Title: Resetting of H3K4me2 during mammalian parental-to-zygote transition

Chong Wang[1,2]†, Yang Li[1,2]†, Yaqian Wang[1,2]†, Yong Shi[1,2]†, Xiangrui Meng[1,2], Wenbo Li[1,2], Jia Guo[1,2], Kaiyue Hu[1,2], Hao Chen[2,3*],Jiawei Xu[1,2*]

[1]The First Affiliated Hospital of Zhengzhou University & Institute of Reproductive Health, Henan Academy of Innovations In Medical Science Zhengzhou, 450000 China

[2] NHC Key Laboratory of Birth Defects Prevention, Zhengzhou, 450000 China

[3]Department of Human Cell Biology and Genetics, School of Medicine; Shenzhen Key Laboratory of Gene Regulation and Systems Biology, Southern University of Science and Technology, Shenzhen, 518000, China

† These authors contribute equally to this work

* Corresponding author: jiawxu@foxmail.com or fccxujw@zzu.edu.cn, chenh7@sustech.edu.cn;

ABSTRACT

Upon sperm and oocyte fertilization, the embryo undergoes drastic histone modification reprogramming during pre-implantation development. In this study, we investigated the erasure and re-establishment of H3K4me2 in mouse GV, MII, and embryos using an improved approach called Cleavage Under Targets and Release Using Nuclease (CUT&RUN) with high-throughput sequencing. We found that H3K4me2 extensively occurs as a non-canonical pattern in mouse GV oocytes and early embryos, and parental H3K4me2 was erased by maternal LSD2 during forming the pronucleus in the zygote. The H3K4me2 was erased from GV to MII oocyte and was re-established in the late two-cell stage, removing this epigenetic burden is crucial for ZGA. We then revealed the H3K4me2 locates widespread in the CpG-rich and hypomethylated regulatory regions in the four-cell stage, and the eight-cell stage, but dramatic changes occurred in the inner cell mass (ICM) in mouse embryos. Then, these CpG-rich H3K4me2 regulatory regions became either activated or repressed. To summarize, in this study, we elucidated the pattern of H3K4me2 transition from parent to zygote and H3K4me2 profile during early embryo development. Our findings might provide deeper insights into epigenetic reprogramming during early development and in vitro fertilization in mammalian.

Highlights

1．H3K4me2 showed differential patterns in sperms and oocytes; non-canonical pattern occurred in GV oocytes, but canonical patterns were found in the sperm cells of mice.

2．The H3K4me2 peaks of GV were almost erased in MII oocytes and reconstructed after the major ZGA occurred. Maternal LSD2 was found to regulate H3K4me2 resetting during the parental-to-zygote transition. H3K4me2 may impose an "Epigenetic Burden" on ZGA.

3．Non-canonical H3K4me2 pattern was found in mammalian embryos，however，pervasive H3K4me2 in CpG-rich regulatory regions were found in four-cell and eight-cell stage embryos; The CpG-rich H3K4me2 regulatory regions were either activated or repressed in the ICM.

INTRODUCTION

Drastic epigenetic reprogramming occurs during early mammalian embryo development. Epigenetic marks, such as DNA methylation and histone modification, play important roles in the epigenome reprogramming process[1, 2]. Epigenome reprogramming during the pre-implantation embryo development stage erases the epigenetic memory of oocytes and sperm cells and allows the zygote to establish a new epigenetic landscape, which is relevant for early development and is essential for maternal zygotic transition (MZT) and differentiation[1, 3]. Deficiencies in the reconstruction of the epigenome can result in developmental defects, such as embryonic arrest[3]. Histone modifications are fundamental epigenetic regulators that control gene expression and many key cellular processes[4–6]. Different types of histone modifications can occur in specific genomic regions and perform distinct functions[3]. Most histone modifications occur around transcription start sites (TSSs) in the embryos; this feature is shared by all somatic cells[7–9]. A similar pattern of histone modification commonly occurs in many species and cell lines[10–12], and thus, this distribution pattern is known as the 'canonical' pattern[13]. Highly pervasive H3K27me3 is found in regions depleted of transcription and DNA methylation in oocytes, whereas extensive loss of promoter H3K27me3 occurs at *Hox* and other developmental genes after fertilization, accompanied by global erasure of H3K27me3 from sperm but inheritance of distal H3K27me3 from oocytes; both H3K4me3/H3K27me3 bivalent promoter marks are restored at developmental genes[14].

The K-residues at the N-terminus of histone H3 undergo mono-(H3K4me1), di-(H3K4me2), or tri-methylation (H3K4me3), catalyzed by a family of proteins that contain a SET domain; the methylation of a different number of methyl groups added represents different chromatin status[15–17]. Different numbers and kinds of histone modification marks are associated with different gene expression states[18]. H3K4me3 is

usually enriched near transcription start sites (TSS) that are associated with the activation of nearby genes[19]. H3K4me1 is a global marker of enhancers; regions enriched with H3K4me1 and H3K27ac are identified as super enhancers[20–22]. H3K4me1/H3K4me2/H3K4me3 is enriched on both sides of the TSS and plays diverse and important roles in regulating gene expression in mammals[23, 24]. Some studies found that the H3K4me2 distribution pattern is like the H3K4me3 distribution pattern in plants[25, 26]. Like LSD1, LSD2 (KDM1B) belongs to the FAD-dependent amine oxidase family and shares similar substrate specificity with LSD1, which can function as a H3K4me2 demethylase independently of LSD1[27–29]. The mechanism underlying H3K4me2 reprogramming in the early development of mammals is still unknown, and the importance of the inheritance of the H3K4me2 programming from parents to offspring needs to be determined.

H3K4me2 can coordinate the recovery of protein biosynthesis and homeostasis following DNA damage, and it is required for the growth and survival of organisms[30]. H3K4me2 enrichment in plants is negatively correlated with the gene transcript level. H3K4me2 in rice and arabidopsis specifically distributes over the gene body region and functions as a repressive epigenetic mark[31]. H3K4me2 was found to be a key epigenetic factor in the chicken PGC cell lineage determination, and a narrow domain activated PGC-specific expression genes and signaling pathways[32]. H3K4me2 can occur as early as the embryonic stem cell stage, and its distribution is dynamic and progressive throughout the mouse brain differentiation and tissue development stages[33]. About 90% of the transcription factor binding regions (TFBRs) overlap with the H3K4me2-enriched regions, and along with the H3K27ac-enriched regions, they can greatly reduce false positive predictions of the TFBRs. Additionally, the stronger signals of H3K4me2 enrichment are more likely to overlap with the TFBRs than with weaker signals[34]. A subgroup of genes enriched with high levels of H3K4me2 was found within their gene body linked to T cell functions, which was different from the TSS-centered enrichment of housekeeping genes[35]. Many regions of the sperm epigenome bearing H3K4me2 are present in spermatogonia, suggesting an early establishment of H3K4me2 in spermatogenesis, which is a highly conserved process[36].

In this study, we elucidated the mechanism underlying H3K4me2 reprogramming during the early development of mammals. We optimized the Cleavage Under Target and Release Using Nuclease (CUT&RUN)[37, 38] methods using H3K4me2 antibody to as few as 50 cells and revealed unique H3K4me2 modification patterns and epigenetic reprogramming that occur during the parental-to-zygote transition in mice. We also showed that H3K4me2 resetting depends on maternal LSD2 and parental H3K4me2 acts as an

88 “Epigenetic Burden” on mammalian ZGA.

89

90 **Results**

91 **Genome-wide profiling of H3K4me2 in mouse oocytes and early embryos**

92 The epigenome underwent significant reprogramming during early development. To determine how
93 H3K4me2 is inherited from the sperm and oocyte by the zygote during early development, we investigated
94 H3K4me2 in mouse gametes and pre-implantation embryos using CUT&RUN methods. We optimized the
95 CUT&RUN protocol (CUT&RUN; Methods) using H3K4me2 antibody. We also acquired high-quality data
96 using only 50 cells. Most peaks could be retrieved using only 50 cells (Figures S1A and S1B). Then, we
97 collected GV oocytes, MII, PN5, early two-cell embryos, late two-cell embryos, four-cell embryos, eight-cell
98 embryos, and the inner cell mass (ICM) of mice. Mouse embryonic stem cells (mESC) were used as control.
99 We profiled H3K4me2 in oocytes and early embryos using CUT&RUN and immunofluorescence (IF)
100 techniques. The results of all replicates were highly consistent (Figure 1A; Figures S1A, B, and C). We
101 identified 57,455, 25,073, 110,517, 126,318, 62,575, and 69,154 H3K4me2 peaks in the GV oocytes, late
102 two-cell embryos, four-cell embryos, eight-cell embryos, ICM, and mESC, respectively. Most of the
103 H3K4me2 was erased from GV to MII during oocyte development. The paternal H3K4me2 was also erased
104 after fertilization, and it was re-established in the late two-cell stage (Figures 1A-B). We also demonstrated
105 the H3K4me2, ATAC-seq, and mRNA expression of several developmental key genes (Figure 1C). We found
106 that H3K4me2 positively correlated with H3K4me3 and chromatin accessibility in early embryos (Figures
107 S1D-E). The H3K4me2 genome-wide pattern of four-cell embryos showed a strong association with eight-
108 cell embryos: the H3K4me2 of four-cell and eight-cell embryos was in one cluster, and the correlation was
109 high (R = 0.93) (Figures S1F-G).

110 **Distinct H3K4me2 pattern between sperm and oocyte**

111 We systematically characterized the epigenome of mouse gametes. In the analysis, we included published
112 data on DNA methylation[39], H3K4me3[40], and ATAC-seq of early embryos[41], as well as the H3K4me2 data of
113 sperm[42]. Non-canonical patterns of H3K4me3, enriched broad domains in promoters and distal regions, are
114 present in mouse oocytes as broad domains in partially methylated domains (PMDs)[43, 44]. H3K4me2 shows a
115 non-canonical pattern in the GV oocytes and sperm of mice (Figures 2A-B and E-F). Mature sperm and
116 oocytes are in a state of inactive transcription[45, 46]; however, strong ChIP-seq and CUT&RUN signal

117 enrichment of H3K4me2 was found to be related to transcription[47]. These findings suggested that H3K4me2
118 might be deposited during gametogenesis and retained in mature gametes. To confirm this speculation, we
119 performed a K-means cluster analysis for H3K4me2 in gametes. By comparing promoter H3K4me2 at the
120 sperm and GV stages, we found that the promoters could be divided into four groups (Figure 2A). The
121 promoters in the first group (sperm-specific H3K4me2) were involved in fertilization and germ cell
122 development, and these markers were erased in the later stages (Figures 2A and 2C). In contrast, the
123 promoters in the second group (GV-specific H3K4me2) were associated with mitosis and gamete generation.
124 They played a role in GV generation during oocyte development and were re-established in the late stages
125 (Figures 2A and 2C). The promoters in the third group (sperm-GV share H3K4me2) were related to mRNA
126 processing and translation. The signals of these regions were erased (MII, zygote, and early two-cell stage)
127 and reconstructed (at the late two-cell stage) during development (Figures 2A and 2C). The promoters in the
128 fourth group had a low H3K4me2 level at all stages (Figure 2A).
129 The H3K4me2 enrichment in these promoters was consistent with the changing dynamics of H3K4me3
130 (Figure 2A). Specifically, promoters with high CpG density were preferentially marked by H3K4me2 in the
131 sperm-GV share group (Figure 2A). Next, we investigated H3K4me2-marked distal regions in gametes.
132 Sperm-specific groups are involved in gamete generation (Figure 2D). GV-specific distal H3K4me2 showed
133 functions in cell differentiation and mitotic cell cycle (Figure 2D). Both were erased after fertilization and not
134 reconstructed during embryo development (Figure 2B). H3K4me3 signals did not occupy the sperm-
135 specific/GV-specific H3K4me2 distal regions (Figure 2B). The GV-sperm-shared distal H3K4me2 regions
136 were also marked by H3K4me3. Those genes participated in system development, RNA processing, and
137 regulation of gene expression (Figures 2B and 2D). Compared to GV, sperm H3K4me2 was enriched in a
138 fully methylated promoter, but the distal H3K4me2 of both sperm and GV were enriched in the
139 hypomethylated region (Figures S2A-D). These results indicated that differential H3K4me2 modification
140 patterns occurred between sperms and oocytes.
141
142 **Resetting H3K4me2 in pre-implanted embryos**
143 Significant epigenetic reprograming and chromosome remodeling associated with H3K4me3, H3K27me3,
144 H3K36me3, and H3K27ac occurs during early mammalian development[44, 48, 49]. In another study, we found
145 that epigenome reboot occurred during the early developmental stage in humans[50]. We investigated how

146 parental histone mark H3K4me2 at regulatory elements are reprogrammed after fertilization. We found that
147 GV and sperm H3K4me2 peaks were almost erased in PN5 and re-established along with zygotic genome
148 activation (ZGA). This result was further confirmed by IF analysis, which showed that the reconstruction of
149 H3K4me2 might start from the late two-cell stage (Figures 1B and 3B). The H3K4me2 demethylases
150 KDM1B (LSD2) were also highly expressed at the MII stage (Figure S3A). The CUT&RUN data, which
151 showed small H3K4me2 peaks (related to other stages), reappeared at the late two-cell stage (Figure 3A). We
152 concluded that the IF analysis could not detect the newly established weak H3K4me2 signal in the late two-
153 cell stage. H3K4me2 was correlated with the activation of Prdm9 and Kmt2c at the four-cell stage (Figure
154 S3A), and the signal was strong enough to be detected at the four-cell and eight-cell stages (Figures 3A-B)
155 using both the CUT&RUN and IF techniques. The four/eight-cell H3K4me2 targets only partially overlapped
156 with those in sperm or oocytes, and these overlapped targets were mainly enriched in development-associated
157 pathways or processes (Figure 3C). Additionally, gamete-specific H3K4me2 targets were involved in the
158 development and generation of gametes, whereas four/eight cell-specific targets were associated with cell
159 communication and metabolic processes (Figure 3C). These findings indicated that the resetting of H3K4me2
160 involves distinct sets of genes that are closely related to the development of embryos from mouse gametes to
161 early embryos. About 1% of the four/eight cell H3K4me2 occurred in the hypomethylated regions as early as
162 the late two-cell stage (Figures S3B-D). These results indicated that the resetting of H3K4me2 starts from the
163 late two-cell stage but is mostly re-established at the four/eight-cell stage.
164
165 **The inhibition of LSD2 increased the expression of H3K4me2, which impaired the development of**
166 **embryos in vitro**
167 Histone post-translational modifications are coordinated by histone methyltransferases and demethylases[51].
168 KDM1B, also known as LSD2, is the only demethylase homologous to LSD1 in mammals[29]. Although they
169 share the H3K4me1/2 demethylase function, our analysis showed that Kdm1a was rarely expressed during
170 mouse embryonic development (Figure 4A). While, LSD2, which is required to establish the maternal
171 imprinted gene[52], was highly expressed before the early two-cell stage and decreased after the late two-cell
172 stage (Figures 4A and S3A).
173 Tranylcypromine (TCP) can inhibit the expression of LSD2[53]. To investigate the effect of the expression of
174 H3K4me2 on embryonic development, mouse PN5 embryos were cultured in vitro with three different TCP

