Novel Cell Intrinsic and Extrinsic Mechanisms of X-chromosome Inactivation

by

Marissa Cloutier

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Genetics and Genomics)
at the University of Michigan
2022

Doctoral Committee:

Associate Professor Sundeep Kalantry, Chair
Associate Professor Shigeki Iwase
Professor Sue Moenter
Associate Professor Jacob Mueller
Professor Maureen Sartor
Professor Yukiko Yamashita, Massachusetts Institute of Technology

Marissa Cloutier

mcloutie@umich.edu

ORCID iD: 0000-0001-7078-542X

Dedication

To my parents, Brenda and Bill, who have provided me with endless support through each step of my life and my academic journey.

Acknowledgements

I would like to first acknowledge my graduate mentor, Dr. Sundeep Kalantry, for his support and guidance throughout my doctoral training. Sundeep approaches his work with unparalleled scientific rigor, dedication, and creativity, and it has been a privilege to work with and learn from him during my time in the lab. I owe Sundeep many thanks for helping me develop into a confident and effective scientist, critical thinker, and communicator.

I also wish to thank my thesis committee members, Dr. Shigeki Iwase, Dr. Sue Moenter, Dr. Jacob Mueller, Dr. Maureen Sartor, and Dr. Yukiko Yamashita. It has been a privilege to learn from each of them during my training. Their feedback and vast knowledge of genetics, epigenetics, reproductive biology, molecular biology, and bioinformatics has been critical to my work and development as a scientist. Each committee member has contributed significantly to discussions during my thesis committee meetings and seminars, which have ultimately helped to shape and strengthen my dissertation research.

I would next like to acknowledge the many past and present Kalantry Lab members who have assisted and supported me in my training. Former graduate student Dr. Michael Hinten laid the groundwork for my investigation of Polycomb proteins and *Xist* in mice, as he contributed to experiments, collected embryos, and derived many of the cell lines that I analyzed. Dr. Emily Maclary, another former graduate student, taught me critical dissection and RNA-Seq analysis techniques during my early days in the lab. Clair Harris, our lab manager, trained me in many technical aspects of my work, including tissue culture, molecular biology techniques, and bioinformatics protocols. Valerie Sponyoe, our lab technician, maintained countless mouse

colonies and was instrumental in generating many of the embryos that were essential to my research. I am also grateful to undergraduate students Holly Ong, Reina Brodeur, and Allison Ingman for their work in maintaining the mouse colonies. I would also like to acknowledge the training and advice I received from our lab's postdoctoral researchers, Dr. Milan Samanta and Dr. Srimonta Gayen. Finally, I owe many, many thanks to fellow Kalantry Lab graduate students Megan Trotter, Rebecca Malcore, Itzaira Mercado-Hernandez, Dana Beseiso, Dr. Aaron Williams, and Kritika Kasliwal for their support, advice, and encouragement – both scientific and otherwise – during my time in the lab.

I would also like to thank the many other individuals, departments, and facilities at the University of Michigan that helped make my research and training possible. I am grateful to the Program in Biomedical Sciences (PIBS), the Genetics Training Program (T32-GM07544), the Career Training in Reproductive Biology Program (T32-HD079342), and the Rackham Predoctoral Fellowship for providing my funding. I also wish to thank Molly Martin for answering my many administrative questions, the Mueller Lab for their valuable feedback in our weekly lab meetings, and Dr. Gary Smith and his lab for their many contributions to my work in human embryonic stem cells. I would also like to acknowledge the many other past and present members of the University of Michigan Department of Human Genetics, Department of Bioinformatics, Center for RNA Biology, Reproductive Sciences Program, and Genetics Training Program for their feedback and encouragement throughout my training. The University of Michigan Sequencing Core, Transgenic Core, and Flow Cytometry Core facilities also deserve acknowledgements for all they contributed to my work.

Lastly, I would like to thank my wonderful support system. Without the encouragement of my family and friends, completing this degree would not have been possible.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

X-chromosome inactivation equalizes X-linked gene expression between XX female and XY male therian mammals by silencing gene transcription from one X chromosome in early female embryos. X-inactivation is a paradigm of epigenetic transcriptional regulation because two genetically equivalent chromosomes are transcriptionally differentiated and maintain these transcriptional states through many cell divisions. Mice undergo two distinct forms of X-inactivation: imprinted and random. Imprinted X-inactivation results in the silencing of genes on the paternal X chromosome in preimplantation female embryos. Notably, imprinted X-inactivation is a paradigm of transgenerational epigenetic regulation due to its stable parent-of-origin pattern of inactivation of the paternal X chromosome. In this body of work, I discovered specific functions for core Polycomb repressive complex 2 (PRC2) components EED and EZH1/2 in mouse imprinted X-inactivation; a role for inhibition of GSK-3 proteins, which mediate intracellular signaling, in X-inactivation erosion of human embryonic stem cells (hESCs); and defined distinct requirements for *Xist* RNA versus *Xist* DNA in X-inactivation.

PRC2 is a protein complex that deposits the histone H3 lysine 27 trimethyl (H3K27me3) chromatin modification that is associated with transcriptional silencing. I identified a role for oocyte-derived (maternal) PRC2 protein EED in preventing inactivation of the maternal X chromosome during imprinted X-chromosome inactivation in mice. I also demonstrate that loss of other PRC2 core proteins, EZH1 and EZH2, in the oocyte results in milder defects in X-inactivation in the embryo, suggesting a role for EED in transcriptional silencing independent of the PRC2 complex.

Unlike mice, all cells in the early female human embryo appear to undergo random X-inactivation, which results in the inactivation of either the maternally or paternally inherited X chromosome in individual cells. I helped identify lithium chloride and other inhibitors of glycogen synthase kinase 3 (GSK-3) proteins in hESC culture media as a cause of X-inactivation erosion via repression of the X-inactivation regulatory long-noncoding RNA XIST. I also discovered that GSK-3 inhibition can repress *Xist* in mouse embryonic stem cells (mESCs), potentially via the activation of Wnt signaling. GSK-3 inhibition is a new mechanism by which *Xist* can be regulated and suggests that extracellular signaling can regulate X-inactivation, which is conventionally thought to be regulated cell autonomously. My findings in this study also inform the culture of hESCs.

My thesis work also interrogates the role of *Xist* RNA versus *Xist* DNA in X-inactivation in mice. X-inactivation has long been thought to be controlled by *Xist* RNA, which is expressed solely from the inactive X chromosome and is thought to trigger gene silencing by recruiting protein complexes to the inactive X chromosome. To distinguish a role for *Xist* RNA from that of the *Xist* genomic locus, I ablated *Xist* in female mouse trophoblast stem cells (TSCs), which normally maintain imprinted X-inactivation of the paternal X chromosome. In mouse TSCs devoid of the PRC2 component EED, *Xist* RNA is not expressed. Despite the absence of *Xist* RNA, most paternal X-linked genes remain silenced. By contrast, I found that deletion of the *Xist* genomic loss resulted in de-repression of most paternal X-linked genes. My findings suggest the *Xist* locus silences X-linked genes by mechanisms other than via *Xist* RNA.

Together, this work identifies novel intra- and extracellular factors and mechanisms underlying X-inactivation in mice and humans.

Chapter 1
Introduction

In 1949, Canadian researchers Murray Barr and Ewart Bertram noticed a dark-staining structure in the nucleus of female but not male feline cells (Barr and Bertram, 1949). Years later, this structure – aptly named the "Barr Body" – was shown to be one of the two X chromosomes present in females (Ohno et al., 1959). In 1961, Mary Lyon reasoned that the Barr Body is in fact a condensed, inactivated X chromosome. This discovery marked the beginning of research into X-chromosome inactivation. X-inactivation is the process by which one of the two X-chromosomes in female therian mammals is transcriptionally silenced (Lyon, 1961). Once silenced, with a few key exceptions, replicated copies of the inactivated X-chromosome are stably transmitted to descendant cells (Lyon, 1961; Takagi and Sasaki, 1975). The stable maintenance of the divergent transcriptional states through mitoses by the two X chromosomes – which are equivalent or nearly-equivalent in sequence – makes X-inactivation a powerful model for investigating non-genetic, or epigenetic, modes of transcriptional regulation.

X-inactivation evolved as a dosage compensation mechanism to equalize X-linked gene expression levels between XX female and XY male mammals. The functional monosomy exhibited by the mammalian X chromosome is thought to have resulted from the divergence of the mammalian X and Y chromosomes from a pair of autosomes ~200 mya (Hughes and Page, 2015). Susumu Ohno reasoned in 1967 that the expression level of X-linked genes should be doubled to compensate for the degeneration of genes along the Y chromosome (Ohno, 1967). *Drosophila*, which also harbor X and Y sex chromosomes, provide a clear example of this type

of dosage compensation, as males upregulate their X chromosome expression level to achieve dosage compensation between the sexes (Birchler et al., 2003). Upregulation of X-linked gene expression in mammals is also observed to occur early in embryonic development, but this process occurs in both males and females (Nguyen and Disteche, 2003). Thus, if mammals were to upregulate X-linked gene expression, they would possess an imbalance of X-linked dosage between females, which have two X chromosomes, and males, which have a single X chromosome. To protect the female embryo from functional tetrasomy of the X chromosome resulting from upregulation of X-linked genes, one X chromosome in females becomes silenced in early embryogenesis. X-inactivation is thus thought to be essential for female embryo viability (Marahrens et al., 1997).

Over the past ~70 years, mice have been the most studied organism in the field of X-inactivation. In mice, X-chromosome inactivation exists in two forms: imprinted and random. Imprinted X-inactivation initiates gradually in all cells of the pre-implantation mouse embryo beginning at the 2-4-cell stage (Kalantry et al., 2009; Namekawa et al., 2010; Patrat et al., 2009) (Figure 1.1). Imprinted X-inactivation is thought to have arisen due to an imprint on the maternal X chromosome that prevents the ectopic silencing of the single maternally inherited X chromosome in males (Sado, 2017; Tada et al., 2000). In imprinted X-inactivation, the paternally-inherited X chromosome is selectively inactivated (Takagi and Sasaki, 1975). The maternally-inherited X chromosome, conversely, remains transcriptionally active. Whereas both imprinted and random X-inactivation are epigenetic processes, only imprinted X-inactivation is a paradigm of transgenerational epigenetic inheritance because of the parent-of-origin-specific inactivation of the paternal X chromosome. At ~128-cell blastocyst stage in female mice, cells in the epiblast lineage, which generates the somatic and germ cells in the developing embryo,

reactivate the paternal-X and subsequently undergo random X-inactivation of either the maternally or paternally inherited X-chromosome in individual cells (Mak et al., 2004) (Figure 1.1). Imprinted X-inactivation is stably maintained in extraembryonic tissues, which give rise to the placenta and the yolk-sac (Takagi and Sasaki, 1975; West et al., 1977). Both imprinted and random X-inactivation are preceded by the expression of the *X-inactive specific transcript* (*Xist*) from the prospective inactive-X (Penny et al., 1996).

Early studies of X-chromosomal truncations and translocations in mouse embryos and embryonic stem cells (ESCs) suggested that a region on the X-chromosome called the X-inactivation center (XIC) is necessary for X-inactivation (Lyon et al., 1964; Russell, 1963). This region was later limited to ~1-2 megabases (Mb) by cytological and molecular investigations (Brown et al., 1991b). The XIC has long been the focal point for identifying sequence fragments required for both imprinted and random X-inactivation. Importantly, the XIC houses the *Xist* gene (Brown et al., 1991a; Marahrens *et al.*, 1997; Penny et al., 1996). *Xist* generates a long noncoding (lnc) RNA that physically coats the prospective inactive-X *in cis* and recruits proteins to that X chromosome that execute gene silencing (Brown et al., 1992; Clemson et al., 1996; Moindrot et al., 2015; Panning and Jaenisch, 1996). *Xist* has long been touted as the primary regulatory factor required for X-inactivation (Penny et al., 1996). It has also been thought that *Xist* is required for female viability, but recent unpublished work by our lab and published work by others found a significant subset of female mice can survive to term in the absence of *Xist* (Yang et al., 2016) and can breed to generate more female mice lacking *Xist* (unpublished data, Kalantry Lab). Importantly, past studies of *Xist* function have not distinguished roles for the *Xist* DNA locus versus the *Xist* RNA transcript. Thus, in Chapter 5, I investigate functional differences in *Xist* RNA and *Xist* DNA in imprinted X-inactivation.

Some of the first investigations into *Xist* function suggested that it is essential for imprinted X-inactivation (Marahrens *et al.*, 1997). However, this early work did not distinguish between a requirement for *Xist* in the initiation versus maintenance phases of imprinted X-inactivation. An analysis of preimplantation embryos lacking *Xist* on the paternal X-chromosome (Xp-*Xist*$^{\Delta}$) suggested that the paternal alleles of a subset of X-linked genes could become silenced despite the absence of *Xist* on the paternal-X (Kalantry et al., 2009). The *Xist*-independent genes, nevertheless, required *Xist* to stably remain silenced in later stage embryos (Kalantry et al., 2009). Given these findings, it is likely that *Xist* is necessary to stabilize, but not initiate, imprinted X-inactivation. A prominent allele-specific single-cell RNA-Seq analysis of Xp-*Xist*$^{\Delta}$ preimplantation embryos concluded that the silencing of paternal X-linked genes absolutely requires *Xist* (Deng et al., 2014). However, a closer analysis of the data suggested that a subset of the paternal X-linked genes may undergo silencing despite the absence of paternal *Xist* gene (Borensztein et al., 2017). Given that single-cell RNA-Seq data are subject to technical biases (Chen et al., 2019; Hicks et al., 2018), which are compounded by allele-specific RNA expression quantitation, testing X-linked gene expression by alternate assays, *e.g.*, RNA FISH, Pyrosequencing, and whole embryo RNA-Seq are necessary for a rigorous comparative analysis of the two studies. Furthermore, it is important to consider none of these studies of Xp-*Xist*$^{\Delta}$ female mouse embryos distinguish between roles for *Xist* DNA vs. *Xist* RNA in X-linked gene silencing.

In later studies of *Xist*, defective imprinted X-inactivation due to *Xist* deletion on the paternal X-chromosome or abrogation of random X-inactivation due to the deletion of *Xist* on both the paternal and maternal X-chromosomes surprisingly did not yield the expected ~2-fold increase in X-linked gene expression (Borensztein *et al.*, 2017; Yang *et al.*, 2016). Instead, *Xist*-

mutant embryos in some studies expressed X-linked genes at levels only slightly higher than those in control wild-type embryos. These results suggest that dosage compensation of X-linked gene expression can occur in the absence of *Xist*. It is worth noting that other analyses (Appendix B) (Shin et al., 2010) demonstrate nearly 2-fold increased expression of X-linked genes when *Xist* is deleted, highlighting a potentially context dependent requirement for *Xist* and the variability in findings when different approaches are used for genetic manipulation and transcriptomic analysis. Thus, undiscovered mechanisms for X-chromosome dosage compensation must exist in mammals. The discovery of these processes and mechanisms promises to contribute significantly not only to our understanding of X-chromosome dosage compensation but also to epigenetic transcriptional regulation broadly.

Another prominent gene that is present within the XIC is *Tsix*, a lncRNA overlapping the *Xist* sequence but oriented antisense to *Xist*. And, in contrast to *Xist*, which is expressed exclusively from the inactive-X, *Tsix* is expressed from the active X-chromosome in mice. Through a deletion analysis, Clerc and Avner first proposed that a sequence element or an encoded factor that blocked *Xist* expression resided 3' to *Xist* (Clerc and Avner, 1998). Shortly thereafter, *Lee et al.* discovered that this 3' region expressed an RNA antisense to *Xist*, which they termed *Tsix* (Lee et al., 1999). Through much work by several groups, *Tsix* was found to block *Xist* expression through its transcription across the *Xist* promoter (Ohhata et al., 2008). Furthermore, *Tsix* transcription across the *Xist* promoter is believed to alter the chromatin landscape to one that prevents *Xist* expression. The maternal transmission of a *Tsix*-mutant X-chromosome remarkably resulted in the ectopic induction of *Xist* and inactivation of the maternal X-chromosome in cells that would normally undergo imprinted X-inactivation of only the paternal X-chromosome (Lee, 2000; Sado et al., 2001). *Tsix* was therefore postulated to repress

Xist on the maternal X-chromosome at the onset of imprinted X-inactivation (Lee, 2000). Moreover, the preferential expression of *Tsix* from the maternal-X in the early embryo was proposed to be due to an epigenetic difference between the maternal and paternal *Tsix* loci (Lee, 2000). A recent reanalysis of *Tsix* function by our group, however, found that *Tsix* is dispensable for inhibiting *Xist* in the preimplantation embryo and in cultured stem cells of extra-embryonic trophectodermal and primitive endodermal lineages, which stably maintain imprinted X-inactivation (Maclary et al., 2014). *Tsix* is instead required to prevent *Xist* expression as the extra-embryonic trophectodermal progenitor cells differentiate both in the embryo and in culture.

Another noncoding RNA present within *Xist* in mice is an antisense lncRNA called *Xist-activating antisense RNA* (*XistAR*) (Sarkar et al., 2015). *XistAR* was discovered by our lab to be co-expressed with *Xist* from the inactive X chromosome. Although the function *XistAR* is not fully understood, it is required to promote *Xist* expression (Sarkar *et al.*, 2015). Chapter 5 discusses my findings about the nonequivalence of *Xist* DNA and *Xist* RNA and the role that other lncRNAs encoded within *Xist*, including *XistAR*, may play in X-inactivation. In addition to *XistAR*, our group has discovered two other lncRNAs that are expressed from within the *Xist* gene body but in the antisense orientation from the inactive-X. The function of these transcripts warrants further examination. I discuss these lncRNAs in greater detail in the Conclusion and Future Directions section of Chapter 5.

Xist RNA has long been hypothesized to effect silencing on the inactive-X by recruiting proteins to the inactive-X (Brockdorff et al., 1992). A series of genetic and protein interaction screens has lent credence to this hypothesis and identified hundreds of proteins that either directly or indirectly interact with *Xist* RNA (Chu et al., 2015; McHugh et al., 2015; Minajigi et al., 2015). The deletion or depletion of a subset of these proteins demonstrated that some but not

all are required for X-inactivation (Chu *et al.*, 2015; McHugh et al., 2015; Minajigi *et al.*, 2015; Moindrot *et al.*, 2015; Monfort et al., 2015).

The Polycomb Group contributes several prominent proteins recruited to the inactive-X chromosome at the onset of X-chromosome inactivation. Polycomb proteins are known for modifying chromatin and silencing key developmental genes. The first Polycomb Group gene, Polycomb (Pc), was discovered in *Drosophila melanogaster* by Pamela Lewis (Lewis, 1947). Over 30 years later, Pamela's husband Ed Lewis discovered that Polycomb mutations transform anterior segments to posterior ones due to the ectopic expression of Homeotic (Hox) genes (Lewis, 1978). Subsequent genetic screens revealed other mutants with genetic de-repression effects like those of *Pc*, thus leading to the formation of the Polycomb Group of proteins. Of relevance to this work are the Polycomb Group genes *Esc* and *Ez* which are homologs of mammalian *Eed* and *Ezh1/2*, respectively. Chapters 2 and 3 thoroughly examine the role of maternally-generated *Eed* and *Ezh1/2* in imprinted X-inactivation.

The Polycomb Group protein EED was the first protein shown to be required for imprinted X-inactivation (Wang et al., 2001). EED is a core component of the Polycomb Repressive Complex 2 (PRC2), which deposits Histone H3 lysine 27 trimethylation (H3K27me3) (Margueron et al., 2009). H3K27me3 is a chromatin modification that is enriched on the inactive-X and is associated with transcriptional silencing (Mak et al., 2002; Okamoto et al., 2004). Such modifications are generally held to be a broad mechanism by which transcriptional states are propagated as epigenetic memories across multiple mitotic divisions (Margueron and Reinberg, 2011; Ragunathan et al., 2015; Zhang et al., 2015). Within PRC2, EED is required for complex stability, substrate recognition, and for promoting the enzymatic activity of catalytic subunits EZH2 and EZH1 (Cao et al., 2002; Cao and Zhang, 2004;

Margueron *et al.*, 2009; Tie et al., 2007). Mouse embryos deficient in EED cannot catalyze H3K27me3 and fail to maintain imprinted X-inactivation in extra-embryonic cells of post-implantation embryos (Kalantry et al., 2006; Wang *et al.*, 2001). Subsequent analyses of *Eed*$^{-/-}$ embryos demonstrated that paternal X-linked genes that are normally silenced are reactivated specifically in differentiating trophoblast cells of the EED-deficient embryos (Kalantry *et al.*, 2006).

In addition to being studied in embryos, imprinted X-inactivation can be interrogated in cultured stem cells of the extraembryonic lineages. Trophoblast stem cells (TSCs) and extraembryonic endoderm (XEN) cells, which are progenitors of the extra-embryonic trophectoderm and the primitive endoderm lineages, respectively, stably maintain imprinted X-inactivation in culture (Kunath et al., 2005; Tanaka et al., 1998). Initial reports indicated that TSCs lacking EED stably maintained imprinted X-inactivation due to silencing in *Eed*$^{-/-}$ TSCs of a paternal X-linked *GFP* (*Xp-GFP*) transgene (Kalantry *et al.*, 2006). Upon differentiation of the *Eed*$^{-/-}$ TSCs, however, the *Xp-GFP* transgene became expressed. A recent chromosome-wide analysis of X-linked gene expression in *Eed*$^{-/-}$ TSCs via allele-specific RNA-sequencing (RNA-Seq), however, indicated that ~20% of endogenous paternal X-linked genes are de-repressed and thus require EED for stable silencing in TSCs (Maclary et al., 2017).

Although earlier studies clearly demonstrated that EED and, hence, PRC2 are required to maintain imprinted X-inactivation, they left open the question of whether PRC2 is also required for the initiation of imprinted X-inactivation. Although *Eed*$^{-/-}$ embryos and TSCs appear to initiate imprinted X-inactivation normally (Kalantry *et al.*, 2006; Maclary *et al.*, 2017; Wang *et al.*, 2001), a role for EED and PRC2 in the initiation of imprinted X-inactivation may be masked by the transmission of oocyte-derived, or maternal, EED protein in *Eed*$^{-/-}$ embryos (Kalantry and

Magnuson, 2006; Plath et al., 2002). Chapter 2 explores the role of maternal EED in the execution of imprinted X-inactivation in the early mouse embryo (Figure 1.2).

The class mammalia is comprised of two subclasses: prototheria and theria. Prototherians are egg-laying mammals that originated ~190 mya and now represent the monotreme species platypus and echidna. The therian mammals evolved later from a common ancestor ~180 mya and comprise the eutherian and metatherian mammalian branches. Eutherians are commonly referred to as placental mammals and metatherians as marsupial mammals, though marsupials display a rudimentary placenta. All mammalian branches display X-Y chromosomal system of sexual differentiation. However, the extant monotreme species contain distinct sex chromosomes compared to the therian species, which share an ancestral X-Y system. Relatedly, monotremes do not appear to undergo chromosome-wide dosage compensation of their X-linked genes. Instead, female monotremes tolerate higher expression of many X-linked genes like that seen in avian species (Cooper, 1971; Deakin et al., 2008; Ellegren et al., 2007; Escamilla-Del-Arenal et al., 2011; Grant et al., 2012; Sharman, 1971).

All therian species examined, by contrast, exhibit X-inactivation as a dosage compensation mechanism. In the therian subclass of metatherian species, e.g., opossum, tammar wallaby, and kangaroo, all somatic cells undergo imprinted X-inactivation. Amongst the therian subclass of eutherian species, some species display both imprinted and random X-inactivation some only undergo random X-inactivation. For example, whereas mice exhibit both imprinted and random forms of X-inactivation, humans appear to only undergo random X-inactivation.

Unlike in imprinted X-inactivation, which is characterized by the pre-determined fate of the paternal-X to become inactivated, random X-inactivation comprises the stochastic inactivation of the maternally or paternally inherited X chromosome in individual cells.

Importantly, random X-inactivation is believed to comprise several stages: sensing/counting the number of X chromosomes per diploid genome, choice of an X chromosome for inactivation, initiation of X-inactivation, and the propagation/maintenance of the inactive state through cell divisions (Escamilla-Del-Arenal *et al.*, 2011). Some factors and mechanisms are thought to be exclusively involved in random X-inactivation, as this form of X-inactivation is inherently different from imprinted X-inactivation.

A key protein that appears to be recruited by *Xist* RNA to the inactive-X at the onset of random X-inactivation is SPEN (Chu *et al.*, 2015; McHugh et al., 2015; Minajigi *et al.*, 2015; Moindrot *et al.*, 2015). SPEN is enriched on the inactive-X at the onset of X-inactivation and is required for silencing a subset of X-linked genes in mice (Jachowicz et al., 2022; Moindrot *et al.*, 2015; Monfort *et al.*, 2015; Nesterova et al., 2019; Robert-Finestra et al., 2021). Furthermore, SPEN has been shown to integrate many repressive complexes including NCOR/SMRT and HDAC3, which may contribute to X-linked gene silencing (Dossin et al., 2020; McHugh et al., 2015; Zylicz et al., 2019). One group has even suggested that SPEN may play a role in the induction of *Xist* (Robert-Finestra et al., 2021). It has also been proposed that SPEN may recruit factors to silence transposable elements, and this function may have been coopted to localize *Xist* RNA to the inactive-X and silence X-linked genes (Carter et al., 2020).

Another key protein involved in random X-inactivation in mice is SMCX/KDM5C, which is encoded by the X chromosome and demethylates histone H3 di- and tri- methylated lysine 4 (H3K4me2 and H3K4me3) (Iwase et al., 2007; Tahiliani et al., 2007). H3K4me2 and H3K4me3 are chromatin marks associated with active transcription (Barski et al., 2007). The removal of H3K4me2/3 by KDM5C may thus contribute to gene silencing on the inactive-X. Recent work by our group has found that KDM5C is both necessary and sufficient to induce *Xist*

in female mice and in differentiating female mESCs, which are a system to model X-inactivation *ex vivo*. Moreover, ectopic expression of mouse and human *Kdm5c* as well as of distantly related mammalian species, e.g., the metatherian species opossum and the prototherian species platypus, is sufficient to induce *Xist* in male mESCs, which normally do not express *Xist*. Thus, eutherian, metatherian, and prototherian KDM5C all harbor an evolutionarily conserved function that can induce *Xist* (Samanta et al., 2022). Of note, KDM5C and SPEN are enriched in the same domain within *Xist*, suggesting that these factors may function to recruit one another at the onset of X-inactivation (Dossin *et al.*, 2020; Robert-Finestra *et al.*, 2021; Samanta *et al.*, 2022).

RBM15, an RNA binding protein, was discovered in an shRNA screen to be important for *Xist*-mediated gene silencing and for the efficient deposition of histone H3 lysine 27 trimethylation H3K27me3 on the inactive X-chromosome (Moindrot *et al.*, 2015). This study showed that RBM15 knockdown leads to reduced intensity and size of H3K27me3 domains on the inactive-X (Moindrot *et al.*, 2015). Subsequent work demonstrated that RBM15 deposits the N6-methyladenosine (m6A) modification at the 5' end of *Xist*, and this modification aids in silencing some X-linked genes (Coker et al., 2020; Nesterova *et al.*, 2019).

Although human and mouse embryos both undergo X-chromosome inactivation, the dynamics of X-inactivation differ during human and mouse embryonic development (Kalantry et al., 2009; Kay et al., 1993; Maclary *et al.*, 2014; Mak *et al.*, 2004; Mandal et al., 2020; Moreira de Mello et al., 2017; Okamoto et al., 2011; Petropoulos et al., 2016; Takagi and Sasaki, 1975). Most notably, mice undergo imprinted X-inactivation in the preimplantation embryo whereas humans do not. In humans, random X-inactivation is established during early embryogenesis and then maintained in all female somatic cells of the organism. Recent work has also indicated that the human placenta exhibits random X-inactivation, although large areas of maternal- or

paternal-X biased inactivation may be present in this tissue (Phung et al., 2022). Conflicting data regarding the onset of random X-inactivation in humans exists, likely due to limitations inherent to the study of human embryos. According to *van den Berg et al., 2009*, based on the investigation of a limited number of X-linked genes, human embryos initiate random X-inactivation in the preimplantation stages, with XIST RNA initially accumulating on one of the two X-chromosomes around the 8-cell stage. However, other studies of the single-cell transcriptome of human preimplantation embryos indicated that female human embryos undergo a dampening of X-linked gene expression prior to the initiation of random X-inactivation at or after the blastocyst stage (Moreira de Mello *et al.*, 2017; Petropoulos *et al.*, 2016).

Studies conducted in human cells and tissues have found that the XIC region that appears to be necessary for X-inactivation in mice is dispensable for the maintenance of X-inactivation in human somatic cells (Brown and Willard, 1994). A potential reason for this difference may lie in the structural variation that exists between the mouse and human XIC (Chureau et al., 2002). Although both humans and mice harbor *XIST*/*Xist* within the XIC, the *TSIX* sequence has become a pseudogene in humans and is thus unable to repress *XIST* (Migeon et al., 2002). Another gene within the mouse XIC, *Ppnx*, is absent in the human XIC, but no evidence exists to suggest that this gene is responsible for differences in mouse and human X-inactivation (Migeon *et al.*, 2002). Although all other XIC genes present in mice appear to be conserved in humans, there is much intergenic variation between the two species (Chang and Brown, 2010), which may harbor species-specific regulatory elements.

Like in mice, different types of embryonic stem cell lines representing different developmental stages can be derived from human embryos. Two key cell types that are derived from human embryos are naïve and primed human embryonic stem cells (hESCs). Naïve hESCs

are derived from 8-cell human embryos and represent the developmentally earliest state described for human established cells (Ware et al., 2014). In these naïve hESC lines, a proportion of cells display XIST RNA coating of both X chromosomes but do not appear to transcriptionally inactivate the XIST RNA-coated Xs, like cells in early female human embryos. Most female hESCs cultured in naïve conditions, however, harbor one XIST RNA-coated X chromosome that is transcriptionally active (Figure 1.3) (Guo et al., 2017; Mandal *et al.*, 2020; Messmer et al., 2019; Sahakyan et al., 2017a; Vallot et al., 2017). The heterogeneity of XIST RNA expression in naïve female hESCs appears to be due to the coexistence in culture of at least two populations of pluripotent cells (An et al., 2020). Blocking autocrine bFGF signaling reduces this heterogeneity and is reported to yield nearly all hESCs with two XIST RNA-coated X chromosomes (An *et al.*, 2020), recapitulating the pattern observed in epiblast cells of preimplantation female human embryos (Okamoto *et al.*, 2011; Petropoulos *et al.*, 2016).

Compared to naïve hESCs, primed pluripotent hESCs capture a later stage of embryonic development (Brons et al., 2007; Hanna et al., 2010; Nichols and Smith, 2009; Stadtfeld and Hochedlinger, 2010; Takahashi et al., 2018; Tesar et al., 2007; Weinberger et al., 2016). Primed female hESCs are derived from the inner cell mass (ICM) of blastocyst-stage embryos and female lines exhibit at least three patterns of X-inactivation: no inactive-X chromosome, one inactive-X, or a leaky inactive-X (Figure 1.3) (Anguera et al., 2012; Barakat et al., 2015; Hall et al., 2008; Hoffman et al., 2005; Kim et al., 2014; Lengner et al., 2010; Mekhoubad et al., 2012; Nazor et al., 2012; Patel et al., 2017; Pomp et al., 2011; Shen et al., 2008b; Silva et al., 2008; Tchieu et al., 2010; Tomoda et al., 2012; Vallot et al., 2015; Xie et al., 2016). When female primed pluripotent hESCs are maintained in culture, they often undergo erosion of X-inactivation in which XIST RNA expression and coating are irreversibly lost through passaging.

The loss of XIST coating of the inactive-X is then followed by the stepwise de-repression of some X-linked genes (Dvash et al., 2010; Hall *et al.*, 2008; Lengner *et al.*, 2010; Mekhoubad *et al.*, 2012; Shen *et al.*, 2008b; Vallot *et al.*, 2015; Xie *et al.*, 2016). Chapter 4 addresses the cause of X-inactivation erosion in primed hESCs and provides recommendations for preventing X-inactivation erosion in these cells.

In summary, my studies further define the influence of key cell intrinsic and extrinsic mechanisms on the initiation and maintenance of X-chromosome inactivation in mice and humans. The following chapters detail my findings in four distinct studies that address the role of PRC2 proteins EED and EZH1/2, the *Xist* DNA locus and its primary RNA transcript, and the effect of culture conditions on X-inactivation. By exploring the role of PRC2 proteins, *Xist* RNA, and *Xist* DNA in imprinted X-inactivation, I have gained insight into how these factors function broadly, including their role in initiating epigenetic transcriptional states during embryonic development. I have also identified GSK-3 inhibition as a novel influence on X-inactivation in humans and mice. These data interestingly suggest that X-inactivation may be influenced by Wnt signaling, a cell-extrinsic mechanism, via the inhibition of GSK-3. Through my dissertation research, I have ascertained key differences and similarities in X-inactivation mechanisms between mice and humans, which contribute to our understanding of epigenetics and dosage compensation in eutherian mammals more broadly.

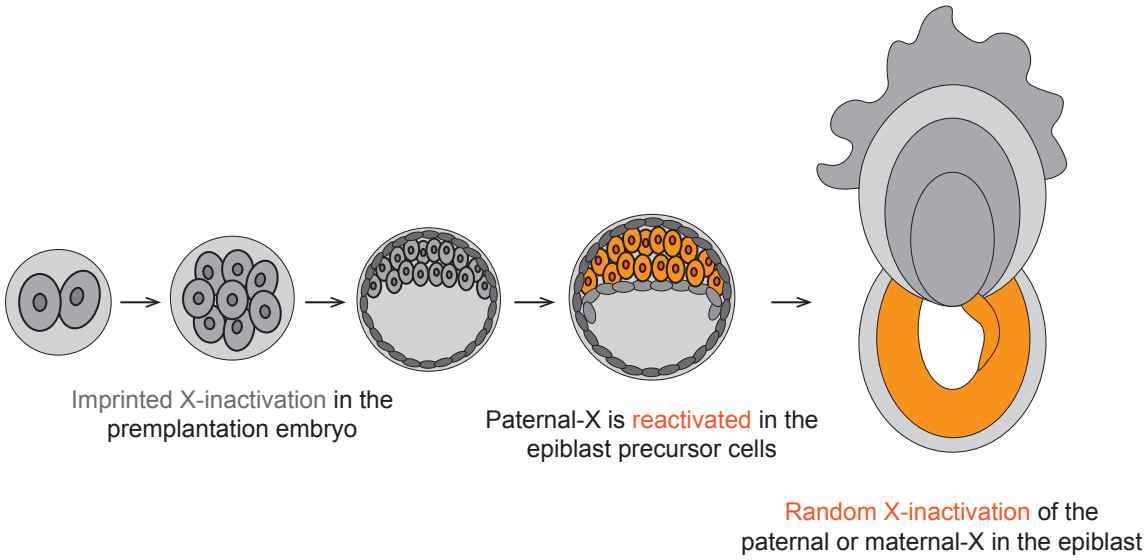


Figure 1.1. Schematic depicting the timing of imprinted and random X-inactivation initiation in the female mouse embryo

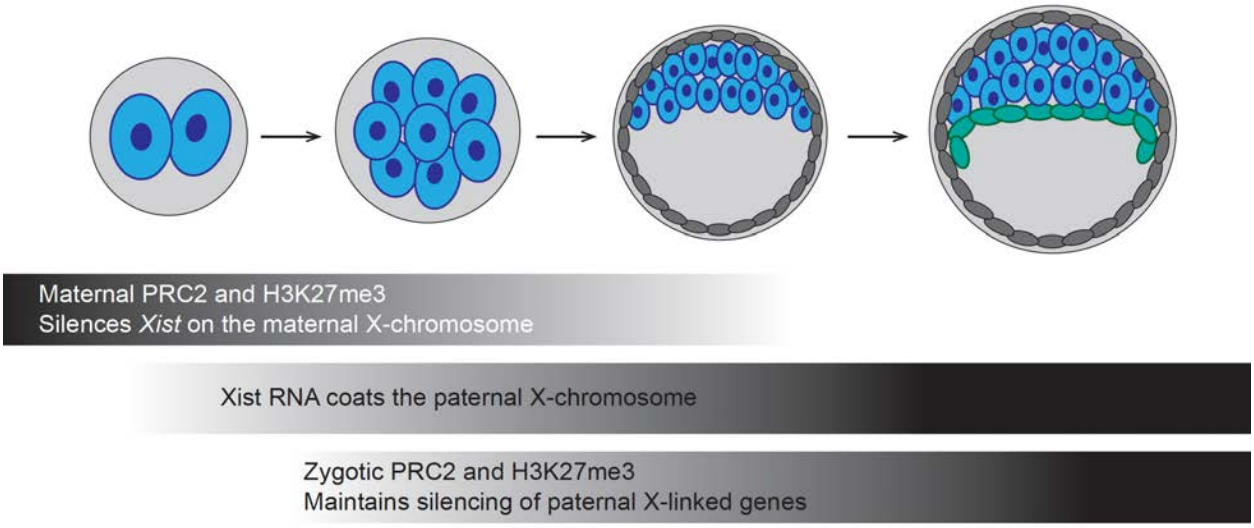


Figure 1.2. Schematic depicting the stages of embryogenesis in which female mice initiate and maintain imprinted X-chromosome inactivation and the role of PRC2 at each stage

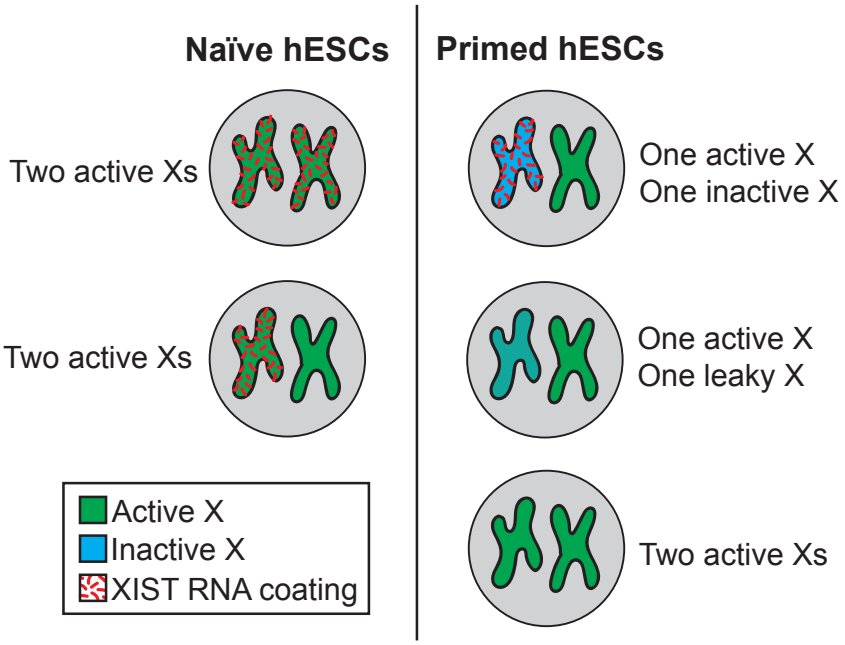


Figure 1.3. X-inactivation patterns observed in female naïve and primed hESCs

Chapter 2

Conversion of Random X-inactivation to Imprinted X-inactivation by Maternal PRC2

Note: This chapter was adopted from a published manuscript describing the role of maternal PRC2 in the initiation of imprinted X-chromosome inactivation:

Harris, C.*, Cloutier, M.*, Trotter, M., Hinten, M., Gayen, S., Du, Z., Xie, W., Sundeep Kalantry (2019). Conversion of random X-inactivation to imprinted X-inactivation by maternal PRC2 *eLife* 8:e44258.

*Denotes equally contributing authors

Abstract

Imprinted X-inactivation silences genes exclusively on the paternally inherited X-chromosome and is a paradigm of transgenerational epigenetic inheritance in mammals. Here, we test the role of maternal vs. zygotic Polycomb repressive complex 2 (PRC2) protein EED in orchestrating imprinted X-inactivation in mouse embryos. In maternal-null (*Eed*$^{m-/-}$) but not zygotic-null (*Eed*$^{-/-}$) early embryos, the maternal X-chromosome ectopically induced *Xist* and underwent inactivation. *Eed*$^{m-/-}$ females subsequently stochastically silenced *Xist* from one of the two X-chromosomes and displayed random X-inactivation. This effect was exacerbated in embryos lacking both maternal and zygotic EED (*Eed*$^{mz-/-}$), suggesting that zygotic EED can also contribute to the onset of imprinted X-inactivation. *Xist* expression dynamics in *Eed*$^{m-/-}$ embryos resemble that of early human embryos, which lack oocyte-derived maternal PRC2 and only undergo random X-inactivation. Thus, expression of PRC2 in the oocyte and transmission of the gene products to the embryo may dictate the occurrence of imprinted X-inactivation in mammals.

Introduction

X-chromosome inactivation results in the mitotically stable transcriptional inactivation of one of the two X-chromosomes in female mammals in order to equalize X-linked gene expression between males and females (Morey and Avner, 2011; Plath *et al.*, 2002). Two different forms of X-inactivation characterize the mouse embryo, imprinted and random. Imprinted X-inactivation results in the exclusive silencing of genes on the paternal X-chromosome and initiates during preimplantation embryogenesis (Huynh and Lee, 2003; Mak *et al.*, 2004; Monk and Kathuria, 1977; Okamoto *et al.*, 2004; Takagi and Sasaki, 1975). In the post-implantation embryo, imprinted X-inactivation is stably maintained in the extraembryonic lineage but reversed in the embryonic lineage (Harper et al., 1982; Mak *et al.*, 2004; Okamoto *et al.*, 2004; Takagi and Sasaki, 1975; West *et al.*, 1977), which subsequently undergoes random inactivation of either the maternal or the paternal X-chromosome (Lyon, 1961). Notably, imprinted X-inactivation is a paradigm for both mitotic as well as meiotic, or transgenerational, epigenetic regulation, due to its stable parent-of-origin-specific inactivation pattern.

X-inactivation is characterized by a well-defined series of epigenetic events (Kalantry, 2011). Both imprinted and random X-inactivation are prefaced by the expression of X-linked non-protein coding *Xist* RNA from the prospective inactive-X (Kay et al., 1994; Penny et al., 1996). During imprinted X-inactivation in the mouse embryo, *Xist* is expressed at the two-cell stage and the RNA visibly begins to coat the paternal-X at the four-cell stage (Kalantry et al., 2009; Namekawa *et al.*, 2010; Patrat *et al.*, 2009). The progressive accumulation of *Xist* RNA coincides with the gradual and stereotyped silencing of paternal X-linked genes that is only completed after the blastocyst stage of embryogenesis (Kalantry et al., 2009; Namekawa *et al.*, 2010; Patrat *et al.*, 2009). Coincident with Xist RNA coating, PRC2 proteins and H3K27me3

accumulate on the inactive-X, correlating with the silencing of X-linked genes (Mak *et al.*, 2004; Okamoto *et al.*, 2004; Plath et al., 2003; Silva et al., 2003). Moreover, the mis-expression of *Xist* results in the concomitant accumulation of PRC2 proteins and H3K27me3 (de la Cruz et al., 2005a; Kohlmaier et al., 2004; Plath *et al.*, 2003; Silva *et al.*, 2003), suggesting that *Xist* RNA directly or indirectly recruits PRC2 to the inactive-X. PRC2 has thus been suggested to contribute to the establishment of X-inactivation (Plath *et al.*, 2003; Silva *et al.*, 2003).

Consistent with a role for PRC2 in X-inactivation, we and others previously showed that post-implantation female mouse embryos mutant for the Polycomb gene *Eed* fail to maintain silencing of paternal X-linked genes during imprinted X-inactivation (Kalantry and Magnuson, 2006; Kalantry *et al.*, 2006; Wang *et al.*, 2001). EED is a non-catalytic component of the PRC2 complex, but EED binding to the PRC2 enzyme EZH2 is required for the methyltransferase activity of EZH2 (Cao *et al.*, 2002; Czermin *et al*., 2002; Kuzmichev *et al*., 2002; Muller *et al.*, 2002). When EED is mutated other core PRC2 proteins are degraded and H3K27me3 is lost (Montgomery et al., 2005). Thus, EED is an essential component of PRC2 and EED function is canonically equated with H3K27me3 catalysis (Margueron and Reinberg, 2010; Montgomery *et al*., 2005).

Although $Eed^{-/-}$ embryos fail to maintain imprinted X-inactivation, they initiate imprinted X-inactivation properly (Kalantry and Magnuson, 2006; Kalantry *et al.*, 2006). A potential answer for this difference is that $Eed^{-/-}$ embryos inherit maternal EED protein that is present in the oocyte (Kalantry and Magnuson, 2006; Plath *et al.*, 2003; Shumacher et al., 1996). The presence of maternally derived EED protein could explain the absence of a defect in establishing imprinted X-inactivation in $Eed^{-/-}$ embryos. Such maternal control of imprinted X-inactivation would also be consistent with a transgenerational epigenetic effect that underlies genomic

imprinting (Barlow, 2011; Ferguson-Smith and Bourc'his, 2018; Lee and Bartolomei, 2013; van Otterdijk and Michels, 2016). Here, we test the hypothesis that oocyte-derived PRC2 orchestrates imprinted X-inactivation in the early embryo.

EED and H3K27me3 Enrichment on the Inactive-X in *Eed*$^{-/-}$ Embryos

PRC2 proteins and H3K27me3 are first enriched on the prospective inactive X-chromosome in the early mouse embryo at the 8-16 cell morula stage (Okamoto *et al.*, 2004). We assessed the accumulation of EED and H3K27me3, and *Xist* RNA by immunofluorescence (IF) combined with fluorescent *in situ* hybridization (FISH) in wild-type (WT) embryonic day (E) 3.5 blastocyst embryos (Cloutier et al., 2018; Hinten et al., 2016), which are in the process of silencing X-linked genes and establishing imprinted X-inactivation (Borensztein *et al.*, 2017; Namekawa *et al.*, 2010; Patrat *et al.*, 2009; Wang et al., 2016). As expected, females displayed coincident accumulation of EED, H3K27me3, and *Xist* RNA in a vast majority of the nuclei (72-100%). Males, by contrast, lacked such enrichment (Figure 2.1A).

Our previous work suggested that zygotically-null preimplantation embryos harbor WT maternal EED protein (Kalantry and Magnuson, 2006; Kalantry *et al.*, 2006). To test for the presence of maternally derived EED protein in *Eed*$^{-/-}$ embryos, we employed our previously generated conditional *Eed* mutation (Supplemental Figure 2.1A) (Maclary *et al.*, 2017). We generated E3.0-E3.5 blastocyst-stage embryos zygotically-null and heterozygous for *Eed* (*Eed*$^{-/-}$ and *Eed*$^{+/-}$, respectively) from a cross of *Eed*$^{+/-}$ females with *Eed*$^{fl/-}$; *Prm-Cre* males. *Prm-Cre* is active during spermatogenesis and catalyzes the deletion of the *lox*p flanked (floxed) *Eed* allele in the mature sperm (Supplemental Figure 2.1B) (O'Gorman et al., 1997). As a result, half of the embryos generated from the above cross are expected to be genotypically *Eed*$^{-/-}$ and the other half *Eed*$^{+/-}$. In the derived embryos, we assayed inactive-X enrichment of EED, H3K27me3, and

Xist RNA by combined IF/FISH (Figure 2.1B). Of the 41 female embryos examined, nine showed coincident accumulation of EED and/or H3K27me3 with *Xist* RNA in over 70% of the nuclei and are not significantly different from WT embryos in Figure 1A ($p > 0.1$). An additional nine embryos were devoid of EED or H3K27me3 enrichment overlapping with the *Xist* RNA coat. We presumed the former to be $Eed^{+/-}$ embryos and the latter to be $Eed^{-/-}$ embryos. An additional 23 embryos displayed 2-70% of nuclei with EED and/or H3K27me3 enrichment. This intermediate class likely represents $Eed^{+/-}$ or $Eed^{-/-}$ embryos that have not yet fully depleted maternally inherited EED protein or $Eed^{+/-}$ embryos that have not yet robustly expressed zygotic EED. Male embryos from the cross, distinguished by a lack of *Xist* RNA coating, did not show enrichment of EED or H3K27me3 in the nucleus, as in the WT male embryos in Figure 1A.

To confirm that there is no bias in the sex ratio or genotype of the embryos, we performed PCR genotyping of embryos derived from the above cross (Figure 2.1C). Embryos from 12 litters showed no statistical difference in the distribution of $Eed^{+/-}$ and $Eed^{-/-}$ male or female embryos ($p > 0.05$), suggesting that the intermediate class of 23 embryos in Figure 2.1A are likely a mixture of $Eed^{+/-}$ or $Eed^{-/-}$ embryos. Together, the results in Figure 1 suggest that genotypically null $Eed^{-/-}$ embryos inherit oocyte-derived maternal EED protein and that expression of EED transitions from maternal to zygotic at or slightly before the blastocyst stage.

To define the kinetics of depletion of maternal EED and induction of zygotic EED prior to the blastocyst stage, we quantified EED and H3K27me3 IF signals in 2-, 4-, 8-, and 16-cell embryos from the following series of crosses. The first cross was $Eed^{fl/fl}$ females crossed to $Eed^{fl/fl}$ males, which yielded control $Eed^{fl/fl}$ embryos. The second was a cross of $Eed^{fl/-}$ females to $Eed^{fl/-}$; *Prm-Cre* males to generate $Eed^{fl/-}$ and $Eed^{-/-}$ embryos ($Eed^{fl/-}$ / $Eed^{-/-}$). Whereas both $Eed^{fl/-}$ and $Eed^{-/-}$ embryos are expected to harbor maternal EED protein, $Eed^{fl/-}$ but not $Eed^{-/-}$

embryos would express zygotic EED. The third cross was of *Eed*$^{fl/fl}$; *Zp3-Cre* females to WT males to yield embryos which are devoid of maternal EED (*Eed*$^{m-/-}$) but are capable of expressing zygotic EED. *Zp3-Cre* is active in the growing oocyte, where it efficiently deletes the *Eed*^{fl} allele and generates embryos devoid of maternal EED (Figure 2.2; Supplemental Figure 2.1A) (Lewandoski et al., 1997). The final cross was a cross of *Eed*$^{fl/fl}$; *Zp3-Cre* females with *Eed*$^{fl/fl}$;*Prm-Cre* males to generate embryos devoid of both maternal and zygotic EED (*Eed*$^{mz-/-}$).

Eed$^{fl/fl}$ and *Eed*$^{fl/-}$ / *Eed*$^{-/-}$ 2-cell embryos exhibited similar levels of EED and H3K27me3, whereas *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos were devoid of both EED and H3K27me3 (Figure 2.2A, 2.2C, and 2.2D; Supplementary File 1). These data are consistent with the 2-cell embryo harboring only maternally derived EED and H3K27me3. Four-cell embryos displayed a similar pattern to 2-cell embryos, although a subset of *Eed*$^{fl/-}$ / *Eed*$^{-/-}$ ~4-cell embryos displayed reduced EED and H3K27me3 levels, consistent with zygotic EED expression beginning at this stage and its failure in *Eed*$^{-/-}$ embryos (Figure 2.2C and Figure 2.2; Supplemental Figure 2.1B). At the ~8-cell stage, *Eed*$^{fl/-}$ / *Eed*$^{-/-}$ embryos showed highly variable EED and H3K27me3 levels, suggesting further differentiation of the two genotypes. In agreement with increasing zygotic *Eed* expression, *Eed*$^{m-/-}$ ~8-cell embryos displayed higher levels of EED and H3K27me3 than the corresponding *Eed*$^{mz-/-}$ embryos (Figure 2.2C and Figure 2.2; Supplemental Figure 2.1B). By the ~16-cell stage, *Eed*$^{fl/-}$ / *Eed*$^{-/-}$ embryos are clearly separated into two categories. One group has statistically lower levels of EED, while the other group is statistically indistinguishable from the *Eed*$^{fl/fl}$ embryos (Figure 2.2B-D). Therefore, the likely genotypes of the two groups are *Eed*$^{-/-}$ and *Eed*$^{fl/-}$, respectively. *Eed*$^{m-/-}$ 16-cell embryos continue to display higher levels of EED and H3K27me3 than the *Eed*$^{mz-/-}$ embryos, but nevertheless harbor significantly lower EED and H3K27me3 levels than *Eed*$^{fl/fl}$ embryos (Figure 2.2B-D). In order to visualize how EED levels

are changing across early embryogenesis, we plotted the mean values of each genotype by embryonic stage (Figure 2.2E). Maternally derived EED starts declining at the 4-cell stage but is still present at the 16-cell stage. Conversely, while zygotic *Eed* transcription initiates at ~4-cell stage, zygotic EED levels are still low at the ~16-cell stage, suggesting that EED in WT $Eed^{fl/fl}$ 16-cell embryos is a combination of maternally derived and zygotically generated protein (Figure 2.2F).

Imprinted X-inactivation Initiation in $Eed^{-/-}$ Embryos

To test if zygotic $Eed^{-/-}$ embryos initiate and establish imprinted X-inactivation, we compared X-linked gene expression in an allele-specific manner in individual hybrid $Eed^{fl/fl}$, $Eed^{fl/-}$, and $Eed^{-/-}$ E3.5 blastocysts by RNA sequencing (RNA-Seq) (Figure 2.3; Supplemental Figure 2.1A). In these embryos, the maternal X chromosome was derived from the *Mus musculus* 129/S1 mouse strain and the paternal-X from the divergent *Mus molossinus* JF1/Ms strain (Materials and Methods). We exploited single nucleotide polymorphisms (SNPs) to assign RNA-Seq reads to either the maternal or paternal X-chromosome in the hybrid embryos (Cloutier *et al.*, 2018; Maclary *et al.*, 2017). A subset of X-linked genes was expressed more robustly from the paternal allele relative to the maternal allele in $Eed^{fl/-}$ and $Eed^{-/-}$ female embryos compared to $Eed^{fl/fl}$ embryos (Figure 2.3A). However, when the allelic expression ratio of all X-linked genes in Figure 2.3A was averaged, paternal X-linked gene expression was not significantly higher in $Eed^{-/-}$ blastocysts compared to $Eed^{fl/-}$ ($p = 0.72$) or $Eed^{fl/fl}$ ($p = 0.76$) female embryos (Figure 2.3B and Figure 2.3; Supplemental Figure 2.1B). X-linked genes were expressed predominantly from the maternal allele in all three genotypes. Thus, the ratio of maternal:paternal X-linked gene expression in $Eed^{-/-}$ female blastocysts was broadly similar to that in $Eed^{fl/fl}$ and $Eed^{fl/-}$ embryos.

We next sought to validate the RNA-Seq data via Pyrosequencing. Pyrosequencing is a low-throughput technique that can accurately capture allelic expression ratios of individual genes (Cloutier *et al.*, 2018; Gayen et al., 2015). We analyzed the expression of *Xist* and three X-linked genes subject to X-inactivation, *Rnf12*, *Atrx*, and *Pgk1*. *Xist* expression analysis by Pyrosequencing was especially important, as there was variability in *Xist* SNP-overlapping read coverage in the RNA-Seq data due potentially to the highly repetitive sequence of *Xist* RNA. We did not detect any significant changes in maternal: paternal allelic expression in hybrid $Eed^{-/-}$ vs. $Eed^{fl/fl}$ and $Eed^{fl/-}$ blastocysts (Figure 2.3C and Figure 2.3; Supplemental Figure 2.1C). Whereas *Xist* was expressed predominantly from the paternal allele, *Rnf12*, *Atrx*, and *Pgk1* were preferentially expressed from the maternal allele in all three genotypes.

As an independent validation of the RNA-Seq and Pyrosequencing results, we also performed RNA FISH to test *Xist* RNA coating and nascent RNA expression of *Rnf12* in $Eed^{-/-}$ and $Eed^{fl/fl}$ female (Figure 2.3D) and male (Figure 2.3; Supplemental Figure 2.1D) blastocysts. RNA FISH has the added benefit of providing single cell expression resolution in embryos (Cloutier *et al.*, 2018; Hinten et al., 2016). We distinguished $Eed^{fl/fl}$ from $Eed^{-/-}$ female embryos by assaying H3K27me3 enrichment by IF on the *Xist* RNA-coated X-chromosome (Figure 2.3D-E). We classified embryos displaying fewer than 5% of the nuclei with this H3K27me3 enrichment as $Eed^{-/-}$ (Figure 2.3E). *Xist* RNA coating and *Rnf12* expression in female $Eed^{-/-}$ embryos did not differ significantly from $Eed^{fl/fl}$ blastocysts (Figure 2.3D and Figure 2.3F). Both sets of embryos displayed *Xist* RNA coating of one X-chromosome and *Rnf12* expression from the other X-chromosome in a majority of the cells. Male $Eed^{-/-}$ or $Eed^{+/-}$ embryos also did not differ significantly from $Eed^{fl/fl}$ embryos in their *Rnf12* expression patterns (Figure 2.3; Supplemental Figure 2.1D). Thus, by three independent assays – allele-specific RNA-Seq,

Pyrosequencing, and RNA FISH – we found that zygotic *Eed* expression is largely dispensable for the initiation and establishment of imprinted X-inactivation.

Defective Imprinted X-inactivation Initiation in *Eed*$^{m-/-}$ Embryos

Since early *Eed*$^{-/-}$ embryos harbor WT maternally derived EED protein, we next examined the role of maternal EED in initiating imprinted X-inactivation in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ blastocysts, which are devoid of maternally-derived EED. *Eed*$^{m-/-}$ blastocysts exhibited a small percentage of nuclei with H3K27me3 enrichment coinciding with the *Xist* RNA coat (Figure 2.4A). *Eed*$^{mz-/-}$ blastocysts, on the other hand, lacked all such overlapping accumulation (Figure 2.4A). H3K27me3 enrichment on the *Xist* RNA-coated X-chromosome in *Eed*$^{m-/-}$ but not *Eed*$^{mz-/-}$ blastocysts is likely due to the expression of zygotic *Eed* in *Eed*$^{m-/-}$ but not *Eed*$^{mz-/-}$ embryos (Figure 2.2).

To test if maternal EED regulates imprinted X-inactivation, we conducted allele-specific RNA-Seq on individual hybrid *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ E3.5 blastocysts (Figure 2.4; Supplemental Figure 2.1A). Strikingly, the RNA-Seq data revealed a relative increase in paternal X-linked gene expression in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos compared to *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, and *Eed*$^{-/-}$ embryos (Figures 2.4B-C; Supplemental Figure 2.1B). Furthermore, *Eed*$^{mz-/-}$ embryos appeared to express paternal X-linked genes to a greater degree compared to *Eed*$^{m-/-}$ embryos (Figure 2.4B). When allelic expression ratios of all X-linked genes in Figure 2.4B were averaged, however, the difference between *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos did not reach statistical significance ($p = 0.14$) (Figure 2.4C).

The shift in the ratio of X-linked gene expression towards the paternal allele in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos could be due to increased paternal X-linked gene expression or to decreased maternal X-linked gene expression. To determine the source of the expression change, we

calculated the normalized expression of genes on the maternal and paternal X-chromosomes for all genotypes (Figure 2.4D; Supplemental Figure 2.1C). Whereas paternal X-linked genes significantly increased in expression, maternal X-linked gene expression decreased in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos compared to *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, and *Eed*$^{-/-}$ embryos. The increase in paternal X-linked gene expression in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos was significant when compared to the three other genotypes. The decrease in maternal X-linked gene expression in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos reached significance only vs. *Eed*$^{fl/fl}$ embryos and not vs. *Eed*$^{fl/-}$ and *Eed*$^{-/-}$ embryos. The lack of a significant decrease between *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos compared to *Eed*$^{fl/-}$ and *Eed*$^{-/-}$ embryos is likely due to the greater variation in maternal X-linked gene expression in *Eed*$^{fl/-}$ and *Eed*$^{-/-}$ embryos. Finally, *Eed*$^{mz-/-}$ embryos displayed a significant increase in paternal X-linked gene expression compared to *Eed*$^{m-/-}$ embryos ($p = 0.02$), suggesting that zygotic EED can contribute to the silencing of a subset of X-linked genes in blastocysts.

To validate the *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ blastocyst RNA-Seq data, we again analyzed allele-specific expression of *Xist*, *Rnf12*, *Atrx*, and *Pgk1* in E3.5 blastocysts by Pyrosequencing. Pyrosequencing also showed a significant defect in the initiation and establishment of imprinted X-inactivation in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos (Figure 2.4E; Supplemental Figure 2.1D). In *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos, *Xist* expression unexpectedly increased from the maternal-X relative to the paternal-X. Conversely, the expression of *Rnf12* and *Atrx* increased from the paternal-X relative to the maternal-X in *Eed*$^{m-/-}$ embryos. In *Eed*$^{mz-/-}$ embryos, in addition to *Rnf12* and *Atrx*, *Pgk1* also displayed nearly equal levels of expression from the maternal and paternal alleles. The Pyrosequencing results thus recapitulate the defects in imprinted X-inactivation observed by RNA-Seq.

Together, the RNA-Seq and Pyrosequencing data lead to several suggestions. The first is that maternal EED depletion induces *Xist* from the maternal X-chromosome in the early embryo. This de-repression is consistent with maternally derived PRC2 repressing the maternal *Xist* locus, which is marked by H3K27me3 in the oocyte [Figure 2.4; Supplemental Figure 2.1E; (Zheng et al., 2016)]. Ectopic *Xist* induction from the maternal-X then results in the silencing of genes on that X-chromosome. The second major suggestion is that loss of maternal EED induces paternal X-linked genes. Finally, the data implicate zygotic EED expression in the silencing of a subset of paternal X-linked genes at the onset of imprinted X-inactivation.

Maternal EED Silences *Xist* on the Maternal-X

To validate the RNA-Seq and Pyrosequencing data from the maternal *Eed* mutants, we performed RNA FISH in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ blastocysts for *Xist* and *Rnf12* (Figure 2.5A). Whereas most nuclei in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ females displayed a single *Xist* RNA coat and monoallelic expression of *Rnf12*, a subset displayed *Xist* RNA coating of both X-chromosomes. The majority of these nuclei also lacked *Rnf12* expression, suggesting silencing of *Rnf12* on both X-chromosomes.

We similarly examined *Eed*$^{mz-/-}$ male blastocysts (Figure 2.5B). A subset of nuclei in *Eed*$^{mz-/-}$ male mutant embryos also exhibited ectopic *Xist* RNA coating of their sole, maternal X-chromosome. Interestingly, *Eed*$^{mz-/-}$ male embryos were present in two distinct morphological classes. The first category was comprised of large, well-developed embryos, which displayed few or no nuclei with *Xist* RNA coating. The second category consisted of underdeveloped embryos, which displayed *Xist* RNA-coating in much higher proportions (20-60% of nuclei). In both sets of embryos, *Xist* RNA coating was often accompanied by a loss of *Rnf12* expression from the X-chromosome. These data suggest that *Xist* RNA coating hinders developmental

progression by silencing genes on the ectopically Xist RNA-coated X-chromosome. *Eed*$^{mz-/-}$ embryos that adaptively repress *Xist* may overcome this developmental deficiency.

The correlation between reduced frequency of ectopic *Xist* RNA-coated nuclei and development of *Eed*$^{mz-/-}$ embryos led us to test the developmental competency of maternal-null *Eed* embryos. We assessed if *Eed*$^{m-/-}$ embryos could yield live born animals. To our surprise, a small number of *Eed*$^{m-/-}$ female as well as male embryos could live to term (Figure 2.5C), suggesting that the ectopic *Xist* RNA expression and coating could be resolved in maternal-null embryos of both sexes. Interestingly, significantly more females were born compared to males ($p = 0.02$, Two-tailed Student's T-test), suggesting that females can more robustly extinguish ectopic *Xist* RNA expression compared to males. These data further suggest that zygotic EED expression is sufficient to compensate for the absence of maternal EED in a subset of the early embryos. *Eed*$^{mz-/-}$ embryos are expected to be inviable, since loss of zygotic *Eed* expression results in lethality of both female and male embryos (Faust et al., 1995; Shumacher *et al.*, 1996; Wang *et al.*, 2001).

Switching of Imprinted to Random X-inactivation in *Eed*$^{m-/-}$ Embryos

The relative paucity of ectopic *Xist* RNA-coated nuclei in female *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ blastocysts observed by RNA FISH in Figure 2.5A-B is inconsistent with the robust ectopic *Xist* RNA expression from and silencing of maternal X-linked genes and the increased expression of paternal X-linked genes that are detected via Pyrosequencing and RNA-Seq (Figure 2.4B-D). We thus postulated that instead of undergoing imprinted inactivation of the paternal X-chromosome, *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ blastocysts switch to random X-inactivation of either the maternal- or the paternal-X in individual cells. Such mosaicism would explain the silencing of

maternal X-linked genes and the induction of paternal X-linked gene expression in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ female embryos detected by RNA-Seq and Pyrosequencing.

To test the above model of X-inactivation mosaicism, we developed and applied an allele-specific *Xist* RNA FISH strategy on hybrid control *Eed*$^{fl/+}$ and test *Eed*$^{m-/-}$ female E3.5 blastocysts (Materials and Methods; Figure 2.6; Supplemental Figure 2.1). Allele-specific *Xist* RNA FISH allowed us to discriminate *Xist* RNA expression from the maternal vs. the paternal X-chromosome in individual cells. Allele-specific *Xist* RNA FISH displayed *Xist* RNA expression from the paternal-X in *Eed*$^{fl/+}$ female blastocysts (Figure 2.6A), as would be expected from embryos stably undergoing imprinted X-inactivation of the paternal-X. In *Eed*$^{m-/-}$ female blastocysts, we saw a mosaic distribution of *Xist* RNA expression and coating. Whereas some *Eed*$^{m-/-}$ blastocyst nuclei displayed *Xist* RNA expression from and coating of the maternal-X, others exhibited *Xist* RNA expression from and coating of the paternal-X. A subset of nuclei in *Eed*$^{m-/-}$ blastocysts exhibited *Xist* RNA expression from both the maternal and paternal X-chromosomes (Figure 2.6A), consistent with the non-allele specific *Xist* RNA FISH data from *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ blastocysts in Figure 2.5A. Male *Eed*$^{m-/-}$ embryos similarly displayed ectopic *Xist* RNA expression from and coating of their sole maternally inherited X-chromosome in approximately 50% of nuclei (Figure 2.6B).

From the blastocyst data, we extrapolated that earlier *Eed*$^{m-/-}$ embryos may harbor a higher proportion of cells with ectopic *Xist* RNA coating of the maternal-X. This pattern is later resolved into the mosaic *Xist* RNA coating pattern observed at the blastocyst stage in females and loss of the *Xist* RNA coat in males. We therefore performed allele-specific *Xist* RNA FISH on 3-16 cell control *Eed*$^{fl/+}$ and test *Eed*$^{m-/-}$ hybrid embryos. In the *Eed*$^{fl/+}$ female embryos, *Xist* RNA was expressed from and coated only the paternal X-chromosome (Figure 2.7A). Most

Eed$^{m-/-}$ female embryos, by contrast, displayed a high percentage of nuclei with *Xist* RNA expression and coating of both X-chromosomes (Figure 2.7A). In male 3-17 cell embryos, *Eed*$^{fl/+}$ embryos did not show any nuclei with *Xist* RNA coating (Figure 2.7B). In *Eed*$^{m-/-}$ male embryos, by contrast, almost every nucleus exhibited ectopic *Xist* expression from and coating of the maternally inherited X-chromosome (Figure 2.7B). Thus, in the absence of maternal EED most cells express *Xist* from both X-chromosomes in early female embryos and from the sole X in early male embryos. By the blastocyst stage, however, one of the two *Xist* alleles is stochastically silenced in most female cells and the sole *Xist* allele is silenced in most male cells.

Lack of Maternal EED in Human Embryos

Intriguingly, the *Xist* RNA coating of both X-chromosomes in female and of the single X in male early preimplantation *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos resemble the pattern observed in preimplantation human female and male embryos (Okamoto et al.; Petropoulos *et al.*, 2016). In early preimplantation human embryos, females display XIST RNA coating of both Xs and males of their sole maternally inherited X. We therefore hypothesized that the XIST RNA expression profile in early human embryos may reflect the absence of maternally derived EED and other core PRC2 proteins in human oocytes. To test this hypothesis, we analyzed RNA-Seq data from mouse and human oocytes to determine the expression levels of core PRC2 genes *Eed*, *Ezh2*, *Ezh1*, and *Suz12* (Kobayashi et al., 2012; Macfarlan et al., 2012; Reich et al., 2011). Compared to mouse oocytes, human oocytes expressed all four genes at negligible levels (Figure 2.8A). This difference in the expression of PRC2 components in oocytes may underlie why early mouse but not human embryos undergo imprinted X-inactivation.

Discussion

Genomic imprinting is a paradigm of transgenerational epigenetic inheritance, since the two parental alleles undergo diametrically divergent transcriptional fates in the embryo. Imprinted X-inactivation is an extreme example of genomic imprinting in that most genes on the paternally inherited X-chromosome undergo silencing. The maternal X-chromosome, by contrast, remains active. Here, we define the transition of maternal to zygotic EED expression in the early embryo and find the presence of maternal but not zygotic EED when imprinted X-inactivation begins. Upon ablation of *Eed* in the oocyte and the absence of maternally derived EED in the embryo, the initiation of imprinted X-inactivation is compromised (Figure 2.8B). Maternal-null ($Eed^{m-/-}$ and $Eed^{mz-/-}$) but not zygotic-null ($Eed^{-/-}$) early preimplantation female and male embryos ectopically induce *Xist* RNA from the maternal X-chromosome. Early $Eed^{m-/-}$ female embryos therefore display *Xist* RNA-coating of both X-chromosomes and mutant males of the sole maternally inherited X-chromosome.

PRC2-catalyzed H3K27me3 marks the *Xist* locus on the maternal X-chromosome during oogenesis (Zheng *et al.*, 2016). In agreement, the injection of the H3K27me3 demethylase *Kdm6b* in the zygote resulted in the de-repression of the *Xist* locus on the maternal X-chromosome in 8-16 cell embryos (Inoue et al., 2017). Female morulas derived from *Kdm6b*-injected zygotes displayed *Xist* RNA coating of both the maternal and the paternal X-chromosome in most blastomeres, suggesting inactivation of both Xs in the embryo. Nullizygosity of X-linked gene expression due to inactivation of both Xs in females or of the single-X in males is expected to result in cell and embryo lethality (Gayen *et al.*, 2015). The conditional deletion of *Eed* in the oocyte, however, yielded live born mice, implying that ectopic *Xist* expression due to H3K27me3 loss and the ensuing inactivation of the maternal-X in the early embryo is resolved later [this chapter; (Prokopuk et al., 2018)]. In agreement, our study

shows that by the blastocyst stage most nuclei in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ female embryos exhibit only one *Xist* RNA coat. However, instead of *Xist* RNA coating exclusively of the paternal X-chromosome as in WT embryos, the maternal *Eed* mutants express *Xist* RNA from and coat either the maternal or the paternal X-chromosome, a hallmark of random X-inactivation. This randomization persists later in development in extraembryonic tissues (data not shown), which normally maintain imprinted inactivation of the paternal-X. Like *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ females, *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ male blastocysts also extinguish ectopic *Xist* induction.

In addition to maternal EED, our data argue that zygotically generated EED contributes to imprinted X-inactivation in the early embryo. In comparison to *Eed*$^{m-/-}$ embryos, *Eed*$^{mz-/-}$ female blastocysts displayed a further increase in paternal X-linked gene expression. One interpretation of these data is that the onset of zygotic EED expression results in the preferential installation of H3K27me3 at the *Xist* locus on the maternal-X in some cells of early *Eed*$^{m-/-}$ embryos. These cells thus forestall or extinguish *Xist* expression from the maternal X-chromosome and inactivate the paternal-X, ultimately resulting in more cells in the embryo in which the paternal-X is inactive compared to the maternal-X. Loss of both maternal and zygotic EED would annul such biased inactivation of the paternal-X and thereby cause a greater increase in paternal X-linked gene expression in *Eed*$^{mz-/-}$ embryos. An alternative possibility is that zygotic EED functions to maintain silencing preferentially of paternal X-linked genes in the early embryo. The differential sensitivity of genes on the maternal vs. paternal X-chromosomes to zygotic EED in *Eed*$^{m-/-}$ embryos may reflect the different kinetics of inactivation of the two X-chromosomes. The ectopic induction of *Xist* and X-linked gene silencing on the maternal-X may occur more slowly compared to that on the paternal-X. Due to this delay, genes on the maternal-X would still be in the process of undergoing silencing in *Eed*$^{m-/-}$ blastocysts. A subset of paternal X-linked genes,

on the other hand, may have established silencing and are now in the maintenance phase of X-inactivation in the blastocysts. In the absence of both maternal and zygotic EED, then, *Eed*$^{mz-/-}$ blastocysts fail to maintain silencing of these paternal X-linked genes. Previous work has shown that zygotic EED is in fact required to maintain silencing of a discrete set of paternal X-linked genes during imprinted X-inactivation (Kalantry and Magnuson, 2006; Kalantry *et al.*, 2006; Maclary *et al.*, 2017).

The ability of the cells of early *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos to resolve *Xist* RNA coating of both Xs in females or of the single X in males implies that the early embryo has an X-chromosome counting mechanism that ensures that a single X-chromosome remain active in females as well as in males, irrespective of its parent of origin. Such a counting mechanism has previously been proposed by Takagi and colleagues to explain the kinetics of *Xist* RNA induction in *XX* and *XY* androgenetic embryos, which harbor only paternal X-chromosomes (Okamoto et al., 2000). Like in *Eed*$^{m-/-}$ embryos, androgenetic 4 and 8-16 cell embryos also initially induce *Xist* RNA from all Xs, which is resolved at the blastocyst stage and results in females displaying a single Xist RNA coat in most nuclei and males exhibiting few or no nuclei with *Xist* RNA coating (Okamoto *et al.*, 2000). Molecular sensing of the X-chromosomal complement in imprinted X-inactivation is also suggested by studies of diploid *XX* parthenogenetic or gynogenetic embryos, which harbor two maternal X-chromosomes. In these preimplantation bi-maternal *XX* embryos, *Xist* expression is delayed and appears to occur stochastically from one or the other X-chromosome (Kay *et al.*, 1994). In agreement, the extraembryonic tissues of post-implantation *XX* parthenogenotes display hallmarks of random X-inactivation instead of the imprinted form observed in WT extraembryonic cells (Rastan et al., 1980). Randomization of X-inactivation in extraembryonic cells of mouse embryos with two

paternal or maternal X-chromosomes led Takagi and colleagues to suggest that imprinted X-inactivation in placental mammals may have arisen from random X-inactivation (Matsui et al., 2001), a notion that our data from *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos agree with.

