Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Jadiel A. Wasson **Date**

David Katz PhD, Advisor

Xiaodong Cheng PhD, Committee Member

Roger Deal PhD, Committee Member

Victor Faundez PhD, Committee Member

Paula Vertino PhD, Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

AN INVESTIGATION OF THE MATERNAL ROLE OF LYSINE SPECIFIC HISTONE DEMETHYLASE 1 (LSD1) AND THE FUNCTIONAL CONSEQUENCES ON DEVELOPMENT

By

Jadiel A. Wasson

B.S., University of Georgia, 2009

Advisor: David J. Katz, Ph.D.

An abstract of

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in the Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell, and Developmental Biology

2016

Abstract

AN INVESTIGATION OF THE MATERNAL ROLE OF LYSINE SPECIFIC HISTONE DEMETHYLASE 1 (LSD1) AND THE FUNCTIONAL CONSEQUENCES ON DEVELOPMENT

By Jadiel A. Wasson

Classically, the concept of inheritance has been restricted to the passage of DNA. However, in the past couple of decades, have come to understand that information associated with DNA independent of sequence – epigenetic information– can also be inherited and affect phenotypic outcomes across cell divisions. In this dissertation, I will discuss these epigenetic mechanisms and their influence on gene expression. I will also discuss the need to reprogram certain epigenetic information in the context of fertilization when genomic and epigenomic information is passed between generations. The main objective of this dissertation is to investigate the maternal contribution of the histone demethlyase LSD1 and to determine its function at fertilization. I argue that not only is LSD1 reprogramming essential for development but that even slight disturbances in reprogramming at fertilization can result in far-reaching consequences on phenotypes including behavioral aberrations in adult animals.

.

AN INVESTIGATION OF THE MATERNAL ROLE OF LYSINE SPECIFIC HISTONE DEMETHYLASE 1 (LSD1) AND THE FUNCTIONAL CONSEQUENCES ON DEVELOPMENT

By

Jadiel A. Wasson

B.S., University of Georgia, 2009

Advisor: David J. Katz, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in the Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell, and Developmental Biology

2016

ACKNOWLEDGEMENTS

I would first like to thank my family and friends who made it possible to achieve what I have. Upon your love I have succeeded. Mom, Thank you for teaching me what it means to be gracious and how to never give up on your dreams. I have learned what perseverance means based on your example. Dad, Thank you for teaching me patience and understanding. Your words of encouragement have kept me afloat. In the darkest times, I always remembered "Go on and get your PhD." I will always remember that you both impressed on me that each generation should be better than the last. These words I have taken to heart and have pushed me to continue to achieve. Mario, Thank you for being my greatest role model. Everything I have pursued and many things I have accomplished were because of the strong example you set. Jahnae, Thank you for reminding me what it takes to be a strong example and for keeping me striving to be best I can. Next thank you to my friends and my classmates, both old and new that made the dark times brighter and hilarious as all hell. Thank you for your encouragement and honesty.

Thank you to the Biochemistry, Cell and Developmental Biology Program and the many advisors that I have had along the way. Thank you to the Katz lab for making my lab life enjoyable. Our daily jokes and conversations not only keep me aware but also kept me thinking and laughing. Dave, Thank you for being the mentor that I needed and for your understanding. I truly would not have made it through without your wisdom.

I'm out!

TABLE OF CONTENTS

Chapter 3:
A resource for analysis of DNA methylation status of genomically imprinted loci

Chapter 4:
LSD1 maternal function and its effect on development

CHAPTER 1:

AN INTRODUCTION TO EPIGENETIC MECHANISMS AND REPROGRAMMING OF CELLULAR IDENTITY

Parts of this section were published here:
Restoring Totipotency Through Epigenetic Reprogramming. Briefings in Functional Genomics 2013 Mar;12(2):118-28. doi: 10.1093/bfgp/els042. Epub 2012 Oct 31. PMID:23117862

Epigenetic modifications are implicated in the maintenance and regulation of transcriptional memory by marking genes that were previously transcribed to facilitate transmission of these expression patterns through cell divisions. However, during two major points in mammalian development, germline specification and fertilization, these epigenetic modifications are reprogrammed. For example, during germline specification and maintenance, extensive epigenetic modifications are acquired while others are erased to ensure proper formation of the gametes, sperm and egg. At fertilization, the fusion of the highly differentiated sperm and egg results in formation of the totipotent zygote. This massive change in cell fate implies that the selective erasure and maintenance of epigenetic modifications at fertilization may be critical for the re-establishment of totipotency. In this section, I will detail studies that have provided insight into the extensive epigenetic reprogramming that occurs around fertilization and the mechanisms that may be involved in the re-establishment of totipotency in the embryo. In addition, I will introduce lysine specific demethylase 1 (LSD1) as a new player in reprogramming at fertilization.

1.1 EPIGENETIC MODIFICATIONS AND THEIR REGULATORY PROTEINS

It is estimated that humans have over 30 trillion cells[1]. Though each of these cells contains the same DNA template, there are over two hundred cell types that are not only morphologically distinct but perform very different functions. These distinct cell types are created by transcriptional differences that are maintained in part by the selective usage of DNA, which is regulated by epigenetic mechanisms.

The term "epigenetic" was first coined by C.H. Waddington in 1942 when he described how gene regulation can influence development[2]. Epigenetics was recently further defined in a Cold Spring Harbor meeting in 2008 as the "stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence"[3]. These "changes in a chromosome" affect accessibility to DNA by either direct modifications to DNA or modifications made to the proteins associated with it. Together, these form the basic unit of chromatin known as the nucleosome[4–6].

Nucleosomes are composed of histones that can be chemically modified and these modifications of the amino-terminal tails are often correlated with the transcription status of the genes that are packaged around them. In addition, certain histone modifications are typically associated with particular histone variants, which contain small numbers of amino acid substitutions compared to canonical histones and can substitute for their corresponding canonical histones within the nucleosome[6,7,8]. Histones have N-terminal tails that extend out from the core, which can be modified through the addition of chemical groups to certain amino acid residues present in the tail. Acetylation, methylation, and phosphorylation are the most characterized modifications, but other modifications exist and continue to be discovered[5].

Transcriptional activity is influenced by several factors: which tail residue is modified, which modification is added, and the number of modifications added (for example, mono-, di-, or tri-methylation). The combination of these factors can sometimes have predictable results on gene expression, leading to the identification of a histone code[9–11]. For example, addition of an acetyl group to lysine 9 on histone 3 (H3K9) correlates with active transcription[12] whereas trimethylation of H3K9 correlates with transcriptional repression[13,14]. The number of modifications to a residue also can alter the biological effect of a particular region of DNA. For example, monomethylation of lysine 4 on histone 3 (H3K4me1)[15] is associated with enhancer regions while trimethylation of the same residue, H3K4me3[16], is associated with actively transcribed promoter regions. Furthermore, histone modifications may have less direct effects on gene expression. For example H3K4me2[17,18] is present in gene bodies and is suggested to act as a type of transcriptional memory modification.

DNA itself can be modified through the addition of methyl groups to the 5' position of cytosine bases. DNA methylation occurs mainly in the context CpG dinucleotides but it has also been demonstrated to occur in a CpA context. Despite occurring mainly in a CpG context, CpG dinucleotides that cluster together to create CpG islands are not methylated. These regions tend to be present at gene promoters, ensuring that genes are not silenced. Classically, DNA methylation was established as a repressive modification as it is known to associate with silenced genes and at repetitive elements. However, DNA methylation is also present within the bodies of genes that have been actively transcribed. It has recently been discovered that methylation in the form of 5-methylcytosine (5mC) can be converted to 5-hydroxymethylcytosine (5hmC), 5-carboxylcytosine (5caC), and 5-formylcytosine (5fC) by the Ten-eleven translocation (TET) enzymes TET1, TET2, and TET3[19]. This conversion is thought to enable demethylation of DNA. Although 5mC has been the most studied and

characterized modifications, 5hmC has begun to be appreciated as having significant biological effects including being involved in neurodevelopment, memory formation, and neurological function[20–23].

Enzymes that regulate the epigenetic landscape of a cell contain domains that allow them to act as readers, writers, and erasers of chromatin modifications, with some enzymes acting in multiple capacities depending on context. Many chromatin modifying enzymes contain reader domains such bromodomains, plant homeodomain (PHD) fingers, and Tudor domains that dock onto various histone modifications and recognize the chromatin state. Some proteins that contain methyl binding domains (MBDs), such as MeCP2, can bind to methylated cytosines in response to DNA methylation status[24,25].

In addition to recognizing specific modifications in chromatin, many histone modifying enzymes will either add or remove a modification. Histone modifying enzymes can contain "reader" domains in addition to either "writer" or "eraser" domains. For example, SET domains, named for their presence in *Drosophila* Su(var)3-9 and Enhancer of zeste enzymes, are present in almost all histone lysine methyltransferases[26,27]. Within the SET domain, various amino acid substitutions confer alterations in the ability to recognize and modify different methylation states on different histone tails[26,27]. One example of a "writer" enzyme is mixed-lineage leukemia protein 2 (MLL2). MLL2 not only recognizes the methylation status of H3K4 through its SET domain, it also catalyzes methylation of K4 using S-adenomethionine (SAM) as a methyl donor[28]. Alternatively, members of the Jumonji C family of histone demethylases act as "erasers"[29]. They are Fe(II)- and a-ketoglutarate dependent dioxygenases that demethylate histones via oxidative hydroxylation[29]. For example, KDM5B, a member of this family, can catalyze the removal of all H3K4 methylation species[30].

Epigenetic modifications on DNA and histones can have significant functional overlap. Both H3K9 methylation and DNA methylation are associated with highly repressed genomic regions like transposable elements, centromeres and telomeres[13,14,31–33]. Furthermore, recent studies have demonstrated an interplay between DNA methylation acquisition and H3K4 methylation[34]. DNMT3L, the cofactor for the DNA methyltransferases DNMT3A/B, reads the modification status of histones[34]. In order for the DNA methyltransferase complex to methylate DNA, it must dock onto the histone tail. H3K4 methylation blocks this docking[34]. This suggests a requirement to remove this modification in order to acquire DNA methylation.

The Amine-Oxidase family of histone demethylases

LSD1 and its paralog, LSD2, are both epigenetic "erasers" that act as lysine specific histone demethylases. LSD1 was first characterized in 2004 by Yang Shi and colleagues[35], and homologs have been identified in taxa as diverse as *S. pombe* and humans. It is one of the most conserved genes at the sequence level and has orthologs in every organism from *S. pombe* to humans[35]. Both LSD1 and LSD2 catalyze the removal of just H3K4me1/2 through a mechanism that is dependent on FAD as a cofactor in an amine oxidase reaction[35]. Because of this chemistry, they can only remove mono- and dimethyl groups from lysines and cannot remove trimethyl groups like the Jumonji C family demethylases[35]. Based on their activity towards H3K4 methylation, both LSD1 and LSD2 function may be required for the establishment of DNA methylation[34–36].

LSD2

Although both LSD1 and LSD2 catalyze demethylation via their amine oxidase domains, they contain different domains that could confer a separation in their functions (Figure 1)[37]. For example, LSD1 associates with its binding partners through its tower domain while LSD2 lacks this domain[36,38]. LSD2 has a zinc-finger domain that can associate directly with DNA[36,38,39]. Although LSD1 is widely expressed in many tissues of the mouse, LSD2 is mainly expressed in the oocyte where it has been demonstrated to function in the establishment of DNA methylation at imprinted loci[40]. In fact, mice that were null for *Lsd2* (*Lsd2*$^{-/-}$) were completely viable. Female mice, however, gave rise to a maternal effect embryonic lethal. Progeny from *Lsd2* mutant mothers do not survive past embryonic day 10.5[40]. Both oocytes from *Lsd2*$^{-/-}$ females and the progeny of these females completely lacked DNA methylation at four out of seven assayed imprinted loci[40] (fully described in Chapter 3). These data suggest a role for LSD2 in the establishment of a subset of imprinted loci. These findings are consistent with the idea that demethylation of H3K4 is required for the acquisition of DNA methylation. However, since only a subset of imprinted loci were affected upon loss of *Lsd2*, another enzyme, perhaps LSD1, may be a required for the acquisition of DNA methylation at the remaining imprinted loci.

LSD1

LSD1 is an enzyme that performs many different biological functions. It operates in a variety of cell types, and its specificity is due to its cell-type dependent interactions with various binding partners. In addition to demethylating K3K4me1/2, LSD1 has also been reported to remove H3K9me1/2 and H4K20me1/2. In complex with CoREST, LSD1 has been reported to remove H3K4me1/2 in order to repress neuronal genes in non-neuronal

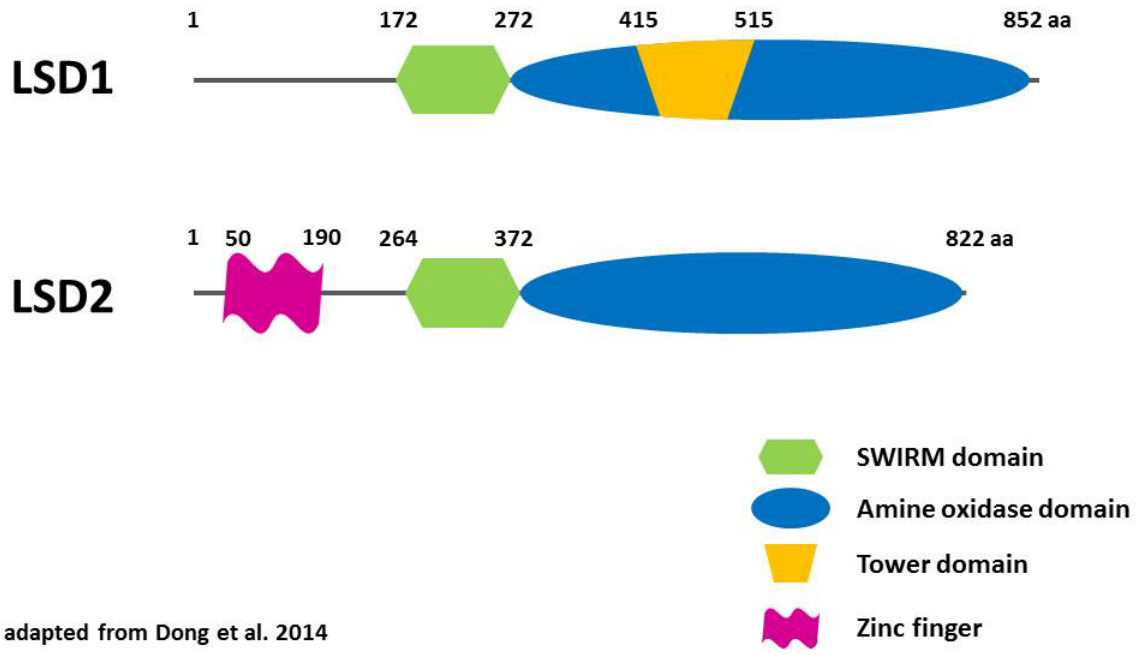


Figure 1: Domain Schematic of LSD1 and LSD2
LSD1 is 852 amino acids long and contains a SWIRM domain, a Tower domain and an amine oxidase domain where it catalyzes lysine demethylation. LSD2 is 822 amino acids in length and contains a Zinc finger, a SWIRM domain and an amine oxidase domain.

cell lineages[41–43]. Through demethylating H3K4me1/2, LSD1 has been implicated in the silencing of endogenous retroviruses and in the formation of heterochromatin[44–46]. When complexed with either the androgen receptor or estrogen receptor, LSD1 can remove H3K9me1/2[47–52]. LSD1 is also alternatively spliced into a neuronal specific isoform of LSD1, nLSD1[53]. This isoform has been reported to demethylate H3K9me1/2 when complexed with Supervillin or H4K20me1/2 when complexed with CoREST[54,55].

Initially, LSD1 was identified as a member of a CtBP2 repressor complex in HeLa cells[56–59]. During B-cell development, LSD1 associates with Blimp-1 to allow plasma cell differentiation through the repression of genes associated with the mature B-cell transcription program[60]. In pituitary gland development, LSD1 regulates cell differentiation through association with two separate complexes; ZEB1, a Kruppel-like repressor, or CSL, a Notch-interacting transcription factor[43]. Other binding partners of LSD1 include, but are not limited to, Snai1, HDAC1/2, BRAF35, PRKCB, AXL1, NuRD, BHC80 and ZFP516[61]. In addition to demethylating histones, LSD1 demethylates non-histone targets including p53 and DNMT1, affecting their stability[62,63]. In general, evidence suggests that the main function of LSD1 is in cell differentiation.

LSD1 is regulated in part by a microRNA cluster miR-137. In colon cancer and in neuroblastoma cells, LSD1 mRNA is targeted for downregulation by miR-137[64,65]. LSD1 can be phosphorylated by protein kinase Cα (PKCα) which allows it to complex with CLOCK and BMAL1 to regulate circadian rhythyms[66]. In addition, the subcellular location of LSD1 is regulated by cell cycle changes[67] and the stability of LSD1 is mediated not only by CoREST but also through deubiquitination by USP28[68,69].

Despite sharing similar enzymatic mechanisms, LSD1 and LSD2 appear to be functionally distinct. Complete loss of *Lsd2* results in viable mice, with females giving rise to

embryonic lethal progeny[40]. Alternatively, complete loss of *Lsd1* results in embryonic lethality by embryonic day 7.5[43,63]. This difference in null phenotype suggests that these enzymes function in different contexts. However, it is unknown whether there may also be functional redundancy between LSD1 and LSD2. As LSD2 functions in the oocyte to regulate DNA methylation specifically at imprinted loci, we hypothesized that LSD1 may have a similar role in DNA methylation establishment. Alternatively, previous work has demonstrated a role for LSD1 in cell differentiation, which may indicate that LSD1 is involved in the fate change that occurs at fertilization.

.

Epigenetic memory: defining cellular identity and that of its daughter cells

Epigenetic information and the enzymes that regulate this information influence gene expression and chromatin accessibility. Together, they create a chromatin landscape that defines a cell's identity. But when a cell changes its identity during differentiation or at fertilization, this chromatin landscape must be regulated and altered. Failure to do so can result in severe consequences for the cell or for the entire organism. In symmetric cell divisions, a cell divides to create two identical daughter cells. Some epigenetic information is retained by the daughters, which acts as an epigenetic memory that informs them of their environment as well as the transcriptional program that they should express[70–72].

Although we do not fully understand this process, one of the most well studied cases of epigenetic memory is in *Arabidopsis thaliana*. The floral repressor locus Flowering Locus C (FLC) is down-regulated by prolonged exposure to cold, which occurs during winter vernalization[73–77]. FLC is initially highly expressed in plants and represses flowering. H3K4me3, H3K36me3 and histone acetylation are all associated with the FLC when it is actively expressed[78,79]. Upon prolonged exposure to cold temperatures, FLC is repressed and

is marked by H3K27me3[74,80–82]. Because this repression is epigenetic in nature (and does not affect the underlying DNA sequence), it is then propagated even when a plant is exposed to warm temperature[79]. The repression of FLC is maintained through a complex that is homologous to Polycomb repressive complex 2 (PRC2) which binds to H3K27 methylation and maintains a repressive state[74,80,81]. This repression continues until the next generation, when the plant embryo begins to develop and the locus is reprogrammed to allow expression. Berry and colleagues were the first to demonstrate that this example of epigenetic memory acts in *cis* at the FLC locus[83]. Two separate fluorescent reporters were used to label each FLC locus in the same cell. Upon exposure to cold temperatures, they saw that each locus was able to exist in its own expression state, allowing one locus to be expressed while the other was repressed. This mixed expression state could be stably propagated upon subsequent cell divisions by daughter cells[83].

Some of the most convincing data regarding epigenetic memory comes from work in *C. elegans*. Loss of function mutations in *spr-5*, an ortholog of the histone demethylase LSD1, results in sterility over multiple generations. This sterility is correlated with increasing H3K4me2 and the increasing expression of spermatogenesis genes over multiple generations[46,84–86]. In this work, epigenetic memory in the form H3K4me2 gets propagated over generations and in turn maintains the expression of sperm genes. This work adds further evidence that chromatin modifications and other *cis* factors can act as epigenetic memory.

The (epigenetic) memory problem: a change in cell fate

Epigenetic memory allows cells to achieve their intended transcriptional states from their initial formation. However, if this process goes awry, in the case of cell fate changes, the inappropriate maintenance of epigenetic memory may lead to severe defects.

In embryonic stem cells, pioneer transcription factors including Oct4 and Sox2 must bind to specific enhancers in order to maintain the totipotent stem cell fate[87–89]. These enhancers are distinctly marked with H3K4me1 which creates an open chromatin environment allowing the loci to be easily accessed by transcriptional machinery[15]. Upon differentiation, these enhancers become inaccessible which prevents binding of the stem cell state-inducing transcription factors. LSD1 is required for the transition from the stem cell fate to the differentiated cell fate[15]. Although LSD1 is not required for the maintenance of stem cell fate, when it is lost, ESCs cannot differentiate; this indicates the importance of histone methylation in the maintenance of cell fate[15]. In this context, H3K4 methylation at enhancers appears to act as an epigenetic memory, allowing cells to maintain their stem fate. Removal of this memory is required to allow a transition in cell fate.

Cancer cells are notorious for reverting their cell fate. In fact, many factors that are associated with stem-like phenotypes are reactivated and overexpressed in cancers[69,90–92]. Overexpression of LSD1 is prevalent in a number of cancers including leukemia, and is often associated with worse outcomes for patients[69,93–95]. LSD1 maintains stem cell-like properties in the blood cells that contribute to leukemia, allowing these cells to over-proliferate[93,96,97]. When overexpressed in mice, an isoform of LSD1 was sufficient to induce lymphoma[95]. Correspondingly, small molecule inhibitors of LSD1 have been developed and show promise in the treatment of cancers[93]. In a distinct type of acute myeloid leukemia,

inhibition of LSD1 increases H3K4me2 at myeloid lineage specific genes[95,96]. This suggests a role for LSD1 and H3K4me2 in cell fate decisions.

Here, we have discussed a role for epigenetic modifications in cell fate. Modifications to chromatin like H3K4 methylation can act as an epigenetic memory that helps maintain cell fate throughout cell divisions. Erasure of this modification is crucial for differentiation, while errors in this process can lead to faulty cell fate transitions and aberrant phenotypes ranging from incomplete differentiation or cancer. In this dissertation, I will elucidate how errors in erasure of H3K4 methylation during epigenetic reprogramming at fertilization can lead to aberrant phenotypes in adult mice.

1.2 NATURAL VERSUS ARTIFICIAL REPROGRAMMING

In 1893, August Weismann was among the first to appreciate that germ cells are specialized cells and that only these cells can give rise to offspring of the subsequent generation. Importantly, from the observation of several organisms, Weismann also concluded that the germ lineage occurs along 'germ-tracks' (the germline lineage) that undergo a high degree of differentiation to become the highly specialized sperm and eggs cells. Indeed, Weismann mused that 'I can see no advantage in objecting to describe a cell of the germ-track as a somatic cell.'[98]. Yet as Weismann also observed, following the completion of the germ track, the egg and sperm must come together at fertilization to produce a totipotent zygote. To reconcile these facts, Weismann proposed that the germ lineage contains a special germ plasm that passively allows the germ-track to proceed through these 'somatic events' while retaining the ability to return to totipotency following fertilization[98]. Today, evidence from a number of model systems is beginning to provide molecular evidence of Weismann's vision.

Since all tissues must be specified from the same set of genes, the process of tissue differentiation is inherently epigenetic. Over the past several years, we have begun to understand epigenetic gene regulation and how epigenetic phenomena are used to control differentiation during development. This understanding has led to an opportunity to elucidate how the germ lineage is specified and maintained at the epigenetic level and how this specification may have to be reversed after fertilization to restore totipotency. In this review, we will not attempt to exhaustively present all that is known about epigenetics in the germline cycle [99]. Rather, we will highlight recent epigenetic germline studies in both invertebrates and vertebrates that hint that extensive epigenetic reprogramming occurs

around fertilization. This reprogramming is likely critical to maintain totipotency from one generation to the next.

Evidence for epigenetic reprogramming at fertilization

Cloning experiments performed in various vertebrates provide evidence for extensive epigenetic reprogramming at fertilization. Beginning with the *Xenopus laevis* cloning experiments performed by John Gurdon in the late 1950's, along with the cloning of Dolly the sheep in 1996, it has been possible to transform a somatic cell nucleus into a cloned animal by transplanting it into enucleated oocyte (Figure 2A) [100,101]. The donor chromosomes from the somatic nucleus inherently contain all of the genetic material necessary to produce all cells of the body, and yet these chromosomes have been epigenetically modified so they produce only the proteins that are necessary to specify and maintain the somatic cell type. Because of this differentiated state, it was once thought that it is impossible to reprogram a somatic nucleus to become totipotent. However, the successful cloning of Dolly and subsequently other animals through somatic cell nuclear transfer (SCNT) suggests, as Weismann correctly proposed, that the oocyte cytoplasm natively contains the necessary factors to perform this reprogramming. Recently, somatic reprogramming has been further revolutionized by Yamanaka and colleagues [102], who demonstrated that somatic reprogramming can be triggered by just four transcription factors (Oct4, Sox2, Nanog and c-Myc) (Figure 2A). Three of these four transcription factors are pluripotency factors thought to play a role in reprogramming the embryo back to pluripotency after fertilization. The success of the induced pluripotent stem cell process (iPS) confirms that natural reprogramming occurs through defined genetic pathways, and yet the mechanisms involved in natural reprogramming and in the induction of pluripotent stem cells remain largely

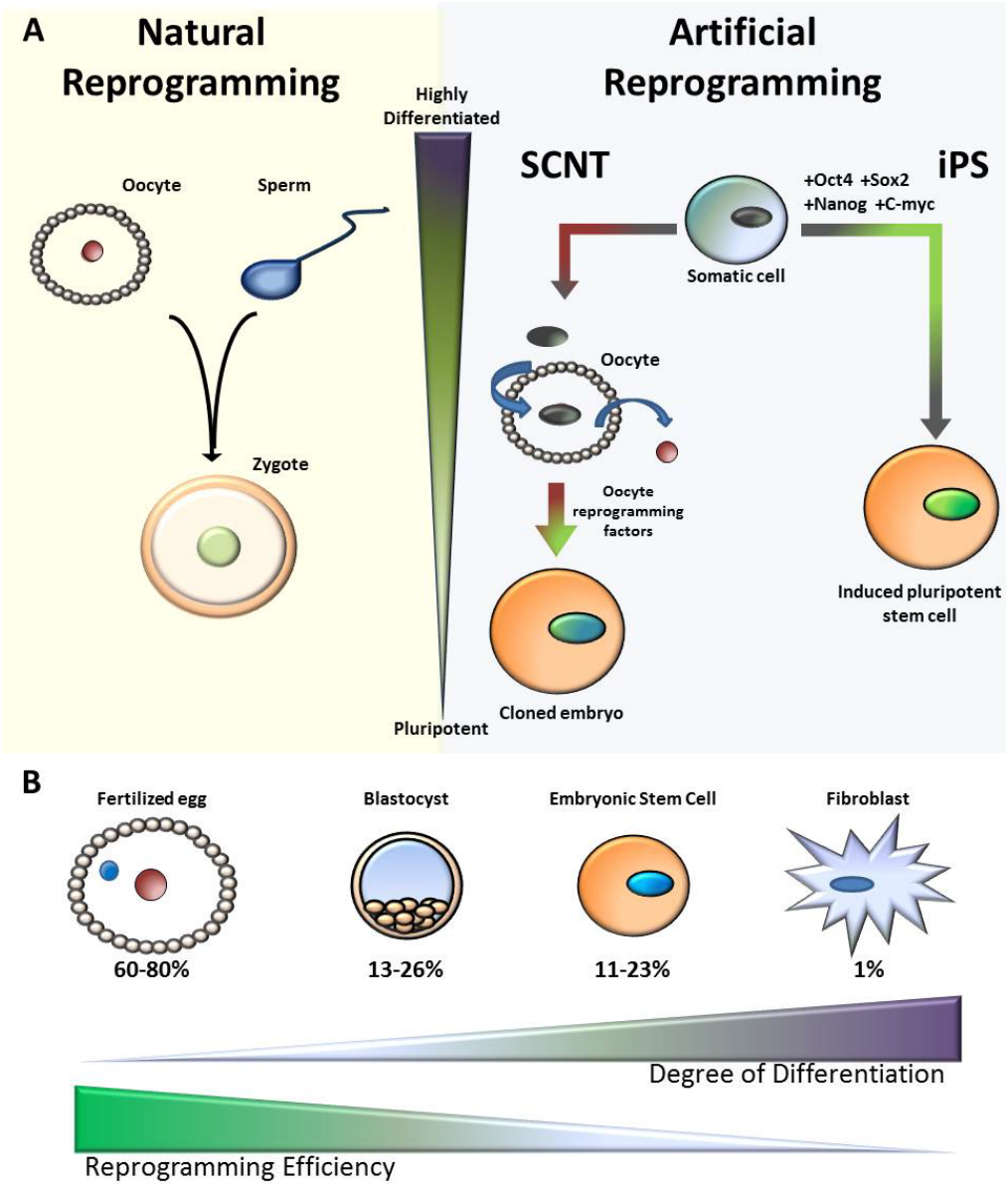


Figure 2: Natural reprogramming versus artificial reprogramming.
(A) The natural reprogramming that occurs when the highly differentiated sperm and egg come together at fertilization to generate the totipotent embryo is analogous to what occurs during SCNT and the iPS process. (B) The reprogramming efficiency of artificial reprogramming processes decreases with the increasing differentiation of the donor cell type.