175 treatment groups (3.6 mM, 5 mM, and 10 mM) and a control group. The inhibition in the presence of 3.6 mM
176 TCP was not prominent, and some embryos could develop to the eight-cell stage at 65 h. When 10 mM TCP
177 was present, embryonic development was strongly inhibited, and the embryo could not develop to the two-
178 cell stage (Figure S4A). When 5 mM TCP was present, the embryos could pass through the two-cell stage
179 with a significant developmental delay. Some embryos developed till the four-cell stage, and then, were
180 completely arrested. These results showed that TCP affected embryonic development in a dose-dependent
181 manner and caused developmental arrest. We concluded that the TCP concentration of 5 mM in the
182 experimental group (5 mM TCP group) could be used to determine whether the abnormal expression of
183 H3K4me2 affects the development of embryonic cells, especially the occurrence of ZGA.
184 We found that adding TCP to culture embryos significantly reduced the expression of LSD2 in embryos
185 (Figure 4B). After 28 h of culture, some of the embryos in the Control group developed to the four-cell stage,
186 while those in the 5 mM TCP group remained at the two-cell stage (Figures 4C and S4B). The development
187 of the embryos was significantly delayed after TCP was added compared to the time taken for the embryos in
188 the Control group to develop. The number of embryos at each stage in the 5 mM TCP group (n = 169) and the
189 Control group (n = 155) was counted and statistically analyzed (Figure 4D). The results showed that 62.96%
190 of embryos in the 5 mM TCP group developed to the two-cell stage (Control: 96.29%), and 36.82% of
191 embryos developed to the four-cell stage (Control: 93.83%). All embryos in the TCP group were arrested at
192 the four-cell stage ($P < 0.001$).
193 To test whether LSD2 specifically functions as H3K4me2 demethylase in mouse embryos,
194 immunofluorescence staining of cells in the 5 mM TCP group and the Control group was performed using
195 H3K4me1, H3K4me2, and H3K4me3 antibodies, respectively. We calculated the immunofluorescence
196 intensity of individual embryos and found no significant difference in the changes in the expression of
197 H3K4me1 between the 5 mM TCP group and the Control group (Figure S4C); also, the changes in the
198 expression of H3K4me3 was not statistically significant (Figures S4C). In contrast, when LSD2 expression
199 was inhibited by TCP, the expression of H3K4me2 at the two-cell stage was significantly higher than the
200 expression of H3K4me2 in the cells of the Control group ($P < 0.001$) (Figure 4E).Elimination of the
201 demethylase action of LSD2 increased the expression of H3K4me2, which delayed and arrested embryonic
202 development. We also found that LSD2 specifically demethylated H3K4me2 in mammalian embryos.
203

204 **Parental H3K4me2 was erased by maternal LSD2 during forming the pronucleus in the zygote**

205 During spermatogenesis, most of the histone modifications are replaced by protamine, but a few histones,
206 including H3K4me2, remain conserved[54]. During the early stages of zygote formation, female and male
207 pronuclei are transcriptionally silent, and the development process is regulated by maternal factors. In our
208 study, during the early stage of zygote formation, the expression of H3K4me2 in PN5 was lower than that in
209 PN3 ($P < 0.001$) (Figures 5A-B), which indicated that H3K4me2 signaling was gradually reduced during the
210 fusion of sperm and egg. When PN3 embryos were cultured to the PN5 stage under 5 mM TCP conditions,
211 the fluorescence signal of H3K4me2 increased significantly ($P < 0.001$) (Figures 5A-B), suggesting that the
212 LSD2 might response the removal of parental H3K4me2 during normal embryonic development. The
213 immunofluorescence values of H3K4me2 in female and male pronuclei were assessed to verify whether the
214 effect of LSD2 inhibition on H3K4me2 was different between male and female pronuclei. When LSD2 was
215 inhibited, the expression of H3K4me2 increased significantly in female and male pronuclei, especially in the
216 male pronucleus (334% increase in the male pronucleus and 155% increase in the female pronucleus) (Figure
217 5C). Sperm carrying H3K4me2 might be more affected by LSD2 demethylase, creating a low H3K4me2
218 expression environment for zygotic genome activation. To summarize, parental H3K4me2 was erased by
219 maternal LSD2, the H3K4me2 in the male pronucleus was erased faster than that in the female pronucleus,
220 and maternal LSD2 specifically demethylated H3K4me2.

221

222 **H3K4me2 participated in the mammalian ZGA and chromatin accessibility regulation**

223 We used the CUT&RUN technology to detect H3K4me2 throughout the genome and performed gene
224 annotation of the differential expression peaks of H3K4me2 in the TCP-treated group and the control group.
225 We found that the H3K4me2 signal increased in 251 genes and decreased in 194 genes (Fold changes > 2, P
226 < 0.01) (Figure 6A). The enrichment analysis of the genes from which H3K4me2 differential peaks
227 originated showed that the molecular function was mainly enriched in the transcription regulation of the
228 promoter of RNA polymerase II, and 35 differential genes were involved in this regulation (Figure 6B).
229 Unlike LSD1, which binds to the promoter region, LSD2 is absent in the promoter region and enriched in the
230 intronic, exonic, and intergenic regions[29]. The distribution of H3K4me2 changed when LSD2 expression was
231 inhibited (Figures 6C-D). The increase in the H3K4me2 signal mainly occurred in the intergenic and
232 genebody regions (intergenic region: 55.05%, genebody region: 32.66%). We speculated that H3K4me2

233 located in these regions might be erased by LSD2. However, the deletion of LSD2 did not increase the
234 H3K4me2 signal in the promoter region.
235 Histone modifications can regulate early embryonic development in mammals by affecting the expression of
236 developmental genes and reprogramming the chromatin structure. RNA-seq data of mouse germ cells and
237 embryos (refer to GEO database: GSE71434) were used to screen minor/major ZGA genes. We examined the
238 gene expression levels of two-cell embryos in the 5 mM TCP group and the Control group (fold changes > 2,
239 $P < 0.05$) and found that 398 genes were up-regulated and 724 genes were down-regulated in the two-cell
240 stage embryos of mice after LSD2 was inhibited (Figures 6E-F). Among these differentially expressed genes,
241 94 down-regulated genes were major ZGA genes, and 18 up-regulated genes were major ZGA genes. The
242 Zfp825 gene played a positive regulatory role in RNA polymerase II promoter transcription. Inhibition of
243 LSD2 can lead to a reduction in the distribution of H3K4me2 on the Zfp825 gene and a weakening of the
244 expression of Zfp825. These changes may affect the frequency, rate, or extent of RNA polymerase II
245 promoter transcription.
246 An increase in the expression of H3K4me2 changed the accessibility of chromatin, of which 1,413 peaks
247 were up-regulation and 914 peaks were down-regulated (fold change > 2; $P < 0.05$) (Figure 6G). Peak-related
248 gene screening of the genes near the promoter region (TSS ±2.5 kb) showed that 240 peak-related genes were
249 up-regulated and 197 peak-related genes were down-regulated. The transcription of genes requires the
250 opening of the tight structure of chromatin to allow the binding of regulatory factors. The up-regulation of
251 H3K4me2 alters the original accessibility state of chromatin, leading to disorders associated with chromatin
252 opening in some developmental genes, which in turn can affect developmental progression. The inhibition of
253 LSD2 expression leads to changes in the distribution of H3K4me2 on genes, abnormal expression of genes
254 associated with ZGA, and disorders related to chromatin opening; these changes together inhibit normal
255 embryonic development.
256
257 **Non-canonical H3K4me2 in mammalian embryos and pervasive H3K4me2 in CpG-rich regulatory**
258 **regions in the four-cell and eight-cell embryos of mice**
259 In GV oocytes and post-ZGA stages, H3K4me2 was enriched in both canonical gene promoters and partially
260 methylated domains (PMDs) (Figures 7A-D). Our results showed that H3K4me2 might be re-established in
261 the late two-cell stage, and strong promoter H3K4me2 enrichment might be completed in the four-cell stage

262 and maintained until the eight-cell stage. We also found that the H3K4me2 signals occupied broader regions
263 than those at other stages (Figure 1A). About 65% of these gene promoters were occupied by H3K4me2 in
264 the four-cell and eight-cell stages and the ICM; these genes played key roles in RNA metabolism and RNA
265 splicing (Figure 7E). We then combined previous H3K4me3 data and chromosome accessibility data and
266 found that these regions overlapped with high CG density regions, H3K4me3 enrichment, and chromatin
267 accessibility regions (Figure 7E). Our results showed that H3K4me2 preferentially occupied hypomethylated
268 and CpG-dense regions (Figures S5A-C).
269 We further examined how the stage-specific distal peaks were reprogrammed and found that about 35% of
270 the promoters of the genes that were associated with cell differentiation and development did not undergo the
271 large-scale re-established of H3K4me2 in the four-cell, eight-cell, and blastocyst stages (Figure 7E). We
272 identified genes expressed in specific stages and examined their promoter H3K4me2 (Figure S5E). We found
273 promoters with constantly high H3K4me2 that are preferentially CpG-rich, whereas promoters with
274 constantly low H3K4me2 were generally CpG-poor (Figures 7F and S5E). A group of genes showed dynamic
275 promoter H3K4me2 that correlated with gene expression (Figure S5E). These genes played key roles in
276 development, differentiation, and morphogenesis. Therefore, our results indicated that H3K4me2 of the
277 promoter during early embryonic development correlated with gene activities and CpG density.
278 Next, we investigated distal H3K4me2 signals and found that four/eight-cell embryos also showed
279 widespread distal H3K4me2 signals (Figure 7C). We first identified the distal H3K4me2 regions in mouse
280 gametes and pre-implantation embryos. These regions were relatively CpG-rich (although lower than
281 promoters) and hypomethylated (Figures S5B and S5C). However, compared to promoters, distal H3K4me2
282 was correlated with CpG density in a continuous and monotonic manner (Figure S5A). To summarize, these
283 results indicated that widespread distal H3K4me2 occurred in CpG-rich and hypomethylated regions in
284 mouse four/eight-cell embryos.
285
286 **Transcription regulation, CcREs priming, and resolving during maternal-to-zygotic transition**
287 A study showed that most of the transcription factor binding regions overlap with H3K4me2-enriched
288 regions[55]. To investigate the association between H3K4me2 and transcription regulation, we identified stage-
289 specific distal regions and the accessibility in these distal peaks of H3K4me2 (Figure 7F). Then, we
290 identified the binding motifs for transcription factors (TFs) enriched in these distal regions (Figure S5D). TFs

291 in mESC and eight-cell embryos include pluripotency factors involved in the development of the system[56, 57].
292 We found that CTF was specifically enriched in eight-cell embryos. The GATA and KLF families are
293 enriched in ICMs.
294 In a study on humans, we found widespread deposition of H3K4me3 at the promoters, including those for
295 developmental genes, before ZGA. These genes were either activated or repressed after ZGA[50]. Such
296 promoter priming is related to embryonic development. To determine whether this also occurs in mice, we
297 identified specific H3K4me2 target genes in the four/eight-cell stage and compared them to genes that are
298 expressed or inactive at all stages. The H3K4me2 of promoters in the four/eight-cell-specific group decreased
299 in ICM (Figures 8A and S6A). The "active" genes showed strong H3K4me2, H3K4me3, and ATAC-seq
300 signals before and after ZGA (Figures 8A and S6A). Next, we investigated if such priming-resolving
301 epigenetic transition also occurs in the distal regions. Distal H3K4me2 sites were enriched on regulatory
302 elements, as 75.4% sites (compared to 16.2% of random) of them overlapped with candidate cis-regulatory
303 elements (CcREs), identified by ENCODE (Figure 8B).
304 Enhancers are DNA sequences that increase promoter activity and, in turn, the frequency of gene
305 transcription. The enhancer region is usually modified by H3K27ac (catalyzed by P300/CBP) and H3K4me1
306 histone 2 (catalyzed by MLL3/4)[58], whereas the H3K4me3 modification level is low. This modification
307 pattern is the opposite of the pattern found in the promoter region. We compared histone modification to
308 H3K4me2 and H3K27ac and found no significant signal overlap between the two histone modifications at
309 specific developmental stages in the distal region. Distal four/eight-cell H3K4me2 marked CcREs showed
310 high enrichment of chromatin accessibility (Figures 8C and S6B). After the four/eight-cell stage, one class
311 ("active") remained accessible and was associated with basic biological processes (GREAT analysis). In
312 contrast, the other class (Figures 8C and S6B; "poised") was associated with H3K27me3 in the ICM and
313 somatic cells; it was located preferentially near developmental genes. These results indicated that the
314 widespread *de novo* H3K4me2 was re-established following ZGA, and H3K4me2 might be globally reset to a
315 permissive state before it is resolved after the four/eight-cell stage (Figure 8D).
316

317 **DISCUSSION**

318 Histone modifications are fundamental epigenetic marks. The reprogramming of histone marks plays an
319 important role in early development[9]. In this study, we profiled the H3K4me2 landscape in mESC, oocytes,

320 and early embryos using a modified low-input, high-sensitivity CUT&RUN method. We found the H3K4me2
321 showed differential patterns in sperms and oocytes and exhibited non-canonical patterns in GV oocytes but
322 canonical patterns in the sperm in mice[36]. Additionally, the GV H3K4me2 peaks were almost erased in MII
323 oocytes, and reconstructed accompanying ZGA occurred. Our results indicated that maternal H3K4me2 was
324 lost during oocyte maturation. The non-canonical H3K4me2 in mammalian embryos and pervasive
325 H3K4me2 in CpG-rich regulatory regions were found in four-cell and eight-cell embryos and in ICM. Hence,
326 we described the first landscape and elucidated the potential function of H3K4me2 during early mammalian
327 development. We also proposed a novel model, known as "Epigenetic Burden", according to which parental
328 H3K4me2 exhibits an "Epigenetic Burden" of maternal-zygote-transition. This burden is erased by maternal
329 factor LSD2 before ZGA. H3K4me2 plays a regulatory role after ZGA (Figure 8 X). H3K4me2 undergoes a
330 transition that starts from an "Epigenetic Burden", and then becomes an "Epigenetic regulator" during early
331 histone modification reprogramming.
332 The paternal epigenetic inheritance is important for development and disease in offspring. Significant
333 epigenome reprogramming occurs and is associated with key processes during early mammalian
334 development[1, 3, 9, 14, 39–41, 59–63]. Genome-wide H3K4me3 exhibited a non-canonical pattern (ncH3K4me3) in
335 full-grown and mature oocytes, i.e., it occurred as broad peaks at promoters and many distal loci[40, 44, 48]. We
336 found that H3K4me2 was a non-canonical pattern in the GV oocyte, but it was weakened in the MII oocyte,
337 indicating that H3K4me2 plays an important role during oocyte maturation. H3K4me3 is inherited in pre-
338 implantation embryos and is erased in the late two-cell embryos when the process of canonical H3K4me3 is
339 initiated[40,44]. The re-establishment of H3K4me2 mainly starts in the late two-cell stage, and it exists as a non-
340 canonical pattern in the following stages. Several studies have shown that in many species, including mice,
341 zebrafish, and humans, the parental epigenome undergoes dramatic reprogramming after terminally
342 differentiated gamete fertilization, followed by subsequent re-establishment of the embryo epigenome, which
343 leads to epigenetic ''rebooting''[9]. We found that H3K4me2 was erased by LSD2 after fertilization and re-
344 established at the late two-cell stage. Inhibition of the deletion of H3K4me2 resulted in the arrest of
345 embryonic development and abnormal expression of genes related to ZGA. This indicated that H3K4me2 is
346 involved in the reactivation of the embryonic epigenome, and parental H3K4me2 is an "Epigenetic Burden".
347 H3K4me3 is an active gene expression marker, which mainly marks the gene promoter region. H3K4me1
348 and H3K4me2 mark enhancer and promoter regions. They have multiple functions for gene activation,

349 according to different enzymes and complexes[22, 30, 35, 64–73]. Some studies have reported a distinct pattern of
350 H3K4me2 enrichment, spanning from the transcriptional start site (TSS) to the gene body, which marks a
351 subset of tissue-specific genes. This pattern is associated with unique H3K4me2 profile mark tissue-specific
352 gene regulation[33, 35]. Our results showed non-canonical H3K4me2 in mammalian embryos and pervasive
353 H3K4me2 in CpG-rich regulatory regions of four-cell and eight-cell embryos of mice. We next investigated
354 whether H3K4me2 is associated with active or repressed genes in embryo development and found that
355 promoter H3K4me2 in early mouse embryo development is correlated with gene activities and high CpG
356 density. We also found that promoter H3K4me2 was consistent with H3K4me3.
357 Histone modification and chromatin accessibility profiles are usually used to identify transcription factor
358 binding regions and motifs. In another study, we found large-scale chromatin accessibility in pre-ZGA human
359 embryos, and these regions contained many transcription factors[3]. Around 90% of the transcription factor
360 binding regions overlap with the H3K4me2-enriched regions. Stronger signals of H3K4me2 enrichment
361 indicate a greater chance of overlap with the transcription factor binding regions[34]. In this study, we identified
362 the binding motifs for transcription factors enriched in these distal regions, and the CCAAT transcription
363 factor (CTF) was found to be specifically enriched in eight-cell embryos. The GATA and KLF families were
364 enriched in ICMs. GATA is a key regulator of primitive endoderm (PE), which is a part of the ICM[74], and
365 KLF factors support naive pluripotency[75]. These results showed that stage-specific distal H3K4me2 involves
366 transcription circuitry during mammalian embryonic development.
367 The state of chromatin influences DNA replication, DNA repair, and gene expression. Accessible chromatin
368 occupies regulatory sequences, such as promoters and enhancers[76]. These elements interact with cell type-
369 specific transcription factors to start transcription processes that determine cell fate[77]. Animal embryos in the
370 early stages undergo extensive histone modification reprogramming, and chromatin remodeling before
371 terminally differentiated gametes develop into pluripotent cells. Some studies have found extensive overlap
372 of accessible promoters and enhancers at the four-cell stage in humans[50]. This finding indicated that crosstalk
373 occurs between histone modification and the chromatin state. In our study, most of the H3K4me2 was re-
374 established in chromatin-accessible regions, including promoters and distal regions at the four-cell stage.
375 Thus, we inferred that H3K4me2 might play an important role in shaping the chromatin state. Our results
376 indicated that crosstalk occurs between accessible chromatin and histone modification.
377 Dynamic histone modification plays an important role during mammalian maternal-to-zygotic transition

