1 **Sex-specific chromatin remodelling safeguards transcription in germ cells**

2

3 **Tien-Chi Huang[1,2], Yi-Fang Wang[1*], Eric Vazquez-Ferrer[1,2*], Cristina E. Requena[1,2], Courtney**
4 **Hanna[3,4], Gavin Kelsey[,3,4] and Petra Hajkova[1,2 †]**

5

6 [1]MRC London Institute of Medical Sciences (LMS), Du Cane Road, London W12 0NN, UK.

7 [2]Institute of Clinical Sciences (ICS), Faculty of Medicine, Imperial College London, Du Cane
8 Road, London, W12 0NN, UK

9 [3]Epigenetics Programme, Babraham Institute, Cambridge, UK

10 [4]Centre for Trophoblast Research, University of Cambridge, Cambridge, UK

11

12 *these authors contributed equally

13

14 [†]Correspondence: petra.hajkova@lms.mrc.ac.uk

15 Tel: +44 (0)20 83838264

16

17

18

19

20 **Abstract:**

21 Stability of the epigenetic landscape underpins maintenance of the cell type specific transcriptional

22 profile. DNA methylation as one of the main repressive epigenetic systems has been shown to be

23 important for long term gene silencing; its loss leading to ectopic and aberrant transcription in

24 differentiated cells and cancer[1]. Interestingly, the developing mouse germ line endures global

25 changes in DNA methylation in the absence of widespread transcriptional activation. Using an ultra-

26 low input native (ULI-n) ChIP approach we show that following DNA demethylation the gonadal

27 primordial germ cells (PGCs) undergo remodelling of repressive histone modifications resulting in a

28 sex specific signature. We further demonstrate that polycomb plays a central role in transcriptional

29 control in the newly hypomethylated germline genome as the genetic loss of *Ezh2* leads to aberrant

30 transcriptional activation, retrotransposon derepression and dramatic loss of developing female

31 germ cells. Last but not least, we show that the base composition of promoters determines

32 H3K27me3 enrichment following the loss of DNA methylation. Overall, our study provides

33 unprecedented insight into the dynamic interplay between repressive chromatin modifications in

34 the context of a developmental reprogramming system.

35

36

37   DNA methylation is functionally associated with transcriptional silencing of tissue-specific genes and

38   repression of transposable elements (TEs). Although globally relatively stable in the soma,

39   development of the mammalian germ line is associated with profound changes in genomic DNA

40   methylation [2-4]. Following the specification of nascent primordial germ cells (PGCs) at embryonic day

41   E6.25 in the mouse, global changes in the germline epigenome encompass both reprogramming of

42   histone modifications as well as erosion of genome-wide DNA methylation [5,6]. Global DNA

43   methylation drops gradually during germ cell migration and further decreases extensively between

44   E10.5 and E12.5 once PGCs have entered the genital ridge, installing the most hypomethylated state

45   ever observed during normal development [7]. The gonadal reprogramming directly precedes sex

46   differentiation followed by the onset of meiosis in female PGCs and proliferation arrest in the male

47   germ line [8] (Fig. 1a).

48   Several lines of evidence indicate that the loss of DNA methylation and TET1-mediated

49   transcriptional activation are crucial for the initiation of the expression of key germline genes at this

50   developmental stage [9]. However, the molecular mechanisms that enable germ cells to maintain and

51   co-ordinate the progressive gene activation and prevent precocious or aberrant wide-spread

52   transcription that could be triggered by genome-wide DNA demethylation are unclear.

53   We have previously shown that the global loss of DNA methylation in the gonadal PGCs temporally

54   overlaps with other major changes in chromatin structure and global histone modifications [6]. In

55   order to understand the temporal changes and to address whether remodeling of the

56   heterochromatin silencing could functionally compensate for the global loss of DNA methylation to

57   provide the means of transcriptional control at this critical stage of PGC development we assessed

58   H3K27me3 and H3K9me3 using an ultra-low input native chromatin immunoprecipitation (ULI-

59   nChIP-seq) [10,11]. The analysis was carried out using 1000 PGCs isolated from E10.5 (pre-gonadal

60   reprogramming) and E13.5 (post-gonadal reprogramming) male and female embryos derived from

61   C57BL/6 X ΔPE-Oct4-GFP crosses [12] (Extended Data Fig.1a, 1b,1c). Similar to previous reports in ES

62   cells [13], more than 50% of H3K27me3 peaks overlapped with promoters and gene body (Fig. 1b) with

63   30% of H3K27me3 peaks located within 1kb of the transcription start sites (TSS) (Extended Data Fig.

64   1d, 1e). In contrast, 75% of H3K9me3 peaks localised to intergenic regions, remote from the

65   annotated TSSs (Fig. 1b, Extended Data Fig. 1d, 1e) with clear peak enrichment at repetitive

66   sequences (58% at E10.5, 64% in E13.5 female and 68% for E13.5 male, respectively) (Extended Data

67   Fig 1f).

68

69 Our ULI-nChIP Seq analysis revealed relatively low correlation between E13.5 male and female PGCs

70 for both H3K27me3 and H3K9me3 documenting sex specific differences at this stage of germline

71 development (Spearman correlation coefficient 0.43 for H3K27me3, 0.5 for H3K9me3) (Fig 1c).

72 Furthermore, for both H3K9me3 and H3K27me3 the correlation between E10.5 and E13.5 PGC

73 samples is even lower than it is between sexes suggesting that the repressive chromatin is globally

74 remodelled between E10.5 and E13.5 alongside the global wave of DNA demethylation (Fig. 1c). In

75 further support of this notion, the peak number of H3K27me3 was greatly reduced from E10.5 to

76 E13.5 in both male and female germ cells (Fig. 1d). Although the loss of H3K27me3 peaks is more

77 pronounced in male PGCs (6933 peaks in the female PGCs vs 2740 identified peaks in the male PGCs),

78 even female cells had only 40% of the peaks persisting from E10.5 with the majority of peaks being

79 newly generated during the reprogramming process (Fig. 1d). Sex specific re-configuration of the

80 signal is also evident upon analysis of the H3K9me3 signal. Despite the comparable number of

81 enriched loci, H3K9me3 seems to be remodelled in female PGCs between E10.5 and E13.5; while the

82 male specific loss of H3K27me3 peaks described above seems to be compensated by a dramatic (2.7

83 fold) increase in the number of H3K9me3 peaks observed in male E13.5 PGCs (Fig. 1e). Of note, the

84 majority of male H3K9me3 peaks are relatively short, compared with peaks observed at E10.5

85 (Extended Data Fig. 1h). Despite a significant change in peaks number, the average ChIP intensity

86 relative to input within peaks are comparable between stages (Extended Data Fig. 1g). Furthermore,

87 the observed sex specific remodelling of heterochromatin is not underpinned by profound changes

88 in the abundance of relevant histone modifications, or differential expression of the relevant histone

89 methyltransferases (Fig. 1f, Supplementary Table 3); but potentially linked to a difference in

90 targeting of histone modifying complexes. Indeed, we observed sex specific expression of the

91 polycomb repressive complex 2 (PRC2) auxiliary subunits AEBP2 and PHF19, known to be implicated

92 in polycomb targeting (Extended Fig. 6a, 6b)[14,15] . Collectively, our analysis shows that global loss of

93 DNA methylation in gonadal germ cells coincides with genome-wide reprograming of repressive

94 histone modifications, resulting in a sex specific heterochromatin configuration in PGCs at E13.5.

95

96 H3K27me3 is catalysed by PRC2 that shows clear preference for binding to unmethylated CG-rich

97 sequences in mammalian cells[16] . We thus asked whether the H3K27me3 remodelling observed in

98 PGCs is due to PRC2 recruitment in response to the global loss of DNA methylation. Our whole

99 genome bisulphite sequencing (WGBS) datasets [9] show that promoter methylation (5mC+5hmC)

100 significantly decreases from E10.5 to E12.5 in both sexes (Extended Data Fig. 2a). However, this does

101 not lead to an overall global increase in H3K27me3 deposition (Extended Data Fig. 2b, 2c).

102   To understand the potential relationship better, we categorised promoters based on the pattern of

103   dynamic changes between DNA methylation and H3K27me3 from E10.5 to E13.5 (See Methods). We

104   observed 4 different patterns: Group A: loss of 5mC, gain of H3K27me3; Group B: loss of 5mC,

105   continuous H3K27me3 enrichment; Group C: devoid of 5mC, continuous high level of H3K27me3;

106   Group D: loss of 5mC, devoid of H3K27me3 (Fig. 2a, Extended Data Fig 3). Of note, all promoters

107   showed general H3K9me3 reduction between E10.5 and E13.5 alongside global DNA demethylation

108   (Extended Data Fig 3a). Promoters showing consistently high H3K27me3 enrichment (Group C) are

109   characterised by high CpG density, high C+G content and lack of DNA methylation at E10.5 (Fig. 2b),

110   consistent with the known characteristics of polycomb targets in mammalian cells[16]. Promoters with

111   lower CpG density contain DNA methylation at E10.5 (Groups A, B and D); however, following

112   methylation loss, CpG density and G+C content determine the polycomb recruitment. As an

113   example, both group A and group D lose DNA methylation, but compared with Group A that attracts

114   H3K27me3 (CpG obs/exp ratio=0.27, GC percentage =46.6%), Group D has significantly lower

115   frequency of CpG (CpG obs/exp ratio=0.22) and G+C (44.8%) (P<2X10-16, pairwise Wilcox test) (Fig.

116   2b). This suggests that critical threshold of CpG density and G+C enrichment is required to render

117   promoters capable of recruiting PRC2 once DNA methylation is removed (Fig. 2c). Interestingly, this

118   relationship is less obvious in the male E13.5 PGCs (Extended Data Fig. 2d, 2e), suggesting a key role

119   for PRC2 subunits that are differentially expressed between sexes in the targeting of polycomb.

120   Promoters in group B have medium to low levels of DNA methylation at E10.5 and are enriched for

121   H3K27me3 even before gonadal DNA demethylation. This group thus likely represents an

122   intermediate state. Similar to group A, group B has higher CpG density and higher GC content,

123   compared with group D (P<2X10-16, pairwise Wilcox test), but higher GC enrichment than group A.

124   Interestingly, promoters in group A and B are enriched for germline-specific genes, such as *Stra8*,

125   *Sycp1, Sycp2, Dnmt3l* and *Tex101* (Extended Data Fig. 3a), whereas promoters in the group C

126   generally are linked to development and lineage-specific regulators (Extended Data Fig. 3a).

127   In order to functionally validate the importance of the observed dynamic changes in H3K27me3, we

128   depleted H3K27me3 in early PGCs using *Blimp1-Cre* (*Prdm1*) driven *Ezh2* deletion [17]. (Extended Data

129   Fig. 4b, 4c, 4d). EZH2 is highly expressed throughout the early germline development, compared

130   with surrounding somatic cells (Extended Data Fig. 4a). The expression of *Blimp1*-Cre begins at E7.0,

131   shortly after PGC specification, allowing us to deplete EZH2 before gonadal reprogramming occurs[17].

132   Consistent with the reported high recombination efficiency of *Blimp1-Cre*[18], EZH2 protein was absent

133   in gonadal *Ezh2* $^{\Delta/\Delta}$*, Blimp1-Cre* PGCs (Fig. 3a) (*Ezh2* conditional knock out is henceforth referred to

134   as *Ezh2* CKO). This coincided with the loss of H3K27me3 in *Ezh2* CKO PGCs, confirming that EZH2 is

135 the main enzyme responsible for catalysing H3K27me3 at this stage of PGC development (Fig. 3b,

136 Extended Data Fig. 6c). Of note, the loss of H3K27me3 did not affect global levels of H2A119ub or

137 H3K9me3, as evaluated by immunofluorescence (Extended Data Fig. 5a, 5b), or the global erasure of

138 DNA demethylation that occurs in gonadal PGCs (Extended Data Fig. 5c, 5d, 5e).

139 Although the observed loss of H3K27me3 did not have any apparent effect on the germ cell numbers

140 in both sexes at E13.5 (Extended Data Fig. 5f), RNA-Seq analysis showed clear separation by sex and

141 genotype (Ctrl vs CKO $Ezh2^{-/-}$ PGCs) (Fig. 3c, Extended Data Fig. 6d). As anticipated based on our

142 H3K27me3 ChIP-Seq analysis (Fig. 1d), loss of EZH2 had a more profound impact on the

143 transcriptome of female PGCs (Fig. 3c, 3d) with 679 and 290 genes differentially expressed in female

144 and male $Ezh2^{-/-}$ PGCs, respectively (adj. P<0.05 FC>2). Consistent with the role of PRC2 as a

145 transcriptional repressor, more than 80% of differentially expressed genes in male (239) and female

146 (565) were upregulated (Fig. 3d) with clear enrichment for cell differentiation, developmental

147 processes and gamete generation using Gene Ontology (GO) analysis (Fig. 3e, Extended Data Fig. 7a).