*Restoring Totipotency Through Epigenetic Reprogramming. Briefings in Functional Genomics 2013 Mar;12(2):118-28. doi: 10.1093/bfgp/els042. Epub 2012 Oct 31. PMID:23117862

unknown. In addition, cloning through SCNT and iPS is highly inefficient compared to the normal process and often results in severe abnormalities, including kidney and liver defects as well as placental overgrowth[103]. These difficulties hint at the extent of epigenetic reprogramming that must occur naturally. Remarkably, when cloned animals survive and have offspring, the distinct abnormalities associated with the type of donor nuclei used are no longer observed[103]. This indicates that passage through the natural germline reprogramming process may be sufficient to revert the abnormal epigenetic state.

A clear illustration of the requirement for epigenetic resetting in somatic reprogramming can be seen by comparing the efficiency of generating a frog or mouse following SCNT from different donor cell types. In Xenopus, SCNT has been performed using embryonic donor cells with an efficiency of 36% versus an efficiency of 1.5% from differentiated cell types[104]. Similarly in mouse, the efficiencies obtained from cloning from a fertilized egg (60–80%), a blastomere (13–26%), an embryonic stem (ES) cell (11–23%) or a fibroblast cell (1%) decrease with increasing degrees of differentiation (Figure 2B)[105]. These data suggest that increasing epigenetic information acquired during differentiation is increasingly more difficult to reprogram. Yet, even in highly differentiated cases, it is possible to generate viable adult animals. This success, despite the limitations of the artificial process, proves that the oocyte is capable of a high degree of epigenetic reprogramming and implies that extensive epigenetic reprogramming takes place during normal reproduction.

A case for reprogramming epigenetic memory

Can the inefficiencies and failures of these processes teach us about the types of epigenetic reprogramming that must occur naturally? In addition to the frequent failure to obtain cloned animals, often cloned *Xenopus* embryos fail to turn on appropriate embryonic

genes and inappropriately express genes from the cell type they were cloned from[106,107]. It is apparent from these difficulties that the donor nuclei have been epigenetically programmed toward specific tissue fates and that this program is not always sufficiently reset during the cloning process. The nature of the epigenetic program in these differentiated donor nuclei is poorly understood. Thus, the mechanisms of reprogramming remain largely unknown. However, over the past several years, parts of the epigenetic program have begun to be elucidated.

One of the most well-characterized histone modifications is methylation of lysine 4 on histone H3 (H3K4me). Histone 3 lysine 4 can be mono-, di- or tri-methylated and these varying levels of methylation may have slightly different functions [108]. However, almost all H3K4me is associated with the histone variant H3.3 and is found at active genes[109,110]. This finding initially led to the hypothesis that H3K4me plays a role in gene activation. Recent work suggests that H3K4me may play a slightly different role in transcription. Accumulating evidence suggests that the acquisition of H3K4me is associated with RNA polymerase II elongation (Figure 3). For example, experiments in yeast have shown that the H3K4 methyltransferase, Set-1 (the yeast ortholog of the mammalian MLL), is in a complex with RNA polymerase II and is recruited to active genes by this interaction[17]. In addition, in *Drosophila melanogaster* S2 cells, high-resolution mapping of H3K4me2 and RNA polymerase II has demonstrated that their distribution closely matches genome wide[18]. These findings have led to a new model where H3K4me acts as an epigenetic memory to maintain transcription patterns during tissue differentiation rather than in de novo transcriptional activation (Figure 3).

Recently, this model has been more directly tested in *Dictyoselium discoideum* cells[111]. Employing a live-cell RNA imaging technique, Muramoto et al.[111] directly examined the

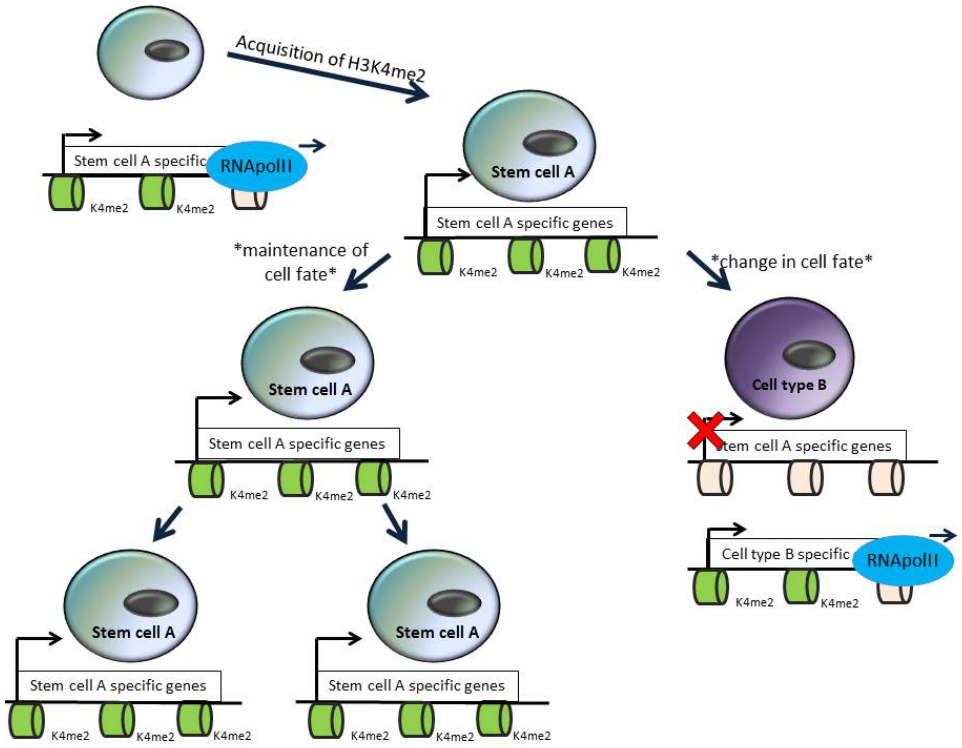


Figure 3: A model for H3K4me2 epigenetic memory.
H3K4me2 is acquired co-transcriptionally as RNA polymerase II elongates. H3K4me2 can be faithfully propagated as a cell divides and may act as an epigenetic transcriptional memory, but may have to be reprogrammed to allow for changes in cell fate.

*Restoring Totipotency Through Epigenetic Reprogramming. Briefings in Functional Genomics 2013 Mar;12(2):118-28. doi: 10.1093/bfgp/els042. Epub 2012 Oct 31. PMID:23117862

expression of genes during inheritance from mother to daughter. These experiments found that expression levels were more faithfully maintained in a cell lineage than in cells that are not lineage related. Furthermore, this epigenetic memory of active transcription was dependent upon H3K4 as well as the H3K4 methyltransferase Set-1.

Could H3K4me comprise part of the epigenetic signal that must be reset during SCNT and natural reproduction? The Gurdon Lab found that the inappropriate expression of endodermal genes was detected in *Xenopus* embryos derived from the transfer of differentiated endodermal nuclei. This inappropriate epigenetic memory was dependent upon lysine 4 of the histone variant H3.3[71]. If H3K4me2 functions in the maintenance of transcriptional patterns, then methylation on lysine 4 of histone H3.3 may have to be reprogrammed during SCNT and natural reproduction.

The best evidence that covalent epigenetic modifications can act as an epigenetic memory comes from the trithorax group (trxG) and Polycomb group (PcG) of genes. The trxG and PcG genes were originally identified in *Drosophila* as genes that are required to maintain the expression pattern of Hox genes after the initial set of transcription factors, which establish their expression, are no longer present (Figure 4A). These findings suggested that trxG and PcG act to maintain transcriptional memory[112]. In the absence of trxG and PcG proteins, flies exhibit homeotic transformations[112].

The founding member of the trxG, trithorax, encodes an H3K4 methyltransferase[113]. This suggests that the role of trithorax in tissue specification may be accomplished through H3K4me. Further evidence can be seen in a transgenic experiment performed in *Drosophila.* trxG and PcG genes act through cis-DNA response elements. One such element, Fab-7 was engineered on a transgene in Drosophila to surround a UAS-lacZ reporter so that the presence of the Fab-7 elements prevented a heat shock-inducible GAL4 driver from

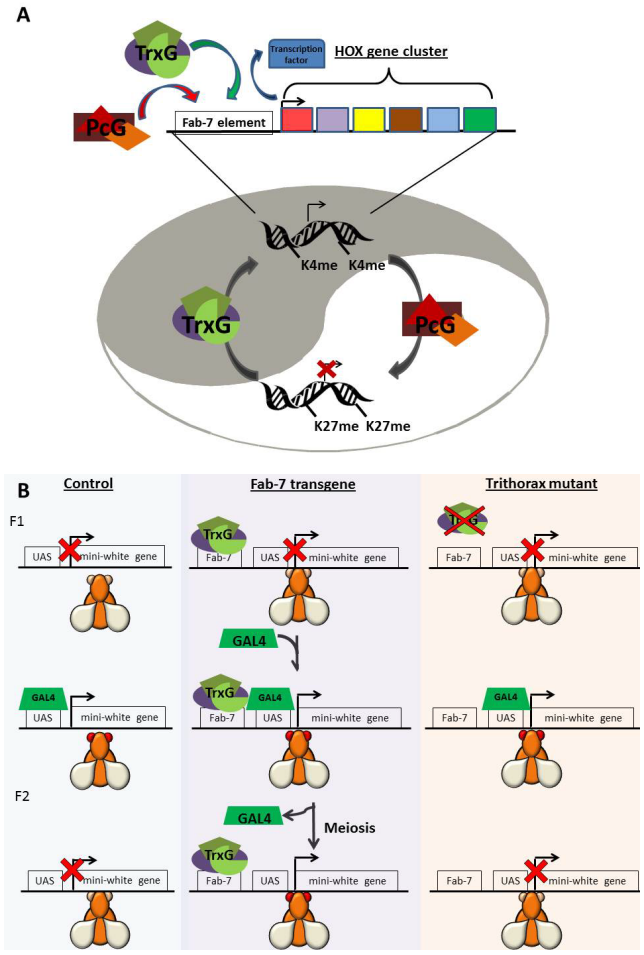


Figure 4: Trithorax and Polycomb group protein complexes maintain epigenetic cellular memory.

(A) Trithorax, the H3K4 methyltransferase, and Polycomb, the H3K27 methyltransferase act antagonistically to maintain active (trxG) or repressed (PcG) chromatin at the Hox cluster in Drosophila. These heritable chromatinstates form a cellular memory of transcription that aids in the specification of proper segments. (B) Cavalli and Paro [19] demonstrated in flies that trithorax binding to the Fab-7 element from the Hox cluster confers epigenetic transcriptional stability. In control flies, a pulse of GAL4 activates the mini-white transgene resulting in red eyes.When the transgene contains the Fab-7 element, this effect is meiotically stable. The meiotic stability is eliminated in trithorax mutants indicating that it is likely dependent upon H3K4me.

*Restoring Totipotency Through Epigenetic Reprogramming. Briefings in Functional Genomics 2013 Mar;12(2):118-28. doi: 10.1093/bfgp/els042. Epub 2012 Oct 31. PMID:23117862

inducing the flanking mini-white gene. When activated by a pulse of GAL4, the expression of the flanking mini-white reporter is mitotically and even meiotically stable, in the absence of GAL4. This stability is dependent upon trithorax, the H3K4 methyltransferase, further suggesting that H3K4me can serve as an epigenetic transcriptional memory (Figure 4B)[114,115].

Epigenetic regulation of pluripotency in embryonic stem cells

The maintenance of pluripotency in ES cells may be regulated by bivalent domains, consisting of H3K4me3 (active) and H3K27me3 (repressive) coexisting at developmentally regulated promoters[116]. At these bivalent promoters, H3K27me3 strongly correlates with the binding of PRC2 and PRC1, encoded by PcG genes, as well as the binding of the pluripotency factors Oct4, Sox2 and Nanog[117,118]. The EZ subunit of PRC2 encodes an H3K27 methyltransferase, while functions as part of the PRC1 complex to bind H3K27me3 and represses transcription[112]. In addition, the repression of many PRC1 target genes requires Oct4[119]. This suggests that the pluripotency factors coordinate H3K27me3 to transcriptionally represses developmental control genes and maintain a pluripotent state[117,120]. At the same time H3K4me3, maintained by the trxG genes, keeps these control regions poised for activation upon differentiation. When ES cells differentiate, the bivalent domains resolve into either H3K4me3 or H3K27me3 domains exclusively[116]. Thus, the bivalent state in ES cells is thought to keep genes poised for rapid conversion to either an 'on' or 'off' state.

Although it is not clear what triggers resolution of bivalent domains, some data provide insight into possible modes of regulation. For example, the H3K4me jumonji-class demethylase KDM5 can interact with the H3K27 methyltransferase complex PRC2 and may function in converting bivalent domains into active domains[121]. Conversely, the H3K27me

demethylase Utx associates with trithorax and may participate in converting bivalent domains into repressed domains[122]. The conversion of these bivalent domains into either repressed or activated domains may be necessary for tissue specification.

1.3 INHERITANCE THROUGH GAMETES: MAINTENANCE AND ERASURE

Although much is known about how pluripotency is maintained in ES cells, how pluripotency is re-established in the embryo from highly differentiated gametes remains largely a mystery. In order to gain insight into this process, we must compare the epigenetic status of the gametes to the reprogrammed pluripotent state of the embryo and ES cells. A recent study in mouse demonstrated that mouse oocytes have a lack of H3K4me3 at methylated CpG islands[120]. However, beyond this relatively little is known about the genome-wide patterns of other histone modifications in oocytes.

In contrast, much more is currently known about the patterns of epigenetic modifications in sperm. In mature vertebrate sperm, much of the genome is repackaged from histones to protamines. Because of this replacement, it was originally thought that vertebrate sperm were lacking epigenetic information contained on histone tails. However, a recent study suggests that this is not the case. In mature human sperm, 4% of the genome remains wrapped around canonical histones, and this chromatin is highly enriched at the promoters of genes that function in embryonic development. Remarkably, many of the promoters that contain canonical histones have both H3K4me2/3 and H3K27me3 and these bivalent promoters significantly overlap with the promoters that are bivalent in ES cells[117]. A highly similar bivalent chromatin state is also found in mature mouse sperm[116]. These findings are especially significant considering that these embryonic development genes with bivalent chromatin are not expressed during spermatogenesis. The presence of bivalent chromatin domains at embryonic development genes in sperm suggests an enticing model where the pluripotent state of the embryo may already be established in sperm and may be faithfully propagated through early development to poise genes for embryonic development (Figure 5). Nevertheless, since the presence of bivalent chromatin has not yet been verified

in early human or mouse embryos, proof of the maintenance and propagation of these domains requires further investigation.

In particular, evidence from zebrafish casts some doubt on the model. Similar to what has been observed in human and mouse, bivalent chromatin domains have also been observed in zebrafish sperm[118]. This suggests that the possibility that the use of bivalent chromatin domains to poise embryonic expression may be a highly conserved mechanism. However, when genome-wide chromatin was assayed in the zebrafish embryo to confirm the propagation of bivalent domains, conflicting results were found. Initially, it was observed that bivalent domains are only acquired following zygotic genome activation (ZGA)[119]. This result suggests that bivalent domains may be re-acquired transcriptionally rather than faithfully maintained from sperm. However, in a subsequent study, bivalent chromatin domains were observed before ZGA[121]. Based on these two results, it is possible that the bivalent chromatin domains are reduced but not fully erased. This could explain why they were not detected in the original study[119]. If this is the case, then the maintenance of reduced bivalent chromatin domains could act as a seed to re-establish larger domains following ZGA. Intriguingly, this type of 'signposting' model is very reminiscent of what occurs at DNA-methylated domains following fertilization in mice.

Maternal effect genes: regulation from one generation to the next

An oocyte has a unique reprogramming ability largely due to maternally stored factors that help kick start embryonic development. The genes that encode these factors are known as maternal effect genes, and they are essential for embryonic survival[123]. They are responsible for regulating the accumulation of RNA and proteins in the oocyte, processing the paternal genome for transcriptional competency, initiating zygotic genome activation

(ZGA), and controlling the progression through pre-implantation development. Defects in maternal effect genes lead to defects in embryonic development that are reflective of aberrations not in the embryo but in the mother. Thus these genes are classified as maternal effect. To date, only a few of these factors have been identified[123].

Mutations in some maternal effect genes cause developmental defects that lead to lethality postnatally when their function is disrupted. For example, mice that are completely null for *Dnmt3a*, the *de novo* DNA methyltransferase, die four weeks after birth[124]. However, females that lack *Dnmt3a* specifically in their oocytes give rise to embryos that die by embryonic day 10.5[125]. These maternal effect embryos have physiological defects that include malformed neural tube and branchial arches, and DNA methylation alterations at a number of imprinted loci[125]. Post-fertilization, histone variant H3.3 is maternally loaded into the zygote and loaded onto the paternal pronucleus right before zygotic genome activation occurs[126]. Some maternal effect genes suggest a link between chromatin organization, histone modifications, and the maternal control over the progression of development. For example, the H3K4 methyltransferase, MLL2, is maternally deposited and responsible for the majority of H3K4me3 that is present at promoters in the early embryo[28]. Maternal loss of MLL2 results in embryonic lethality by the eight-cell stage and inefficient activation of the zygotic genome[28]. Another histone modifier, enhancer of zeste 2 (EZH2), is a polycomb group protein that regulates H3K27 methylation and associates with the oocyte-specific splice variant of DNMT1 (DNMT1o) and DNMT3A during early development[127]. Oocytes that lack EZH2 give rise to embryos that exhibit severe growth retardation. In addition, these embryos completely lose H3K27 and H3K9 methylation suggesting a large-scale loss of gene repression[127]. Maternal effect genes can also influence the three-dimensional organization and accessibility of chromatin. Maternal loss of *Brg1*, a catalytic subunit of

SWI/SNF chromatin remodeling complexes, results in embryonic lethality by the blastocyst stage with the majority of embryos dying at the two- and four-cell stage[128]. Furthermore, BRG1 is linked not only to the initiation of ZGA but also H3K4 methylation[128]. CTCF, an insulator protein, is maternally required to allow proper mitosis in developing embryos[129]. LSD1 is also highly expressed maternally, hinting at its potential role as a maternal effect gene. Taken together, these data begin to paint a picture where epigenetic modifications and regulators are required to prime the zygotic genome in order to allow proper development to occur.

The maternal to zygotic transition: priming development

Many maternal effect genes regulate the transition from maternal gene expression to zygotic gene expression. This process, known as the maternal to zygotic transition, is characterized first by the removal of maternally derived transcripts and protein initiated by maternal products like Argonaute protein 2 (Ago2) and ZFP36L2, a zinc finger protein that destabilizes RNAs[130,131]. In the zygote, maternal proteins are eventually degraded by ubiquitin-proteasome pathways and autophagosomes. MZT is completed once zygotic transcription begins to generating factors that create a positive feedback loop leading to accelerated degradation of maternal transcripts. In mice, a large portion of these transcripts are degraded by the two-cell stage. Maternal depletion of Atg5, a member of the autophagasome, results in developmental arrest at the four- and eight-cell stage[132].

Ultimately, MZT results in zygotic genome activation (ZGA) a process that leads to activation of approximately 15% of the mouse genome[133,134]. In mice, ZGA occurs in a minor wave that is followed by a major wave. In general, the transcriptional differences between mature oocytes and early one-cell embryos are minimal compared to those between

one-cell embryos and two-cell embryos[133,134]. The ZGA minor wave accounts for the minimal difference in transcriptional readout. In fact, inhibition of the first round of DNA replication post-fertilization does not inhibit the initiation of the first wave of ZGA, indicating that maternal factors drive the initiation of zygotic transcription[135]. The first, minor wave of ZGA occurs at the one-cell stage with the male pronucleus four to five times more transcriptionally active than the female pronucleus[136]. Some researchers consider the minor wave of ZGA to be promiscuous, as there is a low level of genome wide transcription and these transcripts remain unprocessed. In this wave, many retroviral elements are also transcribed. But these transcript are unable to give rise to functional proteins[137,138]. It has been suggested that an open chromatin structure at the one-cell stage creates a relaxed transcriptional environment that allows for a global low level of transcription. Subsequently, between the one- and two-cell stage, chromatin is extensively remodeled during the second wave of ZGA. It is here where major reprogramming of the transcriptome occurs, a step that is set up by the minor wave of ZGA[133,136]. Inhibition of HDACs also affects the initiation of zygotic transcription highlighting the need for precise chromatin organization in order for development to occur[133].

Inheritance and reprogramming of DNA methylation

Maternal reprogramming is defined by changes in both histone and DNA modifications. DNA methylation is found at CpG residues in mammals and can be stably maintained by Dnmt1, which selectively methylates hemi-methylated DNA during DNA replication[122]. Most of the CpG methylation in the mammalian genome is present at repeated sequences, such as retrotransposons and their remnants, where it is thought to play a role in stably repressing these sequences[139]. However, some CpG methylation is also found at

developmentally regulated genes. For example, imprinted genes, which are maternally or paternally expressed dependent upon parent-of-origin, are often associated with CpG islands, termed imprinted control regions (ICRs), which are either maternally or paternally methylated (Figure 5)[140].

In oocytes, CpG methylation occurs mainly in CpG island-containing promoters while the most of the genome remains hypomethylated, a pattern that closely resembles the methylation state of the pre-implantation embryo[120,141]. This suggests that oocyte methylation could have a function in the early embryo. The establishment and maintenance of this DNA methylation is dependent on the activity of Dnmt3a and Dnmt1[142,143]. Interestingly, both of these enzymes have CpG island-containing promoters that are hypermethylated in the oocyte, so it possible that DNA methylation is reinforced by a positive feedback loop. Conversely, in mammalian sperm, it has been shown that the methylation distribution in the genome closely resembles the pattern seen in ES cells, with many developmentally regulated genes hypomethylated compared to somatic tissues. These hypomethylated genes correlate strongly with genes bound by the pluripotency factors Oct4, Sox2 and Nanog. However, the pluripotency genes Oct4 and Nanog are themselves hypermethylated in sperm, but become unmethylated and expressed in ES cells[117,144,145]. This suggests that these master pluripotency genes may need to be reprogrammed in the embryo to restore pluripotency.

The reprogramming of DNA methylation at *Oct4* and *Nanog* is thought to occur at fertilization during the genome-wide wave of DNA demethylation that occurs actively in the paternal genome and passively in the maternal genome[146,147]. The mammalian genome contains three TET family proteins that are thought to function in DNA demethylation through a hydroxymethylation intermediate (Figure 5)[19]. The maternal loss of *Tet3* results in the failure to DNA demethylate the paternal genome along with a corresponding increase in

the incidence of embryonic lethality and a decrease in the fertility of mutant females[148]. This suggests that the active demethylation of the paternal genome is catalyzed by maternal Tet3. Furthermore, in Tet3 mutants, Oct4 and Nanog fail to be demethylated and an Oct4-EGFP transgene fails to be expressed from a paternally inherited transgene[148]. These results suggest that Oct4 and Nanog are targets of the Tet3 active paternal demethylation and that this demethylation likely plays a critical role in restoring pluripotency in the embryo after fertilization. Consistent with this model, there is a decrease in the efficiency of reprogramming of DNA methylation and a decrease in Oct4 expression when SCNT is performed using Tet3 mutant oocytes[148]. There is also an increase in the efficiency of iPS when DNA methyltransferase activity is inhibited[149].

Remarkably, ICRs, which must remain methylated to maintain parent-of-origin expression, resist demethylation during the genome-wide DNA demethylation that occurs after fertilization in mammalian embryos. During this demethylation, ICRs do lose some DNA methylation. However, critical CpG residues remain methylated and this methylation is thought to seed the re-acquisition of larger methylated domains at ICRs following demethylation[140]. This mechanism is very reminiscent of what may be occurring at bivalent chromatin domains in the zebrafish embryo. Some retrotransposon sequences such as intra cisternal A-type particle (IAP) elements, which also must remain repressed, are similarly resistant to this DNA demethylation (Figure 5)[141,150].

The maintenance of DNA methylation at critical ICR CpG residues is dependent upon Dnmt1, expressed specifically from a maternal promoter termed Dnmt1o. In the absence of Dnmt1o, embryos die late in gestation due to the failure to maintain allele-specific DNA methylation at ICRs[151]. In addition, the resistance of ICRs to demethylation is dependent upon the protein Stella/PGC7 (Figure 5). PGC7 binds specifically to H3K9me2-

containing chromatin in the mouse maternal pronucleus[152]. In the absence of maternal provided PGC7, the imprinted genes Peg1, Peg5, Peg10, H19 and Rasgrf1 become inappropriately demethylated[153]. Furthermore, loss of the maternal pool TRIM28 leads to variable embryonic lethality. In addition, aberrant demethylation was observed at the paternal H19 locus in addition to other loci, although at a much lower frequency. This suggests that TRIM28 is required to protect some imprints from demethylation at fertilization. ZFP57, a zinc-finger protein, is also involved in the maintenance of DNA methylation at both maternal and paternal imprinted loci post-fertilization[154,155]. Thus, it seems clear that mammals have evolved a complex set of regulatory mechanisms that allow DNA methylation at certain loci to be inherited through fertilization to the next generation while DNA methylation at other loci is reprogrammed.