378 (MZT). Approximately 22% of the oocyte genome associated with broad H3K4me3 is negatively correlated
379 with DNA methylation and becomes confined to transcriptional-start-site (TSS) regions in two-cell embryos,
380 accompanied by the onset of major ZGA[48]. Also, H3K4me2 is erased in MII oocytes, and sperm H3K4me2 is
381 erased after fertilization. Similar to the pattern found for H3K4me3, H3K4me2 of the embryo was also re-
382 established accompanied by the onset of major ZGA. The ncH3K4me3 in oocytes overlaps almost
383 exclusively with partially methylated DNA domains[40]. H3K4me2 is also enriched in canonical gene
384 promoters and partially methylated domains (PMDs).
385 Overall, our findings provided a genome-wide landscape of H3K4me2 modification in mammalian oocytes
386 and pre-implantation embryos. We showed that maternal LSD2 specifically demethylates H3K4me2 after
387 fertilization in mammals. Also, H3K4me2 is erased to a greater extent in the male pronucleus than in the
388 female pronucleus. Based on our results, we proposed a novel model, the "Epigenetic Burden" of histone
389 modification reprogramming during mammalian development. This model can be used to further investigate
390 the epigenetic regulation network in the early developmental stages of mammals.
391
392 MATERIALS AND METHODS
393 Culture of K562 cells
394 The human K562 cell line was cultured at 37 °C, 5% CO_2, and saturated humidity. The cells were cultured in
395 1640 culture medium (GIBCO) containing 10% fetal bovine serum and 90% 1640 culture liquid.
396 Cell culture of mouse ESC
397 Mouse embryo stem cells (ESC) were cultured on 0.1% gelatin-coated plates without irradiated mouse
398 embryonic fibroblasts (MEFs), using standard ESC medium consisting of DMEM (GIBCO) supplemented
399 with 15–20% fetal bovine serum (FBS, Sigma), 0.1 mM β-mercaptoethanol (Sigma), 0.1 mM MEM non-
400 essential amino acids (Cellgro), 2 mM Glutamax (GIBCO), 1% nucleoside (Millipore), 1% EmbryoMax®
401 Penicillin-Streptomycin Solution (Millipore), and 10 μg/L recombinant leukemia inhibitory factor (LIF)
402 (Millipore). The mouse embryo stem cells were cultured at 37 °C, 5% CO_2, and saturated humidity.
403 Collection of mouse gametes and early embryos
404 The mouse oocytes were collected from 4–5-week-old C57BL/6N female mice (purchased from Charles
405 River). The pre-implantation embryos were harvested from 4–5-week-old C57BL/6N females mated with
406 DBA/2 males (eight-week-old when purchased from Charles River). The Germinal vesicle (GV) oocytes

407 were obtained from 4–5-week-old C57BL/6N females, 46–48 h after they were injected with 10 IU of
408 pregnant mare serum gonadotropin (PMSG, Solarbio). The 4–5-week-old C57BL/6N female mice were also
409 administered 10 IU of human chorionic gonadotrophin (HCG) intraperitoneally, 46–48 h after PMSG was
410 administered. Then, they were mated with DBA/2 males. The PN5 zygotes were collected at 28–32 h post-
411 hCG injection. Early two-cell, late two-cell, four-cell, and eight-cell embryos, morulas, and blastocysts were
412 harvested at 32–36 h, 45–48 h, 54–58 h, 60–64 h, 68–75 h, and 84–96 h post-hCG injection. The oocytes in
413 the metaphase of second meiosis (MII) were collected 12–14 h after 10 IU of hCG was injected
414 intraperitoneally. The zona pellucida of all oocytes and embryos was removed using glutathione (GSH
415 Solarbio).

416 Tranylcypromine treatment of mouse embryos

417 One-cell embryos were obtained at 24–26 h (PN3) and 28–32 h (PN5) post hCG injection and cultured in the
418 G1 medium in the presence or absence of Tranylcypromine (5 mM). After 48 h, the embryos were transferred
419 to the G2 medium with or without 5 mM TCP. All embryos were maintained at 37 °C in a humidified
420 atmosphere containing 5% CO_2. Then, morphologically normal PN5, late two-cell, and four-cell embryos
421 were collected for further experiments.

422 Immunofluorescence

423 The samples (mouse oocytes or pre-implantation embryos) were washed twice in MII medium (Sigma), and
424 4% PFA was used as a fixation buffer (BOSTER, AR1068) to fix samples at room temperature for about 60
425 min. Next, the samples were transferred using a mouth pipette and washed once with 1% BSA (Sigma,
426 A1933) in 1×PBS (Corning, 21–040-CVR). The samples were then permeabilized using 0.5% Triton X-100
427 (Sigma, T8787) in 1×PBS at room temperature for 20 min. Then, they were washed with 1% BSA in 1×PBS
428 and blocked using 1% BSA in 1×PBS at room temperature for 1 h. The blocked samples were incubated with
429 H3K4me2 antibody (Sigma Millpore, 07030) in 1% BSA in 1×PBS diluted at a ratio of 1:500, at room
430 temperature for 2 h or 4 °C overnight. After incubation, they were washed at least five times with 1% BSA in
431 1×PBS for 3 min. The samples were incubated with Alexa Fluor 568-conjugated donkey anti-rabbit IgG
432 (Invitrogen, A10042) in 1% BSA in 1×PBS diluted at a ratio of 1:200, at room temperature for 50 min in the
433 dark. After incubation, the samples were washed five times with 1% BSA in 1×PBS for 3 min. Finally, the
434 samples were stained using the ProLong@ Gold Antifade Reagent with DAPI (Cell Signaling, 8961S), using a
435 4-Chamber Glass Bottom Dish (100 per case), 35 mm Dish with 20 mm Bottom Well (Cellvis, D35C4–20–1-

436 N). The samples were incubated at room temperature for about 24 h in the dark. All images were captured
437 using a Zeiss 700 confocal microscope. Raw images were projected along the Z-axis and processed using the
438 ZEN software (Zeiss) with brightness and contrast adjusted (the same parameters were used for each channel
439 of all samples). In the functional test section, we used the LSD2 antibody (Abcam, ab193080), H3K4me1
440 antibody (Abcam, ab176877), H3K4me2 antibody (mAb) (Active motif, 39079), and H3K4me3 antibody
441 (Active motif, 39016) as primary antibodies (diluted at a ratio of 1:500). Secondary antibodies (diluted at a
442 ratio of 1:300) used included Goat Anti-Rabbit lgG H&L (Alexa Fluor@488) (Abcam, ab150077) and Goat
443 Anti-Mouse IgG H&L (Alexa Fluor® 488) (Abcam, ab150113).

444 Zona pellucida removal and ICM/TE isolation

445 Before the CUT&RUN experiment was conducted, to guarantee the effect of permeation, the zona pellucida
446 of all oocytes and embryos was chemically removed by glutathione (GSH Solarbio). The zona pellucida was
447 removed after about 10–30 s. Then, the samples were washed twice in PBS (Corning 21–040-CVR) and
448 washed once with wash buffer. The inner cell mass (ICM)s or TEs of day 4–5 blastocysts were isolated
449 mechanically. Before performing CUT&RUN, the separated cells were washed twice with PBS and once
450 with wash buffer, and then, resuspended in wash buffer.

451 RNA-seq

452 For RNA-seq, 30–40 two-cell embryos were used for one sample. They were transferred to 1× PBS drops at
453 room temperature and washed thrice to avoid any contamination. The cells were lysed in 1 μL of ice-cold
454 lysis buffer for 10 min. Transcriptome amplification was performed using a Single Cell Full Length mRNA-
455 Amplification Kit (Vazyme, N712). The RNA-seq library was constructed using True Prep DNA Library Prep
456 Kit V2 for Illumina® (Vazyme, TD502). Finally, the library was sequenced using an Illumina Nova 6000
457 sequencer (Illumina). The data were analyzed following standard RNA-seq procedures.

458 Generation of miniATAC-seq library

459 The miniATAC-seq libraries were prepared as previously described[3]. Cells for miniATAC-seq were prepared
460 in the same way as those prepared for RNA-seq.

461 CUT&RUN library construction and next-generation sequencing

462 The CUT&RUN method was used following published protocols[37, 38] with modifications in antibody binding,
463 adaptor ligation, and DNA purification steps. All samples (including mouse zygotes, embryos with zona
464 pellucida removed, or ESCs) were transferred into 0.2 mL conventional, non-low-binding PCR tubes

465 (Axygen). The samples were resuspended in 60 μL of wash buffer (20 mM HEPES-KOH, pH = 7.5; 150 mM
466 NaCl; 0.5 mM Spermidine; Roche complete protease inhibitor). Concanavalin-coated magnetic beads
467 (Polyscience, 86057) were gently washed thrice or four times and resuspended in binding buffer (20 mM
468 HEPES-KOH, pH = 7.9; 10 mM KCl; 1 mM $CaCl_2$; 1 mM $MnCl_2$), and then, 10–15 μL of Concanavalin-
469 coated magnetic beads were added to each sample. The samples were incubated at room temperature (about
470 25 °C) at 450 rpm for 20 min in a Thermomixer (Eppendorf). The samples were placed on a magnetic stand
471 for removing the supernatant, and then, resuspended in 50 μL antibody buffer (wash buffer plus digitonin
472 1000:1, 2 mM EDTA, pH = 8.0) with H3K4me2 antibody (Sigma Millpore, 07030) diluted at ratio of 1:100.
473 The samples were incubated at room temperature (23–26 °C) for 2 h at 450 rpm on Thermomixer, or
474 incubated overnight (≥ 16 h) at 4 °C and 450 rpm in a Thermomixer.
475 Next, the samples were held at the magnetic stand and washed with 190 μL of dig wash buffer (wash buffer
476 plus digitonin, freshly pre-heated, 0.005%–0.01%; tested for each batch) once, and then, washed with 100 μL
477 of the buffer again. The samples were then resuspended in 50 μL of dig wash buffer with PA/G-MNase (to a
478 final concentration of 700 ng/mL), and incubated at room temperature (23–26 °C) on Thermomixer for 1 h,
479 or incubated at 4 °C on Thermomixer for 3 h at 450 rpm. The samples were washed with 190 μL of dig wash
480 buffer twice and then washed again with 100 μL of the buffer at the magnetic stand. Next, 100 μL of dig
481 wash buffer was added and incubated on ice for 3 min. Targeted cleavage was performed on ice by adding 2
482 μL of 100 mM $CaCl_2$ for 30 min, and the reaction was stopped by adding 100 μL of 2×stop buffer (340 mM
483 NaCl; 20 mM EDTA, pH = 8.0; 4 mM EGTA, pH = 8.0; 50 μg/mL RNase A; 100 μg/mL glycogen; 0.005%–
484 0.01% preheated digitonin; tested for each batch), and then, the mixture was fully vortexed. The samples
485 were incubated at 37 °C for 30 min for fragment release. The total samples or supernatants were digested by
486 adding 10 ng of carrier RNA (Qiagen, 1068337), 1 μL of 20% SDS, and 2 μL of Proteinase K (Roche). The
487 reaction was performed at 56 °C for 45 min and 72 °C for 20 min in a thermal cycler. DNA was purified
488 using phenol-chloroform, followed by ethanol purification. Then, 200 μL of phenol-chloroform was added,
489 mixed, and centrifuged at 20 °C and 16,000 *g* for 5 min. Next, 200 μL of chloroform was added to each
490 sample. After mixing, the samples were centrifuged at 20 °C and 16,000 *g* for 5 min and transferred to Lo-
491 binding 1.5 mL tubes. Each sample was mixed with 2 μL of glycogen, 20 μL of NaOAC, and 500 μL of
492 100% ethanol (precooled at –20 °C). The mixture was incubated at –20 °C overnight.
493 The samples were centrifuged at 16,000 *g* and 4 °C for 20 min, and then, the supernatant was removed. Next,

494 1 mL of anhydrous ethanol was added to the samples, which were then centrifuged at 4 °C and 16,000 *g* for
495 10 min. The supernatant was removed by opening the lid of the tubes and air-drying. After the liquid
496 disappeared, purified DNA was used for Tru-Seq library construction using NEB Next Ultra II DNA Library
497 Prep Kit for Illumina (NEB, E7645S). For the NEB kit, the DNA was resuspended in 25 μL of ddH_2O,
498 followed by end-repair/A-tailing with 3.5 μL of End Prep buffer and 1.5 μL of an enzyme mix. The reaction
499 was conducted at 20 °C for 30 min and 65 °C for 30 min in a PCR machine. Then, 1.25 μL of diluted
500 adaptors (NEB Multiplex Oligos for Illumina, E7335S) were added and thoroughly vortexed. Next, 15 μL of
501 Ligation master mix and 0.5 μL of Ligation enhancer were added, and the mixture was incubated at 16–20 °C
502 for 16–20 min, or at 4 °C overnight. After the ligation reaction was over, the samples were treated with 1.5
503 μL of USERTM enzyme. The reaction was performed at 37 °C for 15 min in a PCR machine. The samples
504 were purified using 1×AMPure beads. The PCR was performed by adding 25 μL of 2×KAPA HiFi Hot Start
505 Ready Mix (KAPA biosystems, KM2602) with 2.5 μL of primers of the NEB Oligos kit and 2.5 μL of
506 Illumina universal primers. The conditions used were as follows: 98 °C for 45s, (98 °C for 15s and 60 °C for
507 10s) with 16–17 cycles, and 72 °C for 1 min. The final libraries were purified using 0.9×AMPure beads and
508 used for next-generation sequencing.
509
510
511 Data analysis
512 CUT&RUN data processing
513 The single-end CUT&RUN data were mapped to the mm9 genome using Bowtie2 (Version 2.2.9)[78] with the
514 parameters –t –q –N 1 –L 25. The paired-end CUT&RUN data were mapped based on the parameters: -t –q –
515 N 1 –L 25 –X 1000 --nomixed --no-discordant. All unmapped reads, non-uniquely mapped reads, and PCR
516 duplicates were removed. For downstream analysis, the read counts were normalized by calculating the
517 number of reads per kilobase of bin per million reads sequenced (RPKM) for 100-bp bins of the genome. The
518 RPKM values were further Z-score normalized for minimizing the batch effect.
519 ATAC-seq data processing
520 The ATAC-seq reads were mapped to the mm9 genome using Bowtie2 (version 2.2.2)[78] with previously
521 described parameters. All unmapped reads, non-uniquely mapped reads, and PCR duplicates were removed.
522 For downstream analysis, the read counts were normalized by calculating the RPKM for 100 bp-bins of the

genome. The RPKM values were further normalized by Z-score normalization to minimize the batch effect.

DNA methylation data processing

We used the Bismark software for mapping DNA methylation reads to the mm9 genome. PCR duplicates were removed. For each CpG site, the methylation level was computed as the total methylated counts divided by the total counts across all reads covering the CpG site.

Analysis of the promoter H3K4me2 and H3K4me3 enrichment

Annotated promoters (TSS±2.5 kb) of the mm9 refFlat annotation database from the UCSC genome browser were used. Whole-genome Z-score normalized RPKM scores were used for assessing the data.