148 Importantly, loss of EZH2 affected specifically genes with promoters enriched in H3K27me3 in

149 control E13.5 PGCs (Extended Data Fig. 3a, 3b), documenting that following DNA demethylation,

150 promoter H3K27me3 is required for maintaining correct transcriptional regulation of these genes in

151 the developing germ line.

152 Amongst the genes upregulated in female E13.5 $Ezh2$ CKO PGCs, we noticed an enrichment of genes

153 associated with reproduction and meiosis; out of 104 previously identified genes associated with

154 female meiotic prophase [19], 31 genes were differentially expressed following $Ezh2$ deletion (P=4X10^{-}

155 ^{14}) (Fig. 3f, Extended Data Fig. 7c). Similarly, depletion of EZH2 led to the transcriptional

156 derepression of 23 of the previously identified germline reprogramming-responsive (GRR) genes

157 (Extended Data Fig. 7g) [9]. H3K27me3 increased on the promoters of many meiosis-related genes

158 following the loss of DNA methylation between E10.5 to E13.5 (group A promoters), suggesting that

159 H3K27me3 deposition is required to modulate the expression of germline-specific, DNA methylation-

160 sensitive genes in the hypomethylated germ line (Extended Data Fig. 3a, 3b,Extended Data Fig 7d).

161 Although H3K27me3 enrichment on promoters of meiotic genes is observed in both male and

162 female PGCs (Fig. 3f, Extended Data Fig 7d), female PGCs show a much more profound

163 transcriptional response following EZH2 loss (Extended Data Fig. 7c). We reasoned that the presence

164 of signalling-induced, sex-specific transcription factors might be required to drive ectopic or

165 precocious transcription once the functional repressive chromatin mark is removed. To this end, the

166 presence of retinoic acid (RA) is essential for the transcriptional activation of $Stra8$ and entry into

167 meiosis in the female germline. In contrast, RA is degraded in male gonads leading to diversion from

168   the meiotic programme [20]. In agreement with our hypothesis, we found that RA-related transcription

169   factor binding motifs, such as RARB, RXRA and RARG are enriched in genes upregulated in female,

170   but not in male E13.5 *Ezh2* CKO PGCs. (Extended Data Fig. 7e). Furthermore, as expected, the

171   expression levels of these transcription factors are higher in the female germ cells than in the male

172   counterparts at E13.5 (Extended Data Fig. 7f). Taken together, RA signalling and the related

173   transcription factors present in the female germ line at this stage of development potentiate the

174   transcriptional response of the DNA hypomethylated genome upon removal of EZH2.

175   Although the loss of EZH2 led to precocious transcriptional upregulation of meiotic genes in female

176   *Ezh2* CKO PGCs, we did not observe any profound effect on the initial progression of the meiotic

177   prophase (Extended Data Fig. 6e, 8a, 8b). However, despite normal PGC number at E13.5, germ cell

178   numbers in *Ezh2* [f/-], *Blimp1-Cre* ovaries at E16.5 and E18.5 were significantly reduced (Fig. 3g). The

179   accumulation of γH2A.X signal (Fig. 4a) specifically in female *Ezh2* CKO PGCs suggested the presence

180   of DNA damage, prompting us to examine the transcriptional regulation of transposable elements

181   (TEs) that when re-activated could compromise genome integrity. This revealed that many TEs

182   subfamilies were upregulated predominantly in the E13.5 female *Ezh2* CKO PGCs (Fig. 4b), while only

183   very few differentially expressed TE subfamilies were identified in male *Ezh2* CKO PGCs (Extended

184   Data Fig. 9a). The same trend is observed when analysing uniquely mapped, single copy TEs (Fig. 4c,

185   Extended Data Fig. 9b). The derepression of TEs in female *Ezh2* CKO PGCs is further documented by

186   the upregulation of IAP-GAG and LINE1-ORF1 proteins (Fig. 4d) and the enrichment of P53 and

187   interferon-alpha response pathways in Gene Set Enrichment Analysis (GSEA) (Extended Data Fig. 9c,

188   9d). Both of these pathways have been previously linked to retrotransposon activation [21,22]. In

189   addition, while most of RAD51 foci disappeared at E18.5, filament-like RAD51 signal was identified in

190   female *Ezh2* CKO germ cells, suggesting that unrepaired DNA damage persisted along the

191   chromosomes at this developmental stage (Extended Data Fig. 8b).

192   To understand the molecular underpinnings of the TE activation induced by EZH2 loss, we analysed

193   the repressive chromatin profile. This revealed that between E10.5 and E13.5 the levels of H3K9me3

194   increased on many evolutionarily young and potentially transcriptionally active LTR and non-LTR

195   retrotransposons, such as L1Md_A, IAP-d-int, and MMERVK10C, (Fig. 4e). In fact, more than 90% of

196   *de novo* H3K9me3 peaks identified in E13.5 PGCs are located within 1kb to TEs (Extended Data Fig.

197   9e), suggesting that H3K9me3 provides transcriptional repression of TEs in the absence of DNA

198   methylation. TEs tend to gain more H3K9me3 in the male germline, with newly identified peaks

199   often representing the spreading of a repressive domain (Fig. 4e, Extended Data Fig. 9f).

200   Interestingly, many H3K9me3-marked ERVK, ERV1 and LINE1 subfamilies, including L1Md_Gf, T, A,

201 IAP-d-int and RLTR10, also show an increase in H3K27me3 signal between E10.5 and E13.5 (Fig. 4e).

202 H3K9me3 and H3K27me3 do not usually overlap in somatic cells; however, the dynamic changes

203 observed in PGCs indicate that PRC2 is recruited to these TEs following global DNA demethylation

204 (after E10.5), suggesting a possible functional compensation for the DNA methylation loss. The TEs

205 solely enriched in H3K9me3 tend to be evolutionarily young, transcriptionally competent ERVs

206 (IAPEz.int, IAPLTR1.Mm), while the H3K27me3 enriched TEs tend to be TE relics that lost their

207 transcriptional potential (Charlie, Tigger elements) (Extended Data Fig. 10a). Further analysis at the

208 level of single uniquely mapped TEs revealed three types of histone modification enrichment:

209 H3K9me3 and H3K27me3 are enriched separately on different copies of TEs (Extended Data Fig. 10b);

210 with the copies showing pronounced 5mC loss attracting H3K27me3 while TEs associated with

211 higher levels of DNA methylation showing H3K9me3 enrichment (Extended Data Fig. 10b). However,

212 we also detected a surprising co-enrichment of H3K27me3 and H3K9me3 on some copies

213 particularly of LINE1 elements (Extended Data Fig. 10c) as previously suggested[23]. Importantly, the

214 pattern of repressive histone modifications is specific to an individual integrated copy of TE rather

215 than being representative of the whole TE subfamily.

216

217 Soon following their entry into the developing genital ridge, mouse PGCs undergo a wave of global

218 DNA demethylation yielding the most hypomethylated state of genome ever observed during

219 development [9]. Surprisingly, however this loss of DNA methylation does not lead to widespread

220 transcriptional activation raising an important question regarding the regulation of transcriptional

221 programme and fidelity in the almost near absence of this major repressive epigenetic system[9].

222 Using an ULI-nChIP approach, we have investigated the dynamics of alternative epigenetic silencing

223 marks H3K27me3 and H3K9me3 in the context of dynamic global DNA methylation changes

224 occurring in gonadal PGCs. Our results show that genome-wide DNA demethylation is accompanied

225 by global remodelling of both of these repressive histone modifications, resulting in a sex specific

226 heterochromatin signature observed in post-reprogramming E13.5 PGCs (Fig. 1d, 1e). Using a genetic

227 loss of function model, we further show that EZH2- mediated H3K27me3 is required to maintain

228 transcriptional repression in hypomethylated gonadal PGCs (Extended Data Fig. 3), reinstating that

229 PRC2 is necessary to prevent aberrant transcriptional activation when other stable silencing

230 mechanisms, such as DNA methylation, are absent. This is particularly apparent for germline

231 associated and developmental genes (Fig.3, Extended Data Fig. 3, Extended Data Fig. 10f); and

232 especially in the context of the developing early post reprogramming female germ line, where the

233 presence of RA signalling promotes entry into meiosis. Given that robust transcriptional activation

234   has been shown to override polycomb-mediated repression directly or indirectly [24,25], it is thus

235   plausible that the balance between promoter associated H3K27me3 and cell signalling induced

236   transcription factor repertoire regulates the correct timing of meiotic programme initiation in the

237   female germ line at this stage of development.

238   PRC1 has been previously reported to regulate meiosis-related gene expression in male and female

239   germ cells[26]. Intriguingly, only a small number of previously identified PRC1 targets are dysregulated

240   in our *Ezh2* knockout system pointing to the non-redundant roles of PRC1 and PRC2 in the germ line

241   (Extended Data Fig. 10d).

242   Our study importantly reveals that next to silencing of coding genes, EZH2-mediated H3K27me3 is in

243   post reprogramming hypomethylated germ cells required to repress potentially active transposable

244   elements (Fig. 4, Extended Data Fig. 10f). Using a combination of WGBS and ULI-nChIP-Seq data, we

245   show that H3K27me3 is enriched on many TE subfamilies only after gonadal DNA demethylation (Fig.

246   4e) providing an alternative mode of transcriptional silencing. Loss of H3K27me3 results in profound

247   TE derepression, leading to DNA damage and eventually germ cell loss in the female germ line (Fig.

248   3g). Consistent with this, the elevation of TEs has been linked to foetal oocyte attrition [27]. The

249   severity of the TE derepression in the female PGCs upon EZH2 loss is likely due to the sex specific

250   heterochromatin configuration observed in the developing germ line (Fig. 1d,e, 4e). Following the

251   gonadal 5mC loss, the global DNA methylation remains low for a long period of time during

252   oogenesis; whereas *de novo* DNA methylation ensues soon (from E15.5 onwards) in male germ cells

253   [28]. Additionally, the TE associated H3K9me3 enrichment is much more pronounced in male than in

254   female E13.5 PGCs (Fig. 4e), suggesting that the male germ line utilises H3K9me3 and DNA

255   methylation, in connection with SETDB1 and piRNA pathway[29,30] to restrain TEs. In contrast,

256   H3K27me3 is indispensable for the repression of hypomethylated TEs in the female developing germ

257   cells. Our results thus show diversification of the epigenetic repressive systems in controlling the

258   integrated TEs (Extended Data Fig 10e) and explain the previously reported sex specific effect of the

259   Setdb1 deletion (Liu et al,). Our observation is also conceptually consistent with the previous report

260   in cultured mouse ES cells indicating that H3K27me3 relocalise and silence certain groups of

261   retrotransposons upon rapid and extensive loss of DNA methylation[31]; although different groups of

262   TEs seem to be responsive to the H3K27me3 loss in our *in vivo* system (Extended Data Fig. 10e).

263   Intriguingly, the role of polycomb in controlling parasitic DNA is likely to be ancestral as documented

264   by the recent discovery of the role of the Enhancer-of-zeste like protein Ezl1 in the repression of

265   transposable elements in Paramecium[32] .

266   Collectively, our study reveals that PRC2-mediated H3K27me3 is dynamically remodelled and

267   provides a buffering system of transcriptional regulation to modulate the impact of DNA methylation

268   loss during development. We demonstrate that sequence characteristics of promoters play a critical

269   role in PRC2 recruitment following the removal of cytosine methylation with relatively subtle

270   differences in promoter base composition having a profound effect on H3K27me3 enrichment.

271   Although this has been suggested previously using the insertion of synthetic DNA fragments into

272   mESCs[33]; our genome-wide analysis and genetic study *in vivo* provide to our knowledge the first

273   example of the functional interplay between DNA methylation and polycomb silencing systems in

274   the context of development. To this end, aberrant global DNA hypomethylation has been frequently

275   observed in cancer[34]. As both EZH2 and DNA methyltransferases have been identified as therapeutic

276   targets[35], understanding of the crosstalk between epigenetic silencing systems is of a profound

277   importance to allow precise interpretation and prediction of the outcome of clinical interventions.

