-implantation embryos. The siRNA mix was injected into MII oocytes. The injected oocytes then underwent ICSI to initiate further development. Data are presented as the mean ± s.d. si*Ube2i* ($n = 35, 40, 37$), si*Setdb1* ($n = 34, 30, 24$), si*Zfp809* ($n = 33, 34, 20$), si*Sumo2* ($n = 30, 31, 40$), si*Chaf1a* ($n = 30, 30, 28$), si*Trim28* ($n = 25, 26, 36$), control ($n = 35, 35, 21$). *n* refers to the number of 2-cell stage embryos. **g**, Representative images of blastocyst-stage embryos produced from siRNA injection at 4 days after fertilization. This experiment was repeated three times independently with similar results. Scale bar, 100 μm.

In summary, these results suggest that H3K9me3 signals may be linked with transcriptional regulation and intracellular functions in pre-implantation embryos.

As DNA methylation and H3K27me3 are also involved in the formation of heterochromatin[24], we then examined their levels in the three groups of H3K9me3 domains. The 'oocyte-specific domains'

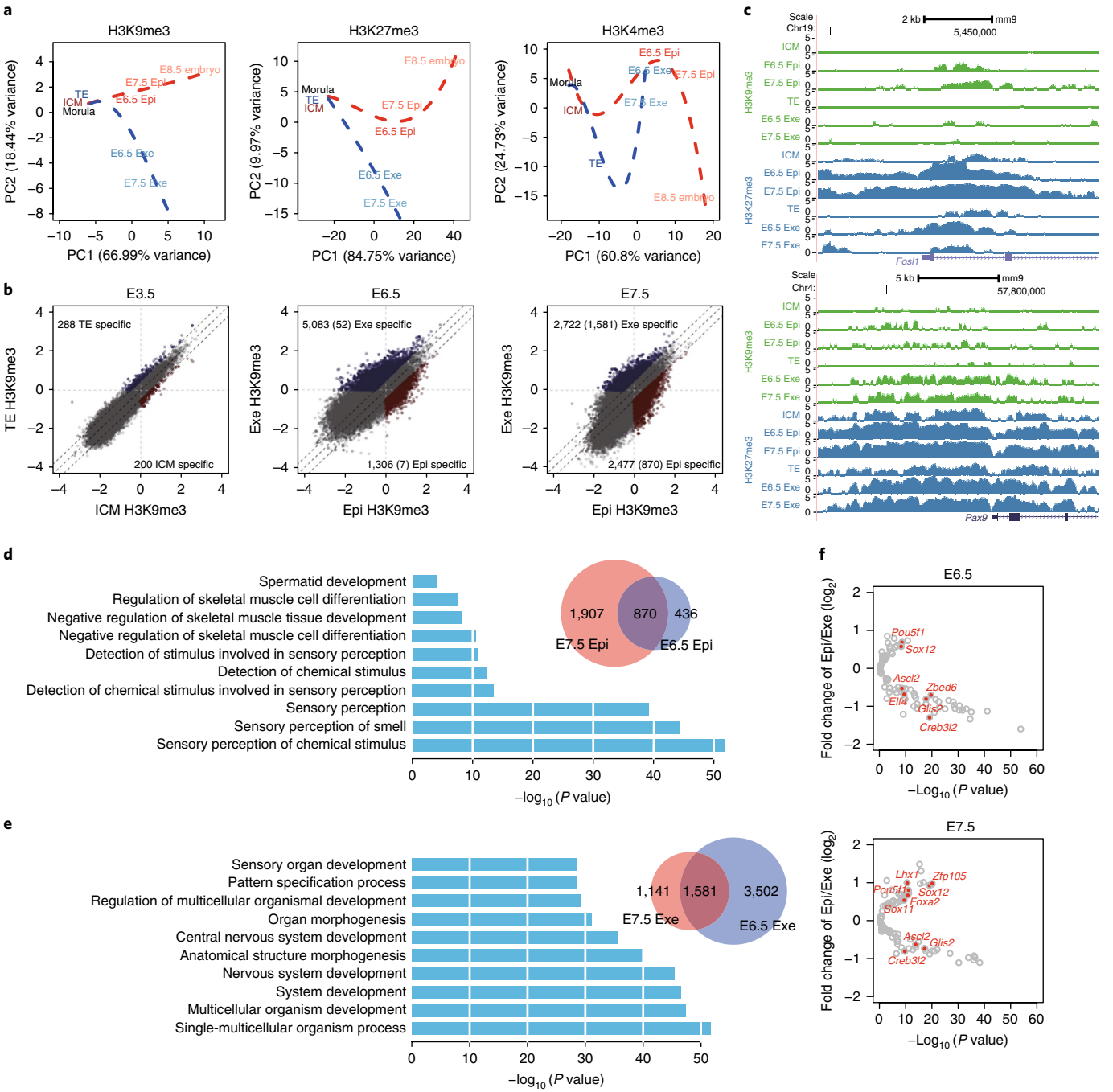

Fig. 7 | Lineage-specific H3K9me3 is established at promoter regions after implantation. a, Principal component analysis of promoter H3K9me3, H3K27me3 and H3K4me3 signals at the blastocyst stage and post-implantation embryos ($n=8$ biologically independent samples). The red dashed line represents the differentiation lineage after the ICM stage, and the blue dashed line represents the differentiation lineage after the trophectoderm stage. **b**, Scatter plots showing the comparison of promoter H3K9me3 signals between two different lineages at 3.5, 6.5 and 7.5 days. The diagonal dashed lines indicate promoters with significant H3K9me3 signal preference ($\log_2$-transformed absolute fold change > 0.5). The numbers of genes with lineage-specific H3K9me3 signal are labelled. **c**, The UCSC genome browser view of H3K9me3 and H3K27me3 levels on representative genes with epiblast-specific H3K9me3 signals (top panel) and extraembryonic-specific H3K9me3 signals (bottom panel). Signals represent the $\log_2$-transformed ChIP/input ratio. **d**,**e**, Right panel, venn diagram shows the overlap between E6.5 and E7.5 epiblast-specific (**d**) or extraembryonic-specific (**e**) H3K9me3-marked genes. Left panel, gene ontology analysis of overlapped epiblast-specific (**d**) or extraembryonic-specific (**e**) genes. *P* values were generated using a modified Fisher's exact test. In **a**–**e**, ChIP-seq for H3K9me3 was performed two times for each indicated stage, except for E7.5 epiblast (four times); data shown represent the average values for each stage. H3K4me3 and H3K27me3 ChIP-seq data for morula-stage embryos, E3.5 ICM and E3.5 trophectoderm were from our previous publication (GSE73952), and others were derived from this study: H3K4me3 ChIP-seq was performed two times, except for E7.5 epiblast and E7.5 extraembryonic (three times), and H3K27me3 ChIP-seq was performed three times, except for E7.5 epiblast (two times). **f**, Scatter plots showing the preference of transcription factor binding in epiblast/extraembryonic-specific H3K9me3 domains based on transcription factor motifs. The *x* axis represents the *P* value of two-sided Fisher's exact tests, and the *y* axis represents the fold change of binding sites in epiblast/extraembryonic-specific H3K9me3 peaks. Factors with significant preferences ($P<1\times10^{-10}$) and with upregulation in the corresponding lineage are labelled on the graph. Significance between groups in **f** was evaluated based on the two-sided fisher's exact test, with the *n* referring to the number of epiblast/extraembryonic-specific H3K9me3 domains and the number of motifs in the epiblast/extraembryonic-specific domains.

were depleted of H3K27me3 signals and the DNA methylation level was severely reduced during development to allow activation of these regions (Fig. 2a and Supplementary Fig. 2a). H3K27me3 signals were reestablished on these developmental gene-enriched regions after implantation, implicating the onset of early lineage commitment[25] (Supplementary Fig. 2b). In comparison, the 'blastocyst-specific domains' still possessed relatively high levels of DNA methylation at the ICM stage (>0.4), together with progressively increased H3K9me3 levels, indicating that these LTR-enriched regions are under tight epigenetic constraint (Fig. 2a and Supplementary Fig. 2a). Most strikingly, the 'cleavage-specific domains' were highly enriched with H3K27me3 signals inherited from the oocyte, suggesting temporal involvement of H3K9me3 on H3K27me3-marked facultative heterochromatin during the first few cell cycles (Fig. 2a and Supplementary Fig. 2a). Taken together, these results indicate that both H3K27me3 and DNA methylation are partially associated with the formation of H3K9me3-dependent heterochromatin during early embryonic development.

Both maternal and paternal genomes undergo substantial reprogramming of H3K9me3 during fertilization. The molecular mechanism underlying the remodelling of maternal and paternal chromatin to the totipotent zygote during fertilization has long been sought[3]. Immunocytochemistry staining experiments suggest that the characteristics of constitutive heterochromatin, including H3K9me3, are only enriched in the maternal pronuclei[10,26]. To understand the allelic reprogramming of H3K9me3 during early development, we regenerated the H3K9me3 profiles using C57BL (oocytes) and DBA (sperm) crossed zygotes at the PN3, early 2-cell, late 2-cell and ICM stages. We first confirmed that the crossed samples were highly consistent with the offspring of the cross between C57BL/6J female × DBA/2J male (BDF1)-background samples (Supplementary Fig. 3a). Then, we took advantage of the single-nucleotide polymorphism (SNP) data from the two strains to measure the allelic-specific enrichment of H3K9me3. We observed allelic-specific signals of H3K9me3, H3K4me3 and H3K27me3 on the *Etv6* gene, and the patterns of the H3K4me3 and H3K27me3 signals were consistent with published data[27], suggesting the precise separation of parent-of-origin-dependent signals (Supplementary Fig. 3b).

