

1      The X-linked intellectual disability gene KDM5C is a  
2      sex-biased brake against germline programs in somatic  
3      lineages

4

5      Katherine M. Bonefas<sup>1,2</sup>, Ilakkiya Venkatachalam<sup>2,3</sup>, and Shigeki Iwase<sup>2</sup>.

6      1. Neuroscience Graduate Program, University of Michigan Medical School, Ann Arbor, MI, 48109, USA.

7      2. Department of Human Genetics, Michigan Medicine, University of Michigan Medical School, Ann Arbor,  
8      MI, 48109, USA.

9      3. Genetics and Genomics Graduate Program, University of Michigan, Ann Arbor, MI, 48109, USA.

10     Correspondence should be addressed to K. Bonefas and S. Iwase (siwase@umich.edu)

## 11 Abstract

12 A pivotal step in the evolution of multicellularity is the division labor among cellular lineages, including the  
13 the distinction of the germline from the soma. In the early embryo, genes that establish germline identity are  
14 repressed in somatic lineages through DNA and histone modifications. Failure to repress germline genes  
15 in somatic lineages is common signature of cancer and observed in select neurodevelopmental disorders,  
16 however it is currently unclear how factors like development and sex influence their repression and somatic  
17 misexpression. Here, we examine how cellular context influences the development of somatic tissue identity  
18 in mice with loss of lysine demethylase 5c (KDM5C), an eraser of histone 3 lysine 4 di and tri-methylation  
19 (H3K4me2/3). We found KDM5C is a crucial regulator of tissue identity, as male *Kdm5c* knockout (-KO) mice  
20 aberrantly express many liver, muscle, ovary, and testis genes within the brain. Late-stage spermatogenesis  
21 genes, but not somatic testicular genes, were highly enriched within the *Kdm5c*-KO brain, indicating an  
22 erosion of soma-germline boundary. Germline genes are typically decommissioned in the post-implantation  
23 epiblast, yet *Kdm5c*-KO epiblast-like cells (EpiLCs) aberrantly expressed key regulators of germline identity  
24 and meiosis, including *Dazl* and *Stra8*. Characterizing germline gene misexpression in males and female  
25 mutants revealed germline gene repression is sexually dimorphic, with female EpiLCs requiring a higher dose  
26 of KDM5C to maintain germline gene suppression. Using a comprehensive list of mouse germline-enriched  
27 genes, we found KDM5C is selectively recruited to a subset of germline gene promoters that contain CpG  
28 islands (CGIs) to facilitate DNA CpG methylation during ESC to EpiLC differentiation. However, late-stage  
29 spermatogenesis genes devoid of promoter CGIs can become expressed in *Kdm5c*-KO cells via ectopic  
30 activation by RFX transcription factors. Together, these data demonstrate KDM5C's fundamental role in  
31 tissue identity and indicate that KDM5C acts as a brake against runaway activation of germline developmental  
32 programs in somatic lineages.

## 33 notes

- 34 • Distinguishing the germline from the soma is a key step in the evolution of multicellularity and sexual  
35 reproduction.
- 36 • Germline gene repression is orchestrated by chromatin regulators and transcription factors.
- 37 • Much of these discoveries have been made looking at the repression of key marker genes for germ cell  
38 specification in the early embryo,
- 39 • While the silencing mechanisms for genes that establish germline identity are well characterized, it is  
40 unclear if other types of germline genes, such as those involved in late oogenesis and spermatogenesis,  
41 employ the same silencing mechanisms and how silencing changes over the course of development.
- 42 • In mammals, genes crucial for early germline specification gain repressive DNA and histone modifica-  
43 tions in the early embryo.

- 44        • Germline genes have not been comprehensively assessed as a whole

## 45 Introduction

46 A single genome holds the instructions to generate the myriad of cell types found within an organism.  
47 This is, in part, accomplished by chromatin regulators that can either promote or impede lineage-specific  
48 gene expression through DNA and histone modifications<sup>1–5</sup>. Human genetic studies revealed mutations in  
49 chromatin regulators are a major cause of neurodevelopmental disorders (NDDs)<sup>6</sup> and many studies have  
50 identified their importance for regulating brain-specific transcriptional programs. Loss of chromatin regulators  
51 can also result in the ectopic expression of tissue-specific genes outside of their target environment, such  
52 as the misexpression of liver-specific genes within adult neurons<sup>7</sup>. However, the mechanisms underlying  
53 ectopic gene expression and its impact upon neurodevelopment are poorly understood.

54 To elucidate the role of tissue identity in chromatin-linked NDDs, it is essential to first characterize the  
55 nature of the dysregulated genes and the molecular mechanisms governing their de-repression. Here we  
56 focus on the X chromosome gene lysine demethylase 5C (KDM5C, also known as SMCX or JARID1C),  
57 which erases histone 3 lysine 4 di- and trimethylation (H3K4me2/3), a permissive chromatin modification  
58 enriched at gene promoters<sup>8</sup>. Pathogenic mutations in *KDM5C* cause Intellectual Developmental Disorder,  
59 X-linked, Syndromic, Claes-Jensen Type (MRXSCJ, OMIM: 300534). MRXSCJ is more common and severe  
60 in males and its neurological phenotypes include intellectual disability, seizures, aberrant aggression, and  
61 autistic behaviors<sup>9–11</sup>. Male *Kdm5c* knockout (-KO) mice recapitulate key MRXSCJ phenotypes, including  
62 hyperaggression, increased seizure propensity, social deficits, and learning impairments<sup>12–14</sup>. RNA sequenc-  
63 ing (RNA-seq) of the *Kdm5c*-KO hippocampus revealed ectopic expression of some germline genes within  
64 the brain<sup>13</sup>. However, it is unclear if other tissue-specific genes are aberrantly transcribed with KDM5C loss,  
65 at what point in development germline gene misexpression begins, and what mechanisms underlie their  
66 dysregulation.

67 Distinguishing between germ cells and somatic cells is a key feature of multicellularity<sup>15</sup> that occurs  
68 during early embryogenesis in many metazoans<sup>16</sup>. In mammals, chromatin regulators are crucial for  
69 decommissioning germline genes during the transition from naïve to primed pluripotency. Initially, germline  
70 gene promoters gain repressive histone H2A lysine 119 monoubiquitination (H2AK119ub1)<sup>17</sup> and histone H3  
71 lysine 9 trimethylation (H3K9me3)<sup>17,18</sup> in embryonic stem cells (ESCs) and are then decorated with DNA  
72 CpG methylation (CpGme) in post-implantation epiblast cells<sup>18–21</sup>. The contribution of KDM5C to this process  
73 remains unclear. Additionally, studies on germline gene repression have primarily been conducted in males  
74 and focused on select marker genes, given the lack of a comprehensive list for germline-enriched genes.  
75 Therefore, it is unknown if the mechanism of repression differs between sexes or for different classes of  
76 germline genes, e.g. meiotic versus spermatid differentiation genes.

77 To illuminate KDM5C's role in tissue identity, here we characterized the aberrant expression of tissue-  
78 enriched genes within the *Kdm5c*-KO brain and epiblast-like stem cells (EpiLCs), an *in vitro* model of the  
79 post-implantation embryo. We curated a list of mouse germline-enriched genes, which enabled genome-wide

80 analysis of germline gene silencing mechanisms for the first time. Additionally, we characterized germline  
81 transcripts expressed in male and female *Kdm5c* mutants to illuminate the impact of sex upon germline  
82 gene suppression. Based on the data presented below, we propose KDM5C plays a fundamental, sexually  
83 dimorphic role in the development of tissue identity during early embryogenesis, including the establishment  
84 of the soma-germline boundary.

## 85 **Results**

### 86 **Tissue-enriched genes are aberrantly expressed in the *Kdm5c*-KO brain**

87 Previous RNA sequencing (RNA-seq) of the adult male *Kdm5c*-KO hippocampus revealed ectopic  
88 expression of some germline genes unique to the testis<sup>13</sup>. It is currently unknown if the testis is the only  
89 tissue type misexpressed in the *Kdm5c*-KO brain. We thus systematically tested whether other tissue-specific  
90 genes are misexpressed in the male brain with constitutive knockout of *Kdm5c* (*Kdm5c*<sup>-y</sup>, 5CKO in figures)<sup>22</sup>  
91 by using a published list of mouse tissue-enriched genes<sup>23</sup>.

92 We found a large proportion of significantly upregulated genes (DESeq2<sup>24</sup>, log2 fold change > 0.5, q <  
93 0.1) within the male *Kdm5c*-KO amygdala and hippocampus are non-brain, tissue-specific genes (Amygdala:  
94 0/0 up DEGs, NaN% ; Hippocampus: 0/0 up DEGs, NaN%) (Figure 1A-B, Supplementary Table 1). For both  
95 the amygdala and hippocampus, the majority of tissue-enriched differentially expressed genes (DEGs) were  
96 testis genes (Figure 1A-B). Even though the testis has the largest total number of tissue-enriched genes  
97 (2,496 genes) compared to any other tissue, testis-enriched DEGs were significantly enriched in both brain  
98 regions (Amygdala p = 1.83e-05, Odds Ratio = 5.13; Hippocampus p = 4.26e-11, Odds Ratio = 4.45, Fisher's  
99 Exact Test). An example of a testis-enriched gene misexpressed in the *Kdm5c*-KO brain is *FK506 binding*  
100 *protein 6 (Fkbp6)*, a known regulator of PIWI-interacting RNAs (piRNAs) and meiosis<sup>25,26</sup> (Figure 1C).

101 Interestingly, we also observed significant enrichment of ovary-enriched genes in both the amygdala  
102 and hippocampus (Amygdala p = 0.00574, Odds Ratio = 18.7; Hippocampus p = 0.048, Odds Ratio = 5.88,  
103 Fisher's Exact Test) (Figure 1A-B). Ovary-enriched DEGs included *Zygotic arrest 1 (Zar1)*, which sequesters  
104 mRNAs in oocytes for meiotic maturation<sup>27</sup> (Figure 1D). Given that the *Kdm5c*-KO mice we analyzed are  
105 male, these data demonstrate that the ectopic expression of gonad-enriched genes is independent of  
106 organismal sex.

107 Although not consistent across brain regions, we also found significant enrichment of genes biased  
108 towards two non-gonadal tissues - the liver (Amygdala p = 0.04, Odds Ratio = 6.58, Fisher's Exact Test)  
109 and muscle (Hippocampus p = 0.01, Odds Ratio = 6.95, Fisher's Exact Test) (Figure 1A-B). These include  
110 *Apolipoprotein C-I (Apoc1)*, a lipoprotein metabolism and transport gene<sup>28</sup> (Figure 1E, see Discussion).

111 Our analysis of oligo(dT)-primed libraries<sup>22</sup> indicates aberrantly expressed mRNAs are polyadenylated  
112 and spliced into mature transcripts in the *Kdm5c*-KO brain (Figure 1C-E). Of note, we observed little to no

113 dysregulation of brain-enriched genes (Amygdala p = 1, Odds Ratio = 1.22; Hippocampus p = 0.74, Odds  
114 Ratio = 1.22, Fisher's Exact Test), despite the fact these are brain samples and the brain has the second  
115 highest total number of tissue-enriched genes (708 genes). Altogether, these results suggest the aberrant  
116 expression of tissue-enriched genes within the brain is a major effect of KDM5C loss.

117 **Germline genes are misexpressed in the *Kdm5c*-KO brain**

118 *Kdm5c*-KO brain expresses testicular germline genes<sup>13</sup> (Figure 1), however the testis also contains  
119 somatic cells that support hormone production and germline functions. To determine if *Kdm5c*-KO results  
120 in ectopic expression of testicular somatic genes, we first evaluated the known functions of testicular  
121 DEGs through gene ontology. We found *Kdm5c*-KO testis-enriched DEGs had high enrichment of germline-  
122 relevant ontologies, including spermatid development (GO: 0007286, p.adjust = 6.2e-12) and sperm axoneme  
123 assembly (GO: 0007288, p.adjust = 2.45e-14) (Figure 2A, Supplementary Table 1).

124 We then evaluated *Kdm5c*-KO testicular DEG expression in wild-type testes versus testes with germ cell  
125 depletion<sup>29</sup>, which was accomplished by heterozygous *W* and *Wv* mutations in the enzymatic domain of *c-Kit*  
126 (*Kit*<sup>W/Wv</sup>)<sup>30</sup>. Almost all *Kdm5c*-KO testis-enriched DEGs lost expression with germ cell depletion (Figure 2B).  
127 We then assessed testis-enriched DEG expression in a published single cell RNA-seq dataset that identified  
128 cell type-specific markers within the testis<sup>31</sup>. Some *Kdm5c*-KO testis-enriched DEGs were classified as  
129 specific markers for different germ cell developmental stages (e.g. spermatogonia, spermatocytes, round  
130 spermatids, and elongating spermatids), yet none marked somatic cells (Figure 2C). Together, these data  
131 demonstrate that the *Kdm5c*-KO brain aberrantly expresses germline genes but not somatic testicular genes,  
132 reflecting an erosion of the soma-germline boundary.

133 As of yet, research on germline gene silencing mechanisms has focused on a handful of key genes rather  
134 than assessing germline gene suppression genome-wide, due to the lack of a comprehensive gene list.  
135 We therefore generated a list of mouse germline-enriched genes using RNA-seq datasets of *Kit*<sup>W/Wv</sup> mice  
136 that included males and females at embryonic day 12, 14, and 16<sup>32</sup> and adult male testes<sup>29</sup>. We defined  
137 genes as germline-enriched if their expression met the following criteria: 1) their expression is greater than  
138 1 FPKM in wild-type gonads 2) their expression in any non-gonadal tissue of adult wild type mice<sup>23</sup> does  
139 not exceed 20% of their maximum expression in the wild-type germline, and 3) their expression in the germ  
140 cell-depleted gonads, for any sex or time point, does not exceed 20% of their maximum expression in the  
141 wild-type germline. These criteria yielded 1,288 germline-enriched genes (Figure 2D), which was hereafter  
142 used as a resource to globally characterize germline gene misexpression with *Kdm5c* loss (Supplementary  
143 Table 2).

144 ***Kdm5c*-KO epiblast-like cells aberrantly express key regulators of germline identity**

145 Germ cells are typically distinguished from somatic cells soon after the embryo implants into the uterine  
146 wall<sup>33,34</sup>, when germline genes are silenced in epiblast stem cells that will form the somatic tissues<sup>35</sup>. This  
147 developmental time point can be modeled *in vitro* through differentiation of naïve embryonic stem cells  
148 (nESCs) into epiblast-like stem cells (EpiLCs) (Figure 3A)<sup>36,37</sup>. While some germline-enriched genes are  
149 also expressed in nESCs and in the 2-cell stage<sup>38–40</sup>, they are silenced as they differentiate into EpiLCs<sup>18,19</sup>.  
150 Therefore, we tested if KDM5C was necessary for the initial silencing of germline genes in somatic lineages  
151 by evaluating the impact of *Kdm5c* loss in male EpiLCs.

