

**Soma-To-Germline Transition and Sexually Dimorphic Transcription in a
Chromatin-Linked Neurodevelopmental Disorder**

by

Katherine M. Bonefas

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Neuroscience)
in The University of Michigan
2024

Doctoral Committee:

Associate Professor Shigeki Iwase, Chair
Professor Gabriel Corfas
Associate Professor Kenneth Y. Kwan
Professor Michael A. Sutton
Associate Natalie C. Tronson

Katherine Bonefas

kbonefas@umich.edu

ORCID iD: [0000-0002-1845-9050](https://orcid.org/0000-0002-1845-9050)

©Katherine Bonefas 2024

Dedication

First and foremost I dedicate this work to the patients and families with Claes-Jensen X-linked intellectual disability. For all patient families I've met over the years, even those I met briefly, you have truly made me a better scientist and advocate. I also dedicate this work to my family, who always encouraged me to follow down every random path of my interests and thus who I blame for my thesis work spanning so many topics. My fiancé, George Hundt, who listens to both my smartest and dumbest ideas, and has always been my go-to person when I'm taking living on a rock hurtling through space too seriously. My cat, Pancake, who moved across the country with me, sits with me every day when I'm working, and reminds me to take breaks in the rudest way her sweet little heart will allow. My friends, especially Drs. Bethany Beekley and Chelsea Phillips, who learned how to be a scientist alongside me and were always the perfect sounding board for both the best and worst days. Finally, I dedicate this thesis to all of the young scientists who's anxiety and fear of being imperfect holds them back from living, including my younger self. Learn to let it all go, it's great on the other side.

Acknowledgements

First and foremost I would like to thank my thesis mentor, Dr. Shigeki Iwase. Thank you for entertaining every crazy idea I've had while simultaneously supplying even crazier ones. You taught me to think big but test rigorously, to listen to what Nature is telling me, and to always remember how much joy there is in being a scientist. This dissertation would not be the same without the invaluable feedback from my thesis committee, Drs. Gabriel Corfas, Michael Sutton, Kenneth Kwan, and Natalie Tronson. During both formal and informal meetings, they have guided me into becoming a better scientist while also teaching me how to exist in academia. I would also like to thank my lab mates, Drs Yumie Nakamura, Christina Valliantos, Robert Porter, and Patricia Garay, for guiding my hands on training. They also provided so much valuable feedback over the years, along with Dr. Masayoshi Nagai, Dr. Bo Zhou, Dr. Takao Tsukahara, Cecilia Gavilan, and Ilakkiya Venkatachalam. I would also like to thank our collaborators on this project, Dr. Jacob Mueller and Dr. Sundeep Kalantry. Dr. Mueller who, in addition to providing his expertise and perspective in germ cell biology, very generously provided unpublished RNA-seq data that was fundamental for classifying what qualifies as a germline gene. Dr. Sundeep Kalantry's expertise in epiblast stem cell culture and chromatin regulation was fundamental for this work and also provided rigorous feedback. Drs Peter and Janice Farrehi and their family who, on top of being wonderful human beings to know, generously provided much of the funding for this work as well as samples for future work based on these findings. The Neuroscience Graduate Program administrative staff, who always saved my life in a crisis and provided so many professional

development opportunities that were instrumental to my future successes. And finally I would like to thank my undergraduate mentors, Dr. Albert Mackrell and Dr. Michael Drew, who respectively taught me the fundamentals of being a molecular biologist and a neuroscientist.

Table of Contents

Dedication	ii
Acknowledgements	iii
List of Figures	vii
Abstract	ix
Chapter 1 - Introduction	1
Shared and sexually dimorphic features of neurodevelopmental disorders.	1
Chromatin regulators, neurodevelopmental disorders, and cellular identity	2
Lysine demethylase 5C (KDM5C) offers a window into the influence of sex and cellular identity on neurodevelopment	4
Chapter 2 - Materials and Methods	7
Mice	7
Cell culture	8
RNA sequencing	9
Published datasets	9
RNA isolation and sequencing	9
Alignment and analysis	10
Chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq)	11

Classifying tissue-enriched and germline-enriched genes	11
Chapter 3 - Shared and Sexually Dimorphic Gene Expression Changes in the Brain with <i>Kdm5c</i> Loss	13
Introduction	13
Shared and sexually dimorphic alterations in gene expression within the cortex and hippocampus with loss of <i>Kdm5c</i>	15
Germline genes are aberrantly transcribed in the brain of male and female <i>Kdm5c</i> mutants	18
Defining germline-enriched genes	21
Discussion	22
Chapter 4 - Characterization of KDM5C-Mediated Germline Gene Repression in Males and Females	25
Introduction	25
Master regulators of germline identity are aberrantly expressed during <i>Kdm5c</i> mutant embryogenesis	27
KDM5C is recruited to germline genes during early embryogenesis	30
Germline gene misexpression occurs in multiple <i>Kdm5c</i> -KO somatic tissues and is influenced by retinoic acid	33
Discussion	37
Chapter 5 - Consequences of Ectopic Germline Gene Expression Upon Neurodevelopment	41
Introduction	41
Dysregulation of genomic imprinting and transposable element expression with <i>Kdm5c</i> loss	44
Genomic imprinting	44
piRNAs and transposable elements	46
Amelioration of <i>Kdm5c</i> -KO brain transcriptome dysregulation through genetic deletion of <i>Dazl</i>	48

Discussion	50
Chapter 6 - Conclusions and Future Directions	54
References	59

List of Figures

1	<i>Kdm5c</i> loss produces both shared and sexually dimorphic gene expression changes in the hippocampus and cortex	17
2	Misexpression of tissue-specific genes, including germline genes, with loss of <i>Kdm5c</i> in males and females	20
3	Male and female <i>Kdm5c</i> mutant epiblast-like cells express master regulators of germline identity	29
4	KDM5C binds to and represses germline gene promoters during embryogenesis	31
5	Misexpression of germline genes in the male <i>Kdm5c</i> -KO postnatal day 6 (P6) brain and liver	34
6	Retinoic acid promotes the prolonged misexpression of germline genes in <i>Kdm5c</i> -KO epiblast-like cells	36
7	Dysregulation of imprinted genes in <i>Kdm5c</i> mutant males and females . . .	45
8	Aberrant expression of piRNA regulatory factors and transposable elements in <i>Kdm5c</i> mutant male and female EpiLCs	47
9	Additional heterozygous loss of <i>Dazl</i> ameliorates the transcriptional dysregulation in the <i>Kdm5c</i> -KO male cortex and hippocampus	49

Abstract

Many male-biased neurodevelopmental disorders (NDDs), including intellectual disability (ID) and autism spectrum disorders (ASDs), are due to or associated with mutations in chromatin regulators. However, it is unclear why chromatin regulation is so essential for neurodevelopment and if this male-bias is due to male-specific functions. While this relationship may be partially explained by their influence upon neuron-specific genes, loss of some chromatin regulators can also lead to the misexpression of tissue-specific genes with unknown consequences. My dissertation work aims to explore the impact of sex and cellular identity in chromatin-linked NDDs by investigating sexually dimorphic transcription and germline gene misexpression in mice lacking lysine demethylase 5c (KDM5C). Loss of KDM5C in male and female mice and humans results in mild to severe intellectual disability and aberrant aggression, yet it is unclear why KDM5C is required for neurodevelopment as it is ubiquitously expressed. KDM5C is thought to suppress gene expression by erasing histone 3 lysine 4 di and tri-methylation (H3K4me2/3), marks enriched at active gene promoters. mRNA sequencing (mRNA-seq) of the male *Kdm5c*-knockout (-KO) brain surprisingly revealed widespread misexpression of germline genes that have no known function in somatic tissues like the brain. It is currently unclear how this ectopic expression of germline genes arises, if they lead to functional consequences, if this phenotype is sex-specific, and whether this apparent soma-to-germline transition contributes to *Kdm5c* mutant neuronal impairments.

To elucidate the shared and sex-specific consequences of *Kdm5c* loss upon neurodevelopment, we performed mRNA-seq in the *Kdm5c* mutant male and female postnatal day 6 (P6)

hippocampus and cortex. We found many sex-specific gene expression changes, including genes relevant to neurodevelopment that may influence behavioral phenotypes. We additionally found many genes dysregulated in both *Kdm5c* mutant male and female brains, including testis and ovary-enriched germline genes. These data indicate male and female somatic cells lacking *Kdm5c* fail to decommission germline identity during early development. We then assessed germline gene misexpression in male and female *Kdm5c* mutant epiblast-like cells (EpiLCs), an *in vitro* model of when germline genes are typically decommissioned. Unlike in the mature brain, *Kdm5c* mutant EpiLCs expressed master regulators of germline identity, including *Dazl*, and their misexpression was exacerbated by the presence of retinoic acid. These data, in conjunction with previous work, suggest germline gene misexpression in the mature *Kdm5c* mutant brain is due to failure to suppress master regulators early in life. In support of this, we found additional heterozygous loss of *Dazl* (*Kdm5c*-KO;*Dazl*-HET) ameliorated the transcriptional dysregulation observed in the *Kdm5c*-KO brain, including rescue of non-germline genes. We additionally found early evidence that germline gene misexpression leads to germline-like phenotypes in *Kdm5c*-KO somatic cells, such as the dysregulation of genomic imprinting and transposable element expression. These preliminary phenotypes offer promising avenues by which the somatic acquisition of germline identity contributes to *Kdm5c*-KO neurodevelopmental impairments. Ultimately, this work provides insight into the molecular demarcation of somatic and germline identity and offers a novel route by which chromatin regulators are essential for proper neurodevelopment in both sexes.

Chapter 1 - Introduction

Shared and sexually dimorphic features of neurodevelopmental disorders.

Neurodevelopmental disorders (NDDs) are broadly classified as disorders that manifest during early development, resulting in varying impairments in social and intellectual functions¹. As of 2020, approximately 2.9% of children are diagnosed with autism spectrum disorder (ASD), 1.4% diagnosed with intellectual disability (ID), and 6.4% diagnosed with learning disability in the United States². Although highly prevalent, there are currently very few options available to treat, care for, and support affected individuals.

It has become increasingly apparent that sex can influence the prevalence and manifestation of NDDs. Males overall are more likely to be diagnosed with a NDD than females, approximately 1.5-fold higher for ID and 4-fold higher for ASD^{3,4}. Biological differences between individuals assigned female or male at birth, such as sex chromosomes and hormones, may contribute to the male-biased prevalence of NDDs. For example, the X chromosome harbors many genes that can cause intellectual disability when mutated, accounting for approximately 16% of all males diagnosed with ID⁵. However, this discrepancy in NDD prevalence may be partially due to sexually dimorphic manifestations of NDD phenotypes that can hinder female diagnosis. For instance, females with ASD are more likely to internalize problems when conflicts arise (e.g. anxiety, depression) while males are more likely to externalize their feelings (e.g. aggression), increasing the likelihood male children are

brought in for testing⁶. Females with ASD also frequently self-report “camouflaging” their social phenotypes by learning and mimicking the social behaviors of their peers⁶. Failure to diagnose females is additionally compounded by the fact that diagnostic criteria for NDDs are primarily based on male patients. More recently studies on other NDDs have also reported sex differences in phenotypic frequency, such as higher likelihood for macrocephaly in males but higher rates of microcephaly in females⁷. However, research in this area overall has been lacking given that many NDD studies are still exclusively conducted in males and the National Institute of Health only required sex to be addressed in research proposals since 2016. Therefore, there is a pressing need to characterize female pathology and phenotypes in NDDs, as these differences will ultimately inform patient outcomes and intervention strategies.

Chromatin regulators, neurodevelopmental disorders, and cellular identity

As next generation sequencing technologies improved, studies unexpectedly found many NDDs are either caused by or linked to mutations in genes that regulate chromatin^{8–10}. Chromatin is the 3D structure of our genome, consisting of DNA wrapped around octamers of histone proteins¹¹. Histone proteins can undergo covalent chemical modifications on their N-terminal tails that can in turn influence DNA accessibility and gene expression^{12,13}. For example, histone 3 lysine 4 (H3K4) can be mono-, di- or tri- methylated (H3K4me1-3), and these marks are enriched at the enhancers and promoters of active genes. This dissertation focuses on H3K4me3, a mark enriched at the transcription start site of active genes^{14,15}. While H3K4me3’s precise role in gene expression is still under debate^{16–18}, there is compelling evidence that it facilitates transcription through recruitment of TFIID - a member of the RNA polymerase preinitiation complex^{19,20}. There is a strong link between H3K4me regulation and NDDs, as mutation of 11 out of 15 genes that read, write, or erase H3K4 mono-, di-, and tri- methylation result in varying neurodevelopmental impairments and syndromic NDDs²¹. Yet it is still unclear why regulation of H3K4 methylation is so crucial for neurodevelopment, as most of H3K4 regulators are ubiquitously expressed across tissues, cell types, and lifespan.

To understand why chromatin regulators are essential for neurodevelopment, many studies have understandably focused on their influence upon neuron-specific transcription, such as genes required for synapse formation and function. Indeed, some chromatin modifiers are now known to have neuron-specific roles in regulating gene expression and chromatin architecture^{22–25}. However, loss of some chromatin regulators can also lead to the misexpression of cell-type specific genes outside of their intended context. In fact, some of the first chromatin regulators, *Trithorax* and *Polycomb*, were initially discovered for their influence on tissue identity through guiding the expression of *Hox* genes during *Drosophila* development^{26–28}. Only later was it discovered *Trithorax* and *Polycomb* are chromatin modifiers - *Trithorax* group proteins place activating histone 3 lysine 4 methylation (H3K4me)^{29–31} while *Polycomb* group proteins install repressive histone 3 lysine 27 methylation (H3K27me) or H2A lysine 119 ubiquitination (H2AK119ub)^{32–34}. Decades of research in model organisms across taxa have made it clear that chromatin regulation is central to Waddington’s proposed ‘epigenetic landscape’, enabling a single genome to encode numerous discrete cellular identities^{35,36}.

As we now view chromatin modifiers as general regulators of transcription, their overarching influence on cellular identity remains underexplored in many fields. While there are some well documented cases of chromatin-linked NDDs with cellular identity defects within the same cell lineage^{37,38}, such as alterations in the neuron-glia developmental transition³⁹ or positional neuronal identity⁴⁰, only a few researchers have explored the misexpression of non-neuronal, tissue-specific genes^{41,42}. The impact of these aberrant transcripts on neurodevelopment and their potential as therapeutic targets for NDDs remains unknown. Furthermore, it is unknown if this erosion of tissue identity would manifest differently between males and females, as even tissues without sexually dimorphic functions show sex-biased gene expression in normal contexts^{43–45}.

Lysine demethylase 5C (KDM5C) offers a window into the influence of sex and cellular identity on neurodevelopment

This dissertation aims to explore the role of sex and cellular identity in chromatin-linked NDDs by characterizing these features in a mouse model of Intellectual Developmental Disorder, X-linked, Syndromic, Claes-Jensen Type (MRXSCJ, OMIM: 300534). Patients diagnosed with MRXSCJ present with varying degrees of neurological impairments, but often display intellectual disability, seizures, aberrant aggression, and autistic behaviors^{46–48}. MRXSCJ was first characterized in the early 2000s^{46,47} and is thought to account for approximately 0.7%–2.8% of all X-linked intellectual disability cases^{49,50}, but many more patients are now being identified as exome sequencing becomes more routine in clinical settings. MRXSCJ is caused by pathogenic mutations in lysine demethylase 5C (*KDM5C*, also known as *SMCX* or *JARID1C*), a chromatin regulator that represses gene expression by erasing histone 3 lysine 4 di- and trimethylation (H3K4me2/3)⁵¹. Pathogenic MRXSCJ mutations are observed throughout *KDM5C* and include nonsense, frameshift, and missense mutations that are predicted to impair KDM5C protein production, chromatin binding, or demethylase activity^{48,52}. However, some patients have missense mutations in unannotated regions with unknown functions⁴⁸ and more research is needed to understand how these variants lead to MRXSCJ phenotypes.

KDM5C is an ideal candidate to investigate the impact of sex and cellular identity in neurodevelopmental disorders, as there is accumulating evidence that KDM5C has both sexually dimorphic functions and is involved in somatic tissue identity. In support of KDM5C having distinct functions in females and males, complete loss of KDM5C in males results in MRXSCJ phenotypes⁵³ while complete loss in females is embryonic lethal⁵⁴. Female lethality may be due to KDM5C’s role in the female-specific epigenetic process of X chromosome inactivation (XCI)⁵⁴, a dosage compensation mechanism in which almost all genes on one X chromosome are silenced^{55,56}. Interestingly, *KDM5C* itself lies on the X chromosome but escapes XCI in mice and humans, resulting in 1.2-2-fold higher *KDM5C* expression in females

compared to males^{57–60}. KDM5C was also recently found to have distinct roles in adiposity⁶¹ and bone mass⁶² in males versus females, indicating KDM5C has sexually dimorphic functions beyond early embryogenesis. The manifestation of MRXSCJ is also sexually dimorphic, as females with heterozygous, pathogenic mutations can either be cognitively normal carriers or display phenotypes that are typically milder than their male counterparts^{48,63}.

Key MRXSCJ phenotypes observed in both male and female patients are recapitulated in mice harboring *Kdm5c* null alleles, in which KDM5C’s catalytic activity and protein stability is disrupted through deletion of exons 11 and 12⁵³. Male mice with hemizygous loss of *Kdm5c* (*Kdm5c* knockout, -KO) display learning impairments, heightened aggression and dominance behaviors, and social impairments^{53,64}. Female mice with heterozygous loss of *Kdm5c* (*Kdm5c*-HET) show impaired social interactions⁶⁵ and milder learning and memory impairments than knockout males⁴². To elucidate why *Kdm5c* loss causes MRXSCJ phenotypes, studies performed next generation RNA sequencing (RNA-seq) in the hippocampus^{42,64} and the amygdala^{53,64}, two brain regions involved in the observed behavioral phenotypes. Surprisingly, they found many of the dysregulated genes are typically only expressed in germ cells within the testis and have no known function in the brain⁴². Ectopic expression of testis genes in the *Kdm5c*-KO brain is potentially due to loss of KDM5C suppression during development, resulting in failure to place repressive DNA methylation at germline gene promoters⁴². However, there is still much unknown about KDM5C’s role in testis-specific gene repression, most pressing is whether ectopic germline transcripts contribute to MRXSCJ phenotypes. Additionally, it is currently unknown if this phenomenon also occurs in females, or if females even display any gene expression changes distinct from males, as the *Kdm5c*-HET female brain has never been sequenced.

This dissertation aims to address these gaps in knowledge by investigating KDM5C’s roles in sexually dimorphic transcription and tissue identity. We aim to illuminate how these factors may influence the manifestation of MRXSCJ phenotypes by addressing the following questions: What is the impact of KDM5C loss on the female brain transcriptome? Does KDM5C suppress other tissue-specific genes, particularly female-specific germline genes? Is the misexpression of tissue-specific genes the same across sexes? Does the identity of ectopic germline transcripts change over time? Is ectopic germline gene transcription unique

to the brain, or does it occur in other *Kdm5c* mutant tissues? Does tissue identity and cellular environment influence germline gene misexpression? Finally and ultimately, do ectopic germline transcripts have functional consequences that contribute MRXSCJ-related phenotypes?

I will address these questions in the following chapters: In Chapter 2, I investigate how KDM5C loss has both a shared and sexually dimorphic consequences on the brain transcriptome. I then characterize the ectopic expression of multiple tissue-specific genes within the male and female *Kdm5c* mutant brain, including the misexpression of germline genes. In Chapter 3, I profile aberrant germline transcripts during male and female *Kdm5c* mutant embryogenesis using an *in vitro* cell culture model. I then explore the influence of tissue identity, development, and retinoic acid upon germline gene misexpression. In Chapter 4, I discuss our preliminary results on the functional impact of aberrant germline gene expression in *Kdm5c* mutants, highlighting potential impairments in genomic imprinting and transposable element regulation. I additionally show preliminary data indicating additional genetic deletion of *Dazl*, a master regulator of germline identity, partially rescues the *Kdm5c* mutant brain transcriptome. Finally, in Chapter 5 I explore the broader impact of this work, including the evolution of sex and cellular identity, related neurodevelopmental disorders, and other potential consequences of ectopic germline expression.

Chapter 2 - Materials and Methods

Mice

As previously described⁵³, mice harboring the *Kdm5c*-KO allele were generated by Cre-mediated deletion of exons 11 and 12. For all postnatal day 6 (P6) RNA sequencing experiments, the *Kdm5c*-KO allele was maintained on a mixed background of C57BL/6J and 129S1/SvImJ. Samples for sequencing were collected across 3 litters. Double mutant male mice harboring both knockout of *Kdm5c* and heterozygous loss of *Dazl* (*Kdm5c*-KO;*Dazl*-HET) were generated by crossing female mice heterozygous for *Kdm5c* loss with male mice heterozygous for *Dazl* loss that were obtained from Jackson Laboratory (Jax Strain #023802). Sex was determined through genotyping primers for *Uba1* on the X and Y chromosomes with the following primers: 5'-TGGATGGTGTGGCCAATG-3', 5'-CACCTGCACGTTGCCCTT-3'. Deletion of *Kdm5c* was determined through the primers 5'-ATGCCCATATTAAGAGTCCCTG-3', 5'-TCTGCCTTGATGGGACTGTT-3', and 5'-GGTTCTCAACACTCACATAGTG-3'. Deletion of *Dazl* was determined through the primers 5'-TGCCAAACTCTTGGAATGA-3', 5'-ACGTGGCTGCACATGATAAG-3', and 5'-CGCCTTCTTGACGAGTTC-3'. All mouse studies complied with the protocols (PRO00010350) by the Institutional Animal Care & Use Committee (IACUC) of The University of Michigan.

Cell culture

We utilized our previously established cultures of male and female wild-type and *Kdm5c* mutant embryonic stem cells⁵⁴. Sex was confirmed by genotyping *Uba1* on the X and Y chromosomes with the following primers: 5'-TGGATGGTGTGGCCAATG-3', 5'-CACCTGCACGTTGCCCTT-3'. Deletion of *Kdm5c* was confirmed through the primers 5'-ATGCCCATATTAAGAGTCCCTG-3', 5'-TCTGCCTTGATGGGACTGTT-3', and 5'-GGTTCTCAACACTCACATAGTG-3'.