278   **Methods**

279   **Mice**

280   All animal experiments were carried out under the UK Home office Project License in the

281   Home Office designated facility.

282   For PGC isolation, Oct4-GFP mice (GOF 18ΔPE-EGFP)[12] were crossed with C57BL/6 mice. The

283   time of mating is determined by the appearance of vaginal plug at noon which is defined as

284   embryonic day 0.5 (E0.5). Germ line deletion of *Ezh2* in PGCs was achieved by crossing mice

285   containing loxP-flanked *Ezh2* allele, *Ezh2*[f/f] mice[36] with *Blimp1 (Prdm1)-Cre* mice. Cre-

286   mediated recombination was detected in 55–76% of PGCs from E8.0 [17], 85% at E9.5 and

287   100% at E13.5[18]. To isolate ctrl and *Ezh2* CKO PGCs for RNA-seq experiments, *Ezh2*[f/f] mice

288   were crossed with Oct4-GFP mice. For genotyping of embryos, embryo tails were digested in

289   genotyping buffer (50mM KCL, 10mM Tris HCL, 0.1mg/ml gelatin, 0.45% Tween 20 dissolved

290   in water) with 0.2 mg/ml Proteinase K (Qiagen) overnight at 55 °C. Proteinase K was

291   inactivated by incubating for 5 min at 95 °C. Genotyping PCR was carried out using REDTaq

292   ready mix (Sigma, R2523). The sex of the embryo from E12.5 onwards was determined by

293   the gross appearance of the gonad and confirmed by PCR. Primers for genotyping are in

294   (**Supplementary Table 5**).

295   **PGC isolation by flow cytometry**

296   The dissected genital ridges (from E11.5) or segments of embryos (E9.5-E10.5) were first

297   digested for 5 min at 37 °C using 0.05% Trypsin-EDTA (Gibco) or TrypLE Express (Gibco).

298   Enzymatic digestion was followed by neutralization with pre-warm DMEM (Gibco, 21969-

299   035), containing 10% foetal bovine serum (Gibco) and pipetting up and down 20-30 times,

300   followed by centrifugation 2,000 rpm for 3min to collect cells. Cells were subsequently

301 resuspended in DMEM/F-12 supplemented with hyaluronidase (300 µg/ml; Sigma) by
302 pipetting up and down 10-20 times to generate single cell suspensions. Following
303 centrifugation at 2,000 rpm for 3 min to collect cells, the cell pellet was resuspended in ice-
304 cold PBS supplemented with poly-vinyl alcohol (10 µg/ml, Sigma) and EGTA (0.4 mg/ml,
305 Sigma). GFP-positive cells were isolated using an Aria IIu (BD Bioscience) or Aria III (BD
306 Bioscience) flow cytometer.

RNA-seq library preparation

308 E13.5 PGCs from each embryo were sorted separately and centrifuged at 2,000 rpm for 3
309 min following liquid nitrogen snap freezing. Total RNA was extracted using ZR-Duet DNA and
310 RNA Mini Prep kit (Zymo). The genotype of embryos was determined by PCR. RNA quality
311 was measured by 2100 bioanalyzer (Agilent technologies). To prepare cDNA libraries, 1 ng
312 total RNA was used for cDNA synthesis amplification (11 cycles) using SMART-seq V4 Ultra
313 Low input RNA kit (Takara), according to the manufacturer's instructions. The cDNA was
314 sheared by Covaris S2 sonicator (peak: 175, Duty: 10, cycle: 200, duration: 240 secs)
315 (Covaris). Fragmented cDNA was indexed and converted to sequencing libraries using
316 NEBNext Ultra II DNA library Prep Kit (NEB) following the manufacturer's instructions. All
317 libraries were purified by AMPure XP beads (Beckman–Coulter) and sequenced on the
318 Illumina HiSeq 2500.

Cryosection immunofluorescence staining

320 The dissected embryonic trunk (E10.5) or genital ridges from embryos (E11.5-E18.5) were
321 fixed by 2% paraformaldehyde (PFA) for 30 min at 4 °C, washed 3x 10min in PBS and
322 incubated overnight in 30% sucrose in PBS at 4 °C. The samples were then mounted in O.C.T.
323 mounting medium (VWR) and stored at -80 °C. The mounted samples were sectioned for 10
324 um using Leica cryostat (Leica CM1950). The cryosections were post-fixed with 2% PFA in
325 PBS for 3 min, washed 3x 5min in PBS and permeabilised with blocking buffer (1% bovine
326 serum albumin (BSA)/0.1% Triton in PBS). The primary antibodies were added in blocking
327 buffer and incubated with slides overnight at 4 °C. The slides were subsequently washed 3
328 times in blocking buffer and incubated with Alexa dyes conjugated secondary antibodies
329 (Molecular Probes) in blocking buffer for 1 hr at room temperature in the dark. Then the
330 slides were washed 2 times for 5 min in blocking buffer and one for 5 min in PBS. Finally, the
331 slides were treated with DAPI (0.1 µg/ml) for 20 min and mounted in Vectashield (Vector
332 Laboratories) and imaged by using a Leica SP5 confocal microscope.

333 For 5mC and 5hmC staining, post-fixed sections were first permeabilized for 30 min with
334 blocking buffer (1% BSA/0.5% Triton in PBS) and subsequently treated with RNase A (10
335 mg/ml, Roche) (in 1% BSA/PBS) for 1 hr at 37 °C. Followed by times for 5 min washes with
336 PBS, sections were incubated with pre-warmed 4N HCl for 10, 15 and 20 min at 37 °C to
337 denature genomic DNA, followed by three washes with PBS. After incubating for 30 min at
338 room temperature in 1% BSA/PBS containing 0.1% Triton X-100, the sections were

339   incubated with primary antibodies at 4 °C overnight in the same buffer. Sections were

340   subsequently washed three times in blocking buffer (1% BSA/0.1% Triton in PBS) for 5 min

341   and incubated with the corresponding combination of secondary antibodies in the same

342   buffer for 1 hr in the dark at room temperature. Secondary antibody incubation was

343   followed by three 5 min washes with PBS. DNA was then stained with propidium iodide (PI)

344   (0.5ug/ml). After a final wash in PBS for 10 min, the sections were mounted with Vectashield

345   (Vector Laboratories). Antibodies are listed in (**Supplementary Table 5**).

346   **Quantification and Image analysis**

347   All IF images were processed and merged by Image J FIJI. For quantification of IAP GAG and

348   LINE1 ORF1, the integrated intensity in cytoplasm was measured by Cell profiler 2.2.0.To

349   identify cytoplasm area, DAPI was used to mark the nucleus and MVH-positive cells were

350   selected. The measured integrated intensity values were normalised to the background of

351   each staining to obtain normalised integrated intensity. For γH2AX, the intensity in the

352   nucleus was measured by Image J. The intensity in PGCs was divided by the intensity in the

353   somatic cells.

354   **Germ cell counting**

355   We counted MVH-positive germ cells in the every fifth cryosection throughout the entire

356   gonad. To calculate the total number of germ cells in each gonad, the numbers counted per

357   section were multiplied by 5. The calculated germ cell numbers in the E16.5 and E18.5 Ctrl

358   gonads are close to the previous report[37].

359   **Western blot for PGCs**

360   5000 Oct4-GFP positive PGCs and 5000 ES cells grown in serum-based medium were sorted

361   and centrifuged at 2,000 rpm for 3 min. Cells were lysed with 10ul RIPA buffer (Sigma,

362   RO278) with protease inhibitor cocktail (Roche Complete tables mini). Total lysate was

363   loaded into 12% acrylamide/bis gel and separated in running buffer (25 mM Tris base, 190

364   mM glycine, 0.1% sodium dodecyl sulfate (SDS)). The protein was transferred to a

365   nitrocellulose membrane (Amersham) after electrophoresis in transferring buffer (25 mM

366   Tris base, 190 mM glycine, 0.1% SDS, 20% methanol). The membrane was blocked by 5%

367   BSA in PBST (0.1% tween in PBS) for 30 min at room temperature. Primary antibodies were

368   added in PBST with 5% BSA and incubated with the membrane overnight at 4 °C (MVH:

369   1:1000, H3K27me3 and H3K9me3: 1:2000, H3: 1:5000). The membrane was washed 3 times

370   with PBST for 10 min. Horseradish peroxidase (HRP) conjugated secondary antibodies

371   (1:10000) incubated for 1 hr at room temperature and then wash 3 times with PBST.

372   Luminata Crescendo Western HRP substrate (WBLUR0100, Milipore) was used for detection

373   on Amersham imager 680. Antibodies are listed in (**Supplementary Table 5**).

374   **5mdC/5hmdC quantification by LC-MS/MS**

375 DNA was extracted from 1000-6000 PGCs using ZR-Duet DNA and RNA Mini Prep kit (Zymo)
376 and digested to nucleosides overnight at 37°C using a nucleoside digestion mix (NEB,
377 M0649). The nucleosides were separated on an Agilent RRHD Eclipse Plus C18 2.1 × 100 mm
378 1.8u column using the HPLC 1290 system (Agilent) and mobile phases 100% water 0.1%
379 formic acids and 80% methanol, 0.1% formic acids. Quantification was carried out in an
380 Agilent 6490 triple quadrupole mass spectrometer on multiple reaction monitoring mode
381 (MRM), by monitoring specific transition pair of m/z 250.1/134.1 for dC, 290.1/174.1 for dG,
382 264.1/148.1 for 5mdC and 280.1/164.1 for 5hmdC. To calculate the concentrations of
383 individual nucleosides, standard curves were generated (dC and dG from Berry and
384 Associated; 5mdC and 5hmdC from CarboSynth). All samples and standard curve points
385 were spiked with a similar amount of isotope-labelled synthetic nucleosides (13C15N-dC
386 and 13C15N-dG purchased from Silantes, and d3-mdC and d215N2-mhdC was obtained
387 from T. Carell, Center for Integrated Protein Science at the Department of Chemistry,
388 Ludwig-Maximilians-Universität München, Germany). The threshold for peak detection is a
389 signal-to-noise ratio (calculated with a peak-to-peak method) above 10. . Final
390 measurements were normalised by dividing by the dG level measured for the same sample.
391 Limit of detection (LOD) was 0.005 -250 fmol for 5mdC and 5hmdC, and 0.1-5000 for dC and
392 dG. Limit of quantification (LOQ) was 0.025 – 50 fmol for 5mdC and 5hmdC, and 1-1000 for
393 dC and dG.

394 **Ultra low-input native chromatin immunoprecipitation (ULI-nChIP)**

395 Ultra low-input nChIP-seq was performed as previously described[10,11]. PGCs were FACS
396 sorted into cold PBS and were lysed by nuclei isolation buffer (Sigma, NUC-101). For E10.5,
397 PGCs were pooled from different embryos. 1000 PGCs were used for each ChIP reaction. For
398 E13.5, embryos from independent litters are used as biological replicates. Samples were
399 permeabilized with 0.1% Triton-X-100 /0.1% deoxycholate in PBS on ice. Samples were
400 digested with 200 U of micrococcal nuclease (NEB, M0247S) in digestion buffer at 21 °C for
401 7.5 mins and the reaction was stopped with 100 mM EDTA. An aliquot of chromatin (10%)
402 was used as an input control. Samples were pre-cleaned by protein A/G beads for 2 hrs at 4
403 °C. Meanwhile, 250 ng anti-H3K9me3 (Diagenode, C15410056) and anti-H3K27me3
404 (Millipore, 07-449) antibodies were incubated with protein A/G beads for 2 hrs at 4 °C.
405 Chromatin samples were incubated with antibody-bound beads overnight at 4 °C. Samples
406 were washed by two low-salt washes and one high salt wash and eventually DNA was eluted
407 from the beads for 1.5 hrs at 60 °C. For sizes selection, AMPure XP beads were used at 1.8:1
408 ratio. Library preparation was performed using a MicroPlex Library Preparation kit v2
409 (Diagenode, C05010013) following manufacturer's recommendations. Libraries were further
410 purified by AMPure XP beads at 1:1 ratio. Library quality was measured with a 2100
411 Bioanalyzer Instrument (Agilent) and library size was assessed with a Kapa library qPCR
412 quantification kit (Roche, KK4824). Library sequencing was performed in a 100bp paired-end
413 mode with a HiSeq 2500 System (Illumina) following manufacturer's recommendations.