To compare the different reprogramming features of multiple epigenetic modifications during fertilization, we first performed hierarchical clustering analysis of the gamete-specific and allelic-specific signals in early embryos. Consistent with previous studies, H3K4me3, H3K27me3 and DNA methylation possess distinct embryonic maternal and paternal signals, with the maternal signal closely resembling oocytes and the paternal signal more loosely resembling sperm (Fig. 3a), suggesting more-extensive reprogramming of the paternal genome[13,27]. However, in contrast to other modifications, neither the maternal nor the paternal H3K9me3 signal in the zygote was clustered with its gamete of origin and further analysis showed that both the maternal and the paternal H3K9me3 regions in the zygote were mostly de novo established, which means that extensive reprogramming of H3K9me3 occurred on both parental genomes (Fig. 3a,b). Compared to the paternal allele, which only inherited a small percentage of sperm H3K9me3 domains (14.4%, $n=4,255$) during fertilization, the maternal allele inherited nearly half of the oocyte H3K9me3 domains at the zygote stage (42.1%, $n=19,519$) (Fig. 3c,d). In addition, there were significantly more new maternal H3K9me3 domains ($n=73,069$) than paternal H3K9me3 domains ($n=18,359$), which led to an overwhelming number of maternal-specific H3K9me3 regions compared to paternal-specific regions at the zygote stage (Fig. 3c,d). Taken together, these results suggest that both maternal and paternal H3K9me3 undergo substantial reprogramming after fertilization and that the imbalance between the parental H3K9me3 signals stem from both epigenetic inheritance from the gametes and reprogramming at the zygote stage.

Allelic-specific H3K9me3 landscape in mouse pre-implantation embryos. Next, we asked whether the allelic imbalance between the parental H3K9me3 was consistent throughout early embryonic development. We used the allelic-specific score (AS score) to define the status of the allelic bias. As expected, significantly more maternal-specific H3K9me3 regions were identified than paternal-specific H3K9me3 regions at the zygote stage and the disequilibrium lasted until the ICM stage (Fig. 4a). However, the percentage of the paternal-specific regions increased, whereas the the maternal-specific regions decreased, suggesting that the allelic-specific H3K9me3 moved towards a relatively balanced state during embryogenesis. Similar balancing trends were also observed for other epigenetic modifications (Fig. 4b). Interestingly, when we clustered the embryonic allelic-specific regions by the AS score, we found that the increase in the number of paternal-specific regions was mainly caused by the loss of maternal-specific H3K9me3 signals during development (cluster 2 in Fig. 4c and Supplementary Fig. 3c; Fig. 4d), which is in contrast to H3K4me3, whose reprogramming was mainly due to the gain of paternal-specific signals[27]. However, most regions were marked by maternal-specific H3K9me3 signals throughout the examined developmental stages (clusters 3–5 in Fig. 4c and Supplementary Fig. 3c), including the promoter of the paternal-imprinted gene *Snrpn* (Fig. 4e).

Allelic-specific expression of imprinted genes is usually controlled by DNA methylation; however, recent studies have shown that histone modifications are also involved in imprinting control[28–31]. According to our ChIP–seq data, we found that only a few of the maternally expressed imprinted genes were marked by paternal-specific H3K9me3 modification. By contrast, a large fraction of paternally expressed imprinted genes were marked by maternal-specific H3K9me3 modification, as well as maternal-specific H3K27me3 modification and DNA methylation (Supplementary Fig. 3d). When comparing the imprinted genes regulated by different allelic-specific modifications at the ICM stage, we found that the H3K9me3-regulated and DNA methylation-regulated imprinted genes were highly overlapped, whereas more H3K27me3-regulated imprinted genes were observed (Supplementary Fig. 3e). Consistent with recent studies[29–31], our results suggested that H3K9me3 might have an overlapping role with DNA methylation in imprinted gene regulation, whereas H3K27me3 could regulate the expression of imprinted genes independently of DNA methylation.

Epigenetic switch from DNA methylation to H3K9me3-regulated transcriptional activity of retro-transposons during development. The global DNA demethylation during mouse early embryonic development results in the activation of retro-transposons at the 2-cell stage[13,14]. These retro-transposons must be properly regulated and silenced in later developmental stages to ensure genome stability, in which we suspect H3K9me3 might play a part. To test that, we performed association tests between H3K9me3 signals and the DNA methylation levels in promoters and 471 LTRs. Compared to promoter regions, in which the associations were mostly positive, 50 LTRs showed significant negative associations (defined as 'H3K9me3-increased LTRs') and only 8 LTRs showed positive associations (defined as 'H3K9me3-decreased LTRs') (Fig. 5a and Supplementary Table 2), suggesting gradual substitution of DNA methylation to H3K9me3, specifically on these LTRs (Fig. 5b). Consistent with the increase in the levels of H3K9me3, these LTRs that were transiently expressed from the 2-cell to the 8-cell stage were downregulated in later stages (Fig. 5b). For example, the 2-cell-specific retro-transposon MERVL, which belongs to the ERVL family and is related to the expanded cell fate potential[32], had increased H3K9me3 signals and decreased expression levels from the 4-cell stage, and IAPEY3 was silenced by H3K9me3 later at the morula stage (Fig. 5c,d). Most H3K9me3-increased LTRs belong to the ERVK families, suggesting that different families of LTRs acquire

epigenetic modifications independently during development[33] (Fig. 5e). We next asked whether H3K27me3 also contributed to the silencing of LTRs. Similar association analyses indicated that 28 LTRs showed significant negative associations, which was markedly less than that of H3K9me3 (Supplementary Fig. 4a–c). Less than one-third of H3K27me3-increased LTRs overlapped with H3K9me3-increased LTRs (Fig. 5f), suggesting that H3K27me3 could also regulate the silencing of a distinct set of LTRs in early embryos. These analyses revealed that the epigenetic silencing factor on LTRs switched from DNA methylation to histone modifications during early embryonic development, in which H3K9me3 played a major part.

We next sought to investigate the factors that were responsible for establishing H3K9me3 on these LTRs. Previous studies have highlighted genes encoding several factors, including CHAF1A, SUMO2, TRIM28 and ZFP80, that can cooperate with the ESET complex to repress endogenous retroviruses (ERVs) in mESCs[18]. We examined the binding status for these factors based on the mESC data, and the results indicated that they are highly enriched in H3K9me3-increased LTRs compared to the gene that encodes the non-related factor REST (Fig. 5g). In addition, most of these factors were highly expressed from the 4-cell stage (Supplementary Fig. 4d), suggesting that they might be responsible for establishing H3K9me3 on LTR retro-transposons in later developmental stages.

CHAF1A functions as the core factor in H3K9me3-mediated LTR silencing in early embryos. To test which factor is involved in H3K9me3-mediated LTR silencing, we knocked down (KD) *Chaf1a*, *Sumo2*, *Setdb1*, *Trim28*, *Ube2i* and *Zfp809* in mouse embryos by injecting small interfering RNA (siRNA) that targeted these genes into MII oocytes, followed by intracytoplasmic sperm injection (ICSI) (Supplementary Fig. 5a). We then analysed the H3K9me3 and expression profiles of KD and control embryos at the morula stage (Supplementary Fig. 5b). Clustering analysis indicated that the H3K9me3 levels and expression profiles on LTR regions were generally distinct from those of control embryos, with *Chaf1a*-KD displaying a marked difference (Fig. 6a). We then identified LTRs with significant changes in H3K9me3 and expression levels (Supplementary Fig. 5c,d). As expected, several LTRs were identified with reduced H3K9me3 levels in *Chaf1a*-, *Sumo2*-, *Setdb1*- and *Trim28*-KD embryos; however, only *Chaf1a*-KD led to a significant reactivation of these LTRs (Fig. 6b,c). Furthermore, we found that the 50 LTRs with increased H3K9me3 signals during normal embryonic development had both descendent H3K9me3 signals and upregulated expression levels in *Chaf1a*-KD embryos (Fig. 6c–e and Supplementary Fig. 5e). These results indicate that *Chaf1a* is the core factor responsible for silencing these LTRs through establishing H3K9me3 modification. Consistent with the defects in silencing of the LTR retro-transposons, we found that, although most KD embryos could develop to the morula stage, a large fraction was arrested at the blastocyst stage; *Chaf1a*-KD had the most severe phenotype, as almost all of the embryos failed to form blastocysts (Fig. 6f,g), suggesting that *Chaf1a*-mediated LTR silencing might be crucial for blastocyst formation and the first lineage differentiation. Taken together, these results demonstrate that *Chaf1a*, *Sumo2*, *Setdb1* and *Trim28* could effect the generation of H3K9me3 on LTRs. Moreover, *Chaf1a* seemed to be the most important factor in H3K9me3-mediated LTR silencing and was essential for proper development to the blastocyst stage.

Lineage-specific heterochromatin on promoters is established in post-implantation embryos. H3K9me3-dependent heterochromatin is believed to be the barrier for cell fate changes, which can repress lineage-incompatible genes and prevent them from being activated by transcription factors[24]. To understand the establishment of lineage-specific histone methylation during development, we compared the distribution of these modifications on the two lineages from E3.5 to E7.5. Clustering analysis at promoter regions indicated that both H3K9me3 and H3K27me3 signals showed striking differences between the two lineages in post-implantation embryos (Fig. 7a), whereas H3K4me3 displayed more similar levels. Moreover, the difference in H3K9me3 signals at the promoters was not widely observed between the ICM and the trophectoderm in E3.5 embryos, but was continually established in E6.5 and E7.5 embryos (Fig. 7b,c). Further analysis indicated that nearly half of the lineage-specific H3K9me3 signals were inherited from E6.5 to E7.5, both of which were closely related to the function of the opposite lineage (Fig. 7d,e). Consistently, the expression level of epiblast or extraembryonic-specific H3K9me3-marked genes were also downregulated in their corresponding lineage (Supplementary Fig. 6a). In summary, these results suggest that lineage-specific H3K9me3-marked promoters are established after implantation to repress the lineage-incompatible genes.