152 *Kdm5c*-KO cell morphology during ESC to EpiLC differentiation appeared normal (Figure 3B) and EpiLCs  
153 properly expressed markers of primed pluripotency, such as *Dnmt3b*, *Fgf5*, *Pou3f1*, and *Otx2* (Figure 3C). We  
154 then identified tissue-enriched DEGs in a RNA-seq dataset of wild-type and *Kdm5c*-KO EpiLCs<sup>41</sup> (DESeq2,  
155 log<sub>2</sub> fold change > 0.5, q < 0.1, Supplementary Table 3). Similar to the *Kdm5c*-KO brain, we observed  
156 general dysregulation of tissue-enriched genes, with the largest number of genes belonging to the brain and  
157 testis, although they were not significantly enriched (Figure 3D). Using our list of mouse germline-enriched  
158 genes assembled above, we identified 68 germline genes misexpressed in male *Kdm5c*-KO EpiLCs.

159 We then compared EpiLC germline DEGs to those expressed in the *Kdm5c*-KO brain to determine if  
160 germline genes are constitutively dysregulated or change over the course of development. The majority of  
161 germline DEGs were unique to either EpiLCs or the brain, with only *D1Pas1* and *Cyct* shared across all  
162 tissue/cell types (Figure 3E-F). EpiLC germline DEGs had particularly high enrichment of meiosis-related  
163 gene ontologies when compared to the brain (Figure 3G, Supplementary Table 3), such as meiotic cell  
164 cycle process (GO:1903046, p.adjust = 2.2e-07) and meiotic nuclear division (GO:0140013, p.adjust  
165 = 1.37e-07). While there was modest enrichment of meiotic gene ontologies in both brain regions, the  
166 *Kdm5c*-KO hippocampus primarily expressed late-stage spermatogenesis genes involved in sperm axoneme  
167 assembly (GO:0007288, p.adjust = 0.00621) and sperm motility (GO:0097722, p.adjust = 0.00612).

168 Notably, DEGs unique to *Kdm5c*-KO EpiLCs included key drivers of germline identity, such as *Stimulated*  
169 *by retinoic acid 8* (*Stra8*: log<sub>2</sub> fold change = 3.73, q = 2.17e-39) and *Deleted in azoospermia like* (*Dazl*:  
170 log<sub>2</sub> fold change = 3.36, q = 3.19e-12) (Figure 3H). These genes are typically expressed when a subset  
171 of epiblast stem cells become primordial germ cells (PGCs) and then again in mature germ cells to trigger  
172 meiotic gene expression programs<sup>42–44</sup>. Of note, some germline genes, including *Dazl*, are also expressed  
173 in the two-cell embryo<sup>39,45</sup>. However, we did not see derepression of two-cell stage-specific genes, like  
174 *Duxf3* (*Dux*) (log<sub>2</sub> fold change = -0.282, q = 0.337) and *Zscan4d* (log<sub>2</sub> fold change = 0.25, q = 0.381) (Figure  
175 3H, Supplementary Table 3), indicating *Kdm5c*-KO EpiLCs do not revert back to a 2-cell state. Altogether,  
176 *Kdm5c*-KO EpiLCs express key drivers of germline identity and meiosis while the brain primarily expresses  
177 spermiogenesis genes, indicating germline gene misexpression mirrors germline development during the  
178 progression of somatic development.

179 **Female epiblast-like cells have heightened germline gene misexpression with *Kdm5c***  
180 **loss**

181 It is currently unknown if the misexpression of germline genes is influenced by sex, as previous studies  
182 on germline gene repressors have focused on male cells<sup>17,18,20,46,47</sup>. Sex is particularly pertinent in the case  
183 of KDM5C because it partially escapes X chromosome inactivation (XCI), resulting in a higher dosage in  
184 females<sup>48–51</sup>. We therefore explored the impact of chromosomal sex upon germline gene suppression by  
185 comparing their dysregulation in male *Kdm5c* hemizygous knockout (*Kdm5c*<sup>-y</sup>, XY *Kdm5c*-KO, XY 5CKO),  
186 female homozygous knockout (*Kdm5c*<sup>-/-</sup>, XX *Kdm5c*-KO, XX 5CKO), and female heterozygous knockout  
187 (*Kdm5c*<sup>-/+</sup>, XX *Kdm5c*-HET, XX 5CHET) EpiLCs<sup>41</sup>.

188 In EpiLCs, homozygous and heterozygous *Kdm5c* knockout females expressed over double the number  
189 of germline-enriched genes than hemizygous males (Figure 4A, Supplementary Table 3). While the majority  
190 of germline DEGs in *Kdm5c*-KO males were also dysregulated in females (74%), many were sex-specific,  
191 such as *Tktl2* and *Esx1* (Figure 4B). We then compared the known functions of germline genes dysregulated  
192 uniquely in males and females or misexpressed in all samples (Figure 4C, Supplementary Table 3). Female-  
193 specific germline DEGs were enriched for meiotic (GO:0051321 - meiotic cell cycle, p.adjust = 7.81E-14) and  
194 flagellar (GO:0003341 - cilium movement, p.adjust = 4.87E-06) functions, while male-specific DEGs had roles  
195 in mitochondrial and cell signaling (GO:0070585 - protein localization to mitochondrion, p.adjust = 0.025).

196 The majority of germline genes expressed in both sexes were more highly dysregulated in females  
197 compared to males (Figure 4D-F). This increased degree of dysregulation in females, along with the  
198 increased total number of germline genes, indicates females are more sensitive to losing KDM5C-mediated  
199 germline gene suppression. Heightened germline gene dysregulation in females could be due to impaired  
200 XCI in *Kdm5c* mutants<sup>41</sup>, as many spermatogenesis genes lie on the X chromosome<sup>52,53</sup>. However, female  
201 germline DEGs were not biased towards the X chromosome (p = 1, Odds Ratio = 0.96, Fisher's Exact Test)  
202 and females had a similar overall proportion of germline DEGs belonging to the X chromosome as males  
203 (XY *Kdm5c*-KO - 10.29%, XX *Kdm5c*-HET - 7.43%, XX *Kdm5c*-KO - 10.59%) (Figure 4G). The majority of  
204 germline DEGs instead lie on autosomes for both male and female *Kdm5c* mutants (Figure 4G). Thus, while  
205 female EpiLCs are more prone to germline gene misexpression with KDM5C loss, it is likely independent of  
206 XCI defects.

207 **Germline gene misexpression in *Kdm5c* mutants is independent of germ cell sex**

208 Although many germline genes have shared functions in the male and female germline, e.g. PGC  
209 formation, meiosis, and genome defense, some have unique or sex-biased expression. Therefore, we  
210 wondered if *Kdm5c* mutant males would primarily express sperm genes while mutant females would primarily  
211 express egg genes. To comprehensively assess whether germline gene sex corresponds with *Kdm5c*  
212 mutant sex, we first filtered our list of germline-enriched genes for egg and sperm-biased genes (Figure 4,

213 Supplementary Table 2). We defined germ cell sex-biased genes as those whose expression in the opposite  
214 sex, at any time point, is no greater than 20% of the gene's maximum expression in a given sex. This  
215 criteria yielded 67 egg-biased, 1,024 sperm-biased, and 197 unbiased germline-enriched genes. We found  
216 regardless of sex, egg, sperm, and unbiased germline genes were dysregulated in all *Kdm5c* mutants at  
217 similar proportions (Figure 4I-J). Furthermore, germline genes dysregulated exclusively in either male or  
218 female mutants were also not biased towards their corresponding germ cell sex (Figure 4I). Altogether, these  
219 results demonstrate sex differences in germline gene dysregulation is not due to sex-specific activation of  
220 sperm or egg transcriptional programs.

## 221 **KDM5C binds to a subset of germline gene promoters during early embryogenesis**

222 KDM5C binds to the promoters of several germline genes in embryonic stem cells (ESCs) but not in  
223 neurons<sup>13,54</sup>. However, due to the lack of a comprehensive list of germline-enriched genes, it is unclear if  
224 KDM5C is enriched at germline gene promoters, what types of germline genes KDM5C regulates, and if its  
225 binding is maintained at any germline genes in neurons.

226 To address these questions, we analyzed KDM5C chromatin immunoprecipitation followed by DNA  
227 sequencing (ChIP-seq) datasets in EpiLCs<sup>41</sup> and primary forebrain neuron cultures (PNCs)<sup>12</sup> (MACS2 q <  
228 0.1, fold enrichment > 1, and removal of false-positive *Kdm5c*-KO peaks). EpiLCs had a higher total number  
229 of high-confidence KDM5C peaks than PNCs (EpiLCs: 5,808, PNCs: 1,276). KDM5C was primarily localized  
230 to gene promoters in both cell types (promoters = transcription start site (TSS) ± 500 bp, EpiLCs: 4,190,  
231 PNCs: 745), although PNCs showed increased localization to non-promoter regions (Figure 5A).

232 The majority of promoters bound by KDM5C in PNCs were also bound in EpiLCs (513 shared promoters),  
233 however a large portion of gene promoters were bound by KDM5C only in EpiLCs (3,677 EpiLC only  
234 promoters) (Figure 5B). Genes bound by KDM5C in both PNCs and EpiLCs were enriched for functions  
235 involving nucleic acid turnover, such as deoxyribonucleotide metabolic process (GO:0009262, p.adjust =  
236 8.28e-05) (Figure 5C, Supplementary Table 4). Germline ontologies were enriched only in EpiLC-specific,  
237 KDM5C-bound promoters, such as meiotic nuclear division (GO: 0007127 p.adjust = 6.77e-16) (Figure 5C).  
238 There were no significant ontologies for PNC-specific KDM5C target genes. Using our mouse germline gene  
239 list, we observed evident KDM5C signal around the TSS of many germline genes in EpiLCs, but not in PNCs  
240 (Figure 5D). Based on our ChIP-seq peak cut-off criteria, KDM5C was highly enriched at 211 germline gene  
241 promoters in EpiLCs (16.4% of all germline genes) (Figure 5E, Supplementary Table 2). Of note, KDM5C  
242 was only bound to about one third of RNA-seq DEG promoters unique to EpiLCs or the brain (EpiLC only  
243 DEGs: 34.9%, Brain only DEGs: 30%) (Supplementary Figure 1A-C). Representative examples of EpiLC  
244 DEGs bound and unbound by KDM5C in EpiLCs are *Dazl* and *Stra8*, respectively (Figure 5F). However,  
245 the four of the five germline genes dysregulated in both EpiLCs and the brain were bound by KDM5C in  
246 EpiLCs (*D1Pas1*, *Hsf2bp*, *Cyct*, and *Stk31*) (Supplementary Figure 1A). Together, these results demonstrate  
247 KDM5C is recruited to a subset of germline genes in EpiLCs, including meiotic genes, but does not directly

248 regulate germline genes in neurons. Furthermore, the majority of germline mRNAs expressed in *Kdm5c*-KO  
249 cells are dysregulated independent of direct KDM5C recruitment to their gene promoters, however genes  
250 dysregulated across *Kdm5c*-KO development are often direct KDM5C targets.

251 Many germline-specific genes are suppressed by the polycomb repressive complex 1.6 (PRC1.6), which  
252 contains the transcription factor heterodimers E2F/DP1 and MGA/MAX that respectively bind E2F and  
253 E-box motifs within germline gene promoters<sup>17,18,20,40,46,47,55–57</sup>. PRC1.6 members may recruit KDM5C to  
254 germline gene promoters<sup>13</sup>, given their association with KDM5C in HeLa cells and ESCs<sup>45,58</sup>. We thus  
255 used HOMER<sup>59</sup> to identify transcription factor motifs enriched at KDM5C-bound or unbound germline gene  
256 promoters (TSS ± 500 bp, q-value < 0.1, Supplementary Table 4). MAX and E2F6 binding sites were  
257 significantly enriched at germline genes bound by KDM5C in EpiLCs (MAX q-value: 0.0068, E2F6 q-value:  
258 0.0673, E2F q-value: 0.0917), but not at germline genes unbound by KDM5C (Figure 5G). One third of  
259 KDM5C-bound promoters contained the consensus sequence for either E2F6 (E2F, 5'-TCCCGC-3'), MGA  
260 (E-box, 5'-CACGTG-3'), or both, but only 17% of KDM5C-unbound genes contained these motifs (Figure 5H).  
261 KDM5C-unbound germline genes were instead enriched for multiple RFX transcription factor binding sites  
262 (RFX q-value < 0.0001, RFX2 q-value < 0.0001, RFX5 q-value < 0.0001) (Figure 5I, Supplementary figure  
263 1D). RFX transcription factors bind X-box motifs<sup>60</sup> to promote ciliogenesis<sup>61,62</sup> and among them is RFX2, a  
264 central regulator of post-meiotic spermatogenesis<sup>63,64</sup>. Although *Rfx2* is also not a direct target of KDM5C  
265 (Supplementary Figure 1E), RFX2 mRNA is derepressed in *Kdm5c*-KO EpiLCs (Figure 5J). Thus, RFX2 is a  
266 candidate transcription factor for driving the ectopic expression of many KDM5C-unbound germline genes in  
267 *Kdm5c*-KO cells.

268 **KDM5C is recruited to CpG islands at germline promoters to facilitate *de novo* DNA  
269 methylation**

270 Previous work found two germline gene promoters have a marked reduction in DNA CpG methylation  
271 (CpGme) in the adult *Kdm5c*-KO hippocampus<sup>13</sup>. Since histone H3K4me2/3 impede *de novo* CpGme<sup>65,66</sup>,  
272 KDM5C's removal of H3K4me2/3 may be required to suppress germline genes. However, KDM5C's catalytic  
273 activity was recently shown to be dispensable for suppressing *Dazl* in undifferentiated ESCs<sup>45</sup>. To reconcile  
274 these observations, we hypothesized KDM5C erases H3K4me2/3 to promote the initial placement of CpGme  
275 at germline gene promoters in EpiLCs.

276 To test this hypothesis, we first characterized KDM5C's expression as naïve ESCs differentiate into  
277 EpiLCs (Figure 6A). While *Kdm5c* mRNA steadily decreased from 0 to 48 hours of differentiation (Figure  
278 6B), KDM5C protein initially increased from 0 to 24 hours and then decreased to near knockout levels by 48  
279 hours (Figure 6C). We then characterized KDM5C's substrates (H3K4me2/3) at germline gene promoters  
280 with *Kdm5c* loss using published ChIP-seq datasets<sup>22,41</sup>. *Kdm5c*-KO samples showed a marked increase in  
281 H3K4me2 in EpiLCs (Figure 6D) and H3K4me3 in the amygdala (Figure 6E) around the TSS of germline

282 genes. Together, these data suggest KDM5C acts during the transition between ESCs and EpiLCs to remove  
283 H3K4me2/3 at germline gene promoters.

