

1 Erosion of somatic tissue identity with loss of the X-linked intellectual  
2 disability factor KDM5C

3

4 **Abstract**

5 **Introduction**

6 Embryonic stem cells employ a single genome to form the myriad of discrete cellular identities that make up a complete  
7 organism. This is, in part, accomplished by chromatin regulators that can either promote or impede lineage-specific  
8 gene expression through histone and DNA modifications<sup>1,2</sup>. Although many chromatin regulators were initially identified  
9 for their roles in cellular identity<sup>3,4</sup>, recent advancements in next generation sequencing unexpectedly revealed many  
10 neurodevelopmental disorders (NDDs) are caused by or linked to mutations in chromatin regulators[CIT]. The connection  
11 between neurodevelopment and chromatin regulation can be partially explained by their regulation of brain-specific genes  
12 or chromatin states, such as orchestrating synaptic maturation<sup>5</sup> or the transition between neuron and glia during cortical  
13 development<sup>6</sup>. However, loss of some chromatin regulators can also result in aberrant transcription of tissue-specific genes  
14 outside of their intended environment<sup>3,4,7</sup>. Very few studies have explored the ectopic expression of tissue-specific genes  
15 in chromatin-linked neurodevelopmental disorders<sup>8,9</sup> and it is unclear if this partial loss of brain identity contributes to  
16 neurodevelopmental impairments.

17 To elucidate if dysregulation of brain tissue identity contributes to neurodevelopmental impairments, it is essential to  
18 first characterize the types of genes misexpressed and the molecular mechanisms underlying their de-repression. In this  
19 study, we characterized the dysregulation of tissue-enriched genes with loss of the chromatin regulator lysine demethylase  
20 5C (KDM5C, also known as SMCX or JARID1C), a histone 3 lysine 4 demethylase. Pathogenic mutations in *KDM5C*  
21 cause Intellectual Developmental Disorder, X-linked, Syndromic, Claes-Jensen Type (MRXSCJ, OMIM: 300534), whose  
22 features include short stature, intellectual disability, seizures, aberrant aggression, and autistic behaviors<sup>10-12</sup>. Previous  
23 work has demonstrated constitutive *Kdm5c* knockout (-KO) mice recapitulate key MRXSCJ patient phenotypes, including  
24 hyperaggression and learning impairments<sup>13</sup>. Next generation RNA sequencing (RNA-seq) in the *Kdm5c*-KO hippocampus  
25 unexpectedly revealed ectopic expression of testis-enriched genes within the brain<sup>9</sup>. However, it is currently unclear if  
26 misexpression in the *Kdm5c*-KO brain is unique to testis genes, as other tissue-enriched genes have not been characterized.  
27 Furthermore, some of the ectopic testis transcripts have known functions in germ cells<sup>9</sup>, potentially reflecting a failure to  
28 distinguish between somatic (non-meiotic) and germline (meiotic) identities with KDM5C loss. While there is emerging  
29 evidence that KDM5C regulates select germline genes, it is unclear which germline *Kdm5c*-KO. - KDM5C's molecular  
30 mechanisms unclear - direct vs indirect - Developmental time point of involvement - How the genes change over time.

- 31 • Systematically characterizing KDM5C's role in repressing germline identity will provide molecular footholds for testing  
32 the impact of ectopic germline gene expression on neurodevelopment. These can then be developed into novel

33 therapeutics for MRXSCJ treatment and other NDDs that share a similar etiology **note: don't want to be too specific**  
34 **about therapeutics**

35 To elucidate KDM5C's role in tissue identity, here we characterized the aberrant transcription of tissue-enriched genes  
36 within the *Kdm5c*-KO brain and epiblast-like cells (EpiLCs), an *in vitro* model of the post-implantation embryo. We identified  
37 general dysregulation of tissue-enriched genes in both the adult *Kdm5c*-KO brain and EpiLCs, including misexpression  
38 of liver, muscle, ovary, and testis genes. Both the *Kdm5c*-KO amygdala and hippocampus had significant enrichment of  
39 testis-biased genes that are typically only expressed in germ cells. While the *Kdm5c*-KO brain primarily expresses genes  
40 important for late germ cell development, *Kdm5c*-KO EpiLCs aberrantly express key drivers of germline identity and meiosis,  
41 such as *Dazl* and *Stra8*. We found that KDM5C only binds a subset of germline genes expressed in the *Kdm5c*-KO brain  
42 and EpiLCs, indicating germline-enriched mRNAs can be aberrantly expressed through indirect mechanisms. Finally, we  
43 found KDM5C loss impairs the placement of DNA methylation at germline gene promoters as ESCs differentiate into EpiLCs.  
44 Therefore, we propose KDM5C plays a crucial role in the development of tissue identity during early embryogenesis, including  
45 establishment of the soma-germline boundary.

- 46 • We then generated a dataset of germline-enriched genes to aid in the characterization of germline gene misexpression  
47 • **note: need a better conclusion sentence - work on when we know what's happening with last figure/functional**  
48 **consequences**  
49 • **note: I feel like I'm missing the key points still on the last paragraph**

## 50 Results

### 51 **Tissue-enriched genes, including testis genes, are aberrantly expressed in the *Kdm5c*-KO brain**

- 52 • **note: should I compare amygdala and hippocampus? scandaglia only looked at hippocampus**

53 Previous RNA sequencing (RNA-seq) in the adult hippocampus ectopic expression of some testis genes within the *Kdm5c*  
54 knockout (-KO) brain<sup>9</sup>. Since tissue-enriched genes have not been systematically characterized in the *Kdm5c*-KO brain,  
55 it is currently unclear if this erosion of brain tissue identity is a major consequence of *Kdm5c* loss and if it is unique to  
56 testis-enriched genes. Therefore, we first globally assessed the expression of previously characterized mouse tissue-enriched  
57 genes<sup>14</sup> in our published mRNA-seq datasets of the amygdala and hippocampus in adult mice with constitutive knockout of  
58 *Kdm5c*<sup>15</sup>.

59 We found a large proportion of genes significantly upregulated within the *Kdm5c*-KO brain (DESeq2<sup>16</sup>, log2 fold change >  
60 0.5, q < 0.1) are typically enriched in non-brain tissues (Amygdala: 35%, Hippocampus: 24%) (Figure 1A-B). The majority  
61 of tissue-enriched differentially expressed genes (tissue-enriched DEGs) were testis genes (Figure 1A-C). Testis-biased  
62 DEGs were significantly enriched for both brain regions (Amygdala p = 1.83e-05; Hippocampus p = 4.26e-11, Fisher's Exact  
63 Test), even though the testis has the largest total number of tissue-biased genes compared to any other tissue (2,496 genes).  
64 Of note, we did not observe enrichment of brain-enriched genes (Amygdala p = 1; Hippocampus p = 0.74, Fisher's Exact),  
65 despite the fact these are brain samples and the brain has the second highest total number of tissue-enriched genes (708  
66 genes).

67 In addition to the high enrichment of testis genes, we also identified aberrant expression of other tissue-enriched genes  
68 within the *Kdm5c*-KO brain. Although the *Kdm5c*-KO mice were male, we also observed significant enrichment of ovary-biased

69 DEGs in both the amygdala and hippocampus (Amygdala p = 0.00574; Hippocampus p = 0.048, Fisher's Exact) (Figure 1D).  
70 Intriguingly, many ovary and testis-biased DEGs have functions specific to germ cells and have no known role in the brain.  
71 For example, the testis-biased DEG *FK506 binding protein 6* (*Fkbp6*) is a known regulator of piRNA expression and meiosis  
72 in germ cells<sup>17,18</sup> (Figure 1C) while the ovary-enriched DEG *Zygotic arrest 1* (*Zar1*) was recently shown to sequester mRNAs  
73 in oocytes for meiotic maturation and early zygote development<sup>19</sup> (Figure 1D). Although not consistent across brain regions,  
74 we also found enrichment of two non-gonadal tissues - the liver (Amygdala p = 0.0398, Fisher's Exact Test) and the muscle  
75 (Hippocampus p = 0.0104, Fisher's Exact Test). An example of a liver-biased DEG is *Apolipoprotein C-I* (*Apoc1*), which is  
76 involved in lipoprotein metabolism (Figure 1E). Testis, ovary, and liver-enriched DEGs showed little to no expression in the  
77 wild-type brain, yet our mRNA-seq data indicate they are polyadenylated and spliced into mature transcripts (Figure 1C-E).  
78 Together, these results suggest misexpression of testis and other tissue-enriched genes within the brain is a major effect of  
79 KDM5C loss.