Embryonic stem cells (ESCs) were initially cultured in primed ESC (pESC) media consisting of KnockOut DMEM (Gibco#10829-018), fetal bovine serum (Gibco#A5209501), KnockOut serum replacement (Invitrogen#10828-028), Glutamax (Gibco#35050-061), Anti-Anti (Gibco#15240-062), MEM Non-essential amino acids (Gibco#11140-050), and beta-mercaptoethanol (Sigma#M7522). They were then transitioned into ground-state “naive” ESCs (nESCs) by culturing for four passages in N2B27 media containing DMEM/F12 (Gibco#11330-032), Neurobasal media (Gibco#21103-049), Gluamax, Anti-Anti, N2 supplement (Invitrogen#17502048), and B27 supplement without vitamin A (Invitrogen#12587-010), and beta-mercaptoethanol. Both pESC and nESC media were supplemented with the GSK3 inhibitor CHIR99021 (Sigma #SML1046-5MG), the MEK inhibitor PD0325901 (Sigma #PZ0162-5MG), and leukemia inhibitory factor (LIF, Millipore#ESG1107).

nESCs were differentiated into epiblast-like cells (EpiLCs, 48 hours) and extended EpiLCs (exEpiLCs, 96 hours) by culturing in N2B27 media containing DMEM/F12, Neurobasal media, Gluamax, Anti-Anti, N2 supplement, B27 supplement (Invitrogen#17504044), beta-mercaptoethanol, fibroblast growth factor 2 (FGF2, R&D Biotechne 233-FB), and activin A (R&D Biotechne 338AC050CF). EpiLC media made with B27 without vitamin A (Invitrogen#12587-010) was used to test the impact of retinoic acid on germline gene expression.

Immunocytochemistry of DAZL in EpiLCs was performed by washing cells thrice with phosphobuffered saline (PBS), fixing cells in 4% paraformaldehyde, washing thrice in PBS, blocking in PBS containing 0.3% Triton X-100, and 5% fetal bovine serum for 1 hour, washing thrice with PBS, and incubating in primary antibody (Rabbit anti DAZL, abcam ab34139,

1:200) in the blocking solution overnight at 4C. The next day cells were rinsed thrice with PBS, incubated in secondary antibody (Alexafluor 488 Invitrogen #710369, 1:1000) in blocking buffer, rinsed thrice in PBS, and then imaged. Images were taken blinded for genotype, chosen based on similar levels of DAPI signal, and then quantified via ImageJ.

RNA sequencing

Published datasets

All published datasets are available at the Gene Expression Omnibus (GEO) <https://www.ncbi.nlm.nih.gov/geo>. Previously published RNA sequencing datasets analyzed in this study included the male *Kdm5c*-KO adult amygdala and hippocampus⁶⁴ (available at GEO: GSE127722) and male and female *Kdm5c* mutant EpiLCs⁵⁴ (available at GSE: GSE96797). We have also published the the male and female *Kdm5c* mutant P6 cortex and hippocampus RNA-seq results regarding sexually dimorphic gene expression⁶⁵ and the data is available at GEO: GSE206346

RNA isolation and sequencing

For all P6 mouse RNA-seq experiments, we dissected out the liver and the brain and then isolated the cortex and hippocampus through micro-dissection. For each animal study we used three to four animals per genotype across at least 3 litters. Total RNA was purified and DNase-treated by the Qiagen RNEasy Midi Kit (Qiagen #75144) after homogenizing tissue samples in Buffer RLT with an electric homogenizer. Purified RNA was sent to Novogene for sample quality control and cDNA library preparation by poly A enrichment. Libraries were prepared using the NEBNext Ultra™ II Directional RNA Library Prep Kit with oligo-dT priming and sequenced through the Illumina NovaSeq 6000 sequencing platform to generate paired-end 150 base-pair reads. For ESC, EpiLC, and exEpiLC cell culture experiments, total RNA was purified and DNase-treated by the Qiagen RNEasy Midi Kit and 151 basepair, poly A-enriched libraries were prepared for Illumina NovaSeq 6000 sequencing. Library preparation and next-generation sequencing for cell culture experiments was carried out in

the Advanced Genomics Core at the University of Michigan.

Alignment and analysis

After ensuring read quality via FastQC (v0.11.8), reads were then mapped to the mm10 *Mus musculus* genome (Gencode) using STAR (v2.5.3a), during which we removed duplicates and kept only uniquely mapped reads. Count files were generated by FeatureCounts (Subread v1.5.0), and BAM files were converted to bigwigs using deeptools (v3.1.3) and visualized by the UCSC genome browser. RStudio (v3.6.0) was then used to analyze counts files by DESeq2 (v1.26.0)⁶⁶ to identify differentially expressed genes (DEGs) with a q-value (p-adjusted via FDR/Benjamini–Hochberg correction) less than 0.1 and a log2 fold change greater than 0.2. For the male and female P6 experiment, the DESeq2 design included a grouping variable to test the interaction between sex and genotype. For the P6 *Kdm5c*-KO;*Dazl*-HET double mutant experiment, DESeq2 design was based on genotype. Batch effect corrections were not required for these two experiments as samples were purified together. For the ESC, EpiLC, and exEpiLC experiment, the DESeq2 design included batch effect correction and grouping variable combining genotype, time point, and RA treatment. For all DESeq2 analyses, log2 fold changes were calculated with lfcShrink using the ashR package⁶⁷. MA-plots were generated by ggpubr (v0.4.0), and Eulerr diagrams were generated by eulerr (v6.1.1). Boxplots and scatterplots were generated by ggpubr and ggplot2 (v3.3.2). Heatmaps of gene expression were generated using the base R functions scale and hclust and visualized using the R package ComplexHeatmap. The Upset plot was generated via the package UpSetR (v1.4.0)⁶⁸.

To quantify transposable element expression, reads were first aligned to mm10 using STAR (v2.5.3a) with the flags “`–winAnchorMultimapNmax 200 –outFilterMultimapNmax 100`” to enable multiple alignments. Transposable element expression compared to wild-type controls was then assessed through TETranscripts (v2.2.3)⁶⁹ using the settings “`–mode multi –minread 1 –i 10 –padj 0.1`”.

Chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq)

We analyzed our previously published KDM5C ChIP-seq experiments in primary neuron cultures from the cortex and hippocampus⁵³ (available at GEO: GSE61036) and EpiLCs⁵⁴ (available at GEO: GSE96797). ChIP-seq of histone 3 lysine 4 dimethylation in male EpiLCs⁵⁴ is also available at GEO: GSE96797. ChIP-seq of histone 3 lysine 4 trimethylation in the male amygdala⁶⁴ are available at GEO: GSE127817.

ChIP-seq reads were aligned to mm10 using Bowtie1 (v1.1.2) allowing up to two mismatches. Only uniquely mapped reads were used for analysis. Peaks were called using MACS2 software (v 2.1.0.20140616) using input BAM files for normalization, with filters for a q-value < 0.1 and a fold enrichment > 1 . Common peak sets were obtained via DiffBind, and count tables for the common peaks were generated with the Bedtools multicov command. We removed “black-listed” genomic regions that often give aberrant signals. Enriched motifs were identified using HOMER⁷⁰. Gene ontology (GO) analyses were performed by Metascape⁷¹ at <http://metascape.org>

Classifying tissue-enriched and germline-enriched genes

Tissue-enriched differentially expressed genes were determined by evaluating their classification in a previously published list of genes enriched in 17 male and female mouse tissues⁷². This study defined expression as greater than 1 Fragments Per Kilobase of transcript per Million mapped read (FPKM) and tissue enrichment as at least 4-fold higher expression than any other tissue.

We curated a list of germline-enriched genes using an RNA-seq dataset from wild-type and germline-depleted male and female mouse embryos from embryonic day 12, 14, and 16 (unpublished data provided by Dr. Jacob Mueller, University of Michigan). Germline-enriched genes met the following criteria: 1) their expression is greater than 1 FPKM in wild-type germline 2) their expression in any wild-type somatic tissues⁷² does not exceed 20% of maximum expression in wild-type germline, and 3) their expression in the germ cell-depleted

(W/W^v) germline, for any sex or time point, does not exceed 20% of maximum expression in wild-type germline. We then categorized germline genes as either female-biased (XX) or male-biased (XY) by filtering for genes whose expression in the opposite sex, at any time point, is no greater than 20% of the gene's maximum expression in a given sex. Germline genes that did not meet these criteria for sex-biased expression were classified as unbiased.

Chapter 3 - Shared and Sexually Dimorphic Gene Expression Changes in the Brain with *Kdm5c* Loss

Introduction

Intellectual developmental disorder, X-linked, syndromic, Claes-Jensen Type (MRXSCJ, OMIM: 300534) is caused by deleterious mutations in the chromatin regulator *KDM5C*. When MRXSCJ was first described in the early 2000s, the majority of patients were males who inherited causative mutations from cognitively normal carrier mothers^{46,47}. However, many more women with MRXSCJ impairments have been recently identified^{48,63,73} as rapid improvements in sequencing technologies now allow for more routine exome sequencing. Women heterozygous for deleterious *KDM5C* mutations can be unaffected carriers or present with a wide spectrum of phenotypes, such as mild memory impairments, developmental delay/intellectual disability, aberrant aggression, impaired language expression, and anxiety. Most females are reported to have milder phenotypes than males, potentially due to skewed X chromosome inactivation preferentially silencing the deleterious allele⁷⁴.

Currently, very little is known about the biological phenomena that influences how MRXSCJ pathology manifests in males and females. Interestingly, there is growing evidence that *KDM5C* itself has sexually dimorphic functions (see Chapter 1) but it is unclear to

what extent this is true in the brain. Patient data suggests at least some females require more KDM5C than males for neurodevelopment, as heterozygous loss in females can result in neuronal impairments, even though their KDM5C expression would be roughly equivalent to unaffected males. Of note, it has been proposed KDM5C's Y chromosome homologue, KDM5D, compensates for the lower KDM5C dosage in unaffected males as their structures are highly similar⁵¹. Nevertheless, their functions are likely divergent in the brain, as KDM5D is insufficient to protect male MRXSCJ patients from deleterious *KDM5C* mutations and KDM5D loss has never been linked to neurodevelopmental impairments. Although there is accumulating evidence KDM5C is differentially regulated in females versus males, it remains unclear if KDM5C has sex-specific roles in neurodevelopment, as RNA sequencing has never been performed in the *Kdm5c* mutant female brain. Uncovering these differences is crucial to understanding how MRXSCJ manifests between the sexes and may ultimately inform sex-specific intervention strategies.

A common model for investigating why KDM5C loss leads to neuronal impairments is mice harboring *Kdm5c* loss-of-function mutations that recapitulate key phenotypes of MRXSCJ⁵³. Surprisingly, mRNA sequencing (mRNA-seq) of the male adult *Kdm5c* knockout (-KO) brain revealed aberrant transcription of many genes that are typically only expressed in the germ cells of the testis (e.g. sperm)⁴². This opens an intriguing possibility that MRXSCJ phenotypes are, in part, caused by impaired development of neuronal identity (see Chapter 5 and Conclusion). The impact of this phenomenon and whether it occurs in a sex-specific manner, such as females expressing ovary-enriched genes, is unknown. Furthermore, whether *Kdm5c* mutants express other tissue-specific genes or if aberrant expression is strictly limited to the testis germline remains to be characterized.

To elucidate shared and sexually dimorphic consequences of KDM5C loss upon neurodevelopment, we performed bulk mRNA sequencing (mRNA-seq) in the postnatal day 6 (P6) hippocampus and cortex of *Kdm5c* mutant male and female mice. Of note, we compared *Kdm5c* knockout males (*Kdm5c*-KO) and *Kdm5c* heterozygous females (*Kdm5c*-HET), as these reflect patient genotypes since complete loss of *Kdm5c* in females is embryonic lethal⁵⁴ (see Chapter 1). We found *Kdm5c* mutants had both shared and sexually dimorphic gene expression changes within the brain, including novel genes with functions pertinent to MRXSCJ

pathology that were not previously identified in adult males. *Kdm5c*-HET females had unique changes in early developmental genes while knockout males had unique changes in genes related to memory and aggression. In both male and female *Kdm5c* mutants, we observed the ectopic transcription of many tissue-specific genes, the majority being testis-enriched germline genes. Interestingly, *Kdm5c* mutant males and females also aberrantly expressed ovary-enriched germline genes, suggesting KDM5C's role in germline gene repression is independent of organismal and gonadal sex. These results provide novel insight into the development of male and female MRXSCJ pathology and characterize KDM5C's role in maintaining tissue-specific gene expression across sexes.

Shared and sexually dimorphic alterations in gene expression within the cortex and hippocampus with loss of *Kdm5c*

To assess the role of KDM5C in neurodevelopment across sexes, we performed next-generation mRNA sequencing in the male *Kdm5c* hemizygous knockout (XY *Kdm5c*-KO), female *Kdm5c* heterozygous knockout (XX *Kdm5c*-HET), male wild-type (XY WT), and female wild-type (XX WT) mouse brain. Currently all *Kdm5c*-KO brain sequencing and phenotyping has been performed in adulthood^{42,53,64}. Thus, it is unknown if KDM5C loss causes transient developmental changes that contribute to phenotypes later in life, such as the reduction in dendritic spine density observed in the *Kdm5c*-KO adult male hippocampus and amygdala^{53,64}. Therefore, we chose to assess gene expression changes in the postnatal day 6 (P6) hippocampus and cortex since high rates of synaptogenesis and synaptic pruning occur during this period⁷⁵.

We first evaluated sex-specific expression of *Kdm5c* in the wild-type brain and found females expressed 1.489-fold higher *Kdm5c* than males ($p = 1.2 \times 10^{-5}$, T-test, XY WT versus XX WT), consistent with previous reports in other tissues⁵⁷⁻⁶⁰ (Figure 1A). We next evaluated the impact of *Kdm5c* loss on the male and female transcriptome by identifying genes that were differentially expressed from their wild-type counterpart with DESeq2⁶⁶ ($q < 0.1$, log2 fold change > 0.2). In both *Kdm5c*-KO males and *Kdm5c*-HET females, the majority of

differentially expressed genes (DEGs) were upregulated compared to wild type expression, consistent with KDM5C’s proposed role in transcriptional repression⁵¹ (Figure 1B). *Kdm5c*-KO males had an overall greater number of DEGs than *Kdm5c*-HET females and while some genes showed significant changes in only one sex, most female upregulated DEGs were also upregulated in males. To assess the impact of sex on the magnitude of a gene’s dysregulation, we compared the log2 fold change (log2FC) from wild type of upregulated DEGs in males versus females. We found that for both males and females, the DEGs with the greatest log2FC were those significantly dysregulated in both sexes (Figure 1C, left). However, *Kdm5c*-HET females exhibited approximately 40% of the extent of male dysregulation for these shared DEGs (Figure 1C, right).

In comparison to our previous work in the adult brain⁶⁴, we found many novel genes dysregulated in the male *Kdm5c*-KO P6 hippocampus and cortex. Indeed, even with identical experimental procedures, genetic mutation, and analytical pipeline, the number of *Kdm5c*-KO DEGs in the P6 cortex and hippocampus is over 3-fold larger than the adult hippocampus (Male adult DEGs: 298, P6 DEGs: 1079). This suggests the brain’s transcriptome may be particularly vulnerable to KDM5C loss during early stages of neurodevelopment. DEGs newly identified in P6 included *Gap junction protein (Gjb1)*, which had the greatest log2FC from wild type in both male and female *Kdm5c* mutants (XX: Log2FC = 1.169; padj = 4.01e-37, XY: Log2FC = 1.96; padj = 1.92e-82) (Figure 1D). *Gjb1* encodes the protein connexin-32, whose expression increases across postnatal cortical development⁷⁶ to form gap junctions in oligodendrocytes and neurons⁷⁷ and is implicated in epilepsy⁷⁸, a phenotype commonly observed in male and female MRXSCJ patients^{46–48}.

In addition to common gene expression changes, we also found 847 male-specific DEGs and 20 female-specific DEGs. Many male-specific DEGs overexpressed in *Kdm5c*-KOs have known roles in aggression and memory. One example is *phenylethanolamine N-methyltransferase (Pnmt)* (Log2FC = 0.849; padj = 3.18e-15), an enzyme that methylates norepinephrine to form epinephrine⁷⁹ and causes extreme aggression when overexpressed in male mice⁸⁰ (Figure 1E). *Kdm5c*-KO males also overexpressed *galanin (Gal)* (Log2FC = 0.373; padj = 0.00214), a neuropeptide that can inhibit excitatory neurotransmission⁸¹ and whose overexpression is linked to memory impairments^{82–84} (Figure 1F).

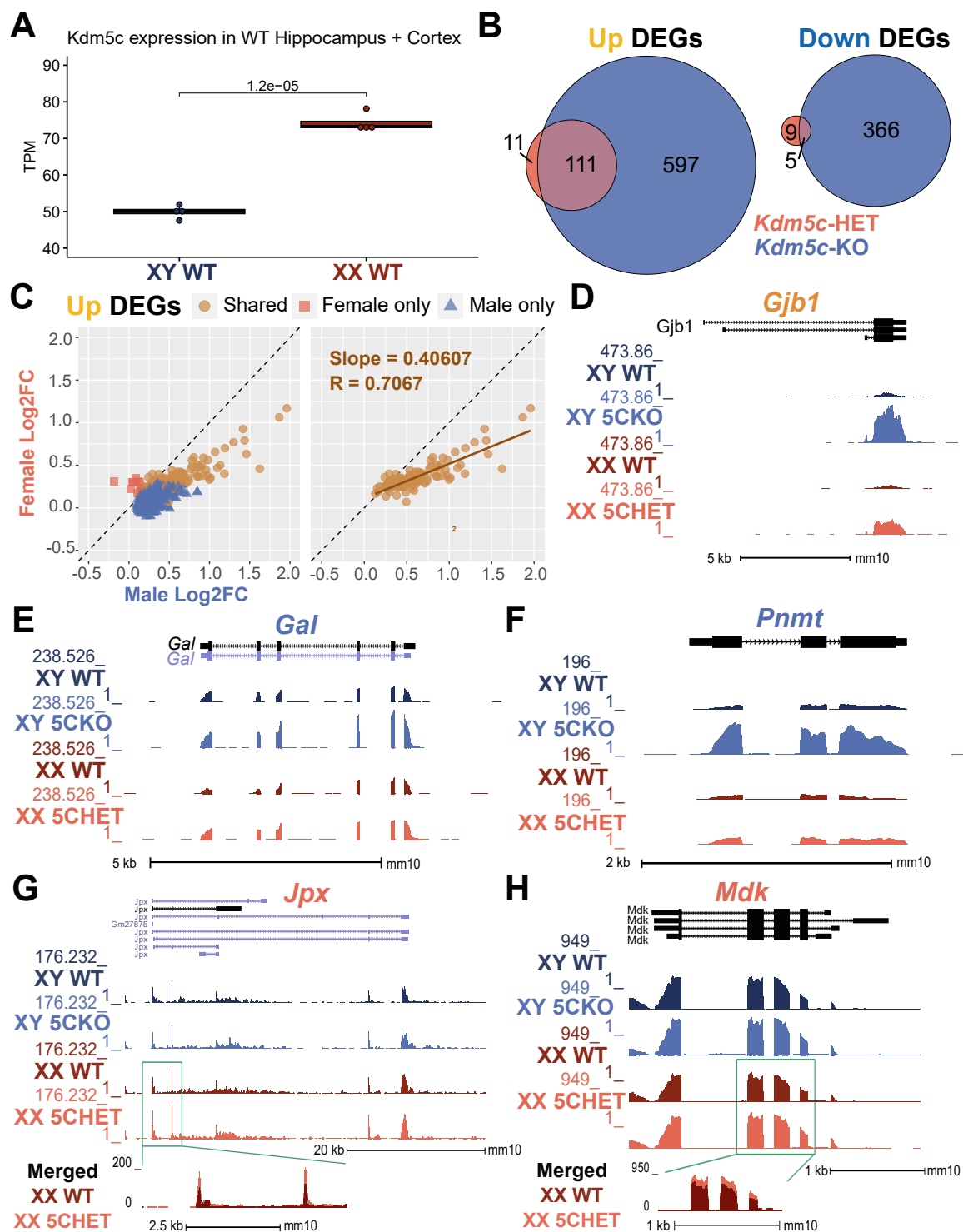


Figure 1: *Kdm5c* loss produces both shared and sexually dimorphic gene expression changes in the hippocampus and cortex. **A.** *Kdm5c* expression in transcripts per million (TPM) in the wild-type male (XY WT) and female (XX WT) brain. **B.** Overlap of upregulated (Up) and downregulated (Down) differentially expressed genes (DEGs), with *Kdm5c* mutant male DEGs in blue and female DEGs in pink. **C.** Left — Scatter plot of Up DEGs comparing their log2 fold change from wild type in males (x-axis) versus females (y-axis). Right—Same as left but displaying only shared DEGs. **D.** Average bigwigs of a DEG shared between males and females, *Gjb1*. **E-F.** Average bigwigs of male-specific DEGs, *Gal* and *Pnmt*. **G-H.** Average bigwigs of female-specific DEGs, *Jpx* and *Mdk*.

Contrastingly, *Kdm5c*-HET females uniquely displayed slight but significant overexpression of genes important for early development. One DEGs of interest is *JPX transcript*, *XIST activator* (*Jpx*) (Log2FC = 0.297; padj = 0.0879), a long non-coding RNA (lncRNA) crucial for X chromosome inactivation (XCI), a females-specific epigenetic process that occurs during early embryogenesis^{85,86} (Figure 1G). Of note, *Jpx* is typically expressed in both males and females, but only promotes XCI in females through induction of *Xist* expression^{85,86}. Another female-specific DEG, *midkine* (*Mdk*) (Log2FC = 0.305; padj = 0.0281), is a neurotrophic factor expressed transiently during embryogenesis to promote neurite outgrowth^{87,88} (Figure 1H).

Altogether, these data demonstrate that loss of *Kdm5c* has both shared and sexually dimorphic consequences on brain transcription during neurodevelopment. We identified novel genes dysregulated during male *Kdm5c*-KO development compared to previous studies in adulthood that are relevant to MRXSCJ phenotypes. We additionally found heterozygous loss of *Kdm5c* in female mice is sufficient to cause readily detectable gene expression changes, even though they have milder behavioral phenotypes than male knockouts⁴². We found many *Kdm5c* mutant DEGs are shared between males and females but have a milder change from wild type in females. While *Kdm5c* mutant females displayed unique upregulation of early developmental genes, mutant males uniquely upregulated aggression and memory-related genes. These sexually dimorphic responses to KDM5C loss may ultimately contribute to how MRXSCJ manifests in male versus female patients, such as contributing to higher symptom severity in males.