284 Germline genes accumulate CpG methylation (CpGme) at CpG islands (CGIs) during the transition  
285 from naïve to primed pluripotency<sup>19,21,67</sup>. We first examined how many of our germline-enriched genes had  
286 promoter CGIs (TSS ± 500 bp) using the UCSC genome browser<sup>68</sup>. Notably, out of 1,288 germline-enriched  
287 genes, only 356 (27.64%) had promoter CGIs (Figure 6F, Supplementary Table 2). CGI-containing germline  
288 genes had higher enrichment of meiotic gene ontologies compared to CGI-free genes, including meiotic  
289 nuclear division (GO:0140013, p.adjust = 2.17e-12) and meiosis I (GO:0007127, p.adjust = 3.91e-10)  
290 (Figure 6G, Supplementary Table 5). Germline genes with promoter CGIs were more highly expressed than  
291 CGI-free genes across spermatogenesis stages, with highest expression in meiotic spermatocytes (Figure  
292 6H). Contrastingly, CGI-free genes only displayed substantial expression in post-meiotic round spermatids  
293 (Figure 6H). Although only a minor portion of germline gene promoters contained CGIs, CGIs strongly  
294 determined KDM5C's recruitment to germline genes ( $p = 2.37e-67$ , Odds Ratio = 17.8, Fisher's Exact Test),  
295 with 79.15% of KDM5C-bound germline gene promoters harboring CGIs (Figure 6F).

296 To assess how KDM5C loss impacts initial CpGme placement at germline gene promoters, we performed  
297 whole genome bisulfite sequencing (WGBS) in male wild-type and *Kdm5c*-KO ESCs and 96-hour extend  
298 EpiLCs (exEpiLCs), when germline genes reach peak methylation levels<sup>18</sup> (Figure 6I). We first identified  
299 which germline gene promoters significantly gained CpGme in wild-type cells during nESC to exEpiLCs  
300 differentiation (methylKit<sup>69</sup>,  $q < 0.01$ ,  $|methylation\ difference| > 25\%$ , TSS ± 500 bp). In wild-type cells, the  
301 majority of germline genes gained substantial CpGme at their promoter during differentiation (60.08%),  
302 regardless if their promoter contained a CGI (Figure 6J, Supplementary Table 5).

303 We then identified promoters differentially methylated in wild-type versus *Kdm5c*-KO exEpiLCs (methylKit,  
304  $q < 0.01$ ,  $|methylation\ difference| > 25\%$ , TSS ± 500 bp, Supplementary Table 5). Of the 48,882 promoters  
305 assessed, 274 promoters were significantly hypomethylated and 377 promoters were significantly hyper-  
306 methylated with KDM5C loss (Supplementary Figure 2A). Many promoters hyper- and hypomethylated  
307 in *Kdm5c*-KO exEpiLCs belonged to genes with unknown functions. However, 10.22% of hypomethyl-  
308 ated promoters belonged to germline genes and germline-relevant ontologies like meiotic nuclear division  
309 (GO:0140013, p.adjust = 0.012) are significantly enriched (Supplementary Figure 2B, Supplementary Table  
310 5). Approximately half of all germline gene promoters hypomethylated in *Kdm5c*-KO exEpiLCs are direct  
311 targets of KDM5C in EpiLCs (13 out of 28 hypomethylated promoters).

312 Promoters that showed the most robust loss of CpGme in *Kdm5c*-KO exEpiLCs (lowest q-values) harbored  
313 CGIs (Figure 6K). CGI promoters, but not CGI-free promoters, had a significant reduction in CpGme with  
314 KDM5C loss as a whole (Figure 6L) (Non-CGI promoters  $p = 0.0846$ , CGI promoters  $p = 0.0081$ , Mann-  
315 Whitney U test). Significantly hypomethylated promoters included germline genes consistently dysregulated  
316 across multiple *Kdm5c*-KO RNA-seq datasets<sup>13</sup>, such as *D1Pas1* (methylation difference = -60.03%, q-value  
317 = 3.26e-153) and *Naa11* (methylation difference = -42.45%, q-value = 1.44e-38) (Figure 6M). Unexpectedly,

318 we observed only a modest reduction in CpGme at *Dazl*'s promoter (methylation difference = -6.525%,  
319 q-value = 0.0159) (Figure 6N). Altogether, these results demonstrate KDM5C is recruited to germline gene  
320 CGIs in EpiLCs to promote CpGme at those promoters. Furthermore, our data suggest while KDM5C's  
321 catalytic activity is required for the repression of some germline genes, CpGme can be placed at others even  
322 with elevated H3K4me2/3 around the TSS.

## 323 Discussion

324 In the above study, we demonstrate KDM5C's pivotal role in the development of tissue identity. We first  
325 characterized tissue-enriched genes expressed within the mouse *Kdm5c*-KO brain and identified substantial  
326 derepression of testis, liver, muscle, and ovary-enriched genes. Testis genes significantly enriched within the  
327 *Kdm5c*-KO amygdala and hippocampus are specific to the germline and absent in somatic cells. *Kdm5c*-  
328 KO epiblast-like cells (EpiLCs) aberrantly express key drivers of germline identity and meiosis, including  
329 *Dazl* and *Stra8*, while the adult brain primarily expresses genes important for late spermatogenesis. We  
330 demonstrated that although sex did not influence whether sperm or egg-specific genes were misexpressed,  
331 female EpiLCs have heightened germline gene de-repression with KDM5C loss. Germline genes can become  
332 aberrantly expressed in *Kdm5c*-KO cells via indirect mechanisms, such as activation through ectopic RFX  
333 transcription factors. Finally, we found KDM5C is dynamically regulated during ESC to EpiLC differentiation  
334 to promote long-term germline gene silencing through CGI DNA methylation. Therefore, we propose KDM5C  
335 plays a fundamental role in the development of tissue identity during early embryogenesis, including the  
336 establishment of the soma-germline boundary. By systematically characterizing KDM5C's role in germline  
337 gene repression, we unveiled distinct mechanisms governing the misexpression of distinct germline gene  
338 classes in somatic lineages. Ultimately, these data provide molecular footholds which can be exploited to  
339 test the overarching contribution of ectopic germline gene expression upon neurodevelopment.

340 By comparing *Kdm5c* mutant males and females, we revealed germline gene suppression is sexually  
341 dimorphic. Female EpiLCs are more severely impacted by loss of KDM5C-mediated germline gene sup-  
342 pression, yet this difference is not due to the large number of germline genes on the X chromosome<sup>52,53</sup>.  
343 Heightened germline gene misexpression in females may be related to females having a higher dose of  
344 KDM5C than males, due to its escape from XCI<sup>48–51</sup>. Intriguingly, heterozygous knockout females (*Kdm5c*<sup>-/+</sup>)  
345 also had over double the number of germline DEGs than hemizygous knockout males (*Kdm5c*<sup>-/Y</sup>), even  
346 though their expression of KDM5C should be roughly equivalent to that of wild-type males (*Kdm5c*<sup>+/Y</sup>). Males  
347 could partially compensate for KDM5C's loss via the Y-chromosome homolog, KDM5D<sup>8</sup>. However, KDM5D  
348 has not been reported to regulate germline gene expression. Nevertheless, these results demonstrate  
349 germline gene silencing mechanisms differ between males and females, which warrants further study to  
350 elucidate the biological ramifications and underlying mechanisms.

351 We found KDM5C is largely dispensable for promoting normal gene expression during development, yet

352 is critical for suppressing ectopic developmental programs. While some germline genes, such as *Dazl*, are  
353 also expressed in the 2-cell stage, the inner cell mass, and naïve ESCs, they are silenced in epiblast stem  
354 cells/EpiLCs<sup>18,40,45,70,71</sup>. Our data suggest the 2-cell-like state reported in *Kdm5c*-KO ESCs<sup>45</sup> likely reflects  
355 KDM5C's primary role in germline gene repression (Figure 3). Germline gene misexpression in *Kdm5c*-  
356 KO EpiLCs may indicate they are differentiating into primordial germ cell-like cells (PGCLCs)<sup>33,34,36</sup>. Yet,  
357 *Kdm5c*-KO EpiLCs had normal cellular morphology and properly expressed markers for primed pluripotency,  
358 including *Otx2* which blocks EpiLC differentiation into PGCs/PGCLCs<sup>72</sup>. In addition to unimpaired EpiLC  
359 differentiation, *Kdm5c*-KO gross brain morphology is overall normal<sup>12</sup> and hardly any brain-specific genes  
360 were significantly dysregulated in the amygdala and hippocampus (Figure 1). Thus, ectopic germline gene  
361 expression occurs in conjunction with overall proper somatic differentiation in *Kdm5c*-KO animals.

362 Our work provides novel insight into the cross-talk between H3K4me2/3 and CpGme, which are gen-  
363 erally mutually exclusive<sup>73</sup>. In EpiLCs, loss of KDM5C binding at a subset of germline gene promoters,  
364 e.g. *D1Pas1*, strongly impaired promoter CGI methylation and resulted in their long-lasting de-repression  
365 into adulthood. Removal of H3K4me2/3 at CGIs is a plausible mechanism for KDM5C-mediated germline  
366 gene suppression<sup>13,54</sup>, given H3K4me2/3 repell DNMT3 activity<sup>65,66</sup>. However, emerging work indicates  
367 many histone-modifying enzymes have non-catalytic functions that influence gene expression, sometimes  
368 even more potently than their catalytic roles<sup>74,75</sup>. Indeed, KDM5C's catalytic activity was recently found to be  
369 dispensable for repressing *Dazl* in ESCs<sup>45</sup>. In our study, *Dazl*'s promoter still gained CpGme in *Kdm5c*-KO  
370 exEpiLCs, even with elevated H3K4me2. *Dazl* and a few other germline genes employ multiple repressive  
371 mechanisms to facilitate CpGme, such as DNMT3A/B recruitment via E2F6 and MGA<sup>17,18,46,47</sup>. Thus, while  
372 some germline CGIs require KDM5C-mediated H3K4me removal to overcome promoter CGI escape from  
373 CpGme<sup>73,76</sup>, others do not. These results also suggest the requirement for KDM5C's catalytic activity can  
374 change depending upon the locus and developmental stage. Further experiments are required to determine  
375 if catalytically inactive KDM5C can suppress germline genes at later developmental stages.

376 By generating a comprehensive list of mouse germline-enriched genes, we revealed distinct derepressive  
377 mechanisms governing early versus late-stage germline programs. Previous work on germline gene silencing  
378 has focused on genes with promoter CGIs<sup>19,73</sup>, and indeed the majority of KDM5C targets in EpiLCs were  
379 germ cell identity genes harboring CGIs. However, over 70% of germline-enriched gene promoters lacked  
380 CGIs, including the many KDM5C-unbound germline genes that are de-repressed in *Kdm5c*-KO cells. CGI-  
381 free, KDM5C-unbound germline genes were primarily late-stage spermatogenesis genes and significantly  
382 enriched for RFX2 binding sites, a central regulator of spermiogenesis<sup>63,64</sup>. These data suggest that once  
383 activated during early embryogenesis, drivers of germline gene expression like *Rfx2*, *Stra8*, and *Dazl* turn  
384 on downstream germline programs, ultimately culminating in the expression of spermiogenesis genes in  
385 the adult *Kdm5c*-KO brain. Therefore, we propose KDM5C is recruited via promoter CGIs to act as a brake  
386 against runaway activation of germline-specific programs. Future studies should address how KDM5C is  
387 targeted to CGIs.

388 The above work provides the mechanistic foundation for KDM5C-mediated repression of tissue and  
389 germline-specific genes. However, the contribution of these ectopic, tissue-specific genes towards neurolog-  
390 ical impairments is still unknown. In addition to germline genes, we also identified significant enrichment  
391 of muscle and liver-enriched transcripts within the *Kdm5c*-KO brain. Intriguingly, select liver and muscle-  
392 enriched DEGs do have known roles within the brain, such as the liver-enriched lipid metabolism gene  
393 *Apolipoprotein C-I (Apoc1)*<sup>28</sup>. *APOC1* dysregulation is implicated in Alzheimer's disease in humans<sup>77</sup> and  
394 overexpression of *Apoc1* in the mouse brain can impair learning and memory<sup>78</sup>. KDM5C may therefore be  
395 crucial for neurodevelopment by fine-tuning the expression of tissue-enriched, dosage-sensitive genes like  
396 *Apoc1*.

397 Given that germline genes have no known functions within the brain, their impact upon neurodevelopment  
398 is currently unknown. In *C. elegans*, somatic misexpression of germline genes via loss of *Retinoblastoma*  
399 (*Rb*) homologs results in enhanced piRNA signaling and ectopic P granule formation in neurons<sup>79,80</sup>. Ectopic  
400 testicular germline transcripts have also been observed in a variety of cancers, including brain tumors in  
401 *Drosophila* and mammals<sup>81,82</sup> and shown to promote cancer progression<sup>83–85</sup>. Intriguingly, mouse models  
402 and human cells for other chromatin-linked NDDs also display impaired soma-germline demarcation<sup>86–88</sup>,  
403 such as mutations in DNA methyltransferase 3b (DNMT3B), H3K9me1/2 methyltransferases G9A/GLP,  
404 and methyl-CpG -binding protein 2 (MECP2). Recently, the transcription factor ZMYM2 (ZNF198), whose  
405 mutation causes a NDD (OMIM #619522), was also shown to repress germline genes by promoting H3K4me  
406 removal and CpGme<sup>89</sup>. Thus, KDM5C is among a growing cohort of neurodevelopmental disorders with  
407 erosion of the germline-soma boundary. Further research is required to determine the impact of these  
408 germline genes upon neuronal functions and the extent to which this phenomenon occurs in humans.

## 409 Materials and Methods

### 410 Classifying tissue-enriched and germline-enriched genes

411 Tissue-enriched differentially expresssd genes (DEGs) were determined by their classification in a previ-  
412 ously published dataset from 17 male and female mouse tissues<sup>23</sup>. This study defined tissue expression as  
413 greater than 1 Fragments Per Kilobase of transcript per Million mapped read (FPKM) and tissue enrichment  
414 as at least 4-fold higher expression than any other tissue.

415 We curated a list of germline-enriched genes using an RNA-seq dataset from wild-type and germline-  
416 depleted (Kit<sup>W/W<sup>v</sup></sup>) male and female mouse embryos from embryonic day 12, 14, and 16<sup>32</sup>, as well as adult  
417 male testes<sup>29</sup>. Germline-enriched genes met the following criteria: 1) their expression is greater than 1  
418 FPKM in wild-type germline 2) their expression in any wild-type somatic tissues<sup>23</sup> does not exceed 20%  
419 of maximum expression in wild-type germline, and 3) their expression in the germ cell-depleted (Kit<sup>W/W<sup>v</sup></sup>)  
420 germline, for any sex or time point, does not exceed 20% of maximum expression in wild-type germline. We

421 defined sperm and egg-biased genes as those whose expression in the opposite sex, at any time point, is no  
422 greater than 20% of the gene's maximum expression in a given sex. Genes that did not meet this threshold  
423 for either sex were classified as 'unbiased'.