Inheritance and reprogramming of histone methylation

In addition to demethylating key pluripotency genes to reactivate the embryonic program, it is likely that other epigenetic information must be removed at fertilization to prevent the germline program from being inappropriately propagated in the embryo. For example, along with bivalent domains which may contain critical information for embryogenesis, mature sperm also have H3K4me2/3 in genes that functioned previously in the germline during spermatogenesis[117]. The acquisition of this histone information in sperm is easier to reconcile than the acquisition of bivalent chromatin, because these genes presumably acquired H3K4me2/3 in their chromatin during transcription. Like bivalent domains, these histone modification domains could potentially be stably inherited in the embryo and this 'epigenetic baggage' could potentially result in the inappropriate expression of spermatogenesis genes in the embryo. However, recent data suggest that mechanisms may

exist to prevent this. As loss of function mutations in LSD1 ortholog in both *Drosophila* and *Caenorhabditis elegans* result in sterility, and in *C. elegans* this sterility is correlated with increasing H3K4me2 and the increasing expression of spermatogenesis genes[46,84–86]. These results suggest that the role of LSD1 in the germline is to demethylate H3K4me2 at spermatogenesis genes and prevent this epigenetic baggage from heritably persisting in the embryo of the next generation. This finding fits nicely with the findings from human and mouse sperm, where H3K4me2 is observed in spermatogenesis genes[116,117]. In flies, worms and mice, LSD1 is maternally deposited in the oocyte[46,84,156]. Thus, it is intriguing to propose that the H3K4me2 in mature vertebrate sperm at spermatogenesis genes may not be transmitted to the embryo due to the maternal demethylation activity of LSD1 acting at or around fertilization (Figure 5).

Although LSD1 appears to function in reprogramming the germline program at fertilization, there is also evidence in *C. elegans* that heritable epigenetic mechanisms exist which may facilitate re-initiation of the germline program in the subsequent generation. The co-transcriptional acquisition of H3K36 methylation by Set-2 in yeast is proposed to block spurious intragenic transcription[157]. However, the function of H3K36 methylation in metazoans is not well understood. In *C. elegans*, H3K36 methylation is carried out by a Set2-related protein, methyltransferase 1 (MET-1) and a second methyltransferase, maternal effect sterile (MES)-4[158,159]. Genome-wide MES-4 mapping in the early embryo revealed that MES-4 associates with germline-specific genes that were previously expressed in the maternal germline and maintains H3K36 methylation at these loci in the absence of transcription[160,161]. Upon specification of the germline precursors Z2 and Z3, MES-4 transitions from global maintenance to exclusive association with Z2/Z3[160,161]. It has been proposed that this H3K36 methylation maintenance by MES-4 serves as an epigenetic memory of

transcriptional patterns by marking genes that were previously transcribed in the maternal germline for reactivation in the embryonic germline[160,161]. Consistent with this hypothesis, maternal loss of MES-4 results in the loss of the embryonic primordial germ cells Z2/Z3 and complete sterility (Figure 5)[162].

Taken together, the data on histone methylation and DNA methylation provide an emerging picture. They suggest that there is a highly specific epigenetic reprogramming event that takes place at fertilization which helps remove epigenetic baggage, while specifically allowing important epigenetic information to be transmitted through to the embryo.

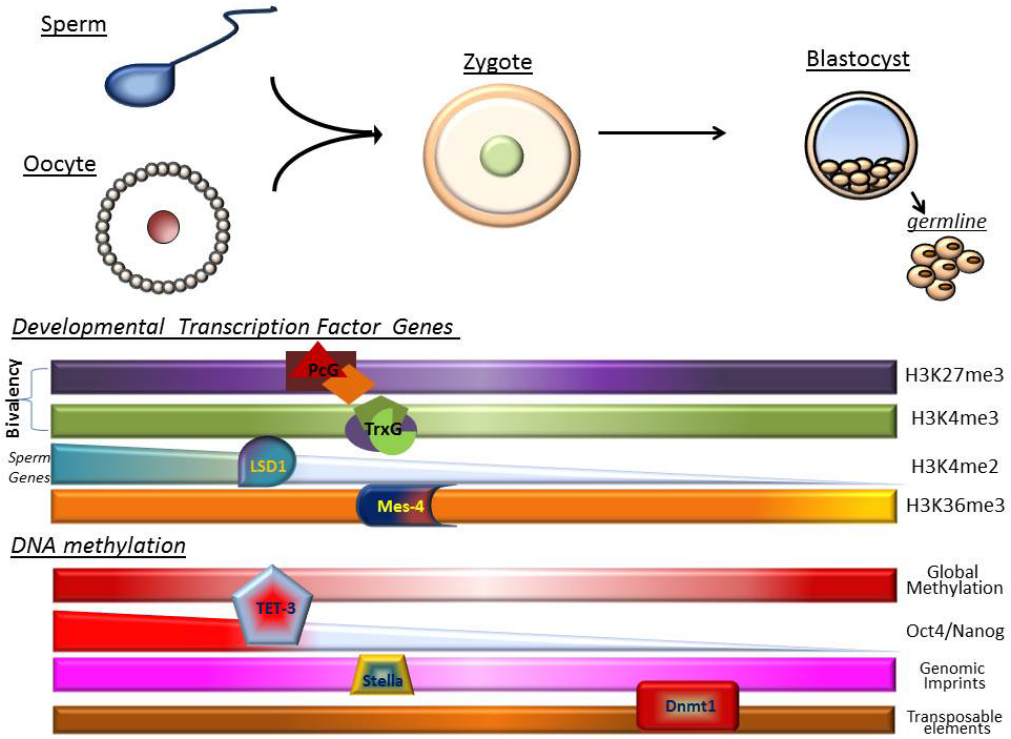


Figure 5: Summary of reprogramming events that occur at fertilization. At fertilization, the highly differentiated gametes undergo a dramatic change in cell fate to form the totipotent zygote. During this process, certain epigenetic information is stably propagated while other epigenetic information is reprogrammed. Bivalent domains (H3K4me3 and H3K27me3) acquired through trithorax (MLL) and Polycomb at developmental transcription factor loci are stably propagated from sperm to the embryo in multiple organisms. H3K4me2 at spermatogenesis genes is erased at fertilization by LSD1 in C. elegans. H3K36me3 at germline genes is stably maintained by MES-4 in the C. elegans embryo.Global DNA methylation is erased byTet3 after fertilization and then returns later in embryogenesis in mice. During this erasure, DNA methylation at *Oct4* and *Nanog* is erased but is maintained at critical CpG residues at imprinted loci due to the protein Stella. Certain retrotransposons, such as IAP elements, are resistant to the global demethylation at fertilization in mice.

*Restoring Totipotency Through Epigenetic Reprogramming. Briefings in Functional Genomics 2013 Mar;12(2):118-28. doi: 10.1093/bfgp/els042. Epub 2012 Oct 31. PMID:23117862

1.4 TRANSGENERATIONAL INHERITANCE: EVIDENCE FOR ERRORS IN REPROGRAMMING?

Based on the complex series of epigenetic reprogramming events that likely occur to restore totipotency following fertilization, it seems probable that there may be instances where epigenetic reprogramming is incomplete, allowing epigenetic information to be inappropriately transmitted through fertilization from the gametes to the embryo of the next generation. These instances could give rise to epialleles with medical implications. Recently, a number of such instances have been documented in model organisms[163]. For example, when a mouse female pronucleus is transplanted into a recipient egg from a different genetic background, the resulting nucleocytoplasmic hybrids have inappropriate transcription and corresponding DNA methylation defects in certain tissues. Remarkably, >50% of the time, these defects are meiotically heritable through sperm[164]. In male rats, exposure to endocrine disruptors during the period of gonadal sex determination causes increased male infertility with corresponding changes in DNA methylation. This reduction in male fertility is meiotically heritable for at least four generations[165]. In viable yellow mice, transcription originating from an IAP retrotransposon results in ectopic expression of the agouti locus and mice with a yellow coat color. This coat color can be preferentially inherited through the mother and correlates with the DNA methylation status of the IAP retrotransposon, suggesting that the preferential inheritance is due to incomplete reprogramming of the DNA methylation in the IAP element. Interestingly, this preferential inheritance of coat color does not occur paternally[166].

1.5 A MIND AT WAR: THE FUNCTIONAL CONSEQUENCES OF ABERRANT REPROGRAMMING:

Loss of maternal factors often leads to lethality, indicating the essential nature of that large scale reprogramming events. Slight defects in epigenetic reprogramming at fertilization could lead to phenotypic aberrations that manifest later in life through the maintenance of aberrant DNA methylation or histone modifications. Evidence of this comes from data that demonstrates DNA methylation aberrations arise in progeny from females lacking *Dnmt3a/3L*, *Trim28*, and *Zfp57*[125,143,167,168]. Furthermore, it has been suggested that aberrations in DNA methylation can be propagated in the few animals that survive lethality. DNA methylation has been shown to be important for proper neuronal functioning, which opens the possibility that defects in epigenetic reprogramming at fertilization can lead to neuronal defects later in development[155].

Evidence that DNA methylation is important for neuronal function comes from the combined deletion of Tet1 and Tet2, methylcytosine dioxygenases[21]. Surprisingly, this double mutant combination does not result in embryonic lethality but the majority of animals do exhibit perinatal lethality. Of the rare double knockout animals that survive to adulthood, there is a significant reduction in the fertility of these animals. Progeny from double knockout females were observed to have aberrations in DNA methylation at a few imprinted loci[21]. This suggests that maternal Tet1 and Tet2 are important both for the establishment of imprints and for the maintenance of DNA methylation in the next generation. Furthermore, hypermethylation was observed in the brains of double knockout adult animals suggesting that the defect in DNA methylation may be stably propagated from early development[21].

Conversely, a protein whose activity depends on proper DNA methylation patterns is MECP2[24]. Mutations in *Mecp2* can lead to Rett syndrome[171]. Recently, it has been

appreciated that MECP2 binds to key 5mC residues in order to prevent improper expression of key neuronal genes during neurogenesis[25,172–174]. Alterations in MECP2 function during neurogenesis leads to the symptoms associated with Rett syndrome including intellectual disability and reduced social interaction[171,173,174]. Interestingly, many of the phenotypes associated with Rett syndrome are also those associated with autism and imprinting disorders such as Angelman syndrome[25]. Taken together, these data suggests that aberrant DNA methylation, which could result from aberrant reprogramming, can have phenotypic outcomes that manifest later in development through altering proteins whose function depends on precise DNA methylation patterns.

1.6 OUTSTANDING QUESTIONS AND OBJECTIVES

In recent years, researchers have come to appreciate the extensive role that epigenetic information plays in tissue differentiation. Based on these data, it has become increasingly clear that a complex set of regulatory mechanisms likely exist to regulate what epigenetic information is stably propagated from the gametes through fertilization to the embryo. However, at the start of my work, the role of histone demethylation reprograming at fertilization, and how it relates to DNA methylation maintenance and/or erasure remained to be determined. In addition, it was unclear whether defects in reprogramming can lead to phenotypic alterations past embryonic development. Thus, the main objectives of this dissertation were to (1) to demonstrate that *Lsd1* acts as a maternal effect gene and is required for epigenetic reprogramming at fertilization, (2) to demonstrate heritable alterations in the epigenetic landscape that results from altered reprogramming, and (3) to reveal phenotypic outcomes that can result from aberrant LSD1 reprogramming.

CHAPTER 2:

MATERNALLY PROVIDED LSD1 ENABLES THE MATERNAL-TO-ZYGOTIC TRANSITION AND PREVENTS DEFECTS THAT MANIFEST POSTNATALLY

Jadiel A. Wasson, Ashley K. Simon, Dexter A. Myrick, Gernot Wolf, Shawn Driscoll, Samuel L. Pfaff, Todd S. Macfarlan and David J. Katz

This Chapter was published in eLife in January 2016.

Citation: eLife 2016;10.7554/eLife.08848. PMID:26814574

Somatic cell nuclear transfer has established that the oocyte contains maternal factors with epigenetic reprogramming capacity. Yet the identity and function of these maternal factors during the gamete to embryo transition remains poorly understood. In *C. elegans*, LSD1/KDM1a enables this transition by removing H3K4me2 and preventing the transgenerational inheritance of transcription patterns. Here we show that loss of maternal LSD1 in mice results in embryonic arrest at the 1-2 cell stage, with arrested embryos failing to undergo the maternal-to-zygotic transition. This suggests that LSD1 maternal reprogramming is conserved. Moreover, partial loss of maternal LSD1 results in striking phenotypes weeks after fertilization; including perinatal lethality and abnormal behavior in surviving adults. These maternal effect hypomorphic phenotypes are associated with alterations in DNA methylation and expression at imprinted genes. These results establish a mammalian disease paradigm where defects in early epigenetic reprogramming can lead to defects that manifest later in development.

2.1 Introduction

At fertilization, the epigenome of the developing zygote undergoes widespread changes in DNA methylation and histone methylation[119,121,128,141,175–178]. This reprogramming is driven by the deposition of maternal proteins and RNA into the zygote[179]. After fertilization, zygotic genes become transcriptionally active, while the maternal transcriptional program is silenced[179]. In mice, this maternal-to-zygotic transition (MZT) occurs between the one- and two-cell (1-2C) stage[175,180,181].

Somatic cell nuclear transfer (SCNT) experiments in *Xenopus* demonstrated that the oocyte has the capacity to reprogram a differentiated somatic nucleus into a cloned embryo[101]. This epigenetic reprogramming capacity of the oocyte is also conserved in mammals[100]. It is thought that the epigenetic reprogramming capacity of the oocyte enables the MZT. Nevertheless, the enzymes involved in maternal epigenetic reprogramming, and the consequences of failure to reprogram, largely remain a mystery.

Recently, work in *C. elegans* implicated LSD1/SPR-5/KDM1a (hereafter referred to as LSD1) in maternal reprogramming at fertilization. Di-methylation of lysine 4 on histone H3 (H3K4me2) is acquired in gamete genes during the specification and maintenance of the germline. The H3K4me2 demethylase LSD1 is maternally deposited into the zygote at fertilization and required to prevent H3K4me2 from being inherited transgenerationally. Without LSD1, H3K4me2 accumulates at gamete genes across generations and results in the inappropriate expression of these genes. This correlates with increasing sterility in the population over time[84].

The requirement to reprogram H3K4 methylation is also suggested to be a critical step in SCNT. During SCNT in *Xenopus*, the cloned embryos often inappropriately express

genes from the tissue where the somatic nucleus was derived[106]. This transcriptional memory can be eliminated by overexpressing histone H3 with a mutated K4 residue[71]. This work implies that H3K4 methylation may be the carrier of transgenerational transcriptional memory, and that complete reprogramming may require the removal of this modification.

Genomic imprinting is the monoallelic expression of a small number of genes based on their parent of origin[182]. Imprinted genes are dependent upon DNA methylation at small imprinting control regions (ICRs) associated with these loci[182]. At ICRs, CpG methylation is established in the gametes and maintained throughout the development of the offspring[140]. In mammals, there are two amine-oxidase histone demethylases, LSD2/KDM1B (hereafter referred to as LSD2) and LSD1[35,40]. LSD2 is expressed mainly in the oocyte and required for the establishment of maternal imprints at a subset of imprinted loci[40]. Without LSD2, embryos derived from these oocytes exhibit a maternal effect embryonic lethality phenotype prior to mid-gestation[40]. This demonstrates that, similar to *C. elegans*, amine oxidase-type histone demethylases can function maternally in mammals.

LSD1 is a component of several protein complexes. As part of the CoREST complex, it specifically demethylates H3K4me1/2, but not H3K4me3[35,41]. Alternatively, when associated with the Androgen Receptor complex LSD1 has been shown to demethylate H3K9me2 *in vitro*[49]. LSD1 is an essential gene in mammalian development, as homozygous mutants fail to develop properly after implantation and die prior to embryonic day 8 (e8)[45,183,184]. To determine whether LSD1 may also be involved in maternal reprogramming, we conditionally deleted *Lsd1* in mouse oocytes with three different maternal *Cre* transgenes. Deletion of *Lsd1* maternally with either *Zp3-Cre* or *Gdf9-Cre* results in embryonic lethality primarily at the 1-2C stage and these embryos fail to undergo the

MZT. This suggests that LSD1 plays a conserved role in maternal reprogramming. Surprisingly, deletion of *Lsd1* maternally with *Vasa-Cre* results in an incomplete loss of LSD1 in oocytes. This uncovers a hypomorphic effect in which some surviving animals exhibit long-range developmental defects, including perinatal lethality and behavioral abnormalities. These defects are associated with disruption of the epigenetic landscape, including aberrant DNA methylation and expression at imprinted loci. These results demonstrate that defects in epigenetic reprogramming between generations can lead to abnormalities later in development.

2.2 RESULTS

LSD1 is expressed throughout oocyte development

Based on the previously demonstrated maternal role of LSD1 in *C. elegans* and LSD2 in mice, we first asked if LSD1 is expressed in mouse oocytes[40,84]. RNA-seq datasets from ovulated oocytes and 2C-stage embryos suggested abundant LSD1 transcripts in these cells[44]. Immunofluorescence (IF) and immunohistochemistry (IHC) with an antibody raised against LSD1 confirms that LSD1 is expressed in the oocyte nucleus and in the surrounding follicle cells throughout oocyte development (Figure 1A, Figure 1-figure supplement 1A-L). Therefore, to determine if LSD1 functions in oocytes, we conditionally deleted *Lsd1* by crossing *Kdm1a/Lsd1*$^{fl/fl}$ mice[184] to three different maternal *Cre* transgenic lines, *Ddx4/Vasa*^{Cre} [185], *Gdf9*^{Cre} [186] , and *Zp3*^{Cre} [187] (oocytes and embryos from *Kdm1a/Lsd1*$^{fl/\lrcorner}$*::Ddx4/Vasa*^{Cre}*, Kdm1a/Lsd1*$^{fl/\lrcorner}$*::Gdf9*^{Cre}*,* and *Kdm1a/Lsd1*$^{fl/\lrcorner}$*::Zp3*^{Cre} mice will be referred to hereafter as *Lsd1*^{Vasa}*, Lsd1*^{Gdf9} and *Lsd1*^{Zp3} respectively). *Ddx4/Vasa*^{Cre} is expressed in the germline beginning at e18 and induces full deletion by birth (Figure 1B)[185]. *Gdf9*^{Cre} is expressed in oocytes beginning at postnatal day 3 (P3), including in primordial follicles[186], while *Zp3*^{Cre} is also expressed in oocytes, but beginning at P5 in primary follicles (Figure 1B)[187]. LSD1 IHC and IF demonstrate that deletion of *Lsd1* with either *Gdf9*^{Cre} or *Zp3*^{Cre} results in the complete loss of LSD1 from the oocyte nucleus (Figure 1C-F, I). Crossing the *Lsd1*$^{fl/fl}$ allele to *Ddx4/Vasa*^{Cre} also results in complete maternal deletion, as can be determined by the 100% segregation of the maternally deleted allele to the offspring (all of offspring from *Lsd1*^{Vasa} mothers crossed to wild-type fathers result in heterozygous -/+ offspring) (Figure 1-figure supplement 2). This demonstrates that, similar to deletion via *Gdf9*^{Cre} or *Zp3*^{Cre}, crossing the *Lsd1*$^{fl/fl}$ allele to *Ddx4/Vasa*^{Cre} also effectively deletes *Lsd1* maternally. However, for reasons

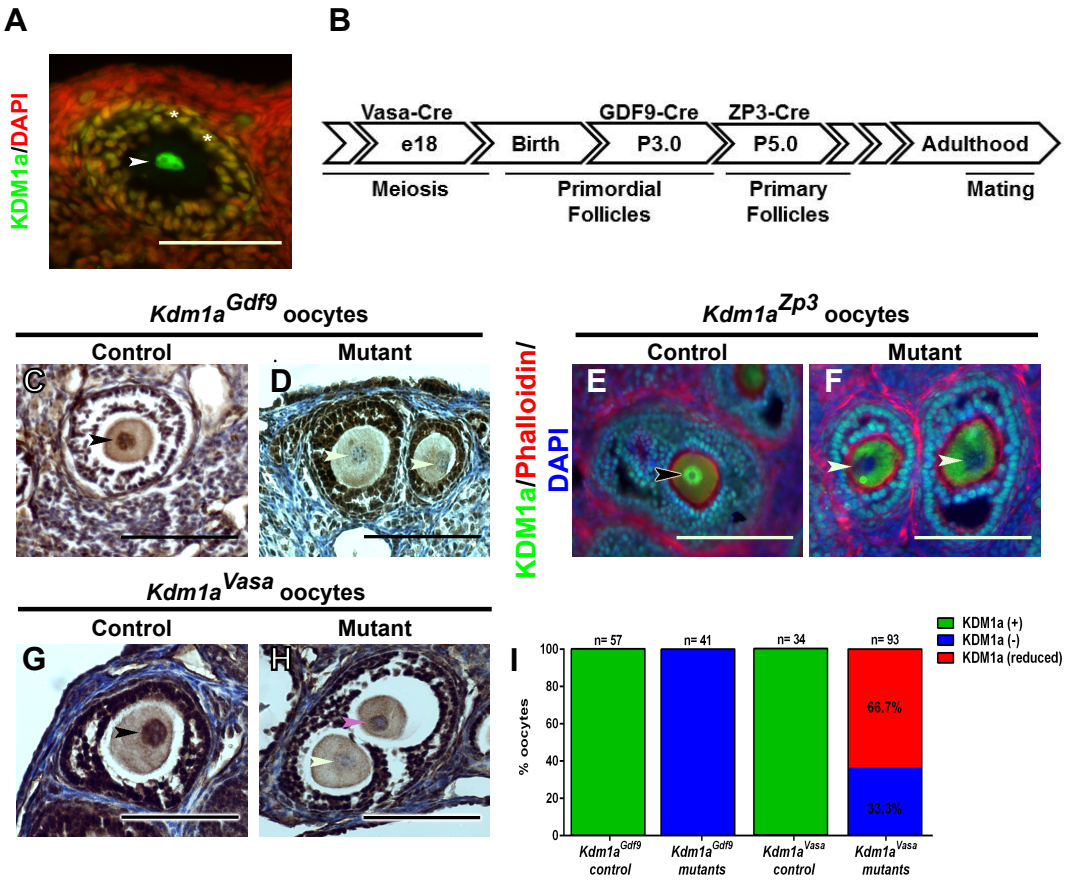


Figure 1.

Maternal Expression and Conditional Deletion of *Lsd1* in Mouse Oocytes

(A) Wild Type mouse oocyte nucleus (white arrowhead) and surrounding follicle cells (white asterisks) stained with anti-LSD1 (green) antibody and DAPI (red). (B) Developmental timeline of maternal Cre expression (*Vasa-Cre*, *Gdf9-Cre* and *Zp3-Cre* transgenes) and corresponding oogenesis stages. (C,D) Immunohistochemistry (IHC) with anti-LSD1 (brown) antibody and hematoxylin (blue) showing LSD1 nuclear expression (black arrowhead) and absence of expression (white arrowheads) in $Lsd1^{Gdf9}$ control (C) and mutant (D) oocytes. (E,F) Immunofluorescence (IF) with anti-LSD1 (green) antibody, phalloidin (red) and DAPI (blue) showing LSD1 nuclear expression (black arrowhead) and absence of expression (white arrowheads) in $Lsd1^{Zp3}$ control (E) and mutant (F) oocytes. (G,H) IHC with anti-LSD1(brown) antibody and hematoxylin (blue) showing LSD1 nuclear expression (black arrowhead), absence of expression (white arrowhead) and reduced expression (pink arrowhead) in $Lsd1^{Vasa}$ control (G) and mutant (H) oocytes. (I) Percentage of oocytes with LSD1 (green), reduced LSD1 (red) or no LSD1 (blue) staining in $Lsd1^{Gdf9}$ and $Lsd1^{Vasa}$ heterozygous control versus mutant oocytes. Scale bars represent 50μm. n=number of oocytes analyzed with percentages indicated for each category.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

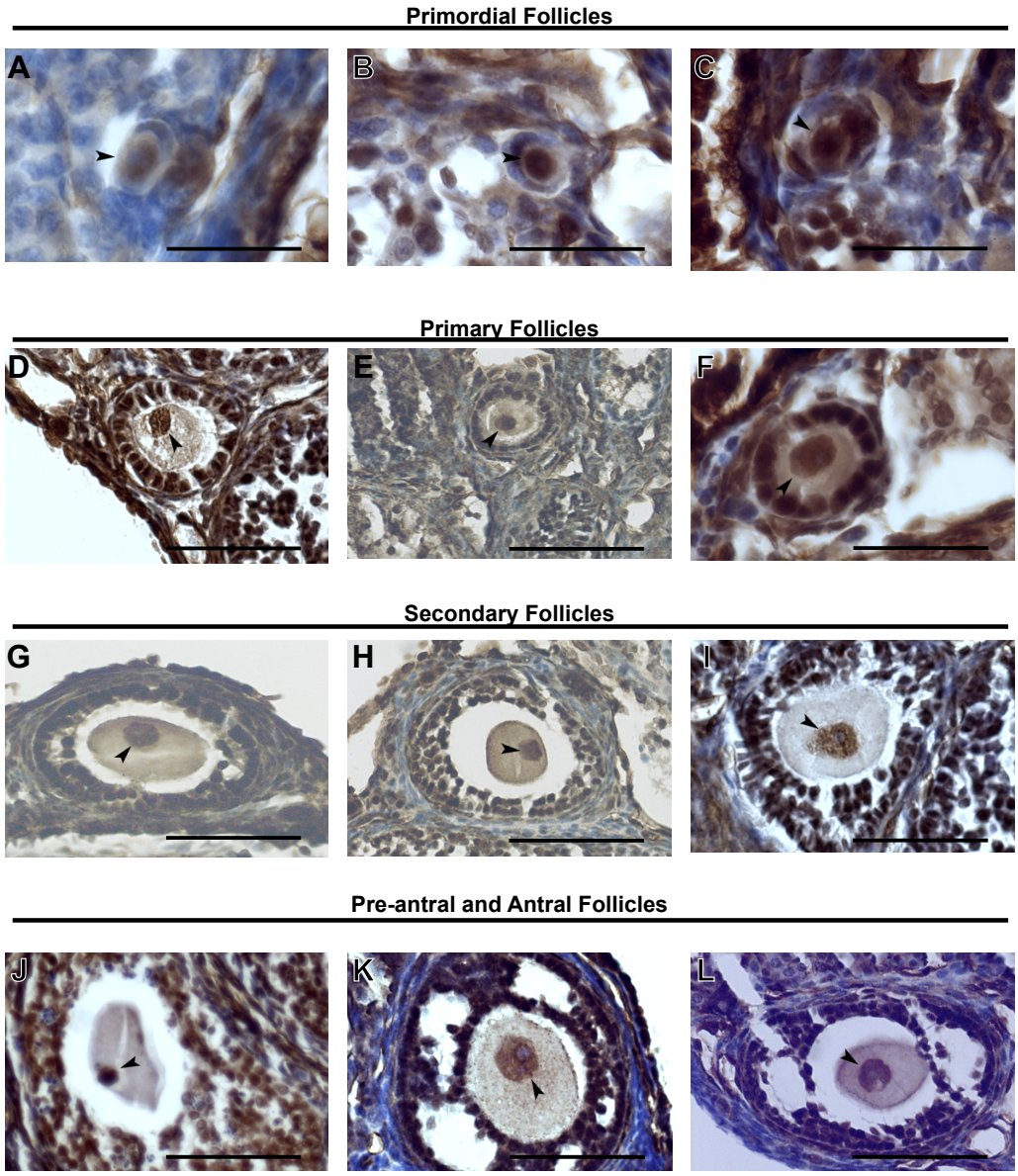


Figure 1-figure supplement 1: LSD1 Expression in Staged Oocytes
(A-L) Immunohistochemistry (IHC) of primordial follicles (A,B,C), primary follicles (D,E,F), secondary follicles (G,H,I) and pre-antral and antral follicles (J,K,L) stained with anti-LSD1(brown) antibody and hematoxylin (blue). The oocyte nucleus is indicated with black arrowheads. Scale bars represent 50μm.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

that are not clear, deletion of *Lsd1* with *Ddx4/Vasa*^{Cre} results in a hypomorphic effect, where 33.3% of oocytes completely lack LSD1 protein, but 66.7% retain a low level of LSD1 (Figure 1G-I). This incomplete effect is surprising because the *Ddx4/Vasa*^{Cre} transgene is reported to be expressed earlier in the germline than either *Gdf9*^{Cre} or *Zp3*^{Cre} (Figure 1B)[185,186]. It is possible that the low level of LSD1 remaining in some oocytes is due to delayed deletion of *Lsd1*, though the reason for this potential delay is unknown.

