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

88 development of mammals. We optimized the Cleavage Under Target and Release Using Nuclease
89 (CUT&RUN)[37,38] methods using H3K4me2 antibody to as few as 50 cells and revealed unique H3K4me2
90 modification patterns and epigenetic reprogramming that occur during the parental-to-zygote
91 transition in mice. We also showed that H3K4me2 resetting depends on maternal LSD2 and parental
92 H3K4me2 acts as an "Epigenetic Burden" on mammalian ZGA.
93
94 **Results**
95 **Genome-wide profiling of H3K4me2 in mouse oocytes and early embryos**
96 The epigenome underwent significant reprogramming during early development. To determine how
97 H3K4me2 is inherited from the sperm and oocyte by the zygote during early development, we
98 investigated H3K4me2 in mouse gametes and pre-implantation embryos using CUT&RUN methods.
99 We optimized the CUT&RUN protocol (CUT&RUN; Methods) using H3K4me2 antibody. We also
100 acquired high-quality data using only 50 cells. Most peaks could be retrieved using only 50 cells
101 (Figures S1A and S1B). Then, we collected GV oocytes, MII, PN5, early two-cell embryos, late two-cell
102 embryos, four-cell embryos, eight-cell embryos, and the inner cell mass (ICM) of mice. Mouse
103 embryonic stem cells (mESC) were used as control. We profiled H3K4me2 in oocytes and early
104 embryos using CUT&RUN and immunofluorescence (IF) techniques. The results of all replicates were
105 highly consistent (Figure 1A; Figures S1A, B, and C). We identified 57,455, 25,073, 110,517, 126,318,
106 62,575, and 69,154 H3K4me2 peaks in the GV oocytes, late two-cell embryos, four-cell embryos,
107 eight-cell embryos, ICM, and mESC, respectively. Most of the H3K4me2 was erased from GV to MII
108 during oocyte development. The paternal H3K4me2 was also erased after fertilization, and it was re-
109 established in the late two-cell stage (Figures 1A-B). We also demonstrated the H3K4me2, ATAC-seq,
110 and mRNA expression of several developmental key genes (Figure 1C). We found that H3K4me2
111 positively correlated with H3K4me3 and chromatin accessibility in early embryos (Figures S1D-E). The
112 H3K4me2 genome-wide pattern of four-cell embryos showed a strong association with eight-cell
113 embryos: the H3K4me2 of four-cell and eight-cell embryos was in one cluster, and the correlation was
114 high (R = 0.93) (Figures S1F-G).
115 **Distinct H3K4me2 pattern between sperm and oocyte**
116 We systematically characterized the epigenome of mouse gametes. In the analysis, we included

117 published data on DNA methylation[39], H3K4me3[40], and ATAC-seq of early embryos[41], as well as the
118 H3K4me2 data of sperm[42]. Non-canonical patterns of H3K4me3, enriched broad domains in
119 promoters and distal regions, are present in mouse oocytes as broad domains in partially methylated
120 domains (PMDs)[43,44]. H3K4me2 shows a non-canonical pattern in the GV oocytes and sperm of mice
121 (Figures 2A-B and E-F). Mature sperm and oocytes are in a state of inactive transcription[45,46]; however,
122 strong ChIP-seq and CUT&RUN signal enrichment of H3K4me2 was found to be related to
123 transcription[47]. These findings suggested that H3K4me2 might be deposited during gametogenesis
124 and retained in mature gametes. To confirm this speculation, we performed a K-means cluster analysis
125 for H3K4me2 in gametes. By comparing promoter H3K4me2 at the sperm and GV stages, we found
126 that the promoters could be divided into four groups (Figure 2A). The promoters in the first group
127 (sperm-specific H3K4me2) were involved in fertilization and germ cell development, and these
128 markers were erased in the later stages (Figures 2A and 2C). In contrast, the promoters in the second
129 group (GV-specific H3K4me2) were associated with mitosis and gamete generation. They played a
130 role in GV generation during oocyte development and were re-established in the late stages (Figures
131 2A and 2C). The promoters in the third group (sperm-GV share H3K4me2) were related to mRNA
132 processing and translation. The signals of these regions were erased (MII, zygote, and early two-cell
133 stage) and reconstructed (at the late two-cell stage) during development (Figures 2A and 2C). The
134 promoters in the fourth group had a low H3K4me2 level at all stages (Figure 2A).
135 The H3K4me2 enrichment in these promoters was consistent with the changing dynamics of H3K4me3
136 (Figure 2A). Specifically, promoters with high CpG density were preferentially marked by H3K4me2 in
137 the sperm-GV share group (Figure 2A). Next, we investigated H3K4me2-marked distal regions in
138 gametes. Sperm-specific groups are involved in gamete generation (Figure 2D). GV-specific distal
139 H3K4me2 showed functions in cell differentiation and mitotic cell cycle (Figure 2D). Both were erased
140 after fertilization and not reconstructed during embryo development (Figure 2B). H3K4me3 signals did
141 not occupy the sperm-specific/GV-specific H3K4me2 distal regions (Figure 2B). The GV-sperm-shared
142 distal H3K4me2 regions were also marked by H3K4me3. Those genes participated in system
143 development, RNA processing, and regulation of gene expression (Figures 2B and 2D). Compared to
144 GV, sperm H3K4me2 was enriched in a fully methylated promoter, but the distal H3K4me2 of both
145 sperm and GV were enriched in the hypomethylated region (Figures S2A-D). These results indicated

146 that differential H3K4me2 modification patterns occurred between sperms and oocytes.

147

148 **Resetting H3K4me2 in pre-implanted embryos**

149 Significant epigenetic reprograming and chromosome remodeling associated with H3K4me3,
150 H3K27me3, H3K36me3, and H3K27ac occurs during early mammalian development[44,48,49]. In another
151 study, we found that epigenome reboot occurred during the early developmental stage in humans[50].
152 We investigated how parental histone mark H3K4me2 at regulatory elements are reprogrammed after
153 fertilization. We found that GV and sperm H3K4me2 peaks were almost erased in PN5 and re-
154 established along with zygotic genome activation (ZGA). This result was further confirmed by IF
155 analysis, which showed that the reconstruction of H3K4me2 might start from the late two-cell stage
156 (Figures 1B and 3B). It is worth noting that previous studies using different immunofluorescence
157 parameters reported detectable H3K4me2 signals at the zygote and two-cell stages[51]. This discrepancy
158 may arise from differences in confocal microscope settings, antibody sensitivity, or sample processing.
159 In our study, we used consistent parameters across all stages from GV to blastocyst; when we
160 specifically optimized the acquisition parameters for zygotes and two-cell embryos, weak signals
161 could be observed. The H3K4me2 demethylases KDM1B (LSD2) were also highly expressed in growing
162 oocytes but showed decreased expression in MII oocytes (Figure S3A, see also revised heatmap). We
163 have re-analyzed the expression data and corrected the heatmap normalization; the revised Figure
164 S3A now accurately shows that Kdm1b is highly expressed in growing oocytes, with lower levels in 8-
165 week and MII oocytes, consistent with its role in maternal imprint establishment. The CUT&RUN data,
166 which showed small H3K4me2 peaks (related to other stages), reappeared at the late two-cell stage
167 (Figure 3A). We concluded that the IF analysis could not detect the newly established weak H3K4me2
168 signal in the late two-cell stage. H3K4me2 was correlated with the activation of Prdm9 and Kmt2c at
169 the four-cell stage (Figure S3A), and the signal was strong enough to be detected at the four-cell and
170 eight-cell stages (Figures 3A-B) using both the CUT&RUN and IF techniques. The four/eight-cell
171 H3K4me2 targets only partially overlapped with those in sperm or oocytes, and these overlapped
172 targets were mainly enriched in development-associated pathways or processes (Figure 3C).
173 Additionally, gamete-specific H3K4me2 targets were involved in the development and generation of
174 gametes, whereas four/eight cell-specific targets were associated with cell communication and

175 metabolic processes (Figure 3C). These findings indicated that the resetting of H3K4me2 involves
176 distinct sets of genes that are closely related to the development of embryos from mouse gametes to
177 early embryos. About 1% of the four/eight cell H3K4me2 occurred in the hypomethylated regions as
178 early as the late two-cell stage (Figures S3B-D). These results indicated that the resetting of H3K4me2
179 starts from the late two-cell stage but is mostly re-established at the four/eight-cell stage.
180

181 **The inhibition of LSD2 increased the expression of H3K4me2, which impaired the development**
182 **of embryos in vitro**

183 Histone post-translational modifications are coordinated by histone methyltransferases and
184 demethylases[52]. KDM1B, also known as LSD2, is the only demethylase homologous to LSD1 in
185 mammals[29]. Although they share the H3K4me1/2 demethylase function, our analysis showed that
186 Kdm1a was rarely expressed during mouse embryonic development (Figure 4A). While, LSD2, which is
187 required to establish the maternal imprinted gene[53], was highly expressed before the early two-cell
188 stage and decreased after the late two-cell stage (Figures 4A and S3A).
189 Tranylcypromine (TCP) can inhibit the expression of LSD2[54]. To investigate the effect of the expression
190 of H3K4me2 on embryonic development, mouse PN5 embryos were cultured in vitro with three
191 different TCP treatment groups (3.6 mM, 5 mM, and 10 mM) and a control group. The inhibition in the
192 presence of 3.6 mM TCP was not prominent, and some embryos could develop to the eight-cell stage
193 at 65 h. When 10 mM TCP was present, embryonic development was strongly inhibited, and the
194 embryo could not develop to the two-cell stage (Figure S4A). When 5 mM TCP was present, the
195 embryos could pass through the two-cell stage with a significant developmental delay. Some embryos
196 developed till the four-cell stage, and then, were completely arrested. These results showed that TCP
197 affected embryonic development in a dose-dependent manner and caused developmental arrest. As
198 no prior study has reported the precise effective concentration of TCP for inhibiting LSD2 in mouse
199 embryos, our dose-response experiment empirically determined that 5 mM TCP was sufficient to
200 cause significant H3K4me2 accumulation and developmental arrest without immediate lethality (see
201 Methods).We concluded that the TCP concentration of 5 mM in the experimental group (5 mM TCP
202 group) could be used to determine whether the abnormal expression of H3K4me2 affects the
203 development of embryonic cells, especially the occurrence of ZGA.

204 We found that adding TCP to culture embryos significantly reduced the expression of LSD2 in
205 embryos (Figure 4B). After 28 h of culture, some of the embryos in the Control group developed to
206 the four-cell stage, while those in the 5 mM TCP group remained at the two-cell stage (Figures 4C
207 and S4B). The development of the embryos was significantly delayed after TCP was added compared
208 to the time taken for the embryos in the Control group to develop. The number of embryos at each
209 stage in the 5 mM TCP group (n = 169) and the Control group (n = 155) was counted and statistically
210 analyzed (Figure 4D). The results showed that 62.96% of embryos in the 5 mM TCP group developed
211 to the two-cell stage (Control: 96.29%), and 36.82% of embryos developed to the four-cell stage
212 (Control: 93.83%). All embryos in the TCP group were arrested at the four-cell stage ($P < 0.001$).
213 To test whether LSD2 specifically functions as H3K4me2 demethylase in mouse embryos,
214 immunofluorescence staining of cells in the 5 mM TCP group and the Control group was performed
215 using H3K4me1, H3K4me2, and H3K4me3 antibodies, respectively. We calculated the
216 immunofluorescence intensity of individual embryos and found no significant difference in the
217 changes in the expression of H3K4me1 between the 5 mM TCP group and the Control group (Figure
218 S4C); also, the changes in the expression of H3K4me3 was not statistically significant (Figures S4C). In
219 contrast, when LSD2 expression was inhibited by TCP, the expression of H3K4me2 at the two-cell
220 stage was significantly higher than the expression of H3K4me2 in the cells of the Control group ($P <$
221 0.001) (Figure 4E).Elimination of the demethylase action of LSD2 increased the expression of H3K4me2,
222 which delayed and arrested embryonic development. We also found that LSD2 specifically
223 demethylated H3K4me2 in mammalian embryos.
224
225 **Parental H3K4me2 was erased by maternal LSD2 during forming the pronucleus in the zygote**
226 During spermatogenesis, most of the histone modifications are replaced by protamine, but a few
227 histones, including H3K4me2, remain conserved[55]. During the early stages of zygote formation, female
228 and male pronuclei are transcriptionally silent, and the development process is regulated by maternal
229 factors. In our study, during the early stage of zygote formation, the expression of H3K4me2 in PN5
230 was lower than that in PN3 ($P < 0.001$) (Figures 5A-B), which indicated that H3K4me2 signaling was
231 gradually reduced during the fusion of sperm and egg. When PN3 embryos were cultured to the PN5
232 stage under 5 mM TCP conditions, the fluorescence signal of H3K4me2 increased significantly ($P <$

233 0.001) (Figures 5A-B), suggesting that the LSD2 might response the removal of parental H3K4me2
234 during normal embryonic development. The immunofluorescence values of H3K4me2 in female and
235 male pronuclei were assessed to verify whether the effect of LSD2 inhibition on H3K4me2 was
236 different between male and female pronuclei. When LSD2 was inhibited, the expression of H3K4me2
237 increased significantly in female and male pronuclei, especially in the male pronucleus (334% increase
238 in the male pronucleus and 155% increase in the female pronucleus) (Figure 5C). Sperm carrying
239 H3K4me2 might be more affected by LSD2 demethylase, creating a low H3K4me2 expression
240 environment for zygotic genome activation. To summarize, parental H3K4me2 was erased by maternal
241 LSD2, the H3K4me2 in the male pronucleus was erased faster than that in the female pronucleus, and
242 maternal LSD2 specifically demethylated H3K4me2.
243
244 **H3K4me2 participated in the mammalian ZGA and chromatin accessibility regulation**
245 We used the CUT&RUN technology to detect H3K4me2 throughout the genome and performed gene
246 annotation of the differential expression peaks of H3K4me2 in the TCP-treated group and the control
247 group. We found that the H3K4me2 signal increased in 251 genes and decreased in 194 genes (Fold
248 changes > 2, $P < 0.01$) (Figure 6A). The enrichment analysis of the genes from which H3K4me2
249 differential peaks originated showed that the molecular function was mainly enriched in the
250 transcription regulation of the promoter of RNA polymerase Ⅱ, and 35 differential genes were
251 involved in this regulation (Figure 6B).
252 Unlike LSD1, which binds to the promoter region, LSD2 is absent in the promoter region and enriched
253 in the intronic, exonic, and intergenic regions[29]. The distribution of H3K4me2 changed when LSD2
254 expression was inhibited (Figures 6C-D). The increase in the H3K4me2 signal mainly occurred in the
255 intergenic and genebody regions (intergenic region: 55.05%, genebody region: 32.66%). We speculated
256 that H3K4me2 located in these regions might be erased by LSD2. However, the deletion of LSD2 did
257 not increase the H3K4me2 signal in the promoter region.
258 Histone modifications can regulate early embryonic development in mammals by affecting the
259 expression of developmental genes and reprogramming the chromatin structure[1,3,9]. RNA-seq data of
260 mouse germ cells and embryos (refer to GEO database: GSE71434) were used to screen minor/major
261 ZGA genes. We examined the gene expression levels of two-cell embryos in the 5 mM TCP group and

262 the Control group (fold changes > 2, $P < 0.05$) and found that 398 genes were up-regulated and 724
263 genes were down-regulated in the two-cell stage embryos of mice after LSD2 was inhibited (Figures
264 6E-F). Among these differentially expressed genes, 94 down-regulated genes were major ZGA genes,
265 and 18 up-regulated genes were major ZGA genes. The Zfp825 gene played a positive regulatory role
266 in RNA polymerase II promoter transcription[34]. Inhibition of LSD2 can lead to a reduction in the
267 distribution of H3K4me2 on the Zfp825 gene and a weakening of the expression of Zfp825. These
268 changes may affect the frequency, rate, or extent of RNA polymerase II promoter transcription.
269 An increase in the expression of H3K4me2 changed the accessibility of chromatin, of which 1,413
270 peaks were up-regulation and 914 peaks were down-regulated (fold change > 2; $P < 0.05$) (Figure 6G).
271 Peak-related gene screening of the genes near the promoter region (TSS ±2.5 kb) showed that 240
272 peak-related genes were up-regulated and 197 peak-related genes were down-regulated. The
273 transcription of genes requires the opening of the tight structure of chromatin to allow the binding of
274 regulatory factors. The up-regulation of H3K4me2 was associated with altered chromatin accessibility
275 in some developmental genes, which in turn could affect developmental progression. The inhibition of
276 LSD2 expression leads to changes in the distribution of H3K4me2 on genes, abnormal expression of
277 genes associated with ZGA, and disorders related to chromatin opening; these changes together
278 may inhibit normal embryonic development.
279
280 **Non-canonical H3K4me2 in mammalian embryos and pervasive H3K4me2 in CpG-rich**
281 **regulatory regions in the four-cell and eight-cell embryos of mice**
282 In GV oocytes and post-ZGA stages, H3K4me2 was enriched in both canonical gene promoters and
283 partially methylated domains (PMDs) (Figures 7A-D). Our results showed that H3K4me2 might be re-
284 established in the late two-cell stage, and strong promoter H3K4me2 enrichment might be completed
285 in the four-cell stage and maintained until the eight-cell stage. We also found that the H3K4me2
286 signals occupied broader regions than those at other stages (Figure 1A). About 65% of these gene
287 promoters were occupied by H3K4me2 in the four-cell and eight-cell stages and the ICM; these genes
288 played key roles in RNA metabolism and RNA splicing (Figure 7E). We then combined previous
289 H3K4me3 data and chromosome accessibility data and found that these regions overlapped with high
290 CG density regions, H3K4me3 enrichment, and chromatin accessibility regions (Figure 7E). Our results