To investigate the factors related to the establishment of lineage-specific H3K9me3, we examined the enrichment of the binding sites of 168 transcription factors and chromatin regulators on the promoter regions of lineage-specific genes based on mESC ChIP–seq data. Interestingly, in addition to the genes encoding H3K9me3-related factors, such as *Cbx7* and *Kdm4c*, we found significant enrichment of the genes encoding the polycomb repressive complex 2 (PRC2) complex, including *Ezh2* and *Suz12*, on extraembryonic-specific H3K9me3-marked genes, which indicates that H3K27me3 might also be involved in the regulation of lineage-specific gene expression (Supplementary Fig. 6b,c). We further predicted the transcription factors that contributed to the establishment of lineage-specific H3K9me3 by performing motif enrichment analysis. Six factors, including *Pou5f1*, *Sox12*, *Sox11*, *Lhx1*, *Zfp105* and *Foxa2*, showed enrichment of both epiblast-specific H3K9me3 domains and higher expression levels in the epiblast, and five factors, including *Zbed6*, *Elf4*, *Glis2*, *Creb3l2* and *Ascl2*, showed enrichment of extraembryonic-specific H3K9me3 domains and higher expression levels in extraembryonic tissue (Fig. 7f and Supplementary Fig. 6d). These results provide further information regarding the potential regulators and building mechanism of lineage-specific H3K9me3-dependent heterochromatins.

Discussion

The inheritance and reprogramming of heterochromatin during mammalian early embryonic development and lineage differentiation represent long-standing challenges[2]. Here, we provide an unprecedented view of the H3K9me3-dependent heterochromatin dynamics during fertilization, early embryonic development and post-implantation embryonic differentiation. Our data confirm the overwhelming genome-wide presence of maternal H3K9me3 signals over paternal allele signals[9,11]. Surprisingly, we found that the de novo-established H3K9me3 further enhanced the asymmetric distributions of H3K9me3 between the two alleles after fertilization. It would be of great interest to investigate the function of zygote-enhanced H3K9me3 asymmetry in future studies, which is possibly related to DNA methylation preservation and imprinting maintainance[34]. Finally, the drastic reprogramming of H3K9me3-dependent heterochromatin upon fertilization removes H3K9me3 barriers from promoter regions, creating a less constrained epigenetic environment for subsequent zygotic genome activation and may be required for the embryo to reach the totipotent state[35].

During pre-implantation embryo development, the H3K9me3-marked genes were relatively few and stable, whereas a large fraction of LTRs were gradually marked by H3K9me3-dependent heterochromatin. The expression levels of 2-cell-specific LTRs, such as MERVL, have been reported to be markers of expanded cell fate potential[32,36], and we demonstrated that the silencing of those LTRs after the 2-cell stage correlated with the increased H3K9me3 level on

those retro-transposons, which is controlled by the CHAF1A complex[37]. Consistently, previous studies have indicated that downregulation of *Chaf1a* in mESCs can induce 2-cell-like cell formation in vitro[38], with highly expressed endogenous retroviruses and low levels of pluripotent protein expression[36]. These results raise the possibility that the CHAF1A-mediated changes in chromatin states may be a prerequisite for the totipotency-to-pluripotency transition.

Embryonic cell differentiation enables mammalian embryos to generate multiple cell types in different organisms as well as in extraembryonic tissues[2,39]. Early lineage specification events are generally believed to be regulated by specific transcription factors[40]. Our analyses suggest that the asymmetric distribution of lineage-specific transcription factors at the blastocyst stage may result in differential permissive chromatin structures in later embryonic developmental stages and that the establishment of lineage-specific H3K9me3-dependent heterochromatins can assure irreversible cell fate changes[24]. Among our predicted transcription factor list, Pou5f1 has been reported to collaborate with Setdb1 to restrict the extra-embryonic trophoblast lineage potential by H3K9me3 in mESCs[41]. Future efforts are needed to elucidate the detailed molecular mechanisms underlying lineage-specific transcription factor binding and heterochromatin formation.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41556-018-0093-4.

Received: 15 May 2017; Accepted: 21 March 2018;
Published online: 23 April 2018

References

1. Hackett, J. A. & Surani, M. A. Beyond DNA: programming and inheritance of parental methylomes. *Cell* **153**, 737–739 (2013).
2. Li, E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat. Rev. Genet.* **3**, 662–673 (2002).
3. Burton, A. & Torres-Padilla, M. E. Epigenetic reprogramming and development: a unique heterochromatin organization in the preimplantation mouse embryo. *Brief. Funct. Genomics* **9**, 444–454 (2010).
4. Okamoto, I., Otte, A. P., Allis, C. D., Reinberg, D. & Heard, E. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303**, 644–649 (2004).
5. Feldman, N. et al. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat. Cell Biol.* **8**, 188–194 (2006).
6. Fukuda, A. et al. The role of maternal-specific H3K9me3 modification in establishing imprinted X-chromosome inactivation and embryogenesis in mice. *Nat. Commun.* **5**, 5464 (2014).
7. Tachibana, M. et al. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* **16**, 1779–1791 (2002).
8. Messerschmidt, D. M. et al. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science* **335**, 1499–1502 (2012).
9. Liu, H., Kim, J. M. & Aoki, F. Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos. *Development* **131**, 2269–2280 (2004).
10. Santos, F., Peters, A. H., Otte, A. P., Reik, W. & Dean, W. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev. Biol.* **280**, 225–236 (2005).
11. Puschendorf, M. et al. PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat. Genet.* **40**, 411–420 (2008).
12. Smith, Z. D. et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
13. Wang, L. et al. Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**, 979–991 (2014).
14. Peaston, A. E. et al. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **7**, 597–606 (2004).
15. Kigami, D., Minami, N., Takayama, H. & Imai, H. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biol. Reprod.* **68**, 651–654 (2003).
16. Leung, D. C. & Lorincz, M. C. Silencing of endogenous retroviruses: when and why do histone marks predominate? *Trends Biochem. Sci.* **37**, 127–133 (2012).
17. Bulut-Karslioglu, A. et al. Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Mol. Cell* **55**, 277–290 (2014).
18. Yang, B. X. et al. Systematic identification of factors for provirus silencing in embryonic stem cells. *Cell* **163**, 230–245 (2015).
19. Brind'Amour, J. et al. An ultra-low-input native ChIP–seq protocol for genome-wide profiling of rare cell populations. *Nat. Commun.* **6**, 6033 (2015).
20. Liu, X. et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**, 558–562 (2016).
21. Zhang, Y. et al. Model-based analysis of ChIP–seq (MACS). *Genome Biol.* **9**, R137 (2008).
22. Wu, J. et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657 (2016).
23. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
24. Becker, J. S., Nicetto, D. & Zaret, K. S. H3K9me3-dependent heterochromatin: barrier to cell fate changes. *Trends Genet.* **32**, 29–41 (2016).
25. Surface, L. E., Thornton, S. R. & Boyer, L. A. Polycomb group proteins set the stage for early lineage commitment. *Cell Stem Cell* **7**, 288–298 (2010).
26. van der Heijden, G. W. et al. Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech. Dev.* **122**, 1008–1022 (2005).
27. Zhang, B. et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**, 553–557 (2016).
28. Reik, W. & Walter, J. Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.* **2**, 21–32 (2001).
29. Inoue, A., Jiang, L., Lu, F., Suzuki, T. & Zhang, Y. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* **547**, 419–424 (2017).
30. Lewis, A. et al. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat. Genet.* **36**, 1291–1295 (2004).
31. Umlauf, D. et al. Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat. Genet.* **36**, 1296–1300 (2004).
32. Choi, Y. J. et al. Deficiency of microRNA miR-34a expands cell fate potential in pluripotent stem cells. *Science* **355**, eaag1927 (2017).
33. Fort, A. et al. Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nat. Genet.* **46**, 558–566 (2014).
34. Nakamura, T. et al. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–419 (2012).
35. Zhou, L. Q. & Dean, J. Reprogramming the genome to totipotency in mouse embryos. *Trends Cell Biol.* **25**, 82–91 (2015).
36. Macfarlan, T. S. et al. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
37. Hatanaka, Y. et al. Histone chaperone CAF-1 mediates repressive histone modifications to protect preimplantation mouse embryos from endogenous retrotransposons. *Proc. Natl Acad. Sci. USA* **112**, 14641–14646 (2015).
38. Ishiuchi, T. et al. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* **22**, 662–671 (2015).
39. Burton, A. & Torres-Padilla, M. E. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 723–734 (2014).
40. Albert, M. & Peters, A. H. Genetic and epigenetic control of early mouse development. *Curr. Opin. Genet. Dev.* **19**, 113–121 (2009).
41. Yuan, P. et al. Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev.* **23**, 2507–2520 (2009).
42. Zheng, H. et al. Resetting epigenetic memory by reprogramming of histone modifications in mammals. *Mol. Cell* **63**, 1066–1079 (2016).

Acknowledgements

This work was primarily supported by the National Key R&D Program of China (2016YFA0100400) and the National Natural Science Foundation of China (31721003). This work was also supported by the Ministry of Science and Technology of China (2017YFA0102602 and 2015CB964800), the National Natural Science Foundation of China (31430056, 81630035, 31401266, 31771646, 31701262, 31401247, 31501196, 31501183, 31571365 and 31501197), the Shanghai Subject Chief Scientist Program (15XD1503500 and 17XD1403600), the Shanghai Rising-Star Program (17QA1402700 and 17QA1404200), the Shanghai Chenguang Program (16CG17, 16CG19 and 15CG19), the Shanghai municipal medical and health discipline construction projects (2017ZZ02015) and the National Postdoctoral Program for Innovative Talents (BX20170174).