#### 80 Germline genes are aberrantly expressed in the male *Kdm5c*-KO brain

81 The testis contains both germ cells (meiotic cells, e.g. spermatogonia) and somatic cells (non-meiotic, e.g. Leydig cells)  
82 that support hormone production and germline functions. Select testis-enriched DEGs that were characterized previously  
83 had germline-specific functions<sup>9</sup>, suggesting *Kdm5c*-KO cells fail to demarcate between the soma and germline. To test if  
84 this holds true for all *Kdm5c*-KO testis-biased DEGs, we first assed their function through gene ontology. We found high  
85 enrichment of germline-relevant ontologies, including spermatid development (GO: 0007286, p.adjust = 6.2e-12) and sperm  
86 axoneme assembly (GO: 0007288, p.adjust = 2.45e-14) (Figure 2A).

87 To further validate if these testis DEGs are truly germline genes, we then compared their expression in somatic versus  
88 germ cells within the testis. We first compared their expression within wild-type versus germ cell-depleted testes<sup>20</sup>. In this  
89 study, germ cell depletion was accomplished by heterozygous *W* and *Wv* mutations in the enzymatic domain of *c-Kit* (*Kit*<sup>W/Wv</sup>),  
90 which prevents the maturation of germ cells and results in overall germline loss<sup>21</sup>. Almost all *Kdm5c*-KO testis-enriched  
91 DEGs lost expression with germ cell depletion (Figure 2B). The only testis-enriched DEG that did not show considerable  
92 downregulation with germline depletion was *FK506 binding protein 6* (*Fkbp6*), the aforementioned testis gene that regulates  
93 piRNA expression and meiosis in germ cells<sup>17,18</sup>. We then assessed testis-enriched DEG expression in a published single  
94 cell RNA-seq dataset that identified cell type-specific markers within the testis<sup>22</sup>. We found that while some testis-enriched  
95 DEGs were classified as specific markers for different germ cell types (e.g. spermatogonia, spermatocytes, round spermatids,  
96 and elongating spermatids) none marked somatic cells (Figure 2C). Together, these data suggest the somatic *Kdm5c*-KO  
97 brain aberrantly expresses germline genes.

98 We wanted to more deeply characterize germline gene misexpression with *Kdm5c* loss, but lacked a comprehensive list of  
99 mouse germline-enriched genes. To facilitate downstream analyses, we generated a list of male and female germline-enriched  
100 genes by evaluating gene expression in germline-depleted (*Kit*<sup>W/Wv</sup>) mice. Current available *Kit*<sup>W/Wv</sup> datasets included males  
101 and females at embryonic day 12, 14, and 16<sup>23</sup>, as well as adult male testes<sup>20</sup>.

102 We defined genes as germline-enriched if their expression met the following criteria: 1) their expression is greater than 1  
103 FPKM in wild-type gonads 2) their expression in any wild-type, non-gonadal tissue<sup>14</sup> does not exceed 20% of their maximum  
104 expression in the wild-type germline, and 3) their expression in the germ cell-depleted germline, for any sex or time point,  
105 does not exceed 20% of their maximum expression in the wild-type germline. These criteria yielded 1,288 germline-enriched  
106 genes (Figure 2D), which was hereafter used as a resource for assessing misexpression of germline genes with loss of  
107 *Kdm5c* (Supplementary table 1).

108 **Kdm5c-KO epiblast-like cells aberrantly express master regulators of germline identity**

109 Misexpression of germline genes in the adult *Kdm5c*-KO brain suggests mutants fail to demarcate germline and somatic  
110 cellular identity. Germ cells are typically distinguished from somatic cells soon after the embryo implants into the uterine  
111 wall<sup>24,25</sup> when a subset of epiblast stem cells become the primordial germ cells (PGCs) while the remainder differentiate into  
112 the ectoderm, mesoderm, and endoderm to form the somatic tissues<sup>26</sup>. This developmental time point can be modeled *in*  
113 *vitro* through differentiation of embryonic stem cells (ESCs) into post-implantation epiblast-like stem cells (EpiLCs) (Figure 3A,  
114 top). Previous studies have demonstrated that while some germline-enriched genes are also expressed in embryonic stem  
115 cells (ESCs), they are silenced as they differentiate into EpiLCs<sup>27</sup>. Therefore, we assessed if KDM5C was necessary for  
116 initial germline gene silencing by evaluating the impact of *Kdm5c* loss in male EpiLCs.

117 We first identified *Kdm5c*-KO EpiLC DEGs through DESeq2 analysis of our previously published RNA-seq dataset<sup>28</sup>  
118 (log2 fold change > 0.5, q < 0.1). Similar to the *Kdm5c*-KO brain, we observed general dysregulation of tissue-enriched  
119 genes, with the largest number of genes belonging to the brain and testis, although they were not significantly enriched  
120 (Figure 3B). Using our curated list, we found *Kdm5c*-KO EpiLCs aberrantly expressed 54germline-enriched genes, including  
121 the previously characterized hippocampal DEG<sup>9</sup> *Cytochrome C, testis-specific (Cyct)* (Figure 3C). Although we observed  
122 aberrant expression of many tissue-enriched genes, we did not observe any significant difference in primed pluripotency  
123 genes (Figure 3D) or gross changes in *Kdm5c*-KO cell morpholgy during differentiation (Figure 3E), indicating KDM5C loss  
124 does not impair EpiLC formation.

125 To evaluate if all germline DEGs are constitutively dysregulated or change over the course of development, we then  
126 compared EpiLC germline DEGs to those expressed in the *Kdm5c*-KO brain. The majority of EpiLC germline DEGs were  
127 unique to EpiLCs, with only *Cyct* shared across all sequencing datasets (Figure 3F). We then compared the function of  
128 EpiLC and brain germline DEGs and found particularly high enrichment of meiosis-related gene ontologies in EpiLCs (Figure  
129 3G), such as meiotic cell cycle (GO: 0051321, p.adjust = 4.44e-07). While there was modest enrichment of meiotic gene  
130 ontologies in both brain regions, the *Kdm5c*-KO hippocampus showed strong enrichment of late-stage sperm genes, such  
131 those involved in the sperm axoneme. This shift from meiotic genes to later spermatogenesis genes in the hippocampus  
132 suggests the germline developmental program could occur ectopically as *Kdm5c*-KO cells progresses through somatic  
133 tissue development. **note: this is strengthened by the ChIP-seq data since KDM5C is not directly bound to many**  
134 **brain/flagellar DEGs. This point might be stronger in the ChIPseq figure**

135 Interestingly, DEGs unique to *Kdm5c*-KO EpiLCs included key drivers of germline identity, such as *Stimulated by retinoic*  
136 *acid 8 (Stra8)* and *Deleted in azoospermia like (Dazl)* (Figure 3H). These genes are typically expressed during embryonic  
137 germ cell development to commit PGCs to the germline fate, but are also expressed later in life to trigger meiotic gene  
138 expression programs<sup>29–31</sup>. Of note, some germline genes, including *Dazl*, are also expressed in the two-cell embryo. However,  
139 we did not see misexpression of two-cell embryo-specific genes, like *Zscan4d* (q = 0.381), indicating *Kdm5c*-KO in EpiLCs do  
140 not revert back to a 2-cell state (Figure 3H).

141 *Dazl* is an RNA-binding protein that regulates the translation of germline mRNAs and is essential for germ cell function<sup>32</sup>.  
142 A significant portion of germline transcripts misexpressed in *Kdm5c*-KO EpiLCs are known binding targets of DAZL, including  
143 *Stra8*<sup>33</sup> (p = 1.698e-07, Fisher's Exact Test). This suggests expression of DAZL protein could promote the translation of other  
144 aberrant germline transcripts, influencing their ability to impact *Kdm5c*-KO cellular function. We thus tested DAZL protein  
145 expression in *Kdm5c*-KO EpiLCs through immunocytochemistry (Figure 3H). We observed about 25% of *Kdm5c*-KO EpiLCs  
146 expressed DAZL protein and it was localized to the cytoplasm (p = 0.0015, Welch's t-test), consistent with the pattern of DAZL  
147 expression in spermatogonia<sup>33</sup>. Altogether these results suggest tissue-specific genes are misexpressed during *Kdm5c*-KO

148 embryogenesis, including key drivers of germline identity that can be translated into protein.