Evidence suggests that the X-linked *Rnf12* gene may be a key component of the X-chromosome counting mechanism during imprinted X-inactivation. The maternal-X allele of *Rnf12* is required to induce *Xist* from the paternal-X in preimplantation mouse embryos (Shin *et al.*, 2010). Upon *Xist* RNA coating, *Rnf12* is rapidly silenced on the paternal X-chromosome (Kalantry et al., 2009; Namekawa *et al.*, 2010; Patrat *et al.*, 2009). In *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos, in addition to the paternal *Rnf12* allele, the maternal *Rnf12* allele is also stringently silenced due to ectopic *Xist* RNA coating of the maternal-X. Since *Rnf12* is required for Xist RNA induction in the preimplantation embryo, the silencing of all *Rnf12* alleles in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ female and male embryos may paradoxically lead to the loss of *Xist* RNA expression from both Xs in females or from the sole X-chromosome in males. In females, this transient state of two active-Xs may then be followed by random X-inactivation, analogously to how differentiating pluripotent epiblast cells undergo random X-inactivation (Gayen *et al.*, 2015; Maclary *et al.*, 2014; Mak *et al.*, 2004). The X-chromosome counting process and randomization of X-inactivation in the early embryo may explain how *Eed*$^{m-/-}$ embryos can yield live born animals [this study and (Prokopuk *et al.*, 2018)].

While preparing this manuscript, a publication reported that extraembryonic tissues of *Eed* maternal-null female post-implantation embryos exhibit random X-inactivation (Inoue et al., 2018). The primary piece of data in the study supporting this conclusion is the expression of maternal and paternal X-linked genes, including *Xist*, in post-implantation E6.5 female *Eed*$^{m-/-}$ extraembryonic tissues by allele-specific RNA-Seq. Although in agreement with our

conclusions, the study does not directly demonstrate when imprinted X-inactivation switches to random X-inactivation and whether loss of zygotic *Eed* would result in a similar outcome. Our study, by contrast, genetically dissects the relative contributions of maternal vs. zygotic EED in the initiation and establishment of imprinted X-inactivation by three different approaches, allele-specific RNA-Seq, Pyrosequencing, and allele-specific *Xist* RNA FISH. We are thus able to pinpoint when and how the loss of maternal EED converts imprinted X-inactivation to random X-inactivation in preimplantation embryos. Genetically testing the requirement of maternal vs. zygotic EED is necessary to determine that the establishment of imprinted X-inactivation in the preimplantation embryo is maternally but not zygotically controlled.

Xist RNA expression in *Eed*$^{m-/-}$ mouse embryos mimics the pattern observed in human embryos, which do not undergo imprinted X-inactivation and ultimately display only random X-inactivation (Okamoto *et al.*; Petropoulos *et al.*, 2016). In agreement, like the *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos, human oocytes do not express *Eed* and other core PRC2 genes, suggesting that the presence or absence of maternal PRC2 proteins may dictate whether placental mammals undergo imprinted X-inactivation.

Conclusion and Future Directions

This work establishes a role for maternally-generated EED in the initiation of imprinted X-inactivation in female mouse embryos. Our group's findings may serve as the basis for many future studies, including the investigation of a role for other PRC2 components in X-linked gene silencing (addressed in Chapter 3), the contribution of zygotically-generated EED to imprinted X-inactivation, a role for EED and other PRC2 components in X-linked gene regulation in other imprinted tissues, *Xist* induction mechanics in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ mouse embryos, and the conservation of imprinted X-inactivation between species.

Recent work by others has suggested that the dynamics of imprinted X-chromosome inactivation differ between the preimplantation embryo and extraembryonic tissues of the mouse (Kunath et al., 2005). Therefore, an important future direction for this work is to examine allele-specific X-linked gene expression in individual female mouse extraembryonic tissues, including the visceral endoderm and extraembryonic ectoderm. I obtained preliminary transcriptomic data in *Eed*$^{fl/fl}$ and *Eed*$^{m-/-}$ E6.5 female mouse extraembryonic tissues (visceral endoderm and extraembryonic ectoderm together) by dissecting and performing allele-specific RNA-Seq on *Eed*$^{fl/fl}$ and *Eed*$^{m-/-}$ E6.5 on these tissues (Supplemental Figure 2.6). My analysis showed defects in paternal X-linked gene silencing in *Eed*$^{m-/-}$ E6.5 tissues compared to *Eed*$^{fl/fl}$ tissues. However, my data did not allow me to discern differences in paternal X-linked gene expression between the visceral endoderm and extraembryonic ectoderm. The presence of a maternal-X expression bias in the *Eed*$^{m-/-}$ pooled tissues I characterized suggests that one of these tissues may not fully undergo random X-inactivation when maternal EED is missing. Thus, examining allele-specific X-linked gene expression in both wild-type and *Eed*-mutant visceral endoderm and extraembryonic ectoderm tissues will further inform our understanding of the maintenance of imprinted X-inactivation.

The mechanism by which *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos initiate random X-inactivation can and should also be explored at a single-cell and resolution using RNA FISH, live cell imaging, and single-cell sequencing approaches. The exact mechanics of the switch from imprinted to random X-inactivation in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos is not yet understood, but performing these high-resolution analysis techniques on *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos will provide valuable insight into how random X-inactivation initiates in these embryos. Live-cell imaging of *Xist* expression dynamics at early embryonic stages will be particularly informative, as we were not

able to determine in this work how biallelic *Xist* expression is resolved between the ~16-cell and blastocyst stages in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ female embryos.

Ectopic expression of EED and/or other PRC2 components in rat oocytes, which lack EED, may inform whether EED and/or PRC2 is sufficient to silence maternal *Xist* and effect imprinted X-inactivation. Generating hybrid rat embryos from these transgenic oocytes could provide valuable insight into the mechanism by which imprinted X-inactivation initiates. If rat embryos harboring maternal EED can initiate imprinted X-inactivation, this would establish the sufficiency of maternal EED to initiate this process.

Materials and Methods

Ethics Statement

This study was performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. All animals were handled according to protocols approved by the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan (protocol #s PRO6455 and PRO8425).

Mice

Mice harboring a conditional mutation in *Eed* were described in our prior publication (Maclary *et al.*, 2017). A *Mus molossinus* JF1 X-chromosome was introgressed to generate *Eed*$^{fl/fl}$; *XJF*^{1}*Y* males. *Mus musculus Eed*$^{fl/fl}$ females were backcrossed onto the 129/S1 background. The X-linked *Gfp* transgenic (*X-Gfp*) and JF1 strains have been described previously (Hadjantonakis et al., 1998; Kalantry and Magnuson, 2006; Kalantry *et al.*, 2006; Kalantry et al., 2009; Maclary *et al.*, 2017).

Embryos generated for the purpose of allele-specific RNA-Seq, Pyrosequencing, or allele-specific RNA fluorescence *in situ* hybridization (FISH) were sired by males harboring the

X^{JF1} X-chromosome. Embryos generated for immunofluorescence (IF) and non-allele specific RNA FISH were sired by males harboring the *X-Gfp* transgene. The paternal *X-Gfp* is only transmitted to daughters. Thus, GFP fluorescence conferred by the paternally transmitted *X-Gfp* transgene was used to sex the embryos.

For derivation of embryos lacking zygotic *Eed*, the *Protamine-Cre* (*Prm-Cre*) transgene was bred into an $Eed^{fl/fl}$ or $Eed^{fl/-}$ background. *Prm-Cre* is expressed only during spermatogenesis (O'Gorman et al., 1997), thus resulting in the deletion of the *Eed* floxed allele in the male germline. For derivation of embryos lacking maternal EED, a *Cre* transgene controlled by the *Zona pellucida 3* gene promoter (*Zp3-Cre*) (Lewandoski *et al.*, 1997), was used to delete the floxed *Eed* alleles in growing oocytes.

Mouse Embryo Dissections and Processing

E3.5 embryos were isolated as described (Maclary *et al.*, 2014). Embryos were flushed from the uterine limbs in 1X PBS (Invitrogen, #14200) containing 6 mg/ml BSA (Invitrogen, #15260037).

Two to sixteen cell embryos were flushed from oviducts of superovulated females with 1X PBS (Invitrogen, #14200) containing 6 mg/ml BSA (Invitrogen, #15260037). For superovulation, 4-5-week-old, or 9-12-week-old females were treated with 5 IU of pregnant mare's serum gonadotropin (PMSG, Sigma, # G-4877) and 46 hours later with 5 IU of human chorionic gonadotropin (hCG, Sigma, #CG-5). Embryos were harvested 48-74 hours post hCG. The zona pellucida surrounding embryos was removed through incubation in cold acidic Tyrode's solution (Sigma, #T1788), followed by neutralization through several transfers of cold M2 medium (Sigma, #M7167).

Isolated E3.5 embryos were either lysed for RNA isolation or plated onto 0.2% gelatin- (Sigma, #G2500) and/or 0.01% Poly-L-Lysine (PLL, Sigma # P4707)-coated glass coverslips in

0.25X PBS for immunofluorescence (IF) coupled with RNA *in situ* hybridization (FISH). 2-16 cell embryos were plated on coverslips coated in 0.01% Poly-L-Lysine for IF. E3.5 or 4-16 cell embryos were plated on coverslips coated with 1X Denhardt's (Sigma, #D9905) solution for allele-specific RNA FISH. For plated embryos, excess solution was aspirated, and coverslips were air-dried for approximately 15-30 mins. After drying, embryos were permeabilized and fixed in 50 µL solution of either 0.05% or 0.1% Tergitol (Sigma, #NP407) with 1% paraformaldehyde (Electron Microscopy Sciences, #15710) in 1X PBS for 5 min, followed by 1% paraformaldehyde in 1X PBS for an additional 5 min. Excess solution was tapped off onto paper towels, and coverslips were rinsed 3X with 70% ethanol and stored in 70% ethanol at -20°C prior to IF or RNA FISH.

PCR

For embryo DNA isolation, embryos were isolated as described above, individual blastocysts were lysed in 15 µL buffer composed of 50 mM KCl, 10 mM Tris-Cl (pH 8.3), 2.5 mM $MgCl_2$, 0.1 mg/mL gelatin, 0.45% NP-40, 0.45% Tween-20, and 0.4 mg/mL Proteinase K (Fisher, #BP1700). Embryos in lysis buffer were incubated at 50°C overnight, then stored at 4°C until use. Genomic PCR used 1-3 µL lysate per sample. Reactions for *Eed* were carried out in ChromaTaq buffer (Denville Scientific) with 2.5 mM $MgCl_2$ added. *XX* vs. *XY* sexing PCR reactions were carried out in Klentherm buffer (670mM Tris pH 9.1, 160mM (NH_4) SO_4, 35mM $MgCl_2$,15mg/ml BSA). Both used RadiantTaq DNA polymerase (Alkali Scientific, #C109). Primer sequences are described in Table 2.1.

Liveborn animals from the cross of $Eed^{fl/fl}$; *Zp3-Cre* female by WT male were genotyped for *Eed* to confirm deletion of the floxed allele. Ear punches were taken after weaning and lysed

in 50 µL of lysis buffer (above). Ear punches were incubated at 50°C overnight, then stored at 4°C until use. 1 µL of DNA lysate was used per reaction. *Eed* PCRs were carried out as above.

Quantification of Allele-specific Expression by Pyrosequencing

Allele-specific expression was quantified using the Qiagen PyroMark sequencing platform, as previously described (Gayen *et al.*, 2015). Briefly, the amplicons containing SNPs were designed using the PyroMark Assay Design software. cDNAs were synthesized using Invitrogen SuperScript III One-Step RT-PCR System (Invitrogen, #12574-026). Following the PCR reaction, 5 µL of the 25 µL reaction was run on a 3% agarose gel to assess the efficacy of amplification. The samples were then prepared for pyrosequencing according to the standard recommendations for use with the PyroMark Q96 ID sequencer. All amplicons spanned intron(s), thus permitting discrimination of RNA vs. any contaminating genomic DNA amplification due to size differences. Control reactions lacking reverse transcriptase for each sample were also performed to rule out genomic DNA contamination. E3.5 embryos of similar sizes for all genotypes were used in the Pyrosequencing assays. Pyrosequencing primer sequences are listed in Table 2.2.

Immunofluorescence (IF)

Embryos mounted on gelatin-, PLL-, and/or PLL/gelatin-coated glass coverslips were washed 3 times in 1X PBS for 3 min each while shaking. Coverslips were then incubated in blocking buffer consisting of 0.5 mg/mL BSA (New England Biolabs, #B9001S), 50 µg/mL yeast tRNA (Invitrogen, #15401-029), 80 units/mL RNAseOUT (Invitrogen, #10777-019), and 0.2% Tween 20 (Fisher, #BP337-100) in 1X PBS in a humid chamber for 30 min at 37°C. The samples were next incubated with primary antibody diluted in blocking buffer for 45 min -2 hr in the humid chamber at 37°C. The samples were then washed 3 times in 1X PBS/0.2% Tween 20 for 3 min

each while shaking. After a 5 min incubation in blocking buffer at 37°C in the humid chamber, the samples were incubated in blocking buffer containing fluorescently conjugated secondary antibody for 30 min in the humid chamber at 37°C, followed by three washes in PBS/0.2% Tween 20 while shaking for 3 min each. For samples undergoing only IF, DAPI was added to the third wash at a 1:250,000 dilution. Coverslips were then mounted on slides in Vectashield (Vector Labs, #H-1000). For samples undergoing IF and RNA FISH, the samples were processed for RNA FISH following the third wash. Antibody information is listed in Table 2.3.

RNA Fluorescence *In Situ* Hybridization (RNA FISH)

RNA FISH with double-stranded and strand-specific probes was performed as previously described (Gayen *et al.*, 2015; Hinten et al., 2016; Kalantry et al., 2009). The *Rnf12* dsRNA FISH probe was made by random-priming using BioPrime DNA Labeling System (Invitrogen, #18094011) and labeled with Cy3-dCTP (GE Healthcare, #PA53021) using a previously described fosmid template (Kalantry et al., 2009). Strand-specific *Xist* probes were generated from templates as described (Maclary *et al.*, 2014; Sarkar *et al.*, 2015). Probes were labeled with Fluorescein-12-UTP (Roche, #11427857910) or Cy5-CTP (GE Healthcare, #25801087). Labeled probes from multiple templates were precipitated in a 0.5M ammonium acetate solution (Sigma, #09691) along with 300 µg of yeast tRNA (Invitrogen, #15401-029) and 150 µg of sheared, boiled salmon sperm DNA (Invitrogen, #15632-011). The solution was then spun at 15,000 rpm for 20 min at 4°C. The pellet was washed consecutively with 70% ethanol and 100% ethanol while spinning at 15,000 rpm at room temperature. The pellet was dried and resuspended in deionized formamide (VWR, #97062-010). The probe was denatured by incubating at 90°C for 10 min followed by an immediate 5 min incubation on ice. A 2X

hybridization solution consisting of 4X SSC and 20% Dextran sulfate (Millipore, #S4030) was added to the denatured solution. All probes were stored in the dark at -20°C until use. Following IF, embryos mounted on coverslips were dehydrated through 2 min incubations in 70%, 85%, 95%, and 100% ethanol solutions and subsequently air-dried. The coverslips were then hybridized to the probe overnight in a humid chamber at 37°C. The samples were then washed 3 times for 7 min each at 37°C with 2X SSC/50% formamide, 2X SSC, and 1X SSC. A 1:250,000 dilution of DAPI (Invitrogen, #D21490) was added to the third 2X SSC wash. Coverslips were then mounted on slides in Vectashield (Vector Labs, #H-1000).

Allele-specific *Xist* RNA FISH

Allele specific *Xist* RNA FISH probes were generated as described (Levesque et al., 2013). Briefly, a panel of short oligonucleotide probes were designed to uniquely detect either the *M. musculus* or the *M. molossinus* alleles of *Xist* (Table 2.4). Five probes were designed for each *Xist* allele. Each probe overlapped a SNP that differs between the two strains, with the SNP located at the fifth base pair position from the 5' end. The same panel of five SNPs was used for both sets of allele-specific probes. The 3' end of each oligonucleotide probe is fluorescently tagged using Quasar dyes (Biosearch technologies). *M. musculus*-specific oligos were labeled with Quasar 570 and *M. molossinus* oligos labeled with Quasar 670. In addition to labeled SNP-overlapping oligonucleotides, a panel of 5 "mask" oligonucleotides were also synthesized. These "mask" probes are complimentary to the 3' end of the labeled allele-specific probes and will hybridize to the allele-specific oligonucleotides, leaving only 9-10 base pairs of sequence surrounding the polymorphic site available to initially hybridize to the target *Xist* RNA. Since this region of complementarity is short, the presence of a single nucleotide polymorphism is sufficient to destabilize the hybridization with the alternate allele. Sequences of detection and

mask probes are listed in Table 2.4. Allele-specific *Xist* RNA FISH probes were combined with a strand-specific *Xist* RNA probe, labeled with Fluorescein-12-UTP (Roche, #11427857910), which served as a guide probe that hybridizes to *Xist* RNA generated from both *Xist* alleles and ensured the fidelity of the allele-specific probes in detecting the cognate *Xist* RNA molecules. The guide *Xist* RNA probe was first ethanol precipitated as previously described, then resuspended in hybridization buffer containing 10% dextran sulfate, 2X saline-sodium citrate (SSC) and 10% formamide. The precipitated guide RNA probe was then mixed with the *M. musculus* and *M. molossinus* detection probes, to a final concentration of 5 nM per allele-specific oligo, and 10 nM mask probe, yielding a 1:1 mask:detection oligonucleotide ratio. Coverslips were hybridized to the combined probe overnight in a humid chamber at 37°C. After overnight hybridization, samples were washed twice in 2X SSC with 10% formamide at 37°C for 30 min, followed by one wash in 2X SSC for 5 mins at room temperature. A 1:250,000 dilution of DAPI (Invitrogen, #D21490) was added to the second 2X SSC with 10% formamide wash. Coverslips were then mounted on slides in Vectashield (Vector Labs, #H-1000).

Microscopy

Stained samples were imaged using a Nikon Eclipse TiE inverted microscope with a Photometrics CCD camera. The images were deconvolved and uniformly processed using NIS-Elements software. For four color images (blue, green, red, and white), the far-red spectrum was employed for the fourth color (AlexaFluor 647 secondary antibody and Cy5-UTP labelled riboprobes for RNA FISH). Additional antibody information is outlined in Table 2.3.

EED and H3K27me3 IF intensity quantification were performed using the "3D Measurement; 3D thresholding, 3D viewing and voxel-based measurements" software package (Nikon Instruments, 77010582). Individual nuclei were marked by creating a binary image,

using the "Threshold" function, over the DAPI stain of the nuclei. Each nucleus was designated as a Region of Interest (ROI) by converting the binary image to an ROI. An additional polygonal ROI was manually created over a non-nuclear region of background stain and then that level of background was subtracted from the entire image. For each channel, average intensity of each nucleus was taken as the intensity measurements from individual ROIs. These intensity values of individual nuclei were then averaged to get the average intensity per embryo. Embryos with 2-3 cells were categorized as being at the 2-cell stage in development. The 4-cell stage encompassed embryos with 4-5 cells. Embryos with 6-10 cells were classified as being at the 8-cell stage in development, and the 16-cell stage encompassed embryos with 14-19 cells. To preserve IF intensities, the images of embryos were not deconvolved. Intensity data for individual nuclei is presented in Figure 2- source data 1.

The Threshold function of the software cannot always distinguish between two nuclei that are overlapping. Similarly, if a single nucleus is an odd shape, it may get counted as multiple nuclei by the software. Some embryos were therefore had different numbers of nuclei measured than how the number of cells in the embryo. If the number of cells in an embryo differs from the number of nuclei listed, the actual number of cells is indicated in parenthesis next to the embryo label in Figure 2- source data 1.

RNA-Seq Sample Preparation

mRNA was isolated from whole embryos using the Dynabeads mRNA DIRECT Kit (Thermo Fisher, # 610.11) according to the manufacturer's instructions. E3.5 embryos of similar sizes of all genotypes were used for RNA-Seq. *Eed*$^{fl/-}$ and *Eed*$^{-/-}$ embryos were genotyped by *Eed* RT-PCR and all embryo genotypes were confirmed by quantifying the relative expression of *Eed* exon 7 to the sample's number of mapped reads (Figure 3- figure supplement 1 and Figure 4-

figure supplement 1). Samples were submitted to the University of Michigan DNA Sequencing Core for Poly-A RNA purification and, separately, strand-specific library preparation using the Takara SMARTer Seq V4 stranded low input kit (Takara, #634889). All libraries were sequenced on the Illumina HiSeq2000 or HiSeq4000 platforms to generate 50 bp paired end reads.

Mapping of RNA-Seq Data

Quality control analysis of the RNA-Seq data was conducted using FastQC. SNP data from whole-genome sequencing of the 129/S1 (*M. musculus*) and JF1/Ms (*M. molossinus*) mouse strains were substituted into the mm9 mouse reference genome build (C57BL/6 J) using VCFtools to generate *in silico* 129/S1 and JF1/Ms reference genomes (Maclary *et al.*, 2017). Sequencing reads were separately mapped to each of the two in silico genomes using STAR (Dobin et al., 2013), allowing 0 mismatches in mapped reads to ensure allele-specific mapping of SNP-containing reads to only one strain-specific genome. STAR was selected for read mapping, in part due to the improved ability to handle structural variability and indels, with the goal of reducing mapping bias to the genome most similar to the reference genome. STAR is a spliced aligner capable of detecting structural variations and is able to handle small insertions and deletions during read mapping. STAR additionally permits soft clipping of reads during mapping, trimming the ends of long reads that cannot be perfectly mapped. This function would permit clipping of reads that end near indels, thus preserving mappability at SNPs near indels.

Prior work showed that the variability due to mapping bias between the 129/S1 and JF1/Ms genomes is minimal (Maclary *et al.*, 2017). Although small biases may affect allelic mapping at a subset of SNP sites within a gene, the effect is mitigated since most genes contain multiple SNPs (Supplemental Figure 2.3).

Allele-specific Analysis of RNA-Seq Data

For allelic expression analysis, only RNA-Seq reads overlapping known SNP sites that differ between the 129/S1 and JF1/Ms genomes were retained. All multi-mapping reads were excluded from the allele-specific analysis. For each SNP site, reads mapping to the 129/S1 and JF1/Ms X chromosomes were counted and the proportion of reads from each X chromosome identified. Allelic expression was calculated individually for each SNP site; for genes containing multiple SNPs, the paternal-X percentage for all SNPs was averaged to calculate gene-level allelic expression. All SNP sites with at least 10 SNP-overlapping reads were retained. Genes containing at least one SNP site with at least 10 SNP-overlapping reads were retained for further analysis and are referred to in the text as informative. In X-linked genes, the SNP frequency is ~1 SNP/250bp in transcribed RNAs (Keane et al., 2011; Maclary *et al.*, 2017; Takada et al., 2013; Yalcin et al., 2011).

RNA-Seq Expression Analysis

To calculate expression from the maternal vs. paternal X-chromosomes, all reads were first merged into a single alignment file and the number of reads per RefSeq annotated gene was counted using HTSeq. To calculate the percentage of expression arising from the paternal X-chromosome, the total read counts from HTSeq were normalized by number of mapped reads. Then, the normalized number of mapped reads for each gene was multiplied by the proportion of SNP-containing reads mapping to the paternal X-chromosome. This analysis was done in R using the following formula:

$$\left\{\text{total reads} \times \left(\frac{\text{paternal reads}}{\text{maternnal reads} + \text{paternal reads}}\right)\right\}$$

Analysis of Human and Mouse Oocyte RNA-Seq Data

For analysis of publicly available oocyte RNA-Seq data, raw Fastq files were obtained from the NCBI Sequence Read Archive. Quality control analysis was conducted using FastQC. Reads were aligned to the mm9 (mouse) or hg19 (human) reference genome using STAR (Dobin et al., 2013) and counted using FeatureCounts (Liao et al., 2014). BioProject and Run numbers for samples analyzed are listed in Table 2.5.

Statistical Analysis & Plots

Welch's two-sample T-tests were used to test for significant differences between the means of Pyrosequencing and RNA-Seq allelic expression data. This test was chosen due to the unequal variance and sample sizes between different genotype groups. In the RNA-Seq allelic expression significance tests, the average percent paternal expression of all informative X-linked genes was calculated for each sample. The total paternal expression value for each genotype group was calculated by calculating the mean of the informative percent paternal values for all samples in that genotype group. A two-tailed Student's T-test was used to determine the significance of RNA FISH and IF data. All histograms and heatmaps were made using the ggplot and Pheatmaps R packages, respectively. Dotplots were made using Python's Seaborn package. Only genes that were informative in all samples were included in the heatmaps.

Data Availability

RNA-Seq data generated for this study have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE123173.

Author Contributions

C.H. and M.H. generated the mouse lines characterized in this study. M.C., C.H., M.T., M.H., and S.G. dissected and processed the mouse embryos characterized. C.H. and M.T. performed the RNA FISH and IF FISH experiments. M.C. and M.H. performed the RNA-Seq and

Pyrosequencing experiments. M.C. performed the sequencing analysis. Z.D. and W.X. performed allele-specific ChIP-Seq experiment. C.H., M.C., and S.K. designed experiments, analyzed data, and prepared the manuscript.

Table 2.1. Genotyping PCR Primers

Gene/Chromosome	Primer Name	Primer Sequence
Eed	Eed 5'	GGACTCATCCTCTGGTAGAGCAGC
	Eed 3’	CCCAAGATCATTACCCCAGA
	Eed R1	TCAATTGGTGGGTTTTGGAT
X/Y Chromosome	XY F	CCGCTGCCAAATTCTTTGG
	XY R	TGAAGCTTTTGGCTTTGAG

Table 2.2. Pyrosequencing Primers

Gene	Primer Name	Primer Sequence
Xist	Xist_RT_F	CAAGAAGAAGGATTGCCTGGATTT
	Xist_RT_R	5′-biotin-GCGAGGACTTGAAGAGAAGTTCTG
	Xist_PyroSeq	CAAACAATCCCTATGTGA
Atrx	Atrx_RT_F	ATAGCTTCAGATTCTGATGAAACC
	Atrx_RT_R	5′-biotin-ACATCGTTGTCACTGCCACTT
	Atrx_PyroSeq	TAAGCTCAGATGAAAAGA
Pgk1	Pgk1_RT_F	TTTCCGAGCCTCACTGTCC
	Pgk1_RT_R	5’-biotin-CTTTAGCGCCTCCCAAGA
	Pgk1_PyroSeq	GTCCAGAGCGACCCT
Rnf12	Rnf12_RT_F	TGCAGCCAACAAGTGAAATTCC
	Rnf12_RT_R	5’-biotin-TATCTGCTGTCTCAGGGTCACATG
	Rnf12_PyroSeq	TAGAACTTCCTTCAGGC

Table 2.3. Immunofluorescence Antibodies

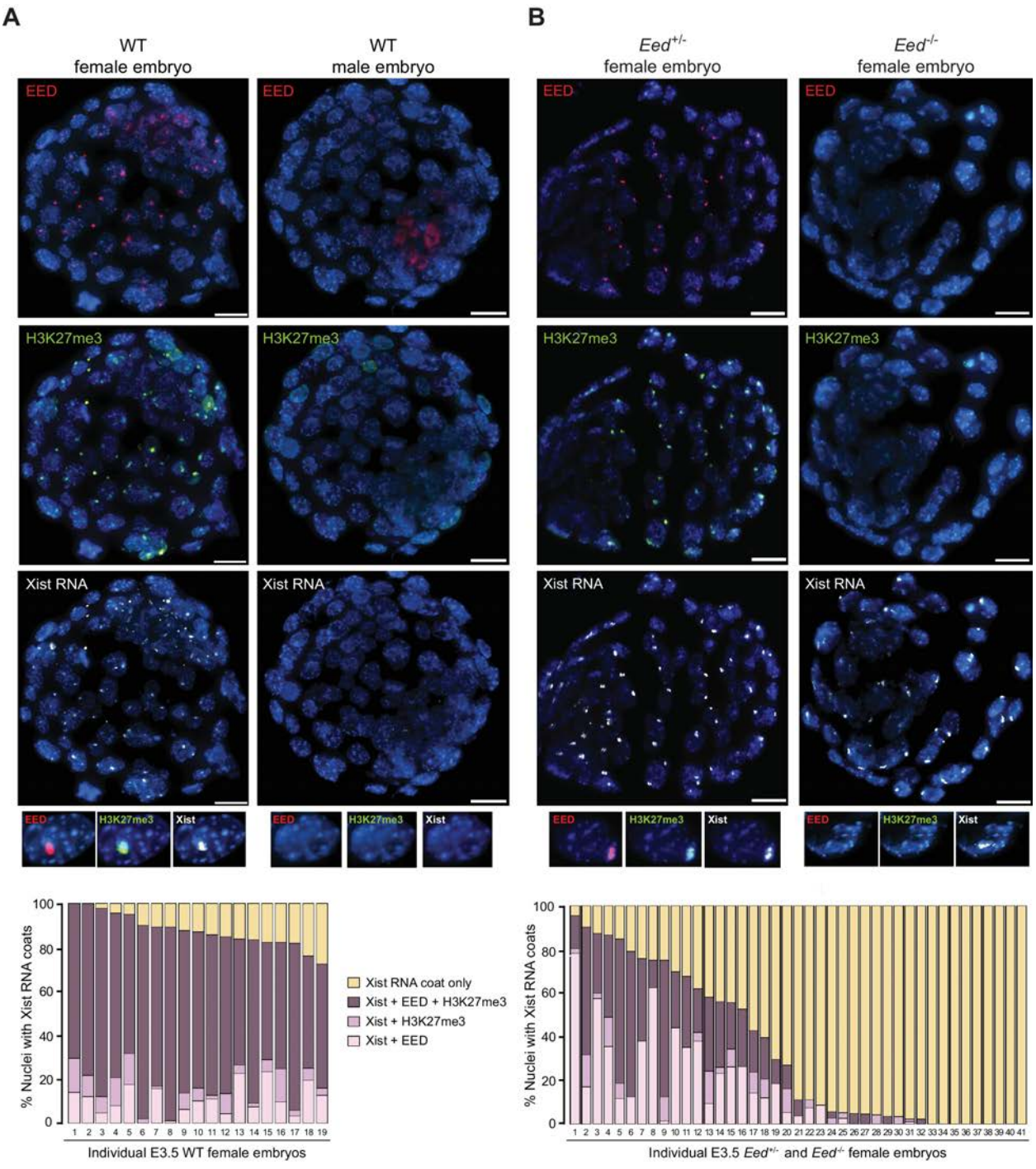
Antibody Name	Company	Catalog #	Dilution	Figure	Type
Monoclonal EED	(Sewalt et al., 1998)		1:1000	2.1	primary
			1:2500	2.2	primary
Polyclonal H3K27me3	Millipore	ABE44	1:5000	2.1	primary
			1:25000	2.2	primary
Alexa Fluor DαM 555	Invitrogen	A32773	1:300	2.1	secondary
			1:500	2.2	secondary
Alexa Fluor DαRb 488	Invitrogen	A21206	1:300	2.1	secondary
			1:500	2.2	secondary
Alexa Fluor DαRb 647	Invitrogen	A31573	1:300	2.3,2.4	secondary

Table 2.4. Allele-specific RNA FISH Probe Coordinates and Sequences

SNP Coordinate (mm9)	*M. musculus* Specific Probe (with 3' Quasar 570)	*M. molossinus* Specific Probe (with 3' Quasar 670)	Mask Probe
chrX:100664254	ATCACGCTGAAGACCCAGTTTTCTG	ATCATGCTGAAGACCCAGTTTTCTG	CAGAAAACTGGGTCTT
chrX:100669174	ATGCTGGGAGAACTGCTGTTGTGATG	ATGCCGGGAGAACTGCTGTTGTGATG	CATCACAACAGCAGTT
chrX:100676048	GCTCGGTGGATGAGTTTGAAAGAAAGTAC	GCTCAGTGGATGAGTTTGAAAGAAAGTAC	GTACTTTCTTTCAAACTCA
chrX:100676261	GTGTCGTTGGCATCCAAAATATTCATTG	GTGTTGTTGGCATCCAAAATATTCATTG	CAATGAATATTTTGGATGC
chrX:100677431	CTGCGGCTTCCGCGCAACACC	CTGCTGCTTCCGCGCAACACC	GGTGTTGCGCGG

Table 2.5. Human and Mouse Oocyte RNA-Seq Data Accession Numbers

Human Oocyte RNA-Seq		Mouse Oocyte RNA-Seq	
BioProject ID	**Run Number**	**BioProject ID**	**Run Number**
PRJNA146903	SRR351336	PRJDB21	DRR001701
PRJNA146903	SRR351337	PRJDB21	DRR001702
PRJEB8994	ERR841204	PRJNA154207	SRR385627

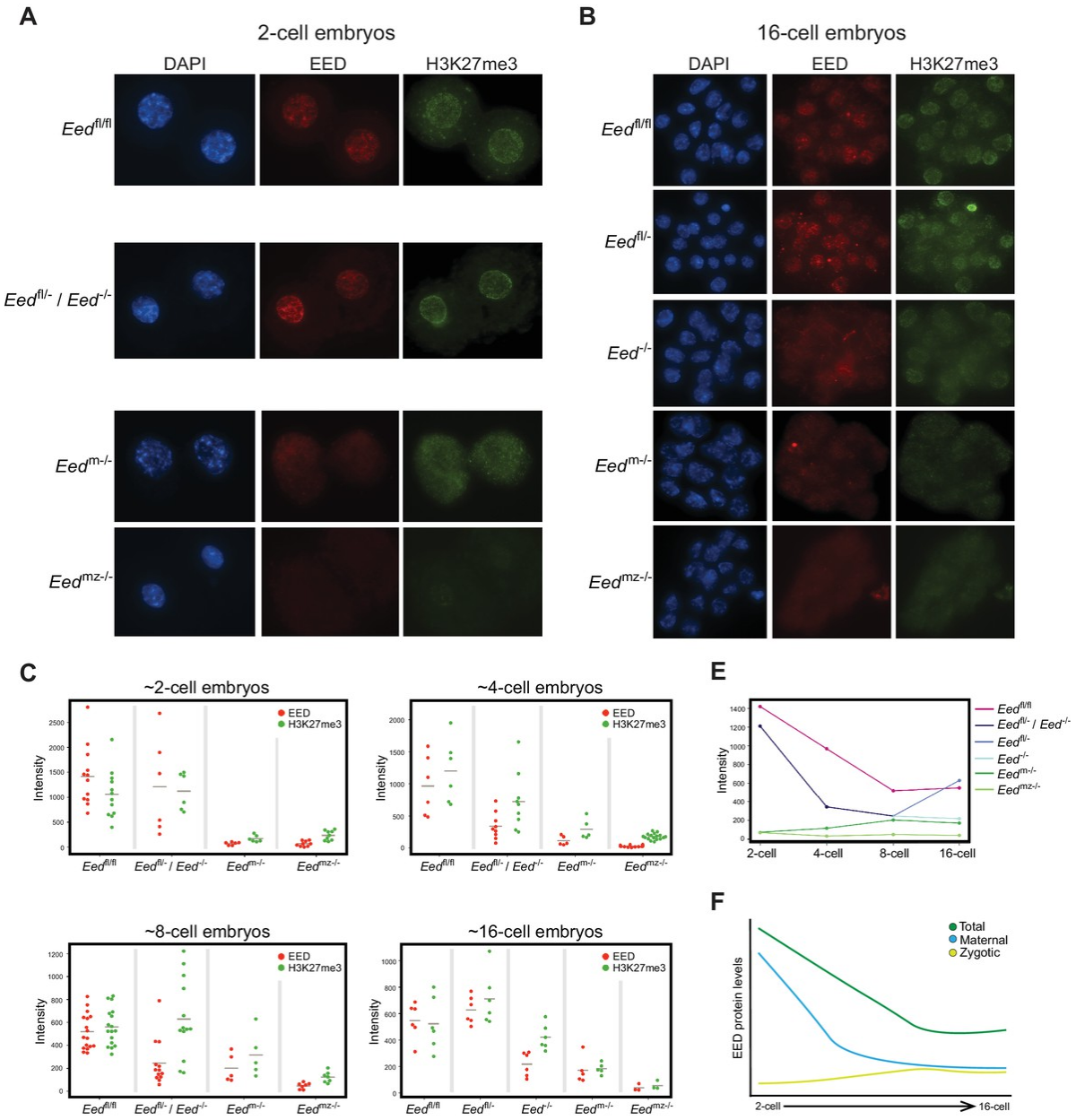


C

Distribution of E3.5 Embryos (12 litters)

Genotype	Females	Males
Eed$^{+/-}$	16	12
Eed$^{-/-}$	9	15

Figure 2.1. Coincident accumulation of EED and H3K27me3 on the inactive X-chromosome in blastocyst-stage WT, *Eed*$^{+/-}$ and *Eed*$^{-/-}$ mouse embryos (A,B) RNA FISH detection of *Xist* RNA (white) and immunofluorescence (IF) detection of EED (red) and H3K27me3 (green) in representative female and male wild-type (WT) (A) or female *Eed*$^{+/-}$ and *Eed*$^{-/-}$ (B) E3.0 – E3.5 blastocyst embryos. Nuclei are stained blue with DAPI. Scale bars, 20 μm. Embryos ranged in size from 23 to 57 nuclei. Bar plots, percentage of nuclei with coincident accumulation of *Xist* RNA and EED and/or H3K27me3 enrichment in individual embryos. (C) Genotype and sex distribution of *Eed*$^{+/-}$ and *Eed*$^{-/-}$ mouse blastocyst embryos from the cross in (B). The difference between the frequency of *Eed*$^{+/-}$ vs *Eed*$^{-/-}$ male and female embryos is not significant ($p > 0.05$, Two-tailed Student's T-test).



D

p-values of ~2-cell embryo EED IF intensity differences between genotypes

	Eed$^{fl/fl}$	*Eed*$^{fl/-}$ / *Eed*$^{-/-}$	*Eed*$^{m-/-}$	*Eed*$^{mz-/-}$
Eed$^{fl/fl}$				
Eed$^{fl/-}$ / *Eed*$^{-/-}$	0.58			
Eed$^{m-/-}$	5×10^{-5}	0.02		
Eed$^{mz-/-}$	2×10^{-6}	3×10^{-3}	0.88	

p-values of ~16-cell embryo EED IF intensity differences between genotypes

	Eed$^{fl/fl}$	*Eed*$^{fl/-}$	*Eed*$^{-/-}$	*Eed*$^{m-/-}$	*Eed*$^{mz-/-}$
Eed$^{fl/fl}$					
Eed$^{fl/-}$	0.29				
Eed$^{-/-}$	6×10^{-4}	3×10^{-5}			
Eed$^{m-/-}$	7×10^{-4}	5×10^{-5}	0.44		
Eed$^{mz-/-}$	5×10^{-4}	4×10^{-5}	0.01	0.08	

Figure 2.2. Assessment of maternal and zygotic EED expression in early preimplantation embryos (A, B) Immunofluorescent (IF) detection of EED (red) and H3K27me3 (green) in 2- and 16-cell *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$ / *Eed*$^{-/-}$, *Eed*$^{m-/-}$, and *Eed*$^{mz-/-}$ embryos. Nuclei are stained blue by DAPI. (C) Dot plots of EED and H3K27me3 IF signals in the five genotypes (*Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, *Eed*$^{-/-}$, *Eed*$^{m-/-}$, *Eed*$^{mz-/-}$) at the ~2-cell, ~4-cell, ~8-cell, and ~16-cell stage. Each dot represents an individual embryo. The gray line indicates mean fluorescence intensity. (D) Significance testing of differences in EED fluorescence intensity in ~2-cell embryos and ~16-cell embryos plotted in (C) (Two-tailed Student's T-test). (E) Mean EED fluorescence intensity from data in (C) plotted across early embryogenesis. (F) Model of change in maternal, zygotic, and total EED expression levels during early embryonic development.

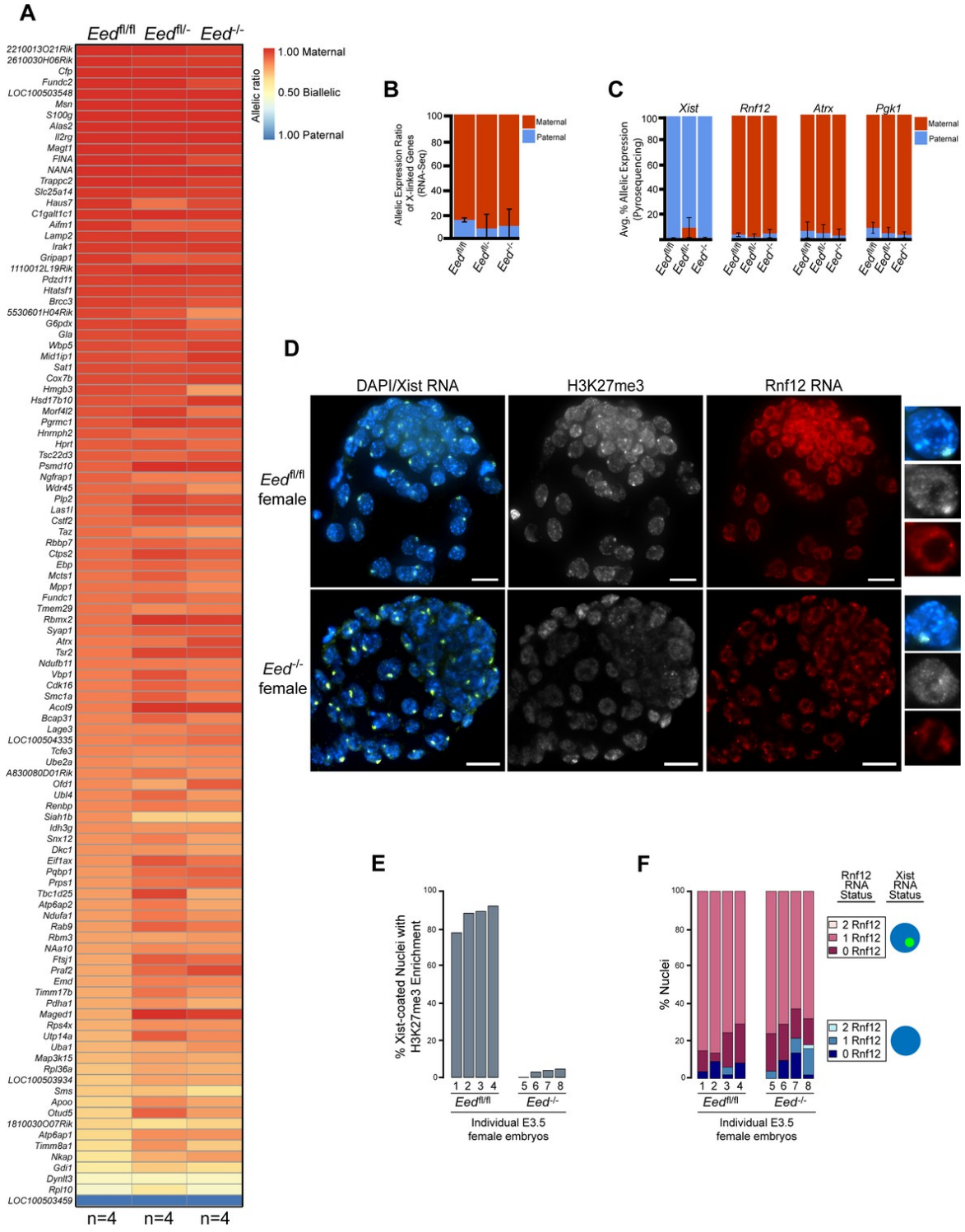

A
Eedfl/fl
Eedfl/-
Eed-/-
1.00 Maternal
0.50 Biallelic
1.00 Paternal
Allelic ratio
2210013O21Rik
2610030H06Rik
Cfp
Fundc2
LOC100503548
Msn
S100g
Alas2
Il2rg
Magt1
FlNA
NANA
Trappc2
Slc25a14
Haus7
C1galt1c1
Aifm1
Lamp2
Irak1
Gripap1
1110012L19Rik
Pdzd11
Htatsf1
Brcc3
5530601H04Rik
G6pdx
Gla
Wbp5
Mid1ip1
Sat1
Cox7b
Hmgb3
Hsd17b10
Morf4l2
Pgrmc1
Hnrnph2
Hprt
Tsc22d3
Psmd10
Ngfrap1
Wdr45
Plp2
Las1l
Cstf2
Taz
Rbbp7
Ctps2
Ebp
Mcts1
Mpp1
Fundc1
Tmem29
Rbmx2
Syap1
Atrx
Tsr2
Ndufb11
Vbp1
Cdk16
Smc1a
Acot9
Bcap31
Lage3
LOC100504335
Tcfe3
Ube2a
A830080D01Rik
Ofd1
Ubl4
Renbp
Siah1b
Idh3g
Snx12
Dkc1
Eif1ax
Pqbp1
Prps1
Tbc1d25
Atp6ap2
Ndufa1
Rab9
Rbm3
NAa10
Ftsj1
Praf2
Emd
Timm17b
Pdha1
Maged1
Rps4x
Utp14a
Uba1
Map3k15
Rpl36a
LOC100503934
Sms
Apoo
Otud5
1810030O07Rik
Atp6ap1
Timm8a1
Nkap
Gdi1
Dynlt3
Rpl10
LOC100503459
n=4
n=4
n=4
B
Allelic Expression Ratio of X-linked Genes (RNA-Seq)
Maternal
Paternal
Eedfl/fl
Eedfl/-
Eed-/-
C
Xist
Rnf12
Atrx
Pgk1
Avg. % Allelic Expression (Pyrosequencing)
Maternal
Paternal
D
DAPI/Xist RNA
H3K27me3
Rnf12 RNA
Eedfl/fl female
Eed-/- female
E
% Xist-coated Nuclei with H3K27me3 Enrichment
1 2 3 4
5 6 7 8
Eedfl/fl
Eed-/-
Individual E3.5 female embryos
F
% Nuclei
Rnf12 RNA Status
Xist RNA Status
2 Rnf12
1 Rnf12
0 Rnf12
2 Rnf12
1 Rnf12
0 Rnf12
1 2 3 4
5 6 7 8
Eedfl/fl
Eed-/-
Individual E3.5 female embryos

Figure 2.3. Lack of defective X-inactivation initiation in *Eed*$^{-/-}$ blastocysts (A) Allele-specific X-linked gene expression heat map of female *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, and *Eed*$^{-/-}$ blastocysts. Four embryos each of *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, and *Eed*$^{-/-}$ genotypes were sequenced individually and only genes with informative allelic expression in all samples are plotted (see Materials and Methods). Genes are ordered based on allelic expression in *Eed*$^{fl/fl}$ embryos. (B) Average allelic expression of the RNA-Seq data shown in (A). The mean allelic expression of X-linked genes lacks significant difference between each combination of the three genotypes ($p > 0.05$, Welch's two-sample T-test). (C) Pyrosequencing-based quantification of allelic expression of X-linked genes *Xist*, *Rnf12*, *Atrx* and *Pgk1* in *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, and *Eed*$^{-/-}$ blastocysts. Error bars represent the standard deviation of data from 3 to 6 independent blastocyst embryos. The mean allelic expression of all four genes lacks significant difference between each combination of the three genotypes ($p > 0.05$, Welch's two-sample T-test). (D) RNA FISH detection of *Xist* RNA (green), Rnf12 RNA (red), and IF detection of H3K27me3 (white) in representative *Eed*$^{fl/fl}$ or *Eed*$^{-/-}$ female blastocysts. Nuclei are stained blue with DAPI. Scale bars, 20 μm. Individual nuclei displaying representative categories of stains are shown to the right of each embryo. Embryos ranged in size from 39 to 100 nuclei. (E) Bar plot of percentage of nuclei with coincident accumulation of *Xist* RNA and H3K27me3 in individual *Eed*$^{fl/fl}$ and *Eed*$^{-/-}$ embryos. Each bar is an individual embryo. Embryo numbers under the bars correspond to the same embryos plotted in F). (F) Bar plots of percentage of nuclei with or without *Xist* RNA-coating and Rnf12 RNA expression in the embryos stained in D) and plotted in E). The numbers under the bars correspond to the same embryos plotted in E).

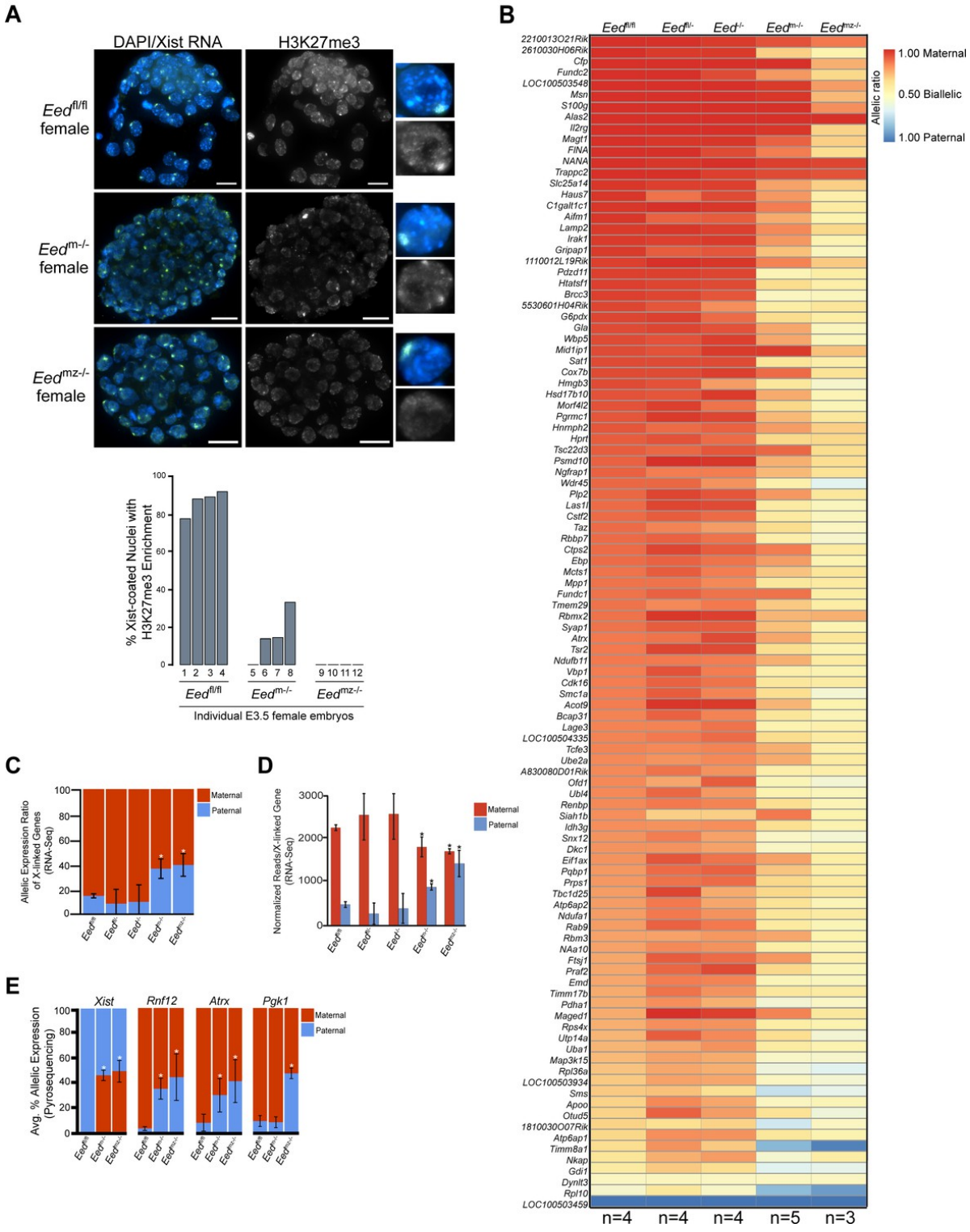
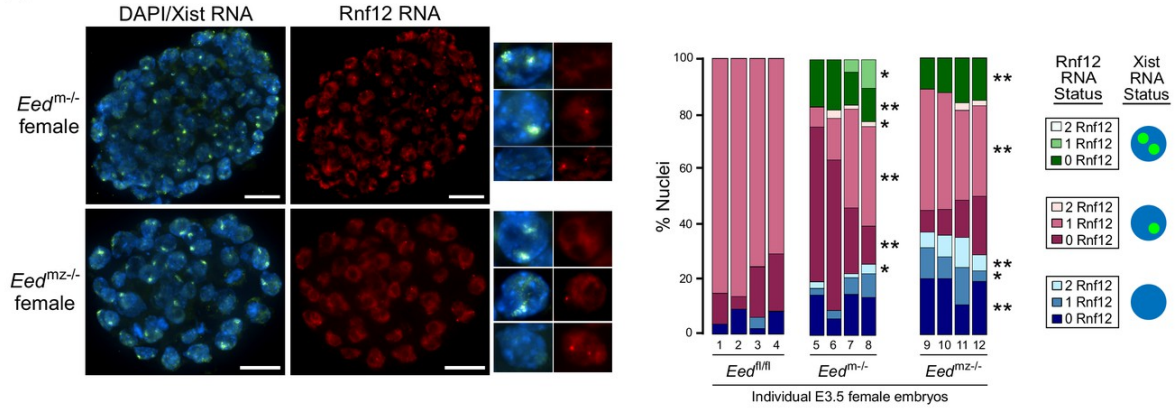
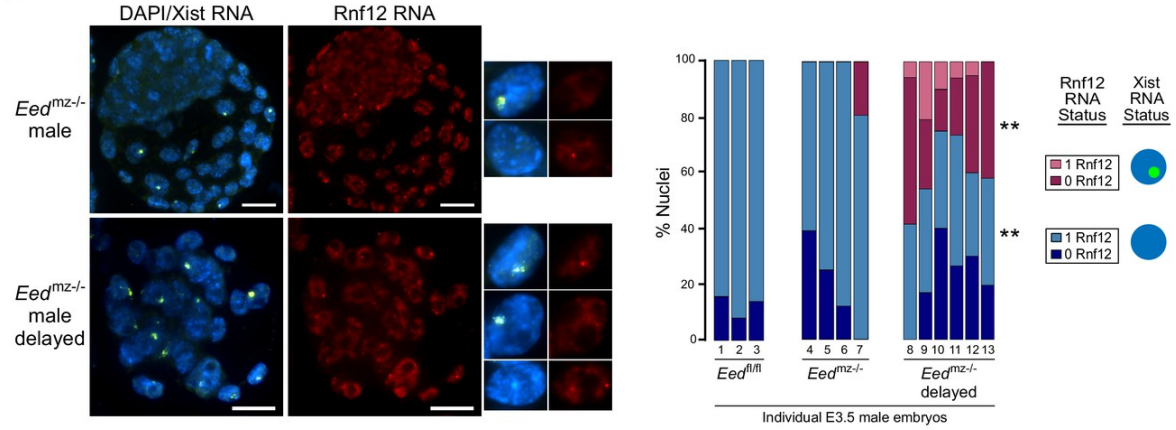

A
DAPI/Xist RNA
H3K27me3
Eed fl/fl female
Eed m-/- female
Eed mz-/- female
% Xist-coated Nuclei with H3K27me3 Enrichment
Individual E3.5 female embryos
B
Allelic ratio
1.00 Maternal
0.50 Biallelic
1.00 Paternal
n=4
n=4
n=4
n=5
n=3
C
Allelic Expression Ratio of X-linked Genes (RNA-Seq)
Maternal
Paternal
D
Normalized Reads/X-linked Gene (RNA-Seq)
E
Xist
Rnf12
Atrx
Pgk1
Avg. % Allelic Expression (Pyrosequencing)

Figure 2.4. Defective imprinted X-inactivation initiation in blastocysts lacking maternal EED (A) RNA FISH detection of *Xist* RNA (green) and IF stain for H3K27me3 (white) in representative *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ female blastocysts. Nuclei are stained blue with DAPI. Scale bars, 20 μm. *Eed*$^{fl/fl}$ blastocyst from Figure 2.3D shown for comparison. Right, individual representative nuclei. Mutant embryos ranged in size from 46 to 80 nuclei. Bar plot shows percentage of nuclei in each embryo analyzed that displayed H3K27me3 enrichment on the *Xist* RNA-coated X-chromosome. (B) Maternal:paternal X-linked gene expression heat map of female *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ blastocysts. Five *Eed*$^{m-/-}$ and three *Eed*$^{mz-/-}$embryos were sequenced individually and only genes with informative allelic expression in all samples are plotted (see Materials and Methods). *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, and *Eed*$^{-/-}$ data from Figure 2.3A shown for comparison. Genes are ordered based on allelic expression in *Eed*$^{fl/fl}$ embryos. (C) Average maternal:paternal X-linked gene expression ratio from the RNA-Seq data shown in B). *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, and *Eed*$^{-/-}$ data from Figure 2.3B shown for comparison. The mean allelic expression of X-linked genes is significantly different between *Eed*$^{m-/-}$ and *Eed*$^{fl/fl}$, and *Eed*$^{mz-/-}$ and *Eed*$^{fl/fl}$ blastocysts. ($p < 0.05$, Welch's two-sample T-test). (D) Average normalized maternal and paternal X-linked gene expression in blastocysts. Maternal and paternal X-linked gene expression is significantly different between *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos compared to *Eed*$^{fl/fl}$ embryos (*, $p < 0.05$, Two-tailed Student's T-test). (E) Pyrosequencing-based quantification of allelic expression of X-linked genes in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ blastocysts. *Eed*$^{fl/fl}$ data from Figure 2.3C are shown for comparison. Error bars represent the standard deviation of data from 3 to 6 independent blastocyst embryos. The mean allelic expression of *Xist*, *Rnf12*, and *Atrx* is significantly different between *Eed*$^{fl/fl}$ and *Eed*$^{m-/-}$ embryos. The mean allelic expression of *Xist*, *Rnf12*, *Pgk1*, and *Atrx* is significantly different between *Eed*$^{fl/fl}$ and *Eed*$^{mz-/-}$ embryos ($p < 0.05$, Welch's two-sample T-test).

A



B



C

Eed$^{m-/-}$ Liveborn Animals

Dam	Sire	Number of litters	Number of pups (avg/litter)	Total females	Total males	*Eed*$^{fl/fl}$ pups	*Eed*$^{+/-}$ pups
Eed$^{fl/fl}$	*Mus musculus*-derived *Eed*$^{fl/fl}$	9	61 (6.8)	31 (53%)	30 (47%)	61	0
Eed$^{fl/fl}$;*Zp3-Cre*	*Mus musculus*-derived WT	5	12 (2.4)	9 (75%)	3 (25%)	0	12
Eed$^{fl/fl}$;*Zp3-Cre*	*Mus molossinus*-derived WT	4	10 (2.5)	7 (70%)	3 (30%)	0	10

Figure 2.5. RNA FISH analysis of X-inactivation in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ blastocysts (A, B) RNA FISH detection of *Xist* RNA (green) and Rnf12 RNA (red) in representative *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ female (A) and *Eed*$^{mz-/-}$ male (B) blastocysts. Nuclei are stained blue with DAPI. Scale bars, 20 μm. Individual nuclei of representative categories of stain are shown to the right of each embryo. *Eed*$^{fl/fl}$ female data from Figure 2.3D shown for comparison. Mutant female embryos ranged in size from 46 to 80 nuclei. Fully developed mutant male embryos ranged in size from 53 to 110 nuclei. Delayed mutant male embryos ranged in size from 30 to 40 nuclei. Bar plot shows percentage of nuclei in each embryo with *Xist* RNA coats and/or Rnf12 RNA expression. Each bar represents an individual embryo and embryo numbers under the bars correspond to the same female embryos plotted in Figure 2.3A. *, $p < 0.05$; **, $p < 0.01$, Two-tailed Student's T-test, between *Eed*$^{m-/-}$ and *Eed*$^{fl/fl}$, or *Eed*$^{mz-/-}$ and *Eed*$^{fl/fl}$. (C) Data showing the number of *Eed*$^{m-/-}$ embryos which can live to term compared to *Eed*$^{fl/fl}$ embryos. WT, wild-type. Table shows *Eed*$^{m-/-}$ litters sired by *Mus musculus*-derived male or *Mus molossinus*-derived male. Male *Eed*$^{m-/-}$ offspring are underrepresented compared to females, $p = 0.02$, Two-tailed Student's T-test.

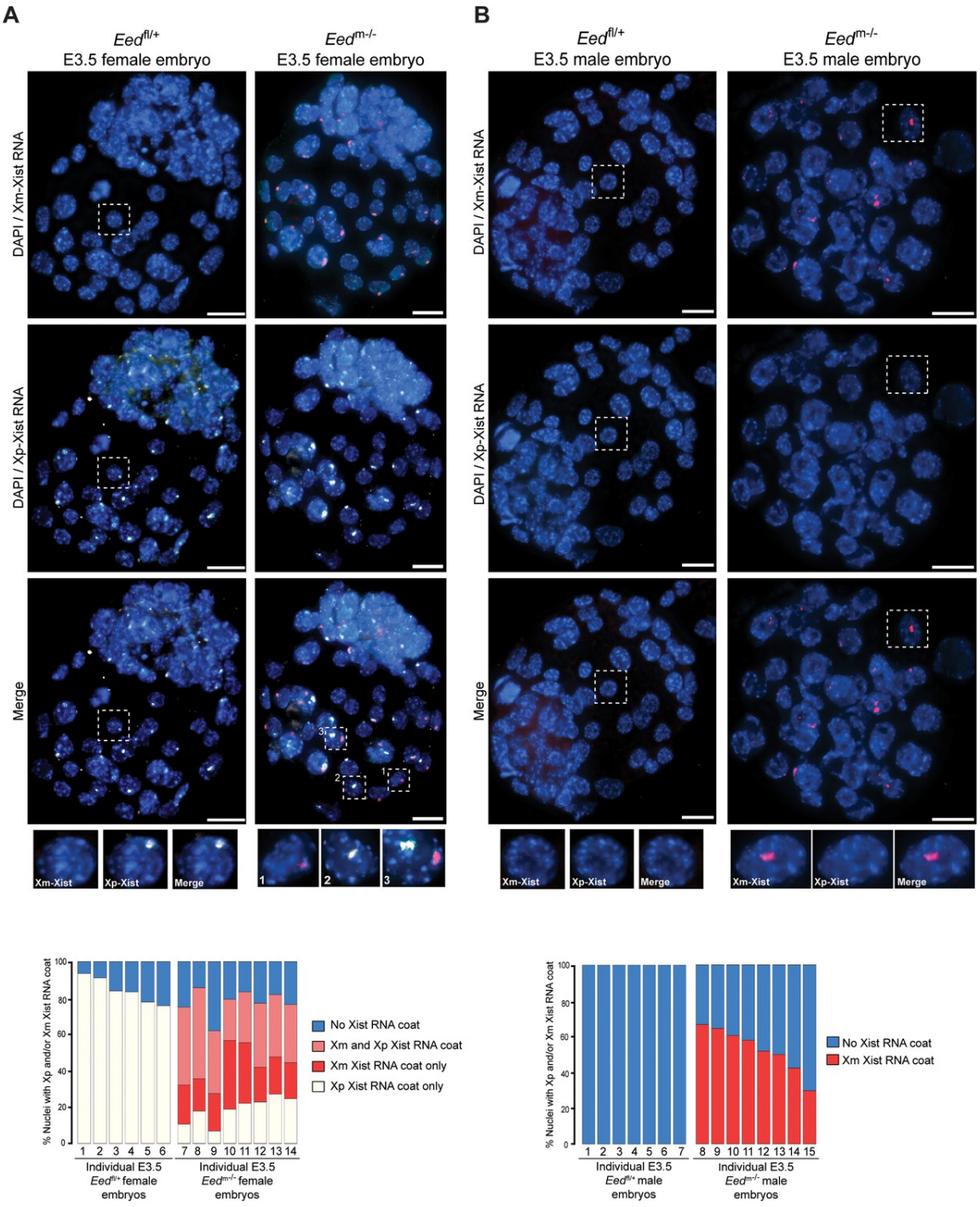


Figure 2.6. Switching of imprinted to random X-inactivation in E3.5 embryos lacking maternal EED (A, B) Allele-Specific *Xist* RNA FISH in *Eed*$^{fl/+}$ and *Eed*$^{m-/-}$ male and female E3.0-E3.5 blastocyst embryos. *Xist* RNA expressed from the maternal X-chromosome is indicated in red and from the paternal X-chromosome in white. Representative embryos are depicted. Nuclei are stained blue with DAPI. Scale bars, 20 μm.

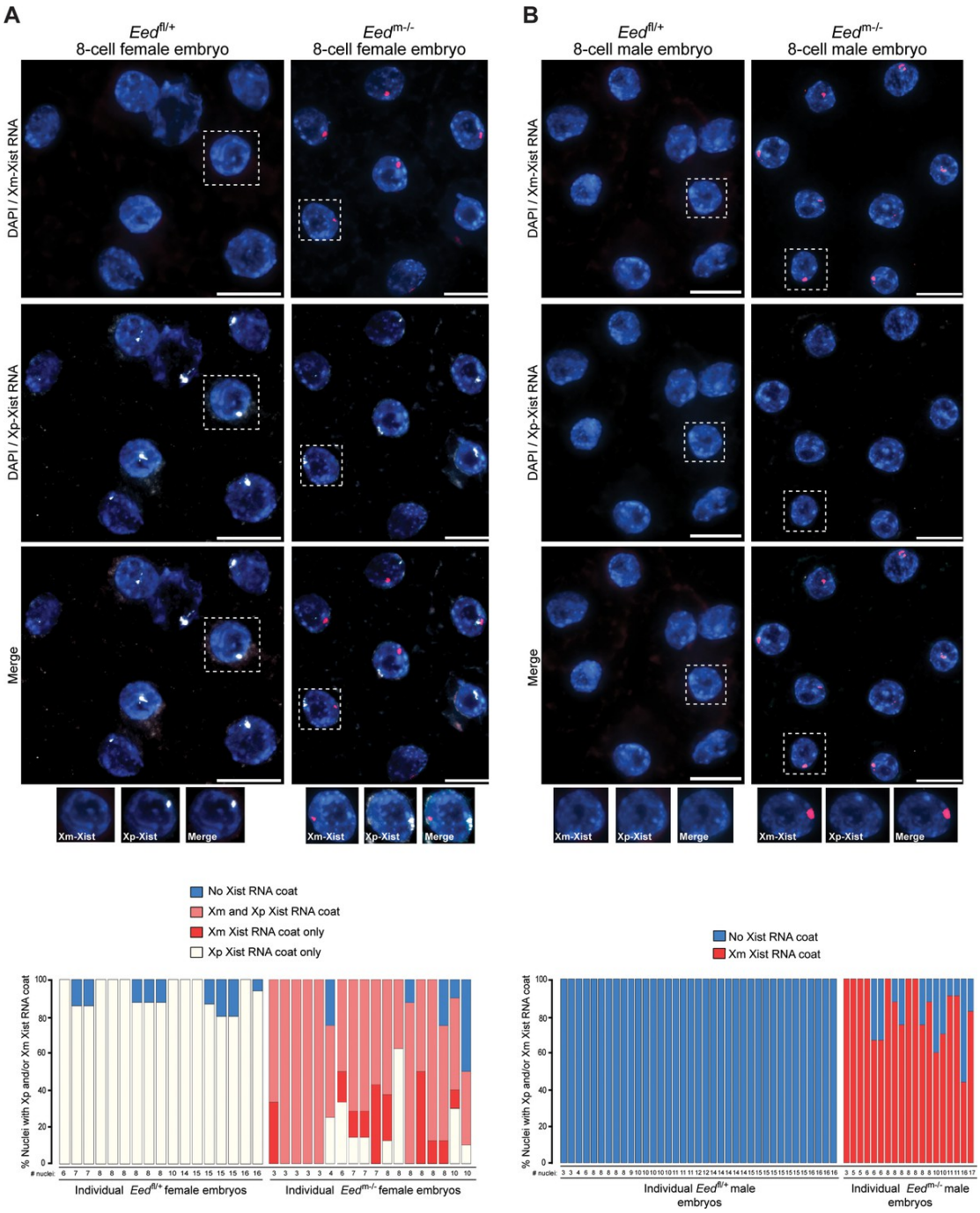


Figure 2.7. Switching of imprinted to random X-inactivation in 3–16 cell embryos lacking maternal EED (A, B) Allele-Specific *Xist* RNA FISH in *Eed*$^{fl/+}$ and *Eed*$^{m-/-}$ female and male 3–16 cell embryos. *Xist* RNA expressed from the maternal X-chromosome is indicated in red and from the paternal X-chromosome in white. Representative embryos are depicted. Nuclei are stained blue with DAPI. Scale bars, 20 μm.

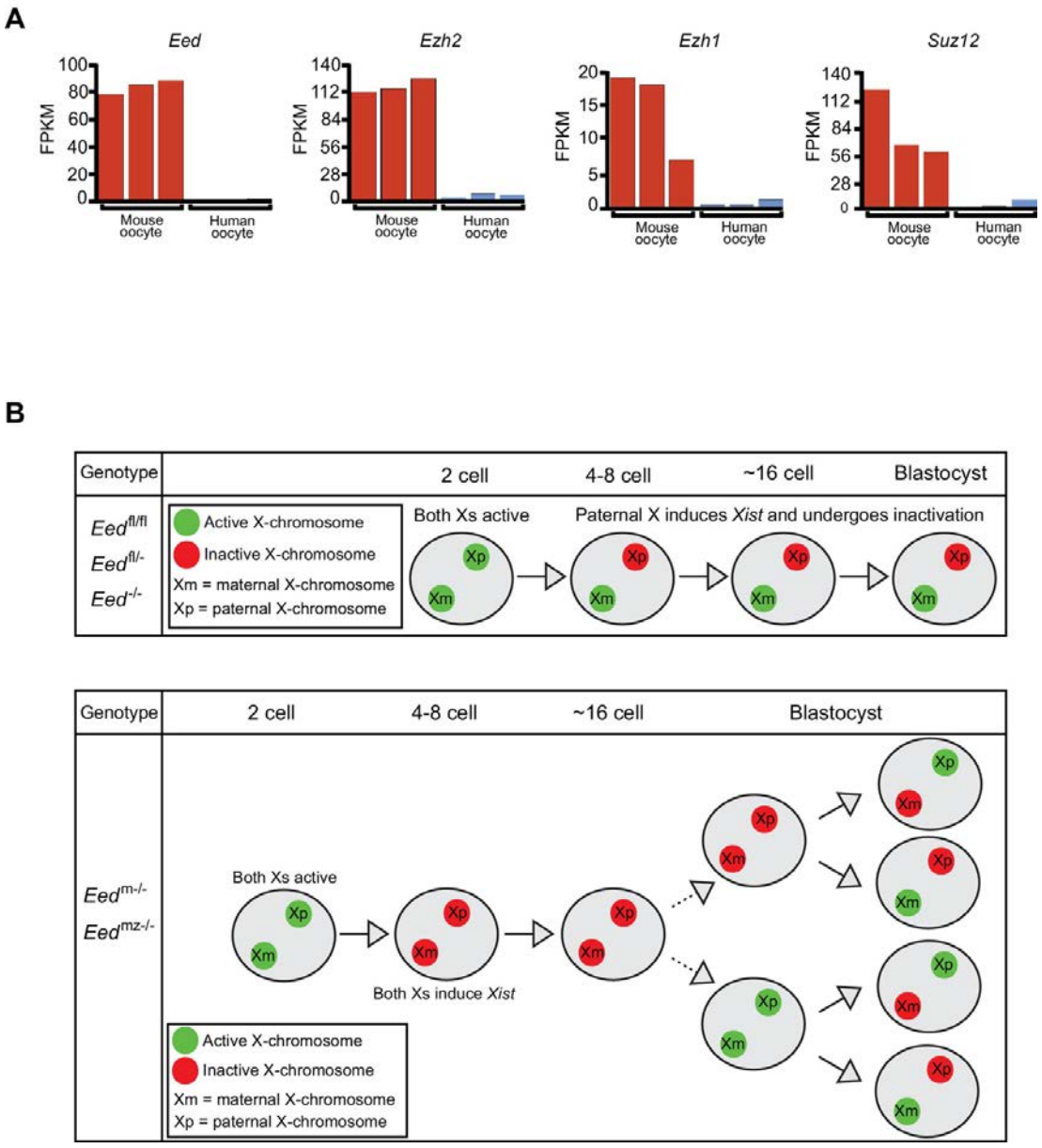
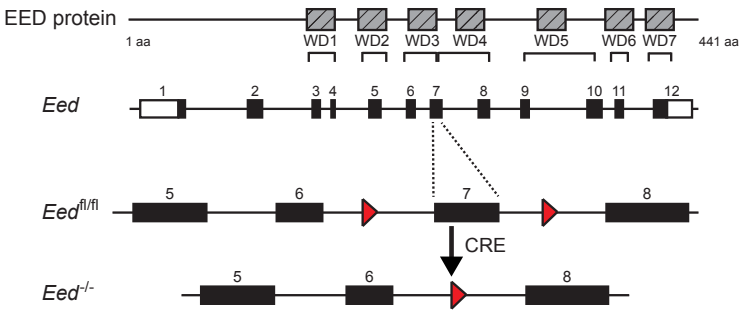


Figure 2.8. Lack of PRC2 expression in human oocytes and a path to randomization of X-inactivation in early embryos (A) Expression levels by RNA-Seq of core PRC2 components in human and mouse oocytes. (B) Model of maternal PRC2 function during preimplantation mouse embryogenesis.