#### 424 **Cell culture**

425 We utilized our previously established cultures of male wild-type and *Kdm5c* knockout (-KO)  
426 embryonic stem cells<sup>41</sup>. Sex was confirmed by genotyping *Uba1/Uba1y* on the X and Y chromo-  
427 somes with the following primers: 5'-TGGATGGTGTGCCAATG-3', 5'-CACCTGCACGTTGCCCTT-  
428 3'. Deletion of *Kdm5c* exons 11 and 12, which destabilize KDM5C protein<sup>12</sup>, was confirmed  
429 through the primers 5'-ATGCCCATATTAAGAGTCCCTG-3', 5'-TCTGCCTTGATGGACTGTT-3', and  
430 5'-GGTTCTAACACTCACATAGTG-3'.

431 Embryonic stem cells (ESCs) and epiblast-like cells were cultured using previously established  
432 methods<sup>37</sup>. Briefly, ESCs were initially cultured in primed ESC (pESC) media consisting of KnockOut  
433 DMEM (Gibco#10829–018), fetal bovine serum (Gibco#A5209501), KnockOut serum replacement  
434 (Invitrogen#10828–028), Glutamax (Gibco#35050-061), Anti-Anti (Gibco#15240-062), MEM Non-essential  
435 amino acids (Gibco#11140-050), and beta-mercaptoethanol (Sigma#M7522). They were then transitioned  
436 into ground-state, "naïve" ESCs (nESCs) by culturing for four passages in N2B27 media containing  
437 DMEM/F12 (Gibco#11330–032), Neurobasal media (Gibco#21103–049), Gluamax (Gibco#35050-061),  
438 Anti-Anti (Gibco#15240-062), N2 supplement (Invitrogen#17502048), and B27 supplement without vitamin  
439 A (Invitrogen#12587-010), and beta-mercaptoethanol (Sigma#M7522). Both pESC and nESC media  
440 were supplemented with 3 µM GSK3 inhibitor CHIR99021 (Sigma #SML1046-5MG), 1 µM MEK inhibitor  
441 PD0325901 (Sigma #PZ0162-5MG), and 1,000 units/mL leukemia inhibitory factor (LIF, Millipore#ESG1107).

442 nESCs were differentiated into epiblast-like cells (EpiLCs, 48 hours) and extendend EpiLCs (exEpiLCs,  
443 96 hours) by culturing in N2B27 media containing DMEM/F12, Neurobasal media, Gluamax, Anti-Anti, N2  
444 supplement, B27 supplement (Invitrogen#17504044), and beta-mercaptoethanol supplemented with 10  
445 ng/mL fibroblast growth factor 2 (FGF2, R&D Biotechne 233-FB), and 20 ng/mL activin A (R&D Biotechne  
446 338AC050CF), as previously described<sup>37</sup>.

#### 447 **Real time quantitative PCR (RT-qPCR)**

448 nESCs were differentiated into EpiLCs as described above. Cells were lysed with Tri-reagent BD (Sigma  
449 #T3809) at 0, 24, and 48 hours of differentiation. RNA was phase separated with 0.1 µL/µL 1-bromo-3-  
450 chloropropane (Sigma #B9673) and then precipitated with with isopropanol (Sigma #I9516) and ethanol puri-  
451 fied. For each sample, 2 µg of RNA was reverse transcribed using the ProtoScript II Reverse transcriptase kit  
452 from New England Biolabs (NEB #M0368S) and primed with oligo dT. Expression of *Kdm5c* was detected us-  
453 ing the primers 5'-CCCATGGAGGCCAGAGAATAAG-3' 5'-CTCAGCGGATAAGAGAATTGCTAC-3' and nor-

454 malized to TBP using the primers 5'-TTCAGAGGATGCTCTAGGAAAGA-3' 5'-CTGTGGAGTAAGTCCTGTGCC-  
455 3' with the Power SYBR™ Green PCR Master Mix (ThermoFisher #4367659).

## 456 **Western Blot**

457 Total protein was extracted during nESC to EpiLC differentiation at 0, 24, and 48 hours by sonicating cells  
458 at 20% amplitude for 15 seconds in 2X SDS sample buffer, then boiling at 100 °C for 10 minutes. Proteins  
459 were separated on a 7.5% SDS page gel, transferred overnight onto a fluorescent membrane, blotted for  
460 rabbit anti-KDM5C (in house, 1:500) and mouse anti-DAXX (Santa Cruz #(H-7): sc-8043, 1:500), and then  
461 imaged using the LiCor Odyssey CLx system. Band intensity was quantified using ImageJ.

## 462 **RNA sequencing (RNA-seq) data analysis**

463 After ensuring read quality via FastQC (v0.11.8), reads were then mapped to the mm10 *Mus musculus*  
464 genome (Gencode) using STAR (v2.5.3a), during which we removed duplicates and kept only uniquely  
465 mapped reads. Count files were generated by FeatureCounts (Subread v1.5.0), and BAM files were  
466 converted to bigwigs using deeptools (v3.1.3) and visualized by the UCSC genome browser<sup>68</sup>. RStudio  
467 (v3.6.0) was then used to analyze counts files by DESeq2 (v1.26.0)<sup>24</sup> to identify differentially expressed  
468 genes (DEGs) with a q-value (p-adjusted via FDR/Benjamini–Hochberg correction) less than 0.1 and a log2  
469 fold change greater than 0.5. For all DESeq2 analyses, log2 fold changes were calculated with IfcShrink  
470 using the ashR package<sup>90</sup>. MA-plots were generated by ggpubr (v0.6.0), and Eulerr diagrams were generated  
471 by eulerr (v6.1.1). Boxplots and scatterplots were generated by ggpubr (v0.6.0) and ggplot2 (v3.3.2). The  
472 Upset plot was generated via the package UpSetR (v1.4.0)<sup>91</sup>. Gene ontology (GO) analyses were performed  
473 by the R package enrichPlot (v1.16.2) using the biological processes setting and compareCluster.

## 474 **Chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) data analysis**

475 ChIP-seq reads were aligned to mm10 using Bowtie1 (v1.1.2) allowing up to two mismatches. Only  
476 uniquely mapped reads were used for analysis. Peaks were called using MACS2 software (v2.2.9.1) using  
477 input BAM files for normalization, with filters for a q-value < 0.1 and a fold enrichment > 1. We removed  
478 “black-listed” genomic regions that often give aberrant signals. Common peak sets were obtained in R via  
479 DiffBind<sup>92</sup> (v3.6.5). In the case of KDM5C ChIP-seq, *Kdm5c*-KO false-positive peaks were then removed from  
480 wild-type samples using bedtools (v2.25.0). Peak proximity to genomic loci was determined by ChIPSeeker<sup>93</sup>  
481 (v1.32.1). Gene ontology (GO) analyses were performed by the R package enrichPlot (v1.16.2) using the  
482 biological processes setting and compareCluster. Enriched motifs were identified using HOMER<sup>59</sup> to search  
483 for known motifs within 500 base pairs upstream and downstream of the transcription start site. Average binding  
484 across genes was visualized using deeptools (v3.1.3). Bigwigs were visualized using the UCSC genome  
485 browser<sup>68</sup>.

486 **CpG island (CGI) analysis**

487 Locations of CpG islands were determined through the mm10 UCSC genome browser CpG island track<sup>68</sup>,  
488 which classified CGIs as regions that have greater than 50% GC content, are larger than 200 base pairs,  
489 and have a ratio of CG dinucleotides observed over the expected amount greater than 0.6. CGI genomic  
490 coordinates were then annotated using ChIPseeker<sup>93</sup> (v1.32.1) and filtered for ones that lie within promoters  
491 of germline-enriched genes (TSS ± 500).

492 **Whole genome bisulfite sequencing (WGBS)**

493 Genomic DNA (gDNA) from male naïve ESCs and extended EpiLCs was extracted using the Wizard  
494 Genomic DNA Purification Kit (Promega A1120), following the instructions for Tissue Culture Cells. gDNA  
495 from two wild-types and two *Kdm5c*-KOs of each cell type was sent to Novogene for WGBS using the  
496 Illumina NovaSeq X Plus platform and sequenced for 150 bp paired-end reads (PE150). All samples had  
497 greater than 99% bisulfite conversion rates. Reads were adapter and quality trimmed with Trim Galore  
498 (v0.6.10) and aligned to the mm10 genome using Bismark<sup>94</sup> (v0.22.1). Analysis of differential methylation at  
499 gene promoters was performed using methylKit<sup>69</sup> (v1.28.0) with a minimum coverage of 3 paired reads, a  
500 percentage greater than 25% or less than -25%, and q-value less than 0.01. methylKit was also used to  
501 calculate average percentage methylation at germline gene promoters. Methylation bedgraph tracks were  
502 generated via Bismark and visualized using the UCSC genome browser<sup>68</sup>.

503 **Data availability**

504 **WGBS in wild-type and *Kdm5c*-KO ESCs and exEpiLCs**

505 Raw fastq files are deposited in the Sequence Read Archive (SRA) <https://www.ncbi.nlm.nih.gov/sra>  
506 under the bioProject PRJNA1165148. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1165148>

507 **Published datasets**

508 All published datasets are available at the Gene Expression Omnibus (GEO) <https://www.ncbi.nlm.nih.gov/geo/>. Published RNA-seq datasets analyzed in this study included the male wild-type and *Kdm5c*-KO  
509 adult amygdala and hippocampus<sup>22</sup>, available at GEO: GSE127722. Male and female wild-type, *Kdm5c*-KO,  
510 and *Kdm5c*-HET EpiLCs<sup>41</sup> are available at GEO: GSE96797.

512 Previously published ChIP-seq experiments included KDM5C binding in wild-type and *Kdm5c*-KO  
513 EpiLCs<sup>41</sup> (available at GEO: GSE96797) and mouse primary neuron cultures (PNCs) from the cortex  
514 and hippocampus<sup>12</sup> (available at GEO: GSE61036). ChIP-seq of histone 3 lysine 4 dimethylation (H3K4me2)  
515 in male wild-type and *Kdm5c*-KO EpiLCs<sup>41</sup> is also available at GEO: GSE96797. ChIP-seq of histone 3 lysine  
516 4 trimethylation (H3K4me3) in wild-type and *Kdm5c*-KO male amygdala<sup>22</sup> are available at GEO: GSE127817.

517 **Data analysis**

518 Scripts used to generate the results, tables, and figures of this study are available via the GitHub  
519 repository: [https://github.com/kbonefas/KDM5C\\_Germ\\_Mechanism](https://github.com/kbonefas/KDM5C_Germ_Mechanism)

520 **Acknowledgements**

521 We thank Drs. Sundeep Kalantry, Milan Samanta, and Rebecca Malcore for providing protocols and  
522 expertise in culturing mouse ESCs and EpiLCs, as well as providing the wild-type and *Kdm5c*-KO ESCs  
523 used in this study. We thank Dr. Jacob Mueller for his insight in germline gene regulation and directing  
524 us to the germline-depleted mouse models. We also thank Drs. Gabriel Corfas, Kenneth Kwan, Natalie  
525 Tronson, Michael Sutton, Stephanie Bielas, Donna Martin, and the members of the Iwase, Sutton, Bielas,  
526 and Martin labs for helpful discussions and critiques of the data. We thank members of the University  
527 of Michigan Reproductive Sciences Program for providing feedback throughout the development of this  
528 work. This work was supported by grants from the National Institutes of Health (NIH) National Institute of  
529 Neurological Disorders and Stroke (NS089896, 5R21NS104774, and NS116008 to S.I.), National Institute  
530 of Mental Health (1R21MH135290 to S.I.), the Simons Foundation Autism Research Initiative (SFARI, SFI-  
531 AN-AR-Pilot-00005721 to S.I.), the Farrehi Family Foundation Grant (to S.I.), the University of Michigan  
532 Career Training in Reproductive Biology (NIH T32HD079342, to K.M.B.), the NIH Early Stage Training in  
533 the Neurosciences Training Grant (NIH T32NS076401 to K.M.B.), and the Michigan Predoctoral Training in  
534 Genetics Grant (NIH T32GM007544, to I.V.)

535 **Author Contributions**

536 K.M.B. and S.I. conceived the study and designed the experiments. I.V. generated the ESC and exEpiLC  
537 WGBS data. K.M.B performed all data analysis and all other experiments. The manuscript was written by  
538 K.M.B and S.I., and edited by K.M.B, S.I., and I.V.

539 **Declaration of Interest**

540 S.I. is a member of the Scientific Advisory Board of KDM5C Advocacy, Research, Education & Support  
541 (KARES). Other authors declare no conflict of interest.

542 **References**

- 543 1. Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* **403**,  
544 41–45. <https://doi.org/10.1038/47412>.
- 545 2. Jenuwein, T., and Allis, C.D. (2001). Translating the Histone Code. *Science* **293**, 1074–1080.  
546 <https://doi.org/10.1126/science.1063127>.
- 547 3. Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570.  
548 <https://doi.org/10.1038/276565a0>.
- 549 4. Kennison, J.A., and Tamkun, J.W. (1988). Dosage-dependent modifiers of polycomb and antennapedia  
550 mutations in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8136–8140. <https://doi.org/10.1073/pnas.8>  
550 5.21.8136.
- 551 5. Kassis, J.A., Kennison, J.A., and Tamkun, J.W. (2017). Polycomb and Trithorax Group Genes in  
552 *Drosophila*. *Genetics* **206**, 1699–1725. <https://doi.org/10.1534/genetics.115.185116>.
- 553 6. Gabriele, M., Lopez Tobon, A., D'Agostino, G., and Testa, G. (2018). The chromatin basis of  
554 neurodevelopmental disorders: Rethinking dysfunction along the molecular and temporal axes. *Prog  
Neuropsychopharmacol Biol Psychiatry* **84**, 306–327. <https://doi.org/10.1016/j.pnpbp.2017.12.013>.
- 555 7. Schaefer, A., Sampath, S.C., Intrator, A., Min, A., Gertler, T.S., Surmeier, D.J., Tarakhovsky, A.,  
556 and Greengard, P. (2009). Control of cognition and adaptive behavior by the GLP/G9a epigenetic  
556 suppressor complex. *Neuron* **64**, 678–691. <https://doi.org/10.1016/j.neuron.2009.11.019>.
- 557 8. Iwase, S., Lan, F., Bayliss, P., De La Torre-Ubieta, L., Huarte, M., Qi, H.H., Whetstine, J.R., Bonni, A.,  
558 Roberts, T.M., and Shi, Y. (2007). The X-Linked Mental Retardation Gene SMCX/JARID1C Defines a  
558 Family of Histone H3 Lysine 4 Demethylases. *Cell* **128**, 1077–1088. <https://doi.org/10.1016/j.cell.200>  
558 7.02.017.
- 559 9. Claes, S., Devriendt, K., Van Goethem, G., Roelen, L., Meireleire, J., Raeymaekers, P., Cassiman,  
560 J.J., and Fryns, J.P. (2000). Novel syndromic form of X-linked complicated spastic paraparesia. *Am J  
Med Genet* **94**, 1–4.
- 561 10. Jensen, L.R., Amende, M., Gurok, U., Moser, B., Gimmel, V., Tzschach, A., Janecke, A.R., Tariverdian,  
562 G., Chelly, J., Fryns, J.P., et al. (2005). Mutations in the JARID1C gene, which is involved in  
transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. *Am J Hum  
Genet* **76**, 227–236. <https://doi.org/10.1086/427563>.
- 563 11. Carmignac, V., Nambot, S., Lehalle, D., Callier, P., Moortgat, S., Benoit, V., Ghoumid, J., Delobel,  
564 B., Smol, T., Thuillier, C., et al. (2020). Further delineation of the female phenotype with KDM5C  
disease causing variants: 19 new individuals and review of the literature. *Clin Genet* **98**, 43–55.  
<https://doi.org/10.1111/cge.13755>.