291 showed that H3K4me2 preferentially occupied hypomethylated and CpG-dense regions (Figures S5A-
292 C).
293 We further examined how the stage-specific distal peaks were reprogrammed and found that about
294 35% of the promoters of the genes that were associated with cell differentiation and development did
295 not undergo the large-scale re-established of H3K4me2 in the four-cell, eight-cell, and blastocyst
296 stages (Figure 7E). We identified genes expressed in specific stages and examined their promoter
297 H3K4me2 (Figure S5E). We found promoters with constantly high H3K4me2 that are preferentially
298 CpG-rich, whereas promoters with constantly low H3K4me2 were generally CpG-poor (Figures 7F and
299 S5E). A group of genes showed dynamic promoter H3K4me2 that correlated with gene expression
300 (Figure S5E). These genes played key roles in development, differentiation, and morphogenesis.
301 Therefore, our results indicated that H3K4me2 of the promoter during early embryonic development
302 correlated with gene activities and CpG density.
303 Next, we investigated distal H3K4me2 signals and found that four/eight-cell embryos also showed
304 widespread distal H3K4me2 signals (Figure 7C). We first identified the distal H3K4me2 regions in
305 mouse gametes and pre-implantation embryos. These regions were relatively CpG-rich (although
306 lower than promoters) and hypomethylated (Figures S5B and S5C). However, compared to promoters,
307 distal H3K4me2 was correlated with CpG density in a continuous and monotonic manner (Figure S5A).
308 To summarize, these results indicated that widespread distal H3K4me2 occurred in CpG-rich and
309 hypomethylated regions in mouse four/eight-cell embryos.
310
311 **Transcription regulation, CcREs priming, and resolving during maternal-to-zygotic transition**
312 A study showed that most of the transcription factor binding regions overlap with H3K4me2-enriched
313 regions[56]. To investigate the association between H3K4me2 and transcription regulation, we identified
314 stage-specific distal regions and the accessibility in these distal peaks of H3K4me2 (Figure 7F). Then,
315 we identified the binding motifs for transcription factors (TFs) enriched in these distal regions (Figure
316 S5D). TFs in mESC and eight-cell embryos include pluripotency factors involved in the development of
317 the system[57,58]. We found that CTF was specifically enriched in eight-cell embryos. The GATA and KLF
318 families are enriched in ICMs.
319 In a study on humans, we found widespread deposition of H3K4me3 at the promoters, including those

320 for developmental genes, before ZGA. These genes were either activated or repressed after ZGA[50].
321 Such promoter priming is related to embryonic development. To determine whether this also occurs in
322 mice, we identified specific H3K4me2 target genes in the four/eight-cell stage and compared them to
323 genes that are expressed or inactive at all stages. The H3K4me2 of promoters in the four/eight-cell-
324 specific group decreased in ICM (Figures 8A and S6A). The "active" genes showed strong H3K4me2,
325 H3K4me3, and ATAC-seq signals before and after ZGA (Figures 8A and S6A). Next, we investigated if
326 such priming-resolving epigenetic transition also occurs in the distal regions. Distal H3K4me2 sites
327 were enriched on regulatory elements, as 75.4% sites (compared to 16.2% of random) of them
328 overlapped with candidate cis-regulatory elements (CcREs), identified by ENCODE (Figure 8B).
329 Enhancers are DNA sequences that increase promoter activity and, in turn, the frequency of gene
330 transcription. The enhancer region is usually modified by H3K27ac (catalyzed by P300/CBP) and
331 H3K4me1 histone 2 (catalyzed by MLL3/4)[59], whereas the H3K4me3 modification level is low. This
332 modification pattern is the opposite of the pattern found in the promoter region. We compared
333 histone modification to H3K4me2 and H3K27ac and found no significant signal overlap between the
334 two histone modifications at specific developmental stages in the distal region. Distal four/eight-cell
335 H3K4me2 marked CcREs showed high enrichment of chromatin accessibility (Figures 8C and S6B).
336 After the four/eight-cell stage, one class ("active") remained accessible and was associated with basic
337 biological processes (GREAT analysis). In contrast, the other class (Figures 8C and S6B; "poised") was
338 associated with H3K27me3 in the ICM and somatic cells; it was located preferentially near
339 developmental genes. These results indicated that the widespread *de novo* H3K4me2 was re-
340 established following ZGA, and H3K4me2 might be globally reset to a permissive state before it is
341 resolved after the four/eight-cell stage (Figure 8D).
342

343 **DISCUSSION**

344 Histone modifications are fundamental epigenetic marks. The reprogramming of histone marks plays
345 an important role in early development[9]. In this study, we profiled the H3K4me2 landscape in mESC,
346 oocytes, and early embryos using a modified low-input, high-sensitivity CUT&RUN method. We found
347 the H3K4me2 showed differential patterns in sperms and oocytes and exhibited non-canonical
348 patterns in GV oocytes but canonical patterns in the sperm in mice[36]. Additionally, the GV H3K4me2

349 peaks were almost erased in MII oocytes, and reconstructed accompanying ZGA occurred. Our results
350 indicated that maternal H3K4me2 was lost during oocyte maturation. The non-canonical H3K4me2 in
351 mammalian embryos and pervasive H3K4me2 in CpG-rich regulatory regions were found in four-cell
352 and eight-cell embryos and in ICM. Hence, we described the first landscape and elucidated the
353 potential function of H3K4me2 during early mammalian development. We also proposed a novel
354 model, known as "Epigenetic Burden", according to which parental H3K4me2 exhibits an "Epigenetic
355 Burden" of maternal-zygote-transition. This burden is erased by maternal factor LSD2 before ZGA.
356 H3K4me2 plays a regulatory role after ZGA (Figure 8 X). H3K4me2 undergoes a transition that starts
357 from an "Epigenetic Burden", and then becomes an "Epigenetic regulator" during early histone
358 modification reprogramming.
359 The paternal epigenetic inheritance is important for development and disease in offspring. Significant
360 epigenome reprogramming occurs and is associated with key processes during early mammalian
361 development[1,3,9,14,39-41,60-64]. Genome-wide H3K4me3 exhibited a non-canonical pattern (ncH3K4me3) in
362 full-grown and mature oocytes, i.e., it occurred as broad peaks at promoters and many distal loci[40,44,48].
363 We found that H3K4me2 was a non-canonical pattern in the GV oocyte, but it was weakened in the
364 MII oocyte, indicating that H3K4me2 plays an important role during oocyte maturation. H3K4me3 is
365 inherited in pre-implantation embryos and is erased in the late two-cell embryos when the process of
366 canonical H3K4me3 is initiated[40,44]. The re-establishment of H3K4me2 mainly starts in the late two-cell
367 stage, and it exists as a non-canonical pattern in the following stages. Several studies have shown that
368 in many species, including mice, zebrafish, and humans, the parental epigenome undergoes dramatic
369 reprogramming after terminally differentiated gamete fertilization, followed by subsequent re-
370 establishment of the embryo epigenome, which leads to epigenetic ''rebooting''[9]. We found that
371 H3K4me2 was erased by LSD2 after fertilization and re-established at the late two-cell stage.
372 Inhibition of the deletion of H3K4me2 resulted in the arrest of embryonic development and abnormal
373 expression of genes related to ZGA. This indicated that H3K4me2 is involved in the reactivation of the
374 embryonic epigenome, and parental H3K4me2 is an "Epigenetic Burden". A previous study reported
375 strong H3K4me2 staining in zygotes and two-cell embryos using different immunofluorescence
376 protocols[51]. This apparent discrepancy likely reflects differences in antibody sensitivity, confocal
377 acquisition parameters, or embryo fixation conditions. In our hands, when we specifically adjusted the

378 laser power and gain for the zygote and two-cell stages, weak H3K4me2 signals could be detected
379 (data not shown). However, to maintain consistent imaging conditions across all stages from GV to
380 blastocyst, we used identical parameters, which may have underestimated the signal at early stages.
381 Therefore, our conclusion that H3K4me2 is largely erased and then re-established remains valid, but
382 the precise dynamics may be more nuanced and dependent on technical settings.
383 H3K4me3 is an active gene expression marker, which mainly marks the gene promoter region.
384 H3K4me1 and H3K4me2 mark enhancer and promoter regions. They have multiple functions for gene
385 activation, according to different enzymes and complexes[22,30,35,65-74]. Some studies have reported a
386 distinct pattern of H3K4me2 enrichment, spanning from the transcriptional start site (TSS) to the gene
387 body, which marks a subset of tissue-specific genes. This pattern is associated with unique H3K4me2
388 profile mark tissue-specific gene regulation[33,35]. Our results showed non-canonical H3K4me2 in
389 mammalian embryos and pervasive H3K4me2 in CpG-rich regulatory regions of four-cell and eight-
390 cell embryos of mice. We next investigated whether H3K4me2 is associated with active or repressed
391 genes in embryo development and found that promoter H3K4me2 in early mouse embryo
392 development is correlated with gene activities and high CpG density. We also found that promoter
393 H3K4me2 was consistent with H3K4me3.
394 Histone modification and chromatin accessibility profiles are usually used to identify transcription
395 factor binding regions and motifs. In another study, we found large-scale chromatin accessibility in
396 pre-ZGA human embryos, and these regions contained many transcription factors[3]. Around 90% of the
397 transcription factor binding regions overlap with the H3K4me2-enriched regions. Stronger signals of
398 H3K4me2 enrichment indicate a greater chance of overlap with the transcription factor binding
399 regions[34]. In this study, we identified the binding motifs for transcription factors enriched in these distal
400 regions, and the CCAAT transcription factor (CTF) was found to be specifically enriched in eight-cell
401 embryos. The GATA and KLF families were enriched in ICMs. GATA is a key regulator of primitive
402 endoderm (PE), which is a part of the ICM[75], and KLF factors support naive pluripotency[76]. These results
403 showed that stage-specific distal H3K4me2 involves transcription circuitry during mammalian
404 embryonic development.
405 The state of chromatin influences DNA replication, DNA repair, and gene expression. Accessible
406 chromatin occupies regulatory sequences, such as promoters and enhancers[77]. These elements interact

407 with cell type-specific transcription factors to start transcription processes that determine cell fate[78].
408 Animal embryos in the early stages undergo extensive histone modification reprogramming, and
409 chromatin remodeling before terminally differentiated gametes develop into pluripotent cells. Some
410 studies have found extensive overlap of accessible promoters and enhancers at the four-cell stage in
411 humans[50]. This finding indicated that crosstalk occurs between histone modification and the chromatin
412 state. In our study, most of the H3K4me2 was re-established in chromatin-accessible regions,
413 including promoters and distal regions at the four-cell stage. Thus, we inferred that H3K4me2 might
414 play an important role in shaping the chromatin state. Our results indicated that crosstalk occurs
415 between accessible chromatin and histone modification.
416 Dynamic histone modification plays an important role during mammalian maternal-to-zygotic
417 transition (MZT). Approximately 22% of the oocyte genome associated with broad H3K4me3 is
418 negatively correlated with DNA methylation and becomes confined to transcriptional-start-site (TSS)
419 regions in two-cell embryos, accompanied by the onset of major ZGA[48]. Also, H3K4me2 is erased in
420 MII oocytes, and sperm H3K4me2 is erased after fertilization. Similar to the pattern found for
421 H3K4me3, H3K4me2 of the embryo was also re-established accompanied by the onset of major ZGA.
422 The ncH3K4me3 in oocytes overlaps almost exclusively with partially methylated DNA domains[40].
423 H3K4me2 is also enriched in canonical gene promoters and partially methylated domains (PMDs).
424 A limitation of the current study is that we did not distinguish between maternal and paternal alleles
425 when analyzing H3K4me2 distribution in pre-implantation embryos. Allele-specific H3K4me2 patterns
426 could reveal asymmetric epigenetic inheritance and potential parental-origin effects on ZGA. Future
427 studies using single-cell or low-input allele-resolved CUT&RUN, combined with SNP-based allele
428 assignment, would be valuable to determine whether certain regulatory regions retain parental-
429 specific H3K4me2 marks and how they contribute to embryonic development.
430 Overall, our findings provided a genome-wide landscape of H3K4me2 modification in mammalian
431 oocytes and pre-implantation embryos. We showed that maternal LSD2 specifically demethylates
432 H3K4me2 after fertilization in mammals. Also, H3K4me2 is erased to a greater extent in the male
433 pronucleus than in the female pronucleus. Based on our results, we proposed a novel model, the
434 "Epigenetic Burden" of histone modification reprogramming during mammalian development. This
435 model can be used to further investigate the epigenetic regulation network in the early developmental

436 stages of mammals.

437

438 MATERIALS AND METHODS

439 Culture of K562 cells

440 The human K562 cell line was cultured at 37 °C, 5% CO_2, and saturated humidity. The cells were
441 cultured in 1640 culture medium (GIBCO) containing 10% fetal bovine serum and 90% 1640 culture
442 liquid.

443 Cell culture of mouse ESC

444 Mouse embryo stem cells (ESC) were cultured on 0.1% gelatin-coated plates without irradiated mouse
445 embryonic fibroblasts (MEFs), using standard ESC medium consisting of DMEM (GIBCO) supplemented
446 with 15–20% fetal bovine serum (FBS, Sigma), 0.1 mM β-mercaptoethanol (Sigma), 0.1 mM MEM non-
447 essential amino acids (Cellgro), 2 mM Glutamax (GIBCO), 1% nucleoside (Millipore), 1% EmbryoMax®
448 Penicillin-Streptomycin Solution (Millipore), and 10 μg/L recombinant leukemia inhibitory factor (LIF)
449 (Millipore). The mouse embryo stem cells were cultured at 37 °C, 5% CO_2, and saturated humidity.

450 Collection of mouse gametes and early embryos

451 The mouse oocytes were collected from 4–5-week-old C57BL/6N female mice (purchased from
452 Charles River). The pre-implantation embryos were harvested from 4–5-week-old C57BL/6N females
453 mated with DBA/2 males (eight-week-old when purchased from Charles River). The Germinal vesicle
454 (GV) oocytes were obtained from 4–5-week-old C57BL/6N females, 46–48 h after they were injected
455 with 10 IU of pregnant mare serum gonadotropin (PMSG, Solarbio). The 4–5-week-old C57BL/6N
456 female mice were also administered 10 IU of human chorionic gonadotrophin (HCG) intraperitoneally,
457 46–48 h after PMSG was administered. Then, they were mated with DBA/2 males. The PN5 zygotes
458 were collected at 28–32 h post-hCG injection. Early two-cell, late two-cell, four-cell, and eight-cell
459 embryos, morulas, and blastocysts were harvested at 32–36 h, 45–48 h, 54–58 h, 60–64 h, 68–75 h,
460 and 84–96 h post-hCG injection. The oocytes in the metaphase of second meiosis (MII) were collected
461 12–14 h after 10 IU of hCG was injected intraperitoneally. The zona pellucida of all oocytes and
462 embryos was removed using glutathione (GSH Solarbio).

463 Tranylcypromine treatment of mouse embryos

464 One-cell embryos were obtained at 24–26 h (PN3) and 28–32 h (PN5) post hCG injection and cultured

465 in the G1 medium in the presence or absence of Tranylcypromine (5 mM). After 48 h, the embryos
466 were transferred to the G2 medium with or without 5 mM TCP. All embryos were maintained at 37 °C
467 in a humidified atmosphere containing 5% CO_2. Then, morphologically normal PN5, late two-cell, and
468 four-cell embryos were collected for further experiments.