Identification of stage-specific distal ATAC-seq peaks

Stage-specific distal ATAC-seq peaks were identified as described previously. Briefly, distal peaks from human embryos and ES cells were merged, and then, the average Z-score was normalized. The RPKM scores were calculated based on them. A Shannon entropy-based method [3] was used to identify stage-specific peaks. Peaks with entropy less than 2 were selected as candidates and were further screened using the following criteria: the distal peak had high enrichment at this stage (normalized RPKM > 1) and positive enrichment (normalized RPKM > 0) at no more than two additional stages.

Minor/major ZGA genes and stage-specific genes were screened using the RNA-seq data from mouse germ cells and embryos. ZGA genes were defined as those unexpressed or underexpressed in oocytes (FPKM < 5) but highly upregulated (FPKM > 5, at least three-fold upregulation) in either the one-cell or early two-cell embryo stage (minor ZGA genes) or the late two-cell embryo stage (major ZGA genes). The genes that were highly expressed in oocytes and considerably upregulated (over five-fold upregulation) during the late two-cell embryo stage were also defined as major ZGA genes.

For stage-specific genes in eight-cell embryos and ICMs, a Shannon entropy-based method [3] was used to identify stage specific-genes, as previously described. Maternally expressed genes, defined as those expressed in GV or MII oocytes (FPKM ≥ 1), were excluded from this analysis to avoid the confounding effects of lingering maternal RNAs.

Clustering analysis

We conducted k-means clustering of gene expression levels at various stages using Cluster 3.0 with the parameters –g 7 (Euclidean distance). Heat maps were generated using Java Treeview 3.0. We conducted hierarchical clustering analysis for H3K4me3 at various stages using the R package “ape” based on Pearson’s

correlation between each pair of stages. The distance between the stages was calculated as 1 − Pearson's correlation.

Identification of oocyte PMDs

We calculated the average DNA methylation levels for 10-kb bins of the genome and counted the number of CpG dyads in each 10-kb bin. Bins with an average DNA methylation level below 0.6 and above 20 CpG dyads were selected and merged into PMDs using BEDtools (Version v2.25.0)[79]. The promoter regions (±2.5 kb) were removed from PMDs.

Identification of H3K4me2 peaks

The H3K4me2 peaks were called using MACS2 (Version 2.0.9)[80] with the parameters --broad --nomodel --nolambda -g mm 1.87e9/-g hs 2.7e9. Noisy peaks with weak signals (summed RPKM < 0) were removed before further analyses were conducted.

Statistical analysis

All statistical analyses in this study were briefly described in the text and performed using R (v3.6.1) (http://CRAN.R-project.org/).

DATA AVAILABILITY

All raw sequencing data generated during this work have been submitted to the Genome Sequence Archive in National Genomics Data Center (China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences). The GSA accession ID of this paper is CRA004978.

FUNDING AND ACKNOWLEDGEMENT

This work was supported by the National Key R&D Program of China (2019YFA0802200 and 2019YFA0110900 to Jiawei Xu), the National Natural Science Foundation of China 32170819 to Jiawei Xu, Henan University science and technology innovation team to Jiawei Xu), Scientific Research Innovation team of the First Affiliated Hospital of Zhengzhou University Excellent innovation team to Jiawei Xu, Henan Zhongyuan Medical Science and Technology Innovation Development Foundation to Jiawei Xu.

Author contributions

Conceptualization: JWX,HC

Methodology: CW， YQW, YL

Experiments performing: CW, YL, YQW, KYH, XRM, WBL

Data NGS (next generation sequencing): CW,YL,YQW

Bioinformatics analysis: YS, JWX, CW, YL, JG

Supervision: JWX

Writing—original draft: JWX, CW, YS， YQW, YL

Writing—review & editing: JWX

Competing interests

The authors declare no competing interests.

REFERENCES

1 Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089–1093, doi:10.1126/science.1063443 (2001).

2 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, doi:10.1038/nrm3885 (2014).

3 Wu, J. *et al.* Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature* **557**, 256–260, doi:10.1038/s41586–018–0080–8 (2018).

4 Stillman, B. Histone Modifications: Insights into Their Influence on Gene Expression. *Cell* **175**, 6–9, doi:10.1016/j.cell.2018.08.032 (2018).

5 Huang, H., Sabari, B. R., Garcia, B. A., Allis, C. D. & Zhao, Y. SnapShot: histone modifications. *Cell* **159**, 458–458.e451, doi:10.1016/j.cell.2014.09.037 (2014).

6 Peterson, C. L. & Laniel, M. A. Histones and histone modifications. *Curr Biol* **14**, R546–551, doi:10.1016/j.cub.2004.07.007 (2004).

7 Hammoud, S. *et al.* Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460**, 473–478, doi:10.1038/nature08162 (2009).

8 Hammoud, S. *et al.* Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. *Cell stem cell* **15**, 239–253, doi:10.1016/j.stem.2014.04.006 (2014).

9 Xia, W. & Xie, W. Rebooting the Epigenomes during Mammalian Early Embryogenesis. *Stem Cell Reports* **15**, 1158–1175, doi:10.1016/j.stemcr.2020.09.005 (2020).

609 10 Goldberg, A. *et al.* Distinct factors control histone variant H3.3 localization at specific genomic regions.
610 *Cell* **140**, 678–691, doi:10.1016/j.cell.2010.01.003 (2010).
611 11 Jin, C. *et al.* H3.3/H2A.Z double variant-containing nucleosomes mark' nucleosome-free regions' of
612 active promoters and other regulatory regions. *Nature genetics* **41**, 941–945, doi:10.1038/ng.409 (2009).
613 12 Mito, Y., Henikoff, J. & Henikoff, S. Genome-scale profiling of histone H3.3 replacement patterns. *Nature*
614 *genetics* **37**, 1090–1097, doi:10.1038/ng1637 (2005).
615 13 Ishiuchi, T. *et al.* Reprogramming of the histone H3.3 landscape in the early mouse embryo. *Nature*
616 *structural & molecular biology* **28**, 38–49, doi:10.1038/s41594–020–00521–1 (2021).
617 14 Zheng, H. *et al.* Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals.
618 *Mol Cell* **63**, 1066–1079, doi:10.1016/j.molcel.2016.08.032 (2016).
619 15 Zentner, G. E. & Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nat Struct*
620 *Mol Biol* **20**, 259–266, doi:10.1038/nsmb.2470 (2013).
621 16 Soares, L. M. *et al.* Determinants of Histone H3K4 Methylation Patterns. *Mol Cell* **68**, 773–785 e776,
622 doi:10.1016/j.molcel.2017.10.013 (2017).
623 17 Klose, R. J. & Zhang, Y. Regulation of histone methylation by demethylimination and demethylation. *Nat*
624 *Rev Mol Cell Biol* **8**, 307–318, doi:10.1038/nrm2143 (2007).
625 18 Eissenberg, J. C. & Shilatifard, A. Histone H3 lysine 4 (H3K4) methylation in development and
626 differentiation. *Dev Biol* **339**, 240–249, doi:10.1016/j.ydbio.2009.08.017 (2010).
627 19 Ruthenburg, A. J., Allis, C. D. & Wysocka, J. Methylation of lysine 4 on histone H3: intricacy of writing
628 and reading a single epigenetic mark. *Mol Cell* **25**, 15–30, doi:10.1016/j.molcel.2006.12.014 (2007).
629 20 Rickels, R. *et al.* Histone H3K4 monomethylation catalyzed by Trr and mammalian COMPASS-like
630 proteins at enhancers is dispensable for development and viability. *Nat Genet* **49**, 1647–1653, doi:10.1038/ng.3965
631 (2017).
632 21 Smith, E. & Shilatifard, A. Enhancer biology and enhanceropathies. *Nat Struct Mol Biol* **21**, 210–219,
633 doi:10.1038/nsmb.2784 (2014).
634 22 Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts
635 developmental state. *Proc Natl Acad Sci U S A* **107**, 21931–21936, doi:10.1073/pnas.1016071107 (2010).
636 23 Hu, D. *et al.* The MLL3/mLL4 branches of the COMPASS family function as major histone H3K4
637 monomethylases at enhancers. *Mol Cell Biol* **33**, 4745–4754, doi:10.1128/mcb.01181–13 (2013).

638 24 Liu, T. *et al.* Broad chromosomal domains of histone modification patterns in C. elegans. *Genome Res* **21**,
639 227–236, doi:10.1101/gr.115519.110 (2011).
640 25 Du, Z. *et al.* Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and
641 H3K27ac in Oryza sativa L. Japonica. *Mol Plant* **6**, 1463–1472, doi:10.1093/mp/sst018 (2013).
642 26 van Dijk, K. *et al.* Dynamic changes in genome-wide histone H3 lysine 4 methylation patterns in response
643 to dehydration stress in Arabidopsis thaliana. *BMC Plant Biol* **10**, 238, doi:10.1186/1471–2229–10–238 (2010).
644 27 Zheng, Y. C. *et al.* A Systematic Review of Histone Lysine-Specific Demethylase 1 and Its Inhibitors.
645 *Med Res Rev* **35**, 1032–1071, doi:10.1002/med.21350 (2015).
646 28 Yang, Z. *et al.* AOF1 is a histone H3K4 demethylase possessing demethylase activity-independent
647 repression function. *Cell Res* **20**, 276–287, doi:10.1038/cr.2010.12 (2010).
648 29 Fang, R. *et al.* Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating intragenic
649 H3K4me2 methylation. *Mol Cell* **39**, 222–233, doi:10.1016/j.molcel.2010.07.008 (2010).
650 30 Wang, S., Meyer, D. H. & Schumacher, B. H3K4me2 regulates the recovery of protein biosynthesis and
651 homeostasis following DNA damage. *Nat Struct Mol Biol* **27**, 1165–1177, doi:10.1038/s41594–020–00513–1
652 (2020).
653 31 Liu, Y. *et al.* H3K4me2 functions as a repressive epigenetic mark in plants. *Epigenetics & chromatin* **12**,
654 40, doi:10.1186/s13072–019–0285–6 (2019).
655 32 Zhang, C. *et al.* Narrow H3K4me2 is required for chicken PGC formation. *J Cell Physiol* **236**, 1391–1400,
656 doi:10.1002/jcp.29945 (2021).
657 33 Zhang, J., Parvin, J. & Huang, K. Redistribution of H3K4me2 on neural tissue specific genes during
658 mouse brain development. *BMC Genomics* **13 Suppl 8**, S5, doi:10.1186/1471–2164–13-S8-S5 (2012).
659 34 Wang, Y., Li, X. & Hu, H. H3K4me2 reliably defines transcription factor binding regions in different cells.
660 *Genomics* **103**, 222–228, doi:10.1016/j.ygeno.2014.02.002 (2014).
661 35 Pekowska, A., Benoukraf, T., Ferrier, P. & Spicuglia, S. A unique H3K4me2 profile marks tissue-specific
662 gene regulation. *Genome Res* **20**, 1493–1502, doi:10.1101/gr.109389.110 (2010).
663 36 Lambrot, R., Siklenka, K., Lafleur, C. & Kimmins, S. The genomic distribution of histone H3K4me2 in
664 spermatogonia is highly conserved in spermdagger. *Biol Reprod* **100**, 1661–1672, doi:10.1093/biolre/ioz055
665 (2019).
666 37 Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency

667 for low cell numbers. *Nat Protoc* **13**, 1006–1019, doi:10.1038/nprot.2018.015 (2018).

668 38 Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA
669 binding sites. *Elife* **6**, doi:10.7554/eLife.21856 (2017).

670 39 Wang, L. *et al.* Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**, 979–
671 991, doi:10.1016/j.cell.2014.04.017 (2014).

672 40 Zhang, B. *et al.* Allelic reprogramming of the histone modification H3K4me3 in early mammalian
673 development. *Nature* **537**, 553–557, doi:10.1038/nature19361 (2016).

674 41 Wu, J. *et al.* The landscape of accessible chromatin in mammalian pre-implantation embryos. *Nature* **534**,
675 652–657, doi:10.1038/nature18606 (2016).

676 42 Siklenka, K. *et al.* Disruption of histone methylation in developing sperm impairs offspring health
677 transgenerationally. *Science (New York, N.Y.)* **350**, aab2006, doi:10.1126/science.aab2006 (2015).

678 43 Hanna, C. W. *et al.* MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. *Nature*
679 *structural & molecular biology* **25**, 73–82, doi:10.1038/s41594–017–0013–5 (2018).

680 44 Liu, X. *et al.* Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation
681 embryos. *Nature* **537**, 558–562, doi:10.1038/nature19362 (2016).

682 45 Li, L., Lu, X. & Dean, J. The maternal to zygotic transition in mammals. *Molecular aspects of medicine*
683 **34**, 919–938, doi:10.1016/j.mam.2013.01.003 (2013).

684 46 Lee, M. T., Bonneau, A. R. & Giraldez, A. J. Zygotic genome activation during the maternal-to-zygotic
685 transition. *Annual review of cell and developmental biology* **30**, 581–613, doi:10.1146/annurev-cellbio-100913–
686 013027 (2014).

687 47 Liu, Y.-W. *et al.* LincRNAFEZF1-AS1 represses p21 expression to promote gastric cancer proliferation
688 through LSD1-Mediated H3K4me2 demethylation. *Molecular cancer* **16**, 39, doi:10.1186/s12943–017–0588–9
689 (2017).

690 48 Dahl, J. A. *et al.* Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic
691 transition. *Nature* **537**, 548–552, doi:10.1038/nature19360 (2016).

692 49 Xu, Q. *et al.* SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development.
693 *Nat Genet* **51**, 844–856, doi:10.1038/s41588–019–0398–7 (2019).

694 50 Xia, W. *et al.* Resetting histone modifications during human parental-to-zygotic transition. *Science* **365**,
695 353–360, doi:10.1126/science.aaw5118 (2019).

696 51 Kaniskan, H. U., Martini, M. L. & Jin, J. Inhibitors of Protein Methyltransferases and Demethylases.
697 *Chem Rev* **118**, 989–1068, doi:10.1021/acs.chemrev.6b00801 (2018).
698 52 Ciccone, D. N. *et al.* KDM1B is a histone H3K4 demethylase required to establish maternal genomic
699 imprints. *Nature* **461**, 415–418, doi:10.1038/nature08315 (2009).
700 53 Dai, X. J. *et al.* Tranylcypromine Based Lysine-Specific Demethylase 1 Inhibitor: Summary and
701 Perspective. *J Med Chem* **63**, 14197–14215, doi:10.1021/acs.jmedchem.0c00919 (2020).
702 54 Jukam, D., Shariati, S. A. M. & Skotheim, J. M. Zygotic Genome Activation in Vertebrates. *Dev Cell* **42**,
703 316–332, doi:10.1016/j.devcel.2017.07.026 (2017).
704 55 Wang, Y., Li, X. & Hu, H. H3K4me2 reliably defines transcription factor binding regions in different cells.
705 *Genomics* **103**, 222–228, doi:10.1016/j.ygeno.2014.02.002 (2014).
706 56 Dominguez, M. H., Ayoub, A. E. & Rakic, P. POU-III transcription factors (Brn1, Brn2, and Oct6)
707 influence neurogenesis, molecular identity, and migratory destination of upper-layer cells of the cerebral cortex.
708 *Cerebral cortex (New York, N.Y. : 1991)* **23**, 2632–2643, doi:10.1093/cercor/bhs252 (2013).
709 57 Sarkar, A. & Hochedlinger, K. The sox family of transcription factors: versatile regulators of stem and
710 progenitor cell fate. *Cell stem cell* **12**, 15–30, doi:10.1016/j.stem.2012.12.007 (2013).
711 58 Field, A. & Adelman, K. Evaluating Enhancer Function and Transcription. *Annual review of biochemistry*
712 **89**, 213–234, doi:10.1146/annurev-biochem-011420–095916 (2020).
713 59 Xavier, M. J., Roman, S. D., Aitken, R. J. & Nixon, B. Transgenerational inheritance: how impacts to the
714 epigenetic and genetic information of parents affect offspring health. *Hum Reprod Update* **25**, 518–540,
715 doi:10.1093/humupd/dmz017 (2019).
716 60 Chen, Q., Yan, W. & Duan, E. Epigenetic inheritance of acquired traits through sperm RNAs and sperm
717 RNA modifications. *Nat Rev Genet* **17**, 733–743, doi:10.1038/nrg.2016.106 (2016).
718 61 Zhang, B. *et al.* Widespread Enhancer Dememorization and Promoter Priming during Parental-to-Zygotic
719 Transition. *Molecular cell* **72**, doi:10.1016/j.molcel.2018.10.017 (2018).
720 62 Chen, Z., Djekidel, M. & Zhang, Y. J. N. g. Distinct dynamics and functions of H2AK119ub1 and
721 H3K27me3 in mouse pre-implantation embryos. **53**, 551–563, doi:10.1038/s41588–021–00821–2 (2021).
722 63 Xia, W. *et al.* Resetting histone modifications during human parental-to-zygotic transition. **365**, 353–360,
723 doi:10.1126/science.aaw5118 (2019).
724 64 Agarwal, S. *et al.* KDM1A maintains genome-wide homeostasis of transcriptional enhancers.