A

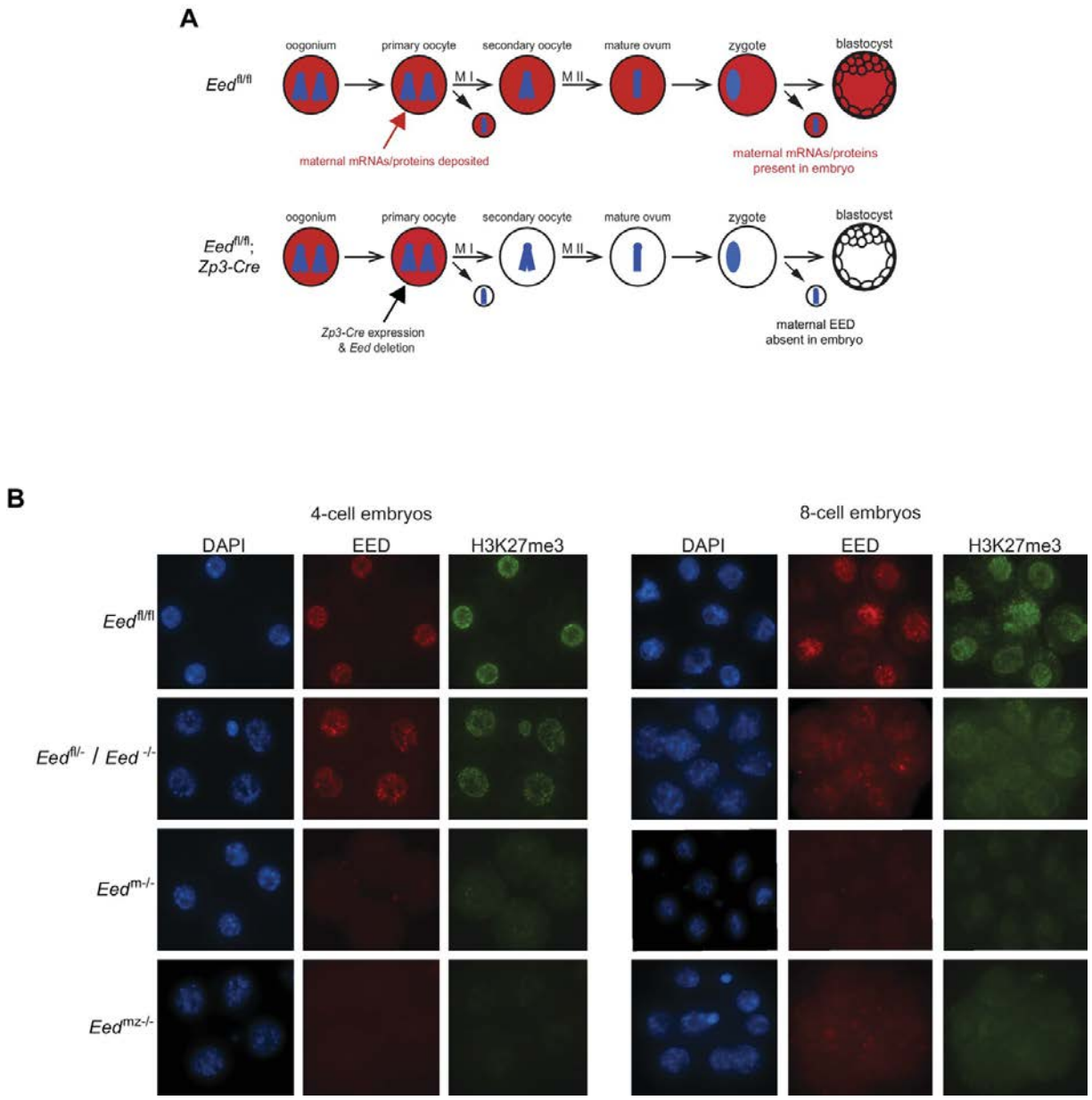


B

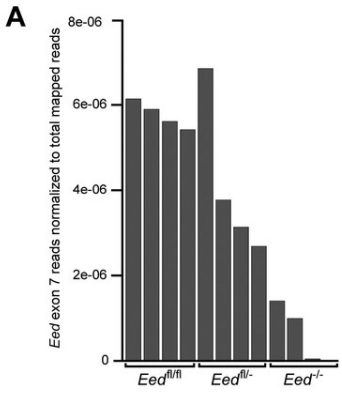
Eed^{fl} deletion efficiency by *Prm-Cre*

Dam	Sire	Number of litters	Number of pups	*Eed*$^{fl/fl}$ pups	*Eed*$^{fl/-}$ or *Eed*$^{+/-}$ pups
Eed$^{fl/fl}$	*Eed*$^{fl/fl}$;*Prm*-Cre	6	36	4 (11%)	32 (89%)
Eed$^{fl/fl}$ or WT	*Eed*$^{fl/-}$;*Prm*-Cre	17	88	8 (9%)	80 (91%)

Supplemental Figure 2.1. Generation of *Eed*$^{-/-}$ embryos (A) Schematic depicting the deletion of floxed *Eed* exon seven by CRE recombinase. (B) Breeding data showing the efficiency of *Prm-Cre* deletion of the *Eed*^{fl} allele.

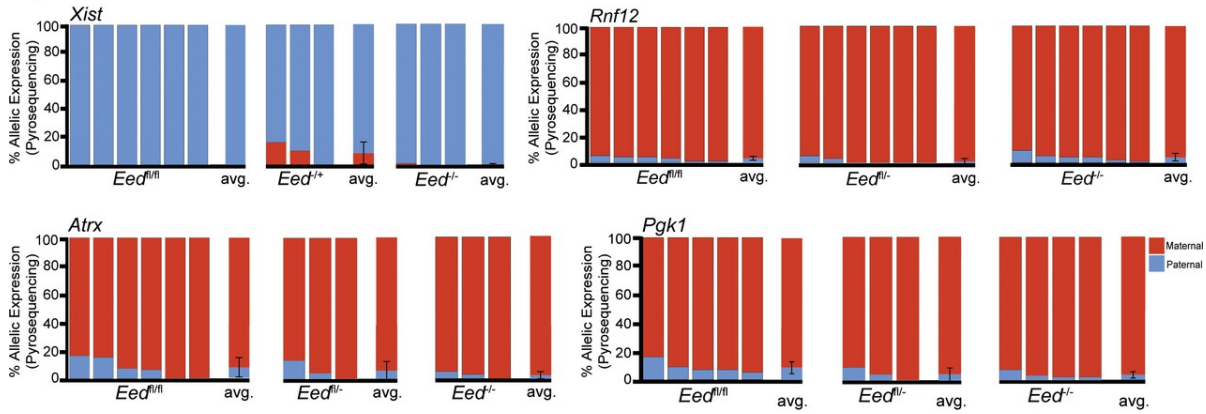


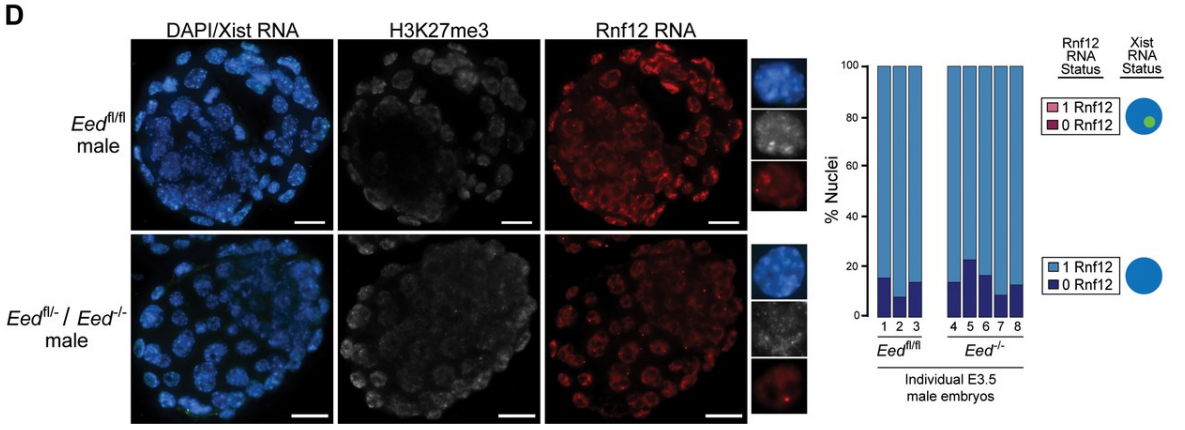
Supplemental Figure 2.2. Analysis of EED and H3K27me3 fluorescence intensity in *Eed* mutants (A) Schematic depicting the deletion of *Eed* exon seven by *Zp3-Cre* used to generate embryos maternally null for *Eed*. (B) Representative images of $Eed^{fl/fl}$, $Eed^{fl/-}$, $Eed^{-/-}$, $Eed^{m-/-}$, and $Eed^{mz-/-}$ 4- and 8-cell embryos stained by IF for EED and H3K27me3. Nuclei are indicated by blue DAPI stain, EED stain is indicated in red, and H3K27me3 stain is indicated in green.



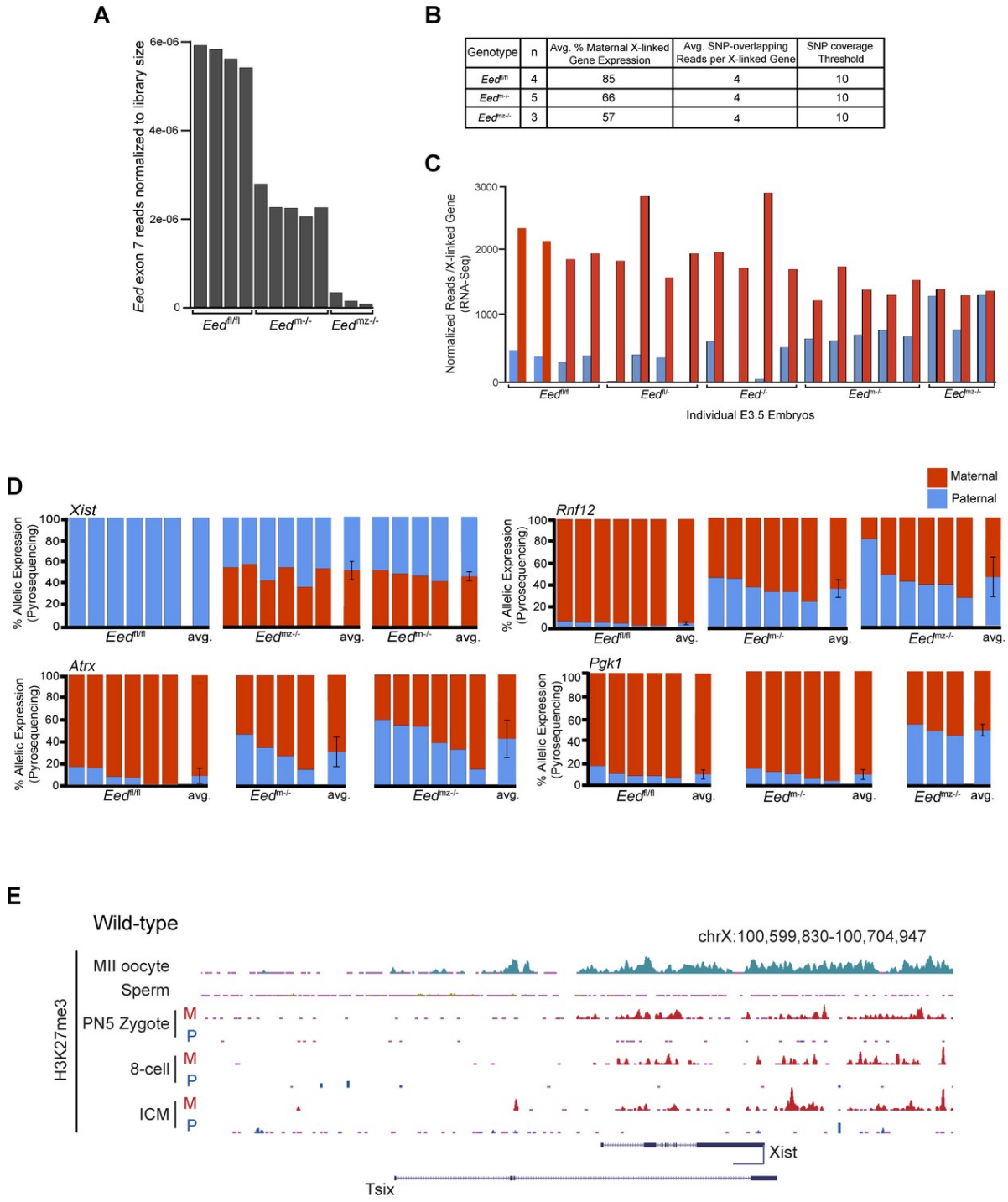
B

Genotype	n	Avg. % Maternal X-linked Gene Expression	Avg. SNP-overlapping Reads per X-linked Gene	SNP Coverage Threshold
Eed$^{fl/fl}$	4	85	4	10
Eed$^{fl/-}$	4	90	4	10
Eed$^{-/-}$	4	88	4	10



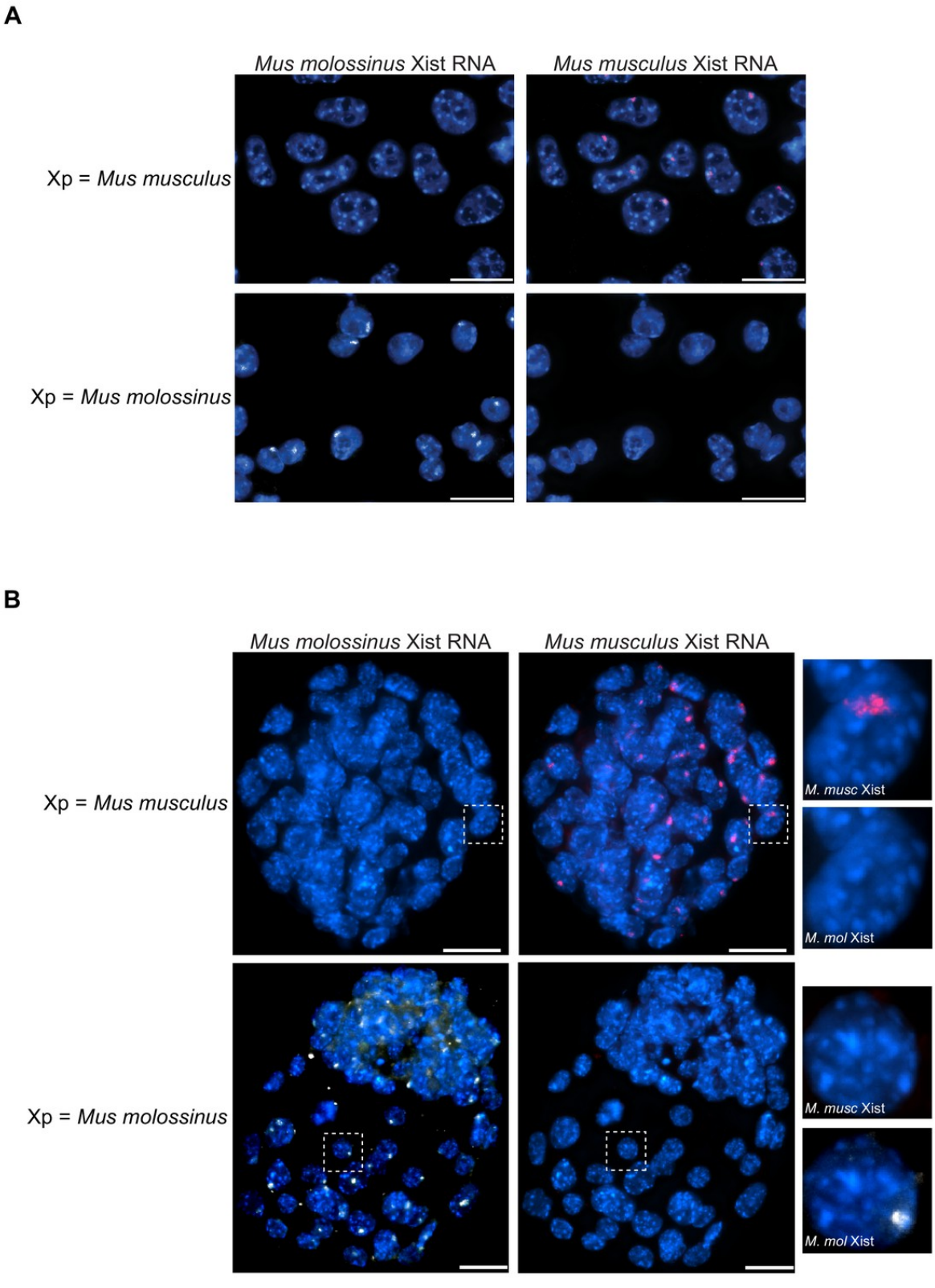


Supplemental Figure 2.3. X-linked gene expression in *Eed*$^{-/-}$ embryos (A) Validation of genotypes of E3.5 *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, and *Eed*$^{-/-}$ female blastocyst embryos. *Eed* exon 7 RNA-Seq reads are normalized to total mapped RNA-Seq reads. (B) Table describing the RNA-Seq genotypes, number of sequenced embryos, average % maternal X-linked gene expression, average number of SNPs per X-linked gene, and the SNP overlapping read coverage threshold. (C) Pyrosequencing-based quantification of allelic expression of X-linked genes Xist, Rnf12, Atrx, and Pgk1 in individual *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, and *Eed*$^{-/-}$ female blastocysts. Error bars, standard deviation of data from 3 to 6 independent embryos. The mean allelic expression of all four genes lacks significant difference between each combination of the three genotypes ($p > 0.05$, Welch's two-sample T-test). (D) RNA FISH detection of *Xist* RNA (green), Rnf12 RNA (red), and IF detection of H3K27me3 (white) in representative *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, or *Eed*$^{-/-}$ male blastocysts. Nuclei are stained blue with DAPI. Scale bars, 20 μm. Right of each embryo, individual nuclei displaying representative categories of stains. Embryos ranged in size from 56 to 65 nuclei. Bar plot, percentage of nuclei with or without *Xist* RNA-coating and Rnf12 RNA expression.

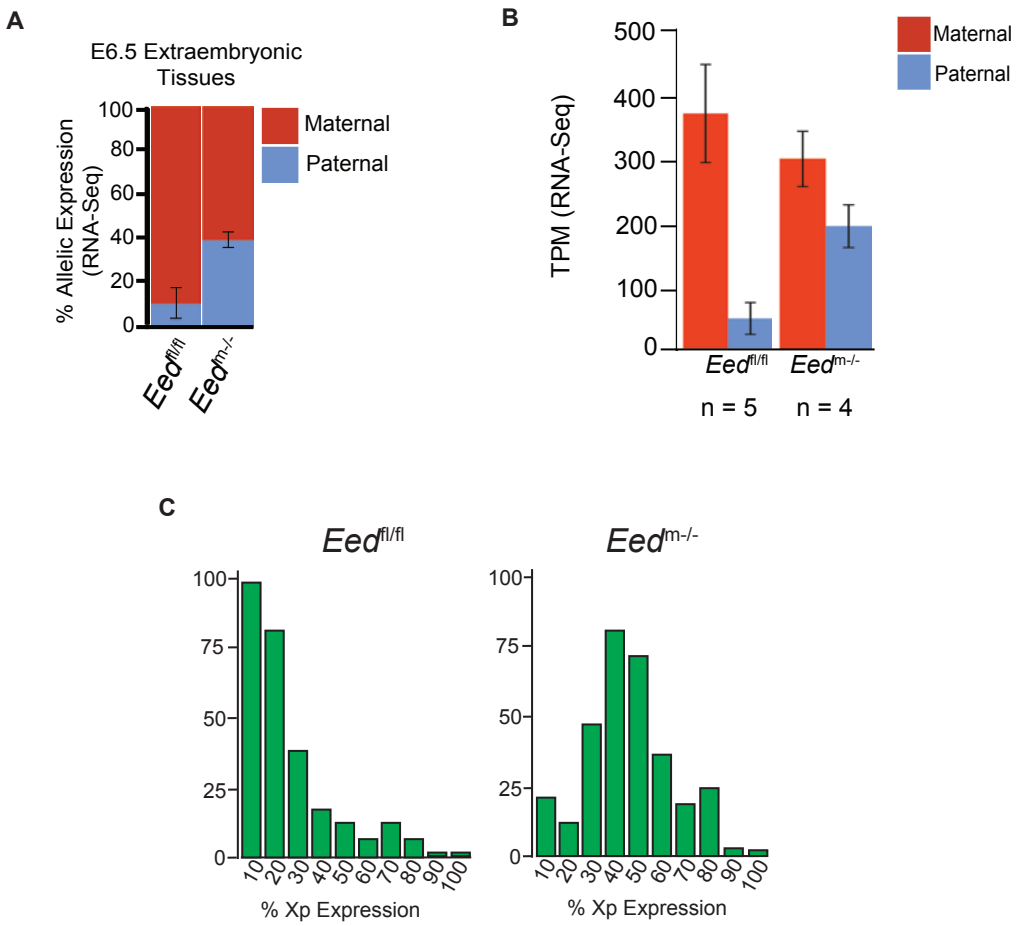


Genotype	n	Avg. % Maternal X-linked Gene Expression	Avg. SNP-overlapping Reads per X-linked Gene	SNP coverage Threshold
Eed$^{fl/fl}$	4	85	4	10
Eed$^{m-/-}$	5	66	4	10
Eed$^{mz-/-}$	3	57	4	10

Supplemental Figure 2.4. Generation and X-linked gene profiling of *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos (A) Validation of genotypes of E3.5 female embryos. *Eed* exon 7 RNA-Seq reads are normalized to total mapped RNA-Seq reads. (B) Table describing the RNA-Seq genotypes, number of sequenced embryos, average percentage maternal X-linked gene expression, average number of SNPs per X-linked gene, and the SNP overlapping read coverage threshold. (C) Normalized maternal or paternal reads per X-linked gene in individual *Eed*$^{fl/fl}$, *Eed*$^{fl/-}$, *Eed*$^{-/-}$, *Eed*$^{m-/-}$, and *Eed*$^{mz-/-}$ female E3.5 blastocysts. (D) Pyrosequencing-based quantification of allelic expression of X-linked genes *Xist*, *Rnf12*, *Atrx*, and *Pgk1* in individual *Eed*$^{fl/fl}$, *Eed*$^{m-/-}$, and *Eed*$^{mz-/-}$ female E3.5 blastocysts. Error bars, standard deviation of data from 3 to 6 independent embryos. The mean allelic expression for *Xist*, *Rnf12*, and *Atrx* is significantly different between *Eed*$^{fl/fl}$ and *Eed*$^{m-/-}$embryos ($p < 0.05$, Welch's two-sample T-test). The mean allelic expression for *Xist*, *Rnf12*, *Atrx*, and *Pgk1* is significantly different between *Eed*$^{fl/fl}$ and *Eed*$^{mz-/-}$ embryos ($p < 0.05$, Welch's two-sample T-test). The mean allelic expression of *Pgk1* is significantly different between *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ embryos ($p < 0.05$, Welch's two-sample T-test). (E) Allele-specific H3K27me3 ChIP-Seq at the *Xist* locus of wild-type MII oocyte, sperm, PN5 zygote, 8 cell embryo, and inner cell mass (ICM) (Zheng *et al.*, 2016).


A
Mus molossinus Xist RNA
Mus musculus Xist RNA
Xp = Mus musculus
Xp = Mus molossinus
B
Mus molossinus Xist RNA
Mus musculus Xist RNA
Xp = Mus musculus
Xp = Mus molossinus
M. musc Xist
M. mol Xist
M. musc Xist
M. mol Xist

Supplemental Figure 2.5. Characterization of allele-specific *Xist* RNA FISH probe in cells and embryos (A) Female Trophoblast stem (TS) cells (top panel) and extraembryonic endoderm (XEN) stem cells (bottom panel) stained with an allele-specific *Xist* RNA FISH probe. Both TS cells and XEN cells express *Xist* from and undergo imprinted X-inactivation of the paternal X-chromosome (Kunath et al., 2005; Tanaka et al., 1998). The TS cells are derived from a cross of JF1 *Mus molossinus* dam with a 129/S1-derived *Mus musculus* sire. The XEN cells are generated from a cross of 129/S1 *Mus musculus* dam and JF1 *Mus molossinus*-derived sire. In the TS cells, the paternal-X is therefore *Mus musculus* derived while in the XEN cells the paternal-X is JF1 *Mus molossinus* derived. *Mus musculus*-specific *Xist* RNA FISH probe detects the complimentary *Xist* RNA in red and the *Mus molossinus*-specific *Xist* RNA FISH probe detects its complimentary *Xist* RNA in white. (B) $Eed^{fl/+}$ female E3.5 embryos stained with the same allele-specific *Xist* RNA FISH probe as in (A). Top panels, representative stained embryo derived from a cross of $Eed^{fl/fl}$; $XJF^{1}X^{JF1}$ *Mus molossinus*-derived dam with a *Mus musculus* sire. Bottom panels, representative stained embryo from an $Eed^{fl/fl}$ *Mus musculus*-derived dam with a JF1 *Mus molossinus*-derived sire (this embryo is also shown in Figure 2.6A). Due to imprinted X-inactivation, both E3.5 embryos are expected to express *Xist* RNA from their paternal X-chromosome.



Supplemental Figure 2.6. Characterization of E6.5 female mouse extraembryonic tissues by allele-specific RNA-Seq (A) Average maternal:paternal X-linked gene expression ratio from allele-specific RNA-Seq of E6.5 female extraembryonic tissues. The mean allelic expression of X-linked genes is significantly different between *Eed*$^{fl/fl}$ and *Eed*$^{m-/-}$ tissues. ($p < 0.05$, Welch's two-sample T-test). (B) Absolute TPM expression of maternal and paternal X-linked genes in E6.5 female extraembryonic tissues. Error bars represent standard deviation. (C) Histograms depicting the average number of X-linked genes undergoing various degrees (in 10% increments) of expression from the paternal-X in *Eed*$^{fl/fl}$ and *Eed*$^{m-/-}$ extraembryonic tissues.

Chapter 3
Distinct Requirements for PRC2 Components EZH1/2 and EED in Imprinted X-inactivation

Abstract

Imprinted X-inactivation results in the silencing of genes on the paternal X chromosome in cells of early female mouse embryos and is a paradigm of transgenerational epigenetic inheritance. Although the paternal-X is inactivated in early female mouse embryos, the epigenetic imprint resides on the maternal-X. During oogenesis, Polycomb repressive complex (PRC)-catalyzed histone modifications repress the *Xist* locus. In the absence of the core PRC2 subunit EED in the oocyte (maternal), PRC2-catalyzed histone H3K27me3 modification is depleted and instead of imprinted X-inactivation early female embryos switch to random X-inactivation. Here, we find that loss of maternal and zygotic PRC2 methyltransferases EZH2 and EZH1 (EZH2/1) does not phenocopy the loss of EED in early embryos. Whereas the absence of maternal/zygotic EED results in nearly equal expression of maternal and paternal X-linked genes, the loss of maternal/zygotic EZH2/1 results in the preferential expression of maternal X-linked genes in the early embryo. Together, these data suggest that EED can function independently of PRC2 and H3K27me3 to silence paternal X-linked genes during imprinted X-inactivation.

Introduction

X-inactivation equalizes X-linked gene expression between XX female and XY male mammals via transcriptional silencing of genes on one of the two X chromosomes in early female embryos (Lyon, 1961). Failure of X-inactivation in females leads to lethality during embryonic

development (Marahrens *et al.*, 1997; Sarkar *et al.*, 2015; Takagi, 1980; Takagi and Abe, 1990). Two distinct forms of X-inactivation characterize the mouse embryo: imprinted and random. Imprinted X-inactivation results in the inactivation exclusively of the paternal X chromosome and initiates in all cells of the preimplantation female embryo (Mak *et al.*, 2004; Takagi et al., 1978). Imprinted inactivation of the paternal-X is then stably maintained in the extra-embryonic lineages of the trophectoderm and the primitive endoderm of post-implantation embryos (Harper *et al.*, 1982; Takagi and Sasaki, 1975; West *et al.*, 1977). In the peri-implantation embryo, the epiblast progenitor cells reactivate silenced genes on the paternal X-chromosome. Subsequently, individual differentiating epiblast cells randomly inactivate either the maternal or the paternal X chromosome (Gardner and Lyon, 1971; Lyon, 1961; Mak *et al.*, 2004).

Both imprinted and random X-inactivation are models of epigenetic regulation. Despite equivalent sequences, one X chromosome is inactivated while the other X remains active in a shared nucleoplasm. Additionally, once an X chromosome is inactivated, replicated copies of that X chromosome remain inactive through many rounds of cell division (Morey and Avner, 2011). Imprinted X-inactivation is also a paradigm of transgenerational epigenetic regulation due to its stable parent-of-origin-specific inactivation pattern (Huynh and Lee, 2003; Mak *et al.*, 2004; Takagi and Sasaki, 1975).

In the developing mouse embryo, X-inactivation is characterized by a set of temporally-specified regulatory events. At the onset of both imprinted and random X-inactivation, the X-linked non-protein coding *Xist* RNA is expressed from the prospective inactive-X (Gayen *et al.*, 2015; Kalantry et al., 2009; Kay *et al.*, 1994; Shiura and Abe, 2019). At the two-cell stage, *Xist* RNA is upregulated from the paternally inherited X chromosome. *Xist* RNA then physically coats the paternal-X beginning at the four-cell stage (Deng *et al.*, 2014; Kalantry et al., 2009;

Kay *et al.*, 1994; Okamoto *et al.*, 2004). The progressive accumulation of *Xist* RNA on the paternal-X coincides with the gradual silencing of paternal X-linked genes. Furthermore, histone deacetylation occurs on the inactive-X at the onset of X-inactivation (Keohane et al., 1998). Acetylation of the lysine residues of core histones is a well-characterized chromatin modification that is associated with transcriptional activation (Hebbes et al., 1988). Histone deacetylation, conversely, is associated with transcriptional repression. By the 8-16 cell morula stage, silencing factors including the Polycomb repressive complex 2 (PRC2) and PRC2-catalyzed chromatin mark histone H3 lysine 27 tri-methylation (H3K27me3) are cytologically enriched on the inactive-X (Erhardt et al., 2003; Okamoto *et al.*, 2004; Silva *et al.*, 2003). Because the mis-expression of *Xist* leads to the ectopic accumulation of PRC2 proteins and H3K27me3, *Xist* RNA is thought to directly or indirectly recruit PRC2 to the inactive-X (de la Cruz et al., 2005b; Kohlmaier *et al.*, 2004; Plath *et al.*, 2003; Silva *et al.*, 2003). PRC2 has thus been suggested to contribute to the silencing of X-linked genes (Kalantry *et al.*, 2006; Plath *et al.*, 2003; Silva *et al.*, 2003).

Although PRC2 and H3K27me3 are enriched on the inactive-X both in cultured stem cells and in the embryo (Kalantry *et al.*, 2006; Plath *et al.*, 2003; Silva *et al.*, 2003), we and others found that embryos and cultured stem cells lacking zygotically expressed core PRC2 protein EED initiate and maintain silencing of many paternal X-linked genes that are subject to imprinted X-inactivation (Harris et al., 2019; Inoue et al., 2018). Furthermore, we demonstrated in female mouse trophoblast stem cells (TSCs), which maintain imprinted X-inactivation of the paternal-X, that the essential PRC2 protein EED is not required to maintain silencing of most paternal X-linked genes (Kalantry *et al.*, 2006; Maclary *et al.*, 2017). The absence of EED

destabilizes PRC2 and results in the degradation of the PRC2 methyltransferase enzyme EZH2 and loss of H3K27me3 (Montgomery et al., 2005).

Although zygotic PRC2 appears to be dispensable for chromosome-wide silencing of paternal X-linked genes, we and others have shown that oocyte-generated PRC2 and H3K27me3 are required for imprinted X-inactivation. Ablation of *Eed* in the oocyte results in loss of H3K27me3 and ectopic induction of *Xist* from the maternal X chromosome in the early embryo (Inoue *et al.*, 2017; Xie *et al.*, 2016). The *Xist* locus on the maternal-X is normally decorated with a broad domain of H3K27me3 that is established during mouse oocyte growth and persists through preimplantation embryonic development (Harris *et al*., 2019; Inoue *et al.*, 2017; Xie *et al.*, 2016). Ablation of *Eed* in the oocyte (maternal *Eed*; $Eed^{m-/-}$) results in undetectable H3K27me3 and ectopic expression of *Xist* from the maternal-X in the early embryo. In $Eed^{m-/-}$ early preimplantation female embryos, *Xist* is thus initially expressed from both the maternal and paternal X chromosomes. As the mutant female embryos develop, this pattern of biallelic *Xist* expression is resolved into expression of one or the other *Xist* allele resulting in a random pattern of inactivation in cells that would normally exhibit imprinted X-inactivation of the paternal-X (Harris *et al.*, 2019; Inoue et al., 2018). This switching of imprinted to random X-inactivation also characterizes embryos that lack maternal PRC1 proteins PCGF1/6 (Mei et al., 2021). PRC1 complexes catalyze histone H2AK119 mono-ubiquitination (H2AK119ub1) (Wang et al., 2004). These results together suggest an interplay between PRC2 and PRC1 in ensuring that the maternal X chromosome resists *Xist* expression and X-inactivation in the early mouse embryo. Thus, although the paternal X chromosome undergoes imprinted inactivation, the imprint resides on the maternal-X and is comprised of PRC-catalyzed histone modifications deposited in the oocyte on the maternal *Xist* locus.

Whereas EED is an essential non-catalytic core component of PRC2, EZH2 and its homolog EZH1 are the PRC2 methyltransferase enzymes that catalyze H3K27me3 (Cao *et al.*, 2002; Czermin *et al.*, 2002; Margueron et al., 2008; Shen et al., 2008a). Here, we test the requirement of EZH2/1 in the regulation of imprinted X-inactivation by generating embryos devoid of oocyte-derived (maternal) and zygotically made EED, EZH2 or EZH1, or both EZH2 and EZH1 (EZH2/1). We find that embryos lacking maternal and zygotic *Ezh2* (*Ezh2*$^{mz-/-}$) and *Ezh1* (*Ezh1*$^{mz-/-}$) display nearly normal imprinted X-inactivation of paternal X-linked genes, similar to wild-type embryos. *Eed* (*Eed*$^{mz-/-}$) and *Ezh2/1* (*Ezh2/1*$^{mz-/-}$), on the other hand, both display randomization of X-inactivation in the early female mouse embryos, suggesting that EZH2 and EZH1 can compensate for each other in the catalysis of H3K27me3 during oogenesis. Unexpectedly, however, whereas *Eed*$^{mz-/-}$ express maternal and paternal alleles of X-linked genes at nearly equal levels, consistent with randomization of X-inactivation, *Ezh2/1*$^{mz-/-}$ preferentially express maternal alleles of X-linked genes. These data suggest that EED functions as a part of PRC2, but also independently of EZH2/1 and H3K27me3 in silencing genes on the paternal X chromosome during imprinted X-inactivation.

Imprinted X-inactivation Initiation in *Ezh2*$^{m-/-}$ and *Ezh2*$^{mz-/-}$ Embryos

To test a requirement for EZH2 in imprinted X-inactivation (Kalantry et al., 2009), we generated and characterized *Ezh2*$^{m-/-}$ female mouse embryonic day (E) E3.5 blastocyst-stage preimplantation embryos, which normally undergo imprinted X-inactivation, through *Cre-lox* mediated deletion of *Ezh2* in the oocyte (Lewandoski *et al.*, 1997). We crossed females harboring *lox*p flanked (floxed) *Ezh2* alleles (*Ezh2*$^{fl/fl}$) and a *Zp3-Cre* transgene that catalyzes the deletion of floxed sequences in the growing oocyte with *Ezh2*$^{fl/fl}$ males (Figure 3.1A;

Supplemental Figure 3.1A; Materials and Methods). This cross yielded embryos lacking maternally-generated EZH2, but still harboring a zygotic *Ezh2* allele.

Wild-type female blastocysts display enrichment of H3K27me3 on the inactive-X that is detected cytologically and monoallelic expression of most X-linked genes from the active maternal-X (Chapter 2). In *Ezh2*$^{fl/fl}$ and *Ezh2*$^{m-/-}$ female embryos, we first assayed H3K27me3 enrichment on the inactive-X, *Xist* RNA coating, which marks the inactive-X, and nascent transcripts of X-linked genes subject to inactivation in individual cells by immunofluorescence (IF) combined with RNA fluorescence *in situ* hybridization (FISH) (Figure 3.1B-C; Supplemental Figure 3.1B). We observed monoallelic RNA FISH signals from *Atrx*, *Gla*, *Pdha1*, *Utx*, *G6Pdx*, *Lamp2*, and *Rnf12* in both *Ezh2*$^{fl/fl}$ and *Ezh2*$^{m-/-}$ blastocysts, indicating silencing of these X-linked genes on the inactive-X (Figure 3.1B). *Utx* RNA was biallelically detected in both genotypes, consistent with *Utx* escaping X-inactivation and being expressed from both the active- and the inactive-X. As expected, cells of *Ezh2*$^{fl/fl}$ embryos also displayed H3K27me3 enrichment coincident with *Xist* RNA coating. *Ezh2*$^{m-/-}$ embryos, by contrast, exhibited *Xist* RNA coating but lacked robust H3K27me3 enrichment, consistent with the EZH2 catalyzing H3K27me3 (Figure 3.1C; Supplementary Figure 3.1B). These results suggest that *Ezh2*$^{m-/-}$ female mouse blastocysts initiate imprinted X-chromosome inactivation normally.

We next generated E3.5 blastocysts lacking both maternal and zygotic *Ezh2* by crossing *Ezh2*$^{fl/fl}$; *Zp3-Cre* females with *Ezh2*$^{fl/fl}$ males harboring a *Prm-Cre* transgene that is expressed during spermatogenesis (O'Gorman *et al*., 1997) (Figure 3.2A-B; Materials and Methods). The *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ embryos we generated were hybrid, harboring a maternal X chromosome derived from the *Mus musculus* 129/S1 mouse strain and a paternal-X from the divergent *Mus molossinus* JF1/Ms strain (Materials and Methods). Hybrid embryos are useful for assaying

allele-specific gene expression because the maternal- and paternal-X each harbor many unique strain-specific SNPs.

We assayed X-linked gene expression in *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ hybrid E3.5 female blastocysts by allele-specific RNA Sequencing (RNA-Seq) (Figure 3.2C-E; Supplemental Figure 3.2). In the RNA-Seq analysis, we exploited single nucleotide polymorphisms (SNPs) to assign RNA-Seq reads to either the maternal or paternal X-chromosome in the hybrid embryos (Materials and Methods) (Cloutier *et al.*, 2018; Maclary *et al.*, 2017). In both *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ embryos, *Xist* RNA was expressed predominantly from the paternal X chromosome (Figure 3.2D). *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ female embryos also displayed similar overall patterns of paternal X-linked gene silencing (Figure 3.2E; Supplemental Figure 3.2). Of note, a small subset of X-linked genes in the *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ female mouse embryos displayed expression from both the maternal and paternal alleles, which is likely due to a combination of genes that escape X-inactivation and genes that are in the process of being inactivated but have not become fully silenced on the paternal-X. We validated the RNA-Seq analyses by performing Pyrosequencing, a low-throughput technique that can accurately capture allelic expression ratios of individual genes (Cloutier *et al.*, 2018; Gayen *et al.*, 2015). Pyrosequencing of the X-linked genes *Atrx*, *Rnf12*, *G6pdx*, and *Pdha1* showed no significant differences in the ratio of allelic expression between *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ blastocysts (Figure 3.2F). Taken together, these data indicate that female mouse embryos lacking maternal and zygotic EZH2 appropriately silence most paternal X-linked genes at the blastocyst stage. Unexpectedly, though, this result does not recapitulate the paternal-X gene silencing patterns observed in *Eed*$^{m-/-}$ and *Eed*$^{mz-/-}$ female blastocysts, which undergo random X-inactivation of either the maternal or paternal-X (Chapter 2).

Imprinted X-inactivation Initiation in *Ezh1*$^{mz-/-}$ Embryos

EZH1 is a mammalian specific homologue of EZH2. Whereas EZH2 is thought to be the predominant H3K27me3 catalyst, EZH1 can catalyze some H3K27me3, as *Ezh2*$^{-/-}$ mouse embryonic stem cells (ESCs) display residual H3K27me3 deposition that is lost in *Ezh2*$^{-/-}$;*Ezh1*$^{-/-}$ ESCs (Hojfeldt et al., 2018; Shen *et al.*, 2008a). These results imply that EZH1 can catalyze H3K27me3. Given that *Ezh2*$^{mz-/-}$ female mouse embryos initiate imprinted X-inactivation properly, we hypothesized that maternal and/or zygotic EZH1 may contribute to the silencing of paternal X-linked genes in female mouse blastocysts normally. To test this hypothesis, we generated and characterized hybrid *Ezh2*$^{fl/fl}$;*Ezh1*$^{mz-/-}$ female mouse blastocysts by intercrossing *Ezh2*$^{fl/fl}$;*Ezh1*$^{-/-}$ mice, which are viable and fertile (Figure 3.3; Materials and Methods). Like embryos lacking maternal EZH2, *Ezh2*$^{fl/fl}$;*Ezh1*$^{mz-/-}$ embryos displayed no significant defect in *Xist* expression from the paternal-X (Figure 3.4A; Supplemental Figure 3.2). Furthermore, we observed no significant difference in the allelic expression ratio of X-linked genes in *Ezh2*$^{fl/fl}$;*Ezh1*$^{mz-/-}$ embryos compared to *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ female mouse embryos by allele-specific RNA-Seq (Figure 3.4B-D; Supplemental Figure 3.2). However, like in *Ezh2*$^{mz-/-}$ female embryos, a small subset of paternal X-linked genes exhibited increased biallelism in *Ezh2*$^{fl/fl}$; *Ezh1*$^{mz-/-}$ compared to *Ezh2*$^{fl/fl}$ embryos. We again validated these data by Pyrosequencing of X-linked genes *Rnf12*, *Atrx*, *G6pdx*, and *Pdha1* and found no significant differences in the maternal:paternal expression ratio between *Ezh2*$^{fl/fl}$ and *Ezh2*$^{fl/fl}$; *Ezh1*$^{mz-/-}$ embryos for all four genes (Figure 3.4E). Taken together, these data suggest that maternal and zygotic EZH1 are largely dispensable for the silencing of most paternal X-linked genes in female mice.

Random X-inactivation in *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ Embryos

Given our finding that maternal and zygotic EZH2 and EZH1 are individually dispensable for the initiation of imprinted X-inactivation, we hypothesized that these factors

exhibit functional redundancy in the oocyte in catalyzing H3K27me3 at the maternal *Xist* locus and thereby prevent inactivation of the maternal-X in the embryo. To test this hypothesis, we generated and characterized E3.5 blastocysts lacking both maternal and zygotic EZH2 and EZH1 (*Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$) (Figure 3.5; Materials and Methods). We first compared allelic *Xist* RNA expression and coating by allele-specific RNA FISH in *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$, *Eed*$^{mz-/-}$, and *Ezh2*$^{fl/fl}$ E3.5 female and male embryos (Figure 3.5A-B). Allele-specific *Xist* RNA FISH allowed us to detect *Xist* RNA expression from the maternal vs. paternal X-chromosome in individual cells in the embryos. Female *Ezh2*$^{fl/fl}$ blastocysts expressed *Xist* only from the paternal-X, as expected. *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ and *Eed*$^{mz-/-}$ female blastocysts, by contrast, exhibited *Xist* RNA expression from either the maternal or paternal-X, from both Xs, or from neither X (Figure 3.5A) (Chapter 2). This pattern of *Xist* expression is also observed in *Eed*$^{m-/-}$ embryos, in which it is ultimately resolved into random inactivation of either the maternal or the paternal X-chromosome (Chapter 2) (Inoue et al., 2018). Moreover, like male *Eed*$^{m-/-}$ embryos *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ and *Eed*$^{mz-/-}$ male blastocysts exhibited ectopic expression of maternal *Xist*, whereas *Ezh2*$^{fl/fl}$ male embryos displayed no *Xist* RNA coating of the maternal-X (Figure 3.5B).

Xist RNA coating of either the maternal- or paternal-X in *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ and *Eed*$^{mz-/-}$ female embryos suggests that, like *Eed*$^{m-/-}$ female embryos, these embryos undergo random instead of imprinted X-inactivation. Characterization of X-linked gene expression in *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ female embryos by allele-specific RNA-Seq determined that *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ female embryos display unbiased allelic expression of many, but not all, X-linked genes, consistent with random X-inactivation (Figure 3.6A-B; Supplemental Figure 3.2). Pyrosequencing of *Rnf12*, *Atrx*, *G6pdx*, and *Pdha1* expression also showed a significant relative increase in the expression

of paternal X-linked genes in *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ female embryos compared to *Ezh2*$^{fl/fl}$ embryos (Figure 3.6C).

Differential Roles for EZH2/1 and EED in X-linked Gene Silencing

Although allele-specific RNA FISH analysis indicated that both *Eed*$^{m-/-}$ and *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ female embryos equally undergo random X-inactivation, the allelic expression profile of X-linked genes in *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ female blastocysts did not fully recapitulate the pattern observed in the *Eed*$^{mz-/-}$ female blastocysts (Chapter 2). Importantly, comparison of expression of pluripotency factors *Nanog*, *Gata6*, and *Sox2* indicated that all the embryos I characterized were indeed of the blastocyst-stage (Supplemental Figure 3.3).

I next compared the allele-specific RNA-Seq data generated from *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ and *Eed*$^{mz-/-}$ female blastocysts to interrogate differences in paternal X-linked gene silencing (Figure 3.7A-C; Supplemental Figure 3.2). This analysis indicated that over half (66%) of the X-linked genes in *Eed*$^{mz-/-}$ embryos exhibited an unbiased allelic expression ratio in (between 37.25 – 62.5% paternal-X expression) (Figure 3.7C; Table 3.1). *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos exhibited unbiased expression of only 15% of X-linked genes, as most X-linked genes were expressed with a significant maternal bias (Figure 3.7C; Table 3.1). Furthermore, only four genes exhibited unbiased expression in *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos that did not undergo unbiased expression in *Eed*$^{mz-/-}$ embryos (Figure 3.7C; Table 3.1). Taken together, these data suggest an allele-specific gene silencing role for maternal EED.

Like *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ and *Eed*$^{mz-/-}$ female blastocysts, the *Eed*$^{mz-/-}$ and *Eed*$^{m-/-}$ embryos characterized in Chapter 2 displayed significant differences in allele-specific X-linked gene silencing. I thus compared the number of X-linked genes exhibiting unbiased allelic expression in *Eed*$^{mz-/-}$ and *Eed*$^{m-/-}$ embryos. As stated above, over half (66%) of the X-linked genes in

Eed$^{mz-/-}$ embryos exhibited an unbiased allelic expression ratio in (between 37.25 – 62.5% paternal-X expression) (Figure 3.7C-D; Table 3.1). However, *Eed*$^{m-/-}$ embryos exhibited unbiased expression of only 26% of X-linked genes, as many X-linked genes were expressed with a significant maternal bias (Figure 3.7D; Table 3.1). Comparison of *Eed*$^{m-/-}$ and *Ezh2*$^{mz-/-}$;*Ezh1*$^{mz-/-}$ embryos showed that most genes that exhibit unbiased expression in these two genotypes are the same (Figure 3.7E; Table 3.1). These results suggest that maternally generated EED plays a role in silencing genes on the paternal-X in early mouse embryos, likely apart from gene silencing mediated by PRC2.

A Potential Role for EED-mediated H3K27 Deacetylation in X-linked Gene Silencing

Previous work demonstrated that EED, but not EZH1/2, interacts with histone deacetylases (HDACs) and that this interaction may enhance histone H3 lysine 27 (H3K27) deacetylation activity (Ai et al., 2017; Cao et al., 2014; Kuzmichev et al., 2005; van der Vlag and Otte, 1999). Acetylation of the lysine residues of core histones is a well-characterized chromatin modification that is associated with transcriptional activation (Hebbes et al., 1988). Histone deacetylation, conversely, is associated with transcriptional repression. Furthermore, histone deacetylation has been shown to occur on the inactive-X at the onset of X-inactivation (Keohane et al., 1998). We thus hypothesized that EED may interact with HDACs in early female mouse embryos to deacetylate X-linked genes and control transcription in an allele-specific manner. To examine this hypothesis, we profiled H3K27ac by allele-specific ChIP-Seq in wild-type 2-cell and 8-cell preimplantation female mouse embryos and compared the H3K27ac profiles of X-linked genes that appeared to be sensitive to maternal and zygotic EED vs. EZH1/2 loss. Interestingly, analysis of these 2-cell embryos indicated that genes exhibiting unbiased expression in *Eed*$^{mz-/-}$ but not *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos harbor a higher level of H3K27ac on

the paternal-X than genes that exhibit unbiased expression in both genotypes (Figure 3.8A). However, this pattern was not recapitulated in 8-cell embryos (Figure 3.8B), suggesting that removal of some but not all acetylation from the paternal-X may occur between the 2- and 8-cell stages.

In addition to examining H3K27ac marking of the paternal-X, we mined published allele-specific ChIP-Seq data for another histone modification – histone H3 lysine 4 trimethylation (H3K4me3) – which closely correlates with H3K27ac marking (Zubek et al., 2016). These data were generated from wild-type 8-cell female preimplantation mouse embryos (Dahl et al., 2016). We again analyzed the H3K4me3 profiles of X-linked genes exhibiting unbiased expression in *Eed*$^{mz-/-}$ but not *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos vs. genes exhibiting unbiased expression in both genotypes. These data indicate that genes exhibiting maternal-X expression are more likely to harbor H3K4me3 than those that exhibit unbiased allelic expression in *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos (Figure 3.8C). Interestingly, these data do not recapitulate the paternal-X H3K27ac data described above, which may be due to H3K4me3 remaining on the paternal-X after H3K27ac removal or to technical differences between these analyses. These data further support a role for maternal EED in the deacetylation of paternal X-linked genes and the regulation of X-linked transcription in an allele-specific manner, although this correlation remains to be mechanistically examined.

Another histone modification that has previously been suggested to be involved in imprinting on the paternal-X is histone H2A Lysine 119 mono-ubiquitination (H2AK119ub1), which is deposited by the Polycomb Repressive Complex 1 (PRC1) and is associated with gene silencing (Mei *et al.*, 2021; Simon and Kingston, 2009; Wang *et al.*, 2004). We thus mined and examined the allele-specific H2AK119ub1 profiles of paternal X-linked genes in 8-cell mouse

embryos exhibiting unbiased expression in *Eed*$^{mz-/-}$ but not *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos and found no significant difference in H2AK119ub1on the paternal-X in these genes (Figure 3.8D) (Chen et al., 2021). These data suggest that the paternal X-linked silencing differences observed between *Eed*$^{mz-/-}$ and *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ are likely not due to paternal-X H2AK119ub1 deposition.

Discussion

Chapter 2 demonstrated the oocyte-derived PRC2 protein EED is required to ensure that the maternal X chromosome does not become inactivated in the early mouse embryo. PRC2-catalyzed H3K27me3 is enriched at the maternal *Xist* locus in oocytes, and this H3K27me3 remains enriched at and silences maternal *Xist* in the preimplantation female mouse embryo (Inoue *et al.*, 2017; Zhang et al., 2016). When the maternal *Xist* allele is silenced, paternal *Xist* is expressed exclusively. In the absence of oocyte-derived EED and maternal H3K27me3 *Xist* marking, early female mouse embryos undergo random X-inactivation instead of imprinted X-inactivation.

EZH1/2 in gene silencing and imprinted X-inactivation initiation

EZH2 is the primary histone methyltransferase component of PRC2 (Di Croce and Helin, 2013; Margueron and Reinberg, 2011; Schuettengruber et al., 2007; Zhang *et al.*, 2015). Previous work showed that EZH2 homolog EZH1 can catalyze a low level of H3K27me3 in *Ezh2*$^{-/-}$ mouse embryonic stem cells (ESCs) (Hojfeldt *et al.*, 2018; Shen *et al.*, 2008a). This H3K27me3 catalysis by EZH1 suggests that EZH1 can compensate somewhat for the loss of EZH2 function and catalyze a low level of H3K27me3 in early mouse embryos. Our data further support this finding, as female blastocysts lacking either EZH1 or EZH2 undergo imprinted X-inactivation resulting in the preferential silencing of paternal X-linked genes compared to embryos lacking both maternal EZH1 and EZH2, which undergo random instead of imprinted X-inactivation.

A PRC2-independent role for EED in X-linked gene silencing

The allele-specific RNA FISH analysis in this chapter and in Chapter 2 showed that *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ and *Eed*$^{m-/-}$ female embryos, like *Eed*$^{mz-/-}$ female embryos, undergo random *Xist* induction and thus random X-inactivation. Despite random X-inactivation, I observed differences in the expression ratios of maternal:paternal X-linked genes between *Eed*$^{mz-/-}$ and *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ female embryos. Specifically, *Eed*$^{mz-/-}$ embryos displayed unbiased expression of maternal and paternal X-linked genes, whereas *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ female embryos display a slight maternal-X gene expression bias. *Eed*$^{m-/-}$ embryos also display a similar difference in the expression ratios of maternal:paternal X-linked genes compared to *Eed*$^{mz-/-}$ female mouse embryos. The preferential silencing of some paternal-X genes in *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ and *Eed*$^{m-/-}$ embryos is likely due to one of two possibilities: 1) expression of zygotic EED in the early embryo functions independently of EZH2/1, and thus of PRC2 and PRC2-catalyzed H3K27me3, to preferentially silence a subset of paternal X-linked genes independently of *Xist* RNA coating of the paternal X chromosome; or, 2) zygotic EED functions to silence genes on *Xist* RNA-coated maternal and paternal X chromosomes but does so more efficiently on the paternal-X as compared to the maternal-X. The maternal-X is known to harbor a more complex set of chromatin modifications compared to the paternal X chromosome, which is inherited without many histones and any inherited histones appear to be removed shortly after fertilization (Wu et al., 2016). The second possibility predicts a failure of silencing of X-linked genes on the Xist RNA-coated X chromosomes in the *Eed*$^{mz-/-}$ female embryos due to the loss of zygotic EED-mediated gene silencing. In these embryos, some dosage compensation of X-linked genes may nevertheless occur due to intrinsic transcriptional balancing mechanisms that

ensure proper stoichiometries of gene products (Birchler et al., 2007). Below, I discuss my rationale for each of these hypotheses and existing data that support each possibility.

In support of the first hypothesis above that zygotic EED preferentially silences some paternal X-linked genes independently of *Xist* RNA coating, one group found that paternal X-linked gene silencing is epigenetically programmed during spermatogenesis (Sun *et al*., 2015). In a study investigating how *Xist* is imprinted, the authors created transgenic mice carrying a 200 kb sequence from the X-inactivation Center and observed the consequences of transgene zygosity on *Xist* imprinting in the male germ line. In this study, the preferential silencing of paternal X-linked genes in the early embryo was proposed to be due to the paucity of pairing of the X chromosome to its homolog in the male germline. However, the authors present no mechanistic data in support of epigenetic marking or imprinting of the paternal-X in the male germline nor the perdurance of such mark(s) in the early embryo, which undergoes a dramatic chromatin reprogramming by the morula stage of embryogenesis (Wu et al., 2016). In fact, the paternal-X has been shown to be devoid of most epigenetic marks in early embryogenesis (Xu et al., 2021). Interestingly, we did find that the histone modification H3K27ac is enriched on the paternal-X in two cell female mouse embryos, but this differential marking is resolved by the 8-cell stage. It is possible that EED-mediated histone deacetylation in the early embryo is a mechanism underlying the preferential silencing of paternal X-linked genes. The preferential histone deacetylation of genes on the paternal-X but not the maternal-X maybe due to the maternal-X harboring a more complex set of chromatin modifications that are deposited during oogenesis compared to that during spermatogenesis, during which most histones are replaced by protamines (Ward et al., 1991; Wykes et al., 2003). Furthermore, imprinting of the paternal-X argues against findings by our group and others that the maternal-X carries the imprint (e.g., the

PRC2-catalyzed histone H3K27me3 chromatin modification of the *Xist* locus) that prevents the maternal-X from being inactivated in the early embryo (Chapter 2; Harris *et al*., 2019; Inoue *et al*., 2018). The paternal-X is devoid of this imprint and thus is subject to inactivation.

The second hypothesis proposes that zygotic EED participates with HDAC proteins that help remove acetyl marks on chromatin on the *Xist* RNA-coated X chromosomes, contributing to silencing of genes on the *Xist* RNA coated X chromosomes in wild-type female embryos (Figure 3.8E). If this histone deacetylation does not occur, then these X-linked genes may fail to undergo complete silencing as is normally observed for most genes on the inactive-X. Acetylation of the lysine residues of core histones is a well-characterized chromatin modification that is associated with transcriptional activation (Hebbes et al., 1988). Histone deacetylation, by contrast, is associated with transcriptional repression and has been found to be one of the first factors to be enriched on the inactive-X (Keohane et al., 1998; Zylicz et al., 2019). Mass spectrometry data from van der Vlag & Otte, 1999, Kuzmichez et al., 2005, and Cao et al., 2014 showed that EED interacts with essential histone deacetylases HDAC1 and HDAC2. *Hdac1* deletion in mice leads to embryonic lethality (Lagger et al., 2002) whereas *Hdac2* deficiency results in perinatal lethality stemming from heart defects (Montgomery et al., 2007). Furthermore, van der Vlag & Otte, 1999 indicate that EED, but not EZH2, interacts with HDAC2, consistent with the notion that EED may have a PRC2- and H3K27me3-independent role in gene silencing through histone deacetylation. In support of this idea, 2017 study by Ai *et al.* showed that EED interacts with histone deacetylases to enhance their catalytic activity.

When coupled with findings that EED can interact with HDACs, a potential explanation for the differences in silencing of genes on the paternal-X we see in *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ and *Eed*$^{m-/-}$ vs. *Eed*$^{mz-/-}$ embryos may lie in a role for EED-mediated histone deacetylation in silencing X-

linked genes on the *Xist* RNA-coated X chromosomes. *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ and *Eed*$^{m-/-}$ but not *Eed*$^{mz-/-}$ embryos express zygotic EED, which may impart gene silencing through a partnership with HDACs. Notably, this hypothesis also explains the mild gene silencing defect we observed in Chapter 2 in *Eed*$^{-/-}$ embryos that harbor maternal EED but lack zygotically produced EED. *Eed*$^{mz-/-}$ embryos, which lack both maternal and zygotic EED and which display ~1:1 maternal:paternal X-linked gene expression, may fail to undergo proper X-linked gene silencing on the *Xist* RNA coated X chromosomes due to deficient histone deacetylation. A failure of X-inactivation is expected to result in 2X gene expression compared to female samples that have undergone X-inactivation. *Eed*$^{mz-/-}$ embryos, though, do not display a marked increase in X-linked gene expression. The lack of a 2X level of X-linked gene expression in *Eed*$^{mz-/-}$ embryos, however, may be due to intrinsic dosage compensation mechanisms that protect against dramatic increases in gene expression despite increase or decrease in copy number of genes (Birchler et al., 2007). The idea of 'intrinsic dosage compensation' upon the loss of X-inactivation is supported by experiments in which the *Xist* gene is deleted from both X chromosomes and thus neither X chromosome can produce any Xist RNA (Kalantry Lab, unpublished data; Yang et al., 2016).

The data presented here do not differentiate between a direct or indirect interaction between zygotically generated EED and HDACs. Furthermore, these data do not provide direct evidence that EED interacts with HDACs to silence X-linked genes in female mouse preimplantation embryos, nor do our data distinguish which HDACs may interact with EED in these embryos. Previously published work has demonstrated that the interaction of EED with various HDACs is cell type specific, so future analyses must be performed in embryo-derived cells to test for EED-HDAC interactions at this stage of development. It is of course possible

that EED does not interact with HDACs in mouse embryos but has a different PRC2-independent role. Future work will mechanistically dissect this PRC2-independent function of EED in gene silencing, in addition to any tissue-specific roles it may have.

Conclusion and Future Directions

This study demonstrates that PRC2 components EED and EZH1/2 play different roles in X-linked gene silencing in female mouse embryos. More broadly, this work suggests that EED can function apart from the canonical PRC2 complex to effect gene silencing. A key future experiment to expand upon this finding is a comparative study of the EED and EZH1/2 interactomes in the genotypes described in this chapter via immunoprecipitation-mass spectrometry (IP-MS). Although individual mouse blastocyst-stage embryos will not yield enough protein for such an experiment, the interactomes of embryo-derived imprinted cell lines such as TSCs can be investigated. An IP-MS analysis will determine whether EED interacts with HDACs independently of PRC2, as suggested in this study. Investigation of the EED and EZH1/2 interactomes in wild-type cells will also inform whether PRC2 variant(s) may be involved in imprinted X-inactivation initiation, as it will allow for the identification of variant-specific accessory proteins that interact with EED and/or EZH1/2 in cells undergoing imprinted X-inactivation. A growing body of work has revealed that specialized variants of PRC2 – namely, PRC2.1 and PRC2.2 – coordinate to deposit H3K27me3 in different contexts (Hauri et al., 2016; Healy et al., 2019). Thus, one or both variant complexes may play a role in the silencing of paternal X-linked genes.

Although we have found that some *Eed*$^{m-/-}$ female mice can survive to term and are fertile, not all *Eed*$^{m-/-}$ embryos are viable. A possible explanation for why some female *Eed*$^{m-/-}$ embryos survive whereas others do not is the preferential inactivation of the paternal X-

chromosome in some but not all female embryos. Although mouse embryos lacking paternal *Xist* and maternal *Tsix* can undergo inactivation of the maternal-X, very few of these embryos are viable (Sado et al., 2001). Thus, inactivation of the paternal-X may confer a survival advantage in female mouse embryos. Future work will investigate whether preferential inactivation of the paternal X chromosome occurs in female mice lacking maternally-generated EED. Investigation of this question can be accomplished by allele-specific single-cell RNA-Seq and allele-specific RNA FISH of mouse embryos and allele-specific RNA Seq of female adult tissues to examine whether maternal X-linked gene expression bias exists.

Materials and Methods

Mice

Mice harboring a conditional mutation in *Eed* were generated by the University of Michigan Transgenic Animal Model Core using $Eed^{tm1a(EUCOMM)Wtsi}$ targeted ES cells (EUCOMM). ES cells were injected into blastocysts and implanted into pseudo-pregnant females. Mice with high percentages of chimerism were bred and assessed for germline transmission. To generate homozygous *Eed* mutant mice harboring polymorphic X-chromosomes, first, male and female mice on a B6 *Mus musculus* background carrying the conditional mutant allele for *Eed* were intercrossed ($Eed^{fl/+}$ x $Eed^{fl/+}$) to achieve homozygosity. To obtain mice conditionally mutant for *Eed* and on the JF1 *Mus molossinus* divergent background, we bred $Eed^{fl/fl}$ males (B6 *Mus musculus* background) to *WT* JF1 *Mus molossinus* females. This gave us F1 hybrid $Eed^{fl/+}$ males that possessed an X-chromosome from the JF1 *Mus molossinus* background (X^{JF1}/Y). Such males were backcrossed to *WT* JF1 *Mus molossinus* females to derive $Eed^{fl/+}$ females that were a mix of B6 *Mus musculus* and JF1 *Mus molossinus* and harbored two X-chromosomes from the JF1 *Mus molossinus* background (X^{JF1}/X^{JF1}). $Eed^{fl/+}$;X^{JF1}/X^{JF1} females were bred against $Eed^{fl/+}$;*X*

$^{JF1}/Y$ males to derive *Eed*$^{fl/fl}$; X^{JF1}/Y males. To obtain the female embryos used in our experiments (*Eed*$^{m-/-;z-/-}$;X^{Lab}/X^{JF1} and *Eed*$^{fl/fl}$;X^{Lab}/X^{JF1}), we crossed *Eed*$^{fl/fl}$ females with or without the *Zp3-Cre* transgene on the B6 *Mus musculus* background with an *Eed*$^{fl/fl}$ with or without the *Prm-Cre* male that was a mix of B6 *Mus musculus* and JF1 *Mus molossinus* but possessed an X-chromosome from the JF1 *Mus molossinus* background (X^{JF1}/Y). The JF1/Ms strain has been described previously.

Ezh2 mice were maintained on a 129 background and gifted from Alexander Tarakhovsky. Mice were crossed in a similar manner as the *Eed* mice above for deriving the *Ezh2*$^{m-/-;z-/-}$;X^{Lab}/X^{JF1} and *Ezh2*$^{fl/fl}$;X^{Lab}/X^{JF1} blastocysts, using females with and without the *Zp3-Cre* transgene and males with and without the *Stra8-Cre* transgene.

Ezh1 mice were gifted from Alexander Tarakhovsky, originally bred by Dønal O'Carroll in Thomas Jenuwein's laboratory, maintained on a BL/6 background.

Ezh2$^{fl/fl}$ and *Ezh1*$^{-/-}$ mice were intercrossed and bred to generate our *Ezh2*$^{fl/fl}$; *Ezh1* $^{-/-}$;$X^{Lab/}X^{JF1}$ mice and our *Ezh*$^{m-/-;z-/-}$;*Ezh1*$^{-/-}$;$X^{Lab/}X^{JF1}$ *Ezh2*$^{fl/fl}$;*Ezh1*$^{-/-}$;$X^{Lab/}X^{JF1}$ blastocysts in a similar manner as described for the generation of our *Eed* and *Ezh2* mice/embryos. We used females with and without the *Zp3-Cre* transgene and males with and without the *Prm-Cre* transgene.

Mouse Embryo Dissection and Processing

Embryonic day (E) 3.5 embryos were isolated essentially as described (Maclary *et al.*, 2014). Embryos were flushed from the uterine limbs in 1X PBS (Invitrogen, #14200) containing 6 mg/ml BSA (Invitrogen, #15260037).

The zona pellucida surrounding embryos was removed through incubation in cold acidic Tyrode's solution (Sigma, #T1788), followed by neutralization through several transfers of cold M2 medium (Sigma, #M7167).

Isolated E3.5 embryos were either lysed for RNA isolation or plated onto 0.2% gelatin- (Sigma, #G2500) and/or 0.01% Poly-L-Lysine (PLL, Sigma, # P4707)-coated glass coverslips (22mm X 22mm, Thermo Fisher Scientific, #12548B) in 0.25X PBS for immunofluorescence (IF) coupled with RNA in situ hybridization (FISH). 2–16 cell embryos were plated on coverslips coated in 0.01% Poly-L-Lysine for IF. E3.5 or 4–16 cell embryos were plated on coverslips coated with 1X Denhardt's (Sigma, #D9905) solution for allele-specific RNA FISH. For plated embryos, excess solution was aspirated, and coverslips were air-dried for approximately 15–30 min. After drying, embryos were permeabilized and fixed in 50 µL solution of either 0.05% or 0.1% Tergitol (Sigma, #NP407) with 1% paraformaldehyde (Electron Microscopy Sciences, #15710) in 1X PBS for 5 min, followed by 1% paraformaldehyde in 1X PBS for an additional 5 min. Excess solution was gently tapped off onto paper towels, and coverslips were rinsed 3X with 70% ethanol and stored in 70% ethanol at −20°C prior to IF and/or RNA FISH.

PCR Genotyping

For embryo DNA isolation, embryos were isolated as described above. Individual blastocysts were lysed in 15 µL buffer composed of 50 mM KCl, 10 mM Tris-Cl (pH 8.3), 2.5 mM $MgCl_2$, 0.1 mg/mL gelatin, 0.45% NP-40, 0.45% Tween-20, and 0.4 mg/mL Proteinase K (Fisher, #BP1700). Embryos in lysis buffer were incubated at 50°C overnight, then stored at 4°C until use. Genomic PCR used 1–3 µL lysate per sample. PCR Reactions to detect *Ezh1* and *Ezh2* were carried out in ChromaTaq buffer (Denville Scientific) with 2.5 mM $MgCl_2$ added. XX vs. XY

sexing PCR reactions were carried out in Klentherm buffer (670 mM Tris pH 9.1, 160 mM (NH4)SO4, 35 mM $MgCl_2$,15mg/ml BSA). Both used RadiantTaq DNA polymerase (Alkali Scientific, #C109). Primer sequences are listed in Table 3.2.

Liveborn animals from the cross of *Ezh2/1*$^{fl/fl}$; *Zp3-Cre* females by WT males were genotyped for *Ezh1* and/or *Ezh2* to confirm deletion of the floxed allele. Ear punches were taken after weaning and lysed in 50 µL of lysis buffer (above). Ear punches were incubated at 50°C overnight, then stored at 4°C until use. 1 µL of DNA lysate was used per reaction. *Ezh1* and *Ezh2* PCRs were carried out as described above.

Quantification of allele-specific RNA expression by Pyrosequencing

Allele-specific expression was quantified using the Qiagen PyroMark sequencing platform, as previously described (Gayen *et al.*, 2015). Briefly, the amplicons containing SNPs were designed using the PyroMark Assay Design software. cDNAs were synthesized using Invitrogen SuperScript III One-Step RT-PCR System (Invitrogen, #12574–026). Following the PCR reaction, 5 µL of the 25 µL reaction was run on a 3% agarose gel to assess the efficacy of amplification. The samples were then prepared for Pyrosequencing according to the standard recommendations for use with the PyroMark Q96 ID sequencer. All amplicons spanned intron(s), thus permitting discrimination of RNA vs. any contaminating genomic DNA amplification due to size differences. Control reactions lacking reverse transcriptase for each sample were also performed to rule out genomic DNA contamination. E3.5 embryos of similar sizes for all genotypes were used in the Pyrosequencing assays. Pyrosequencing primers are listed in Table 3.3.

Immunofluorescence (IF)

Embryos mounted on gelatin-, PLL-, and/or PLL/gelatin-coated glass coverslips were washed 3 times in 1X PBS for 3 min each while shaking. Coverslips were then incubated in blocking buffer consisting of 0.5 mg/mL BSA (New England Biolabs, #B9001S), 50 µg/mL yeast tRNA (Invitrogen, #15401–029), 80 units/mL RNAseOUT (Invitrogen, #10777–019), and 0.2% Tween 20 (Fisher, #BP337-100) in 1X PBS in a humid chamber for 30 min at 37°C. The samples were next incubated with primary antibody diluted in blocking buffer for 45 min −2 hr in the humid chamber at 37°C. The samples were then washed 3 times in 1X PBS/0.2% Tween 20 for 3 min each while shaking. After a 5 min incubation in blocking buffer at 37°C in the humid chamber, the samples were incubated in blocking buffer containing fluorescently-conjugated secondary antibody for 30 min in the humid chamber at 37°C, followed by three washes in PBS/0.2% Tween 20 while shaking for 3 min each. For samples undergoing only IF staining, DAPI was added to the third wash at a 1:250,000 dilution. Coverslips were then mounted on slides in Vectashield (Vector Labs, #H-1000). For samples undergoing IF combined with RNA FISH, the samples were processed for RNA FISH following the third wash.

RNA fluorescence *in situ* hybridization (RNA FISH)

RNA FISH with double-stranded and strand-specific probes was performed as previously described (Gayen *et al.*, 2015; Hinten et al., 2016; Kalantry et al., 2009). The *Rnf12* dsRNA FISH probe was made by random-priming using BioPrime DNA Labeling System (Invitrogen, #18094011) and labeled with Cy3-dCTP (GE Healthcare, #PA53021) using a previously described fosmid template (Kalantry et al., 2009). Strand-specific *Xist* probes were generated from templates as described (Maclary *et al.*, 2014; Sarkar *et al.*, 2015). Probes were labeled with Fluorescein-12-UTP (Roche, #11427857910) or Cy5-CTP (GE Healthcare, #25801087). Labeled probes from multiple templates were precipitated in a 0.5M ammonium acetate solution

(Sigma, #09691) along with 300 µg of yeast tRNA (Invitrogen, #15401–029) and 150 µg of sheared, boiled salmon sperm DNA (Invitrogen, #15632–011). The solution was then spun at 15,000 rpm for 20 min at 4°C. The pellet was washed consecutively with 70% ethanol and 100% ethanol while spinning at 15,000 rpm at room temperature. The pellet was dried and resuspended in deionized formamide (VWR, #97062–010). The probe was denatured by incubating at 90°C for 10 min followed by an immediate 5 min incubation on ice. A 2X hybridization solution consisting of 4X SSC and 20% Dextran sulfate (Millipore, #S4030) was added to the denatured solution. All probes were stored in the dark at −20°C until use. Following IF, embryos mounted on coverslips were dehydrated through 2 min incubations in 70%, 85%, 95%, and 100% ethanol solutions and subsequently air-dried. The coverslips were then hybridized to the probe overnight in a humid chamber at 37°C. The samples were then washed 3 times for 7 min each at 37°C with 2X SSC/50% formamide, 2X SSC, and 1X SSC. A 1:250,000 dilution of DAPI (Invitrogen, #D21490) was added to the third 2X SSC wash. Coverslips were then mounted on slides in Vectashield (Vector Labs, #H-1000).

Allele-specific *Xist* RNA FISH

Allele-specific *Xist* RNA FISH probes were generated as described (Levesque *et al.*, 2013). Briefly, a panel of short oligonucleotide probes were designed to uniquely detect either the *M. musculus* or the *M. molossinus* alleles of *Xist* (Table 3.4). Five probes were designed for each *Xist* allele. Each probe overlapped a single nucleotide polymorphism (SNP) that differs between the two strains, with the SNP located at the fifth base pair position from the 5' end. The same panel of five SNPs was used for both sets of allele-specific probes. The 3' end of each oligonucleotide probe is fluorescently tagged using Quasar dyes (Biosearch technologies). *M. musculus*-specific oligos were labeled with Quasar 570 and *M. molossinus*-specific oligos

labeled with Quasar 670. In addition to labeled SNP-overlapping oligonucleotides, a panel of 5 'mask' oligonucleotides were also synthesized. These 'mask' probes are complimentary to the 3' end of the labeled allele-specific probes and will hybridize to the allele-specific oligonucleotides, leaving only 9–10 base pairs of sequence surrounding the polymorphic site available to initially hybridize to the target *Xist* RNA. Since this region of complementarity is short, the presence of a single nucleotide polymorphism is sufficient to destabilize the hybridization with the alternate allele. Sequences of detection and mask probes are listed in Table 3.4. Allele-specific *Xist* RNA FISH probes were combined with a strand-specific *Xist* RNA probe, labeled with Fluorescein-12-UTP (Roche, #11427857910), which served as a guide probe that hybridizes to *Xist* RNA generated from both *Xist* alleles and ensured the fidelity of the allele-specific probes in detecting the cognate *Xist* RNA molecules. The guide *Xist* RNA probe was first ethanol precipitated as previously described, then resuspended in hybridization buffer containing 10% dextran sulfate, 2X saline-sodium citrate (SSC) and 10% formamide. The precipitated guide RNA probe was then mixed with the *M. musculus* and *M. molossinus* detection probes, to a final concentration of 5 nM per allele-specific oligo, and 10 nM mask probe, yielding a 1:1 mask:detection oligonucleotide ratio. Coverslips were hybridized to the combined probe overnight in a humid chamber at 37°C. After overnight hybridization, samples were washed twice in 2X SSC with 10% formamide at 37°C for 30 min, followed by one wash in 2X SSC for five min at room temperature. A 1:250,000 dilution of DAPI (Invitrogen, #D21490) was added to the second 2X SSC with 10% formamide wash. Coverslips were then mounted on slides in Vectashield (Vector Labs, #H-1000).

Microscopy

Stained samples were imaged using a Nikon Eclipse TiE inverted microscope with a Photometrics CCD camera. The images were deconvolved and uniformly processed using NIS-Elements software. For four color images (blue, green, red, and white), the far-red spectrum was employed for the fourth color (AlexaFluor 647 secondary antibody and Cy5-UTP labelled riboprobes for RNA FISH).

RNA-Seq sample preparation

mRNA was isolated from whole embryos using the Dynabeads mRNA DIRECT Kit (Thermo Fisher, # 610.11) according to the manufacturer's instructions. E3.5 embryos of similar sizes of all genotypes were used for RNA-Seq. Samples were submitted to the University of Michigan DNA Sequencing Core for strand-specific library preparation using the Takara SMARTer Seq V4 stranded low input kit (Takara, #634889). All libraries were sequenced on the Illumina HiSeq2000 or HiSeq4000 platforms to generate 50 bp paired-end reads.

Mapping of RNA-Seq data

Quality control analysis of the RNA-Seq data was conducted using FastQC. SNP data from whole-genome sequencing of the 129/S1 (*M. musculus*) and JF1/Ms (*M. molossinus*) mouse strains were substituted into the mm9 mouse reference genome build (C57BL/6 J) using VCFtools to generate in silico 129/S1 and JF1/Ms reference genomes (Keane *et al.*, 2011; Maclary *et al.*, 2017; Takada *et al.*, 2013; Yalcin *et al.*, 2011). Sequencing reads were separately mapped to each of the two in silico genomes using STAR (Dobin *et al.*, 2013), allowing 0 mismatches in mapped reads to ensure allele-specific mapping of SNP-containing reads to only one strain-specific genome. STAR was selected for read mapping, in part due to the improved ability to handle structural variability and indels, with the goal of reducing mapping bias to the genome most like the reference genome. STAR is a spliced aligner capable of detecting

structural variations and can handle small insertions and deletions during read mapping. STAR additionally permits soft-clipping of reads during mapping, trimming the ends of long reads that cannot be perfectly mapped. This function would permit clipping of reads that end near indels, thus preserving mappability at SNPs near indels.