Author contributions

Y.G. and S.G. conceived and designed the experiments. C.W. performed computational analysis. X.L. and L.Y. performed the ChIP experiments. C.W., X.L., Y.Z. and Y.G. designed and performed the data analysis. W.L., C.C., X.K., J.C., Y.H.Z., Y.W., R.L., H.W. and T.D. assisted with the sample preparation. C.W., X.L., Y.G., Y.Z. and S.G. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41556-018-0093-4.

Correspondence and requests for materials should be addressed to Y.G. or Y.Z. or S.G.

Methods

Animals and collection of mouse embryos. The specific pathogen-free grade mice were housed in the animal facility of Tongji University, Shanghai, China.
To get MII oocytes and pre-implantation embryos, B6D2F1 or C57BL/6 female mice (8–10-weeks old) were superovulated by injection with 7 IU each of pregnant mare serum gonadotropin (PMSG), followed by injection of 5 IU of human chorionic gonadotropin (hCG) (San-Sheng Pharmaceutical) 48 h later. The superovulated female mice were mated with B6D2F1 or DBA2 male mice. Then, the zygotes, early 2-cell- and late 2-cell-stage embryos were collected from the oviducts of the female mice at 24 h, 32 h and 44 h after hCG injection, respectively. For zygotes, we collected PN3-stage zygotes based on the microscopic observation of the size of the two pronuclei and the distance between them for further experiments[43]. MII oocytes were collected from the oviducts of unmated female mice. To get 4-cell-, 8-cell-, morula- and blastocyst-stage embryos, late 2-cell-stage embryos were cultured in Q2 medium to reach the corresponding stage.
To get post-implantation embryos, female mice at 6.5–8.5 d.p.c. (days post-coitum) were executed and the uteri were dissected and transferred to a petri dish with PBS. Next, each decidua was carefully freed from the uterine muscle layers using properly sharpened forceps. 6.5 d.p.c., 7.5 d.p.c. and 8.5 d.p.c. embryos were carefully separated from deciduas. Reichart's membrane and the ectoplacental cone were also removed from the embryos. All experiments were performed in accordance with the University of Health Guide for the Care and Use of Laboratory Animals and were approved by the Biological Research Ethics Committee of Tongji University.

Mouse sperm extraction. Two cauda epididymus were collected from one male, the cut and the semen was squeezed out. Next, they were placed in a 1.5-ml eppendorf tube with 500 µl warm HCZB medium[44] and incubated at 37°C for 10–15 min to allow sperm to swim out.

Sample harvest for ChIP–seq, total RNA-seq and WGBS. For the MII oocytes and cleavage-stage embryos, the zona pellucidae of the embryos were removed with 0.5% pronase E (Sigma) and the embryos were then incubated in Ca^{2+}-free CZB medium[45] for 5 min. Polar bodies were removed by gentle pipetting using a fire-polished glass needle with an inner diameter of 120 µm. For ICM and trophectoderm isolation, the zona pellucidae of blastocysts were removed with 0.5% pronase E (Sigma). The embryos were then incubated in Ca^{2+}-free CZB for 20 min and the tight junctions of trophectoderm cells and ICM cells were separated by gently pipetting using a pipette with a diameter of 40–60 µm. For epiblast and extraembryonic ectoderm isolation, embryos were cut at the embryonic/extraembryonic boundary using properly sharpened forceps.
A fire-polished glass needle with an inner diameter that is slightly smaller than the width of the embryonic/extraembryonic fragment was used for separating epiblast or extraembryonic ectoderm from the visceral endoderm, and the epiblast or extraembryonic ectoderm was then incubated in Ca^{2+}-free CZB medium for 10 min. Single cells were separated by gentle pipetting using a fire-polished glass needle with an inner diameter of 15–20 µm.

Cell culture. The R1 ESCs were purchased from the American Type Culture Collection (ATCC) and not further authenticated. The RO-7 TSCs were derived in our laboratory and have been published previously, which have been verified by immunofluorescent staining and real-time PCR. All cell lines have been regularly tested negatively for mycoplasm contamination. The R1 ESCs were cultured on mitomycin C-treated mouse embryonic fibroblasts in ES medium containing DMEM (Merck Millipore) supplemented with 15% (v/v) FBS (Hyclone), 1 mM L-glutamine (Merck Millipore), 0.1 mM mercaptoethanol (Merck Millipore), 1% non-essential amino acid stock (Merck Millipore), penicillin/streptomycin (100×, Merck Millipore), nucleosides (100×, Merck Millipore) and 1,000 U ml^{-1} leukaemia inhibitory factor (Merck Millipore). The RO-7 TSCs were cultured in 70% FCM (Feeder condition medium, RPMI1640 supplemented with 20% FBS, 1% non-essential amino acid stock, 1 mM sodium pyruvate, 100 µM β-mercaptoethanol, 2 mM l-glutamine (Merck Millipore), collected from cultured feeder cells), F4H (25 ng ml^{-1} fibroblast growth factor 4 (Invitrogen) and 1.0 µg ml^{-1} heparin (Sigma)) and 30% TS medium (RPMI1640 supplemented with 20% FBS, 1 mM sodium pyruvate, 100 µM β-mercaptoethanol, 2 mM l-glutamine (Merck Millipore)).

ULI-NChIP–seq. For ULI-NChIP–seq, 500 cells or 5×10^6 sperm were used per reaction, and two or three replicates were performed for each stage. All isolated cells were washed three times in 0.5% BSA in PBS solution (Sigma) to avoid possible contamination. The ULI-NChIP procedure was performed as previously described[19]. One microgram of histone H3K4me3 antibody (9727, Cell Signaling Technology), histone H3K27me3 antibody (pAb-069-050, Diagnode) or histone H3K9me3 antibody (39161, Active Motif) was used for each immunoprecipitation reaction.
The sequence libraries were generated using the KAPA Hyper Prep Kit for the Illumina platform (kk8504), following the manufacturer's instructions. Paired-end 125-bp or 150-bp sequencing was performed on a HiSeq 2500 or X10 (Illumina) at Berry Genomics and Cloudhealth Medical Group, respectively.

WGBS. For WGBS, 100 cells were used per reaction. All isolated cells were washed three times in 0.5% BSA-PBS solution to avoid possible contamination. The sequencing libraries were generated using the Pico Methyl-Seq Library Prep Kit (D5456, Zymo Research) following the manufacturer's instructions. Paired-end 150-bp sequencing was performed on a HiSeq X10 (Illumina) at the Cloudhealth Medical Group.

Total RNA-seq. For total RNA-seq, 200 cells were used per reaction, and two or three replicates were performed for each stage. All isolated cells were washed three times in 0.5% BSA-PBS solution to avoid possible contamination. The cDNA was prepared using the Ovation RNA-Seq System V2 (7102, NuGEN) following the manufacturer's instructions. Then, the sequence libraries were generated using the KAPA Hyper Prep Kit for the Illumina platform (kk8504), following the manufacturer's instructions. Paired-end 125-bp or 150-bp sequencing was performed on a HiSeq 2500 or X10 (Illumina) at the Berry Genomics and Cloudhealth Medical Group, respectively.

siRNA KD in the early embryo. siRNAs against mouse *Chaf1a*, *Sumo2*, *Setdb1*, *Trim28*, *Ube2i* and *Zfp809* were diluted in nuclease-free water to 20 µM. The siRNAs for a given gene were mixed for KD, with a working concentration of 5 µM. Oocytes were injected with approximately 10 pl of siRNAs using a Piezo-driven micromanipulator. Oocytes were allowed to recover for at least 1 h in the incubator prior to ICSI. The sperm head was then injected into the oocyte according to the method described previously[44]. The injected embryos were observed and summarized from the 2-cell stage to the blastocyst stage. HEPES-buffered CZB (HCZB) medium was used for gamete handling and ICSI in air. CZB medium was used for embryo culture in an atmosphere of 5% CO_2. For embryo incubation, CZB was overlaid with sterile mineral oil (Sigma). The siRNA sequences were designed using GE Dharmacon siDESIGN Center (http://dharmacon.gelifesciences.com/design-center/) and were synthesized by the Biological Resource Center at NIBS. siRNA sequences are shown in Supplementary Table 3.

Reverse transcription and quantitative RT–PCR analysis. For quantitative RT–PCR analysis of siRNA-KD efficiency, total RNA of 20 4-cell-stage embryos were purified using the RNeasy Mini Kit (74104, Qiagen) according to the manufacturer's instruction. cDNA was synthesized by a reverse transcription system (Promega). Quantitative RT–PCR was performed using a SYBR Premix Ex Taq (Takara) and signals were detected with the ABI7500 Real-Time PCR System (Applied BioSystems). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as an endogenous control. Primers are listed in Supplementary Table 3.

ChIP–seq, RNA-seq and bisulfite (BS)-seq data processing and normalization. ChIP–seq reads were aligned to the mouse genome build mm9 using the bwa (v0.7.12) mem command[46]. Signal tracks for each sample were generated using the MACS2 (v2.0.10.20131216) pile-up function and were normalized to 1 million reads (RPM). To examine the reproducibility of the ChIP–seq experiments, we calculated the correlation of the normalized signal intensity between biological replicates on merged H3K9me3 peaks across all stages. As the replicates were highly correlated to each other (Pearson's correlation > 0.8), we then pooled the biological replicates together for each stage. H3K4me3 and H3K27me3 ChIP–seq data in mouse pre-implantation embryos from our previous publications were used in the analysis[20]. To minimize the effect of chromatin structure and sequencing bias, we re-normalized the ChIP–seq signal by input samples. We first divided the genome into 1-kb consecutive bins and calculated the normalized input signal (RPM) for each bin. Regions with extremely low input signal were assigned with a genomic average to prevent over amplification, then all the histone modification signal tracks were divided by the input signal in its corresponding bins and the histone modification/input ratio were $\log_2$ transformed to generate the input normalized histone signal tracks. The RNA-seq reads were mapped to the mm9 reference genome using TopHat (v1.3.3)[47]. Expression levels for all RefSeq genes were quantified to fragments per kilobase million (FPKM) using Cufflinks (v1.2.0), and FPKM values of replicates were averaged[48]. All the BS-seq reads were first processed using TrimGalore (v0.3.3) to trim adaptor and low-quality reads. Adaptor-trimmed reads were then mapped to a combined genome with mm9 and 48052 lambda sequence using bsmap (v2.89)[49]. The methylation level of each CpG site was estimated using mcall (v1.3.0)[50].