Germline genes are aberrantly transcribed in the brain of male and female *Kdm5c* mutants

Previous sequencing in the adult male *Kdm5c*-KO brain surprisingly identified many genes that are typically only expressed in germ cells of the male testis⁴². It is currently unknown whether *Kdm5c* mutants express other tissue-specific genes and if germline gene de-repression is occurs in a sex-specific manner. We first broadly assessed de-repression of tissue-specific

genes by evaluating whether DEGs upregulated in *Kdm5c*-KO males or *Kdm5c*-HET females overlapped with mRNAs enriched in 17 male and female mouse tissues⁷². This study defined tissue-specific enrichment as genes with more than 4-fold higher expression in a given tissue over any other tissues. We found loss of *Kdm5c* in both males and females led to de-repression of many tissue-enriched genes, with the majority of male and female DEGs belonging to the testis (Figure 2A-B). Even though the testis has the largest number of tissue-specific genes compared to any other tissue (2496 genes), testis enrichment was significant for *Kdm5c*-KO male DEGs, although not female DEGs (P6 XY $p = 2.09 \times 10^{-4}$; P6 XX $p = 0.37$, Fisher's Exact). Testis genes were also significantly enriched in our previously published mRNA sequencing of the mature *Kdm5c*-KO male hippocampus and amygdala (Adult XY Amygdala $p = 2.67 \times 10^{-5}$; Adult XY Hippocampus $p = 1.1 \times 10^{-11}$, Fisher's Exact, data not shown). Of note, we did not observe enrichment of brain-specific genes in any of our *Kdm5c* mutant brain samples (P6 XY $p = 0.136$; P6 XX $p = 0.186$; Adult XY Amygdala $p = 0.299$; Adult XY Hippocampus $p = 0.564$, Fisher's Exact), even though the brain has the second highest total number of tissue-enriched genes (708 genes). Together, these results suggest the high number of testis genes expressed in the *Kdm5c* mutant brain are not due to random de-repression of all genes. Many of *Kdm5c* mutant testis-enriched DEGs have functions specific to male germ cell development, such as *Transcription factor-like 5 (Tcfl5)*, a transcription factor that promotes the meiotic gene expression in males⁸⁹ (Figure 2C-D). Interestingly, we also found a few ovary-enriched genes expressed in *Kdm5c* mutant males and females, such as *Zygotic arrest 1 (Zar1)* who was recently shown to sequester mRNAs in oocytes⁹⁰ (Figure 2E-F). Currently, these testis and ovary-enriched genes have no known role in neurodevelopment, even though some show weak expression in the wild-type brain (Figure 2C-F).

The testis contains both germ cells (e.g. spermatogonia) that undergo meiosis for reproduction as well as somatic cells (e.g. Leydig cells) that support hormone production and germline functions. To validate if these testis DEGs are truly germline genes, we first compared their expression in wild-type versus germ cell-depleted (W/Wv) mouse testes⁹¹ (Figure 2G). We plotted all testis-enriched DEGs identified in our adult male *Kdm5c*-KO brain sequencing along with our P6 male and female *Kdm5c* mutant testis DEGs. We found almost all testis-enriched DEGs lose their expression with germ cell depletion (Figure 2G). We

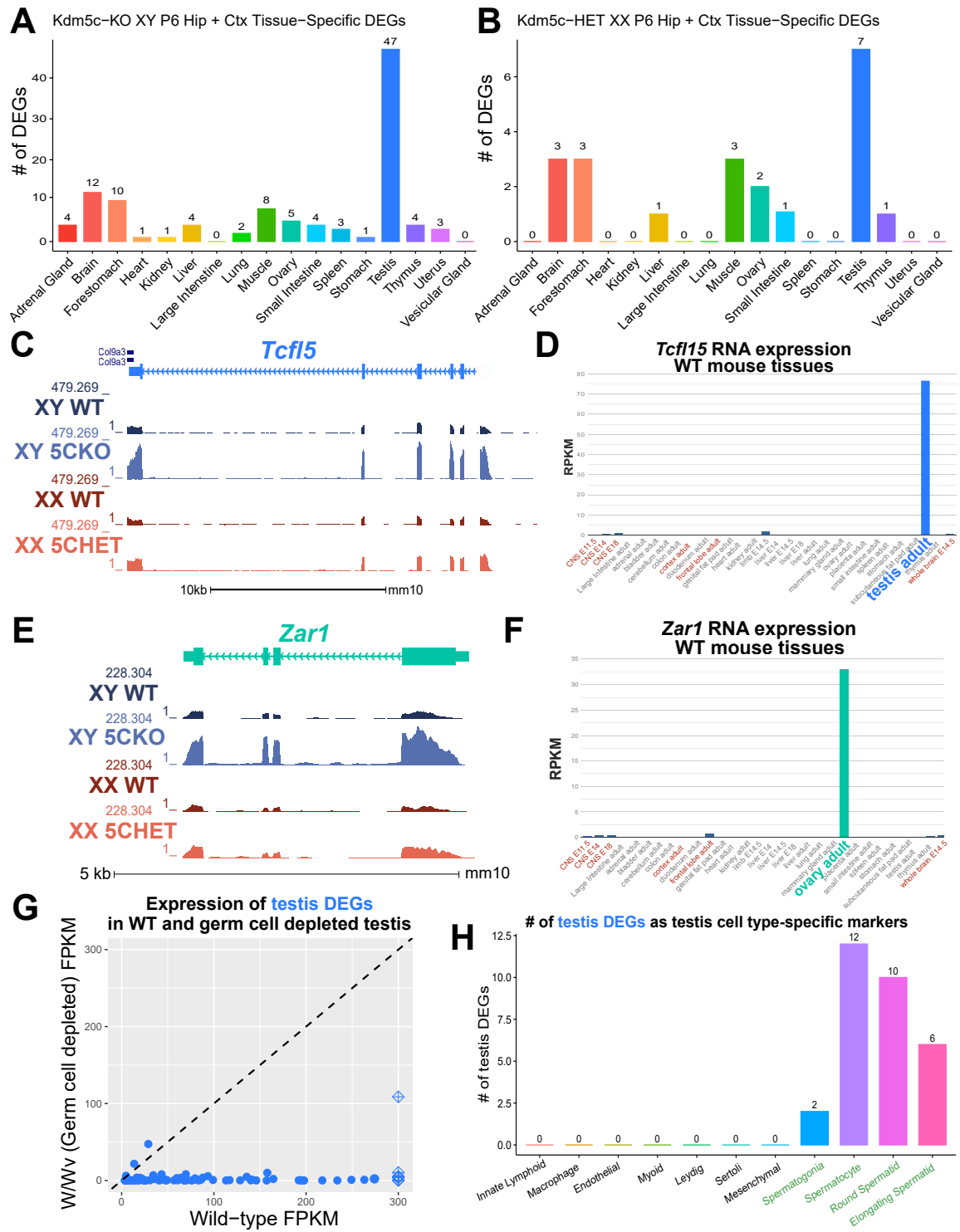


Figure 2: Misexpression of tissue specific genes, including germline genes, with loss of *Kdm5c* in males and females. A-B. The number of tissue-enriched differentially expressed genes (DEGs) in the male (A) and female (B) *Kdm5c* mutant postnatal day 6 (P6) cortex and hippocampus. **C.** Bigwigs of an example aberrantly expressed testis-enriched gene, *Tcf15*. **D.** Expression of *Tcf15* in wild-type tissues from NCBI Gene, with testis highlighted in blue and brain tissues highlighted in red. (Legend continued on next page.)

Figure 2: Misexpression of tissue specific genes, including germline genes, with loss of *Kdm5c* in males and females. (Legend continued.) **E.** Bigwigs of an example ovary-enriched germline DEG, *Zar1* **F.** Expression of *Zar1* in wild-type tissues from NCBI Gene, with ovary highlighted in teal and brain tissues highlighted in red. **G.** Expression of testis DEGs from the *Kdm5c* mutant adult male hippocampus, adult male amygdala, P6 male hippocampus and cortex, and P6 female hippocampus and cortex in wild-type versus germ cell depleted testis. Expression is in Fragments Per Kilobase of transcript per Million mapped reads (FPKM) **H.** Number of testis DEGs that were classified as cell-type specific markers in a single cell RNA-seq dataset of the testis. Germline cell types are highlighted in green, somatic cell types in black.

additionally assessed testis DEG expression in a published testis single cell RNA-seq dataset that identified cell type-specific markers⁹². We found some *Kdm5c* mutant testis DEGs were classified as specific markers for multiple stages of germ cell development - including spermatogonia, spermatocytes, round spermatids, and elongating spermatids - yet none marked somatic cells (Figure 2H). Misexpression of germline genes appears to be independent of germ cell sex, as male and female *Kdm5c* mutant brains also express some ovary-enriched germline genes, even though there are relatively few ovary genes (100 genes) compared to the testis (2496 genes). Ovary genes included the aforementioned oocyte translational regulator *Zygote arrest 1 (Zar1)*⁹⁰ that is important for silencing ovarian transcription until postfertilization (Figure 2E-F). Therefore, we concluded both male and female *Kdm5c* mutant brains aberrantly express germline genes that have functions specific to both the male and female germline.

Defining germline-enriched genes

To facilitate downstream analyses, we generated a list of male and female germline-enriched genes by evaluating gene expression in wild-type and germline-depleted male and female mouse embryos from embryonic day 12, 14, and 16 (unpublished RNA-seq data provided by Dr. Jacob Mueller, University of Michigan). We defined genes as germline-enriched if their expression met the following criteria: 1) their expression is greater than 1 FPKM in wild-type germline 2) their expression in any wild-type somatic tissues does not exceed 20% of maximum expression in wild-type germline, and 3) their expression in the germ cell-depleted (W/Wv) germline, for any sex or time point, does not exceed 20% of their maximum expression in the wild-type germline. These criteria yielded 1287 germline-enriched

genes. To more comprehensively assess the impact of sex on germline gene misexpression, we additionally categorized if germline-enriched genes displayed female-biased (XX), male-biased (XY), or unbiased expression. We defined sex-biased genes as those whose expression in the opposite sex, at any time point, is no greater than 20% of the gene’s maximum expression in a given sex. This yielded 65 XX-biased, 1023 XY-biased, and 199 unbiased germline-enriched genes, which is consistent with the testis overall having a more unique transcriptome than the ovary⁷².

Discussion

In the above work, we characterized gene expression changes in the male *Kdm5c*-KO and female *Kdm5c*-HET brain during neurodevelopment. Our data indicate that the loss of *Kdm5c* leads to both common and distinct gene expression changes between sexes, which may underlie shared and unique manifestations of MRXSCJ phenotypes. We additionally found loss of KDM5C in males and females led to the ectopic expression of many tissue-specific genes in the brain, including the misexpression of both testis and ovary-enriched germline genes in both male and female *Kdm5c* mutants.

This study provides the first RNA sequencing dataset of *Kdm5c* mutant brains during neurodevelopment, enabling us to identify many dysregulated genes that were not previously found in the mature brain. We found the male *Kdm5c*-KO P6 cortex and hippocampus had over 3-fold more DEGs than previous reports in the mature hippocampus, indicating that KDM5C may play an important role during early brain maturation. In support of this, KDM5C’s global protein expression decreases as the brain matures⁹³ and conditional knockout of *Kdm5c* in adult excitatory neurons leads to far milder behavioral and gene expression changes⁴². However, experiments directly assessing the KDM5C’s role over time are warranted, as developmental decrease in KDM5C expression could be explained by the increased ratio of astrocytes to neurons over development and decreased adult hippocampal DEG number could be due to the P6 sequencing including the cortex. Nevertheless, many of the novel DEGs we uncovered in the P6 brain had known functions related to MRXSCJ phenotypes. Out of these genes we highlighted *Gjb1*, the most highly dysregulated gene in both

Kdm5c-KO males and *Kdm5c*-Het females. *Gjb1* encodes the protein connexin-32 that forms gap-junctions in oligodendrocytes and neurons⁷⁷. While loss of *Gjb1* causes the sensorimotor neuropathy Charcot–Marie–Tooth type 1⁹⁴, upregulation of *Gjb1* is associated with seizures, since bicuculline-induced epileptic activity increases *Gjb1* mRNA and connexin-32 protein in the hippocampus⁷⁸. Many male and female MRXSCJ patients have epileptic seizures⁴⁸ and *Kdm5c*-KO male mice have an increased propensity for kainic acid-induced seizures⁴². Further work is required to elucidate if developmental *Gjb1* overexpression promotes seizure propensity or if it is instead a consequence of a seizure-prone state. Additionally, it is currently unknown if a developmental increase in *Gjb1* expression occurs in human MRXSCJ patients.

This study also provided the first RNA sequencing dataset for the *Kdm5c*-HET female brain, enabling us to characterize shared and sexually dimorphic gene expression changes. *Kdm5c*-HET mice overall display milder behavioral phenotypes than knockout males⁴², yet we found readily detectable gene expression changes in the *Kdm5c*-HET brain. Along with many genes dysregulated in both sexes, we also found some female-specific DEGs, such as *Jpx* and *Mdk*, that are involved in early development. Female-specific dysregulation of developmental genes may explain why homozygous loss of *Kdm5c* in females is embryonic lethal and may contribute to previously observed X chromosome inactivation defects⁵⁴. Meanwhile, some *Kdm5c* mutant male-specific DEGs, such as *Pnmt* and *Galanin*, have known roles in aggression and memory - processes often impaired in MRXSCJ patients. *Pnmt* encodes an epinephrine-synthesis enzyme and transgenic overexpression of *Pnmt* in mice results in elevated fighting behaviors in males but no changes in female behavior⁸⁰. *Galanin* encodes a neuropeptide that is thought to influence learning and memory, as it is overexpressed in the limbic system of post-mortem Alzheimer’s disease patients⁹⁵ and *galanin* overexpression in the mouse forebrain impairs spatial learning^{82–84}. *Pnmt* and *Gal* being male-specific DEGs imply that MRXSCJ phenotypes common between the sexes, such as aberrant aggression and learning impairments, may involve distinct mechanisms downstream of *Kdm5c* loss.

In both sexes, we identified widespread dysregulation of tissue-specific genes with *Kdm5c* loss, indicating KDM5C plays a global role in tissue identity. While it is currently unknown if these tissue-specific transcripts are translated into functional protein, they appear to be processed into mature mRNAs, since they are spliced and polyadenylated. We found both sexes

primarily expressed testis-enriched genes that have no known function in the brain, although tissue-specific gene dysregulation was overall weaker in *Kdm5c*-HET females. Misexpression of testis genes was initially described in the adult *Kdm5c*-KO male hippocampus⁴², in which they identified many testis genes with functions specific to male gametogenesis, such as the RNA helicase *DNA segment, Chr 1, Pasteur Institute 1 (D1Pas1)*⁹⁶ and the mitochondrial gene *Cytochrome c, testis (Cyt)*⁹⁷. We confirmed these testis DEGs are in fact germline genes through published RNA sequencing datasets since they are uniquely expressed in germ cells and not somatic cells. Interestingly, we found both *Kdm5c*-HET females and *Kdm5c*-KO males also aberrantly express a few ovary-specific germline genes, such as *Zar1*. These results indicate KDM5C role in germline gene repression is both independent of organismal sex and independent of the gonadal sex in which the genes typically function. However, if this erosion of brain tissue identity contributes to MRXSCJ impairments is currently unknown (See Chapter 5 and Conclusions).

Chapter 4 - Characterization of KDM5C-Mediated Germline Gene Repression in Males and Females

Introduction

While aberrant germline transcripts are found in all *Kdm5c* mutant mouse brain datasets generated thus far^{42,53,64,65}, very little is known about the underlying mechanism. KDM5C likely contributes to germline gene repression during early development, as KDM5C occupancy at germline gene promoters decreases across cellular differentiation and *Kdm5c*-KO neuron cultures fail to suppress germline genes when KDM5C is restored⁴². These data indicate the sustained expression of germline genes within the *Kdm5c* mutant brain is due to a failure to fully demarcate somatic (i.e., non-meiotic cells) and germ cell identity during early embryogenesis.

Germ cells are typically distinguished from somatic cells soon after the embryo implants into the uterine wall^{98,99}. Around embryonic day (E) 6.5 in mice, a subset of epiblast stem cells proximal to the extra-embryonic ectoderm receive bone morphogenic protein signals, such as BMP4, that initiate their development into primordial germ cells (PGCs). While the remaining epiblast differentiates into the ectoderm, mesoderm, and endoderm to form the body's somatic tissues¹⁰⁰, PGCs migrate to the genital ridge which sends signals to commit

PGCs to a germ cell fate. When male and female germ cells arrive at the genital ridge around E11, they express *Dazl*, an RNA binding protein that regulates the translation of germline-specific genes¹⁰¹. Once germ cells express DAZL, retinoic acid can then trigger them to undergo meiosis by turning on the transcription factor *Stra8*^{102–104}. Up until this point, germ cells are bipotential and can become oocytes or spermatocytes, regardless of chromosomal sex. Germ cells become sexually dimorphic around E12.5, when female germ cells initiate meiosis while males undergo mitotic arrest and postpone meiosis until postnatal day (P) 10^{102–104}. This difference in meiotic timing is due to differential regulation of retinoic acid signaling: embryonic male germ cells express the enzyme CYP26B1 to degrade retinoic acid and block the meiotic signal until after birth^{102–104}.

Researchers investigating the molecular mechanisms behind somatic versus germ cell specification often employ an *in vitro* model of early embryogenesis, in which embryonic stem cells (ESCs) are differentiated into post-implantation epiblast-like stem cells (EpiLCs). Chromatin regulators play a key role in decommissioning germline genes as ESCs differentiate into EpiLCs by placing multiple repressive chromatin marks at germline gene promoters. In ESCs, the non-canonical polycomb repressive complex 1.6 (PRC1.6) places repressive histone H2A lysine 119 monoubiquitination (H2AK119ub1)¹⁰⁵ and SET domain bifurcated histone lysine methyltransferase 1 (SETDB1)¹⁰⁶ along with euchromatic histone methyltransferase 1 and 2 (EHMT1/2 also known as GLP/G9a)¹⁰⁵ place repressive histone 3 lysine 9 trimethylation (H3K9me3). When ESCs become EpiLCs, DNA methyltransferase 3A and B (DNMT3A/B) place repressive DNA CpG methylation to enable long-term silencing of germline genes^{106–108}. It has been proposed KDM5C may be important for this DNA methylation placement, as KDM5C's substrate (H3K4me3) is known to impair DNMT3A enzymatic activity¹⁰⁹, DNMT3A binds unmethylated H3K4¹¹⁰, and DNA methylation is eroded at the promoters of at least two germline genes in the *Kdm5c*-KO hippocampus⁴². However, the dynamics behind this interaction and when KDM5C is required have yet to be elucidated.

As aberrant germline gene expression has yet to be characterized during *Kdm5c* mutant embryogenesis, it is currently unclear how KDM5C loss impacts the somatic transcriptome across time. For example, are the same germline genes constitutively dysregulated or do they change depending upon cellular context? It is also unclear if embryonic dysregulation

will be the same between *Kdm5c* mutant males and females, as some germline genes, like *Stra8*, are differentially regulated during male and female embryonic germline development. Furthermore, if germline genes are aberrantly expressed during *Kdm5c* mutant embryogenesis, it is unclear if long-term misexpression is unique to the brain or found throughout somatic tissues. As transcription early in life sets the stage for later brain development, identifying which germline genes are dysregulated and when will offer insight into their potential impact upon *Kdm5c* mutant neurodevelopment and possibly reveal novel avenues for therapeutic intervention.

Master regulators of germline identity are aberrantly expressed during *Kdm5c* mutant embryogenesis

We first assessed transcription of germline genes in *Kdm5c* mutant mouse epiblast-like cells (EpiLCs), an *in vitro* model of epiblast stem cells within the post-implantation embryo (Figure 3A). EpiLCs are an ideal model for interrogating embryonic misexpression of germline genes, as some germline-enriched genes are also expressed in embryonic stem cells (ESCs) but are silenced as they differentiate into EpiLCs¹⁰⁶. Using our previously published RNA-seq dataset⁵⁴, we assessed differential expression of germline genes in *Kdm5c* mutant male (XY) and female (XX) EpiLCs in comparison to wild-type controls through DESeq2. Because complete loss of KDM5C in females is not lethal at this time point, we were able to characterize germline transcripts in both *Kdm5c* heterozygous (XX *Kdm5c*-HET) and homozygous null (XX *Kdm5c*-KO) female EpiLCs.

We found both XX and XY *Kdm5c* mutant EpiLCs aberrantly express many germline transcripts and many are shared between all mutants (Figure 3B-C). Interestingly, female heterozygous and homozygous mutants expressed about twice as many germline genes than *Kdm5c*-KO males (XX *Kdm5c*-HET: 212 XX *Kdm5c*-KO: 186, XY *Kdm5c*-KO: 90), potentially suggesting they are more prone to adopting a germ cell-like state (Figure 3C). We then asked if the identity of ectopic germline transcripts was influenced by *Kdm5c* mutant sex, as some germline genes show biased transcription in the wild-type XX or XY gonads (see Chapter 3).

We found XX *Kdm5c*-HET, XX *Kdm5c*-KO, and XY *Kdm5c*-KO EpiLCs all expressed similar proportions of XX-biased, XY-biased, and unbiased germline genes that roughly reflected the proportion of genes shared and unique within the male and female wild-type gonads (In WT gonads: 6% XX-biased, 79% XY-biased, 15% unbiased) (Figure 3D). This suggests KDM5C is particularly important for embryonic repression of germline genes in females, however the identity of aberrant germline mRNAs is largely independent of sex.