469 Immunofluorescence

470 The samples (mouse oocytes or pre-implantation embryos) were washed twice in MII medium (Sigma),
471 and 4% PFA was used as a fixation buffer (BOSTER, AR1068) to fix samples at room temperature for
472 about 60 min. Next, the samples were transferred using a mouth pipette and washed once with 1%
473 BSA (Sigma, A1933) in 1×PBS (Corning, 21–040-CVR). The samples were then permeabilized using
474 0.5% Triton X-100 (Sigma, T8787) in 1×PBS at room temperature for 20 min. Then, they were washed
475 with 1% BSA in 1×PBS and blocked using 1% BSA in 1×PBS at room temperature for 1 h. The blocked
476 samples were incubated with H3K4me2 antibody (Sigma Millpore, 07030) in 1% BSA in 1×PBS diluted
477 at a ratio of 1:500, at room temperature for 2 h or 4 °C overnight. After incubation, they were washed
478 at least five times with 1% BSA in 1×PBS for 3 min. The samples were incubated with Alexa Fluor 568-
479 conjugated donkey anti-rabbit IgG (Invitrogen, A10042) in 1% BSA in 1×PBS diluted at a ratio of 1:200,
480 at room temperature for 50 min in the dark. After incubation, the samples were washed five times with
481 1% BSA in 1×PBS for 3 min. Finally, the samples were stained using the ProLong@ Gold Antifade
482 Reagent with DAPI (Cell Signaling, 8961S), using a 4-Chamber Glass Bottom Dish (100 per case), 35
483 mm Dish with 20 mm Bottom Well (Cellvis, D35C4–20–1-N). The samples were incubated at room
484 temperature for about 24 h in the dark. All images were captured using a Zeiss 700 confocal
485 microscope. Raw images were projected along the Z-axis and processed using the ZEN software (Zeiss)
486 with brightness and contrast adjusted (the same parameters were used for each channel of all
487 samples). In the functional test section, we used the LSD2 antibody (Abcam, ab193080), H3K4me1
488 antibody (Abcam, ab176877), H3K4me2 antibody (mAb) (Active motif, 39079), and H3K4me3 antibody
489 (Active motif, 39016) as primary antibodies (diluted at a ratio of 1:500). Secondary antibodies (diluted
490 at a ratio of 1:300) used included Goat Anti-Rabbit IgG H&L (Alexa Fluor@488) (Abcam, ab150077)
491 and Goat Anti-Mouse IgG H&L (Alexa Fluor® 488) (Abcam, ab150113).

492 Zona pellucida removal and ICM/TE isolation

493 Before the CUT&RUN experiment was conducted, to guarantee the effect of permeation, the zona

494 pellucida of all oocytes and embryos was chemically removed by glutathione (GSH Solarbio). The zona
495 pellucida was removed after about 10–30 s. Then, the samples were washed twice in PBS (Corning 21–
496 040-CVR) and washed once with wash buffer. The inner cell mass (ICM)s or TEs of day 4–5 blastocysts
497 were isolated mechanically. Before performing CUT&RUN, the separated cells were washed twice with
498 PBS and once with wash buffer, and then, resuspended in wash buffer.

499 RNA-seq

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

507 Generation of miniATAC-seq library

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

510 CUT&RUN library construction and next-generation sequencing

511 The CUT&RUN method was used following published protocols[37,38] with modifications in antibody
512 binding, adaptor ligation, and DNA purification steps. All samples (including mouse zygotes, embryos
513 with zona pellucida removed, or ESCs) were transferred into 0.2 mL conventional, non-low-binding
514 PCR tubes (Axygen). The samples were resuspended in 60 μL of wash buffer (20 mM HEPES-KOH,
515 pH = 7.5; 150 mM NaCl; 0.5 mM Spermidine; Roche complete protease inhibitor). Concanavalin-
516 coated magnetic beads (Polyscience, 86057) were gently washed thrice or four times and resuspended
517 in binding buffer (20 mM HEPES-KOH, pH = 7.9; 10 mM KCl; 1 mM $CaCl_2$; 1 mM $MnCl_2$), and then, 10–
518 15 μL of Concanavalin-coated magnetic beads were added to each sample. The samples were
519 incubated at room temperature (about 25 °C) at 450 rpm for 20 min in a Thermomixer (Eppendorf).
520 The samples were placed on a magnetic stand for removing the supernatant, and then, resuspended
521 in 50 μL antibody buffer (wash buffer plus digitonin 1000:1, 2 mM EDTA, pH = 8.0) with H3K4me2
522 antibody (Sigma Millpore, 07030) diluted at ratio of 1:100. The samples were incubated at room

523 temperature (23–26 °C) for 2 h at 450 rpm on Thermomixer, or incubated overnight (⩾ 16 h) at 4 °C
524 and 450 rpm in a Thermomixer.
525 Next, the samples were held at the magnetic stand and washed with 190 μL of dig wash buffer (wash
526 buffer plus digitonin, freshly pre-heated, 0.005%–0.01%; tested for each batch) once, and then, washed
527 with 100 μL of the buffer again. The samples were then resuspended in 50 μL of dig wash buffer with
528 PA/G-MNase (to a final concentration of 700 ng/mL), and incubated at room temperature (23–26 °C)
529 on Thermomixer for 1 h, or incubated at 4 °C on Thermomixer for 3 h at 450 rpm. The samples were
530 washed with 190 μL of dig wash buffer twice and then washed again with 100 μL of the buffer at the
531 magnetic stand. Next, 100 μL of dig wash buffer was added and incubated on ice for 3 min. Targeted
532 cleavage was performed on ice by adding 2 μL of 100 mM $CaCl_2$ for 30 min, and the reaction was
533 stopped by adding 100 μL of 2×stop buffer (340 mM NaCl; 20 mM EDTA, pH = 8.0; 4 mM EGTA,
534 pH = 8.0; 50 μg/mL RNase A; 100 μg/mL glycogen; 0.005%–0.01% preheated digitonin; tested for each
535 batch), and then, the mixture was fully vortexed. The samples were incubated at 37 °C for 30 min for
536 fragment release. The total samples or supernatants were digested by adding 10 ng of carrier RNA
537 (Qiagen, 1068337), 1 μL of 20% SDS, and 2 μL of Proteinase K (Roche). The reaction was performed at
538 56 °C for 45 min and 72 °C for 20 min in a thermal cycler. DNA was purified using phenol-chloroform,
539 followed by ethanol purification. Then, 200 μL of phenol-chloroform was added, mixed, and
540 centrifuged at 20 °C and 16,000 *g* for 5 min. Next, 200 μL of chloroform was added to each sample.
541 After mixing, the samples were centrifuged at 20 °C and 16,000 *g* for 5 min and transferred to Lo-
542 binding 1.5 mL tubes. Each sample was mixed with 2 μL of glycogen, 20 μL of NaOAC, and 500 μL of
543 100% ethanol (precooled at –20 °C). The mixture was incubated at –20 °C overnight.
544 The samples were centrifuged at 16,000 *g* and 4 °C for 20 min, and then, the supernatant was
545 removed. Next, 1 mL of anhydrous ethanol was added to the samples, which were then centrifuged at
546 4 °C and 16,000 *g* for 10 min. The supernatant was removed by opening the lid of the tubes and air-
547 drying. After the liquid disappeared, purified DNA was used for Tru-Seq library construction using NEB
548 Next Ultra II DNA Library Prep Kit for Illumina (NEB, E7645S). For the NEB kit, the DNA was
549 resuspended in 25 μL of ddH_2O, followed by end-repair/A-tailing with 3.5 μL of End Prep buffer and
550 1.5 μL of an enzyme mix. The reaction was conducted at 20 °C for 30 min and 65 °C for 30 min in a
551 PCR machine. Then, 1.25 μL of diluted adaptors (NEB Multiplex Oligos for Illumina, E7335S) were

552 added and thoroughly vortexed. Next, 15 μL of Ligation master mix and 0.5 μL of Ligation enhancer
553 were added, and the mixture was incubated at 16–20 °C for 16–20 min, or at 4 °C overnight. After the
554 ligation reaction was over, the samples were treated with 1.5 μL of USERTM enzyme. The reaction was
555 performed at 37 °C for 15 min in a PCR machine. The samples were purified using 1×AMPure beads.
556 The PCR was performed by adding 25 μL of 2×KAPA HiFi Hot Start Ready Mix (KAPA biosystems,
557 KM2602) with 2.5 μL of primers of the NEB Oligos kit and 2.5 μL of Illumina universal primers. The
558 conditions used were as follows: 98 °C for 45s, (98 °C for 15s and 60 °C for 10s) with 16–17 cycles, and
559 72 °C for 1 min. The final libraries were purified using 0.9×AMPure beads and used for next-
560 generation sequencing.
561
562
563 Data analysis
564 CUT&RUN data processing
565 The single-end CUT&RUN data were mapped to the mm9 genome using Bowtie2 (Version 2.2.9)[79]
566 with the parameters –t –q –N 1 –L 25. The paired-end CUT&RUN data were mapped based on the
567 parameters: -t –q –N 1 –L 25 –X 1000 --nomixed --no-discordant. All unmapped reads, non-uniquely
568 mapped reads, and PCR duplicates were removed. For downstream analysis, the read counts were
569 normalized by calculating the number of reads per kilobase of bin per million reads sequenced (RPKM)
570 for 100-bp bins of the genome. The RPKM values were further Z-score normalized for minimizing the
571 batch effect.
572 ATAC-seq data processing
573 The ATAC-seq reads were mapped to the mm9 genome using Bowtie2 (version 2.2.2)[79] with previously
574 described parameters. All unmapped reads, non-uniquely mapped reads, and PCR duplicates were
575 removed. For downstream analysis, the read counts were normalized by calculating the RPKM for 100
576 bp-bins of the genome. The RPKM values were further normalized by Z-score normalization to
577 minimize the batch effect.
578 DNA methylation data processing
579 We used the Bismark software for mapping DNA methylation reads to the mm9 genome. PCR
580 duplicates were removed. For each CpG site, the methylation level was computed as the total

581 methylated counts divided by the total counts across all reads covering the CpG site.

582 Analysis of the promoter H3K4me2 and H3K4me3 enrichment

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

586 Identification of stage-specific distal ATAC-seq peaks

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

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

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

603 Clustering analysis

604 We conducted k-means clustering of gene expression levels at various stages using Cluster 3.0 with
605 the parameters –g 7 (Euclidean distance). Heat maps were generated using Java Treeview 3.0. We
606 conducted hierarchical clustering analysis for H3K4me3 at various stages using the R package "ape"
607 based on Pearson's correlation between each pair of stages. The distance between the stages was
608 calculated as 1 − Pearson's correlation.

609 Identification of oocyte PMDs

610 We calculated the average DNA methylation levels for 10-kb bins of the genome and counted the
611 number of CpG dyads in each 10-kb bin. Bins with an average DNA methylation level below 0.6 and
612 above 20 CpG dyads were selected and merged into PMDs using BEDtools (Version v2.25.0)[80]. The
613 promoter regions (±2.5 kb) were removed from PMDs.

614 Identification of H3K4me2 peaks

615 The H3K4me2 peaks were called using MACS2 (Version 2.0.9)[81] with the parameters --broad --
616 nomodel --nolambda -g mm 1.87e9/-g hs 2.7e9. Noisy peaks with weak signals (summed RPKM < 0)
617 were removed before further analyses were conducted.

618 Statistical analysis

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

621

622 DATA AVAILABILITY

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

626 FUNDING AND ACKNOWLEDGEMENT

627 This work was supported by the National Key R&D Program of China (2019YFA0802200 and
628 2019YFA0110900 to Jiawei Xu), the National Natural Science Foundation of China 32170819 to Jiawei
629 Xu, Henan University science and technology innovation team to Jiawei Xu), Scientific Research
630 Innovation team of the First Affiliated Hospital of Zhengzhou University Excellent innovation team to
631 Jiawei Xu, Henan Zhongyuan Medical Science and Technology Innovation Development Foundation to
632 Jiawei Xu.

633

634 Author contributions

635 Conceptualization: JWX,HC

636 Methodology: CW, YQW, YL

637 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).

10 Goldberg, A. *et al.* Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**, 678-691, doi:10.1016/j.cell.2010.01.003 (2010).

11 Jin, C. *et al.* H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nature genetics* **41**, 941-945, doi:10.1038/ng.409 (2009).

12 Mito, Y., Henikoff, J. & Henikoff, S. Genome-scale profiling of histone H3.3 replacement patterns. *Nature genetics* **37**, 1090-1097, doi:10.1038/ng1637 (2005).

13 Ishiuchi, T. *et al.* Reprogramming of the histone H3.3 landscape in the early mouse embryo. *Nature structural & molecular biology* **28**, 38-49, doi:10.1038/s41594-020-00521-1 (2021).

14 Zheng, H. *et al.* Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. *Mol Cell* **63**, 1066-1079, doi:10.1016/j.molcel.2016.08.032 (2016).

676 15 Zentner, G. E. & Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nat*
677 *Struct Mol Biol* **20**, 259-266, doi:10.1038/nsmb.2470 (2013).
678 16 Soares, L. M. *et al.* Determinants of Histone H3K4 Methylation Patterns. *Mol Cell* **68**, 773-785 e776,
679 doi:10.1016/j.molcel.2017.10.013 (2017).
680 17 Klose, R. J. & Zhang, Y. Regulation of histone methylation by demethylimination and demethylation.
681 *Nat Rev Mol Cell Biol* **8**, 307-318, doi:10.1038/nrm2143 (2007).
682 18 Eissenberg, J. C. & Shilatifard, A. Histone H3 lysine 4 (H3K4) methylation in development and
683 differentiation. *Dev Biol* **339**, 240-249, doi:10.1016/j.ydbio.2009.08.017 (2010).
684 19 Ruthenburg, A. J., Allis, C. D. & Wysocka, J. Methylation of lysine 4 on histone H3: intricacy of
685 writing and reading a single epigenetic mark. *Mol Cell* **25**, 15-30, doi:10.1016/j.molcel.2006.12.014
686 (2007).
687 20 Rickels, R. *et al.* Histone H3K4 monomethylation catalyzed by Trr and mammalian COMPASS-like
688 proteins at enhancers is dispensable for development and viability. *Nat Genet* **49**, 1647-1653,
689 doi:10.1038/ng.3965 (2017).
690 21 Smith, E. & Shilatifard, A. Enhancer biology and enhanceropathies. *Nat Struct Mol Biol* **21**, 210-219,
691 doi:10.1038/nsmb.2784 (2014).
692 22 Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts
693 developmental state. *Proc Natl Acad Sci U S A* **107**, 21931-21936, doi:10.1073/pnas.1016071107
694 (2010).
695 23 Hu, D. *et al.* The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4
696 monomethylases at enhancers. *Mol Cell Biol* **33**, 4745-4754, doi:10.1128/MCB.01181-13 (2013).
697 24 Liu, T. *et al.* Broad chromosomal domains of histone modification patterns in C. elegans. *Genome*
698 *Res* **21**, 227-236, doi:10.1101/gr.115519.110 (2011).
699 25 Du, Z. *et al.* Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and
700 H3K27ac in Oryza sativa L. Japonica. *Mol Plant* **6**, 1463-1472, doi:10.1093/mp/sst018 (2013).
701 26 van Dijk, K. *et al.* Dynamic changes in genome-wide histone H3 lysine 4 methylation patterns in
702 response to dehydration stress in Arabidopsis thaliana. *BMC Plant Biol* **10**, 238, doi:10.1186/1471-
703 2229-10-238 (2010).
704 27 Zheng, Y. C. *et al.* A Systematic Review of Histone Lysine-Specific Demethylase 1 and Its Inhibitors.
705 *Med Res Rev* **35**, 1032-1071, doi:10.1002/med.21350 (2015).
706 28 Yang, Z. *et al.* AOF1 is a histone H3K4 demethylase possessing demethylase activity-independent
707 repression function. *Cell Res* **20**, 276-287, doi:10.1038/cr.2010.12 (2010).
708 29 Fang, R. *et al.* Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating intragenic
709 H3K4me2 methylation. *Mol Cell* **39**, 222-233, doi:10.1016/j.molcel.2010.07.008 (2010).
710 30 Wang, S., Meyer, D. H. & Schumacher, B. H3K4me2 regulates the recovery of protein biosynthesis
711 and homeostasis following DNA damage. *Nat Struct Mol Biol* **27**, 1165-1177, doi:10.1038/s41594-
712 020-00513-1 (2020).
713 31 Liu, Y. *et al.* H3K4me2 functions as a repressive epigenetic mark in plants. *Epigenetics & chromatin*
714 **12**, 40, doi:10.1186/s13072-019-0285-6 (2019).
715 32 Zhang, C. *et al.* Narrow H3K4me2 is required for chicken PGC formation. *J Cell Physiol* **236**, 1391-
716 1400, doi:10.1002/jcp.29945 (2021).
717 33 Zhang, J., Parvin, J. & Huang, K. Redistribution of H3K4me2 on neural tissue specific genes during
718 mouse brain development. *BMC Genomics* **13 Suppl 8**, S5, doi:10.1186/1471-2164-13-S8-S5
719 (2012).