Prior work showed that the variability due to mapping bias between the 129/S1 and JF1/Ms genomes is minimal in our RNA-Seq analysis pipeline (Maclary *et al.*, 2017). Although small biases may affect allelic mapping at a subset of SNP sites within a gene, the effect is mitigated since most genes contain multiple SNPs.

Allele-specific analysis of RNA-Seq data

For allelic expression analysis, only RNA-Seq reads overlapping known SNP sites that differ between the 129/S1 and JF1/Ms genomes were retained. All multi-mapping reads were excluded from the allele-specific analysis. For each SNP site, reads mapping to the 129/S1 and JF1/Ms X chromosomes were counted and the proportion of reads from each X chromosome identified. Allelic expression was calculated individually for each SNP site; for genes containing multiple SNPs, the paternal-X percentage for all SNPs was averaged to calculate gene-level allelic expression. All SNP sites with at least 10 SNP-overlapping reads were retained. Genes containing at least one SNP site with at least 10 SNP-overlapping reads were retained for further analysis and are referred to in the text as informative. In X-linked genes, the SNP frequency is ~1 SNP/250 bp in transcribed RNAs (Keane *et al.*, 2011; Maclary *et al.*, 2017; Takada *et al.*, 2013; Yalcin *et al.*, 2011).

RNA-Seq expression analysis

To calculate expression from the maternal vs. paternal X-chromosomes, all reads were first merged into a single alignment file and the number of reads per RefSeq annotated gene was

counted using HTSeq. To calculate the percentage of expression arising from the paternal X-chromosome, the total read counts from HTSeq were normalized by number of mapped reads. Then, the normalized number of mapped reads for each gene was multiplied by the proportion of SNP-containing reads mapping to the paternal X-chromosome. This analysis was done in R using the following formula:

$$\left\{\text{total reads} \times \left(\frac{\text{paternal reads}}{\text{maternal reads} + \text{paternal reads}}\right)\right\}$$

Statistical Analysis & Plots

Welch's two-sample T-tests were used to test for significant differences between the means of Pyrosequencing and RNA-Seq allelic expression data. This test was chosen due to the unequal variance and sample sizes between different genotype groups. In the RNA-Seq allelic expression significance tests, the average percent paternal expression of all informative X-linked genes was calculated for each sample. The total paternal expression value for each genotype group was determined by calculating the mean of the informative percent paternal values for all samples in that genotype group. Histograms and heatmaps were made using the ggplot and Pheatmaps R packages, respectively. Only genes that were informative in all samples were included in the heatmaps.

Author Contributions

C.H. and M.H. generated the mouse lines characterized in this chapter and conducted the IF FISH experiments. M.C. and M.H. dissected and processed the mouse embryos and conducted the RNA-Seq and Pyrosequencing experiments. M.C. performed the allele-specific RNA FISH experiments. Z.D. and W.X. performed the allele-specific ChIP-Seq experiment. M.C. performed the RNA-Seq and ChIP-Seq data analyses. C.H., M.C., M.H. and S.K. designed

experiments. M.C and S.K. prepared the figures and text for this chapter, which is in preparation to be submitted as a manuscript to a peer-reviewed journal.

Table 3.1. Percent paternal-X expression values for all genes with informative expression in all *Eed*$^{mz-/-}$, *Ezh1/2*$^{mz-/-}$, and *Eed*$^{m-/-}$ embryos sequenced. Table also indicates whether each gene exhibited unbiased allelic expression (between 37.5 – 62.5% paternal expression).

Gene Name	Average *Eed*$^{mz-/-}$ % Paternal Expression	*Eed*$^{mz-/-}$ Unbiased Expression	Average *Ezh1/2*$^{mz-/-}$ % Paternal Expression	*Ezh2/1*$^{mz-/-}$ Unbiased Expression	Average *Eed*$^{m-/-}$ % Paternal Expression	*Eed*$^{m-/-}$ Unbiased Expression
1110012L19Rik	0.293	NO	0.085	NO	0.160	NO
1600014K23Rik	0.635	NO	0.207	NO	0.302	NO
1810030O07Rik	0.509	YES	0.385	YES	0.692	NO
2210013O21Rik	0.139	NO	0.022	NO	0.065	NO
2610030H06Rik	0.444	YES	0.111	NO	0.308	NO
2900062L11Rik	0.344	NO	0.030	NO	0.000	NO
5530601H04Rik	0.370	NO	0.359	NO	0.411	YES
A830080D01Rik	0.429	YES	0.314	NO	0.427	YES
Abcb7	0.339	NO	0.204	NO	0.025	NO
Abcd1	0.293	NO	0.279	NO	0.446	YES
Acot9	0.444	YES	0.181	NO	0.253	NO
Aifm1	0.442	YES	0.302	NO	0.237	NO
Alas2	0.003	NO	0.000	NO	0.000	NO
Apex2	0.424	YES	0.230	NO	0.199	NO
Apoo	0.477	YES	0.451	YES	0.501	YES
Apool	0.398	YES	0.249	NO	0.346	NO
Armcx5	0.608	YES	0.305	NO	0.598	YES
Atg4a-2	0.221	NO	0.046	NO	0.078	NO
Atp6ap1	0.453	YES	0.294	NO	0.384	YES
Atp6ap2	0.466	YES	0.341	NO	0.362	NO
Atrx	0.376	YES	0.351	NO	0.224	NO
BC023829	0.423	YES	0.250	NO	0.073	NO
Bcap31	0.465	YES	0.367	NO	0.374	NO
Bex1	0.537	YES	0.404	YES	0.327	NO
Brcc3	0.499	YES	0.299	NO	0.509	YES
Btk	0.148	NO	0.181	NO	0.028	NO
C1galt1c1	0.397	YES	0.246	NO	0.149	NO
C430049B03Rik	0.299	NO	0.223	NO	0.267	NO
Capn6	0.325	NO	0.150	NO	0.091	NO
Ccdc120	0.419	YES	0.281	NO	0.556	YES
Ccdc22	0.509	YES	0.310	NO	0.243	NO
Cdk16	0.407	YES	0.329	NO	0.381	YES
Cenpi	0.459	YES	0.165	NO	0.391	YES

Cfp	0.243	NO	0.104	NO	0.001	NO
Chm	0.345	NO	0.283	NO	0.217	NO
Cox7b	0.365	NO	0.256	NO	0.112	NO
Cstf2	0.491	YES	0.276	NO	0.391	YES
Ctps2	0.404	YES	0.172	NO	0.145	NO
Cul4b	0.458	YES	0.346	NO	0.389	YES
Cybb	0.252	NO	0.107	NO	0.040	NO
Ddx3x	0.589	YES	0.264	NO	0.315	NO
Diap2	0.409	YES	0.482	YES	0.388	YES
Dkc1	0.465	YES	0.363	NO	0.515	YES
Dlg3	0.413	YES	0.233	NO	0.486	YES
Dusp9	0.488	YES	0.423	YES	0.257	NO
Dynlt3	0.500	YES	0.501	YES	0.500	YES
E530001F21Rik	1.000	NO	0.885	NO	0.628	NO
Ebp	0.452	YES	0.265	NO	0.202	NO
Efnb1	0.339	NO	0.149	NO	0.000	NO
Eif1ax	0.418	YES	0.261	NO	0.220	NO
Eif2s3x	0.883	NO	0.703	NO	0.777	NO
Elk1	0.200	NO	0.196	NO	0.201	NO
Emd	0.491	YES	0.447	YES	0.479	YES
Ercc6l	0.472	YES	0.415	YES	0.239	NO
F8a	0.301	NO	0.136	NO	0.072	NO
Fam122b	0.442	YES	0.232	NO	0.098	NO
Fam3a	0.426	YES	0.311	NO	0.474	YES
Fgd1	0.351	NO	0.294	NO	0.144	NO
Fhl1	0.196	NO	0.081	NO	0.000	NO
Flna	0.340	NO	0.128	NO	0.150	NO
Fmr1nb	0.389	YES	0.259	NO	0.577	YES
Ftsj1	0.435	YES	0.337	NO	0.197	NO
Fundc1	0.416	YES	0.256	NO	0.122	NO
Fundc2	0.324	NO	0.169	NO	0.195	NO
G6pdx	0.419	YES	0.227	NO	0.332	NO
Gata1	0.413	YES	0.247	NO	0.467	YES
Gdi1	0.625	YES	0.355	NO	0.646	NO
Gla	0.499	YES	0.329	NO	0.229	NO
Gm4779	0.541	YES	0.431	YES	0.454	YES
Gm6222	0.686	NO	0.414	YES	0.516	YES
Gm6880	0.691	NO	0.723	NO	0.735	NO
Gm9	0.433	YES	0.093	NO	0.075	NO

Gnl3l	0.436	YES	0.166	NO	0.604	YES
Gpc3	0.357	NO	0.306	NO	0.000	NO
Gpkow	0.670	NO	0.290	NO	0.529	YES
Gprasp1	0.414	YES	0.428	YES	0.596	YES
Gripap1	0.497	YES	0.207	NO	0.275	NO
Gyk	0.563	YES	0.307	NO	0.000	NO
Haus7	0.420	YES	0.279	NO	0.199	NO
Hccs	0.477	YES	0.404	YES	0.448	YES
Hcfc1	0.464	YES	0.246	NO	0.385	YES
Hdac6	0.417	YES	0.322	NO	0.212	NO
Hmgb3	0.539	YES	0.424	YES	0.387	YES
Hmgn5	0.334	NO	0.303	NO	0.068	NO
Hnrnph2	0.303	NO	0.151	NO	0.203	NO
Hprt	0.321	NO	0.228	NO	0.322	NO
Hsd17b10	0.464	YES	0.109	NO	0.231	NO
Htatsf1	0.428	YES	0.149	NO	0.330	NO
Huwe1	0.489	YES	0.457	YES	0.572	YES
Idh3g	0.454	YES	0.370	NO	0.359	NO
Ikbkg	0.660	NO	0.303	NO	0.474	YES
Il2rg	0.315	NO	0.099	NO	0.019	NO
Irak1	0.451	YES	0.154	NO	0.196	NO
Kdm5c	0.542	YES	0.311	NO	0.494	YES
Kdm6a	0.438	YES	0.474	YES	0.208	NO
Kif4	0.624	YES	0.292	NO	0.583	YES
Klhl15	0.512	YES	0.279	NO	0.430	YES
Lage3	0.439	YES	0.373	NO	0.458	YES
Lamp2	0.329	NO	0.143	NO	0.160	NO
Las1l	0.492	YES	0.287	NO	0.406	YES
Ldoc1	0.298	NO	0.227	NO	0.105	NO
LOC100503426	0.440	YES	0.220	NO	0.311	NO
LOC100503459	0.997	NO	0.998	NO	0.999	NO
LOC100503548	0.205	NO	0.002	NO	0.026	NO
LOC100503934	0.550	YES	0.409	YES	0.508	YES
LOC100504335	0.408	YES	0.300	NO	0.335	NO
LOC100504463	0.389	YES	0.284	NO	0.125	NO
Mageb16	0.432	YES	0.429	YES	0.703	NO
Maged1	0.380	YES	0.166	NO	0.153	NO
Maged2	0.325	NO	0.181	NO	0.014	NO
Magt1	0.307	NO	0.156	NO	0.092	NO

Maoa	0.237	NO	0.253	NO	0.122	NO
Map3k15	0.531	YES	0.343	NO	0.520	YES
Mbnl3	0.477	YES	0.354	NO	0.599	YES
Mcts1	0.379	YES	0.295	NO	0.280	NO
Mecp2	0.481	YES	0.304	NO	0.693	NO
Med14	0.411	YES	0.442	YES	0.348	NO
Mid1ip1	0.270	NO	0.059	NO	0.029	NO
Mmgt1	0.448	YES	0.352	NO	0.235	NO
Morf4l2	0.478	YES	0.299	NO	0.327	NO
Mospd1	0.408	YES	0.304	NO	0.334	NO
Mpp1	0.399	YES	0.329	NO	0.389	YES
Msl3	0.441	YES	0.301	NO	0.504	YES
Msn	0.264	NO	0.050	NO	0.015	NO
Mtm1	0.254	NO	0.181	NO	0.168	NO
Naa10	0.469	YES	0.378	YES	0.413	YES
Ndufa1	0.402	YES	0.309	NO	0.372	NO
Ndufb11	0.390	YES	0.291	NO	0.246	NO
Ngfrap1	0.360	NO	0.291	NO	0.265	NO
Nkap	0.411	YES	0.360	NO	0.544	YES
Nono	0.046	NO	0.027	NO	0.035	NO
Nsdhl	0.270	NO	0.187	NO	0.017	NO
Ofd1	0.558	YES	0.430	YES	0.478	YES
Ogt	0.519	YES	0.331	NO	0.424	YES
Otud5	0.577	YES	0.316	NO	0.366	NO
Pdha1	0.510	YES	0.467	YES	0.559	YES
Pdzd11	0.416	YES	0.239	NO	0.462	YES
Pgk1	0.470	YES	0.186	NO	0.064	NO
Pgrmc1	0.380	YES	0.204	NO	0.261	NO
Phf16	0.307	NO	0.117	NO	0.259	NO
Pim2	0.529	YES	0.441	YES	0.344	NO
Plac1	0.488	YES	0.435	YES	0.277	NO
Plp2	0.385	YES	0.254	NO	0.191	NO
Pls3	0.500	YES	0.298	NO	0.145	NO
Pola1	0.466	YES	0.514	YES	0.423	YES
Porcn	0.411	YES	0.365	NO	0.249	NO
Pqbp1	0.418	YES	0.287	NO	0.286	NO
Praf2	0.473	YES	0.211	NO	0.321	NO
Prkx	0.258	NO	0.236	NO	0.095	NO
Prps1	0.397	YES	0.250	NO	0.326	NO

Prps2	0.353	NO	0.102	NO	0.469	YES
Psmd10	0.348	NO	0.223	NO	0.246	NO
Rab9	0.477	YES	0.355	NO	0.394	YES
Rap2c	0.420	YES	0.210	NO	0.275	NO
Rbbp7	0.531	YES	0.374	NO	0.424	YES
Rbm10	0.549	YES	0.295	NO	0.396	YES
Rbm3	0.442	YES	0.376	YES	0.227	NO
Rbmx2	0.242	NO	0.278	NO	0.255	NO
Renbp	0.466	YES	0.363	NO	0.347	NO
Rlim	0.422	YES	0.206	NO	0.151	NO
Rnf128	0.418	YES	0.182	NO	0.225	NO
Rp2h	0.438	YES	0.179	NO	0.208	NO
Rpgr	0.505	YES	0.306	NO	0.730	NO
Rpl10	0.905	NO	0.608	YES	0.832	NO
Rpl36a	0.647	NO	0.480	YES	0.582	YES
Rps4x	0.458	YES	0.360	NO	0.419	YES
S100g	0.174	NO	0.061	NO	0.000	NO
Sash3	0.000	NO	0.035	NO	0.026	NO
Sat1	0.428	YES	0.280	NO	0.376	YES
Sept6	0.233	NO	0.127	NO	0.000	NO
Sh3bgrl	0.161	NO	0.022	NO	0.075	NO
Sh3kbp1	0.000	NO	0.019	NO	0.000	NO
Siah1b	0.446	YES	0.446	YES	0.133	NO
Slc10a3	0.259	NO	0.128	NO	0.188	NO
Slc25a14	0.304	NO	0.193	NO	0.210	NO
Slc25a43	0.453	YES	0.062	NO	0.000	NO
Slc35a2	0.540	YES	0.364	NO	0.328	NO
Slc9a6	0.388	YES	0.285	NO	0.185	NO
Smc1a	0.554	YES	0.276	NO	0.342	NO
Sms	0.629	NO	0.536	YES	0.710	NO
Snx12	0.555	YES	0.429	YES	0.516	YES
Stag2	0.450	YES	0.179	NO	0.494	YES
Suv39h1	0.568	YES	0.322	NO	0.352	NO
Syap1	0.455	YES	0.316	NO	0.270	NO
Taz	0.503	YES	0.266	NO	0.335	NO
Tbc1d25	0.382	YES	0.489	YES	0.320	NO
Tbl1x	0.420	YES	0.125	NO	0.540	YES
Tceal8	0.343	NO	0.151	NO	0.023	NO
Tcfe3	0.422	YES	0.218	NO	0.244	NO

Thoc2	0.399	YES	0.341	NO	0.454	YES
Timm17b	0.406	YES	0.255	NO	0.382	YES
Timm8a1	0.961	NO	0.741	NO	0.843	NO
Timp1	0.372	NO	0.244	NO	0.133	NO
Tmem29	0.411	YES	0.393	YES	0.293	NO
Tmsb15b1	0.393	YES	0.284	NO	0.196	NO
Tmsb15b1-Tmsb15b2	0.318	NO	0.080	NO	0.044	NO
Tmsb15b2	0.275	NO	0.044	NO	0.017	NO
Trap1a	0.158	NO	0.036	NO	0.071	NO
Trappc2	0.066	NO	0.042	NO	0.057	NO
Trmt2b	0.398	YES	0.215	NO	0.128	NO
Tsc22d3	0.339	NO	0.142	NO	0.108	NO
Tsr2	0.458	YES	0.404	YES	0.334	NO
Uba1	0.457	YES	0.324	NO	0.402	YES
Ube2a	0.394	YES	0.232	NO	0.232	NO
Ubl4	0.487	YES	0.380	YES	0.457	YES
Utp14a	0.566	YES	0.330	NO	0.341	NO
Vbp1	0.459	YES	0.350	NO	0.427	YES
Vsig4	0.000	NO	0.089	NO	0.000	NO
Wbp5	0.486	YES	0.219	NO	0.226	NO
Wdr45	0.647	NO	0.265	NO	0.439	YES
Xiap	0.380	YES	0.091	NO	0.298	NO
Xist	0.451	YES	0.359	NO	0.535	YES
Xlr3a	0.394	YES	0.252	NO	0.393	YES
Xlr3b	0.813	NO	0.228	NO	0.359	NO
Xlr4c	0.465	YES	0.251	NO	0.022	NO
Yipf6	0.575	YES	0.244	NO	0.164	NO
Zdhhc9	0.387	YES	0.129	NO	0.089	NO
Zfp280c	0.429	YES	0.285	NO	0.833	NO
Zfx	0.588	YES	0.284	NO	0.112	NO
Zmym3	0.375	NO	0.215	NO	0.643	NO
Zrsr2	0.482	YES	0.364	NO	0.303	NO

Table 3.2. Genotyping PCR Primers

Gene/Chromosome	Primer Name	Primer Sequence
Ezh1	WT-R2	GCTCCTGTCCTCATAGCAAGA
	SA-1	GTACTCTTAACCACTGGACTG
	LACZ-2	AAGCGCCATTCGCCATTCAGG
Ezh2	Enx1-3-loxP	CTGCTCTGAATGGCAACTCC
	Ezh2-5-loxp-3	CTGGCTCTGTGGAACCAAAC
	Ezh2-L5-loxp-1	ATGGGCCTCATAGTGACAGG
X/Y Chromosome	XY F	CCGCTGCCAAATTCTTTGG
	XY R	TGAAGCTTTTGGCTTTGAG

Table 3.3. Pyrosequencing Primers

Gene	Primer Name	Primer Sequence
G6pdx	G6pdx_RT_F	GGAGTCCAGGGGCAGACTGATA
	G6pdx_RT_R	5'-biotin-CCACCCATCTTTCCACAAGACC
	G6pdx_PyroSeq	TGATAGGCATTCTTTCTG
Atrx	Atrx_RT_F	ATAGCTTCAGATTCTGATGAAACC
	Atrx_RT_R	5′-biotin-ACATCGTTGTCACTGCCACTT
	Atrx_PyroSeq	TAAGCTCAGATGAAAAGA
Pdha1	Pdha1_RT_F	5'-biotin-AGCAATCTTGCAAGTGTTGAAGAA
	Pdha1_RT_R	TTTTCAAGCCTTTTGTTGTCTGG
	Pdha1_PyroSeq	TAGAACTTCCTTCAGGC
Rnf12	Rnf12_RT_F	TGCAGCCAACAAGTGAAATTCC
	Rnf12_RT_R	5'-biotin-TATCTGCTGTCTCAGGGTCACATG
	Rnf12_PyroSeq	TAGAACTTCCTTCAGGC

Table 3.4. Allele-specific RNA FISH Probe Coordinates and Sequences

SNP Coordinate (mm9)	*M. musculus* Specific Probe (with 3' Quasar 570)	*M. molossinus* Specific Probe (with 3' Quasar 670)	Mask Probe
chrX:100664254	ATCACGCTGAAGACCCAGTTTTCTG	ATCATGCTGAAGACCCAGTTTTCTG	CAGAAAACTGGGTCTT
chrX:100669174	ATGCTGGGAGAACTGCTGTTGTGATG	ATGCCGGGAGAACTGCTGTTGTGATG	CATCACAACAGCAGTT
chrX:100676048	GCTCGGTGGATGAGTTTGAAAGAAAGTAC	GCTCAGTGGATGAGTTTGAAAGAAAGTAC	GTACTTTCTTTCAAACTCA
chrX:100676261	GTGTCGTTGGCATCCAAAATATTCATTG	GTGTTGTTGGCATCCAAAATATTCATTG	CAATGAATATTTTGGATGC
chrX:100677431	CTGCGGCTTCCGCGCAACACC	CTGCTGCTTCCGCGCAACACC	GGTGTTGCGCGG

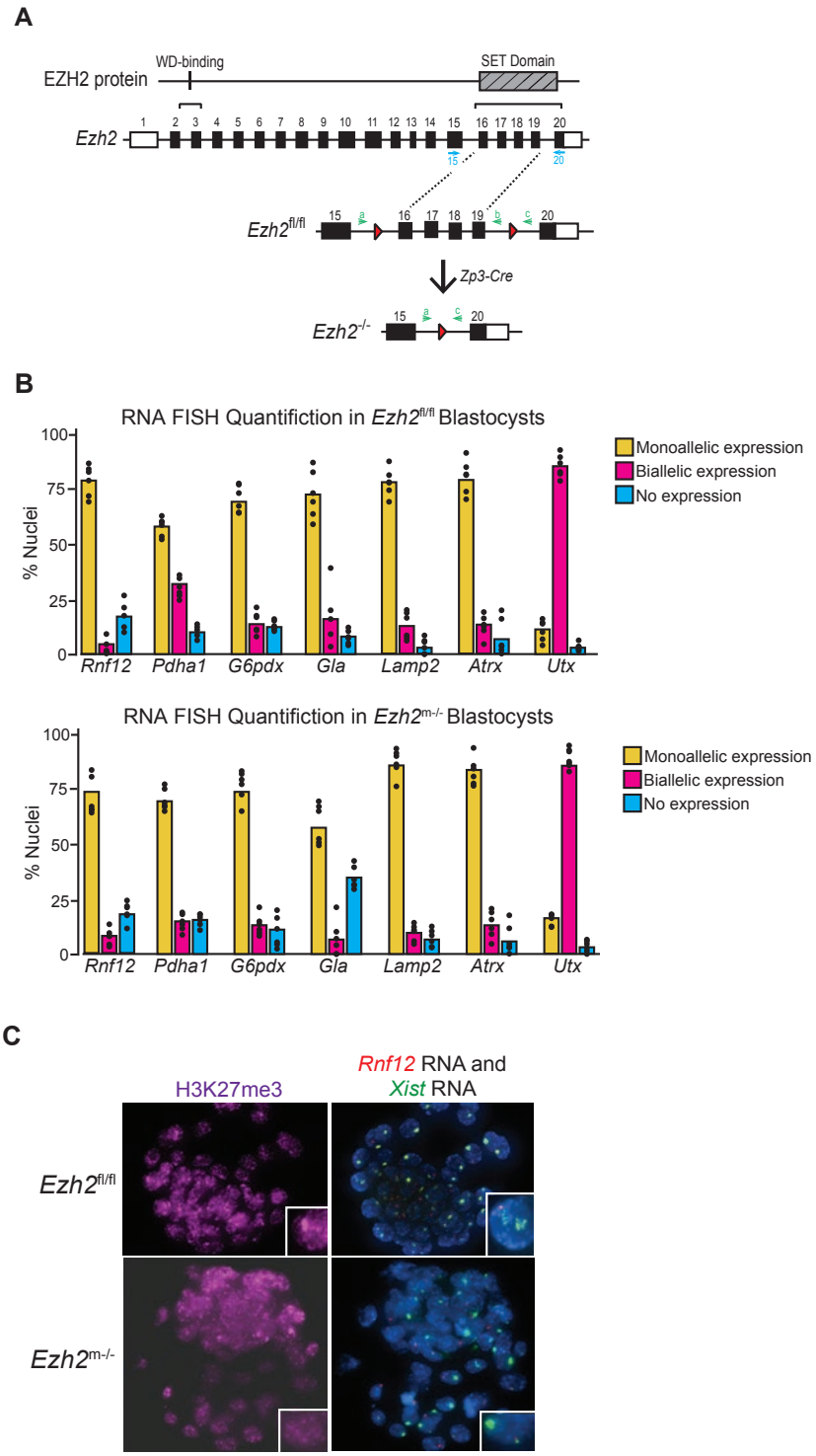

A
EZH2 protein
WD-binding
SET Domain
Ezh2
Ezh2fl/fl
Zp3-Cre
Ezh2-/-
B
RNA FISH Quantifiction in Ezh2fl/fl Blastocysts
% Nuclei
Monoallelic expression
Biallelic expression
No expression
Rnf12
Pdha1
G6pdx
Gla
Lamp2
Atrx
Utx
RNA FISH Quantifiction in Ezh2m-/- Blastocysts
C
H3K27me3
Rnf12 RNA and
Xist RNA
Ezh2fl/fl
Ezh2m-/-

Figure 3.1. Imprinted X-inactivation in mouse embryos lacking maternal EZH2 (A) Schematic depicting the *Ezh2* deletion strategy used in this study. (B) Quantification X-linked gene expression by RNA FISH in *Ezh2*$^{fl/fl}$ and *Ezh2*$^{m-/-}$ female mouse blastocysts. Individual dots represent embryos stained and bars represent the average % of nuclei with monoallelic (yellow), biallelic (pink), or no expression (blue) in all embryos. (C) Representative image of *Ezh2*$^{fl/fl}$ and *Ezh2*$^{m-/-}$ female mouse blastocysts stained to detect *Xist* RNA coating (in green), *Rnf12* RNA (in red), and histone H3 lysine 27 tri-methylation (H3-K27me3; in purple). Nuclei are stained blue with DAPI.

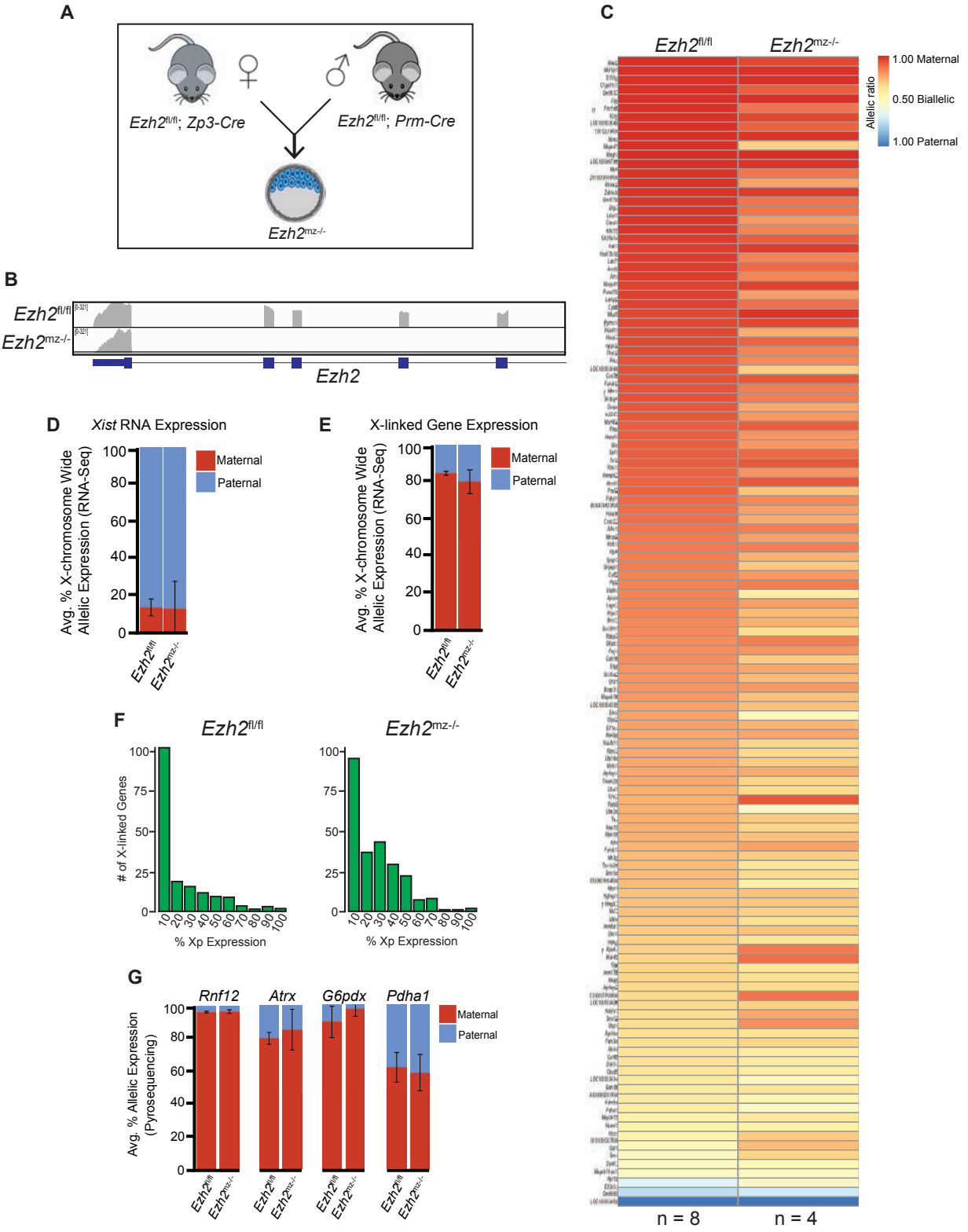
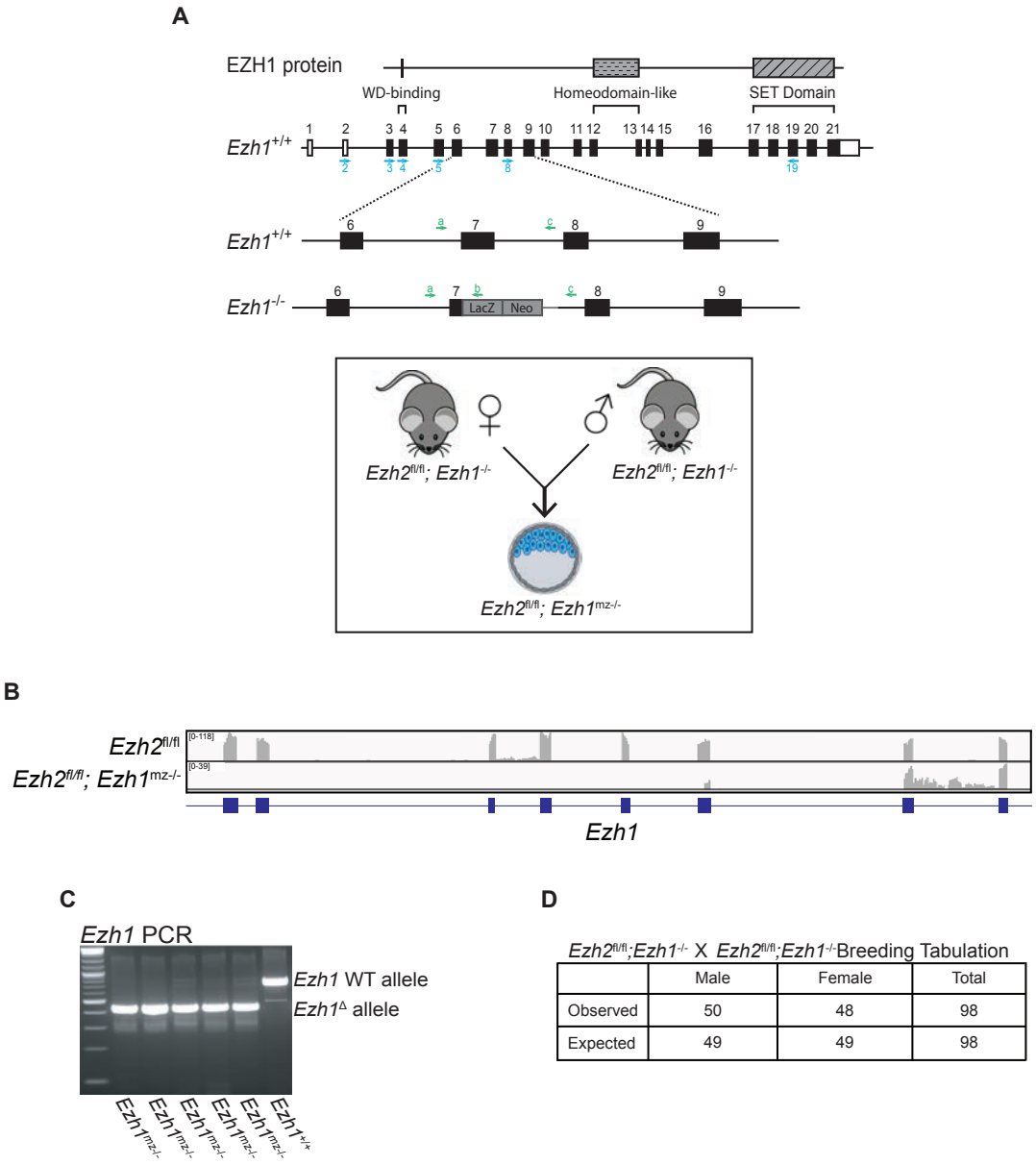

A
Ezh2fl/fl; Zp3-Cre
Ezh2fl/fl; Prm-Cre
Ezh2mz-/-
B
Ezh2fl/fl
Ezh2mz-/-
Ezh2
C
Ezh2fl/fl
Ezh2mz-/-
Allelic ratio
1.00 Maternal
0.50 Biallelic
1.00 Paternal
n = 8
n = 4
D
Xist RNA Expression
Avg. % X-chromosome Wide Allelic Expression (RNA-Seq)
Maternal
Paternal
E
X-linked Gene Expression
F
# of X-linked Genes
% Xp Expression
G
Rnf12
Atrx
G6pdx
Pdha1
Avg. % Allelic Expression (Pyrosequencing)

Figure 3.2. Imprinted X-inactivation in blastocysts lacking maternal and zygotic EZH2 (A) Breeding scheme used to generate *Ezh2*$^{mz-/-}$ embryos. (B) Representative Integrative Genome Browser snapshot of RNA-Seq data demonstrating successful *Ezh2* deletion. (C) Allele-specific X-linked gene expression heat map of female *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ blastocysts. Eight embryos of the *Ezh2*$^{fl/fl}$ genotype and four embryos of the *Ezh2*$^{mz-/-}$ genotype were sequenced individually and only genes with informative allelic expression in all samples are plotted (see Materials and Methods). (D) Average maternal:paternal *Xist* expression ratio from the RNA-Seq data. *Xist* is not included in (C) because informative allelic expression was not achieved in all samples sequenced. The mean allelic expression of *Xist* is not significantly different between *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ blastocysts ($p > 0.05$, Welch's two-sample T-test). (E) Average maternal:paternal X-linked gene expression ratio from the RNA-Seq data shown in (C). The mean allelic expression of X-linked genes is not significantly different between *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ blastocysts. ($p > 0.05$, Welch's two-sample T-test). (F) Histograms depicting the average number of X-linked genes undergoing various degrees (in 10% increments) of expression from the paternal-X in *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ blastocysts. (G) Pyrosequencing-based quantification of allelic expression of X-linked genes in *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ blastocysts. Error bars represent the standard deviation of data from 3 to 6 independent blastocyst embryos. The mean allelic expression of *Rnf12*, *Atrx, G6pdx,* and *Pdha1* is not significantly different between *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$ embryos ($p > 0.05$, Welch's two-sample T-test).



	Male	Female	Total
Observed	50	48	98
Expected	49	49	98

Figure 3.3. Generating mouse embryos lacking maternal and zygotic EZH1 (A) Schematic of the *Ezh1* deletion strategy and breeding scheme to generate *Ezh2*$^{fl/fl}$; *Ezh1*$^{mz-/-}$ embryos. (B) Representative Integrative Genome Browser snapshot of RNA-Seq data demonstrating successful *Ezh1* deletion. (C) Representative PCR gel demonstrating successful *Ezh1* deletion. (D) Breeding tabulation from an intercross of *Ezh2*$^{fl/fl}$; *Ezh1*$^{-/-}$ female mice with *Ezh2*$^{fl/fl}$; *Ezh1*$^{-/-}$ male mice demonstrating that *Ezh2*$^{fl/fl}$; *Ezh1*$^{mz-/-}$ offspring are viable.

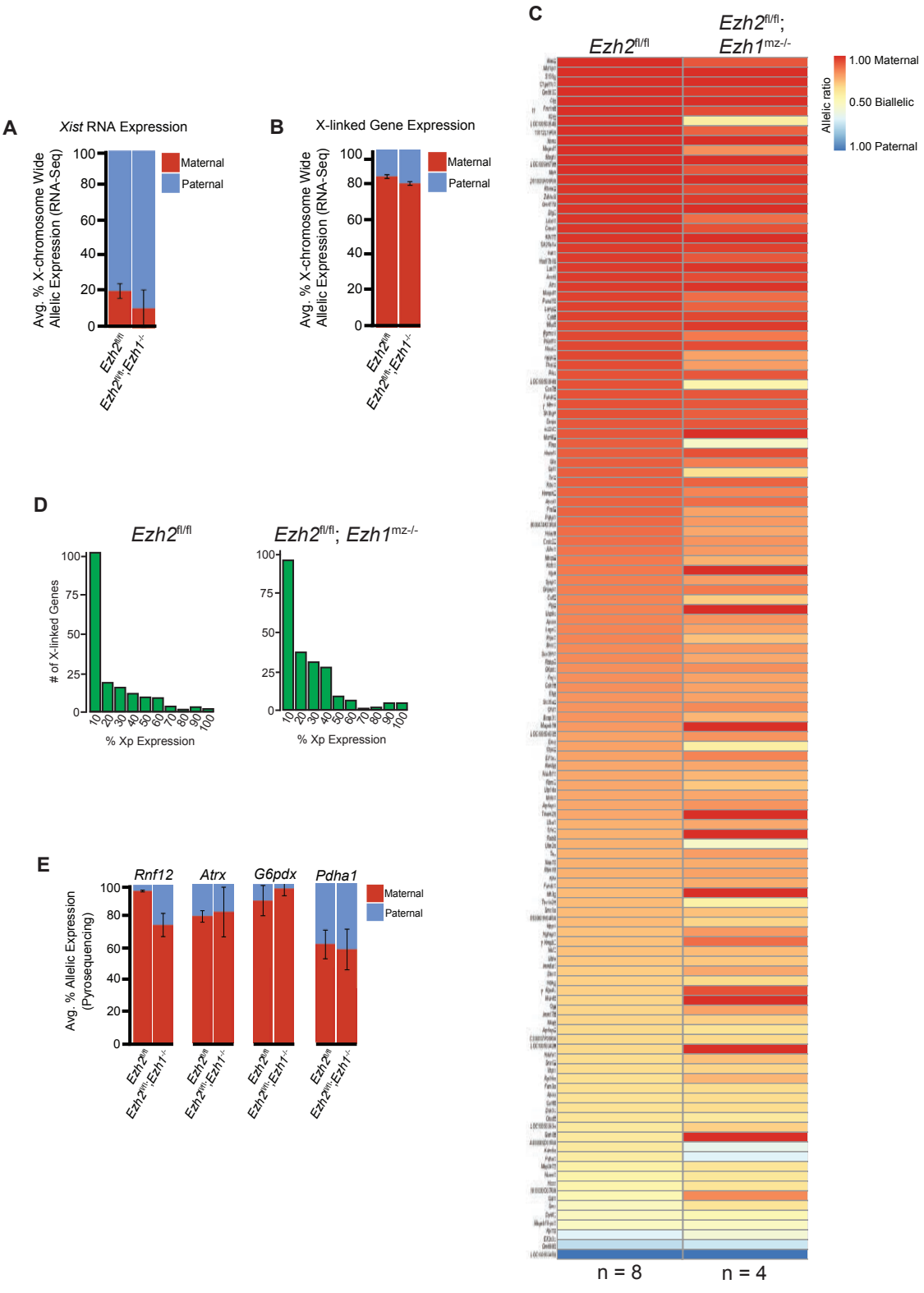

A
Xist RNA Expression
Avg. % X-chromosome Wide Allelic Expression (RNA-Seq)
Maternal
Paternal
Ezh2fl/fl
Ezh2fl/fl;Ezh1-/-
B
X-linked Gene Expression
C
Ezh2fl/fl
Ezh2fl/fl; Ezh1mz-/-
1.00 Maternal
0.50 Biallelic
1.00 Paternal
Allelic ratio
n = 8
n = 4
D
# of X-linked Genes
% Xp Expression
E
Rnf12
Atrx
G6pdx
Pdha1
Avg. % Allelic Expression (Pyrosequencing)

Figure 3.4. Imprinted X-inactivation in blastocysts lacking maternal and zygotic EZH1 (A) Average maternal:paternal *Xist* expression ratio from the RNA-Seq data. The mean allelic expression of *Xist* is not significantly different between *Ezh2*$^{fl/fl}$ and *Ezh1*$^{mz-/-}$ blastocysts ($p > 0.05$, Welch's two-sample T-test). (B) Average maternal:paternal X-linked gene expression ratio from the RNA-Seq data. The mean allelic expression of X-linked genes is not significantly different between *Ezh2*$^{fl/fl}$ and *Ezh1*$^{mz-/-}$ blastocysts. ($p > 0.05$, Welch's two-sample T-test). (C) Allele-specific X-linked gene expression heat map of female *Ezh2*$^{fl/fl}$ and *Ezh1*$^{mz-/-}$ blastocysts. Eight embryos of the *Ezh2*$^{fl/fl}$ genotype and four embryos of the *Ezh1*$^{mz-/-}$ genotype were sequenced individually and only genes with informative allelic expression in all samples are plotted (see Materials and Methods). *Xist* is not included in the heat map because informative allelic expression was not achieved for this gene in all samples sequenced. (D) Histograms depicting the average number of X-linked genes undergoing various degrees (in 10% increments) of expression from the paternal-X in *Ezh2*$^{fl/fl}$ and *Ezh1*$^{mz-/-}$ blastocysts. (E) Pyrosequencing-based quantification of allelic expression of X-linked genes in *Ezh2*$^{fl/fl}$ and *Ezh1*$^{mz-/-}$ blastocysts. Error bars represent the standard deviation of data from 3 to 6 independent blastocyst embryos. The mean allelic expression of *Rnf12*, *Atrx, G6pdx,* and *Pdha1* is not significantly different between *Ezh2*$^{fl/fl}$ and *Ezh1*$^{mz-/-}$ embryos ($p > 0.05$, Welch's two-sample T-test).

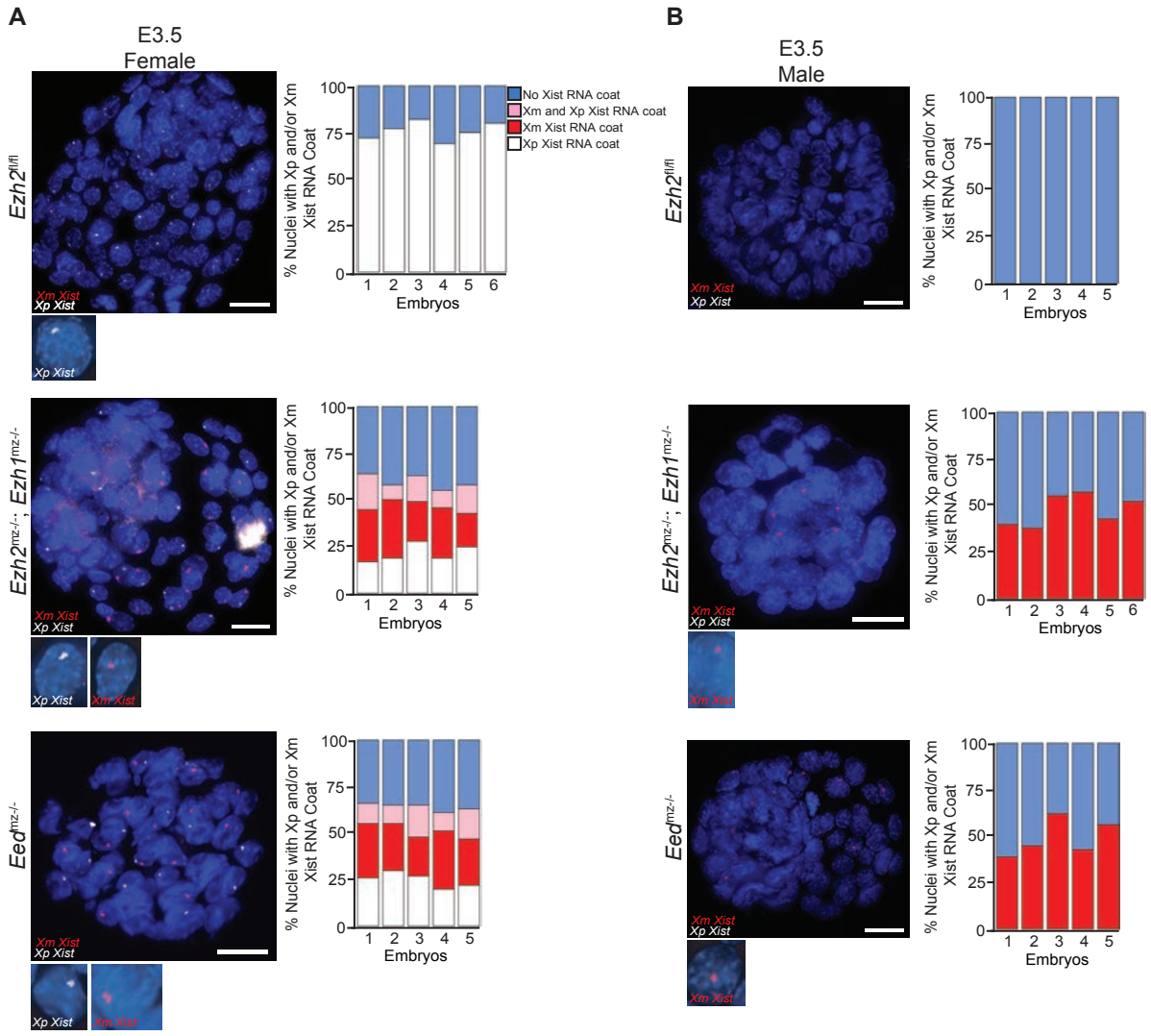


Figure 3.5. Switching of imprinted to random X-inactivation in E3.5 embryos lacking maternal and zygotic EZH1/2 and maternal and zygotic EED (A,B) Allele-Specific *Xist* RNA FISH in *Ezh2*$^{fl/fl}$, *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$, and *Eed*$^{mz-/-}$ female and male E3.5 blastocyst-stage embryos. *Xist* RNA expressed from the maternal X-chromosome is indicated in red and from the paternal X-chromosome in white. Representative embryos are depicted, in addition to representative images of individual nuclei depicting all gene expression patterns exhibited in each genotype and quantification bar graphs representing all embryos stained. Nuclei are stained blue with DAPI. Scale bars, 20 µm.

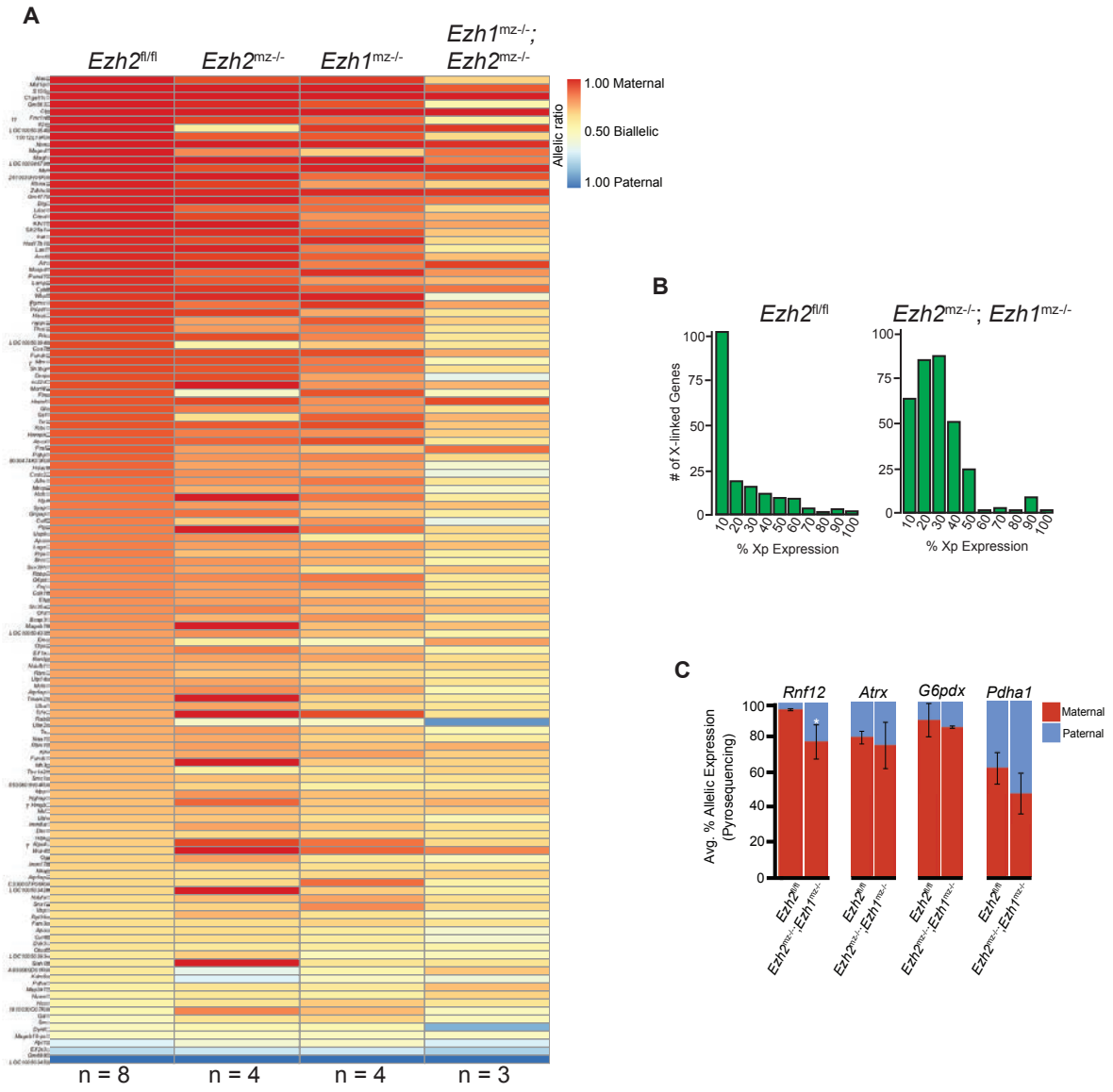
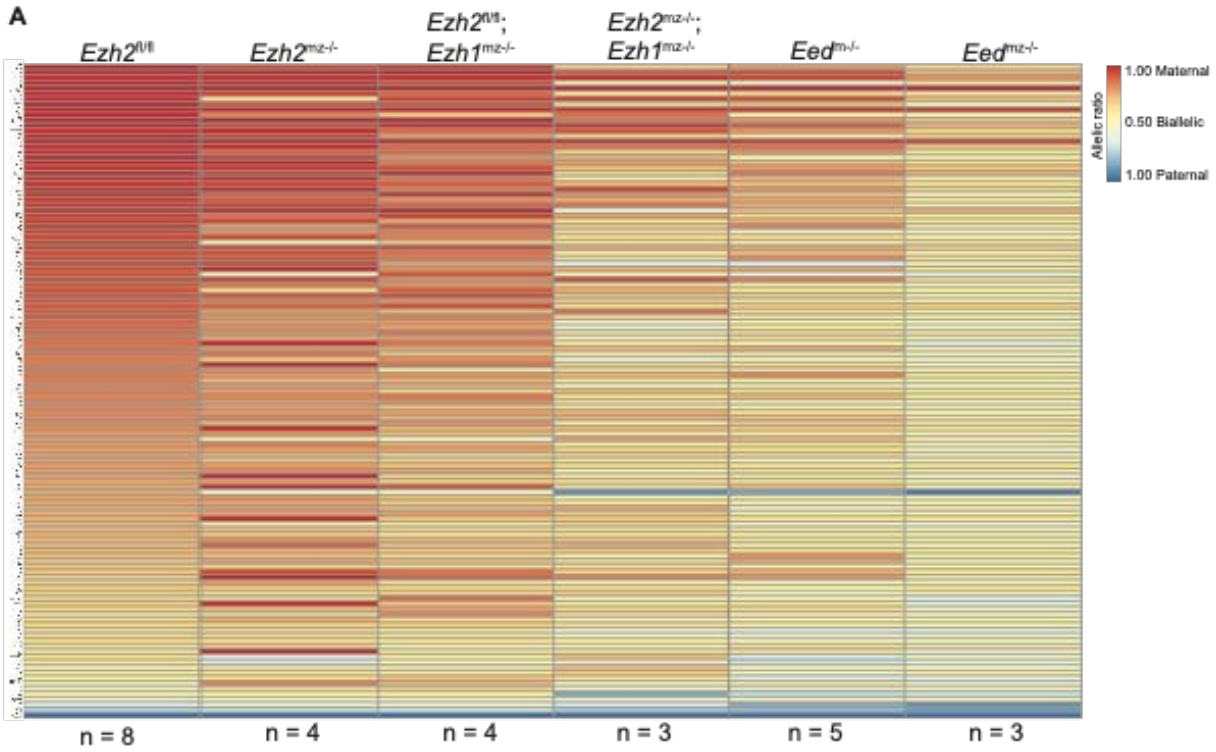
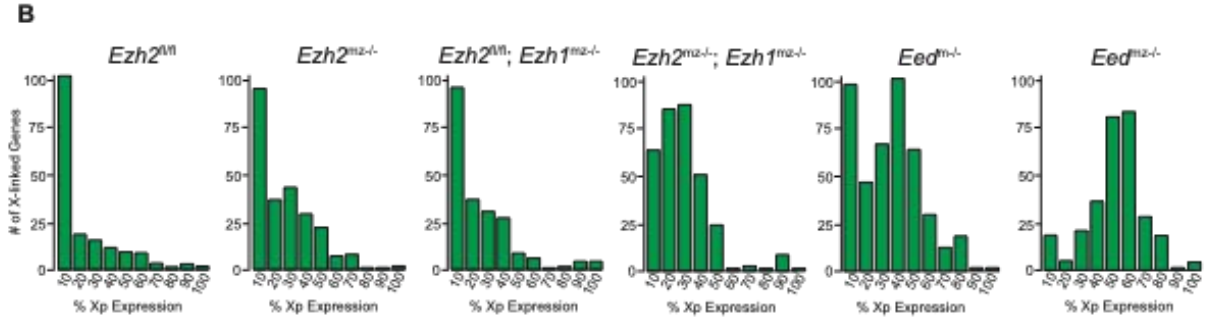
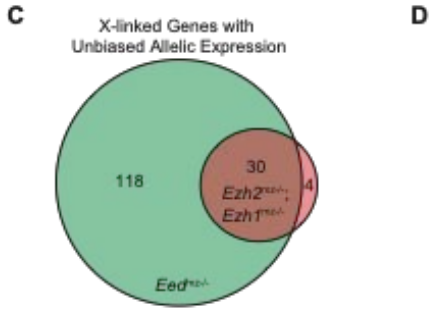
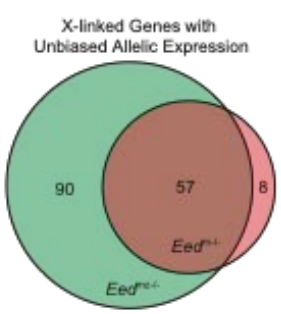

A
Ezh2fl/fl
Ezh2mz-/-
Ezh1mz-/-
Ezh1mz-/-; Ezh2mz-/-
Allelic ratio
1.00 Maternal
0.50 Biallelic
1.00 Paternal
n = 8
n = 4
n = 4
n = 3
B
Ezh2fl/fl
Ezh2mz-/-; Ezh1mz-/-
# of X-linked Genes
% Xp Expression
% Xp Expression
C
Rnf12
Atrx
G6pdx
Pdha1
Maternal
Paternal
Avg. % Allelic Expression (Pyrosequencing)

Figure 3.6. Defective paternal X-linked gene silencing in embryos lacking both maternal and zygotic EZH1 and EZH2 (A) Allele-specific X-linked gene expression heat map of female *Ezh2*$^{fl/fl}$, *Ezh1*$^{mz-/-}$, *Ezh2*$^{mz-/-}$, and *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ blastocysts. Eight embryos of the *Ezh2*$^{fl/fl}$ genotype, four embryos of the *Ezh1*$^{mz-/-}$ and *Ezh2*$^{mz-/-}$ genotypes, and three embryos of the *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ genotype were sequenced individually and only genes with informative allelic expression in all samples are plotted (see Materials and Methods). *Xist* is not included in the heat map because informative allelic expression was not achieved for this gene in all samples sequenced. (D) Histograms depicting the average number of X-linked genes undergoing various degrees (in 10% increments) of expression from the paternal-X in *Ezh2*$^{fl/fl}$, *Ezh1*$^{mz-/-}$, *Ezh2*$^{mz-/-}$, and *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ blastocysts. (E) Pyrosequencing-based quantification of allelic expression of X-linked genes in *Ezh2*$^{fl/fl}$, *Ezh1*$^{mz-/-}$, *Ezh2*$^{mz-/-}$, and *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ blastocysts. Error bars represent the standard deviation of data from 3 to 6 independent blastocyst embryos. The mean allelic expression of *Atrx, G6pdx,* and *Pdha1* is not significantly different between *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos ($p > 0.05$, Welch's two-sample T-test). The mean allelic expression of *Rnf12* is significantly different between *Ezh2*$^{fl/fl}$ and *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos ($p < 0.05$, Welch's two-sample T-test).


A
Ezh2fl/fl
Ezh2mz-/-
Ezh2fl/fl; Ezh1mz-/-
Ezh2mz-/-; Ezh1mz-/-
Eedm-/-
Eedmz-/-
1.00 Maternal
0.50 Biallelic
1.00 Paternal
Allelic ratio
n = 8
n = 4
n = 4
n = 3
n = 5
n = 3


B
Ezh2fl/fl
Ezh2mz-/-
Ezh2fl/fl; Ezh1mz-/-
Ezh2mz-/-; Ezh1mz-/-
Eedm-/-
Eedmz-/-
# of X-linked Genes
% Xp Expression


C
X-linked Genes with
Unbiased Allelic Expression
118
30
4
Ezh2mz-/-;
Ezh1mz-/-
Eedmz-/-


D
X-linked Genes with
Unbiased Allelic Expression
90
57
8
Eedm-/-
Eedmz-/-

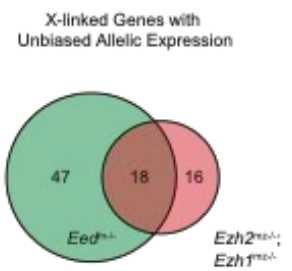

E
X-linked Genes with
Unbiased Allelic Expression
47
18
16
Eedm-/-
Ezh2mz-/-;
Ezh1mz-/-

Figure 3.7. Different paternal-X silencing profiles in embryos lacking maternal and zygotic EED and EZH1/2 (A) Allele-specific X-linked gene expression heat map of female *Ezh2*$^{fl/fl}$, *Ezh1*$^{mz-/-}$, *Ezh2*$^{mz-/-}$, *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$, *Eed*$^{m-/-}$, and *Eed*$^{mz-/-}$ blastocysts. Eight embryos of the *Ezh2*$^{fl/fl}$ genotype, five embryos of the *Eed*$^{m-/-}$ genotype, four embryos of the *Ezh1*$^{mz-/-}$ and *Ezh2*$^{mz-/-}$ genotypes, and three embryos of the *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ and *Eed*$^{mz-/-}$ genotype were sequenced individually and only genes with informative allelic expression in all samples are plotted (see Materials and Methods). *Xist* is not included in the heat map because informative allelic expression was not achieved for this gene in all samples sequenced. (B) Histograms depicting the average number of X-linked genes undergoing various degrees (in 10% increments) of expression from the paternal-X in *Ezh2*$^{fl/fl}$, *Ezh1*$^{mz-/-}$, *Ezh2*$^{mz-/-}$, *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$, *Eed*$^{m-/-}$, and *Eed*$^{mz-/-}$ blastocysts. (C-E) Comparison of X-linked genes with unbiased allelic expression in (C) *Eed*$^{mz-/-}$ vs. *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos, (D) *Eed*$^{mz-/-}$ vs. *Eed*$^{m-/-}$ embryos, and (E) *Eed*$^{m-/-}$ vs. *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ embryos.

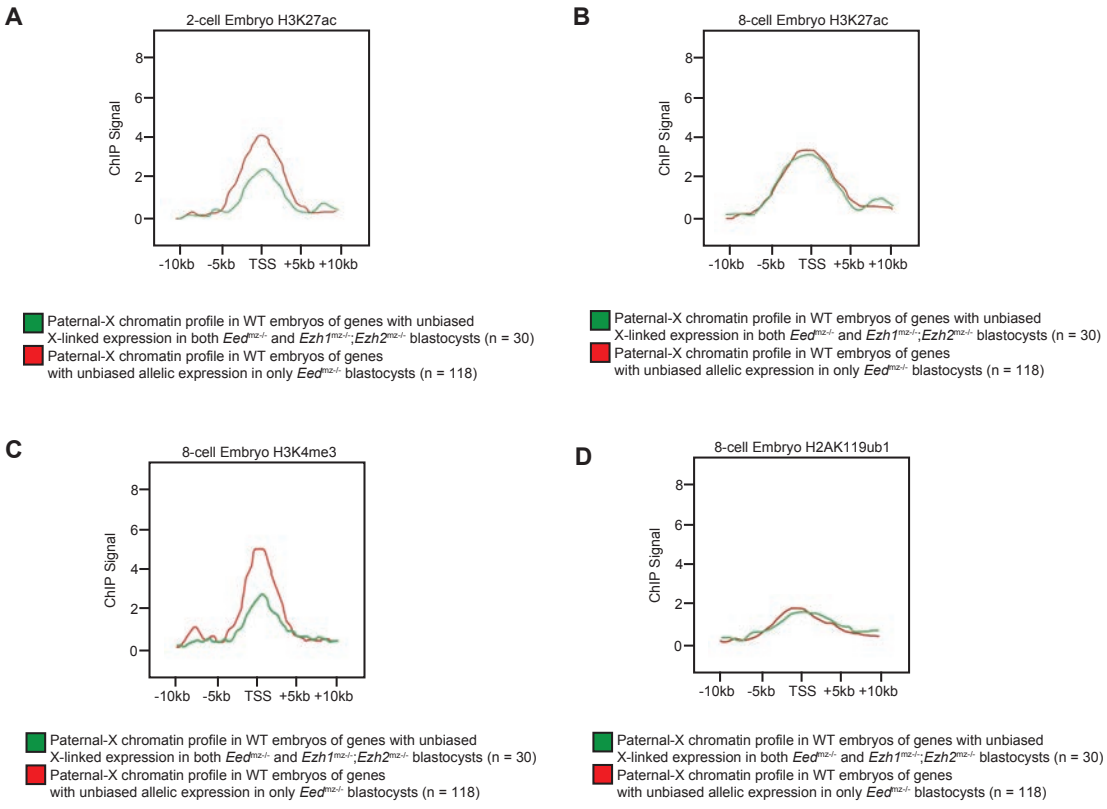
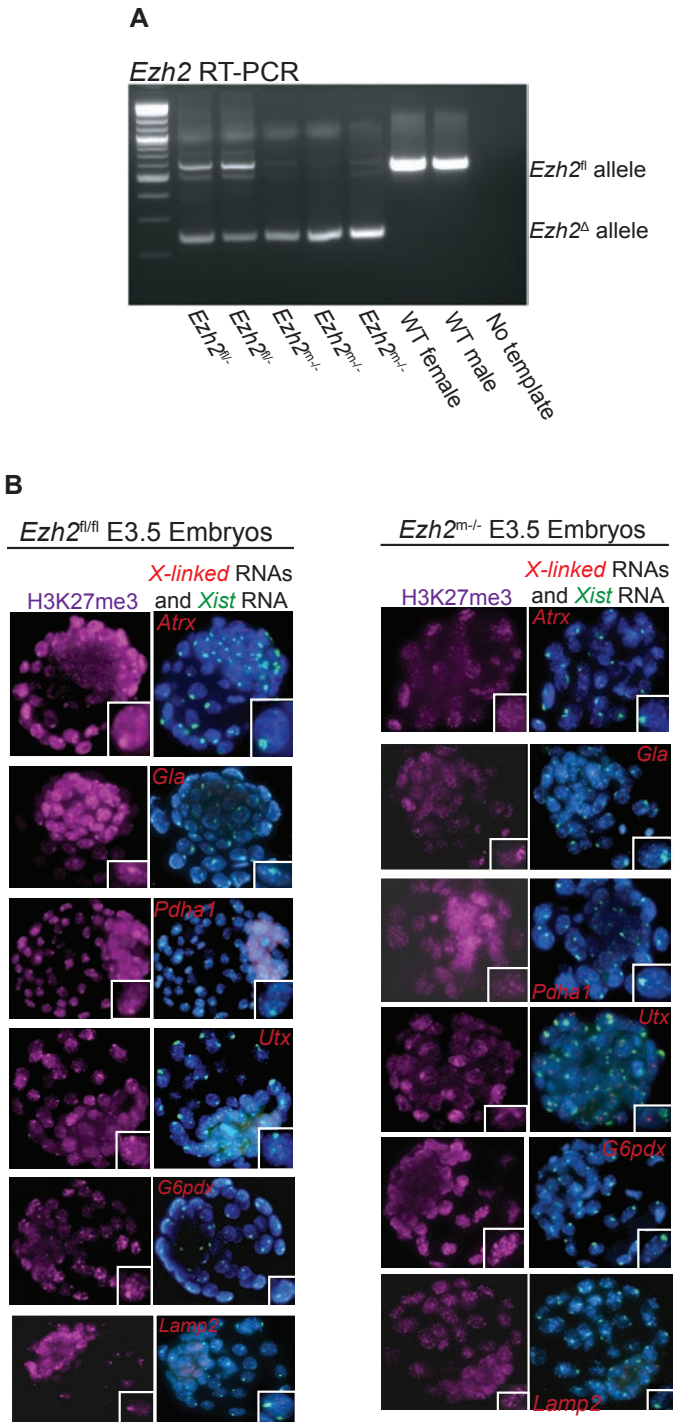
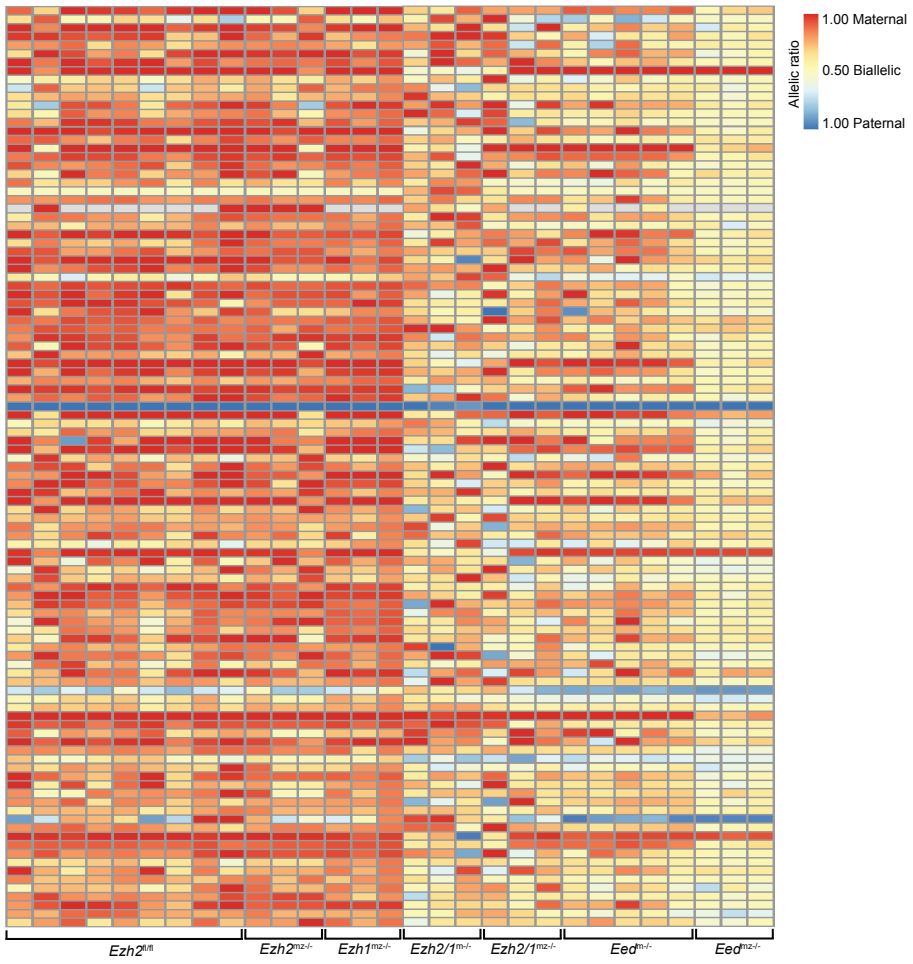

A
2-cell Embryo H3K27ac
ChIP Signal
8
6
4
2
0
-10kb -5kb TSS +5kb +10kb
Paternal-X chromatin profile in WT embryos of genes with unbiased X-linked expression in both Eed mz-/- and Ezh1 mz-/-;Ezh2 mz-/- blastocysts (n = 30)
Paternal-X chromatin profile in WT embryos of genes with unbiased allelic expression in only Eed mz-/- blastocysts (n = 118)
B
8-cell Embryo H3K27ac
ChIP Signal
8
6
4
2
0
-10kb -5kb TSS +5kb +10kb
Paternal-X chromatin profile in WT embryos of genes with unbiased X-linked expression in both Eed mz-/- and Ezh1 mz-/-;Ezh2 mz-/- blastocysts (n = 30)
Paternal-X chromatin profile in WT embryos of genes with unbiased allelic expression in only Eed mz-/- blastocysts (n = 118)
C
8-cell Embryo H3K4me3
ChIP Signal
8
6
4
2
0
-10kb -5kb TSS +5kb +10kb
Paternal-X chromatin profile in WT embryos of genes with unbiased X-linked expression in both Eed mz-/- and Ezh1 mz-/-;Ezh2 mz-/- blastocysts (n = 30)
Paternal-X chromatin profile in WT embryos of genes with unbiased allelic expression in only Eed mz-/- blastocysts (n = 118)
D
8-cell Embryo H2AK119ub1
ChIP Signal
8
6
4
2
0
-10kb -5kb TSS +5kb +10kb
Paternal-X chromatin profile in WT embryos of genes with unbiased X-linked expression in both Eed mz-/- and Ezh1 mz-/-;Ezh2 mz-/- blastocysts (n = 30)
Paternal-X chromatin profile in WT embryos of genes with unbiased allelic expression in only Eed mz-/- blastocysts (n = 118)
E
Ezh1 mz-/-; Ezh2 mz-/- and Eed m-/- Embryos
EED Sensitive Genes
EED
HDAC
X-linked Gene
H3K27ac
Eed mz-/- Embryos
EED Sensitive Genes
X-linked Gene
H3K27ac

Figure 3.8. Profiling paternal-X H3K27ac, H3K4me3, and H2AK119ub1 in preimplantation mouse embryos (A-B) Profiling H3K27ac in (A) 2-cell and (B) 8-cell stage mouse embryos. Green line shows the paternal-X chromatin profile in WT embryos of genes with unbiased X-linked expression in both *Eed*$^{mz-/-}$ and *Ezh1*$^{mz-/-}$; *Ezh2*$^{mz-/-}$ blastocysts. Red line shows paternal-X chromatin profile in WT embryos of genes with unbiased allelic expression in only *Eed*$^{mz-/-}$ blastocysts. (C) Profiling H3K4me3 in 8-cell stage mouse embryos. Green line shows the paternal-X chromatin profile in WT embryos of genes with unbiased X-linked expression in both *Eed*$^{mz-/-}$ and *Ezh1*$^{mz-/-}$; *Ezh2*$^{mz-/-}$ blastocysts. Red line shows paternal-X chromatin profile in WT embryos of genes with unbiased allelic expression in only *Eed*$^{mz-/-}$ blastocysts. (D) Profiling H2AK119ub1 in 8-cell stage mouse embryos. Green line shows the paternal-X chromatin profile in WT embryos of genes with unbiased X-linked expression in both *Eed*$^{mz-/-}$ and *Ezh1*$^{mz-/-}$; *Ezh2*$^{mz-/-}$ blastocysts. Red line shows paternal-X chromatin profile in WT embryos of genes with unbiased allelic expression in only *Eed*$^{mz-/-}$ blastocysts. (E) Model depicting the possible PRC2-independent role for EED in mediating histone deacetylation on the paternal-X in early female mouse embryos to silence some paternal X-linked genes.