RNA-seq data analysis

Published RNA-Seq datasets for E10.5, E12.5 male, E12.5 female, E14.5 male and E14.5 female PGCs[9] were retrieved from Gene Expression Omnibus (GEO) under accession GSE76958.

Paired-end 100bp sequencing reads were processed with trimmomatic (0.33)[38] for trimming adapters and low quality reads, then aligned against Ensembl mouse genome (NCBIM37) with Tophat2 (2.0.11)[39]. Reference sequence assembly and transcript annotation were obtained from Illumina iGenomes (https://support.illumina.com/sequencing/sequencing_software/igenome.html). Gene-based read counts were obtained using featureCounts function from Rsubread Bioconductor package (1.24.2)[40,41]. Normalisation was performed in DESeq2 Bioconductor package (1.14.1)[42] and data was rlog transformed to allow for visualisation by PCA and heat maps. Differentially expressed gene (DEG) analysis was also performed with DESeq2 and DEGs were defined with Benjamini-Hochberg adjusted P-value<0.05 and fold change > 2. Z score for heat maps was calculated based on rlog transformed data[42]. Gene ontology analysis was performed with goseq Bioconductor package (1.24)[43]. After converting mouse gene symbol to human gene symbol using the report of Human and Mouse Homology retrieved from Mouse Genome Informatics (MGI, http://www.informatics.jax.org), Gene Set Enrichment Analysis (GSEA 2.2.0)[44] was then performed with GseaPreranked tool using Hallmarks gene set (h.all.v5.2.symbols.gmt). Motif analysis was performed with PWMEnrich Bioconductor package (4.20.0). Sample distance matrix was calculated using Euclidean distance and then hierarchical clustering was performed using the complete linkage method.

In order to compare our RNA-seq data with WT PGC RNA-seq data[9] we used RUVseq Bioconductor package (1.18.0)[45] to correct the batch effect with RUVg function and argument k=1 before producing the PCA plot.

ChIP-seq data analysis

ChIP-seq and input libraries were sequenced and 100 bp paired-end reads were aligned to UCSC mm9 mouse genome with BWA (0.7.5a)[46]. Aligned reads were sorted by Picard SortSam tool (1.9; Picard), and duplicated reads were removed by Picard MarkDuplicates tool (1.9; Picard). Using samtools view (0.1.18) with arguments -F 4 -f 0x02, we only kept properly paired reads for downstream analysis. Properly paired reads were also further separated into uniquely mapped reads and multiple mapped reads (MAPQ=0) for repeat sequence analysis. (Multiple mapped reads were randomly assigned to one of their mapped genomic loci).

We created 2 kb bin windows along the mm9 genome and subsequently removed windows overlap with mm9 blacklist regions[47,48]. Read counts for each windows for ChIP and input were obtained using featureCounts function from Rsubread. The ChIP signal was normalised to input using DESeq2. Bins Per Million (BPM) was calculated same as TPM in RNAseq[49,50].

455 Peak calling was performed using MACS2 (2.1.0)[51] with arguments -f BAMPE --broad -g mm.
456 To annotate the peaks to genomic regions, we used ChIPseeker in the Bioconductor package
457 (1.18.0)[52] with default settings, i.e. if peak overlapped with multiple genomic features, the
458 annotation was assigned in the following priority: Promoter, 5' UTR, 3' UTR, Exon, Intron,
459 Downstream and Intergenic. Only L1, ERV1, ERVK and ERVL repetitive families were used to
460 calculate the distance between peaks and transposable elements.
461
462 **Transposable element analysis**
463
464 To annotate ChIP-seq and RNA-seq read counts to repetitive sequences, we used the
465 repeatmasker track without repeat classes 'Low_complexity' or 'Simple_repeat' from mm9
466 UCSC genome table browser.
467
468 For TE subfamily level (eg. L1Md_T, IAP-d-int…), both uniquely and multiple mapped reads
469 (multiple mapped reads were randomly assigned to one of their mapped genomic loci).
470 Differentially expressed transposable elements were identified using edgeR Bioconductor
471 package (3.26.5)[53]. For single copy level, only uniquely mapped reads were considered.
472
473 **WGBS analysis**
474
475 Published WGBS datasets from wild type E10.5, E12.5 male and female PGCs[9] were
476 retrieved from Gene Expression Omnibus (GEO) under accession GSE76971.
477 Low quality reads and adaptors were trimmed off using trim_galore (v0.4.4;
478 https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Bismark [54] was then
479 applied with –bowtie2 –ambiguous -p 8 –multicore 4 –ambig_bam settings, and trimmed
480 reads were mapped to mm9 BS genome. Cytosine methylated and unmethylated coverage
481 reports were created by applying bismark_methylation_extractor. We focused on CpG sites
482 from the reports with at least 5 fold coverage, and methylation rate was calculated as count
483 methylated/(count methylated+count unmethylated).
484
485 **Promoter definition and classification**
486
487 Promoters were defined as −2000 bp to +500 bp relative TSS. For the heatmaps, we used 2
488 kb windows with at least 1 kb overlapped with defined promoters. Promoter-associated
489 transcripts were obtained from R Bioconductor package TxDb.
490 Mmusculus.UCSC.mm9.knownGene (v 3.2.2). We classified total 33158 promoters based on
491 the pattern of dynamic changes in DNA methylation and H3K27me3.
492 Group A: Loss of DNA methylation, Gain of H3K27me3 at E13.5.
493 $DNA\ methylation\ rate\ E10.5 > 0.2, DNA\ methylation\ rate\ E10.5 - E12.5 > 0,$
494 $Log_2((E13.5\ H3K27me3)/(ChIP\ input\ )) \geq 0$
$Log_2((E13.5\ H3K27me3)/(ChIP\ input\ )) - Log_2((E10.5\ H3K27me3)/(ChIP\ input\ )) \geq 0.5$
495 Group B: Median loss of DNA methylation, High H3K27me3 at both stages.
$DNA\ methylation\ rate\ E10.5 > 0.1, DNA\ methylation\ rate\ E10.5 - E12.5 > 0$
$Log_2((E13.5\ H3K27me3)/(ChIP\ input\ )) \geq 1$
$Log_2((E13.5\ H3K27me3)/(ChIP\ input\ )) - Log_2((E10.5\ H3K27me3)/(ChIP\ input\ )) < 0.5$
496 Group C: Low DNA methylation rate, High H3K27me3.
$DNA\ methylation\ rate\ E10.5 < 0.1$
$Log_2((E13.5\ H3K27me3)/(ChIP\ input\ )) \geq 1$
$Log_2((E13.5\ H3K27me3)/(ChIP\ input\ )) - Log_2((E10.5\ H3K27me3)/(ChIP\ input\ )) < 0.5$

497   Group D: Loss of DNA methylation, Low H3K27me3 at E13.5.

$DNA\ methylation\ rate\ E10.5 > 0.2, or\ DNA\ methylation\ rate\ E10.5 - E12.5 > 0$

$Log_2((E13.5\ H3K27me3)/(ChIP\ input\ )) < 0.5$

498

499   Non-classified: the rest of promoters.

500

501   CpG obs/exp ratio, percentage of CpG dinucleotide and percentage of C or G were

502   calculated within 2kb window as previously reported[55].

503

Statistical analysis

504

505   All statistical analysis were performed using R or Graphpad software and are described in

506   the figure legends. Biological replicates for all experiments were based on embryos from

507   independent litters. Specifically, P values were calculated by Wilcoxon rank sum test

508   (Extended Data Fig. 2c, Fig. 3b, 7c). Adjusted P values were calculated by Wilcoxon rank sum

509   test and adjusted by Bonferroni correction (Fig. 2). P values were calculate by two-tailed

510   unpaired Student's t test (Fig.3g and Fig. 4a, 4d). Adjusted P values were calculated by

511   ANOVA and Tukey's post-hoc multiple comparison test (Extended Data Fig. 5e).

512   Box plots were plotted using Tukey's method. The upper and lower hinges represent the

513   first and the third quartiles. The central line represents the median. The upper end of the

514   whisker represents the lowest value among either the third quartile plus 1.5 X IQR or the

515   maximum value from the data set. The lower end of the whisker represents the largest

516   value among either the first quartile minus 1.5 X (IQR) or the minimum value from the data

517   set.

518   Box plots were generated by R package ggplot2 with argument outlier.shape = NA in (Fig. 2b

519   and extended data Fig.3b).

Data availability

520

521   ChIP-seq and RNA-seq data have been deposited in Gene Expression Omnibus (GEO) under

522   GSE141182.

523

524

525

526

527

528

529

Figure Legends

Figure 1: Sex-specific remodelling of repressive histone modifications during genome-wide DNA demethylation in PGCs.

a, Overview of chromatin dynamics during mouse PGC development. *: The assessment is based on immunofluorescence staining. **b**, Genomic distribution of H3K27me3 and H3K9me3 peaks. **c,** Genome-wide correlation (Spearman's correlation coefficient, 2 kb bins, whole genome) of H3K27me3 and H3K9me3 enrichment after normalised to the input. **d**, **e,** Total H3K27me3 and H3K9me3 peak numbers at each PGC stage. Venn diagram showing peaks overlapping between E10.5, E13.5 male and E13.5 female PGCs. Bar charts show retained and *de novo* peaks. **f**, Western blot analysis of H3K27me3 (left) and H3K9me3 (right) abundance based on 5000 mESCs and 5000 GFP-positive E13.5 PGCs. MVH: germ cell marker. P value was calculated from 3 independent experiments by two-tailed Student's *t* test. Error bar: standard deviation.

Figure 2: Base composition determines H3K27me3 enrichment during gonadal DNA demethylation.

a, Heat map depicting the dynamics of DNA methylation and the enrichment of H3K27me3 at promoters. 4 groups of promoters were identified based on the dynamic patterns between DNA methylation and H3K27me3 (See Methods). **b,** Sequence characteristics of the promoters described in (Fig. 2a). adj. P values were calculated by pairwise comparison using Wilcoxon rank sum test and adjusted by Bonferroni correction. ***: P<0.001. **c,** Dynamics of DNA methylation and H3K27me3 at all promoters with respect to their sequence characteristics. Methylation rate and H3K27me3 enrichment are shown by the colour gradient. Distribution of each dot's value is shown using rug plot along x axis and y axis.

Figure 3: Conditional *Ezh2* KO leads to wide-spread transcriptional derepression and loss of germ cells in the female germ line.

a, b, Immunofluorescence staining of EZH2 and H3K27me3 in gonadal sections (n=2). MVH: germ cell marker. Yellow arrowheads indicate PGCs. DNA stained by DAPI (blue). Scale bar: 10 um. **c,** Principal component analysis (PCA) of RNA-seq data from E13.5 male and female Ctrl and *Ezh2* CKO (*Ezh2*[-/-], *Blimp1-Cre*) PGCs at E13.5. **d,** Number of differentially expressed (DE) genes using different fold change (FC) thresholds. adj. P<0.05. **e,** Gene ontology (GO) terms associated with DE genes (CKO vs Ctrl). **f,** Heat map of gene expression and chromatin dynamics of promoters associated with meiotic prophase genes. P value was calculated using Fisher exact test. **g,** Representative immunostaining of E18.5 Ctrl and *Ezh2* CKO female gonad sections. DNA was stained by DAPI. Scale bar: 100 um. Dot plot shows the total

566 number of germ cells per ovary. Each dot represents one biological replicates. Error bars
567 indicate standard deviation. P values were calculated by two-tailed unpaired Student's t test.

568 **Figure 4 EZH2-mediated H3K27me3 regulates retrotransponson repression.**

569 **a,** Immunofluorescence staining for γH2AX (biological replicates n=2). Quantification of the
570 staining per PGC nucleus shown on the right. Numbers indicate the numbers of analysed
571 PGC nuclei. DNA was stained by DAPI. Scale Bar: 10 um. **b,** Expression of TEs (Multiple
572 mapped plus uniquely mapped reads) in E13.5 PGCs (*Ezh2* CKO vs Ctrl). Significantly
573 upregulated TEs are labelled in red. (Fold change > 1, FDR< 0.1). **c,** Dot plot represents the
574 expression levels of individual TE copies. Each dot represents a single element in indicated
575 subfamilies. Only uniquely mapped reads were considered. **d,** Immunofluorescence staining
576 (top) and signal quantification (bottom) of IAP GAG and LINE1 ORF1 proteins. Biological
577 replicates n=2. Statistical analysis was carried out using two-tailed unpaired Student's *t* test.
578 **e,** Heat map showing H3K9me3 and H3K27me3 enrichment of TE subfamilies ranked by DNA
579 methylation levels at E10.5. Each row represents one TE subfamily. Both multiple mapped
580 and uniquely mapped reads were considered.