725 doi:10.1101/gr.234559.118 (2021).

726 65 Ye, R., Cao, C. & Xue, Y. J. E. i. b. Enhancer RNA: biogenesis, function, and regulation. **64**, 883–894,
727 doi:10.1042/ebc20200014 (2020).

728 66 Kang, Y., Kim, Y., Kang, J. & Kim, A. J. F. j. o. p. o. t. F. o. A. S. f. E. B. Histone H3K4me1 and H3K27ac
729 play roles in nucleosome eviction and eRNA transcription, respectively, at enhancers. **35**, e21781,
730 doi:10.1096/fj.202100488R (2021).

731 67 Kang, H. *et al.* Dynamic regulation of histone modifications and long-range chromosomal interactions
732 during postmitotic transcriptional reactivation. **34**, 913–930, doi:10.1101/gad.335794.119 (2020).

733 68 Ishihara, H. *et al.* Primed histone demethylation regulates shoot regenerative competency. **10**, 1786,
734 doi:10.1038/s41467–019–09386–5 (2019).

735 69 Murphy, P., Wu, S., James, C., Wike, C. & Cairns, B. J. C. Placeholder Nucleosomes Underlie Germline-
736 to-Embryo DNA Methylation Reprogramming. **172**, 993–1006.e1013, doi:10.1016/j.cell.2018.01.022 (2018).

737 70 Yan, J. *et al.* Histone H3 lysine 4 monomethylation modulates long-range chromatin interactions at
738 enhancers. **28**, 204–220, doi:10.1038/cr.2018.1 (2018).

739 71 Soldi, M. *et al.* Chromatin proteomics reveals novel combinatorial histone modification signatures that
740 mark distinct subpopulations of macrophage enhancers. **45**, 12195–12213, doi:10.1093/nar/gkx821 (2017).

741 72 Dorighi, K. *et al.* Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters
742 Independently of H3K4 Monomethylation. **66**, 568–576.e564, doi:10.1016/j.molcel.2017.04.018 (2017).

743 73 Kaikkonen, M. *et al.* Remodeling of the enhancer landscape during macrophage activation is coupled to
744 enhancer transcription. **51**, 310–325, doi:10.1016/j.molcel.2013.07.010 (2013).

745 74 Morrisey, E. E. *et al.* GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in
746 the mouse embryo. *Genes & development* **12**, 3579–3590 (1998).

747 75 Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to primed
748 pluripotency in rodents and humans. *Nature reviews. Molecular cell biology* **17**, 155–169,
749 doi:10.1038/nrm.2015.28 (2016).

750 76 Gross, D. S. & Garrard, W. T. Nuclease hypersensitive sites in chromatin. *Annual review of biochemistry*
751 **57**, 159–197 (1988).

752 77 Rivera, C. M. & Ren, B. Mapping human epigenomes. *Cell* **155**, 39–55, doi:10.1016/j.cell.2013.09.011
753 (2013).

754 78 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359,
755 doi:10.1038/nmeth.1923 (2012).
756 79 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features.
757 *Bioinformatics* **26**, 841–842, doi:10.1093/bioinformatics/btq033 (2010).
758 80 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008–
759 9–9-r137 (2008).
760
761

Figure 1. Mapping H3K4me2 in mouse gametes and pre-implantation embryos.

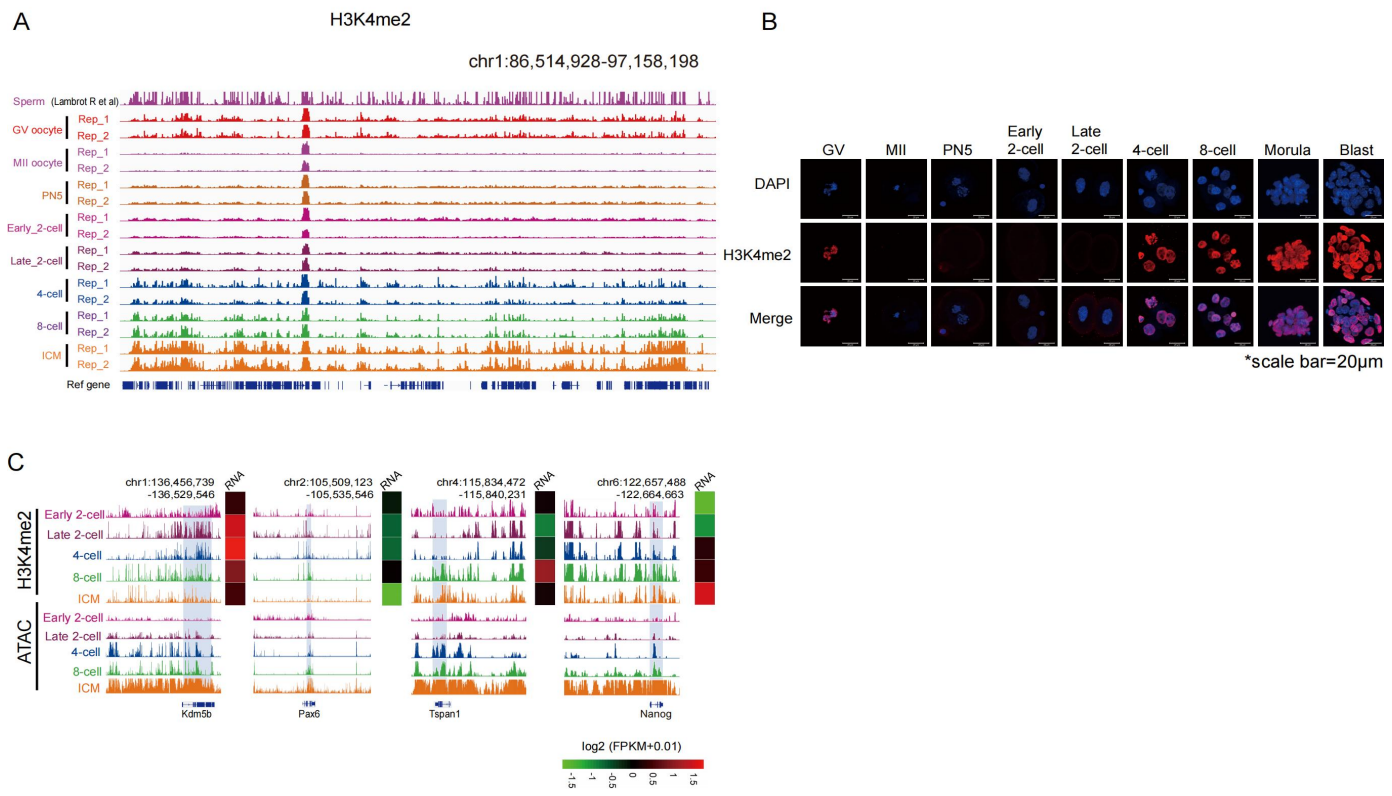


Figure 1. Mapping H3K4me2 in mouse (C57BL6/N) gametes and pre-implantation embryos.

(A) The IGV view shows the global landscape of H3K4me2 signals in mouse sperm, GV oocytes, MII oocytes, and pre-implantation embryos. The H3K4me2 data in sperm was obtained from ENCODE (Lambrot et al.). ICM was isolated from the blastocyst.

(B) Immunofluorescence of H3K4me2 in mouse (C57BL6/N) GV oocytes, MII oocytes, PN5 zygote, early two-cell embryo, late two-cell embryo, four-cell embryo, eight-cell embryo, morula, and blastocyst; the scale bar is also shown.

(C) H3K4me2 enrichment near typical genes and their expression levels are shown. H3K4me2 signal, ATAC-seq, and gene expression are shown.

Figure 2. Comparison of H3K4me2 between sperm and oocyte

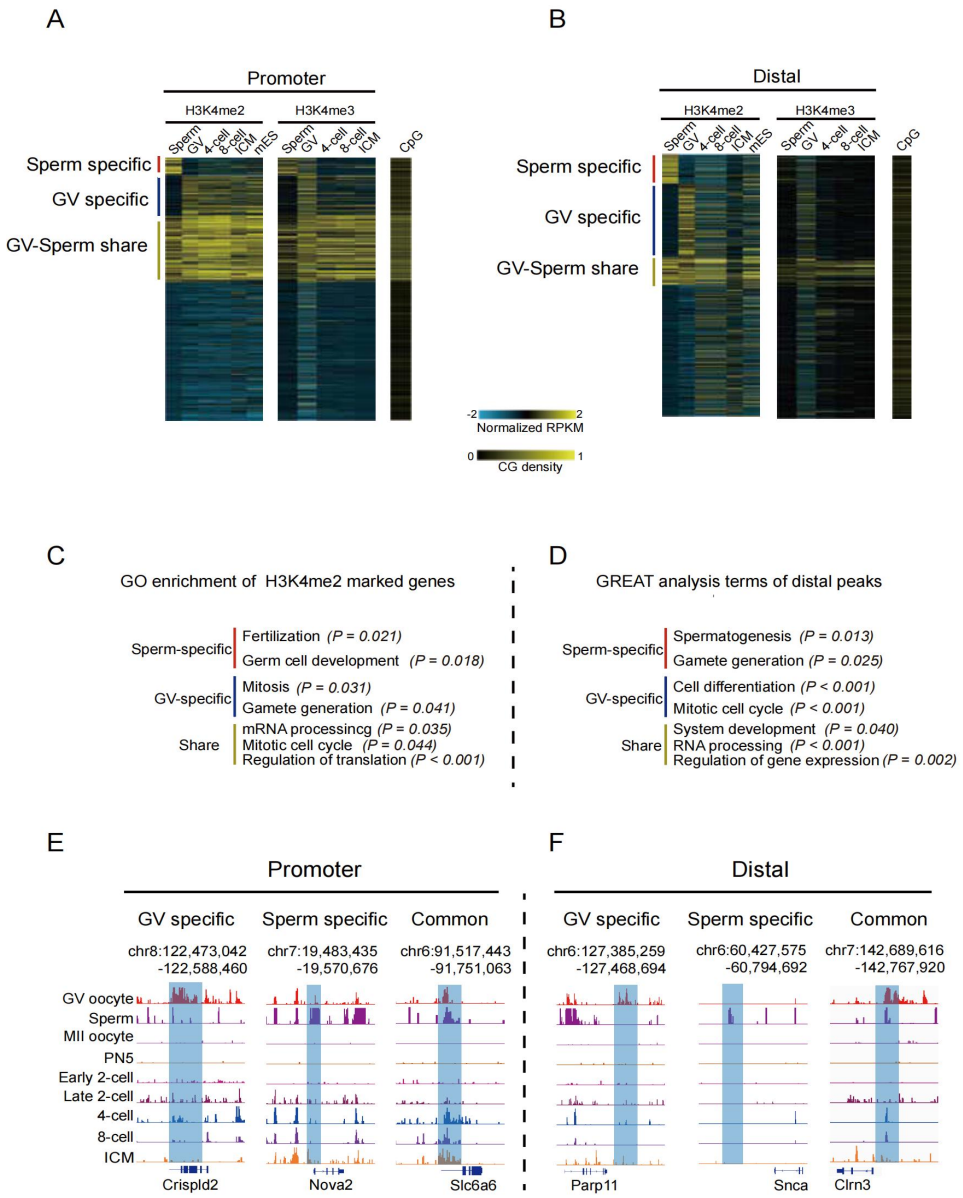


Figure 2. Comparison of H3K4me2 between sperms and oocytes.

(A-B) Heatmaps showing H3K4me2 and H3K4me3 expressed genes between sperm and GV oocyte in the promoter and distal region at different stages. CpG densities are also mapped and shown.

(C) The enriched GO terms of H3K4me2 marked genes in GV and sperm-specific or shared genes are shown along with the associated p-values.

(D) The results of the GREAT analysis for distal peaks in GV and sperm-specific or shared genes are shown

along with the associated p-values.

(E-F) The IGV views show H3K4me2 expressions in GV and sperm-specific or shared genes at the representative promoters (left, shaded) or distal accessible regions (right, shaded).

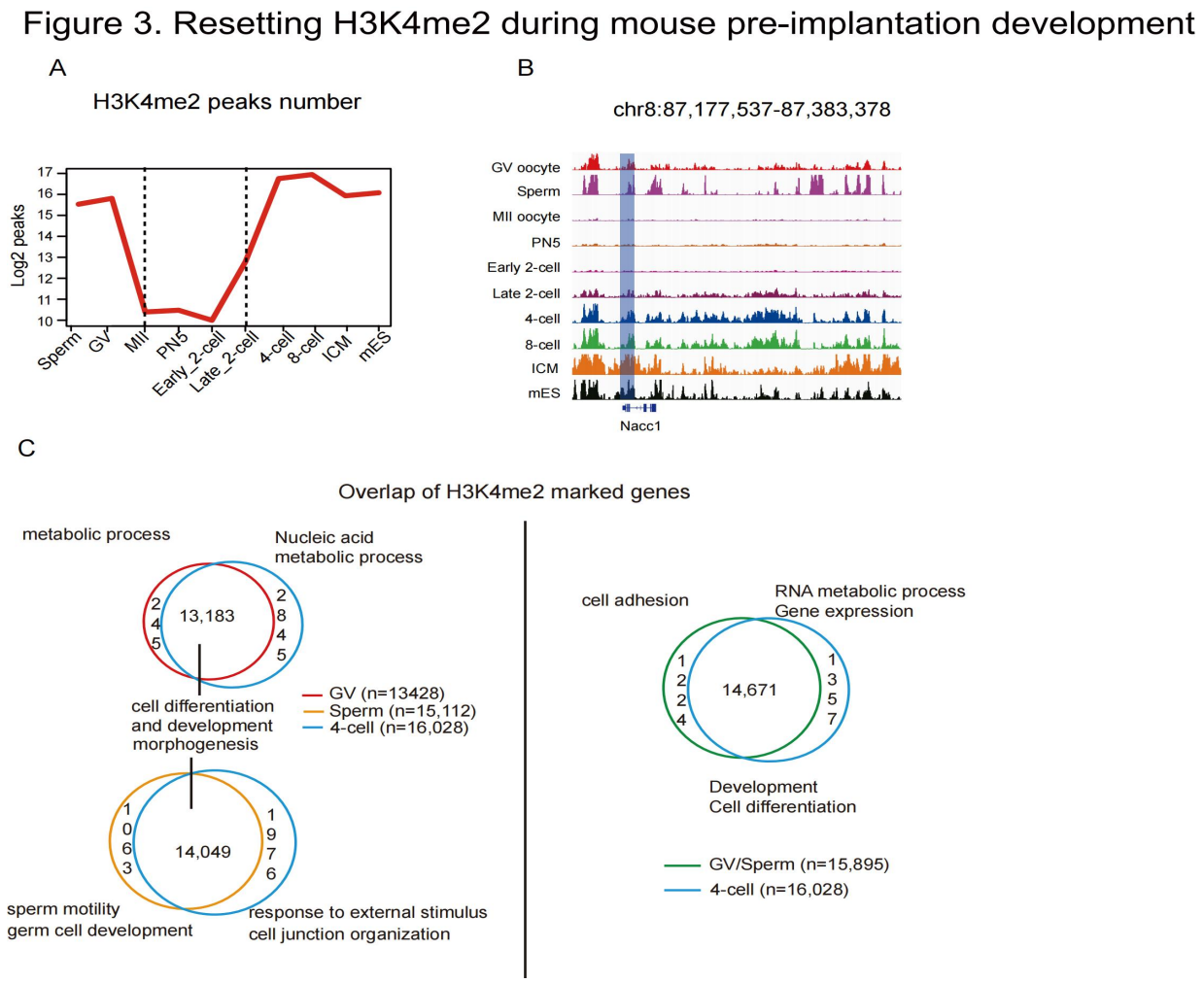


Figure 3. Resetting of H3K4me2 during mouse pre-implantation development.