Loss of maternal LSD1 results in 1-2 cell embryonic arrest

To determine if there is a functional requirement for maternal LSD1 in mice, we crossed *Lsd1*^{Vasa}, *Lsd1*^{Gdf9} and *Lsd1*^{Zp3} females to wild-type males to generate heterozygous offspring (Figure 1-figure supplement 2). In mice, zygotic transcription begins in the 1C embryo just prior to the first cleavage to the 2C stage[175,180,181]. The heterozygous offspring from the maternally deleted mothers have a normal *Lsd1* gene on the paternal allele. Thus, crossing maternally deleted mothers to wild-type fathers enables us to isolate the maternal function of LSD1 (Maternal-, Zygotic+, hereafter referred to as M-Z+). M-Z+ heterozygous embryos derived from *Lsd1*^{Gdf9} mutant mothers are hereafter referred to as *Lsd1*^{Gdf9} M-Z+ embryos, while M+Z+ heterozygous embryos derived from littermate control mothers that are *Cre* minus are hereafter referred to as *Lsd1*^{Gdf9} M+Z+ embryos. *Lsd1*^{Gdf9} M-Z+ embryos exhibit embryonic arrest at the 1-2C stage (Figure 2-figure supplement 1A-I). Specifically, in control *Lsd1*^{Gdf9} M+Z+ embryos at embryonic day 1.5 (e1.5), we observe 7% fragmented/degraded embryos, 65% 1-cell embryos and 28% 2-cell embryos (n=135, Figure 2-figure supplement 1I). In contrast, in *Lsd1*^{Gdf9} M-Z+ embryos at e1.5 we observe 40% fragmented/degraded embryos, 59% unfertilized oocytes or 1-cell embryos, and only 1% 2C embryos (n=134, Figure 2-figure supplement 1I). The vast majority of the non-degraded *Lsd1*^{Gdf9} M-Z+

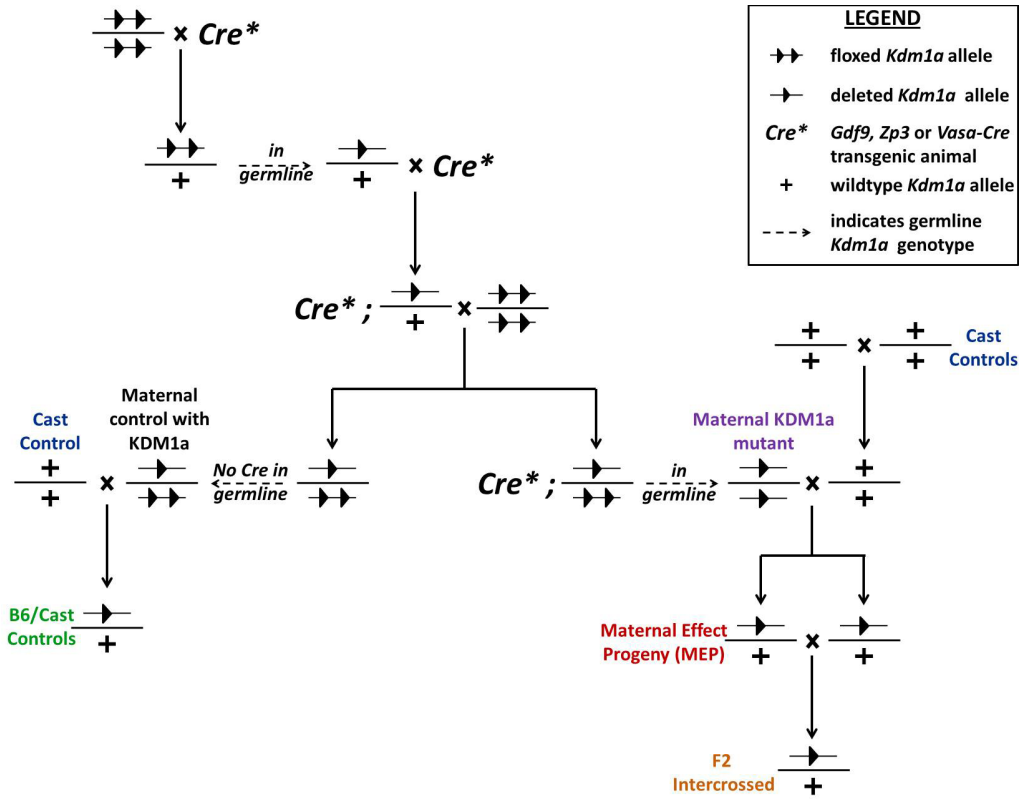


Figure 1-figure supplement 2:
Generation of *Lsd1* mutant and control animals

Lsd1 animals were generated by crossing multiple generations of *Lsd1*$^{fl/fl}$ animals with either *Gdf9*-, *Zp3*-, or *Vasa-Cre* transgenic animals. Blue indicates *Mus castaneus* control animals. Purple indicates *Lsd1* mutant females. Green indicates B6/Cast hybrid control progeny. Red indicates *Lsd1* maternal effect progeny (MEP). Orange indicates progeny resulting from intercrossing 2 MEP adult animals. All labelled progeny were used in crosses and assays presented in subsequent figures (color-coding matches animals used and graphed in each figure).

*eLife 2016;10.7554/eLife.08848. PMID:26814574

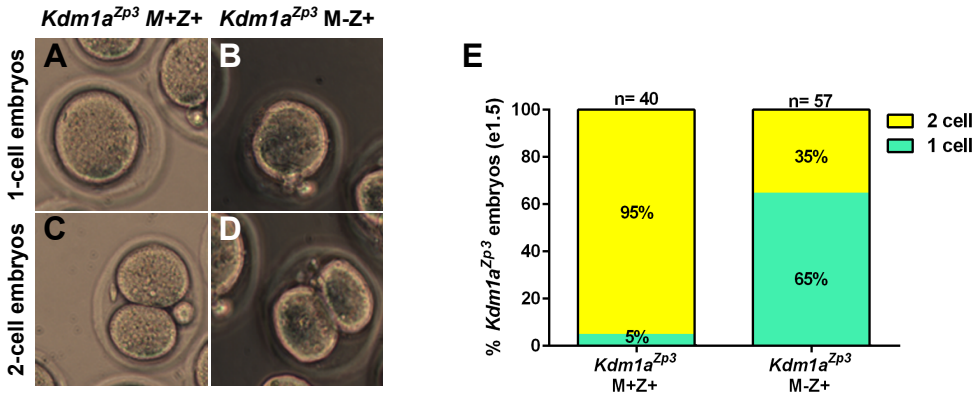


Figure 2.

***Lsd1*^{Zp3} Embryos Arrest at the 1-2 cell Stage**

(A,B,C,D) Brightfield images of (A,C) M+Z+ and (B,D) M-Z+ 1- and 2-cell embryos derived from *Lsd1*^{Zp3} control and mutant mothers at e1.5. (E) Percentage of 1-cell (green) and 2-cell (yellow) embryos derived from *Lsd1*^{Zp3} control and mutant mothers at e1.5. n = 40 for *Lsd1*^{Zp3} M+Z+ embryos from 3 litters. n = 57 for *Lsd1*^{Zp3} M-Z+ embryos from 6 litters.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

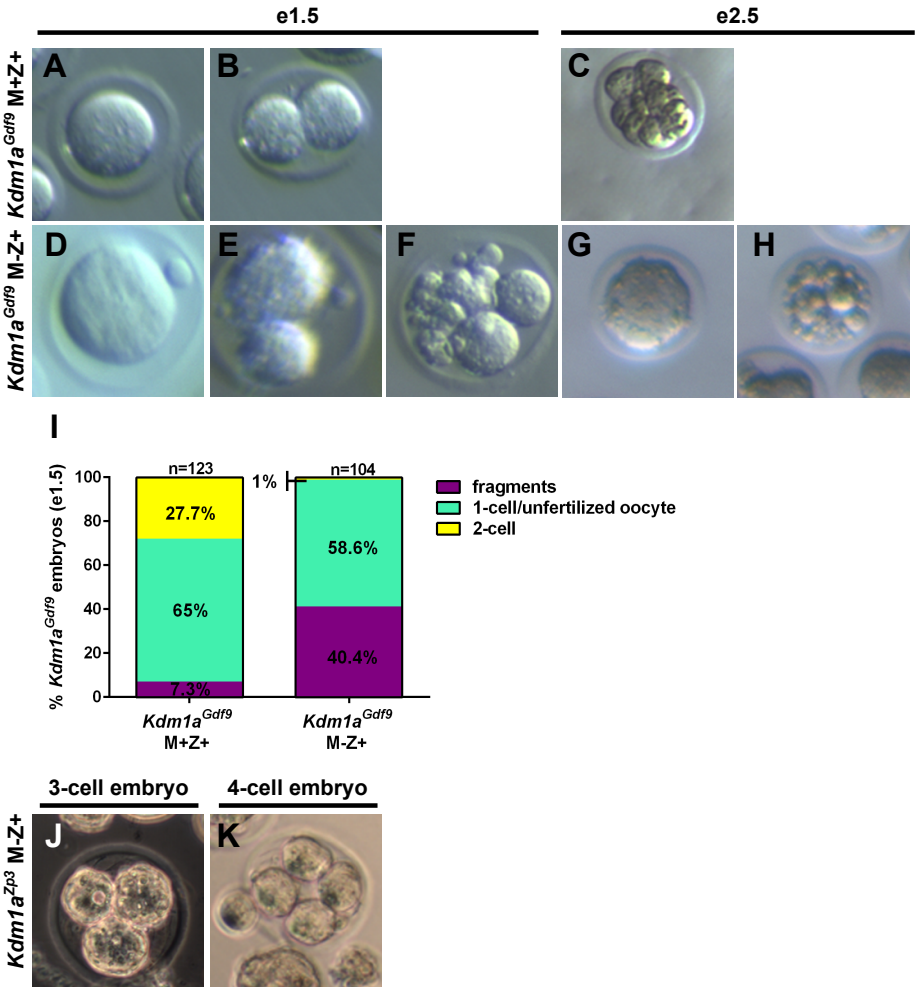


Figure 2-figure supplement 1:
Lack of Normal *Lsd1*^{Gdf9} and *Lsd1*^{Zp3} Embryos at embryonic day 1.5 and 2.5
(A,B,D,E,F) Brightfield images of embryonic day 1.5 (e1.5) M+Z+ 1-cell (A) and 2-cell (B) embryos and M-Z+ 1-cell (D),2-cell (E), and fragmented (F) embryos derived from *Lsd1*^{Gdf9} heterzygous control and mutant mothers. (C,G,H) Brightfield images of e2.5 M+Z+ 8-cell (C) embryo and M-Z+ abnormal 1-cell (G), and fragmented (H) embryos derived from *Lsd1*^{Gdf9} heterozygous control and mutant mothers.(I) Percentage of fragmented (purple), unfertilized oocyte or 1C (green), and 2C (yellow) embryos from *Lsd1*^{Gdf9} heterozygous control and mutant mothers. n = 123 for *Lsd1*^{Gdf9} M+Z+ control embryos from 8 litters. n = 104 for *Lsd1*^{Gdf9} M-Z+ embryos from 8 litters. (J) Brightfield image of 3-cell M-Z+ embryo derived from a *Lsd1*^{Zp3} mutant mother. (K) Brightfield image of 4-cell M-Z+ embryo derived from a *Lsd1*^{Zp3} mutant mother.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

embryos are clearly fertilized and arrested at the 1C stage. However, we do occasionally observe unfertilized oocytes. In addition, we sometimes observe embryos that are highly abnormal morphologically and are difficult to clearly assign to a particular category. As a result, we quantified these *Lsd1*^{Gdf9} M-Z+ embryos together. Nevertheless, compared to *Lsd1*^{Gdf9} M+Z+ embryos, *Lsd1*^{Gdf9} M-Z+ embryos have a large increase in the number of fragmented/degraded embryos at the expense of normal 2C embryos (Figure 2-figure supplement 1I). Also, the remaining 1C and 2C *Lsd1*^{Gdf9} M-Z+ embryos do not progress beyond the 1-2C stage, as we never observe any later stage embryos even at e2.5 (Figure 2-figure supplement 1C, G, H).

M-Z+ heterozygote embryos derived from *Lsd1*^{Zp3} mothers (hereafter referred to as *Lsd1*^{Zp3} M-Z+ embryos) undergo a phenotype that is similar to *Lsd1*^{Gdf9} M-Z+ embryos. At e1.5, 95% (N=20) of control *Lsd1*^{Zp3} M+Z+ embryos have reached the 2C stage (Figure 2E). In contrast, at e1.5 only 35% (N=57) of *Lsd1*^{Zp3} M-Z+ embryos reach the 2C stage (Figure 2E), though this is still much higher than the 1% of *Lsd1*^{Gdf9} M-Z+ embryos that reach the 2C stage (Figure 2-figure supplement 1I).

It is not clear why there is a difference in the number of *Lsd1*^{Zp3} versus *Lsd1*^{Gdf9} M-Z+ embryos that cleave to the 2C stage. However, this may be due to subtle differences in developmental timing, strain background, or *Cre* specific differences. In addition, we have observed two *Lsd1*^{Zp3} M-Z+ embryos that develop past the 2C stage to the 3C- and 4C- stage (Figure 2-figure supplement 1J,K), though like the *Lsd1*^{Gdf9} M-Z+ embryos, no *Lsd1*^{Zp3} M-Z+ embryos survive to the blastocyst stage (data not shown). Taken together, the results from *Lsd1*^{Gdf9} and *Lsd1*^{Zp3} mice suggest that the embryonic arrest, primarily at the 1-2C stages is the LSD1 maternal loss-of-function phenotype.

Loss of maternal LSD1 results in a failure to undergo the MZT

To determine the maternal function of LSD1, we considered two possibilities; (1) that LSD1 affects transcription in the oocyte, or (2) that maternal LSD1 affects transcription post-fertilization during the MZT. To determine if LSD1 affects transcription in the oocyte, we first compared the transcriptome of control *Lsd1*$^{fl/fl}$ [44] versus *Lsd1*^{Zp3} mutant oocytes. We found relatively few transcriptional changes, with 195 genes over expressed and 281 genes under expressed in *Lsd1*^{Zp3} oocytes (Figure 3A, Figure 3-source data 1A, Figure 3-figure supplement 5). In addition, amongst the 875 repeat types extracted from the RepeatMasker database (including LINEs, SINEs, etc.) there were only 2 repeat families that were activated and 2 repressed in *Lsd1*^{Zp3} oocytes (Figure 3A, Figure 3-source data 1A, Figure 3-figure supplement 5). Manual inspection revealed that few of these genes/repeats had both large fold-changes and low p-values, with the exception of the Mt1 gene, which was very highly activated in *Lsd1*^{Zp3} oocytes. Furthermore, we found several mitochondrially encoded genes (like mtCo3 and Atpase6) that had small but significantly reduced levels in LSD1^{Zp3} oocytes. These mitochondrial genes are unlikely to be directly regulated by LSD1, but their reduced levels could indicate a subtle metabolic defect in the mutants. Taken together, these results demonstrate that loss of LSD1 has little effect on transcription in oocytes.

Since we did not see a significant effect on oocytes, we considered the possibility that loss of LSD1 affects transcription post-fertilization during the MZT. To determine if this is the case, we isolated fertilized *Lsd1*^{Zp3} M-Z+ embryos that cleaved to the 2C stage, and compared them to *Lsd1*$^{fl/fl}$ M+Z+ 2C embryos[44]. We chose to focus on embryos from *Lsd1*^{Zp3} mothers because a substantially higher proportion of these embryos reach the 2C stage. We found a dramatic alteration of the transcriptome, with 1527 genes over expressed

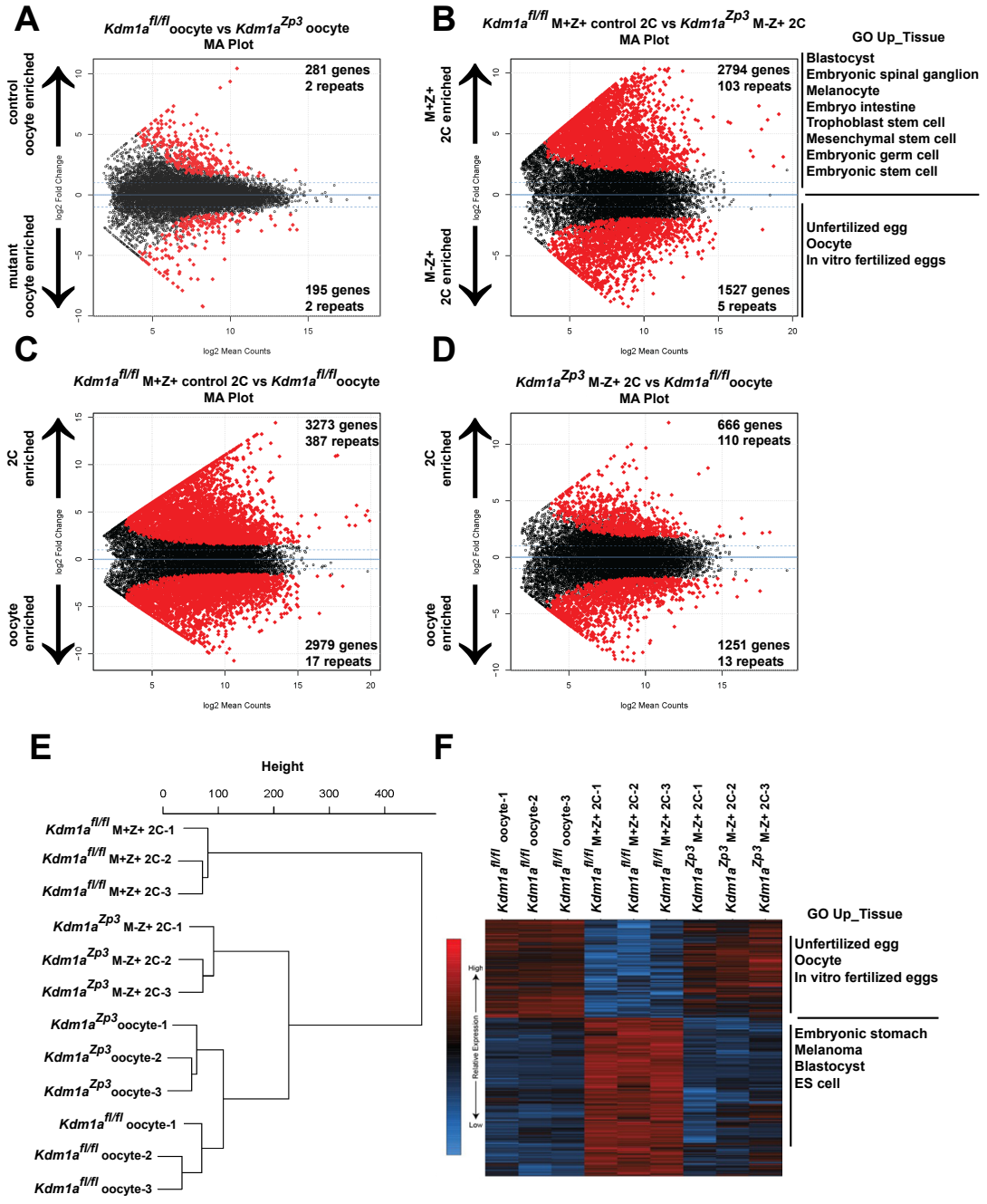

A
Kdm1a fl/fl oocyte vs Kdm1a Zp3 oocyte
MA Plot
control oocyte enriched
mutant oocyte enriched
281 genes
2 repeats
195 genes
2 repeats
log2 Fold Change
log2 Mean Counts
B
Kdm1a fl/fl M+Z+ control 2C vs Kdm1a Zp3 M-Z+ 2C
MA Plot
M+Z+ 2C enriched
M-Z+ 2C enriched
2794 genes
103 repeats
1527 genes
5 repeats
GO Up_Tissue
Blastocyst
Embryonic spinal ganglion
Melanocyte
Embryo intestine
Trophoblast stem cell
Mesenchymal stem cell
Embryonic germ cell
Embryonic stem cell
Unfertilized egg
Oocyte
In vitro fertilized eggs
C
Kdm1a fl/fl M+Z+ control 2C vs Kdm1a fl/fl oocyte
MA Plot
2C enriched
oocyte enriched
3273 genes
387 repeats
2979 genes
17 repeats
D
Kdm1a Zp3 M-Z+ 2C vs Kdm1a fl/fl oocyte
MA Plot
2C enriched
oocyte enriched
666 genes
110 repeats
1251 genes
13 repeats
E
Height
0
100
200
300
400
Kdm1a fl/fl M+Z+ 2C-1
Kdm1a fl/fl M+Z+ 2C-2
Kdm1a fl/fl M+Z+ 2C-3
Kdm1a Zp3 M-Z+ 2C-1
Kdm1a Zp3 M-Z+ 2C-2
Kdm1a Zp3 M-Z+ 2C-3
Kdm1a Zp3 oocyte-1
Kdm1a Zp3 oocyte-2
Kdm1a Zp3 oocyte-3
Kdm1a fl/fl oocyte-1
Kdm1a fl/fl oocyte-2
Kdm1a fl/fl oocyte-3
F
Kdm1a fl/fl oocyte-1
Kdm1a fl/fl oocyte-2
Kdm1a fl/fl oocyte-3
Kdm1a fl/fl M+Z+ 2C-1
Kdm1a fl/fl M+Z+ 2C-2
Kdm1a fl/fl M+Z+ 2C-3
Kdm1a Zp3 M-Z+ 2C-1
Kdm1a Zp3 M-Z+ 2C-2
Kdm1a Zp3 M-Z+ 2C-3
High
Relative Expression
Low
GO Up_Tissue
Unfertilized egg
Oocyte
In vitro fertilized eggs
Embryonic stomach
Melanoma
Blastocyst
ES cell

Figure 3.
The MZT is Impaired in *Lsd1*^{Zp3} Mutants
(A,B) Differential expression of mRNAs in *Lsd1*$^{fl/fl}$ versus *Lsd1*^{Zp3} oocytes (A) or *Lsd1*^{Zp3} M+Z+ versus *Lsd1*^{Zp3} M-Z+ 2C embryos (B) as determined by RNA-seq. Genes/repeats highlighted in red are significant with the number of significant gene/repeats show. GO enrichment using the Up_tissue database was performed on *Lsd1*^{Zp3} M+Z+ 2C enriched and *Lsd1*^{Zp3} M-Z+ 2C enriched mRNAs, with a list of the most enriched categories displayed. (C,D) Differential expression of mRNAs in *Lsd1*^{Zp3} M+Z+ 2C embryos versus *Lsd1*$^{fl/fl}$ oocytes (C) or *Lsd1*^{Zp3} M-Z+ 2C embryos versus *Lsd1*$^{fl/fl}$ oocytes (D). The numbers of zygotically activated (2C enriched) genes/repeats and zygotically repressed (oocyte enriched) genes/repeats are highlighted in each comparison. (E) Hierarchical cluster dendrogram of transcriptomes in *Lsd1*$^{fl/fl}$ oocytes, *Lsd1*^{Zp3} oocytes, *Lsd1*^{Zp3} M+Z+ 2C embryos, and *Lsd1*^{Zp3} M-Z+ 2C embryos. (F) Heat map of gene expression of principal component 1 (PC1) genes in *Lsd1*$^{fl/fl}$ oocytes, *Lsd1*^{Zp3} M+Z+ 2C embryos, and *Lsd1*^{Zp3} M-Z+ 2C embryos. The most GO Up_tissue enriched terms are displayed for the 2 categories of PC1 genes.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

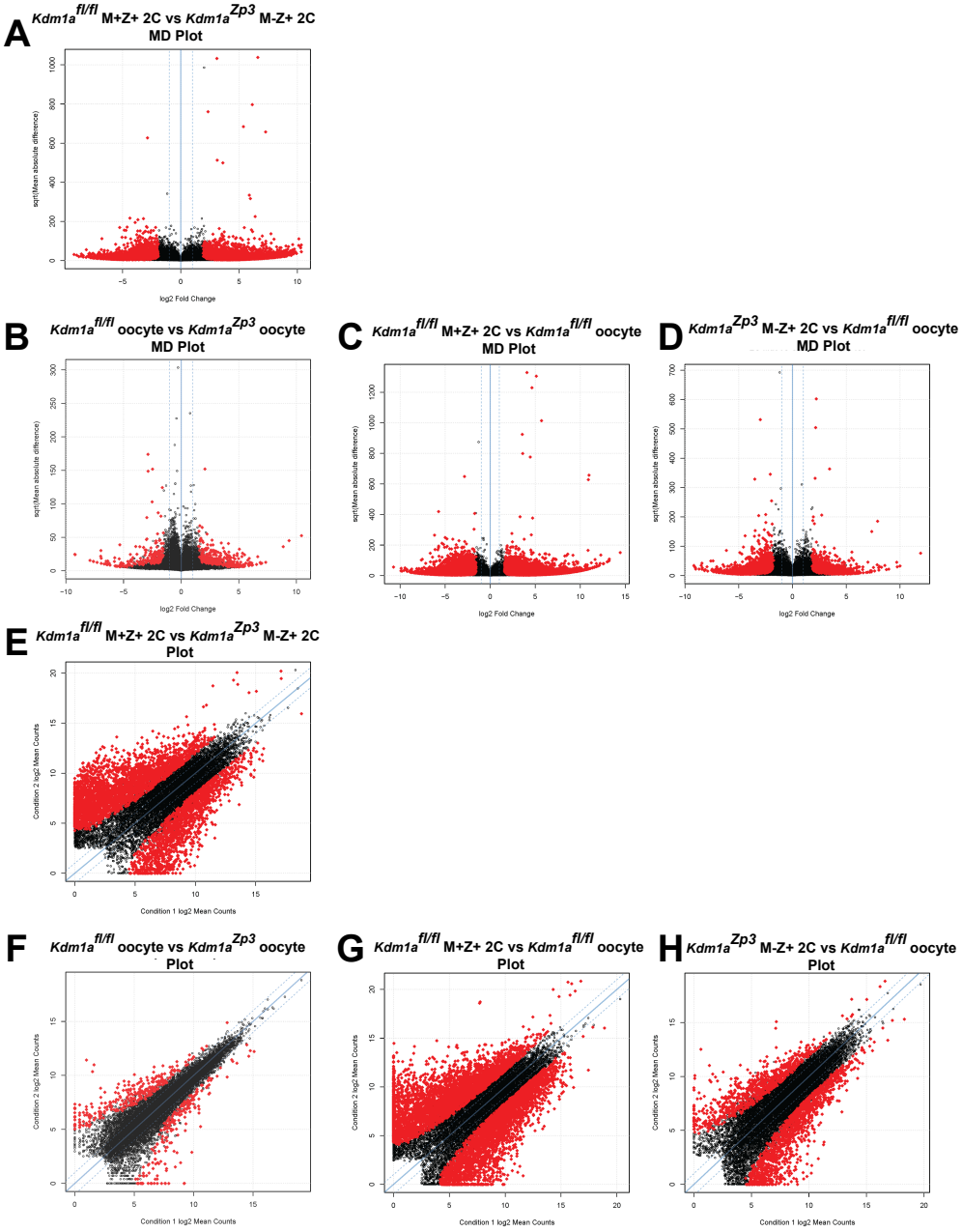


Figure 3-figure supplement 1:
***Lsd1*^{Zp3} Embryos Arrest at the 1-2 cell Stage**