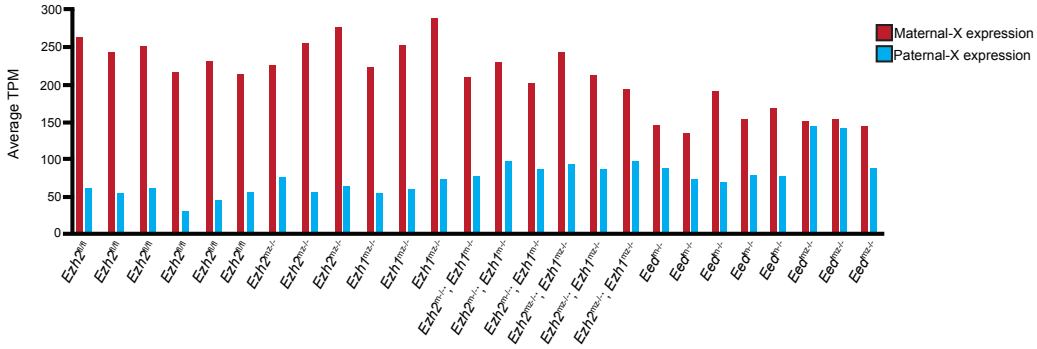


Supplemental Figure 3.1. Characterizing embryos lacking maternal EZH2 by PCR and IF-FISH (A) Representative gel indicating successful *Ezh2* deletion in *Ezh2*$^{m-/-}$ embryos. (B) Representative images of *Ezh2*$^{fl/fl}$ and *Ezh2*$^{m-/-}$ female mouse blastocysts stained to detect *Xist* RNA coating (in green), other X-linked RNAs (in red), and histone H3 lysine 27 tri-methylation (H3-K27me3; in purple). Nuclei re stained blue with DAPI. Quantification of these RNA FISH are included in Figure 3.1

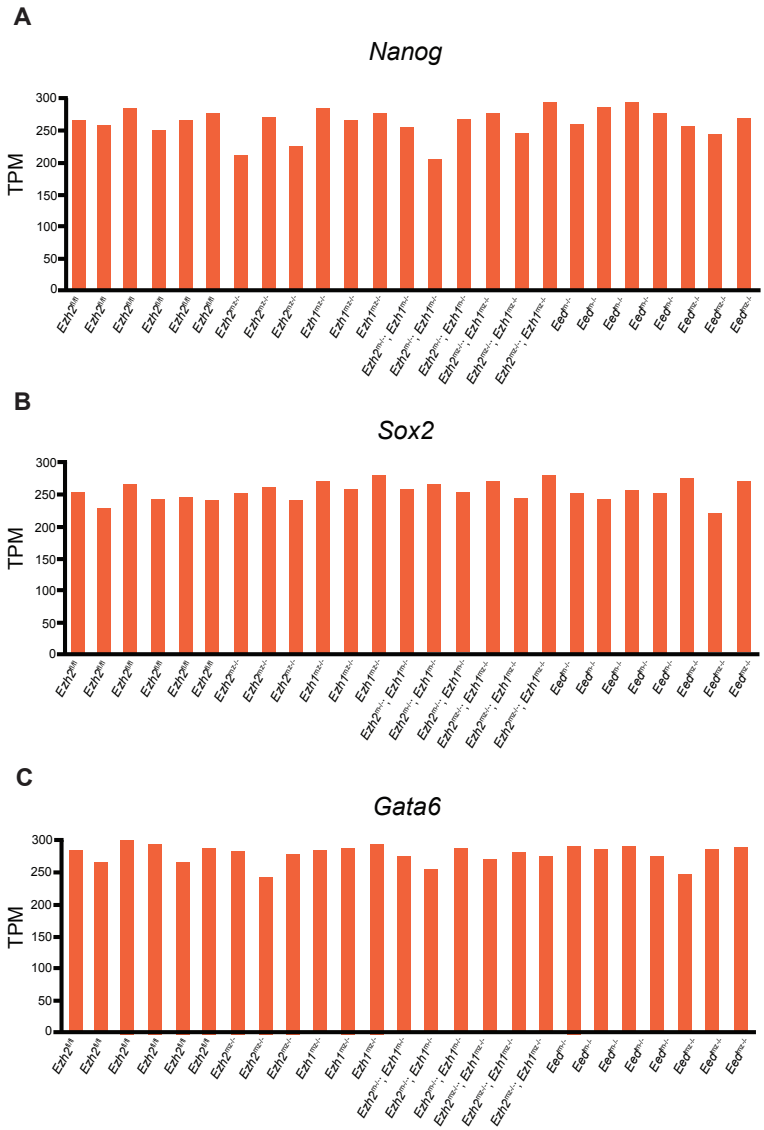
A


Individual E3.5 Female Embryos
1.00 Maternal
0.50 Biallelic
1.00 Paternal
Allelic ratio
Ezh2fl/fl
Ezh2mz-/-
Ezh1mz-/-
Ezh2/1m-/-
Ezh2/1mz-/-
Eedm-/-
Eedmz-/-

B


Maternal vs. Paternal X-linked Gene Expression
Average TPM
300
250
200
150
100
50
0
Maternal-X expression
Paternal-X expression
Ezh2fl/fl
Ezh2mz-/-
Ezh1mz-/-
Ezh2m-/-; Ezh1m-/-
Ezh2mz-/-; Ezh1m-/-
Ezh2m-/-; Ezh1mz-/-
Ezh2mz-/-; Ezh1mz-/-
Eedm-/-
Eedmz-/-

Supplemental Figure 3.2. Allelic expression profiles of individually sequenced embryos (A) Allele-specific X-linked gene expression heat map of individual female *Ezh2*$^{fl/fl}$, *Ezh1*$^{mz-/-}$, *Ezh2*$^{mz-/-}$, *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$, *Eed*$^{m-/-}$, and *Eed*$^{mz-/-}$ blastocysts. Eight embryos of the *Ezh2*$^{fl/fl}$ genotype, five embryos of the *Eed*$^{m-/-}$ genotype, four embryos of the *Ezh1*$^{mz-/-}$ and *Ezh2*$^{mz-/-}$ genotypes, and three embryos of the *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$ and *Eed*$^{mz-/-}$ genotype were sequenced individually and only genes with informative allelic expression in all samples are plotted (see Materials and Methods). *Xist* is not included in the heat map because informative allelic expression was not achieved for this gene in all samples sequenced. (B) Histogram showing the average maternal and paternal X-linked gene TPM in individual female *Ezh2*$^{fl/fl}$, *Ezh1*$^{mz-/-}$, *Ezh2*$^{mz-/-}$, *Ezh2*$^{mz-/-}$; *Ezh1*$^{mz-/-}$, *Eed*$^{m-/-}$, and *Eed*$^{mz-/-}$ blastocysts.



Supplemental Figure 3.3. Quantification of pluripotency factors by RNA-Seq (A-C) RNA-Seq quantification of (A) Nanog, (B) Sox2, and (C) Gata6 in individual embryos.

Chapter 4

Preventing Erosion of X-chromosome Inactivation in Human Embryonic Stem Cells

Note: This chapter was adopted from a published manuscript detailing the investigation of culture conditions that lead to X-chromosome inactivation erosion in hESCs:

Cloutier, M.*, Kumar, S.*, Buttigieg, E.*, Keller, L., Lee, B., Williams, A., Mojica-Perez, S., Erliandri, I., Monteiro Da Rocha, A., Cadigan, K., Smith, G. D., Kalantry, S. Preventing Erosion of X-chromosome Inactivation in Human Embryonic Stem Cells (2022). ***Nature Communications*** **13:2516.**

*Denotes equally contributing authors

Abstract

X-chromosome inactivation is a paradigm of epigenetic transcriptional regulation. Female human embryonic stem cells (hESCs) often undergo erosion of X-inactivation upon prolonged culture. Here, we investigate the sources of X-inactivation instability by deriving new primed pluripotent hESC lines. We find that culture media composition dramatically influenced the expression of XIST lncRNA, a key regulator of X-inactivation. hESCs cultured in a defined xenofree medium stably maintained XIST RNA expression and coating, whereas hESCs cultured in the widely-used mTeSR1 medium lost XIST RNA expression. We pinpointed lithium chloride in mTeSR1 as a cause of XIST RNA loss. The addition of lithium chloride or inhibitors of glycogen synthase kinase 3 (GSK-3) proteins that are targeted by lithium to the defined hESC culture medium impeded XIST RNA expression. GSK-3 inhibition in differentiating female mouse embryonic stem cells and epiblast stem cells also resulted in a loss of XIST RNA expression. Together, these data may reconcile observed variations in X-inactivation in hESCs and inform the faithful culture of pluripotent stem cells.

Introduction

Human pluripotent stem cells (hPSCs) offer the possibility to model early human development *in vitro* and are substrates for regenerative medicine (Nichols and Smith, 2009; Rossant and Tam, 2017; Thomson et al., 1998). The promise of hPSCs relies on their ability to faithfully maintain their epigenetic and transcriptional profiles in culture.

X-chromosome inactivation is a paradigm of epigenetic transcriptional regulation that equalizes X-linked gene expression between female and male mammals (Disteche, 2016; Morey, 2011; Plath *et al.*, 2002). Once inactivated, with a few key exceptions, replicated copies of the silenced X chromosome remain stably inactive in descendant cells (Lyon, 1961). X-inactivation is an experimentally tractable system to interrogate epigenetic transcriptional regulation because two equivalent X chromosomes become transcriptionally divergent, and these divergent transcriptional states are subsequently stably transmitted across mitotic cell division.

X-inactivation has been studied extensively in mouse embryos and mouse embryonic stem cells (mESCs). In the female preimplantation mouse embryo, all cells undergo imprinted inactivation of the paternal X-chromosome (Kalantry et al., 2009; Kay *et al.*, 1993; Takagi and Sasaki, 1975). At the peri-implantation stage, the inactivated paternal X-chromosome is reactivated in the pluripotent epiblast progenitor cells (Maclary *et al.*, 2014; Mak *et al.*, 2004). Conventionally cultured mESCs capture this transient population of pluripotent cells and harbor two active X-chromosomes (Rastan and Robertson, 1985). Upon differentiation, pluripotent mouse embryonic epiblast cells as well as mESCs inactivate either the maternal or the paternal X-chromosome in individual cells in a process termed random X-inactivation (Chen et al., 2016; Gayen *et al.*, 2015; Penny et al., 1996; Pintacuda and Cerase, 2015; Rastan and Robertson, 1985; Shiura and Abe, 2019).

In contrast to the mouse, human female preimplantation embryos do not undergo imprinted inactivation of the paternal X-chromosome (Mandal *et al.*, 2020; Moreira de Mello *et al.*, 2017; Okamoto *et al.*, 2011; Petropoulos *et al.*, 2016). Instead, both X chromosomes in female preimplantation human embryos appear to initiate some degree of silencing (Mandal *et al.*, 2020; Moreira de Mello *et al.*, 2017; Okamoto *et al.*, 2011; Petropoulos *et al.*, 2016). The X-linked long noncoding XIST RNA, which is a hallmark of the inactive X-chromosome and is required for stable X-inactivation in mice (Kalantry et al., 2009; Marahrens *et al.*, 1997), is expressed from and coats in *cis* both X-chromosomes in most cells of female human blastocyst-stage embryos (Moreira de Mello *et al.*, 2017; Okamoto *et al.*, 2011; Petropoulos *et al.*, 2016). XIST RNA coating in turn recruits a diverse array of proteins to the X chromosome that silence gene expression (Chu *et al.*, 2015; McHugh et al., 2015; Minajigi *et al.*, 2015; Moindrot *et al.*, 2015; Monfort *et al.*, 2015). Despite XIST RNA coating of the X-chromosomes, however, X-linked genes are not fully silenced in preimplantation human embryos (Moreira de Mello *et al.*, 2017; Okamoto *et al.*, 2011; Petropoulos *et al.*, 2016).

Unlike female mESCs, which harbor two active-Xs (Rastan and Robertson, 1985), female human ESC (hESC) lines display variable patterns of X-inactivation (An *et al.*, 2020; Anguera *et al.*, 2012; Bar et al., 2019; Barakat *et al.*, 2015; Guo *et al.*, 2017; Hall *et al.*, 2008; Hoffman *et al.*, 2005; Kim *et al.*, 2014; Lengner *et al.*, 2010; Mandal *et al.*, 2020; Mekhoubad *et al.*, 2012; Nazor *et al.*, 2012; Patel *et al.*, 2017; Pomp *et al.*, 2011; Sahakyan *et al.*, 2017a; Shen *et al.*, 2008b; Silva *et al.*, 2008; Tchieu *et al.*, 2010; Tomoda *et al.*, 2012; Vallot *et al.*, 2015; Vallot *et al.*, 2017; Xie *et al.*, 2016). This variability appears to reflect both differences in X-inactivation patterns in early mouse vs. human female embryos and how hESCs are derived and cultured.

The pattern of X-inactivation in early preimplantation human embryos is partially recapitulated *in vitro* through the derivation of 'naïve' pluripotent female hESCs (An *et al.*, 2020; Guo *et al.*, 2017; Mandal *et al.*, 2020; Messmer *et al.*, 2019; Sahakyan *et al.*, 2017a; Vallot *et al.*, 2017). In these naïve hESC lines, a proportion of cells display XIST RNA coating of both X chromosomes but do not appear to transcriptionally inactivate the XIST RNA-coated Xs, like cells in early female human embryos. Most female hESCs cultured in naïve conditions, however, harbor one XIST RNA-coated X chromosome that is transcriptionally active (Guo *et al.*, 2017; Mandal *et al.*, 2020; Messmer *et al.*, 2019; Sahakyan *et al.*, 2017a; Vallot *et al.*, 2017). The heterogeneity of XIST RNA expression in naïve female hESCs appears to be due to the coexistence in culture of at least two populations of pluripotent cells (An *et al.*, 2020). Blocking autocrine bFGF signaling reduces this heterogeneity and is reported to yield nearly all hESCs with two XIST RNA-coated X chromosomes (An *et al.*, 2020), recapitulating the pattern observed in epiblast cells of preimplantation female human embryos (Okamoto *et al.*, 2011; Petropoulos *et al.*, 2016). Differentiation of these naïve hESCs into the 'primed' pluripotent hESCs results in most of the cells undergoing random X-inactivation and exhibiting XIST RNA coating of a single X chromosome that is transcriptionally inactive (An *et al.*, 2020).

Compared to naïve hESCs, primed pluripotent hESCs capture a later stage of embryonic development and may be analogous to mouse epiblast stem cells (mEpiSCs) (Brons et al., 2007; Hanna et al., 2010; Nichols and Smith, 2009; Stadtfeld and Hochedlinger, 2010; Takahashi et al., 2018; Tesar et al., 2007; Weinberger et al., 2016). Female mEpiSCs contain one inactivated, XIST RNA-coated X chromosome (Gayen *et al.*, 2015). Primed female hESCs, by contrast, exhibit at least three patterns of X-inactivation. Primed female hESCs can harbor no inactive-X chromosome, one inactive-X, or a leaky inactive-X (Anguera *et al.*, 2012; Barakat *et al.*, 2015;

Hall *et al.*, 2008; Hoffman *et al.*, 2005; Kim *et al.*, 2014; Lengner *et al.*, 2010; Mekhoubad *et al.*, 2012; Nazor *et al.*, 2012; Patel *et al.*, 2017; Pomp *et al.*, 2011; Shen *et al.*, 2008b; Silva *et al.*, 2008; Tchieu *et al.*, 2010; Tomoda *et al.*, 2012; Vallot *et al.*, 2015; Xie *et al.*, 2016). Upon prolonged culture, many primed female hESC lines lose XIST RNA coating, which is accompanied by the expression of a subset of previously silenced genes from the inactive X-chromosome (Anguera *et al.*, 2012; Bar and Benvenisty, 2019; Dvash *et al.*, 2010; Hall *et al.*, 2008; Mekhoubad *et al.*, 2012; Nazor *et al.*, 2012; Patel *et al.*, 2017; Shen *et al.*, 2008b; Silva *et al.*, 2008; Xie *et al.*, 2016). This loss of XIST RNA coating and re-expression of silenced X-linked genes has been termed X-inactivation erosion (Dvash *et al.*, 2010; Hall *et al.*, 2008; Lengner *et al.*, 2010; Mekhoubad *et al.*, 2012; Shen *et al.*, 2008b; Vallot *et al.*, 2015; Xie *et al.*, 2016). Loss of XIST RNA coating and the leaky expression of inactive X-linked genes also characterize cultured human induced pluripotent stem cells (hiPSCs) and are thought to be irreversible (Bar and Benvenisty, 2019; Fukuda et al., 2021; Mekhoubad *et al.*, 2012; Nazor *et al.*, 2012; Patel *et al.*, 2017; Vallot *et al.*, 2015). Increased X-linked gene expression, due to X-inactivation erosion or failure, can be deleterious to development and differentiation (Kalantry *et al.*, 2006; Marahrens *et al.*, 1997; Patel *et al.*, 2017). The instability of the epigenetically inactivated X-chromosome lends caution to the use of pluripotent female human cells in disease modeling and regenerative medicine.

The underlying reasons for XIST RNA loss and X-inactivation instability in cultured hESCs are unclear. To gain insights into changes in X-inactivation in hESCs, we derived new hESC lines and analyzed expression of XIST RNA and other X-linked genes in these hESCs under diverse culture conditions and across many passages. Our results implicate the presence of

lithium chloride or other GSK-3 inhibitors in the culture media as a cause of XIST RNA expression loss in female hESCs.

Loss of XIST RNA Coating in Cultured hESCs

To test the kinetics of X-inactivation in hESCs, we derived and characterized a series of new female hESC lines under primed pluripotency conditions on human fibroblast feeder cells (HFFs) (Materials and Methods; Figure 4.1A; Supplemental Figure 4.1). Upon passaging via mechanical splitting of the hESC colonies, we performed RNA fluorescence *in situ* hybridization (FISH) to capture XIST RNA coating at each passage in these cells. RNA FISH permits visualization of XIST RNA coating and nascent X-linked gene expression in nuclei of individual cells (Cloutier *et al.*, 2018). To minimize bias due to clonal expansion of the mechanically passaged hESCs, we quantified and stratified the percentage of nuclei with XIST RNA coats on a per colony basis with a minimum of 100 nuclei counted per colony.

After derivation on HFFs, we passaged the first hESC line analyzed, UM33-4, onto a Matrigel-coated surface and grew the cells at atmospheric oxygen concentration (20%) in mTeSR1 medium. Matrigel is a widely used extracellular matrix substrate that bypasses the need for feeder cells in culturing hESCs and mTeSR media are commonly used to culture hESCs on Matrigel (Hey et al., 2018; Hughes et al., 2010; International Stem Cell Initiative et al., 2010; Ludwig et al., 2006; Nengqing et al., 2020). In the UM33-4 hESCs, we quantified the percentage of nuclei with XIST RNA coats per colony at each passage (P) from P1 to P29 (Figure 4.1B-D). From P1-7, all nuclei in all colonies harbored single XIST RNA coats. From P8 onwards, however, the percentage of XIST RNA coated nuclei per colony significantly decreased (general linear model, $p = 0.002$) and at P27 XIST RNA coating disappeared altogether in all nuclei in all colonies. During subsequent passaging, hESC line UM33-4 stably

maintained the complete absence of nuclei with XIST RNA coats. Of note, nuclei lacking XIST RNA coats were devoid of any XIST RNA FISH signals, suggesting that a lack of XIST RNA coating reflected an absence of XIST RNA expression since the RNA FISH assay detects stabilized as well as nascent RNAs (Cloutier *et al.*, 2018).

XIST RNA Coating in hESCs Cultured in Atmospheric vs. Physiological O_2 Concentration

To investigate the underlying causes of XIST RNA coating loss in hESCs, we next cultured hESCs in atmospheric (20%) and physiological (5%) O_2 concentrations. Atmospheric O_2 has been reported to cause loss of XIST RNA expression in cultured female hESCs in some studies (An *et al.*, 2020; Lengner *et al.*, 2010; O'Leary et al., 2012; Vallot *et al.*, 2015) but not others (Anguera *et al.*, 2012; de Oliveira Georges et al., 2014; Diaz Perez et al., 2012; Mekhoubad *et al.*, 2012; Patel *et al.*, 2017; Theunissen et al., 2014; Xie *et al.*, 2016). We derived two new female hESC lines on HFFs, UM63-1 and UM77-2, and passaged them from P5 to P14 onto Matrigel-coated surfaces under atmospheric and physiological O_2 concentrations in parallel. We quantified the proportion of XIST RNA-coated nuclei per colony in the UM63-1 and UM77-2 hESCs from P5 to P14 since UM33-4 hESCs in Figure 4.1 lost XIST RNA coating in a significant proportion of nuclei per colony in these passages. Under atmospheric O_2 culture conditions, both the UM63-1 and UM77-2 hESC lines displayed decreasing percentages of nuclei with XIST RNA coats per colony during passaging, recapitulating the pattern observed in the UM33-4 hESC line (Figure 4.2A-B). When cultured at physiological O_2 concentration, the UM63-1 and UM77-2 hESC lines exhibited a similar decrease in the proportion of XIST RNA-coated nuclei per colony with successive passaging (Fig. 4.2C-D). The difference in the frequency of nuclei per colony without XIST RNA coating in either cell line when cultured in physiological vs. atmospheric O_2 levels was not significant (general linear model, $p = 0.1$).

Together with XIST RNA coating, in the hESCs we assayed the expression of two additional X-linked genes, *ATRX* and *USP9X*, by RNA FISH. Under both O_2 concentrations, nuclei that had lost XIST RNA-coating maintained silencing of one *ATRX* allele but displayed increased frequency of biallelic expression of *USP9X* across passaging (Supplemental Figure 4.2). These results led us to conclude that culturing primed pluripotent hESCs in atmospheric vs. physiological O_2 concentration does not alter the loss of XIST RNA expression and coating in hESCs, in agreement with other reports (Anguera *et al.*, 2012; de Oliveira Georges *et al.*, 2014; Diaz Perez *et al.*, 2012; Mekhoubad *et al.*, 2012; Patel *et al.*, 2017; Theunissen *et al.*, 2014; Xie *et al.*, 2016). Furthermore, some X-linked genes become derepressed when XIST RNA is lost in cultured hESCs, also in agreement with previously published results (Dvash *et al.*, 2010; Hall *et al.*, 2008; Lengner *et al.*, 2010; Mekhoubad *et al.*, 2012; Shen *et al.*, 2008b; Vallot *et al.*, 2015; Xie *et al.*, 2016).

Impact of hESC Culture Surface on XIST RNA Coating and Expression

We next sought to examine other variables that could explain the loss of XIST RNA coating in cultured hESCs. After their derivation and initial culture on HFFs, the hESC lines in Figures 4.1 and 4.2, UM33-4, UM63-1, and UM77-2, were transitioned to culture and passaging on a Matrigel-coated surface. The loss of XIST RNA coating coincided with continued passaging of the hESCs on Matrigel (Figures 4.1 and 4.2). We therefore examined whether the surface on which hESCs are cultured could contribute to X-inactivation erosion. In addition to Matrigel, hESCs can also be propagated on HFFs. We therefore tested whether continued culturing of hESCs on HFFs would maintain XIST RNA expression and coating. At nearly all passages, strikingly >80% of the nuclei in nearly all colonies of the UM63-1 and UM77-2 hESC lines cultured on HFFs exhibited XIST RNA coating from P1-14 (Figure 4.3). By contrast, as shown

in Figure 4.2, UM63-1 and UM77-2 hESCs cultured on Matrigel displayed an increasing percentage of colonies without any XIST RNA-coated nuclei through passaging (general linear model comparison, $p < 0.001$).

Analysis of hESC Culture Media Effect on XIST RNA Coating

hESCs grown on Matrigel vs. HFFs differ both in the culture surface as well as in the culture media used. hESCs grown on Matrigel are typically cultured in mTeSR media (Hughes *et al.*, 2010), (Hey *et al.*, 2018; International Stem Cell Initiative *et al.*, 2010; Ludwig *et al.*, 2006; Nengqing *et al.*, 2020), whereas our hESC lines grown on HFFs were cultured with a XenoFree (XF) medium (see Materials and Methods for detailed culture media composition). We therefore investigated whether culture media could underlie the differences in XIST RNA coating observed in hESCs cultured on Matrigel vs. HFFs. We cultured hESC lines UM63-1 and UM77-2 on HFFs in mTeSR1 medium and in XF medium in parallel. Unexpectedly, we found that culturing hESCs on HFFs with XF medium stably maintained nuclei with XIST RNA coating, whereas hESCs cultured on HFFs with mTeSR1 medium displayed a significant decrease in the proportion of nuclei with XIST RNA coating per colony during equivalent passaging (general linear model comparison, $p < 0.001$) (Figure 4.4). We could not culture hESCs on Matrigel with XF medium, as the hESCs failed to grow and attach to the Matrigel surface under these conditions.

To further investigate the effect of culture media on XIST RNA coating in hESCs, we next asked whether XIST RNA coating patterns would change if hESCs cultured on HFFs in mTeSR1 medium were switched onto XF medium and *vice versa*. We therefore cultured the hESC line UM63-1 on HFFs initially in XF medium and then switched the culture medium to mTeSR1 medium (Supplemental Figure 4.3). When cultured in the XF medium, UM63-1

displayed a small percentage of colonies without any XIST RNA-coated nuclei (Figure 4.5A). This slight loss of XIST RNA coating may be due to a freeze-thaw cycle prior to culture, which has previously been linked to loss of XIST RNA coating in hESCs (Lengner *et al.*, 2010). Whereas the two hESC lines maintained XIST RNA coating when cultured in the XF medium, both hESC lines displayed a progressive decrease in the percentage of XIST RNA coated nuclei per colony after being switched to culturing on mTeSR1 medium (general linear model, $p = 0.01$) (Fig. 4.5A-B). Figure 4.5 displays the three most informative categories of percentage of XIST RNA coated nuclei per colony (80-100%; 20-79%; 0-19%) (see Supplementary Figure 4.4 for further stratification of the data).

When hESCs were cultured on HFFs first in mTeSR1 medium and then switched to XF medium, the frequency of XIST RNA-coated nuclei per colony stabilized (Figure 4.5C-D; Supplementary Figure 4.4). Although their proportion did not decrease significantly, XIST RNA-coated nuclei per colony also did not increase in frequency across passaging when hESCs were switched from mTeSR1 to XF medium (general linear model, $p = 0.1$), suggesting that loss of XIST RNA expression is irreversible. From these findings, we conclude that culturing hESCs in XF medium stably maintains XIST RNA coating and that culturing hESCs in mTeSR1 medium causes an irreversible loss of XIST RNA expression.

Lithium Chloride in mTeSR1 Medium as a Cause of XIST RNA Loss

To determine the underlying reasons for the loss of XIST RNA coating in hESCs cultured in mTeSR1 vs. XF media, we compared the chemical composition of the two media (Materials and Methods) (International Stem Cell Initiative *et al.*, 2010; Ludwig *et al.*, 2006). We noted that whereas mTeSR1 contains lithium chloride (LiCl), XF medium does not. Lithium is known to intersect with a number of intracellular signaling pathways (Jakobsson et al., 2017), including,

importantly, with the Wnt pathway by inhibiting GSK-3 proteins (Freland and Beaulieu, 2012; Klein and Melton, 1996; Stambolic et al., 1996). We thus tested if addition of LiCl to XF medium (XF with LiCl) at the concentration present in mTeSR1 (0.98mM) would result in loss of XIST RNA coating. We derived an independent hESC line, UM90-14, and cultured the cells in parallel in XF, mTeSR1, and XF with LiCl media on HFFs under physiological oxygen concentration (5% O_2). As in Figure 4.5, we plotted the three most informative categories of percentage of XIST RNA-coated nuclei per colony (80-100%; 20-79%; 0-19%). As expected, hESCs cultured in XF medium did not exhibit a significant decrease in the percentage of XIST RNA coated nuclei per colony across passaging (general linear model, $p = 0.3$) (Figure 4.6A). hESCs cultured in mTeSR1 medium displayed a significant reduction in the percentage of XIST RNA coated nuclei per colony across passaging (general linear model, $p < 0.001$) (Figure 4.6B). Strikingly, hESCs cultured in the XF medium with LiCl also lost XIST RNA coating during passaging in a significant percentage of nuclei per colony compared to cells cultured in XF medium (general linear model, $p < 0.001$) (Figure 4.6C).

Stability of XIST RNA Coating in Differentiated hESCs

The erosion of X-inactivation in cultured hESCs compelled the examination of X-inactivation stability during the differentiation of female hESCs. X-inactivation is reported to be stable upon differentiation of human PSCs in some (Patel *et al.*, 2017) but not other studies (Geens et al., 2016; Hall *et al.*, 2008; Hoffman *et al.*, 2005; Silva *et al.*, 2008; Tomoda *et al.*, 2012; Vallot *et al.*, 2015). We assayed XIST RNA coating in hESCs differentiated into embryoid bodies (EBs) with a commercial medium, AggreWell™ that is based on the mTeSR1 medium and contains LiCl (Ludwig *et al.*, 2006), and with XF medium lacking the pluripotency promoting factor basic FGF (bFGF). bFGF is widely used as a supplement to promote growth of

stem cells in vitro, including in the mTeSR1 and XF media used in this study (Xu et al., 2005). XIST RNA coating decreased in a significant proportion of nuclei in embryoid bodies differentiated with the AggreWell™ medium (general linear model, $p < 0.001$) but not with the XF medium lacking bFGF (general linear model, $p = 0.3$). We next assayed XIST RNA coating in hESCs differentiated into embryoid bodies in XF medium lacking bFGF but supplemented with LiCl at the concentration present in mTeSR1 (0.98 mM) and again observed loss of XIST RNA coating in a significant proportion of nuclei (general linear model, $p < 0.01$) (Figure 4.7). These data suggest that media containing LiCl causes loss of XIST RNA expression during the differentiation of hESCs, like during the culture of undifferentiated hESCs.

GSK-3 Inhibition and Loss of XIST RNA Coating in hESCs

A prominent mode of action of lithium is inhibition of the GSK-3 pathway (Freland and Beaulieu, 2012; Jakobsson *et al.*, 2017). GSK-3 is a multifaceted kinase with over 100 known substrates and exists as two common isoforms, GSK-3α and GSK-3β (Beurel et al., 2015; Embi et al., 1980; Grimes and Jope, 2001). GSK-3β function appears to overlap with that of GSK-3α (Doble et al., 2007), suggesting partial functional redundancy of the two proteins. A major function of GSK-3 is the negative regulation of β-catenin, which is a key mediator of the canonical Wnt signaling pathway (Behrens et al., 1998). Wnt signaling, in turn, plays an important role in hESC proliferation and cell survival (Doble and Woodgett, 2003; Patel et al., 2004; Singh et al., 2012). Based on the ability of lithium to inhibit the GSK-3 proteins and the loss of XIST RNA expression in LiCl-treated cells, we focused our subsequent analyses on the effects of specific GSK-3 inhibitors on XIST RNA expression in hESCs.

We tested the effects of three GSK-3 inhibitors, Alsterpaullone, BIO, and Ly2090314, on XIST RNA coating and X-linked gene expression in hESC line UM90-14. Alsterpaullone is an

ATP-competitive inhibitor of GSK-3α/β proteins (Leost et al., 2000). BIO is a highly potent, selective, and reversible ATP-competitive inhibitor of GSK-3α/β (Sato et al., 2004). Ly2090314 is a small-molecule inhibitor of GSK-3α/β isoforms (Eldar-Finkelman and Martinez, 2011). We cultured the hESC line UM90-14 for 10 passages (P6-P15) in XF medium supplemented with each of the three GSK-3 inhibitors individually. The 90-14 hESC line at P6 exhibited a slightly reduced frequency of nuclei with XIST RNA coating compared to earlier experiments (Figure 4.8), due possibly to the cells having been frozen and thawed prior to this series of hESC cultures (Lengner *et al.*, 2010). The addition of each of the three GSK-3 inhibitors to the XF medium resulted in a significant reduction in nuclei with XIST RNA coating per colony upon continued passaging compared to those cultured in XF medium alone (general linear model comparison, $p < 0.001$), like hESCs cultured in XF with LiCl and mTeSR1 media (Figure 4.8). A comparison of the transcriptome by RNA sequencing (RNA-Seq) of the hESCs cultured in the different media formulations above suggests that hESCs cultured in the different culture conditions above are similar but not identical to one another and to other hESCs and human epiblast cells and distinct from human somatic cell types (Supplemental Figure 4.5). Differential expression analysis of the RNA-Seq data revealed no significantly differentially expressed genes after adjusting for multiple testing, although this is likely due to sequencing only one replicate per culture condition.

GSK-3 Inhibition and Loss of XIST RNA Coating in Differentiating mESCs

We next tested if GSK-3 inhibition would lead to a loss of Xist RNA expression in differentiating female mouse (m) ESCs. Female mESCs harbor two active X-chromosomes and undergo stochastic inactivation of one of the two Xs upon differentiation (Samanta and Kalantry, 2020). We differentiated the mESCs into primed pluripotent epiblast-like cells (mEpiLCs)

(Samanta and Kalantry, 2020) (Materials and Methods). We then cultured the mEpiLCs in the presence of the GSK-3 inhibitor CHIR99021 (CHIR) for 48 hrs and assessed Xist RNA coating. The mEpiLCs cultured without CHIR stably maintained Xist RNA coating whereas mEpiLCs cultured with CHIR lost a significant percent of Xist RNA coating in a significant number of nuclei per colony (general linear model comparison, $p < 0.001$) (Figure 4.9). These data suggest that the effect of GSK-3 inhibition on XIST RNA expression is conserved between hESCs and mESCs.

GSK-3 Inhibition and Loss of XIST RNA Coating in mEpiSCs

We next assessed the effects of GSK-3 inhibition on *Xist* RNA expression in mouse epiblast stem cells (mEpiSCs). mEpiSCs are primed pluripotent stem cells that are analogous to primed pluripotent hESCs (Brons *et al.*, 2007; Tesar *et al.*, 2007). Like primed hESCs, mEpiSCs harbor a randomly inactivated X chromosome (Gayen *et al.*, 2015). We found that a significant number of mEpiSCs also lost *Xist* RNA coating upon culture in media supplemented with CHIR (general linear model, $p < 0.001$) (Figure 4.10A-C).

Conserved TCF Binding Sites Upstream of Human and Mouse *XIST*/*Xist*

The effect of GSK-3 inhibition on XIST/*Xist* RNA coating in PSCs may be due directly or indirectly to the activation of Wnt signaling. Canonical Wnt signaling regulates transcription through the binding of T-cell factor (TCF) and β-catenin to regulatory sequences of the target genes (Behrens *et al.*, 1998; Blauwkamp et al., 2008; Zhang et al., 2014). The activation of Wnt signaling is conventionally thought to only activate transcription (de Jaime-Soguero et al., 2018), but recent studies suggest that Wnt signaling may also repress the expression of some target genes (Blauwkamp *et al.*, 2008; Zhang *et al.*, 2014). We therefore examined human and mouse genomic sequence 5' of the XIST/*Xist* transcription start sites (TSSs) and found three conserved

putative TCF binding motifs within this region (Figure 4.10D). We also found that sequences surrounding these motifs are also conserved between humans and mice, suggesting that these sites may serve as platforms for the binding of other transcription factors that may in turn contribute to the silencing of XIST/*Xist* in the two species.

Discussion

PSCs are a model to investigate epigenetic mechanisms that underlie cell fate transitions. X-inactivation is an experimentally tractable system to dissect epigenetic transcriptional regulation in PSCs. Much prior work has demonstrated that prolonged culture of female hPSCs results in loss of XIST RNA coating and erosion of X-inactivation (Anguera *et al.*, 2012; Bar and Benvenisty, 2019; Fukuda *et al.*, 2021; Hall *et al.*, 2008; Mekhoubad *et al.*, 2012; Nazor *et al.*, 2012; Patel *et al.*, 2017; Shen *et al.*, 2008b; Silva *et al.*, 2008; Vallot *et al.*, 2015; Xie *et al.*, 2016). Consistent with previous findings, our data rule out atmospheric O_2 concentration as a cause of XIST RNA loss in hESCs (de Oliveira Georges *et al.*, 2014; Diaz Perez *et al.*, 2012; Mekhoubad *et al.*, 2012; Patel *et al.*, 2017; Theunissen *et al.*, 2014; Xie *et al.*, 2016). Instead, we found that a primary source of XIST RNA loss in hESCs is the composition of the hESC culture medium. The loss of XIST RNA expression in the hESCs is irreversible, in agreement with prior findings (Mekhoubad *et al.*, 2012; Shen *et al.*, 2008b; Silva *et al.*, 2008). Of note, a recent report suggests that DNA methylation may contribute to the irreversibility of XIST RNA loss in cultured female hESCs that have undergone X-inactivation erosion (Fukuda *et al.*, 2021).

Female hESCs lacking XIST RNA coating have been suggested to expand in culture due to a proliferation advantage compared to female hESCs with XIST RNA coating (Anguera *et al.*, 2012). Our data, though, argue against this possibility because non-XIST RNA coated hESCs did not increase in proportion when hESCs cultured in mTeSR1 medium were switched to

culture with XF medium, which stably maintained but did not lead to increased frequency of nuclei with XIST RNA coating. Instead, our data suggest that LiCl in hESC culture medium can actively cause loss of XIST RNA expression. The loss of XIST RNA coating, however, did not appear to result in chromosome-wide reactivation of all silenced X-linked genes. XIST RNA loss in the hESCs, though, has been reported to result in the reactivation of a subset of silenced X-linked genes (Adewumi et al., 2007; Mandal *et al.*, 2020; Patel *et al.*, 2017; Patrat et al., 2020; Shen *et al.*, 2008b; Vallot *et al.*, 2017), consistent with our findings. Through increased expression of X-linked genes, the erosion of X-inactivation can negatively impact hESC physiology, for example through compromised differentiation (Patel *et al.*, 2017).

One factor suggested to cause loss of XIST RNA expression in hESCs is the X-linked XACT RNA. XACT RNA is expressed from and coats the active X-chromosome at the onset of X-inactivation (Vallot et al., 2013; Vallot *et al.*, 2015; Vallot *et al.*, 2017). *XACT* expression also precedes loss of XIST RNA coating from the inactive-X during prolonged hESC culture (Sahakyan *et al.*, 2017a; Sahakyan et al., 2017b; Vallot *et al.*, 2015). An analysis of published female hESC RNA-Seq data, however, suggests that *XACT* expression is likely to be transient in cultured hESCs (Guo *et al.*, 2017; Sun et al., 2018; Syrett et al., 2018).

That culture media components like LiCl can alter X-inactivation counters the conventional notion that X-inactivation is a cell autonomous process that is immune to extracellular influences. Lithium interferes with a number of cell signaling pathways, most prominently through the inhibition of GSK-3 proteins (Freland and Beaulieu, 2012; Jakobsson *et al.*, 2017). The inclusion of GSK-3 inhibitors in hESC culture media recapitulated the loss of XIST RNA coating observed with the addition of LiCl to the culture medium, suggesting that

inhibition of GSK-3 proteins by lithium may be responsible for loss of XIST RNA coating in hESCs cultured in mTeSR1 medium.

A primary mode of action of GSK-3 is the negative regulation of the canonical Wnt signaling pathway (Doble and Woodgett, 2003; Patel *et al.*, 2004; Singh *et al.*, 2012). GSK-3 phosphorylates TCF/LEF coactivator β-catenin, which prevents β-catenin from entering nucleus (Cadigan and Nusse, 1997). GSK3 inhibition permits β-catenin to enter the nucleus where it associates with TCF/LEF transcription factors to activate gene expression (Schaefer and Peifer, 2019). But, recent work has found that Wnt signaling can also repress gene expression in some contexts (Blauwkamp *et al.*, 2008; Zhang *et al.*, 2014). TCF binding at target sites together with other cofactors that silence gene expression may repress transcription of Wnt targets (Blauwkamp *et al.*, 2008). The putative TCF binding sites and conserved surrounding sequence upstream of *XIST* suggests that TCF together with other cofactors that may directly repress *XIST* in PSCs. It's also possible that GSK-3β and other components of the Wnt signaling pathway indirectly repress *XIST* by inducing the expression of genes that in turn silence *XIST* or by crosstalk with other signaling pathways that silence *XIST* (An *et al.*, 2020; Del Rosario et al., 2017; Sripathy et al., 2017). Because the cell types examined in this work all exhibit stable maintenance of X-inactivation under normal conditions, our findings suggest a role for Wnt signaling in the maintenance of X-inactivation in PSCs. Furthermore, given the nearly ubiquitous nature of Wnt signaling, there are likely other compensatory mechanisms (additional pathways, Wnt signaling dose, etc.) that allow for X-inactivation maintenance in female adult somatic cells. Future work will dissect the direct vs. indirect regulation of XIST/*Xist* RNA expression by GSK-3 and Wnt signaling in PSCs.

Naïve hESCs are cultured with a cocktail of inhibitors that target components of several cellular signaling pathways, including Wnt signaling via the inhibition of GSK-3 (Gafni et al., 2013; Takashima et al., 2014; Ware *et al.*, 2014). GSK-3 inhibition and Wnt activation are both suggested to promote naïve hESC pluripotency (de Jaime-Soguero *et al.*, 2018; Guo *et al.*, 2017; Singh *et al.*, 2012; Theunissen *et al.*, 2014; Xu et al., 2016). In naïve hESCs, XIST RNA can be expressed from and coat both X chromosomes (Sahakyan *et al.*, 2017a). However, the XIST RNA coats in naïve hESCs appear more diffuse (An *et al.*, 2020), suggesting decreased *XIST* expression compared to that in primed pluripotent hESCs. This lower XIST RNA expression in naïve hESCs may potentially be due to the inclusion of GSK-3 inhibitors in naïve hESC culture media.

Although human and mouse embryos both undergo X-chromosome inactivation, the dynamics of X-inactivation differ during human and mouse embryonic development (Kalantry et al., 2009; Kay *et al.*, 1993; Maclary *et al.*, 2014; Mak *et al.*, 2004; Mandal *et al.*, 2020; Moreira de Mello *et al.*, 2017; Okamoto *et al.*, 2011; Petropoulos *et al.*, 2016; Takagi and Sasaki, 1975). Notably, XIST appears to be less crucial to the maintenance of X-inactivation in humans than in mice (Brown & Willard, 1994), which may explain why many X-linked genes remain inactive upon the loss of XIST RNA in hESCs. Furthermore, whereas human *XIST* and mouse *Xist* transcripts share some sequence homology, the genomic sequences surrounding human *XIST* and mouse *Xist* are considerably divergent (Chureau *et al.*, 2002). The identification of conserved TCF binding motifs and surrounding sequences upstream of the human *XIST* and mouse *Xist* TSSs suggest a possible conserved role for Wnt signaling in *XIST*/*Xist* regulation and dosage compensation in humans and mice.

Our findings suggest a careful assessment and selection of culture media to faithfully maintain X-inactivation in female hESCs. Given that lithium chloride and other GSK-3 inhibitors can cause loss of XIST RNA expression, the use of media without LiCl and other GSK-3 inhibitors may be more appropriate in the maintenance and differentiation of primed pluripotent hESCs. The inclusion of GSK-3 inhibitors in naïve hESC culture media and their impact on the cellular epigenome may also require careful investigation.

Although our study has shown that the inclusion of LiCl and GSK-3 inhibitors in culture media impedes expression of XIST RNA in PSCs, it does not exclude other causes of XIST RNA expression loss in the PSCs. In addition to GSK-3 inhibition, sources of cellular stress, such as freeze-thaw cycles of hESCs, may also alter XIST RNA expression (Lengner *et al.*, 2010).

Conclusion and Future Directions

The data presented in this chapter suggest a novel, cell-extrinsic mechanism for *XIST*/*Xist* induction and X-chromosome inactivation regulation in humans and mice. Future work will further define the molecular process(es) by which GSK-3 inhibition regulates *XIST*/*Xist* expression in each of these species. As this project progresses, I propose investigating the putative TCF binding sites we identified upstream of *XIST* to determine if direct Wnt-mediated silencing of *XIST* occurs in hESCs, mEpiLCs, and/or mEpiSCs. Using CRISPR-Cas9 to systematically remove or block each putative TCF binding site and RNA FISH to detect XIST RNA in these cells will reveal if these genomic regions are essential to *XIST* regulation. It has been postulated that conserved genomic sequences flanking *XIST* may be involved in its regulation (Chang and Brown, 2010), and the putative TCF binding sites we identified are strong candidates to investigate this idea. A ChIP-Seq analysis in both WT and mutant cells surveying

TCF occupancy in the *XIST*-flanking conserved regions will also shed light on whether these sites may be involved in direct or indirect *XIST* regulation. Findings from these analyses will not only be important for our understanding of *XIST*/*Xist* regulation but may also provide a new example of Wnt-mediated gene silencing.

The experiments outlined in this chapter suggest that Wnt pathway activation can silence *XIST*/*Xist* in human and mouse stem cells, but I did not investigate the converse: whether Wnt inhibition can induce *Xist*. Thus, the culture of undifferentiated female mouse embryonic stem cells (mESCs) in the presence of Wnt inhibitors will be important to investigate the sufficiency of Wnt inhibition to induce *Xist*. Determining if Wnt signaling can induce *Xist* and silence X-linked genes in mESCs will also inform whether Wnt signaling is involved in the initiation of X-inactivation in mice. Under normal mESC culture conditions, undifferentiated mESCs harbor two active X-chromosomes, and thus do not express *Xist* RNA (Rastan and Robertson, 1985). Preliminary RNA FISH experiments carried out by others in the lab have shown mESCs cultured with the Wnt inhibitor IWP-2 induce *Xist* expression. Further analysis of these mESCs by allele-specific RNA-Seq will help determine whether Wnt inhibition can induce *Xist* and silence X-linked genes. RNA-Seq analysis and characterization of these mESCs by IF FISH will also allow us to determine whether *Xist* expression is accompanied by differentiation.

Because naïve female hESCs induce *XIST* from both X chromosomes, the aforementioned experiment is not feasible in pre-X-inactivation female hESCs. However, the inclusion of Wnt inhibitors in the culture of male hESCs, which do not normally express XIST RNA, may show whether Wnt inhibition is sufficient to induce *XIST* in humans. Culturing and characterizing female hESCs that have undergone erosion of X-inactivation with Wnt inhibitors may also be performed to investigate this question.

Related to the experiments outlined above, perturbation of the Wnt pathway in the presence of GSK-3 inhibitors will also be useful to assess whether GSK-3 inhibition is sufficient to activate Wnt signaling and silence *XIST*/*Xist* in humans and mice. Preliminary evidence I gathered from culturing two independent EpiSC lines with both a GSK-3 inhibitor and a Wnt inhibitor suggests that GSK-3 inhibition coupled with Wnt inhibition results in stable *Xist* expression (Supplemental Figure 4.6). These preliminary data support the idea that Wnt activation silences *Xist*, but EpiSCs should be cultured and characterized for more passages to fully assess the ability of Wnt inhibition to rescue the effects of GSK-3 inhibition.

Another future direction for this project involves a systematic investigation of X-linked gene de-repression upon the loss of XIST RNA in hESCs. Currently available RNA-Seq datasets containing hESC X-linked gene expression information do not investigate the stepwise reawakening of X-linked genes from the inactive-X in hESCs undergoing X-inactivation erosion. A more thorough analysis of genes that are expressed vs. silenced on the inactive-X following XIST RNA loss may shed light on requirements for X-linked gene silencing broadly. Female hESCs harboring inactive X chromosomes without XIST RNA coating also provide a valuable model for the study of factors beyond XIST that control X-linked gene silencing.

A final caveat I did not address in this study is whether the maintenance of proper X-linked gene dose in female hESCs lacking XIST RNA is due to selection against cells that undergo upregulated X-linked gene expression. To my eye, there was no apparent difference in cell growth or number of cells that were treated with GSK-3 inhibitors vs. those that were not. Furthermore, the number of colonies between GSK-3 treated and untreated hESCs did not vary significantly. However, future experiments quantifying cell death in varying culture conditions will systematically address the possibility that the X-linked gene dose maintenance in cells

lacking XIST RNA is due to cell selection. These experiments can and should also be conducted in female mEpiSCs and mEpiLCs.

Materials and Methods

Written informed consent for human embryo donation was obtained from both gamete providers as outlined by NIH guidelines and hESC line derivation was performed under University of Michigan's Institutional Review Board approved study, "Derivation of human Embryonic Stem Cells" (IRB-Med; HUM00028742). This study was also approved by the Human Pluripotent Stem Cell Research Oversight (HPSCRO) Committee (HPSCRO record #1035).

hESC Derivation, Expansion, and Characterization

The sources and identifying information for all hESC and feeder cell lines used in this study are included in Supplemental Table 4.1. All hESC lines used in this study were derived from human embryos generated for infertility treatments and were donated to the University of Michigan for one of two reasons: 1) they were no longer needed for reproductive purposes by the donating couple; or 2) they were considered unsuitable for implantation following preimplantation genetic testing (PGT). Voluntary IRB-approved informed consent was obtained from each gamete provider of the embryos donated for hESC derivation. Each donor was informed, in writing and verbally, that donated embryos would be used for attempted derivation of hESCs. Each donor couple was informed of other disposition options, and that neither consenting nor refusing to donate embryos for research would affect the quality of care provided to the potential donors. There was a clear separation between the prospective donors' decision to create human embryos for reproductive purposes and the prospective donors' decision to donate human embryos for research purposes. No payments, cash or in kind, were offered for the donated embryos. Finally, all human embryo donations were documented with evidence of compliance with each

of the fifteen (15) elements of Section II (A) of NIH Guidelines, submitted to, reviewed, and accepted on the NIH hESC registry (https://grants.nih.gov/stem_cells/registry/current.htm).

Vitrified human embryos were warmed according to the donating fertility healthcare provider vitrification/warming protocol specific for the donated embryos. Warmed day (d) 5 embryos were cultured in equilibrated 50ul drops of G2 Plus medium (Vitrolife, #10132), overlaid with Ovoil (Vitrolife, #10028) in a 37˚C humidified incubator at 6% CO_2/5% O_2/89% N_2 (low O_2) until blastocysts were fully expanded or had completed the hatching process from the zona pellucida. The inner cell mass (ICM) of each embryo was either removed immediately or the embryo was incubated for several hrs prior to removal of the ICM, depending on the degree of blastocoele expansion. Blastocyst morphology was scored according to the Gardner and Schoolcraft criteria (Gardner et al., 2000).

The ICM of each expanded blastocyst was isolated via laser-assisted microdissection (Hamilton Thorne Biosciences) in 10ul drops G-MOPS Plus medium (Vitrolife, #10130), overlaid with Ovoil (Supplemental Figure 1A). The blastocysts were immobilized with a holding pipette (20µm I.D., 40µm O.D.). Once adequate tension was established, a calibrated visual target was employed to align the region of the blastocyst to be breached. Several pulses of infrared laser (300mV, 0.2ms) were delivered directly to the hatched embryo or through the zona pellucida opposite the location of the ICM, until the trophectoderm (TE) was exposed. The number of laser pulses necessary to microdissect the ICM varied for each individual blastocyst.

The blastocysts with breached zona pellucida were repositioned to a holding pipet on the opposite side (30µm I.D., 60µm O.D.), and were immobilized with adequate tension on the exposed, laser-breached area of the TE. On the opposite side, adequate tension was applied to the zona pellucida with the smaller (20µm I.D., 40µm O.D.) holding pipet and the TE/ICM

complex was extracted from the zona pellucida with gentle pulling motion of the holding pipets. Each ICM was then microdissected from the overlying TE cells with infrared laser pulses, ensuring that the distance of the laser pulses to the ICM was adequate (300mW 0.2ms, Hamilton Thorne Biosciences).

The dissected ICM were plated on neonatal HFFs (Global Stem, #GSC-3002), inactivated by g-irradiation at 3.7 x 104 cells/cm^2, in a 35mm tissue culture dish on a coverslip (Falcon, BD) pre-coated with 0.1% recombinant human gelatin (Fibrogen) (Supplemental Figure 4.1A; considered passage 0, or P0). The plated ICM was cultured in hESC XenoFree (XF) culture medium [Knock-out DMEM (Gibco, #10829)] containing 20% XF Knockout Serum Replacement (KSR) (Gibco, #12618012), 1mM Glutamax™ (Gibco, #35050-061), 0.1mM β-mercaptoethanol (Sigma-Aldrich, #M6250), 10mM non-essential Amino Acid 100x (Gibco, #11140-05), 4 ng/ml animal-free basic human recombinant Fibroblast Growth Factor (bFGF) (MilliporeSigma, #GF003AF-100UG)]. The ICM and cell outgrowths were cultured under 5% CO_2/5% O_2/90% N_2 (low O_2) in a humidified incubator at 37°C. Attachment of ICM was evaluated 48 hrs after plating onto HFFs and the hESCs were first passaged (P1) 4-6 days after initial evaluation (Supplemental Figure 4.1A). Following manually/mechanically splitting the initial hESC colony, portions were transferred to new HFFs with the above culture medium or as specified for individual experiments.

The culture conditions for cell expansion during P1-3 on HFFs were 5% CO_2/5% O_2/90% N_2 (low O_2) at 37°C. During this early hESC derivation passages, epiblast-like structures were identified, grown, split, and expanded until well-recognized hESC colonies were observed [P2-3 (~d14-21); Supplemental Figure 4.1A]. Passaging was performed without enzymatic treatment and with manual and mechanical cutting of hESC colonies with glass microtools into ~100µM

pieces. At an early passage, hESCs were confirmed to be female (XX chromosome complement) by comparative genomic hybridization array (aCGH) following whole genome amplification (Supplemental Figure 4.1D; See below). Specific experimental culture condition decision branching was performed at P1-3 with plating of hESCs at each indicated passage for analysis of XIST RNA coating and X-linked gene expression. If hESCs continued to be passaged on HFFs, culture conditions were as indicated in the Results section with low O_2 or 5% CO_2/20% O_2/remainder air (high O_2) at 37°C. In some experiments, as described in the Results section, hESCs were transferred onto feeder-free matrix, Matrigel (Corning, #354277), and cultured either with the mTeSR1 medium (International Stem Cell Initiative *et al.*, 2010; Ludwig *et al.*, 2006) (StemCell Technologies, #85850) together with 20% KSR or continued with the XF medium. A full list of the contents of mTeSR1 medium can be found in ref. (Ludwig *et al.*, 2006). Early passage (P4-10) cryopreservation of lines was performed using vitrification (Wu et al., 2005). Following experiments comparing XF culture medium and mTeSR1 medium, the composition of the two culture media identified the presence of lithium chloride (LiCl; 0.98mM) in mTeSR1 and its absence in the XF culture medium. To test the impact of LiCl, 0.98mM LiCl was added XF culture medium. The experiments outlined in Figure 4.8 were performed with addition of GSK-3 inhibitors to XF medium at the following reported optimized concentrations: LY2090314, 1.5 nM; BIO, 5.0nM; Alsterpaullone, 4.0nM (Eldar-Finkelman and Martinez, 2011; Leost *et al.*, 2000; Sato *et al.*, 2004). The 90-14 hESC line was vitrified and warmed prior to culturing for this experiment.

All four hESC lines used in this study (UM33-4, UM63-1, UM77-2, and UM90-14) were characterized for: 1) absence of mycoplasma; 2) presence of pluripotency markers (Supplemental Figure 1B-C); 3) normal XX female karyotype (Supplemental Figure 1E; see below); 4) short

tandem repeat (STR) analysis (see below); and, 5) embryoid body formation and presence of lineage markers for endoderm, mesoderm, and ectoderm (Supplemental Figure 1F).

Testing hESCs for Mycoplasma Contamination

All four hESC lines used in the study were tested for mycoplasma contamination using a PCR-based assay (Venor™ GeM Mycoplasma Detection Kit, Sigma, #MP0025) and gel electrophoresis. The assay contained a PCR-negative control (no polymerase), a positive control (non-infectious DNA fragments of *Mycoplasma orale* genome, band size 267bp), and an internal control (internal sequence of HTLV-I tax gene, presence of amplification band at 191bp and no band at 267bp).

Immunocytochemistry

All four hESC lines were cultured on coverslips prior to immunocytochemistry. Culture media was aspirated, and cells were washed with 1X PBS (Fisher Scientific, #10010023) and fixed in a solution containing 4% paraformaldehyde (Electron Microscopy Sciences, #15710) and 4% Sucrose (ThermoFisher Scientific, #BP220) for 15 min at room temperature. The cells were then permeabilized with 0.1% Triton X-100 (Fisher Scientific, #EP151) for 5 min at room temperature and were blocked in a 5% Normal Goat Serum (Invitrogen # 31872) solution for 1 hr at room temperature. The cells were stained with primary antibodies overnight at 4°C. Immunocytochemical analysis was performed on all hESC lines with specific antibodies against pluripotency markers SOX2 (rabbit; EMD Millipore #AB5603; 1:800 dilution); NANOG (rabbit; Abcam #ab21624; 1:150 dilution); OCT3/4 (goat; Santa Cruz Biotechnology #sc-8628; 1:300 dilution); TRA1-60 (mouse; Millipore #MAB4360; 1:200 dilution); and, SSSEA-4 (mouse; Millipore #MAB4304; 1:100 dilution). Primary antibody staining was followed by three washes in 1× PBS (Fisher Scientific, #10010023) for 5 min each. After washing, samples were

incubated with fluorescent secondary antibodies: donkey anti-rabbit secondary antibody coupled to the Cy3 fluorophore (ImmunoResearch #711-165-152) was used at 1:200 dilution with the rabbit anti-NANOG primary antibody and at 1:800 dilution with the rabbit anti-SOX2 primary antibody; donkey anti-goat secondary antibody coupled to the FITC fluorophore (Jackson ImmunoResearch #705-096-147) was used at 1:800 dilution with the goat anti-OCT3/4 primary antibody; donkey anti-mouse secondary antibody coupled to the Cy3 fluorophore (Jackson ImmunoResearch #715-165-150) was used at the following dilutions with the primary antibodies: mouse anti-TRA 1-60, 1:100 dilution; and, mouse anti-SSEA4, 1:800 dilution. Nuclei were labelled with Hoechst co-staining (blue; Hoechst ThermoFisher Scientific #33258) and imaged using an Olympus IX71 microscope. Colony morphology was simultaneously assessed by phase-contrast brightfield microscopy.

Quantitative Real-Time PCR for Pluripotency and Cell Lineage Markers

RNA was isolated from all four hESC lines using an RNA isolation kit (QIAGEN RNeasy Kit, #74104) and reverse transcribed into cDNA using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, #4368814). Expression of pluripotency markers, *OCT3/4*, *NANOG*, and *SOX2*, and tissue-specific markers, Alpha-fetoprotein (*AFP*) and *GATA-4* (endoderm), Brachyury (*BRACHY*) and vascular endothelial cadherin (*VE-CAD*; mesoderm), and neuron-specific class III β-tubulin (*TUJ-1*) and type I intermediate filament chain keratin 18 (*KRT-18*; ectoderm) were measured in triplicate by quantitative real-time PCR using primer sets listed in Supplemental Table 4.2. Relative fold expression for genes of interest were calculated using the comparative CT method with *β-ACTIN* (*ACTB*) as the internal control. Expression levels of pluripotency markers in the hESCs were normalized to *ACTB* expression.

Karyotyping

G-banding was performed on 20 metaphase spreads of all four UM hESCs at passages ranging from P6-31 (Cell Line Genetics, Madison, WI). Metaphase spreads were evaluated at 100X with a Leica GSL Scanner (100X objective, Leica GSL 120 CytoVision (Leica MicroSystems, Buffalo Grove, IL) with band–count resolution of ~475. Cytogenic analysis demonstrated all hESC lines analyzed were female 46XX, as shown in Supplemental Table 4.3.

Short Tandem Repeat Analysis

Cell Line DNA fingerprinting was performed on all four hESC lines used in the study (Cell Line Genetics, Madison, Wisconsin). Analysis of fifteen short-tandem repeat (STR) loci, plus the gender determining locus, *Amelogenin*, confirmed the presence of a single human cell line that is unique from lines published in the ATCC, NIH, or DSMZ websites.

Characterizing Differentiation Potential of hESCs

hESCs were cultured on Matrigel (Corning, #354277) in mTeSR1 medium (StemCell Technologies, #85850) as described above with the following culture conditions: 37°C, 5% CO_2/20% O_2/remainder air until reaching ~80% confluency and harvested mechanically with scrapers. Detached cells were transferred and cultured in a 60mm petri dish (BD Falcon, #353652) containing 6 ml of AggreWell^{TM} medium for embryoid body (EB) differentiation (Stem Cell Technologies, #05893) with media changes every other day for 14d. After 8d of differentiation, EBs were collected and RNA was extracted using QIAGEN RNeasy Kit (#74104) from all four hESC lines. The reverse transcription (RT) of total RNA to single-stranded cDNA was performed with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, #4368814). Cell differentiation was assessed by profiling expression of molecular markers associated with the three somatic germ layers using real time reverse

transcription PCR using primer sets listed under Quantitative Real-Time PCR for Pluripotency and Lineage Markers (Bio Rad SsoAdvance SYBR Green Supermix Bio-Rad, #1725272).

hESC Differentiation into Embryoid Bodies for XIST RNA Coating Analyses

hESC lines UM77-2 and UM 63-1 were grown, maintained, and expanded on HFFs in XF medium as described above. Colonies were cut, detached, and transferred to a 60mm culture dish (BD Falcon, #353652) and subsequently cultured with one of three designated EB culture media: 1) XF medium without bFGF; 2) XF medium without bFGF and with 0.98mM LiCl; and, 3) Aggrewell^{TM} medium (Stem Cell Technologies, #05893) (mTeSR1-based medium containing LiCl). EBs were cultured for 9 days at 37^{o}C, 5% CO_2 and 20% O_2. EBs were dissociated with L7 passage solution (Lonza, #FP-5013) prior to collection at days 3, 6, and 9. Dissociated EBs were plated on Matrigel-coated (Corning, #354277) coverslips in each of the above media formulations for 48 hrs before processing for RNA FISH.

mEpiLC Generation and Culture

mESC lines were derived from individual E3.5 pre-implantation mouse embryos, which were plated on quiescent mouse embryonic fibroblast (MEF) feeder cells in ESC derivation media consisting of KnockOut DMEM (Gibco, #10829–018), KnockOut serum replacement (Invitrogen, #10828–028), l-glutamine (Gibco, #25030), MEM non-essentials amino acids (Gibco, #11140–050), β-Mercaptoethanol (Sigma, #M7522), Penicillin-streptomycin (100×) (Gibco, #15070–063), GSK3 inhibitor CHIR99021 (Stemgent, #04–0004), MEK inhibitor PD0325901 (Stemgent, #04–0006) and LIF (10^7/mL) (Millipore, #ESG1106). Plates were incubated at 37°C, 5% CO_2 for 48 hrs. On day 3 post-plating, mESC derivation media was replaced with fresh mESC derivation media. Outgrowths became prominent at 4–5 days post-plating, and were dissociated in 0.05% trypsin (Invitrogen, #25300-054). Dissociated embryos

were plated individually into wells of a MEF-plated 96-well plate with ESC derivation media. mESC colonies became evident over the next 2–3 days and were maintained in mESC culture media consisting of KnockOut DMEM (Gibco, #10829–018), fetal bovine serum embryonic stem cell qualified (ES-FBS) (Bio-Techne, #S10250), KnockOut serum replacement (Invitrogen, #10828–028), l-glutamine (Gibco, #25030), MEM non-essential amino acids (Gibco, #11140–050), β-Mercaptoethanol (Sigma, #M7522), GSK3 inhibitor CHIR99021 (Stemgent, #04–0004), MEK inhibitor PD0325901 (Stemgent, #04–0006) and LIF (10^7/mL) (Millipore, #ESG1106).

To generate mEpiLCs, mESC lines were passaged into 2i culture conditions [N2B27 medium consisting 50% DMEM/F12 (Thermo Fisher Scientific, # 11320033), 50% neurobasal medium (Thermo Fisher Scientific, #21103049), 2 mM L-glutamine (GIBCO, #25030), 0.1 mM β-mercaptoethanol (Sigma,#M7522), N2 supplement (Invitrogen #17502048), B27 supplement (Invitrogen #17504-044), supplemented with 3 μM GSK-3 inhibitor CHIR99021 (CHIR) (Stemgent #04-0004), 1 μM MEK inhibitor PD0325901 (Stemgent #04-0006), and 1000 U/ml LIF (Millipore #ESG1106)] at 5% CO_2 and grown in gelatin-coated tissue culture dishes for 4 passages (Buecker et al., 2014; Hayashi et al., 2011).

To differentiate mESCs into EpiLCs, mESCs were cultured in N2B27 medium supplemented with 10 ng/ml bFGF (R&D Systems, #233-FB) and 20 ng/ml Activin A (R&D Systems, #338-AC) in Fibronectin (15μg/ml) (Sigma #F1141) coated tissue culture dishes for 48 hrs. The mEpiLCs were then cultured without or with 3 μM CHIR99021 (CHIR) (Stemgent #04-0004) for an additional 48 hrs in N2B27 medium. A concentration of 3 μM was selected for CHIR because this is the CHIR concentration used in naïve mESC culture (Samanta and Kalantry, 2020). mEpiLCs utilized for RNA FISH staining were cultured, permeabilized, and fixed on fibronectin-coated (15μg/ml) (Sigma #F1141) coverslips.

mEpiSC Derivation and Culture

mEpiSCs were derived from individual E3.5 pre-implantation mouse embryos, which were plated on quiescent mouse embryonic fibroblast (MEF) feeder cells in K15F5 medium containing Knockout DMEM (GIBCO, #10829-018) supplemented with 15% Knockout Serum Replacement (Gibco, #A1099201), 5% ES-FBS (GIBCO, #104390924), 2 mM L-glutamine (Gibco, #25030), 1X nonessential amino acids (Gibco, #11140-050), and 0.1 mM β-mercaptoethanol (Sigma, #M7522). After 5–6 days, blastocyst outgrowths were dissociated partially with 0.05% trypsin (Invitrogen, #25300-054). The partial dissociates were plated individually into a 1.9-cm^2 well containing a MEF feeder layer and cultured for an additional 4–6 days in K15F5 medium. The culture was then passaged by a brief exposure (2–3 min) to 0.05% trypsin/EDTA with gentle pipetting to prevent complete single-cell dissociation of pluripotent clusters and plated into a 9.6-cm2 well containing MEF feeders in K15F5 medium. Morphologically distinct mEpiSC colonies became evident over the next 4–8 days and were subcloned from a mixed population of cells, including mESCs. mEpiSC colonies were manually dissociated into small clusters using a glass needle and plated into 1.9-cm^2 wells containing MEF feeders in mEpiSC medium consisting of Knockout DMEM (Gibco, #10829018) supplemented with 20% Knockout Serum Replacement (Gibco, #10828010), 2 mM Glutamax™ (Gibco, #35050061), 1× nonessential amino acids (Gibco, #11140050), 0.1 mM β-mercaptoethanol (Sigma, #M7522), and 10-ng/ml FGF2 (R&D Systems, #233-FB).

After derivation, mEpiSCs were cultured in mEpiSC medium and passaged every third day using 1.5 mg/ml collagenase type IV (GIBCO, #17104-019) with pipetting into small clumps. mEpiSC medium used to culture CHIR-treated mEpiSCs was supplemented with 3 μM

CHIR99021 (CHIR) (Stemgent #04-0004). mEpiSCs generated for RNA FISH staining were cultured, permeabilized, and fixed on fibronectin-coated (15μg/ml) (Sigma #F1141) coverslips.

RNA Fluorescence *In Situ* Hybridization (FISH) Probe Labeling & Precipitation

Human and mouse probes [(*XIST* BAC (BAC PAC, RP13-183A17); *ATRX* BAC (BAC PAC, RP11-42M11); *USP9X* BAC (Invitrogen, CTD 3174G14); *XIST* fosmid (BAC PAC, G135P63425C4)] were labeled with Fluorescein-12-dUTP (Invitrogen), Cy3-dCTP (GE Healthcare, #PA53021), or Cy5-dCTP (GE Healthcare, #PA55031). Labeled probes for multiple genes were precipitated in a 3M sodium acetate (Teknova, #S0298) solution along with 300 μg of yeast tRNA (Invitrogen, #15401–029), and 150 μg of sheared, boiled salmon sperm DNA (Invitrogen, #15632–011). The solution was then centrifuged at 21,130 X g for 20 min at 4°C. The resulting pellet was washed in 70% ethanol, then washed in 100% ethanol, dried, and re-suspended in deionized formamide (ISC Bioexpress, #0606–500ML). The re-suspended probe was denatured via incubation at 90°C for 10 min followed by an immediate 5 min incubation on ice. A 2X hybridization solution consisting of 4X SSC, 20% Dextran sulfate (Millipore, #S4030), and 2.5 mg/ml purified BSA (New England Biolabs, #B9001S) was added to the denatured probe/formamide solution. Probes were stored at −20°C until use.

RNA FISH Staining

hESCs, differentiated mEpiLCs, and mEpiSCs grown on coverslips were permeabilized through sequential treatment with ice-cold cytoskeletal extraction (CSK) buffer containing 100 mM NaCl (ThermoFisher Scientific, #BP358), 300 mM sucrose (ThermoFisher Scientific, #BP220), 3 mM $MgCl_2$ (ThermoFisher Scientific, #AA12315A7), and 10 mM PIPES buffer (Sigma Aldrich, #P6757), pH 6.8 for 30 sec; ice-cold CSK buffer containing 0.4% Triton X-100 (Fisher Scientific, #EP151) for 30 sec; followed twice with ice-cold CSK buffer for 30 sec each. After

permeabilization, cells were fixed by incubation in 4% paraformaldehyde (Electron Microscopy Sciences, #15710) for 10 min. Cells were then rinsed 3 times in 70% ethanol and stored in 70% ethanol at −20°C prior to RNA FISH staining. Prior to RNA FISH probe hybridization, coverslips were dehydrated through 2 min incubations in 70%, 85%, 95%, and 100% ethanol solutions and subsequently air-dried for 15 min. The coverslips were then hybridized to the FISH probe overnight in a humid chamber at 37°C. The samples were then washed 3 times for 7 min each at 37°C with 2X SSC (Invitrogen, #AM9765)/50% deionized formamide (ISC Bioexpress, VWR:#0606), 2X SSC, and 1X SSC. A 1:250,000 dilution of DAPI (Invitrogen, #D21490) was added to the third 2X SSC wash. Coverslips were then mounted on slides in Vectashield (Vector Labs, #H-1000) and sealed with nail polish.

Microscopy

Coverslips containing stained cells were imaged using a Nikon Eclipse TiE inverted microscope with a Photometrics CCD camera. The images were deconvolved and uniformly processed using NIS-Elements software (Version 4.60.00).

Quantification of RNA FISH Stains

Expression of X-linked genes was quantified at 10X to 100X resolution starting at the upper left corner of each image. Percent expression of monoallelic XIST RNA was calculated for each colony on the coverslip and recorded in one of the following expression categories for all experiments except Fig. 5: 0%, 1-19%, 20-39%, 40-59%, 60-79%, 80-99%, 100%. Figs. 5 and 6 display XIST RNA coating data in only three key expression categories (0-19%, 20-79%, 80-100%), as these categories most clearly demonstrated XIST RNA coating loss. Expression of X-linked genes *ATRX* and *USP9X* in individual nuclei was quantified in the following categories: nuclei with XIST RNA coats and monoallelic ATRX or USP9X expression; nuclei with XIST

RNA coats and biallelic ATRX or USP9X expression; nuclei without XIST RNA coats and monoallelic ATRX or USP9X expression; and, nuclei without XIST RNA coats and biallelic *ATRX* or *USP9X* expression. Colonies with fewer than 100 cells or colonies with excessive cell overlaps were not quantified. All measurements were taken from distinct samples. Raw quantification data for all RNA FISH experiments in this study are included in the Source Data file that accompanies this publication.

RNA Sequencing (RNA-Seq) Sample Preparation

Total RNA was isolated from TRIzol (Life Technologies, #15596-018) according to the manufacturer's instructions. cDNA libraries were generated and sequenced on the Illumina HiSeq2000 platform to generate 75 bp single-end reads.

Analysis of RNA-Seq Data

Quality control analysis was conducted using FastQC (Version 0.11.9). Reads were aligned to the hg19 (human) reference genome using STAR (Dobin *et al.*, 2013) (Version 2.7.10a) and counted using FeatureCounts (Liao *et al.*, 2014) (Version 1.22.2). Differential expression analysis was conducted using DESeq2 (Version 1.34.0).

Statistics

General linear model regression analysis was performed in R to determine statistically significant differences in XIST RNA coating through passaging (Figs. 1-10). For experiments comparing different culture conditions (Figs. 2-10), p-values for individual linear models or p-values comparing the linear models are reported, as indicated. A threshold of $p = 0.05$ was used to test for statistical significance. Principal Component Analysis (PCA) was conducted using R.

Data Availability

All hESC lines generated for this study have been deposited to the NIH Human Embryonic Stem Cell Registry [https://grants.nih.gov/stem_cells/registry/current.htm]. The raw and processed RNA-Seq data generated in this study have been deposited in the Gene Expression Omnibus under accession code GSE157809 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE157809]. Source data, which include raw RNA FISH quantification and differential expression datasets, are provided online with this publication [https://www.nature.com/articles/s41467-022-30259-x].

Author Contributions

L.K., S.M.P., I.E., and A.M.D.R. derived and cultured hESCs. M.C., S. Kumar, E.B., B.L., and A.W. performed RNA FISH in hESCs and quantified the hESC RNA FISH data. M.C. conducted the hESC RNA-Seq sample preparation and analysis. M.C. also performed all mouse cell culture, RNA FISH staining, stain quantification, and molecular biology experiments. K.C identified TCF binding sites near human XIST and mouse Xist. G.D.S., A.M.D.R., M.C., S. Kumar, and S. Kalantry designed experiments and analyzed data. M.C., S. Kumar, G.D.S., and S. Kalantry prepared the manuscript.

Table 4.1. hESC and Feeder Cell Sources and Identifiers.

Cell Line	Source	Identifier
Inactive Human Foreskin Fibroblasts (HFFs)	Global Stem	Global Stem: #GSC-3002
Mouse Embryonic Fibroblasts (MEFs)	Global Stem	Global Stem: #GSC-6001G
UM33-4 hESC line (NIH approval number NIH hESC-14-0279)	MStem Cell Laboratories	hESC Line: UM33-4
UM63-1 hESC line (NIH approval number NIH hESC-14-0277)	MStem Cell Laboratories	hESC Line: UM63-1
UM77-2 hESC line (NIH approval number NIH hESC-14-0278)	MStem Cell Laboratories	hESC Line: UM77-2
UM90-14 hESC line (NIH approval number NIH hESC-15-0306)	MStem Cell Laboratories	hESC Line: UM90-14

Table 4.2. Quantitative Real-Time PCR Primers.

Gene	Forward Primer	Reverse Primer
OCT3/4	GATGGCGTACTGTGGGCCC	TGGGACTCCTCCGGGTTTTG
NANOG	TCCTCCTCTTCCTCTATACTAAC	CCCACAAATCACAGGCATAG
SOX2	GAGAGAAAGAAAGGGAGAGAAG	GAGAGAGGCAAACTGGAATC
β-ACTIN	GCCGAGGACTTTGATTGC	GTGTGGACTTGGGAGAGG
AFP	AAACTATTGGCCTGTGGCGA	GGCCAACACCAGGGTTTACT
GATA-4	CAGATGCCTTTACACGCTGA	TCCGCTTGTTCTCAGATCCT
BRACHY	ACCCAGTTCATAGCGGTGAC	GGATTGGGAGTACCCAGGTT
VE-CAD	CCTACCAGCCCAAAGTGTGT	GAGATGACCACGGGTAGGAA
TUJ-1	ATGCGGGAGATCGTGCACAT	CCCTGAGCGGACACTGT
KRT-18	CACAGTCTGCTGAGGTTGGA	GAGCTGCTCCATCTGTAGGG

Table 4.3. hESC Line Karyotyping Results.