(A) The number of H3K4me2 peaks in mouse embryo stem cells (mESCs), mouse gametes, and pre-implantation embryos was calculated.

(B) The IGV view shows H3K4me2 signals in mESCs, mouse gametes, and pre-implantation embryos; Ncc1 gene sites are also shown.

(C) Venn diagrams illustrate the overlap of H3K4me2 marked genes between sperm/GV oocytes, separately (left) or jointly (right), and four-cell embryos. The enriched GO terms for the shared and non-overlapping genes are also shown.

Figure 4. Inhibition of LSD2 resulted in abnormal embryonic development

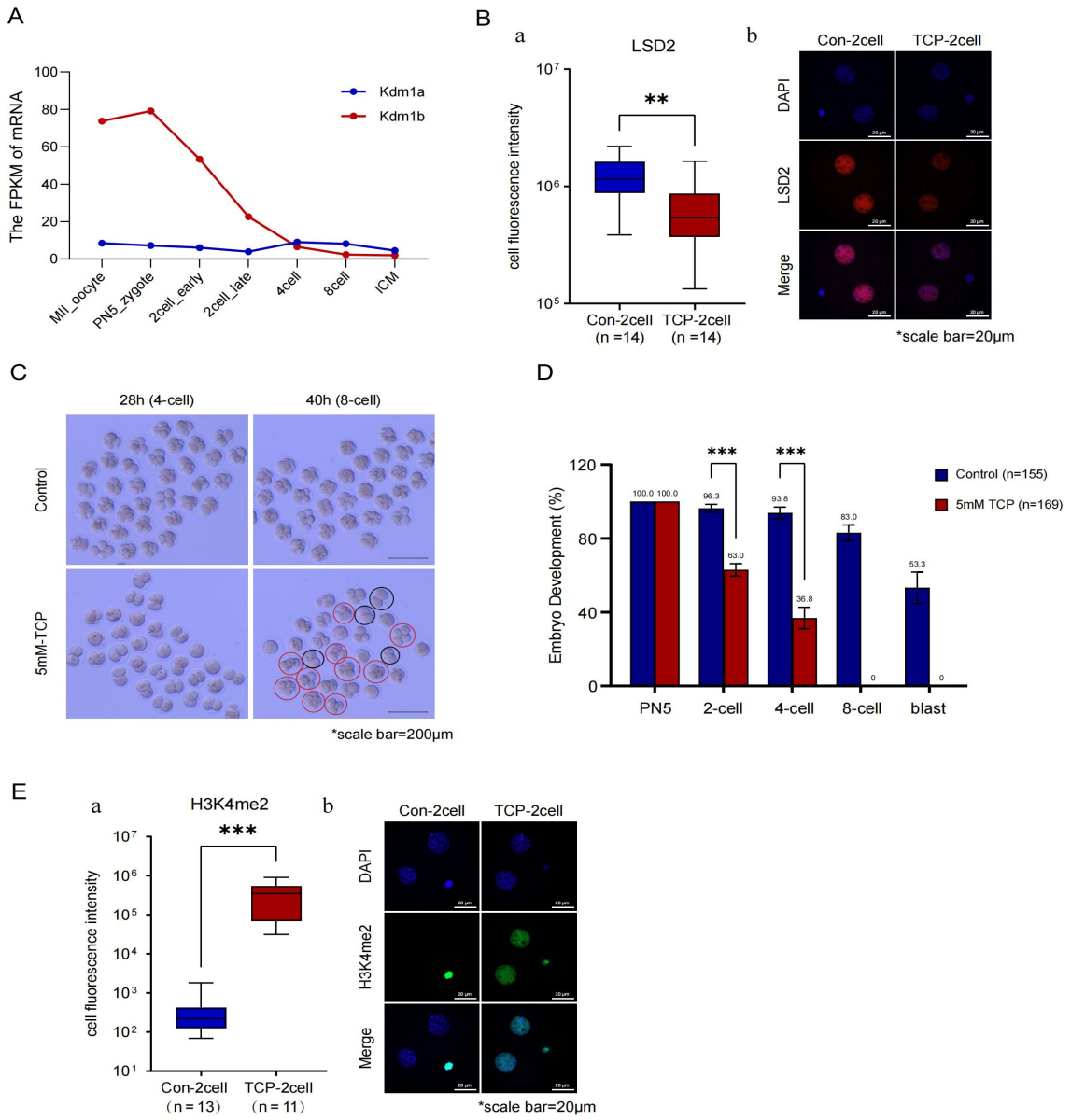


Figure 4. The inhibition of LSD2 resulted in abnormal embryonic development

(A) The dynamic changes in the Kdm1a (blue line) and Kdm1b (red line) levels in oocytes and early embryos (based on published RNA-seq data43).

(B) a.The immunofluorescence intensity of LSD2 in the experimental group (5 mM TCP) and the control group are shown using a box diagram. b. Immunofluorescence images showed the immunofluorescence intensity of LSD2 of two-cell embryos in the experimental group (5 mM TCP) (right) and the control group (left) ; the scale bar is also shown.

(C) The developmental morphology of mouse embryos in the experimental group (5 mM TCP) and the control group after 28 h and 40 h of culture are shown along with the scale bar.

(D) The figure illustrates the embryonic development of mice at different stages after culture in vitro in the

experimental group (5 mM TCP) and the control group.

(E) a.The immunofluorescence intensity of H3K4me2 in the experimental group (5 mM TCP) and the control group are shown using a box diagram. b. Immunofluorescence images showed the immunofluorescence intensity of H3K4me2 of two-cell embryos in the experimental group (5 mM TCP) (right) and the control group (left) ; the scale bar is also shown.

Figure 5. Parental H3K4me2 is cleared by maternal LSD2 during sperm-egg fusion

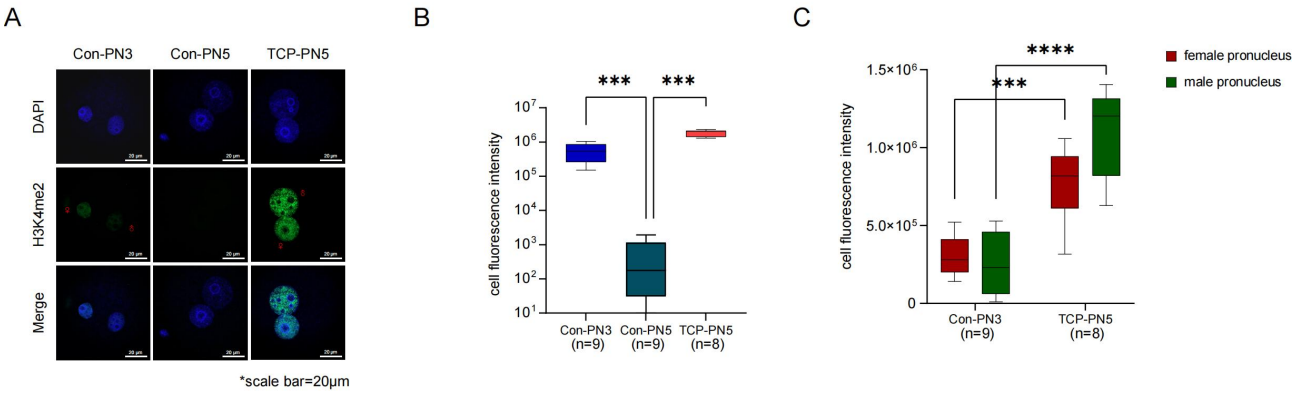


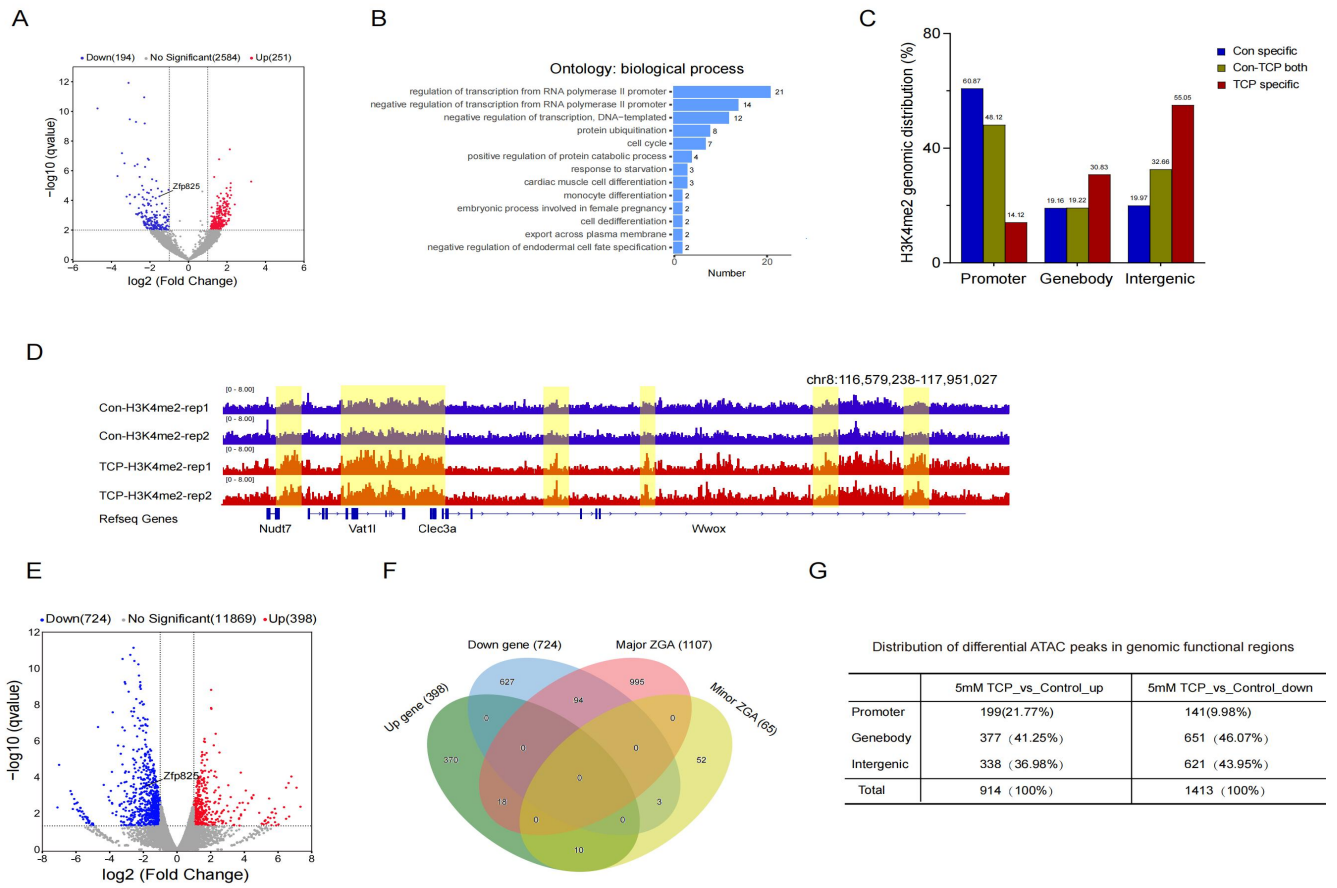
Figure 5. Parental H3K4me2 was erased by maternal LSD2 during sperm-egg fusion

(A) Immunofluorescence images showed the immunofluorescence intensity of H3K4me2 in single-cell embryos in the experimental group (5 mM TCP) and the control group (C); the scale bar is also shown.

(B) The immunofluorescence intensity of H3K4me2 in single-cell embryos in the experimental group (5 mM TCP) and the control group are represented as box and whisker plots.

(C) The immunofluorescence intensity of H3K4me2 in male and female embryonic prokaryotes of the experimental group (5 mM TCP) and the control group are represented as box and whisker plots.

Figure 6. Inhibition of LSD2 affects gene expression



	5mM TCP_vs_Control_up	5mM TCP_vs_Control_down
Promoter	199(21.77%)	141(9.98%)
Genebody	377（41.25%）	651（46.07%）
Intergenic	338（36.98%）	621（43.95%）
Total	914（100%）	1413（100%）

Figure 6. Inhibition of LSD2 affects gene expression

(A) A volcano map of genes with differential H3K4me2 expression after TCP inhibition of LSD2 expression.

(B) The GO analysis showed enrichment of biological processes after an increase in the expression of H3K4me2.

(C) The peak distribution area of H3K4me2 in the experimental group (5 mM TCP) and the control group

(D) The IGV view shows the H3K4me2 signals in typical genes in the experimental group (5 mM TCP) and the control group.

(E) A volcano map of the differential expression of H3K4me2 genes in the experimental group (5 mM TCP) and the control group.

(F) The Venn diagram illustrates the intersection of differentially expressed genes and minor ZGA and major ZGA genes between the experimental group (5 mM TCP) and the control group

(G) The distribution of differential ATAC peaks in the functional regions of the genome.

Figure 7. Pervasive H3K4me2 in CpG-rich regulatory regions in mouse embryos

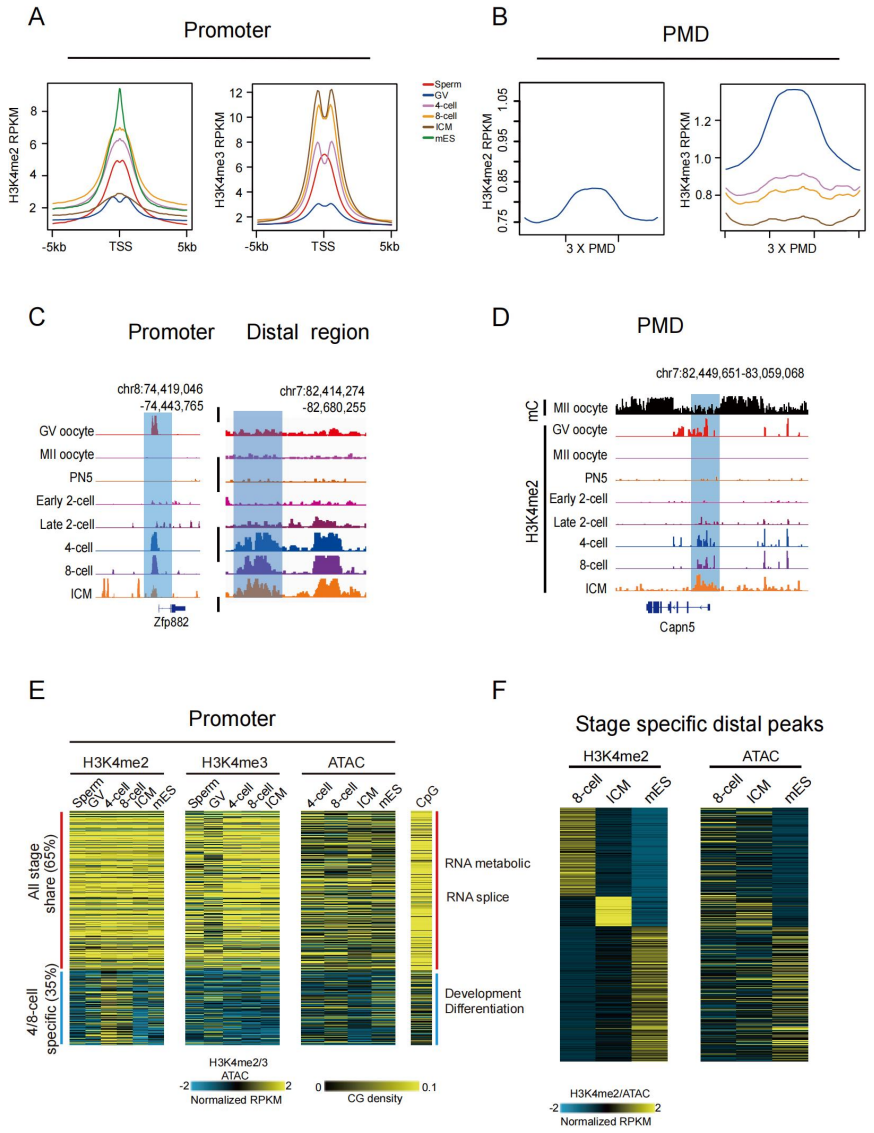


Figure 7. Pervasive H3K4me2 in CpG-rich regulatory regions in mouse embryos.