- 565 12. Iwase, S., Brookes, E., Agarwal, S., Badeaux, A.I., Ito, H., Vallianatos, C.N., Tomassy, G.S., Kasza, T.,  
Lin, G., Thompson, A., et al. (2016). A Mouse Model of X-linked Intellectual Disability Associated with  
Impaired Removal of Histone Methylation. *Cell Reports* *14*, 1000–1009. <https://doi.org/10.1016/j.celrep.2015.12.091>.
- 566
- 567 13. Scandaglia, M., Lopez-Atalaya, J.P., Medrano-Fernandez, A., Lopez-Cascales, M.T., Del Blanco, B.,  
Lipinski, M., Benito, E., Olivares, R., Iwase, S., Shi, Y., et al. (2017). Loss of Kdm5c Causes Spurious  
Transcription and Prevents the Fine-Tuning of Activity-Regulated Enhancers in Neurons. *Cell Rep* *21*,  
47–59. <https://doi.org/10.1016/j.celrep.2017.09.014>.
- 568
- 569 14. Bonefas, K.M., Vallianatos, C.N., Raines, B., Tronson, N.C., and Iwase, S. (2023). Sexually Dimorphic  
Alterations in the Transcriptome and Behavior with Loss of Histone Demethylase KDM5C. *Cells* *12*,  
637. <https://doi.org/10.3390/cells12040637>.
- 570
- 571 15. Devlin, D.K., Ganley, A.R.D., and Takeuchi, N. (2023). A pan-metazoan view of germline-soma  
distinction challenges our understanding of how the metazoan germline evolves. *Current Opinion in  
Systems Biology* *36*, 100486. <https://doi.org/10.1016/j.coisb.2023.100486>.
- 572
- 573 16. Lehmann, R. (2012). Germline Stem Cells: Origin and Destiny. *Cell Stem Cell* *10*, 729–739.  
<https://doi.org/10.1016/j.stem.2012.05.016>.
- 574
- 575 17. Endoh, M., Endo, T.A., Shinga, J., Hayashi, K., Farcas, A., Ma, K.W., Ito, S., Sharif, J., Endoh, T.,  
Onaga, N., et al. (2017). PCGF6-PRC1 suppresses premature differentiation of mouse embryonic  
stem cells by regulating germ cell-related genes. *eLife* *6*. <https://doi.org/10.7554/eLife.21064>.
- 576
- 577 18. Mochizuki, K., Sharif, J., Shirane, K., Uranishi, K., Bogutz, A.B., Janssen, S.M., Suzuki, A., Okuda,  
A., Koseki, H., and Lorincz, M.C. (2021). Repression of germline genes by PRC1.6 and SETDB1  
in the early embryo precedes DNA methylation-mediated silencing. *Nat Commun* *12*, 7020. <https://doi.org/10.1038/s41467-021-27345-x>.
- 578
- 579 19. Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T., and Weber, M. (2010).  
Targets and dynamics of promoter DNA methylation during early mouse development. *Nat Genet* *42*,  
1093–1100. <https://doi.org/10.1038/ng.708>.
- 580
- 581 20. Velasco, G., Hubé, F., Rollin, J., Neuillet, D., Philippe, C., Bouzinba-Segard, H., Galvani, A., Viegas-  
Péquignot, E., and Francastel, C. (2010). Dnmt3b recruitment through E2F6 transcriptional repressor  
mediates germ-line gene silencing in murine somatic tissues. *Proc Natl Acad Sci U S A* *107*, 9281–  
9286. <https://doi.org/10.1073/pnas.1000473107>.
- 582
- 583 21. Hackett, J.A., Reddington, J.P., Nestor, C.E., Dunican, D.S., Branco, M.R., Reichmann, J., Reik,  
W., Surani, M.A., Adams, I.R., and Meehan, R.R. (2012). Promoter DNA methylation couples  
genome-defence mechanisms to epigenetic reprogramming in the mouse germline. *Development*  
139, 3623–3632. <https://doi.org/10.1242/dev.081661>.
- 584

- 585 22. Vallianatos, C.N., Raines, B., Porter, R.S., Bonefas, K.M., Wu, M.C., Garay, P.M., Collette, K.M.,  
Seo, Y.A., Dou, Y., Keegan, C.E., et al. (2020). Mutually suppressive roles of KMT2A and KDM5C  
in behaviour, neuronal structure, and histone H3K4 methylation. *Commun Biol* 3, 278. <https://doi.org/10.1038/s42003-020-1001-6>.
- 586
- 587 23. Li, B., Qing, T., Zhu, J., Wen, Z., Yu, Y., Fukumura, R., Zheng, Y., Gondo, Y., and Shi, L. (2017). A  
Comprehensive Mouse Transcriptomic BodyMap across 17 Tissues by RNA-seq. *Sci Rep* 7, 4200.  
<https://doi.org/10.1038/s41598-017-04520-z>.
- 588
- 589 24. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for  
590 RNA-seq data with DESeq2. *Genome Biol* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
- 591 25. Crackower, M.A., Kolas, N.K., Noguchi, J., Sarao, R., Kikuchi, K., Kaneko, H., Kobayashi, E., Kawai, Y.,  
Kozieradzki, I., Landers, R., et al. (2003). Essential Role of Fkbp6 in Male Fertility and Homologous  
592 Chromosome Pairing in Meiosis. *Science* 300, 1291–1295. <https://doi.org/10.1126/science.1083022>.
- 593 26. Xiol, J., Cora, E., Koglgruber, R., Chuma, S., Subramanian, S., Hosokawa, M., Reuter, M., Yang, Z.,  
Berninger, P., Palencia, A., et al. (2012). A Role for Fkbp6 and the Chaperone Machinery in piRNA  
594 Amplification and Transposon Silencing. *Molecular Cell* 47, 970–979. <https://doi.org/10.1016/j.molcel.2012.07.019>.
- 595 27. Cheng, S., Altmeppen, G., So, C., Welp, L.M., Penir, S., Ruhwedel, T., Menelaou, K., Harasimov, K.,  
Stützer, A., Blayney, M., et al. (2022). Mammalian oocytes store mRNAs in a mitochondria-associated  
596 membraneless compartment. *Science* 378, eabq4835. <https://doi.org/10.1126/science.abq4835>.
- 597 28. Rouland, A., Masson, D., Lagrost, L., Vergès, B., Gautier, T., and Bouillet, B. (2022). Role of  
598 apolipoprotein C1 in lipoprotein metabolism, atherosclerosis and diabetes: A systematic review.  
*Cardiovasc Diabetol* 21, 272. <https://doi.org/10.1186/s12933-022-01703-5>.
- 599 29. Mueller, J.L., Skaletsky, H., Brown, L.G., Zaghlul, S., Rock, S., Graves, T., Auger, K., Warren,  
600 W.C., Wilson, R.K., and Page, D.C. (2013). Independent specialization of the human and mouse X  
chromosomes for the male germ line. *Nat Genet* 45, 1083–1087. <https://doi.org/10.1038/ng.2705>.
- 601 30. Handel, M.A., and Eppig, J.J. (1979). Sertoli Cell Differentiation in the Testes of Mice Genetically  
Deficient in Germ Cells. *Biology of Reproduction* 20, 1031–1038. <https://doi.org/10.1095/biolreprod20.5.1031>.
- 602
- 603 31. Green, C.D., Ma, Q., Manske, G.L., Shami, A.N., Zheng, X., Marini, S., Moritz, L., Sultan, C.,  
Gurczynski, S.J., Moore, B.B., et al. (2018). A Comprehensive Roadmap of Murine Spermatogenesis  
Defined by Single-Cell RNA-Seq. *Dev Cell* 46, 651–667.e10. <https://doi.org/10.1016/j.devcel.2018.07.025>.
- 604
- 605 32. Soh, Y.Q., Junker, J.P., Gill, M.E., Mueller, J.L., van Oudenaarden, A., and Page, D.C. (2015). A  
Gene Regulatory Program for Meiotic Prophase in the Fetal Ovary. *PLoS Genet* 11, e1005531.  
<https://doi.org/10.1371/journal.pgen.1005531>.

- 606
- 607 33. Magnúsdóttir, E., and Surani, M.A. (2014). How to make a primordial germ cell. *Development* *141*,  
608 245–252. <https://doi.org/10.1242/dev.098269>.
- 609 34. Günesdogan, U., Magnúsdóttir, E., and Surani, M.A. (2014). Primordial germ cell specification: A  
610 context-dependent cellular differentiation event [corrected]. *Philos Trans R Soc Lond B Biol Sci* *369*.  
<https://doi.org/10.1098/rstb.2013.0543>.
- 611 35. Bardot, E.S., and Hadjantonakis, A.-K. (2020). Mouse gastrulation: Coordination of tissue patterning,  
612 specification and diversification of cell fate. *Mechanisms of Development* *163*, 103617. <https://doi.org/10.1016/j.mod.2020.103617>.
- 613 36. Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the  
614 mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* *146*, 519–532.  
<https://doi.org/10.1016/j.cell.2011.06.052>.
- 615 37. Samanta, M., and Kalantry, S. (2020). Generating primed pluripotent epiblast stem cells: A methodol-  
616 ogy chapter. *Curr Top Dev Biol* *138*, 139–174. <https://doi.org/10.1016/bs.ctdb.2020.01.005>.
- 617 38. Welling, M., Chen, H., Muñoz, J., Musheev, M.U., Kester, L., Junker, J.P., Mischerikow, N., Arbab, M.,  
618 Kuijk, E., Silberstein, L., et al. (2015). DAZL regulates Tet1 translation in murine embryonic stem cells.  
*EMBO Reports* *16*, 791–802. <https://doi.org/10.15252/embr.201540538>.
- 619 39. Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O.,  
620 Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus  
activity. *Nature* *487*, 57–63. <https://doi.org/10.1038/nature11244>.
- 621 40. Suzuki, A., Hirasaki, M., Hishida, T., Wu, J., Okamura, D., Ueda, A., Nishimoto, M., Nakachi, Y.,  
622 Mizuno, Y., Okazaki, Y., et al. (2016). Loss of MAX results in meiotic entry in mouse embryonic and  
germline stem cells. *Nat Commun* *7*, 11056. <https://doi.org/10.1038/ncomms11056>.
- 623 41. Samanta, M.K., Gayen, S., Harris, C., Maclary, E., Murata-Nakamura, Y., Malcore, R.M., Porter, R.S.,  
624 Garay, P.M., Vallianatos, C.N., Samollow, P.B., et al. (2022). Activation of Xist by an evolutionarily  
conserved function of KDM5C demethylase. *Nat Commun* *13*, 2602. <https://doi.org/10.1038/s41467-022-30352-1>.
- 625 42. Koubova, J., Menke, D.B., Zhou, Q., Capel, B., Griswold, M.D., and Page, D.C. (2006). Retinoic  
acid regulates sex-specific timing of meiotic initiation in mice. *Proc. Natl. Acad. Sci. U.S.A.* *103*,  
626 2474–2479. <https://doi.org/10.1073/pnas.0510813103>.
- 627 43. Lin, Y., Gill, M.E., Koubova, J., and Page, D.C. (2008). Germ Cell-Intrinsic and -Extrinsic Factors  
628 Govern Meiotic Initiation in Mouse Embryos. *Science* *322*, 1685–1687. <https://doi.org/10.1126/science.1166340>.

- 629 44. Endo, T., Mikedis, M.M., Nicholls, P.K., Page, D.C., and De Rooij, D.G. (2019). Retinoic Acid and Germ  
630 Cell Development in the Ovary and Testis. *Biomolecules* 9, 775. <https://doi.org/10.3390/biom9120775>.
- 631 45. Gupta, N., Yakhou, L., Albert, J.R., Azogui, A., Ferry, L., Kirsh, O., Miura, F., Battault, S., Yamaguchi,  
632 K., Laisné, M., et al. (2023). A genome-wide screen reveals new regulators of the 2-cell-like cell state.  
*Nat Struct Mol Biol.* <https://doi.org/10.1038/s41594-023-01038-z>.
- 633 46. Pohlers, M., Truss, M., Frede, U., Scholz, A., Strehle, M., Kuban, R.-J., Hoffmann, B., Morkel, M.,  
Birchmeier, C., and Hagemeyer, C. (2005). A Role for E2F6 in the Restriction of Male-Germ-Cell-  
634 Specific Gene Expression. *Current Biology* 15, 1051–1057. <https://doi.org/10.1016/j.cub.2005.04.060>.
- 635 47. Dahlet, T., Truss, M., Frede, U., Al Adhami, H., Bardet, A.F., Dumas, M., Vallet, J., Chicher, J.,  
Hammann, P., Kottnik, S., et al. (2021). E2F6 initiates stable epigenetic silencing of germline genes  
636 during embryonic development. *Nat Commun* 12, 3582. <https://doi.org/10.1038/s41467-021-23596-w>.
- 637 48. Agulnik, A.I., Mitchell, M.J., Mattei, M.G., Borsani, G., Avner, P.A., Lerner, J.L., and Bishop, C.E.  
(1994). A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation in mouse  
638 and human. *Hum Mol Genet* 3, 879–884. <https://doi.org/10.1093/hmg/3.6.879>.
- 639 49. Carrel, L., Hunt, P.A., and Willard, H.F. (1996). Tissue and lineage-specific variation in inactive  
X chromosome expression of the murine Smcx gene. *Hum Mol Genet* 5, 1361–1366. <https://doi.org/10.1093/hmg/5.9.1361>.
- 640 50. Sheardown, S., Norris, D., Fisher, A., and Brockdorff, N. (1996). The mouse Smcx gene exhibits  
developmental and tissue specific variation in degree of escape from X inactivation. *Hum Mol Genet*  
642 5, 1355–1360. <https://doi.org/10.1093/hmg/5.9.1355>.
- 643 51. Xu, J., Deng, X., and Disteche, C.M. (2008). Sex-Specific Expression of the X-Linked Histone  
Demethylase Gene Jarid1c in Brain. *PLoS ONE* 3, e2553. <https://doi.org/10.1371/journal.pone.0002553>.
- 644 52. Wang, P.J., McCarrey, J.R., Yang, F., and Page, D.C. (2001). An abundance of X-linked genes  
expressed in spermatogonia. *Nat Genet* 27, 422–426. <https://doi.org/10.1038/86927>.
- 645 53. Khil, P.P., Smirnova, N.A., Romanienko, P.J., and Camerini-Otero, R.D. (2004). The mouse X  
chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome  
646 inactivation. *Nat Genet* 36, 642–646. <https://doi.org/10.1038/ng1368>.
- 647 54. Al Adhami, H., Vallet, J., Schaal, C., Schumacher, P., Bardet, A.F., Dumas, M., Chicher, J., Hammann,  
P., Daujat, S., and Weber, M. (2023). Systematic identification of factors involved in the silencing  
648 of germline genes in mouse embryonic stem cells. *Nucleic Acids Research* 51, 3130–3149. <https://doi.org/10.1093/nar/gkad071>.
- 649 650