581 Extended data Figure 1

582 **Summary of ULI-nChIP-seq and genomic distribution of H3K27me3 and H3K9me3**
583 **enrichment.**

584 **a,** Experimental scheme of PGC isolation using ΔPE Oct4-GFP mice (GOF 18ΔPE-GFP)[12].
585 ~1000 PGCs were used for ULI-nChIP-seq. **b,** Characteristics of the H3K27me3 and H3K9me3
586 ULI-nChIP-seq. Fraction of paired-end reads based on the mappability. Uniquely aligned
587 (dark green). Uniquely aligned duplicates (light green). Multiple aligned (blue). Unaligned
588 (grey). **c,** Scatter plot showing the correlation between 2 biological replicates using 2kb
589 window. Higher variability observed between E13.5 samples relates to slight difference in
590 developmental progression between different embryos and litters (embryos from
591 independent litters used as biological replicates). Pearson correlation coefficient is shown
592 on the top left. **d**, Genomic distribution of H3K27me3 and H3K9me3 peaks. More than 60%
593 of H3K27me3 peaks are associated with promoters or gene bodies. H3K9me3 peaks are
594 located mostly in distal intergenic regions and introns. **e,** Distribution of H3K27me3 and
595 H3K9me3 peaks relative to transcription start site (TSS). **f,** Bar chart showing proportion of
596 H3K27me3 and H3K9me3 peaks associated with TEs in the genome. **g,** Violin plot showing
597 the distribution of peak intensity in E10.5, E13.5 male and female PGCs. Peaks were
598 identified by MACS2 peak calling pipeline with broad peak setting (See Methods). **h,**
599 Distribution of H3K27me3 and H3K9me3 peak length. The number of short peaks increases
600 at E13.5 male PGCs, compared with peaks at E10.5.

601

602

603

604   Extended Data Figure 2

605   **Dynamics of DNA methylation and H3K27me3 at promoters during gonadal**
606   **reprogramming.**

607   **a,** Whole-genome bisulphite sequencing (WGBS) data from E10.5, E12.5 female and E12.5
608   male PGCs[9]. Density plot depicting DNA methylation levels at all promoters. **b,** Density plot
609   depicting H3K27me3 enrichment at all promoters. **c,** Violin plot of H3K27me3 levels at
610   promoters that lost DNA methylation (DNA methylation >0.2 at E10.5). **d,** Dynamics of DNA
611   methylation and H3K27me3 at all promoters. DNA methylation and H3K27me3 enrichment
612   are shown by colour gradient. Distribution of each dot's value is shown using rug plot along
613   x axis and y axis. **e,** Box plot shows H3K27me3 enrichment of low CpG density (CpG <4.1%)
614   promoters which gained H3K27me3 in the female PGCs following global loss of DNA
615   methylation. Box plots were presented by Tukey method. P values were calculated by
616   Wilcoxon rank sum test. ***: P <0.001.

617   Extended data Figure 3

618   **Dynamics of DNA methylation, H3K9me3 and H3K27me3 at promoters during gonadal**
619   **reprogramming.**

620   **a,** Heat map depicting the H3K27me3 and H3K9me3 enrichment (ULI-nChIP-seq), and DNA
621   methylation rate (WGBS) at promoters. The promoters were grouped based on the pattern
622   of dynamic change between DNA methylation and H3K27me3 (See also Fig. 2). Group A: loss
623   of DNA methylation, gain H3K27me3 at E13.5. Group B: median loss of DNA methylation,
624   high H3K27me3. Group C: low DNA methylation, High H3K27me3. Group D: loss of DNA
625   methylation, low H3K27me3 at E13.5. Promoters that did not meet the criteria were
626   grouped into non-classified. The expression levels of promoter-associated genes from RNA-
627   seq are presented by TPM (Transcripts Per Kilobase Million) or z-score. The total number of
628   promoters in each group are shown on the left. **b,** Box plot showing the quantitative
629   measurement in each category, female and male PGCs, respectively. P values were
630   calculated by Wilcoxon rank sum test. **c,** Venn diagram showing the number of overlapped
631   promoters between male and female PGCs.

632   Extended Data Figure 4

633   **Generation of the germline specific *Ezh2* conditional knock out.**

634   **a,** Representative immunofluorescence (IF) staining for EZH2 using cryosectioned genital
635   ridges. EZH2 is highly expressed during PGC development, compared with surrounding
636   somatic cells. Biological replicates n=3. OCT4: PGC marker. DAPI indicates DNA. Scale bar:
637   10um. **b,** Functional domains of EZH2 protein and targeting strategy of *Ezh2* allele. Open
638   boxes: exons. Black arrowhead: loxP sites **c,** Breeding scheme for germline *Ezh2* knockout.
639   *Ezh2* $^{\Delta/\Delta}$, Tg (*Blimp1-Cre*) refers to CKO in the figures. f: allele flanked by loxP sites (floxed).

640   Δ: Deleted allele generated using Cre-mediated recombination. Tg (*Blimp1-Cre*): transgenic

641   mice express Cre recombinase under the control of *Blimp1* (*Prdm1)* promoter. **d,** Deleted

642   alleles were confirmed by genotyping using the primers shown by black arrows in (**b**). #1:

643   *Ezh2*^f/Δ^, Tg (*Blimp1-Cre*)+/-. #2: *Ezh2*^f/+^

644   Extended Data Figure 5

645   **Global H3K9me3, H2A119ub and DNA methylation are not altered in PGCs following the**

646   **loss of EZH2.**

647   **a, b, c,** Representative IF staining for H3K9me3, H2A119ub and TET1 using cryosectioned

648   genital ridges. MVH: PGC marker. Biological replicates n=2. **d,** IF staining for 5-

649   methylcytosine (5mC). 5mC is enriched in pericentromeric regions in the nucleus of somatic

650   cells but depleted in both Ctrl and *Ezh2* CKO germ cells. **e,** Global 5mC and 5hmC levels were

651   measured by LC-MS/MS. Each dot represents one biological replicate. Mean values of

652   5mdC/dG or 5hmdC/dG are shown. adj. P values were calculated using ANOVA and Tukey's

653   post-hoc multiple comparison test. **f,** Representative IF images of E13.5 Ctrl and *Ezh2* CKO

654   female and male gonads. DNA was stained by DAPI. Scale bar: 100 um. The bar chart shows

655   the total number of germ cells per female gonad. Error bars indicate standard deviation.

656   Biological replicates n=2.

657   Extended data Figure 6

658   **Transcriptome analysis of Ctrl and *Ezh2* CKO PGCs.**

659   **a,** Catalytic core and accessory subunits of mammalian PRC2. **b,** RNA expression of PRC2

660   components during germ cell development. **c,** RNA expression of *Ezh1* in the Ctrl and *Ezh2*

661   CKO germ cells. **d,** Sample distance matrix of RNA-seq samples by non-supervising cluster

662   (see Methods). **e,** PCA Plot shows the distance of transcriptomes from different PGC

663   developmental stages. Dash line circle indicates samples of the same developmental stage.

664   Extended Data Figure 7

665   **Sex-specific transcription factor repertoire determines transcriptional activation upon loss**

666   **of EZH2.**

667   **a,** Gene ontology (GO) terms associated with E13.5 ♂ DE genes (Ctrl vs CKO). **b,** Integrative

668   Genomics Viewer (IGV) plot shows the H3K27me3 enrichment and RNA-seq read counts of

669   *Stra8*. Mouse genome: mm9. **c,** Heat map depicting gene expression and the chromatin

670   dynamics at promoters of meiotic differentially expressed (DE) genes. Box plot shows the

671   H3K27me3 enrichment and RNA expression (TPM) of meiotic DE genes in male and female

672   PGCs. Z scores were calculated for male and female separately. P values were calculated by

673   Wilcoxon rank sum test. ***: P<0.01 **d,** Bar chart showing the odds ratio of 104 meiosis

674   prophase genes[19] in each groups of promoters (Fig. 2). P values were calculated by Fisher

675   exact test. *: P<0.05. **: P<0.01. ***: P<0.001. **e,** Promoters of upregulated genes in female

676   *Ezh2* CKO are significantly enriched for transcription factor motifs that relate to retinoid acid
677   signalling pathway. Motif analysis was performed using Bioconductor package PWMEnrich. **f,**
678   Heat map shows the relative gene expression of identified transcription factors in male and
679   female PGC samples. **g,** Heat map shows gene expression of 45 Germline Reprogramming
680   Responsive (GRR) genes in Ctrl and *Ezh2* CKO samples. Differentially expressed genes are
681   shown on the top (adj. P<0.05).

682   Extended Data 8

683   **Loss of Ezh2 does not lead to precocious meiotic prophase**

684   **a,** Representative IF images of the meiotic specific synaptonemal complex protein SCP3 in
685   embryonic ovaries. Axial element alignment was observed in female E18.5 germ cells. **b,**
686   Representative IF images of cryosectioned gonads. γH2AX signal shows DNA double strand
687   breaks (DSBs) occurring during homologous recombination. Accumulation of γH2AX signal in
688   E16.5 ctrl and *Ezh2* CKO germ cells, this is greatly reduced at E18.5. A number RAD51 foci
689   can be identified at E16.5 but greatly decreases at E18.5. Filament-like, RAD51-positive
690   structure was identified in the *Ezh*2 CKO germ cells but not in Ctrl germ cells. Germ cells are
691   indicated by yellow arrowhead and are positive for MVH. DAPI indicates DNA. Scale bar:
692   10um.

693   Extended Data Figure 9

694   **EZH2-mediated H3K27me3 regulates TE repression.**

695   **a,** Differential expression analysis of TE expression in E13.5 male PGCs (CKO vs Ctrl).
696   Significantly upregulated TEs are labelled in red. Fold change > 1, FDR< 0.1. **b,**
697   Multidimensional scaling (MDS) plot showing distance of Ctrl and *Ezh2* CKO samples based
698   on TE expression. Only uniquely mapped reads were considered. **c,** Gene Set Enrichment
699   Analysis (GSEA) of DE genes(E13.5♀ CKO vs Ctrl). Number of genes enriched in each gene
700   set is shown by the circle size. **d,** GSEA plot showing genes upregulated in female *Ezh2* CKO
701   PGCs are enriched in p53 pathway and interferon alpha response. FDR q value < 0.25 was
702   considered significant. NES: normalized enrichment score. **e,** Analysis of distance between
703   H3K9me3 *de novo* peaks, transcription start sites (TSS) and transposable elements (TEs).**f,**
704   Representative IGV plot showing H3K9me3 enrichment on IAP Ez elements.

705   Extended Data Figure 10

706   **H3K9me3 enrichment on TEs and co-localisation with H3K27me3.**

707   **a,** TE subfamilies enriched predominantly for H3K9me3 or H3K27me3. Each row represents
708   one TE subfamily. Multiple mapped and uniquely mapped reads were taken into account. **b,**
709   **c,** Heat map showing DNA methylation, H3K9me3 and H3K27me3 enrichment at individual
710   copies of IAPLTR2_Mm and L1Md_Gf. Each row represents one uniquely mapped, single TE
711   copy belonging to the respective TE subfamily. **d,** Venn diagram depicting overlap of genes

712 upregulated in the *Ezh2* CKO female PGCs and the *Rnf2* CKO PGCs[26]. **e,** Venn diagram
713 showing overlap between the EZH2 and SETDB1 regulated TEs in mouse germ cells and
714 between the EZH2 regulated TEs in germ cells and the EED regulated TEs in mESCs [23,31]. **f,**
715 Model depicting the relationship between DNA demethylation and heterochromatin
716 changes in gonadal PGCs undergoing epigenetic reprogramming.