To evaluate whether germline mRNAs are constitutively dysregulated or change over the course of development, we compared EpiLC germline DEGs to those expressed in the *Kdm5c* mutant brain. The majority of EpiLC germline DEGs were not dysregulated in the brain, with only a few shared across sequencing datasets (Figure 3E). Only two germline genes were consistently misexpressed in all male and female datasets, *DNA segment, Chr 1, Pasteur Institute 1 (D1Pas1)* and *Testis expressed gene 14 (Tex14)*. Surprisingly, germline DEGs unique to *Kdm5c* mutant EpiLCs included master regulators of germline identity, such as *Stra8* and *Dazl* (Figure 3F). As discussed above, these genes are typically expressed during embryonic germ cell development to commit PGCs to the germline fate, but they are also expressed later in life to trigger meiotic gene expression programs^{102–104}. *Stra8*'s expression is typically sexually dimorphic, as female germ cells initiate meiosis during embryogenesis while males begin meiosis after birth^{102–104}. Yet we found both XX and XY *Kdm5c* mutant EpiLCs ectopically express *Stra8* RNA. XX and XY *Kdm5c* mutants also aberrantly express *Dazl*, a translational regulator essential for germ cell function¹⁰¹. Interestingly, a significant portion of germline transcripts expressed in *Kdm5c* mutant EpiLCs are known binding targets of DAZL, including *Stra8*¹¹¹ (XY *Kdm5c*-KO EpiLCs $p = 1.698\text{e-}07$, Fisher's Exact Test). This suggests expression of DAZL protein would enable the translation of other aberrant germline transcripts, influencing their ability to impact on *Kdm5c* mutant cellular function. Therefore, we then stained XY *Kdm5c*-KO EpiLCs for DAZL protein expression through immunocytochemistry (Figure 3G). We found about 25% of *Kdm5c*-KO EpiLCs expressed DAZL protein in their cytoplasm ($p = 0.0015$, Welch's t-test, XY KO versus WT), consistent with the subcellular localization observed when DAZL is stained in spermatogonia¹¹¹. Expression of DAZL protein in EpiLCs gives credence to the possibility that ectopic germline transcripts have functional consequences within *Kdm5c* mutant cells. However, their impact may be limited to early

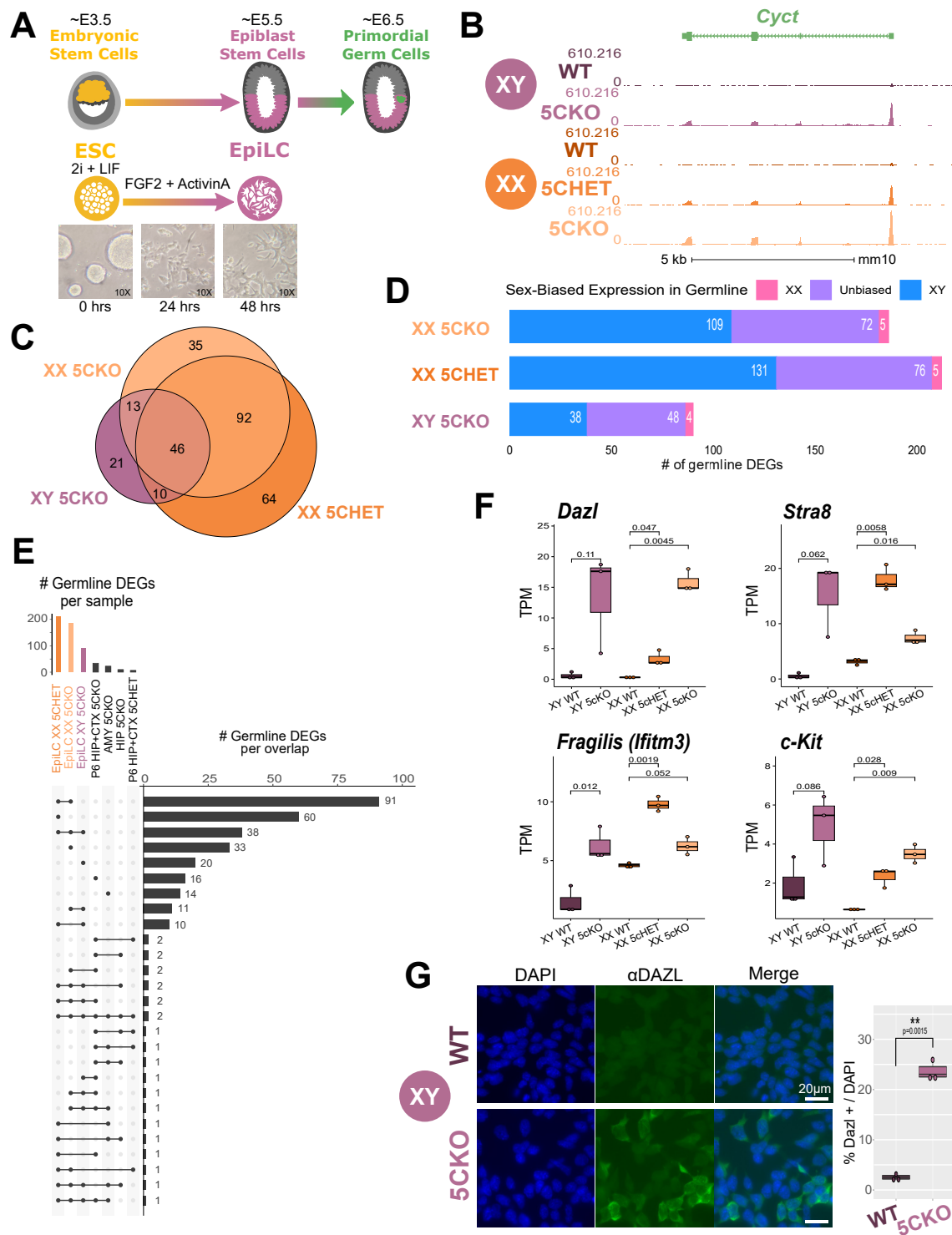


Figure 3: Male and female *Kdm5c* mutant epiblast-like cells express master regulators of germline identity **A.** Top - Diagram of *in vivo* differentiation of embryonic stem cells (ESCs) of the inner cell mass into epiblast stem cells and primordial germ cells. Bottom - *in vitro* differentiation of ESCs into epiblast-like cells (EpiLCs). **B.** Bigwigs of an example germline gene, *Cyct* that is dysregulated in male and female *Kdm5c* mutant EpiLCs **C.** Overlap of all germline DEGs between male and female EpiLCs **D.** Number of EpiLC germline DEGs that show sex-biased expression in the wild-type gonads. (Legend continued on next page.)

Figure 3: Male and female *Kdm5c* mutant epiblast-like cells express master regulators of germline identity. (Legend continued.) **E.** Upset plot displaying the overlap of germline DEGs expressed in male and female *Kdm5c* mutant EpiLC and brain RNA-seq datasets. **F.** Expression of genes key for early germ cell development in *Kdm5c* mutant EpiLCs in transcripts per million (TPM) with p-values for Welch’s t-test compared to wild type of the same sex. **G.** Immunocytochemistry of DAZL in male wild-type (WT) and *Kdm5c*-KO (5CKO) EpiLCs.

development, as we did not observe aberrant *Dazl* expression in the *Kdm5c* mutant brain.

KDM5C is recruited to germline genes during early embryogenesis

To further characterize the mechanism behind KDM5C-mediated germline gene repression, we evaluated the accumulation of KDM5C’s substrates, histone 3 lysine 4 di and tri-methylation (H3K4me2/3), at germline gene promoters using our previously published chromatin immunoprecipitation followed by sequencing (ChIP-seq) datasets^{54,64}. In *Kdm5c*-KO male EpiLCs, we observed aberrant accumulation of H3K4me2 at transcription start site (TSS) of many germline genes (Figure 4A). Accumulation of H3K4 methylation was even more apparent in the adult *Kdm5c*-KO male amygdala, where we observed ectopic peaks of H3K4me3 at germline TSSs that are completely absent in the wild-type brain (Figure 4B). Together, these data indicate KDM5C loss can lead to aberrant accumulation of H3K4 methylation at germline gene promoters that can last across the lifespan.

To evaluate KDM5C’s role in germline gene repression across development, we compared KDM5C binding in EpiLCs and primary neuron cultures from the cortex and hippocampus (PNCs) using our previously published ChIP-seq datasets^{53,54}. Although we identified many KDM5C ChIP-seq peaks in both cell types (EpiLCs: 2,437, PNCs: 1,610), KDM5C was only bound to germline gene promoters in EpiLCs and not in PNCs (Figure 4C). In EpiLCs, KDM5C was bound to germline genes transcribed in *Kdm5c*-KO EpiLCs as well as those dysregulated in the brain, such as *D1Pas1* (Figure 4C). This indicates long-term misexpression of germline genes is due to developmental loss of KDM5C suppression. KDM5C was also bound to some germline gene promoters that are only transcribed in EpiLCs, such as the master regulator *Dazl* (Figure 4C). To further delineate KDM5C’s role in suppressing germline

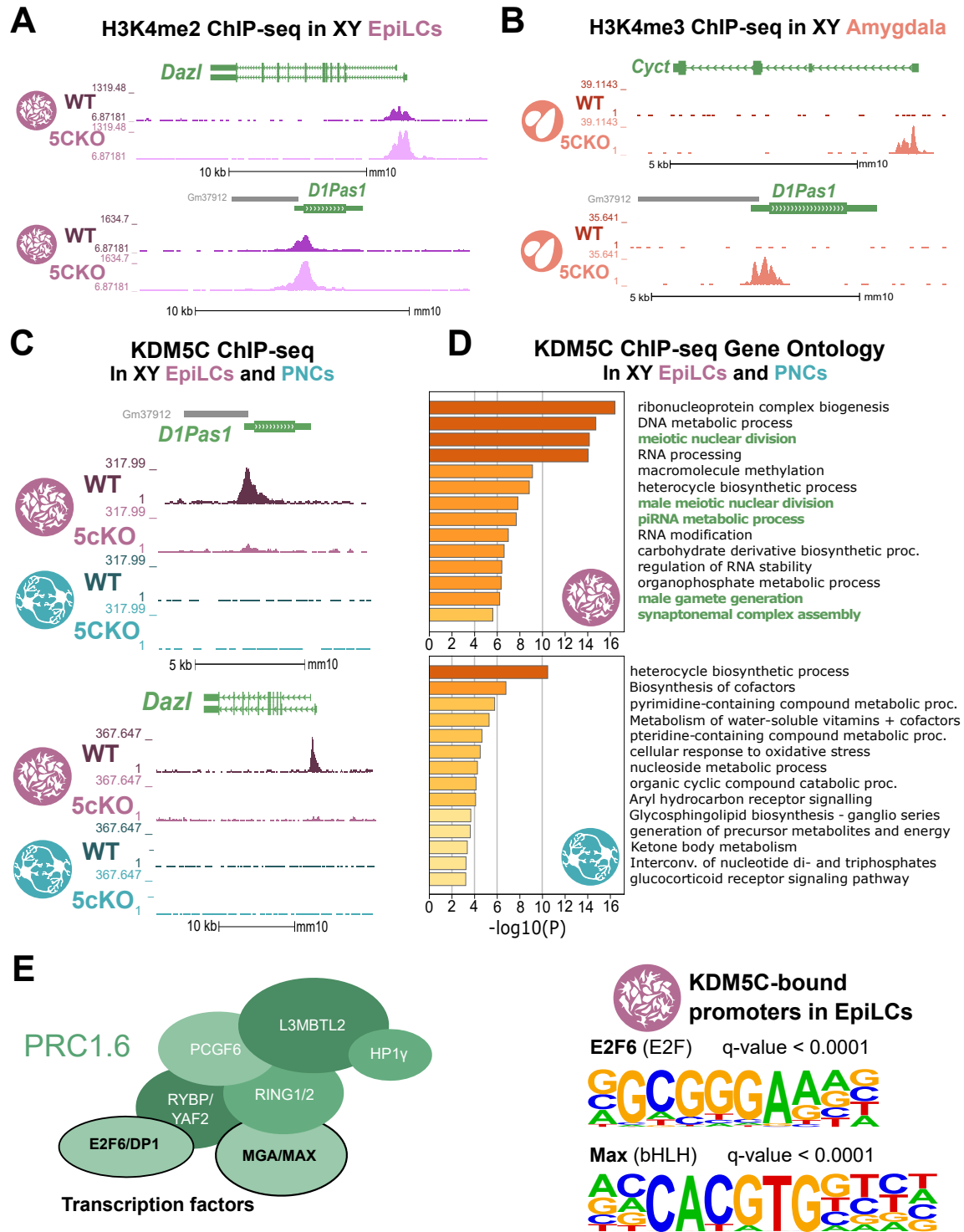


Figure 4: KDM5C binds to and represses germline gene promoters during embryogenesis. **A.** Bigwigs of chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) for histone 3 lysine 4 dimethylation (H3K4me2) in wild-type (WT) and *Kdm5c*-KO (5CKO) male epiblast-like cells (XY EpiLCs) at two germline gene promoters. **B.** Bigwigs of H3K4 trimethylation (H3K4me3) ChIP-seq in male wild-type and *Kdm5c*-KO adult amygdala at two germline gene promoters. **C.** Bigwigs of KDM5C ChIP-seq in wild-type and *Kdm5c*-KO EpiLCs (purple) and primary neuron cultures from the hippocampus and cortex (blue). (Legend continued on next page.)

Figure 4: KDM5C binds to and represses germline gene promoters during embryogenesis. (Legend continued.) **D.** Metascape gene ontology analysis of all genes with KDM5C-bound promoters in XY EpiLCs (top) and primary neuron cultures (PNCs, bottom). Germline-relevant ontologies highlighted in green. **E.** Left - diagram of polycomb repressive complex 1.6 (PRC1.6) members. Right - Two motifs enriched from HOMER analysis of KDM5C-bound promoters in EpiLCs, E2F6 and Max.

identity, we assessed the function of all genes with KDM5C bound to their promoter in EpiLCs and PNCs through gene ontology (GO) analysis (Figure 4D). We found KDM5C-bound genes in EpiLCs are significantly enriched for ontologies related to meiosis, including meiotic nuclear division (GO:0140013) and synaptonemal complex assembly (GO:0007130). Contrastingly, KDM5C appears to have a distinct, non-germline function in PNCs and regulates genes involved in heterocycle synthesis (GO: 0018130) and pyrimidine metabolism (GO: 0072527).

Although KDM5C is bound to germline genes in EpiLCs, it is unclear how KDM5C is recruited to these loci as KDM5C itself does not contain domains known for sequence-specific binding. Some evidence suggests KDM5C is recruited by PRC1.6, as KDM5C binds to PRC1.6 members in HeLa cells¹¹² and co-immunoprecipitates with PRC1.6's core member (PGCF6) in ESCs¹¹³. To elucidate the means by which KDM5C is recruited to germline genes, we evaluated which transcription factor motifs are enriched among KDM5C-bound genes through HOMER (Figure 4E). We found KDM5C-bound genes in EpiLCs, but not PNCs, displayed significant enrichment of MAX and E2F6 motifs, two transcription factors with known roles in germline gene repression^{114,115} and are the DNA-binding members of PRC1.6^{105,116,117}. Altogether, these results indicate ectopic transcription of germline genes in the mature *Kdm5c* mutant brain is due to loss of KDM5C repression during embryogenesis, in concordance with previous findings⁴². Additionally, we found KDM5C suppresses many meiotic genes during early embryogenesis that are not transcribed in the *Kdm5c* mutant brain, indicating they are suppressed through secondary, KDM5C-independent mechanisms.

Germline gene misexpression occurs in multiple *Kdm5c*-KO somatic tissues and is influenced by retinoic acid

Our *in vitro* results in EpiLCs indicate germline genes are aberrantly expressed very early on in *Kdm5c* mutant embryogenesis. As epiblast stem cells give rise to all somatic cells in the body, we then wondered if germline mRNAs are found in other *Kdm5c* mutant somatic tissues or if the brain is uniquely sensitive. To compare germline gene expression directly between neural and non-neural tissues in an unbiased manner, we then performed mRNA sequencing in the *Kdm5c*-KO male P6 liver along with the cortex and hippocampus harvested from the same animals (Figure 5A). The *Kdm5c*-KO brain had many more DEGs than the liver overall, (Brain: 1,079, Liver: 70), supporting KDM5C's uniquely important role in neurodevelopment (Figure 5B). While we found aberrant germline transcripts in both tissues, the majority of germline DEGs were tissue-specific with only one gene dysregulated in both datasets (*Testis-expressed gene 14*, *Tex14*) (Figure 5C). The liver had far fewer germline DEGs than the P6 brain and had the lowest number of germline DEGs when compared to all of our previous RNA-seq studies in *Kdm5c*-KO males (Figure 5D). Germline gene misexpression was highest in EpiLCs (when KDM5C is bound to their promoters), second-highest in the P6 brain, and further decreased in the adult amygdala and hippocampus. Of note, the number of germline DEGs identified in each tissue is likely influenced by cell heterogeneity, as bulk sequencing can mask discordant, cell type-specific changes. Yet, we found many more germline DEGs in the P6 hippocampus and cortex than in the liver, even though these brain regions contain a many more distinct cell types than the liver. Altogether, these results demonstrate the identity of germline genes that are misexpressed long-term is tissue-specific and the brain may be particularly sensitive to long-term, aberrant germline gene expression with *Kdm5c* loss.

Differential misexpression of germline genes within *Kdm5c*-KO somatic tissues suggests cellular environment greatly influences how this phenotype manifests. We were particularly interested in the contribution of retinoic acid (RA) on long-term germline gene misexpression, as it is already known to regulate the expression of some germline genes¹⁰², including the meiotic transcription factor *Stra8*. RA also influences body patterning and tissue specific

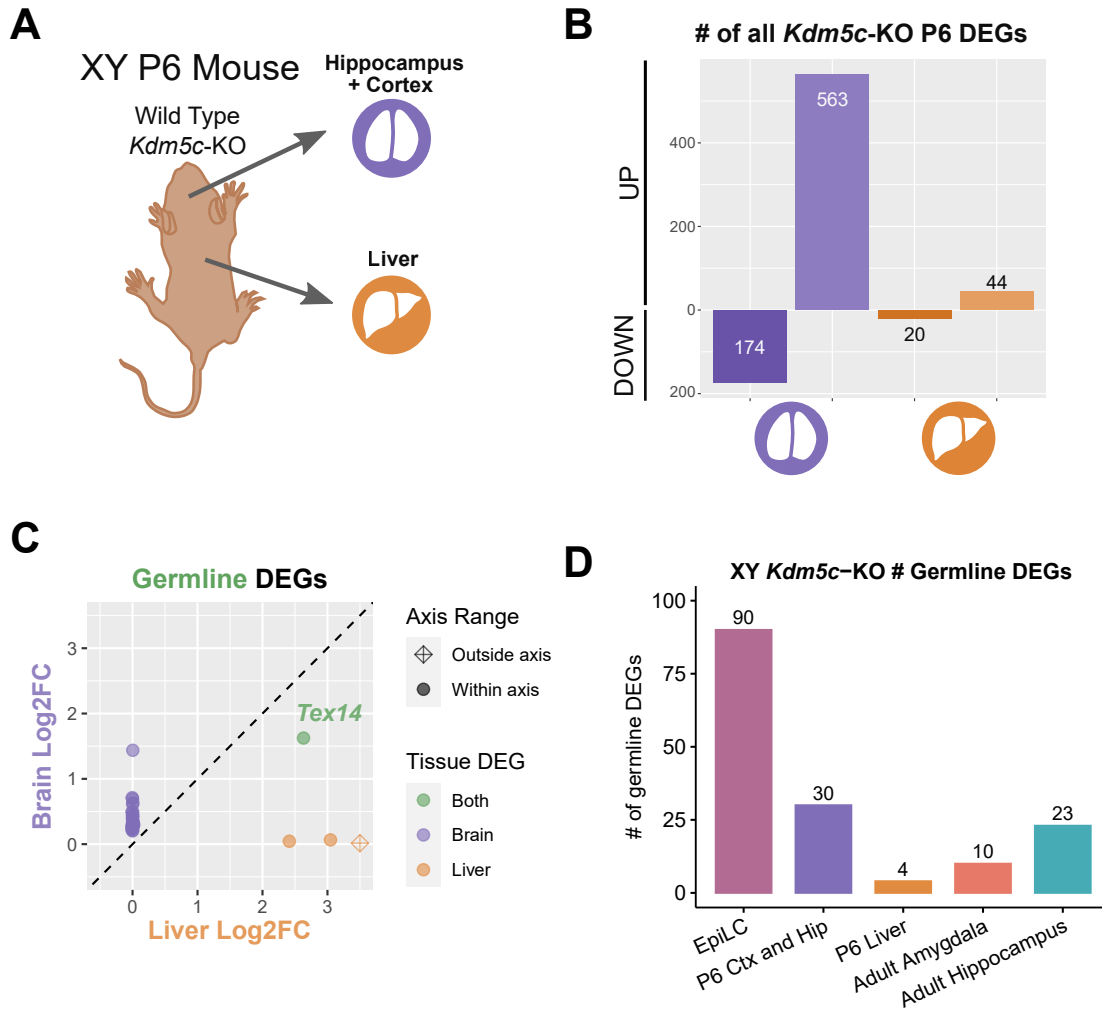


Figure 5: Misexpression of germline genes in the male *Kdm5c*-KO postnatal day 6 (P6) brain and liver. **A.** Diagram of tissues harvested for RNA sequencing. **B.** Total number of all up and downregulated differentially expressed genes (DEGs) in the P6 brain (purple) and liver (orange). **C.** Scatter plot of germline-enriched DEGs comparing the log2 fold change (L2FC) from wild-type in the respective tissues. Only one gene, *Tex14* was dysregulated in both tissues. **D.** Histogram of the number of germline-enriched DEGs in all male *Kdm5c*-KO RNA-seq datasets analyzed in this study.

gene expression for many other organs, including neural differentiation and anterior-posterior neural patterning^{118,119}. To assess the if RA promotes the aberrant transcription of germline genes, we differentiated wild-type and *Kdm5c*-KO male ESCs into EpiLCs in media with or without vitamin A, the precursor for RA (Figure 6A). Naive, “ground-state” ESCs were cultured using serum-free media and thus only exposed to vitamin A upon EpiLC differentiation. We performed mRNA-seq on cells differentiated for either 0 hours (ESCs, no RA), 48 hours (EpiLCs), or 96 hours (extended EpiLCs, exEpiLCs) to evaluate RA’s influence over time. We then called genes differentially expressed between *Kdm5c*-KO and wild-type cells exposed to the same conditions using DESeq2 ($q < 0.1$, $\log_2$ fold change > 0.5) and filtered for germline-enriched genes. We found that ESCs and EpiLCs differentiated for 48 hours had a similar total number of germline DEGs regardless of RA status (Figure 6B). However, at 96 hours *Kdm5c*-KO exEpiLCs cultured with RA had almost double the number of DEGs than exEpiLCs cultured without RA, with 116 germline genes uniquely dysregulated when exposed to RA long-term.

