

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^{1,2}, Ilakkiya Venkatachalam^{2,3}, and Shigeki Iwase².

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 The division labor among cellular lineages is a pivotal step in the evolution of multicellularity. In mammals,
13 the soma-germline boundary is formed during early embryogenesis, when genes that establish germline
14 identity are repressed in somatic lineages through DNA and histone modifications at promoter CpG is-
15 lands (CGI). Somatic misexpression of germline genes is a signature of cancer and observed in select
16 neurodevelopmental disorders. However it is currently unclear if all germline genes use the same repressive
17 mechanisms and if factors like development and sex influence their dysregulation. Here, we examine how
18 cellular context influences the development of somatic tissue identity in mice lacking lysine demethylase
19 5c (KDM5C), an X chromosome eraser of histone 3 lysine 4 di and tri-methylation (H3K4me2/3). We found
20 KDM5C is a crucial regulator of tissue identity, as male *Kdm5c* knockout (-KO) mice aberrantly express
21 many liver, muscle, and germline genes within the brain. By developing a comprehensive list of mouse
22 germline-enriched genes, we found the *Kdm5c*-KO brain aberrantly expresses late-stage spermatogenesis
23 genes while *Kdm5c*-KO epiblast-like cells (EpiLCs) primarily expressed key regulators of germline identity
24 and meiosis, such as *Dazl* and *Stra8*. KDM5C-mediated germline gene repression is sexually dimorphic, as
25 female EpiLCs require a higher dose of KDM5C to maintain germline silencing. KDM5C represses germline
26 genes during ESC to EpiLC differentiation, during which it is recruited to a subset of CGI-containing germline
27 identity gene promoters to facilitate CpG methylation. However the majority of germline gene promoters
28 did not harbor CGIs and were unbound by KDM5C, including late-stage spermatogenesis genes that are
29 expressed in *Kdm5c*-KO cells via ectopic activation by RFX transcription factors. Altogether, these data
30 demonstrate KDM5C's fundamental role in tissue identity and indicate that KDM5C acts as a sex-biased
31 brake against runaway activation of germline developmental programs in somatic lineages.

32 **notes** - How does sex and development influence germline gene repression? - How does KDM5C interact
33 with known silencing machinery? - Do all germline genes use teh same repressive mechanisms
34 aberrantly expressed , including *Dazl* and *Stra8*. - Germline genes are typically decommissioned in the
35 post-implantation epiblast, yet *Kdm5c*-KO epiblast-like cells (EpiLCs) aberrantly expressed key regulators of
36 germline identity and meiosis, including *Dazl* and *Stra8*.

- 37 • Distinguishing the germline from the soma is a key step in the evolution of multicellularity and sexual
38 reproduction.
- 39 • Germline gene repression is orchestrated by chromatin regulators and transcription factors.
- 40 • Much of these discoveries have been made looking at the repression of key marker genes for germ cell
41 specification in the early embryo,
- 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^{1–5}. Human genetic studies revealed impaired
49 chromatin regulation commonly occurs in both cancer^{6–8} and neurodevelopmental disorders (NDDs)⁹. While
50 many studies have identified their importance for regulating tumor suppressor genes and brain-specific
51 transcriptional programs, loss of chromatin regulators can also cause ectopic expression of tissue-specific
52 genes outside of their target environment, such as the misexpression of testis genes in colon tumors¹⁰ or
53 liver-specific genes within adult neurons¹¹. However, the mechanisms driving ectopic gene expression and
54 its impact upon cancer and neurodevelopment are still poorly understood.

55 Separation between germline and somatic cellular identity is a pivotal step in the evolution of multicellu-
56 larity and sexual reproduction^{12–14},[hanschenMulticellularityDrivesEvolution2018?](#) In mammals, chromatin regulators
57 decommission germline genes in somatic lineages when the early embryo transitions from naïve to primed
58 pluripotency. Initially, germline gene promoters gain repressive histone H2A lysine 119 monoubiquitination
59 (H2AK119ub1)¹⁵ and histone H3 lysine 9 trimethylation (H3K9me3)^{15,16} in embryonic stem cells (ESCs) and
60 are then decorated with DNA CpG methylation (CpGme) at their CpG islands (CGIs) in post-implantation
61 epiblast cells^{16–19}. While the silencing mechanisms for genes that establish germline identity are well charac-
62 terized, it is unclear if other types of germline genes, such as those involved in the later stages of oogenesis
63 and spermatogenesis, employ the same silencing mechanisms. Furthermore, because many studies have
64 focused on the silencing of key marker genes during early male embryonic development, much is unknown
65 about how cellular context (i.e. sex and tissue environment) influences the manifestation of germline gene
66 misexpression. Intriguingly, impaired soma-germline demarcation is a signature of aggressive cancers
67 and observed in select neurodevelopmental disorders (NDDs)^{8,20–22}. Thus, elucidating how cell context
68 contributes to germline gene dysregulation will reveal novel mechanisms governing these pathologies.

69 Here, we employed genome-wide analyses to explore loss of tissue identity in mice lacking the chromatin
70 regulator lysine demethylase 5C (KDM5C, also known as SMCX or JARID1C). KDM5C lies on the X
71 chromosome and erases histone 3 lysine 4 di- and trimethylation (H3K4me2/3), a permissive chromatin
72 modification enriched at gene promoters²³. Somatic loss of KDM5C promotes tumorigenicity in a variety
73 of cancer types^{24–26}, while germline mutations cause the NDD Intellectual Developmental Disorder, X-
74 linked, Syndromic, Claes-Jensen Type (MRXSCJ, OMIM: 300534). MRXSCJ is more common and severe
75 in males and its neurological phenotypes include intellectual disability, seizures, aberrant aggression,
76 and autistic behaviors^{27–29}. Male *Kdm5c* knockout (-KO) mice recapitulate key MRXSCJ phenotypes,
77 including hyperaggression, increased seizure propensity, social deficits, and learning impairments^{30–32}. RNA
78 sequencing (RNA-seq) of the *Kdm5c*-KO hippocampus revealed ectopic expression of some testis germline
79 genes within the brain³¹. However, it is unclear if other tissue-specific genes are aberrantly transcribed with

80 KDM5C loss, at what point in development germline gene misexpression begins, what mechanisms underlie
81 their dysregulation, and how KDM5C interacts with other known germline silencing mechanisms.

82 To illuminate KDM5C's role in tissue identity, we characterized the aberrant expression of tissue-enriched
83 genes within the *Kdm5c*-KO brain and epiblast-like stem cells (EpiLCs), an *in vitro* model of the post-
84 implantation embryo. We curated a list of mouse germline-enriched genes, enabling genome-wide analysis
85 of germline gene silencing mechanisms for the first time. Additionally, we characterized germline transcripts
86 expressed in male and female *Kdm5c* mutants to illuminate the impact of sex upon germline gene suppression.
87 Based on the data presented below, we propose KDM5C plays a fundamental, sexually dimorphic role in the
88 development of tissue identity during early embryogenesis, including the establishment of the soma-germline
89 boundary.

90 Results

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

92 Previous RNA sequencing (RNA-seq) of the adult male *Kdm5c*-KO hippocampus revealed ectopic
93 expression of some germline genes unique to the testis³¹. It is currently unknown if the testis is the only
94 tissue type misexpressed in the *Kdm5c*-KO brain. We first systematically tested whether other tissue-specific
95 genes are misexpressed in the male brain with constitutive knockout of *Kdm5c* (*Kdm5c*^y, 5CKO in figures)³³
96 by using a published list of mouse tissue-enriched genes³⁴.

97 We found a large proportion of significantly upregulated genes (DESeq2³⁵, log2 fold change > 0.5, q <
98 0.1) within the male *Kdm5c*-KO amygdala and hippocampus are non-brain, tissue-specific genes (Amygdala:
99 0/0 up DEGs, NaN% ; Hippocampus: 0/0 up DEGs, NaN%) (Figure 1A-B, Supplementary Table 1). For both
100 the amygdala and hippocampus, the majority of tissue-enriched differentially expressed genes (DEGs) were
101 testis genes (Figure 1A-B). Even though the testis has the largest total number of tissue-enriched genes
102 (2,496 genes) compared to any other tissue, testis-enriched DEGs were significantly enriched in both brain
103 regions (Amygdala p = 1.83e-05, Odds Ratio = 5.13; Hippocampus p = 4.26e-11, Odds Ratio = 4.45, Fisher's
104 Exact Test). An example of a testis-enriched gene misexpressed in the *Kdm5c*-KO brain is *FK506 binding*
105 *protein 6* (*Fkbp6*), a known regulator of PIWI-interacting RNAs (piRNAs) and meiosis^{36,37} (Figure 1C).

106 Interestingly, we also observed significant enrichment of ovary-enriched genes in both the amygdala
107 and hippocampus (Amygdala p = 0.00574, Odds Ratio = 18.7; Hippocampus p = 0.048, Odds Ratio = 5.88,
108 Fisher's Exact Test) (Figure 1A-B). Ovary-enriched DEGs included *Zygotic arrest 1* (*Zar1*), which sequesters
109 mRNAs in oocytes for meiotic maturation³⁸ (Figure 1D). Given that the *Kdm5c*-KO mice we analyzed are
110 male, these data demonstrate that the ectopic expression of gonad-enriched genes is independent of
111 organismal sex.

112 Although not consistent across brain regions, we also found significant enrichment of genes biased

113 towards two non-gonadal tissues - the liver (Amygdala p = 0.04, Odds Ratio = 6.58, Fisher's Exact Test)
114 and muscle (Hippocampus p = 0.01, Odds Ratio = 6.95, Fisher's Exact Test) (Figure 1A-B). These include
115 *Apolipoprotein C-I* (*Apoc1*), a lipoprotein metabolism and transport gene³⁹ (Figure 1E, see Discussion).

116 Our analysis of oligo(dT)-primed libraries³³ indicates aberrantly expressed mRNAs are polyadenylated
117 and spliced into mature transcripts in the *Kdm5c*-KO brain (Figure 1C-E). Of note, we observed little to no
118 dysregulation of brain-enriched genes (Amygdala p = 1, Odds Ratio = 1.22; Hippocampus p = 0.74, Odds
119 Ratio = 1.22, Fisher's Exact Test), despite the fact these are brain samples and the brain has the second
120 highest total number of tissue-enriched genes (708 genes). Altogether, these results suggest the aberrant
121 expression of tissue-enriched genes within the brain is a major effect of KDM5C loss.

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

123 *Kdm5c*-KO brain expresses testicular germline genes³¹ (Figure 1), however the testis also contains
124 somatic cells that support hormone production and germline functions. To determine if *Kdm5c*-KO results
125 in ectopic expression of testicular somatic genes, we first evaluated the known functions of testicular
126 DEGs through gene ontology. We found *Kdm5c*-KO testis-enriched DEGs had high enrichment of germline-
127 relevant ontologies, including spermatid development (GO: 0007286, p.adjust = 6.2e-12) and sperm axoneme
128 assembly (GO: 0007288, p.adjust = 2.45e-14) (Figure 2A, Supplementary Table 1).