Cell Line	Passage	Result	Notes
UM77-2	19	46,XX	Karyotyped
UM33-4	31	46,XX	Karyotyped
UM33-4	23	46,XX	24-Chromosome Molecular/Microarray PGS
UM63-1	22	46,XX	24-Chromosome Molecular/Microarray PGS
UM63-1	8	46,XX	Karyotyped; 18 cells with normal karyotype, 2 cells with non-clonal aberrations
UM63-1	20	46,XX	Karyotyped; 11 cells with normal karyotype, 7 cells had 3 X chromosomes, 2 cells with non-clonal aberrations
UM90-14	6	46,XX	Karyotyped; 19 cells with a normal karyotype, 1 cell with a non-clonal chromosomal aberration

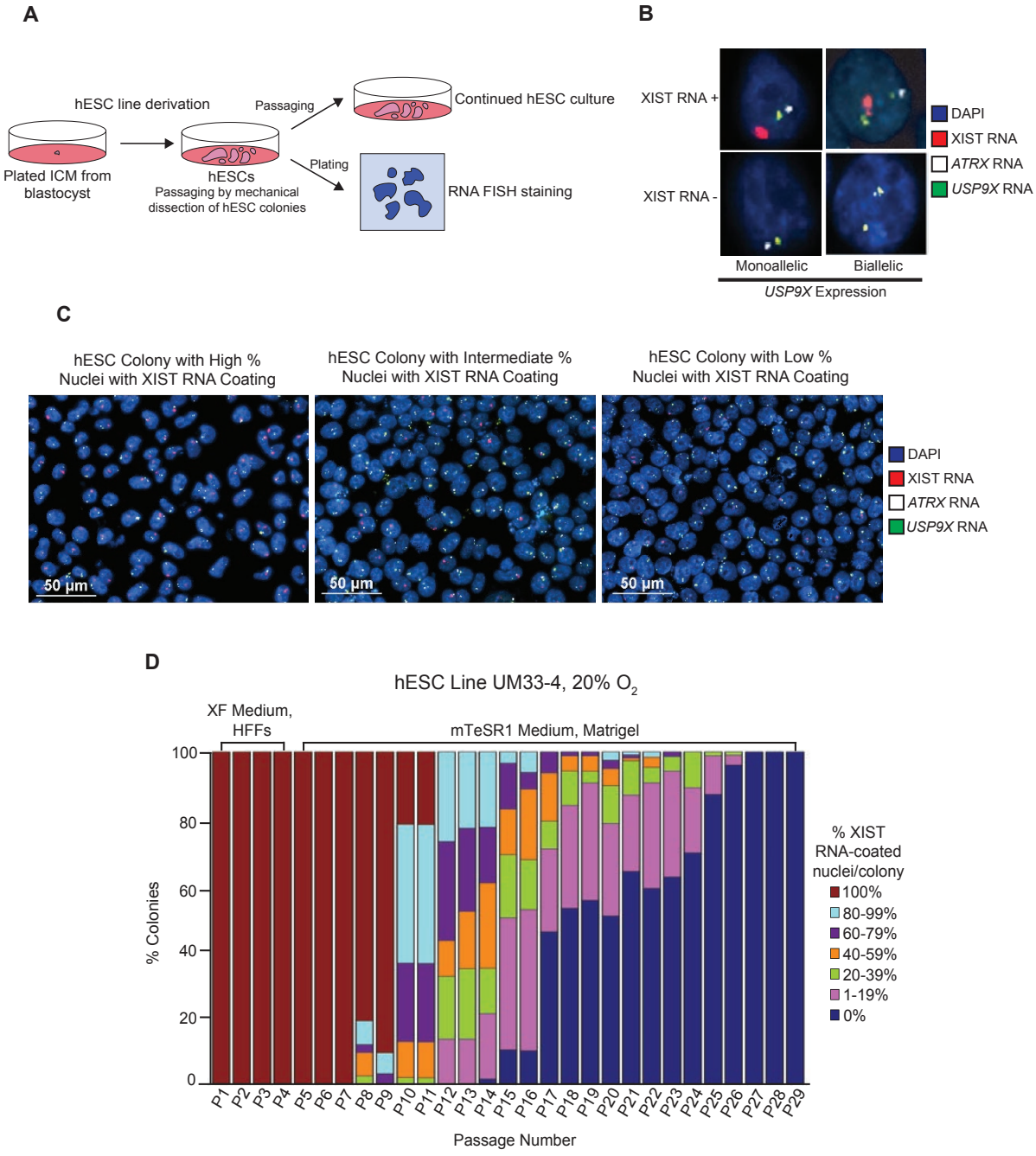

A
hESC line derivation
Passaging
Continued hESC culture
Plated ICM from blastocyst
hESCs
Passaging by mechanical dissection of hESC colonies
Plating
RNA FISH staining
B
XIST RNA +
XIST RNA -
DAPI
XIST RNA
ATRX RNA
USP9X RNA
Monoallelic
Biallelic
USP9X Expression
C
hESC Colony with High % Nuclei with XIST RNA Coating
hESC Colony with Intermediate % Nuclei with XIST RNA Coating
hESC Colony with Low % Nuclei with XIST RNA Coating
50 µm
50 µm
50 µm
DAPI
XIST RNA
ATRX RNA
USP9X RNA
D
hESC Line UM33-4, 20% O2
XF Medium, HFFs
mTeSR1 Medium, Matrigel
% Colonies
100
80
60
40
20
0
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 P13 P14 P15 P16 P17 P18 P19 P20 P21 P22 P23 P24 P25 P26 P27 P28 P29
Passage Number
% XIST RNA-coated nuclei/colony
100%
80-99%
60-79%
40-59%
20-39%
1-19%
0%

Figure 4.1. Loss of XIST RNA coating upon prolonged passaging of female hESCs (A) Schematic depicting the derivation, culture, passaging, and RNA FISH staining of hESCs in this study. (B) Representative nuclei stained to detect XIST RNA (red), RNAs from X-linked genes ATRX (white) and USP9X (green), and the nucleus with DAPI (blue). Top, representative nuclei with XIST RNA coating. Bottom, representative nuclei without XIST RNA coating. At least 100 nuclei were counted per colony for hESC RNA FISH quantification. The total number of colonies quantified at each passage range from 2-138 and are cataloged in source data. (C) Representative hESC colonies stained to detect XIST RNA (red), ATRX RNA (white), and USP9X RNA (green). Nuclei are stained blue with DAPI. At least 100 nuclei were counted per colony for hESC RNA FISH quantification. (D) Percentage of nuclei with XIST RNA coats per colony in hESC line UM33-4 derived in XenoFree (XF) medium on human fibroblast feeders (HFFs) and cultured subsequently on Matrigel under atmospheric oxygen levels (20%). The percentage of nuclei with XIST RNA coats in individual hESC colonies were stratified into 20% increments. 100% value indicates that all nuclei in a colony harbored XIST RNA coats, whereas 0% indicates that all nuclei lacked XIST RNA coats in a colony. The percentage of colonies harboring nuclei with XIST RNA coats decreased significantly with passage number (general linear model, $p = 0.002$). See also Supplementary Figure 4.2.

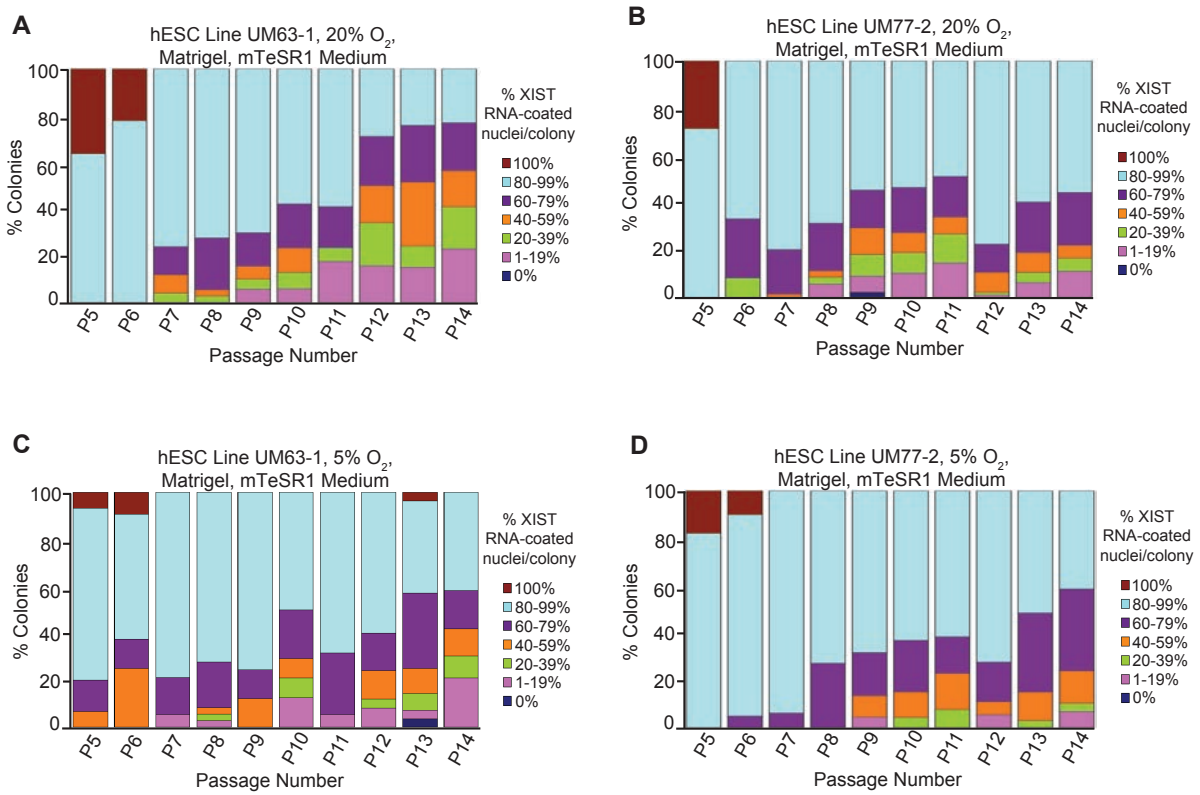


Figure 4.2. XIST RNA coating in female hESCs cultured in atmospheric vs. physiological O_2 concentration (A-D) Percentage of nuclei with XIST RNA coats per colony of hESC lines UM63-1 (A,C) and UM77-2 (B,D) cultured in parallel under 20% (A,B) and 5% (C,D) O_2 concentration on Matrigel in mTeSR1 medium. The difference in the frequency of nuclei without XIST RNA coats per colony in either cell line when cultured at physiological vs. atmospheric O_2 concentration is not significant (general linear model comparison, $p = 0.1$). See also Supplementary Figure 4.2. At least 100 nuclei were counted per colony for hESC RNA FISH quantification. The total number of colonies quantified at each passage range between 11-93.

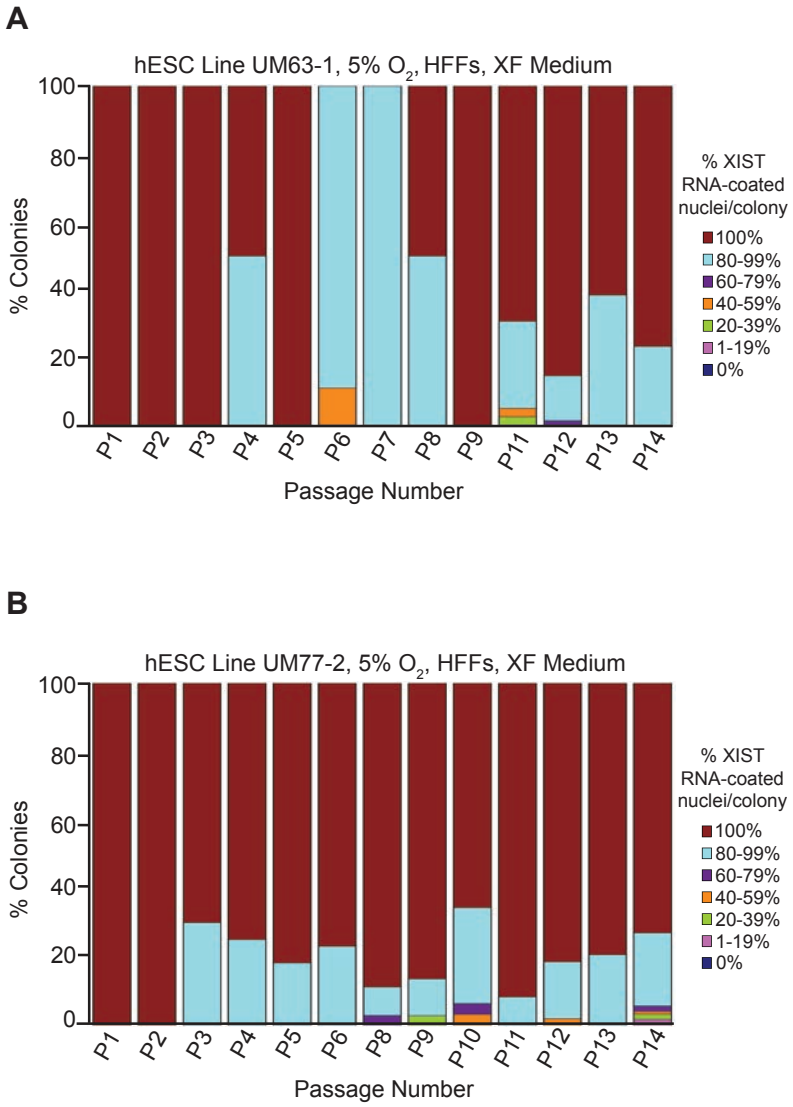


Figure 4.3. Impact of culture surface on XIST RNA coating in female hESCs (A,B) Percentage of nuclei with XIST RNA coats in colonies of hESC line UM63-1 (A) and UM77-2 (B) cultured on HFFs in XF medium under 5% O_2 concentration. The frequency of nuclei harboring XIST RNA coats per colony in either cell line when cultured on HFFs did not decrease significantly (general linear model, $p = 0.09$). At least 100 nuclei were counted per colony for hESC RNA FISH quantification. The total number of colonies quantified at each passage range between 1-126.

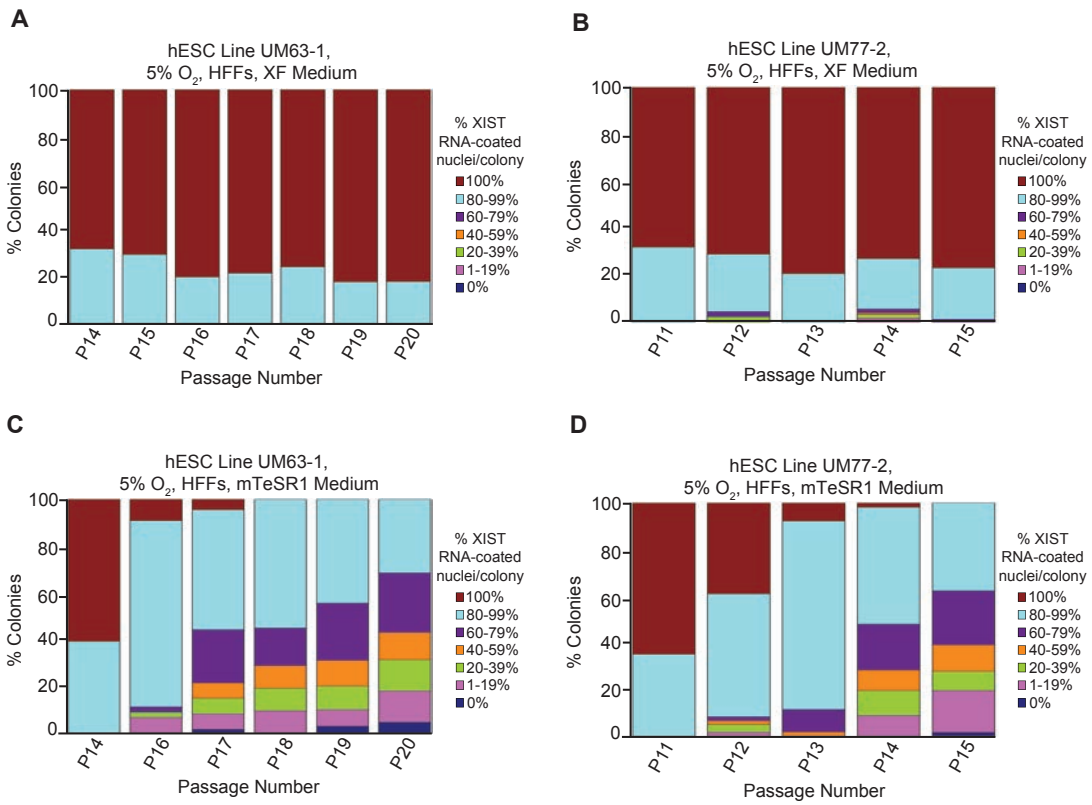


Figure 4.4. Impact of culture medium on XIST RNA coating in female hESCs (A-D) Percentage of nuclei with XIST RNA coats per colony in hESC lines UM63-1 (A,C) and UM77-2 (B,D) cultured in parallel in XF medium (A,B) and mTeSR1 medium (C,D) on HFFs. hESCs cultured with mTeSR1 medium displayed a significant decrease in nuclei with XIST RNA coating compared to hESCs cultured in XF medium during passaging (general linear model comparison; $p < 0.001$). All hESCs in this experiment were cultured in 5% O_2 on HFFs. The quantification data for P13-14 in B are taken from Figure 4.3B. At least 100 nuclei were counted per colony for hESC RNA FISH quantification. The total number of colonies quantified at each passage range between 10-180.

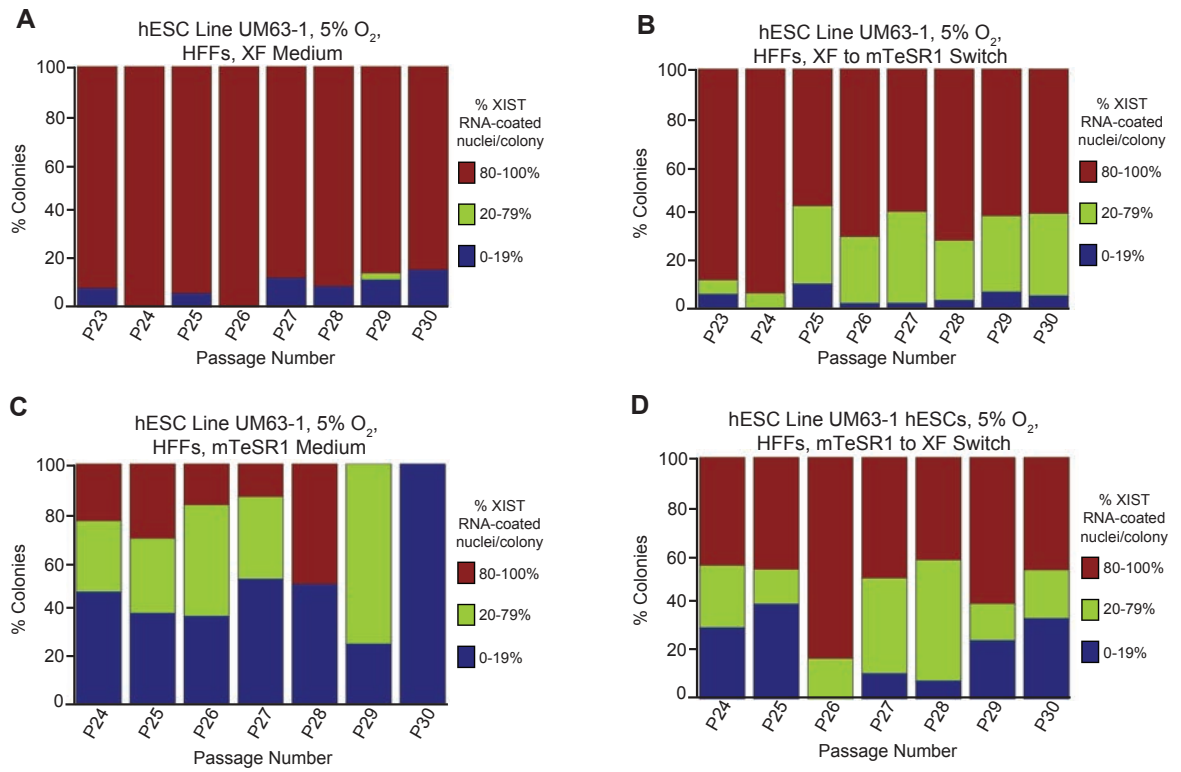
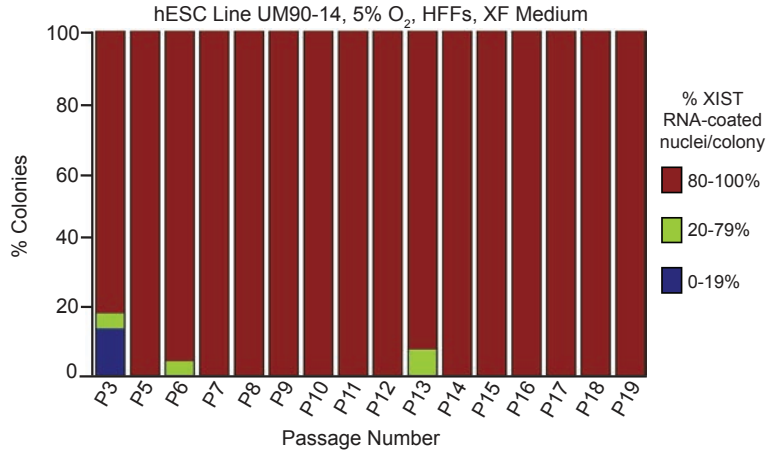
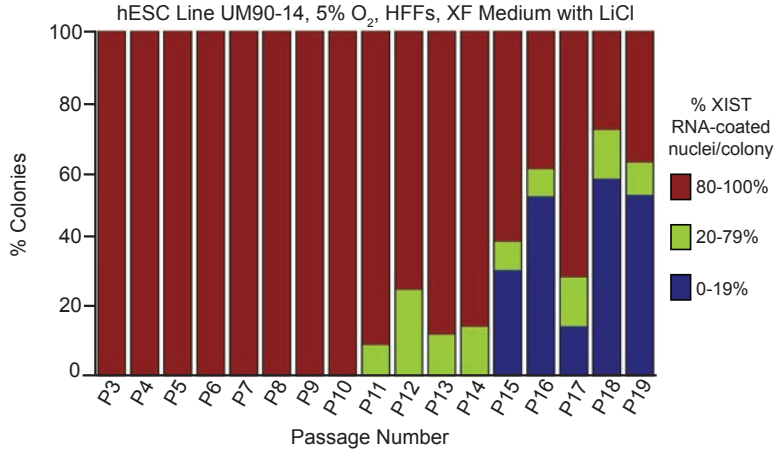


Figure 4.5. Analysis of culture media switching on XIST RNA coating in female hESCs(A,B) Percentage of nuclei with XIST RNA coats per colony of hESC line UM63-1 (A) cultured continuously in XF medium and (B) cultured initially in XF medium and subsequently switched to mTeSR1 medium. Percentage of nuclei with XIST RNA coats per colony of hESC line UM63-1 (C) continuously cultured in mTeSR1 medium and (D) cultured initially in mTeSR1 medium and then switched to XF medium. hESCs cultured initially in XF medium and subsequently switched to mTeSR1 medium displayed a significant decrease in nuclei with XIST RNA coating during passaging compared to hESCs cultured continuously in XF medium (general linear model comparison, $p = 0.01$). hESCs cultured continuously in mTeSR1 medium displayed a significant decrease in nuclei with XIST RNA coating during passaging compared to hESCs cultured initially in mTeSR1 medium and then switched to XF medium (general linear model comparison, $p < 0.001$). All hESCs in this experiment were cultured in 5% O_2 on HFFs. See also Supplemental Figure 4.3 and Supplementary Figure 4.4. At least 100 nuclei were counted per colony for hESC RNA FISH quantification. The total number of colonies quantified at each passage range between 2-148.

A


hESC Line UM90-14, 5% O2, HFFs, XF Medium
% Colonies
100
80
60
40
20
0
P3
P5
P6
P7
P8
P9
P10
P11
P12
P13
P14
P15
P16
P17
P18
P19
Passage Number
% XIST RNA-coated nuclei/colony
80-100%
20-79%
0-19%

B


hESC Line UM90-14, 5% O2, HFFs, XF Medium with LiCl
% Colonies
100
80
60
40
20
0
P3
P4
P5
P6
P7
P8
P9
P10
P11
P12
P13
P14
P15
P16
P17
P18
P19
Passage Number
% XIST RNA-coated nuclei/colony
80-100%
20-79%
0-19%

C

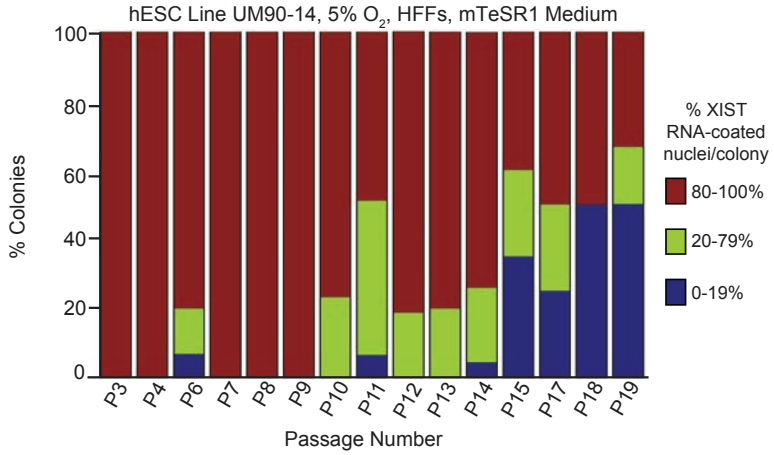

hESC Line UM90-14, 5% O2, HFFs, mTeSR1 Medium
% Colonies
100
80
60
40
20
0
P3
P4
P6
P7
P8
P9
P10
P11
P12
P13
P14
P15
P17
P18
P19
Passage Number
% XIST RNA-coated nuclei/colony
80-100%
20-79%
0-19%

Figure 4.6. LiCl in mTeSR1 medium as a cause of XIST RNA loss in female hESCs (A-C) Percentage of nuclei with XIST RNA coating in colonies of hESC line UM90-14 cultured in XF medium (A), XF medium supplemented with 0.98mM LiCl (XF with LiCl) (B), and mTeSR1 medium (C). hESCs cultured in XF medium did not display a significant decrease in nuclei with XIST RNA coats across passaging (general linear model, $p = 0.2$). hESCs cultured in XF medium with LiCl and mTeSR1 medium lost XIST RNA coating in a significant percentage of nuclei per colony during passaging compared to cells cultured in XF medium (general linear model, $p < 0.001$). All hESCs in this experiment were cultured in 5% O_2 on HFFs. At least 100 nuclei were counted per colony for hESC RNA FISH quantification. The total number of colonies quantified at each passage range between 5-78.

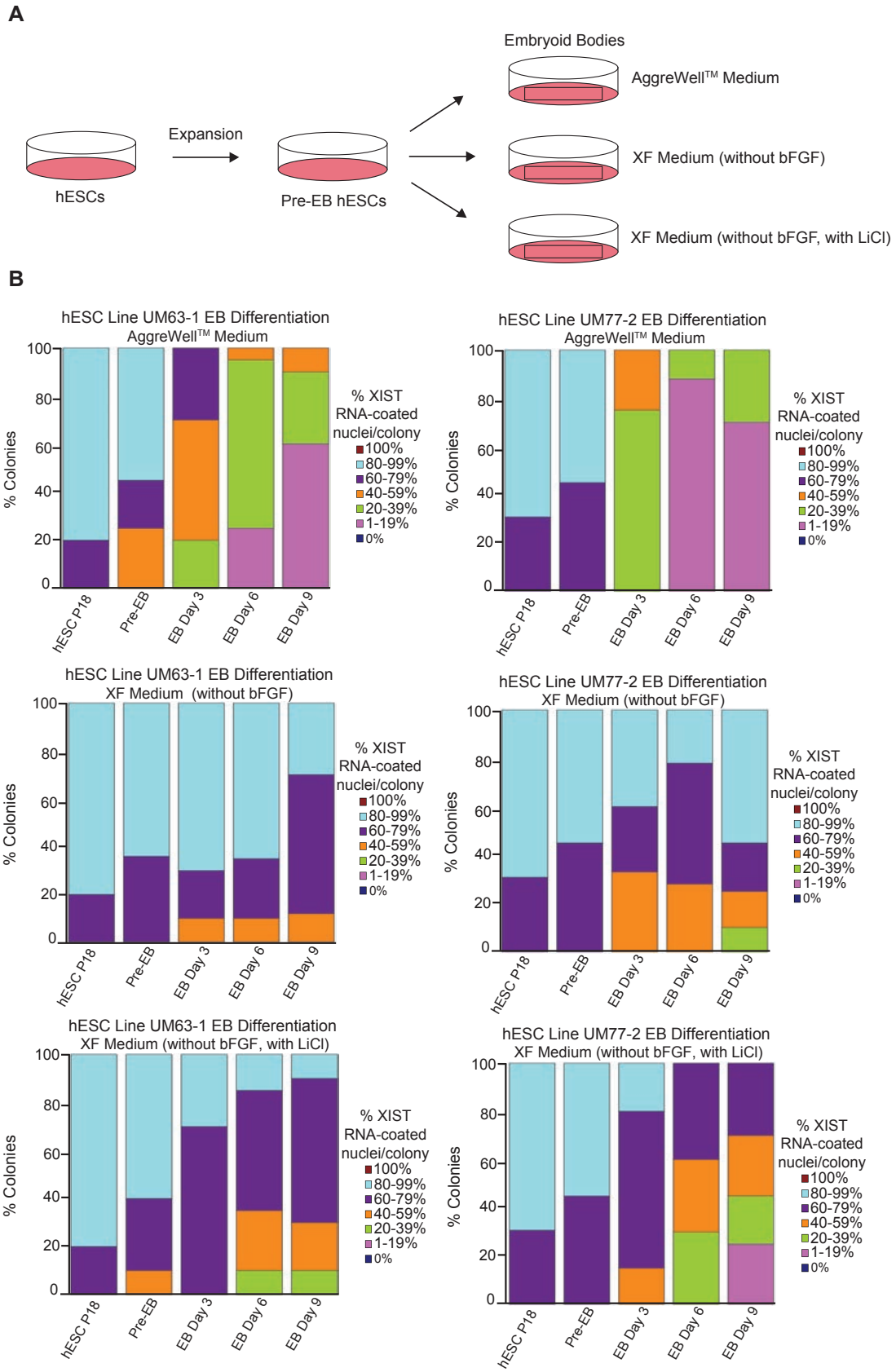

A
Embryoid Bodies
AggreWell™ Medium
Expansion
hESCs
Pre-EB hESCs
XF Medium (without bFGF)
XF Medium (without bFGF, with LiCl)
B
hESC Line UM63-1 EB Differentiation
AggreWell™ Medium
hESC Line UM77-2 EB Differentiation
AggreWell™ Medium
hESC Line UM63-1 EB Differentiation
XF Medium (without bFGF)
hESC Line UM77-2 EB Differentiation
XF Medium (without bFGF)
hESC Line UM63-1 EB Differentiation
XF Medium (without bFGF, with LiCl)
hESC Line UM77-2 EB Differentiation
XF Medium (without bFGF, with LiCl)
% Colonies
100
80
60
40
20
0
hESC P18
Pre-EB
EB Day 3
EB Day 6
EB Day 9
% XIST RNA-coated nuclei/colony
100%
80-99%
60-79%
40-59%
20-39%
1-19%
0%

Figure 4.7. Analysis of XIST RNA coating during differentiation of female hESCs (A) Schematic of hESC differentiation into embryoid bodies (EBs) with three different media formulations: a commercially available AggreWell™ medium; XF medium lacking BFGF; and, XF medium lacking bFGF but containing 0.98mM LiCl. (B) Percentage of nuclei with XIST RNA coating in EBs generated from hESC lines UM77-2 and UM63-1. EBs generated and cultured in XF medium with LiCl and AggreWell™ medium lost a significant proportion of XIST RNA coating per colony compared to EBs generated and cultured in XF medium (general linear model comparison; $p < 0.001$). At least 100 nuclei were counted per colony for hESC RNA FISH quantification. The total number of colonies quantified at each passage range between 10-17.

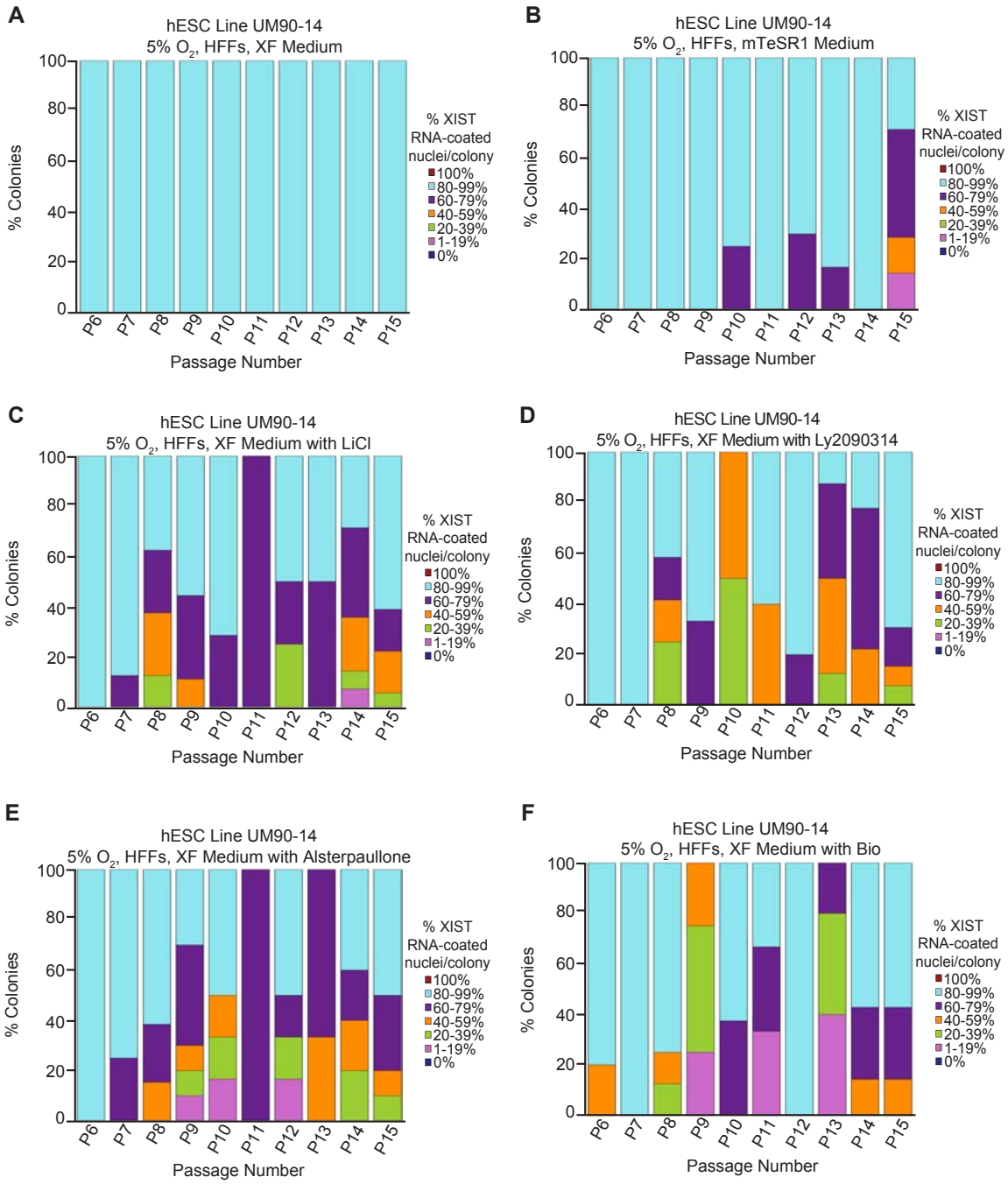

A
hESC Line UM90-14
5% O2, HFFs, XF Medium
B
hESC Line UM90-14
5% O2, HFFs, mTeSR1 Medium
C
hESC Line UM90-14
5% O2, HFFs, XF Medium with LiCl
D
hESC Line UM90-14
5% O2, HFFs, XF Medium with Ly2090314
E
hESC Line UM90-14
5% O2, HFFs, XF Medium with Alsterpaullone
F
hESC Line UM90-14
5% O2, HFFs, XF Medium with Bio
% Colonies
Passage Number
P6 P7 P8 P9 P10 P11 P12 P13 P14 P15
% XIST RNA-coated nuclei/colony
100%
80-99%
60-79%
40-59%
20-39%
1-19%
0%

Figure 4.8. GSK-3 inhibition and loss of XIST RNA coating in female hESCs (A-F) Percentage of nuclei with XIST RNA coating in colonies of hESC line UM90-14 cultured in XF medium (A); mTeSR1 medium (B); XF medium supplemented with 0.98mM LiCl (XF with LiCl) (C); XF medium supplemented with 1.5 nM Ly2090314 (XF with Ly2090314) (D); XF medium supplemented with 5.0 nM Alsterpaullone (XF with Alsterpaullone) (E); and, XF medium supplemented with 5.0 nM BIO (XF with Bio) (F). hESCs cultured in mTeSR1, XF with LiCl, XF with Ly2090314, XF with Alsterpaullone, and XF with BIO media lost XIST RNA coating during passaging in a significant percentage of nuclei compared to hESCs cultured in XF medium (general linear model comparison, $p < 0.001$). All hESCs in this experiment were cultured in 5% O_2 on HFFs. At least 100 nuclei were counted per colony for hESC RNA FISH quantification. The total number of colonies quantified at each passage range between 1-18.

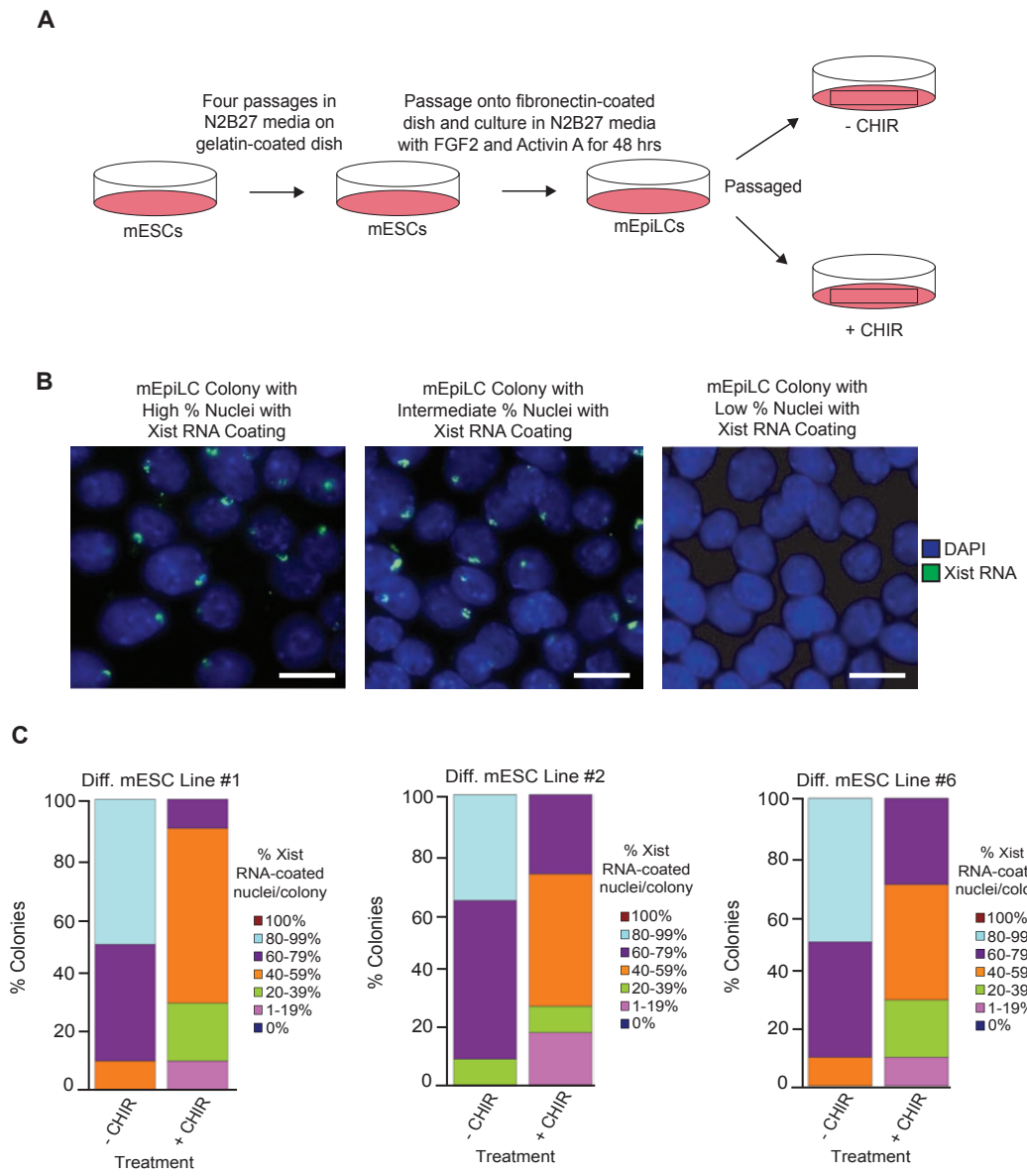


Figure 4.9. GSK-3 Inhibition and loss of *Xist* RNA coating in differentiating female mESCs
(A) Strategy for the differentiation of mESC lines into mEpiLCs and culture of mEpiLCs with and without the GSK-3 inhibitor CHIR99021 (CHIR; 3 μM). (B) Representative images of mEpiLCs with high, intermediate, and low percent of nuclei with *Xist* RNA coating (green). Nuclei are stained blue with DAPI. At least 10 nuclei were counted per colony in mEpiLC RNA FISH quantification. Scale bars are ~100 microns. (C) Percentage of nuclei with *Xist* RNA coating in differentiating mEpiLCs with and without CHIR generated from three independent ESC lines. mEpiLCs cultured with CHIR lost a significant proportion of *Xist* RNA coating compared to mEpiLCs cultured without CHIR in all three mEpiLC replicates (general linear model comparison, $p < 0.001$). At least 10 nuclei were counted per colony for mEpiLC RNA FISH quantification. The total number of colonies quantified at each passage range between 12-20.

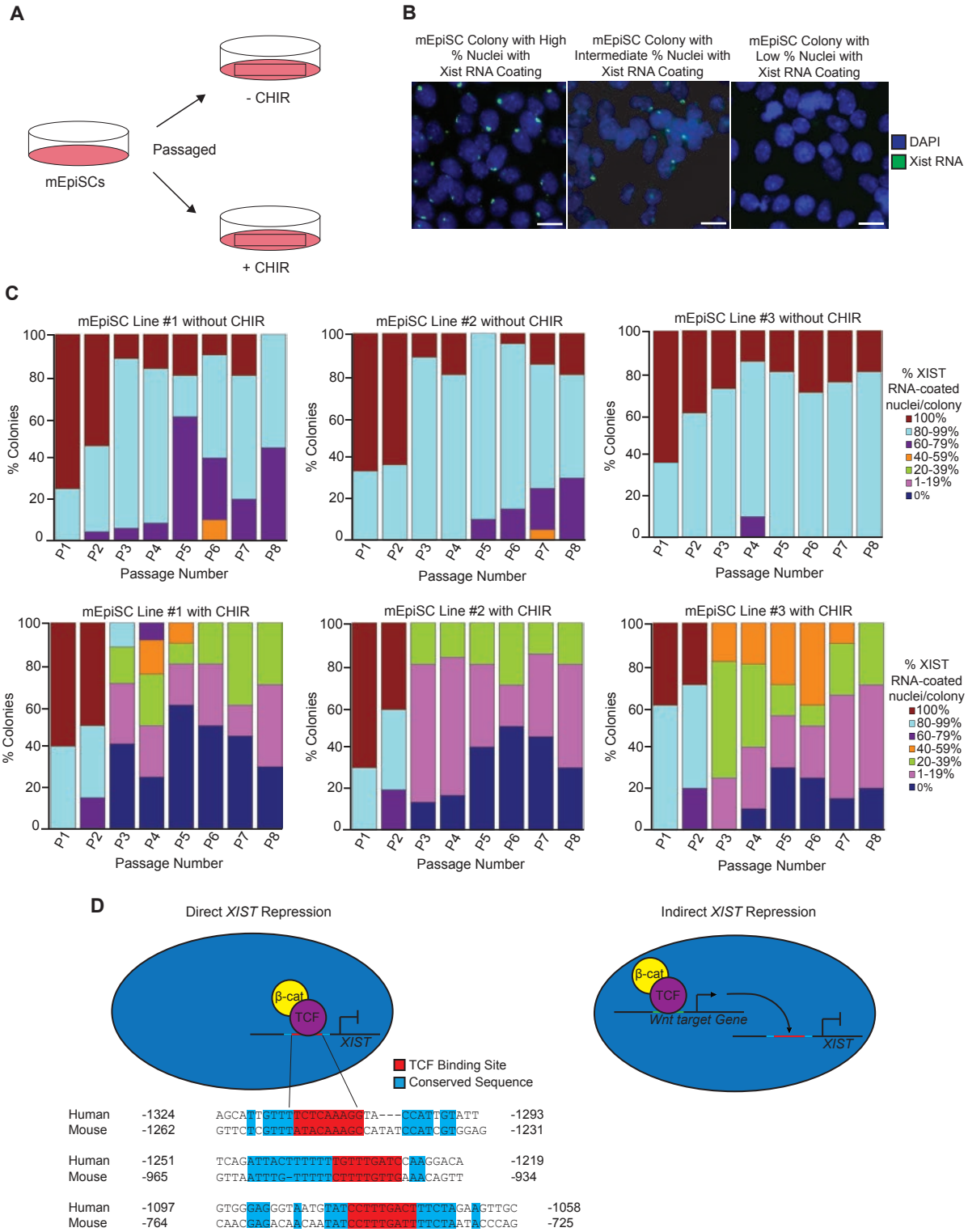
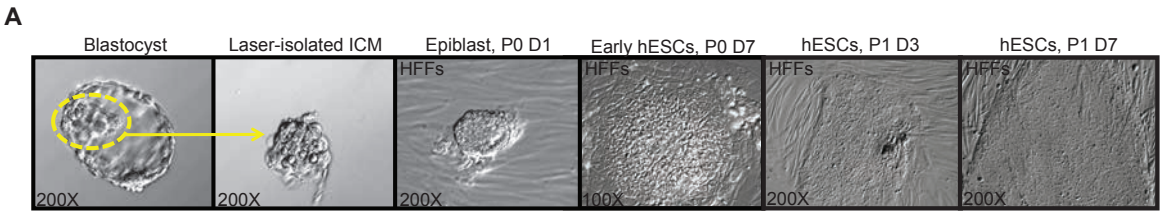
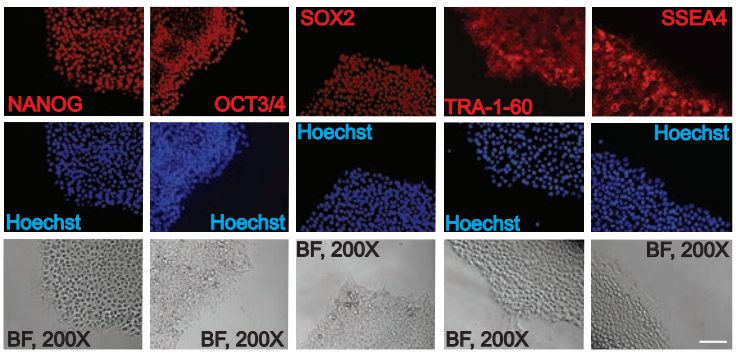
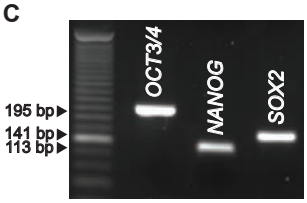

A
- CHIR
Passaged
mEpiSCs
+ CHIR
B
mEpiSC Colony with High % Nuclei with Xist RNA Coating
mEpiSC Colony with Intermediate % Nuclei with Xist RNA Coating
mEpiSC Colony with Low % Nuclei with Xist RNA Coating
DAPI
Xist RNA
C
mEpiSC Line #1 without CHIR
mEpiSC Line #2 without CHIR
mEpiSC Line #3 without CHIR
mEpiSC Line #1 with CHIR
mEpiSC Line #2 with CHIR
mEpiSC Line #3 with CHIR
% Colonies
Passage Number
P1 P2 P3 P4 P5 P6 P7 P8
% XIST RNA-coated nuclei/colony
100%
80-99%
60-79%
40-59%
20-39%
1-19%
0%
D
Direct XIST Repression
Indirect XIST Repression
β-cat
TCF
XIST
Wnt target Gene
TCF Binding Site
Conserved Sequence
Human -1324 AGCATTGTTTTCTCAAAGGTA---CCATTGTATT -1293
Mouse -1262 GTTCTTCGTTTATACAAAGCCATATCCATCGTGGAG -1231
Human -1251 TCAGATTACTTTTTTTGTTTGATCCAAGGACA -1219
Mouse -965 GTTAATTTG-TTTTTCTTTTGTTGAAACAGTT -934
Human -1097 GTGGGAGGGTAATGTATCCTTTGACTTTCTAGAAGTTGC -1058
Mouse -764 CAACGAGACAACAATATCCTTTGATTTTCTAATACCCAG -725

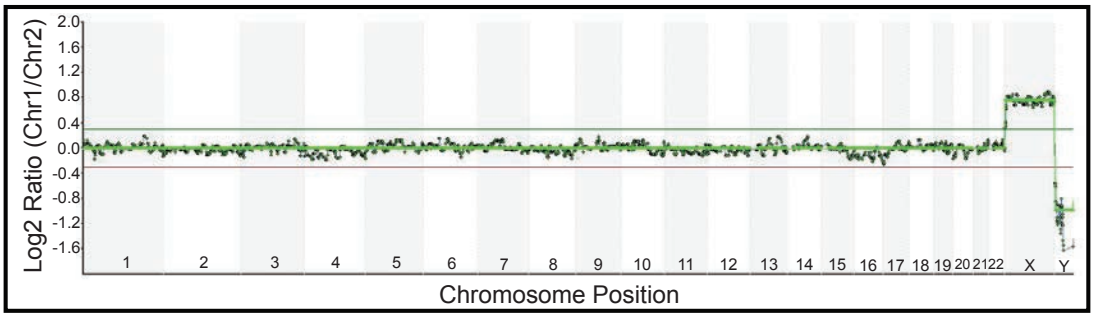
Figure 4.10. GSK-3 inhibition and loss of *Xist* RNA coating in female mEpiSCs (A) Strategy to test the impact of GSK-3 inhibition on *Xist* RNA coating in mEpiSCs. (B) Representative RNA FISH images of colonies with high, intermediate, and low percentage of *Xist* RNA coating (green). Nuclei are stained blue with DAPI. At least 10 nuclei were counted per colony in mEpiSC RNA FISH quantification. Scale bars are ~100 microns. (C) Percentage of nuclei with *Xist* RNA coating in three independent mEpiSC lines cultured with and without the GSK-3 inhibitor CHIR99021 (CHIR; 3 μM). All three mEpiSC lines cultured with CHIR lost a significant proportion of *Xist* RNA coating compared to mEpiSCs cultured without CHIR (general linear model comparison, $p < 0.001$). (D) Model of direct or indirect repression of human XIST and mouse Xist expression. Conserved mouse and human *Xist/XIST* sequences upstream of the *Xist/XIST* TSSs shown, with putative TCF binding motifs in red and surrounding conserved sequence in blue. At least 10 nuclei were counted per colony for mEpiSC RNA FISH quantification. The total number of colonies quantified at each passage range between 15-25.

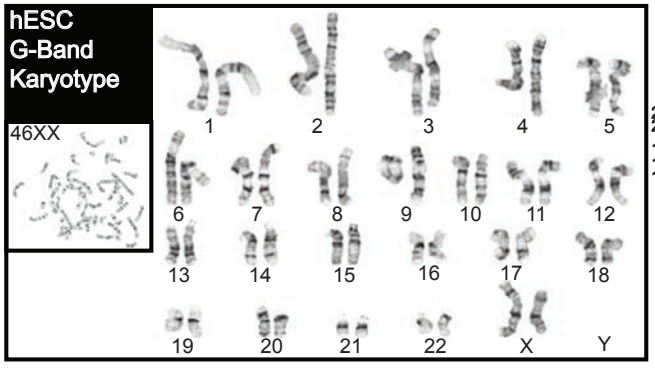


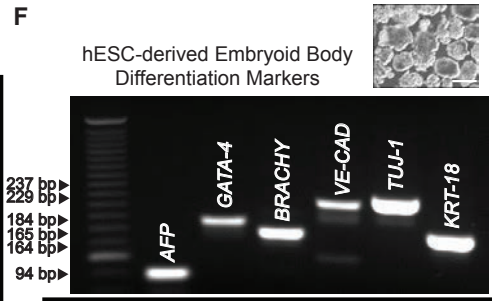




qPCR for hESC Pluripotency Markers	
Transcript	**CT***
NANOG	6.98
OCT3/4	5.25
SOX2	7.16
*CT calculated in relation to *ActB*	

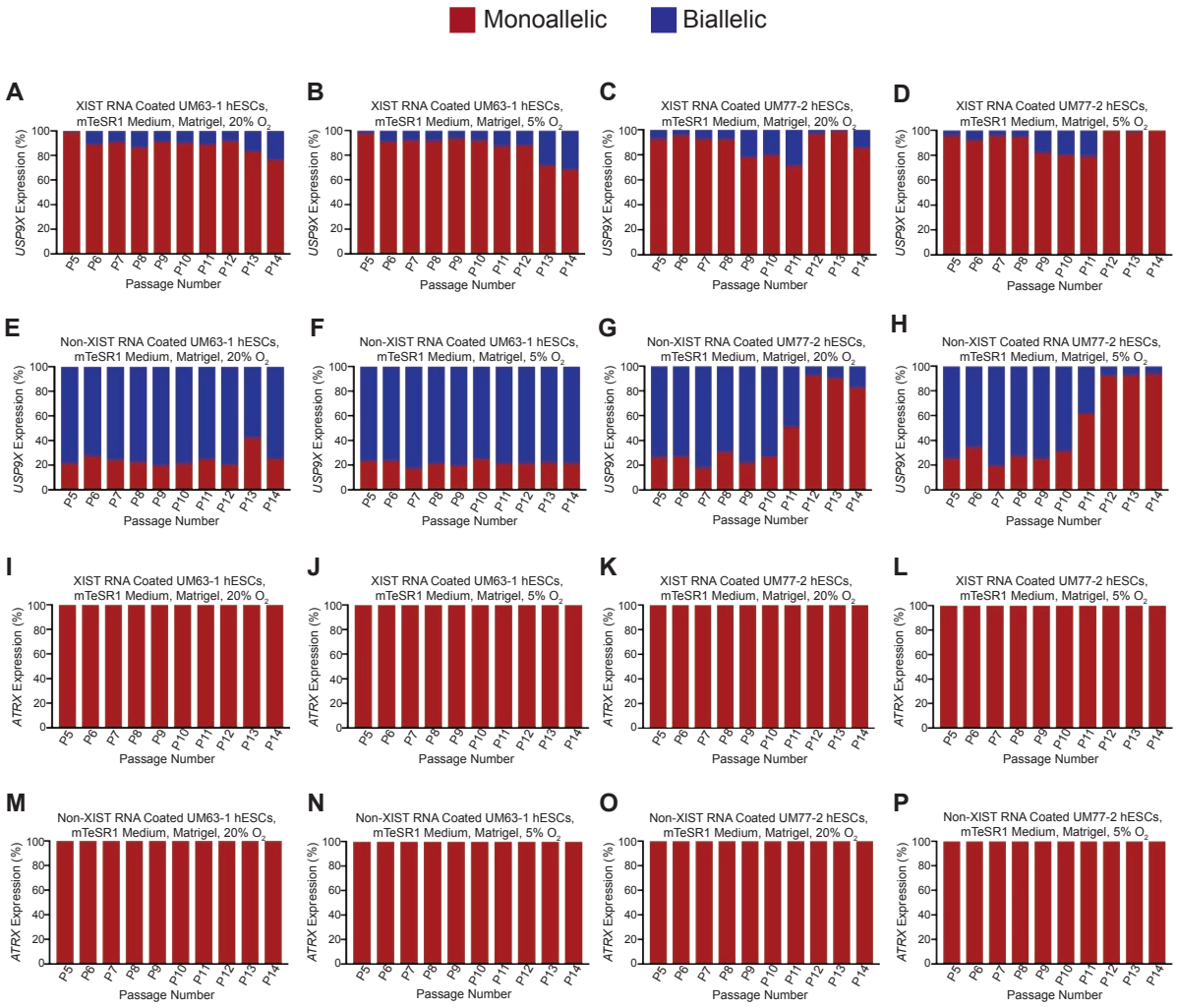




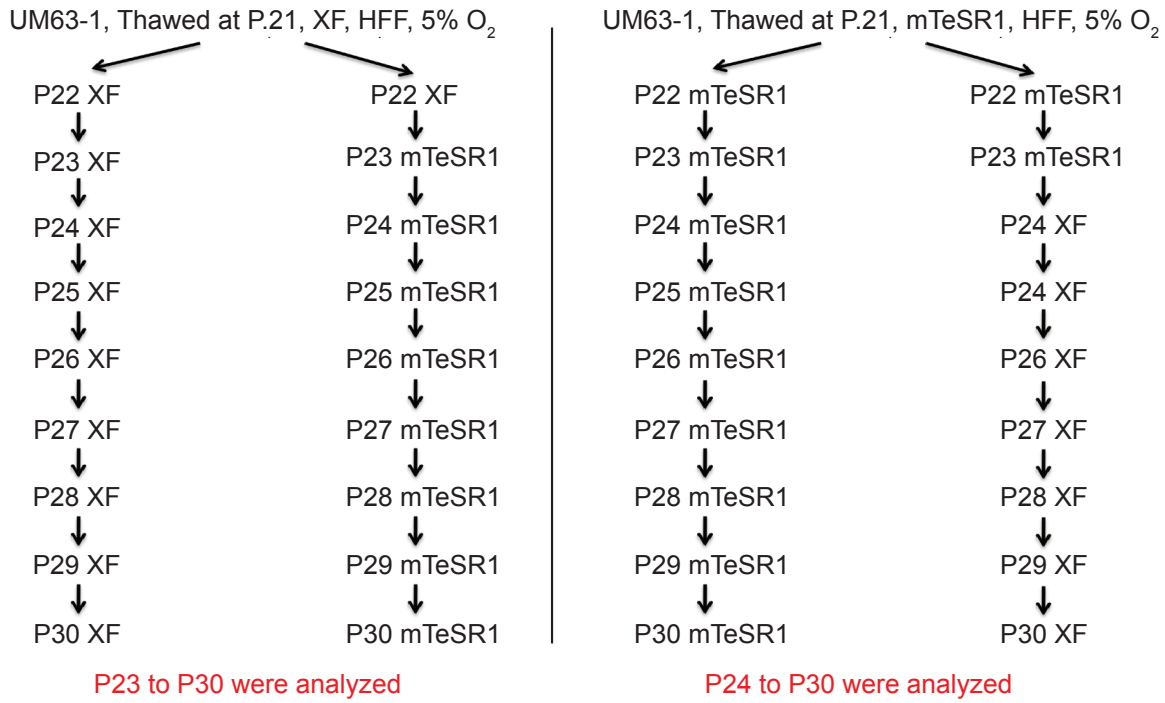


qPCR for D8 hESC-derived EB Differentiation Markers	
Transcript	**CT***
AFP (endoderm)	15.26
GATA-4 (endoderm)	17.38
BRACHY (mesoderm)	14.07
VE-CAD (mesoderm)	17.52
TUJ-1 (ectoderm)	6.56
KRT-18 (ectoderm)	5.23

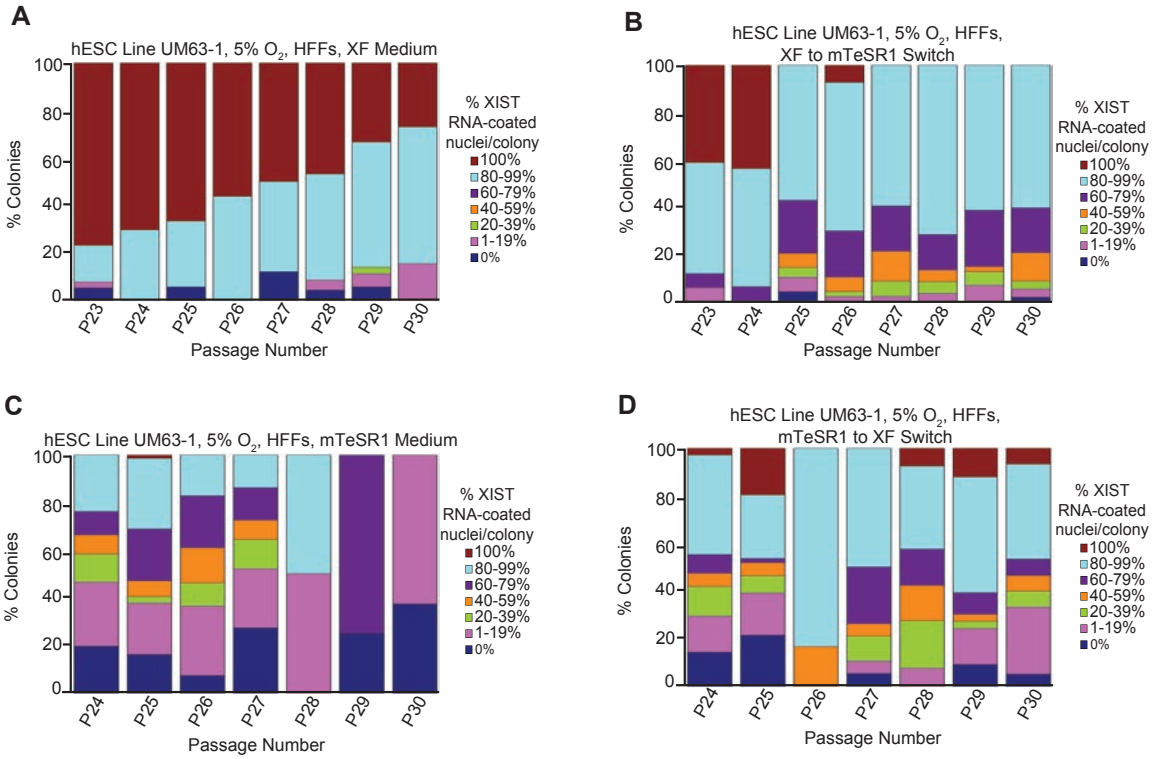
Supplemental Figure 4.1. Derivation and characterization of hESCs (A) Representative micrographs of day (D) 5 human blastocyst embryos following thawing; laser dissection of inner cell mass (ICM); plating of the ICM on human foreskin fibroblast feeder cells (HFFs) resulting in epiblast outgrowth; and, the resultant derivation of hESC colonies. Four blastocysts were used to generate four independent hESC lines for this study; micrographs were generated for each embryo and hESC line. (B) Representative micrographs of hESCs stained by immunofluorescence to detect pluripotency markers NANOG, OCT3/4, SOX2, TRA-1-60, and SSEA4 (red) with Hoechst nuclear staining (blue), and brightfield (BF) images to assess morphology. All four independent hESC lines generated for this study were subjected to immunofluorescent staining. Scale bar ~50 microns. (C) Top, RT-qPCR detection of the pluripotency markers *OCT3/4*, *NANOG*, and *SOX2*. Bottom, relative quantitation by RT-qPCR of *NANOG*, *OCT3/4*, and *SOX2* in relation to *β-ACTIN* (*ACTB*). RT-qPCR was conducted in triplicate for every transcript in each hESC line. Collectively these data (b, c top and bottom) demonstrate reproducibility of protein/transcript assays showing hESC line pluripotency. (D) Representative low passage number 24-chromosome molecular/microarray demonstrating normal karyotype (46,XX) of a female hESC line. (E) Representative G-band karyotype of metaphase spreads of a female hESC line. (F) Representative micrograph of hESC-derived embryoid bodies (EB, Day 8, inset, scale bar ~100 microns) and RT-qPCR detection of lineage markers (*AFT*, *GATA-4* – two independent endoderm markers; *BRACHY*, *VE-CAD* – two independent mesoderm markers; and *TUJ-1*, KRT-18 – two independent ectoderm markers). Bottom, relative quantitation by RT-qPCR of lineage markers in hESC-derived EBs normalized to *ACTB*. Collectively these data (F, top and bottom) demonstrate reproducibility of the transcription assays of the ability of hESC line-derived EBs to form endoderm, mesoderm, and ectoderm. RT-qPCR was conducted in triplicate for each transcript for every hESC line. Each hESC line employed in this study was subjected to these analyses and only the female hESC lines with the expected complement of autosomes and X chromosomes were analyzed further.



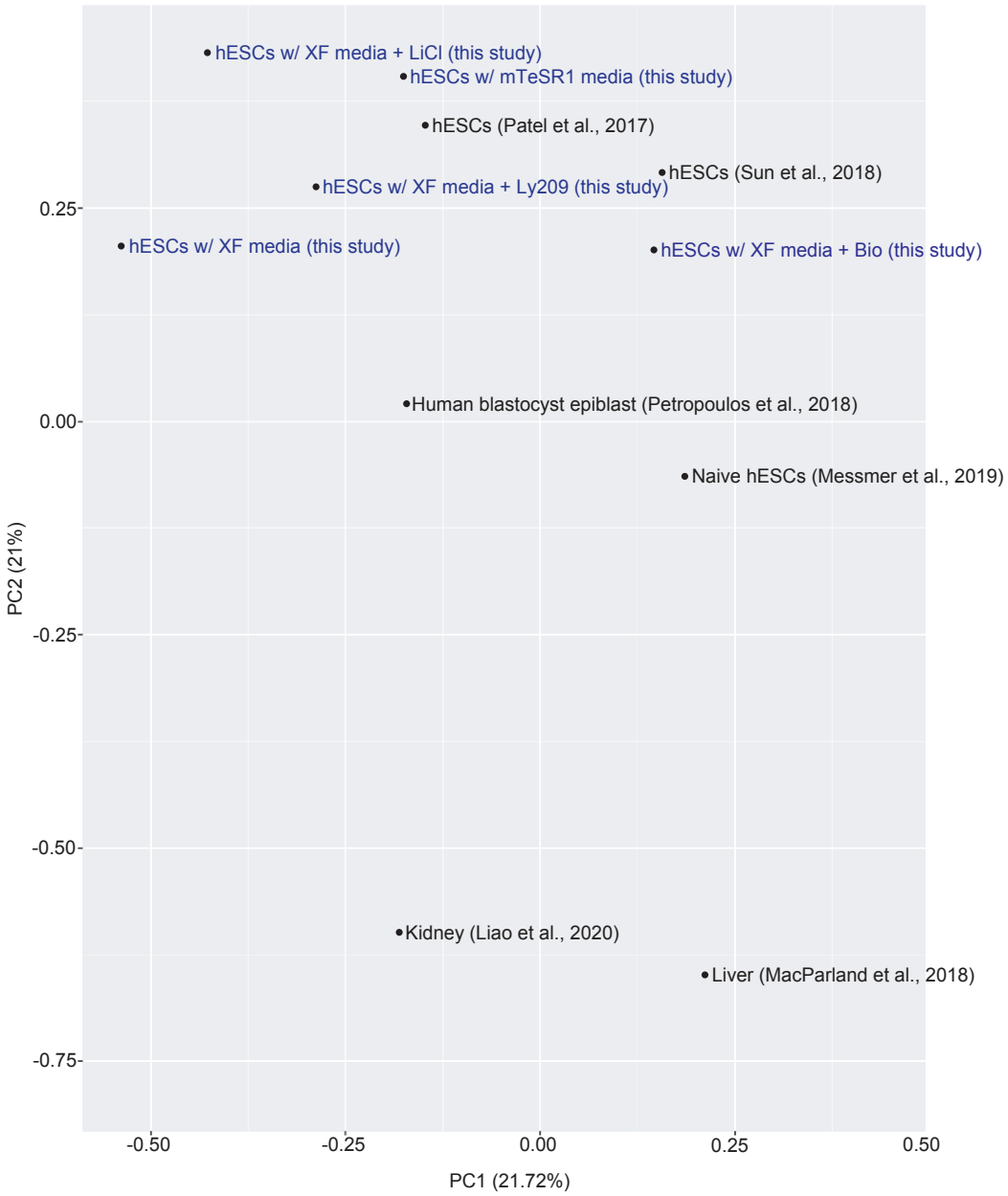
Supplemental Figure 4.2. Impact of atmospheric (20%) and physiological (5%) O_2 concentration on expression of X-linked genes USP9X and ATRX in female hESCs (A-B) Expression pattern of nascent *USP9X* RNA in nuclei with XIST RNA coating in UM63-1 hESCs cultured under 20% (A) or 5% (B) O_2 detected by RNA FISH. (C-D) As in A-B, but with UM77-2 hESCs. (E-F) Expression pattern of nascent *USP9X* RNA in nuclei without XIST RNA coating in UM63-1 hESCs cultured under 20% (E) or 5% (F) O_2. (G-H) As in E-F, but with UM77-2 hESCs. (I-J) Expression pattern of nascent *ATRX* RNA in nuclei with XIST RNA coating in UM63-1 hESCs, cultured under 20% (I) or 5% (J) O_2. (K-L) As in I-J, but with UM77-2 hESCs. (M-N) Expression pattern of nascent *ATRX* RNA in nuclei without XIST RNA coating in UM63-1 hESCs cultured under 20% (M) or 5% (N) O_2. (O-P) As in M-N, but with UM77-2 hESCs. See also Figure 4.2.



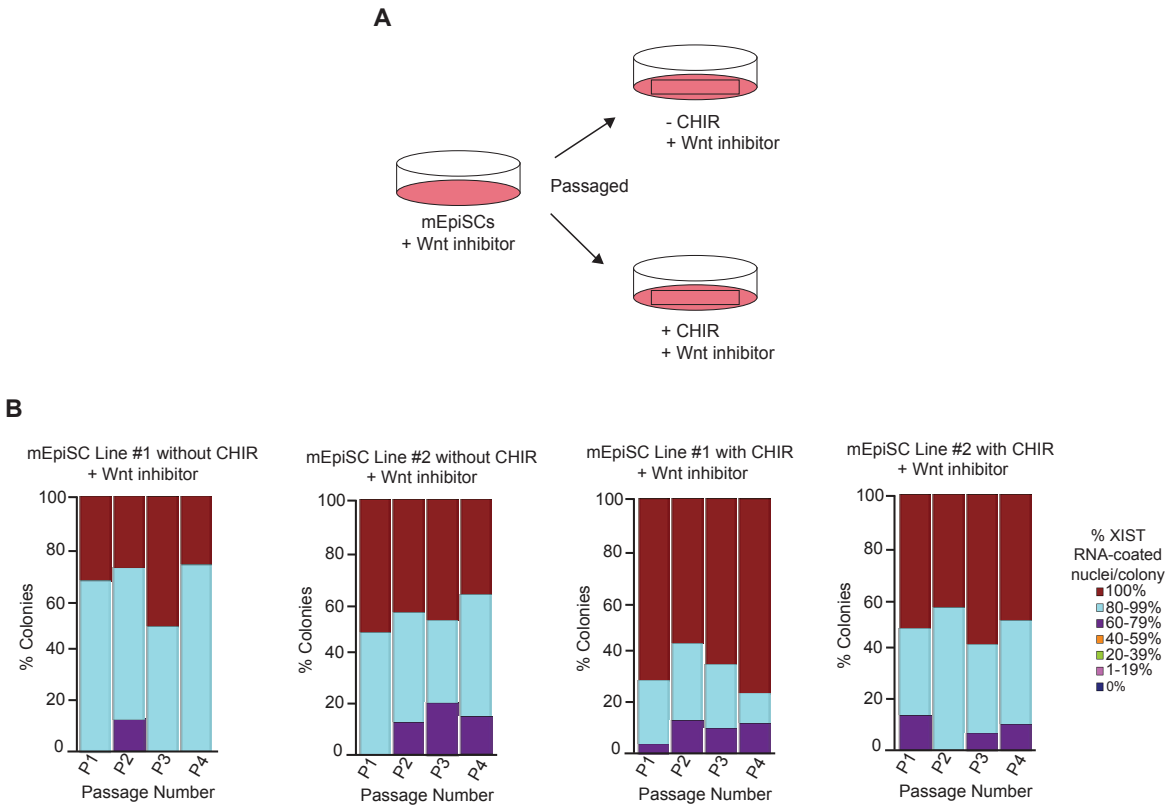
Supplemental Figure 4.3. Strategy for culture media switch experiment in Figure 4.5



Supplemental Figure 4.4. Detailed analysis of culture media switching on XIST RNA coating stratification of XIST RNA FISH data from Figure 5 into seven categories of percentage of nuclei with XIST RNA coats per colony of hESC line UM63-1 (A) cultured continuously in XF medium and (B) cultured initially in XF medium and subsequently switched to mTeSR1 medium. Percentage of nuclei with XIST RNA coats per colony of hESC line UM63-1 (C) continuously cultured in mTeSR1 medium and (D) cultured initially in mTeSR1 medium and then switched to XF medium. hESCs cultured initially in XF medium and subsequently switched to mTeSR1 medium displayed a significant decrease in the proportion of nuclei with XIST RNA coating per colony during passaging compared to hESCs cultured continuously in XF medium (general linear model comparison, $p = 0.01$). hESCs cultured continuously in mTeSR1 medium displayed a significant decrease in nuclei with XIST RNA coating during passaging compared to those cultured initially in mTeSR1 medium and then switched to XF medium (general linear model comparison, $p < 0.001$). See also Figure 4.5 and Supplemental Figure 4.3.



Supplemental Figure 4.5. Transcriptome comparison of hESCs, human blastocyst epiblast, and differentiated cell types. Principal component analysis of RNA-Seq data generated from P28 hESC line UM90-14 cultured using the following media formulations: XF medium; mTeSR1 medium; XF medium with 0.98mM LiCl; XF medium with 1.5 nM Ly2090314; XF medium with 5.0 nM BIO. Data generated for this study was compared to published RNA-Seq datasets (Messmer *et al.*, 2019; Patel *et al.*, 2017; Petropoulos *et al.*, 2016; Sun *et al.*, 2018).



Supplemental Figure 4.6. mEpiSCs cultured with Wnt inhibitors and GSK-3 inhibitors maintain *Xist* RNA coating (A) Strategy to test the impact of GSK-3 inhibition coupled with Wnt inhibition on *Xist* RNA coating in mEpiSCs. (B) Percentage of nuclei with *Xist* RNA coating in two independent mEpiSC lines cultured with and without the GSK-3 inhibitor CHIR99021 (CHIR; 3 μM) and with Wnt inhibitor (IWP-2; 2 μM). None of the mEpiSC lines lost a significant proportion of *Xist* RNA coating through passaging.

Chapter 5

Differential Roles for *Xist* RNA vs. *Xist* DNA in X-Chromosome Inactivation

Abstract

X-chromosome inactivation is a paradigm of epigenetic transcriptional regulation that results in the silencing of genes on one of the two X chromosomes in female mammalian cells. X-inactivation is thought to be controlled by the long noncoding (lnc) X-inactive specific transcript (*Xist*). *Xist* RNA is expressed solely from the inactive X chromosome and is thought to trigger gene silencing by recruiting protein complexes to the inactive-X. We previously found female trophoblast stem cells (TSCs) devoid of the Polycomb group protein EED lacked *Xist* RNA expression. Despite the absence of *Xist* RNA, many genes remained silenced on the inactive X chromosome in *Eed*$^{-/-}$ TSCs. This observation contrasted with previous findings in other cells where deletion of *Xist* activates silenced X-linked genes. To distinguish the role of *Xist* RNA from that of the underlying *Xist* genomic DNA sequence, we deleted most of the *Xist* genomic locus in female TSCs. Our data demonstrate that *Xist* genomic deletion reactivates X-linked genes that remain silenced upon loss of *Xist* RNA expression in *Eed*$^{-/-}$ TSCs. These results thus suggest that the *Xist* locus may silence X-linked genes by mechanisms other than via the production of *Xist* RNA.