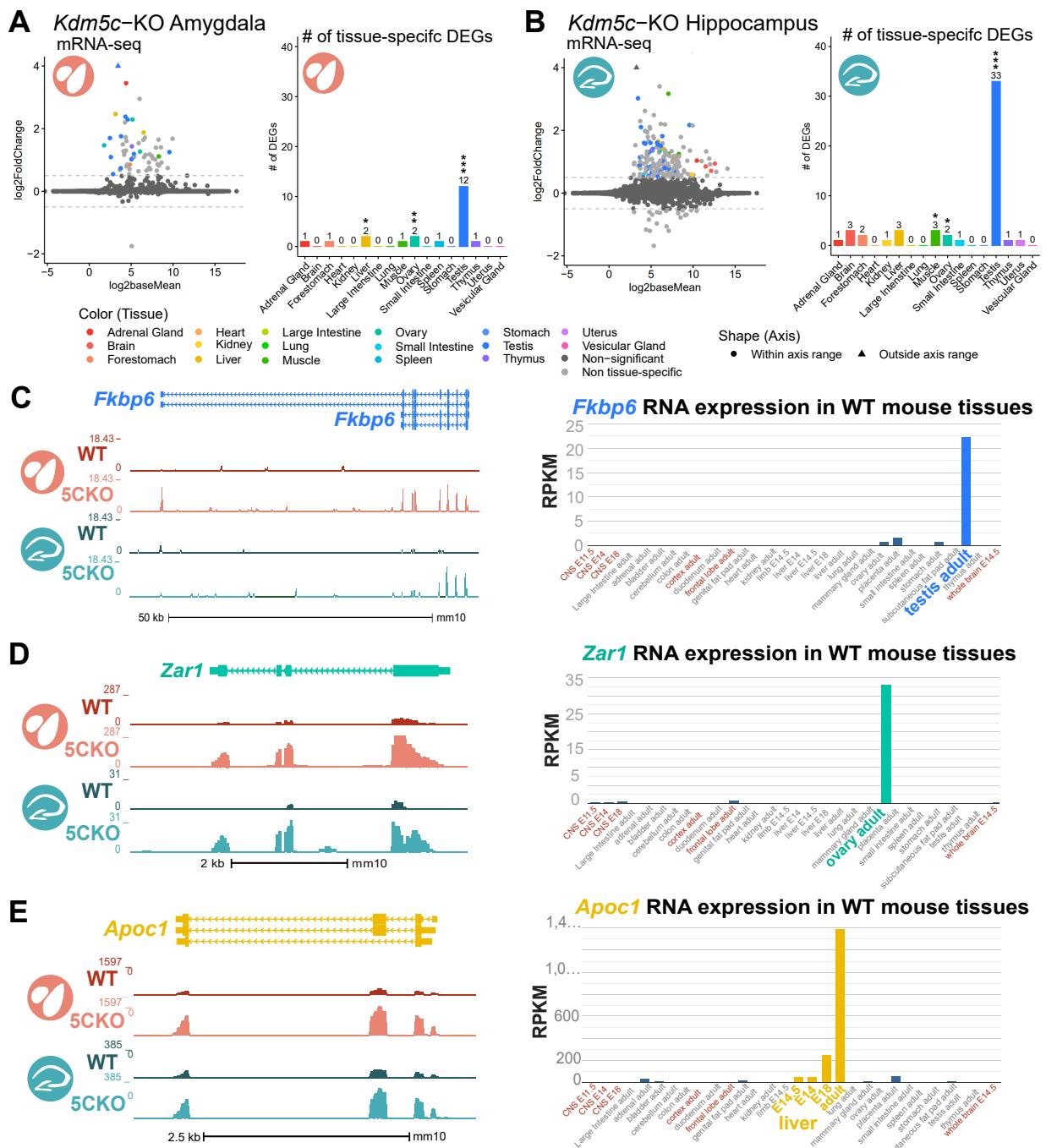
- 651 55. Hurlin, P.J. (1999). Mga, a dual-specificity transcription factor that interacts with Max and contains a  
652 T-domain DNA-binding motif. *The EMBO Journal* *18*, 7019–7028. <https://doi.org/10.1093/emboj/18.2.4.7019>.
- 653 56. Stielow, B., Finkernagel, F., Stiewe, T., Nist, A., and Suske, G. (2018). MGA, L3MBTL2 and E2F6  
654 determine genomic binding of the non-canonical Polycomb repressive complex PRC1.6. *PLoS Genet*  
655 *14*, e1007193. <https://doi.org/10.1371/journal.pgen.1007193>.
- 656 57. Tatsumi, D., Hayashi, Y., Endo, M., Kobayashi, H., Yoshioka, T., Kiso, K., Kanno, S., Nakai, Y., Maeda,  
I., Mochizuki, K., et al. (2018). DNMTs and SETDB1 function as co-repressors in MAX-mediated  
repression of germ cell-related genes in mouse embryonic stem cells. *PLoS ONE* *13*, e0205969.  
657 <https://doi.org/10.1371/journal.pone.0205969>.
- 658 58. Tahiliani, M., Mei, P., Fang, R., Leonor, T., Rutenberg, M., Shimizu, F., Li, J., Rao, A., and Shi, Y.  
(2007). The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation.  
Nature *447*, 601–605. <https://doi.org/10.1038/nature05823>.
- 659 59. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H.,  
and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime  
cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell* *38*, 576–589.  
660 <https://doi.org/10.1016/j.molcel.2010.05.004>.
- 661 60. Gajiwala, K.S., Chen, H., Cornille, F., Roques, B.P., Reith, W., Mach, B., and Burley, S.K. (2000).  
Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. *Nature* *403*,  
662 916–921. <https://doi.org/10.1038/35002634>.
- 663 61. Swoboda, P., Adler, H.T., and Thomas, J.H. (2000). The RFX-Type Transcription Factor DAF-19  
Regulates Sensory Neuron Cilium Formation in *C. elegans*. *Molecular Cell* *5*, 411–421. [https://doi.org/10.1016/S1097-2765\(00\)80436-0](https://doi.org/10.1016/S1097-2765(00)80436-0).
- 664 62. Ashique, A.M., Choe, Y., Karlen, M., May, S.R., Phamluong, K., Solloway, M.J., Ericson, J., and  
Peterson, A.S. (2009). The Rfx4 Transcription Factor Modulates Shh Signaling by Regional Control of  
665 Ciliogenesis. *Sci. Signal.* *2*. <https://doi.org/10.1126/scisignal.2000602>.
- 666 63. Kistler, W.S., Baas, D., Lemeille, S., Paschaki, M., Seguin-Estevez, Q., Barras, E., Ma, W., Duteyrat, J.-  
L., Morlé, L., Durand, B., et al. (2015). RFX2 Is a Major Transcriptional Regulator of Spermiogenesis.  
667 *PLoS Genet* *11*, e1005368. <https://doi.org/10.1371/journal.pgen.1005368>.
- 668 64. Wu, Y., Hu, X., Li, Z., Wang, M., Li, S., Wang, X., Lin, X., Liao, S., Zhang, Z., Feng, X., et al.  
(2016). Transcription Factor RFX2 Is a Key Regulator of Mouse Spermiogenesis. *Sci Rep* *6*, 20435.  
669 <https://doi.org/10.1038/srep20435>.
- 670 65. Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis  
for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L  
domain. *EMBO Reports* *10*, 1235–1241. <https://doi.org/10.1038/embor.2009.218>.

- 672
- 673 66. Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., et al. (2015).  
674 Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* **517**,  
640–644. <https://doi.org/10.1038/nature13899>.
- 675 67. Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein,  
676 B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and  
differentiated cells. *Nature* **454**, 766–770. <https://doi.org/10.1038/nature07107>.
- 677 68. Nassar, L.R., Barber, G.P., Benet-Pagès, A., Casper, J., Clawson, H., Diekhans, M., Fischer, C.,  
678 Gonzalez, J.N., Hinrichs, A.S., Lee, B.T., et al. (2023). The UCSC Genome Browser database: 2023  
update. *Nucleic Acids Research* **51**, D1188–D1195. <https://doi.org/10.1093/nar/gkac1072>.
- 679 69. Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueiroa, M.E., Melnick, A., and Mason,  
680 C.E. (2012). methylKit: A comprehensive R package for the analysis of genome-wide DNA methylation  
profiles. *Genome Biol* **13**, R87. <https://doi.org/10.1186/gb-2012-13-10-r87>.
- 681 70. Torres-Padilla, M.-E. (2020). On transposons and totipotency. *Philos Trans R Soc Lond B Biol Sci*  
682 **375**, 20190339. <https://doi.org/10.1098/rstb.2019.0339>.
- 683 71. Yang, M., Yu, H., Yu, X., Liang, S., Hu, Y., Luo, Y., Izsvák, Z., Sun, C., and Wang, J. (2022). Chemical-  
induced chromatin remodeling reprograms mouse ESCs to totipotent-like stem cells. *Cell Stem Cell*  
684 **29**, 400–418.e13. <https://doi.org/10.1016/j.stem.2022.01.010>.
- 685 72. Zhang, J., Zhang, M., Acampora, D., Vojtek, M., Yuan, D., Simeone, A., and Chambers, I. (2018).  
686 OTX2 restricts entry to the mouse germline. *Nature* **562**, 595–599. <https://doi.org/10.1038/s41586-018-0581-5>.
- 687 73. Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007).  
688 Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human  
genome. *Nat Genet* **39**, 457–466. <https://doi.org/10.1038/ng1990>.
- 689 74. Aubert, Y., Egolf, S., and Capell, B.C. (2019). The Unexpected Noncatalytic Roles of Histone Modifiers  
in Development and Disease. *Trends in Genetics* **35**, 645–657. <https://doi.org/10.1016/j.tig.2019.06.004>.
- 690 75. Morgan, M.A.J., and Shilatifard, A. (2020). Reevaluating the roles of histone-modifying enzymes  
and their associated chromatin modifications in transcriptional regulation. *Nat Genet* **52**, 1271–1281.  
691 <https://doi.org/10.1038/s41588-020-00736-4>.
- 692 76. Long, H.K., King, H.W., Patient, R.K., Odom, D.T., and Klose, R.J. (2016). Protection of CpG  
islands from DNA methylation is DNA-encoded and evolutionarily conserved. *Nucleic Acids Res* **44**,  
693 6693–6706. <https://doi.org/10.1093/nar/gkw258>.
- 694

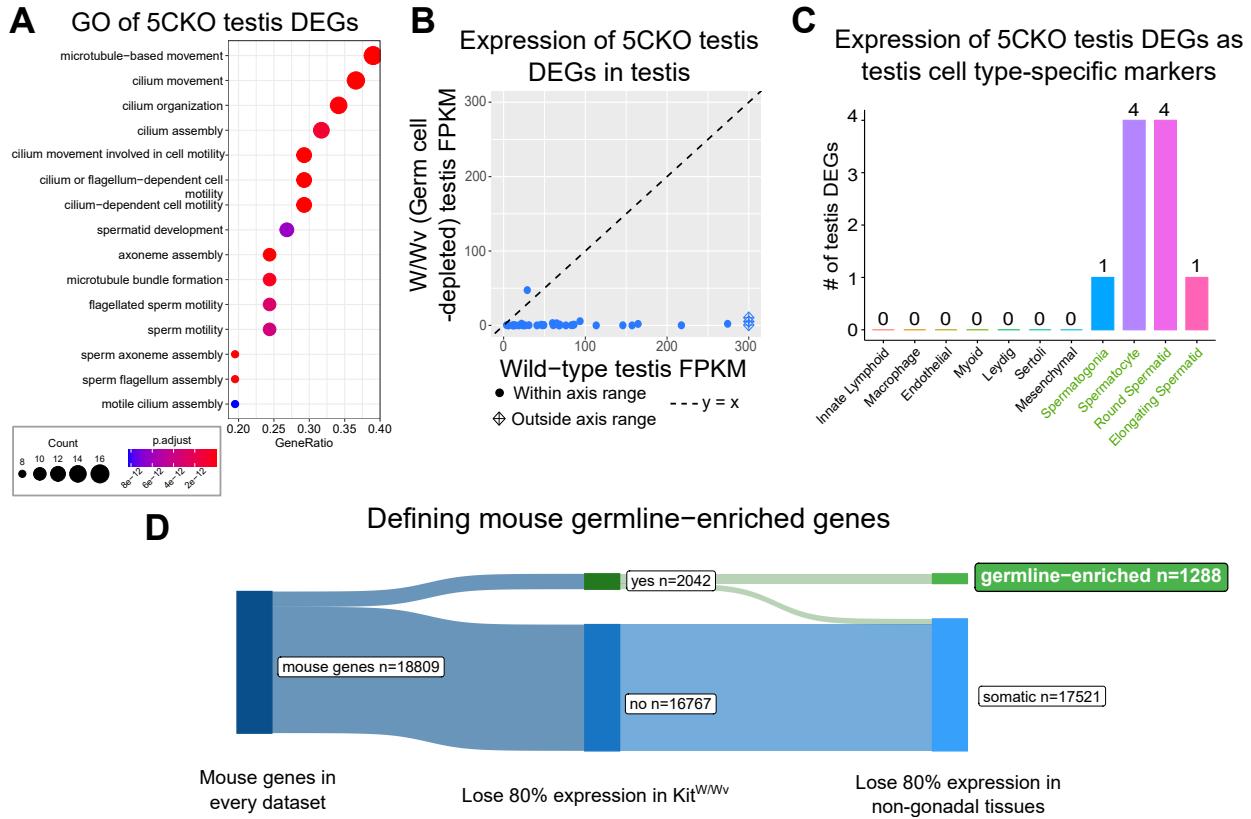
- 695 77. Leduc, V., Jasmin-Bélanger, S., and Poirier, J. (2010). APOE and cholesterol homeostasis in  
Alzheimer's disease. *Trends in Molecular Medicine* 16, 469–477. <https://doi.org/10.1016/j.molmed.2010.07.008>.
- 696
- 697 78. Abildayeva, K., Berbée, J.F.P., Blokland, A., Jansen, P.J., Hoek, F.J., Meijer, O., Lütjohann, D., Gautier,  
T., Pillot, T., De Vente, J., et al. (2008). Human apolipoprotein C-I expression in mice impairs learning  
and memory functions. *Journal of Lipid Research* 49, 856–869. <https://doi.org/10.1194/jlr.M700518JLR200>.
- 698
- 699 79. Wang, D., Kennedy, S., Conte, D., Kim, J.K., Gabel, H.W., Kamath, R.S., Mello, C.C., and Ruvkun,  
G. (2005). Somatic misexpression of germline P granules and enhanced RNA interference in  
700 retinoblastoma pathway mutants. *Nature* 436, 593–597. <https://doi.org/10.1038/nature04010>.
- 701 80. Wu, X., Shi, Z., Cui, M., Han, M., and Ruvkun, G. (2012). Repression of germline RNAi pathways  
in somatic cells by retinoblastoma pathway chromatin complexes. *PLoS Genet* 8, e1002542. <https://doi.org/10.1371/journal.pgen.1002542>.
- 702
- 703 81. Nielsen, A.Y., and Gjerstorff, M.F. (2016). Ectopic Expression of Testis Germ Cell Proteins in Cancer  
704 and Its Potential Role in Genomic Instability. *Int J Mol Sci* 17. <https://doi.org/10.3390/ijms17060890>.
- 705 82. Adebayo Babatunde, K., Najafi, A., Salehipour, P., Modarressi, M.H., and Mobasher, M.B. (2017).  
Cancer/Testis genes in relation to sperm biology and function. *Iranian Journal of Basic Medical  
706 Sciences* 20. <https://doi.org/10.22038/ijbms.2017.9259>.
- 707 83. Janic, A., Mendizabal, L., Llamazares, S., Rossell, D., and Gonzalez, C. (2010). Ectopic expression  
of germline genes drives malignant brain tumor growth in *Drosophila*. *Science* 330, 1824–1827.  
708 <https://doi.org/10.1126/science.1195481>.
- 709 84. Ghafouri-Fard, S., and Modarressi, M.-H. (2012). Expression of Cancer–Testis Genes in Brain Tumors:  
710 Implications for Cancer Immunotherapy. *Immunotherapy* 4, 59–75. <https://doi.org/10.2217/imt.11.145>.
- 711 85. Nin, D.S., and Deng, L.-W. (2023). Biology of Cancer-Testis Antigens and Their Therapeutic Implica-  
712 tions in Cancer. *Cells* 12, 926. <https://doi.org/10.3390/cells12060926>.
- 713 86. Bonefas, K.M., and Iwase, S. (2021). Soma-to-germline transformation in chromatin-linked neurode-  
714 velopmental disorders? *FEBS J.* <https://doi.org/10.1111/febs.16196>.
- 715 87. Velasco, G., Walton, E.L., Sterlin, D., Héduouin, S., Nitta, H., Ito, Y., Fouyssac, F., Mégarbané, A.,  
Sasaki, H., Picard, C., et al. (2014). Germline genes hypomethylation and expression define a  
molecular signature in peripheral blood of ICF patients: Implications for diagnosis and etiology.  
716 *Orphanet J Rare Dis* 9, 56. <https://doi.org/10.1186/1750-1172-9-56>.
- 717 88. Walton, E.L., Francastel, C., and Velasco, G. (2014). Dnmt3b Prefers Germ Line Genes and Cen-  
718 tromeric Regions: Lessons from the ICF Syndrome and Cancer and Implications for Diseases. *Biology  
(Basel)* 3, 578–605. <https://doi.org/10.3390/biology3030578>.