720 34 Wang, Y., Li, X. & Hu, H. H3K4me2 reliably defines transcription factor binding regions in different
721 cells. *Genomics* **103**, 222-228, doi:10.1016/j.ygeno.2014.02.002 (2014).
722 35 Pekowska, A., Benoukraf, T., Ferrier, P. & Spicuglia, S. A unique H3K4me2 profile marks tissue-
723 specific gene regulation. *Genome Res* **20**, 1493-1502, doi:10.1101/gr.109389.110 (2010).
724 36 Lambrot, R., Siklenka, K., Lafleur, C. & Kimmins, S. The genomic distribution of histone H3K4me2 in
725 spermatogonia is highly conserved in spermdagger. *Biol Reprod* **100**, 1661-1672,
726 doi:10.1093/biolre/ioz055 (2019).
727 37 Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high
728 efficiency for low cell numbers. *Nat Protoc* **13**, 1006-1019, doi:10.1038/nprot.2018.015 (2018).
729 38 Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of
730 DNA binding sites. *Elife* **6**, doi:10.7554/eLife.21856 (2017).
731 39 Wang, L. *et al.* Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**,
732 979-991, doi:10.1016/j.cell.2014.04.017 (2014).
733 40 Zhang, B. *et al.* Allelic reprogramming of the histone modification H3K4me3 in early mammalian
734 development. *Nature* **537**, 553-557, doi:10.1038/nature19361 (2016).
735 41 Wu, J. *et al.* The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature*
736 **534**, 652-657, doi:10.1038/nature18606 (2016).
737 42 Siklenka, K. *et al.* Disruption of histone methylation in developing sperm impairs offspring health
738 transgenerationally. *Science (New York, N.Y.)* **350**, aab2006, doi:10.1126/science.aab2006 (2015).
739 43 Hanna, C. W. *et al.* MLL2 conveys transcription-independent H3K4 trimethylation in oocytes.
740 *Nature structural & molecular biology* **25**, 73-82, doi:10.1038/s41594-017-0013-5 (2018).
741 44 Liu, X. *et al.* Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation
742 embryos. *Nature* **537**, 558-562, doi:10.1038/nature19362 (2016).
743 45 Li, L., Lu, X. & Dean, J. The maternal to zygotic transition in mammals. *Molecular aspects of*
744 *medicine* **34**, 919-938, doi:10.1016/j.mam.2013.01.003 (2013).
745 46 Lee, M. T., Bonneau, A. R. & Giraldez, A. J. Zygotic genome activation during the maternal-to-
746 zygotic transition. *Annual review of cell and developmental biology* **30**, 581-613,
747 doi:10.1146/annurev-cellbio-100913-013027 (2014).
748 47 Liu, Y.-W. *et al.* LincRNAFEZF1-AS1 represses p21 expression to promote gastric cancer
749 proliferation through LSD1-Mediated H3K4me2 demethylation. *Molecular cancer* **16**, 39,
750 doi:10.1186/s12943-017-0588-9 (2017).
751 48 Dahl, J. A. *et al.* Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic
752 transition. *Nature* **537**, 548-552, doi:10.1038/nature19360 (2016).
753 49 Xu, Q. *et al.* SETD2 regulates the maternal epigenome, genomic imprinting and embryonic
754 development. *Nat Genet* **51**, 844-856, doi:10.1038/s41588-019-0398-7 (2019).
755 50 Xia, W. *et al.* Resetting histone modifications during human parental-to-zygotic transition. *Science*
756 **365**, 353-360, doi:10.1126/science.aaw5118 (2019).
757 51 Ancelin, K. *et al.* Maternal LSD1/KDM1A is an essential regulator of chromatin and transcription
758 landscapes during zygotic genome activation. *Elife* **5**, doi:10.7554/eLife.08851 (2016).
759 52 Kaniskan, H. U., Martini, M. L. & Jin, J. Inhibitors of Protein Methyltransferases and Demethylases.
760 *Chem Rev* **118**, 989-1068, doi:10.1021/acs.chemrev.6b00801 (2018).
761 53 Ciccone, D. N. *et al.* KDM1B is a histone H3K4 demethylase required to establish maternal genomic
762 imprints. *Nature* **461**, 415-418, doi:10.1038/nature08315 (2009).
763 54 Dai, X. J. *et al.* Tranylcypromine Based Lysine-Specific Demethylase 1 Inhibitor: Summary and

764 Perspective. *J Med Chem* **63**, 14197-14215, doi:10.1021/acs.jmedchem.0c00919 (2020).
765 55 Jukam, D., Shariati, S. A. M. & Skotheim, J. M. Zygotic Genome Activation in Vertebrates. *Dev Cell*
766 **42**, 316-332, doi:10.1016/j.devcel.2017.07.026 (2017).
767 56 Wang, Y., Li, X. & Hu, H. H3K4me2 reliably defines transcription factor binding regions in different
768 cells. *Genomics* **103**, 222-228, doi:10.1016/j.ygeno.2014.02.002 (2014).
769 57 Dominguez, M. H., Ayoub, A. E. & Rakic, P. POU-III transcription factors (Brn1, Brn2, and Oct6)
770 influence neurogenesis, molecular identity, and migratory destination of upper-layer cells of the
771 cerebral cortex. *Cerebral cortex (New York, N.Y. : 1991)* **23**, 2632-2643, doi:10.1093/cercor/bhs252
772 (2013).
773 58 Sarkar, A. & Hochedlinger, K. The sox family of transcription factors: versatile regulators of stem
774 and progenitor cell fate. *Cell stem cell* **12**, 15-30, doi:10.1016/j.stem.2012.12.007 (2013).
775 59 Field, A. & Adelman, K. Evaluating Enhancer Function and Transcription. *Annual review of*
776 *biochemistry* **89**, 213-234, doi:10.1146/annurev-biochem-011420-095916 (2020).
777 60 Xavier, M. J., Roman, S. D., Aitken, R. J. & Nixon, B. Transgenerational inheritance: how impacts to
778 the epigenetic and genetic information of parents affect offspring health. *Hum Reprod Update* **25**,
779 518-540, doi:10.1093/humupd/dmz017 (2019).
780 61 Chen, Q., Yan, W. & Duan, E. Epigenetic inheritance of acquired traits through sperm RNAs and
781 sperm RNA modifications. *Nat Rev Genet* **17**, 733-743, doi:10.1038/nrg.2016.106 (2016).
782 62 Zhang, B. *et al.* Widespread Enhancer Dememorization and Promoter Priming during Parental-to-
783 Zygotic Transition. *Molecular cell* **72**, doi:10.1016/j.molcel.2018.10.017 (2018).
784 63 Chen, Z., Djekidel, M. & Zhang, Y. J. N. g. Distinct dynamics and functions of H2AK119ub1 and
785 H3K27me3 in mouse preimplantation embryos. **53**, 551-563, doi:10.1038/s41588-021-00821-2
786 (2021).
787 64 Xia, W. *et al.* Resetting histone modifications during human parental-to-zygotic transition. **365**,
788 353-360, doi:10.1126/science.aaw5118 (2019).
789 65 Agarwal, S. *et al.* KDM1A maintains genome-wide homeostasis of transcriptional enhancers.
790 doi:10.1101/gr.234559.118 (2021).
791 66 Ye, R., Cao, C. & Xue, Y. J. E. i. b. Enhancer RNA: biogenesis, function, and regulation. **64**, 883-894,
792 doi:10.1042/ebc20200014 (2020).
793 67 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
794 play roles in nucleosome eviction and eRNA transcription, respectively, at enhancers. **35**, e21781,
795 doi:10.1096/fj.202100488R (2021).
796 68 Kang, H. *et al.* Dynamic regulation of histone modifications and long-range chromosomal
797 interactions during postmitotic transcriptional reactivation. **34**, 913-930,
798 doi:10.1101/gad.335794.119 (2020).
799 69 Ishihara, H. *et al.* Primed histone demethylation regulates shoot regenerative competency. **10**,
800 1786, doi:10.1038/s41467-019-09386-5 (2019).
801 70 Murphy, P., Wu, S., James, C., Wike, C. & Cairns, B. J. C. Placeholder Nucleosomes Underlie
802 Germline-to-Embryo DNA Methylation Reprogramming. **172**, 993-1006.e1013,
803 doi:10.1016/j.cell.2018.01.022 (2018).
804 71 Yan, J. *et al.* Histone H3 lysine 4 monomethylation modulates long-range chromatin interactions at
805 enhancers. **28**, 204-220, doi:10.1038/cr.2018.1 (2018).
806 72 Soldi, M. *et al.* Chromatin proteomics reveals novel combinatorial histone modification signatures
807 that mark distinct subpopulations of macrophage enhancers. **45**, 12195-12213,

808 doi:10.1093/nar/gkx821 (2017).
809 73 Dorighi, K. *et al.* Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters
810 Independently of H3K4 Monomethylation. **66**, 568-576.e564, doi:10.1016/j.molcel.2017.04.018
811 (2017).
812 74 Kaikkonen, M. *et al.* Remodeling of the enhancer landscape during macrophage activation is
813 coupled to enhancer transcription. **51**, 310-325, doi:10.1016/j.molcel.2013.07.010 (2013).
814 75 Morrisey, E. E. *et al.* GATA6 regulates HNF4 and is required for differentiation of visceral endoderm
815 in the mouse embryo. *Genes & development* **12**, 3579-3590 (1998).
816 76 Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to
817 primed pluripotency in rodents and humans. *Nature reviews. Molecular cell biology* **17**, 155-169,
818 doi:10.1038/nrm.2015.28 (2016).
819 77 Gross, D. S. & Garrard, W. T. Nuclease hypersensitive sites in chromatin. *Annual review of*
820 *biochemistry* **57**, 159-197 (1988).
821 78 Rivera, C. M. & Ren, B. Mapping human epigenomes. *Cell* **155**, 39-55,
822 doi:10.1016/j.cell.2013.09.011 (2013).
823 79 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-
824 359, doi:10.1038/nmeth.1923 (2012).
825 80 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features.
826 *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).
827 81 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-
828 2008-9-9-r137 (2008).
829

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

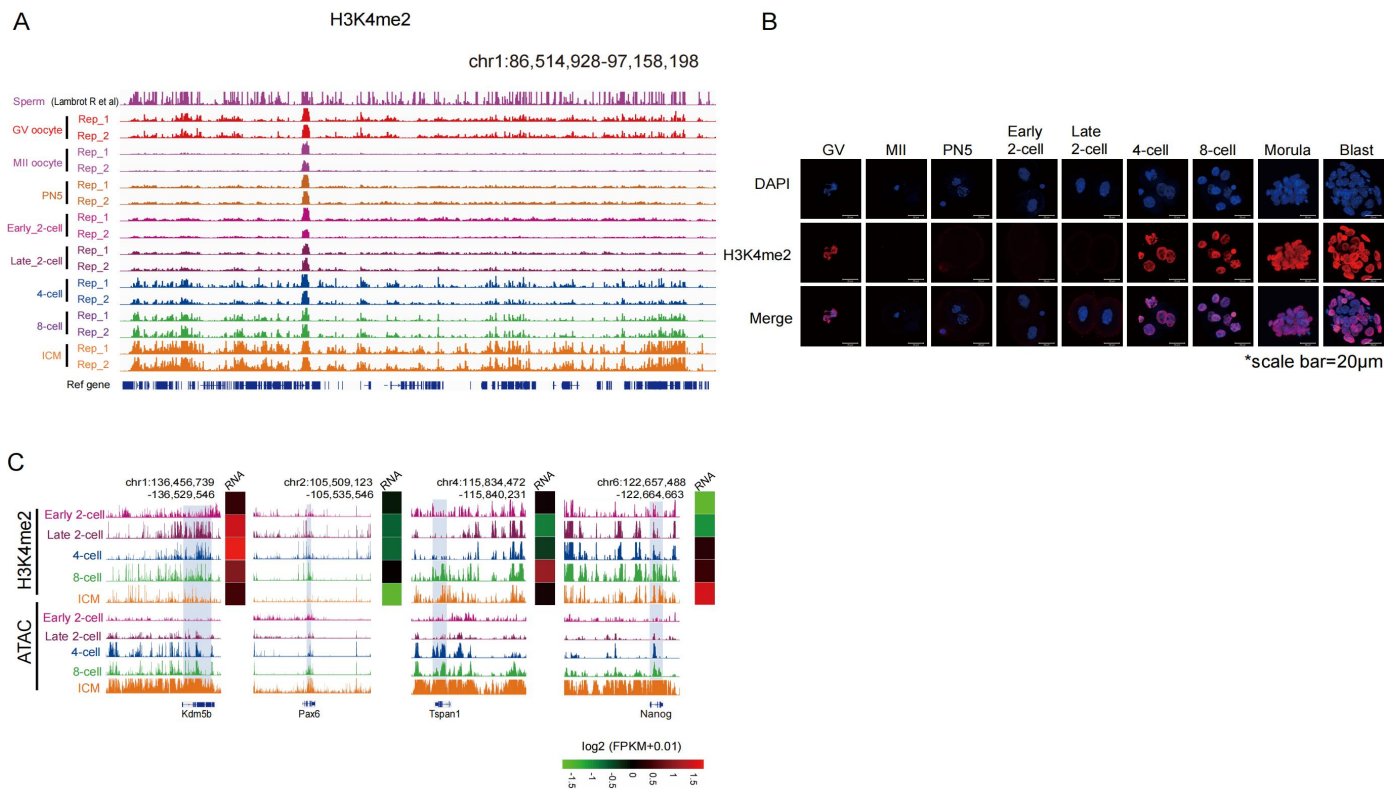


Figure 1. Mapping H3K4me2 in mouse (C57BL6/N) gametes and pre-implantation embryos.

(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

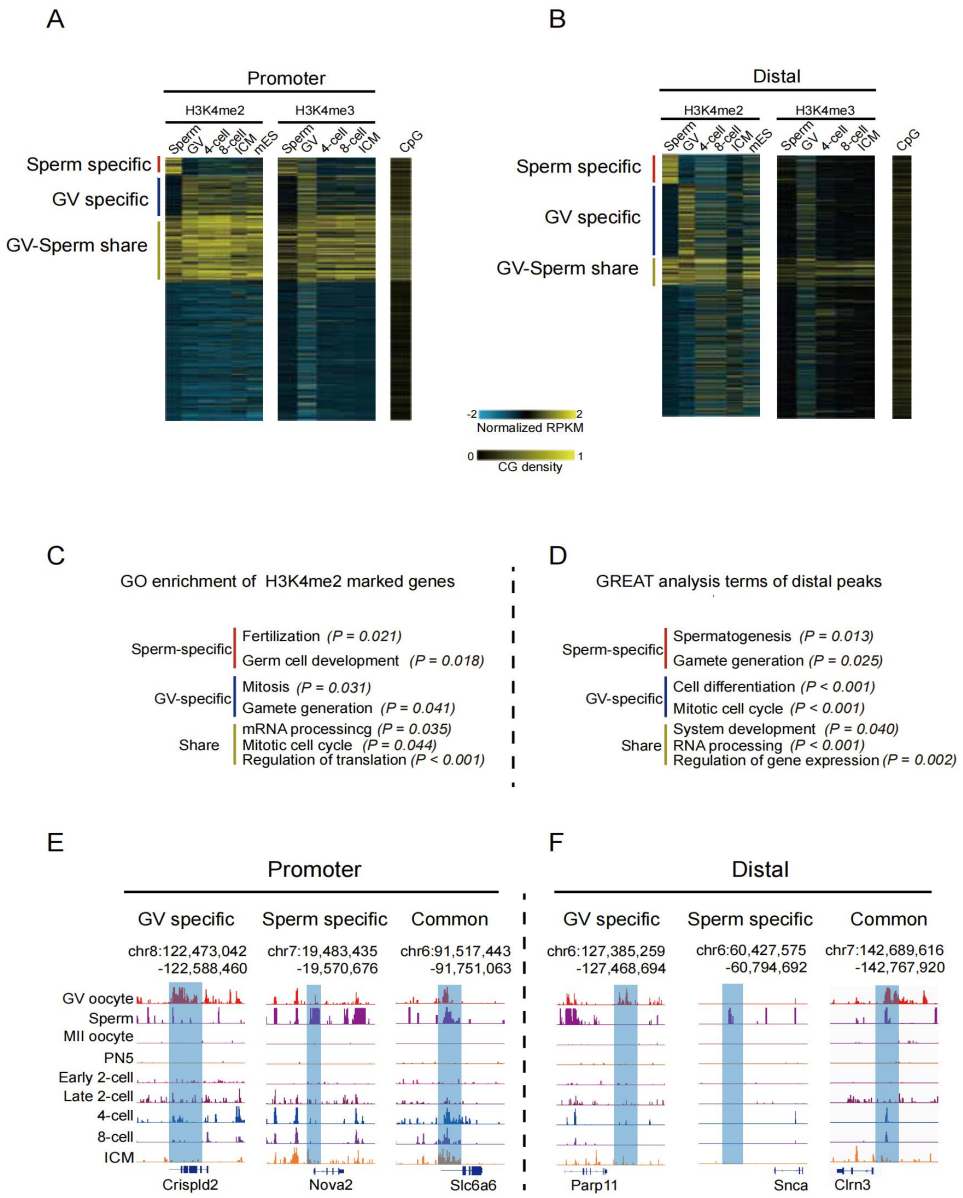


Figure 2. Comparison of H3K4me2 between sperms and oocytes.