129 We then evaluated *Kdm5c*-KO testicular DEG expression in wild-type testes versus testes with germ cell
130 depletion⁴⁰, which was accomplished by heterozygous *W* and *Wv* mutations in the enzymatic domain of *c-Kit*
131 (*Kit*^{W/Wv})⁴¹. Almost all *Kdm5c*-KO testis-enriched DEGs lost expression with germ cell depletion (Figure 2B).
132 We then assessed testis-enriched DEG expression in a published single cell RNA-seq dataset that identified
133 cell type-specific markers within the testis⁴². Some *Kdm5c*-KO testis-enriched DEGs were classified as
134 specific markers for different germ cell developmental stages (e.g. spermatogonia, spermatocytes, round
135 spermatids, and elongating spermatids), yet none marked somatic cells (Figure 2C). Together, these data
136 demonstrate that the *Kdm5c*-KO brain aberrantly expresses germline genes but not somatic testicular genes,
137 reflecting an erosion of the soma-germline boundary.

138 As of yet, research on germline gene silencing mechanisms has focused on a handful of key genes rather
139 than assessing germline gene suppression genome-wide, due to the lack of a comprehensive gene list.
140 We therefore generated a list of mouse germline-enriched genes using RNA-seq datasets of *Kit*^{W/Wv} mice
141 that included males and females at embryonic day 12, 14, and 16⁴³ and adult male testes⁴⁰. We defined
142 genes as germline-enriched if their expression met the following criteria: 1) their expression is greater than
143 1 FPKM in wild-type gonads 2) their expression in any non-gonadal tissue of adult wild type mice³⁴ does
144 not exceed 20% of their maximum expression in the wild-type germline, and 3) their expression in the germ
145 cell-depleted gonads, for any sex or time point, does not exceed 20% of their maximum expression in the
146 wild-type germline. These criteria yielded 1,288 germline-enriched genes (Figure 2D), which was hereafter
147 used as a resource to globally characterize germline gene misexpression with *Kdm5c* loss (Supplementary

148 Table 2).

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

150 Germ cells are typically distinguished from somatic cells soon after the embryo implants into the uterine
151 wall^{44,45}, when germline genes are silenced in epiblast stem cells that will form the somatic tissues⁴⁶. This
152 developmental time point can be modeled *in vitro* through differentiation of naïve embryonic stem cells
153 (nESCs) into epiblast-like stem cells (EpiLCs) (Figure 3A)^{47,48}. While some germline-enriched genes are
154 also expressed in nESCs and in the 2-cell stage^{49–51}, they are silenced as they differentiate into EpiLCs^{16,17}.
155 Therefore, we tested if KDM5C was necessary for the initial silencing of germline genes in somatic lineages
156 by evaluating the impact of *Kdm5c* loss in male EpiLCs.

157 *Kdm5c*-KO cell morphology during ESC to EpiLC differentiation appeared normal (Figure 3B) and EpiLCs
158 properly expressed markers of primed pluripotency, such as *Dnmt3b*, *Fgf5*, *Pou3f1*, and *Otx2* (Figure 3C). We
159 then identified tissue-enriched DEGs in a RNA-seq dataset of wild-type and *Kdm5c*-KO EpiLCs⁵² (DESeq2,
160 log2 fold change > 0.5, q < 0.1, Supplementary Table 3). Similar to the *Kdm5c*-KO brain, we observed
161 general dysregulation of tissue-enriched genes, with the largest number of genes belonging to the brain and
162 testis, although they were not significantly enriched (Figure 3D). Using our list of mouse germline-enriched
163 genes assembled above, we identified 68 germline genes misexpressed in male *Kdm5c*-KO EpiLCs.

164 We then compared EpiLC germline DEGs to those expressed in the *Kdm5c*-KO brain to determine if
165 germline genes are constitutively dysregulated or change over the course of development. The majority of
166 germline DEGs were unique to either EpiLCs or the brain, with only *D1Pas1* and *Cyct* shared across all
167 tissue/cell types (Figure 3E-F). EpiLC germline DEGs had particularly high enrichment of meiosis-related
168 gene ontologies when compared to the brain (Figure 3G, Supplementary Table 3), such as meiotic cell
169 cycle process (GO:1903046, p.adjust = 2.2e-07) and meiotic nuclear division (GO:0140013, p.adjust
170 = 1.37e-07). While there was modest enrichment of meiotic gene ontologies in both brain regions, the
171 *Kdm5c*-KO hippocampus primarily expressed late-stage spermatogenesis genes involved in sperm axoneme
172 assembly (GO:0007288, p.adjust = 0.00621) and sperm motility (GO:0097722, p.adjust = 0.00612).

173 Notably, DEGs unique to *Kdm5c*-KO EpiLCs included key drivers of germline identity, such as *Stimulated*
174 *by retinoic acid 8* (*Stra8*: log2 fold change = 3.73, q = 2.17e-39) and *Deleted in azoospermia like* (*Dazl*:
175 log2 fold change = 3.36, q = 3.19e-12) (Figure 3H). These genes are typically expressed when a subset
176 of epiblast stem cells become primordial germ cells (PGCs) and then again in mature germ cells to trigger
177 meiotic gene expression programs^{53–55}. Of note, some germline genes, including *Dazl*, are also expressed
178 in the two-cell embryo^{50,56}. However, we did not see derepression of two-cell stage-specific genes, like
179 *Duxf3* (*Dux*) (log2 fold change = -0.282, q = 0.337) and *Zscan4d* (log2 fold change = 0.25, q = 0.381) (Figure
180 3H, Supplementary Table 3), indicating *Kdm5c*-KO EpiLCs do not revert back to a 2-cell state. Altogether,
181 *Kdm5c*-KO EpiLCs express key drivers of germline identity and meiosis while the brain primarily expresses
182 spermiogenesis genes, indicating germline gene misexpression mirrors germline development during the

183 progression of somatic development.

184 **Female epiblast-like cells have heightened germline gene misexpression with *Kdm5c***
185 **loss**

186 It is currently unknown if the misexpression of germline genes is influenced by sex, as previous studies
187 on germline gene repressors have focused on male cells^{15,16,18,57,58}. Sex is particularly pertinent in the case
188 of KDM5C because it partially escapes X chromosome inactivation (XCI), resulting in a higher dosage in
189 females⁵⁹⁻⁶². We therefore explored the impact of chromosomal sex upon germline gene suppression by
190 comparing their dysregulation in male *Kdm5c* hemizygous knockout (*Kdm5c*^{-y}, XY *Kdm5c*-KO, XY 5CKO),
191 female homozygous knockout (*Kdm5c*^{-/-}, XX *Kdm5c*-KO, XX 5CKO), and female heterozygous knockout
192 (*Kdm5c*^{-/+}, XX *Kdm5c*-HET, XX 5CHET) EpiLCs⁵².

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

201 The majority of germline genes expressed in both sexes were more highly dysregulated in females
202 compared to males (Figure 4D-F). This increased degree of dysregulation in females, along with the
203 increased total number of germline genes, indicates females are more sensitive to losing KDM5C-mediated
204 germline gene suppression. Heightened germline gene dysregulation in females could be due to impaired
205 XCI in *Kdm5c* mutants⁵², as many spermatogenesis genes lie on the X chromosome^{63,64}. However, female
206 germline DEGs were not biased towards the X chromosome (p = 1, Odds Ratio = 0.96, Fisher's Exact Test)
207 and females had a similar overall proportion of germline DEGs belonging to the X chromosome as males
208 (XY *Kdm5c*-KO - 10.29%, XX *Kdm5c*-HET - 7.43%, XX *Kdm5c*-KO - 10.59%) (Figure 4G). The majority of
209 germline DEGs instead lie on autosomes for both male and female *Kdm5c* mutants (Figure 4G). Thus, while
210 female EpiLCs are more prone to germline gene misexpression with KDM5C loss, it is likely independent of
211 XCI defects.

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

213 Although many germline genes have shared functions in the male and female germline, e.g. PGC
214 formation, meiosis, and genome defense, some have unique or sex-biased expression. Therefore, we
215 wondered if *Kdm5c* mutant males would primarily express sperm genes while mutant females would primarily

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

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

227 KDM5C binds to the promoters of several germline genes in embryonic stem cells (ESCs) but not in
228 neurons^{31,65}. However, due to the lack of a comprehensive list of germline-enriched genes, it is unclear if
229 KDM5C is enriched at germline gene promoters, what types of germline genes KDM5C regulates, and if its
230 binding is maintained at any germline genes in neurons.

231 To address these questions, we analyzed KDM5C chromatin immunoprecipitation followed by DNA
232 sequencing (ChIP-seq) datasets in EpiLCs⁵² and primary forebrain neuron cultures (PNCs)³⁰ (MACS2 q <
233 0.1, fold enrichment > 1, and removal of false-positive *Kdm5c*-KO peaks). EpiLCs had a higher total number
234 of high-confidence KDM5C peaks than PNCs (EpiLCs: 5,808, PNCs: 1,276). KDM5C was primarily localized
235 to gene promoters in both cell types (promoters = transcription start site (TSS) ± 500 bp, EpiLCs: 4,190,
236 PNCs: 745), although PNCs showed increased localization to non-promoter regions (Figure 5A).

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

251 EpiLCs (*D1Pas1*, *Hsf2bp*, *Cyct*, and *Stk31*) (Supplementary Figure 1A). Together, these results demonstrate
252 KDM5C is recruited to a subset of germline genes in EpiLCs, including meiotic genes, but does not directly
253 regulate germline genes in neurons. Furthermore, the majority of germline mRNAs expressed in *Kdm5c*-KO
254 cells are dysregulated independent of direct KDM5C recruitment to their gene promoters, however genes
255 dysregulated across *Kdm5c*-KO development are often direct KDM5C targets.

256 Many germline-specific genes are suppressed by the polycomb repressive complex 1.6 (PRC1.6), which
257 contains the transcription factor heterodimers E2F6/DP1 and MGA/MAX that respectively bind E2F and
258 E-box motifs within germline gene promoters^{15,16,18,51,57,58,66–68}. PRC1.6 members may recruit KDM5C to
259 germline gene promoters³¹, given their association with KDM5C in HeLa cells and ESCs^{56,69}. We thus
260 used HOMER⁷⁰ to identify transcription factor motifs enriched at KDM5C-bound or unbound germline gene
261 promoters (TSS ± 500 bp, q-value < 0.1, Supplementary Table 4). MAX and E2F6 binding sites were
262 significantly enriched at germline genes bound by KDM5C in EpiLCs (MAX q-value: 0.0068, E2F6 q-value:
263 0.0673, E2F q-value: 0.0917), but not at germline genes unbound by KDM5C (Figure 5G). One third of
264 KDM5C-bound promoters contained the consensus sequence for either E2F6 (E2F, 5'-TCCCGC-3'), MGA
265 (E-box, 5'-CACGTG-3'), or both, but only 17% of KDM5C-unbound genes contained these motifs (Figure 5H).
266 KDM5C-unbound germline genes were instead enriched for multiple RFX transcription factor binding sites
267 (RFX q-value < 0.0001, RFX2 q-value < 0.0001, RFX5 q-value < 0.0001) (Figure 5I, Supplementary figure
268 1D). RFX transcription factors bind X-box motifs⁷¹ to promote ciliogenesis^{72,73} and among them is RFX2, a
269 central regulator of post-meiotic spermatogenesis^{74,75}. Although *Rfx2* is also not a direct target of KDM5C
270 (Supplementary Figure 1E), RFX2 mRNA is derepressed in *Kdm5c*-KO EpiLCs (Figure 5J). Thus, RFX2 is a
271 candidate transcription factor for driving the ectopic expression of many KDM5C-unbound germline genes in
272 *Kdm5c*-KO cells.