- 719 89. Graham-Paquin, A.-L., Saini, D., Sirois, J., Hossain, I., Katz, M.S., Zhuang, Q.K.-W., Kwon, S.Y.,  
Yamanaka, Y., Bourque, G., Bouchard, M., et al. (2023). ZMYM2 is essential for methylation of  
germline genes and active transposons in embryonic development. *Nucleic Acids Research*, gkad540.  
720 <https://doi.org/10.1093/nar/gkad540>.
- 721 90. Stephens, M. (2016). False discovery rates: A new deal. *Biostat*, kxw041. <https://doi.org/10.1093/biostatistics/kxw041>.
- 723 91. Conway, J.R., Lex, A., and Gehlenborg, N. (2017). UpSetR: An R package for the visualization of  
intersecting sets and their properties. *Bioinformatics* 33, 2938–2940. <https://doi.org/10.1093/bioinformatics/btx364>.
- 725 92. Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D.,  
Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated  
726 with clinical outcome in breast cancer. *Nature* 481, 389–393. <https://doi.org/10.1038/nature10730>.
- 727 93. Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: An R/Bioconductor package for ChIP peak  
annotation, comparison and visualization. *Bioinformatics* 31, 2382–2383. <https://doi.org/10.1093/bioinformatics/btv145>.
- 729 94. Krueger, F., and Andrews, S.R. (2011). Bismark: A flexible aligner and methylation caller for Bisulfite-  
730 Seq applications. *Bioinformatics* 27, 1571–1572. <https://doi.org/10.1093/bioinformatics/btr167>.

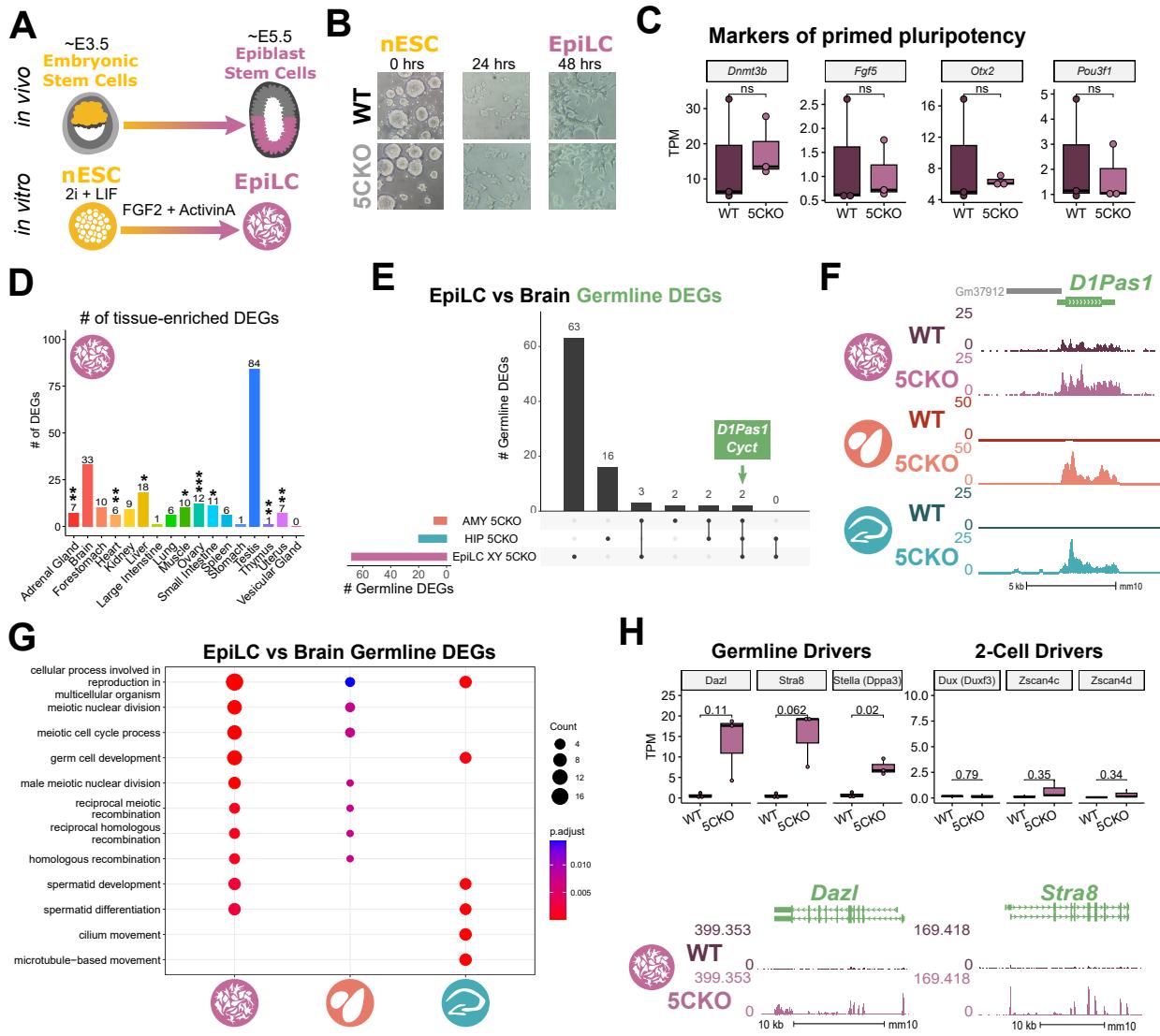
731 **Figures and Tables**



**Figure 1: Tissue-enriched genes are misexpressed in the *Kdm5c*-KO brain.** **A-B.** Expression of tissue-enriched genes (Li et al 2017) in the male *Kdm5c*-KO amygdala (A) and hippocampus (B). Left - MA plot of mRNA-sequencing. Right - Number of tissue-enriched differentially expressed genes (DEGs). \* p<0.05, \*\* p<0.01, \*\*\* p<0.001, Fisher's Exact Test. **C.** Left - UCSC browser view of an example aberrantly expressed testis-enriched DEG, *FK506 binding protein 6* (*Fkbp6*) in the wild-type (WT) and *Kdm5c*-KO (5CKO) amygdala (red) and hippocampus (teal) (Average, n = 4). Right - Expression of *Cyct* in wild-type tissues from NCBI Gene, with testis highlighted in blue and brain tissues highlighted in red. **D.** Left - UCSC browser view of an example ovary-enriched DEG, *Zygotic arrest 1* (*Zar1*). Right - Expression of *Zar1* in wild-type tissues from NCBI Gene, with ovary highlighted in teal and brain tissues highlighted in red. **E.** Left - UCSC browser view of an example liver-enriched DEG, *Apolipoprotein C-I* (*Apoc1*). Right - Expression of *Apoc1* in wild-type tissues from NCBI Gene, with liver highlighted in orange and brain tissues highlighted in red.



**Figure 2: Aberrant transcription of germline genes in the *Kdm5c*-KO in the brain.** **A.** enrichPlot gene ontology (GO) of *Kdm5c*-KO amygdala and hippocampus testis-enriched DEGs **B.** Expression of testis DEGs in wild-type (WT) testis versus germ cell-depleted (W/Wv) testis (Mueller et al 2013). Expression is in Fragments Per Kilobase of transcript per Million mapped reads (FPKM). **C.** Number of testis DEGs that were classified as cell-type specific markers in a single cell RNA-seq dataset of the testis (Green et al 2018). Germline cell types are highlighted in green, somatic cell types in black. **D.** Sankey diagram of mouse genes filtered for germline enrichment based on their expression in wild-type and W/Wv mice and in adult mouse non-gonadal tissues (Li et al 2017).



**Figure 3: Kdm5c-KO epiblast-like cells express key drivers of germline identity**

**A.** Top - Diagram of *in vivo* differentiation of embryonic stem cells (ESCs) of the inner cell mass into epiblast stem cells. Bottom - *in vitro* differentiation of ESCs into epiblast-like cells (EpiLCs).

**B.** Representative images of male wild-type (WT) and *Kdm5c*-KO (5CKO) cells during ESC to EpiLC differentiation. Brightfield images taken at 20X.

**C.** No significant difference in primed pluripotency marker expression in wild-type versus *Kdm5c*-KO EpiLCs. Welch's t-test, expression in transcripts per million (TPM).

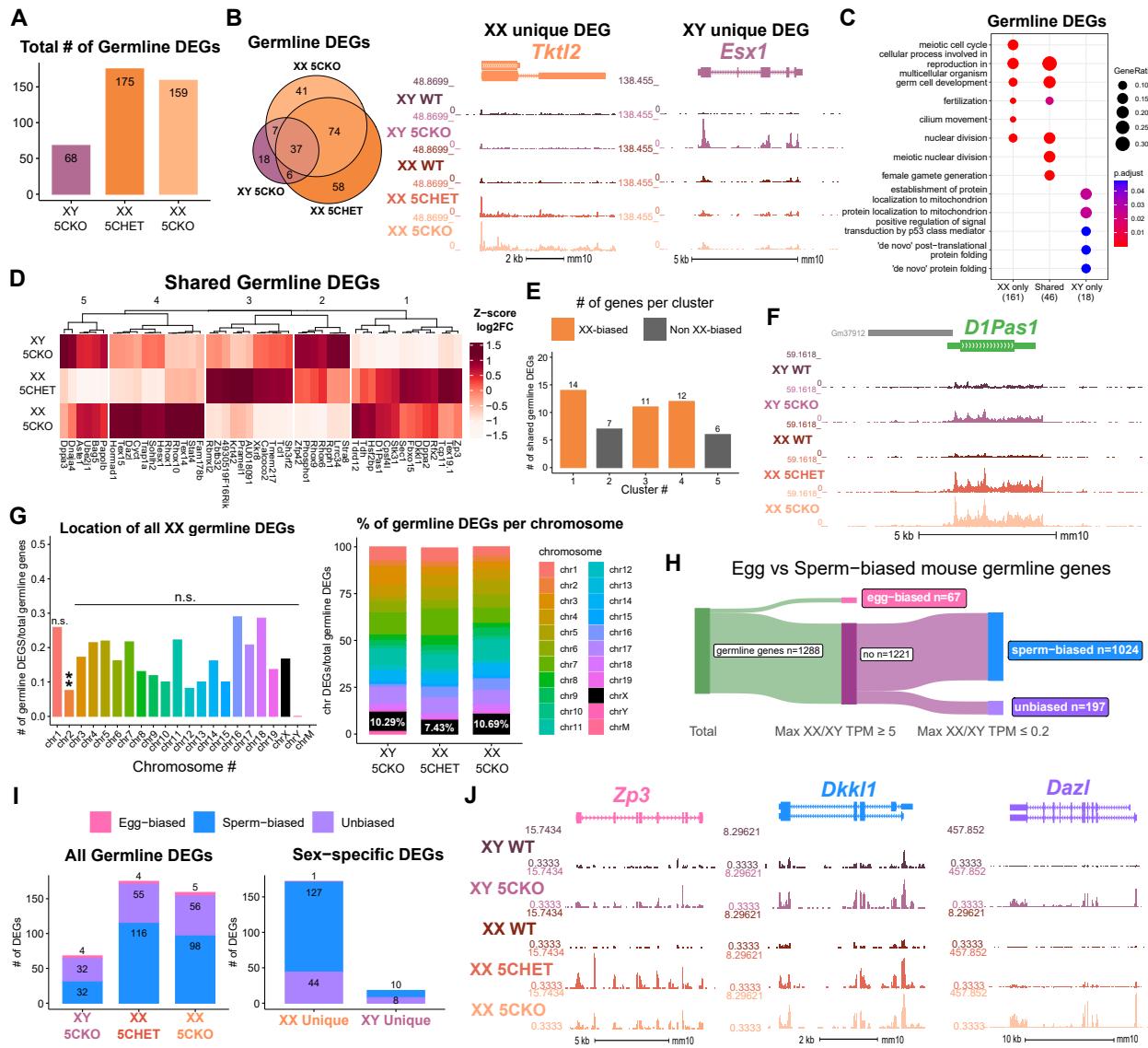
**D.** Number of tissue-enriched differentially expressed genes (DEGs) in *Kdm5c*-KO EpiLCs, amygdala (AMY), and hippocampus (HIP) RNA-seq datasets. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , Fisher's Exact Test.