(A-H) Differential expression of mRNAs in *Lsd1*$^{fl/fl}$ versus *Lsd1*^{Zp3} oocytes (A,E), *Lsd1*^{Zp3} M+Z+ versus *Lsd1*^{Zp3} M-Z+ 2C embryos (B,F), *Lsd1*^{Zp3} M+Z+ versus *Lsd1*$^{fl/fl}$ oocytes (C,G), and *Lsd1*^{Zp3} M-Z+ versus *Lsd1*$^{fl/fl}$ oocytes (D,H) as determined by RNA-seq. Genes/repeats highlighted in red are significant.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

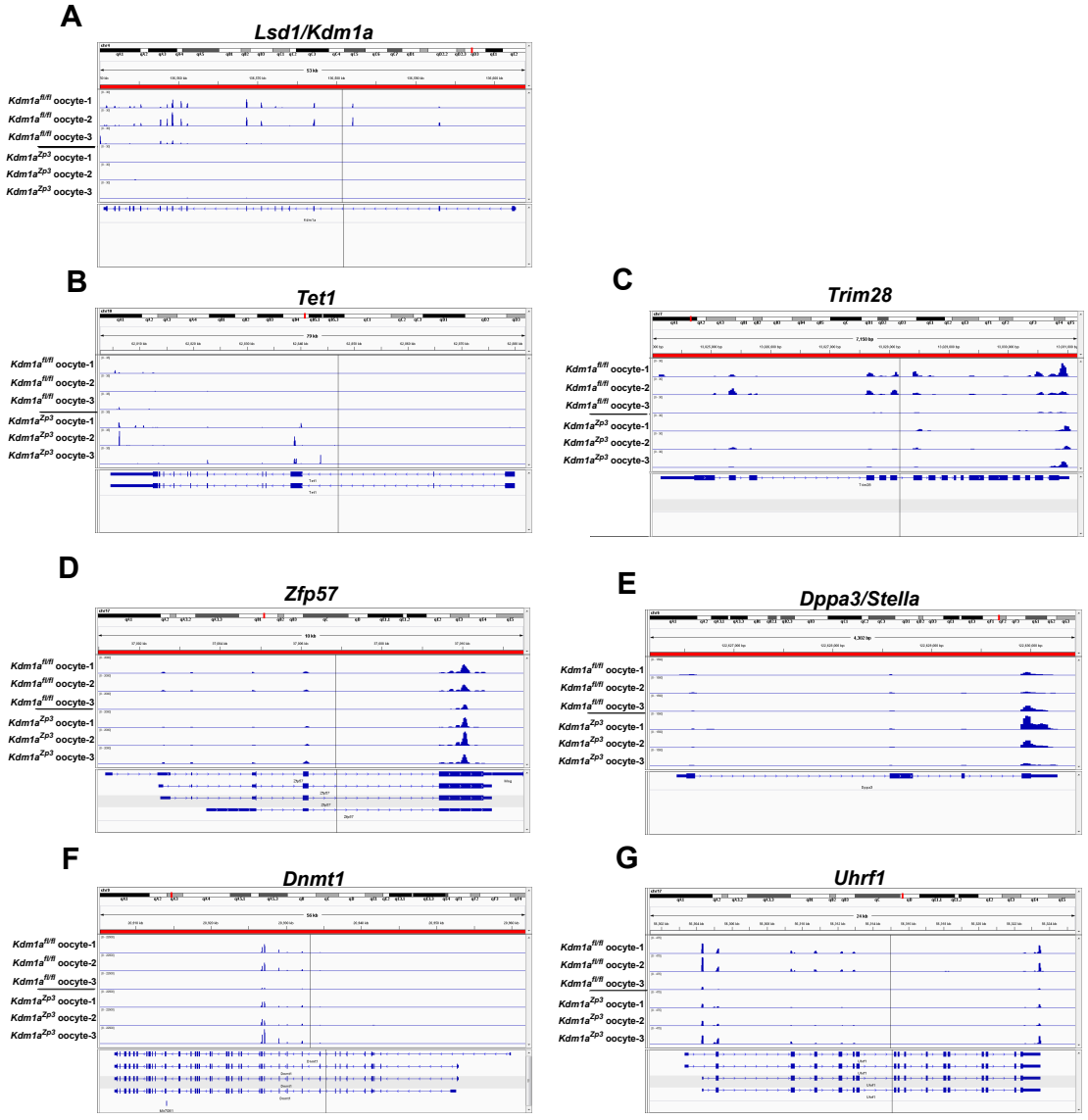


Figure 3-figure supplement 5:
Expression of Epigenetic regulators in *Lsd1*$^{fl/fl}$ and *Lsd1*^{Zp3} oocytes
Sequenced RNA-seq reads showing relative expression from *Lsd1*$^{fl/fl}$ oocytes and *Lsd1*^{Zp3} mutant oocytes aligned to the genome for *Lsd1/Kdm1a*(A), *Tet1*(B), *Trim28*(C), *Zfp57*(D), *Dppa3/stella*(E), *Dnmt1*(F) and *Uhrf1*(G). Gene tracks visualized using Integrative Genomics Viewer.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

and 2794 genes under expressed in *Lsd1*^{Zp3} M-Z+ 2C embryos (Figure 3B, Figure 3-source data 1B). In addition, there were 103 repeat families that were repressed in *Lsd1*^{Zp3} 2C M-Z+ embryos, compared with only 5 that were activated (Figure 3B, Figure 3-source data 1B). These findings are in contrast to what has been observed in *Lsd1* mutant ES cells that display a general de-repression of both genes and repeats[44,45]. GO analysis of over and under expressed genes in *Lsd1*^{Zp3} 2C M-Z+ embryos demonstrated an enrichment of genes normally expressed in unfertilized oocytes, and a reduction in genes associated with embryonic development and tissue specific gene expression. This suggests that *Lsd1*^{Zp3} M-Z+ 2C embryos may fail to undergo the MZT (Figure 3B, Figure 3-source data 1B).

To test this hypothesis further, we compared the expression of mRNAs from control *Lsd1*$^{fl/fl}$ M+Z+ 2C embryos[44] or *Lsd1*^{Zp3} M-Z+ 2C embryos with *Lsd1*$^{fl/fl}$ oocytes[44]. If the *Lsd1*^{Zp3} M-Z+ 2C embryos fail to undergo the MZT, then we would expect them to be more similar to *Lsd1*$^{fl/fl}$ oocytes than *Lsd1*$^{fl/fl}$ M+Z+ 2C embryos. In *Lsd1*$^{fl/fl}$ M+Z+ 2C embryos relative to *Lsd1*$^{fl/fl}$ oocytes, there are massive changes in the mRNA profile, with >3,000 genes (and 387 repeat families) becoming zygotically activated and ~3,000 maternal genes (and 17 repeat families) which are suppressed. This demonstrates that MZT has occurred (Figure 3C, Figure 3-source data 1C). In contrast, *Lsd1*^{Zp3} M-Z+ 2C embryos fail to properly activate the zygotic genome, with only 666 genes becoming activated (and 110 repeat families), and only ~1200 maternal genes repressed (13 repeat families) (Figure 3D, Figure 3-source data 1D). Hierarchical clustering of *Lsd1*$^{fl/fl}$ oocytes, *Lsd1*^{Zp3} oocytes, *Lsd1*$^{fl/fl}$ M+Z+ 2C embryos and *Lsd1*^{Zp3} M-Z+ 2C embryos confirms this failure to activate the zygotic genome. Specifically, though *Lsd1*$^{fl/fl}$ oocytes are similar to *Lsd1*^{Zp3} oocytes, *Lsd1*^{Zp3} M-Z+ 2C embryos are more similar to *Lsd1*$^{fl/fl}$ oocytes than *Lsd1*$^{fl/fl}$ M+Z+ 2C embryos (Figure 3E). In addition, principal component analysis (PCA) demonstrated that ~40% of the variance

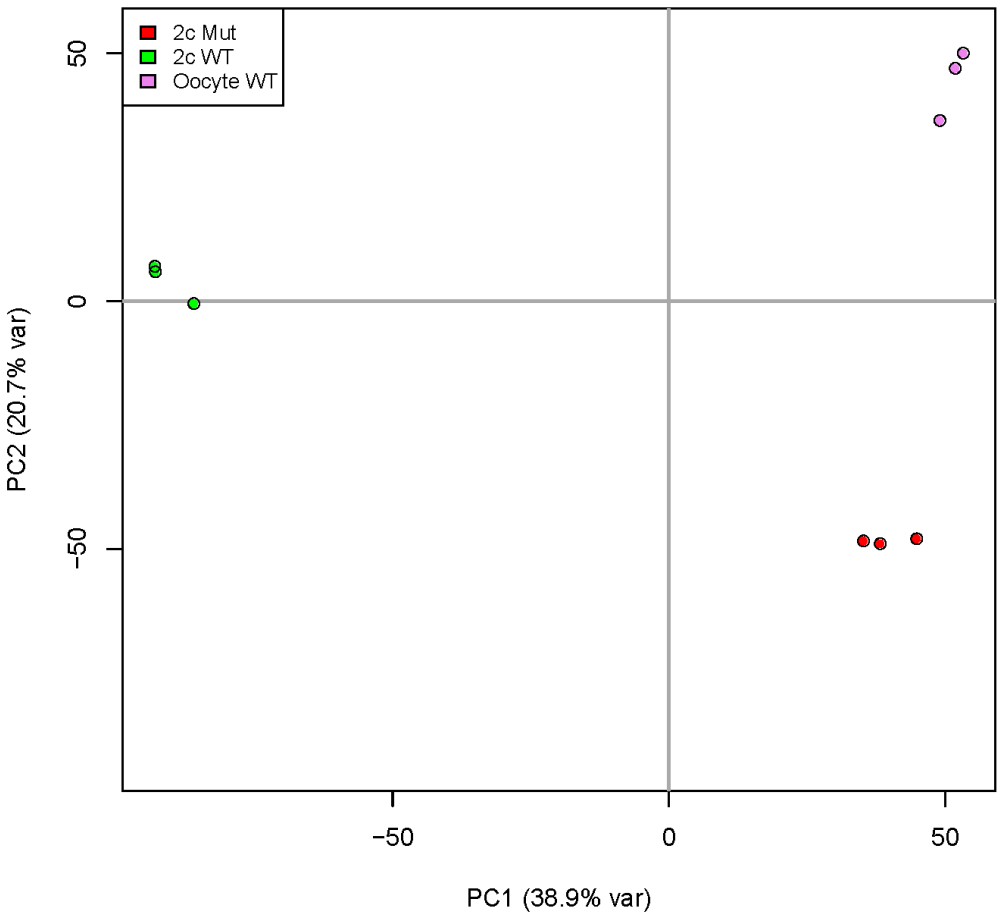


Figure 3-figure supplement 2:
Principal Component Analysis of *Lsd1*^{Zp3} 2C Embryos
(A)Principal Component 1 is plotted on x-axis and Principal Component 2 is plotted on y-axis. Variance due to each component for *Lsd1*^{Zp3} M-Z+ embryos (red), *Lsd1*^{Zp3} M+Z+ control embryos (green), and *Lsd1*$^{fl/fl}$ oocytes (purple) are shown.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

between *Lsd1*$^{fl/fl}$ oocytes, *Lsd1*$^{fl/fl}$ M+Z+ 2C embryos, and *Lsd1*^{Zp3} M-Z+ 2C embryos can be explained by a single component (PC1, Figure 3-figure supplement 2), and the heat maps generated from PC1 genes confirm that *Lsd1*^{Zp3} M-Z+ 2C embryos have expression profiles more similar to *Lsd1*$^{fl/fl}$ oocytes than *Lsd1*$^{fl/fl}$ M+Z+ 2C embryos (Figure 3F). GO analysis demonstrated that PC1 genes fall into two categories, those associated with unfertilized oocytes and are highly expressed in oocytes and *Lsd1*^{Zp3} M-Z+ 2C embryos, and those associated with early embryonic development which are enriched only in *Lsd1*$^{fl/fl}$ M+Z+ 2C embryos (Figure 3F). Taken together, these data confirm that the major defect in the transcriptome of *Lsd1*^{Zp3} M-Z+ 2C embryos is a failure to undergo the MZT.

Deletion of *Lsd1* maternally with *Vasa-Cre* results in a hypomorphic phenotype

Deletion of *Lsd1* with either *Gdf9-Cre* or *Zp3-Cre* results in 100% of oocytes completely lacking LSD1 (Figure 1C-F). In contrast, 66.7% of *Lsd1*^{Vasa} mutant oocytes retain some LSD1 protein, though the amount is much lower than heterozygous control oocytes (Figure 1G-I). To determine if this lower amount of LSD1 retained in some *Lsd1*^{Vasa} mutant oocytes gives rise to a hypomorphic phenotype, we mated *Lsd1*^{Vasa} females to wild-type males to generate heterozygous M-Z+ offspring (Figure 1-figure supplement 2). As is the case with *Lsd1*^{Gdf9} and *Lsd1*^{Zp3} M-Z+ offspring, these heterozygous M-Z+ offspring (hereafter referred to as *Lsd1*^{Vasa} M-Z+ embryos) have a normal paternal *Lsd1* allele. Thus, any phenotypic effects in these *Lsd1*^{Vasa} heterozygous M-Z+ offspring are due to reduced maternal LSD1.

Similar to *Lsd1*^{Gdf9} and *Lsd1*^{Zp3} M-Z+ embryos, the majority of *Lsd1*^{Vasa} M-Z+ embryos arrest prior to the blastocyst stage (Figure 4A-E). In control *Lsd1*^{Vasa} M+Z+ embryos at e3.5, 55% of embryos have reached the blastocyst stage, 35% are multicellular

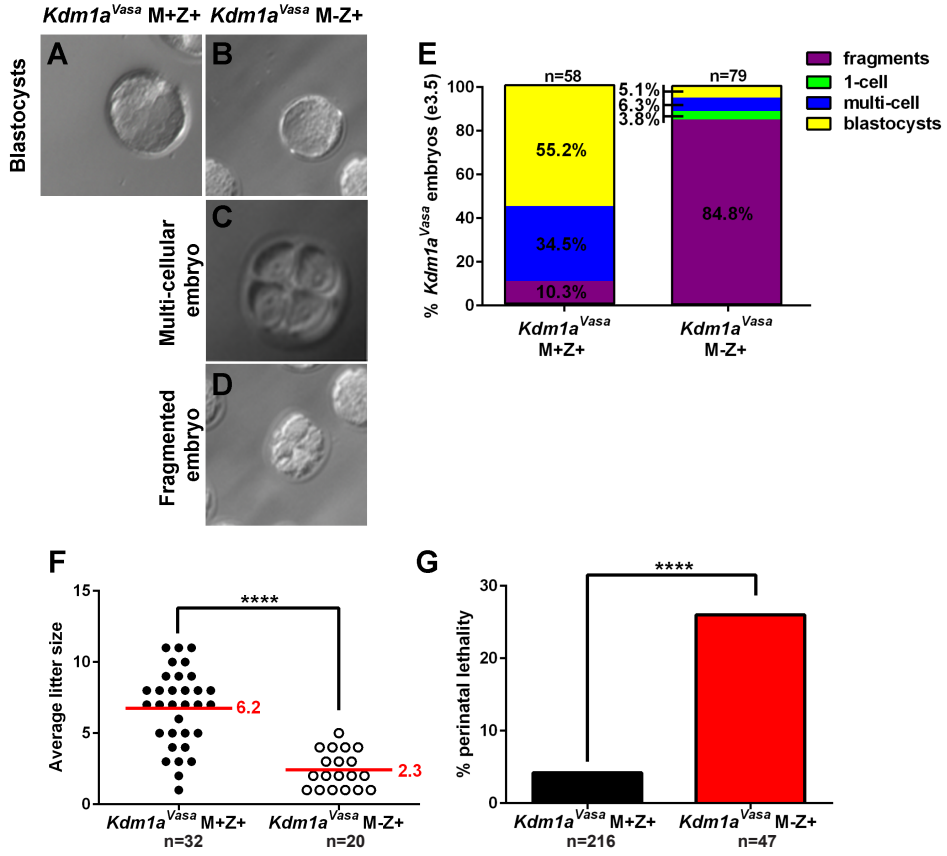


Figure 4.

Hypomorphic Phenotype in *Lsd1*^{Vasa} Progeny

(A,B,C,D) Brightfield images of M+Z+ (A) and M-Z+ (B,C,D) embryos derived from *Lsd1*^{Vasa} heterozygous control and mutant mothers at embryonic day 3.5 (e3.5). Panels show blastocysts (A,B), a multicellular embryo (C) and a fragmented embryo (D). (E) Percentage of fragmented (purple), 1-cell (green), multi-cellular (blue) and blastocyst (yellow) embryos from *Lsd1*^{Vasa} heterozygous control and mutant mothers at e3.5. n =58 for *Lsd1*^{Vasa} M+Z+ embryos from 7 litters. n = 79 for *Lsd1*^{Vasa} M-Z+ embryos from 10 litters. (F) Litter sizes of *Lsd1*^{Vasa} heterozygous control and mutant mothers. Average litter size for each indicated by red line. Each circle indicates one litter and n=number of litters analyzed. p-values calculated using an unpaired t-test with **** = p<0.0001 indicating statistical significance. (G) Percentage of newborn pups from *Lsd1*^{Vasa} heterozygous control and mutant mothers that died perinatally. n = number of litters analyzed. p-values calculated using an unpaired t-test with **** = p<0.0001 indicating statistical significance.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

(>2C) and 10% are fragmented/degraded (n=37, Figure 4E). In contrast, at e3.5 85% of the *Lsd1*^{Vasa} M-Z+ embryos have fragmented/degraded (n=35, Figure 4E). However, unlike *Lsd1*^{Gdj9} and *Lsd1*^{Zp3} M-Z+ embryos, which undergo complete embryonic arrest at the 1-2C stage, by e3.5 6% of *Lsd1*^{Vasa} M-Z+ embryos have reached the multicellular stage (>2C) and 5% are blastocysts (n=35, Figure 4E). Remarkably, some of these *Lsd1*^{Vasa} M-Z+ embryos survive to birth. The average litter size born from *Lsd1*^{Vasa} mutant mothers is 2.3 (n=20), versus 6.2 (n=37) born from littermate control mothers (Figure 4F). However, the vast majority of time vaginal plugged *Lsd1*^{Vasa} mutant mothers give rise to no viable progeny. Thus, this average litter size is undoubtedly a vast overestimate of the survival of *Lsd1*^{Vasa} M-Z+ embryos overall.

The transition from *in utero* development to postnatal development, during the first 48 hours immediately following birth, is a highly stressful time for mice. If a mouse has a subtle defect during embryogenesis, the newborn pup may die perinatally. This is reflected in the large number of mouse models that exhibit perinatal lethality[188]. Therefore, to determine if lower maternal LSD1 might give rise to more subtle embryonic defects, we carefully monitored the first 48 hours following birth in *Lsd1*^{Vasa} M-Z+ progeny. Consistent with this being a major developmental transition, even in genotypically normal mice, we observe that 5% of offspring (N=216) born from littermate *Lsd1*^{Vasa} control mothers die perinatally during the first 48 hours after birth (Figure 4G). However, in *Lsd1*^{Vasa} M-Z+ progeny, the percentage of perinatal lethality is significantly increased (26%, n=47) (Figure 4G). It is unclear why these animals are dying perinatally. However, this increase in perinatal lethality is consistent with these animals having more subtle developmental defects. Of note, on the rare occasions that more than one animal was born from *Lsd1*^{Vasa} mutant mothers, the full litter either all died or all survived. The significance of this phenomenon is unclear.

Lsd1^{Vasa} M-Z+ progeny exhibit abnormal behavior

Of the small number of M-Z+ animals that are born from *Lsd1*^{Vasa} mutant mothers, nearly 1/3 die perinatally in the first 48 hours after birth. However, the remaining 2/3 survive to adulthood and appear to be morphologically normal. Notably, because of the embryonic arrest and the perinatal lethality, this remaining 2/3 is a very small number of animals (n=10 from 8 crosses). Since progeny from *Lsd1*^{Vasa} mutant mothers have defects that can lead to perinatal lethality, we considered the possibility that the surviving animals from these *Lsd1*^{Vasa} mutant mothers (hereafter referred to as *Lsd1*^{Vasa} M-Z+ adults) could have phenotypic defects as well. To analyze these surviving animals, we crossed *Lsd1*^{Vasa} mutant mothers (C57BL/6 background) to *M. castaneus* (CAST) fathers. This enabled us to use polymorphisms to examine subsequent molecular defects in a parent of origin specific fashion. Importantly, F1 hybrid (B6/CAST) offspring from *Lsd1*^{Vasa} mutant mothers have the same hypomorphic phenotype that we observe in a B6 background, where a small number of animals survive until birth (average litter size=1.6, n=6 litters from 8 crosses).

Mating the surviving *Lsd1*^{Vasa} M-Z+ adults produced progeny, suggesting no gross defects in the specification of the germline or mating behavior. However, when we observed these *Lsd1*^{Vasa} M-Z+ adults closely, we noticed that they exhibit abnormal behaviors (Figure 5-figure supplement 1A, Video 1). These abnormal behaviors include excessive scratching and digging, along with food harassing, where animals grind up all of the remaining food in the food hopper and incorporate it into the bedding (Figure 5A-C). To quantify the food harassing behavior, we monitored both the weight of the food in the food hopper and the height of the bedding in the cage. Over a 3-day period, each individual *Lsd1*^{Vasa} M-Z+ adult almost completely depleted the food in the food hopper due to food grinding and harassing

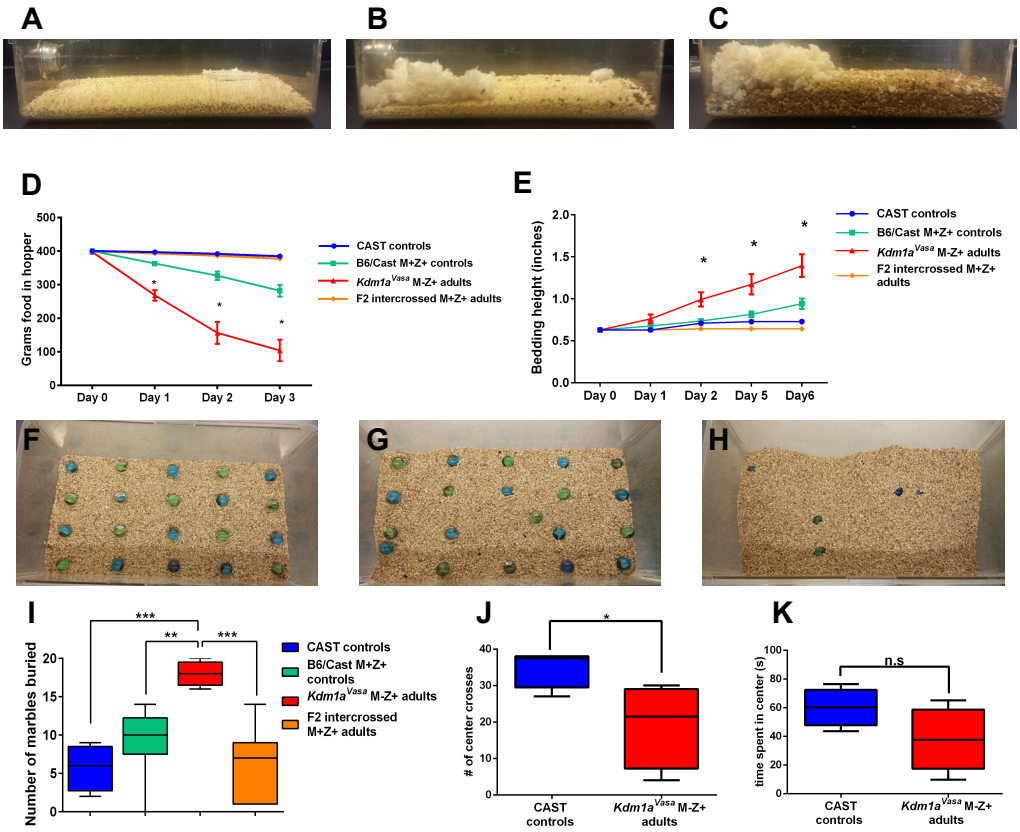


Figure 5.