273 **KDM5C is recruited to CpG islands at germline promoters to facilitate *de novo* DNA
274 methylation**

275 Previous work found two germline gene promoters have a marked reduction in DNA CpG methylation
276 (CpGme) in the adult *Kdm5c*-KO hippocampus³¹. Since histone H3K4me2/3 impede *de novo* CpGme^{76,77},
277 KDM5C's removal of H3K4me2/3 may be required to suppress germline genes. However, KDM5C's catalytic
278 activity was recently shown to be dispensable for suppressing *Dazl* in undifferentiated ESCs⁵⁶. To reconcile
279 these observations, we hypothesized KDM5C erases H3K4me2/3 to promote the initial placement of CpGme
280 at germline gene promoters in EpiLCs.

281 To test this hypothesis, we first characterized KDM5C's expression as naïve ESCs differentiate into
282 EpiLCs (Figure 6A). While *Kdm5c* mRNA steadily decreased from 0 to 48 hours of differentiation (Figure
283 6B), KDM5C protein initially increased from 0 to 24 hours and then decreased to near knockout levels by 48
284 hours (Figure 6C). We then characterized KDM5C's substrates (H3K4me2/3) at germline gene promoters

285 with *Kdm5c* loss using published ChIP-seq datasets^{33,52}. *Kdm5c*-KO samples showed a marked increase in
286 H3K4me2 in EpiLCs (Figure 6D) and H3K4me3 in the amygdala (Figure 6E) around the TSS of germline
287 genes. Together, these data suggest KDM5C acts during the transition between ESCs and EpiLCs to remove
288 H3K4me2/3 at germline gene promoters.

289 Germline genes accumulate CpG methylation (CpGme) at CpG islands (CGIs) during the transition
290 from naïve to primed pluripotency^{17,19,78}. We first examined how many of our germline-enriched genes had
291 promoter CGIs (TSS ± 500 bp) using the UCSC genome browser⁷⁹. Notably, out of 1,288 germline-enriched
292 genes, only 356 (27.64%) had promoter CGIs (Figure 6F, Supplementary Table 2). CGI-containing germline
293 genes had higher enrichment of meiotic gene ontologies compared to CGI-free genes, including meiotic
294 nuclear division (GO:0140013, p.adjust = 2.17e-12) and meiosis I (GO:0007127, p.adjust = 3.91e-10)
295 (Figure 6G, Supplementary Table 5). Germline genes with promoter CGIs were more highly expressed than
296 CGI-free genes across spermatogenesis stages, with highest expression in meiotic spermatocytes (Figure
297 6H). Contrastingly, CGI-free genes only displayed substantial expression in post-meiotic round spermatids
298 (Figure 6H). Although only a minor portion of germline gene promoters contained CGIs, CGIs strongly
299 determined KDM5C's recruitment to germline genes ($p = 2.37e-67$, Odds Ratio = 17.8, Fisher's Exact Test),
300 with 79.15% of KDM5C-bound germline gene promoters harboring CGIs (Figure 6F).

301 To assess how KDM5C loss impacts initial CpGme placement at germline gene promoters, we performed
302 whole genome bisulfite sequencing (WGBS) in male wild-type and *Kdm5c*-KO ESCs and 96-hour extend
303 EpiLCs (exEpiLCs), when germline genes reach peak methylation levels¹⁶ (Figure 6I). We first identified
304 which germline gene promoters significantly gained CpGme in wild-type cells during nESC to exEpiLCs
305 differentiation (methylKit⁸⁰, $q < 0.01$, $|methylation\ difference| > 25\%$, TSS ± 500 bp). In wild-type cells, the
306 majority of germline genes gained substantial CpGme at their promoter during differentiation (60.08%),
307 regardless if their promoter contained a CGI (Figure 6J, Supplementary Table 5).

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

317 Promoters that showed the most robust loss of CpGme in *Kdm5c*-KO exEpiLCs (lowest q-values) harbored
318 CGIs (Figure 6K). CGI promoters, but not CGI-free promoters, had a significant reduction in CpGme with
319 KDM5C loss as a whole (Figure 6L) (Non-CGI promoters $p = 0.0846$, CGI promoters $p = 0.0081$, Mann-
320 Whitney U test). Significantly hypomethylated promoters included germline genes consistently dysregulated

321 across multiple *Kdm5c*-KO RNA-seq datasets³¹, such as *D1Pas1* (methylation difference = -60.03%, q-value
322 = 3.26e-153) and *Naa11* (methylation difference = -42.45%, q-value = 1.44e-38) (Figure 6M). Unexpectedly,
323 we observed only a modest reduction in CpGme at *Dazl*'s promoter (methylation difference = -6.525%,
324 q-value = 0.0159) (Figure 6N). Altogether, these results demonstrate KDM5C is recruited to germline gene
325 CGIs in EpiLCs to promote CpGme at those promoters. Furthermore, our data suggest while KDM5C's
326 catalytic activity is required for the repression of some germline genes, CpGme can be placed at others even
327 with elevated H3K4me2/3 around the TSS.

328 Discussion

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

345 By comparing *Kdm5c* mutant males and females, we revealed germline gene suppression is sexually
346 dimorphic. Female EpiLCs are more severely impacted by loss of KDM5C-mediated germline gene sup-
347 pression, yet this difference is not due to the large number of germline genes on the X chromosome^{63,64}.
348 Heightened germline gene misexpression in females may be related to females having a higher dose of
349 KDM5C than males, due to its escape from XCI^{59–62}. Intriguingly, heterozygous knockout females (*Kdm5c*^{+/−})
350 also had over double the number of germline DEGs than hemizygous knockout males (*Kdm5c*^{−/−}), even
351 though their expression of KDM5C should be roughly equivalent to that of wild-type males (*Kdm5c*^{+/+}). Males
352 could partially compensate for KDM5C's loss via the Y-chromosome homolog, KDM5D²³. However, KDM5D
353 has not been reported to regulate germline gene expression. Nevertheless, these results demonstrate
354 germline gene silencing mechanisms differ between males and females, which warrants further study to

355 elucidate the biological ramifications and underlying mechanisms.

356 We found KDM5C is largely dispensable for promoting normal gene expression during development, yet
357 is critical for suppressing ectopic developmental programs. While some germline genes, such as *Dazl*, are
358 also expressed in the 2-cell stage, the inner cell mass, and naïve ESCs, they are silenced in epiblast stem
359 cells/EpiLCs^{16,51,56,81,82}. Our data suggest the 2-cell-like state reported in *Kdm5c*-KO ESCs⁵⁶ likely reflects
360 KDM5C's primary role in germline gene repression (Figure 3). Germline gene misexpression in *Kdm5c*-
361 KO EpiLCs may indicate they are differentiating into primordial germ cell-like cells (PGCLCs)^{44,45,47}. Yet,
362 *Kdm5c*-KO EpiLCs had normal cellular morphology and properly expressed markers for primed pluripotency,
363 including *Otx2* which blocks EpiLC differentiation into PGCs/PGCLCs⁸³. In addition to unimpaired EpiLC
364 differentiation, *Kdm5c*-KO gross brain morphology is overall normal³⁰ and hardly any brain-specific genes
365 were significantly dysregulated in the amygdala and hippocampus (Figure 1). Thus, ectopic germline gene
366 expression occurs in conjunction with overall proper somatic differentiation in *Kdm5c*-KO animals.

367 Our work provides novel insight into the cross-talk between H3K4me2/3 and CpGme, which are gen-
368 erally mutually exclusive⁸⁴. In EpiLCs, loss of KDM5C binding at a subset of germline gene promoters,
369 e.g. *D1Pas1*, strongly impaired promoter CGI methylation and resulted in their long-lasting de-repression
370 into adulthood. Removal of H3K4me2/3 at CGIs is a plausible mechanism for KDM5C-mediated germline
371 gene suppression^{31,65}, given H3K4me2/3 repell DNMT3 activity^{76,77}. However, emerging work indicates
372 many histone-modifying enzymes have non-catalytic functions that influence gene expression, sometimes
373 even more potently than their catalytic roles^{85,86}. Indeed, KDM5C's catalytic activity was recently found to be
374 dispensible for repressing *Dazl* in ESCs⁵⁶. In our study, *Dazl*'s promoter still gained CpGme in *Kdm5c*-KO
375 exEpiLCs, even with elevated H3K4me2. *Dazl* and a few other germline genes employ multiple repressive
376 mechanisms to facilitate CpGme, such as DNMT3A/B recruitment via E2F6 and MGA^{15,16,57,58}. Thus, while
377 some germline CGIs require KDM5C-mediated H3K4me removal to overcome promoter CGI escape from
378 CpGme^{84,87}, others do not. These results also suggest the requirement for KDM5C's catalytic activity can
379 change depending upon the locus and developmental stage. Further experiments are required to determine
380 if catalytically inactive KDM5C can suppress germline genes at later developmental stages.

381 By generating a comprehensive list of mouse germline-enriched genes, we revealed distinct derepressive
382 mechanisms governing early versus late-stage germline programs. Previous work on germline gene silencing
383 has focused on genes with promoter CGIs^{17,84}, and indeed the majority of KDM5C targets in EpiLCs were
384 germ cell identity genes harboring CGIs. However, over 70% of germline-enriched gene promoters lacked
385 CGIs, including the many KDM5C-unbound germline genes that are de-repressed in *Kdm5c*-KO cells. CGI-
386 free, KDM5C-unbound germline genes were primarily late-stage spermatogenesis genes and significantly
387 enriched for RFX2 binding sites, a central regulator of spermiogenesis^{74,75}. These data suggest that once
388 activated during early embryogenesis, drivers of germline gene expression like *Rfx2*, *Stra8*, and *Dazl* turn
389 on downstream germline programs, ultimately culminating in the expression of spermiogenesis genes in
390 the adult *Kdm5c*-KO brain. Therefore, we propose KDM5C is recruited via promoter CGIs to act as a brake

391 against runaway activation of germline-specific programs. Future studies should address how KDM5C is
392 targeted to CGIs.

393 The above work provides the mechanistic foundation for KDM5C-mediated repression of tissue and
394 germline-specific genes. However, the contribution of these ectopic, tissue-specific genes towards neurolog-
395 ical impairments is still unknown. In addition to germline genes, we also identified significant enrichment
396 of muscle and liver-enriched transcripts within the *Kdm5c*-KO brain. Intriguingly, select liver and muscle-
397 enriched DEGs do have known roles within the brain, such as the liver-enriched lipid metabolism gene
398 *Apolipoprotein C-I (Apoc1)*³⁹. *APOC1* dysregulation is implicated in Alzheimer's disease in humans⁸⁸ and
399 overexpression of *Apoc1* in the mouse brain can impair learning and memory⁸⁹. KDM5C may therefore be
400 crucial for neurodevelopment by fine-tuning the expression of tissue-enriched, dosage-sensitive genes like
401 *Apoc1*.