To evaluate the global impact of RA on germline gene misexpression, we performed hierarchical clustering of germline DEG expression in *Kdm5c*-KO cells at each time point and RA condition (Figure 6C-E). This clustering included all germline genes that were called as significantly different from wild-type controls for any time point or RA status. Based on their relative difference in expression (calculated in transcripts per million, TPM) across differentiation and RA exposure, RA exacerbated the expression of genes in clusters 1, 2, or 4 at one or multiple time points (Figure 6C-D). Genes in cluster 1 showed increased expression at 48 hours regardless of RA presence, but only displayed sustained activation at 96 hours with RA (Figure 6E, top left). This cluster included the meiotic gene *Sycp3*, an essential component of the synaptonemal complex¹²⁰. Contrastingly, RA transiently promoted the expression of genes in cluster 2 at 48 hours, but their expression was dampened by 96 hours independent of RA (Figure 6E, top right). Cluster 2 included the meiotic transcription factor *Stra8*, which is known to be induced by RA. The majority of RA-influenced germline DEGs fell into cluster 4 and showed highest expression at 96 hours with RA (Figure 6E, bottom left). Interestingly, this cluster included some germline genes that are de-repressed in the mature *Kdm5c*-KO brain, such as *Spag16*, that were not previously identified in 48 hour

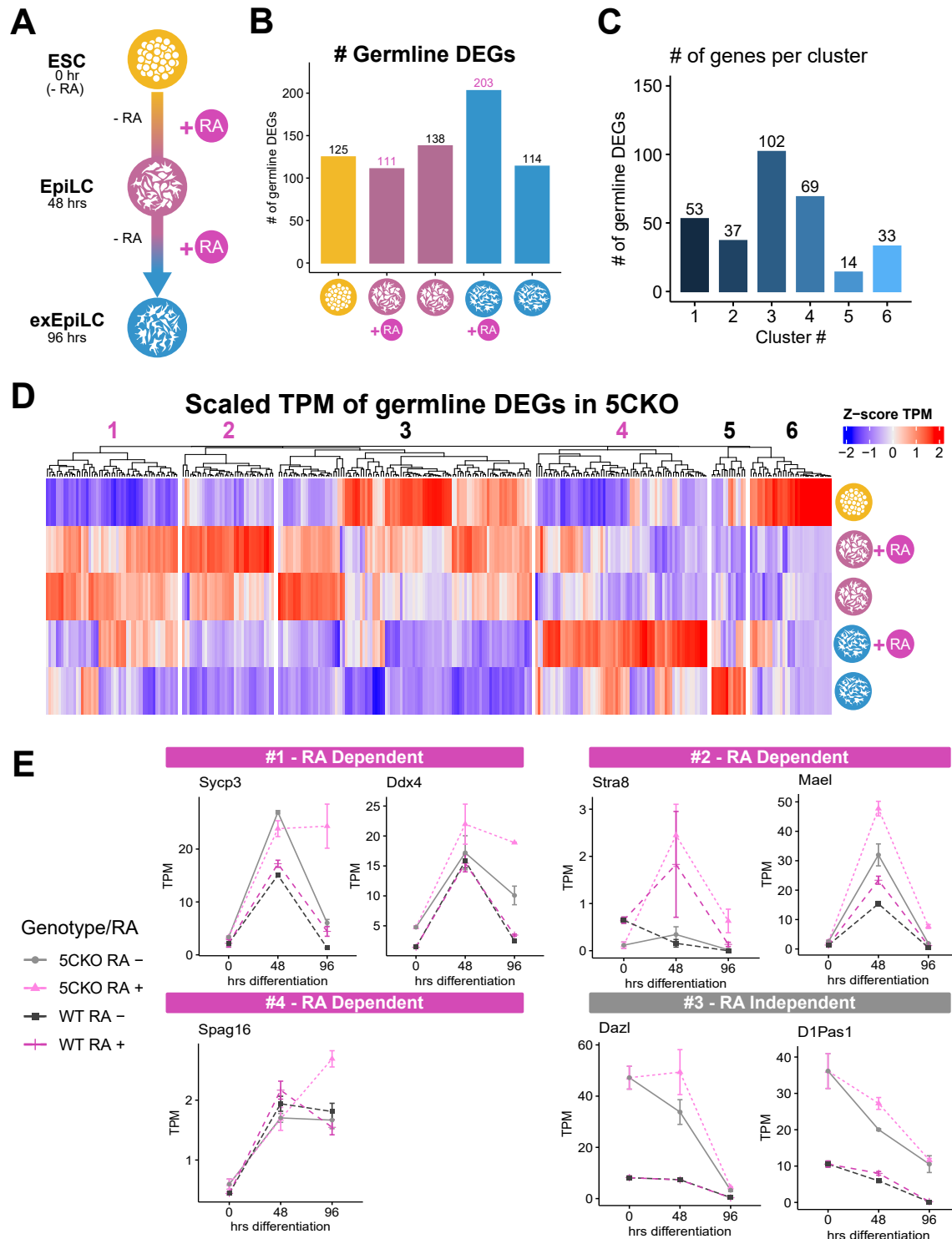


Figure 6: Retinoic acid promotes the prolonged misexpression of germline genes in *Kdm5c*-KO epiblast-like cells. **A.** Diagram of male embryonic stem cell (ESC, yellow) differentiation into epiblast-like cells (EpiLCs, purple) and extended EpiLCs (exEpiLCs, blue), either with or without vitamin A (pink) in the media, the precursor for retinoic acid (RA). **B.** Number of *Kdm5c*-KO male germline-DEGs in the different cell types and culture conditions. **C.** Hierarchical clustering of germline gene expression in *Kdm5c*-KO cells, the number of genes in each cluster identified. (Legend continued on next page.)

Figure 6: Retinoic acid promotes the prolonged misexpression of germline genes in *Kdm5c*-KO epiblast-like cells. (Legend continued.) **D.** Heatmap of germline DEG expression in *Kdm5c*-KO (5CKO) cells at different differentiation time points and culture conditions, grouped by hierarchical clustering. Colors indicate the scaled transcripts per million (TPM) expression of germline DEGs, with red being highest expression and blue being lowest. **E.** Line plots of example germline genes in each cluster, displaying their TPM at the different time points. Black - wild-type cells differentiated without retinoic acid, dark pink - wild-type cells differentiated with retinoic acid, gray - *Kdm5c*-KO cells differentiated without retinoic acid, pale pink - *Kdm5c*-KO cells differentiated with retinoic acid.

EpiLCs. Surprisingly, many germline genes within clusters 1, 2, and 4 also displayed increased expression in wild-type cells at 48 hours, with or without RA, but are then repressed by 96 hours. This suggests differentiation itself turns on germline gene expression programs in wild-type cells that are then decommissioned with prolonged differentiation. KDM5C appears to be involved in this decommissioning process, as *Kdm5c*-KO exEpiLCs either fail to turn off germline genes entirely (clusters 1 and 4) or have dampened but significant germline gene overexpression compare to wild-type cells (cluster 2). Combined, clusters 1, 2, and 4 contained about half of all *Kdm5c*-KO germline DEGs (RA-dependent: 159 genes, RA-independent: 149 genes), indicating RA greatly influences the manifestation of *Kdm5c*-KO germline gene misexpression. However, we also found many germline DEGs whose aberrant expression is independent of RA. Cluster 3 was the single largest cluster identified and these genes decreased their expression across cellular differentiation regardless of RA status, but were still significantly upregulated from wild-type cells at 96 hours (Figure 6E, bottom right). Cluster 3 included germline genes that were only dysregulated in *Kdm5c* mutant EpiLCs, such as *Dazl*, but also germline DEGs expressed in EpiLCs and the mature brain, like *D1Pas1*. Therefore, additional, unknown mechanisms influence which germline genes display long-term de-repression in the *Kdm5c* mutant brain and which are ultimately silenced.

Discussion

In the above work we characterized germline gene misexpression in *Kdm5c* mutants during the embryonic period when the soma-germline boundary is typically established using an *in vitro* cell culture model. This work supports previous evidence that proposed KDM5C suppresses germline genes during embryogenesis due to loss of KDM5C binding at germline gene

promoters in differentiated cell types⁴². However, embryonic germline gene misexpression had not yet been characterized in *Kdm5c* mutants and it was unclear if the identity of misexpressed germline transcripts would change over the course of development. Additionally, it was unknown if the impact of *Kdm5c* loss during embryogenesis would be the same in males and females, since the expression of some germline genes is sexually dimorphic during typical embryonic germ cell development. Furthermore, it was unknown if tissue identity and cellular environment influences the manifestation of germline gene misexpression in *Kdm5c* mutants. Characterizing these distinctions through the experiments above illuminates the molecular mechanisms underlying this unique phenotype. These results not only elucidate fundamental mechanisms governing germline and somatic identity but also provides the intellectual framework for testing the functional consequences of germline gene misexpression (see Chapter 5).

Unlike our findings in the brain, we found female *Kdm5c* mutant EpiLCs misexpressed about twice as many germline genes than *Kdm5c*-KO males. Females expressing more germline genes than males during embryogenesis then fewer later in life could be due to the cells that are most severely impacted failing to survive and proliferate. Higher rates of cell death in *Kdm5c* mutants could also explain why we observe many meiotic genes dysregulated in male and female EpiLCs but not in the brain. While we currently do not know if *Kdm5c* mutant cells have increased rates of cell death, it has been observed in ESCs lacking the transcription factor *Max*¹¹⁴, which aberrantly express many of the same germline genes as *Kdm5c* mutant ESCs and EpiLCs. In addition to *Kdm5c* mutant EpiLCs expressing a handful of germline genes also found within the brain, we found both male and female EpiLCs expressed *Stra8* and *Dazl*, two master regulators of germline identity and meiosis. Excitingly, we found *Dazl* is translated into protein that is localized to the cytoplasm within male *Kdm5c*-KO EpiLCs, supporting our hypothesis that aberrant germline transcripts can be functional. However, their functional impact may be limited to early development, as DAZL regulates the translation of other germline transcripts and we did not observe *Dazl* expression in the brain. If germline gene misexpression does contribute to MRXSCJ-related phenotypes, further work determining if germline genes are translated within the brain will ultimately inform the therapeutic window for treatment intervention.

Of note, some germline genes, including DAZL, are also expressed in the 2-cell embryo, the first developmental stage after fertilization when cells are still totipotent¹¹³. One recent publication screening for suppressors of the 2-cell state in ESCs through a *Dazl* reporter and also identified KDM5C as a DAZL suppressor along with other novel proteins¹¹³. While *Kdm5c*-KO ESCs primarily displayed overexpression of germline genes, some of which are also expressed in the 2-cell state, they also found misexpression of a few 2-cell state-specific genes. *Kdm5c*-KO DEGs specific to the 2-cell state included the transcription factor *Dux*, a key regulator of totipotency in mice. However, unlike the three other 2-cell regulators they investigated, misexpression of *Dazl* in *Kdm5c*-KOs was independent of *Dux* expression. We did not observe *Dux* dysregulation in our *Kdm5c*-KO ESC or EpiLC RNA-seq datasets, but we did observe overexpression of *Zscan4a-f* in EpiLCs, a group of 2-cell genes that are also expressed in the mouse ovary and testis (data not shown). Altogether, these data indicate KDM5C primarily acts as a suppressor of germ cell genes, but the overlap between 2-cell and germ cell genes can also lead to a 2-cell-like state in ESCs with *Kdm5c* loss.

Interestingly, this study also found KDM5C's suppression of DAZL in ESCs was independent of its catalytic activity¹¹³. *Dazl*, along with other germline genes, is initially silenced in ESCs by repressive histone 2A lysine 119 monoubiquitination (H2AK119ub1) and histone 3 lysine 9 trimethylation (H3K9me3), but then later fully decommissioned in EpiLCs by *de novo* DNA methylation by DNMT3A and B^{105,106}. Since KDM5C's substrate (H3K4me3) can repress *de novo* DNA methylation¹¹⁰, KDM5C's catalytic activity may only be required once ESCs undergo differentiation for long-term silencing. Indeed, another study found DNA methylation is eroded at the promoter of at least two germline genes transcribed in the mature *Kdm5c*-KO brain⁴². If this is the case, the observed rescue of *Dazl* repression by catalytic dead KDM5C in ESCs could be due to KDM5C stabilizing the binding of other chromatin regulators at DAZL's promoter. Further experiments testing catalytic dead KDM5C during ESC and EpiLC differentiation are required to tease apart these two possibilities. Nevertheless, we did observe aberrant accumulation of H3K4me2/3 at *Kdm5c*-KO germline gene promoters in EpiLCs and in the mature amygdala, suggesting KDM5C typically erases these marks in wild-type cells. However, it is still unclear if this H3K4me2/3 accumulation directly impairs long-term germline gene silencing or if their transcriptional repression occurs independent of

KDM5C's catalytic activity.

We additionally demonstrated misexpression of germline genes during *Kdm5c* mutant embryogenesis leads to their transcription in multiple *Kdm5c*-KO somatic tissues later in life, indicating all somatic tissues and cell types are impacted to some degree. Out of all our *Kdm5c*-KO RNA-seq datasets, we observed the most robust dysregulation of germline transcripts in EpiLCs. The P6 forebrain had the second highest number of germline-enriched mRNAs and this further decreased in the adult *Kdm5c*-KO brain. The P6 liver had the lowest number of germline DEGs, potentially suggesting the brain is more sensitive to prolonged germline gene misexpression. While cell heterogeneity likely masks the dysregulation of some germline DEGs given that RNA sequencing was done in bulk, it is striking that very few germline DEGs were shared between datasets. These results suggest developmental cues can greatly influence the long-term misexpression of germline genes. We tested the influence of vitamin A, the precursor for retinoic acid (RA), as it is an important morphogen for cell growth and differentiation in multiple tissues^{118,119}. RA exacerbated the misexpression of many germline genes at one or multiple time points as *Kdm5c*-KO ESCs differentiated into exEpiLCs, including the meiotic genes *Stra8* and *Sycp3*. Interestingly, some germline DEGs that we had previously classified as brain-specific were expressed only when EpiLCs were exposed to RA for 96 hours, like *Spag16*. However, many germline DEGs also displayed a steady decrease in expression over the course of ESC to exEpiLC differentiation regardless of RA exposure. While these genes included some that are not dysregulated later in life, like *Dazl*, it also included genes consistently misexpressed in the *Kdm5c* mutant brain, like *D1Pas1*. Thus, other, unknown cues influence which germline genes persist expression over the course of *Kdm5c* mutant development and which were decommissioned. Meiotic genes in particular were almost exclusively observed in EpiLCs and not identified in later *Kdm5c* mutant tissues. Given the misexpression of these genes can lead to genome instability¹²¹, their loss later in life could be due to repression by unknown, secondary surveillance mechanisms or death of the cells the most severely impacted, as discussed above.

Chapter 5 - Consequences of Ectopic Germline Gene Expression Upon Neurodevelopment

Introduction

“Structure dictates function” is a foundational principle of biology that also holds true for the chromatin landscape. Chromatin marks across the genome influence which genes are transcribed, ultimately determining what tasks a cell can perform and how a cell responds to external stimuli^{11,12}. Germ cells have a highly specialized transcriptome^{72,122} because of their unique functions, including but not limited to the meiotic reorganization of genetic material, maintaining genome integrity for future generations, and re-establishing sex-specific genomic imprinting. Ectopic transcription of germline genes with the loss *Kdm5c* suggests the chromatin landscape in somatic cells has fundamentally shifted towards a germline fate. Therefore, this deviation in chromatin structure opens the possibility that *Kdm5c* mutant somatic cells acquire aberrant germline related functions. It is currently unknown if ectopic germline genes cause functional changes in *Kdm5c* mutant cells and if these changes impact the development of the *Kdm5c* mutant brain’s transcriptome.

While the vast majority of germline genes have never been studied in neurodevelopment, some studies have investigated their functional consequences when expressed in other, non-

traditional contexts. In ESCs, loss of the transcription factor Myc-Associated factor X (MAX), a key member of the PRC1.6 complex¹⁰⁵, results in aberrant transcription of *Stra8*, *Dazl*, and *Sycp3*¹¹⁴. These genes are fundamental for meiosis in germ cells and are also expressed in *Kdm5c* mutant ESCs and EpiLCs. Even in stringent culture conditions to maintain ESC identity, *Max*-null ESCs displayed meiotic leptotene and zygotene-like SYCP3 staining, increase in the double strand break marker gamma H2AX, and the increased rates of apoptosis¹¹⁴. Interestingly, additional loss of *Stra8* in *Max*-null ESCs rescued both meiotic morphology and cell survival¹¹⁴, demonstrating that the additional loss of master germline regulators can rescue the deleterious effects caused by misexpression of germline genes. A few studies in non-mammalian models have also observed functional consequences to germline gene misexpression outside of early embryogenesis. In *Drosophila* and *C. elegans*, somatic cells repress germline gene expression through the chromatin-regulatory complex Drosophila RBF, E2F, and Myb or Myb-MuvB (dREAM/MMB)¹²³. Loss of dREAM/MMB complex members in *C. elegans* results in the somatic misexpression of P granule genes¹²⁴, a subcellular RNA and protein structure unique to germ cells that are essential for germline specification in *C. elegans*^{125,126}. These somatic cells also displayed germline-like features such as enhanced RNA interference (RNAi) activity¹²⁴, a mechanism of genome defense abundant within germ cells¹²⁷. In *Drosophila*, loss of dREAM/MMB complex member *lethal (3) malignant brain tumor (l(3)mbt)* leads to brain tumors that express germline-specific proteins, such as PIWI, VASA, and AUB¹²⁸. Additional deletion of these germline genes suppressed tumor growth, suggesting the runaway proliferation of these cancerous cells is due to acquisition of ectopic, germline-like proliferation and immortality. Together, these studies indicate misexpression of germline-enriched genes can manifest germline-like functions within somatic cells which can then be rescued by their suppression. However, it is currently unknown if somatic acquisition of germline characteristics contributes to neurodevelopmental impairments, as the impact of ectopic germline transcription has yet to be investigated within the mammalian brain.

In the preliminary work below, we interrogated manifestation of two germline-like characteristics within *Kdm5c* mutant somatic cells - dysregulation of genomic imprinting and transposable element repression. Canonically imprinted genes are those exclusively transcribed from either the maternal or paternal chromosome and this expression pattern is

established in the germline through imprint erasure and then placement of sex-specific DNA methylation¹²⁹. We observed widespread dysregulation of imprinted gene expression within the *Kdm5c* mutant EpiLCs and brain tissues, including some whose dysregulation is linked to MRXSCJ-reminiscent phenotypes. Our ongoing work will investigate if overexpression of imprinted genes is caused by a germline-like loss of genomic imprinting or upregulation of the already expressed allele through allele-specific RNA sequencing. We additionally found *Kdm5c* mutant cells also express germline genes important for the biogenesis of PIWI-interacting RNAs (piRNAs), a class of short, non-coding RNAs enriched in germ cells. piRNAs are thought to help maintain germ cell genome integrity by repressing the expression of transposable elements (TEs), stretches of DNA that can move throughout the genome. We found male *Kdm5c*-KO EpiLCs displayed global downregulation of TEs, indicating hyperactivity of the piRNA pathway. Surprisingly, *Kdm5c* mutant females showed both up and downregulation of TEs. Ongoing work will evaluate if piRNAs themselves are overexpressed in male and female *Kdm5c* mutant EpiLCs and determine if TE dysregulation is due to aberrant germline gene expression through shRNA knockdown of the master regulator *Dazl*.

Finally, assessed the overall impact of aberrant germline gene expression upon the *Kdm5c*-KO brain transcriptome by genetically blunting the translation of germline mRNAs through additional *Dazl* deletion. We found the additional heterozygous loss of *Dazl* during *Kdm5c*-KO development ameliorated the transcriptional dysregulation caused by *Kdm5c* loss alone. We found this rescue was not specific to germline transcripts, indicating *Kdm5c*-KO;*Dazl*-heterozygous double mutation partially restored the brain's overall developmental trajectory. We are currently investigating if this transcriptional rescue ameliorates *Kdm5c*-KO neuronal impairments, such as impaired dendritic arborization, homeostatic synaptic plasticity, and MRXSCJ-related behavioral changes. Ultimately, the completion of this work will provide insight into the consequences of blurring the soma-germline boundary and potentially provide the intellectual basis for pursuing novel therapeutic interventions for MRXSCJ and related NDDs.

Dysregulation of genomic imprinting and transposable element expression with *Kdm5c* loss

Genomic imprinting

One potential consequence of somatic cells adopting a germ cell-like fate is the misexpression of imprinted genes. In somatic cells, canonically imprinted genes are expressed exclusively from either the maternal or paternal chromosome through the placement of allele-specific DNA methylation¹²⁹. This parental-specific expression is achieved by erasing previous genomic imprints in germ cells and establishing sexually dimorphic DNA methylation at 20 imprinting control regions (ICRs). Tight regulation of imprinted gene expression is particularly crucial for proper neurodevelopment, as dysregulation of genomic imprinting at one or multiple loci can cause severe NDDs, such as Angelman’s syndrome and Prader–Willi syndrome¹³⁰. Therefore, misregulation of imprinted genes offers a clear route by which germline-like characteristics could contribute to neurodevelopmental impairments.

We first globally evaluated imprinted gene expression in our *Kdm5c* mutant bulk RNA sequencing datasets. We found significant dysregulation of imprinted genes in all of our *Kdm5c* mutant brain datasets, however EpiLCs displayed the most widespread dysregulation (Figure 7A). While imprinted DEGs were primarily upregulated in male and female brain samples, XX and XY EpiLCs showed significant up and downregulation of many imprinted genes. Both XX and XY *Kdm5c* mutant EpiLC and brain samples showed dysregulation of maternally and paternally expressed genes, indicating their dysregulation is independent of organismal sex. We then evaluated which imprinted DEGs were shared or unique between *Kdm5c* mutant RNA-seq datasets. EpiLCs displayed unique dysregulation of many famous imprinted genes, including *H19* and *Igf2* (Figure 7B). Imprinted genes upregulated across multiple *Kdm5c* mutant datasets included *Dlk1* and *Cdkn1c* (Figure 7C-D), two genes with reported neuronal functions^{131–133}.