Abnormal Behaviors in *Lsd1*^{Vasa} M-Z+ Adults

(A,B,C) Mouse cages at day 0 (A) and day 8 (B) from *M. castaneus* (CAST) controls compared to day 6 (C) from a *Lsd1*^{Vasa} M-Z+ adult. (D) Quantification of change in weight of food in the hopper from CAST controls, B6/CAST hybrid M+Z+ controls, and F2 intercrossed M+Z+ adults versus *Lsd1*^{Vasa} M-Z+ adults. Data are shown as mean for each day with error bars indicating ±S.E.M. (E) Quantification of change in bedding height from CAST controls, B6/CAST hybrid M+Z+ controls, and F2 intercrossed M+Z+ adults versus *Lsd1*^{Vasa} M-Z+ adults. Data are shown as mean for each day with error bars indicating ±S.E.M. (F,G,H) Mouse cages before (F) and after (G,H) the marble burying assay was performed on a CAST control (G) compared to a *Lsd1*^{Vasa} M-Z+ adult (H). (I) Quantification of the number of marbles buried during the marble burying assay performed on CAST controls, B6/CAST hybrid M+Z+ controls, and F2 intercrossed M+Z adults versus *Lsd1*^{Vasa} M-Z+ adults. Data are shown as quartiles with error bars indicating ±S.E.M. (J,K) Open field test performance in CAST controls versus *Lsd1*^{Vasa} M-Z+ adults scored by number of center crosses (J) and time spent in center of cage (K). Data are shown as quartiles with error bars indicating ±S.E.M. p-values calculated using an unpaired t-test with n.s. indicating $p>0.05$, * = $p<0.05$, ** = $p<0.005$, *** = $p<0.0005$. All asterisks indicate statistical significance.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

(Figure 5D and Figure 5-figure supplement 1B-D). This resulted in a correspondingly large increase in the height of the bedding over a 6-day period (Figure 5E). These effects are not observed in *Lsd1*^{Vasa} mutant mothers or the *M. castaneus* animals to which they were mated (Figure 5D, E and Figure 5-figure supplement 1B, E), indicating that the behavior may be due to a maternal effect. To determine if this behavior is due to a maternal effect, we intercrossed the affected *Lsd1*^{Vasa} M-Z+ adults. If the food-grinding behavior is due to a maternal effect, then mating two affected *Lsd1*^{Vasa} M-Z+ adults should not produce progeny with abnormal behavior. This is because the *Lsd1*^{Vasa} M-Z+ adults used as mothers in these crosses have a normal *Lsd1* allele (-/+) so the resulting intercrossed progeny (hereafter referred to as F2 intercrossed M+Z+ adults) are all M+. Intercrossing affected M-Z+ F1 hybrids produced no F2 intercrossed M+Z+ adults with the food-grinding behavior, suggesting that the food-grinding behavior is dependent upon a maternal effect (Figure 5D,E and Figure 5-figure supplement 1D,G). Nevertheless, it remains possible that the F1 hybrid background also contributes to the abnormal behavior. To test this possibility, we crossed control littermates of *Lsd1*^{Vasa} mothers to CAST fathers to generate control B6/CAST hybrids (hereafter referred to as B6/CAST M+Z+ controls). These control hybrids exhibit modest food-grinding behavior, suggesting that the hybrid strain background likely also contributes to the maternal food-grinding defect (Figure 5D, E and Figure 5-figure supplement 1C, F). However, *Lsd1*^{Vasa} M-Z+ adults are significantly more affected than B6/CAST M+Z+ hybrid controls (Figure 5D,E and Figure 5-figure supplement 1C,F), suggesting that the abnormal food-grinding behavior is predominantly due to a maternal effect. Importantly, we also do not observe any *Lsd1* haploinsufficiency defects. For example, the F2 intercrossed M+Z+ adults that are heterozygous for *Lsd1* do not exhibit this abnormal behavior (Figure 5D, E; Figure 5-figure supplement 1D, G). This suggests that

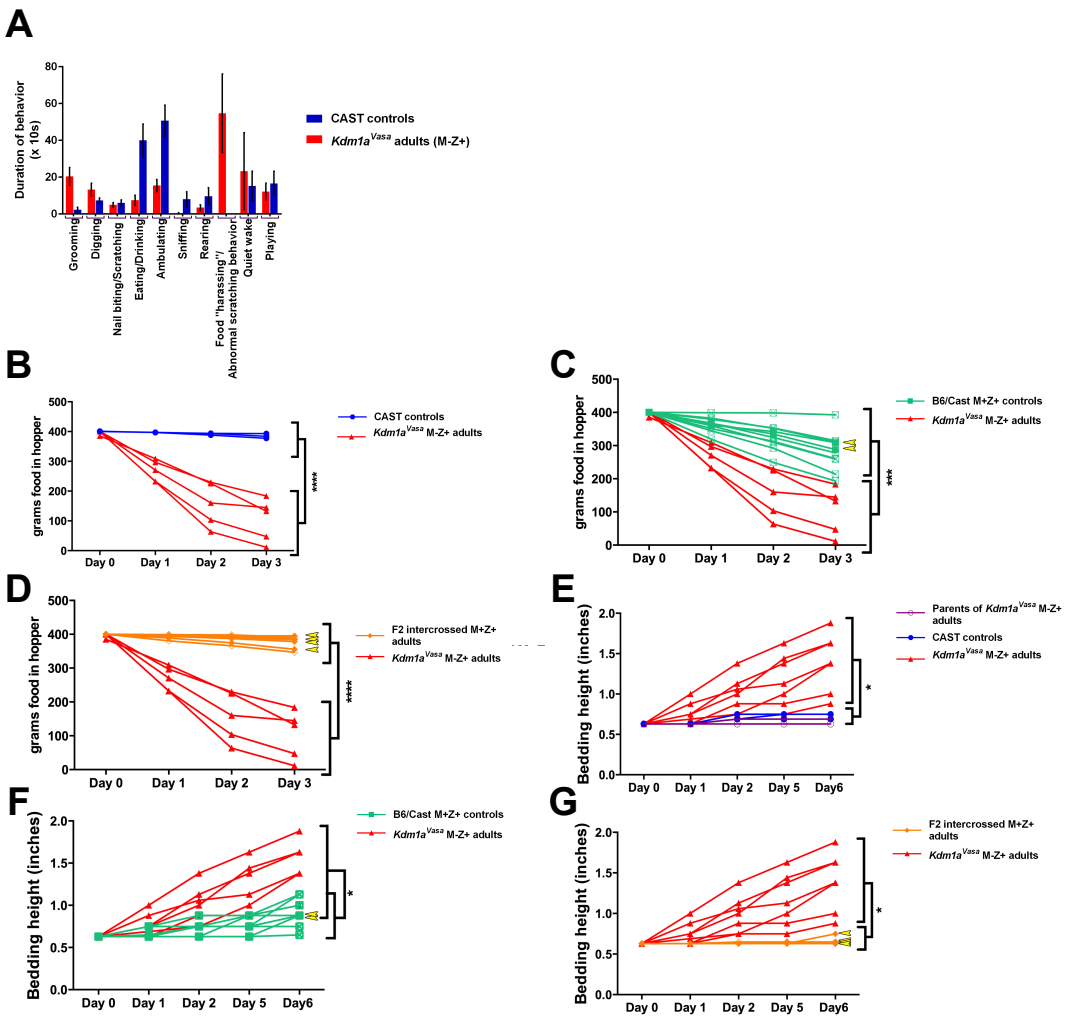

A
Duration of behavior (x 10s)
80
60
40
20
0
Grooming
Digging
Nail biting/Scratching
Eating/Drinking
Ambulating
Sniffing
Rearing
Food "harassing"/ Abnormal scratching behavior
Quiet wake
Playing
CAST controls
Kdm1a Vasa adults (M-Z+)
B
grams food in hopper
500
400
300
200
100
0
Day 0
Day 1
Day 2
Day 3
****
CAST controls
Kdm1a Vasa M-Z+ adults
C
grams food in hopper
***
B6/Cast M+Z+ controls
Kdm1a Vasa M-Z+ adults
D
grams food in hopper
****
F2 intercrossed M+Z+ adults
Kdm1a Vasa M-Z+ adults
E
Bedding height (inches)
2.0
1.5
1.0
0.5
0.0
Day 0
Day 1
Day 2
Day 5
Day6
*
Parents of Kdm1a Vasa M-Z+ adults
CAST controls
Kdm1a Vasa M-Z+ adults
F
Bedding height (inches)
*
B6/Cast M+Z+ controls
Kdm1a Vasa M-Z+ adults
G
Bedding height (inches)
*
F2 intercrossed M+Z+ adults
Kdm1a Vasa M-Z+ adults

Figure 5-figure supplement 1:
Abnormal Behaviors in individual *Lsd1*^{Vasa} M-Z+ Adults
(A) Behavioral ethogram of *M. castaneus* (CAST) controls versus *Lsd1*^{Vasa} M-Z+ adults. (B) Quantification of change in weight of food in the hopper of parents of *Lsd1*^{Vasa} M-Z+ adults and CAST controls versus *Lsd1*^{Vasa} M-Z+ adults. (C) Quantification of change in weight of food in the hopper of B6/CAST M+Z+ controls versus *Lsd1*^{Vasa} M-Z+ adults. (D) Quantification of change in weight of food in the hopper of F2 intercrossed M+Z+ adults versus *Lsd1*^{Vasa} M-Z+ adults. (E) Quantification of change in bedding height of parents of *Lsd1*^{Vasa} M-Z+ adults and CAST controls versus *Lsd1*^{Vasa} M-Z+ adults. (F) Quantification of change in bedding height of B6/CAST M+Z+ controls versus *Lsd1*^{Vasa} M-Z+ adults. (G) Quantification of change in bedding height of F2 intercrossed M+Z+ adults versus *Lsd1*^{Vasa} M-Z+ adults. The measurements for each individual animal (B,C,D) and (E,F,G) correspond to the averages shown in Figure 5 (D,E). Yellow arrowheads represent animals heterozygous for *Lsd1*. Data shown as mean for each day. p-values calculated using an unpaired t-test with * = $p<0.05$, *** = $p<0.0005$, **** = $p<0.0001$. All asterisks indicate statistical significance. p-values calculated using an unpaired t-test with n.s. indicating $p>0.05$, * = $p<0.05$, ** = $p<0.005$, *** = $p<0.0005$. All asterisks indicate statistical significance.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

the abnormal behavior is not due to *Lsd1* haploinsufficiency. Finally, behavioral defects were observed in *Lsd1*^{Vasa} M-Z+ survivors obtained from litters ranging from 1-4 animals. This suggests that the maternal behavioral abnormalities are not simply due to reduced litter size (Table 1).

Observations of *Lsd1*^{Vasa} M-Z+ adults indicated that these mice also exhibit excessive digging (Figure 5-figure supplement 1A, Video 1). To quantify this behavior, we assayed *Lsd1*^{Vasa} M-Z+ adults and controls in a marble-burying assay. In this assay, individual mice are placed in a cage in which 20 marbles are arrayed on top of the bedding. Mice exhibiting excessive digging behavior will bury higher numbers of marbles. For example, *Slitrk5* obsessive-compulsive mice bury approximately 50% of the marbles in 30 minutes, compared to approximately 25% in controls[189].

Lsd1^{Vasa} M-Z+ adults exhibit striking behavior in the marble-burying assay. Compared to *M. castaneus* controls, which bury 25% in 25 minutes (Figure 5G, I, Video 2), *Lsd1*^{Vasa} M-Z+ adults bury 90% in the same time period (Figure 5H, I, Video 2). Similar to what was observed in the food-grinding assays, we do not observe this behavior in F2 intercrossed M+Z+ adults (average 7 marbles, Figure 5I), but we do see slightly elevated marble-burying in B6/CAST hybrid controls (average 10, Figure 5I). Nevertheless, as in the food-grinding assay, *Lsd1*^{Vasa} M-Z+ adults are significantly more affected than B6/CAST hybrid M+Z+ controls (Figure 5D,E and Figure 5-figure supplement 1C,F). Taken together, these data indicate that the excessive digging behavior is also predominantly due to a maternal effect.

Finally, we also performed the open-field test on *Lsd1*^{Vasa} M-Z+ adults. In this assay, animals are placed in a much larger space than they are accustomed to. Animals that are

more anxious or fearful cross the middle of cage fewer times and spend less time overall in the center (Video 3). Consistent with the behavioral defects that we observe in food-grinding and marble-burying, *Lsd1*^{Vasa} M-Z+ adults perform significantly fewer movements across the middle of the cage (Figure 5J) and spend less time in the center overall than CAST controls (Figure 5K).

***Lsd1*^{Vasa} M-Z+ progeny have imprinting defects**

The perinatal lethality and behavioral defects observed in *Lsd1*^{Vasa} M-Z+ progeny are remarkable because these animals have a normal *Lsd1* allele. Thus the defects that we observe in *Lsd1*^{Vasa} M-Z+ progeny must be due to a heritable effect originating from the low level of maternal LSD1. As a result, we sought to determine the nature of this heritable defect. The *Lsd1* homolog *Lsd2* is expressed maternally in mice, and loss of *Lsd2* results in a heritable embryonic lethality defect associated with a failure to maternally acquire DNA methylation at a subset of imprinted genes[40]. Therefore, we considered the possibility that the heritable defects could be due to DNA methylation defects at these loci, or in maternally methylated imprinted loci unaffected by the loss of LSD2. In addition, in our RNA-seq data, we observe the misregulation of multiple genes that could potentially affect DNA methylation at imprinted loci. For example, four genes that are known to maternally affect DNA methylation at imprinted loci, *Tet1, Trim28*, *Zfp57* and *Dppa3/Stella*[21,153,168,190,191], are all misregulated in *Lsd1*^{Zp3} M-Z+ mutants (*Tet1:* -3.8fold**,** *Trim28:* -3.3fold, *Zfp57:* +3.2fold and *Dppa3/Stella:* +.58 fold, Figure 3-figure supplement 3; Figure 3-source data 1B, quantitative RT-PCR validation of *Trim28, Zfp57* and *Dppa3/Stella* in Figure 3-figure supplement 4). Additionally, the DNA methyltransferase DNMT1 is overexpressed in *Lsd1*^{Zp3} M-Z+ mutants (*Dnmt1:* +2.1fold, Figure 3-figure supplement 3; Figure 3-source data 1B). DNMT1

is thought to primarily act as the maintenance DNA methyltransferase[142,192], but when overexpressed, DNMT1 has been shown to have *de novo* DNA methyltransferase activity[193]. Taken together, these observations raise the possibility that DNA methylation and monoallelic expression at imprinted genes could be altered in *Lsd1*^{Vasa} M-Z+ progeny, either directly or indirectly. Interestingly, despite the clear loss of LSD1 in *Lsd1*^{Zp3} mutant oocytes, the regulatory proteins that are affected by the loss of LSD1 are either not expressed in oocytes (Tet1) or not affected until after fertilization (*Trim28, Zfp57, Dppa3*/*Stella* and *Dnmt1,* Figure 3-figure supplement 4,5). This is consistent with our previous observations that LSD1 primarily affects the MZT.

To investigate DNA methylation and expression of imprinted genes in *Lsd1*^{Vasa} M-Z+ offspring, we performed bisulfite analysis and qRT-PCR on perinatal lethal *Lsd1*^{Vasa} M-Z+ pups. We analyzed imprinted genes that were both affected (*Zac1, Impact, Grb10, Mest*), and unaffected in *Lsd2* mutants (*H19, Igf2r, Snrpn*), as well as those that are both maternally (*Zac1, Impact, Grb10, Igf2r, Mest, Snrpn*) and paternally methylated (*H19*). For these analyses, we took advantage of B6/CAST polymorphisms to determine whether individual alleles in F1 hybrid offspring came from the mother or the father. These F1 hybrid pups, derived from hypomorphic B6 *Lsd1*^{Vasa} mothers mated to wild-type CAST fathers, are denoted as maternal effect progeny 1 and 2 (MEP1 and MEP2), and were compared to two stage matched B6/CAST hybrid controls.

Two imprinted loci, *Zac1* and *Impact*, are severely affected in *Lsd1*^{Vasa} offspring. At *Zac1,* DNA methylation is inappropriately acquired on the paternal allele in both MEP, though to a slightly lesser extent in MEP2 (Figure 6A). This is associated with significantly less expression of *Zac1* in both MEP (Figure 6B). Moreover, *Zac1* is normally expressed

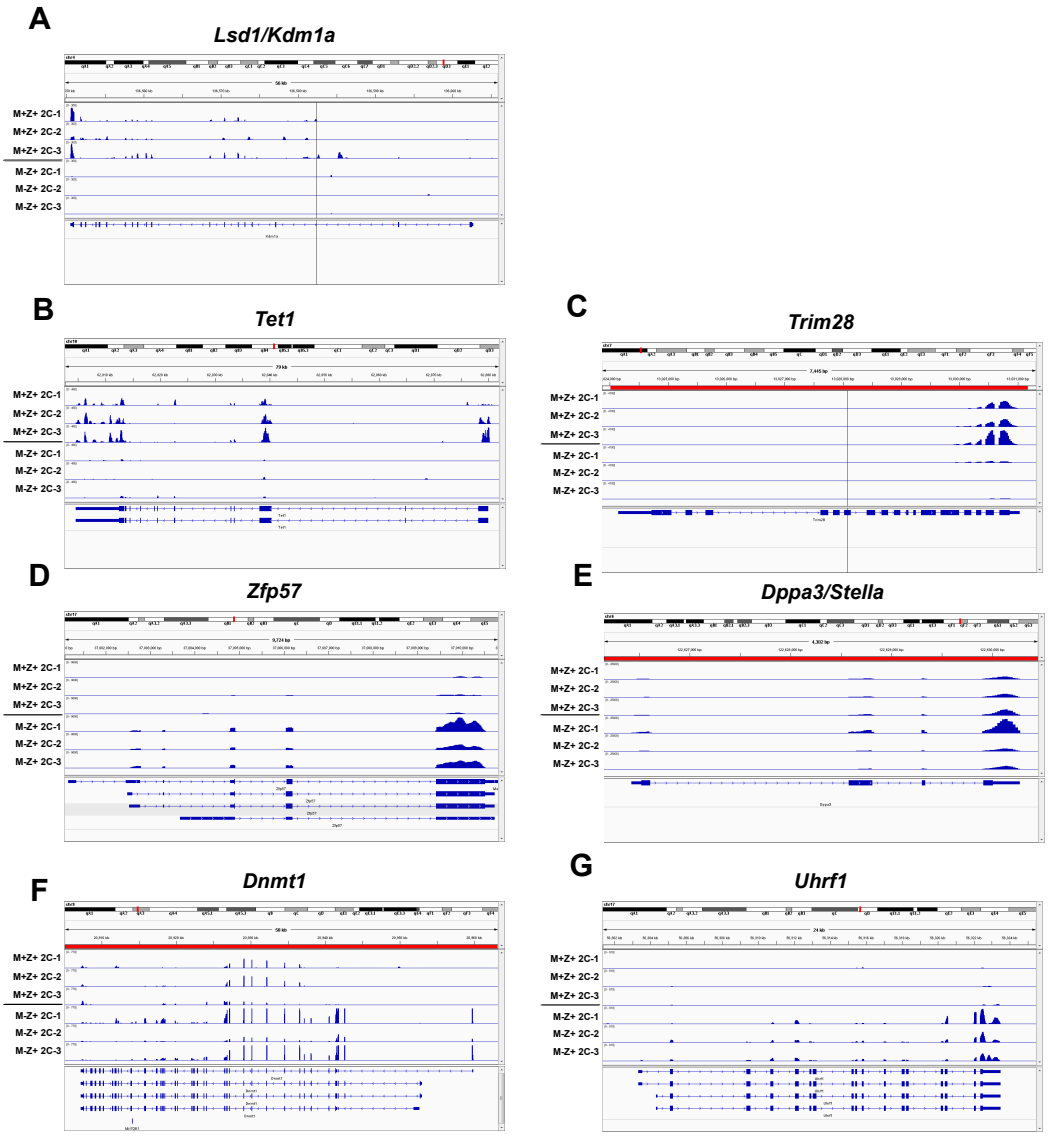


Figure 3-figure supplement 3:
Expression of Epigenetic regulators in *Lsd1*^{Zp3} 2C Embryos
Sequenced RNA-seq reads showing relative expression from *Lsd1*^{Zp3} M+Z+ and M-Z+ 2C embryos aligned to the genome for *Lsd1/Kdm1a*(A), *Tet1*(B), *Trim28*(C), *Zfp57*(D), *Dppa3/stella*(E), *Dnmt1*(F) and *Uhrf1*(G). Gene tracks visualized using Integrative Genomics Viewer.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

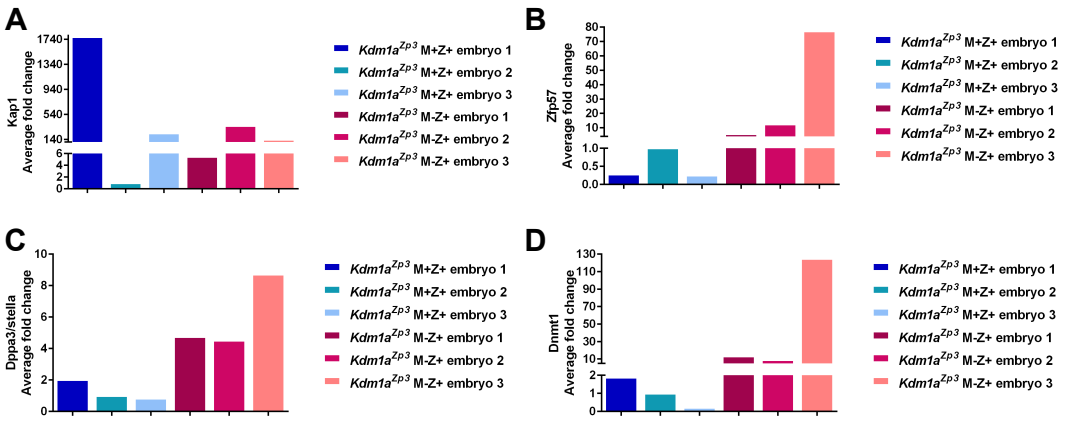


Figure 3-figure supplement 4:
Relative expression of epigenetic regulators in *Lsd1*^{Zp3} 2C embryos
Relative expression analysis of epigenetic regulators including *Trim28*(A), *Zfp57*(B) *Dppa3/stella*(C), and *Dnmt1*(D) in *Lsd1*^{Zp3} M+Z+ 2C embryos compared to *Lsd1*^{Zp3} M-Z+ 2C embryos. Y-axis represents average fold change. All gene expression was normalized to hprt expression.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

exclusively from the paternal allele (Figure 6C). However, in MEP1 *Zac1* is now expressed exclusively from the maternal allele, while in MEP2, the slightly lower level of inappropriate paternal methylation is associated with biallelic expression (Figure 6C). At *Impact*, the overall DNA methylation pattern is unchanged **(**Figure 6D), but there is a dramatic change in expression. *Impact* is normally expressed exclusively from the paternal allele, but in both MEP1 and 2, *Impact* is now expressed predominantly from the maternal allele (Figure 6F). In addition, in both MEP1 and MEP2, there is a significant decrease in the overall expression level of *Impact* (Figure 6E).

In addition to the altered imprinting at *Zac1* and *Impact*, we observe a general disruption of several additional imprinted genes. At three of these additional imprinted loci (*H19*, *Grb10* and *Igf2r*), DNA methylation is altered (Figure 6G and Figure 6-figure supplement 1A, C). For example, at *H19* DNA methylation is inappropriately lost on the paternal allele in MEP1 (Figure 6G). There is also a slight acquisition of inappropriate methylation on the maternal allele of both MEP (Figure 6G). At two additional imprinted genes (*Mest* and *Snrpn*), DNA methylation appears normal (Figure 6-figure supplement 1E, G). Furthermore, at three of the additional imprinted loci, we observe an overall decrease in gene expression. These include *Igf2r*, *Mest*, and *Snrpn* (Figure 6-figure supplement 1D, F, H), while at *H19*, expression levels are unaffected (Figure 6H). However, the allele-specific expression of all of these additional imprinted loci remains unaffected (data not shown). Finally, at *Grb10* MEP1 and MEP2 are affected differently. DNA methylation is normal in MEP1, but inappropriately lost from the paternal allele in MEP2 (Figure 6-figure supplement 1A). This is associated with no change in expression in MEP1, but a loss of expression in MEP2 (Figure 6-figure supplement 1B). However, it should be noted that the interpretation of methylation and expression patterns of *Grb10* is complicated by the fact

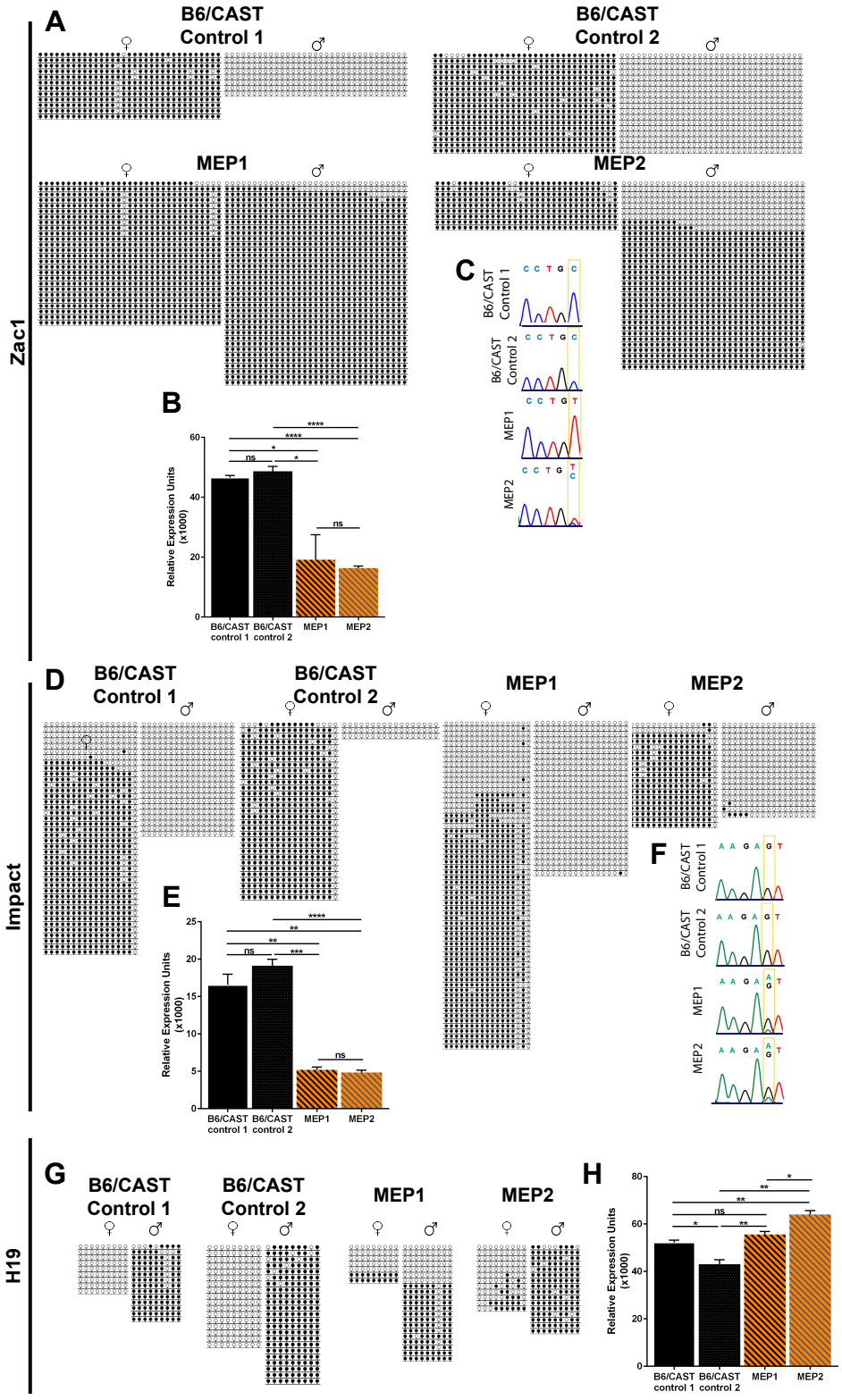

Zac1
A
B6/CAST Control 1
B6/CAST Control 2
MEP1
MEP2
B
Relative Expression Units (x1000)
B6/CAST control 1
B6/CAST control 2
MEP1
MEP2
C
Impact
D
E
F
H19
G
H

Figure 6.
Imprinting Defects in *Lsd1*^{Vasa} Progeny
(A,D,G) Allele-specific bisulfite analysis of *Zac1* (A), *Impact* (D), and *H19* (G). Each line represents the clone of an allele. Each circle represents a CpG dinucleotide where closed circles indicate methylation and open circles indicate no methylation. Maternal and paternal alleles are indicated. (B,E,H) Relative expression analysis of *Zac1* (B), *Impact* (E), and *H19* (H). Expression normalized to *β-actin*. Error bars indicate S.E.M. p-values calculated using an unpaired t-test with n.s. indicating $p>0.05$, * = $p<0.05$, ** = $p<0.005$, **** = $p<0.0001$. All asterisks indicate statistical significance. (C,F) Allele-specific expression of *Zac1* (C) and *Impact* (F). The polymorphic base is highlighted in yellow. For *Zac1*, the maternal allele SNP is T (red) in highlighted position and paternal allele SNP is C (blue) in electrophoretogram. For *Impact*, the maternal allele SNP is A (green) in highlighted position and paternal allele SNP is G (black) in electrophoretogram. All analyses were performed on 2 staged matched B6/CAST hybrid control pups and 2 maternal effect progeny (MEP) exhibiting perinatal lethality. p-values calculated using an unpaired t-test with n.s. indicating $p>0.05$, * = $p<0.05$, ** =$p<0.005$, *** = $p<0.0005$. All asterisks indicate statistical significance.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