Identification of ChIP–seq peaks and comparison between stages. All the ChIP–seq peaks were called by MACS14 (v1.4.2)[21] with the parameters --nomodel --nolambda --shiftsize = 73 over input files. As the H3K9me3 peak number detected at each stage could be affected by the sequencing depths, we use the same number of reads (80 million for the H3K9me3 samples and 20 million for the input samples) when available that were randomly selected from samples of each stage. To define the established and disappeared H3K9me3 domains, we split the H3K9me3 peaks into 1-kb domains and compared them between consecutive stages: both ICM domains and trophectoderm domains were compared with the morula stage and domains of the E6.5 epiblast and extraembryonic ectoderm were compared with the ICM and the trophectoderm separately. Domains with no surrounding domains within 300 bp in the previous stage were defined as established domains, otherwise they were defined as maintained domains. Domains with no surrounding domains within 300 bp in the next stage were defined as disappeared domains. Enrichment

of histone modification peaks in promoter, short interspersed nuclear element (SINE), long interspersed nuclear element (LINE) and LTR regions was calculated using observed versus expected probability. The observed probability was calculated using the length of the histone modification peaks that covers the related genomic regions versus the length of the total histone modification peaks, and the expected probability was calculated using the length of the total related genomic regions versus the length of the mouse genome.

Clustering analysis of stage-specific H3K9me3 domains. To classify the stage-specific H3K9me3 domains, we first split the H3K9me3 peaks into 1-kb domains. We merged the domains from MII oocytes to the ICM stage and filtered the domains that cover less than two stages. The input normalized H3K9me3 signal was calculated for each domain. We then performed k-means clustering on the filtered domains, setting k = 7. Cluster 1 was defined as 'oocyte-specific domains'; clusters 2–4 were merged as 'cleavage-specific domains', as they were formed temporarily during the developmental process; and clusters 5–7 were merged as 'blastocyst-specific domains'. Heatmaps of input normalized H3K27me3 signal and DNA methylation level of each domain were generated according the order of the H3K9me3 domains. Promoters of the genes (defined as ±2 kb around the transcription start site (TSS)) that only overlapped with 'oocyte-specific domains' were defined as 'oocyte-specific genes', only overlapped with 'cleavage-specific domains' were defined as 'cleavage-specific genes' and only overlapped with 'blastocyst-specific domains' were defined as 'blastocyst-specific genes'.

Gene ontology analysis. Functional annotation was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resource[51]. Gene ontology terms for each functional cluster were summarized to a representative term, and *P* values were plotted to show the significance.

Allele assignment of sequencing reads. To assign each read to its parental origins, we first obtained the SNPs between C57BL strains and DBA strains, as well as C57BL strains and PWK strains from the Sanger Institute (http://www.sanger.ac.uk/science/data/mouse-genomes-project). We then examined all the SNPs in each read that showed high-quality base calling (Phred score ≥ 30). For paired-end reads, SNP information from both reads was summed, and when multiple SNPs were present in the same read, the parental origin was determined by votes from all SNPs. The read was assigned to the allele that at least two-thirds of the total votes supported. The public SNP-trackable H3K4me3, H3K27me3 and ATAC data sets in mouse pre-implantation development were included in the analysis[22,27,34].

Definition of lost, inherited and gained regions during fertilization. To analyse the inheritance and establishment of allelic-specific epigenetic modification from gamete to zygote, we first divided the mouse genome into 1-kb consecutive bins and only those bins that were covered by at least 10 SNP-trackable reads were considered. For histone modifications, we calculated the normalized read counts of the gametes, maternal and paternal zygotes on these regions. For maternal signals, regions with a >0.2 MII oocyte signal and a <0.2 maternal zygote signal were defined as 'lost regions'; regions with both a >0.2 MII oocyte signal and a >0.2 maternal zygote signal were defined as 'inherited regions'; regions with a <0.2 MII oocyte signal and a >0.2 maternal zygote signal were defined as 'gained regions'. Similar calculations were made for paternal signals. For DNA methylation, the cut-off for gametes was regarded as 0.7, whereas the cut-off for maternal and paternal zygotes was regarded as 0.5; similar calculations were performed for DNA methylation levels on these regions.

Identification of allelic-specific regions and genes. To identify the allelic-specific regions after fertilization, we first divided the mouse genome into 1-kb consecutive bins, and only those bins covered by at least 10 SNP-trackable reads were considered. For histone modifications, the significance of allele bias was assessed by the binomial test. For DNA methylation, the significance was assessed using the Fisher's exact test. The AS score was defined as $-\log_{10}$ (*P* value), which was defined as positive for maternal-specific regions and negative for paternal-specific regions. Allelic-specific regions were identified using a cut-off of 3 (the absolute value). Similar AS scores were calculated for RefSeq genes, with the scores calculated on the promoter regions (defined as ±2 kb around the TSS) and the minimum SNP-trackable reads as 20. Allelic-specific genes were identified using a cut-off of 2 (the absolute value). We obtained 150 imprinted genes from the mouse book (http://www.mousebook.org/mousebook-catalogs/imprinting-resource) and compared them with the allelic-specific genes defined based on H3K9me3, H3K27me3 and DNA methylation data. Paternally imprinted and maternally imprinted genes were analysed separately.

Expression, DNA methylation and histone modification level quantification of repeats elements. To assess the expression level of repeats elements, all the RNA-seq files were re-mapped to the mm9 genome using the STAR aligner software, allowing up to three mismatches and filtering out reads that mapped to >500 positions in the genome[52]. Mapped files were then processed using the makeTagDirectory script of HOMER with the -keepOne option[53]. The tag directories of the mapped files were analysed using the analyzeRepeats.pl script of HOMER with the option 'repeat' and -noadj. This script adds the reads that map multiple loci to the expression of the repeat class they represent, which were summarized to 1,221 repeat class. The total read counts of each sample were normalized to 1 million, and replicates were averaged for comparison.

To analyse the methylation level and histone modification level of repeat elements, we downloaded the repeat annotations from the UCSC table browser. DNA methylation level and normalized histone modification signals were calculated for each repeat annotations and the values of the same repeat class were summed and then averaged by the number of copies in the genome.

Association analysis of histone modification and methylation level on repeats. We performed association analysis between the DNA methylation level and the input normalized H3K9me3 signal on samples from the oocyte stage to the ICM stage. Pearson's correlation was calculated for DNA methylation and the input normalized H3K9me3 signal on genes and LTRs, and association tests were performed based on weighted Pearson's correlation coefficient[54]. Genes or LTRs with a negative correlation and significant associations ($P \leq 0.05$) have increased H3K9me3 signal and decreased DNA methylation level along development and were defined as H3K9me3-increased genes or LTRs; genes or LTRs with a positive correlation and significant associations have both decreased H3K9me3 signal and DNA methylation level along development and were defined as H3K9me3-decreased genes or LTRs. We filtered the genes and LTRs with extremely low H3K9me3 signals ($\log_2$-transformed H3K9me3/input ratio of <0.1). Similar calculations were made between the DNA methylation level and the input normalized H3K27me3 signal.

Enrichment of transcription factor-binding sites in LTRs. To calculate the enrichment of transcription factor-binding sites in LTR regions, we obtained the ChIP–seq data of *Chaf1a*, *Sumo2*, *Trim28*, *Zfp809*, *Setdb1*, *Suz12* and *Rest* in mESCs from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/)[18]. All the ChIP–seq data were processed as previously described. The Jaccard index of the transcription factor ChIP–seq peaks and LTRs were calculated using bedtools (v2.20.1)[55]. The Jaccard index of the same repeat class were summed and averaged by the copy of the class. Enrichment of the transcription factor-binding sites in LTRs was evaluated based on the Jaccard index of different groups.

Identification of LTRs with differential H3K9me3 and expression levels. The H3K9me3 level and RNA-seq read counts of all KD experiments were first normalized by the sequence depth and then compared with control experiments to identify LTRs with differential H3K9me3 and expression levels. We used the Combat function from the R sva package to remove the potential batch effect for the H3K9me3 and RNA-seq samples from different batches[56]. LTRs with an input normalized H3K9me3 signal of >0 ($\log_2$ transformed) in both control and KD experiments and an absolute fold change >0.25 ($\log_2$ transformed) were defined as differential H3K9me3 LTRs. LTRs with normalized RNA-seq read counts greater than −4 ($\log_2$ transformed) in both control and KD experiments and an absolute fold change greater than 1 ($\log_2$ transformed) were defined as differentially expressed LTRs.

Identification of genes with lineage-specific H3K9me3 level. To identify the genes with lineage-specific H3K9me3 level, we calculated the input normalized H3K9me3 level on all RefSeq gene promoters and enhancers (defined as ±10 kb around the TSS). We compared the H3K9me3 signal of the ICM and trophectoderm stage, the E6.5 epiblast and extraembryonic stage, as well as the E7.5 epiblast and extraembryonic stage. Genes with lineage-specific H3K9me3 level were defined as genes with an input normalized H3K9me3 signal of >0 ($\log_2$-transformed scaled) in both stages and an absolute fold change of >0.5 ($\log_2$ transformed).

Differential gene expression analysis. To perform differential gene expression analysis, we first calculated the read counts of each RNA-seq sample using HTSeq (v0.6.0). Then, the results were fed into edgeR to perform differential analysis. Genes with a Benjamini and Hochberg-adjusted $P \leq 0.05$ and a fold change > 1 were defined as differentially expressed between compared stages.