Because H3K4me3 can impair the placement of *de novo* DNA methylation, imprinted genes could be upregulated due to KDM5C loss directly rather than due to somatic acquisition of germline identity. To evaluate if KDM5C plays a direct role in genomic imprinting, we

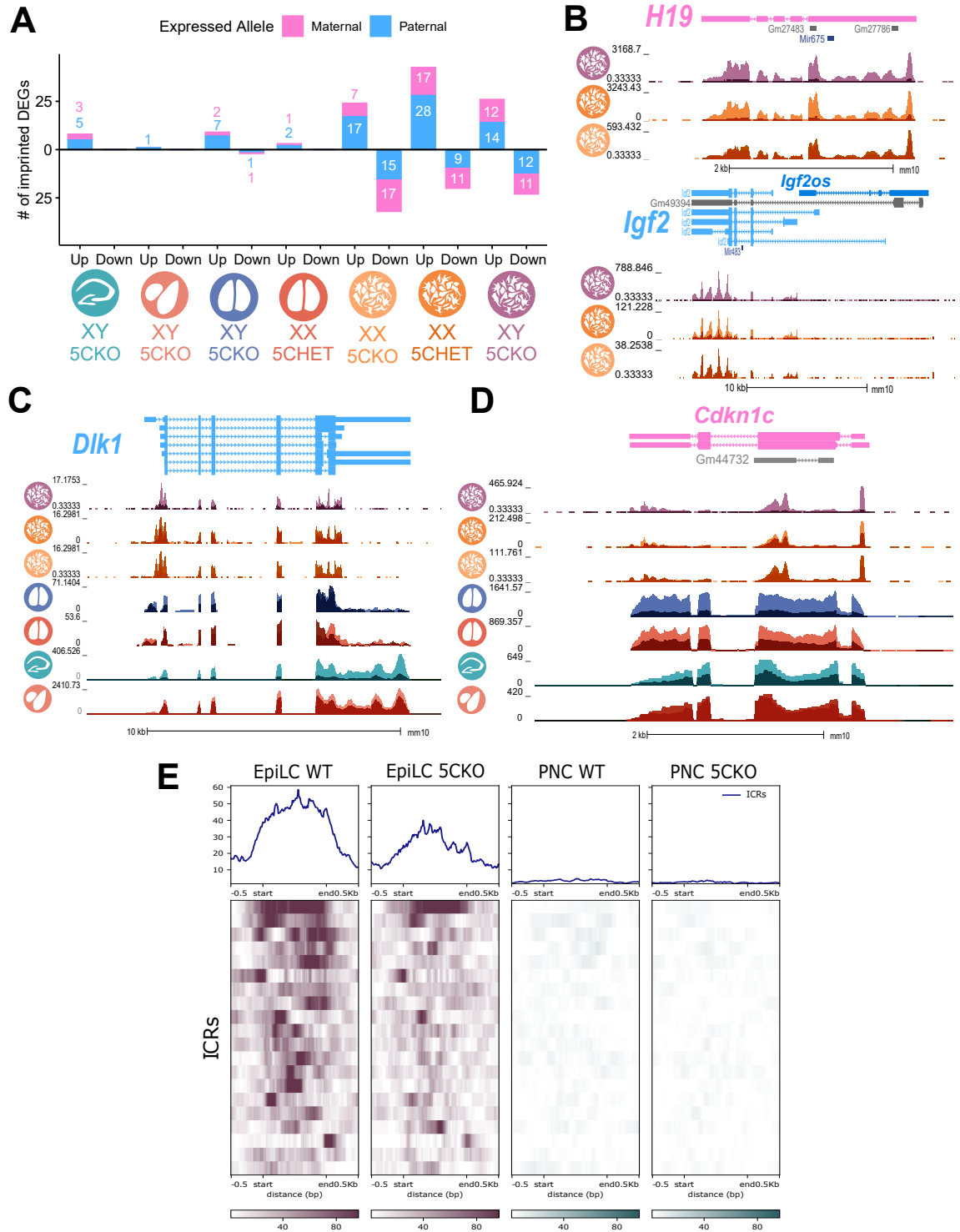


Figure 7: Dysregulation of imprinted genes in *Kdm5c* mutant males and females. **A.** Canonically imprinted genes dysregulated in our *Kdm5c* mutant RNA-seq datasets. Paternally expressed genes are in blue and maternally expressed genes in pink. **B.** Bigwigs of two imprinted genes dysregulated in *Kdm5c* mutant EpiLCs, *H19* and *Igf2*. Wild-type and *Kdm5c* mutant tracks are overlaid, with the wild-type being the darker color for each sex and sample. **C.** Bigwigs of *Dlk1*, a paternally expressed imprinted genes significantly dysregulated in *Kdm5c*-KO male EpiLCs, *Kdm5c*-HET female EpiLCs, and the adult male *Kdm5c*-KO hippocampus. (Legend continued on next page.)

Figure 7: Dysregulation of imprinted genes in *Kdm5c* mutant males and females. (Legend continued.) **D.** Bigwigs of *Cdkn1c*, a maternally expressed imprinted genes significantly dysregulated in *Kdm5c*-KO male EpiLCs, *Kdm5c*-HET female EpiLCs, and the postnatal day 6 male and female *Kdm5c* mutant cortex and hippocampus. **E.** Heatmap of KDM5C ChIP-seq signals in EpiLCs (left, purple) and primary neuron cultures (PNCs, right, blue) at imprinting control regions (ICRs) in wild-type and *Kdm5c*-KO (5CKO) negative controls. Graphs generated by deepTools.

assessed KDM5C binding at canonical imprinting control regions (ICRs), the 20 central hubs that are differentially methylated in the male and female germline to regulate imprinted gene expression (Figure 7E). KDM5C ChIP-seq in male EpiLCs and primary hippocampal and cortical neuron cultures (PNCs) revealed some modest enrichment of KDM5C at ICRs in EpiLCs but not PNCs. However, we did observe high background KDM5C signal in *Kdm5c*-KO EpiLC negative controls. These results indicate KDM5C may have a modest role in regulating ICRs during early embryogenesis. However, further experiments and analyses are warranted to distinguish if *Kdm5c* mutant imprinting dysregulation is a direct effect of KDM5C loss, reflecting a shift towards germline-like identity, or a combination of the two. If imprinted gene dysregulation in *Kdm5c* mutants is truly a consequence of germline-like identity, we would rescue of imprinted gene dysregulation when germline genes are suppressed through alternative mechanisms, such as *Dazl* shRNA knockdown. Of note, all of our current *Kdm5c* mutant RNA-seq datasets cannot distinguish if upregulated imprinted genes are biallelically expressed or due to upregulation of the already active allele. Therefore, we are currently generating F1 hybrids of JF1 and C57/Bl6 mice, which have frequent SNPs that enable sensitive, allele-specific RNA sequencing, to elucidate if genomic imprinting is truly lost in the *Kdm5c* mutant brain.

piRNAs and transposable elements

Another avenue by which germline genes could impact *Kdm5c* mutant somatic cell function is through the aberrant regulation of PIWI-interacting RNAs (piRNAs). piRNAs are short non-coding RNAs (24-32 nucleotides long) that are highly enriched in the germline and were originally considered germ cell-specific¹³⁴. However, emerging evidence suggests piRNAs are lowly expressed in somatic cells, including neurons. piRNAs are thought to promote genome integrity by repressing the expression of transposable elements (TEs), selfish pieces of DNA

that can move and replicate throughout the genome. When germ cells reset the chromatin landscape, the genome becomes highly vulnerable to TE expansion and deleterious insertions could be inherited in future generations. We found multiple regulators of piRNA biogenesis are aberrantly expressed in *Kdm5c* mutant EpiLCs and a few are also dysregulated in brain samples (data not shown). Both male and female *Kdm5c* mutant EpiLCs also aberrantly express many well known regulators of piRNA biogenesis, including *Piwi2*, and *Mael*¹³⁵ (Figure 8A-B, only male exEpiLC data shown).

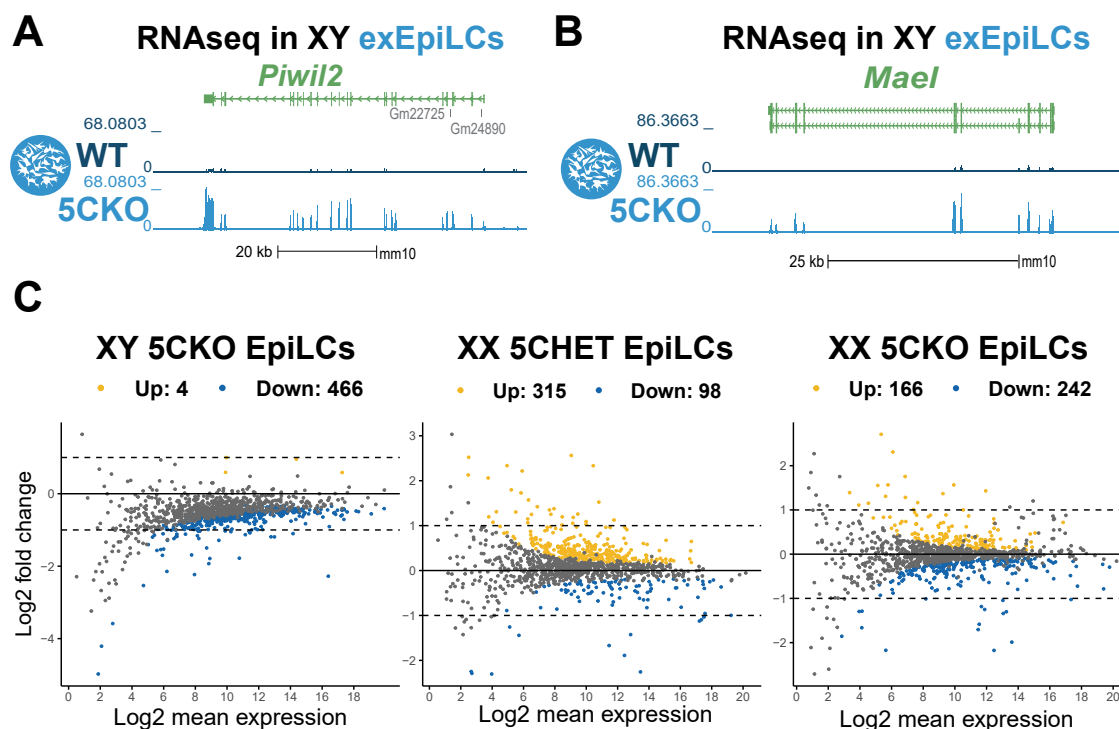


Figure 8: Aberrant expression of piRNA regulatory factors and transposable elements in *Kdm5c* mutant male and female EpiLCs. **A.** Bigwigs of *Piwi Like RNA-Mediated Gene Silencing 2* (*Piwi2*, also known as *Mili*) expression in male wild-type and *Kdm5c*-KO extended epiblast-like cells (exEpiLCs). **B.** Bigwigs of *Maelstrom spermatogenic transposon silencer* (*Mael*) expression in male wild-type and *Kdm5c*-KO extended epiblast-like cells (exEpiLCs). **C.** MA plots of transposable element expression in male and female *Kdm5c* mutant EpiLCs.

Since piRNAs are known to repress TEs in germ cells, we then assessed TE expression in *Kdm5c* mutant cells through TETRANSCRIPTS⁶⁹ (Figure 8C). While we did not observe significant changes in TE expression in any of our brain RNAseq datasets (data not shown), we did find widespread dysregulation of TEs in *Kdm5c* mutant EpiLCs. Interestingly, *Kdm5c*-KO male EpiLCs displayed global downregulation of TE expression, consistent with our prediction

of piRNA hyperactivity. However, the impact of *Kdm5c* loss on TE expression is sexually dimorphic, as *Kdm5c*-HET and KO females showed both up and downregulation of TEs. Further investigation is required to determine if dysregulation of TE expression in *Kdm5c* mutants is a true reflection of germline-like features, a distinct phenotype caused by KDM5C loss, or a combination of the two.

Amelioration of *Kdm5c*-KO brain transcriptome dysregulation through genetic deletion of *Dazl*

To test if misexpressed germline genes negatively impact *Kdm5c*-KO neurodevelopment as a whole, we blunted the impact of ectopic germline transcripts through the genetic deletion of *Dazl*. As discussed above, DAZL is an RNA binding protein that regulates the translation of many germline-specific transcripts¹⁰¹ and we found its protein is aberrantly expressed in *Kdm5c* mutant EpiLCs. Since a significant portion of *Kdm5c*-KO germline DEGs are DAZL targets (see Chapter 4), we predict loss of DAZL will prevent their translation into protein, ultimately dampening their impact upon somatic cell function. We generated double mutant male mice harboring both knockout of *Kdm5c* and heterozygous loss of *Dazl* (*Kdm5c*-KO;*Dazl*-HET) by crossing female *Kdm5c*-HET mice with male *Dazl*-HET mice obtained from Jackson Laboratory (Jax: 023802) (Figure 9A). Double mutant mice generated from this cross have decreased expression of DAZL throughout life, including during the embryonic time period when we predict DAZL protein to be aberrantly expressed based on our results in EpiLCs (see Chapter 4).

We assessed the impact of additional *Dazl* loss on the *Kdm5c*-KO brain's transcriptome by performing RNA-seq in the P6 cortex and hippocampus of male wild-type (WT), *Kdm5c*-KO (5C), *Dazl*-HET (DZ), and *Kdm5c*-KO;*Dazl*-HET double mutant (DM) mice. We then identified which genes are differentially expressed from wild type in each mutant genotype using DESeq2 ($q < 0.1$, $\log_2$ Fold Change > 0.2). As predicted given that *Dazl* has no known function in the brain, we saw very few changes in mice with heterozygous loss of *Dazl* alone. However, heterozygous loss of *Dazl* coupled with *Kdm5c*-KO blunted the transcription of

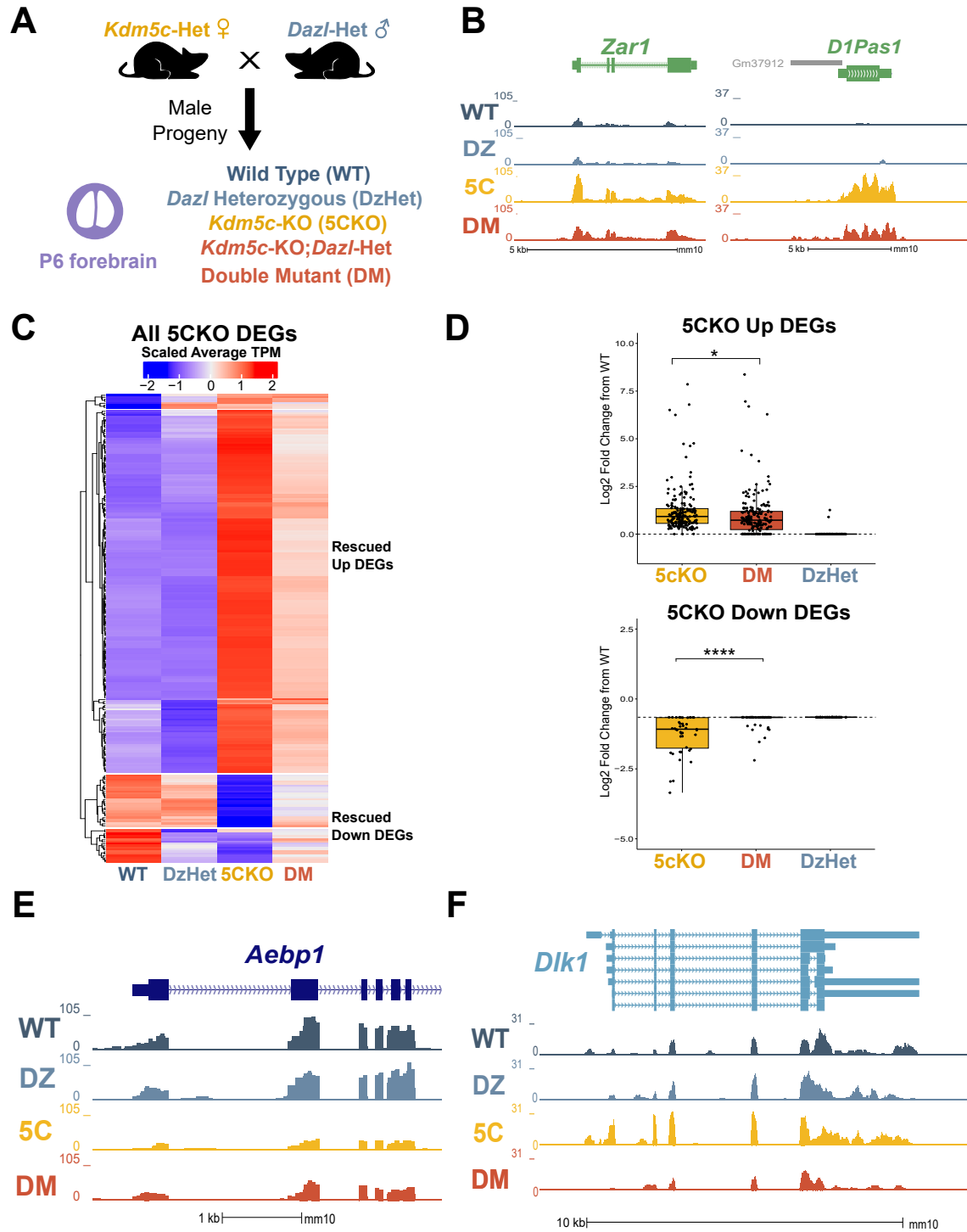


Figure 9: Additional heterozygous loss of *Dazl* ameliorates the transcriptional dysregulation in the *Kdm5c*-KO male cortex and hippocampus. **A.** Breeding diagram to generate the male progeny harvested for RNA-seq. **B.** Two example *Kdm5c*-KO germline DEGs whose expression is reduced in the double mutant brain **C.** Heatmap of the scaled expression (transcripts per million, TPM) in all genotypes of all genes significantly dysregulated in the *Kdm5c*-KO brain. **D.** Boxplots of *Kdm5c*-KO upregulated (top) and downregulated (bottom) DEGs, plotting their log2 fold change (L2FC) from wild-type in the different genotypes. Welch's t-test, *Kdm5c*-KO versus double mutant, * = $p < 0.5$, ****, $p < 0.0001$. (Legend continued on next page.)

Figure 9: Additional heterozygous loss of *Dazl* ameliorates the transcriptional dysregulation in the *Kdm5c*-KO male cortex and hippocampus. (Legend continued.) **E.** Bigwigs of the gene *Aebp1* that is downregulated in the *Kdm5c*-KO brain but shows rescue in the double mutant brain. **F.** Bigwigs of the upregulated imprinted gene *Dlk1* that shows rescue in the double mutant brain.

many *Kdm5c*-KO germline DEGs (Figure 9B).

To assess the overall impact of *Dazl* mutation upon the *Kdm5c*-KO brain transcriptome, we performed hierarchical clustering on all *Kdm5c*-KO DEGs to assess how their relative expression changes in the four genotypes (Figure 9C). Excitingly, we saw global amelioration of *Kdm5c*-KO transcriptional dysregulation in *Kdm5c*-KO;*Dazl*-HET double mutants, including suppression of non-germline genes upregulated upregulated within the brain. Most surprisingly, genes downregulated in the *Kdm5c*-KO brain showed the most robust shift towards wild-type expression levels (Figure 9C-E). Dampening of non-germline genes downregulated and upregulated in the *Kdm5c*-KO brain indicates additional heterozygous loss of *Dazl* partially restores the *Kdm5c*-KO brain’s somatic identity and shifts its development towards a typical wild-type course. In support of *Dazl* loss restoring the *Kdm5c*-KO brain’s somatic identity, we also observed partial rescue of overexpressed imprinted genes (Figure 9F). However, more experiments are required to determine if this rescue truly reflects a restoration of somatic cellular identity or just further rescue of neurodevelopment, as *Dlk1* can be biallelically expressed in wild-type hippocampal neurons¹³³. Nevertheless, these results do suggest aberrant germline genes contribute to *Kdm5c* mutant neurodevelopmental impairments, as DAZL itself has no known function in the brain yet dysregulation of non-germline transcripts was dampened with additional heterozygous loss of *Dazl*. Ongoing experiments will determine if this transcriptional rescue is sufficient to ameliorate previously characterized *Kdm5c*-KO neuronal phenotypes, such as reduced dendritic spine density and behavioral impairments^{53,64}.

Discussion

In the above work we provide the first, preliminary evidence that aberrant transcription of germline genes can lead to functional consequences within *Kdm5c* mutant cells. We first characterized the manifestation of two germline-related phenotypes: dysregulation

of genomic imprinting and transposable element expression. Many canonically imprinted genes are dysregulated in *Kdm5c* mutant EpiLCs and some of these genes show sustained upregulation in the developing and mature *Kdm5c* mutant brain (Figure 7). Imprinted genes uniquely dysregulated in EpiLCs included *H19* and *Igf2*, which oppose one another to regulate body weight¹³⁶. Many imprinted genes are also known to play important roles in neurodevelopment¹³⁷ and dysregulation of genomic imprinting causes the NDDs Angelman’s syndrome and Prader-Willi syndrome¹³⁰. Among imprinted DEGs genes was *cyclin-dependent kinase inhibitor 1c (Cdkn1c)*, a paternally imprinted gene that encodes the cell cycle regulator p57^{KIP2}¹³¹. Interestingly, mice with global, transgenic overexpression of *Cdkn1c* have phenotypes similar to *Kdm5c*-KO mice, including small body size and enhanced social dominance in the tube test¹³⁸. Another imprinted DEG consistently dysregulated across datasets is *Delta-like homolog 1 (Dlk1)*, encodes a transmembrane regulator of cell growth and differentiation. However, its possible *Dlk1* overexpression in our brain bulk RNA-seq data reflects a developmental stage transition in the brain rather a dysregulation of genomic imprinting itself, given that *Dlk1* is biallelically expressed in neurons during hippocampal development¹³³. Our ongoing work employing allele-specific RNA sequencing will help distinguish these possibilities. Nevertheless, known roles for imprinted DEGs in neurodevelopment makes loss of genomic imprinting a plausible route by which *Kdm5c* mutant germ cell-like characteristics could functionally impact neurodevelopment.

We also found *Kdm5c* mutant cells aberrantly expressed many regulators of piRNA biogenesis (Figure 8). piRNAs highly expressed in germ cells where they maintain genome fidelity by repressing transposable element (TE) expression^{134,139}. However, some TEs are expressed during germ cell development in a sexually dimorphic manner and may actually aid in germ cell development¹⁴⁰. We found that while TEs were unchanged in the *Kdm5c* mutant brain, *Kdm5c* mutant EpiLCs displayed widespread dysregulation of TEs. Interestingly, TE dysregulation appears to be sexually dimorphic, as we found global suppression of TE expression in *Kdm5c*-KO males while *Kdm5c*-KO and HET females had both up and downregulated TEs.