717

718

729

730 **Author Contributions**

731 T.C.H and P.H. conceived the study; T.C.H performed the experiments and analysed the data; Y.F.W.
732 analysed the next-generation sequencing data; E.V-F. carried out the ULI-nChIP with the help of C.H.
733 and G.K.; C.E.R. carried out LC–MS/MS; T.C.H. and P.H. wrote the manuscript.

734

735

736

737 1   Robertson, K. D. DNA methylation and human disease. *Nat Rev Genet* **6**, 597-610,
738   doi:10.1038/nrg1655 (2005).
739 2   Guibert, S., Forne, T. & Weber, M. Global profiling of DNA methylation erasure in mouse
740   primordial germ cells. *Genome research* **22**, 633-641, doi:10.1101/gr.130997.111 (2012).
741 3   Hackett, J. A. *et al.* Germline DNA demethylation dynamics and imprint erasure through 5-
742   hydroxymethylcytosine. *Science* **339**, 448-452 (2013).
743 4   Seisenberger, S. *et al.* The Dynamics of Genome-wide DNA Methylation Reprogramming in
744   Mouse Primordial Germ Cells. *Molecular Cell* **48**, 849-862 (2012).
745 5   Seki, Y. *et al.* Cellular dynamics associated with the genome-wide epigenetic reprogramming
746   in migrating primordial germ cells in mice. *Development* **134**, 2627-2638,
747   doi:10.1242/dev.005611 (2007).
748 6   Hajkova, P. *et al.* Chromatin dynamics during epigenetic reprogramming in the mouse germ
749   line. *Nature* **452**, 877-881, doi:nature06714 [pii]

750 10.1038/nature06714 (2008).

751 7 Hajkova, P. Epigenetic reprogramming in the germline: towards the ground state of the
752   epigenome. *Philos Trans R Soc Lond B Biol Sci* **366**, 2266-2273, doi:10.1098/rstb.2011.0042
753   (2011).

754 8 Spiller, C., Koopman, P. & Bowles, J. Sex Determination in the Mammalian Germline. *Annu*
755   *Rev Genet* **51**, 265-285, doi:10.1146/annurev-genet-120215-035449 (2017).

756 9 Hill, P. W. S. *et al.* Epigenetic reprogramming enables the transition from primordial germ
757   cell to gonocyte. *Nature* **555**, 392-396, doi:10.1038/nature25964 (2018).

758 10 Brind'Amour, J. *et al.* An ultra-low-input native ChIP-seq protocol for genome-wide profiling
759   of rare cell populations. *Nat Commun* **6**, 6033, doi:10.1038/ncomms7033 (2015).

760 11 Hanna, C. W. *et al.* MLL2 conveys transcription-independent H3K4 trimethylation in oocytes.
761   *Nature structural & molecular biology* **25**, 73-82, doi:10.1038/s41594-017-0013-5 (2018).

762 12 Yoshimizu, T. *et al.* Germline-specific expression of the Oct-4/green fluorescent protein (GFP)
763   transgene in mice. *Dev Growth Differ* **41**, 675-684 (1999).

764 13 Juan, A. H. *et al.* Roles of H3K27me2 and H3K27me3 Examined during Fate Specification of
765   Embryonic Stem Cells. *Cell Rep* **18**, 297, doi:10.1016/j.celrep.2016.12.036 (2017).

766 14 Li, H. *et al.* Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature* **549**, 287-
767   291, doi:10.1038/nature23881 (2017).

768 15 Wang, X. *et al.* Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition
769   by RNA. *Nature structural & molecular biology* **24**, 1028-1038, doi:10.1038/nsmb.3487
770   (2017).

771 16 Mendenhall, E. M. *et al.* GC-rich sequence elements recruit PRC2 in mammalian ES cells.
772   *PLoS genetics* **6**, 1-10 (2010).

773 17 Ohinata, Y. *et al.* Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**,
774   207-213, doi:10.1038/nature03813 (2005).

775 18 Li, Z. *et al.* The Sm protein methyltransferase PRMT5 is not required for primordial germ cell
776   specification in mice. *EMBO J* **34**, 748-758, doi:10.15252/embj.201489319 (2015).

777 19 Soh, Y. Q. *et al.* A Gene Regulatory Program for Meiotic Prophase in the Fetal Ovary. *PLoS*
778   *genetics* **11**, e1005531, doi:10.1371/journal.pgen.1005531 (2015).

779 20 Bowles, J. & Koopman, P. Retinoic acid, meiosis and germ cell fate in mammals.
780   *Development* **134**, 3401-3411, doi:10.1242/dev.001107 (2007).

781 21 De Cecco, M. *et al.* L1 drives IFN in senescent cells and promotes age-associated
782   inflammation. *Nature* **566**, 73-78, doi:10.1038/s41586-018-0784-9 (2019).

783 22 Jackson-Grusby, L. *et al.* Loss of genomic methylation causes p53-dependent apoptosis and
784   epigenetic deregulation. *Nat Genet* **27**, 31-39, doi:10.1038/83730 (2001).

785 23 Liu, S. *et al.* Setdb1 is required for germline development and silencing of H3K9me3-marked
786   endogenous retroviruses in primordial germ cells. *Genes and Development* **28**, 2041-2055,
787   doi:10.1101/gad.244848.114 (2014).

788 24 Riising, E. M. *et al.* Gene silencing triggers polycomb repressive complex 2 recruitment to
789   CpG islands genome wide. *Mol Cell* **55**, 347-360, doi:10.1016/j.molcel.2014.06.005 (2014).

790 25 Schmitges, F. W. *et al.* Histone methylation by PRC2 is inhibited by active chromatin marks.
791   *Mol Cell* **42**, 330-341, doi:10.1016/j.molcel.2011.03.025 (2011).

792 26 Yokobayashi, S. *et al.* PRC1 coordinates timing of sexual differentiation of female primordial
793   germ cells. *Nature* **495**, 236-240, doi:10.1038/nature11918 (2013).

794 27 Malki, S., van der Heijden, G. W., O'Donnell, K. A., Martin, S. L. & Bortvin, A. A role for
795   retrotransposon LINE-1 in fetal oocyte attrition in mice. *Dev Cell* **29**, 521-533,
796   doi:10.1016/j.devcel.2014.04.027 (2014).

797 28 Saitou, M., Kagiwada, S. & Kurimoto, K. Epigenetic reprogramming in mouse pre-
798   implantation development and primordial germ cells. *Development* **139**, 15-31, doi:139/1/15
799   [pii]

800 10.1242/dev.050849 (2012).

801   29   Pezic, D., Manakov, S. A., Sachidanandam, R. & Aravin, A. A. piRNA pathway targets active
802   LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes and
803   Development* **28**, 1410-1428, doi:10.1101/gad.240895.114 (2014).
804   30   Inoue, K., Ichiyanagi, K., Fukuda, K., Glinka, M. & Sasaki, H. Switching of dominant
805   retrotransposon silencing strategies from posttranscriptional to transcriptional mechanisms
806   during male germ-cell development in mice. *PLoS genetics* **13**, e1006926,
807   doi:10.1371/journal.pgen.1006926 (2017).
808   31   Walter, M., Teissandier, A., Perez-Palacios, R. & Bourc'his, D. An epigenetic switch ensures
809   transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *Elife*
810   **5**, doi:10.7554/eLife.11418 (2016).
811   32   Frapporti, A. *et al.* The Polycomb protein Ezl1 mediates H3K9 and H3K27 methylation to
812   repress transposable elements in Paramecium. *Nat Commun* **10**, 2710, doi:10.1038/s41467-
813   019-10648-5 (2019).
814   33   Wachter, E. *et al.* Synthetic CpG islands reveal DNA sequence determinants of chromatin
815   structure. *Elife* **3**, e03397, doi:10.7554/eLife.03397 (2014).
816   34   Ehrlich, M. DNA hypomethylation in cancer cells. *Epigenomics* **1**, 239-259,
817   doi:10.2217/epi.09.33 (2009).
818   35   Jones, P. A., Issa, J. P. & Baylin, S. Targeting the cancer epigenome for therapy. *Nat Rev
819   Genet* **17**, 630-641, doi:10.1038/nrg.2016.93 (2016).
820   36   Su, I. H. *et al.* Ezh2 controls B cell development through histone H3 methylation and Igh
821   rearrangement. *Nature immunology* **4**, 124-131, doi:10.1038/ni876 (2003).
822   37   Malki, S., Tharp, M. E. & Bortvin, A. A Whole-Mount Approach for Accurate Quantitative and
823   Spatial Assessment of Fetal Oocyte Dynamics in Mice. *Biology of reproduction* **93**, 113,
824   doi:10.1095/biolreprod.115.132118 (2015).
825   38   Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence
826   data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).
827   39   Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions,
828   deletions and gene fusions. *Genome Biol* **14**, R36, doi:10.1186/gb-2013-14-4-r36 (2013).
829   40   Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping
830   by seed-and-vote. *Nucleic Acids Res* **41**, e108, doi:10.1093/nar/gkt214 (2013).
831   41   Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for
832   assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930,
833   doi:10.1093/bioinformatics/btt656 (2014).
834   42   Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
835   RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
836   43   Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-
837   seq: accounting for selection bias. *Genome Biol* **11**, R14, doi:10.1186/gb-2010-11-2-r14
838   (2010).
839   44   Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for
840   interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550,
841   doi:10.1073/pnas.0506580102 (2005).
842   45   Risso, D., Ngai, J., Speed, T. P. & Dudoit, S. Normalization of RNA-seq data using factor
843   analysis of control genes or samples. *Nat Biotechnol* **32**, 896-902, doi:10.1038/nbt.2931
844   (2014).
845   46   Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform.
846   *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).
847   47   Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome.
848   *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).
849   48   Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE Blacklist: Identification of
850   Problematic Regions of the Genome. *Sci Rep* **9**, 9354, doi:10.1038/s41598-019-45839-z
851   (2019).

852   49   Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data
853   analysis. *Nucleic Acids Res* **44**, W160-165, doi:10.1093/nar/gkw257 (2016).
854   50   Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or
855   without a reference genome. *BMC Bioinformatics* **12**, 323, doi:10.1186/1471-2105-12-323
856   (2011).
857   51   Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137,
858   doi:10.1186/gb-2008-9-9-r137 (2008).
859   52   Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak
860   annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383,
861   doi:10.1093/bioinformatics/btv145 (2015).
862   53   Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for
863   differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140,
864   doi:10.1093/bioinformatics/btp616 (2010).
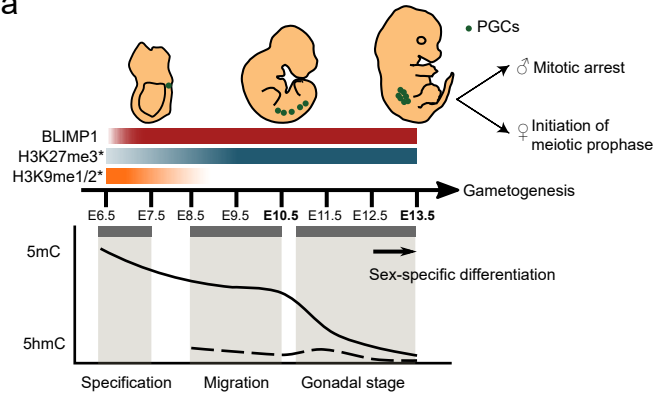865   54   Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-
866   Seq applications. *Bioinformatics* **27**, 1571-1572, doi:10.1093/bioinformatics/btr167 (2011).
867   55   Gardiner-Garden, M. & Frommer, M. CpG islands in vertebrate genomes. *J Mol Biol* **196**,
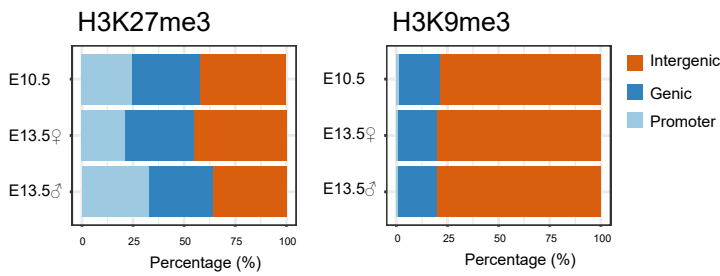868   261-282, doi:10.1016/0022-2836(87)90689-9 (1987).
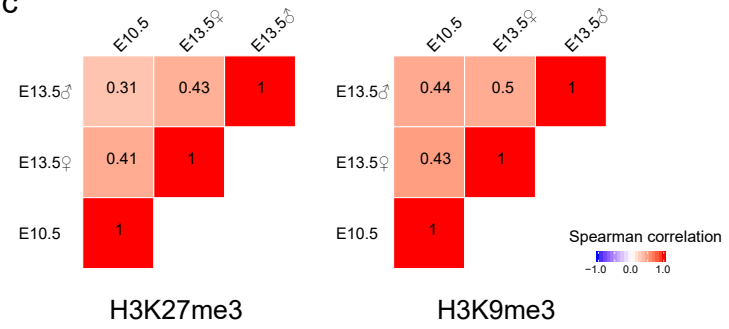
869