Introduction

X-chromosome inactivation is a dosage compensation mechanism that equalizes X-linked gene expression between XX female and XY male mammals via transcriptional silencing of genes on one of the two X chromosomes in early female embryos (Lyon, 1961). Early studies of X-

chromosomal truncations and translocations in mouse embryos and embryonic cells suggested that a region on the X chromosome called the X-inactivation center (XIC) was necessary for X-inactivation (Lyon *et al.*, 1964; Russell, 1963). Cytological and molecular examinations later narrowed this region to ~1-2 mb (Brown *et al.*, 1991b) and this genomic segment became the focal point to identify sequence elements required for X-inactivation. The XIC houses the *Xist* gene (Brown *et al.*, 1991a; Marahrens *et al.*, 1997; Penny et al., 1996). *Xist* RNA is known to physically coat the prospective inactive-X in *cis* and recruit gene silencing proteins to that X chromosome (Brown et al., 1992; Clemson et al., 1996; Panning & Jaenisch, 1996; Moindrot & Brockdorff, 2015). Thus, *Xist* has been thought to be necessary for X-inactivation (Marahrens et al., 1997; Penny et al., 1996; Stavropoulos et al., 2001).

Female mice undergo two forms of X chromosome inactivation: imprinted and random. Imprinted X-inactivation initiates around the 2-cell stage of embryogenesis and is characterized by the inactivation of the paternally-inherited X chromosome (Kalantry et al., 2009; Namekawa *et al.*, 2010; Patrat *et al.*, 2009). Imprinted X-inactivation is subsequently maintained in the extra-embryonic tissues, the trophectoderm and the primitive endoderm lineages (Harper *et al.*, 1982; Takagi and Sasaki, 1975). In the epiblast of the blastocyst-stage embryo, the paternal-X is reactivated, and these cells subsequently undergo random inactivation of either the maternal- or paternal-X in individual cells. Epiblast cells that undergo random X-inactivation give rise to the embryonic tissues, whereas the extraembryonic tissues maintain imprinted X-inactivation throughout development. One of the extraembryonic lineages that stably maintains imprinted X-inactivation is the trophectoderm, from which trophoblast stem cells (TSCs) can be derived and utilized to study the maintenance of imprinted X-inactivation *in vitro*.

Functional studies of *Xist* have provided insight into the mechanisms underlying X-inactivation. The observation that *Xist* RNA is induced from and physically coats the X chromosome from which it is transcribed suggests that *Xist* RNA is a key factor in effecting X-linked gene silencing (Okamoto et al., 2004; Kalantry et al., 2009; Namekawa et al., 2010; Patrat et al., 2009; Mak et al., 2004; Rastan et al., 1982; MacMahon et al., 1983). In support of this claim, embryos that inherit a paternal X chromosome harboring an *Xist* mutation die at early post-implantation stages due to extra-embryonic developmental defects resulting from defective imprinted X-inactivation (Marahrens et al., 1997; Kalantry et al., 2009). Furthermore, epiblasts in $Xist^{+/-}$ embryos were found to have biased random X-inactivation such that all embryonic derived cells possessed a wild-type inactive X chromosome (Kalantry et al., 2009; Marahrens *et al.*, 1997). Similar results were obtained *in vitro* with mouse embryonic stem cells (mESCs) (Penny et al., 1996). Multicopy transgenes present on autosomes have also been shown to be sufficient to ectopically induce *Xist* expression (Wutz et al., 2002). Gene silencing in regions harboring these *Xist* transgenes has also been observed (Lee et al., 1996). Based on these findings and others, it has been widely believed that *Xist* is both necessary and sufficient to induce X-inactivation.

Although *Xist* has been thought to be necessary and sufficient for initiating X-inactivation, some evidence suggests that *Xist* RNA is dispensable for X-inactivation in some contexts. For example, mouse embryos that inherit a paternal X chromosome harboring an *Xist* deletion were still able to silence a subset of X-linked genes during the pre-implantation phase of embryogenesis (Kalantry et al., 2009). However, these embryos display defects in the post-implantation maintenance phase of imprinted X-inactivation (Kalantry et al., 2009). Furthermore, *Xist* RNA expression is fully abrogated in $Eed^{-/-}$ trophoblast stem cells (TSCs),

which are an *in vitro* model of imprinted X-inactivation and express *Xist* RNA only from the paternal X chromosome. Despite this loss of *Xist* RNA expression, our group found that only a small subset of normally silenced paternal X-linked genes is derepressed from the inactive-X (19%) in the $Eed^{-/-}$ TSCs (Maclary *et al.*, 2017). These data support the findings from the characterization of a separate TSC line harboring a point mutation in the Polycomb protein EED (Kalantry et al., 2006). Taken together, these results suggest that *Xist* RNA may not be necessary for stable silencing of many genes on the inactive X chromosome. Thus, there are likely other factors at work that are integral to the X-inactivation process.

Although much work has pointed to an important role for *Xist* in X-inactivation, past studies examining *Xist* function have assumed that any observable defects resulting from *Xist* DNA perturbation must be due to the loss of canonical *Xist* RNA. However, some work by our group suggests that factors besides *Xist* RNA that are present within the *Xist* genomic locus may be responsible for gene silencing in imprinted X-chromosome inactivation (Maclary et al., 2017; Kalantry et al., 2006). Should *Xist* RNA be shown to not serve an important role in X-inactivation, this would not be the first instance where a lncRNA arising from a locus was mistakenly thought to be a functional entity (Selleri et al., 2016; Bassett et al., 2014). To test whether the *Xist* locus may serve an *Xist* RNA-independent function in imprinted X-inactivation, we generated female TSCs harboring a 17kb deletion within *Xist* on the paternal X chromosome ($Xist^{+/fl}$:Tam), which abrogates *Xist* RNA expression. We then compared allele-specific gene silencing in the *Xist*-mutant TSCs we generated with that in $Eed^{-/-}$ TSCs that lack *Xist* RNA expression but harbor an intact *Xist* locus. Our data support the notion that *Xist* RNA-independent factors that are encoded within the *Xist* locus are required to silence X-linked genes and maintain imprinted X-inactivation.

TSCs Lacking *Xist* RNA Display Minor Defects in Paternal X-linked Gene Silencing

Previous work by others in the Kalantry Lab demonstrated that female mouse TSCs lacking EED also lack H3K27me3 enrichment, *Xist* expression, and *Xist* RNA coating of the inactive-X (Kalantry *et al.*, 2006; Maclary *et al.*, 2017). Because these TSCs contain intact *Xist* loci but lack *Xist* RNA expression and coating, $Eed^{-/-}$ TSCs provide a valuable model for testing the role of *Xist* DNA vs. *Xist* RNA in imprinted X-inactivation. The $Eed^{-/-}$ TSCs previously generated by our group were hybrid, which allowed us to exploit single nucleotide polymorphisms (SNPs) to distinguish expression of genes from either the maternal or paternal X-chromosome using allele-specific profiling methods (Cloutier *et al.*, 2018; Maclary *et al.*, 2017). Interestingly, when examined by allele-specific RNA-Seq, $Eed^{-/-}$ TSCs displayed significant silencing defects in only a fraction (19%) of paternal X-linked genes, with the remainder of the genes maintaining silencing on the paternal-X despite the absence of *Xist* RNA (Figure 5.1) (Maclary *et al.*, 2017). These data therefore suggest that *Xist* RNA is dispensable for silencing most paternal X-linked genes in female mouse TSCs.

Broad De-repression of Paternal X-linked Genes in TSCs lacking *Xist* DNA

To test requirements for *Xist* RNA vs. *Xist* DNA in X-linked gene silencing, I generated hybrid female TSCs lacking a critical region of *Xist* using an inducible *Cre*-lox approach (Materials and Methods). The TSC lines used to delete *Xist* harbor *loxp* sites flanking paternal-X *Xist* exons 1-3, which can be excised via the induction of an endogenous *Cre* with Tamoxifen treatment (Materials and Methods). Notably, this deletion begins ~5kb upstream of the *Xist* TSS and has previously been shown to abrogate *Xist* RNA expression (Csankovszki et al., 1999). *Xist* exons 1-3 have previously been proposed to be crucial for *Xist*-mediated X-linked gene silencing (Csankovszki *et al.*, 1999; Marahrens *et al.*, 1997; Penny et al., 1996). Upon *Cre*-mediated

deletion in *Xist*, we could not establish a constitutive *Xist*$^{+/-}$ TSC line due to strong selection for cells that still harbored an intact paternal-X *Xist* allele. I therefore employed a transient *Xist* deletion strategy by inducing CRE through treatment of the *Xist*$^{+/fl}$ TSCs with 9uM Tamoxifen for 12 hours (Figure 5.2A). The 12-hour Tamoxifen treatment time was optimized through a series of time course experiments I conducted (Tables 5.1 and 5.2) and which yielded a significant number of nuclei with the *Xist* deletion (~30%). For simplicity, I will refer to Tamoxifen-treated *Xist*$^{+/fl}$ TSCs as *Xist*$^{+/fl}$:Tam. Upon generating these *Xist*$^{+/fl}$:Tam TSCs, I characterized X-linked gene expression in these cells and untreated *Xist*$^{+/fl}$ controls by fluorescence *in situ* hybridization (RNA FISH), Sanger sequencing, and allele-specific RNA-Seq.

RNA FISH analysis of *Xist* RNA coating in *Xist*$^{+/fl}$:Tam TSCs indicated that *Xist* was deleted in ~30% of nuclei, on average (Tables 5.1 and 5.2). In contrast, *Xist*$^{+/fl}$ control TSCs displayed ~6% of nuclei that lacked *Xist* RNA coating. In addition to quantifying *Xist* expression, I assessed expression of two other X-linked genes, *Rnf12* and *Atrx*, in *Xist*$^{+/fl}$:Tam TSCs and *Xist*$^{+/fl}$ controls (Figure 5.2B-D). Both *Rnf12* and *Atrx* are subject to X-inactivation and are expressed almost exclusively from the active-X and not from the inactive-X in TSCs (Maclary *et al.*, 2017). By RNA FISH, therefore, RNAs arising from the *Rnf12* and *Atrx* genes are normally detected as monoallelic signals in a vast majority of cells. Analysis of *Atrx* expression showed monoallelic expression in the majority of *Xist*$^{+/fl}$ control TSCs and significantly higher biallelic expression in both *Xist* RNA coated and non-*Xist* RNA coated *Xist*$^{+/fl}$:Tam TSCs (Figure 5.2B). I validated the biallelic expression of *Atrx* in *Xist*$^{+/fl}$:Tam TSCs by RT-PCR followed by Sanger sequencing in a region of *Atrx* that harbors a strain-specific SNP (Figure 5.2C). When *Rnf12* expression was assessed by RNA FISH, *Xist*$^{+/fl}$:Tam TSCs also

exhibited higher levels of biallelic expression in compared to *Xist*$^{+/fl}$ TSC controls, suggesting defective *Rnf12* silencing upon loss of *Xist* expression from the paternal-X (Figure 5.2D). These data indicate a failure of imprinted X-inactivation in a subset of *Xist*$^{+/fl}$:Tam TSCs, but not in *Xist*$^{+/fl}$ TSC controls.

Following characterization of *Xist*$^{+/fl}$ and *Xist*$^{+/fl}$:Tam TSCs by RNA FISH, I examined these cell lines by allele-specific RNA-Seq to quantify the relative allelic expression of many X-linked genes. Once again, the hybrid nature of these TSCs allowed for allele-specific expression analysis. Allele-specific RNA-Seq of the *Xist*$^{+/fl}$:Tam TSCs revealed a significant shift in the overall maternal:paternal allelic expression ratio when compared to *Xist*$^{+/fl}$ controls (Figure 5.3). Furthermore, when compared to the allelic expression ratio of *Eed*$^{-/-}$ TSCs, *Xist*$^{+/fl}$:Tam TSCs displayed a significantly greater number of genes (73%) to be de-repressed paternal-X (Figure 5.3). As a control, I also generated and characterized both Tamoxifen-treated and untreated *Xist*$^{fl/+}$ TSCs, with maternal-X harboring the floxed *Xist* allele (Materials and Methods). I refer to these cells as *Xist*$^{fl/+}$:Tam. Due to imprinted X-inactivation, *Xist* is not expressed from the maternal-X in TSCs (Maclary et al., 2017). Consistent with previous work (Mak *et al.*, 2004; Okamoto *et al.*, 2004), maternal-X *Xist* deletion in *Xist*$^{fl/+}$:Tam TSCs resulted in no significant change in paternal X-linked gene silencing (Figure 5.3). Taken together, these data suggest that loss of paternal-X *Xist* DNA in some *Xist*$^{+/fl}$:Tam TSCs leads to a much greater defect in imprinted X-inactivation than loss of *Xist* RNA expression in *Eed*$^{-/-}$ TSCs.

Xist$^{+/fl}$:Tam; _Eed_$^{-/-}$ TSCs Recapitulate Paternal-X Silencing Observed in _Xist_$^{+/fl}$:Tam TSCs

Given the significant difference in paternal X-linked gene silencing in *Eed*$^{-/-}$ TSCs vs. *Xist*$^{+/fl}$:Tam TSCs, I sought to generate TSC lines lacking both *Eed* and paternal-X *Xist* (Figure 5.4A). These *Xist*$^{+/fl}$:Tam; *Eed*$^{-/-}$ TSCs served as an important control because they allowed me

to observe the transcriptional consequence of *Eed* deletion in the same cell lines in which I could subsequently delete *Xist*. To generate *Xist*$^{+/fl}$; *Eed*$^{-/-}$ cells, I used a CRISPR-Cas9 approach to induce a frameshift mutation in exon 7 of *Eed* in the *Xist*$^{+/fl}$ TSC lines characterized above (Materials and Methods). *Eed* exon 7 harbors a WD40 domain that is essential for proper EED function (Montgomery et al., 2005; Sathe and Harte, 1995). I confirmed *Eed* deletion by PCR, Sanger sequencing, and RNA-Seq (Figure 5.4B-D). Furthermore, I coupled RNA FISH with Immunofluorescent (IF) staining to show that EED, H3K27me3 enrichment, and *Xist* expression and RNA coating was lost in *Xist*$^{+/fl}$; *Eed*$^{-/-}$ TSCs but not in *Xist*$^{+/fl}$ controls (Figure 5.4E). The loss of H3K27me3 enrichment, *Xist* expression, and *Xist* RNA coating in *Xist*$^{+/fl}$; *Eed*$^{-/-}$ TSCs recapitulated the pattern observed in the *Eed*$^{-/-}$ TSCs generated by Maclary *et al.*, 2017. Furthermore, characterization of allelic expression by allele-specific RNA-Seq in *Xist*$^{+/fl}$; *Eed*$^{-/-}$ TSCs largely recapitulated the expression pattern observed in *Eed*$^{-/-}$ TSCs, with only a handful of genes becoming de-repressed from the paternal-X (Figure 5.5).

After confirming *Eed* deletion in the *Xist*$^{+/fl}$; *Eed*$^{-/-}$ TSCs, I treated these cells with Tamoxifen to remove exons 1-3 of *Xist* via *Cre*-mediated deletion. Again, I was unable to obtain a constitutive *Xist* knockout line, but *Xist* deletion still occurred in ~30% of cells. Allele-specific RNA-Seq of these *Xist*$^{+/fl}$:Tam; *Eed*$^{-/-}$ TSCs showed that their X-linked allelic expression ratio recapitulated the ratios observed in *Xist*$^{+/fl}$:Tam TSCs, suggesting no cell line-specific silencing effect. Taken together, these experiments highlight a differential role for *Xist* DNA vs. *Xist* RNA in maintaining imprinted X-inactivation in female mouse TSCs.

Discussion

Here I show that *Xist* RNA is not functionally equivalent to *Xist* DNA in the maintenance of imprinted mouse X-chromosome inactivation. Prior work by Maclary *et al.*, 2017 showed that

Xist RNA loss is dispensable to maintain repression of most paternal X-linked genes in female mouse TSCs. Only a fraction (19%) of genes is upregulated from the paternal-X when *Xist* RNA is absent. However, characterization of $Xist^{+/fl}$:Tam TSCs illustrate that if exons 1-3 of *Xist* DNA are deleted, many genes on the previously inactivated X-chromosome become de-repressed. I found that ~four times as many genes are de-repressed on the paternal-X in $Xist^{+/fl}$:Tam TSCs compared to $Eed^{-/-}$ TSCs, pointing to an essential role for other factors contained within *Xist* DNA in maintaining paternal X-linked gene silencing. Taken together, these data suggest additional functions for the *Xist* locus which contribute to X-inactivation independently of *Xist* RNA.

While conducting this study, I was also unable to generate a constitutive $Xist^{+/fl}$:Tam TSC line, which prevented me from assaying X-linked gene expression in a cell line fully lacking *Xist* DNA. Future work will further dissect the roles for *Xist* RNA vs. *Xist* DNA, including generating and characterizing allele-specific X-linked gene expression in a constitutive $Xist^{+/-}$ TSC line and investigating roles for additional transcripts arising from the *Xist* locus in controlling imprinted X-inactivation.

Conclusion and Future Directions

The work outlined in this chapter establishes the nonequivalence of *Xist* DNA and *Xist* RNA in imprinted X-chromosome inactivation. However, the reason(s) underlying this nonequivalence remain a mystery. A potential explanation for the functional difference between *Xist* DNA and *Xist* RNA may lie in transcripts that arise from the *Xist* locus that were deleted in this study and other studies investigating *Xist*. Using single-stranded RNA FISH probes tiling across *Xist*, our group has identified three noncoding RNAs that are embedded within and expressed in the antisense orientation from the *Xist* locus (Sarkar *et al*., 2015; Kalantry Lab, unpublished work).

These RNAs have been validated by RT-PCR, although we have not been able to detect them by RNA-Seq because they may be chromatin-bound and not easily purified, they are repetitive in nature and may not map correctly, or they may be lowly transcribed. Furthermore, although these RNAs overlap the *Xist* locus, the double-stranded RNA FISH probe we use to detect *Xist* is not sensitive enough to detect these antisense RNAs transcribed from the *Xist* locus. The first of these noncoding RNAs, *XistAR*, likely contributes to inducing *Xist* expression and has been characterized by others our group (Sarkar *et al.*, 2015). Roles for the two additional *Xist*-encoded transcripts have not yet been thoroughly investigated.

The first uncharacterized transcript within *Xist* the Kalantry Lab has identified, *Xist antisense transcript 2* (*XistAS2*), is transcribed exclusively from the inactive-X and has been found to interact in *cis* with the inactive X chromosome in wild-type female TSCs. Its adjacent localization to the *Xist* RNA domain and its ability to physically interact with the inactive-X (unpublished data, Kalantry Lab) suggest that *XistAS2* may participate gene silencing on the X chromosome, perhaps separately from *Xist* RNA. Although this transcript can be detected using RNA FISH and RT-PCR, it undetectable by current RNA-Seq approaches. The reason for this could be that *XistAS2* is either lowly transcribed or bound to chromatin and is therefore not easily isolated. A potential role for this transcript in X-linked gene silencing should be investigated by genetic perturbation followed by sequencing and cell imaging analyses considering the findings outlined in this chapter.

The second uncharacterized transcript our group identified is *Xist antisense transcript 3* (*XistAS3*). Others in the lab preliminarily found that *XistAS3* begins in the last *Xist* exon and ends at the 3'end of *Xist* (unpublished data, Kalantry Lab). Intriguingly, preliminary data suggest that *XistAS3* has the same splice structure as *Xist*, but in the antisense orientation,

although further validation is required to determine if this structure of *XistAS3* is accurate. Ultimately, like the other novel ncRNAs our group identified, *XistAS3* expression should be perturbed genetically and analyzed using molecular sequencing and staining techniques to elucidate its potential role in X-inactivation.

To test the function for each of the above uncharacterized *Xist*-encoded RNAs, targeted inhibition approaches should be undertaken to disrupt each of the RNAs without disturbing *Xist* expression or the underlying *Xist* locus. The deletion of *Xist* exons 1-3 outlined in this chapter also fully ablated *XistAR* and *XistAS2*, so we were unable to distinguish between roles for the *Xist* DNA locus and these novel *Xist*-encoded transcripts in maintaining imprinted X-inactivation. Targeting these novel transcripts using Cas13 or homologous recombination approaches coupled with characterization of allele-specific X-linked gene expression will allow our group to determine if any of these *Xist*-encoded transcripts play a role in the maintenance of imprinted X-inactivation. These approaches will allow us to target the antisense RNAs within *Xist* without disrupting *Xist* RNA expression. If any of these transcripts do play a role in paternal X-linked gene silencing, one would expect a significant shift in the allelic ratio of maternal:paternal X-linked gene expression upon transcript inhibition.

In addition to testing the function of the *Xist*-encoded antisense RNAs in wild-type TSCs, it will be informative to test the expression of the *Xist* antisense transcripts in female mouse *Eed*$^{-/-}$ TSCs lacking *Xist* RNA expression. The work outlined in this chapter and in Maclary *et al*., 2017 thoroughly demonstrate that *Xist* RNA expression is ablated in *Eed*$^{-/-}$ female mouse TSCs. We hypothesize that the X-linked gene expression differences between cells lacking *Xist* RNA and those lacking *Xist* DNA and RNA may be due to the presence of antisense RNAs expressed from the *Xist* locus. However, we have not tested whether these antisense RNAs are expressed

in *Eed*$^{-/-}$ female mouse TSCs. Our group can test the expression of these antisense RNAs in *Eed*$^{-/-}$ TSCs by RT-PCR and RNA FISH, much like how we identified these transcripts. We can also test the function of these transcripts in *Eed*$^{-/-}$ TSCs using the approaches discussed above.

Should a role for any of these novel *Xist* antisense transcripts be discovered, further investigation of their contributions to the establishment, initiation, and maintenance phases of X-inactivation should be undertaken by profiling these transcripts in various cell types and embryo stages. Additionally, the role of *Xist* RNA vs. *Xist* DNA in these different X-inactivation phases should be investigated, as my work in this chapter has only addressed the distinction between *Xist* RNA and *Xist* DNA in TSCs. The findings I have outlined here open new avenues for the study of *Xist* and X-linked gene regulation, as prior work has only addressed roles for *Xist* DNA or RNA in different cellular and developmental contexts.

Materials and Methods

Ethics Statement

This study was performed in strict accordance with the recommendations in the guide for the Care and Use of Laboratory Animals of the National Institutes of Health. All animals were handled according to protocols approved by the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan (protocol #PRO00006455).

Mice

Mice harboring a conditional mutation in *Eed* were generated for a previous study (Maclary *et al.*, 2017) by the University of Michigan Transgenic Animal Model Core using *Eed*$^{tm1a(EUCOMM)Wtsi}$ targeted ES cells (EUCOMM). Briefly, embryonic stem cells (ESCs) were injected into blastocysts and implanted into pseudo-pregnant females. Mice with high percentages of chimerism were bred and assessed for germline transmission. To generate

homozygous *Eed* mutant mice harboring polymorphic X chromosomes, first, male and female mice on a B6 *Mus musculus* background carrying the conditional mutant allele for *Eed* were intercrossed ($Eed^{fl/+}$ x $Eed^{fl/+}$) to achieve homozygosity. To obtain mice conditionally mutant for *Eed* and on the JF1 *Mus molossinus* divergent background, $Eed^{fl/fl}$ males (B6 *Mus musculus* background) were bred to *WT* JF1 *Mus molossinus* females. This cross yielded F1 hybrid $Eed^{fl/+}$ males that possessed an X chromosome from the JF1 *Mus molossinus* background (X^{JF1}/Y). Such males were backcrossed to *WT* JF1 *Mus molossinus* females to derive $Eed^{fl/+}$ females that were a mix of B6 *Mus musculus* and JF1 *Mus molossinus* and harbored two X chromosomes from the JF1 *Mus molossinus* background (X^{JF1}/X^{JF1}). $Eed^{fl/+};X^{JF1}/X^{JF1}$ females were bred with $Eed^{fl/+};X^{JF1}/Y$ males to derive $Eed^{fl/fl};X^{JF1}/Y$ males. To obtain female embryos to be used for TSC derivation, an $Eed^{fl/fl}$ female on the B6 *Mus musculus* background was crossed with an $Eed^{fl/fl}$ male that was a mix of B6 *Mus musculus* and JF1 *Mus molossinus* but possessed an X chromosome from the JF1 *Mus molossinus* background (X^{JF1}/Y). The JF1/*Mus molossinus* strain has been described previously. $Xist^{+/fl};X^{GFP}/Y$ *M. musculus* males (maintained on a 129 background) and JF1 *M. molossinus* females were bred in house.

TS Cell Derivation and Culture

Blastocysts were dissected from pregnant mice at embryonic day (E) 3.5 and plated in four well dishes pre-seeded with mouse embryonic fibroblasts (MEFs). Hatched embryos were cultured in standard TSC medium supplemented with 1.5x FGF4 and Heparin for 4-5 days until blastocyst outgrowths reached ideal size. Blastocysts were then trypsinized in 0.05% Trypsin-EDTA, neutralized with TSC media supplemented with 1.5x FGF4 and Heparin, and cultured on MEFs in 96 well dishes. Once TSC lines were well established, XX/XY PCRs confirmed female lines and PCRs for *Eed* and *Xist* confirmed $Eed^{fl/fl};X^{Lab}/X^{JF1}$ and $Xist^{+/fl};X^{JF1}/X^{Lab}$ lines, respectively.

Cell lines were then cultured in standard TSC media supplemented with FGF4 and Heparin. RNA was harvested from TSCs using TRIzol (Invitrogen, #15596-018) and RT-PCR was performed as described below. For RNA-FISH, TSCs were split onto gelatin-coated glass coverslips and allowed to grow for 2-3 days. The cells were then permeabilized through sequential treatment with ice-cold cytoskeletal extraction buffer (CSK; 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, and 10 mM PIPES buffer, pH 6. 8) for 30 seconds, ice-cold CSK buffer containing 0.4% Triton X-100 (Fisher Scientific, #EP151) for 30 seconds, followed twice with ice-cold CSK for 30 seconds. After permeabilization, cells were fixed by incubation in 4% paraformaldehyde at room temperature for 10 minutes. Cells were then rinsed three times each in 70% ethanol and stored in 70% ethanol at -20°C prior to RNA-FISH.

Generating Stable *Eed*$^{-/-}$ TSCs

Eed$^{fl/fl}$ TSCs were plated at a 1:24-1:48 dilution into six well dishes pre-seeded with MEFs and allowed to adhere to the MEFs until the next day. Cells were then transduced with Ad5-CMV-Cre (Adenovirus serotype type 5, University of Michigan Viral Vector Core adenoviral construct, 4 x 10^{12} particles/mL) at multiplicity of infection (MOI) of 1000. Once cell colonies were large enough following the initial transduction, they were subcloned into 96 well dishes pre-seeded with MEFs and re-transduced 24 hours later with Adeno-Cre at a MOI of 1000. Following this, expanded 96 well samples were split to six well dishes pre-seeded MEFs and again transduced 24 hours later. A portion of each 96 well samples was lysed for DNA genotyping to assess the efficiency of *Cre*-mediated deletion of the *Eed* floxed alleles. Subcloning, transduction, and genotyping procedures were repeated until a pure population of *Eed*$^{-/-}$ TSCs was achieved. *Eed*$^{-/-}$ TSCs were maintained in culture as described above.

Generating *Xist*$^{+/fl}$:Tam TSCs

Xist$^{+/fl}$ TSCs were plated at a 1:24 dilution on gelatinized coverslips in six well dishes. Cells were transduced with Tamoxifen for 48 hours. Cells adhering to coverslips were then CSK-treated and fixed with 4% PFA and stored for immunofluorescence and/or RNA-FISH. The remaining adherent cells on the edges of each well of the six well dishes were washed once with 1 mL cold 1X PBS, followed by aspiration of PBS. Cells were then incubated in 1mL TRIzol at 4°C for 5 minutes. Lysates were stored in TRIzol at -80°C until RNA extraction. Of note, we discovered while optimizing this protocol that Tamoxifen loses its efficacy after being opened for >1 month, even when stored in proper conditions.

Generating *Xist*$^{+/fl}$:Tam; *Eed*$^{-/-}$ TSCs

Xist$^{+/fl}$ TSCs were plated on a 60mm dish containing puromycin-resistant MEFs obtained from the University of Michigan Transgenics Core. TSCs were cultured with TSC media and allowed to grow to ~50% confluence. TSC Media was then changed to TSC media containing 10ug/mL polybrene (Milipore, TR-1003-G) and 200 uL lentivirus containing pwCas9. After 1-1.5 days, media was changed to TSC media containing 2ug/ml puromycin. Media was then changed daily (~5-6 days) until Puromycin-resistant colonies appeared. Puromycin-resistant clones were selected and subcloned into a 96 well plate containing puromycin resistant MEF cells. TSC media containing 2ug/ml puromycin was changed daily and clones were monitored closely. Each clone was split into a 24-well well and then into 6-well wells, each containing MEFs.

Once established, *Xist*$^{+/fl}$; pcwCas9 TSCs were again cultured in a 60mm dish on MEFs. Guide RNAs targeting exon 7 of *Eed* were then introduced via a zeomycin-resistant plasmid. I determined via titration that the concentration of zeomycin that is lethal to nonresistant TSCs is not lethal to MEF cells, so antibiotic resistant MEFs were not used for this selection step. Once TSCs reached ~50% confluence, their media was supplemented with

2ug/mL zeomycin, and media was changed daily. After 5-6 days, zeomycin-resistant colonies appeared and were subcloned into 96-well wells. TSC media containing 2ug/ml zeomycin was changed daily and clones were monitored closely. Each clone was split into a 24-well well and then into 6-well wells, each containing MEFs.

Inducible expression of pcwCas9 was achieved by treating the cells generated above with 2ug/mL doxycycline (Sigma, D5207) for 3 days. Colonies were then selected and subcloned into a 96 well plate containing MEF cells. TSC media was changed daily, and clones were monitored closely. Each clone was split into a 24-well well and then into 6-well wells, each containing MEFs. PCR validation of *Eed* deletion was then performed using the primers listed in Table 5.3.

PCR

For DNA isolation, cell pellets from TSCs were lysed in buffer composed of 50mM KCl, 10mM Tris-Cl (pH 8.3), 2.5mM $MgCl_2$, 0.1mg/ml gelatin, 0.45%NP-40, and 0.45% Tween-20. Cells in lysis buffer were incubated at 50^0C overnight, and then stored at 4^0C until use. Genomic PCR reactions were carried out in ChromaTaq buffer (Denville Scientific) with 1.5mM Magnesium Chloride using RadiantTaq DNA polymerase (Alkali Scientific, #C109). PCR primer sequences are listed in Table 5.3.

RT-PCR

Total RNA was isolated from TRIzol following the manufacturer's instructions, then Poly-A+ selected using DynaBeads mRNA Direct kit (Life Technologies, #61012). SuperScript III One-Step RT-PCR Kit with Platinum Taq enzyme mixture (Life Technologies, #12574-035) was used to prepare and amplify the complementary DNA (cDNA). Primer sequences and SNP information for each amplicon are included in Table 5.4. Amplified cDNAs were run on agarose

gels and purified using the Clontech NucleoSpin Kit (Clontech, #740609). The purified cDNAs were then Sanger sequenced and sequencing traces were examined for SNPs characteristic of the *M. molossinus*-derived X chromosome and the *M. musculus*-derived X chromosome.

Allele Specific RNA-Seq Sample Preparation

Total RNA from TSCs was isolated from TRIzol (Life Technologies, #15596-018) according to the manufacturer's instructions. cDNA libraries were generated and sequenced on the Illumina HiSeq2000 platform to generate 75 bp single-end reads.

Mapping of RNA-Seq data

Quality control analysis of the RNA-Seq data was conducted using FastQC. SNP data from whole-genome sequencing of the 129/S1 (M. musculus) and JF1/Ms (M. molossinus) mouse strains were substituted into the mm9 mouse reference genome build (C57BL/6 J) using VCFtools to generate in silico 129/S1 and JF1/Ms reference genomes (Keane et al., 2011; Maclary et al., 2017; Takada et al., 2013; Yalcin et al., 2011). Sequencing reads were separately mapped to each of the two in silico genomes using STAR (Dobin et al., 2013), allowing zero mismatches in mapped reads to ensure allele-specific mapping of SNP-containing reads to only one strain-specific genome. STAR was selected for read mapping, in part due to the improved ability to handle structural variability and indels, with the goal of reducing mapping bias to the genome most like the reference genome (Dobin *et al.*, 2013). STAR is a spliced aligner capable of detecting structural variations and can handle small insertions and deletions during read mapping. STAR additionally permits soft-clipping of reads during mapping, trimming the ends of long reads that cannot be perfectly mapped. This function would permit clipping of reads that end near indels, thus preserving mapping capability at SNPs near indels.

Prior work showed that the variability due to mapping bias between the 129/S1 and JF1/Ms genomes is minimal in our RNA-Seq analysis pipeline (Maclary et al., 2017). However, small biases may affect allelic mapping at a subset of SNP sites within a gene, but this effect is mitigated since most genes contain multiple SNPs.

Allele-specific analysis of RNA-Seq data

For allelic expression analysis, only RNA-Seq reads overlapping known SNP sites that differ between the 129/S1 and JF1/Ms genomes were retained. All multi-mapping reads were excluded from the allele-specific analysis. For each SNP site, reads mapping to the 129/S1 and JF1/Ms X chromosomes were counted and the proportion of reads from each X chromosome identified. Allelic expression was calculated individually for each SNP site; for genes containing multiple SNPs, the paternal-X percentage for all SNPs was averaged to calculate gene-level allelic expression. All SNP sites with at least 10 SNP-overlapping reads were retained. Genes containing at least one SNP site with at least 10 SNP-overlapping reads were retained for further analysis and are referred to in the text as informative. In X-linked genes, the SNP frequency is ~1 SNP/250 bp in transcribed RNAs (Keane et al., 2011; Maclary et al., 2017; Takada et al., 2013; Yalcin et al., 2011).

RNA-Seq expression analysis

To calculate expression from the maternal vs. paternal X chromosomes, all reads were first merged into a single alignment file and the number of reads per RefSeq annotated gene was counted using HTSeq. To calculate the percentage of expression arising from the paternal X chromosome, the total read counts from HTSeq were normalized by number of mapped reads. Then, the normalized number of mapped reads for each gene was multiplied by the proportion of

SNP-containing reads mapping to the paternal X chromosome. This analysis was done in R using the following formula:

$$\left\{ \text{total reads} \times \left(\frac{\text{paternal reads}}{\text{maternnal reads} + \text{paternal reads}} \right) \right\}$$

RNA-FISH

Samples were dehydrated through room temperature ethanol series (five minutes each for 70%, 85%, 95%, and 100% ethanol). Coverslips were allowed to dry for 15 minutes at room temperature after the 100% ethanol wash, followed by hybridizing the samples overnight with the appropriate RNA-FISH probe. After the hybridization, samples were washed for seven minutes at 39°C, three times each in 2X SSC/50% formamide. This was followed by three seven-minute washes at 39°C in 2X SSC (1:100,000-1:200,000 dilution of DAPI added at third wash of 2X SSC), followed by two seven-minute washes at 39°C, in 1X SSC. Sample coverslips were then mounted onto glass microscope slides with Vectashield. Coverslips were sealed to the glass slides with clear nail polish.

Microscopy

Images of all stained samples were captured using a Nikon Eclipse TiE inverted microscope build with a Photometrics CCD camera. The images were analyzed after deconvolution using NIS-Elements software. All images were processed uniformly.

Statistical Analysis

All statistical analyses were performed using Welch's two sample t-tests with a significance level of $\alpha = 0.05$.

Author Contributions

Michael Hinten, PhD derived the *Xist*$^{+/fl}$ cell lines described in this chapter and Arushi Varshney, PhD performed initial RNA FISH staining in these lines. M.C. conducted all other cell derivations, optimization of Xist deletion, RNA FISH staining, and other analyses in this chapter. M.C. also generated the figures and text for this chapter, which is currently in preparation for submission as a manuscript to a peer-reviewed journal.

Table 5.1. RNA FISH quantification of *Xist* deletion via Tamoxifen induction in *Xist*$^{+/fl}$ TSC line 1. Each Tamoxifen induction time point was conducted in tandem with an EtOH induced control.

Treatment	0 *Xist* RNA Coats	1 *Xist* RNA Coat	2 *Xist* RNA Coats
12hr 9uM Tamoxifen	39	64	0
12hr EtOH	2	96	0
12hr 9uM Tamoxifen	37	65	0
12hr EtOH	7	88	0
12hr 9uM Tamoxifen	36	69	0
12hr EtOH	4	97	0
12hr 9uM Tamoxifen	23	60	0
12hr EtOH	5	89	0
24hr 9uM Tamoxifen	22	93	0
24hr EtOH	7	96	0
36hr 9uM Tamoxifen	16	89	0
36hr EtOH	2	99	0
48hr 9uM Tamoxifen	10	97	0
48hr EtOH	3	99	0

Table 5.2. RNA FISH quantification of *Xist* deletion via Tamoxifen induction in *Xist*$^{+/fl}$ TSC line 2. Each Tamoxifen induction time point was conducted in tandem with an EtOH induced control.

Treatment	0 *Xist* RNA Coats	1 *Xist* RNA Coat	2 *Xist* RNA Coats
12hr 9uM Tamoxifen	66	71	0
12hr EtOH	9	96	0
12hr 9uM Tamoxifen	52	66	0
12hr EtOH	4	98	0
12hr 9uM Tamoxifen	56	59	0
12hr EtOH	3	105	0
12hr 9uM Tamoxifen	49	63	0
12hr EtOH	6	101	0
24hr 9uM Tamoxifen	54	58	0
24hr EtOH Treated	6	96	0
48hr 9uM Tamoxifen	11	91	0
48hr EtOH Treated	5	94	0

Table 5.3. Genotyping PCR Primers

Gene	Primer Name	Primer Sequence
Xist	5LoxR_LW	ACCCTTGCCTTTTCCATTTT
	Xist3R_LW	CACTGGCAAGGTGAATAGCA
	XpromL_LW	TTTCTGGTCTTTGAGGGCAC
Eed	Eed_5’	GGACTCATCCTCTGGTAGAGCAGC
	Eed_3’	CCCAAGATCATTACCCCAGA
	Eed_R1	TCAATTGGTGGGTTTTGGAT

Table 5.4. RT-PCR Primers

Gene	Primer Name	Primer Sequence
Atrx	*Atrx* RT Forward	5’ - GGGATTGCTGCTGTGAGTCT
	Atrx RT Reverse	5’ – CCACCATCTTCTTGCCATCT

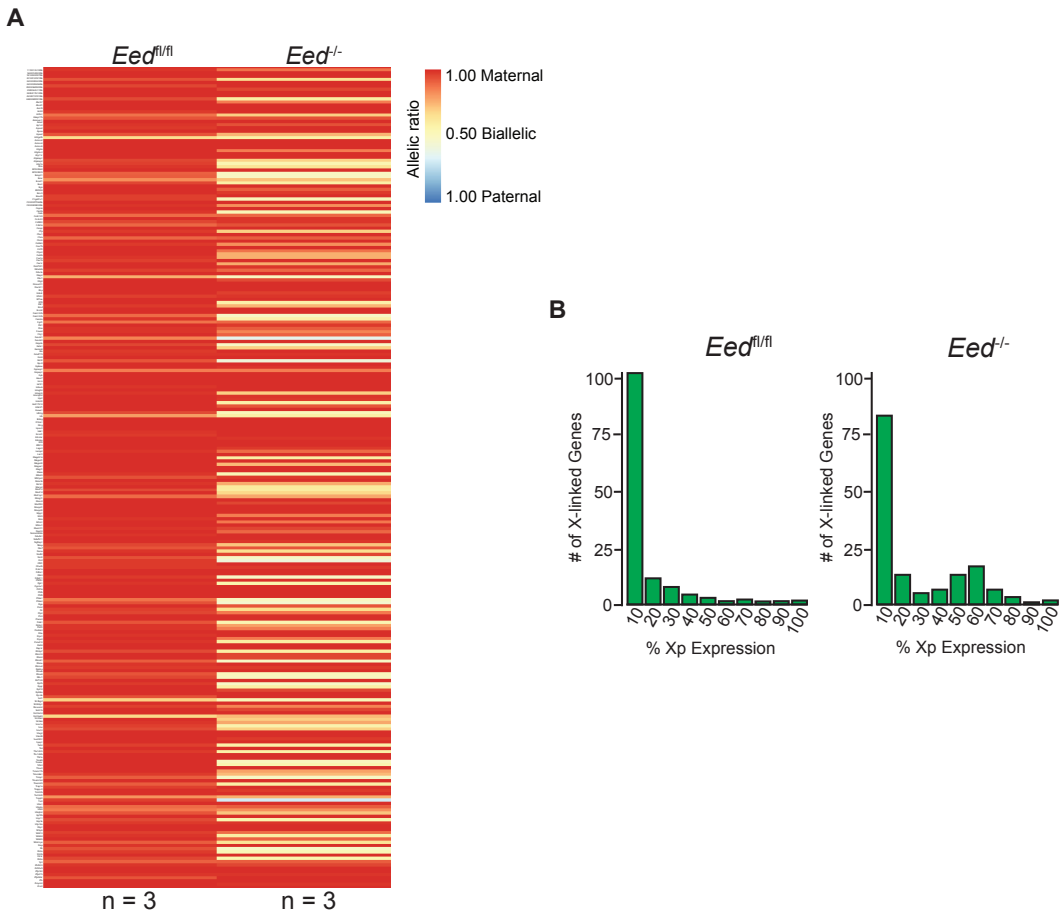


Figure 5.1. *Eed*$^{-/-}$ TSCs display minor defects in paternal X-linked gene silencing (A) Allele-specific X-linked gene expression heat map of female *Eed*$^{fl/fl}$ TSCs and female *Eed*$^{-/-}$ TSCs. Three TSC lines each of the *Eed*$^{fl/fl}$ and *Eed*$^{-/-}$ genotypes were sequenced individually and only genes with informative allelic expression in all samples are plotted (Materials and Methods). (B) Histograms depicting the average number of X-linked genes undergoing various degrees (in 10% increments) of expression from the paternal-X in *Eed*$^{fl/fl}$ and *Eed*$^{-/-}$ TSCs.

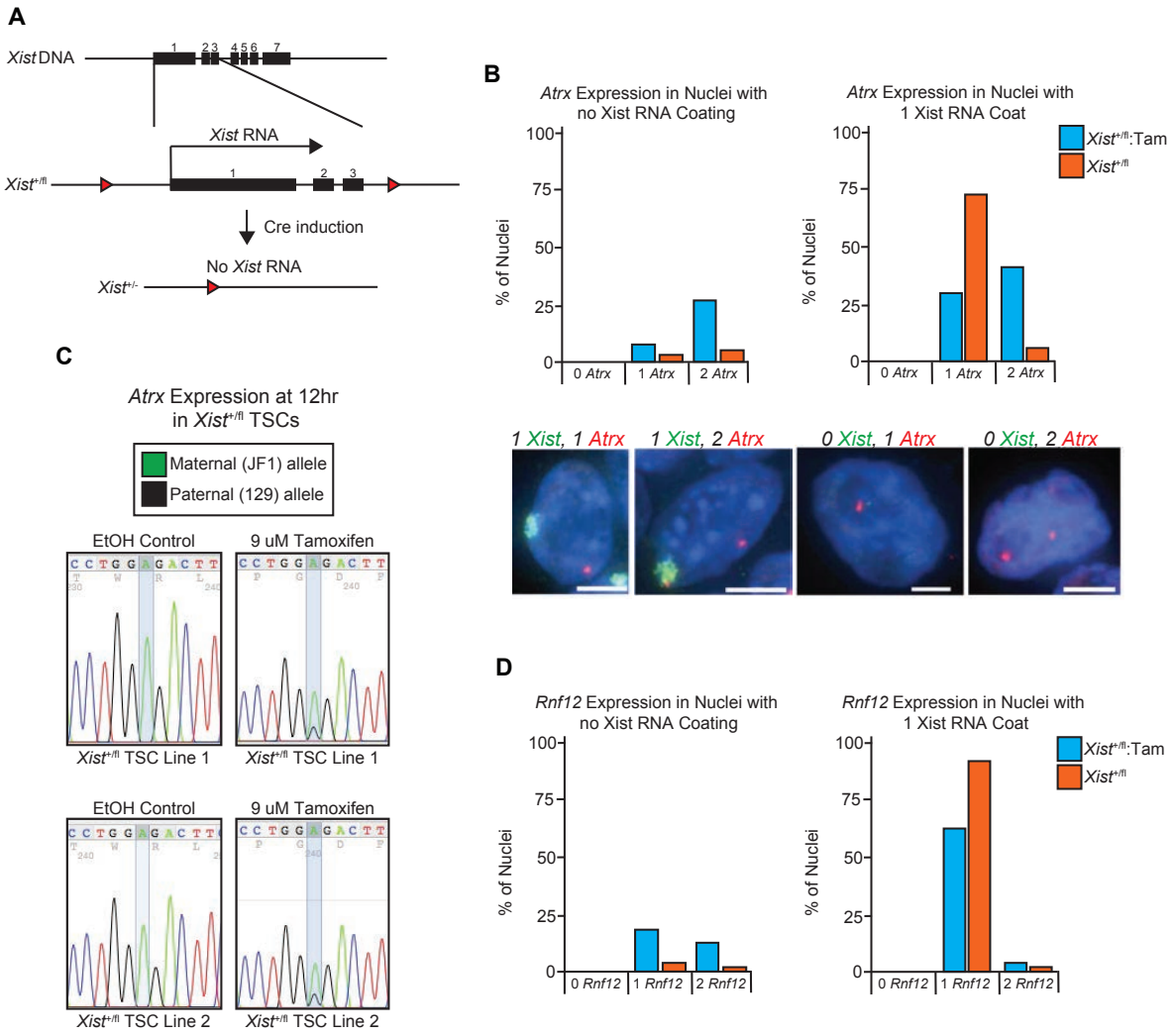


Figure 5.2. Generating and characterizing *Xist*$^{+/fl}$:Tam TSCs (A) Schematic depicting the *Xist* mutation. (B) Proportion of *Atrx* expression in nuclei with no *Xist* RNA coat or one *Xist* RNA coat in Cre-transduced and non-transduced TSCs (*Xist*$^{+/fl}$ Mock, no Cre transduction; *Xist*$^{+/fl}$:Tam Cre, transduction with Tamoxifen). Below, representative single nuclei images of each class of *Xist*$^{+/fl}$ TSCs observed. RNA-FISH for *Xist* is in green and nascent transcription detection of *Atrx* in red. Nuclei stained blue with DAPI. Scale bars are 2μm. (C) Representative snapshots of Sanger sequencing chromatograms from RT-PCR amplification of *Atrx*. (D) Proportion of *Atrx* expression in nuclei with no *Xist* RNA coat or one *Xist* RNA coat in Cre-transduced and non-transduced TSCs (*Xist*$^{+/fl}$ Mock, no Cre transduction; *Xist*$^{+/fl}$:Tam Cre, transduction with Tamoxifen).

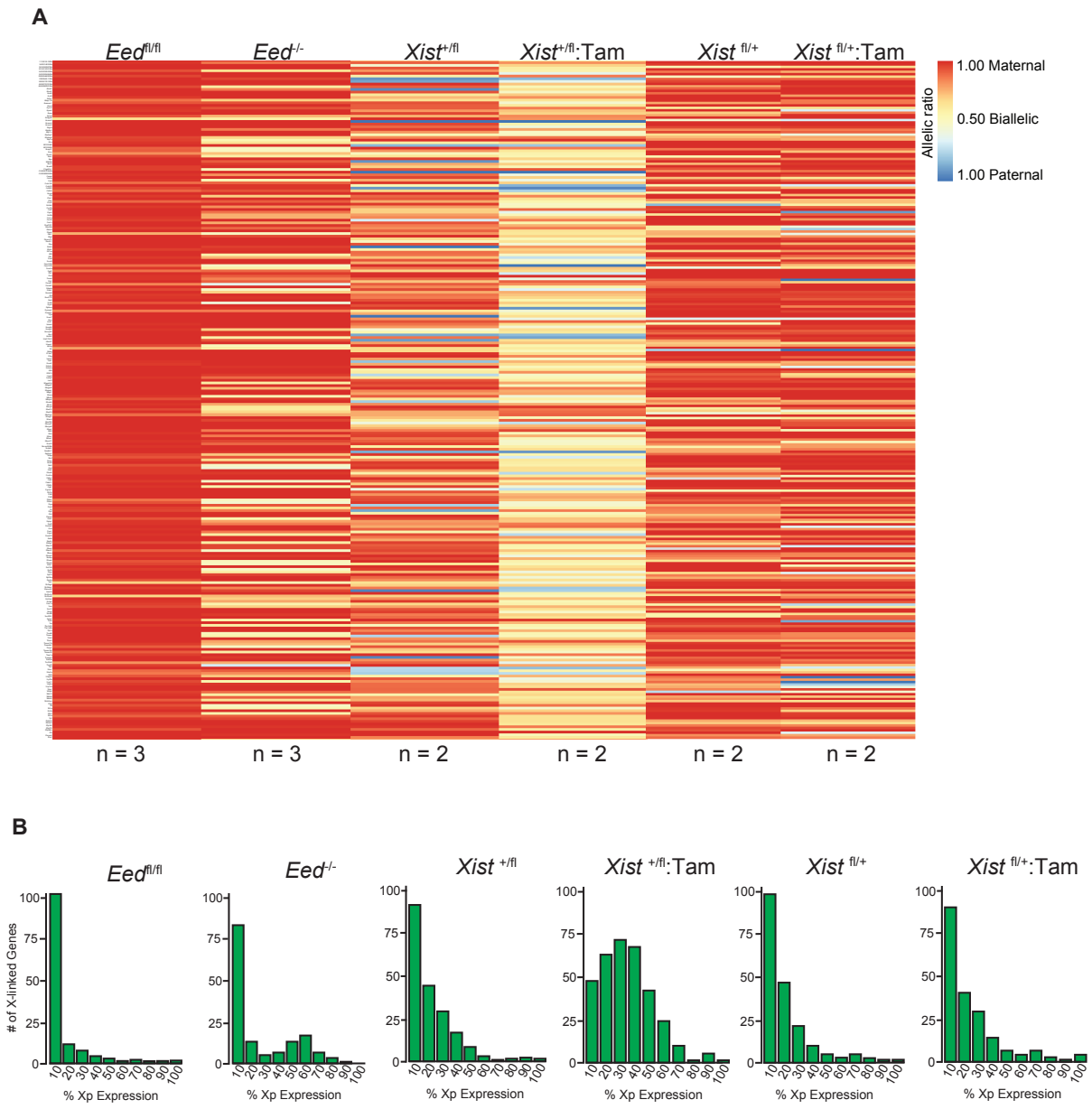


Figure 5.3. *Xist*$^{+/fl}$:Tam TSCs display significant defects in paternal X-linked gene silencing
(A) Allele-specific X-linked gene expression heat map of female *Eed*$^{fl/fl}$, *Eed*$^{-/-}$, *Xist*$^{+/fl,}$ and *Xist*$^{+/fl}$:Tam, *Xist*$^{fl/+}$, and *Xist*$^{fl/+}$:Tam TSCs. Three TSC lines each of the *Eed*$^{fl/fl}$ and *Eed*$^{-/-}$ genotypes and two TSC lines each of the *Xist*$^{+/fl,}$ and *Xist*$^{+/fl}$:Tam, *Xist*$^{fl/+}$, and *Xist*$^{fl/+}$:Tam genotypes were sequenced individually and only genes with informative allelic expression in all samples are plotted (Materials and Methods). (B) Histograms depicting the average number of X-linked genes undergoing various degrees of expression (in 10% increments) from the paternal-X in *Eed*$^{fl/fl}$, *Eed*$^{-/-}$, *Xist*$^{+/fl,}$ and *Xist*$^{+/fl}$:Tam, *Xist*$^{fl/+}$, and *Xist*$^{fl/+}$:Tam TSCs.

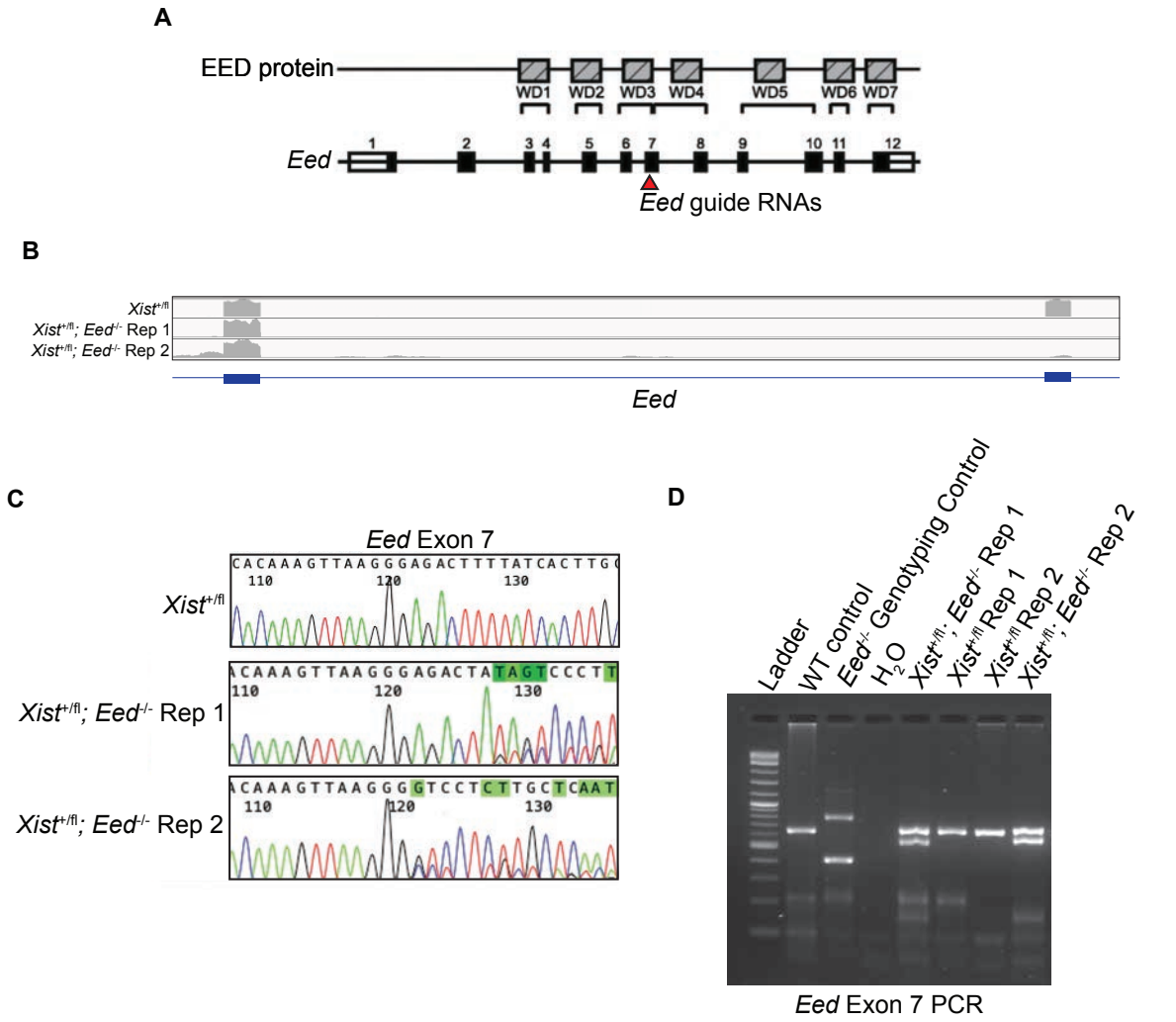

A
EED protein
WD1
WD2
WD3
WD4
WD5
WD6
WD7
Eed
1
2
3 4
5
6 7
8
9
10 11
12
Eed guide RNAs
B
Xist+/fl
Xist+/fl; Eed-/- Rep 1
Xist+/fl; Eed-/- Rep 2
Eed
C
Eed Exon 7
Xist+/fl
Xist+/fl; Eed-/- Rep 1
Xist+/fl; Eed-/- Rep 2
110
120
130
D
Ladder
WT control
Eed-/- Genotyping Control
H2O
Xist+/fl; Eed-/- Rep 1
Xist+/fl Rep 1
Xist+/fl Rep 2
Xist+/fl; Eed-/- Rep 2
Eed Exon 7 PCR
E
Merged
Xist RNA
EED
H3K27me3
Xist+/fl
Xist+/fl; Eed-/-
% Nuclei with Xist RNA Coating
% Nuclei with H3K27me3 Enrichment
100
80
60
40
20
0
Xist+/fl
Xist+/fl; Eed-/-

Figure 5.4. Generating and validating *Xist*$^{+/fl}$; *Eed*$^{-/-}$ TSCs (A) Schematic depicting the location for *Eed* disruption via CRISPR-Cas9. (B) Representative RNA-Seq Integrative Genome Viewer tracks showing *Eed* exon 7 deletion in *Xist*$^{+/fl}$; *Eed*$^{-/-}$ TSCs vs. *Xist*$^{+/fl}$ control TSCs. (C) Chromatogram snapshot of PCR amplification of the *Eed* exon 7 region disrupted via CRISPR-Cas9. (D) PCR gel depicting *Eed* exon 7 deletion in both *Xist*$^{+/fl}$; *Eed*$^{-/-}$ replicates. (E) representative IF-FISH images and quantification depicting *Xist* (green), EED (red), and H3K72me3 (white) in *Xist*$^{+/fl}$; *Eed*$^{-/-}$ TSCs vs. *Xist*$^{+/fl}$ control TSCs. Cells were stained blue with DAPI.

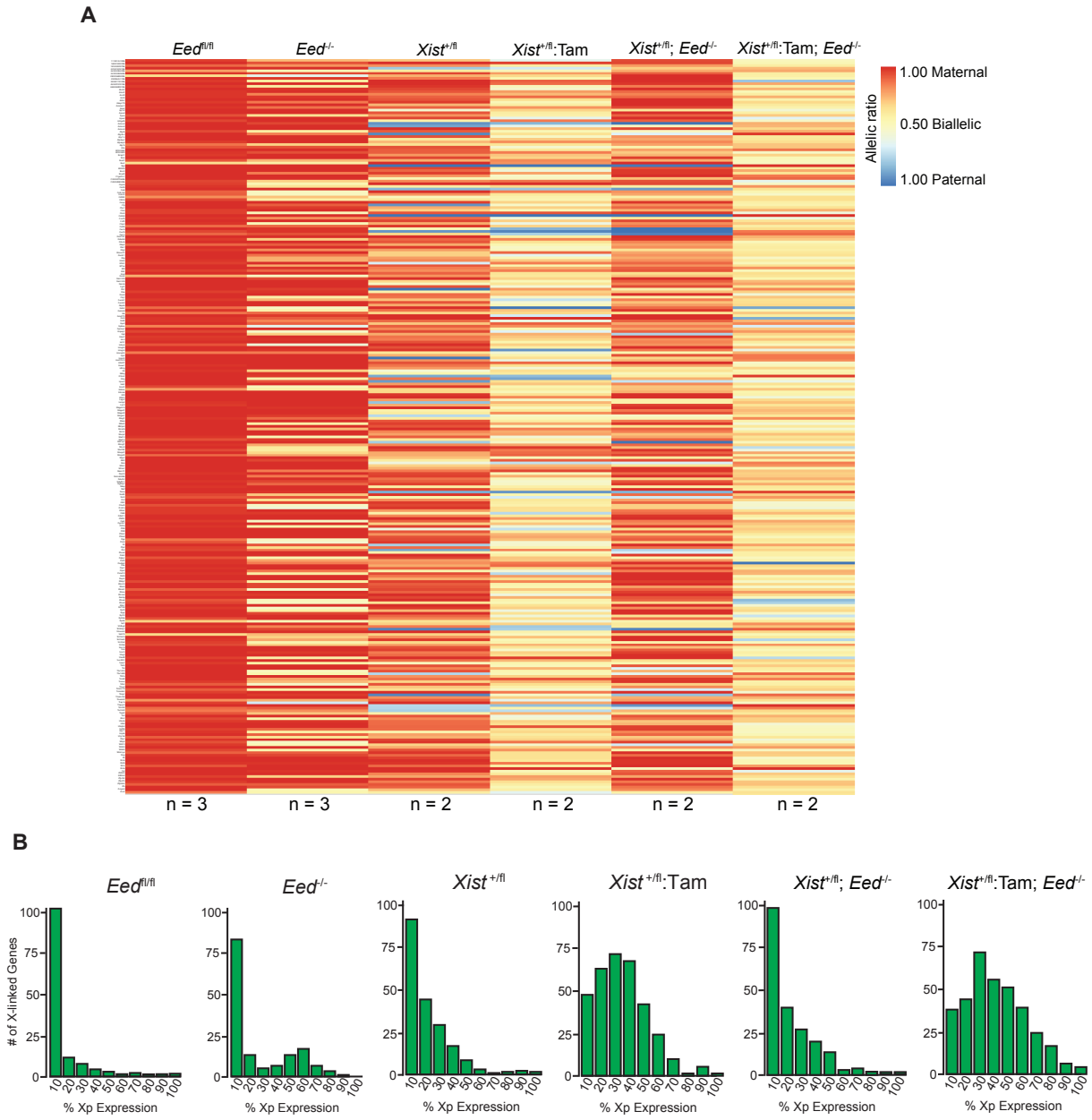


Figure 5.5. *Xist*$^{+/fl}$:Tam; *Eed*$^{-/-}$ TSCs display similar allelic expression ratios to *Xist*$^{+/fl}$:Tam TSCs (A) Allele-specific X-linked gene expression heat map of female *Eed*$^{fl/fl}$, *Eed*$^{-/-}$, *Xist*$^{+/fl}$, *Xist*$^{+/fl}$:Tam, *Xist*$^{+/fl}$; *Eed*$^{-/-}$, and *Xist*$^{+/fl}$:Tam; *Eed*$^{-/-}$ TSCs. Three TSC lines each of the *Eed*$^{fl/fl}$ and *Eed*$^{-/-}$ genotypes and two TSC lines each of the *Xist*$^{+/fl,}$ and *Xist*$^{+/fl}$:Tam, *Xist*$^{+/fl}$; *Eed*$^{-/-}$, and *Xist*$^{+/fl}$:Tam; *Eed*$^{-/-}$ genotypes were sequenced individually and only genes with informative allelic expression in all samples are plotted (Materials and Methods). (B) Histograms depicting the average number of X-linked genes undergoing various degrees of expression (in 10% increments) from the paternal-X in *Eed*$^{fl/fl}$, *Eed*$^{-/-}$, *Xist*$^{+/fl,}$ and *Xist*$^{+/fl}$:Tam, *Xist*$^{+/fl}$; *Eed*$^{-/-}$, and *Xist*$^{+/fl}$:Tam; *Eed*$^{-/-}$ TSCs.

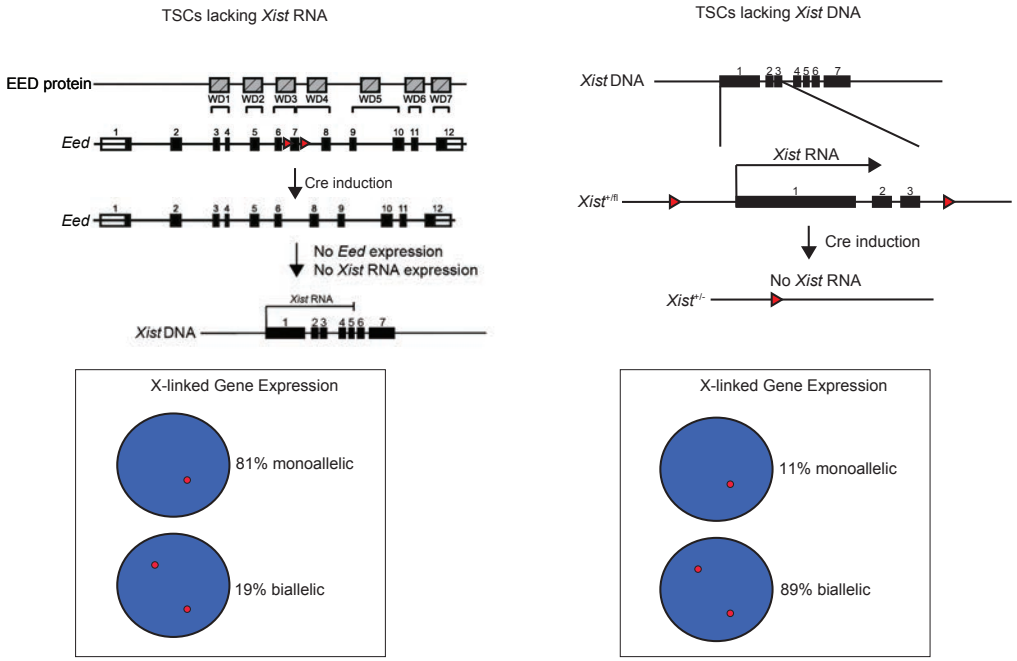


Figure 5.6. X-linked gene expression patterns in TSCs lacking *Xist* DNA versus *Xist* RNA.

Chapter 6
Concluding Remarks

Prior work has been insufficient to explain the complex nature of X-inactivation in female mice and humans. In my dissertation research, I have provided evidence for novel intra- and extracellular mechanisms of X-inactivation. This work also defines future directions for the investigation of these factors and mechanisms. Chapters 2 and 3 genetically dissected the contributions of maternally- and zygotically-generated PRC2 components to X-linked gene silencing. Chapter 3 also provided evidence for a novel function for the PRC2 core protein EED apart from its canonical function as part of PRC2 and histone H3K27me3 deposition. Chapter 4 implicates GSK-3 inhibition and Wnt signaling in *XIST* induction in hESCs and demonstrates conservation of this mechanism between humans and mice. Chapter 5 dissects the relative contributions of *Xist* RNA vs. *Xist* DNA to imprinted X-inactivation in female mouse cells. Each of these chapters provides key insights into how proper dosage compensation is achieved in mammals and each study invites the investigation of exciting new questions. In the Conclusion and Future Directions section of each chapter, I have provided a framework to further interrogate each question I have addressed. In this final chapter, I discuss more broadly some key open questions and limitations that remain in the field of X-inactivation and transcriptional regulation as well as techniques and approaches that can be employed to investigate these questions.

In Chapter 2, I helped discover that preimplantation female mouse embryos undergo random X-inactivation instead of imprinted X-inactivation in the absence of maternally-derived EED (Harris *et al.*, 2019). In a separate study, others in the Kalantry Lab found that female

mouse embryos lacking SMCX/KDM5C display normal imprinted X-inactivation but fail to undergo random X-inactivation in the epiblast lineage (Samanta *et al.*, 2022). Considering our group's findings, I hypothesize that the switch of imprinted X-inactivation to random X-inactivation in embryos lacking maternal EED requires KDM5C. Thus, the generation and characterization of mouse embryos lacking both maternal EED and KDM5C is an important and informative future direction for the Kalantry Lab. I expect that PRC2-catalyzed H3K27me3 and KDM5C-mediated demethylation of H3K4me2/3 will mechanistically connect imprinted to random X-inactivation, but how these processes relate remains to be experimentally addressed. If both imprinted and random X-inactivation are lacking, I hypothesize that some intrinsic dosage compensation will still occur, as has been suggested by others (Birchler *et al*., 2007).

Investigating the connection between imprinted and random X-inactivation through EED and KDM5C loss will shed light on another prominent question in the field: the evolutionary origin of imprinted and random X-inactivation. Imprinted X-inactivation has been proposed to be the ancestral form of X-inactivation in mammals and that this process may have arisen independently in the marsupial and eutherian lineages (Renfree et al., 2009). Random X-inactivation has been postulated to have arisen later in placental mammals due to the relaxation of *Xist* imprinting during placental radiation (Gribnau and Grootegoed, 2012). My findings in Chapter 2 provide a different view, since embryos that fail to undergo imprinted X-inactivation resort to random X-inactivation as a dosage compensation mechanism. This result agrees with findings by others (Matsui *et al.*, 2001) that imprinted X-inactivation in extraembryonic cells of androgenetic or gynogenetic embryos that harbor a complement of only either maternal or paternal chromosomes can switch to random X-inactivation. Further investigation into this question is sure to provide a more complete picture of how and when imprinted and random X-

inactivation evolved. Investigating the interplay between EED and KDM5C in imprinted and random X-inactivation will certainly provide insight into the evolutionary origins of X-inactivation. Additionally, testing if KDM5C can induce expression of *Rsx*, which encodes an RNA in metatherians with *Xist*-like expression and functional properties, will contribute to our understanding of X-inactivation evolution (Grant *et al.*, 2012; Sprague et al., 2019).

Another question that warrants further investigation is the interplay between the PRC1 and PRC2 complexes and their associated chromatin modifications. PRC1 components along with the PRC1 catalytic readout, H2AK119ub1, are co-enriched on the inactive X chromosome in early development and have been suggested to silence X-linked genes (Simon and Kingston, 2009). PRC2 and H3K27me3 are also enriched on the inactive-X and have been suggested to contribute to X-linked gene silencing (Mak *et al.*, 2002; Okamoto *et al.*, 2004). PRC1 has historically been thought to function in X-inactivation after PRC2 is recruited to the inactive-X, as PRC1 often targets and binds H3K27me3 (Fischle et al., 2003; Min et al., 2003). However, some data suggest that PRC1 may be directly recruited to the inactive-X by *Xist* RNA (Schoeftner et al., 2006) and that PRC1 may recruit PRC2 to the inactive-X (Chen *et al.*, 2021; Mei *et al.*, 2021). A possibility is that PRC1 and PRC2 can reciprocally recruit each other to the inactive-X and silence X-linked genes in varying contexts, but the interplay of these complexes with one another and with other factors warrants further investigation. In future work, a more complete and stepwise genetic analysis of PRC1 and PRC2 recruitment timing should be undertaken to better understand how and in what order these two factors may contribute to X-inactivation. This work can be carried out using a combination of genetic manipulation and chromatin profiling approaches.

In addition to the interplay between PRC1 and PRC2, variant PRC1 and PRC2 complexes and their roles in X-linked gene silencing should be further investigated. PRC1 has been shown to exist in the canonical form, which includes a CBX subunit, and non-canonical versions, with RYBP/YAF2 subunits (Gao et al., 2012; Morey et al., 2013; Tavares et al., 2012). Both canonical and non-canonical PRC1 mediate lysine 119 monoubiquitinated histone H2A (H2AK119ub1), which leads to facultative heterochromatin formation and transcriptional inactivation, although they do so in different contexts (Blackledge et al., 2020; Fursova et al., 2019). Several variants of PRC2 have also been identified, which deposit H3K27me3 in different contexts (Hauri *et al.*, 2016; Healy *et al.*, 2019). Because my dissertation work probed the core components of PRC2, which exist in both PRC2.1 and PRC2.2, I have been unable to determine whether each of these variants play different roles in X-linked gene silencing. To investigate this question, I propose generating mouse embryos lacking the PRC2.1 accessory protein PCL and, separately, embryos lacking the PRC2.2 accessory protein JARID2, as these factors are specific and essential to each of the PRC2 variant complexes (Healy *et al.*, 2019). Allele-specific analyses of X-linked gene expression in these blastocyst-stage embryos, like my analyses in Chapters 2 and 3, will shed light on which of these two complex variants may be necessary to silence maternal *Xist* and/or to silence X-linked genes in early embryos. Future analyses of similar knockouts of PRC1 variant-specific factors will also shed light on the functions of each of these complexes.

Although much of my dissertation research has investigated the effects of histone modifiers and the modifications they deposit on transcription, I have not addressed related questions like how histone turnover may influence gene regulation or the balance that exists between the reading, writing, and erasure of post-translational modifications. Recent work in

epigenetics has focused increasingly on these topics and they are worthy of examination in the context of X-inactivation. In future work, the use of techniques such as stable isotope labeling of amino acid (SILAC) pulse experiments with quantitative mass spectrometry-based proteomics holds promise to further elucidate the contributions of histone turnover to X-inactivation initiation and maintenance. For example, work by others in the proteomics space has demonstrated variable rates of turnover for acetylation vs. methylation (Zee et al., 2010). Thus, the analysis of *Eed* mutant vs. *Ezh1/2* mutant embryos or cells using techniques to gauge histone turnover may shed new light on the roles of acetylation vs. methylation in gene regulation on the X chromosome.

In accordance with the future directions outlined in Chapter 5, RNA purification and sequencing techniques should continue to be refined to capture the transcriptome more accurately. Throughout my dissertation research, I consistently encountered problems with variable read coverage and inaccurate sequencing of repetitive and chromatin-associated transcripts, particularly *Xist*. Although I was able to assess *Xist* expression using other techniques, more accurate quantification of *Xist* and other problematic transcripts by RNA-Seq would greatly benefit future X-inactivation studies. I suspect that this could be achieved by utilizing long-read RNA-Seq techniques, as these approaches cn eliminate the mis-mapping of many repetitive reads. Furthermore, as mentioned in Chapter 5, many transcripts likely exist that are not detectable using current sequencing methodologies and purification techniques due to their repetitive nature, low expression, and/or chromatin association. Thus, advances in sequencing and purification technologies are likely to unlock parts of the transcriptome that have yet to be explored.

Some key analyses outlined in this body of work leverage allele-specific molecular techniques to gain insight into parent-of-origin gene regulation. These techniques are powerful tools to study imprinting, but future analyses will benefit from the use of single cell RNA-Seq (Tang et al., 2010) and single molecule imaging (Liu et al., 2015) to assess transcriptional dynamics more precisely. The allele-specific RNA-Seq, Pyrosequencing, and Sanger sequencing approaches I have employed in my dissertation research to analyze hybrid cells and embryos provide a glimpse into how genes are allelically expressed, but these techniques only possess the resolution to study cells in aggregate. Future use of single cell RNA-Seq to investigate the effects of epigenetic factors in the context of X-inactivation will provide a clearer view of transcriptional dynamics at the single-cell level (Stuart and Satija, 2019). Likewise, the allele-specific RNA FISH technique that I used in Chapters 2 and 3 to interrogate maternal vs. paternal *Xist* expression in various Polycomb mutants is a powerful tool to study X-linked gene expression at an allele-specific single-cell resolution. Future use of this approach to examine the expression dynamics of additional genes will provide a more precise view of transcriptional dynamics in individual cells. In addition, the adoption of live-cell single molecule imaging techniques will allow our group to study the activity of individual chromatin-associated factors in living cells (Lionnet and Wu, 2021). Single molecule imaging techniques have already been used by others to study the function of various epigenetic factors, including Polycomb proteins (Brown et al., 2021), but much work remains in profiling the activity of chromatin regulators in varying contexts. Future work employing these high-resolution approaches in combination with genetic manipulation will dissect more precisely how individual cells, chromosomes, and molecules function under various conditions.