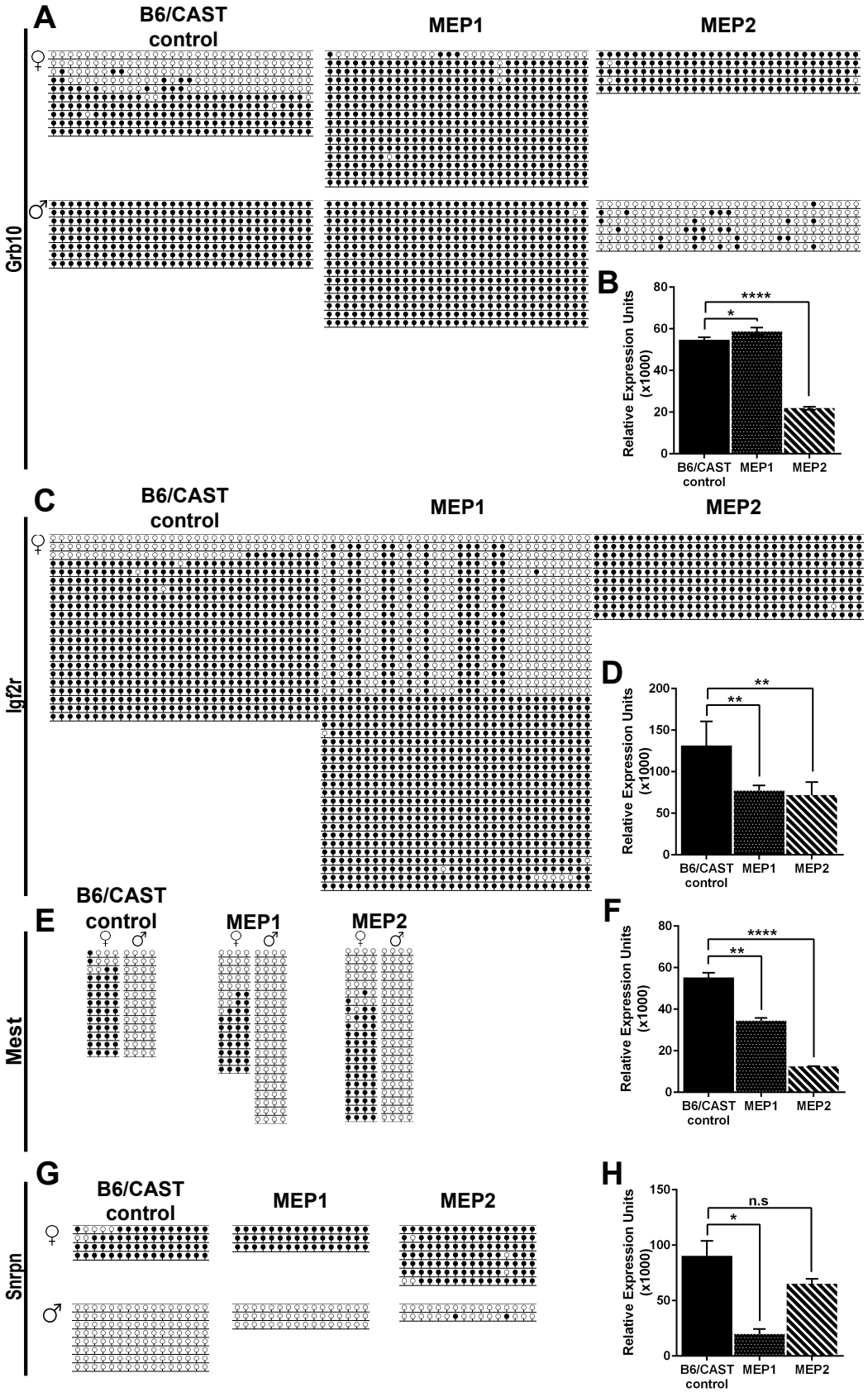

A
B6/CAST control
MEP1
MEP2
Grb10
♀
♂
B
Relative Expression Units (x1000)
0
20
40
60
80
*
****
B6/CAST control
MEP1
MEP2
C
B6/CAST control
MEP1
MEP2
Igf2r
♀
D
Relative Expression Units (x1000)
0
50
100
150
200
**
**
B6/CAST control
MEP1
MEP2
E
B6/CAST control
MEP1
MEP2
Mest
♀
♂
F
Relative Expression Units (x1000)
0
20
40
60
80
**
****
B6/CAST control
MEP1
MEP2
G
B6/CAST control
MEP1
MEP2
Snrpn
♀
♂
H
Relative Expression Units (x1000)
0
50
100
150
*
n.s
B6/CAST control
MEP1
MEP2

Figure 6-figure supplement 1:
Imprinting Analysis of *Lsd1*^{Vasa} Progeny
(A,C,E,G) Allele-specific bisulfite analysis of *Grb10* (A), *Igf2r* (C), *Mest* (E), and *Snrpn* (G). Each line represents the clone of an allele. Each circle represents a CpG dinucleotide where closed circles indicate methylation and open circles indicate no methylation. Maternal and paternal alleles are indicated. (B,D,F,H) Relative expression analysis of *Grb10* (A), *Igf2r* (C), *Mest* (E), and *Snrpn* (G). Expression normalized to *β-actin*. Error bars indicate ±S.E.M. p-values calculated using an unpaired t-test with n.s. indicating $p>0.05$, * = $p<0.05$, ** = $p<0.005$, **** = $p<0.0001$. All asterisks indicate statistical significance. All analyses were performed on a staged matched B6/CAST hybrid control pup and 2 maternal effect progeny (MEP) exhibiting perinatal lethality.

*eLife 2016;10.7554/eLife.08848. PMID:26814574

that *Grb10* switches from being expressed maternally to being expressed paternally in the brain[194]. Taken together, the changes in DNA methylation and expression indicate a disruption in the normal regulation of genomic imprinting.

2.3 DISCUSSION OF RESULTS

A major question in the field of epigenetics is whether histone methylation is regulated between generations, and whether potential defects in this process can lead to abnormalities in adults. To address this question, we asked if the maternal reprogramming function of LSD1, first demonstrated in *C. elegans*[84], is conserved in mammals. As in *C. elegans*, LSD1 is heavily deposited maternally in the mouse oocyte, suggesting that LSD1 could play a conserved role in maternal reprogramming (Figure 1). To address this hypothesis directly, we conditionally deleted *Lsd1* maternally in mouse oocytes using three separate *Cre* drivers. Deletion of *Lsd1* in developing oocytes via *Gdf9-Cre* or *Zp3-Cre* completely eliminates maternal LSD1 and results in embryonic arrest prior to the blastocyst stage, with the vast majority of embryos never developing past the 2C stage (Figure 1,2). Thus, maternal LSD1 is essential for the progression of mammalian embryogenesis.

The observed 1-2C embryonic arrest does not rule out an additional role for LSD1 in oocyte maturation, as we observe that a small fraction of LSD1 deficient oocytes are unfertilized, fertilized with multiple sperm, or display other gross morphological defects, including loss of the zona pellucida. However, in all cases the majority of oocytes are fertilized normally, and many embryos undergo cleavage to reach the 2C stage (Figure 2). Also, in hypomorphic $Lsd1^{Vasa}$ animals we observe defects on the paternal allele of imprinted loci (Figure 6). These results suggest a requirement for LSD1 post-fertilization in the zygote. In mammals, zygotic transcription initiates between the one and two cell stage[175,180,181]. As a result, the heterozygous offspring derived from our crosses likely begin to express LSD1 from paternal allele at the 2C stage. Taken together, these results pinpoint a requirement for maternal LSD1 reprogramming between fertilization and cleavage. Consistent with this

conclusion, the accompanying paper by Ancelin et al.[195] demonstrates that chemically inhibiting LSD1 function post-fertilization also results in a 2C arrest. This further suggests a zygotic requirement for LSD1 immediately post-fertilization.

Importantly, the observation that many maternally deleted embryos undergo cleavage, suggests that LSD1 is not just required for cell-cycle progression or cell viability (Figure 2). This conclusion is consistent with the observation that LSD1 null embryos survive until e7[184]. In addition, the clear difference in phenotype between maternally deleted embryos and embryonically deleted embryos, which survive past e7, clearly delineates the maternal requirement for LSD1 from the embryonic requirement.

To determine the potential cause of the LSD1 maternal embryonic arrest, we profiled the expression of *Lsd1*$^{fl/fl}$ oocytes and *Lsd1*^{Zp3} M-Z+ embryos that underwent cleavage to the 2C stage. Compared to control 2C embryos that have robustly activated the zygotic transcriptional program, *Lsd1*^{Zp3} M-Z+ 2C embryos retain the expression of nearly 2 out of 3 oocyte genes, and fail to activate approximately 4 out of 5 zygotically activated embryonic genes (Figure 3). This indicates that maternally deleted *Lsd1* embryos fail to undergo the MZT. Consistent with this finding, we observe relatively few transcriptional changes in *Lsd1*$^{fl/fl}$ oocytes (Figure 3). Taken together, these results suggest that LSD1 primarily enables the progression of embryogenesis by facilitating the MZT. This requirement for LSD1 in facilitating the MZT is also demonstrated, using a different *Lsd1* conditional allele, in the accompanying paper by Ancelin et al[195].

For reasons that are unclear, deletion of *Lsd1* maternally with *Vasa-Cre* results in an incomplete loss of LSD1 (Figure 1). Similar to *Lsd1*^{Zp3} and *Lsd1*^{Gdf9} M-Z+ embryos, many of the resulting *Lsd1*^{Vasa} M-Z+ embryos undergo early embryonic arrest, but some survive past

the 2C stage (Figure 4). This uncovers a hypomorphic effect due to partial loss of LSD1 maternal reprogramming (Figures 4,5,6). Amongst the embryos that survive past the 2C stage, some survive until birth. However, of these animals which are born, a significantly increased fraction die perinatally (Figure 4). Remarkably, these perinatal lethal pups are derived from crosses between *Lsd1*^{Vasa} mothers and wild-type fathers (M-Z+). This suggests that the increased perinatal lethality is due to hypomorphic maternal LSD1 in the mother's oocyte.

Subtle defects in maternal LSD1 reprogramming may also contribute to behavioral abnormalities later in life. To investigate this possibility, we performed a number of behavioral assays on surviving *Lsd1*^{Vasa} M-Z+ adult animals. These include monitoring food disappearance and bedding height to measure excessive food grinding, the marble burying assay to measure excessive digging behavior, and the open-field test, to measure anxiety. In all of these assays, we observe a maternal effect on behavior due to hypomorphic maternal LSD1 from *Lsd1*^{Vasa} mothers (Figure 5). This suggests that hypomorphic LSD1 in the mother's oocyte can lead to altered behavior in the adult.

To investigate the potential molecular nature of the heritable defects, we analyzed DNA methylation and expression at imprinted loci. Both DNA methylation and expression are altered at imprinted genes in perinatal lethal *Lsd1*^{Vasa} pups. In particular, we observe a striking disruption in the allele-specific expression of both *Zac1* and *Impact*, along with decreased expression and altered DNA methylation at other imprinted loci (Figure 6). This suggests that subtle defects in maternal LSD1 can alter the epigenetic landscape in a heritable fashion. Also, though LSD1 and LSD2 may have some functional overlap, the distinct

imprinting defects observed in each mutant demonstrate that they have functions that are not redundant.

There are two mechanisms that we can envision as potentially contributing to the observed imprinting defects. Loss of maternal LSD1 may lead indirectly to the disruption of imprinting due to the misexpression of genes that regulate imprinted loci. Alternatively, loss of maternal LSD1 could affect imprinted genes directly, due to the failure to regulate histone methylation during a critical window following fertilization. Evidence for an indirect mechanism can be inferred from the misregulation of several genes that are known to regulate imprinting in the *Lsd1*^{Zp3} M-Z+ 2C RNA-seq data set. For example, we observe a 2-fold increase in the expression of *Dnmt1* (Figure 3-figure supplement 3,4). We also observe the ectopic expression of *Uhrf1*, a DNMT1 cofactor that enhances its activity (Figure 3-figure supplement 3,4)[196]. Previously the overexpression of DNMT1 has been shown to result in *de novo* DNA methyltransferase activity[193]. It is possible that the increased DNA methylation on paternal *Zac1* alleles and/or maternal *H19* alleles in *Lsd1*^{Vasa} pups, could be due to the ectopic effects of transient DNMT1 overexpression in early embryogenesis. We also observe the misexpression of four other genes that are known to maternally affect DNA methylation at imprinted loci; *Tet1*, *Trim28*, *Zfp57* and *Dppa3/Stella* (Figure 3-figure supplement 3,4)[21,153,154,168,191]. The misregulation of these factors may also contribute to the observed disruption of imprinting. Alternatively, it is possible that the misregulation of these genes, or DNMT1, could contribute to the perinatal lethality and behavioral defects observed in *Lsd1*^{Vasa} progeny by affecting DNA methylation at non-imprinted loci.

In addition to potential indirect effects, LSD1 may directly affect imprinted genes by regulating histone methylation during a critical window following fertilization. Histone

methylation marks, particularly H3K9me3 and H3K4me3, mark active and silenced imprinted genes, respectively[197,198]. Previous studies have suggested that the time period immediately following fertilization is a critical window for epigenetic reprogramming events[177]. The accompanying paper by Ancelin et al.[195] demonstrates that loss of LSD1 maternally leads to global changes in both H3K4 and H3K9 methylation during this critical time period, thus providing a potential mechanism for imprint disruption. Consistent with this possibility, loss of LSD1 in ES cells has been shown to lead to genome-wide changes in both H3K4 and H3K9 methylation, as well as changes in the expression of a number of imprinted genes[45].

Global changes in H3K4 and H3K9 methylation could lead to the disruption of imprinting through the following potential mechanisms. H3K4me2 has been shown to block the docking and activity of the DNMT3a/3b *de novo* methyltransferase complex[34], and in maternal LSD2 mutants, the lack of H3K4me2 demethylation in the oocyte has been proposed to account for the failure to maternally acquire DNA methylation at the ICRs of genomically imprinted loci[40]. Thus, it is possible that some of the DNA methylation and expression defects observed in *Lsd1* mutants could result from failure to maintain DNA methylation during a critical window following fertilization, due to the failure to demethylate H3K4me2. Alternatively, LSD1 has been shown to demethylate H3K9 methylation when complexed with the androgen receptor[49], and a neuronal splice variant of *Lsd1* can demethylate H3K9 methylation when associated with *Supervillain*[54]. H3K9 methylation is highly associated with, and mechanistically linked to, DNA methylation[199], so defects in the regulation of H3K9 methylation could also contribute to the observed imprinting defects by affecting DNA methylation at ICRs. Importantly, some imprinted loci are regulated directly, while others are regulated indirectly[200]. Thus the combination of direct and indirect effects

outlined above, due to loss of maternal LSD1, could disrupt imprinting through a large range of combinations.

Intriguingly, defects in some of the critical maternally provided regulators that are misexpressed in the *Lsd1*^{Zp3} 2C RNA-seq data set have been associated with heritable defects. For example, *Tet1/Tet2* double mutants exhibit a partial perinatal lethality phenotype that is reminiscent of what we observe in *Lsd1*^{Vasa} progeny[21]. Also, mice that are haploinsufficient for *Trim28* display abnormal exploratory behaviors[201]. Furthermore, both *Zac1* and *Impact* have been associated with nervous system defects, and loss of *Zac1* paternally results in perinatal lethality[202–207]. Thus, defects in these genes could potentially contribute to the abnormalities that we observe in *Lsd1*^{Vasa} M-Z+ adults. In order to interrogate the heritable mechanisms that we observe in *Lsd1*^{Vasa} progeny further, we are engineering a hypomorphic maternal *Lsd1* allele. This allele will be used to generate larger numbers of hypomorphic animals that can be used to further elucidate the mechanism.

Taken together, our data highlight the fundamental importance of LSD1-mediated epigenetic reprogramming at fertilization. In addition, our results establish a mammalian disease paradigm where defects in early epigenetic reprogramming between generations can lead to defects that manifest later in development, including perinatal lethality and altered behavior(Model Figure 7).

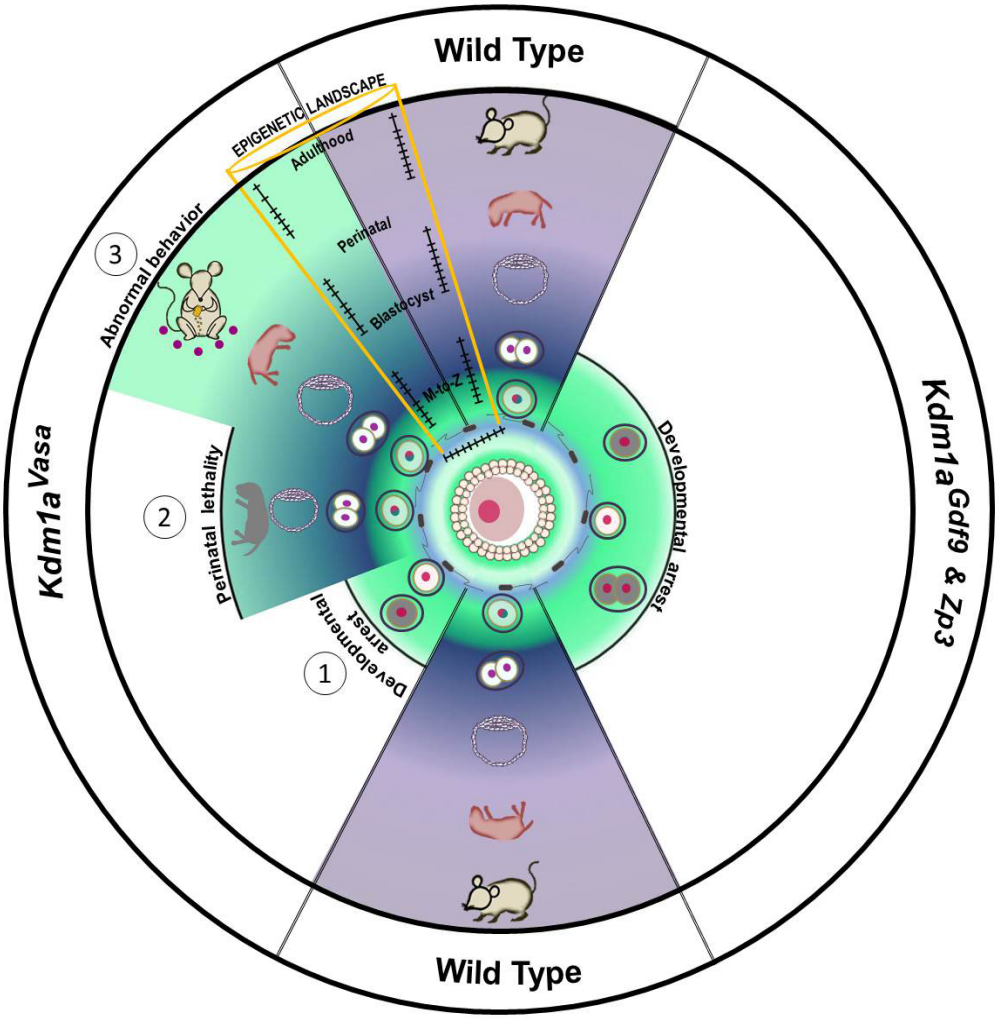


Figure 7.
Model: Loss of Maternal LSD1 Results in Defects Later in Development
In Wild Type oocytes, after fertilization (denoted by blue sperm encircling oocyte) the fertilized egg undergoes the maternal to zygotic transition (MZT; green to blue/purple) at the 1-2 cell stage. These M+Z+ embryos proceed normally through development (indicated by blastocyst, perinatal stage pup, and adult mouse). In contrast, when *Lsd1* is deleted with either *Gdf9*- or *Zp3-Cre*, the resulting $Lsd1^{Gdf9}$ and $Lsd1^{Zp3}$ progeny become arrest at the 1-2 cell stage and never undergo the MZT (green). When *Lsd1* is deleted with *Vasa-Cre*, we observe 3 hypomorphic outcomes in resulting $Lsd1^{Vasa}$ progeny: (1) developmental arrest at the 1-2 cell stage, (2) perinatal lethality and (3) abnormal behavior in surviving adult animals. These outcomes are due to reduced LSD1 in the mothers oocyte, suggesting that lowered maternal LSD1 can result in defects much later in development. These long-range outcomes are associated with imprinting defects (depicted as Wild Type versus mutant changes in DNA methylation within the yellow region).

*eLife 2016;10.7554/eLife.08848. PMID:26814574

2.4 APPENDIX: CHAPTER 2

This Appendix includes unpublished data that supports the data presented in Chapter 2.

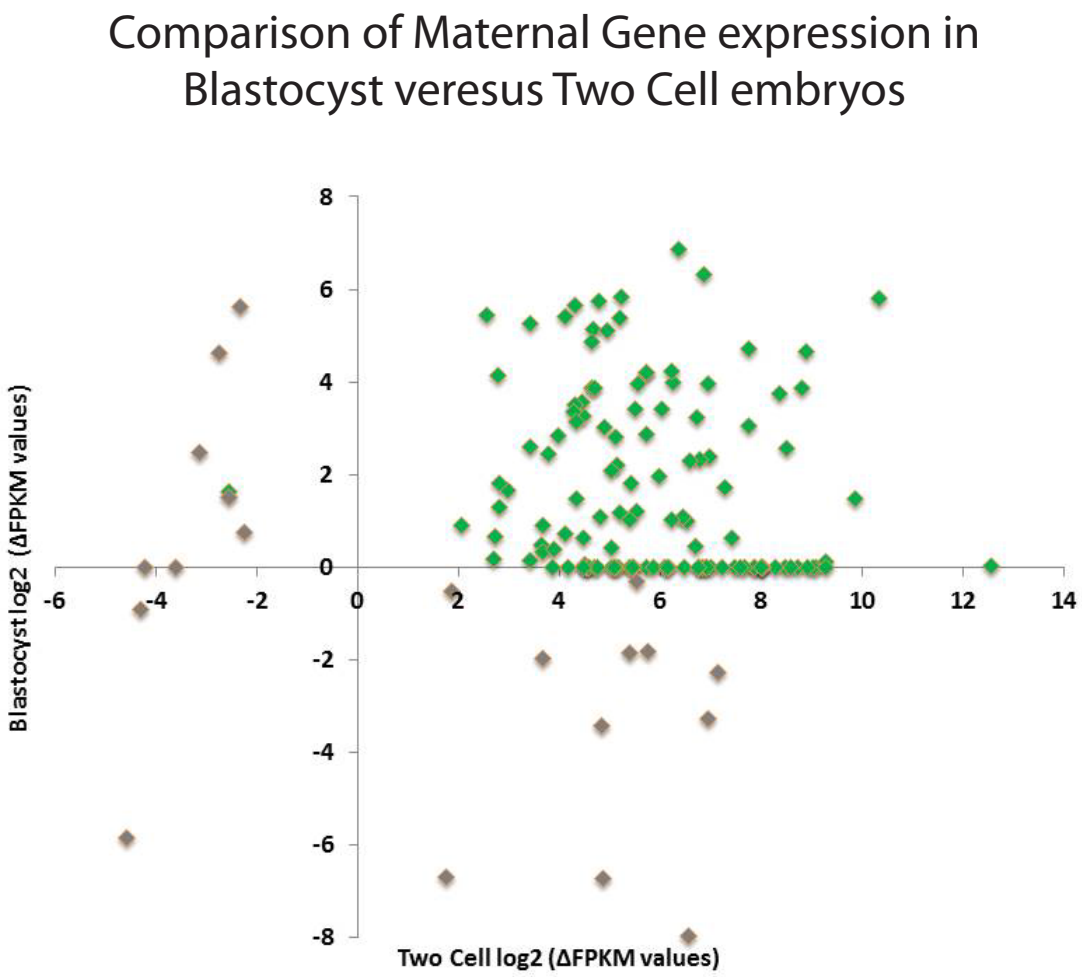


Figure 1: Meta-analysis of Blastocyst and 2C RNA-seq data for maternally expressed genes

We compared expression of 146 maternally expressed genes from two RNA-seq data sets. Data points are represented as $\log_2$(change in FPKM) from blastocyst RNA-seq data and two-cell embryo RNA-seq data.

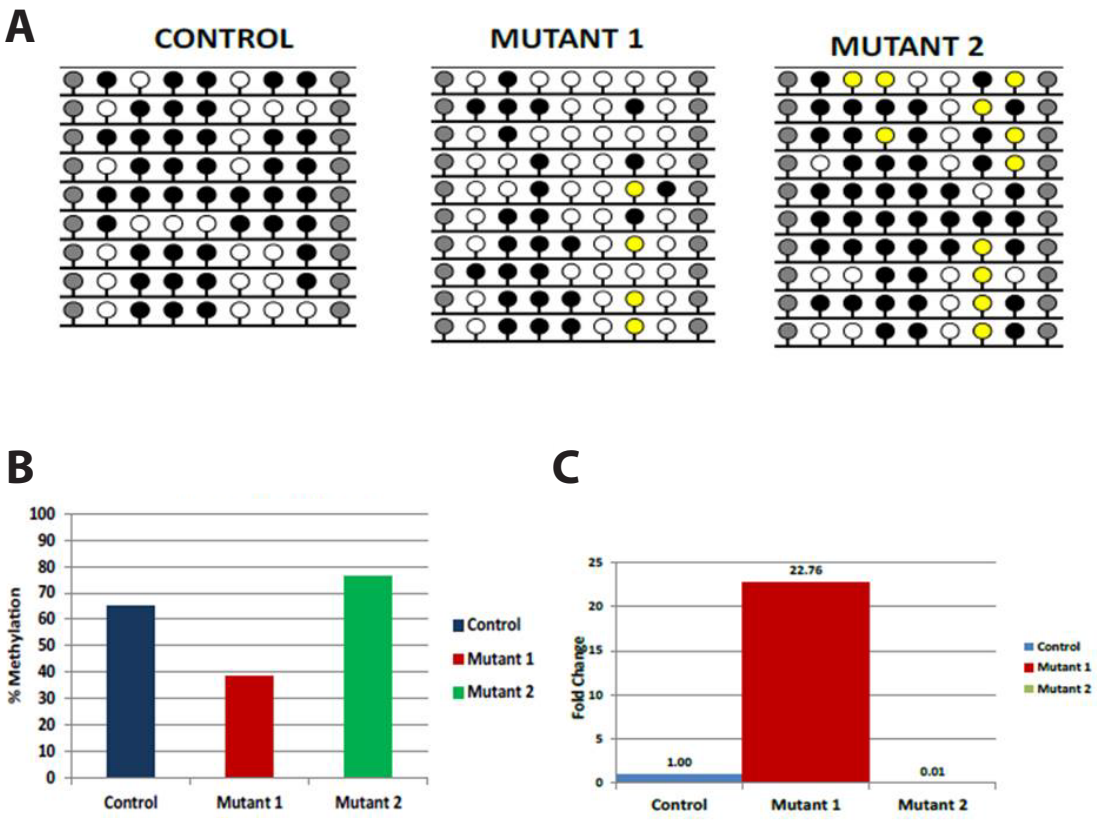
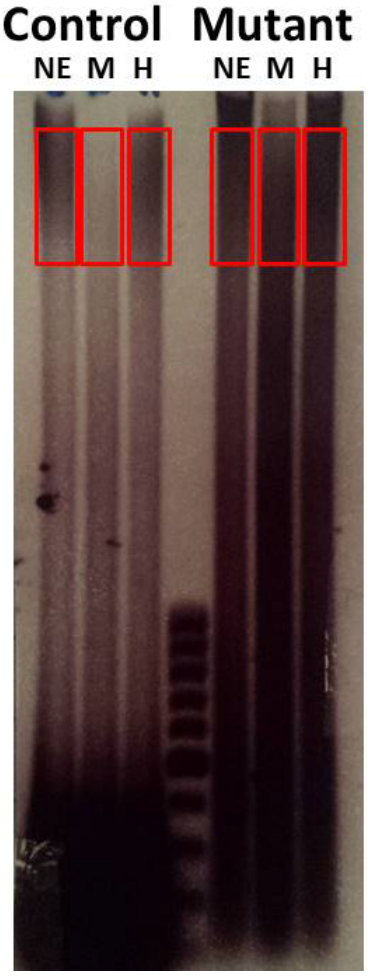


Figure 2: DNA methylation and expression alterations at *Oct4*
A. Lollipop diagrams representing DNA methylation status at the *Oct4* promoter. White circles indicate unmethylated cytosines; black circles indicate methylated cytosines; yellow circles represent mutated CpG dinucleotides. B. Bar graph representing percent methylation of *Oct4* promoter for each animal. C. Relative expression of *Oct4* in each animal. Fold change indicated above each bar.