(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).

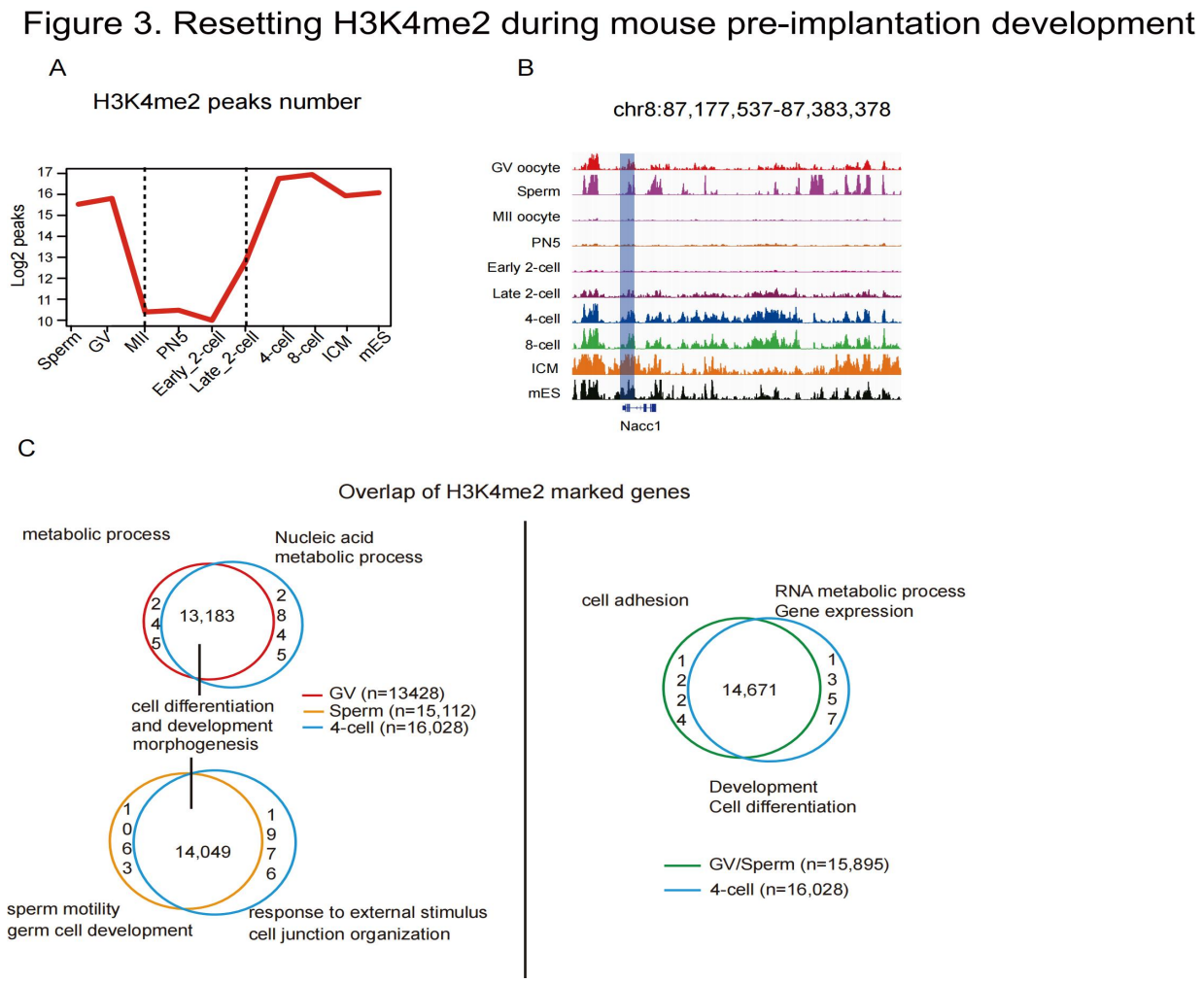


Figure 3. Resetting of H3K4me2 during mouse pre-implantation development.

(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

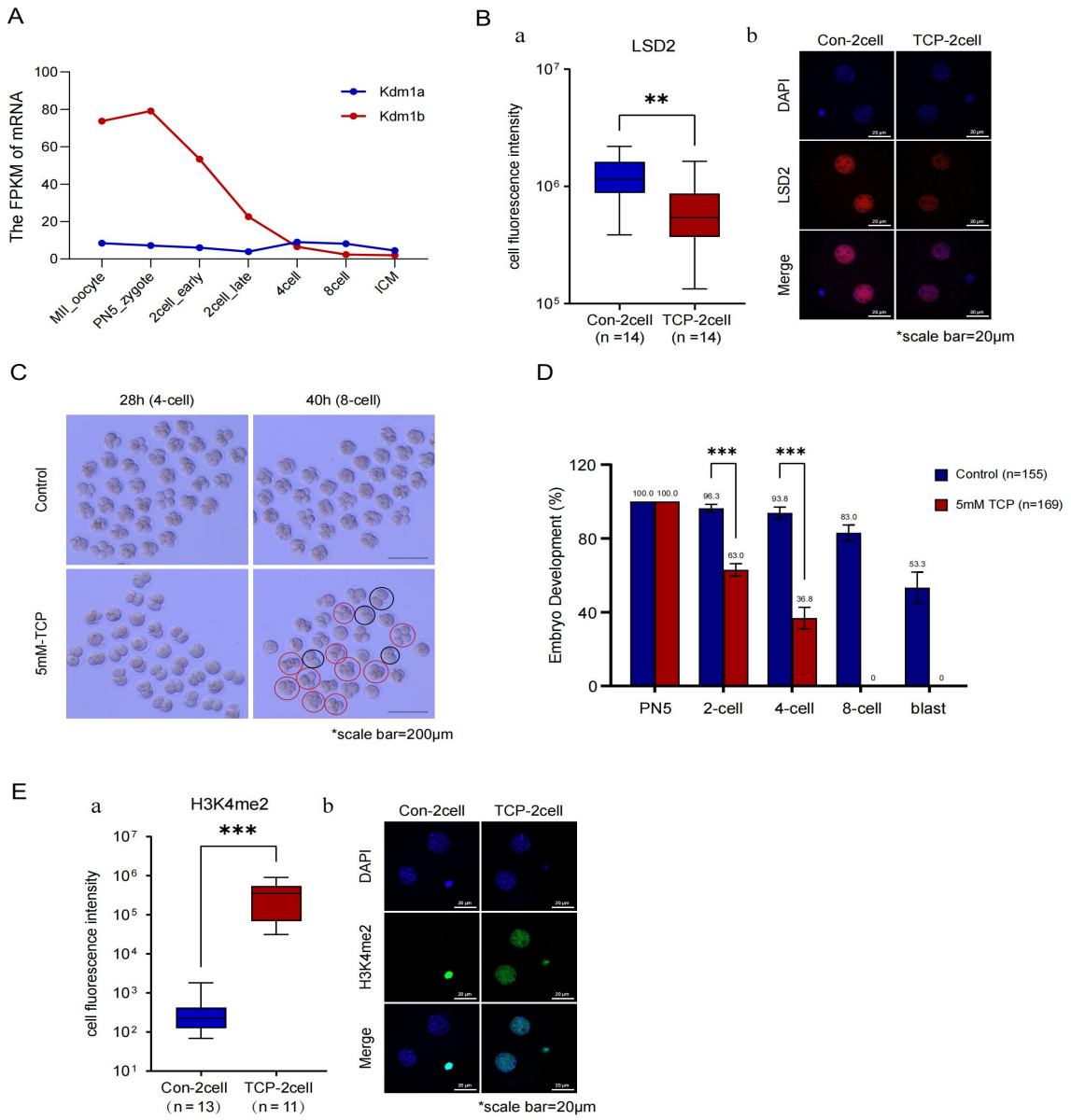


Figure 4. The inhibition of LSD2 resulted in abnormal embryonic development

(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

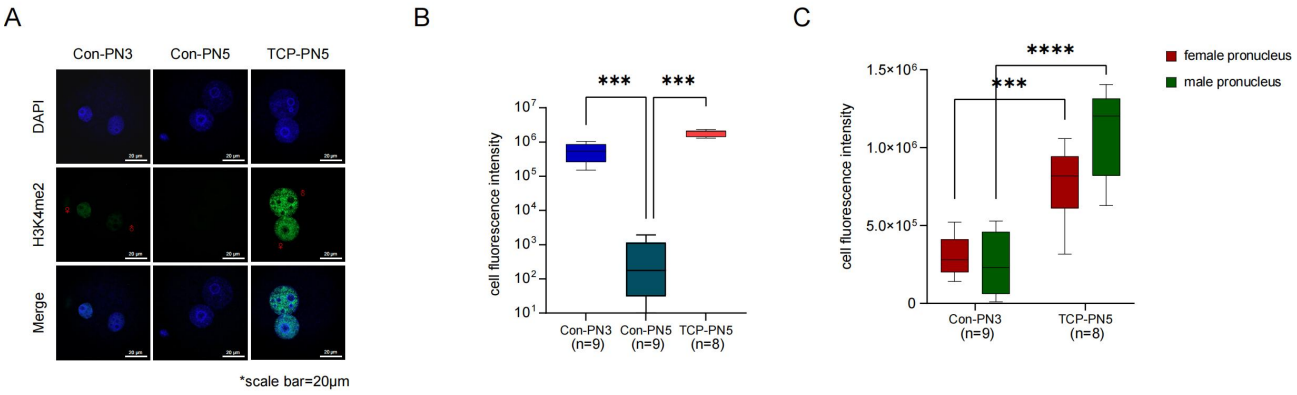


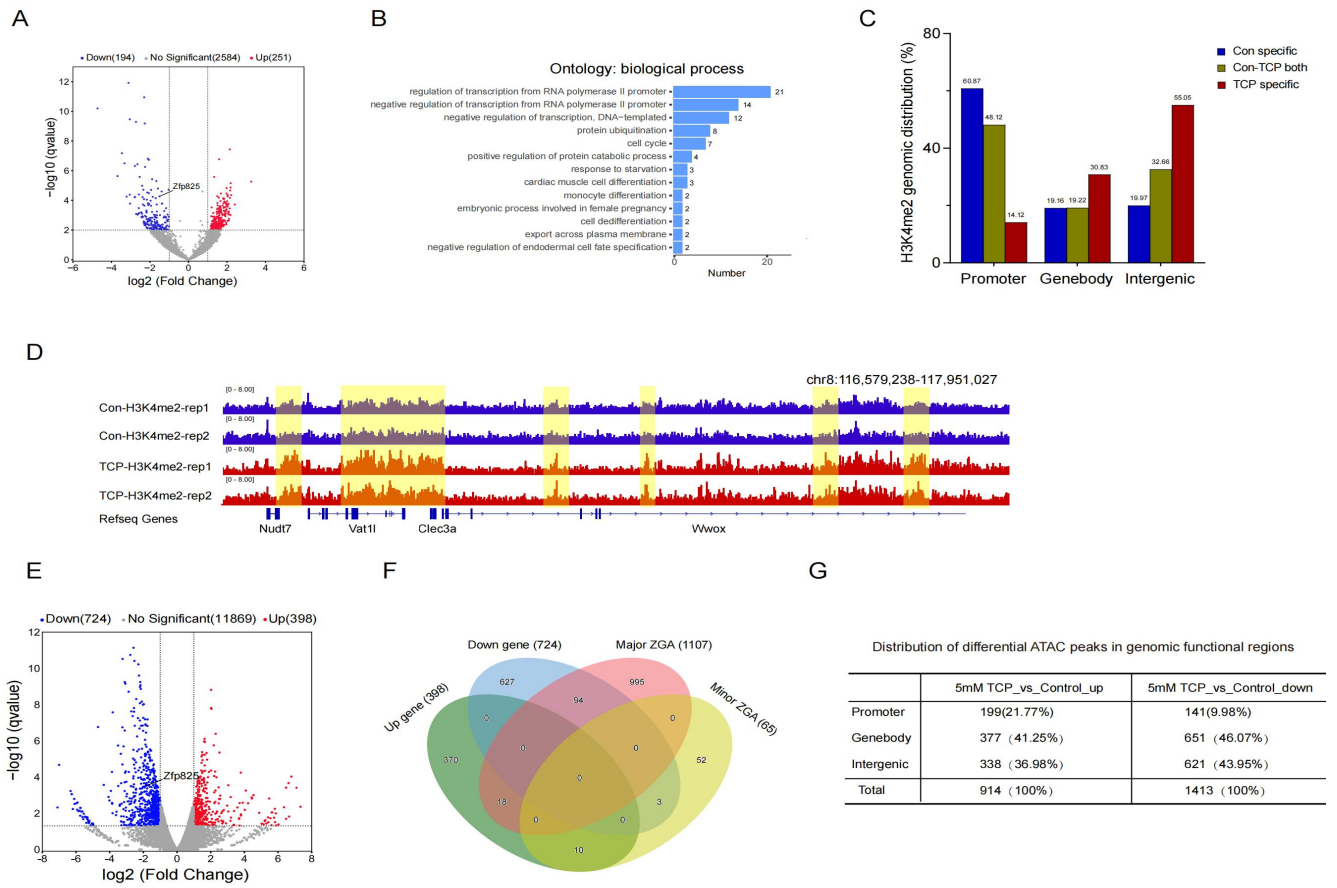
Figure 5. Parental H3K4me2 was erased by maternal LSD2 during sperm-egg fusion

(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

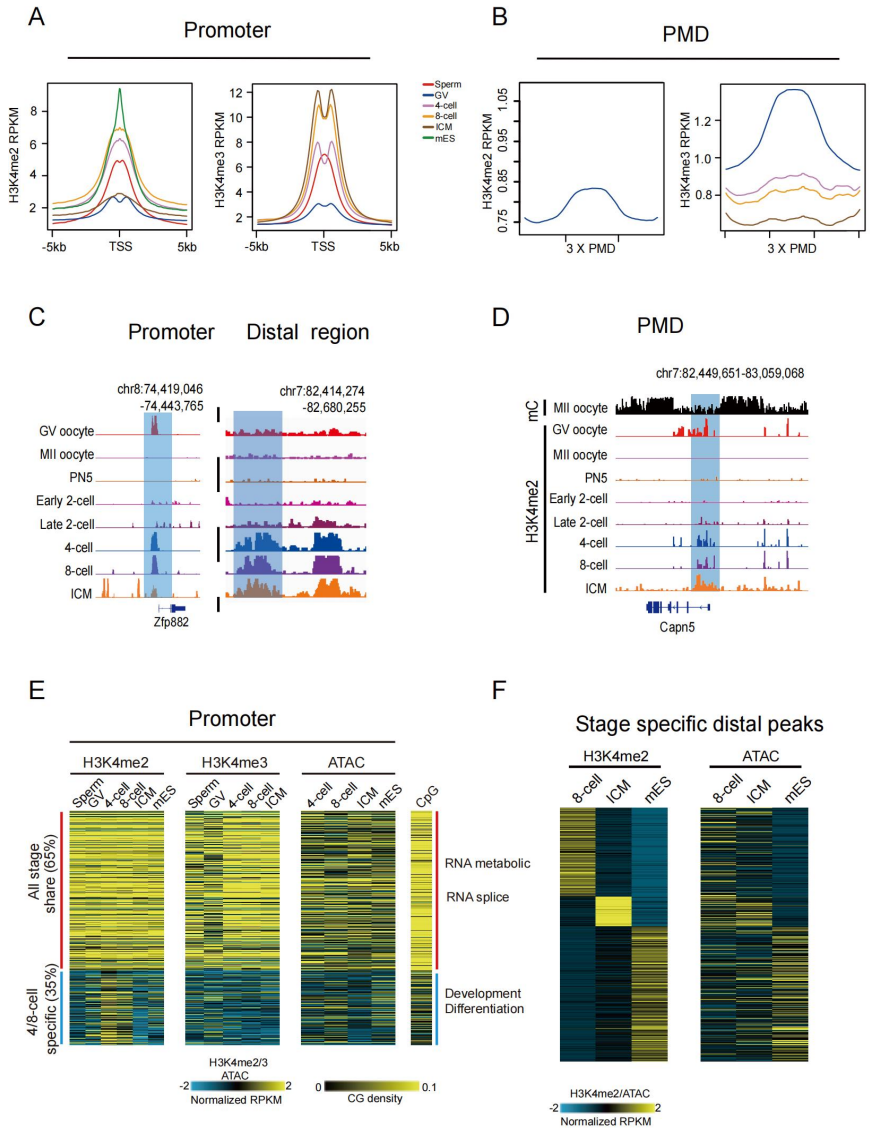


Figure 7. Pervasive H3K4me2 in CpG-rich regulatory regions in mouse embryos.