402 Given that germline genes have no known functions within the brain, their impact upon neurodevelopment
403 is currently unknown. In *C. elegans*, somatic misexpression of germline genes via loss of *Retinoblastoma*
404 (*Rb*) homologs results in enhanced piRNA signaling and ectopic P granule formation in neurons^{90,91}. Ectopic
405 testicular germline transcripts have also been observed in a variety of cancers, including brain tumors in
406 *Drosophila* and mammals and shown to promote cancer progression^{20,21,92-94}. Intriguingly, mouse models
407 and human cells for other chromatin-linked NDDs also display impaired soma-germline demarcation^{22,95,96},
408 such as mutations in DNA methyltransferase 3b (DNMT3B), H3K9me1/2 methyltransferases G9A/GLP,
409 and methyl-CpG -binding protein 2 (MECP2). Recently, the transcription factor ZMYM2 (ZNF198), whose
410 mutation causes a NDD (OMIM #619522), was also shown to repress germline genes by promoting H3K4me
411 removal and CpGme⁹⁷. Thus, KDM5C is among a growing cohort of neurodevelopmental disorders with
412 erosion of the germline-soma boundary. Further research is required to determine the impact of these
413 germline genes upon neuronal functions and the extent to which this phenomenon occurs in humans.

414 Materials and Methods

415 Classifying tissue-enriched and germline-enriched genes

416 Tissue-enriched differentially expressed genes (DEGs) were determined by their classification in a previ-
417 ously published dataset from 17 male and female mouse tissues³⁴. This study defined tissue expression as
418 greater than 1 Fragments Per Kilobase of transcript per Million mapped read (FPKM) and tissue enrichment
419 as at least 4-fold higher expression than any other tissue.

420 We curated a list of germline-enriched genes using an RNA-seq dataset from wild-type and germline-
421 depleted (*Kit^{W/W^v}*) male and female mouse embryos from embryonic day 12, 14, and 16⁴³, as well as adult
422 male testes⁴⁰. Germline-enriched genes met the following criteria: 1) their expression is greater than 1
423 FPKM in wild-type germline 2) their expression in any wild-type somatic tissues³⁴ does not exceed 20%

424 of maximum expression in wild-type germline, and 3) their expression in the germ cell-depleted (*Kit*^{W/W^v})
425 germline, for any sex or time point, does not exceed 20% of maximum expression in wild-type germline. We
426 defined sperm and egg-biased genes as those whose expression in the opposite sex, at any time point, is no
427 greater than 20% of the gene's maximum expression in a given sex. Genes that did not meet this threshold
428 for either sex were classified as 'unbiased'.

429 **Cell culture**

430 We utilized our previously established cultures of male wild-type and *Kdm5c* knockout (-KO)
431 embryonic stem cells⁵². Sex was confirmed by genotyping *Uba1/Uba1y* on the X and Y chromo-
432 somes with the following primers: 5'-TGGATGGTGTGCCAATG-3', 5'-CACCTGCACGTTGCCCTT-
433 3'. Deletion of *Kdm5c* exons 11 and 12, which destabilize KDM5C protein³⁰, was confirmed
434 through the primers 5'-ATGCCCATATTAAGAGTCCCTG-3', 5'-TCTGCCTTGATGGACTGTT-3', and
435 5'-GGTTCTAACACTCACATAGTG-3'.

436 Embryonic stem cells (ESCs) and epiblast-like cells were cultured using previously established
437 methods⁴⁸. Briefly, ESCs were initially cultured in primed ESC (pESC) media consisting of KnockOut
438 DMEM (Gibco#10829–018), fetal bovine serum (Gibco#A5209501), KnockOut serum replacement
439 (Invitrogen#10828–028), Glutamax (Gibco#35050-061), Anti-Anti (Gibco#15240-062), MEM Non-essential
440 amino acids (Gibco#11140-050), and beta-mercaptoethanol (Sigma#M7522). They were then transitioned
441 into ground-state, "naïve" ESCs (nESCs) by culturing for four passages in N2B27 media containing
442 DMEM/F12 (Gibco#11330–032), Neurobasal media (Gibco#21103–049), Gluamax (Gibco#35050-061),
443 Anti-Anti (Gibco#15240-062), N2 supplement (Invitrogen#17502048), and B27 supplement without vitamin
444 A (Invitrogen#12587-010), and beta-mercaptoethanol (Sigma#M7522). Both pESC and nESC media
445 were supplemented with 3 μM GSK3 inhibitor CHIR99021 (Sigma #SML1046-5MG), 1 μM MEK inhibitor
446 PD0325901 (Sigma #PZ0162-5MG), and 1,000 units/mL leukemia inhibitory factor (LIF, Millipore#ESG1107).
447 nESCs were differentiated into epiblast-like cells (EpiLCs, 48 hours) and extendend EpiLCs (exEpiLCs,
448 96 hours) by culturing in N2B27 media containing DMEM/F12, Neurobasal media, Gluamax, Anti-Anti, N2
449 supplement, B27 supplement (Invitrogen#17504044), and beta-mercaptoethanol supplemented with 10
450 ng/mL fibroblast growth factor 2 (FGF2, R&D Biotechne 233-FB), and 20 ng/mL activin A (R&D Biotechne
451 338AC050CF), as previously described⁴⁸.

452 **Real time quantitative PCR (RT-qPCR)**

453 nESCs were differentiated into EpiLCs as described above. Cells were lysed with Tri-reagent BD (Sigma
454 #T3809) at 0, 24, and 48 hours of differentiation. RNA was phase separated with 0.1 μL/μL 1-bromo-3-
455 chloropropane (Sigma #B9673) and then precipitated with isopropanol (Sigma #I9516) and ethanol puri-
456 fied. For each sample, 2 μg of RNA was reverse transcribed using the ProtoScript II Reverse transcriptase kit

457 from New England Biolabs (NEB #M0368S) and primed with oligo dT. Expression of *Kdm5c* was detected us-
458 ing the primers 5'-CCCATGGAGGCCAGAGAATAAG-3' 5'-CTCAGCGGATAAGAGAATTGCTAC-3' and nor-
459 malized to TBP using the primers 5'-TTCAGAGGATGCTCTAGGGAAAGA-3' 5'-CTGTGGAGTAAGTCCTGTGCC-
460 3' with the Power SYBR™ Green PCR Master Mix (ThermoFisher #4367659).

461 **Western Blot**

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

467 **RNA sequencing (RNA-seq) data analysis**

468 After ensuring read quality via FastQC (v0.11.8), reads were then mapped to the mm10 *Mus musculus*
469 genome (Gencode) using STAR (v2.5.3a), during which we removed duplicates and kept only uniquely
470 mapped reads. Count files were generated by FeatureCounts (Subread v1.5.0), and BAM files were
471 converted to bigwigs using deeptools (v3.1.3) and visualized by the UCSC genome browser⁷⁹. RStudio
472 (v3.6.0) was then used to analyze counts files by DESeq2 (v1.26.0)³⁵ to identify differentially expressed
473 genes (DEGs) with a q-value (p-adjusted via FDR/Benjamini–Hochberg correction) less than 0.1 and a log2
474 fold change greater than 0.5. For all DESeq2 analyses, log2 fold changes were calculated with IfcShrink
475 using the ashR package⁹⁸. MA-plots were generated by ggpibr (v0.6.0), and Eulerr diagrams were generated
476 by eulerr (v6.1.1). Boxplots and scatterplots were generated by ggpibr (v0.6.0) and ggplot2 (v3.3.2). The
477 Upset plot was generated via the package UpSetR (v1.4.0)⁹⁹. Gene ontology (GO) analyses were performed
478 by the R package enrichPlot (v1.16.2) using the biological processes setting and compareCluster.

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

480 ChIP-seq reads were aligned to mm10 using Bowtie1 (v1.1.2) allowing up to two mismatches. Only
481 uniquely mapped reads were used for analysis. Peaks were called using MACS2 software (v2.2.9.1)
482 using input BAM files for normalization, with filters for a q-value < 0.1 and a fold enrichment > 1. We
483 removed “black-listed” genomic regions that often give aberrant signals. Common peak sets were obtained
484 in R via DiffBind¹⁰⁰ (v3.6.5). In the case of KDM5C ChIP-seq, *Kdm5c*-KO false-positive peaks were then
485 removed from wild-type samples using bedtools (v2.25.0). Peak proximity to genomic loci was determined
486 by ChIPSeeker¹⁰¹ (v1.32.1). Gene ontology (GO) analyses were performed by the R package enrichPlot
487 (v1.16.2) using the biological processes setting and compareCluster. Enriched motifs were identified using
488 HOMER⁷⁰ to search for known motifs within 500 base pairs up and downstream of the transcription start site.

489 Average binding across genes was visualized using deeptools (v3.1.3). Bigwigs were visualized using the
490 UCSC genome browser⁷⁹.

491 **CpG island (CGI) analysis**

492 Locations of CpG islands were determined through the mm10 UCSC genome browser CpG island track⁷⁹,
493 which classified CGIs as regions that have greater than 50% GC content, are larger than 200 base pairs,
494 and have a ratio of CG dinucleotides observed over the expected amount greater than 0.6. CGI genomic
495 coordinates were then annotated using ChIPseeker¹⁰¹ (v1.32.1) and filtered for ones that lie within promoters
496 of germline-enriched genes (TSS ± 500).

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

498 Genomic DNA (gDNA) from male naïve ESCs and extended EpiLCs was extracted using the Wizard
499 Genomic DNA Purification Kit (Promega A1120), following the instructions for Tissue Culture Cells. gDNA
500 from two wild-types and two *Kdm5c*-KOs of each cell type was sent to Novogene for WGBS using the
501 Illumina NovaSeq X Plus platform and sequenced for 150 bp paired-end reads (PE150). All samples had
502 greater than 99% bisulfite conversion rates. Reads were adapter and quality trimmed with Trim Galore
503 (v0.6.10) and aligned to the mm10 genome using Bismark¹⁰² (v0.22.1). Analysis of differential methylation at
504 gene promoters was performed using methylKit⁸⁰ (v1.28.0) with a minimum coverage of 3 paired reads, a
505 percentage greater than 25% or less than -25%, and q-value less than 0.01. methylKit was also used to
506 calculate average percentage methylation at germline gene promoters. Methylation bedgraph tracks were
507 generated via Bismark and visualized using the UCSC genome browser⁷⁹.

508 **Data availability**

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

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

512 **Published datasets**

513 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
514 adult amygdala and hippocampus³³, available at GEO: GSE127722. Male and female wild-type, *Kdm5c*-KO,
515 and *Kdm5c*-HET EpiLCs⁵² are available at GEO: GSE96797.