**E.** Upset plot displaying the overlap of germline DEGs expressed in *Kdm5c*-KO EpiLCs, amygdala (AMY), and hippocampus (HIP) RNA-seq datasets.

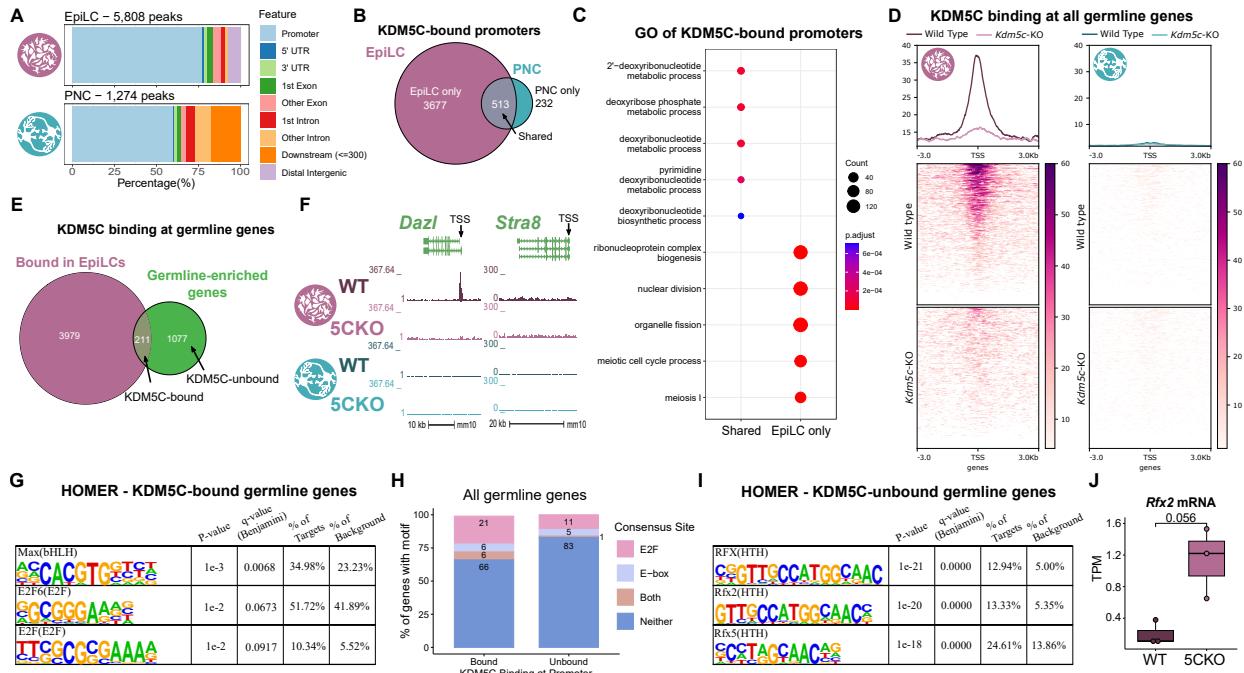
**F.** UCSC browser view of an example germline gene, *D1Pas1*, that is dysregulated *Kdm5c*-KO EpiLCs (top, purple. Average,  $n = 3$ ), amygdala (middle, red. Average,  $n = 4$ ), and hippocampus (bottom, blue. Average,  $n = 4$ ).

**G.** enrichPlot gene ontology analysis comparing enriched biological processes for *Kdm5c*-KO EpiLC, amygdala, and hippocampus germline DEGs.

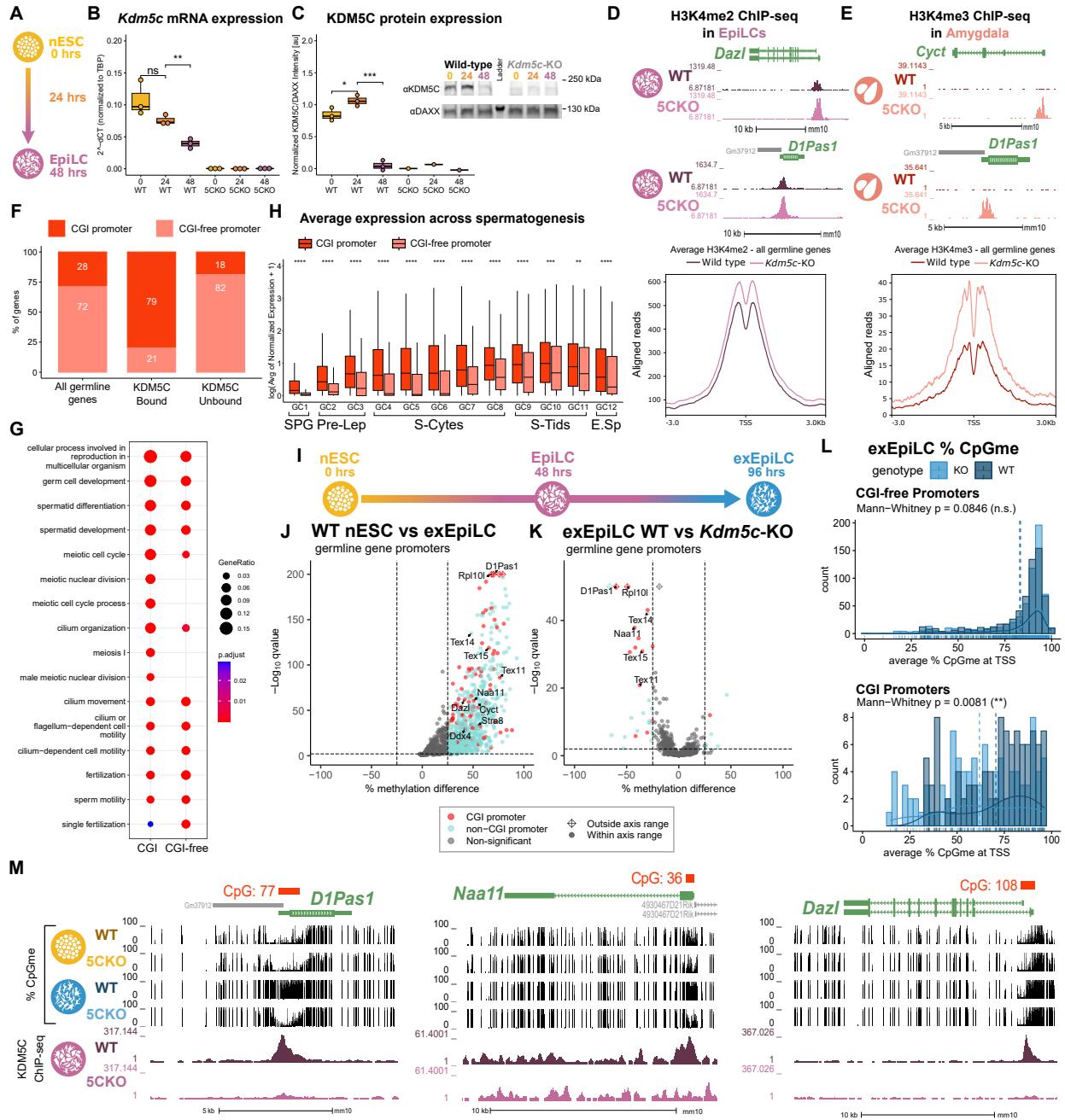
**H.** Top left - Example germline identity DEGs unique to EpiLCs. Top right - Example 2-cell genes that are not dysregulated in *Kdm5c*-KO EpiLCs. p-values for Welch's t-test. Bottom - UCSC browser view of *Dazl* and *Stra8* expression in wild-type and *Kdm5c*-KO EpiLCs (Average,  $n = 3$ ).



**Figure 4: Chromosomal sex influences *Kdm5c*-KO germline gene misexpression.** **A.** Total number of germline-enriched RNA-seq DEGs for male hemizygous *Kdm5c* knockout EpiLCs (XY 5CKO, purple), female heterozygous *Kdm5c* knockout (XX 5CHET, orange), female homozygous *Kdm5c* knockout (XX 5CKO, light orange) EpiLCs. **B.** Left - Eulerr overlap of *Kdm5c* mutant male and female EpiLC germline DEGs. Right - Example of germline DEGs unique to females or males, *Tktl2* and *Esx1*. **C.** enrichPlot gene ontology analysis comparing enriched biological processes for germline DEGs shared between *Kdm5c* mutant males and females (Shared), or unique to one sex (XX only or XY only). **D.** Heatmap of germline DEGs shared between male and female mutants. Color is the average log 2 fold change from sex-matched wild-type, z-scored across rows. **E.** Number of genes within each cluster from D. Clusters with higher expression in females compared to males (XX-biased) highlighted in orange. **F.** UCSC browser view of a male and female shared germline DEG *D1Pas1* that is more highly expressed in female mutants (Average, n = 3). **G.** Left - Number of all female germline DEGs located on each chromosome over the total number of germline-enriched genes on that chromosome. P-values for Fisher Exact Test, \*\* p < 0.01, n.s. non-significant. Germline DEGs were only significant for chromosome 2, in which they were significantly depleted. Right - Percentage of germline DEGs that lie on each chromosome for each *Kdm5c* mutant. X chromosome highlighted in black. **H.** Sankey diagram classifying egg-biased (pink) and sperm-biased (blue) and unbiased (purple) mouse germline-enriched genes. **I.** Number of egg, sperm, or unbiased germline DEGs for male and female *Kdm5c* mutants. **J.** UCSC browser view of egg-biased (*Zp3*), sperm-biased (*Dkk1*), and unbiased (*Dazl*) germline genes dysregulated in both male and female *Kdm5c* mutants (Average of n = 3).



**Figure 5: KDM5C binds to a subset of germline gene promoters during early embryogenesis.** **A.** ChIPseeker localization of KDM5C peaks at different genomic regions in EpiLCs (top) and hippocampal and cortex primary neuron cultures (PNCs, bottom). **B.** Overlap of genes with KDM5C bound to their promoters ( $TSS \pm 500$ ) in EpiLCs (purple) and PNCs (blue). **C.** Gene ontology (GO) comparison of genes with KDM5C bound to their promoter in EpiLCs and PNCs. Genes were classified as either bound in EpiLCs only (EpiLC only), unique to PNCs (PNC only, no significant ontologies) or bound in both PNCs and EpiLCs (shared). **D.** Average KDM5C binding around the transcription start site (TSS) of all germline-enriched genes in EpiLCs (left) and PNCs (right). **E.** Eulerr overlap of germline-enriched genes (green) with significant KDM5C binding at their promoter in EpiLCs (purple). **F.** Example KDM5C ChIP-seq signal around the *Dazl* TSS but not *Stra8* in EpiLCs. **G.** HOMER motif analysis of all KDM5C-bound germline gene promoters, highlighting significant enrichment of MAX, E2F6, and E2F motifs. **H.** Number of all gene promoters bound or unbound by KDM5C with instances of the E2F or E-box consensus sequence. **I.** HOMER motif analysis of all KDM5C-unbound germline gene promoters, highlighting significant enrichment of RFX family transcription factor motifs. **J.** Expression of RNA-seq DEG *Rfx2* in wild-type and *Kdm5c*-KO EpiLCs. P-value of Welch's t-test, expression in transcripts per million (TPM).



**Figure 6: KDM5C promotes long-term silencing of germline genes via DNA methylation at CpG islands.** **A.** Diagram of embryonic stem cell (ESC) to epiblast-like cell (EpiLC) differentiation and collection time points. **B.** Real time quantitative PCR (RT-qPCR) of *Kdm5c* mRNA expression in wild-type (WT) and *Kdm5c*-KO (5CKO) ESCs at 0, 24, and 48 hours of differentiation into EpiLCs. Expression calculated in comparison to TBP mRNA expression ( $2^{-\Delta\Delta CT}$ ). \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , Welch's t-test. **C.** KDM5C protein expression normalized to DAXX. Quantified intensity using ImageJ (artificial units - au). Right - representative lanes of Western blot for KDM5C and DAXX. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , Welch's t-test. **D.** Top - Representative UCSC browser view of histone 3 lysine 4 dimethylation (H3K4me2) ChIP-seq signal at two germline genes in wild-type and *Kdm5c*-KO EpiLCs. Bottom - Average H3K4me2 signal at the TSS of all germline-enriched genes in wild-type (dark purple) and *Kdm5c*-KO (light purple) EpiLCs. **E.** Top - Representative UCSC browser view of histone 3 lysine 4 trimethylation (H3K4me3) ChIP-seq signal at two germline genes in the wild-type (WT) and *Kdm5c*-KO (5CKO) adult amygdala. Bottom - Average H3K4me3 signal at the transcription start site (TSS) of all germline-enriched genes in wild-type (dark red) and *Kdm5c*-KO (light red) amygdala. **F.** Percentage of germline genes that harbor CpG islands (CGIs) in their promoters (CGI promoter, red), based on UCSC annotation. Left - percentage of all germline-enriched genes, middle - KDM5C-bound germline genes, and right - KDM5C-unbound germline genes. (Legend continued on next page.)

**Figure 6: KDM5C promotes long-term silencing of germline genes via DNA methylation at CpG islands.** (Legend continued.) **G.** enrichPlot gene ontology analysis of germline genes with (CGI-promoter) or without (CGI-free) CGIs in their promoter. **H.** Expression of germline genes with (CGI-promoter, red) or without (CGI-free, salmon) CGIs in their promoter across stages of spermatogenesis from Green et al 2018. SPG - spermatogonia, Pre-Lep - preleptotene spermatocytes, S-Cytes - meiotic spermatocytes, S-Tids - post-meiotic haploid round spermatids, E.Sp - elongating spermatids. Wilcoxon test, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ . **I.** Diagram of ESC to extended EpiLC (exEpiLC) differentiation. **J.** Volcano plot of whole genome bisulfite sequencing (WGBS) comparing CpG methylation (CpGme) at germline gene promoters ( $TSS \pm 500$ ) in wild-type (WT) ESCs versus exEpiLCs. Significantly differentially methylated promoters ( $q < 0.01$ ,  $|methyl\text{ation difference}| > 25\%$ ) with CGIs in red, CGI-free promoters in light blue **K.** Volcano plot of WGBS of wild-type (WT) versus *Kdm5c*-KO exEpiLCs for germline gene promoters. Promoters with CGIs in red, CGI-free promoters in light blue. **L.** Histogram of average percent CpGme at the promoter of germline genes with or without CGIs. Wild-type in navy and *Kdm5c*-KO (KO) in light blue. Dashed lines are average methylation for each genotype, p-values for Mann-Whitney U test. **M.** UCSC browser view of germline genes, showing UCSC annotated CGI with number of CpGs, representative CpGme in wild-type (WT) and *Kdm5c*-KO (5CKO) ESCs and exEpiLCs, and KDM5C ChIP-seq signal in EpiLCs.