(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.

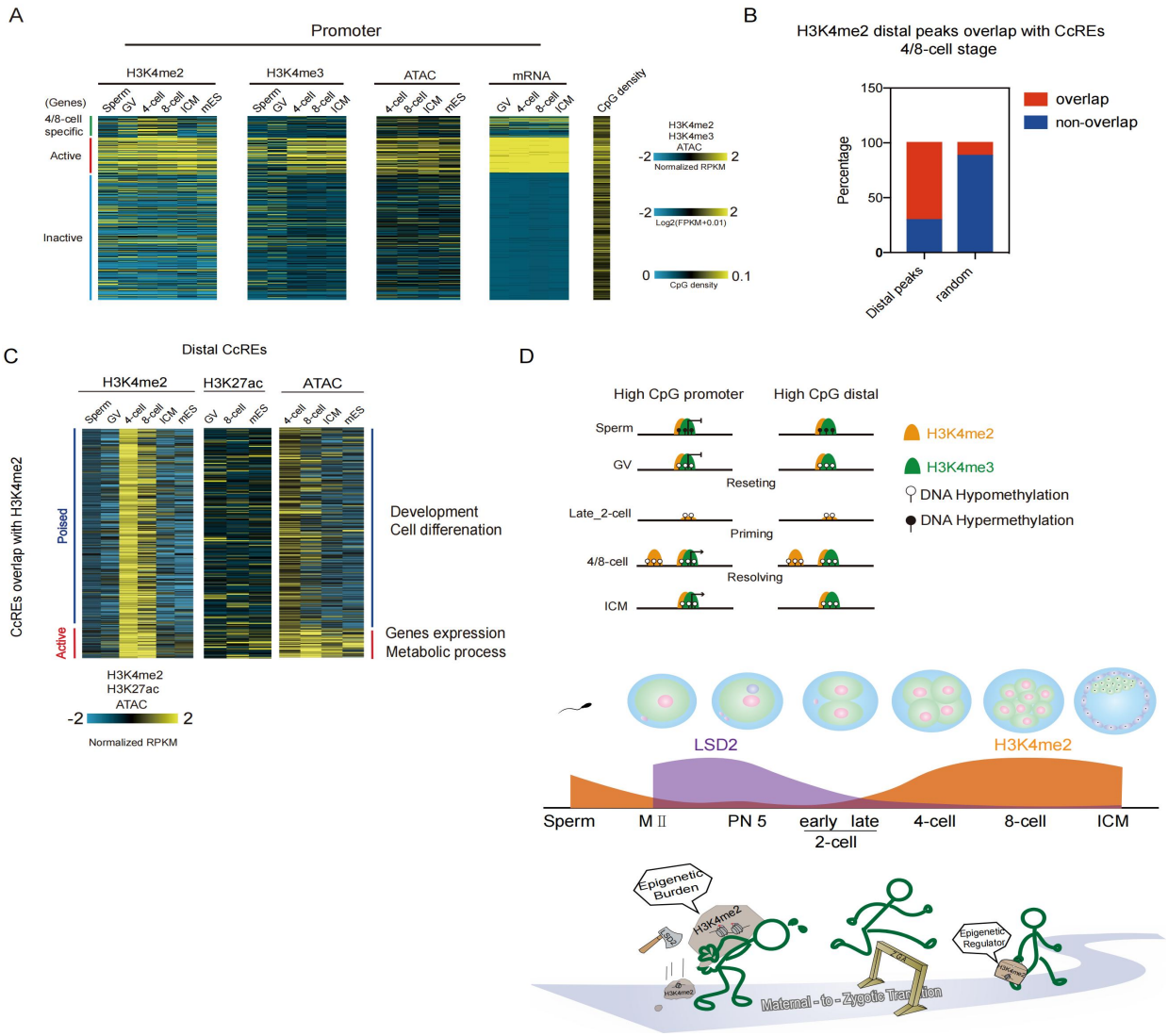


Figure 8. Chromatin states at specific stages and the H3K4me2 mode chart during mouse maternal-to-zygotic transition.

(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.

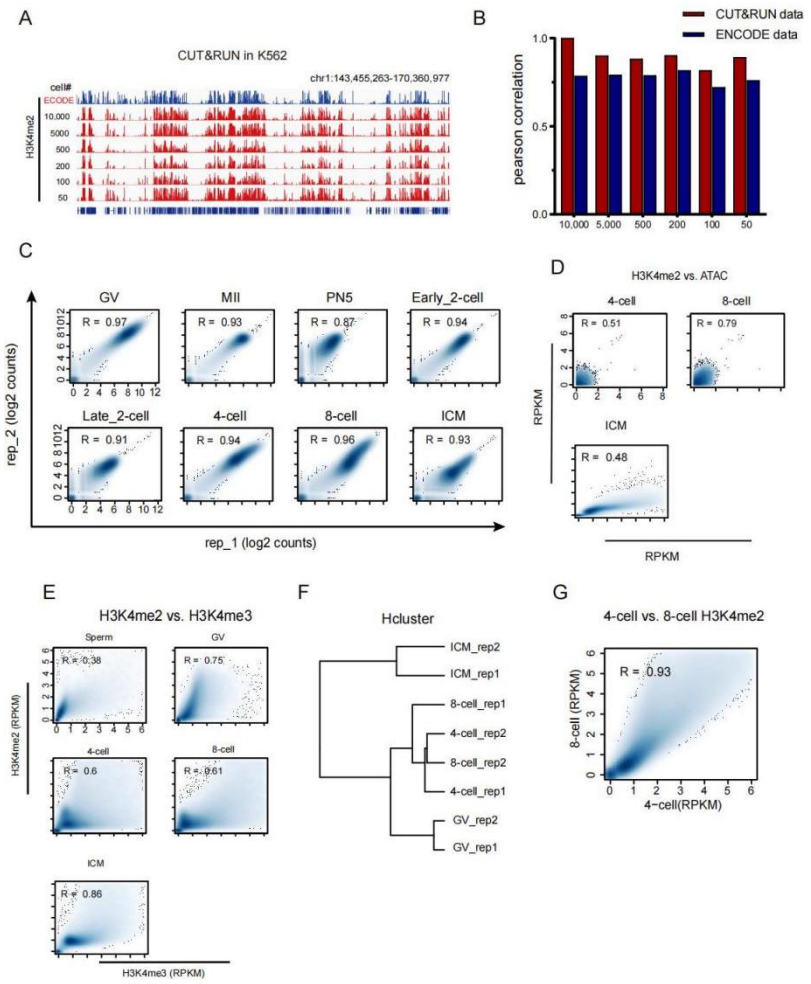
(B) The bar graph shows that the H3K4me2 distal peaks overlapped with the CcREs four/eight-cell stage. Random peaks are shown as a control.

(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.

Extended Data



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

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

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

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

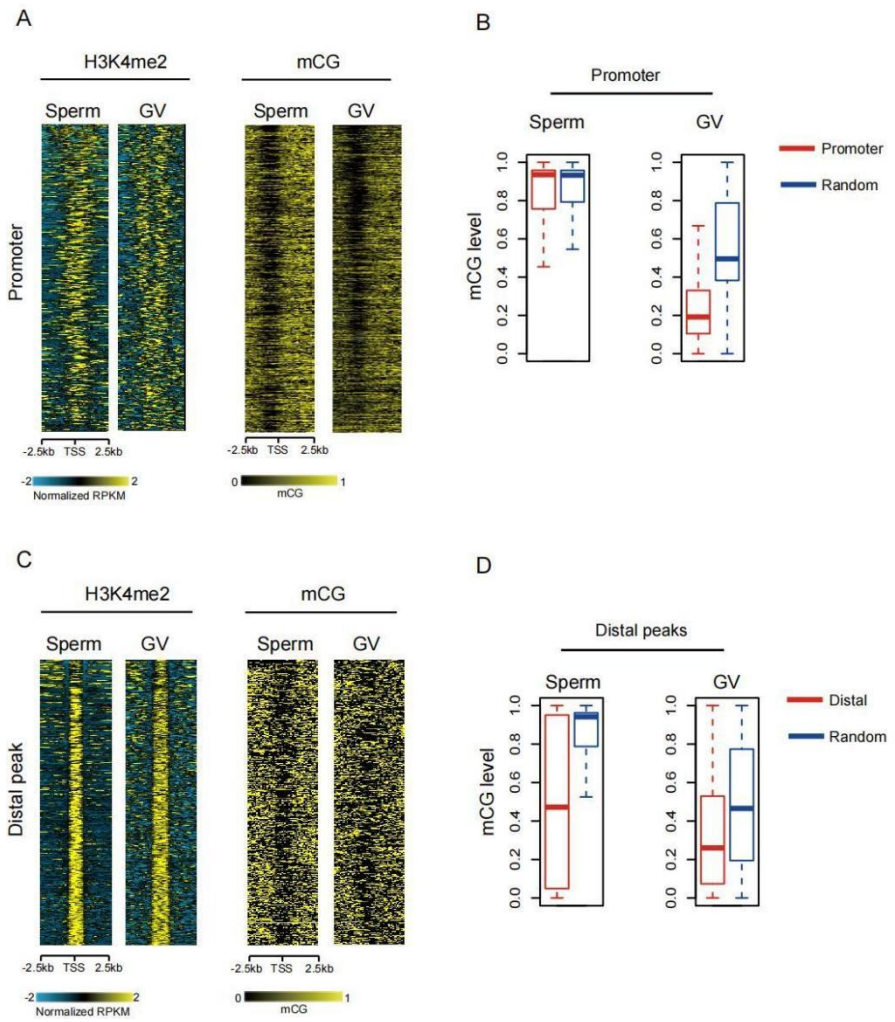
(D) Scatter plots comparing the H3K4me2 signals with ATAC-seq (Assay for

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.



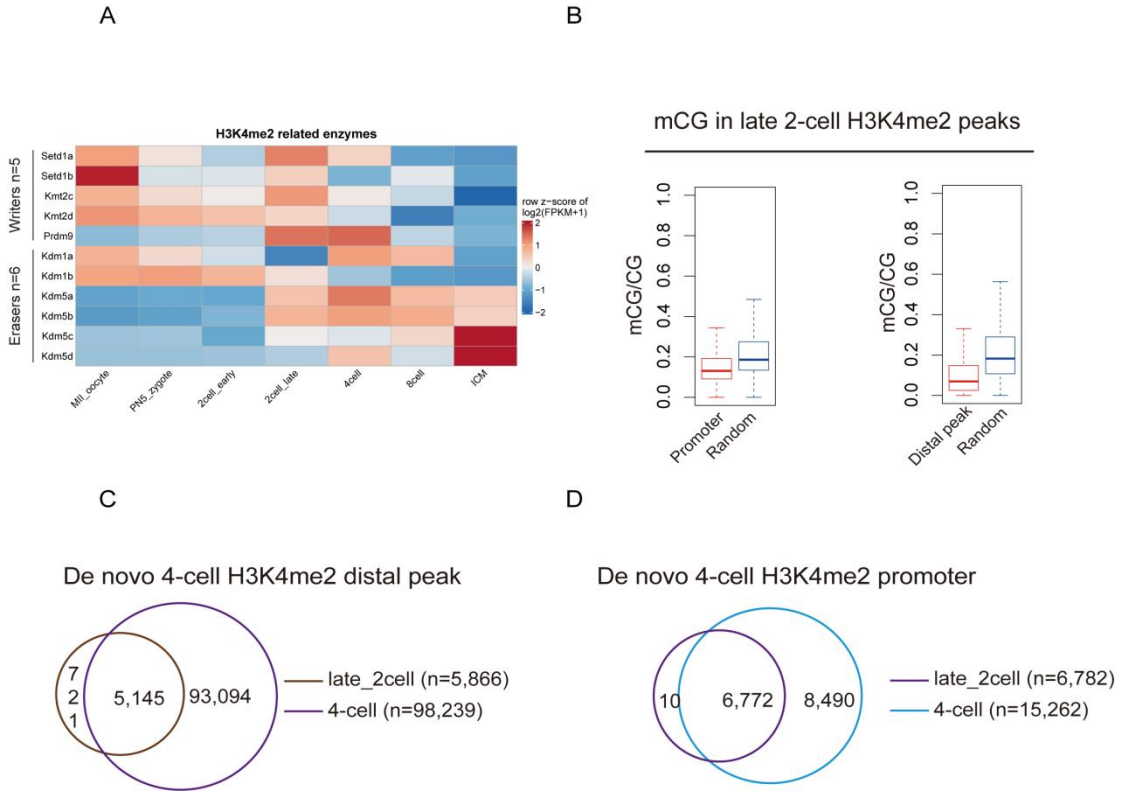
Extended Data Figure 2. H3K4me2 in mouse gametes.

(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.

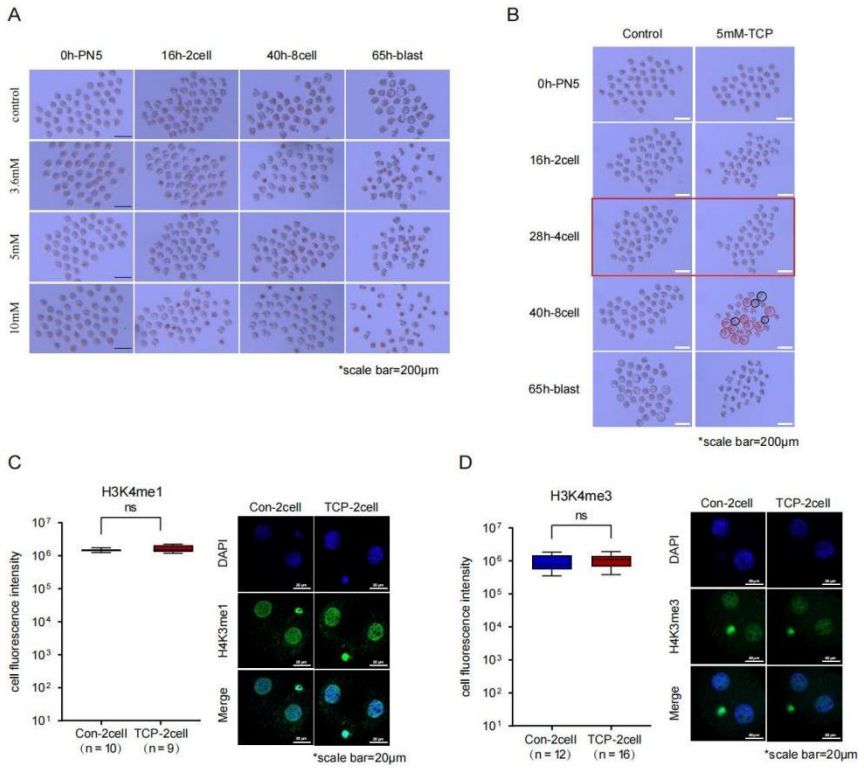


Extended Data Figure 3. H3K4me2 in late 2-cell embryos.

(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.