Although piRNAs are heavily enriched in the germline, emerging evidence suggests weak piRNA expression in somatic cells, including neurons. One study found neuronal piRNAs

influence synaptic plasticity in *Aplysia*¹⁴¹, although to what extent this is true for mammalian neurons remains unclear¹⁴². While there is currently very little known about the relationship between piRNAs, TEs, and neurodevelopment, some studies have implicated piRNA and TE dysregulation in NDDs^{143,144}. Independent studies have found piRNAs are aberrantly expressed in people with ASD¹⁴⁵ and Rett Syndrome, another X-linked NDD caused by loss of a chromatin regulator¹⁴⁶. It is currently unknown if piRNA dysregulation is a cause or consequence of neurodevelopmental impairments, as their roles in unimpaired neurons are still unclear. One hypothesis linking piRNAs to neurodevelopment proposes that wild-type neurons actually permit TE expression during neurodevelopment to generate beneficial genetic diversity^{147,148}. However, this strategy is risky as too much TE expression can lead to deleterious mutations that impair neuronal function. Active research is ongoing to untangle the role of piRNAs and TEs in neuronal development and further characterization of their expression in *Kdm5c* mutants may shed light onto this phenomenon.

Importantly, although these results provide early evidence of impaired genomic imprinting and TE regulation in *Kdm5c* mutant somatic cells, it is currently unclear if these germline-related phenotypes are due to the misexpression of germline genes, a direct consequence of KDM5C loss itself, or a synergistic effect between the two. Both genomic imprinting and TE repression depend upon *de novo* placement of CpG DNA methylation, which can be blocked by KDM5C's substrate H3K4me3. While it is unknown if DNA methylation is impacted at *Kdm5c* mutant ICRs or TE repeats, DNA methylation is lost at the promoters of at least two germline genes in the *Kdm5c*-KO brain⁴². Additional experiments are required to tease apart these contributions, such as investigating imprinted gene and TE expression when germline genes are repressed in *Kdm5c*-KO cells through alternative approaches.

We also evaluated the overall impact of germline gene misexpression upon the developing *Kdm5c*-KO brain transcriptome by performing RNA-seq on the postnatal day 6 hippocampus and cortex of *Kdm5c*-KO;*Dazl*-HET double mutant mice. Given that DAZL promotes the translation of many *Kdm5c*-KO germline DEGs, we predicted additional *Dazl* loss would blunt the impact of aberrant germline transcripts during *Kdm5c*-KO development. We found *Kdm5c*-KO;*Dazl*-HET double mutation globally ameliorated the transcriptional dysregulation caused by *Kdm5c*-KO alone, dampening the dysregulation of both germline and

non-germline DEGs. Surprisingly, *Dazl* loss appeared to have the strongest effect upon genes downregulated in the *Kdm5c*-KO brain. As KDM5C is primarily viewed as a transcriptional repressor⁵¹, downregulated genes are thought to be caused by secondary effects of KDM5C loss. Therefore, their rescue in *Kdm5c*-KO;*Dazl*-HET double mutants suggests *Dazl* loss shifted the *Kdm5c*-KO brain towards a wild-type developmental trajectory. *Dazl* mutation also dampened the overexpression of the imprinted gene *Dlk1*, indicating blunting germline gene misexpression can potentially rescue *Kdm5c*-KO germline-related phenotypes. However, it is currently unknown if this partial restoration of the *Kdm5c*-KO brain transcriptome is sufficient to rescue *Kdm5c*-KO neuronal phenotypes. We have previously shown *Kdm5c*-KO mice have decreased dendritic arborization in the amygdala and hippocampus and have behavioral impairments in learning and memory, dominance and aggression, and social interactions^{64,65}. The manifestation of these phenotypes in *Kdm5c*-KO;*Dazl*-HET double mutants and whether germline gene de-repression occurs in human patients will ultimately inform the therapeutic potential of suppressing germline genes in MRXSCJ patients and related disorders. Additionally, because our data suggests DAZL is only expressed very early in *Kdm5c* mutant development, therapeutic interventions would likely involve targeting particularly harmful germline genes expressed in the brain later in life. Given these germline genes have no function in the brain, future work identifying particularly harmful germline genes could offer ideal drug targets with little predicted off-target effects.

Chapter 6 - Conclusions and Future Directions

The above work interrogates the impact of sex and cellular identity in chromatin-linked neurodevelopmental disorders through the investigation of KDM5C. We found KDM5C loss impacts the expression of both shared and distinct genes within the male and female mouse brain. Some genes with sexually dimorphic dysregulation have known roles in learning and aggression and may impact the manifestation of MRXSCJ-related phenotypes between males and females. We additionally found both male and female *Kdm5c* mutant brains misexpress multiple tissue-specific genes, many of which only have known functions in germ cells of the testis and ovary. Misexpression of germline genes within in the brain is due to loss of KDM5C-mediated germline gene suppression during early embryogenesis, resulting in the aberrant expression of master regulators of germline identity in male and females. This embryonic dysregulation of germline genes is exacerbated by retinoic acid signaling, leading to germline gene misexpression in multiple *Kdm5c* mutant somatic tissues in a tissue-dependent manner. We provided early evidence of germline-like phenotypes within *Kdm5c* mutant somatic cells, including the dysregulation of imprinted genes and transposable element expression. Finally, we found additional loss of the germline master regulator *Dazl* dampened the transcriptional dysregulation observed in the developing *Kdm5c*-KO brain, including rescue of non-germline genes. The culmination of this work demonstrates KDM5C loss leads to both shared and sexually dimorphic consequences during neurodevelopment,

provides insight into the demarcation between somatic and germline identity, and offers a novel route by which chromatin regulators are essential for neurodevelopment through their maintenance of tissue identity.

Shared and distinct gene expression changes between *Kdm5c* mutant males and females highlights the unique relationship between the X chromosome, sexual dimorphism, reproduction, and neurodevelopment. The mammalian X and Y chromosomes evolved from an ancestral autosomal pair that diverged and acquired sex-influencing genes over time, such as the Y chromosome gene *Sry* that triggers testis development¹⁴⁹. Given only males inherit the Y chromosome, it is not surprising many genes beneficial for male-specific development were fixed on the Y chromosome across evolution. However, genes important for male-specific reproduction are also enriched on the X chromosome, including genes fundamental for the early stages of spermatogenesis. One potential explanation for male gene enrichment on the X chromosome is hemizygous exposure - newly evolved recessive genes beneficial for male reproduction are exposed in XY males but would be masked by the dominant allele if evolved on autosomes¹⁵⁰. These genes can then become rapidly fixed on the X chromosome due to improved male reproductive fitness. However, genes don't have to have sexually dimorphic functions to become fixed on the X chromosome. While most genes are distributed among the autosomes at random, genes important for brain function in both males and females are also significantly enriched on the X chromosome¹⁵¹. Mammals harboring novel mutations that improve cognition would also have improved reproductive fitness due to enhanced social communication and problem solving. X chromosome hemizygosity in males increases the likelihood that these brain genes cause phenotypic consequences, enabling positive selection to occur more rapidly and ultimately become fixed within the genome¹⁵². Therefore, both sex-specific and sex-independent genes that are beneficial for fitness and reproduction can quickly accumulate on the X chromosome. Enrichment of both reproductive and neuronal genes on the X chromosome may explain why KDM5C, the only H3K4 demethylase on the mammalian X chromosome, has evolved to regulate these processes. Further work investigating the evolution of KDM5s and whether ancestral KDM5C initially emerged on the X chromosome or autosomes will further illuminate this relationship.

Due to erosion of the Y chromosome over evolution, relatively few X chromosome genes

have surviving Y chromosome homologs. However, the functions of X-Y pairs are not always conserved over time and their divergence could contribute to sexual dimorphism. For example, KDM5C's Y chromosome homolog, KDM5D¹⁵³, is a much weaker H3K4 demethylase than KDM5C⁵¹. Additionally, if KDM5D has a role in neurodevelopment, its function is likely distinct from KDM5C given that it is unable to compensate for hemizygous loss of *KDM5C* in males with MRXSCJ. Differences in function between KDM5C and KDM5D makes KDM5C's partial escape from X chromosome inactivation particularly interesting. Since KDM5D is unable to compensate for KDM5C loss in males, increased KDM5C expression in females suggests females require more KDM5C than males for normal brain function. This is supported by the fact that heterozygous loss of KDM5C in females results in gene expression changes and behavioral phenotypes, even though their KDM5C level should be equivalent to unaffected males. Females also appear to require more KDM5C than males for germline gene repression, since *Kdm5c*-HET female EpiLCs have about double the number of germline DEGs than *Kdm5c*-KO males, even though they still express some functional KDM5C. Interestingly, this indicates female EpiLCs require more KDM5C to maintain somatic cellular identity and are therefore more prone to transitioning to a germ cell-like state than males. Thus, KDM5C's repression of germline identity may be directly linked to its sexually dimorphic expression.

While KDM5C does have some sexually dimorphic functions, our data indicates KDM5C acts during early embryogenesis in both males and females to repress germline genes when the initial soma-germline boundary is established in epiblast stem cells. Intriguingly, KDM5C was recently demonstrated to facilitate X chromosome inactivation (XCI) in females during this developmental stage through the activation of an *Xist* enhancer⁵⁴. Expression of both X chromosomes in female *Kdm5c* -KO cells could itself be a germline phenotype, as female germ cells re-activate their inactive X chromosome to undergo meiosis⁹⁹. Thus, KDM5C may act a central repressor of the germ cell state within somatic tissues, rather than germline phenotypes being a secondary consequence of somatic cells expressing germ cell gene expression programs. Elucidating if KDM5C plays direct roles in imprinted gene regulation and transposable element silencing will shed light on this distinction. Of note, even though many germline genes are on the X and Y chromosomes, germline DEGs in both males and females are

were found on sex chromosomes and autosomes. This indicates germline genes expression in *Kdm5c* mutant females is not simply a consequence of failed XCI. If KDM5C does repress the germ cell state, KDM5C should either be unexpressed in germ cells or its function changed through alternative splicing or association with distinct co-factors compared to somatic cells. Further investigation into KDM5C's role in somatic versus germ cells may illuminate key mechanisms underlying multicellularity and help improve *in vitro* germ cell culture models and fertility treatments.

KDM5C's role in tissue identity is not limited to germline gene repression, as we observed the misexpression of many other somatic tissue-specific genes within the male and female *Kdm5c* mutant brain. Interestingly, neurons and germ cells are currently the only known cell types with a specific mechanism to prevent their genes' ectopic transcription. Analogous to PRC1.6's role in decommissioning germline genes, neuronal genes are actively repressed outside the nervous system by the REST (RE-1 silencing transcription factor) complex^{154,155}. The importance of decommissioning neuronal and germline genes may lie in the fact that the brain and the testes have the two most specialized transcriptomes in the body^{72,122}, and thus all the other organs may suppress these germline or neuronal genes for proper tissue function. Alternatively, other tissue or cell-specific complexes may exist but have gone unnoticed given the misexpression of their relatively few genes would be less apparent compared to neuronal and germline genes. Nevertheless, our data indicates KDM5C plays a pivotal role in suppressing multiple tissue-specific genes. In support of this, KDM5C is also associated with the REST complex to repress neuronal genes in non-neuronal cells¹¹². Therefore, KDM5C may have a currently unexplored, ancestral role in the development of tissue identity and multicellularity.

Currently, the most pressing question unaddressed by the above work is: does loss of *KDM5C* lead to germline gene misexpression in humans? Ongoing experiments in our lab aim to address this question through MRXSCJ patient blood samples and human ESC cultures. While it is unclear if KDM5C's role in germline gene repression is conserved between mice and humans, there are other chromatin-linked neurodevelopmental disorders that display misexpression of germline genes in humans¹⁵⁶. Loss of the *de novo* DNA methyltransferase DNMT3B causes the NDD Immunodeficiency, Centromere instability, Facial anomalies

syndrome 1 (ICF1 MIM: #242860)¹⁵⁷. Phenotypes of ICF1 are highly heterogeneous with varying degrees of penetrance but commonly include facial dysmorphism, recurrent infections, and intellectual disability^{158,159}. Mouse models of ICF syndrome have ectopic germline gene expression throughout the body, including the brain¹⁰⁷, and ICF1 patient blood cells show hypomethylation and expression of several meiotic germline genes. Aberrant germline transcripts are also found in other mouse models of chromatin-linked NDDs, such as *Mecp2*, *G9a/Glp*¹⁵⁶. However, it is currently unknown if these aberrant germline transcripts contribute to patient symptoms or mouse phenotypes. Thus, elucidating the impact of ectopic germline transcripts with loss of KDM5C will provide a novel lens with which to view these related disorders.

References

1. Diagnostic and statistical manual of mental disorders: DSM-5 (2013). 5th ed. (American psychiatric association).
2. Yang, Y., Zhao, S., Zhang, M., Xiang, M., Zhao, J., Chen, S., Wang, H., Han, L., and Ran, J. (2022). Prevalence of neurodevelopmental disorders among US children and adolescents in 2019 and 2020. *Front. Psychol.* *13*, 997648. 10.3389/fpsyg.2022.997648.
3. Loke, H., Harley, V., and Lee, J. (2015). Biological factors underlying sex differences in neurological disorders. *The International Journal of Biochemistry & Cell Biology* *65*, 139–150. 10.1016/j.biocel.2015.05.024.
4. Van Bokhoven, H. (2011). Genetic and Epigenetic Networks in Intellectual Disabilities. *Annu. Rev. Genet.* *45*, 81–104. 10.1146/annurev-genet-110410-132512.
5. Stevenson, R.E., and Schwartz, C.E. (2009). X-linked intellectual disability: Unique vulnerability of the male genome. *Dev Disabil Res Revs* *15*, 361–368. 10.1002/ddrr.81.
6. Hull, L., Petrides, K.V., and Mandy, W. (2020). The Female Autism Phenotype and Camouflaging: A Narrative Review. *Rev J Autism Dev Disord* *7*, 306–317. 10.1007/s40489-020-00197-9.
7. Cuppens, T., Shatto, J., Mangnier, L., Kumar, A.A., Ng, A.C.-H., Kaur, M., Bui, T.A., Leclercq, M., Droit, A., Dunham, I., et al. (2023). Sex difference contributes to phenotypic diversity in individuals with neurodevelopmental disorders. *Front Pediatr* *11*, 1172154. 10.3389/fped.2023.1172154.

8. De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* *515*, 209–215. 10.1038/nature13772.
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 Neuropsychopharmacol Biol Psychiatry* *84*, 306–327. 10.1016/j.pnpbp.2017.12.013.
10. Iwase, S., Bérubé, N.G., Zhou, Z., Kasri, N.N., Battaglioli, E., Scandaglia, M., and Barco, A. (2017). Epigenetic Etiology of Intellectual Disability. *J. Neurosci.* *37*, 10773–10782. 10.1523/JNEUROSCI.1840-17.2017.
11. Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* *389*, 251–260. 10.1038/38444.
12. Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* *403*, 41–45. 10.1038/47412.
13. Jenuwein, T., and Allis, C.D. (2001). Translating the Histone Code. *Science* *293*, 1074–1080. 10.1126/science.1063127.
14. Liu, C.L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S.L., Friedman, N., and Rando, O.J. (2005). Single-Nucleosome Mapping of Histone Modifications in *S. cerevisiae*. *PLoS Biol* *3*, e328. 10.1371/journal.pbio.0030328.
15. Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* *129*, 823–837. 10.1016/j.cell.2007.05.009.
16. Howe, F.S., Fischl, H., Murray, S.C., and Mellor, J. (2017). Is H3K4me3 instructive for transcription activation? *Bioessays* *39*, 1–12. 10.1002/bies.201600095.

17. Ding, Y., Ndamukong, I., Xu, Z., Lapko, H., Fromm, M., and Avramova, Z. (2012). ATX1-Generated H3K4me3 Is Required for Efficient Elongation of Transcription, Not Initiation, at ATX1-Regulated Genes. *PLoS Genet* 8, e1003111. 10.1371/journal.pgen.1003111.
18. Wang, H., Fan, Z., Shliaha, P.V., Miele, M., Hendrickson, R.C., Jiang, X., and Helin, K. (2023). H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. *Nature* 615, 339–348. 10.1038/s41586-023-05780-8.
19. Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W., van Schaik, F.M., Varier, R.A., Baltissen, M.P., Stunnenberg, H.G., Mann, M., and Timmers, H.T. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131, 58–69. 10.1016/j.cell.2007.08.016.
20. Lauberth, S.M., Nakayama, T., Wu, X., Ferris, A.L., Tang, Z., Hughes, S.H., and Roeder, R.G. (2013). H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* 152, 1021–1036. 10.1016/j.cell.2013.01.052.
21. Vallianatos, C.N., and Iwase, S. (2015). Disrupted intricacy of histone H3K4 methylation in neurodevelopmental disorders. *Epigenomics* 7, 503–519. 10.2217/epi.15.1.
22. Sugino, K., Hempel, C.M., Okaty, B.W., Arnson, H.A., Kato, S., Dani, V.S., and Nelson, S.B. (2014). Cell-Type-Specific Repression by Methyl-CpG-Binding Protein 2 Is Biased toward Long Genes. *Journal of Neuroscience* 34, 12877–12883. 10.1523/JNEUROSCI.2674-14.2014.
23. Gabel, H.W., Kinde, B., Stroud, H., Gilbert, C.S., Harmin, D.A., Kastan, N.R., Hemberg, M., Ebert, D.H., and Greenberg, M.E. (2015). Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature* 522, 89–93. 10.1038/nature14319.
24. Dhar, S.S., Zhao, D., Lin, T., Gu, B., Pal, K., Wu, S.J., Alam, H., Lv, J., Yun, K., Gopalakrishnan, V., et al. (2018). MLL4 Is Required to Maintain Broad H3K4me3 Peaks and Super-Enhancers at Tumor Suppressor Genes. *Molecular Cell* 70, 825–841.e6. 10.1016/j.molcel.2018.04.028.

25. Tillotson, R., Cholewa-Waclaw, J., Chhatbar, K., Connelly, J.C., Kirschner, S.A., Webb, S., Koerner, M.V., Selfridge, J., Kelly, D.A., De Sousa, D., et al. (2021). Neuronal non-CG methylation is an essential target for MeCP2 function. *Mol Cell* *81*, 1260–1275.e12. 10.1016/j.molcel.2021.01.011.
26. Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* *276*, 565–570. 10.1038/276565a0.
27. 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. 10.1073/pnas.85.21.8136.
28. Kassis, J.A., Kennison, J.A., and Tamkun, J.W. (2017). Polycomb and Trithorax Group Genes in *Drosophila*. *Genetics* *206*, 1699–1725. 10.1534/genetics.115.185116.
29. Ingham, P.W. (1981). Trithorax: A new homoeotic mutation of *Drosophila melanogaster* : II. The role of trx. *Wilehm Roux Arch Dev Biol* *190*, 365–369. 10.1007/BF00863275.
30. Ingham, P.W. (1985). A clonal analysis of the requirement for the trithorax gene in the diversification of segments in *Drosophila*. *J Embryol Exp Morphol* *89*, 349–365.
31. Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell* *10*, 1107–1117. 10.1016/s1097-2765(02)00741-4.
32. Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* *111*, 185–196. 10.1016/s0092-8674(02)00975-3.
33. Lindroth, A.M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I., et al. (2004). Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. *EMBO J* *23*, 4286–4296. 10.1038/sj.emboj.7600430.

34. Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* *441*, 349–353. 10.1038/nature04733.
35. CH, W. (1957). *The Strategy of Genes* (George Allen & Unwin Ltd).
36. Noble, D. (2015). Conrad Waddington and the origin of epigenetics. *Journal of Experimental Biology* *218*, 816–818. 10.1242/jeb.120071.
37. Kwan, K.Y., Sestan, N., and Anton, E.S. (2012). Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development* *139*, 1535–1546. 10.1242/dev.069963.
38. Lee, T.W., and Katz, D.J. (2020). Hansel, Gretel, and the Consequences of Failing to Remove Histone Methylation Breadcrumbs. *Trends Genet* *36*, 160–176. 10.1016/j.tig.2019.12.004.
39. 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 Astrogenic Fate Transition. *Neuron* *63*, 600–613. 10.1016/j.neuron.2009.08.021.
40. Delgado, R.N., Mansky, B., Ahanger, S.H., Lu, C., Andersen, R.E., Dou, Y., Alvarez-Buylla, A., and Lim, D.A. (2020). Maintenance of neural stem cell positional identity by *Mixed-Lineage Leukemia 1*. *Science* *368*, 48–53. 10.1126/science.aba5960.
41. 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. 10.1016/j.neuron.2009.11.019.
42. 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. 10.1016/j.celrep.2017.09.014.

43. Yang, X., Schadt, E.E., Wang, S., Wang, H., Arnold, A.P., Ingram-Drake, L., Drake, T.A., and Lusis, A.J. (2006). Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res.* *16*, 995–1004. 10.1101/gr.5217506.
44. Lowe, R., Gemma, C., Rakyan, V.K., and Holland, M.L. (2015). Sexually dimorphic gene expression emerges with embryonic genome activation and is dynamic throughout development. *BMC Genomics* *16*, 295. 10.1186/s12864-015-1506-4.
45. Oliva, M., Muñoz-Aguirre, M., Kim-Hellmuth, S., Wucher, V., Gewirtz, A.D.H., Cotter, D.J., Parsana, P., Kasela, S., Balliu, B., Viñuela, A., et al. (2020). The impact of sex on gene expression across human tissues. *Science* *369*, eaba3066. 10.1126/science.aba3066.
46. 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 paraplegia. *Am J Med Genet* *94*, 1–4.
47. 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. 10.1086/427563.
48. 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. 10.1111/cge.13755.
49. Ropers, H.-H., and Hamel, B.C.J. (2005). X-linked mental retardation. *Nat Rev Genet* *6*, 46–57. 10.1038/nrg1501.
50. De Brouwer, A.P.M., Yntema, H.G., Kleefstra, T., Lugtenberg, D., Oudakker, A.R., De Vries, B.B.A., Van Bokhoven, H., Van Esch, H., Frints, S.G.M., Froyen, G., et al. (2007). Mutation frequencies of X-linked mental retardation genes in families from the EuroMRX consortium. *Hum. Mutat.* *28*, 207–208. 10.1002/humu.9482.