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

150 • **note: do Direct vs indirect DEGs motif analysis**

151 • However, it is currently unclear if KDM5C binds to all germline DEGs and if its binding is maintained at any germline  
152 genes in neurons.

153 Previous work suggests KDM5C represses germline genes during early development, as re-expression of KDM5C in  
154 knockout neuronal cultures fails to suppress hippocampal germline DEGs and KDM5C is not bound to their promoters  
155 in neurons<sup>9</sup>. There is some evidence KDM5C binds to select germline gene promoters in ESCs<sup>9</sup>, including two recent  
156 independent screens that found KDM5C binds to Dazl's promoter<sup>34,35</sup>. As KDM5C's binding at germline gene promoters has  
157 not been systematically characterized, it is currently unclear what types of germline genes KDM5C regulates and if its binding  
158 is maintained at any germline genes in neurons.

159 To further characterize the mechanism behind *Kdm5c*-KO germline gene misexpression, we analyzed KDM5C chromatin  
160 immunoprecipitation followed by DNA sequencing (ChIP-seq) datasets in EpiLCs<sup>28</sup> and primary neuron cultures (PNCs) from  
161 the cortex and hippocampus<sup>13</sup>. EpiLCs had a higher total number of KDM5C peaks than PNCs (EpiLCs: 5,808, PNCs: 1,276,  
162 MACS2 q < 0.1 and fold enrichment > 1) and KDM5C was primarily localized to gene promoters in both cell types (EpiLCs:  
163 4,190, PNCs: 745 +/- 500kb from TSS), although PNCs showed increased localization to non-promoter regions (Figure 4A).

164 The majority of promoters bound by KDM5C in PNCs were also bound in EpiLCs (513 shared promoters), however a  
165 large portion of gene promoters were only bound by KDM5C in EpiLCs (3677 EpiLC only promoters) (Figure 4B). We then  
166 performed gene ontology analysis to compare the functions of genes bound by KDM5C at the promoter in different cell  
167 types. While there were no significant ontologies for genes with KDM5C only bound in PNCs, gene ontologies for peaks  
168 shared between PNCs and EpiLCs were enriched for functions involving nucleic acid turnover, such as deoxyribonucleotide  
169 metabolic process (GO:0009262, p.adjust = 8.28e-05) (Figure 4C). Germline-specific ontologies were only enriched in  
170 peaks unique to EpiLCs, such as meiotic nuclear division (GO: 0007127 p.adjust = 6.77e-16) and meiotic cell cycle process  
171 (GO:1903046, p.adjust = 5.05e-16) (Figure 3C). When comparing KDM5C binding at all germline gene promoters, KDM5C  
172 was only bound to a subset of germline gene promoters in EpiLCs and was not bound to any in PNCs (Figure 4D). Together,  
173 this suggests KDM5C is recruited to a subset of germline gene promoters in EpiLCs, including meiotic genes, but does not  
174 regulate germline genes in neurons.

175 To determine if the germline mRNAs expressed in the *Kdm5c*-KO brain and EpiLCs are direct targets of KDM5C, we then  
176 compared KDM5C binding at RNA-seq DEG promoters (Figure 4E). About one third of EpiLC specific and brain-specific  
177 (hippocampus, amygdala, or both) germline DEGs were bound by KDM5C in EpiLCs (EpiLC only: 36%, Brain only: 33.3%).  
178 Some notable differences in KDM5C binding for EpiLC-specific DEGs included *Dazl* and *Stra8*, two key drivers of germline  
179 identity discussed above. Although both mRNAs are expressed in *Kdm5c*-KO EpiLCs, KDM5C is only bound to *Dazl*'s  
180 promoter and not *Stra8*'s in EpiLCs (Figure 4F). In contrast to the unique DEGs, 3 out of the 4 genes dysregulated in both the  
181 brain and EpiLCs bound by KDM5C (Figure 4G). Again, we did not observe any KDM5C binding at germline gene promoters  
182 in PNCs, even for brain-specific DEGs (Figure 4H). Altogether, this suggests the majority of germline mRNAs expressed in  
183 *Kdm5c*-KO cells are dysregulated independent of direct KDM5C recruitment to their promoters during embryogenesis.

184 **KDM5C erases H3K4me3 to promote long-term repression of germline genes via DNA methylation**

185 Although KDM5C is generally thought to suppress transcription through erasure of histone 3 lysine 4 di- and trimethylation  
186 (H3K4me2/3)<sup>36</sup>, recent studies in ESCs have suggested KDM5C's repression *Dazl* is independent of its catalytic activity<sup>34</sup>.  
187 Somatic repression of germline genes is typically established during the transition between naive and primed pluripotency,  
188 which modeled in vitro as ESCs to EpiLC differentiation. In ESCs, chromatin regulators place repressive histone modifications  
189 at germline gene promoters, including histone 2A 119 monoubiquitination (H2AK119ub1) and histone 3 lysine 9 trimethylation  
190 (H3K9me3)<sup>27,37,38</sup>. Germline genes are then silenced long-term in EpiLCs by *de novo* placement of DNA CpG methylation<sup>27</sup>.  
191 It has been proposed KDM5C may promote germline gene silencing via H3K4me3 removal since DNA methylation is lost at  
192 select germline gene promoters in the hippocampus<sup>9</sup> and H3K4me3 can impair DNA methylation placement<sup>39,40</sup>. However,  
193 KDM5C was recently shown to repress *Dazl* in ESCs independent of its catalytic activity<sup>34</sup>. Because KDM5C's role in germline  
194 gene repression has only been characterized in ESCs and in the mature brain, it is currently unclear to what extent KDM5C is  
195 involved during transition between ESCs and EpiLCs and if its catalytic activity is required for long-term silencing.

196 To elucidate KDM5C's role in germline gene silencing, we first characterized KDM5C's substrates, histone 3 lysine 4  
197 di- and trimethylation (H3K4me2/3) in our previously published ChIP-seq datasets in wild type and *Kdm5c*-KO amygdala<sup>15</sup>  
198 and EpiLCs<sup>28</sup>. In congruence with previous work in the *Kdm5c*-KO hippocampus<sup>9</sup>, we observed aberrant accumulation of  
199 H3K4me3 around the transcription start site (TSS) of germline genes in the *Kdm5c*-KO amygdala (Figure 5A). We additionally  
200 found a marked increase in H3K4me2 germline gene TSSs in *Kdm5c*-KO EpiLCs (Figure 5B). Increase in H3K4me2 and  
201 H3K4me3 was highest in *Kdm5c*-KO cells that were highest genes that are bound by KDM5C at their promoter in EpiLCs  
202 (**note: do analysis to check if true**).

203 To assess KDM5C's embryonic role in germline gene silencing, we first characterized KDM5C's expression in ESCs and  
204 EpiLCs by harvesting RNA and protein at 0 hours (ESCs), 24 hours, and 48 hours (EpiLCs). While *Kdm5c* mRNA steadily  
205 decreased from 0 to 48 hours of differentiation, KDM5C protein initially increased from 0 to 24 hours but then decreased to  
206 near knockout levels by 48 hours (Figure 5C).

207 We then determined the role of KDM5C in the initial placement of DNA methylation at germline gene promoters by  
208 performing whole genome bisulfite sequencing in wild-type and *Kdm5c*-KO ESCs and 96 hour extended EpiLCs (exEpiLCs)  
209 (**note: check**). - While wild-type cells accumulate high levels of DNA methylation over the course of ESCs to exEpiLC  
210 differentiation, DNA methylation is markedly reduced in *Kdm5c*-KO

211 **Discussion**

- 212 • The demarcation of the germ vs soma is a key feature of multicellularity
- 213 • Other H3K4me regulators, anything known about tissue-biased gene expression?
- 214 • This suggests the germline developmental program is occurring ectopically as *Kdm5c*-KOs progresses through somatic  
215 tissue development
- 216 • tissue-biased gene expression:
  - 217 – However unlike the gonadal-biased DEGs, many liver and muscle-biased DEGs have a known involvement in brain  
218 function. For example, the liver-biased DEG is *Apolipoprotein C-I* (*Apoc1*), is important for lipoprotein metabolism  
219 but has also been shown to influence learning and memory (Figure 1E).