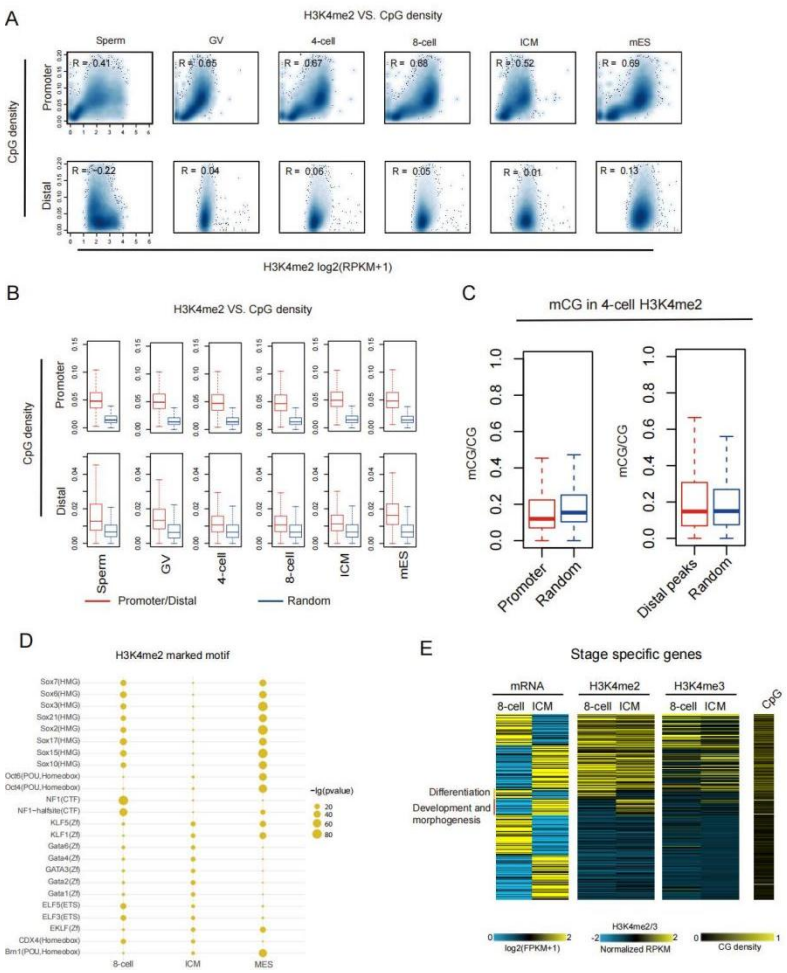


Extended Data Figure 4. Cell culture experiment under TCP treatment

(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.



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

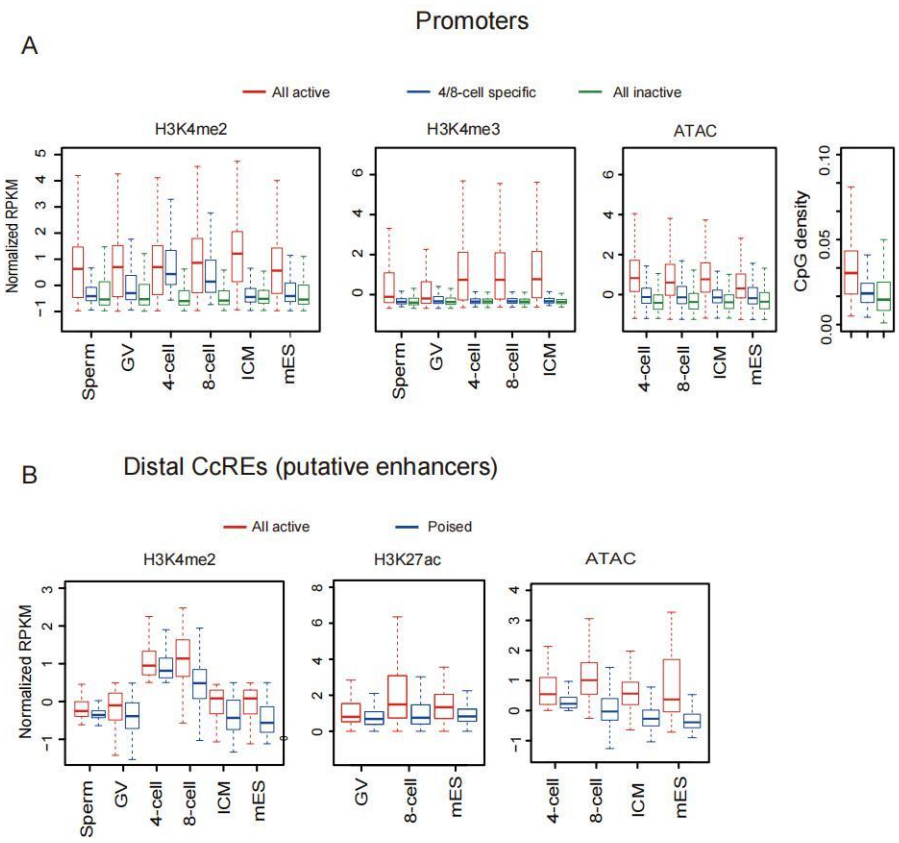
(A) Scatter plots comparing the promoter and distal H3K4me2 signals with CpG densities in mouse gametes, 4-cell, 8-cell, ICM and mESC. The Spearman correlation coefficients are also shown.

(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.



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.

Extended Data Table

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

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

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

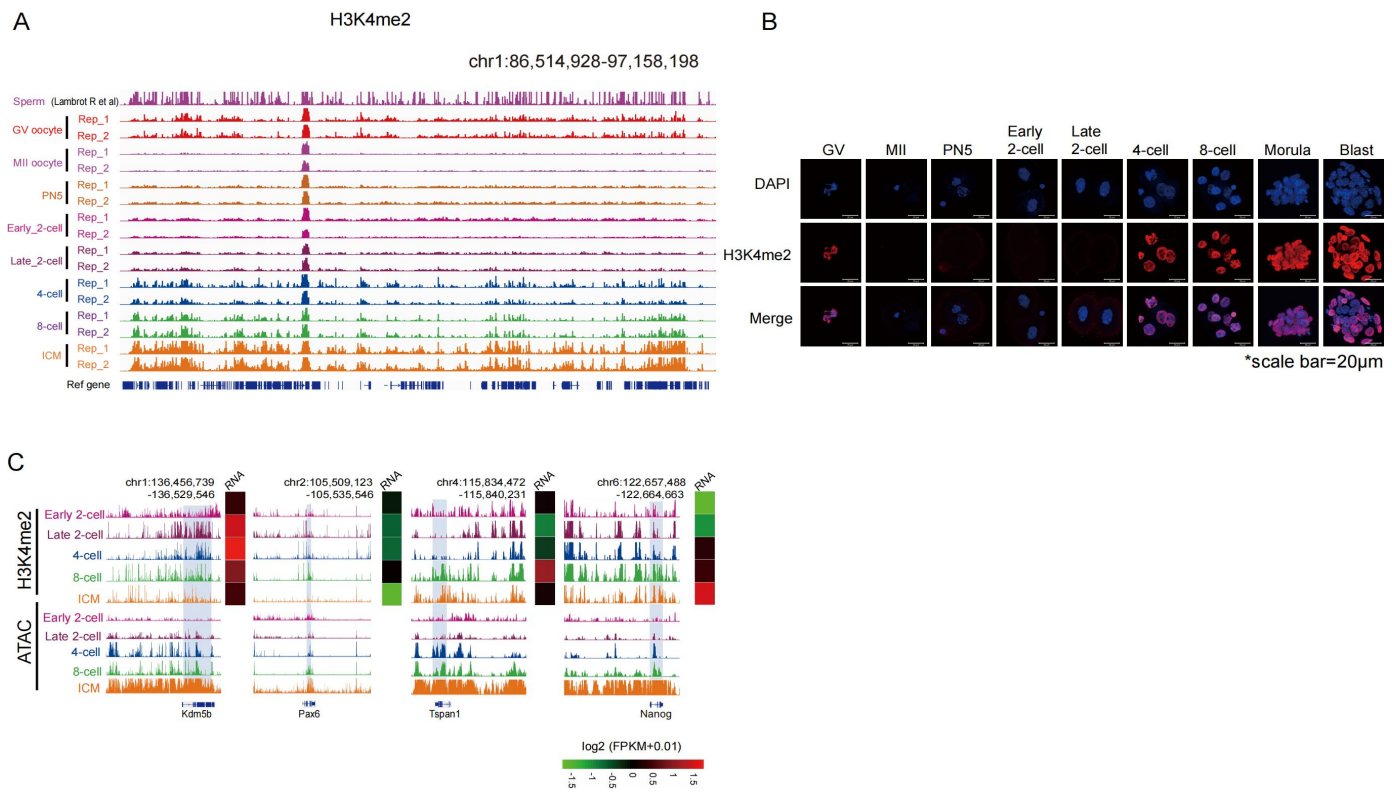


Figure 1. Mapping H3K4me2 in mouse (C57BL6/N) gametes and pre-implantation embryos.

(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

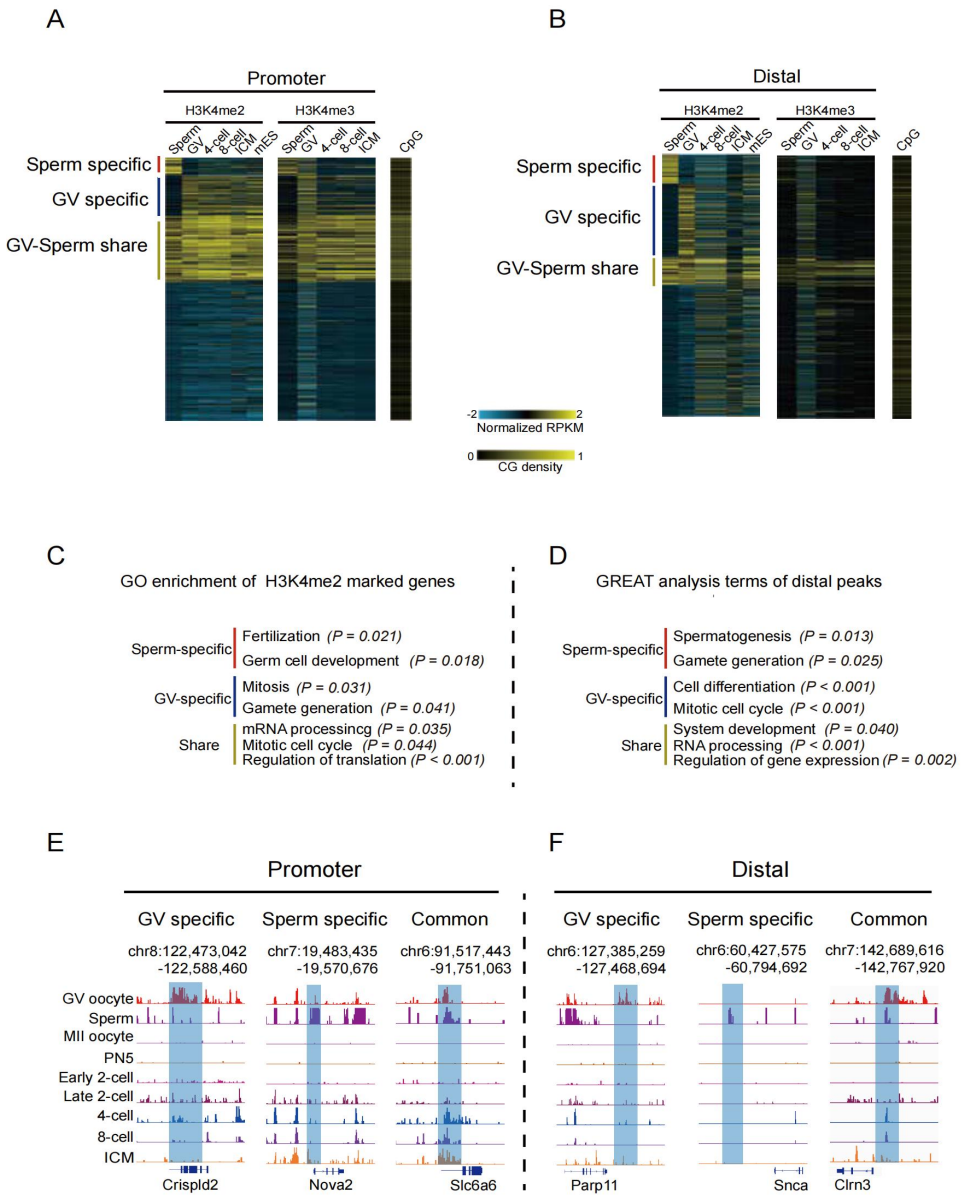


Figure 2. Comparison of H3K4me2 between sperms and oocytes.

(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).

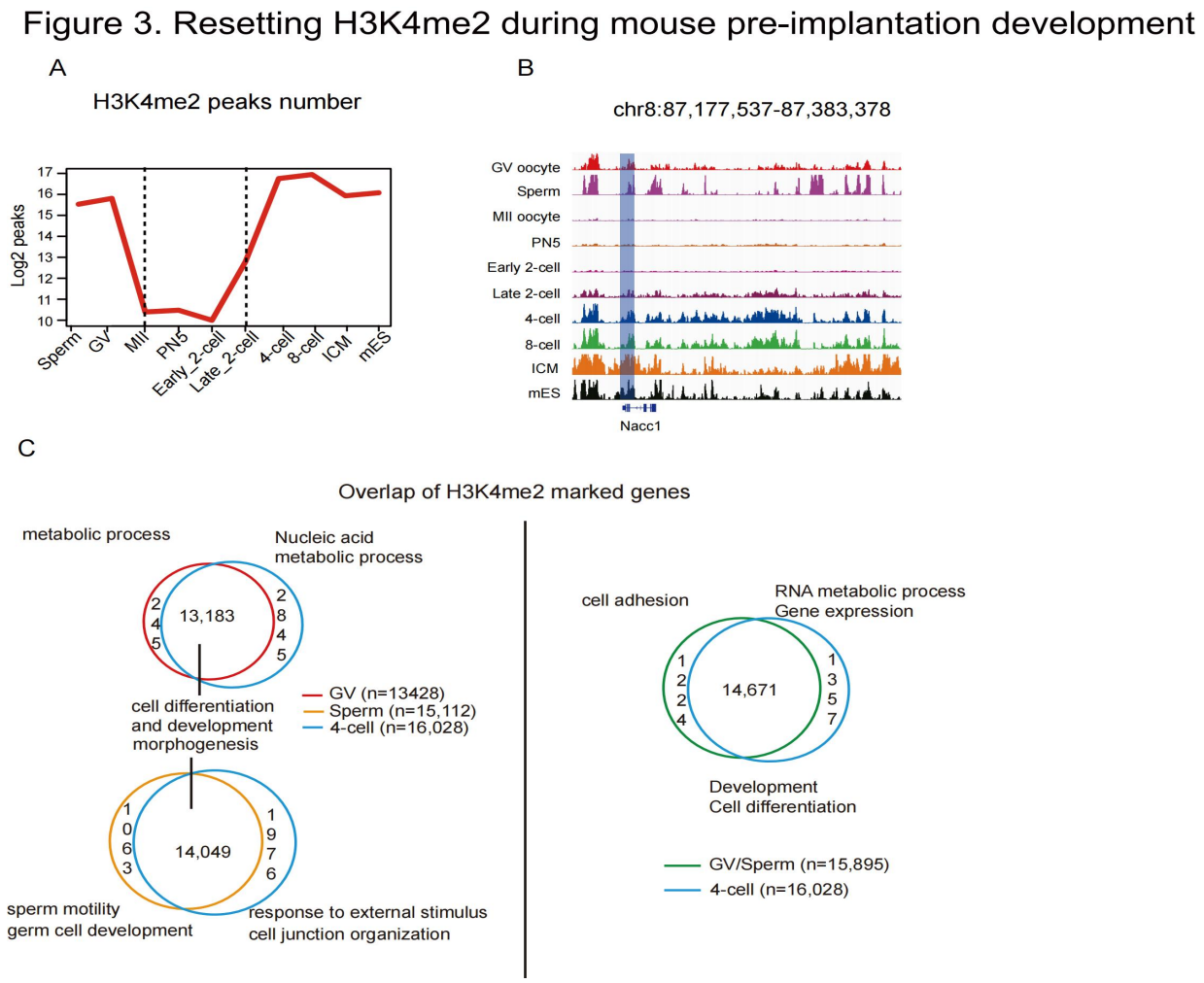


Figure 3. Resetting of H3K4me2 during mouse pre-implantation development.