51. 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. 10.1016/j.cell.2007.02.017.
52. Brookes, E., Laurent, B., Öunap, K., Carroll, R., Moeschler, J.B., Field, M., Schwartz, C.E., Gecz, J., and Shi, Y. (2015). Mutations in the intellectual disability gene KDM5C reduce protein stability and demethylase activity. *Human Molecular Genetics* *24*, 2861–2872. 10.1093/hmg/ddv046.
53. 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. 10.1016/j.celrep.2015.12.091.
54. 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. 10.1038/s41467-022-30352-1.
55. Ohno, S. (1967). *Sex Chromosomes and Sex-Linked Genes* 1966th ed. (Springer).
56. Pessia, E., Makino, T., Bailly-Bechet, M., McLysaght, A., and Marais, G.A.B. (2012). Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. *Proc. Natl. Acad. Sci. U.S.A.* *109*, 5346–5351. 10.1073/pnas.1116763109.
57. 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 and human. *Hum Mol Genet* *3*, 879–884. 10.1093/hmg/3.6.879.
58. 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. 10.1093/hmg/5.9.1361.

59. 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* *5*, 1355–1360. 10.1093/hmg/5.9.1355.
60. 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. 10.1371/journal.pone.0002553.
61. Link, J.C., Wiese, C.B., Chen, X., Avetisyan, R., Ronquillo, E., Ma, F., Guo, X., Yao, J., Allison, M., Chen, Y.-D.I., et al. (2020). X chromosome dosage of histone demethylase KDM5C determines sex differences in adiposity. *Journal of Clinical Investigation* *130*, 5688–5702. 10.1172/JCI140223.
62. Liu, H., Zhai, L., Liu, Y., Lu, D., Vander Ark, A., Yang, T., and Krawczyk, C.M. (2023). The histone demethylase KDM5C controls female bone mass by promoting energy metabolism in osteoclasts. *Sci. Adv.* *9*, eadg0731. 10.1126/sciadv.adg0731.
63. Simensen, R.J., Rogers, R.C., Collins, J.S., Abidi, F., Schwartz, C.E., and Stevenson, R.E. (2012). Short-term memory deficits in carrier females with KDM5C mutations. *Genet Couns* *23*, 31–40.
64. 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. 10.1038/s42003-020-1001-6.
65. Bonefas, K.M., Vallianatos, C.N., Raines, B., Tronson, N.C., and Iwase, S. (2023). Sexually Dimorphic Alterations in the Transcriptome and Behavior with Loss of Histone Demethylase KDM5C. *Cells* *12*, 637. 10.3390/cells12040637.
66. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* *15*, 550. 10.1186/s13059-014-0550-8.
67. Stephens, M. (2016). False discovery rates: A new deal. *Biostat*, kxw041. 10.1093/biostatistics/kxw041.

68. 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. 10.1093/bioinformatics/btx364.
69. Jin, Y., Tam, O.H., Paniagua, E., and Hammell, M. (2015). TEtranscripts: A package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* *31*, 3593–3599. 10.1093/bioinformatics/btv422.
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. 10.1016/j.molcel.2010.05.004.
71. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* *10*, 1523. 10.1038/s41467-019-09234-6.
72. Li, B., Qing, T., Zhu, J., Wen, Z., Yu, Y., Fukumura, R., Zheng, Y., Gondo, Y., and Shi, L. (2017). A Comprehensive Mouse Transcriptomic BodyMap across 17 Tissues by RNA-seq. *Sci Rep* *7*, 4200. 10.1038/s41598-017-04520-z.
73. Shen, R., Li, Y., Liang, A., Li, S., Yang, C., and Huang, H. (2022). A female case with novel KDM5C heterozygous variation presenting with Claes-Jensen type-like phenotype. *BMC Neurol* *22*, 491. 10.1186/s12883-022-03023-3.
74. Vallianatos, C.N., Farrehi, C., Friez, M.J., Burmeister, M., Keegan, C.E., and Iwase, S. (2018). Altered Gene-Regulatory Function of KDM5C by a Novel Mutation Associated With Autism and Intellectual Disability. *Front Mol Neurosci* *11*, 104. 10.3389/fnmol.2018.00104.
75. Thion, M.S., and Garel, S. (2017). On place and time: Microglia in embryonic and perinatal brain development. *Current Opinion in Neurobiology* *47*, 121–130. 10.1016/j.conb.2017.10.004.

76. Nadarajah, B., Thomaïdou, D., Evans, W.H., and Parnavelas, J.G. (1996). Gap junctions in the adult cerebral cortex: Regional differences in their distribution and cellular expression of connexins. *J Comp Neurol* *376*, 326–342. 10.1002/(SICI)1096-9861(19961209)376:2<326::AID-CNE13>3.0.CO;2-J.
77. Nadarajah, B., Jones, A.M., Evans, W.H., and Parnavelas, J.G. (1997). Differential expression of connexins during neocortical development and neuronal circuit formation. *J Neurosci* *17*, 3096–3111.
78. Li, J., Shen, H., Naus, C.C., Zhang, L., and Carlen, P.L. (2001). Upregulation of gap junction connexin 32 with epileptiform activity in the isolated mouse hippocampus. *Neuroscience* *105*, 589–598. 10.1016/s0306-4522(01)00204-4.
79. Goldstein, M., Fuxe, K., and Hökfelt, T. (1972). Characterization and tissue localization of catecholamine synthesizing enzymes. *Pharmacol Rev* *24*, 293–309.
80. Sørensen, D.B., Johnsen, P.F., Bibby, B.M., Böttner, A., Bornstein, S.R., Eisenhofer, G., Pacak, K., and Hansen, A.K. (2005). PNMT transgenic mice have an aggressive phenotype. *Horm Metab Res* *37*, 159–163. 10.1055/s-2005-861301.
81. Hökfelt, T., and Tatemoto, K. (2008). Galanin – 25 years with a multitasking neuropeptide: Galanin – 25 years with a multitasking neuropeptide. *Cell. Mol. Life Sci.* *65*, 1791–1795. 10.1007/s00018-008-8152-9.
82. Steiner, R.A., Hohmann, J.G., Holmes, A., Wrenn, C.C., Cadd, G., Juréus, A., Clifton, D.K., Luo, M., Gutshall, M., Ma, S.Y., et al. (2001). Galanin transgenic mice display cognitive and neurochemical deficits characteristic of Alzheimer’s disease. *Proc Natl Acad Sci U S A* *98*, 4184–4189. 10.1073/pnas.061445598.
83. Wrenn, C.C., Marriott, L.K., Kinney, J.W., Holmes, A., Wenk, G.L., and Crawley, J.N. (2002). Galanin peptide levels in hippocampus and cortex of galanin-overexpressing transgenic mice evaluated for cognitive performance. *Neuropeptides* *36*, 413–426. 10.1016/s0143-4179(02)00111-7.

84. Rustay, N.R., Wrenn, C.C., Kinney, J.W., Holmes, A., Bailey, K.R., Sullivan, T.L., Harris, A.P., Long, K.C., Saavedra, M.C., Starosta, G., et al. (2005). Galanin impairs performance on learning and memory tasks: Findings from galanin transgenic and GAL-R1 knockout mice. *Neuropeptides* *39*, 239–243. 10.1016/j.npep.2004.12.026.
85. Tian, D., Sun, S., and Lee, J.T. (2010). The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* *143*, 390–403. 10.1016/j.cell.2010.09.049.
86. Sun, S., Del Rosario, B.C., Szanto, A., Ogawa, Y., Jeon, Y., and Lee, J.T. (2013). Jpx RNA activates Xist by evicting CTCF. *Cell* *153*, 1537–1551. 10.1016/j.cell.2013.05.028.
87. Winkler, C., and Yao, S. (2014). The midkine family of growth factors: Diverse roles in nervous system formation and maintenance. *Br J Pharmacol* *171*, 905–912. 10.1111/bph.12462.
88. Muramatsu, H., Shirahama, H., Yonezawa, S., Maruta, H., and Muramatsu, T. (1993). Midkine, a retinoic acid-inducible growth/differentiation factor: Immunohistochemical evidence for the function and distribution. *Dev Biol* *159*, 392–402. 10.1006/dbio.1993.1250.
89. Siep, M. (2004). Basic helix-loop-helix transcription factor Tcf15 interacts with the Calmegin gene promoter in mouse spermatogenesis. *Nucleic Acids Research* *32*, 6425–6436. 10.1093/nar/gkh979.
90. Cheng, S., Altmeyden, G., So, C., Welp, L.M., Penir, S., Ruhwedel, T., Menelaou, K., Harasimov, K., Stützer, A., Blayney, M., et al. (2022). Mammalian oocytes store mRNAs in a mitochondria-associated membraneless compartment. *Science* *378*, eabq4835. 10.1126/science.abq4835.
91. Mueller, J.L., Skaletsky, H., Brown, L.G., Zaghlul, S., Rock, S., Graves, T., Auger, K., Warren, 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. 10.1038/ng.2705.

92. Green, C.D., Ma, Q., Manske, G.L., Shami, A.N., Zheng, X., Marini, S., Moritz, L., Sultan, C., Gurczynski, S.J., Moore, B.B., et al. (2018). A Comprehensive Roadmap of Murine Spermatogenesis Defined by Single-Cell RNA-Seq. *Dev Cell* *46*, 651–667.e10. 10.1016/j.devcel.2018.07.025.
93. Poeta, L., Padula, A., Lioi, M.B., Van Bokhoven, H., and Miano, M.G. (2021). Analysis of a Set of KDM5C Regulatory Genes Mutated in Neurodevelopmental Disorders Identifies Temporal Coexpression Brain Signatures. *Genes* *12*, 1088. 10.3390/genes12071088.
94. Kleopa, K.A., Abrams, C.K., and Scherer, S.S. (2012). How do mutations in GJB1 cause X-linked Charcot-Marie-Tooth disease? *Brain Res* *1487*, 198–205. 10.1016/j.brainres.2012.03.068.
95. Bowser, R., Kordower, J.H., and Mufson, E.J. (1997). A confocal microscopic analysis of galaninergic hyperinnervation of cholinergic basal forebrain neurons in Alzheimer’s disease. *Brain Pathol* *7*, 723–730. 10.1111/j.1750-3639.1997.tb01058.x.
96. Inoue, H., Ogonuki, N., Hirose, M., Hatanaka, Y., Matoba, S., Chuma, S., Kobayashi, K., Wakana, S., Noguchi, J., Inoue, K., et al. (2016). Mouse D1Pas1, a DEAD-box RNA helicase, is required for the completion of first meiotic prophase in male germ cells. *Biochemical and Biophysical Research Communications* *478*, 592–598. 10.1016/j.bbrc.2016.07.109.
97. Goldberg, E., Sberna, D., Wheat, T.E., Urbanski, G.J., and Margoliash, E. (1977). Cytochrome c: Immunofluorescent Localization of the Testis-Specific Form. *Science* *196*, 1010–1012. 10.1126/science.193188.
98. Magnúsdóttir, E., and Surani, M.A. (2014). How to make a primordial germ cell. *Development* *141*, 245–252. 10.1242/dev.098269.
99. 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*. 10.1098/rstb.2013.0543.
100. 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. 10.1016/j.mod.2020.103617.

101. Li, H., Liang, Z., Yang, J., Wang, D., Wang, H., Zhu, M., Geng, B., and Xu, E.Y. (2019). DAZL is a master translational regulator of murine spermatogenesis. *Natl Sci Rev* 6, 455–468. 10.1093/nsr/nwy163.
102. 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. 10.1073/pnas.0510813103.
103. 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. 10.1126/science.1166340.
104. 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. 10.3390/biom9120775.
105. 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. 10.7554/eLife.21064.
106. 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. 10.1038/s41467-021-27345-x.
107. 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. 10.1073/pnas.1000473107.
108. 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. 10.1242/dev.081661.

109. 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 autoinhibition and histone H3-induced activation of DNMT3A. *Nature* *517*, 640–644. 10.1038/nature13899.
110. Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. *EMBO Reports* *10*, 1235–1241. 10.1038/embor.2009.218.
111. 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). DAZL mediates a broad translational program regulating expansion and differentiation of spermatogonial progenitors. *eLife* *9*, e56523. 10.7554/eLife.56523.
112. Tahiliani, M., Mei, P., Fang, R., Leonor, T., Rutenberg, M., Shimizu, F., Li, J., Rao, A., and Shi, Y. (2007). The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. *Nature* *447*, 601–605. 10.1038/nature05823.
113. Gupta, N., Yakhou, L., Albert, J.R., Azogui, A., Ferry, L., Kirsh, O., Miura, F., Battault, S., Yamaguchi, K., Laisné, M., et al. (2023). A genome-wide screen reveals new regulators of the 2-cell-like cell state. *Nat Struct Mol Biol.* 10.1038/s41594-023-01038-z.
114. 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. 10.1038/ncomms11056.
115. Dahlet, T., Truss, M., Frede, U., Al Adhami, H., Bardet, A.F., Dumas, M., Vallet, J., Chicher, J., Hammann, P., Kottnik, S., et al. (2021). E2F6 initiates stable epigenetic silencing of germline genes during embryonic development. *Nat Commun* *12*, 3582. 10.1038/s41467-021-23596-w.

116. Huang, Y., Zhao, W., Wang, C., Zhu, Y., Liu, M., Tong, H., Xia, Y., Jiang, Q., and Qin, J. (2018). Combinatorial Control of Recruitment of a Variant PRC1.6 Complex in Embryonic Stem Cells. *Cell Rep* *22*, 3032–3043. 10.1016/j.celrep.2018.02.072.
117. 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* *14*, e1007193. 10.1371/journal.pgen.1007193.
118. Maden, M. (2007). Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nat Rev Neurosci* *8*, 755–765. 10.1038/nrn2212.
119. Duester, G. (2008). Retinoic Acid Synthesis and Signaling during Early Organogenesis. *Cell* *134*, 921–931. 10.1016/j.cell.2008.09.002.
120. Yuan, L., Liu, J.-G., Hoja, M.-R., Wilbertz, J., Nordqvist, K., and Höög, C. (2002). Female Germ Cell Aneuploidy and Embryo Death in Mice Lacking the Meiosis-Specific Protein SCP3. *Science* *296*, 1115–1118. 10.1126/science.1070594.
121. Engel, E., and DeLozier-Blanchet, C.D. (1991). Uniparental disomy, isodisomy, and imprinting: Probable effects in man and strategies for their detection. *Am J Med Genet* *40*, 432–439. 10.1002/ajmg.1320400411.
122. Consortium, G. (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* *348*, 648–660. 10.1126/science.1262110.
123. van den Heuvel, S., and Dyson, N.J. (2008). Conserved functions of the pRB and E2F families. *Nat Rev Mol Cell Biol* *9*, 713–724. 10.1038/nrm2469.
124. 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. 10.1038/nature04010.
125. Kawasaki, I., Shim, Y.-H., Kirchner, J., Kaminker, J., Wood, W.B., and Strome, S. (1998). PGL-1, a Predicted RNA-Binding Component of Germ Granules, Is Essential for Fertility in *C. elegans*. *Cell* *94*, 635–645. 10.1016/S0092-8674(00)81605-0.

126. Pitt, J.N., Schisa, J.A., and Priess, J.R. (2000). P Granules in the Germ Cells of *Caenorhabditis elegans* Adults Are Associated with Clusters of Nuclear Pores and Contain RNA. *Developmental Biology* 219, 315–333. 10.1006/dbio.2000.9607.
127. Sijen, T., and Plasterk, R.H.A. (2003). Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* 426, 310–314. 10.1038/nature02107.
128. 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. 10.1126/science.1195481.
129. Barlow, D.P., and Bartolomei, M.S. (2014). Genomic Imprinting in Mammals. *Cold Spring Harbor Perspectives in Biology* 6, a018382–a018382. 10.1101/cshperspect.a018382.
130. Cassidy, S.B., and Schwartz, S. (1998). Prader-Willi and Angelman syndromes. Disorders of genomic imprinting. *Medicine (Baltimore)* 77, 140–151. 10.1097/00005792-199803000-00005.
131. Andrews, S.C., Wood, M.D., Tunster, S.J., Barton, S.C., Surani, M.A., and John, R.M. (2007). Cdkn1c (p57Kip2) is the major regulator of embryonic growth within its imprinted domain on mouse distal chromosome 7. *BMC Dev Biol* 7, 53. 10.1186/1471-213X-7-53.
132. Imaizumi, Y., Furutachi, S., Watanabe, T., Miya, H., Kawaguchi, D., and Gotoh, Y. (2020). Role of the imprinted allele of the Cdkn1c gene in mouse neocortical development. *Sci Rep* 10, 1884. 10.1038/s41598-020-58629-9.
133. Montalbán-Loro, R., Lassi, G., Lozano-Ureña, A., Perez-Villalba, A., Jiménez-Villalba, E., Charalambous, M., Vallortigara, G., Horner, A.E., Saksida, L.M., Bussey, T.J., et al. (2021). *Dlk1* dosage regulates hippocampal neurogenesis and cognition. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2015505118. 10.1073/pnas.2015505118.
134. Tóth, K.F., Pezic, D., Stuwe, E., and Webster, A. (2016). The piRNA Pathway Guards the Germline Genome Against Transposable Elements. In *Non-coding RNA and the Reproductive System*, D. Wilhelm and P. Bernard, eds. (Springer Netherlands), pp. 51–77. 10.1007/978-94-017-7417-8_4.

135. Li, Y., Zhang, Y., and Liu, M. (2021). Knockout Gene-Based Evidence for PIWI-Interacting RNA Pathway in Mammals. *Front Cell Dev Biol* *9*, 681188. 10.3389/fcell.2021.681188.
136. Nordin, M., Bergman, D., Halje, M., Engström, W., and Ward, A. (2014). Epigenetic regulation of the *Igf2/H19* gene cluster. *Cell Proliferation* *47*, 189–199. 10.1111/cpr.12106.
137. Ho-Shing, O., and Dulac, C. (2019). Influences of genomic imprinting on brain function and behavior. *Current Opinion in Behavioral Sciences* *25*, 66–76. 10.1016/j.cobeha.2018.08.008.
138. McNamara, G.I., Davis, B.A., Browne, M., Humby, T., Dalley, J.W., Xia, J., John, R.M., and Isles, A.R. (2018). Dopaminergic and behavioural changes in a loss-of-imprinting model of *Cdkn1c*. *Genes Brain Behav* *17*, 149–157. 10.1111/gbb.12422.
139. Zamudio, N., and Bourc’his, D. (2010). Transposable elements in the mammalian germline: A comfortable niche or a deadly trap? *Heredity* *105*, 92–104. 10.1038/hdy.2010.53.
140. Maupetit-Mehouas, S., and Vaury, C. (2020). Transposon Reactivation in the Germline May Be Useful for Both Transposons and Their Host Genomes. *Cells* *9*, 1172. 10.3390/cells9051172.
141. Rajasethupathy, P., Antonov, I., Sheridan, R., Frey, S., Sander, C., Tuschl, T., and Kandel, E.R. (2012). A Role for Neuronal piRNAs in the Epigenetic Control of Memory-Related Synaptic Plasticity. *Cell* *149*, 693–707. 10.1016/j.cell.2012.02.057.
142. Nandi, S., Chandramohan, D., Fioriti, L., Melnick, A.M., Hébert, J.M., Mason, C.E., Rajasethupathy, P., and Kandel, E.R. (2016). Roles for small noncoding RNAs in silencing of retrotransposons in the mammalian brain. *Proc. Natl. Acad. Sci. U.S.A.* *113*, 12697–12702. 10.1073/pnas.1609287113.
143. Chavda, V., Madhwani, K., and Chaurasia, B. (2022). PiWi RNA in Neurodevelopment and Neurodegenerative Disorders. *CMP* *15*, 517–531. 10.2174/1874467214666210629164535.

144. Zhang, S.-F., Gao, J., and Liu, C.-M. (2019). The Role of Non-Coding RNAs in Neurodevelopmental Disorders. *Front. Genet.* *10*, 1033. 10.3389/fgene.2019.01033.
145. Salloum-Asfar, S., Elsayed, A.K., Elhag, S.F., and Abdulla, S.A. (2021). Circulating Non-Coding RNAs as a Signature of Autism Spectrum Disorder Symptomatology. *IJMS* *22*, 6549. 10.3390/ijms22126549.
146. Saxena, A., Tang, D., and Carninci, P. (2012). piRNAs Warrant Investigation in Rett Syndrome: An Omics Perspective. *Disease Markers* *33*, 261–275. 10.1155/2012/396737.
147. Bodea, G.O., McKelvey, E.G.Z., and Faulkner, G.J. (2018). Retrotransposon-induced mosaicism in the neural genome. *Open Biol.* *8*, 180074. 10.1098/rsob.180074.
148. Erwin, J.A., Marchetto, M.C., and Gage, F.H. (2014). Mobile DNA elements in the generation of diversity and complexity in the brain. *Nat Rev Neurosci* *15*, 497–506. 10.1038/nrn3730.
149. Lahn, B.T., and Page, D.C. (1999). Four Evolutionary Strata on the Human X Chromosome. *Science* *286*, 964–967. 10.1126/science.286.5441.964.
150. Vallender, E.J., and Lahn, B.T. (2004). How mammalian sex chromosomes acquired their peculiar gene content. *BioEssays* *26*, 159–169. 10.1002/bies.10393.
151. Brand, B.A., Blesson, A.E., and Smith-Hicks, C.L. (2021). The Impact of X-Chromosome Inactivation on Phenotypic Expression of X-Linked Neurodevelopmental Disorders. *Brain Sciences* *11*, 904. 10.3390/brainsci11070904.
152. Vicoso, B., and Charlesworth, B. (2006). Evolution on the X chromosome: Unusual patterns and processes. *Nat Rev Genet* *7*, 645–653. 10.1038/nrg1914.
153. Agulnik, A.I., Longepied, G., Ty, M.T., Bishop, C.E., and Mitchell, M. (1999). Mouse H-Y encoding Smcy gene and its X chromosomal homolog Smcx. *Mamm Genome* *10*, 926–929. 10.1007/s003359901116.
154. Chong, J.A., Tapia-Ramírez, J., Kim, S., Toledo-Aral, J.J., Zheng, Y., Boutros, M.C., Altshuler, Y.M., Frohman, M.A., Kraner, S.D., and Mandel, G. (1995). REST: A mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell* *80*, 949–957. 10.1016/0092-8674(95)90298-8.