- 220 • Otx2 is properly expressed in *Kdm5c*-KO EpiLCs and is known to prevent pgc identity <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6485399/>
- 222 • Altogether, these data indicate that while some germline genes are misexpressed due to direct loss of KDM5C binding  
223 during emryogenesis, secondary downstream mechanisms can also promote their aberrant transcription.
- 224 • KDM5C is dynamically regulated during the window of embryonic germline gene silencing, our EpiLC ChIP-seq is likely  
225 catching the tail end of KDM5C's main involvement.
- 226 • Papers to read/reference:
- 227 – Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells: [https://www.cell.com/fulltext/S0092-8674\(11\)00771-9](https://www.cell.com/fulltext/S0092-8674(11)00771-9)
- 228 – two cell gene list used by Suzuki et al Max paper is based on 2 cell sequencing: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3395470/>

## 231 References

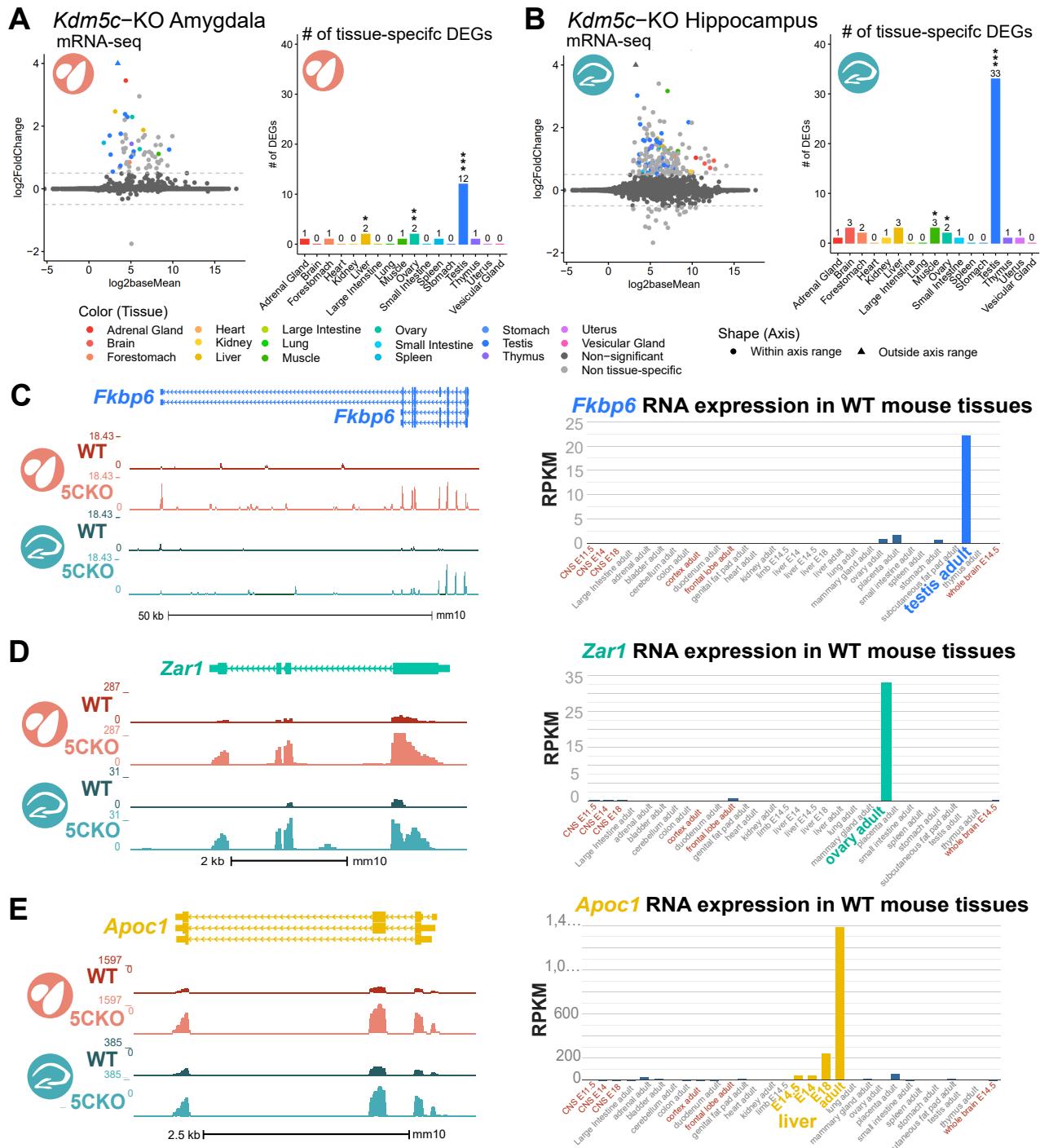
- 232 1. Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* *403*, 41–45. <https://doi.org/10.1038/47412>.
- 233 2. Jenuwein, T., and Allis, C.D. (2001). Translating the Histone Code. *Science* *293*, 1074–1080. <https://doi.org/10.1126/science.1063127>.
- 234 3. Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* *276*, 565–570. <https://doi.org/10.1038/276565a0>.
- 235 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.85.21.8136>.
- 236 5. Zhou, Z., Hong, E.J., Cohen, S., Zhao, W.-N., Ho, H.-Y.H., Schmidt, L., Chen, W.G., Lin, Y., Savner, E., Griffith, E.C., et al. (2006). Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. *Neuron* *52*, 255–269. <https://doi.org/10.1016/j.neuron.2006.09.037>.
- 237 6. Hirabayashi, Y., Suzuki, N., Tsuboi, M., Endo, T.A., Toyoda, T., Shinga, J., Koseki, H., Vidal, M., and Gotoh, Y. (2009). Polycomb Limits the Neurogenic Competence of Neural Precursor Cells to Promote Astrocytic Fate Transition. *Neuron* *63*, 600–613. <https://doi.org/10.1016/j.neuron.2009.08.021>.
- 238 7. Kassis, J.A., Kennison, J.A., and Tamkun, J.W. (2017). Polycomb and Trithorax Group Genes in *Drosophila*. *Genetics* *206*, 1699–1725. <https://doi.org/10.1534/genetics.115.185116>.
- 239 8. 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 suppressor complex. *Neuron* *64*, 678–691. <https://doi.org/10.1016/j.neuron.2009.11.019>.
- 240 9. Scandaglia, M., Lopez-Atalaya, J.P., Medrano-Fernandez, A., Lopez-Cascales, M.T., Del Blanco, B., Lipinski, M., Benito, E., Olivares, R., Iwase, S., Shi, Y., et al. (2017). Loss of Kdm5c Causes Spurious Transcription and Prevents the Fine-Tuning of Activity-Regulated Enhancers in Neurons. *Cell Rep* *21*, 47–59. <https://doi.org/10.1016/j.celrep.2017.09.014>.

- 250 10. Claes, S., Devriendt, K., Van Goethem, G., Roelen, L., Meireleire, J., Raeymaekers, P., Cassiman, J.J., and Fryns, J.P.  
251 (2000). Novel syndromic form of X-linked complicated spastic paraplegia. *Am J Med Genet* **94**, 1–4.
- 252 11. Jensen, L.R., Amende, M., Gurok, U., Moser, B., Gimmel, V., Tzschach, A., Janecke, A.R., Tariverdian, G., Chelly, J.,  
253 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>.
- 254 12. Carmignac, V., Nambot, S., Lehalle, D., Callier, P., Moortgat, S., Benoit, V., Ghoumid, J., Delobel, B., Smol, T., Thuillier,  
255 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>.
- 256 13. Iwase, S., Brookes, E., Agarwal, S., Badeaux, A.I., Ito, H., Vallianatos, C.N., Tomassy, G.S., Kasza, T., Lin, G.,  
257 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>.
- 258 14. Li, B., Qing, T., Zhu, J., Wen, Z., Yu, Y., Fukumura, R., Zheng, Y., Gondo, Y., and Shi, L. (2017). A Comprehensive  
259 Mouse Transcriptomic BodyMap across 17 Tissues by RNA-seq. *Sci Rep* **7**, 4200. <https://doi.org/10.1038/s41598-017-04520-z>.
- 260 15. 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.,  
261 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>.
- 262 16. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data  
263 with DESeq2. *Genome Biol* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
- 264 17. Crackower, M.A., Kolas, N.K., Noguchi, J., Sarao, R., Kikuchi, K., Kaneko, H., Kobayashi, E., Kawai, Y., Kozieradzki, I.,  
265 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>.
- 266 18. Xirol, J., Cora, E., Koglgruber, R., Chuma, S., Subramanian, S., Hosokawa, M., Reuter, M., Yang, Z., Berninger, P.,  
267 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>.
- 268 19. Cheng, S., Altmeppen, G., So, C., Welp, L.M., Penir, S., Ruhwedel, T., Menelaou, K., Harasimov, K., Stützer, A.,  
269 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>.
- 270 20. Mueller, J.L., Skaletsky, H., Brown, L.G., Zaghloul, S., Rock, S., Graves, T., Auger, K., Warren, W.C., Wilson, R.K., and  
271 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>.
- 272 21. Handel, M.A., and Eppig, J.J. (1979). Sertoli Cell Differentiation in the Testes of Mice Genetically Deficient in Germ  
273 Cells. *Biology of Reproduction* **20**, 1031–1038. <https://doi.org/10.1095/biolreprod20.5.1031>.
- 274 22. Green, C.D., Ma, Q., Manske, G.L., Shami, A.N., Zheng, X., Marini, S., Moritz, L., Sultan, C., Gurczynski, S.J., Moore,  
275 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>.