(A-B) H3K4me2 and H3K4me3 enrichment near TSS (TSS ±5 kb) and PMD (3× PMD) in mouse gametes, pre-implantation embryos, and mESC.

(C) The IGV view shows promoter H3K4me2 signals and widespread distal H3K4me2 signals in mouse oocytes and early embryos.

(D) The IGV view shows H3K4me2 signals and the distribution of DNA methylation in mouse oocytes and early embryos at the PMD (partially methylation district) region.

(E) Heat maps show the enrichment of H3K4me2 and H3K4me3 at the promoters, including that in all stages shared and four-cell/eight-cell specific. ATAC-seq, promoter CpG density, and GO analysis results are presented.

(F) Heat maps of the distal peaks of H3K4me2 expressed in specific stages among eight-cell, ICM, and mESC. ATAC-seq are shown.

Figure 8. Chromatin states at specific stages and H3K4me2 mode chart during maternal-to-zygotic transition.

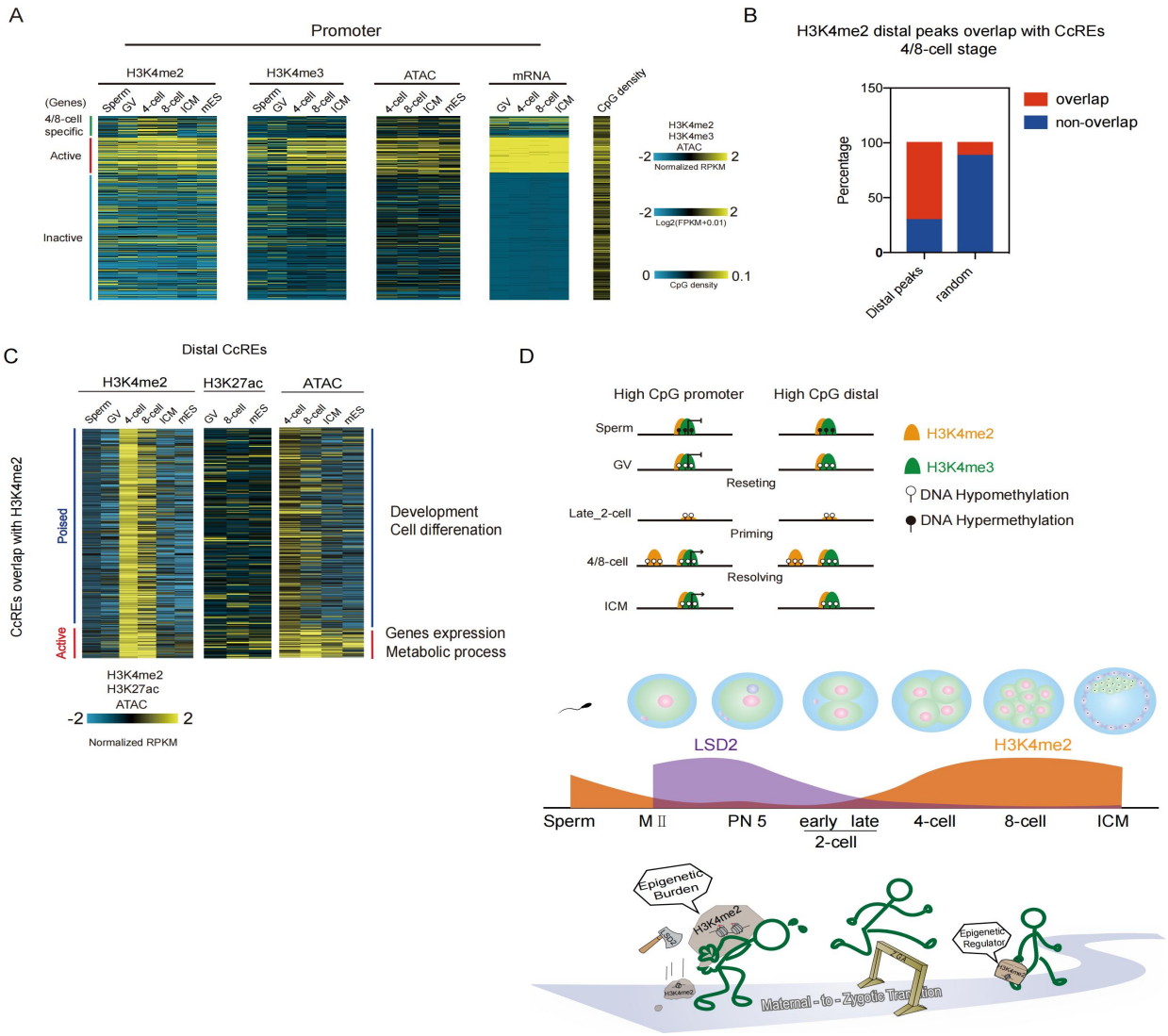


Figure 8. Chromatin states at specific stages and the H3K4me2 mode chart during mouse maternal-to-zygotic transition.

(A) Heat maps of H3K4me2, H3K4me3, and ATAC-seq are shown, “All active” and “All inactive” promoters and the corresponding gene expression are presented. The CpG densities are also shown.

(B) The bar graph shows that the H3K4me2 distal peaks overlapped with the CcREs four/eight-cell stage. Random peaks are shown as a control.

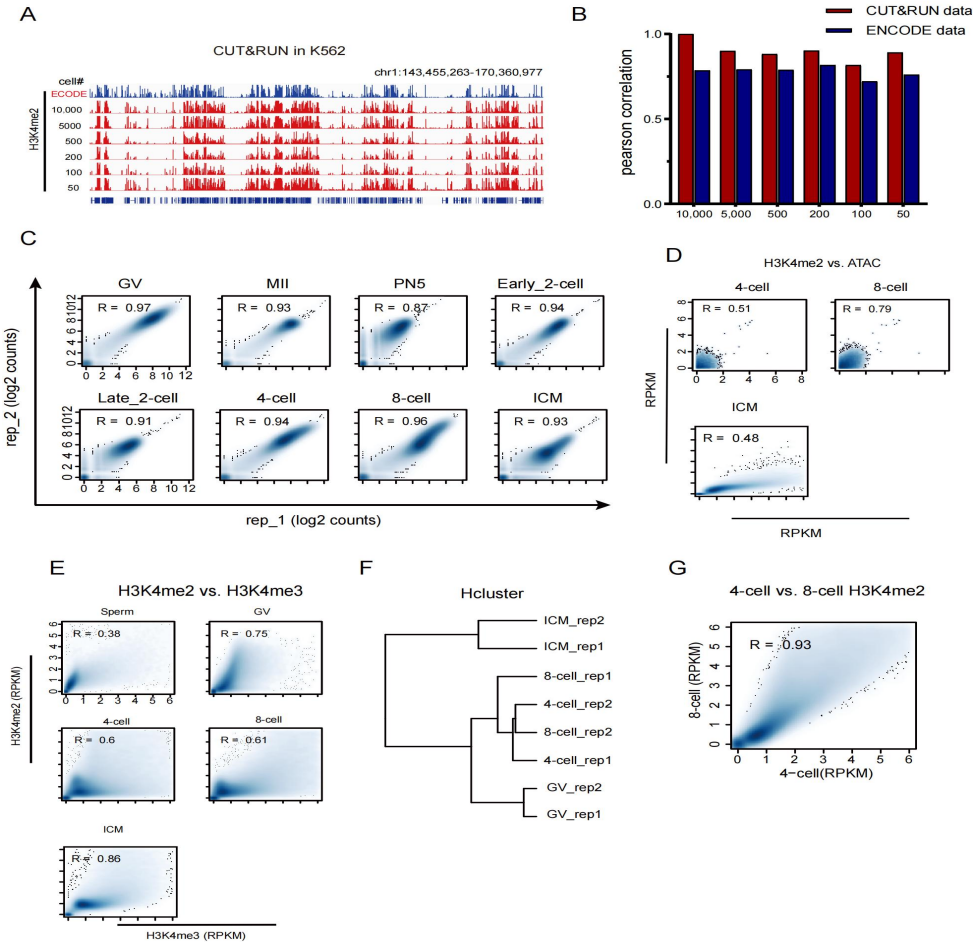
(C) Heat maps of the chromatin states of the overlap between CcREs and the distal peaks of H3K4me2 and H3K27ac are presented. Clustering was conducted using ATAC-seq, H3K4me2 in the four-cell, eight-cell, and ICM stages. The results of the GREAT analysis for each cluster are shown (right). Data on histone modification of H3K27ac in mice was obtained from the GEO database, GSE72784.

(D) Dynamic reprogramming of H3K4me2 during mouse parental-to-zygotic transition. A schematic model of

84 “H3K4me2 resetting” during parental-to-zygotic transition is shown. After fertilization, the parental
85 H3K4me2 is almost globally erased. *De novo* H3K4me2 promoters and distal regions in CpG-rich and DNA
86 hypomethylated regions are re-established at the late two-cell stage, probably as a priming chromatin state.
87 This leads to the resolution of primed promoters to “active” promoters with H3K4me3, or “poised”. A similar
88 transition occurs for putative enhancers.

1 **Extended Data**

Figure S1. Validation of CUT&RUN data



2 **Extended Data Figure 1. Validation of H3K4me2 CUT&RUN data in Human K562 cell**
3 **line, oocytes and early embryos.**

4 (A) The IGV views showing H3K4me2 distributions by CUT&RUN using various numbers
5 of Human K562 cell line. The ENCODE references are added for comparison. GEO
6 accession: GSM733651 .

7 (B) The Pearson correlation coefficients showing the comparison between CUT&RUN of
8 H3K4me2 using various numbers of Human K562 cells and conventional ChIP–seq data from
9 ENCODE.

10 (C) Scatter plots showing the correlations of biological replicates (n=2) of H3K4me2 signals
11 (2kb-bin, whole genome) by CUT&RUN for each stage of mouse oocytes and early embryos.
12 The Pearson correlation coefficients are also shown.

13 (D) Scatter plots comparing the H3K4me2 signals with ATAC-seq (Assay for

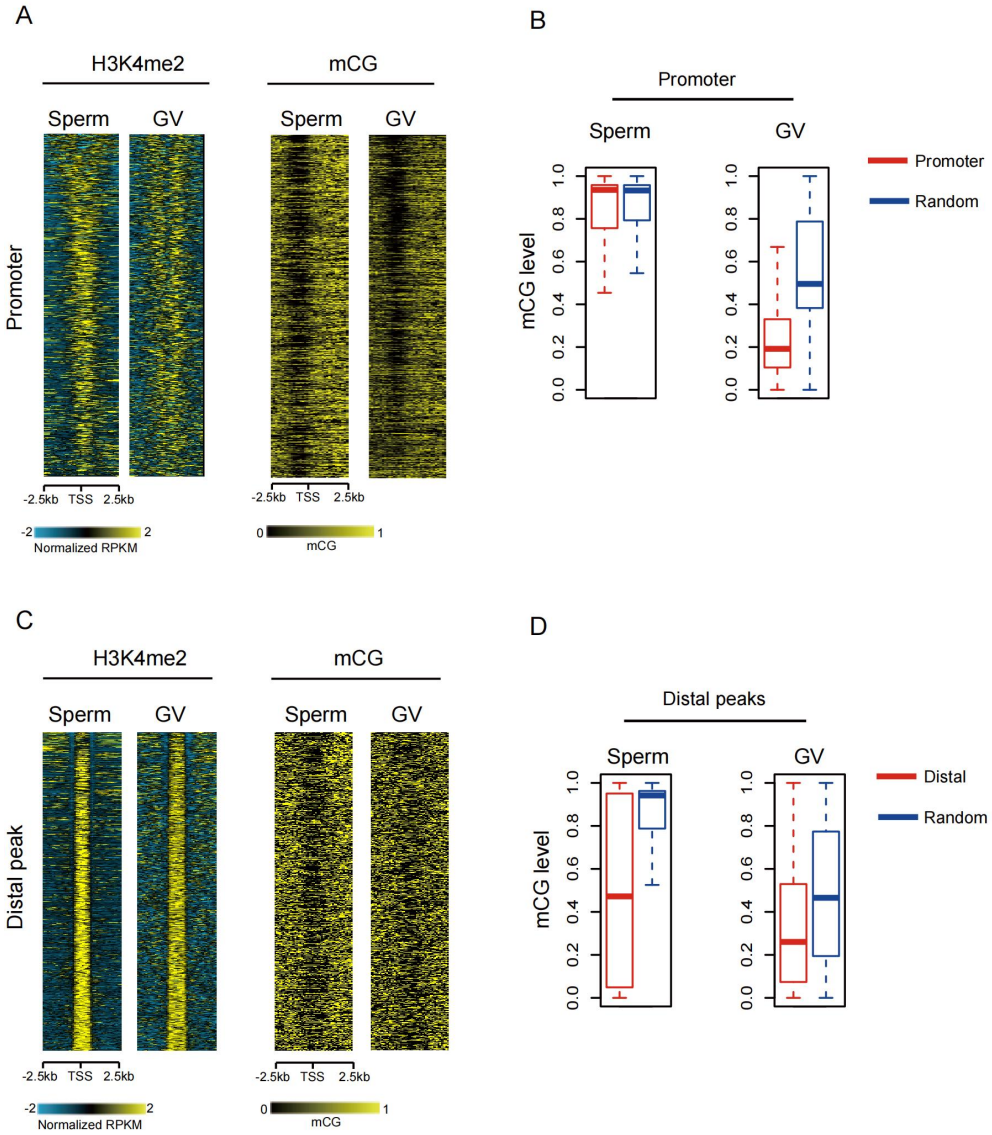
Transposase-Accessible Chromatin with high throughput sequencing) at promoters or distal accessible regions in mouse oocytes and early embryos. Spearman correlation coefficients are also shown.

(E) Scatter plots comparing the H3K4me2 signals with H3K4me3 signals in mouse sperm, oocytes and early embryos. Spearman correlation coefficients are also shown.

(F) Hierarchical clustering of H3K4me2 mouse oocytes and early embryos. Pearson correlation was used to measure distances.

(G) Scatter plots comparing the H3K4me2 signals between 4-cell and 8-cell in mouse embryos. Spearman correlation coefficients are also shown.

Figure S2. H3K4me2 in mouse gametes



Extended Data Figure 2. H3K4me2 in mouse gametes.

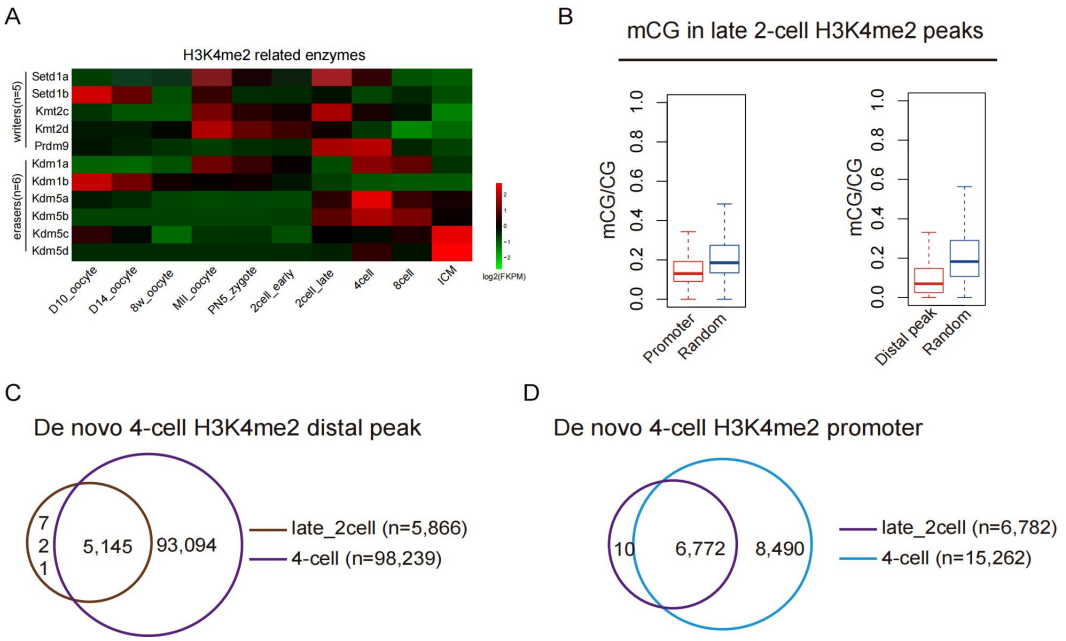
(A) Heatmaps showing H3K4me2 signals at all promoter regions (TSS ± 2.5kb) in mouse gametes. DNA methylation (TSS ± 2.5kb) is also mapped. DNA methlation data cited in GEO dataset：GSE56697.