517 Previously published ChIP-seq experiments included KDM5C binding in wild-type and *Kdm5c*-KO
518 EpiLCs⁵² (available at GEO: GSE96797) and mouse primary neuron cultures (PNCs) from the cortex

519 and hippocampus³⁰ (available at GEO: GSE61036). ChIP-seq of histone 3 lysine 4 dimethylation (H3K4me2)
520 in male wild-type and *Kdm5c*-KO EpiLCs⁵² is also available at GEO: GSE96797. ChIP-seq of histone 3 lysine
521 4 trimethylation (H3K4me3) in wild-type and *Kdm5c*-KO male amygdala³³ are available at GEO: GSE127817.

522 **Data analysis**

523 Scripts used to generate the results, tables, and figures of this study are available via the GitHub
524 repository: https://github.com/kbonefas/KDM5C_Germ_Mechanism

525 **Acknowledgements**

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

540 **Author Contributions**

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

544 **Declaration of Interest**

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

547 **References**

- 548 1. Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* **403**,
549 41–45. <https://doi.org/10.1038/47412>.
- 550 2. Jenuwein, T., and Allis, C.D. (2001). Translating the Histone Code. *Science* **293**, 1074–1080.
551 <https://doi.org/10.1126/science.1063127>.
- 552 3. Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570.
553 <https://doi.org/10.1038/276565a0>.
- 554 4. Kennison, J.A., and Tamkun, J.W. (1988). Dosage-dependent modifiers of polycomb and antennapedia
mutations in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8136–8140. <https://doi.org/10.1073/pnas.8>
555 5.21.8136.
- 556 5. Kassis, J.A., Kennison, J.A., and Tamkun, J.W. (2017). Polycomb and Trithorax Group Genes in
557 *Drosophila*. *Genetics* **206**, 1699–1725. <https://doi.org/10.1534/genetics.115.185116>.
- 558 6. Zhao, S., Allis, C.D., and Wang, G.G. (2021). The language of chromatin modification in human
559 cancers. *Nat Rev Cancer* **21**, 413–430. <https://doi.org/10.1038/s41568-021-00357-x>.
- 560 7. Chi, P., Allis, C.D., and Wang, G.G. (2010). Covalent histone modifications—miswritten, misinterpreted
561 and mis-erased in human cancers. *Nat Rev Cancer* **10**, 457–469. <https://doi.org/10.1038/nrc2876>.
- 562 8. Berdasco, M., and Esteller, M. (2010). Aberrant Epigenetic Landscape in Cancer: How Cellular
563 Identity Goes Awry. *Developmental Cell* **19**, 698–711. <https://doi.org/10.1016/j.devcel.2010.10.005>.
- 564 9. Gabriele, M., Lopez Tobon, A., D'Agostino, G., and Testa, G. (2018). The chromatin basis of
neurodevelopmental disorders: Rethinking dysfunction along the molecular and temporal axes. *Prog
565 Neuropsychopharmacol Biol Psychiatry* **84**, 306–327. <https://doi.org/10.1016/j.pnpbp.2017.12.013>.
- 566 10. Feichtinger, J., Aldeailej, I., Anderson, R., Almutairi, M., Almatrafi, A., Alsiwiehri, N., Griffiths, K.,
Stuart, N., Wakeman, J.A., Larcombe, L., et al. (2012). Meta-analysis of clinical data using human
meiotic genes identifies a novel cohort of highly restricted cancer-specific marker genes. *Oncotarget*
567 **3**, 843–853. <https://doi.org/10.18632/oncotarget.580>.
- 568 11. Schaefer, A., Sampath, S.C., Intrator, A., Min, A., Gertler, T.S., Surmeier, D.J., Tarakhovsky, A.,
and Greengard, P. (2009). Control of cognition and adaptive behavior by the GLP/G9a epigenetic
569 suppressor complex. *Neuron* **64**, 678–691. <https://doi.org/10.1016/j.neuron.2009.11.019>.
- 570 12. Michod, R.E. (2011). Evolutionary Transitions in Individuality: Multicellularity and Sex. In *The Major
Transitions in Evolution Revisited*, B. Calcott and K. Sterelny, eds. (The MIT Press), pp. 169–198.
571 <https://doi.org/10.7551/mitpress/8775.003.0015>.
- 572 13. Chen, L., and Wiens, J.J. (2021). Multicellularity and sex helped shape the Tree of Life. *Proc. R. Soc.
573 B* **288**, 20211265. <https://doi.org/10.1098/rspb.2021.1265>.

- 574 14. 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>.
- 575
- 576 15. 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>.
- 577
- 578 16. 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>.
- 579
- 580 17. 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>.
- 581
- 582 18. 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>.
- 583
- 584 19. 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>.
- 585
- 586 20. Nielsen, A.Y., and Gjerstorff, M.F. (2016). Ectopic Expression of Testis Germ Cell Proteins in Cancer
and Its Potential Role in Genomic Instability. *Int J Mol Sci* 17. <https://doi.org/10.3390/ijms17060890>.
- 587
- 588 21. Adebayo Babatunde, K., Najafi, A., Salehipour, P., Modarressi, M.H., and Mobasheri, M.B. (2017).
Cancer/Testis genes in relation to sperm biology and function. *Iranian Journal of Basic Medical
Sciences* 20. <https://doi.org/10.22038/ijbms.2017.9259>.
- 589
- 590 22. Bonefas, K.M., and Iwase, S. (2021). Soma-to-germline transformation in chromatin-linked neurode-
velopmental disorders? *FEBS J.* <https://doi.org/10.1111/febs.16196>.
- 591
- 592 23. Iwase, S., Lan, F., Bayliss, P., De La Torre-Ubieta, L., Huarte, M., Qi, H.H., Whetstine, J.R., Bonni, A.,
Roberts, T.M., and Shi, Y. (2007). The X-Linked Mental Retardation Gene SMCX/JARID1C Defines a
Family of Histone H3 Lysine 4 Demethylases. *Cell* 128, 1077–1088. <https://doi.org/10.1016/j.cell.2007.02.017>.
- 593
- 594 24. Shen, H., Xu, W., Guo, R., Rong, B., Gu, L., Wang, Z., He, C., Zheng, L., Hu, X., Hu, Z., et al.
(2016). Suppression of Enhancer Overactivation by a RACK7-Histone Demethylase Complex. *Cell*
165, 331–342. <https://doi.org/10.1016/j.cell.2016.02.064>.

- 595
- 596 25. Chen, X., Loo, J.X., Shi, X., Xiong, W., Guo, Y., Ke, H., Yang, M., Jiang, Y., Xia, S., Zhao, M., et al. (2018). E6 Protein Expressed by High-Risk HPV Activates Super-Enhancers of the *EGFR* and *c-MET* Oncogenes by Destabilizing the Histone Demethylase KDM5C. *Cancer Research* 78, 1418–1430. <https://doi.org/10.1158/0008-5472.CAN-17-2118>.
- 597
- 598 26. Zheng, Q., Li, P., Zhou, X., Qiang, Y., Fan, J., Lin, Y., Chen, Y., Guo, J., Wang, F., Xue, H., et al. (2021). Deficiency of the X-inactivation escaping gene *KDM5C* in clear cell renal cell carcinoma promotes tumorigenicity by reprogramming glycogen metabolism and inhibiting ferroptosis. *Theranostics* 11, 599 8674–8691. <https://doi.org/10.7150/thno.60233>.
- 600 27. Claes, S., Devriendt, K., Van Goethem, G., Roelen, L., Meireleire, J., Raeymaekers, P., Cassiman, J.J., and Fryns, J.P. (2000). Novel syndromic form of X-linked complicated spastic paraparesis. *Am J Med Genet* 94, 1–4.
- 601
- 602 28. Jensen, L.R., Amende, M., Gurok, U., Moser, B., Gimmel, V., Tzschach, A., Janecke, A.R., Tariverdian, 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>.
- 603
- 604 29. Carmignac, V., Nambot, S., Lehalle, D., Callier, P., Moortgat, S., Benoit, V., Ghoumid, J., Delobel, 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>.
- 605
- 606 30. 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>.
- 607
- 608 31. 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, 609 47–59. <https://doi.org/10.1016/j.celrep.2017.09.014>.
- 610 32. 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, 611 637. <https://doi.org/10.3390/cells12040637>.
- 612 33. 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>.
- 613

- 614 34. Li, B., Qing, T., Zhu, J., Wen, Z., Yu, Y., Fukumura, R., Zheng, Y., Gondo, Y., and Shi, L. (2017). A
615 Comprehensive Mouse Transcriptomic BodyMap across 17 Tissues by RNA-seq. *Sci Rep* 7, 4200.
<https://doi.org/10.1038/s41598-017-04520-z>.
- 616 35. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for
617 RNA-seq data with DESeq2. *Genome Biol* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
- 618 36. Crackower, M.A., Kolas, N.K., Noguchi, J., Sarao, R., Kikuchi, K., Kaneko, H., Kobayashi, E., Kawai, Y.,
619 Kozieradzki, I., Landers, R., et al. (2003). Essential Role of Fkbp6 in Male Fertility and Homologous
Chromosome Pairing in Meiosis. *Science* 300, 1291–1295. <https://doi.org/10.1126/science.1083022>.
- 620 37. Xiol, J., Cora, E., Koglgruber, R., Chuma, S., Subramanian, S., Hosokawa, M., Reuter, M., Yang, Z.,
621 Berninger, P., Palencia, A., et al. (2012). A Role for Fkbp6 and the Chaperone Machinery in piRNA
Amplification and Transposon Silencing. *Molecular Cell* 47, 970–979. <https://doi.org/10.1016/j.molcel.2012.07.019>.
- 622 38. Cheng, S., Altmeppen, G., So, C., Welp, L.M., Penir, S., Ruhwedel, T., Menelaou, K., Harasimov, K.,
623 Stützer, A., Blayney, M., et al. (2022). Mammalian oocytes store mRNAs in a mitochondria-associated
membraneless compartment. *Science* 378, eabq4835. <https://doi.org/10.1126/science.abq4835>.
- 624 39. Rouland, A., Masson, D., Lagrost, L., Vergès, B., Gautier, T., and Bouillet, B. (2022). Role of
625 apolipoprotein C1 in lipoprotein metabolism, atherosclerosis and diabetes: A systematic review.
Cardiovasc Diabetol 21, 272. <https://doi.org/10.1186/s12933-022-01703-5>.
- 626 40. Mueller, J.L., Skaletsky, H., Brown, L.G., Zaghloul, S., Rock, S., Graves, T., Auger, K., Warren,
627 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>.
- 628 41. 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>.
- 630 42. Green, C.D., Ma, Q., Manske, G.L., Shami, A.N., Zheng, X., Marini, S., Moritz, L., Sultan, C.,
631 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>.
- 632 43. 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>.
- 633 44. Magnúsdóttir, E., and Surani, M.A. (2014). How to make a primordial germ cell. *Development* 141,
635 245–252. <https://doi.org/10.1242/dev.098269>.