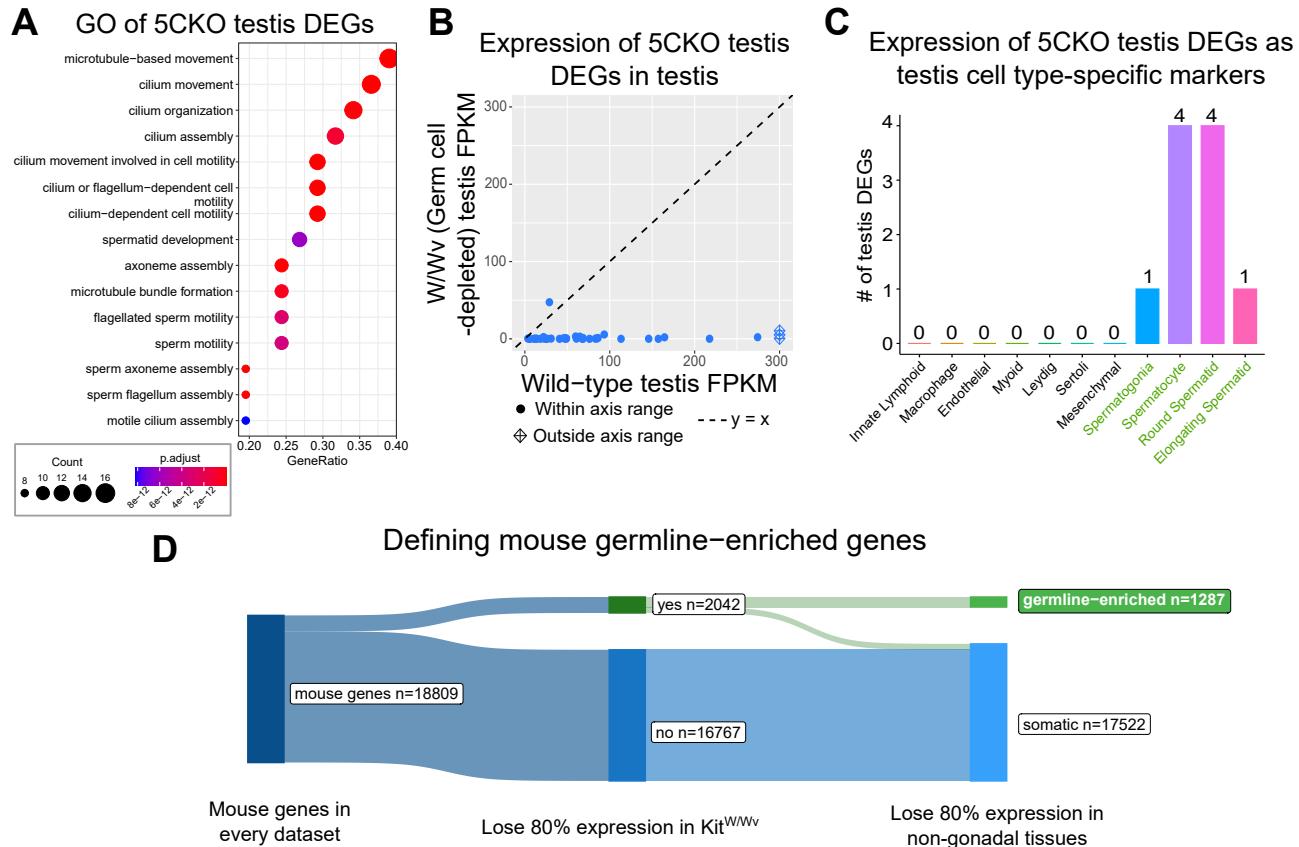
- 276 23. Soh, Y.Q., Junker, J.P., Gill, M.E., Mueller, J.L., van Oudenaarden, A., and Page, D.C. (2015). A Gene Regulatory  
277 Program for Meiotic Prophase in the Fetal Ovary. *PLoS Genet* 11, e1005531. <https://doi.org/10.1371/journal.pgen.1005531>.
- 278 24. Magnúsdóttir, E., and Surani, M.A. (2014). How to make a primordial germ cell. *Development* 141, 245–252.  
279 <https://doi.org/10.1242/dev.098269>.
- 280 25. Günesdogan, U., Magnúsdóttir, E., and Surani, M.A. (2014). Primordial germ cell specification: A context-dependent  
281 cellular differentiation event [corrected]. *Philos Trans R Soc Lond B Biol Sci* 369. <https://doi.org/10.1098/rstb.2013.0543>.
- 282 26. Bardot, E.S., and Hadjantonakis, A.-K. (2020). Mouse gastrulation: Coordination of tissue patterning, specification and  
283 diversification of cell fate. *Mechanisms of Development* 163, 103617. <https://doi.org/10.1016/j.mod.2020.103617>.
- 284 27. Mochizuki, K., Sharif, J., Shirane, K., Uranishi, K., Bogutz, A.B., Janssen, S.M., Suzuki, A., Okuda, A., Koseki, H.,  
285 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>.
- 286 28. Samanta, M.K., Gayen, S., Harris, C., Maclary, E., Murata-Nakamura, Y., Malcore, R.M., Porter, R.S., Garay, P.M.,  
287 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>.
- 288 29. Koubova, J., Menke, D.B., Zhou, Q., Capel, B., Griswold, M.D., and Page, D.C. (2006). Retinoic acid regulates  
289 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>.
- 290 30. Lin, Y., Gill, M.E., Koubova, J., and Page, D.C. (2008). Germ Cell-Intrinsic and -Extrinsic Factors Govern Meiotic  
291 Initiation in Mouse Embryos. *Science* 322, 1685–1687. <https://doi.org/10.1126/science.1166340>.
- 292 31. Endo, T., Mikedis, M.M., Nicholls, P.K., Page, D.C., and De Rooij, D.G. (2019). Retinoic Acid and Germ Cell  
293 Development in the Ovary and Testis. *Biomolecules* 9, 775. <https://doi.org/10.3390/biom9120775>.
- 294 32. Li, H., Liang, Z., Yang, J., Wang, D., Wang, H., Zhu, M., Geng, B., and Xu, E.Y. (2019). DAZL is a master translational  
295 regulator of murine spermatogenesis. *Natl Sci Rev* 6, 455–468. <https://doi.org/10.1093/nsr/nwy163>.
- 296 33. Mikedis, M.M., Fan, Y., Nicholls, P.K., Endo, T., Jackson, E.K., Cobb, S.A., De Rooij, D.G., and Page, D.C. (2020).  
297 DAZL mediates a broad translational program regulating expansion and differentiation of spermatogonial progenitors.  
eLife 9, e56523. <https://doi.org/10.7554/eLife.56523>.
- 298 34. Gupta, N., Yakhou, L., Albert, J.R., Azogui, A., Ferry, L., Kirsh, O., Miura, F., Battault, S., Yamaguchi, K., Laisné,  
299 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>.
- 300 35. Al Adhami, H., Vallet, J., Schaal, C., Schumacher, P., Bardet, A.F., Dumas, M., Chicher, J., Hammann, P., Daujat,  
301 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>.
- 302 36. Iwase, S., Lan, F., Bayliss, P., De La Torre-Ubieta, L., Huarte, M., Qi, H.H., Whetstone, J.R., Bonni, A., Roberts, T.M.,  
303 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>.