The factors and mechanisms I have examined in my thesis work likely affect cellular processes beyond X-inactivation and regions of the genome beyond the X chromosome. For example, it is unlikely that X-inactivation is the only cellular process controlled by maternally generated PRC2. Thus, future use of the data I generated to screen for other cellular processes affected by the loss of maternal EED and/or EZH1/2 (Chapters 2 and 3) is likely to yield new knowledge. Additionally, the inclusion of GSK-3 inhibitors in culture media (Chapter 4) likely has effects beyond *Xist* regulation in humans and mice. Even the investigation of *Xist* RNA vs. *Xist* DNA (Chapter 5) likely has implications beyond those in X-linked gene silencing. For example, *Xist* has recently been shown to exert oncogenic effects by acting as a competing endogenous RNA (ceRNA) which sequesters micro-RNAs (miRNAs) to upregulate target genes (Thomson and Dinger, 2016). In agreement, *Xist* has been shown to attract a plethora of miRNAs and to be upregulated in various tumors (Madhi and Kim, 2019). *Xist* has also been shown to be downregulated in some types of cancer (Richart et al., 2022; Yildirim et al., 2013). To better understand the oncogenic and other regulatory roles for *Xist*, further investigation of the *Xist* gene and its various transcripts will be necessary. Thus, examination of the datasets I have generated to address new questions is likely to reveal more about *Xist*-related and other as-yet-unknown biological processes.

My work in basic science has been motivated by a desire to learn about how the world works, in addition to the hope that my findings will someday contribute to preserving human health. I decided to pursue a PhD in Genetics and Genomics after completing a degree in Epidemiology, which allowed me to investigate associations between biological exposures and outcomes, but not the mechanistic underpinnings of biological states. By delving into the field of X-inactivation in my PhD work, I have gained insights into mechanisms that contribute to

wide-scale gene silencing. I have also learned that these mechanisms are far more complex and context-dependent than I could have ever imagined. As epigenetics research progresses, it is vital to bear in mind that many variables – both cell intrinsic and extrinsic – may influence transcriptional states. Factors such as environmental/culture conditions, cell type, and age/developmental stage likely contribute more to biology than we currently realize. It is through conducting highly controlled stage- and cell type-specific investigations such as those outlined in this work that we will continue to piece together a better understanding of transcriptional regulation. And it is through the advances made by these investigations that the X-inactivation field has and will continue to contribute in significant ways to our understanding of human biology.

Appendices

Appendix A

Transposable Element Expression in *Eed*-mutant Mouse Embryos

Transposable elements (TEs) were first discovered by Barbara McClintock in 1950 when her studies in maize indicated that eukaryotes contain repetitive genomic elements that can move from one chromosomal location to another (McClintock, 1950). Over time, other studies have shown that virtually all organisms harbor TEs, some of which have been amplified and co-opted over evolutionary time to perform important cellular functions (Agrawal et al., 1998; Feschotte, 2008; Levin and Moran, 2011; Levis et al., 1993). Early research into TE regulation suggested that these loci are exclusively marked and silenced by DNA and/or histone 3 lysine 9 (H3K9) methylation (Karimi et al., 2011; Klenov et al., 2011). However, more recent work has suggested that the Polycomb-mediated H3K27me3 mark is also enriched at and may silence some TEs (Deleris et al., 2021).

In this analysis, I used RetroSeq (Keane et al., 2013), a TE discovery tool for next-generation sequencing data, to analyze transposable element expression in various *Eed*-mutant and control female blastocyst-stage embryos generated for the experiments in Chapter 2. In agreement with other work (Deleris *et al.*, 2021), I discovered increased expression of TEs in embryos lacking *Eed*. Interestingly, I observed expression of more TEs in $Eed^{mz-/-}$ embryos than in any other genotype (Table A.1), suggesting that maternal and zygotic *Eed* may function to silence different TEs. Future work will address how maternal and zygotic EED is targeted to specific classes of TEs as well as how PRC2-controlled TE expression contributes to X-inactivation and other early embryonic processes.

Table A.1. Quantification of transposable elements identified in female mouse blastocysts

	Eed$^{fl/fl}$	*Eed*$^{fl/-}$	*Eed*$^{-/-}$	*Eed*$^{m-/-}$	*Eed*$^{mz-/-}$
SINEs	87	93	105	148	167
ALUs	84	86	89	128	132
MIRs	3	7	16	20	35
LINEs	96	95	125	135	128
LINE1	84	82	97	123	108
LINE2	12	13	28	12	20
LTR elements	39	41	45	43	42
MaLRs	23	23	21	23	22
Retrov.	14	15	18	13	17
MER4 group	2	3	6	7	3
DNA elements	7	8	6	10	10
MER1 type	5	6	6	8	7
MER2 type	2	2	0	2	3
Mariners	0	0	0	0	0
Unclassified	9	10	22	15	25
Total	238	247	303	351	372

Appendix B

Investigating Imprinted X-inactivation in *Xist*$^{+/-}$ Female Mouse Embryos

A small project that I completed during my early days in the lab involved analyzing allele-specific RNA-Seq data generated from three hybrid wild-type female mouse blastocysts and three hybrid *Xist*$^{+/-}$ female mouse blastocysts harboring a null *Xist* allele on the paternal X chromosome. The purpose of this analysis was to determine whether paternal X-linked gene silencing in *Xist*$^{+/-}$ female blastocysts recapitulated the pattern observed in *Xist*$^{+/-}$ female trophoblast stem cells (TSCs) analyzed in Chapter 5. In the female *Xist*$^{+/-}$ TSCs analyzed in Chapter 5, I observed loss of *Xist* in only some cells, as the tamoxifen-inducible Cre-lox method I used to delete *Xist* did not remove *Xist* in all cells. However, in the TSCs lacking *Xist*, I observed increased biallelic expression of X-linked genes by RNA FISH, suggesting a failure of X-inactivation. These data were supported by the allele-specific RNA-Seq analyses I conducted in *Xist*$^{+/-}$ TSCs, which showed a significant shift in the maternal:paternal allelic expression ratio compared to that observed in wild-type embryos.

The blastocyst stage in mice represents an intermediate stage where X-linked gene silencing due to imprinted X-inactivation is being established (Kalantry et al., 2009; Namekawa *et al.*, 2010; Patrat *et al.*, 2009). Previous work by Dr. Kalantry suggested that the paternal alleles of a subset of X-linked genes could become silenced despite the absence of *Xist* on the paternal-X (Kalantry et al., 2009). These genes, nevertheless, appeared to require *Xist* to remain stably silenced in later stage embryos (Kalantry et al., 2009). This 2009 study, however, only examined a handful of genes by low throughput approaches including RNA FISH and

Pyrosequencing. A later study analyzed female *Xist*$^{+/-}$ mouse embryos by allele-specific single-cell RNA-Seq and concluded that the silencing of paternal X-linked genes absolutely requires *Xist* (Deng et al., 2014), but reanalysis of this data by others contradicted this claim (Borensztein *et al.*, 2017). Given that single-cell RNA-Seq data are subject to technical biases (Chen *et al.*, 2019; Hicks *et al.*, 2018), which are compounded by allele-specific RNA expression quantitation, testing X-linked gene expression by allele-specific RNA-Seq of whole embryos would appear necessary for a rigorous comparative analysis of the two studies. Thus, a robust analysis of whole *Xist*$^{+/-}$ female blastocysts by allele-specific RNA-Seq allowed me to observe *in vivo* the effect of paternal *Xist* loss at a key developmental stage.

Using the allele-specific RNA-Seq pipeline outlined in the Materials and Methods of Chapters 2, 3, and 5, I observed that loss of *Xist* on the paternal-X in female blastocysts led to nearly equal expression of maternal and paternal X-linked genes (Figure B.1). In contrast, wild-type female blastocysts displayed biased expression of X-linked genes from the maternal allele, consistent with the onset of imprinted X-inactivation. These data indicate that paternal *Xist* is necessary for the establishment of imprinted X-inactivation in female mouse blastocysts. My data also show a nearly 2-fold increase in X-linked gene expression in *Xist*$^{+/-}$ mouse embryos vs. WT mice. This overall increase in X-linked gene expression suggests a failure of dosage compensation in female blastocysts upon paternal *Xist* loss. In future work, allele-specific examination of X-linked gene expression in individual 2-, 4-, 8-, and 16-cell embryos will test the earlier requirement of paternal *Xist* on the initiation of X-linked gene silencing across the X chromosome.

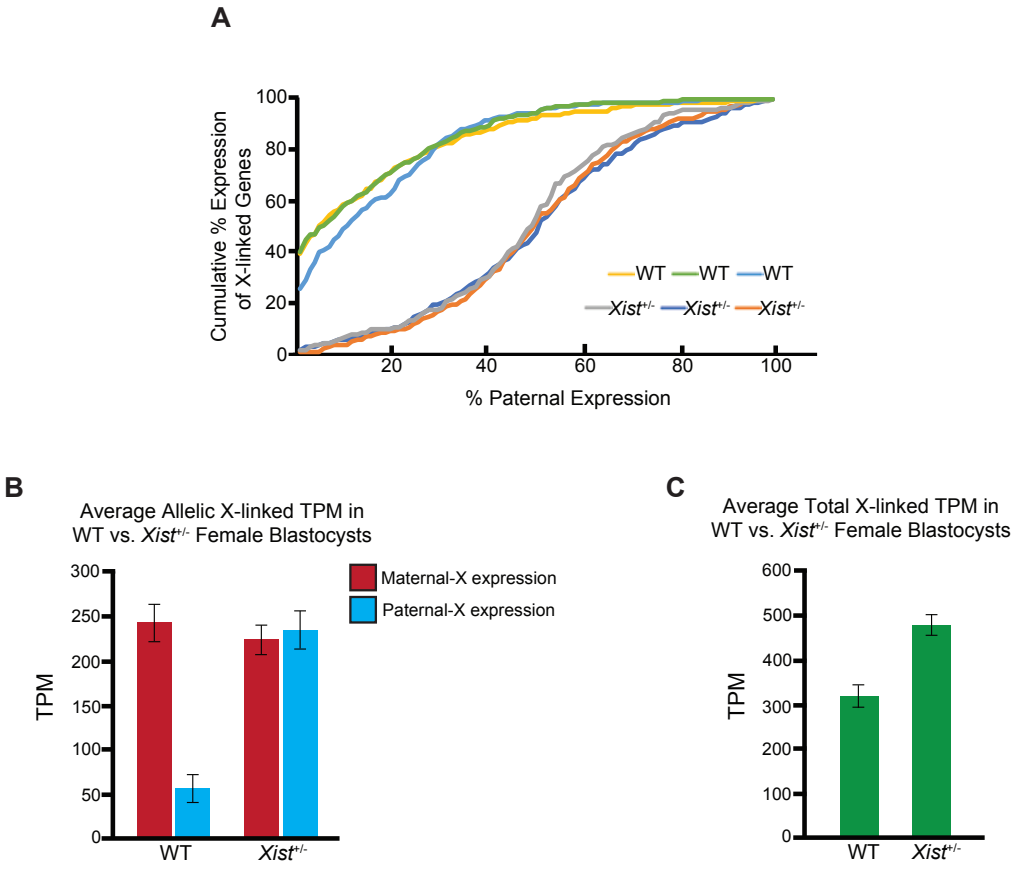


Figure B.1. Allelic X-linked gene expression in three WT and three *Xist*$^{+/-}$ female mouse embryos (A) Cumulative frequency plot of percent paternal X-linked gene expression in WT and *Xist*$^{+/-}$ female mouse embryos. (B) Average transcripts per kilobase million (TPM) from the maternal vs. paternal X chromosome in WT and *Xist*$^{+/-}$ female mouse embryos. Error bars represent standard deviation. (C) Average Total X-linked TPM from WT and *Xist*$^{+/-}$ female mouse embryos. Error bars represent standard deviation.

References

Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S., Bevan, S., et al. (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. Nat Biotechnol *25*, 803-816. 10.1038/nbt1318.

Agrawal, A., Eastman, Q.M., and Schatz, D.G. (1998). Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature *394*, 744-751. 10.1038/29457.

Ai, S., Peng, Y., Li, C., Gu, F., Yu, X., Yue, Y., Ma, Q., Chen, J., Lin, Z., Zhou, P., et al. (2017). EED orchestration of heart maturation through interaction with HDACs is H3K27me3-independent. Elife *6*. 10.7554/eLife.24570.

An, C., Feng, G., Zhang, J., Cao, S., Wang, Y., Wang, N., Lu, F., Zhou, Q., and Wang, H. (2020). Overcoming Autocrine FGF Signaling-Induced Heterogeneity in Naive Human ESCs Enables Modeling of Random X Chromosome Inactivation. Cell Stem Cell. 10.1016/j.stem.2020.06.002.

Anguera, M.C., Sadreyev, R., Zhang, Z., Szanto, A., Payer, B., Sheridan, S.D., Kwok, S., Haggarty, S.J., Sur, M., Alvarez, J., et al. (2012). Molecular signatures of human induced pluripotent stem cells highlight sex differences and cancer genes. Cell Stem Cell *11*, 75-90. 10.1016/j.stem.2012.03.008.

Bar, S., and Benvenisty, N. (2019). Epigenetic aberrations in human pluripotent stem cells. EMBO J *38*. 10.15252/embj.2018101033.

Bar, S., Seaton, L.R., Weissbein, U., Eldar-Geva, T., and Benvenisty, N. (2019). Global Characterization of X Chromosome Inactivation in Human Pluripotent Stem Cells. Cell Rep *27*, 20-29 e23. 10.1016/j.celrep.2019.03.019.

Barakat, T.S., Ghazvini, M., de Hoon, B., Li, T., Eussen, B., Douben, H., van der Linden, R., van der Stap, N., Boter, M., Laven, J.S., et al. (2015). Stable X chromosome reactivation in female human induced pluripotent stem cells. Stem Cell Reports *4*, 199-208. 10.1016/j.stemcr.2014.12.012.

Barlow, D.P. (2011). Genomic imprinting: a mammalian epigenetic discovery model. Annu Rev Genet *45*, 379-403. 10.1146/annurev-genet-110410-132459.

Barr, M.L., and Bertram, E.G. (1949). A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. Nature *163*, 676. 10.1038/163676a0.

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.

Bassett AR, Akhtar A, Barlow DP, Bird AP, Brockdorff N, Duboule D, Ephrussi A, Ferguson-Smith AC, Gingeras TR, Haerty W, Higgs DR, Miska EA, Ponting CP. Considerations when investigating lncRNA function in vivo. Elife. 2014 Aug 14;3:e03058. doi: 10.7554/eLife.03058. PMID: 25124674; PMCID: PMC4132285.

Behrens, J., Jerchow, B.A., Wurtele, M., Grimm, J., Asbrand, C., Wirtz, R., Kuhl, M., Wedlich, D., and Birchmeier, W. (1998). Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. Science *280*, 596-599. 10.1126/science.280.5363.596.

Beurel, E., Grieco, S.F., and Jope, R.S. (2015). Glycogen synthase kinase-3 (GSK3): regulation, actions, and diseases. Pharmacol Ther *148*, 114-131. 10.1016/j.pharmthera.2014.11.016.

Birchler, J.A., Pal-Bhadra, M., and Bhadra, U. (2003). Dosage dependent gene regulation and the compensation of the X chromosome in Drosophila males. Genetica *117*, 179-190. 10.1023/a:1022935927763.

Birchler, J.A., Veitia, R.A (2007). The Gene Balance Hypothesis: From Classical Genetics to Modern Genomics, The Plant Cell, Volume 19, Issue 2, February 2007, Pages 395–402, https://doi.org/10.1105/tpc.106.049338

Blackledge, N.P., Fursova, N.A., Kelley, J.R., Huseyin, M.K., Feldmann, A., and Klose, R.J. (2020). PRC1 Catalytic Activity Is Central to Polycomb System Function. Mol Cell *77*, 857-874 e859. 10.1016/j.molcel.2019.12.001.

Blauwkamp, T.A., Chang, M.V., and Cadigan, K.M. (2008). Novel TCF-binding sites specify transcriptional repression by Wnt signalling. EMBO J *27*, 1436-1446. 10.1038/emboj.2008.80.

Borensztein, M., Syx, L., Ancelin, K., Diabangouaya, P., Picard, C., Liu, T., Liang, J.B., Vassilev, I., Galupa, R., Servant, N., et al. (2017). Xist-dependent imprinted X inactivation and the early developmental consequences of its failure. Nat Struct Mol Biol *24*, 226-233. 10.1038/nsmb.3365.

Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., and Rastan, S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell *71*, 515-526. 10.1016/0092-8674(92)90519-i.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., and Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448*, 191-195. 10.1038/nature05950.

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991a). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature *349*, 38-44. 10.1038/349038a0.

Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J., and Willard, H.F. (1992). The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell *71*, 527-542. 10.1016/0092-8674(92)90520-m.

Brown, C.J., Lafreniere, R.G., Powers, V.E., Sebastio, G., Ballabio, A., Pettigrew, A.L., Ledbetter, D.H., Levy, E., Craig, I.W., and Willard, H.F. (1991b). Localization of the X inactivation centre on the human X chromosome in Xq13. Nature *349*, 82-84. 10.1038/349082a0.

Brown, C.J., and Willard, H.F. (1994). The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. Nature *368*, 154-156. 10.1038/368154a0.

Brown, K., Andrianakos, H., Ingersoll, S., and Ren, X. (2021). Single-molecule imaging of epigenetic complexes in living cells: insights from studies on Polycomb group proteins. Nucleic Acids Res *49*, 6621-6637. 10.1093/nar/gkab304.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell *14*, 838-853. 10.1016/j.stem.2014.04.003.

Cadigan, K.M., and Nusse, R. (1997). Wnt signaling: a common theme in animal development. Genes Dev *11*, 3286-3305. 10.1101/gad.11.24.3286.

Cao, Q., Wang, X., Zhao, M., Yang, R., Malik, R., Qiao, Y., Poliakov, A., Yocum, A.K., Li, Y., Chen, W., et al. (2014). The central role of EED in the orchestration of polycomb group complexes. Nat Commun *5*, 3127. 10.1038/ncomms4127.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043. 10.1126/science.1076997.

Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol Cell *15*, 57-67. 10.1016/j.molcel.2004.06.020.

Carter, A.C., Xu, J., Nakamoto, M.Y., Wei, Y., Zarnegar, B.J., Shi, Q., Broughton, J.P., Ransom, R.C., Salhotra, A., Nagaraja, S.D., et al. (2020). Spen links RNA-mediated endogenous retrovirus silencing and X chromosome inactivation. Elife *9*. 10.7554/eLife.54508.

Chang, S.C., and Brown, C.J. (2010). Identification of regulatory elements flanking human XIST reveals species differences. BMC Mol Biol *11*, 20. 10.1186/1471-2199-11-20.

Chen, G., Ning, B., and Shi, T. (2019). Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. Front Genet *10*, 317. 10.3389/fgene.2019.00317.
Chen, G., Schell, J.P., Benitez, J.A., Petropoulos, S., Yilmaz, M., Reinius, B., Alekseenko, Z., Shi, L., Hedlund, E., Lanner, F., et al. (2016). Single-cell analyses of X Chromosome inactivation dynamics and pluripotency during differentiation. Genome Res *26*, 1342-1354. 10.1101/gr.201954.115.

Chen, Z., Djekidel, M.N., and Zhang, Y. (2021). Distinct dynamics and functions of H2AK119ub1 and H3K27me3 in mouse preimplantation embryos. Nat Genet *53*, 551-563. 10.1038/s41588-021-00821-2.

Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., and Chang, H.Y. (2015). Systematic discovery of Xist RNA binding proteins. Cell *161*, 404-416. 10.1016/j.cell.2015.03.025.

Chureau, C., Prissette, M., Bourdet, A., Barbe, V., Cattolico, L., Jones, L., Eggen, A., Avner, P., and Duret, L. (2002). Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. Genome Res *12*, 894-908. 10.1101/gr.152902.

Clemson, C.M., McNeil, J.A., Willard, H.F., and Lawrence, J.B. (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. J Cell Biol *132*, 259-275. 10.1083/jcb.132.3.259.

Clerc, P., and Avner, P. (1998). Role of the region 3' to Xist exon 6 in the counting process of X-chromosome inactivation. Nat Genet *19*, 249-253. 10.1038/924.

Cloutier, M., Harris, C., Gayen, S., Maclary, E., and Kalantry, S. (2018). Experimental Analysis of Imprinted Mouse X-Chromosome Inactivation. Methods Mol Biol *1861*, 177-203. 10.1007/978-1-4939-8766-5_14.

Coker, H., Wei, G., Moindrot, B., Mohammed, S., Nesterova, T., and Brockdorff, N. (2020). The role of the Xist 5' m6A region and RBM15 in X chromosome inactivation. Wellcome Open Res *5*, 31. 10.12688/wellcomeopenres.15711.1.

Cooper, D.W. (1971). Directed genetic change model for X chromosome inactivation in eutherian mammals. Nature *230*, 292-294. 10.1038/230292a0.

Csankovszki, G., Panning, B., Bates, B., Pehrson, J.R., and Jaenisch, R. (1999). Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. Nat Genet *22*, 323-324. 10.1038/11887.

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.

Dahl, J.A., Jung, I., Aanes, H., Greggains, G.D., Manaf, A., Lerdrup, M., Li, G., Kuan, S., Li, B., Lee, A.Y., et al. (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. Nature *537*, 548-552. 10.1038/nature19360.

de Jaime-Soguero, A., Abreu de Oliveira, W.A., and Lluis, F. (2018). The Pleiotropic Effects of the Canonical Wnt Pathway in Early Development and Pluripotency. Genes (Basel) *9*. 10.3390/genes9020093.

de la Cruz, C.C., Fang, J., Plath, K., Worringer, K.A., Nusinow, D.A., Zhang, Y., and Panning, B. (2005a). Developmental regulation of Suz 12 localization. Chromosoma *114*, 183-192.

de la Cruz, X., Lois, S., Sanchez-Molina, S., and Martinez-Balbas, M.A. (2005b). Do protein motifs read the histone code? Bioessays *27*, 164-175. 10.1002/bies.20176.

de Oliveira Georges, J.A., Vergani, N., Fonseca, S.A., Fraga, A.M., de Mello, J.C., Albuquerque, M.C., Fujihara, L.S., and Pereira, L.V. (2014). Aberrant patterns of X chromosome inactivation in a new line of human embryonic stem cells established in physiological oxygen concentrations. Stem Cell Rev Rep *10*, 472-479. 10.1007/s12015-014-9505-4.

Deakin, J.E., Hore, T.A., Koina, E., and Marshall Graves, J.A. (2008). The status of dosage compensation in the multiple X chromosomes of the platypus. PLoS Genet *4*, e1000140. 10.1371/journal.pgen.1000140.

Del Rosario, B.C., Del Rosario, A.M., Anselmo, A., Wang, P.I., Sadreyev, R.I., and Lee, J.T. (2017). Genetic Intersection of Tsix and Hedgehog Signaling during the Initiation of X-Chromosome Inactivation. Dev Cell *43*, 359-371 e356. 10.1016/j.devcel.2017.09.027.

Deleris, A., Berger, F., and Duharcourt, S. (2021). Role of Polycomb in the control of transposable elements. Trends Genet *37*, 882-889. 10.1016/j.tig.2021.06.003.

Deng, Q., Ramskold, D., Reinius, B., and Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science *343*, 193-196. 10.1126/science.1245316.

Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. Nat Struct Mol Biol *20*, 1147-1155. 10.1038/nsmb.2669.

Diaz Perez, S.V., Kim, R., Li, Z., Marquez, V.E., Patel, S., Plath, K., and Clark, A.T. (2012). Derivation of new human embryonic stem cell lines reveals rapid epigenetic progression in vitro that can be prevented by chemical modification of chromatin. Hum Mol Genet *21*, 751-764. 10.1093/hmg/ddr506.

Disteche, C.M. (2016). Dosage compensation of the sex chromosomes and autosomes. Semin Cell Dev Biol *56*, 9-18. 10.1016/j.semcdb.2016.04.013.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21. 10.1093/bioinformatics/bts635.

Doble, B.W., Patel, S., Wood, G.A., Kockeritz, L.K., and Woodgett, J.R. (2007). Functional redundancy of GSK-3alpha and GSK-3beta in Wnt/beta-catenin signaling shown by using an allelic series of embryonic stem cell lines. Dev Cell *12*, 957-971. 10.1016/j.devcel.2007.04.001.

Doble, B.W., and Woodgett, J.R. (2003). GSK-3: tricks of the trade for a multi-tasking kinase. J Cell Sci *116*, 1175-1186. 10.1242/jcs.00384.

Dossin, F., Pinheiro, I., Zylicz, J.J., Roensch, J., Collombet, S., Le Saux, A., Chelmicki, T., Attia, M., Kapoor, V., Zhan, Y., et al. (2020). SPEN integrates transcriptional and epigenetic control of X-inactivation. Nature *578*, 455-460. 10.1038/s41586-020-1974-9.

Dvash, T., Lavon, N., and Fan, G. (2010). Variations of X chromosome inactivation occur in early passages of female human embryonic stem cells. PLoS One *5*, e11330. 10.1371/journal.pone.0011330.

Eldar-Finkelman, H., and Martinez, A. (2011). GSK-3 Inhibitors: Preclinical and Clinical Focus on CNS. Front Mol Neurosci *4*, 32. 10.3389/fnmol.2011.00032.

Ellegren, H., Hultin-Rosenberg, L., Brunstrom, B., Dencker, L., Kultima, K., and Scholz, B. (2007). Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. BMC Biol *5*, 40. 10.1186/1741-7007-5-40.

Embi, N., Rylatt, D.B., and Cohen, P. (1980). Glycogen synthase kinase-3 from rabbit skeletal muscle. Separation from cyclic-AMP-dependent protein kinase and phosphorylase kinase. Eur J Biochem *107*, 519-527.

Erhardt, S., Su, I.H., Schneider, R., Barton, S., Bannister, A.J., Perez-Burgos, L., Jenuwein, T., Kouzarides, T., Tarakhovsky, A., and Surani, M.A. (2003). Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. Development *130*, 4235-4248. 10.1242/dev.00625.

Escamilla-Del-Arenal, M., da Rocha, S.T., and Heard, E. (2011). Evolutionary diversity and developmental regulation of X-chromosome inactivation. Hum Genet *130*, 307-327. 10.1007/s00439-011-1029-2.

Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. Development *121*, 273-285.

Ferguson-Smith, A.C., and Bourc'his, D. (2018). The discovery and importance of genomic imprinting. eLife *7*. 10.7554/eLife.42368.

Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. Nat Rev Genet *9*, 397-405. 10.1038/nrg2337.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev *17*, 1870-1881. 10.1101/gad.1110503.

Freland, L., and Beaulieu, J.M. (2012). Inhibition of GSK3 by lithium, from single molecules to signaling networks. Front Mol Neurosci *5*, 14. 10.3389/fnmol.2012.00014.

Fukuda, A., Hazelbaker, D.Z., Motosugi, N., Hao, J., Limone, F., Beccard, A., Mazzucato, P., Messana, A., Okada, C., San Juan, I.G., et al. (2021). De novo DNA methyltransferases DNMT3A and DNMT3B are essential for XIST silencing for erosion of dosage compensation in pluripotent stem cells. Stem Cell Reports *16*, 2138-2148. 10.1016/j.stemcr.2021.07.015.

Fursova, N.A., Blackledge, N.P., Nakayama, M., Ito, S., Koseki, Y., Farcas, A.M., King, H.W., Koseki, H., and Klose, R.J. (2019). Synergy between Variant PRC1 Complexes Defines Polycomb-Mediated Gene Repression. Mol Cell *74*, 1020-1036 e1028. 10.1016/j.molcel.2019.03.024.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature *504*, 282-286. 10.1038/nature12745.

Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., and Reinberg, D. (2012). PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. Mol Cell *45*, 344-356. 10.1016/j.molcel.2012.01.002.

Gardner, D.K., Lane, M., Stevens, J., Schlenker, T., and Schoolcraft, W.B. (2000). Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. Fertil Steril *73*, 1155-1158. 10.1016/s0015-0282(00)00518-5.

Gardner, R.L., and Lyon, M.F. (1971). X chromosome inactivation studied by injection of a single cell into the mouse blastocyst. Nature *231*, 385-386. 10.1038/231385a0.

Gayen, S., Maclary, E., Buttigieg, E., Hinten, M., and Kalantry, S. (2015). A Primary Role for the Tsix lncRNA in Maintaining Random X-Chromosome Inactivation. Cell Rep *11*, 1251-1265. 10.1016/j.celrep.2015.04.039.

Geens, M., Seriola, A., Barbe, L., Santalo, J., Veiga, A., Dee, K., Van Haute, L., Sermon, K., and Spits, C. (2016). Female human pluripotent stem cells rapidly lose X chromosome inactivation marks and progress to a skewed methylation pattern during culture. Mol Hum Reprod *22*, 285-298. 10.1093/molehr/gaw004.

Grant, J., Mahadevaiah, S.K., Khil, P., Sangrithi, M.N., Royo, H., Duckworth, J., McCarrey, J.R., VandeBerg, J.L., Renfree, M.B., Taylor, W., et al. (2012). Rsx is a metatherian RNA with Xist-like properties in X-chromosome inactivation. Nature *487*, 254-258. 10.1038/nature11171.

Gribnau, J., and Grootegoed, J.A. (2012). Origin and evolution of X chromosome inactivation. Curr Opin Cell Biol *24*, 397-404. 10.1016/j.ceb.2012.02.004.

Grimes, C.A., and Jope, R.S. (2001). The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling. Prog Neurobiol *65*, 391-426. 10.1016/s0301-0082(01)00011-9.

Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W., et al. (2017). Epigenetic resetting of human pluripotency. Development *144*, 2748-2763. 10.1242/dev.146811.

Hadjantonakis, A.K., Gertsenstein, M., Ikawa, M., Okabe, M., and Nagy, A. (1998). Non-invasive sexing of preimplantation stage mammalian embryos. Nat Genet *19*, 220-222.

Hall, L.L., Byron, M., Butler, J., Becker, K.A., Nelson, A., Amit, M., Itskovitz-Eldor, J., Stein, J., Stein, G., Ware, C., and Lawrence, J.B. (2008). X-inactivation reveals epigenetic anomalies in most hESC but identifies sublines that initiate as expected. J Cell Physiol *216*, 445-452. 10.1002/jcp.21411.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc Natl Acad Sci U S A *107*, 9222-9227. 10.1073/pnas.1004584107.

Harper, M.I., Fosten, M., and Monk, M. (1982). Preferential paternal X inactivation in extraembryonic tissues of early mouse embryos. J Embryol Exp Morphol *67*, 127-135.

Harris, C., Cloutier, M., Trotter, M., Hinten, M., Gayen, S., Du, Z., Xie, W., and Kalantry, S. (2019). Conversion of random X-inactivation to imprinted X-inactivation by maternal PRC2. Elife *8*. 10.7554/eLife.44258.

Hauri, S., Comoglio, F., Seimiya, M., Gerstung, M., Glatter, T., Hansen, K., Aebersold, R., Paro, R., Gstaiger, M., and Beisel, C. (2016). A High-Density Map for Navigating the Human Polycomb Complexome. Cell Rep *17*, 583-595. 10.1016/j.celrep.2016.08.096.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell *146*, 519-532. 10.1016/j.cell.2011.06.052.

Healy, E., Mucha, M., Glancy, E., Fitzpatrick, D.J., Conway, E., Neikes, H.K., Monger, C., Van Mierlo, G., Baltissen, M.P., Koseki, Y., et al. (2019). PRC2.1 and PRC2.2 Synergize to Coordinate H3K27 Trimethylation. Mol Cell *76*, 437-452 e436. 10.1016/j.molcel.2019.08.012.

Hebbes, T.R., Thorne, A.W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J *7*, 1395-1402.

Hey, C.A.B., Saltokova, K.B., Bisgaard, H.C., and Moller, L.B. (2018). Comparison of two different culture conditions for derivation of early hiPSC. Cell Biol Int *42*, 1467-1473. 10.1002/cbin.10966.

Hicks, S.C., Townes, F.W., Teng, M., and Irizarry, R.A. (2018). Missing data and technical variability in single-cell RNA-sequencing experiments. Biostatistics *19*, 562-578. 10.1093/biostatistics/kxx053.

Hinten, M., Maclary, E., Gayen, S., Harris, C., and Kalantry, S. (2016). Visualizing Long Noncoding RNAs on Chromatin. Methods Mol Biol *1402*, 147-164. 10.1007/978-1-4939-3378-5_12.

Hoffman, L.M., Hall, L., Batten, J.L., Young, H., Pardasani, D., Baetge, E.E., Lawrence, J., and Carpenter, M.K. (2005). X-inactivation status varies in human embryonic stem cell lines. Stem Cells *23*, 1468-1478. 10.1634/stemcells.2004-0371.

Hojfeldt, J.W., Laugesen, A., Willumsen, B.M., Damhofer, H., Hedehus, L., Tvardovskiy, A., Mohammad, F., Jensen, O.N., and Helin, K. (2018). Accurate H3K27 methylation can be established de novo by SUZ12-directed PRC2. Nat Struct Mol Biol *25*, 225-232. 10.1038/s41594-018-0036-6.

Hughes, C.S., Postovit, L.M., and Lajoie, G.A. (2010). Matrigel: a complex protein mixture required for optimal growth of cell culture. Proteomics *10*, 1886-1890. 10.1002/pmic.200900758.

Hughes, J.F., and Page, D.C. (2015). The Biology and Evolution of Mammalian Y Chromosomes. Annu Rev Genet *49*, 507-527. 10.1146/annurev-genet-112414-055311.

Huynh, K.D., and Lee, J.T. (2003). Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. Nature *426*, 857-862.

Inoue, A., Chen, Z., Yin, Q., and Zhang, Y. (2018). Maternal Eed knockout causes loss of H3K27me3 imprinting and random X inactivation in the extraembryonic cells. Genes Dev *32*, 1525-1536. 10.1101/gad.318675.118.

Inoue, A., Jiang, L., Lu, F., and Zhang, Y. (2017). Genomic imprinting of Xist by maternal H3K27me3. Genes Dev *31*, 1927-1932. 10.1101/gad.304113.117.

International Stem Cell Initiative, C., Akopian, V., Andrews, P.W., Beil, S., Benvenisty, N., Brehm, J., Christie, M., Ford, A., Fox, V., Gokhale, P.J., et al. (2010). Comparison of defined culture systems for feeder cell free propagation of human embryonic stem cells. In Vitro Cell Dev Biol Anim *46*, 247-258. 10.1007/s11626-010-9297-z.

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.

Jachowicz, J.W., Strehle, M., Banerjee, A.K., Blanco, M.R., Thai, J., and Guttman, M. (2022). Xist spatially amplifies SHARP/SPEN recruitment to balance chromosome-wide silencing and specificity to the X chromosome. Nat Struct Mol Biol *29*, 239-249. 10.1038/s41594-022-00739-1.

Jakobsson, E., Arguello-Miranda, O., Chiu, S.W., Fazal, Z., Kruczek, J., Nunez-Corrales, S., Pandit, S., and Pritchet, L. (2017). Towards a Unified Understanding of Lithium Action in Basic Biology and its Significance for Applied Biology. J Membr Biol *250*, 587-604. 10.1007/s00232-017-9998-2.

Kalantry, S. (2011). Recent advances in X-chromosome inactivation. J Cell Physiol *226*, 1714-1718. 10.1002/jcp.22673.

Kalantry, S., and Magnuson, T. (2006). The Polycomb group protein EED is dispensable for the initiation of random X-chromosome inactivation. PLoS Genet *2*, e66. 10.1371/journal.pgen.0020066.

Kalantry, S., Mills, K.C., Yee, D., Otte, A.P., Panning, B., and Magnuson, T. (2006). The Polycomb group protein Eed protects the inactive X-chromosome from differentiation-induced reactivation. Nat Cell Biol *8*, 195-202. 10.1038/ncb1351.

Kalantry, S., Purushothaman, S., Bowen, R.B., Starmer, J., and Magnuson, T. (2009). Evidence of Xist RNA-independent initiation of mouse imprinted X-chromosome inactivation. Nature *460*, 647-651. nature08161 [pii]
10.1038/nature08161.

Karimi, M.M., Goyal, P., Maksakova, I.A., Bilenky, M., Leung, D., Tang, J.X., Shinkai, Y., Mager, D.L., Jones, S., Hirst, M., and Lorincz, M.C. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. Cell Stem Cell *8*, 676-687. 10.1016/j.stem.2011.04.004.

Kay, G.F., Barton, S.C., Surani, M.A., and Rastan, S. (1994). Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. Cell *77*, 639-650. 10.1016/0092-8674(94)90049-3.

Kay, G.F., Penny, G.D., Patel, D., Ashworth, A., Brockdorff, N., and Rastan, S. (1993). Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. Cell *72*, 171-182.

Keane, T.M., Goodstadt, L., Danecek, P., White, M.A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M., et al. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. Nature *477*, 289-294. 10.1038/nature10413.

Keane, T.M., Wong, K., and Adams, D.J. (2013). RetroSeq: transposable element discovery from next-generation sequencing data. Bioinformatics *29*, 389-390. 10.1093/bioinformatics/bts697.

Keohane, A.M., Lavender, J.S., O'Neill, L.P., and Turner, B.M. (1998). Histone acetylation and X inactivation. Dev Genet *22*, 65-73. 10.1002/(SICI)1520-6408(1998)22:1<65::AID-DVG7>3.0.CO;2-5.

Kim, K.Y., Hysolli, E., Tanaka, Y., Wang, B., Jung, Y.W., Pan, X., Weissman, S.M., and Park, I.H. (2014). X Chromosome of female cells shows dynamic changes in status during human somatic cell reprogramming. Stem Cell Reports *2*, 896-909. 10.1016/j.stemcr.2014.04.003.

Klein, P.S., and Melton, D.A. (1996). A molecular mechanism for the effect of lithium on development. Proc Natl Acad Sci U S A *93*, 8455-8459. 10.1073/pnas.93.16.8455.

Klenov, M.S., Sokolova, O.A., Yakushev, E.Y., Stolyarenko, A.D., Mikhaleva, E.A., Lavrov, S.A., and Gvozdev, V.A. (2011). Separation of stem cell maintenance and transposon silencing functions of Piwi protein. Proc Natl Acad Sci U S A *108*, 18760-18765. 10.1073/pnas.1106676108.

Kobayashi, H., Sakurai, T., Imai, M., Takahashi, N., Fukuda, A., Yayoi, O., Sato, S., Nakabayashi, K., Hata, K., Sotomaru, Y., et al. (2012). Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. PLoS Genet *8*, e1002440. 10.1371/journal.pgen.1002440.

Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., and Wutz, A. (2004). A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. PLoS Biol *2*, E171. 10.1371/journal.pbio.0020171.

Kunath, T., Arnaud, D., Uy, G.D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R.L., Avner, P., and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. Development *132*, 1649-1661. 10.1242/dev.01715.

Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T.S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P., and Reinberg, D. (2005). Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. Proc Natl Acad Sci U S A *102*, 1859-1864. 10.1073/pnas.0409875102.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev *16*, 2893-2905. 10.1101/gad.1035902.

Lagger, G. et al. (2002) Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. The EMBO journal 21, 2672–2681, https://doi.org/10.1093/emboj/21.11.2672.

Lee, J.T. (2000). Disruption of imprinted X inactivation by parent-of-origin effects at Tsix. Cell *103*, 17-27. 10.1016/s0092-8674(00)00101-x.

Lee, J.T., and Bartolomei, M.S. (2013). X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell *152*, 1308-1323. 10.1016/j.cell.2013.02.016.

Lee, J.T., Davidow, L.S., and Warshawsky, D. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. Nat Genet *21*, 400-404. 10.1038/7734.

Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., Alagappan, R., Frampton, G.M., Xu, P., Muffat, J., et al. (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell *141*, 872-883. 10.1016/j.cell.2010.04.010.

Leost, M., Schultz, C., Link, A., Wu, Y.Z., Biernat, J., Mandelkow, E.M., Bibb, J.A., Snyder, G.L., Greengard, P., Zaharevitz, D.W., et al. (2000). Paullones are potent inhibitors of glycogen synthase kinase-3beta and cyclin-dependent kinase 5/p25. Eur J Biochem *267*, 5983-5994. 10.1046/j.1432-1327.2000.01673.x.

Levesque, M.J., Ginart, P., Wei, Y., and Raj, A. (2013). Visualizing SNVs to quantify allele-specific expression in single cells. Nat Methods *10*, 865-867. 10.1038/nmeth.2589.

Levin, H.L., and Moran, J.V. (2011). Dynamic interactions between transposable elements and their hosts. Nat Rev Genet *12*, 615-627. 10.1038/nrg3030.

Levis, R.W., Ganesan, R., Houtchens, K., Tolar, L.A., and Sheen, F.M. (1993). Transposons in place of telomeric repeats at a Drosophila telomere. Cell *75*, 1083-1093. 10.1016/0092-8674(93)90318-k.

Lewandoski, M., Wassarman, K.M., and Martin, G.R. (1997). Zp3-cre, a transgenic mouse line for the activation or inactivation of loxP-flanked target genes specifically in the female germ line. Curr Biol *7*, 148-151. 10.1016/s0960-9822(06)00059-5.

Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565-570. 10.1038/276565a0.

Lewis, P.M. (1947). New Mutants Report. Drosophila. . Inf. Serv. *21*.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930. 10.1093/bioinformatics/btt656.

Lionnet, T., and Wu, C. (2021). Single-molecule tracking of transcription protein dynamics in living cells: seeing is believing, but what are we seeing? Curr Opin Genet Dev *67*, 94-102. 10.1016/j.gde.2020.12.001.

Liu, Z., Lavis, L.D., and Betzig, E. (2015). Imaging live-cell dynamics and structure at the single-molecule level. Mol Cell *58*, 644-659. 10.1016/j.molcel.2015.02.033.

Ludwig, T.E., Levenstein, M.E., Jones, J.M., Berggren, W.T., Mitchen, E.R., Frane, J.L., Crandall, L.J., Daigh, C.A., Conard, K.R., Piekarczyk, M.S., et al. (2006). Derivation of human embryonic stem cells in defined conditions. Nat Biotechnol *24*, 185-187. 10.1038/nbt1177.

Lyon, M.F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature *190*, 372-373. 10.1038/190372a0.

Lyon, M.F., Searle, A.G., Ford, C.E., and Ohno, S. (1964). A Mouse Translocation Suppressing Sex-Linked Variegation. Cytogenetics *3*, 306-323. 10.1159/000129820.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57-63. 10.1038/nature11244.

Maclary, E., Buttigieg, E., Hinten, M., Gayen, S., Harris, C., Sarkar, M.K., Purushothaman, S., and Kalantry, S. (2014). Differentiation-dependent requirement of Tsix long non-coding RNA in imprinted X-chromosome inactivation. Nat Commun *5*, 4209. 10.1038/ncomms5209.

Maclary, E., Hinten, M., Harris, C., Sethuraman, S., Gayen, S., and Kalantry, S. (2017). PRC2 represses transcribed genes on the imprinted inactive X chromosome in mice. Genome Biol *18*, 82. 10.1186/s13059-017-1211-5.

Madhi, H., and Kim, M. (2019). Beyond X-Chromosome Inactivation: The Oncogenic Facet of XIST in Human Cancers. Biomed Sci Letters *25*, 113-122.

Mak, W., Baxter, J., Silva, J., Newall, A.E., Otte, A.P., and Brockdorff, N. (2002). Mitotically stable association of polycomb group proteins eed and enx1 with the inactive x chromosome in trophoblast stem cells. Curr Biol *12*, 1016-1020. 10.1016/s0960-9822(02)00892-8.

Mak, W., Nesterova, T.B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. (2004). Reactivation of the paternal X chromosome in early mouse embryos. Science *303*, 666-669.

Mandal, S., Chandel, D., Kaur, H., Majumdar, S., Arava, M., and Gayen, S. (2020). Single-Cell Analysis Reveals Partial Reactivation of X Chromosome instead of Chromosome-wide

Dampening in Naive Human Pluripotent Stem Cells. Stem Cell Reports *14*, 745-754. 10.1016/j.stemcr.2020.03.027.

Marahrens, Y., Panning, B., Dausman, J., Strauss, W., and Jaenisch, R. (1997). Xist-deficient mice are defective in dosage compensation but not spermatogenesis. Genes Dev *11*, 156-166. 10.1101/gad.11.2.156.

Margueron, R., Justin, N., Ohno, K., Sharpe, M.L., Son, J., Drury, W.J., 3rd, Voigt, P., Martin, S.R., Taylor, W.R., De Marco, V., et al. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. Nature *461*, 762-767. 10.1038/nature08398.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D., and Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol Cell *32*, 503-518. 10.1016/j.molcel.2008.11.004.

Margueron, R., and Reinberg, D. (2010). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343-349. nature09784 [pii]
10.1038/nature09784.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343-349. 10.1038/nature09784.

Matsui, J., Goto, Y., and Takagi, N. (2001). Control of Xist expression for imprinted and random X chromosome inactivation in mice. Hum Mol Genet *10*, 1393-1401.

McClintock, B. (1950). The origin and behavior of mutable loci in maize. Proc Natl Acad Sci U S A *36*, 344-355. 10.1073/pnas.36.6.344.

McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature *521*, 232-236. 10.1038/nature14443.

Mei, H., Kozuka, C., Hayashi, R., Kumon, M., Koseki, H., and Inoue, A. (2021). H2AK119ub1 guides maternal inheritance and zygotic deposition of H3K27me3 in mouse embryos. Nat Genet *53*, 539-550. 10.1038/s41588-021-00820-3.

Mekhoubad, S., Bock, C., de Boer, A.S., Kiskinis, E., Meissner, A., and Eggan, K. (2012). Erosion of dosage compensation impacts human iPSC disease modeling. Cell Stem Cell *10*, 595-609. 10.1016/j.stem.2012.02.014.

Messmer, T., von Meyenn, F., Savino, A., Santos, F., Mohammed, H., Lun, A.T.L., Marioni, J.C., and Reik, W. (2019). Transcriptional Heterogeneity in Naive and Primed Human Pluripotent Stem Cells at Single-Cell Resolution. Cell Rep *26*, 815-824 e814. 10.1016/j.celrep.2018.12.099.

Migeon, B.R., Lee, C.H., Chowdhury, A.K., and Carpenter, H. (2002). Species differences in TSIX/Tsix reveal the roles of these genes in X-chromosome inactivation. Am J Hum Genet *71*, 286-293. 10.1086/341605.

Min, J., Zhang, Y., and Xu, R.M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev *17*, 1823-1828. 10.1101/gad.269603.

Minajigi, A., Froberg, J.E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., Lessing, D., Payer, B., Boukhali, M., Haas, W., and Lee, J.T. (2015). Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. Science *349*. 10.1126/science.aab2276.

Moindrot, B., Cerase, A., Coker, H., Masui, O., Grijzenhout, A., Pintacuda, G., Schermelleh, L., Nesterova, T.B., and Brockdorff, N. (2015). A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. Cell Rep *12*, 562-572. 10.1016/j.celrep.2015.06.053.

Monfort, A., Di Minin, G., Postlmayr, A., Freimann, R., Arieti, F., Thore, S., and Wutz, A. (2015). Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells. Cell Rep *12*, 554-561. 10.1016/j.celrep.2015.06.067.

Monk, M., and Kathuria, H. (1977). Dosage compensation for an X-linked gene in pre-implantation mouse embryos. Nature *270*, 599-601.

Montgomery, N.D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S.J., Otte, A.P., and Magnuson, T. (2005). The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. Curr Biol *15*, 942-947. 10.1016/j.cub.2005.04.051.

Montgomery, R. L. et al. (2007) Histone deacetylases 1 and 2 redundantly regulate cardiac morphogenesis, growth, and contractility. Genes & development 21, 1790–1802, https://doi.org/10.1101/gad.1563807.

Moreira de Mello, J.C., Fernandes, G.R., Vibranovski, M.D., and Pereira, L.V. (2017). Early X chromosome inactivation during human preimplantation development revealed by single-cell RNA-sequencing. Sci Rep *7*, 10794. 10.1038/s41598-017-11044-z.

Morey, C., and Avner, P. (2011). The demoiselle of X-inactivation: 50 years old and as trendy and mesmerising as ever. PLoS Genet *7*, e1002212. 10.1371/journal.pgen.1002212.
Morey, C.a.A., P (2011). The Demoiselle of X-Inactivation: 50 Years Old and As Trendy and Mesmerising As Ever. PLoS Genetics *7*.

Morey, L., Aloia, L., Cozzuto, L., Benitah, S.A., and Di Croce, L. (2013). RYBP and Cbx7 define specific biological functions of polycomb complexes in mouse embryonic stem cells. Cell Rep *3*, 60-69. 10.1016/j.celrep.2012.11.026.

Muller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell *111*, 197-208. 10.1016/s0092-8674(02)00976-5.

Namekawa, S.H., Payer, B., Huynh, K.D., Jaenisch, R., and Lee, J.T. (2010). Two-step imprinted X inactivation: repeat versus genic silencing in the mouse. Mol Cell Biol *30*, 3187-3205. 10.1128/MCB.00227-10.

Nazor, K.L., Altun, G., Lynch, C., Tran, H., Harness, J.V., Slavin, I., Garitaonandia, I., Muller, F.J., Wang, Y.C., Boscolo, F.S., et al. (2012). Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. Cell Stem Cell *10*, 620-634. 10.1016/j.stem.2012.02.013.

Nengqing, L., Dian, L., Yingjun, X., Yi, C., Lina, H., Diyu, C., Yinghong, Y., Bing, S., and Xiaofang, S. (2020). Generation of induced pluripotent stem cell GZHMCi001-A and GZHMCi001-B derived from peripheral blood mononuclear cells of epileptic patients with KCNC1 mutation. Stem Cell Res *47*, 101897. 10.1016/j.scr.2020.101897.

Nesterova, T.B., Wei, G., Coker, H., Pintacuda, G., Bowness, J.S., Zhang, T., Almeida, M., Bloechl, B., Moindrot, B., Carter, E.J., et al. (2019). Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. Nat Commun *10*, 3129. 10.1038/s41467-019-11171-3.

Nguyen, D., and Disteche, C.M. (2003). Dosage compensation of the active X chromosome in mammals. Nature Genetics *38*, 47-53.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487-492. 10.1016/j.stem.2009.05.015.

O'Gorman, S., Dagenais, N.A., Qian, M., and Marchuk, Y. (1997). Protamine-Cre recombinase transgenes efficiently recombine target sequences in the male germ line of mice, but not in embryonic stem cells. Proc Natl Acad Sci U S A *94*, 14602-14607.

O'Leary, T., Heindryckx, B., Lierman, S., van Bruggen, D., Goeman, J.J., Vandewoestyne, M., Deforce, D., de Sousa Lopes, S.M., and De Sutter, P. (2012). Tracking the progression of the human inner cell mass during embryonic stem cell derivation. Nat Biotechnol *30*, 278-282. 10.1038/nbt.2135.

Ohhata, T., Hoki, Y., Sasaki, H., and Sado, T. (2008). Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. Development *135*, 227-235. 10.1242/dev.008490.

Ohno, S. (1967). Sex Chromosomes and Sex Linked Genes. Berlin: Springer Verlag.

Ohno, S., Kaplan, W.D., and Kinosita, R. (1959). Formation of the sex chromatin by a single X-chromosome in liver cells of Rattus norvegicus. Exp Cell Res *18*, 415-418. 10.1016/0014-4827(59)90031-x.

Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. Science *303*, 644-649. 10.1126/science.1092727.

Okamoto, I., Patrat, C., Thepot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J.P., Renard, J.P., Duranthon, V., and Heard, E. Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature *472*, 370-374. nature09872 [pii] 10.1038/nature09872.

Okamoto, I., Patrat, C., Thepot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J.P., Renard, J.P., Duranthon, V., and Heard, E. (2011). Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature *472*, 370-374. 10.1038/nature09872.

Okamoto, I., Tan, S., and Takagi, N. (2000). X-chromosome inactivation in XX androgenetic mouse embryos surviving implantation. Development *127*, 4137-4145.

Panning, B., and Jaenisch, R. (1996). DNA hypomethylation can activate Xist expression and silence X-linked genes. Genes Dev *10*, 1991-2002. 10.1101/gad.10.16.1991.

Patel, S., Bonora, G., Sahakyan, A., Kim, R., Chronis, C., Langerman, J., Fitz-Gibbon, S., Rubbi, L., Skelton, R.J.P., Ardehali, R., et al. (2017). Human Embryonic Stem Cells Do Not Change Their X Inactivation Status during Differentiation. Cell Rep *18*, 54-67. 10.1016/j.celrep.2016.11.054.

Patel, S., Doble, B., and Woodgett, J.R. (2004). Glycogen synthase kinase-3 in insulin and Wnt signalling: a double-edged sword? Biochem Soc Trans *32*, 803-808. 10.1042/BST0320803.

Patrat, C., Okamoto, I., Diabangouaya, P., Vialon, V., Le Baccon, P., Chow, J., and Heard, E. (2009). Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. Proc Natl Acad Sci U S A *106*, 5198-5203. 10.1073/pnas.0810683106.

Patrat, C., Ouimette, J.F., and Rougeulle, C. (2020). X chromosome inactivation in human development. Development *147*. 10.1242/dev.183095.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature *379*, 131-137. 10.1038/379131a0.

Petropoulos, S., Edsgard, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., Reyes, A.P., Linnarsson, S., Sandberg, R., and Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. Cell *167*, 285. 10.1016/j.cell.2016.08.009.

Phung, J., Wang, C.A., Reeders, J., Chan, E.C., Riveros, C., Zakar, T., Paul, J.W., Pennell, C.E., and Smith, R. (2022). Preterm labor is a distinct process from term labor following computational analysis of human myometrium. Am J Obstet Gynecol *226*, 106 e101-106 e116. 10.1016/j.ajog.2021.07.002.

Pintacuda, G., and Cerase, A. (2015). X Inactivation Lessons from Differentiating Mouse Embryonic Stem Cells. Stem Cell Rev Rep *11*, 699-705. 10.1007/s12015-015-9597-5.

Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, C.C., Otte, A.P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. Science *300*, 131-135. 10.1126/science.1084274.

Plath, K., Mlynarczyk-Evans, S., Nusinow, D.A., and Panning, B. (2002). Xist RNA and the mechanism of X chromosome inactivation. Annu Rev Genet *36*, 233-278. 10.1146/annurev.genet.36.042902.092433.

Pomp, O., Dreesen, O., Leong, D.F., Meller-Pomp, O., Tan, T.T., Zhou, F., and Colman, A. (2011). Unexpected X chromosome skewing during culture and reprogramming of human somatic cells can be alleviated by exogenous telomerase. Cell Stem Cell *9*, 156-165. 10.1016/j.stem.2011.06.004.

Prokopuk, L., Stringer, J.M., White, C.R., Vossen, R., White, S.J., Cohen, A.S.A., Gibson, W.T., and Western, P.S. (2018). Loss of maternal EED results in postnatal overgrowth. Clin Epigenetics *10*, 95. 10.1186/s13148-018-0526-8.

Ragunathan, K., Jih, G., and Moazed, D. (2015). Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. Science *348*, 1258699. 10.1126/science.1258699.

Rastan, S., Kaufman, M.H., Handyside, A.H., and Lyon, M.F. (1980). X-chromosome inactivation in extra-embryonic membranes of diploid parthenogenetic mouse embryos demonstrated by differential staining. Nature *288*, 172-173.

Rastan, S., and Robertson, E.J. (1985). X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. J Embryol Exp Morphol *90*, 379-388.

Reich, A., Klatsky, P., Carson, S., and Wessel, G. (2011). The transcriptome of a human polar body accurately reflects its sibling oocyte. J Biol Chem *286*, 40743-40749. 10.1074/jbc.M111.289868.

Renfree, M.B., Hore, T.A., Shaw, G., Graves, J.A., and Pask, A.J. (2009). Evolution of genomic imprinting: insights from marsupials and monotremes. Annu Rev Genomics Hum Genet *10*, 241-262. 10.1146/annurev-genom-082908-150026.

Richart, L., Picod-Chedotel, M.L., Wassef, M., Macario, M., Aflaki, S., Salvador, M.A., Hery, T., Dauphin, A., Wicinski, J., Chevrier, V., et al. (2022). XIST loss impairs mammary stem cell

differentiation and increases tumorigenicity through Mediator hyperactivation. Cell *185*, 2164-2183 e2125. 10.1016/j.cell.2022.04.034.

Robert-Finestra, T., Tan, B.F., Mira-Bontenbal, H., Timmers, E., Gontan, C., Merzouk, S., Giaimo, B.D., Dossin, F., van, I.W.F.J., Martens, J.W.M., et al. (2021). SPEN is required for Xist upregulation during initiation of X chromosome inactivation. Nat Commun *12*, 7000. 10.1038/s41467-021-27294-5.

Rossant, J., and Tam, P.P.L. (2017). New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. Cell Stem Cell *20*, 18-28. 10.1016/j.stem.2016.12.004.

Russell, L.B. (1963). Mammalian X-chromosome action: inactivation limited in spread and region of origin. Science *140*, 976-978. 10.1126/science.140.3570.976.

Sado, T. (2017). What makes the maternal X chromosome resistant to undergoing imprinted X inactivation? Philos Trans R Soc Lond B Biol Sci *372*. 10.1098/rstb.2016.0365.

Sado, T., Wang, Z., Sasaki, H., and Li, E. (2001). Regulation of imprinted X-chromosome inactivation in mice by Tsix. Development *128*, 1275-1286. 10.1242/dev.128.8.1275.

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T.W., Kuoy, E., Langerman, J., Clark, A.T., Jaenisch, R., and Plath, K. (2017a). Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. Cell Stem Cell *20*, 87-101. 10.1016/j.stem.2016.10.006.

Sahakyan, A., Plath, K., and Rougeulle, C. (2017b). Regulation of X-chromosome dosage compensation in human: mechanisms and model systems. Philos Trans R Soc Lond B Biol Sci *372*. 10.1098/rstb.2016.0363.

Samanta, M., and Kalantry, S. (2020). Generating primed pluripotent epiblast stem cells: A methodology chapter. Curr Top Dev Biol *138*, 139-174. 10.1016/bs.ctdb.2020.01.005.

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 *138*, 2602. 10.1038/s41467-022-30352-1.

Sarkar, M.K., Gayen, S., Kumar, S., Maclary, E., Buttigieg, E., Hinten, M., Kumari, A., Harris, C., Sado, T., and Kalantry, S. (2015). An Xist-activating antisense RNA required for X-chromosome inactivation. Nat Commun *6*, 8564. 10.1038/ncomms9564.

Sathe, S.S., and Harte, P.J. (1995). The Drosophila extra sex combs protein contains WD motifs essential for its function as a repressor of homeotic genes. Mech Dev *52*, 77-87. 10.1016/0925-4773(95)00392-e.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat Med *10*, 55-63. 10.1038/nm979.

Schaefer, K.N., and Peifer, M. (2019). Wnt/Beta-Catenin Signaling Regulation and a Role for Biomolecular Condensates. Dev Cell *48*, 429-444. 10.1016/j.devcel.2019.01.025.

Schoeftner, S., Sengupta, A.K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., and Wutz, A. (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. EMBO J *25*, 3110-3122. 10.1038/sj.emboj.7601187.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. Cell *128*, 735-745. 10.1016/j.cell.2007.02.009.

Selleri L, Bartolomei MS, Bickmore WA, He L, Stubbs L, Reik W, Barsh GS. A Hox-Embedded Long Noncoding RNA: Is It All Hot Air? PLoS Genet. 2016 Dec 15;12(12):e1006485. doi: 10.1371/journal.pgen.1006485. PMID: 27977680; PMCID: PMC5157941.

Sewalt, R.G., van der Vlag, J., Gunster, M.J., Hamer, K.M., den Blaauwen, J.L., Satijn, D.P., Hendrix, T., van Driel, R., and Otte, A.P. (1998). Characterization of interactions between the mammalian polycomb-group proteins Enx1/EZH2 and EED suggests the existence of different mammalian polycomb-group protein complexes. Mol Cell Biol *18*, 3586-3595. 10.1128/MCB.18.6.3586.

Sharman, G.B. (1971). Late DNA replication in the paternally derived X chromosome of female kangaroos. Nature *230*, 231-232. 10.1038/230231a0.

Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C., and Orkin, S.H. (2008a). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell *32*, 491-502. 10.1016/j.molcel.2008.10.016.

Shen, Y., Matsuno, Y., Fouse, S.D., Rao, N., Root, S., Xu, R., Pellegrini, M., Riggs, A.D., and Fan, G. (2008b). X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. Proc Natl Acad Sci U S A *105*, 4709-4714. 10.1073/pnas.0712018105.

Shin, J., Bossenz, M., Chung, Y., Ma, H., Byron, M., Taniguchi-Ishigaki, N., Zhu, X., Jiao, B., Hall, L.L., Green, M.R., et al. (2010). Maternal Rnf12/RLIM is required for imprinted X-chromosome inactivation in mice. Nature *467*, 977-981. nature09457 [pii] 10.1038/nature09457.

Shiura, H., and Abe, K. (2019). Xist/Tsix expression dynamics during mouse peri-implantation development revealed by whole-mount 3D RNA-FISH. Sci Rep *9*, 3637. 10.1038/s41598-019-38807-0.

Shumacher, A., Faust, C., and Magnuson, T. (1996). Positional cloning of a global regulator of anterior-posterior patterning in mice. Nature *383*, 250-253.

Shvetsova, E., Sofronova, A., Monajemi, R. et al. Skewed X-inactivation is common in the general female population. Eur J Hum Genet 27, 455–465 (2019). https://doi.org/10.1038/s41431-018-0291-3

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T.B., Webster, Z., Peters, A.H., Jenuwein, T., Otte, A.P., and Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. Dev Cell *4*, 481-495. 10.1016/s1534-5807(03)00068-6.

Silva, S.S., Rowntree, R.K., Mekhoubad, S., and Lee, J.T. (2008). X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc Natl Acad Sci U S A *105*, 4820-4825. 0712136105 [pii]
10.1073/pnas.0712136105.

Simon, J.A., and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol *10*, 697-708. 10.1038/nrm2763.

Singh, A.M., Bechard, M., Smith, K., and Dalton, S. (2012). Reconciling the different roles of Gsk3beta in "naive" and "primed" pluripotent stem cells. Cell Cycle *11*, 2991-2996. 10.4161/cc.21110.

Sprague, D., Waters, S.A., Kirk, J.M., Wang, J.R., Samollow, P.B., Waters, P.D., and Calabrese, J.M. (2019). Nonlinear sequence similarity between the Xist and Rsx long noncoding RNAs suggests shared functions of tandem repeat domains. RNA *25*, 1004-1019. 10.1261/rna.069815.118.

Sripathy, S., Leko, V., Adrianse, R.L., Loe, T., Foss, E.J., Dalrymple, E., Lao, U., Gatbonton-Schwager, T., Carter, K.T., Payer, B., et al. (2017). Screen for reactivation of MeCP2 on the inactive X chromosome identifies the BMP/TGF-beta superfamily as a regulator of XIST expression. Proc Natl Acad Sci U S A *114*, 1619-1624. 10.1073/pnas.1621356114.

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes Dev *24*, 2239-2263. 10.1101/gad.1963910.

Stambolic, V., Ruel, L., and Woodgett, J.R. (1996). Lithium inhibits glycogen synthase kinase-3 activity and mimics wingless signalling in intact cells. Curr Biol *6*, 1664-1668. 10.1016/s0960-9822(02)70790-2.

Stuart, T., and Satija, R. (2019). Integrative single-cell analysis. Nat Rev Genet *20*, 257-272. 10.1038/s41576-019-0093-7.

Sun, C., Zhang, J., Zheng, D., Wang, J., Yang, H., and Zhang, X. (2018). Transcriptome variations among human embryonic stem cell lines are associated with their differentiation propensity. PLoS One *13*, e0192625. 10.1371/journal.pone.0192625.

Sun, S., Payer, B., Namekawa, S., An, J.Y., Press, W., Catalan-Dibene, J., Sunwoo, H., Lee, J.T (2015). Xist imprinting is promoted by the hemizygous (unpaired) state in the male germ line. Proc Natl Acad Sci U S A. 2015 Nov 24;112(47):14415-22. doi: 10.1073/pnas.1519528112. Epub 2015 Oct 21. PMID: 26489649; PMCID: PMC4664331.

Syrett, C.M., Sierra, I., Berry, C.L., Beiting, D., and Anguera, M.C. (2018). Sex-Specific Gene Expression Differences Are Evident in Human Embryonic Stem Cells and During In Vitro Differentiation of Human Placental Progenitor Cells. Stem Cells Dev *27*, 1360-1375. 10.1089/scd.2018.0081.

Tada, T., Obata, Y., Tada, M., Goto, Y., Nakatsuji, N., Tan, S., Kono, T., and Takagi, N. (2000). Imprint switching for non-random X-chromosome inactivation during mouse oocyte growth. Development *127*, 3101-3105. 10.1242/dev.127.14.3101.

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.

Takada, T., Ebata, T., Noguchi, H., Keane, T.M., Adams, D.J., Narita, T., Shin, I.T., Fujisawa, H., Toyoda, A., Abe, K., et al. (2013). The ancestor of extant Japanese fancy mice contributed to the mosaic genomes of classical inbred strains. Genome Res *23*, 1329-1338. 10.1101/gr.156497.113.

Takagi, N. (1980). Primary and secondary nonrandom X chromosome inactivation in early female mouse embryos carrying Searle's translocation T(X; 16)16H. Chromosoma *81*, 439-459. 10.1007/BF00368155.

Takagi, N., and Abe, K. (1990). Detrimental effects of two active X chromosomes on early mouse development. Development *109*, 189-201.

Takagi, N., and Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature *256*, 640-642.

Takagi, N., Wake, N., and Sasaki, M. (1978). Cytologic evidence for preferential inactivation of the paternally derived X chromosome in XX mouse blastocysts. Cytogenet Cell Genet *20*, 240-248. 10.1159/000130856.

Takahashi, S., Kobayashi, S., and Hiratani, I. (2018). Epigenetic differences between naive and primed pluripotent stem cells. Cell Mol Life Sci *75*, 1191-1203. 10.1007/s00018-017-2703-x.
Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. Cell *158*, 1254-1269. 10.1016/j.cell.2014.08.029.

Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072-2075. 10.1126/science.282.5396.2072.

Tang, F., Barbacioru, C., Nordman, E., Li, B., Xu, N., Bashkirov, V.I., Lao, K., and Surani, M.A. (2010). RNA-Seq analysis to capture the transcriptome landscape of a single cell. Nat Protoc *5*, 516-535. 10.1038/nprot.2009.236.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., et al. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. Cell *148*, 664-678. 10.1016/j.cell.2011.12.029.

Tchieu, J., Kuoy, E., Chin, M.H., Trinh, H., Patterson, M., Sherman, S.P., Aimiuwu, O., Lindgren, A., Hakimian, S., Zack, J.A., et al. (2010). Female human iPSCs retain an inactive X chromosome. Cell Stem Cell *7*, 329-342. 10.1016/j.stem.2010.06.024.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196-199. 10.1038/nature05972.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. Cell Stem Cell *15*, 524-526. 10.1016/j.stem.2014.09.003.

Thomson, D.W., and Dinger, M.E. (2016). Endogenous microRNA sponges: evidence and controversy. Nat Rev Genet *17*, 272-283. 10.1038/nrg.2016.20.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147. 10.1126/science.282.5391.1145.

Tie, F., Stratton, C.A., Kurzhals, R.L., and Harte, P.J. (2007). The N terminus of Drosophila ESC binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3 lysine 27. Mol Cell Biol *27*, 2014-2026. 10.1128/MCB.01822-06.

Tomoda, K., Takahashi, K., Leung, K., Okada, A., Narita, M., Yamada, N.A., Eilertson, K.E., Tsang, P., Baba, S., White, M.P., et al. (2012). Derivation conditions impact X-inactivation status in female human induced pluripotent stem cells. Cell Stem Cell *11*, 91-99. 10.1016/j.stem.2012.05.019.

Vallot, C., Huret, C., Lesecque, Y., Resch, A., Oudrhiri, N., Bennaceur-Griscelli, A., Duret, L., and Rougeulle, C. (2013). XACT, a long noncoding transcript coating the active X chromosome in human pluripotent cells. Nat Genet *45*, 239-241. 10.1038/ng.2530.

Vallot, C., Ouimette, J.F., Makhlouf, M., Feraud, O., Pontis, J., Come, J., Martinat, C., Bennaceur-Griscelli, A., Lalande, M., and Rougeulle, C. (2015). Erosion of X Chromosome Inactivation in Human Pluripotent Cells Initiates with XACT Coating and Depends on a Specific Heterochromatin Landscape. Cell Stem Cell *16*, 533-546. 10.1016/j.stem.2015.03.016.

Vallot, C., Patrat, C., Collier, A.J., Huret, C., Casanova, M., Liyakat Ali, T.M., Tosolini, M., Frydman, N., Heard, E., Rugg-Gunn, P.J., and Rougeulle, C. (2017). XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. Cell Stem Cell *20*, 102-111. 10.1016/j.stem.2016.10.014.

van der Vlag, J., and Otte, A.P. (1999). Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. Nat Genet *23*, 474-478. 10.1038/70602.

van Otterdijk, S.D., and Michels, K.B. (2016). Transgenerational epigenetic inheritance in mammals: how good is the evidence? FASEB J *30*, 2457-2465. 10.1096/fj.201500083.

Wang, F., Shin, J., Shea, J.M., Yu, J., Boskovic, A., Byron, M., Zhu, X., Shalek, A.K., Regev, A., Lawrence, J.B., et al. (2016). Regulation of X-linked gene expression during early mouse development by Rlim. Elife *5*. 10.7554/eLife.19127.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature *431*, 873-878. 10.1038/nature02985.

Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A., and Magnuson, T. (2001). Imprinted X inactivation maintained by a mouse Polycomb group gene. Nat Genet *28*, 371-375.

Ward W. and Coffey D (1991). DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. Biol Reprod 44:569–74.

Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., et al. (2014). Derivation of naive human embryonic stem cells. Proc Natl Acad Sci U S A *111*, 4484-4489. 10.1073/pnas.1319738111.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. Nat Rev Mol Cell Biol *17*, 155-169. 10.1038/nrm.2015.28.

West, J.D., Frels, W.I., Chapman, V.M., and Papaioannou, V.E. (1977). Preferential expression of the maternally derived X chromosome in the mouse yolk sac. Cell *12*, 873-882. 10.1016/0092-8674(77)90151-9.

Wu, C.F., Tsung, H.C., Zhang, W.J., Wang, Y., Lu, J.H., Tang, Z.Y., Kuang, Y.P., Jin, W., Cui, L., Liu, W., and Cao, Y.L. (2005). Improved cryopreservation of human embryonic stem cells with trehalose. Reprod Biomed Online *11*, 733-739. 10.1016/s1472-6483(10)61692-6.

Wu, J., Huang, B., Chen, H. et al. (2016) The landscape of accessible chromatin in mammalian preimplantation embryos. Nature 534, 652–657. https://doi.org/10.1038/nature18606

Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat Genet *30*, 167-174. 10.1038/ng820.

Wykes S. and Krawetz S (2003). The structural organization of sperm chromatin. J Biol Chem. 278:29471–7.

Xie, P., Ouyang, Q., Leng, L., Hu, L., Cheng, D., Tan, Y., Lu, G., and Lin, G. (2016). The dynamic changes of X chromosome inactivation during early culture of human embryonic stem cells. Stem Cell Res *17*, 84-92. 10.1016/j.scr.2016.05.011.

Xu, C., Rosler, E., Jiang, J., Lebkowski, J.S., Gold, J.D., O'Sullivan, C., Delavan-Boorsma, K., Mok, M., Bronstein, A., and Carpenter, M.K. (2005). Basic fibroblast growth factor supports undifferentiated human embryonic stem cell growth without conditioned medium. Stem Cells *23*, 315-323. 10.1634/stemcells.2004-0211.

Xu R, Li C, Liu X, Gao S (2021). Insights into epigenetic patterns in mammalian early embryos. Protein Cell. Jan;12(1):7-28. doi: 10.1007/s13238-020-00757-z. PMID: 32671792; PMCID: PMC7815849.

Xu, Z., Robitaille, A.M., Berndt, J.D., Davidson, K.C., Fischer, K.A., Mathieu, J., Potter, J.C., Ruohola-Baker, H., and Moon, R.T. (2016). Wnt/beta-catenin signaling promotes self-renewal and inhibits the primed state transition in naive human embryonic stem cells. Proc Natl Acad Sci U S A *113*, E6382-E6390. 10.1073/pnas.1613849113.

Yalcin, B., Wong, K., Agam, A., Goodson, M., Keane, T.M., Gan, X., Nellaker, C., Goodstadt, L., Nicod, J., Bhomra, A., et al. (2011). Sequence-based characterization of structural variation in the mouse genome. Nature *477*, 326-329. 10.1038/nature10432.

Yang, L., Kirby, J.E., Sunwoo, H., and Lee, J.T. (2016). Female mice lacking Xist RNA show partial dosage compensation and survive to term. Genes Dev *30*, 1747-1760. 10.1101/gad.281162.116.

Yildirim, E., Kirby, J.E., Brown, D.E., Mercier, F.E., Sadreyev, R.I., Scadden, D.T., and Lee, J.T. (2013). Xist RNA is a potent suppressor of hematologic cancer in mice. Cell *152*, 727-742. 10.1016/j.cell.2013.01.034.

Zee, B.M., Levin, R.S., DiMaggio, P.A., and Garcia, B.A. (2010). Global turnover of histone post-translational modifications and variants in human cells. Epigenetics Chromatin *3*, 22. 10.1186/1756-8935-3-22.

Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., Ming, J., Wu, X., Zhang, Y., Xu, Q., et al. (2016). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature *537*, 553-557. 10.1038/nature19361.

Zhang, C.U., Blauwkamp, T.A., Burby, P.E., and Cadigan, K.M. (2014). Wnt-mediated repression via bipartite DNA recognition by TCF in the Drosophila hematopoietic system. PLoS Genet *10*, e1004509. 10.1371/journal.pgen.1004509.

Zhang, T., Cooper, S., and Brockdorff, N. (2015). The interplay of histone modifications - writers that read. EMBO Rep *16*, 1467-1481. 10.15252/embr.201540945.

Zheng, H., Huang, B., Zhang, B., Xiang, Y., Du, Z., Xu, Q., Li, Y., Wang, Q., Ma, J., Peng, X., et al. (2016). Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. Mol Cell *63*, 1066-1079. 10.1016/j.molcel.2016.08.032.

Zubek, J., Stitzel, M.L., Ucar, D., and Plewczynski, D.M. (2016). Computational inference of H3K4me3 and H3K27ac domain length. PeerJ *4*, e1750. 10.7717/peerj.1750.

Zylicz, J.J., Bousard, A., Zumer, K., Dossin, F., Mohammad, E., da Rocha, S.T., Schwalb, B., Syx, L., Dingli, F., Loew, D., et al. (2019). The Implication of Early Chromatin Changes in X Chromosome Inactivation. Cell *176*, 182-197 e123. 10.1016/j.cell.2018.11.041.