Figure 1

Figure 2

a

Promoters: (n=7599)

♀

5mC+5hmC   H3K27me3

E10.5 E12.5♀   E10.5 E13.5♀

5mC+5hmC   H3K27me3

Group A (n=764)

Group B (n=821)

Group C (n=2900)

Group D (n=3114)

Methylation rate

1
0.8
0.6
0.4
0.2
0

H3K27me3
log$_2$(ChIP/Input)

2
1
0
-1
-2

E10.5 E12.5   E10.5 E13.5

H3K27me3

Group A   Group B   Group C   Group D

E10.5 WT

E13.5 ♀ WT

Brdt   Zp3   Hoxc13 Hoxc12   Rbm44

b

CpG obs/exp ratio

♀ Promoter
♂ Promoter
WG 2kb Bin

P<2e-16 ***
P<2e-16 ***
P>0.99

Group A Group B Group C Group D All promoters WG 2kb bin

CpG enrichment

Percentage of CpG dinucleotide (%)

♀ promoter
♂ promoter
WG.2kb Bin

P=1e-07 ***
P<2e-16 ***
P<2e-16 ***

Group A Group B Group C Group D All promoters WG 2kb bin

CpG content

Percentage of C or G (%)

♀ Promoter
♂ Promoter
WG 2kb Bin

P=2.3e-15 ***
P<2e-16 ***
P<2e-16 ***

Group A Group B Group C Group D All promoters WG 2kb bin

G+C content

c

DNA methylation

E10.5   E12.5♀

Methylation rate

0.8
0.6
0.4
0.2
0

H3K27me3

E10.5   E13.5♀

H3K27me3
log$_2$(ChIP/Input)

6
3
0
-3
-6

Percentage of CpG (%)

Percentage of C or G (%)

Figure 3

a

	OCT4	EZH2	Merge

Ctrl E11.5 / CKO

Ctrl E13.5 ♂ / CKO

b

	PGC marker	H3K27me3	Merge

Ctrl E11.5 / CKO

Ctrl E13.5 ♀ / CKO

c

PC2: 7% variance vs PC1: 78% variance

- E13.5 ♂ Ctrl
- E13.5 ♂ CKO
- E13.5♀ Ctrl
- E13.5♀ CKO

d

Number of differentailly expressed genes (CKO vs Ctrl, adj. P<0.05)

Up-regulation / Down-regulation

Fold change: FC 1.5, FC 2.0

E13.5 ♀   E13.5 ♂

e

CKO vs Ctrl DE (adj. P<0.01, FC>1)

- cell differentiation — 227/2940
- cellular developmental process — 239/3224
- multi-organism reproductive process — 67/646
- gamete generation — 51/475
- cell migration — 81/947

-log10(P-value)

E13.5 ♀ CKO vs Ctrl DE (adj. P<0.05, FC>2)

- multi-organism reproductive process — 55/627
- antigen processing and presentation — 14/51
- multicellular organismal process — 225/4624
- reproduction — 68/985
- cell differentiation — 153/2885

-log10(P-value)

f

E13.5 ♀ CKO vs Ctrl DE (adj. P<0.05)

1914 / 31 / 70

P=4X10⁻¹⁴

Page_meiotic propase (104)

5mC+5hmC | H3K27me3 | Ezh2 Ctrl | Ezh2 CKO

E10.5 E12.5♀ E12.5♂ E10.5 E13.5♀ E13.5♂ E13.5♀ E13.5♀

Methylation rate
H3K27me3 log₂(ChIP/Input)
Ezh2 RNA-Seq (z−score)

g

E18.5 ♀   Ctrl / CKO

DAPI / MVH

Total number of germ cells (MVH+)/ Ovary

* P=0.0246

6225 / 3930 / 4089.5 / 1545.5

E16.5 Ctrl   E16.5 CKO   E18.5 Ctrl   E18.5 CKO

Figure 4

Extended Data Figure 1

Extended Data Figure 2

a Methylation Rate distribution for all promoters

b H3K27me3 intensity distribution for all promoters

Extended Data Figure 3

Extended Data Figure 4

a

b

EZH2 | SANT | SANT | CXXC | SET | 746 a.a.

Exons

Ezh2 Floxed locus

15 | *Loxp* | 16 | 17 | 18 | 19 | *Loxp* | 20

gen-Enx1 F   gen-Enx1 R

Cre-mediated deletion

Exons

Ezh2 Floxed locus

15 | *Loxp* | 20

5' left loxP   3' loxP

c

♀*Ezh2* f/f X ♂Tg(*Blimp1-Cre*)+/-

↓

♂*Ezh2* f/+,Tg(*Blimp1-Cre*)+/-

♀*Ezh2* f/f X ♂*Ezh2* f/+,Tg(*Blimp1-Cre*)+/-

↓

Fetus genotype   Germline deletion

Ezh2 f/△,Tg(*Blimp1-Cre*)+/- → **Ezh2△/△,Tg(*Blimp1-Cre*)+/-**

Ezh2 f/+,Tg(*Blimp1-Cre*)+/-

Ezh2 f/△

Ezh2 f/+

d

#1 #2

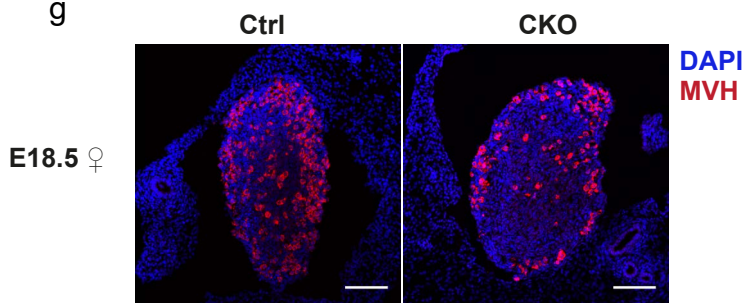440 bp — Floxed allele
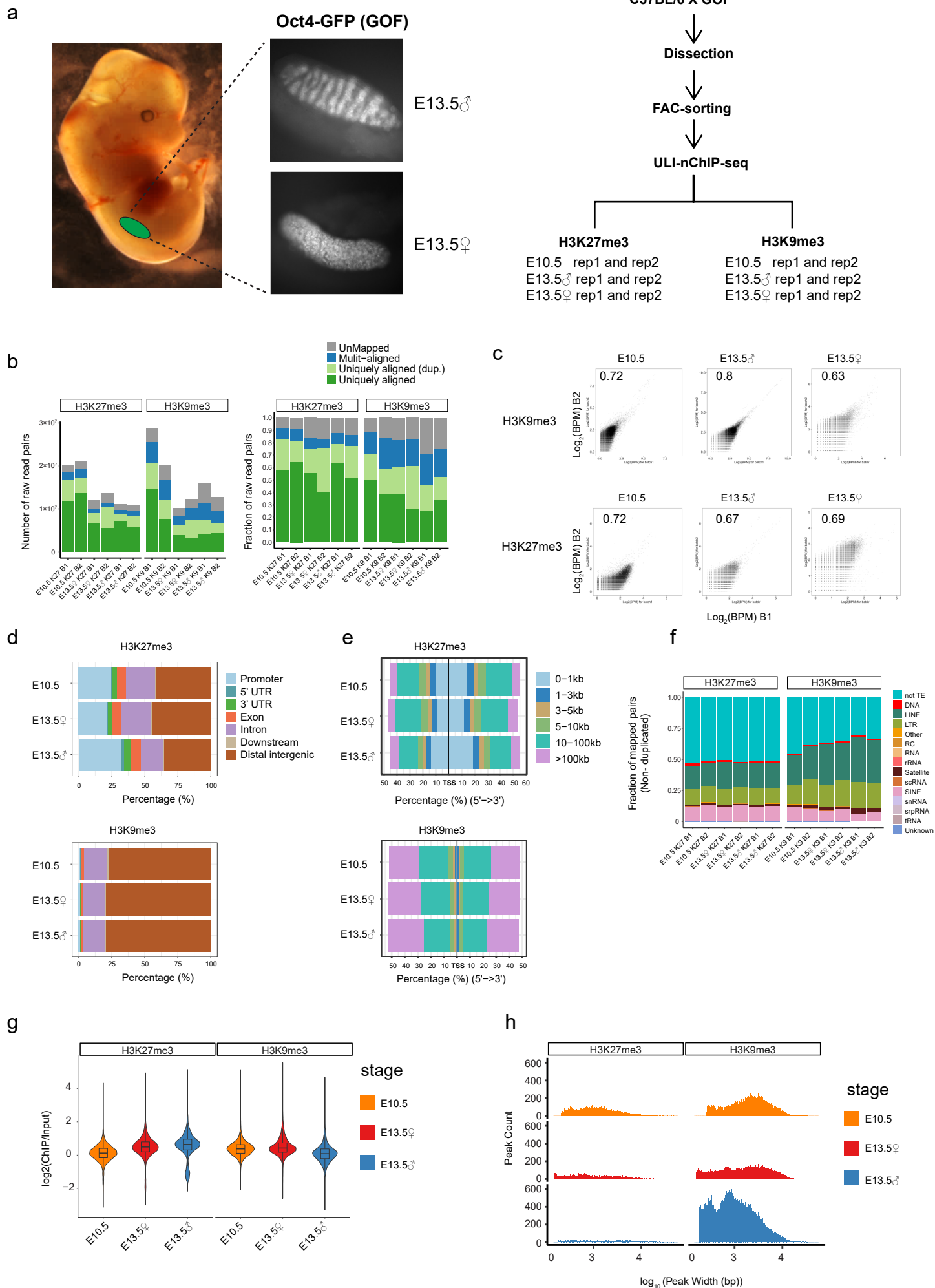410 bp — Wt allele)

370 bp — Deleted allele(△)

200 bp — *Blimp1-Cre*

Extended Data Figure 5

Extended Data Figure 6

a

PRC2.1

PCL1/PHF1
PCL2/MTF2
PCL3/PHF19

LCOR/PALI1   PCL

EPOP   SUZ12   EED

EZH1/2

RBBP4/7   EZHIP

PRC2.2

SUZ12   EED   AEBP2

EZH1/2   JARID2

RBBP4/7   EZHIP

b

Ezh1   *Ezh2*   ♂ PGC   ♀ PGC

Eed   *Suz12*

Rbbp4   *Rbbp7*

Jarid2   *Aebp2*   ♂ PGC   ♀ PGC

Mtf2   *Phf1*

Phf19   *Ezhip*

Epop   *Lcor/Pali1*

c

Ezh1

○ *Ezh2* Ctrl
● *Ezh2* CKO

E13.5♂   E13.5♀

d

Sample Distance Matrix

E13.5 ♂ CKO
E13.5 ♂ Ctrl
E13.5♀ Ctrl
E13.5♀ CKO

E13.5♀ CKO   E13.5♀ Ctrl   E13.5♂ Ctrl   E13.5♂ CKO

e

PCA

● WT
● E13.5 ♂ Ctrl
● E13.5 ♂ CKO
● E13.5♀ Ctrl
● E13.5 ♀ CKO

E10.5 WT
E12.5 WT   E12.5 ♀ WT   E13.5 ♀
E13.5 ♂
E14.5 ♀ WT
E14.5 ♂ WT

PC1: 21.09% variance
PC2: 20.83% variance

Extended Data Figure 7

a E13.5 ♂ CKO vs Ctrl DE (adj. P<0.05, FC>2)

spermatid development — 9/115
phosphatidylinositol 3-kinase signaling — 7/73
multi-organism reproductive process — 22/588
spermatogenesis — 15/348
multi-organism process — 34/1168

-log10(P-value)

b chr6:34880000-34890000

H3K27me3: E10.5 WT, E13.5 ♀ WT, E13.5 ♂ WT
RNA: E13.5 ♀ Ctrl, E13.5 ♀ CKO, E13.5 ♂ Ctrl, E13.5 ♂ CKO