- 636 45. Günesdogan, U., Magnúsdóttir, E., and Surani, M.A. (2014). Primordial germ cell specification: A
context-dependent cellular differentiation event [corrected]. *Philos Trans R Soc Lond B Biol Sci* **369**.
<https://doi.org/10.1098/rstb.2013.0543>.
- 637
- 638 46. Bardot, E.S., and Hadjantonakis, A.-K. (2020). Mouse gastrulation: Coordination of tissue patterning,
specification and diversification of cell fate. *Mechanisms of Development* **163**, 103617. <https://doi.org/10.1016/j.mod.2020.103617>.
- 639
- 640 47. Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the
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>.
- 641
- 642 48. Samanta, M., and Kalantry, S. (2020). Generating primed pluripotent epiblast stem cells: A methodol-
ogy chapter. *Curr Top Dev Biol* **138**, 139–174. <https://doi.org/10.1016/bs.ctdb.2020.01.005>.
- 643
- 644 49. Welling, M., Chen, H., Muñoz, J., Musheev, M.U., Kester, L., Junker, J.P., Mischerikow, N., Arbab, M.,
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>.
- 645
- 646 50. Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O.,
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>.
- 647
- 648 51. Suzuki, A., Hirasaki, M., Hishida, T., Wu, J., Okamura, D., Ueda, A., Nishimoto, M., Nakachi, Y.,
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>.
- 649
- 650 52. Samanta, M.K., Gayen, S., Harris, C., Maclary, E., Murata-Nakamura, Y., Malcore, R.M., Porter, R.S.,
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>.
- 651
- 652 53. 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**,
2474–2479. <https://doi.org/10.1073/pnas.0510813103>.
- 653
- 654 54. Lin, Y., Gill, M.E., Koubova, J., and Page, D.C. (2008). Germ Cell-Intrinsic and -Extrinsic Factors
Govern Meiotic Initiation in Mouse Embryos. *Science* **322**, 1685–1687. <https://doi.org/10.1126/science.1166340>.
- 655
- 656 55. Endo, T., Mikedis, M.M., Nicholls, P.K., Page, D.C., and De Rooij, D.G. (2019). Retinoic Acid and Germ
Cell Development in the Ovary and Testis. *Biomolecules* **9**, 775. <https://doi.org/10.3390/biom9120775>.
- 657

- 658 56. Gupta, N., Yakhou, L., Albert, J.R., Azogui, A., Ferry, L., Kirsh, O., Miura, F., Battault, S., Yamaguchi,
659 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>.
- 660 57. Pohlers, M., Truss, M., Frede, U., Scholz, A., Strehle, M., Kuban, R.-J., Hoffmann, B., Morkel, M.,
Birchmeier, C., and Hagemann, C. (2005). A Role for E2F6 in the Restriction of Male-Germ-Cell-Specific
661 Gene Expression. Current Biology 15, 1051–1057. <https://doi.org/10.1016/j.cub.2005.04.060>.
- 662 58. Dahlet, T., Truss, M., Frede, U., Al Adhami, H., Bardet, A.F., Dumas, M., Vallet, J., Chicher, J.,
Hamann, P., Kottnik, S., et al. (2021). E2F6 initiates stable epigenetic silencing of germline genes
663 during embryonic development. Nat Commun 12, 3582. <https://doi.org/10.1038/s41467-021-23596-w>.
- 664 59. 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
665 and human. Hum Mol Genet 3, 879–884. <https://doi.org/10.1093/hmg/3.6.879>.
- 666 60. 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>.
- 667 61. 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
669 5, 1355–1360. <https://doi.org/10.1093/hmg/5.9.1355>.
- 670 62. 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>.
- 671 63. 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>.
- 672 64. 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
675 inactivation. Nat Genet 36, 642–646. <https://doi.org/10.1038/ng1368>.
- 676 65. Al Adhami, H., Vallet, J., Schaal, C., Schumacher, P., Bardet, A.F., Dumas, M., Chicher, J., Hamann,
P., Daujat, S., and Weber, M. (2023). Systematic identification of factors involved in the silencing
of germline genes in mouse embryonic stem cells. Nucleic Acids Research 51, 3130–3149. <https://doi.org/10.1093/nar/gkad071>.
- 677 66. Hurlin, P.J. (1999). Mga, a dual-specificity transcription factor that interacts with Max and contains a
T-domain DNA-binding motif. The EMBO Journal 18, 7019–7028. <https://doi.org/10.1093/emboj/18.2.7019>.

- 680 67. Stielow, B., Finkernagel, F., Stiewe, T., Nist, A., and Suske, G. (2018). MGA, L3MBTL2 and E2F6
determine genomic binding of the non-canonical Polycomb repressive complex PRC1.6. *PLoS Genet*
681 14, e1007193. <https://doi.org/10.1371/journal.pgen.1007193>.
- 682 68. 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.
683 <https://doi.org/10.1371/journal.pone.0205969>.
- 684 69. 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.
685 *Nature* 447, 601–605. <https://doi.org/10.1038/nature05823>.
- 686 70. 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.
687 <https://doi.org/10.1016/j.molcel.2010.05.004>.
- 688 71. 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,
689 916–921. <https://doi.org/10.1038/35002634>.
- 690 72. 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).
- 692 73. 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
693 Ciliogenesis. *Sci. Signal.* 2. <https://doi.org/10.1126/scisignal.2000602>.
- 694 74. 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.
695 *PLoS Genet* 11, e1005368. <https://doi.org/10.1371/journal.pgen.1005368>.
- 696 75. 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.
697 <https://doi.org/10.1038/srep20435>.
- 698 76. 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
699 domain. *EMBO Reports* 10, 1235–1241. <https://doi.org/10.1038/embor.2009.218>.

- 700 77. Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., et al. (2015).
701 Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* **517**,
640–644. <https://doi.org/10.1038/nature13899>.
- 702 78. Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein,
703 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>.
- 704 79. Nassar, L.R., Barber, G.P., Benet-Pagès, A., Casper, J., Clawson, H., Diekhans, M., Fischer, C.,
705 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>.
- 706 80. Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueiroa, M.E., Melnick, A., and Mason,
707 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>.
- 708 81. Torres-Padilla, M.-E. (2020). On transposons and totipotency. *Philos Trans R Soc Lond B Biol Sci*
709 **375**, 20190339. <https://doi.org/10.1098/rstb.2019.0339>.
- 710 82. Yang, M., Yu, H., Yu, X., Liang, S., Hu, Y., Luo, Y., Izsvák, Z., Sun, C., and Wang, J. (2022). Chemical-
711 induced chromatin remodeling reprograms mouse ESCs to totipotent-like stem cells. *Cell Stem Cell*
29, 400–418.e13. <https://doi.org/10.1016/j.stem.2022.01.010>.
- 712 83. Zhang, J., Zhang, M., Acampora, D., Vojtek, M., Yuan, D., Simeone, A., and Chambers, I. (2018).
OTX2 restricts entry to the mouse germline. *Nature* **562**, 595–599. [018-0581-5](https://doi.org/10.1038/s41586-
713 018-0581-5).
- 714 84. Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007).
715 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>.
- 716 85. Aubert, Y., Egolf, S., and Capell, B.C. (2019). The Unexpected Noncatalytic Roles of Histone Modifiers
717 in Development and Disease. *Trends in Genetics* **35**, 645–657. [004](https://doi.org/10.1016/j.tig.2019.06.
004).
- 718 86. Morgan, M.A.J., and Shilatifard, A. (2020). Reevaluating the roles of histone-modifying enzymes
719 and their associated chromatin modifications in transcriptional regulation. *Nat Genet* **52**, 1271–1281.
<https://doi.org/10.1038/s41588-020-00736-4>.
- 720 87. 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**,
6693–6706. <https://doi.org/10.1093/nar/gkw258>.
- 722 88. 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](https://doi.org/10.1016/j.molmed.2
010.07.008).

- 723
- 724 89. 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>.
- 725
- 726 90. 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 retinoblastoma pathway mutants. *Nature* *436*, 593–597. <https://doi.org/10.1038/nature04010>.
- 727
- 728 91. 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>.
- 729
- 730 92. 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. <https://doi.org/10.1126/science.1195481>.
- 731
- 732 93. Ghafouri-Fard, S., and Modarressi, M.-H. (2012). Expression of Cancer–Testis Genes in Brain Tumors: Implications for Cancer Immunotherapy. *Immunotherapy* *4*, 59–75. <https://doi.org/10.2217/imt.11.145>.
- 733
- 734 94. Nin, D.S., and Deng, L.-W. (2023). Biology of Cancer–Testis Antigens and Their Therapeutic Implications in Cancer. *Cells* *12*, 926. <https://doi.org/10.3390/cells12060926>.
- 735
- 736 95. 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. *Orphanet J Rare Dis* *9*, 56. <https://doi.org/10.1186/1750-1172-9-56>.
- 737
- 738 96. Walton, E.L., Francastel, C., and Velasco, G. (2014). Dnmt3b Prefers Germ Line Genes and Centromeric Regions: Lessons from the ICF Syndrome and Cancer and Implications for Diseases. *Biology* (Basel) *3*, 578–605. <https://doi.org/10.3390/biology3030578>.
- 739
- 740 97. 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. <https://doi.org/10.1093/nar/gkad540>.
- 741
- 742 98. Stephens, M. (2016). False discovery rates: A new deal. *Biostat*, kxw041. <https://doi.org/10.1093/biostatistics/kxw041>.
- 743
- 744 99. 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>.
- 745

- 746 100. Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D.,
747 Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated
with clinical outcome in breast cancer. *Nature* *481*, 389–393. <https://doi.org/10.1038/nature10730>.
- 748 101. Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: An R/Bioconductor package for ChIP peak
749 annotation, comparison and visualization. *Bioinformatics* *31*, 2382–2383. <https://doi.org/10.1093/bioinformatics/btv145>.
- 750 102. Krueger, F., and Andrews, S.R. (2011). Bismark: A flexible aligner and methylation caller for Bisulfite-
751 Seq applications. *Bioinformatics* *27*, 1571–1572. <https://doi.org/10.1093/bioinformatics/btr167>.