- 304 37. Endoh, M., Endo, T.A., Shinga, J., Hayashi, K., Farcas, A., Ma, K.W., Ito, S., Sharif, J., Endoh, T., Onaga, N., et  
305 al. (2017). PCGF6-PRC1 suppresses premature differentiation of mouse embryonic stem cells by regulating germ  
306 cell-related genes. *eLife* 6. <https://doi.org/10.7554/eLife.21064>.
- 307 38. Liu, M., Zhu, Y., Xing, F., Liu, S., Xia, Y., Jiang, Q., and Qin, J. (2020). The polycomb group protein PCGF6  
308 mediates germline gene silencing by recruiting histone-modifying proteins to target gene promoters. *J Biol Chem* 295,  
309 9712–9724. <https://doi.org/10.1074/jbc.RA119.012121>.
- 310 39. Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition  
311 of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. *EMBO Reports* 10,  
312 1235–1241. <https://doi.org/10.1038/embor.2009.218>.
- 313 40. Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., et al. (2015). Structural insight into  
314 autoinhibition and histone H3-induced activation of DNMT3A. *Nature* 517, 640–644. <https://doi.org/10.1038/nature13899>.
- 315 41. 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>.
42. 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>.

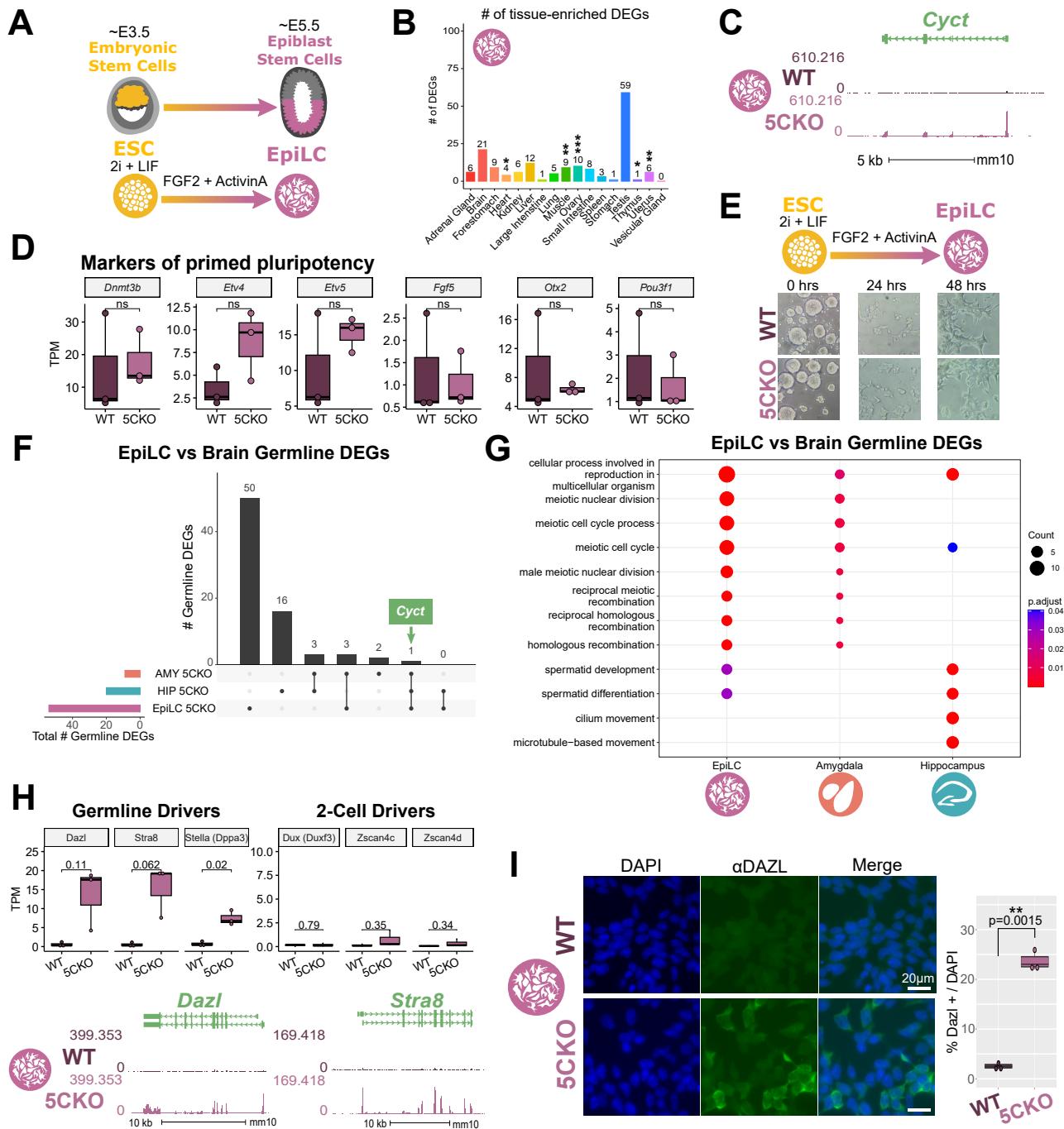
316 **Figures and Tables**



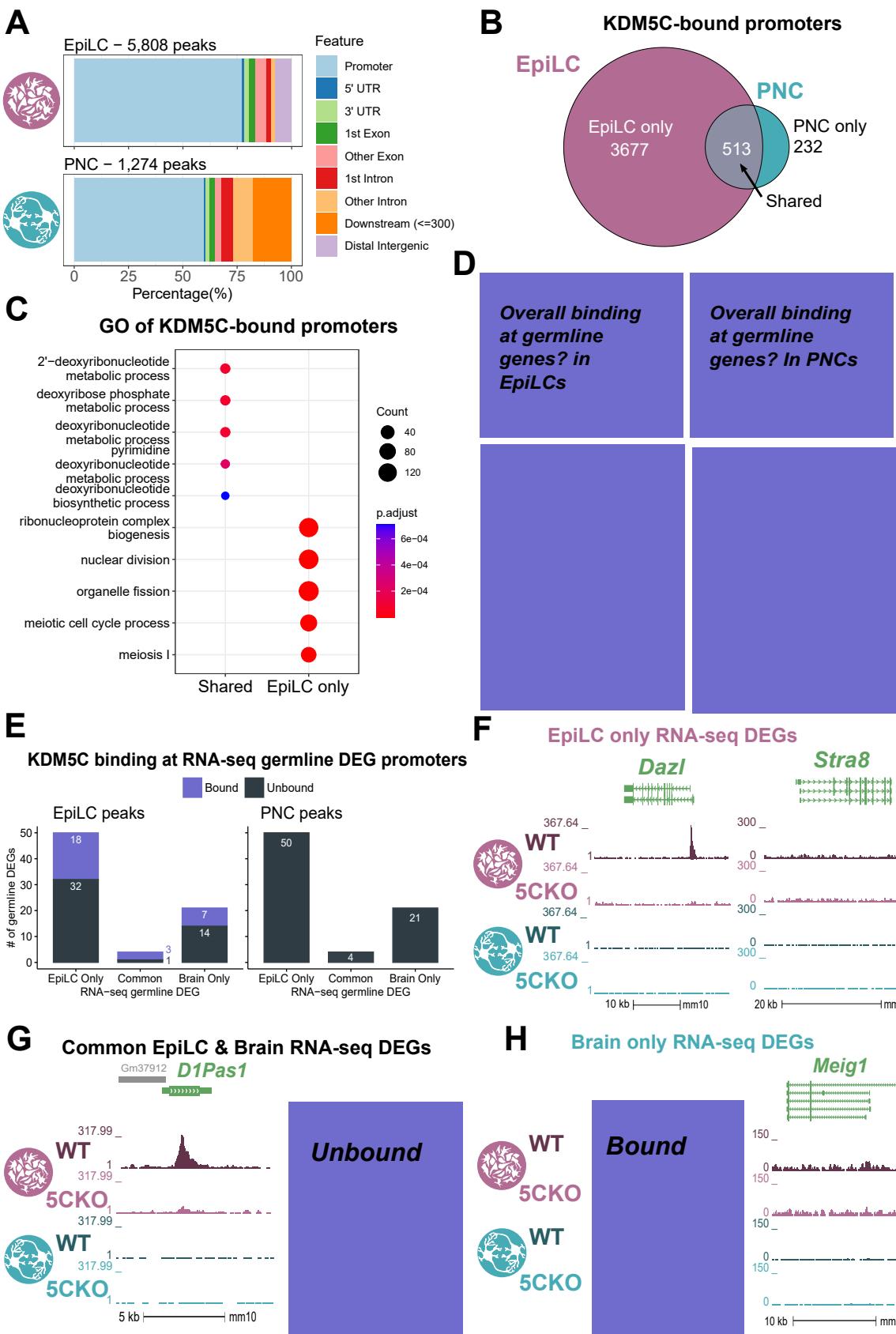
**Figure 1: Tissue-enriched genes are misexpressed in the *Kdm5c-KO* brain.** **A.** Expression of tissue-enriched genes in the male *Kdm5c-KO* amygdala. Left - MA plot of mRNA-seq. Right - Number of tissue-enriched differentially expressed genes (DEGs). \*  $p<0.05$ , \*\*  $p<0.01$ , \*\*\*  $p<0.001$ , Fisher's exact test. **B.** Expression of tissue-enriched genes in the male *Kdm5c-KO* hippocampus. **C.** Left - Average bigwigs of an example aberrantly expressed testis-enriched DEG, *Cytochrome C, testis-specific (Cyclt)* in the wild-type (WT) and *Kdm5c-KO* (5CKO) amygdala (red) and hippocampus (teal). Right - Expression of *Cyclt* in wild-type tissues from NCBI Gene, with testis highlighted in blue and brain tissues highlighted in red. **D.** Left - Average bigwigs 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 - Average bigwigs 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. 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 germline-depleted mice and in adult mouse non-gonadal tissues (Li et al 2017).