(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

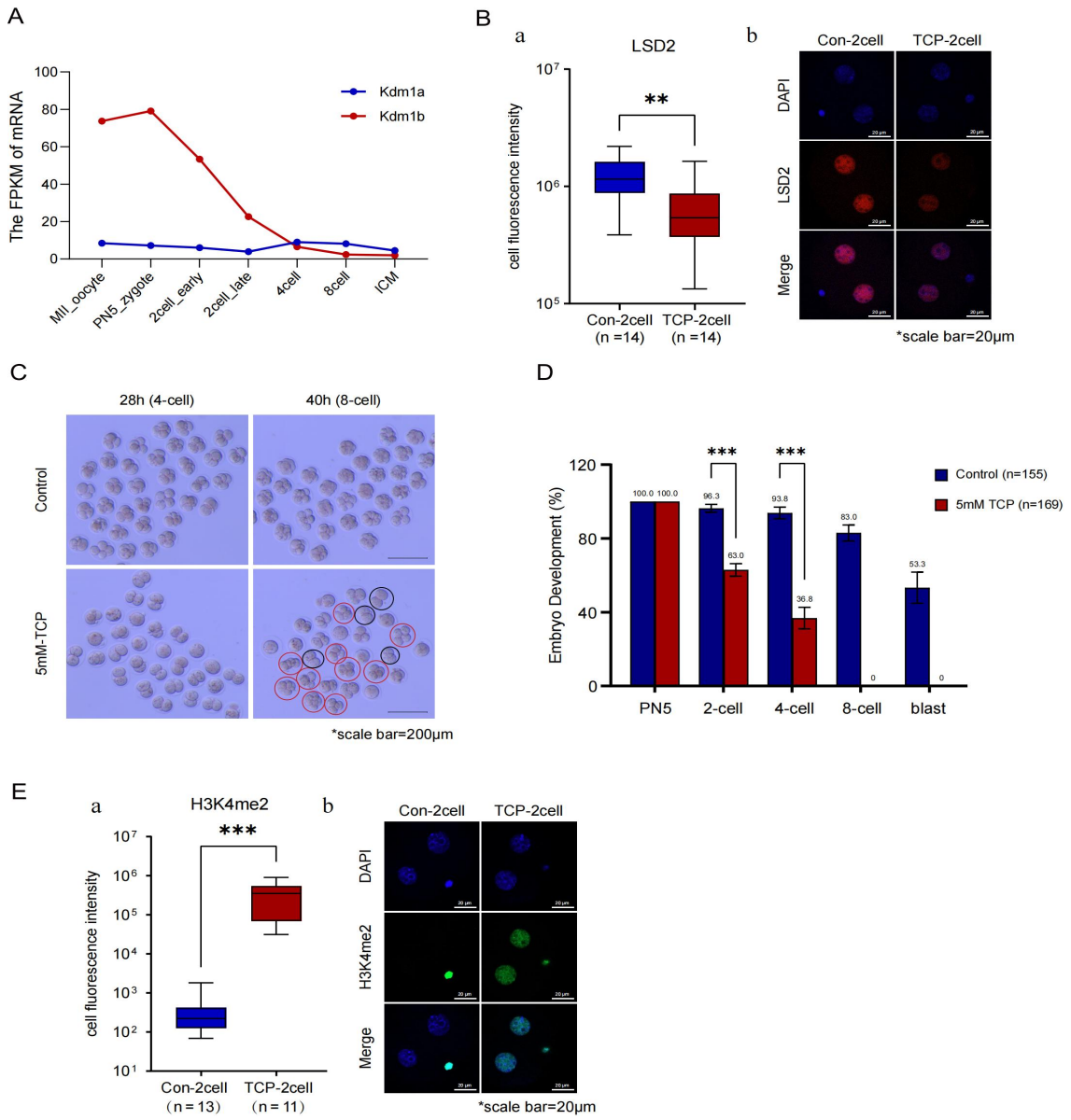


Figure 4. The inhibition of LSD2 resulted in abnormal embryonic development

(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

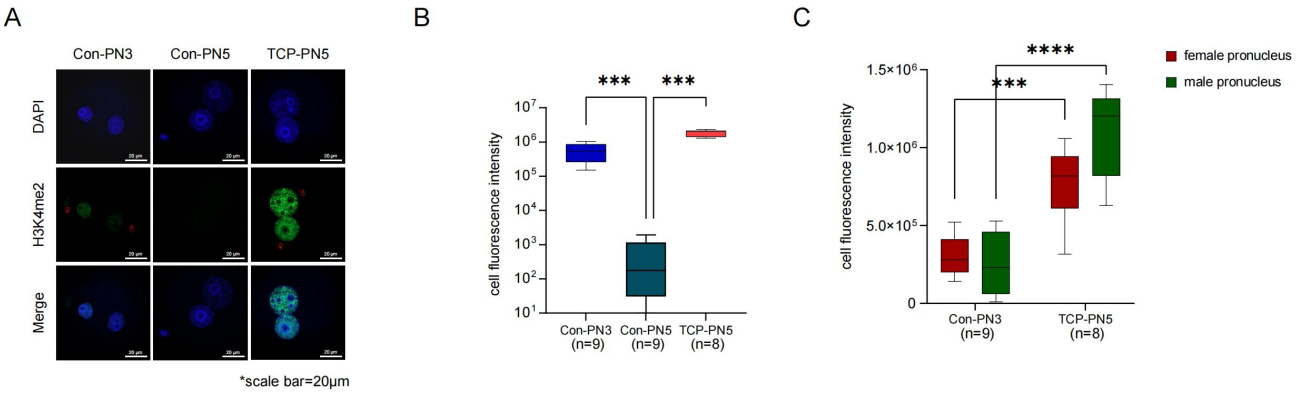


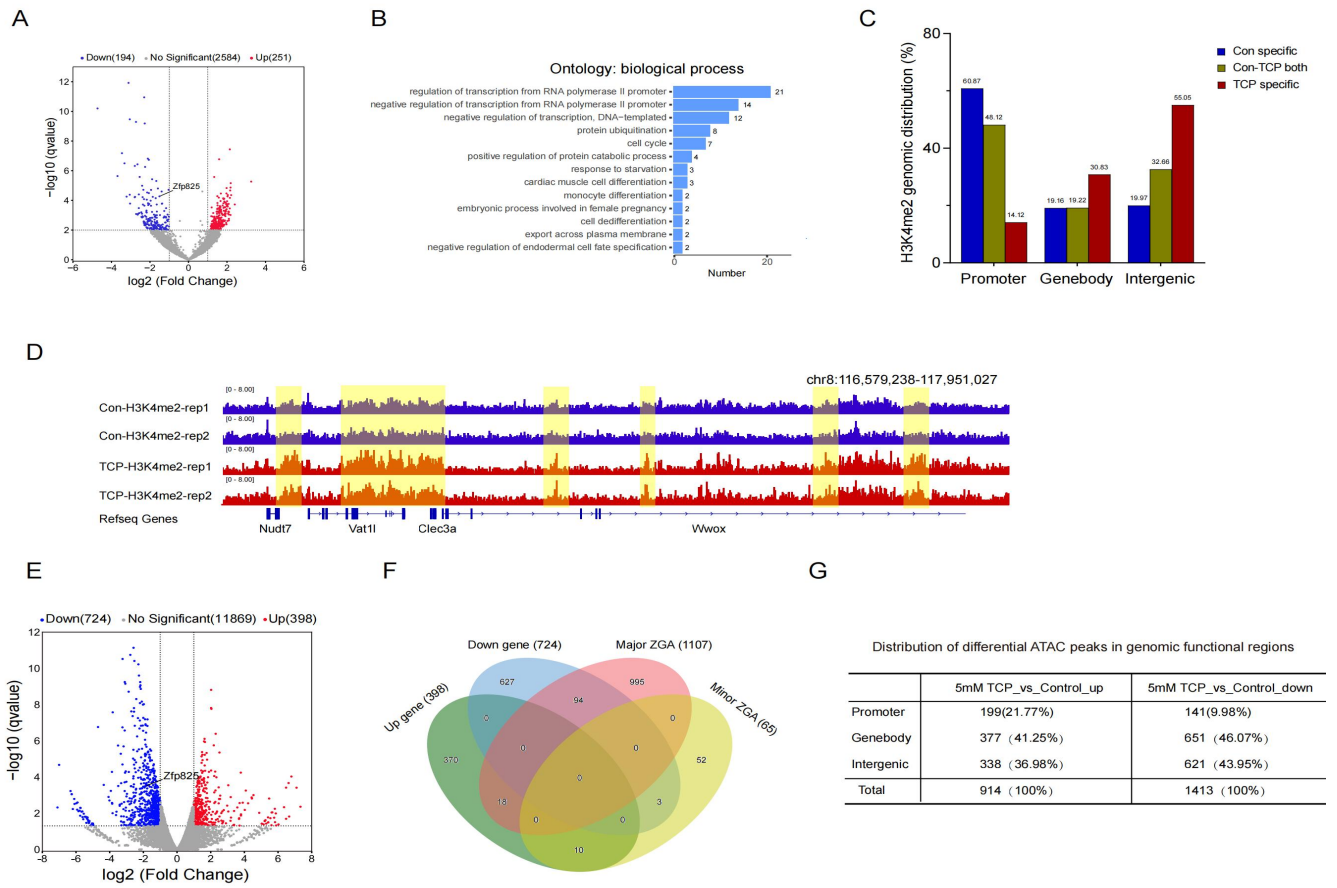
Figure 5. Parental H3K4me2 was erased by maternal LSD2 during sperm-egg fusion

(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

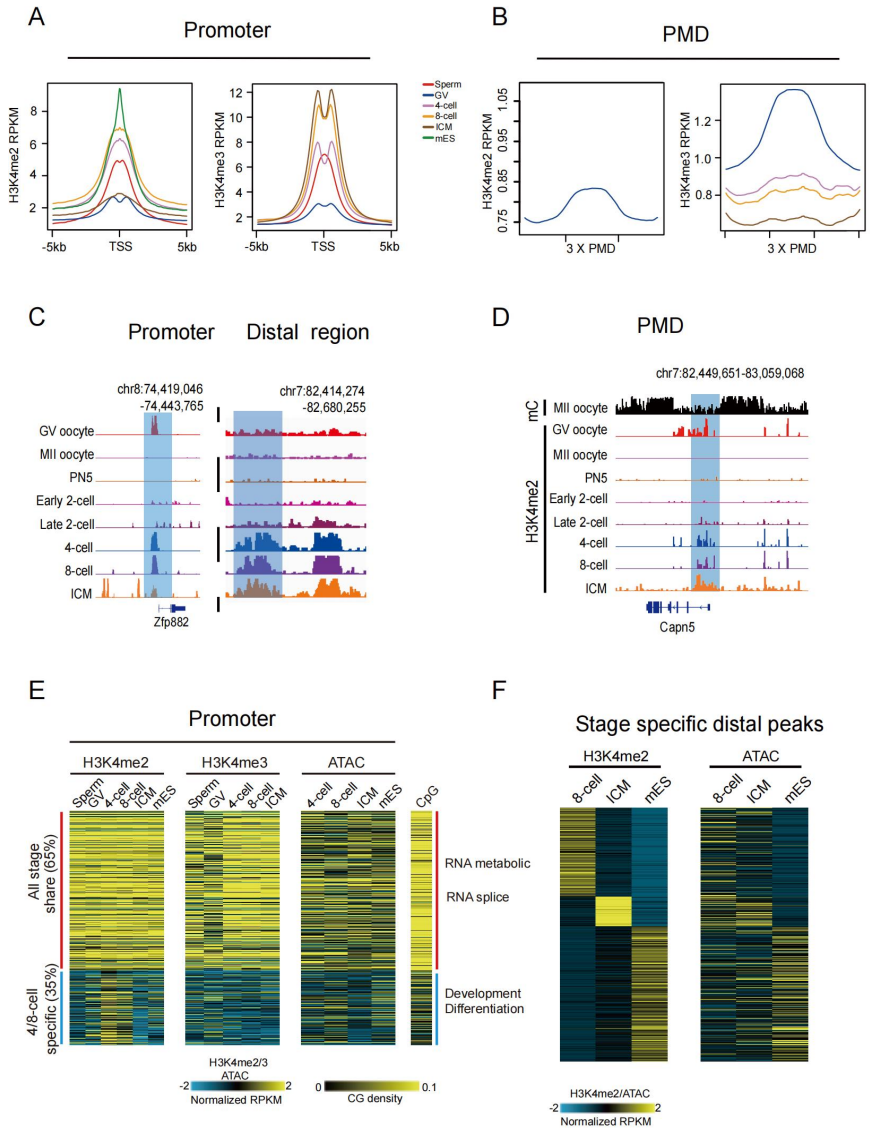


Figure 7. Pervasive H3K4me2 in CpG-rich regulatory regions in mouse embryos.

(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.

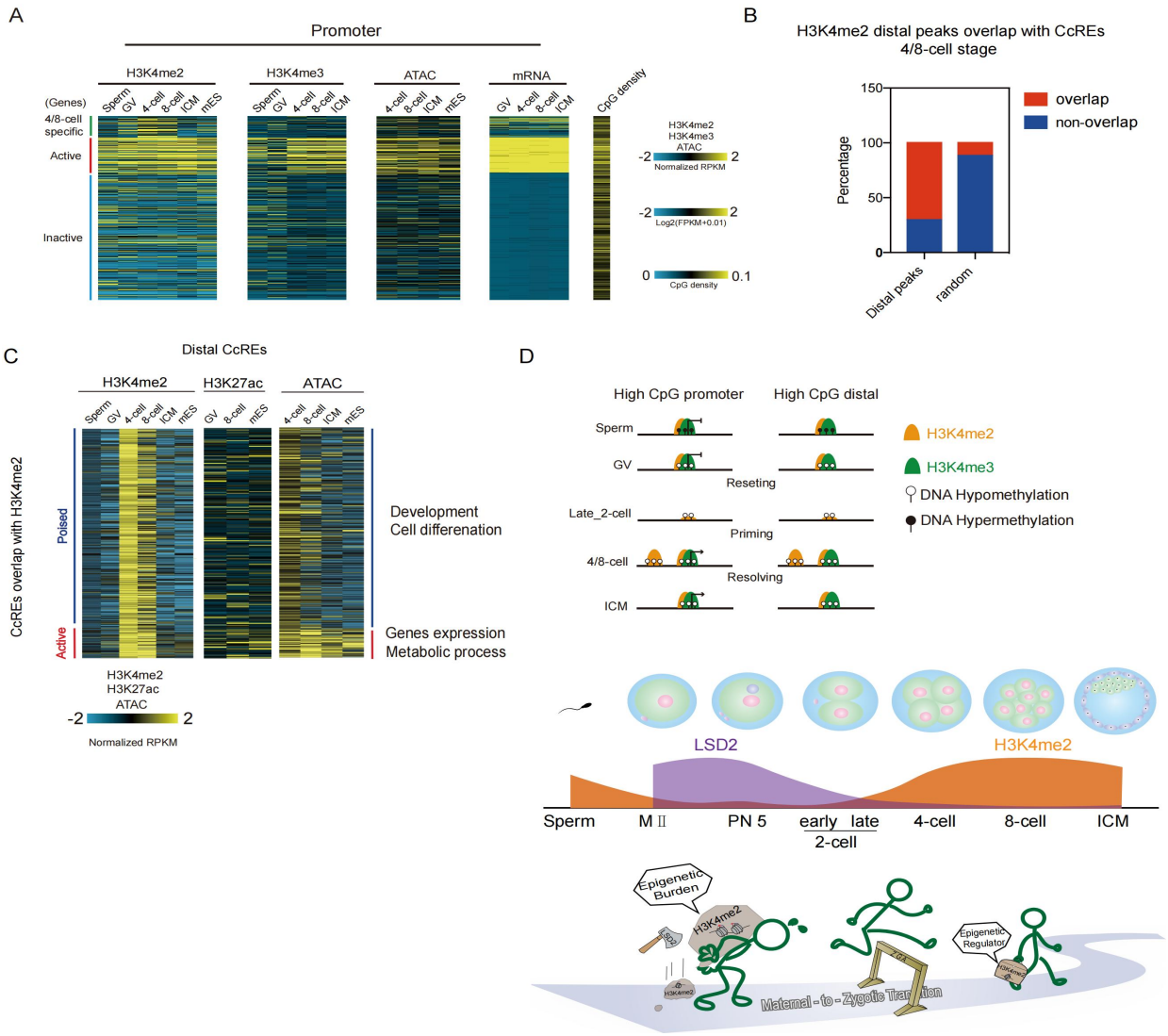


Figure 8. Chromatin states at specific stages and the H3K4me2 mode chart during mouse maternal-to-zygotic transition.

(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.