Transcription factor-binding analysis on lineage-specific H3K9me3 peaks around genes. H3K9me3 peaks from epiblast samples (E6.5 and E7.5) that overlapped with previously defined epiblast-specific genes but not presented in extraembryonic-specific genes were defined as epiblast-specific H3K9me3 peaks; H3K9me3 peaks from extraembryonic tissues (E6.5 and E7.5) that overlapped with extraembryonic-specific genes but not presented in epiblast-specific genes were defined as extraembryonic-specific H3K9me3 peaks. To investigate the potential factors responsible for the establishment of lineage-specific H3K9me3, we collected the publicly available ChIP–seq data in mESCs from the GEO database for 168 factors, including multiple transcription factors, histone enzymes and chromatin remodellers. We identified ChIP–seq peaks using MACS2, and for each factor, we removed the peaks with <5-fold enrichment over control and merged the peaks from different data sets to generate the final binding sites. For each factor, we calculated the number of binding sites that overlapped with epiblast-specific H3K9me3 peaks and extraembryonic-specific H3K9me3 peaks separately, the

preference of binding between epiblast-specific and extraembryonic-specific peaks was evaluated using fold change and significance was calculated using the Fisher's exact test, using the total number of epiblast-specific and extraembryonic-specific H3K9me3 peaks as background. To enlarge the potential transcription factor list, we collected 335 mouse transcription factor motifs from the Cistrome Data Collection[57] and performed genome-wide motif scanning using BINOCh. The identified motif sites were treated as potential binding sites of transcription factors and similar analyses were made to evaluate the preference of transcription factor-binding sites in lineage-specific H3K9me3 peaks.

Statistics and reproducibility. Error bars in the graphical data represent the s.d. For all the box plots presented in the analysis, the centre represents the median value and the lower and upper lines represent the 5% and 95% quantile, respectively. The number (n) for the box plots has been indicated in the corresponding figure legends. Significance between different groups was determined using one-sided Wilcoxon test, with $P<0.05$ considered to be statistical significant. For the association analysis between H3K9me3/H3K27me3 and DNA methylation signal, we choose eight time points ($n=8$, MII oocyte, zygote, 2 cell, 4 cell, 8 cell, morula, ICM and trophectoderm) during development, significance was determined using the two-sided association test. We used DAVID online tools to determine the enrichment of genes in each functional term, and the P value was determined based on a modified Fisher's exact test. The enrichment of transcription factor-binding sites in epiblast-specific/extraembryonic-specific peaks were evaluated using Fisher's exact test, and the transcription factors with $P<1\times10^{-10}$ were defined as significant candidates.

ChIP–seq for H3K9me3, H3K4me3 and H3K27me3 and RNA-seq were performed two to four times and WGBS was performed once, with the precise number of replicates and mapping quality for these data shown in Supplementary Table 1. For all experiments described above, all attempts at replication were successful, with similar results.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Codes availability. All the analysis was made based on custom python and R codes and can be available upon request.

Data availability. All the ChIP–seq, RNA-seq and WGBS data generated in this study are summarized in Supplementary Table 1 and have been deposited in the GEO database under the accession number GSE97778. H3K4me3 and H3K27me3 ChIP–seq data for pre-implantation embryos were from our previous publication (GSE73952)[20]. Allelic H3K4me3, H3K27me3 and ATAC-seq data were downloaded from previous publications (GSE66390, GSE71434 and GSE76687)[22,27,34]. ENCODE mESC H3K9me3 data were downloaded from the ENCODE website (https://www.encodeproject.org/experiments/ENCSR000CFZ/). All other data supporting the findings of this study are available from the corresponding author on reasonable request.

References

43. Adenot, P. G., Mercier, Y., Renard, J. P. & Thompson, E. M. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* **124**, 4615–4625 (1997).
44. Kimura, Y. & Yanagimachi, R. Intracytoplasmic sperm injection in the mouse. *Biol. Reprod.* **52**, 709–720 (1995).
45. Chatot, C. L., Ziomek, C. A., Bavister, B. D., Lewis, J. L. & Torres, I. An improved culture medium supports development of random-bred 1-cell mouse embryos in vitro. *J. Reprod. Fertil.* **86**, 679–688 (1989).
46. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
47. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25**, 1105–1111 (2009).
48. Trapnell, C. et al. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
49. Xi, Y. & Li, W. BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* **10**, 232 (2009).
50. Sun, D. et al. MOABS: model based analysis of bisulfite sequencing data. *Genome Biol.* **15**, R38 (2014).
51. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
52. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
53. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
54. Angermueller, C. et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* **13**, 229–232 (2016).
55. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
56. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
57. Mei, S. et al. Cistrome Data Browser: a data portal for ChIP-seq and chromatin accessibility data in human and mouse. *Nucleic Acids Res.* **45**, 658–662 (2017).

nature research

Corresponding author(s): Shaorong Gao

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No computer code were used to collect the data.
Data analysis	ChIP-seq reads were aligned to the mouse genome build mm9 using the bwa (v0.7.12) mem command. Signal tracks for each sample were generated using the MACS2 (v2.0.10.20131216) pile up function and were normalized to 1 million reads for visualization. The RNA-seq reads were mapped to the mm9 reference genome using Tophat (v1.3.3). Expression levels for all Refseq genes were quantified to FPKM using Cufflinks (v1.2.0), and FPKM values of replicates were averaged. All the BS-seq reads were first processed using TrimGalore (v0.3.3) to trim adaptor and low-quality reads. Adaptor-trimmed reads were then mapped to a combined genome with mm9 and 48052 lambda sequence using bsmap (v2.89). The methylation level of each CpG site was estimated using mcall (v1.3.0).The ChIP-seq peaks were called by MACS14 (v1.4.2)(Zhang et al., 2008) with the parameters --nomodel --nolambda --shiftsize=73 over input files. Functional annotation was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resource. To assess the expression level of repeats elements, all the RNA-seq files were re-mapped to the mm9 genome using the STAR aligner software allowing up to 3 mismatches and filtering out reads mapping to more than 500 positions in the genome(Dobin et al., 2013). Mapped files were then processed using the makeTagDirectory script of HOMER with -keepOne option. To perform differential gene expression analysis, we first calculated the reads count of each RNA-seq sample using HTSeq (v0.6.0). Then the results were fed into edgeR to perform differential analysis. Bedtools (v2.20.1) were used to calculate the Jaccard index in the analysis. We retrieved the genome wide motif sites based on the results of BINOCh (v1.0.0). All custom codes were generated using Python or R, and can be available upon request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All the ChIP-seq, RNA-seq and WGBS data generated in this study have been deposited in the Gene Expression Omnibus under the accession number GSE97778. H3K4me3 and H3K27me3 data were from GSE73952. Allelic H3K4me3, H3K27me3 and ATAC-seq data were downloaded from GSE66390, GSE71434 and GSE76687.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	500 cells or 5 × 10*6 sperms per reaction for ChIP-seq, 100 cells per eaction for WGBS and 200 cells per reaction for total RNA-sequencing. This is the minimum number of cells for the corresponding experiments to get high quality sequencing data under current conditions in our lab. As the sequencing depth of samples in this study is less than 50x, thus the cell number is sufficient for detecting the majority of the modification signal. No statistical method were used to predetermine the sample size.
Data exclusions	No data were excluded from the analyses.
Replication	ChIP-seq and RNA-seq were performed 2-4 times for each stage with the number of replications shown in Supplementary Table 1. WGBS was performed once for each stage, the number of replicates is sufficient to determine the corresponding modification level. All attempts at replication were successful.
Randomization	All oocytes or embryos were collected from the oviducts of female mice and separated to single cells . Single cells were randomly allocated for ChIP-seq, RNA-seq and WGBS.
Blinding	During the siRNA injection , investigators injected the siRNA to the embryos without knowing the types of the siRNA.

Materials & experimental systems

Policy information about availability of materials

n/a	Involved in the study
☒	☐ Unique materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☐	☒ Research animals
☒	☐ Human research participants

Antibodies

Antibodies used	Histone H3K4me3 antibody (Cell signaling Technology, #9727) ,1ug per ChIP reaction Histone H3K27me3 antibody (Diagenode, pAb-069-050) ,1ug per ChIP reaction Histone H3K9me3 antibody (Active Motif, 39161),1ug per ChIP reaction
Validation	Histone H3K4me3 antibody (Cell signaling Technology, #9727) This antibody can be used for ChIP, immunoprecipitation, Western Blotting and Immunofluorescence in human ,mouse, rat and monkey cells according to the manufacture's description. There are validation data for ChIP with Hela cells on the manufactures's website. In our experiment, we validated this antibody by performing ChIP-seq with Mouse R1 ES cells. The sequencing result is consistent with data in ENCODE database.

Histone H3K27me3 antibody (Diagenode, pAb-069-050) This antibody can be used for ChIP, ELISA, Dot Bloting ,Western Blotting and Immunofluorescence in human ,mouse and rat cells according to the manufacture's description. There are validation data for ChIP with Hela cells and HelaS3 cells on the manufactures's website. In our experiment, we validated this antibody by performing ChIP-seq with Mouse R1 ES cells. The sequencing result is consistent with data in ENCODE database.

Histone H3K9me3 antibody (Active Motif, 39161)This antibody can be used for ChIP, Dot Bloting ,Western Blotting and Immunofluorescence with wide range reactivity according to the manufacture's description. There are validation data for ChIP with D425 human Medulloblastoma cells on the manufactures's website. In our experiment, we validated this antibody by performing ChIP-seq with Mouse R1 ES cells. The sequencing result is consistent with data in ENCODE database.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	The R1 ESCs were purchased from the American Type Culture Collection (ATCC) . The RO-7 TSCs were derived in our lab and published before.
Authentication	For TS cell line, we performed immunofluorescent staining of Cdx2 and found the TS cell line is positive for Cdx2. We also perform RT-PCR of marker genes in TS cells, such as Esrrb, Sox2 and Eomes, and these genes are positively expresssed in our TS cell line. The R1 ES cell line was purchased and used in several paper published by our lab, no further authentication were performed for this cell line.
Mycoplasma contamination	All cell lines tested negetive for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Research animals

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Animals/animal-derived materials	B6D2F1 or C57BL/6 female mice (8–10 weeks old) and B6D2Fi or DBA2 male mice (10-16 week old) were used in our study.