752 **Figures and Tables**

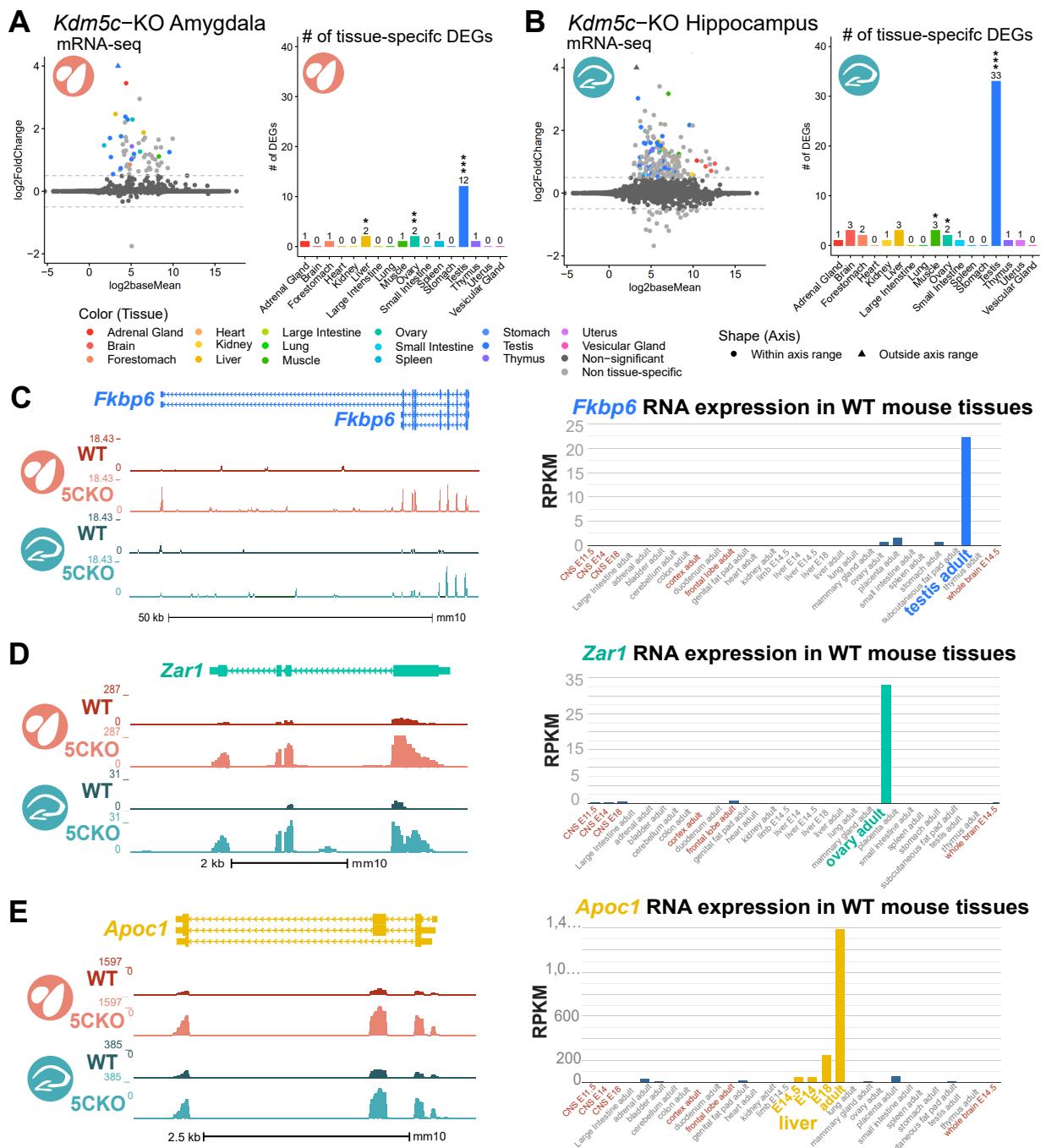


Figure 1: Tissue-enriched genes are misexpressed in the *Kdm5c*-KO brain. **A-B.** Expression of tissue-enriched genes (Li et al 2017) in the male *Kdm5c*-KO amygdala (A) and hippocampus (B). Left - MA plot of mRNA-sequencing. Right - Number of tissue-enriched differentially expressed genes (DEGs). * 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.

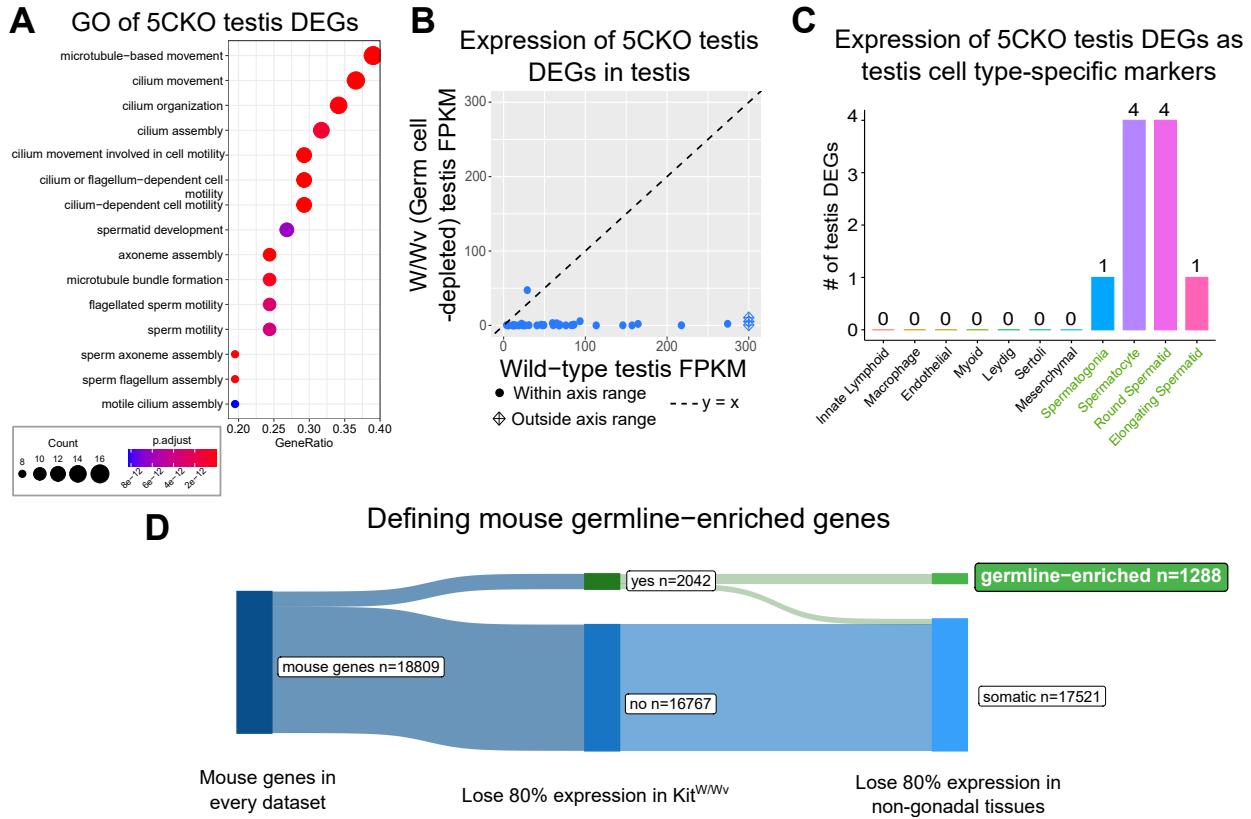


Figure 2: Aberrant transcription of germline genes in the *Kdm5c*-KO in the brain. **A.** enrichPlot gene ontology (GO) of *Kdm5c*-KO amygdala and hippocampus testis-enriched DEGs **B.** Expression of testis DEGs in wild-type (WT) testis versus germ cell-depleted (W/Wv) testis (Mueller et al 2013). Expression 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).

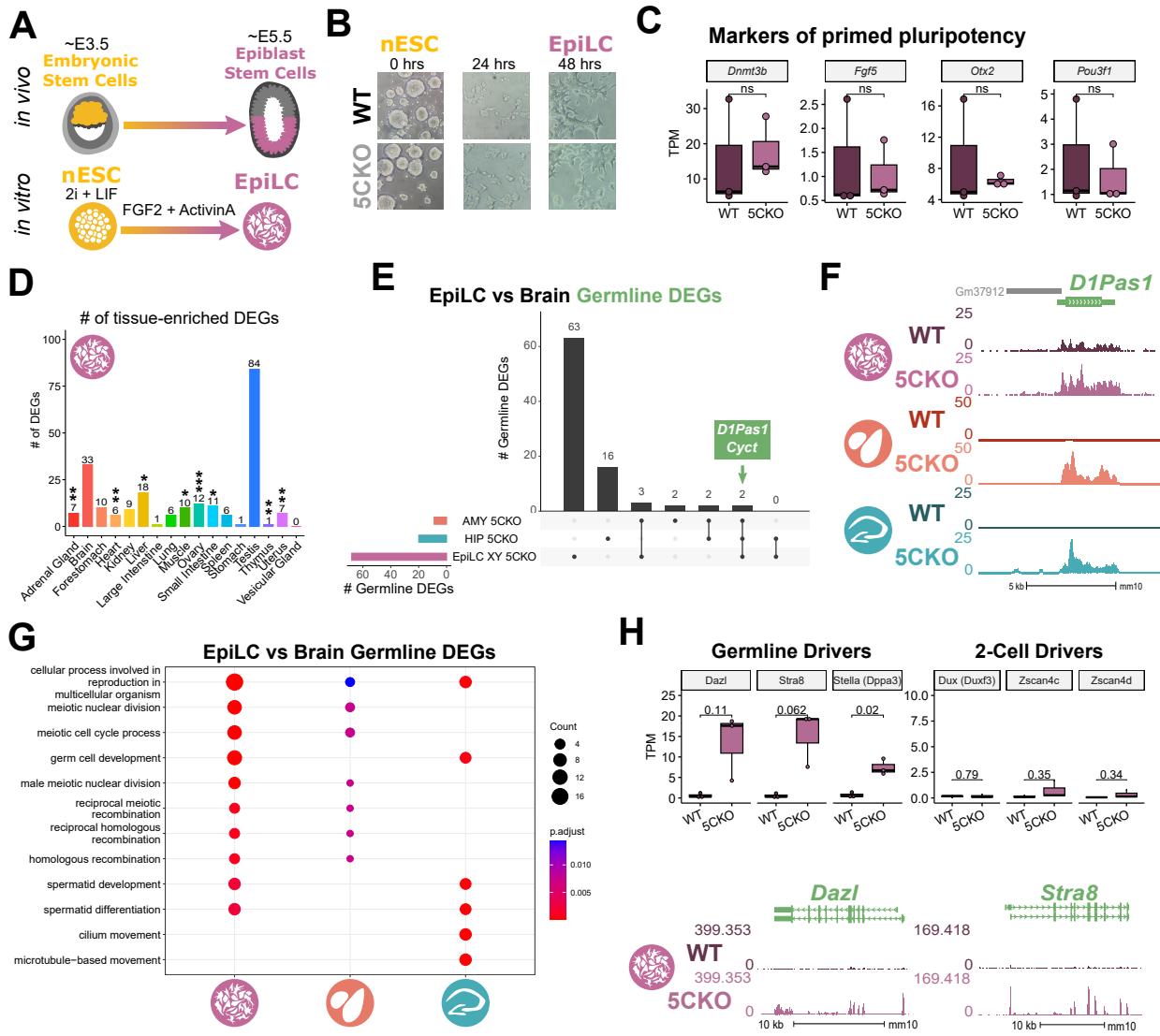


Figure 3: Kdm5c-KO epiblast-like cells express key drivers of germline identity **A.** Top - Diagram of *in vivo* differentiation of embryonic stem cells (ESCs) of the inner cell mass into epiblast stem cells. Bottom - *in vitro* differentiation of ESCs into epiblast-like cells (EpiLCs). **B.** Representative 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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, 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, $n = 3$).