(B) The box plots showing DNA methylation level at all promoter regions in mouse gametes. Random peaks are shown as a control. DNA methlation data cited in GEO dataset: GSE56697.

(C) Heatmaps showing H3K4me2 signals at all distal regions (peak ± 2kb) in mouse gametes. DNA methylation (peak ± 2kb) is also mapped. DNA methlation data cited in GEO dataset: GSE56697.

(D) The box plots showing DNA methylation level at all distal regions in mouse gametes. Random peaks are shown as a control. DNA methlation data cited in GEO dataset：GSE56697.

Figure S3. H3K4me2 in late 2-cell embryos



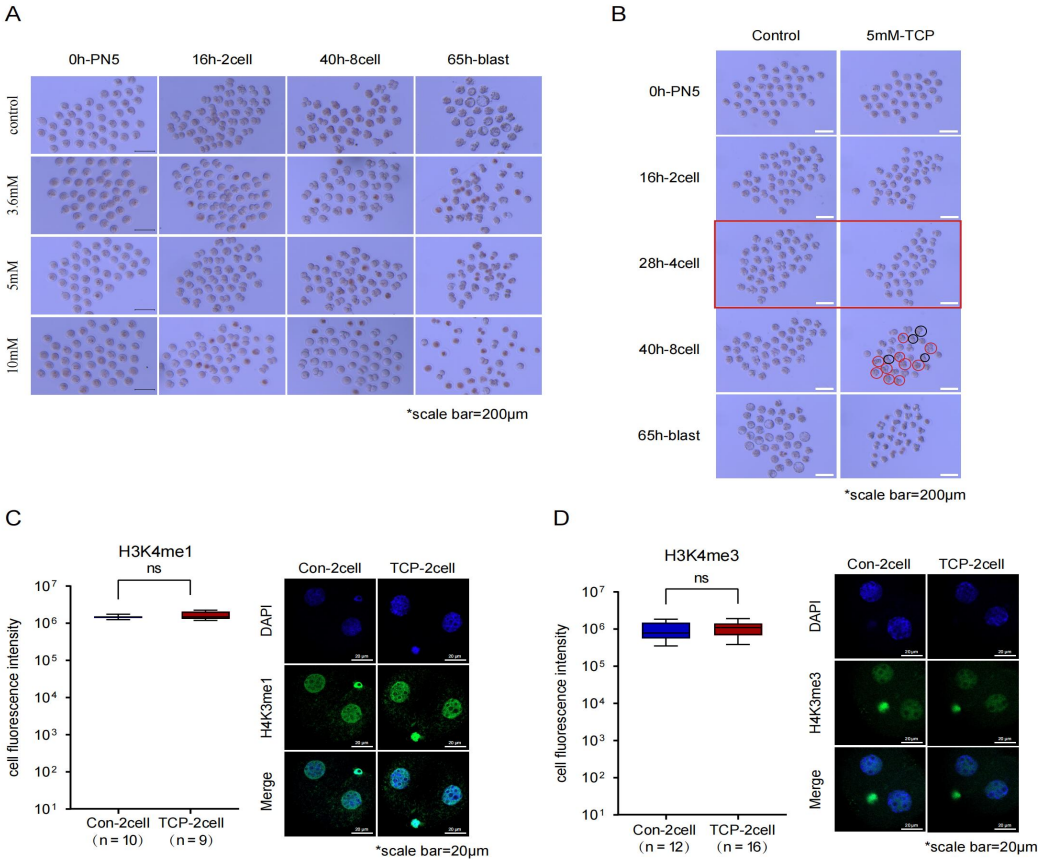
Extended Data Figure 3. H3K4me2 in late 2-cell embryos.

(A) Heatmaps showing the expression of H3K4me2 related regulators in mouse (based on published RNA-seq data43).

(B) The box plots showing DNA methylation level at all promoter regions and distal region in mouse 2-cell stage. Random peaks are shown as a control.

(C-D) Venn diagrams showing De novo 4-cell H3K4me2 promoter and distal peaks number.

Figure S4. Cell culture experiment under TCP treatment



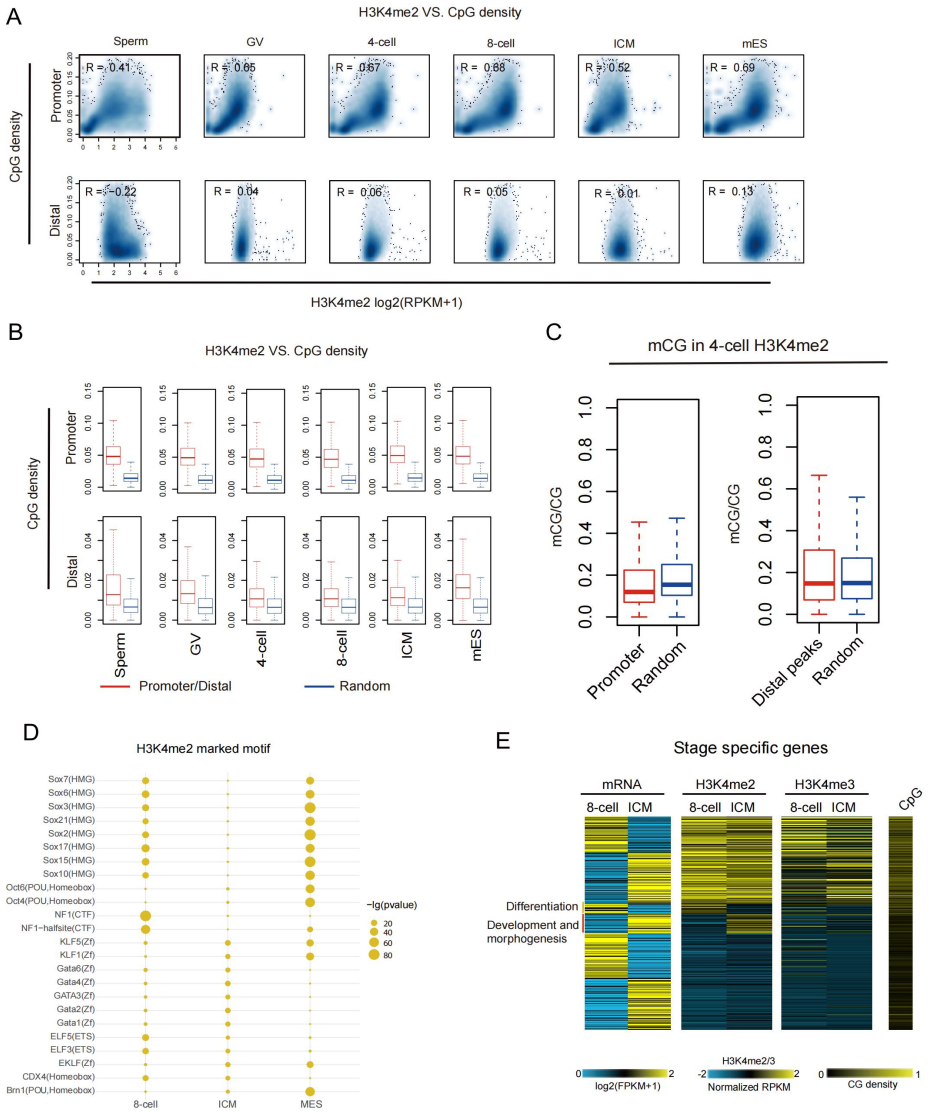
Extended Data Figure 4. Cell culture experiment under TCP treatment

(A)Morphological observation of cells under different concentration of TCP culture, scale bar is also shown.

(B)Morphological observation of embryo development after 5mM TCP inhibition of LSD2 expression.

(C-D)Immunofluorescence images showed the immunofluorescence intensity of H3K4me1 (C) and H3K4me3 (D) of 2-cell embryos in the experimental group (5mM TCP) and control group, scale bar is also shown.

Figure S5. H3K4me2 in mouse 4/8-cell embryos



Extended Data Figure 5. H3K4me2 in mouse 4/8-cell embryos.

(A) Scatter plots comparing the promoter and distal H3K4me2 signals with CpG densities in mouse gametes, 4-cell, 8-cell, ICM and mESC. The Spearman correlation coefficients are also shown.

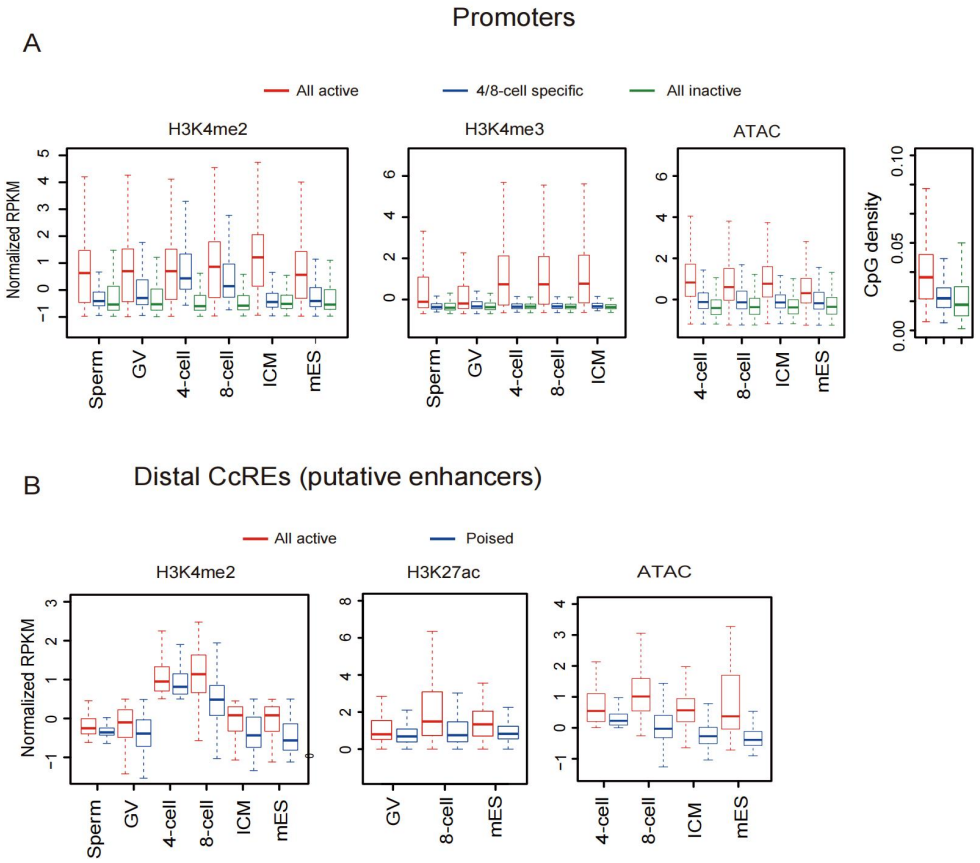
(B) The box plots showing the CpG density and promoter/distal H3K4me2 peaks. Random peaks are shown as a control.

(C) The box plot showing DNA methylation in promoter and distal 4-cell H3K4me2 peaks. Random peaks are shown as a control.

(D) TF motifs identified from active enhancers in mouse early embryos (8-cell, ICM) and mESC. showing motif enrichment p value < 1e-20 at least at one stage were included. Circle size showing TF enrichment and the expression of the TF is color coded.

(E) Heatmaps showing stage-specifically expressed genes further classified by their H3K4me2 states (high, dynamic, and low). RNA-seq, H3K4me3 state, CpG densities and GO analysis results are also mapped and shown.

Figure S6. Chromatin states at specific stages and distal CcREs.



Extended Data Figure 6. Chromatin states at specific stages and distal CcREs.

(A) Box plots showing H3K4me2, H3K4me3 and ATAC-seq signals, “All active” and “All inactive” promoters in mouse gametes, embryos and mESC. CpG density for each group is also shown (right).

(B) The box plot showing the ATAC-seq, H3K4me2 and H3K27ac signals at putative active, poised or random enhancers in mouse gametes, early embryos or mESC, Mouse histone modified H3K27ac data refer to GEO database : GSE72784.

71 **Extended Data Table**

72 The sample information in this study was listed in the table below.

73 Table1. CUT&RUN experimental sample information

Sample name	Experiment	Replicate	Sample number	Cell number
K562 cell line	H3K4me2CUT&RUN	Rep1	50	50
K562 cell line	H3K4me2CUT&RUN	Rep2	50	50
K562 cell line	H3K4me2CUT&RUN	Rep1	100	100
K562 cell line	H3K4me2CUT&RUN	Rep2	100	100
K562 cell line	H3K4me2CUT&RUN	Rep1	200	200
K562 cell line	H3K4me2CUT&RUN	Rep2	200	200
K562 cell line	H3K4me2CUT&RUN	Rep1	500	500
K562 cell line	H3K4me2CUT&RUN	Rep2	500	500
K562 cell line	H3K4me2CUT&RUN	Rep1	5000	5000
K562 cell line	H3K4me2CUT&RUN	Rep2	5000	5000
K562 cell line	H3K4me2CUT&RUN	Rep1	10000	10000
K562 cell line	H3K4me2CUT&RUN	Rep2	10000	10000
MES cell line	H3K4me2CUT&RUN	Rep1	50	50
MES cell line	H3K4me2CUT&RUN	Rep2	50	50
MES cell line	H3K4me2CUT&RUN	Rep1	100	100
MES cell line	H3K4me2CUT&RUN	Rep2	100	100

Sample name	Experiment	Replicate	Sample number	Cell number
MES cell line	H3K4me2CUT&RUN	Rep1	200	200
MES cell line	H3K4me2CUT&RUN	Rep2	200	200
MES cell line	H3K4me2CUT&RUN	Rep1	500	500
MES cell line	H3K4me2CUT&RUN	Rep2	500	500
MES cell line	H3K4me2CUT&RUN	Rep1	5000	5000
MES cell line	H3K4me2CUT&RUN	Rep2	5000	5000
MES cell line	H3K4me2CUT&RUN	Rep1	10000	10000
MES cell line	H3K4me2CUT&RUN	Rep2	10000	10000
GV oocyte	H3K4me2CUT&RUN	Rep1	190	190
GV oocyte	H3K4me2CUT&RUN	Rep2	190	190
MII oocyte	H3K4me2CUT&RUN	Rep1	165	165
MII oocyte	H3K4me2CUT&RUN	Rep2	165	165
PN5 zygote	H3K4me2CUT&RUN	Rep1	250	250
PN5 zygote	H3K4me2CUT&RUN	Rep2	250	250
Early_2cell	H3K4me2CUT&RUN	Rep1	86	172
Early_2cell	H3K4me2CUT&RUN	Rep2	86	172
Late_2cell	H3K4me2CUT&RUN	Rep1	98	196
Late_2cell	H3K4me2CUT&RUN	Rep2	98	196
4cell	H3K4me2CUT&RUN	Rep1	65	260
4cell	H3K4me2CUT&RUN	Rep2	65	260

Sample name	Experiment	Replicate	Sample number	Cell number
8cell	H3K4me2CUT&RUN	Rep1	25	200
8cell	H3K4me2CUT&RUN	Rep2	25	200
ICM	H3K4me2CUT&RUN	Rep1	14	≈420
ICM	H3K4me2CUT&RUN	Rep2	14	≈420