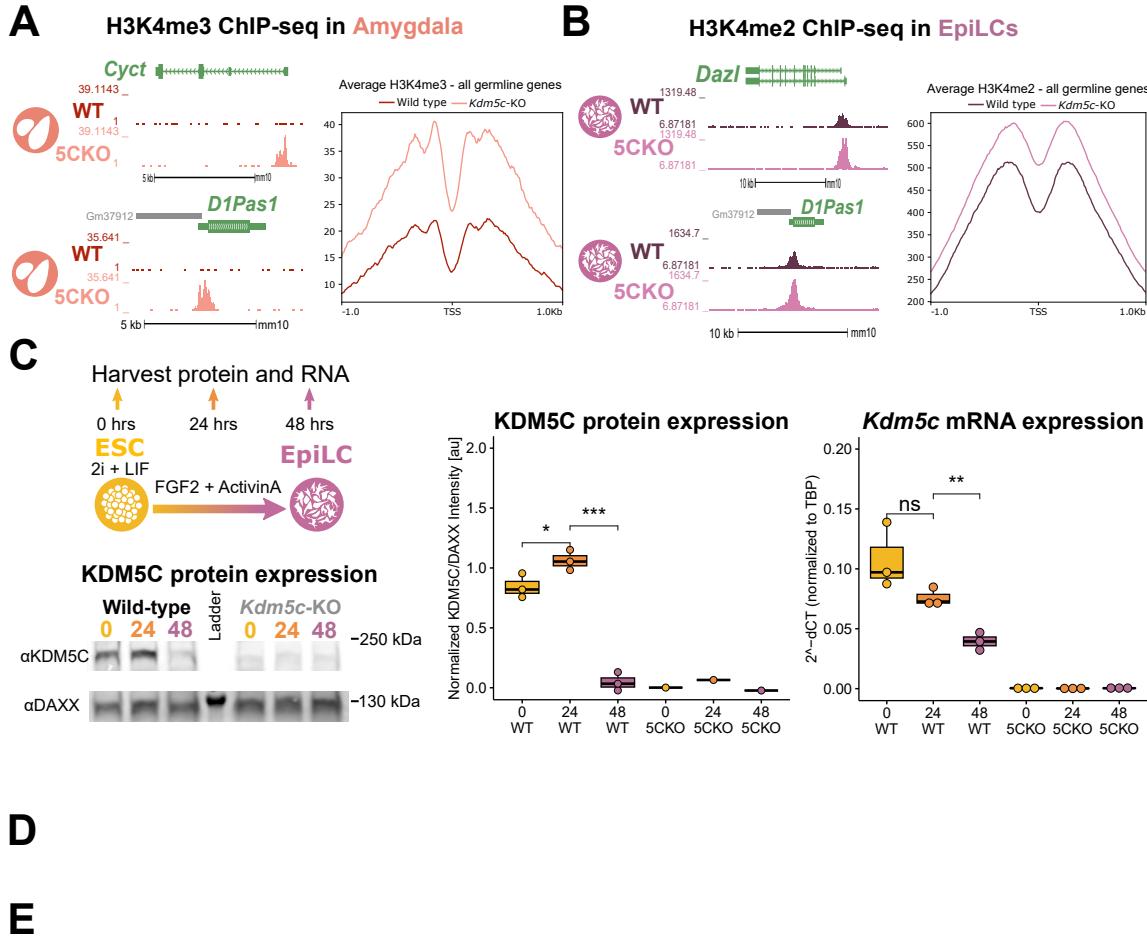


**Figure 3: Kdm5c-KO epiblast-like cells express key drivers regulators of germline identity** **A.** Top - Diagram of *in vivo* differentiation of embryonic stem cells (ESCs) of the inner cell mass into epiblast stem cells. Middle - *in vitro* differentiation of ESCs into epiblast-like cells (EpiLCs). Bottom - representative images of wild-type (WT) and *Kdm5c*-KO ESC to EpiLC differentiation. Brightfield images taken at 20X. **B.** 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). **C.** Number of tissue-enriched differentially expressed genes (DEGs). \* p<0.05, \*\* p<0.01, \*\*\* p<0.001, Fisher's exact test. **D.** Average bigwigs of an example germline gene, *Cyct*, that is dysregulated *Kdm5c*-KO EpiLCs. **E.** Upset plot displaying the overlap of germline DEGs expressed in *Kdm5c*-KO EpiLCs, amygdala (AMY), and hippocampus (HIP) RNA-seq datasets. **F.** enrichPlot comparing enriched gene ontologies for *Kdm5c*-KO EpiLC, amygdala, and hippocampus germline DEGs. **G.** Top - Example germline identity DEGs unique to EpiLCs, p-values for Welch's t-test. Bottom - Average bigwigs of *Dazl* and *Stra8* expression in wild-type and *Kdm5c*-KO EpiLCs **H.** Immunocytochemistry of DAZL in male wild-type (WT) and *Kdm5c*-KO (5CKO) EpiLCs. Percentage of DAZL-positive cells normalized to DAPI, p-value for Welch's t-test.



**Figure 4: 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 in EpiLCs (purple) and PNCs (blue). **C.** Gene ontology (GO) comparison of genes with KDM5C bound to their promoter. 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 at all germline-enriched genes in EpiLCs (left) and PNCs (right). **E.** KDM5C binding at the promoters of RNA-seq germline DEGs. Genes were classified as either only dysregulated in EpiLCs (EpiLC only), genes dysregulated in the hippocampus or amygdala but not EpiLCs (brain only), or genes dysregulated in both EpiLCs and the brain (common). (Legend continued on next page.)

**Figure 4: KDM5C binds to a subset of germline gene promoters during early embryogenesis.** (Legend continued.) **F.** Example ChIP-seq bigwigs of DEGs unique to EpiLCs. Although both are expressed in EpiLCs, KDM5C is bound to the *Dazl* promoter but not the *Stra8* promoter in EpiLCs. **F.** Bigwigs of the upregulated imprinted gene *Dlk1* that shows rescue in the double mutant brain. **G.** Example ChIP-seq bigwigs of DEGs common between brain and EpiLCs. KDM5C is bound to the *D1Pas1* promoter but not the *XXX* promoter in EpiLCs. **H.** Example ChIP-seq bigwigs of DEGs unique to the brain. KDM5C is bound to the *XXX* promoter but not the *Meig1* promoter



**Figure 5: KDM5C's catalytic activity promotes long-term silencing of germline genes via DNA methylation.** **A.** Left - Bigwigs of representative histone 3 lysine 4 trimethylation (H3K4me3) ChIP-seq peaks at two germline genes in the wild-type (WT) and *Kdm5c*-KO (5CKO) adult amygdala. Right - Average H3K4me3 at the transcription start site (TSS) of all germline-enriched genes in wild-type (dark red) and *Kdm5c*-KO amygdala. **B.** Left - Bigwigs of representative histone 3 lysine 4 dimethylation (H3K4me2) ChIP-seq peaks at representative germline genes in wild-type and *Kdm5c*-KO EpiLCs. **C.** RNA and protein expression of KDM5C across ESC to EpiLC differentiation. Top left - diagram of differentiation protocol and collection time points. Bottom left - representative lanes of Western blot for KDM5C and DAXX. Middle - KDM5C protein expression normalized to DAXX. Quantified intensity using ImageJ (artificial units - au). Right - RT-qPCR of *Kdm5c* RNA expression, calculated in comparison to TBP expression ( $2^{-\Delta\Delta CT}$ ). **D.** XXX. **E.** XXX.