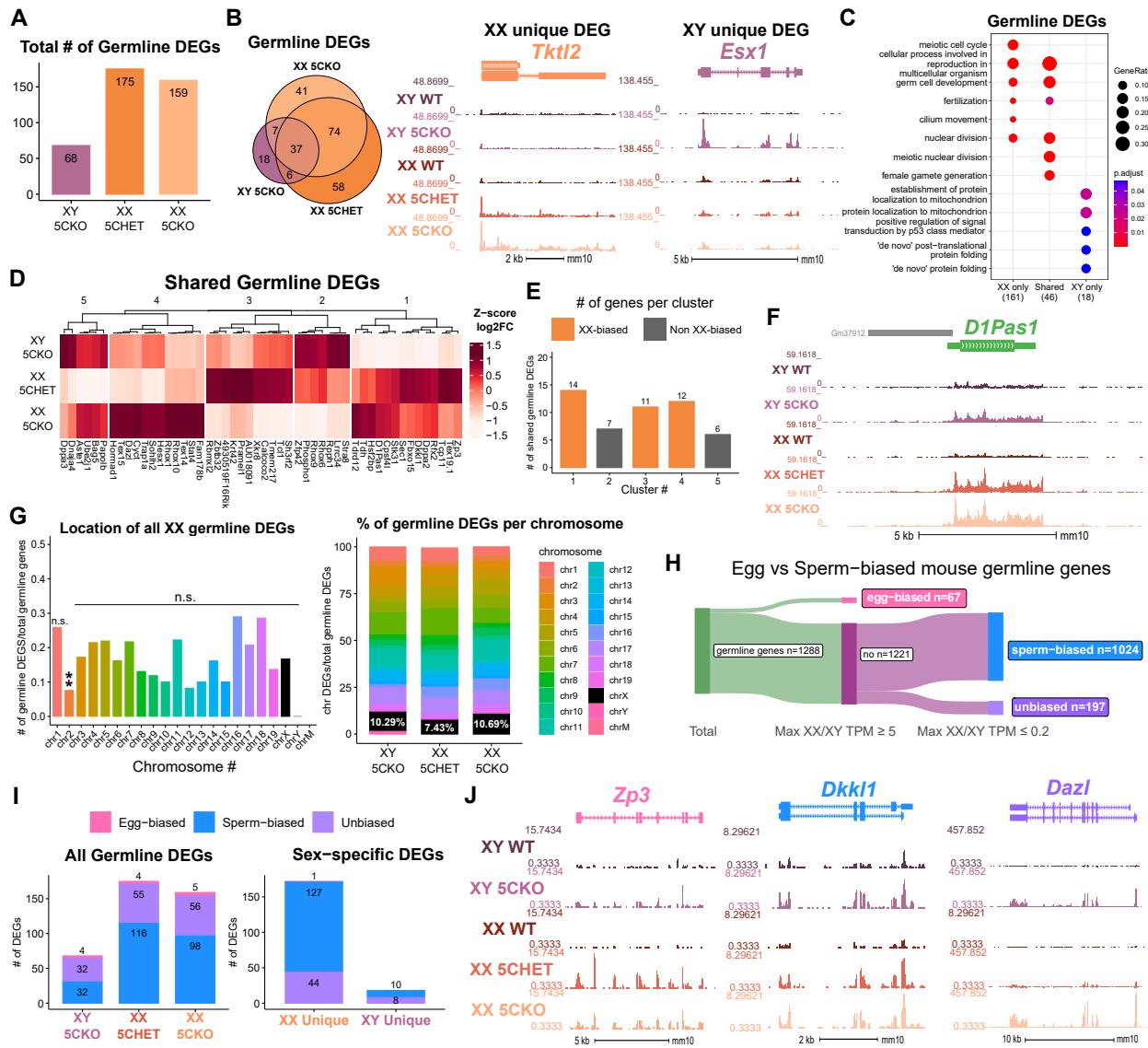


Figure 4: Chromosomal sex influences *Kdm5c*-KO germline gene misexpression. **A.** Total number of germline-enriched RNA-seq DEGs for male hemizygous *Kdm5c* knockout EpiLCs (XY 5CKO, purple), female heterozygous *Kdm5c* knockout (XX 5CHET, orange), female homozygous *Kdm5c* knockout (XX 5CKO, light 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).

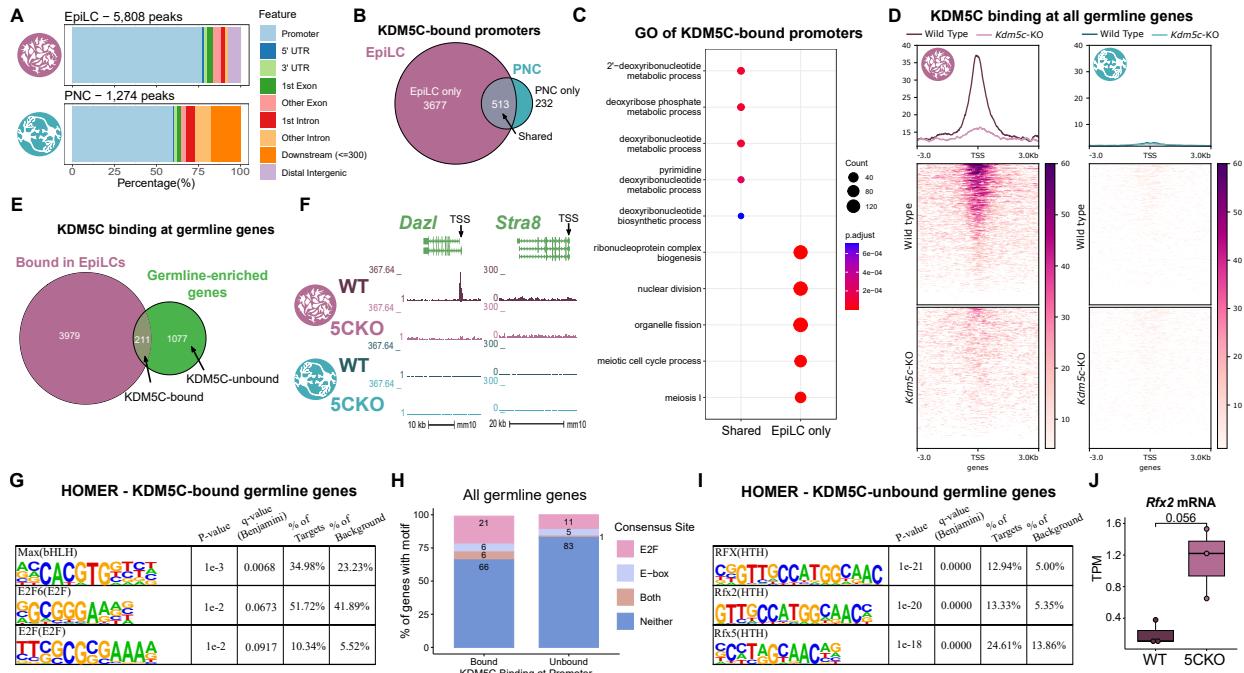


Figure 5: KDM5C binds to a subset of germline gene promoters during early embryogenesis. **A.** ChIPseeker localization of KDM5C peaks at different genomic regions in EpiLCs (top) and hippocampal and cortex primary neuron cultures (PNCs, bottom). **B.** Overlap of genes with KDM5C bound to their promoters ($TSS \pm 500$) 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).

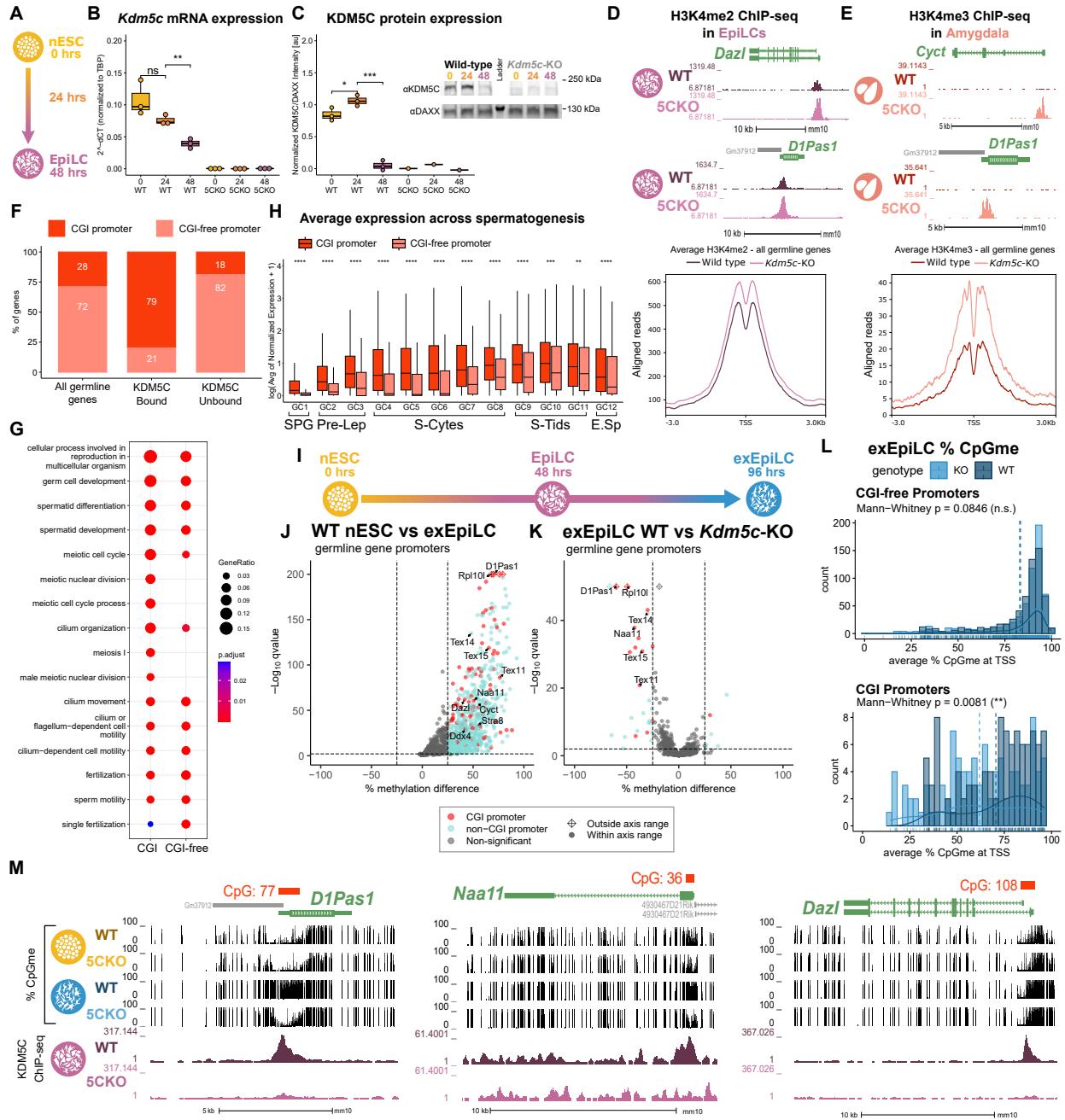


Figure 6: KDM5C promotes long-term silencing of germline genes via DNA methylation at CpG islands. **A.** Diagram of embryonic stem cell (ESC) to epiblast-like cell (EpiLC) differentiation and collection time points. **B.** Real time quantitative PCR (RT-qPCR) of *Kdm5c* mRNA expression in wild-type (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.