Method-specific reporting

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ Magnetic resonance imaging

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE97778
Files in database submission	MIIOocyte.H3K9me3.1.1.fq.gz MIIOocyte.H3K9me3.1.2.fq.gz MIIOocyte.H3K9me3.2.1.fq.gz MIIOocyte.H3K9me3.2.2.fq.gz sperm.input.1.fq.gz sperm.input.2.fq.gz sperm.H3K9me3.1.1.fq.gz sperm.H3K9me3.1.2.fq.gz sperm.H3K9me3.2.1.fq.gz sperm.H3K9me3.2.2.fq.gz zygote.input.1.fq.gz zygote.input.2.fq.gz zygote.H3K4me3.1.1.fq.gz zygote.H3K4me3.1.2.fq.gz zygote.H3K4me3.2.1.fq.gz zygote.H3K4me3.2.2.fq.gz zygote.H3K27me3.1.1.fq.gz zygote.H3K27me3.1.2.fq.gz zygote.H3K27me3.2.1.fq.gz zygote.H3K27me3.2.2.fq.gz

zygote.H3K27me3.3.1.fq.gz
zygote.H3K27me3.3.2.fq.gz
zygote.H3K9me3.1.1.fq.gz
zygote.H3K9me3.1.2.fq.gz
zygote.H3K9me3.2.1.fq.gz
zygote.H3K9me3.2.2.fq.gz
2cell.H3K9me3.1.1.fq.gz
2cell.H3K9me3.1.2.fq.gz
2cell.H3K9me3.2.1.fq.gz
2cell.H3K9me3.2.2.fq.gz
4cell.H3K9me3.1.1.fq.gz
4cell.H3K9me3.1.2.fq.gz
4cell.H3K9me3.2.1.fq.gz
4cell.H3K9me3.2.2.fq.gz
8cell.H3K9me3.1.1.fq.gz
8cell.H3K9me3.1.2.fq.gz
8cell.H3K9me3.2.1.fq.gz
8cell.H3K9me3.2.2.fq.gz
morula.H3K9me3.1.1.fq.gz
morula.H3K9me3.1.2.fq.gz
morula.H3K9me3.2.1.fq.gz
morula.H3K9me3.2.2.fq.gz
ICM.H3K9me3.1.1.fq.gz
ICM.H3K9me3.1.2.fq.gz
ICM.H3K9me3.2.1.fq.gz
ICM.H3K9me3.2.2.fq.gz
TE.H3K9me3.1.1.fq.gz
TE.H3K9me3.1.2.fq.gz
TE.H3K9me3.2.1.fq.gz
TE.H3K9me3.2.2.fq.gz
ESC.H3K4me3.1.1.fq.gz
ESC.H3K4me3.1.2.fq.gz
ESC.H3K4me3.2.1.fq.gz
ESC.H3K4me3.2.2.fq.gz
ESC.H3K27me3.1.1.fq.gz
ESC.H3K27me3.1.2.fq.gz
ESC.H3K27me3.2.1.fq.gz
ESC.H3K27me3.2.2.fq.gz
ESC.H3K9me3.1.1.fq.gz
ESC.H3K9me3.1.2.fq.gz
ESC.H3K9me3.2.1.fq.gz
ESC.H3K9me3.2.2.fq.gz
TSC.H3K9me3.1.1.fq.gz
TSC.H3K9me3.1.2.fq.gz
TSC.H3K9me3.2.1.fq.gz
TSC.H3K9me3.2.2.fq.gz
TSC.H3K9me3.3.1.fq.gz
TSC.H3K9me3.3.2.fq.gz
E65Epi.input.1.fq.gz
E65Epi.input.2.fq.gz
E65Epi.H3K4me3.1.1.fq.gz
E65Epi.H3K4me3.1.2.fq.gz
E65Epi.H3K4me3.2.1.fq.gz
E65Epi.H3K4me3.2.2.fq.gz
E65Epi.H3K27me3.1.1.fq.gz
E65Epi.H3K27me3.1.2.fq.gz
E65Epi.H3K27me3.2.1.fq.gz
E65Epi.H3K27me3.2.2.fq.gz
E65Epi.H3K27me3.3.1.fq.gz
E65Epi.H3K27me3.3.2.fq.gz
E65Epi.H3K9me3.1.1.fq.gz
E65Epi.H3K9me3.1.2.fq.gz
E65Epi.H3K9me3.2.1.fq.gz
E65Epi.H3K9me3.2.2.fq.gz
E65Exe.input.1.fq.gz
E65Exe.input.2.fq.gz
E65Exe.H3K4me3.1.1.fq.gz
E65Exe.H3K4me3.1.2.fq.gz
E65Exe.H3K4me3.2.1.fq.gz
E65Exe.H3K4me3.2.2.fq.gz
E65Exe.H3K27me3.1.1.fq.gz
E65Exe.H3K27me3.1.2.fq.gz
E65Exe.H3K27me3.2.1.fq.gz
E65Exe.H3K27me3.2.2.fq.gz
E65Exe.H3K27me3.3.1.fq.gz

E65Exe.H3K27me3.3.2.fq.gz
E65Exe.H3K9me3.1.1.fq.gz
E65Exe.H3K9me3.1.2.fq.gz
E65Exe.H3K9me3.2.1.fq.gz
E65Exe.H3K9me3.2.2.fq.gz
E75Epi.input.1.fq.gz
E75Epi.input.2.fq.gz
E75Epi.H3K4me3.1.1.fq.gz
E75Epi.H3K4me3.1.2.fq.gz
E75Epi.H3K4me3.2.1.fq.gz
E75Epi.H3K4me3.2.2.fq.gz
E75Epi.H3K4me3.3.1.fq.gz
E75Epi.H3K4me3.3.2.fq.gz
E75Epi.H3K27me3.1.1.fq.gz
E75Epi.H3K27me3.1.2.fq.gz
E75Epi.H3K27me3.2.1.fq.gz
E75Epi.H3K27me3.2.2.fq.gz
E75Epi.H3K9me3.1.1.fq.gz
E75Epi.H3K9me3.1.2.fq.gz
E75Epi.H3K9me3.2.1.fq.gz
E75Epi.H3K9me3.2.2.fq.gz
E75Epi.H3K9me3.3.1.fq.gz
E75Epi.H3K9me3.3.2.fq.gz
E75Epi.H3K9me3.4.1.fq.gz
E75Epi.H3K9me3.4.2.fq.gz
E75Exe.input.1.fq.gz
E75Exe.input.2.fq.gz
E75Exe.H3K4me3.1.1.fq.gz
E75Exe.H3K4me3.1.2.fq.gz
E75Exe.H3K4me3.2.1.fq.gz
E75Exe.H3K4me3.2.2.fq.gz
E75Exe.H3K4me3.3.1.fq.gz
E75Exe.H3K4me3.3.2.fq.gz
E75Exe.H3K27me3.1.1.fq.gz
E75Exe.H3K27me3.1.2.fq.gz
E75Exe.H3K27me3.2.1.fq.gz
E75Exe.H3K27me3.2.2.fq.gz
E75Exe.H3K27me3.3.1.fq.gz
E75Exe.H3K27me3.3.2.fq.gz
E75Exe.H3K9me3.1.1.fq.gz
E75Exe.H3K9me3.1.2.fq.gz
E75Exe.H3K9me3.2.1.fq.gz
E75Exe.H3K9me3.2.2.fq.gz
E85Epi.input.1.fq.gz
E85Epi.input.2.fq.gz
E85Epi.H3K4me3.1.1.fq.gz
E85Epi.H3K4me3.1.2.fq.gz
E85Epi.H3K4me3.2.1.fq.gz
E85Epi.H3K4me3.2.2.fq.gz
E85Epi.H3K27me3.1.1.fq.gz
E85Epi.H3K27me3.1.2.fq.gz
E85Epi.H3K27me3.2.1.fq.gz
E85Epi.H3K27me3.2.2.fq.gz
E85Epi.H3K27me3.3.1.fq.gz
E85Epi.H3K27me3.3.2.fq.gz
E85Epi.H3K9me3.1.1.fq.gz
E85Epi.H3K9me3.1.2.fq.gz
E85Epi.H3K9me3.2.1.fq.gz
E85Epi.H3K9me3.2.2.fq.gz
ctrl.H3K9me3.1.1.fq.gz
ctrl.H3K9me3.1.2.fq.gz
ctrl.H3K9me3.2.1.fq.gz
ctrl.H3K9me3.2.2.fq.gz
ctrl.input.1.fq.gz
ctrl.input.2.fq.gz
siChaf1a.H3K9me3.1.1.fq.gz
siChaf1a.H3K9me3.1.2.fq.gz
siChaf1a.H3K9me3.2.1.fq.gz
siChaf1a.H3K9me3.2.2.fq.gz
siChaf1a.input.1.fq.gz
siChaf1a.input.2.fq.gz
siSumo2.H3K9me3.1.1.fq.gz
siSumo2.H3K9me3.1.2.fq.gz
siSumo2.H3K9me3.2.1.fq.gz
siSumo2.H3K9me3.2.2.fq.gz