317 **Figure outline:**

318     **Figure 1: Misexpression of tissue-specific genes in the *Kdm5c*-KO brain** \* MA-plot and bar graphs of tissue-enriched  
319     genes \* Example testis-specific genes (NCBI and bigwigs) \* An example ovary tissue-specific gene \* An example muscle/liver  
320     tissue-specific gene (NCBI and bigwigs)

321     **Figure 2: The male *Kdm5c*-KO brain expresses male and female germline-enriched genes** \* Gene ontology of  
322     testis DEGs in the amygdala and hippocampus - ontologies are germline ontologies \* Expression of testis DEGs in germline-  
323     depleted testis (this is adult testis data) \* scRNAseq of testis - # of testis DEGs that are germline-specific markers \* Although  
324     far fewer, 5CKO brain also expresses ovary-enriched genes(NCBI and bigwigs of Zar1) \* These ovary enriched genes are  
325     also germline specific (NCBI/Li tissues are in adult ovary, so it would be best to show they're oocyte-specific in adult ovary.  
326     But I don't think there's a published adult female W/Wv dataset. Could try looking at scRNAseq or just do TPM in embryonic  
327     W/Wv data since oocytes are developed at this point? Or both?) \* Defining what is/isn't a germline gene, and which are  
328     male/female biased using embryonic W/Wv data

329     **Figure 3: *Kdm5c*-KO epiblast-like cells express key drivers regulators of germline identity** \* A) ESC to EpiLC  
330     differentiation Left - Morphology is unchanged, \* B) 5CKO EpiLCs express EpiLC differentiation genes similar to WT lvs \*  
331     C) Male EpiLCs express germline genes (example Cyct again) \* Overlap between brain and EpiLC germline genes - show  
332     they're mostly unique \* GO of Brain and EpiLC germline genes (meiotic enriched) \* Bigwigs or TPM of master regulators  
333     \* Show that while some are also 2-cell genes (Dazl), 2-cell specific genes aren't dysregulated (Zscan4). Important point  
334     because published KDM5C dazl paper is saying KDM5C is a 2-cell regulator, but as far as I can tell only genes shared  
335     between germline and 2-cell are dysregulated.

336     Staining of Dazl (+ Stra8 if I can get it to work)

337     **Figure 4: Loss of KDM5C's catalytic activity impairs DNAme placement and long-term silencing of germline  
338     genes** \* Increase in H3K4me3 in *Kdm5c*-KO amygdala at germline genes \* Increase in H3K4me2 in EpiLCs at germline  
339     genes \* *Kdm5c* binding in EpiLCs vs PNCs to show that germline repression is happening in early embryo \* Previous studies  
340     only looked at ESCs, unknown if catalytic activity is required for long-term repression, especially since DNA methylation is  
341     placed later). KDM5C RNA and protein ESC → EpiLC (increasing then decreasing) \* RNA expression of germline genes with  
342     catalytic dead rescue (Ilakkiya) \* DNA methylation in WT and 5CKO EpiLCs (Ilakkiya)

343     **Figure 5: Ectopic, germline-like phenotypes in *Kdm5c*-KO ESCs/EpiLCs** \* Sycp3 staining \* DDX4 staining and  
344     repression of retrotransposons \* Cilia??

345 **Notes**

346     Gaps in knowledge addressed: \* Are other tissue-enriched genes dysregulated, or only testis, germline genes? \* Curating  
347     a robust list of male and female germline genes \* Should talk about 2-cell genes vs germline genes - way to systematically  
348     categorize? \* Mechanism behind long-term germline gene misexpression \* Recent evidence suggests loss of KDM5C  
349     in ESCs express some germline genes \* Unclear if catalytic activity is required for long-term silencing \* Unclear if their  
350     dysregulation lasts throughout life or the same between brain or not \* When in development does it begin? - Recent evidence  
351     suggests some germline genes expressed in 5CKO ESCs but unclear if their dysregulation lasts throughout differentiation  
352     and if the identity of germline genes are different compared to the brain \* Are there functional consequences to germline  
353     gene misexpression?

354     Introduction: \* Chromatin regulators are important for cellular identity \* H3K4me1-3 linked to active gene promoters and

355 enhancers \* Surprisingly, mutations in Chromatin regulators lead to many NDDs (including many H3K4 regulators) \* Recent  
356 studies have shown some chromatin regulators are important for regulating neuron-specific gene expression/chromatin  
357 stat\_compare\_means \* However, loss of some chromatin regulators can also lead to ectopic expression of tissue-enriched  
358 genes \* Very few studies have looked at these genes and it's unclear if these genes contribute to NDD impairments. \*  
359 Necessary to first characterize the mechanism behind their derepression to identify molecular footholds into testing their  
360 contribution to neuronal impairments and potential for therapeutic intervention

- 361 • Loss of KDM5C can result in the misexpression of genes typically only found in the testis
- 362     – Misexpression of tissue-enriched genes hasn't been systematically characterized - Unclear if these genes are  
363         exceptions or if other tissue-specific genes are dysregulated
- 364     – Interestingly, these genes (Cyct, D1pas1) typically function in the germline
- 365     – Germ cells (meiotic cells) are typically distinguished from somatic cells very early on in embryogenesis and is a  
366         key feature of multicellularity
- 367     – Chromatin regulators are very important for decommissioning germline genes and act successively the embryo  
368         implants into the uterine wall
- 369         \* Most studies have focused on ESCs, which have a similar transcriptome to germ cells / 2-cells
- 370         \* recently, KDM5C was shown to repress DAZL in ESCs, independent of its catalytic activity
- 371         \* However, DNA methylation is lost in the mature 5CKO brain, DNA methylation is placed later and it's Unclear if  
372         it's required for long-term repression (maybe too specific, just trying to go into the fact that the mechanism is  
373         partially understood but unclear)
- 374     – Systematic characterization of ectopic germline genes hasn't been done
- 375         \* unknown if other germline-enriched genes are dysregulated, including oocyte-specific genes
- 376         \* Crucially, it's unknown if misexpression of the germline program leads to functional consequences in 5CKO  
377         cells.

378 **Germline gene repression background:**

379 Interestingly, some of the ectopic testis transcripts identified in the *Kdm5c*-KO brain are typically expressed in germ cells<sup>9</sup>.  
380 Unlike somatic cells, germ cells (e.g. sperm and eggs) undergo meiosis and pass on their genetic material to the next genera-  
381 tion. The germline and the soma are typically distinguished during early embryogenesis, when germline genes are silenced  
382 in epiblast stem cells soon after implantation and only reactivated in a subset to form the germline. Chromatin regulators  
383 play a key role in decommissioning germline genes as the embryo transitions from naive to primed pluripotency by placing  
384 repressive histone H2A lysine 119 monoubiquitination (H2AK119ub1)<sup>37</sup>, histone 3 lysine 9 trimethylation (H3K9me3)<sup>27,37</sup>,  
385 and DNA CpG methylation<sup>27,41,42</sup> at germline gene promoters. KDM5C may also be involved in this early decommissioning of  
386 germline genes, as re-expression of KDM5C in neurons fails to suppress their dysregulation<sup>9</sup>. In support of this, KDM5C  
387 was very recently shown to repress *Deleted in azoospermia like* (*Dazl*), a key regulator of germline development, in mouse  
388 embryonic stem cells (ESCs)<sup>34,35</sup>. In support of this, two independent screens in mouse embryonic stem cells (ESCs) recently  
389 identified KDM5C as a repressor of *Deleted in azoospermia like* (*Dazl*), a key regulator of germline development. However,  
390 KDM5C's role in embryonic germline gene repression is currently unclear, given that *Dazl* is also expressed in ESCs and in  
391 the 2-cell stage and germline gene misexpression has yet to be globally characterized during *Kdm5c*-KO embryogenesis.