Stra8

c Meiotic DE genes

p = 0.6147
P= 0.3064
P= 0.493

H3K27me3 log₂(ChIP/Input): E10.5, E13.5 ♀, E13.5 ♂

**
P = 0.008534
P = 0.2868

Gene expression (log2(TPM)): E13.5 ♀ Ctrl, E13.5 ♀ CKO, E13.5 ♂ Ctrl, E13.5 ♂ CKO

Methylation rate
H3K27me3 log₂(ChIP/Input)
WT.RNAseq (TPM)
Expression (z-score)

d 104 meiosis prophase genes

♀
Group A — P=7.89E-07 ***
Group B — P=0.001104 **
Group C — P=0.971044
Group D — P=0.012323 *

Odds Ratio

♂
Group A — P=2.66E-05 ***
Group B — P=0.069418
Group C — P=0.961657
Group D — P=0.033084 *

Odds Ratio

e

Name	Motif	P-value
RARB		9.42E-13
RXRA		3.83E-12
RARG		4.82E-12
ESRRA		2.22E-11
MSX2		2.69E-09
CREB1		0.00017
CEBPB		0.00187
CEBPD		0.0291
CDX2		0.043
MSX1		0.047

f Rarb, Rxra, Rarg, Esrra, Msx2, Creb1, Cebpb, Cebpd, Cdx2, Msx1
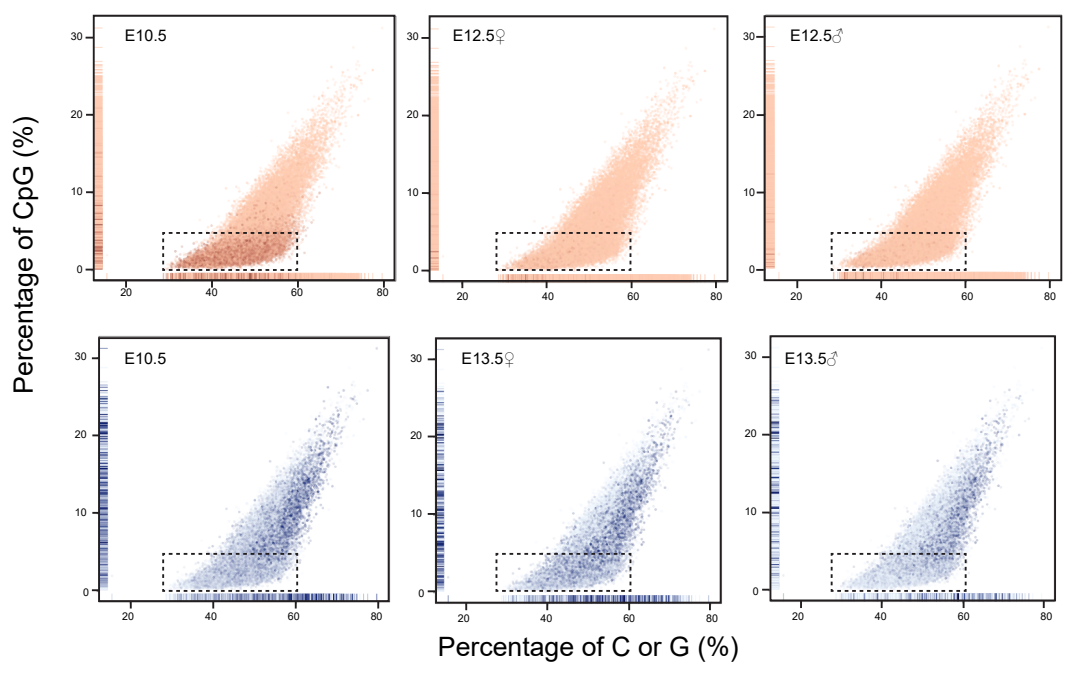
g GRR genes

z-score

Extended Data Figure 8

Extended Data Figure 9

a

E13.5♂ CKO vs Ctrl

LINE

L1P5
L1MA6

LTR

MER34

IAP-d-int

log2FC

log10 P-Value

b

RNA-seq: retrotransposon
Uniquely mapped reads only

Leading logFC dim1

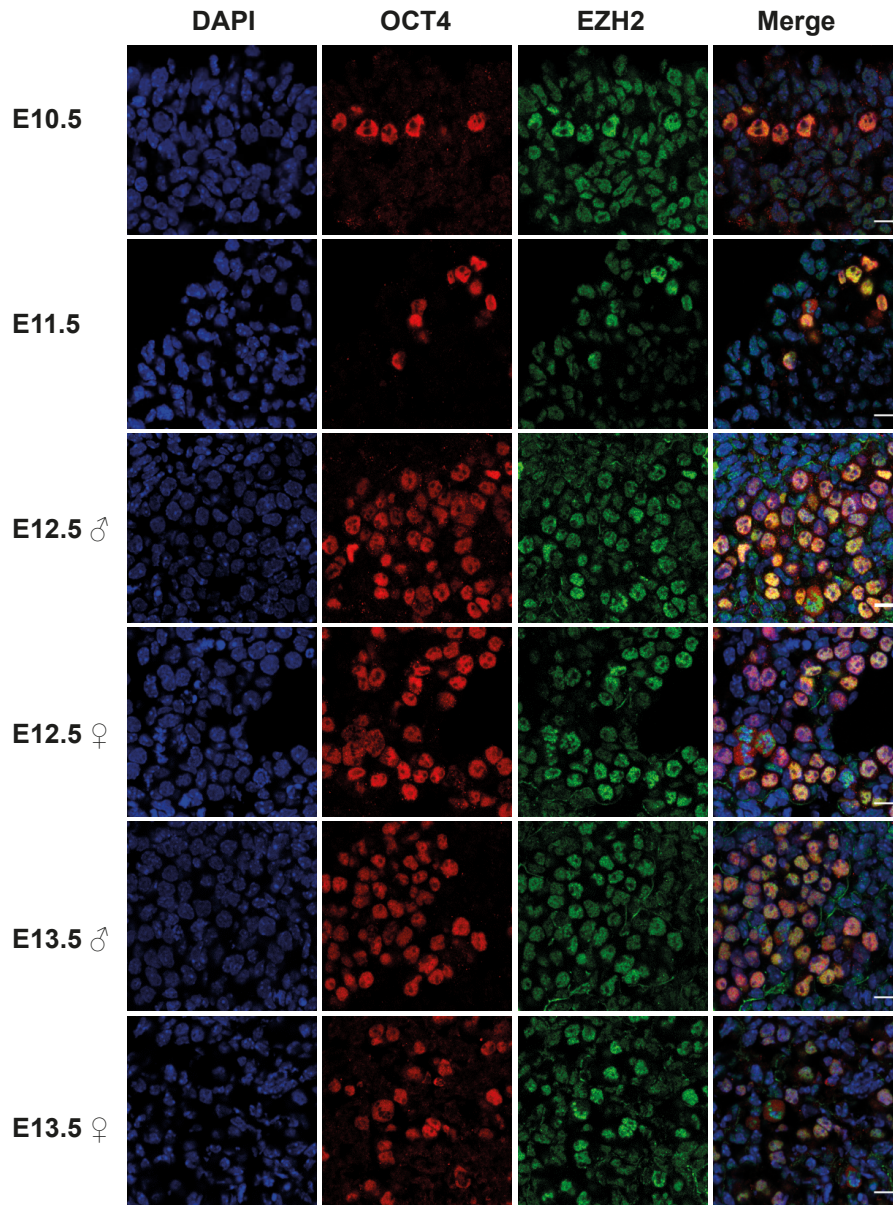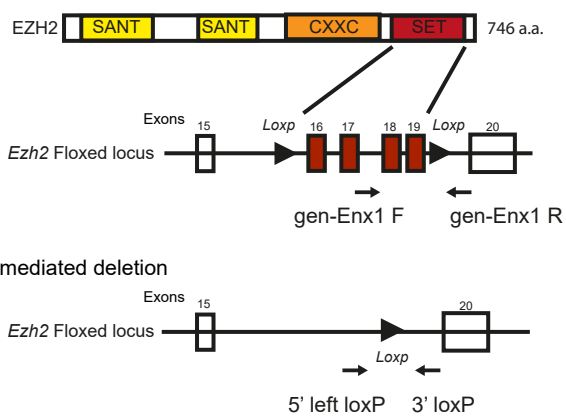Leading logFC dim 2

- E13.5 ♂ Ctrl
- E13.5 ♂ CKO
- E13.5♀ Ctrl
- E13.5♀ CKO

c

COAGULATION
EPITHELIAL_MESENCHYMAL_TRANSITION
ALLOGRAFT_REJECTION
SPERMATOGENESIS
IL2_STAT5_SIGNALING
KRAS_SIGNALING_DN
INFLAMMATORY_RESPONSE
P53_PATHWAY
XENOBIOTIC_METABOLISM
TNFA_SIGNALING_VIA_NFKB
UV_RESPONSE_DN
COMPLEMENT
ESTROGEN_RESPONSE_LATE
KRAS_SIGNALING_UP
ESTROGEN_RESPONSE_EARLY
CHOLESTEROL_HOMEOSTASIS
HYPOXIA
MYOGENESIS
TGF_BETA_SIGNALING
INTERFERON_ALPHA_RESPONSE
MYC_TARGETS_V1
G2M_CHECKPOINT
E2F_TARGETS
MYC_TARGETS_V2
UNFOLDED_PROTEIN_RESPONSE
MITOTIC_SPINDLE
MTORC1_SIGNALING

NES

No. of genes
- 20
- 40
- 60
- 80

FDR q value
0.20
0.15
0.10
0.05
0.00

d

P53 pathway

Enrichment plot: HALLMARK_P53_PATHWAY

Interferon alpha response

Enrichment plot:
HALLMARK_INTERFERON_ALPHA_RESPONSE

e

Distance of H3K9me3 peaks to TSS

E13.5♂ De novo peaks

E13.5♀ De novo peaks

Percentage (%) (5'−>3')

- 0−1kb
- 1−3kb
- 3−5kb
- 5−10kb
- 10−100kb
- >100kb

Distance of H3K9me3 peaks to TEs

E13.5♂ De novo peaks

E13.5♀ De novo peaks

Percentage (%) (5'−>3')

- 0−1kb
- 1−3kb
- 3−5kb
- 5−10kb
- 10−100kb
- >100kb

f

chr3:125830kb-125840kb

E10.5 WT

Peak calling

Broad_E10_KG_S113_WithInput_E10_KG_INPUT_S125_peak_1709

E13.5 ♀ WT

Peak calling

Broad_E13_F_KG_S122_WithInput_E13_F_INPUT_S131_peak_6232

E13.5 ♂ WT
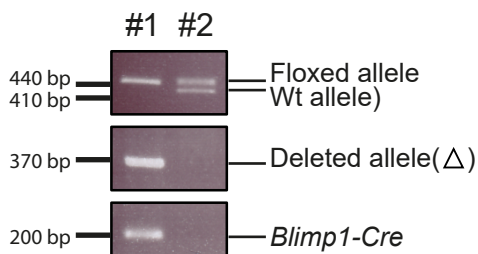
Peak calling

_M_INPUT_S130_peak_30996   Broad_E13_M_KG_S119_WithInput_E13_M_INPUT_S130_peak_30997

IAP Ez-int

Extended Data Figure 10

a

H3K9me3 enriched   H3K27me3 enriched

b

Mutually exclusive
(uniquely mapped, single copy TEs)

IAPLTR2_Mm

c

Co-localised
(uniquely mapped, single copy TEs)

L1Md_Gf

d

Rnf2 CKO up FC1.5, P<0.05   Ezh2 CKO up P<0.05

450 | 74 | 851

e

E13.5♀ PGCs
Ezh2 CKO
Upregulated TEs   E13.5♀ PGCs
Setdb1 CKO
upregulated TEs

20 | 3 | 14

IAPLTR1_Mm
IAPEz.int
MMERVK10C.int

E13.5♀ PGCs
Ezh2 CKO
Upregulated TEs   ES cells 2i+VitC
Eed KO
upregulated TEs at D13

18 | 5 | 18

IAPEz.int
MMERVK10C.int
L1Md_A
L1Md_Gf
L1Md_T

f

Meiosis related genes

Transposable elements