siSumo2.input.1.fq.gz
siSumo2.input.2.fq.gz
siSetdb1.H3K9me3.1.1.fq.gz
siSetdb1.H3K9me3.1.2.fq.gz
siSetdb1.H3K9me3.2.1.fq.gz
siSetdb1.H3K9me3.2.2.fq.gz
siSetdb1.input.1.fq.gz
siSetdb1.input.2.fq.gz
siUbe2i.H3K9me3.1.1.fq.gz
siUbe2i.H3K9me3.1.2.fq.gz
siUbe2i.H3K9me3.2.1.fq.gz
siUbe2i.H3K9me3.2.2.fq.gz
siUbe2i.input.1.fq.gz
siUbe2i.input.2.fq.gz
siZfp809.H3K9me3.1.1.fq.gz
siZfp809.H3K9me3.1.2.fq.gz
siZfp809.H3K9me3.2.1.fq.gz
siZfp809.H3K9me3.2.2.fq.gz
siZfp809.input.1.fq.gz
siZfp809.input.2.fq.gz
siTrim28.input.1.1.fq.gz
siTrim28.input.1.2.fq.gz
siTrim28.H3K9me3.1.1.fq.gz
siTrim28.H3K9me3.1.2.fq.gz
siTrim28.H3K9me3.2.1.fq.gz
siTrim28.H3K9me3.2.2.fq.gz
siTrim28.H3K9me3.3.1.fq.gz
siTrim28.H3K9me3.3.2.fq.gz
zygote.C57BL/DBA.input.1.fq.gz
zygote.C57BL/DBA.input.2.fq.gz
zygote.C57BL/DBA.H3K9me3.1.1.fq.gz
zygote.C57BL/DBA.H3K9me3.1.2.fq.gz
zygote.C57BL/DBA.H3K9me3.2.1.fq.gz
zygote.C57BL/DBA.H3K9me3.2.2.fq.gz
early2cell.C57BL/DBA.input.1.fq.gz
early2cell.C57BL/DBA.input.2.fq.gz
early2cell.C57BL/DBA.H3K9me3.1.1.fq.gz
early2cell.C57BL/DBA.H3K9me3.1.2.fq.gz
early2cell.C57BL/DBA.H3K9me3.2.1.fq.gz
early2cell.C57BL/DBA.H3K9me3.2.2.fq.gz
late2cell.C57BL/DBA.input.1.fq.gz
late2cell.C57BL/DBA.input.2.fq.gz
late2cell.C57BL/DBA.H3K9me3.1.1.fq.gz
late2cell.C57BL/DBA.H3K9me3.1.2.fq.gz
late2cell.C57BL/DBA.H3K9me3.2.1.fq.gz
late2cell.C57BL/DBA.H3K9me3.2.2.fq.gz
late2cell.C57BL/DBA.H3K9me3.3.1.fq.gz
late2cell.C57BL/DBA.H3K9me3.3.2.fq.gz
ICM.C57BL/DBA.input.1.fq.gz
ICM.C57BL/DBA.input.2.fq.gz
ICM.C57BL/DBA.H3K9me3.1.1.fq.gz
ICM.C57BL/DBA.H3K9me3.1.2.fq.gz
ICM.C57BL/DBA.H3K9me3.2.1.fq.gz
ICM.C57BL/DBA.H3K9me3.2.2.fq.gz

MIIOocyte.H3K9me3.1.bw
MIIOocyte.H3K9me3.2.bw
sperm.H3K9me3.1.bw
sperm.H3K9me3.2.bw
zygote.H3K4me3.1.bw
zygote.H3K4me3.2.bw
zygote.H3K27me3.1.bw
zygote.H3K27me3.2.bw
zygote.H3K27me3.3.bw
zygote.H3K9me3.1.bw
zygote.H3K9me3.2.bw
2cell.H3K9me3.1.bw
2cell.H3K9me3.2.bw
4cell.H3K9me3.1.bw
4cell.H3K9me3.2.bw
8cell.H3K9me3.1.bw
8cell.H3K9me3.2.bw
morula.H3K9me3.1.bw
morula.H3K9me3.2.bw
ICM.H3K9me3.1.bw

ICM.H3K9me3.2.bw
TE.H3K9me3.1.bw
TE.H3K9me3.2.bw
ESC.H3K4me3.1.bw
ESC.H3K4me3.2.bw
ESC.H3K27me3.1.bw
ESC.H3K27me3.2.bw
ESC.H3K9me3.1.bw
ESC.H3K9me3.2.bw
TSC.H3K9me3.1.bw
TSC.H3K9me3.2.bw
TSC.H3K9me3.3.bw
E65Epi.K4me3.1.bw
E65Epi.K4me3.2.bw
E65Epi.K27me3.1.bw
E65Epi.K27me3.2.bw
E65Epi.K27me3.3.bw
E65Epi.K9me3.1.bw
E65Epi.K9me3.2.bw
E65Exe.K4me3.1.bw
E65Exe.K4me3.2.bw
E65Exe.K27me3.1.bw
E65Exe.K27me3.2.bw
E65Exe.K27me3.3.bw
E65Exe.K9me3.1.bw
E65Exe.K9me3.2.bw
E75Epi.K4me3.1.bw
E75Epi.K4me3.2.bw
E75Epi.K4me3.3.bw
E75Epi.K27me3.1.bw
E75Epi.K27me3.2.bw
E75Epi.K9me3.1.bw
E75Epi.K9me3.2.bw
E75Epi.K9me3.3.bw
E75Epi.K9me3.4.bw
E75Exe.K4me3.1.bw
E75Exe.K4me3.2.bw
E75Exe.K4me3.3.bw
E75Exe.K27me3.1.bw
E75Exe.K27me3.2.bw
E75Exe.K27me3.3.bw
E75Exe.K9me3.1.bw
E75Exe.K9me3.2.bw
E85Epi.K4me3.1.bw
E85Epi.K4me3.2.bw
E85Epi.K27me3.1.bw
E85Epi.K27me3.2.bw
E85Epi.K27me3.3.bw
E85Epi.K9me3.1.bw
E85Epi.K9me3.2.bw
ctrl.H3K9me3.1.bw
ctrl.H3K9me3.2.bw
siChaf1a.H3K9me3.1.bw
siChaf1a.H3K9me3.2.bw
siSumo2.H3K9me3.1.bw
siSumo2.H3K9me3.2.bw
siSetdb1.H3K9me3.1.bw
siSetdb1.H3K9me3.2.bw
siUbe2i.H3K9me3.1.bw
siUbe2i.H3K9me3.2.bw
siZfp809.H3K9me3.1.bw
siZfp809.H3K9me3.2.bw
siTrim28.H3K9me3.1.bw
siTrim28.H3K9me3.2.bw
siTrim28.H3K9me3.3.bw
zygote.C57BL-DBA.H3K9me3.1.bw
zygote.C57BL-DBA.H3K9me3.2.bw
early2cell.C57BL-DBA.H3K9me3.1.bw
early2cell.C57BL-DBA.H3K9me3.2.bw
late2cell.C57BL-DBA.H3K9me3.1.bw
late2cell.C57BL-DBA.H3K9me3.2.bw
late2cell.C57BL-DBA.H3K9me3.3.bw
ICM.C57BL-DBA.H3K9me3.1.bw
ICM.C57BL-DBA.H3K9me3.2.bw

Genome browser session (e.g. UCSC)	Not available.

Methodology

Replicates	ChIP-seq were performed 2-4 times for each stage with the number of replications shown in Supplementary Table 1. Replicates were highly agreed and the pearson's correlation coefficient on merged peaks for replicates are greater than 0.8.
Sequencing depth	PCR were performed with KAPA HiFi HotStart enzyme and amplify following the manufactures ,the cycle number for PCR amplification is 12-14. The total number of reads for each experiments is between 15M~50M, and the samples were sequenced using Illumina HiSeq 2000 with 150bp read length and paired-end sequenced. Detailed mapping summary could be referenced in Supplementary Table S1.
Antibodies	Histone H3K4me3 antibody (Cell signaling Technology, #9727) This antibody can be used for ChIP, immunoprecipitation, Western Blotting and Immunofluorescence in human ,mouse, rat and monkey cells according to the manufacture's description. There are validation data for ChIP with Hela cells on the manufactures's website. In our experiment, we validated this antibody by performing ChIP-seq with Mouse R1 ES cells. The sequencing result is consistent with data in ENCODE database. We use 1ug antibody for each ChIP reaction. Histone H3K27me3 antibody (Diagenode, pAb-069-050) This antibody can be used for ChIP, ELISA, Dot Bloting ,Western Blotting and Immunofluorescence in human ,mouse and rat cells according to the manufacture's description. There are validation data for ChIP with Hela cells and HelaS3 cells on the manufactures's website. In our experiment, we validated this antibody by performing ChIP-seq with Mouse R1 ES cells. The sequencing result is consistent with data in ENCODE database. We use 1ug antibody for each ChIP reaction. Histone H3K9me3 antibody (Active Motif, 39161)This antibody can be used for ChIP, Dot Bloting ,Western Blotting and Immunofluorescence with wide range reactivity according to the manufacture's description. There are validation data for ChIP with D425 human Medulloblastoma cells on the manufactures's website. In our experiment, we validated this antibody by performing ChIP-seq with Mouse R1 ES cells. The sequencing result is consistent with data in ENCODE database. We use 1ug antibody for each ChIP reaction.
Peak calling parameters	All ChIP and control files were mapped to mouse genome mm9 using bwa(v0.7.12) mem command with the paremeter bwa mem -t 10 mm9.fa sample.fq > sample.sam, the mapped sam files for replicates were merged and then normalized to 80M reads for comparison between stages. ChIP-seq peaks were generated using MACS14 (v1.4.2) with the parameters --nomodel --nolambda --shiftsize=73 over input files.
Data quality	All ChIP-seq peaks were filtered using 5% FDR and greater than 5-fold over input files. The number of ChIP-seq peaks are between 50K-200K after filter.
Software	DAVID Bioinformatics Resource was used to perform functional analysis. Custom analysis codes were generated using python and R, and can be available upon request.