A



B

		Band Intensity Values
Control	no enzyme	2180.25
	Msp	171.97
	HpaII	1367.25
Mutant	no enzyme	1180
	Msp	46.08
	HpaII	847.18

C

	Ratio	Percentage
Control	1.59	63%
Mutant	1.39	72%

Figure 3: Analysis of Global DNA metylation in perinatal lethality animals
A. Gel image of DNA from control P1.0 pups and maternal *Lsd1* mutant perinatal lethal pups was treated with either no enzyme, Msp or 5mC-sensitive HpaII to assess global DNA methylation. Red boxes indicated areas analyzed by Kodak image software for intensity. NE: no enzyme control; M: Msp; H: HpaII. B. Band intensity values for each control and maternal mutant animals. C. Ratio and percentage values calculated from band intensities for control and mutant samples.

CHAPTER 3:

A RESOURCE FOR ANALYSIS OF DNA METHYLATION STATUS OF GENOMICALLY IMPRINTED LOCI

2.1 Design of Resource for probing DNA methylation at imprinting control regions

Abstract

Genomically imprinted loci are expressed mono-allelically dependent upon the parent of origin. Their regulation and structure not only illuminates how the environment can interact with chromatin but also how chromatin can be reprogrammed every generation. Because of their distinct parent of origin regulation, analysis of imprinted loci can be difficult. Single nucleotide polymorphisms (SNPs) are required to accurately assess these elements allele-specifically. However, SNPs that have been reported lack robust verification thus making analysis of imprinting difficult. In addition, the few allele-specific imprinting assays that have been developed employ different mouse strains making it difficult to systemically analyze these loci. Here, we have generated a resource that will allow the systematic analysis of many significant imprinted loci in an allele specific manner. This resource includes comprehensive identification and verification of SNPs present within ten of the most widely used imprinting control regions and verified allele-specific DNA methylation assays for each gee in a C57BL6/CAST strain background.

Introduction

Public resources exist to facilitate experimental research by providing experimental animals, sorting information about particular genes, and cross referencing independent sources to integrate the information located in disparate locations. But the process of combing through these resources can be unnecessarily unwieldly. Assaying imprinted genes can be particularly hard when utilizing these resources due to their complex molecular nature. In addition, there is a large amount of information that must be gathered in order to effectively design experiments to interrogate them. Thus, we identified a need for a protocol that outlines the process of analyzing imprinted genes.

Genomically imprinted loci highlight how DNA methylation and chromatin structure can regulate gene expression. Many of the mechanism that regulate imprinted loci are involved in other contexts including cancer biology and stem cell reprogramming. Because of their distinct molecular nature, any alterations that are present at these loci can be indicative of other molecular defects.

Genomically imprinted genes are a specialized class of genes that are monoallelically expressed[140]. This single allele expression is dependent upon the parent-of-origin of that allele. Therefore some imprinted genes are solely expressed from the paternal allele while others are expressed only from the maternal allele. To date, approximately 150 imprinted genes have been identified in mice and about 100 in humans[208]. However, in recent years high-throughput techniques lead to estimates of as many as 1,000 imprinted loci[209,210]. Monoallelic expression of these genes is essential for normal mammalian development, which emphasizes the importance of these genes as models of epigenetic inheritance[211].

Imprinted genes were first hypothesized to exist based on experiments that sought to generate uniparental embryos. These androgenetic and parthenogenetic embryos did not survive past pre-implantation development and had serious developmental defects[212–215]. Thus, these experiments suggested that expression from one parental genome was insufficient for normal embryonic development. The first imprinted genes, which include *Igf2r*, *Igf2*, and *H19*, were discovered in 1991[216–219]. Prior to this time however, the human syndromes Beckwith-Wiedemann Syndrome (BWS) and Silver-Russell Syndrome (PWS), were described to have parent-of-origin effects[220,221]. Both of these syndromes, which are associated with intellectual disabilities and developmental defects, are caused by large-scale chromosomal abnormalities including uniparental disomy of chromosome 11 in BWS and a deletion of a portion of chromosome 15[220,221]. The chromosomal alterations highlight the functional significance of proper parental expression at these loci.

Imprinted genes tend to be organized on chromosomes in clusters which allows them to be regulated as a single unit[222]. The monoallelic expression of multiple imprinted genes is under the control of differentially methylated regions within the clusters called imprinting control regions (ICRs)[222]. The differential methylation of ICRs is established in the parental germline and maintained throughout development. ICRs are typically between 100 and 3700kb long and are CpG rich[182,223,224].. There are currently about twenty known ICRs that regulate the expression status of imprinted genes. Of these, three are paternally methylated while the remainder are maternally methylated[223]. Maternally methylated ICRs tend to exist in intragenic CpG islands while paternal methylated ICRs tend to be located intergenically[211]. Mechanistically, histone modifications including methylation of lysine 4 on histone 3 (H3K4me) is associated with active imprinted alleles, and methylation of lysine 9 on histone 3 (H3K9me) is associated with silenced imprinted alleles[211,225]. Also, large non-coding RNAs

have been implicated in the establishment of imprinted loci[226]. However, the methylation of ICR regions is largely accepted to be the primary imprint controlling the mono-allelic expression of imprinted loci[182].

DNA methylation in mammals occurs mainly in the context of CpG dinucleotides[227]. In the case of imprinted genes, the density of these CpG dinucleotides is enriched in ICRs[228]. The methylation status of the ICR is what determines the expression status of clusters of imprinted genes[228]. This DNA methylation at ICRs is protected during the large scale reprogramming events that occur after fertilization[120,229,230]. However, during germ cell specification imprints are erased and reestablished in a sex-specific manner[231,232]. The *de novo* DNA methyltransferase DNMT3A and its cofactor DNMT3L are responsible for the establishment of the differential methylation of ICRs in the mouse germline[143,233]. Loss of these enzymes affects DNA methylation of imprinted genes resulting in defective developmental phenotypes[124,233]. For example deletion of DNMT3A results in postnatal lethality around four weeks of age and infertility in both male and female mice[124]. Loss of DNMT3L affects DNA methylation not only at imprinted loci but also at retrotransposable elements[143]. This highlights the fact that many of the same elements that regulate imprinted genes also regulate other genomic loci. Thus detecting error in gene imprinting can inform other molecular alterations to the genome.

The gold standard in probing the DNA methylation status of any locus is bisulfite analysis[234,235]. This technique allows specific regions to be interrogated for their methylation status, making bisulfite analysis a great tool for probing imprinted genes. DNA is treated with sodium bisulfite which converts unmethylated cytosines to uracils. Methylated cytosines are protected from conversion. Following bisulfite conversion, bisulfite PCR is

performed to detect these base pair changes. Primers sets used for this amplification cannot contain any CpG dinucleotides because of the uncertainty of whether a cytosine base in the primer annealing sequence may be methylated. As a result, generating these bisulfite-specific primer sets in these high CpG density regions can be difficult. After, PCR amplification, each PCR product is cloned in order to generate enough material for accurate sequencing, a step that is necessary when detecting not only base pair changes but also for identifying parent of origin sequences. Each clone is subsequently sequenced and then mapped back to a non-bisulfite-converted genome for verification. Lollipop are then generated to show the methylation status of the region of interest. For imprinting control regions, the readout of the bisulfite analysis is a characteristic 50-50 methylated/unmethylated state where the actively expressed allele typically (but not always) unmethylated and the silenced allele methylated. Thus, it is important to know which allele came from which parent in order to determine any allele specific perturbations in DNA methylation. This requires the verification of polymorphisms that are present within the amplified regions.

C57BL/6 (hereafter referred to as B6) mice are the most commonly used mouse strain and were the first mouse strain to be fully sequenced[236]. In addition, other popular mouse strains including BALB/c and FVB, are derived from strains first cultivated by mouse fanciers[236]. Decades of inbreeding has led to a reduction in the prevalence of inter-strain polymorphisms between these popular strains. Thus, an inbred mouse strain derived from a different source must be used for them to be genetically distinct from traditional mouse strains to introduce SNPs in the genome. *Mus castaneus* (CAST) mice originate from a well-defined sub group of wild mice[236]. There is a 45% allelic difference between B6 and CAST mice, making hybrid progeny especially useful for analyzing imprinted loci[237].

The Database of Single Nucleotide Polymorphisms (dbSNP) is the result of a collaboration between the National Center for Biotechnology Information (NCBI) and the National Human Genome Research Institute (NHGRI)[238,239]. The dbSNP is a collection of genetic variation which reports SNPs that have been observed in various assays performed by individual researchers, private businesses, large consortiums, and large genome sequencing centers. It catalogs genetic variation in multiple organisms for the purpose of facilitating large genome-wide association studies. Entries in the dbSNP are very thorough and include information such as unique SNP identification numbers, flanking sequences, descriptions of the population in which the polymorphism was found, as well as occasionally polymorphism frequency[238,239]. However, there are several problems that dbSNP difficult to use. Since it is a public repository, many reported SNPs have not been additionally verified[240,241]. The information provided for particular strains of mice at one site may not be consistent across multiple polymorphic sites, especially in the context of ICRs. The database currently has no minimum requirement for allelic frequencies, which also contributes to the lack of verification for many SNPs. As a result, false positives have been reported at a rate of between 15-17% and many researchers have questioned the quality of the dbSNP[240,241]. In addition, to find particular gene information, one must know specific information such as gene identification number, accession number associated with the gene, or know the reference number for the SNP that is to be assayed. Thus, the database does not allow for straightforward interrogation of specific imprinted loci.

In this paper, we have provided a streamlined approach to assaying the methylation status of a number of the most studied imprinted genes including primer information and optimal conditions, verification of SNPs present in ICRs to allow ease of identification of

parental alleles, and sequence specific information of each SNP. This resource will help enable the systematic interrogation of many significant imprinted genes.

Results

In order to begin the process of interrogating specific imprinted loci, we generated a workflow to streamline the process (Figure 1). Our first criterion was to identify well-defined imprinting control regions (ICRs) that have been extensively studied in the literature. We focused on the following ICRs due to their prevalence in the literature: Grb10, Igf2r, Impact, Lit1, Mest/Peg1, Peg3, Peg10, Snrpn, and Plagl1/Zac1. These ICRs also had well-defined locations in the genome and are associated with differentially methylated regions (DMRs) which allowed for us to probe their methylation status via bisulfite analysis. We then utilized the UCSC Genome Browser in conjunction with dbSNP to determine reported SNPs within a 10kb window surrounding and including the ICRs. Following this *in silico* analysis, we designed bisulfite specific primers to the regions of interest. These regions were under 1kb and were within our 10kb defined window including a significant portion of the ICR and at least one SNP. These bisulfite primers could not contain any CpG dinucleotides, which reduced the availability of genomic regions to amplify. Bisulfite primers were optimized on bisulfite converted DNA using a twelve-step optimization protocol (detailed in Methods). After optimization, bisulfite PCR was performed on B6 female, Castaneus male, and on the hybrid progeny resulting from the mating. Each PCR product was processed to ensure purity. Processing included gel purification, TA cloning, clone amplification by culture, purification of transformed plasmid, digestion and isolation of cloned PCR product. This isolated cloned product was then sequenced. Cloned sequences were processed using AnnHyb software and subsequently mapped back to the genome using

BiQ Analyzer. All clones that were less than 95% bisulfite-converted or with a sequence identity less than 80% of original genomic sequence were not used. To determine the validity of each SNP, sequences were aligned using CLC sequence viewer. Reported SNPs were compared in B6 and Castaneus sequences. If validated in this initial comparison, further validation was performed via analysis of the methylation status in hybrid B6/CAST progeny.

Using this workflow, we validated SNPs in ten ICRs and identified PCR conditions for analysis of each. Grb10 is regulated by an ICR that is approximately 1.4kb and located on chromosome 11 in mouse (Figure 2A). We validated one SNP within a 391bp region (Figure 2A). The polymorphic base is A in the B6 background and a G in the Castaneus background (Figure 2B). Grb10 is methylated on the maternal allele and unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 2C). Within our probed region, we validated one SNP out of 11 reported SNPs from the dbSNP database (Figure 2D). Optimized PCR cycle information was also found (Figure 2E).

H19 is regulated by an ICR on chromosome 7(Figure 3A). We validated three SNPs within a 293bp region (Figure 3A). These polymorphic bases include: a G in the B6 background and a deletion in the Castaneus background; a G in the B6 background and an A in the Castaneus background; an A in the B6 background and a G in the Castaneus background (Figure 3B). H19 is methylated on the maternal allele and unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 3C). Within our probed region, we

validated one SNP out of 11 reported SNPs from the dbSNP database (Figure 3D). Optimized PCR cycle information was also found (Figure 3E).

Igf2r is regulated by an ICR on chromosome 17 (Figure 4A). In Igf2r, we validated two SNPs within a 522bp region (Figure 4A). These polymorphic bases include the following: a G in the B6 background and an A in the Castaneus background; an A in the B6 background and a G in the Castaneus background (Figure 4B). Igf2r is methylated on the maternal allele and unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 4C). Within our probed region, we validated two SNPs out of fourteen reported SNPs from the dbSNP database (Figure 4D). Optimized PCR cycle information was also found (Figure 4E).

Impact is regulated by an ICR on chromosome 18 (Figure 5A). We validated three SNPs within a 433bp region (Figure 5A). These polymorphic bases include: T in the B6 background and A in the Castaneus background; A in the B6 background and a G in the Castaneus background; T in the B6 background and an A in the Castaneus background (Figure 5B). Impact is methylated on the maternal allele and unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 5C). Within our probed region, we validated three SNPs out of twelve reported SNPs from the dbSNP database (Figure 5D). Optimized PCR cycle information was also found (Figure 5E).

Lit1 is regulated by an ICR on chromosome 7 (Figure 6A). We validated one SNP within a 338bp region (Figure 6A). The polymorphic base is G in the B6 background and an A in the Castaneus background (Figure 6B). Lit1 is methylated on the maternal allele and

unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 5C). Within our probed region, we validated one SNP out of eight reported SNPs from the dbSNP database (Figure 6D). Optimized PCR cycle information was also found (Figure 6E).

Mest is regulated by an ICR on chromosome 6(Figure 7A). We validated one SNP within a 136bp region (Figure 7A). The polymorphic base is T in the B6 background and a G in the Castaneus background (Figure 7B). Mest is methylated on the maternal allele and unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 7C). Within our probed region, we validated one SNP out of four reported SNPs from the dbSNP database (Figure7D). Optimized PCR cycle information was also found (Figure 7E).

Peg3 is regulated by an ICR on chromosome 7(Figure 8A). We validated one SNP within a 229bp region (Figure 8A). The polymorphic base is T in the B6 background and a G in the Castaneus background (Figure 8B). Peg3 is methylated on the maternal allele and unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 8C). Within our probed region, we validated one SNP out of thirteen reported SNPs from the dbSNP database (Figure8D). Optimized PCR cycle information was also found (Figure 8E).

Peg10 is regulated by an ICR on chromosome 6(Figure 9A). We validated one SNP within a 551bp region (Figure 9A). The polymorphic base is C in the B6 background and an A in the Castaneus background (Figure 9B). Peg10 is methylated on the maternal allele and unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 9C). Within our

probed region, we validated one SNP out of twenty reported SNPs from the dbSNP database (Figure9D). Optimized PCR cycle information was also found (Figure 9E).

Snrpn is regulated by an ICR on chromosome 7(Figure 10A). We validated two SNPs within a 359bp region (Figure 10A). These polymorphic bases include: T in the B6 background and an A in the Castaneus background; G in the B6 background and a T in the Castaneus background (Figure 10B). Snrpn is methylated on the maternal allele and unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 10C). Within our probed region, we validated two SNPs out of fifteen reported SNPs from the dbSNP database (Figure10D). Optimized PCR cycle information was also found (Figure 10E).

Zac1 is regulated by an ICR on chromosome 10(Figure 11A). We validated one SNP within a 584bp region (Figure 11A). The polymorphic base is A in the B6 background and a G in the Castaneus background (Figure 11B). Zac1 is methylated on the maternal allele and unmethylated on the paternal allele. This polymorphism was further confirmed with the observation of the proper methylation pattern in the hybrid progeny (Figure 11C). Within our probed region, we validated one SNP out of nine reported SNPs from the dbSNP database (Figure11D). Optimized PCR cycle information was also found (Figure 11E).

Of the SNPs that we analyzed we were able to validated 16, while we failed to validate 114 SNPs within those same regions. This SNP verification and statistically analysis of invalidated SNPs will help improve dbSNP utility and also highlight the need for studies such as these.

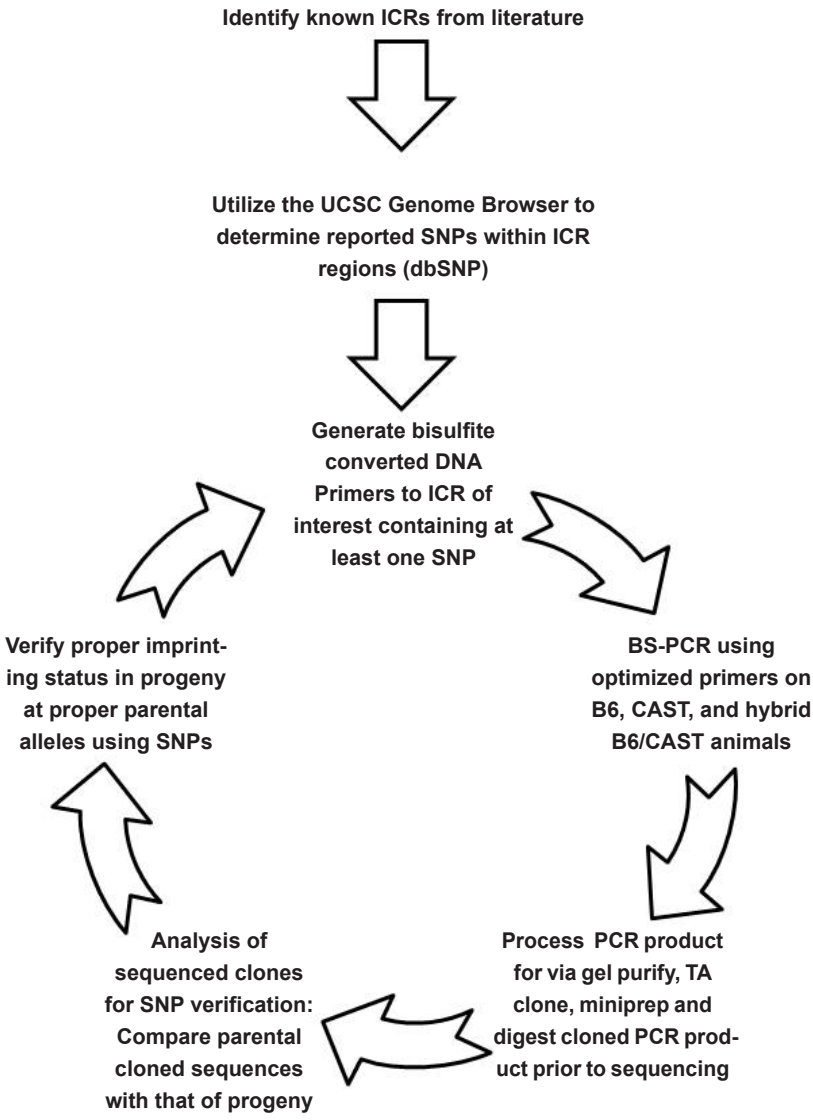


Figure 1: Workflow for SNP verification within ICRs
Known ICRs were first pulled from the literature followed by identificataion of putative SNPs present within each region. These SNPs then underwent a verification process trough bisulfite analysis of both parental and hybrid progreny strains. SNPs that were unverified resulted in repetition of the verification process.

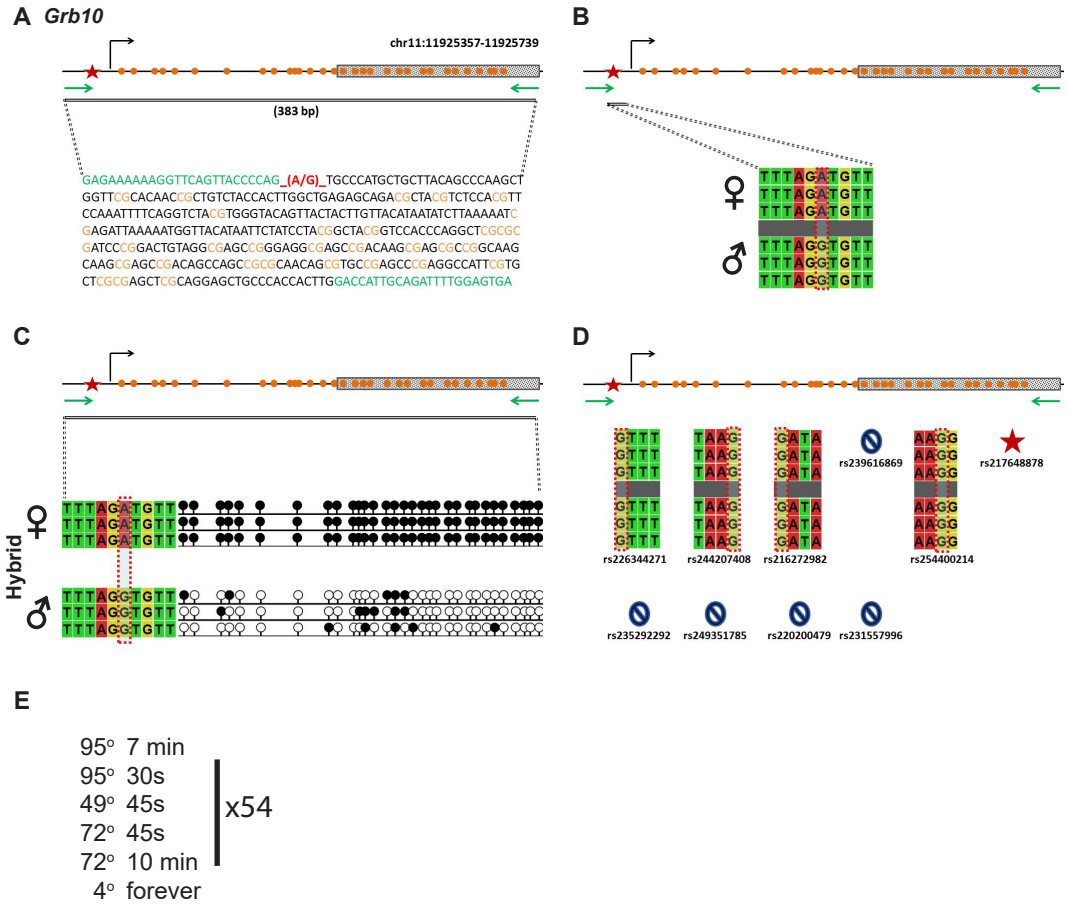


Figure 2: SNP verification within *Grb10* ICR

A. Schematic of *Grb10* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; Orange indicates CpG dinucleotides; Red star and bases indicate verified SNP. B. Verified SNP presented as sequences from B6 female and CAST male. A-to-G SNP is highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosine; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue crossed circle indicates C to T polymorphism that cannot be assayed in bisulfite analysis. E. Optimal PCR conditions for probed region.

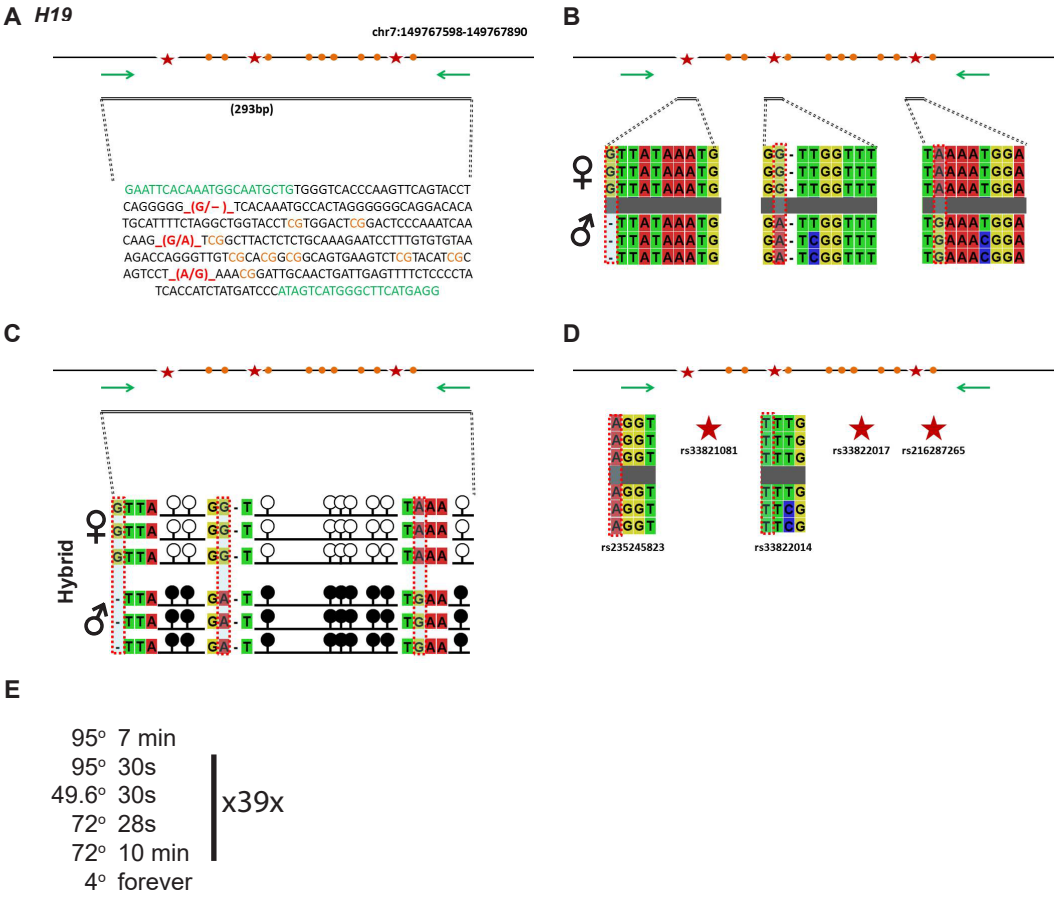


Figure 3: SNP verification within *H19* ICR

A. Schematic of *H19* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. Green indicates primer sequences; Orange indicates CpG dinucleotides; Red star and bases indicate verified SNP. B. Verified SNPs presented as sequences from B6 female and CAST male. G-to-del, G-to-A, and A-to-G SNPs are highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNPs highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red stars indicate validated SNPs. E. Optimal PCR conditions for probed region.

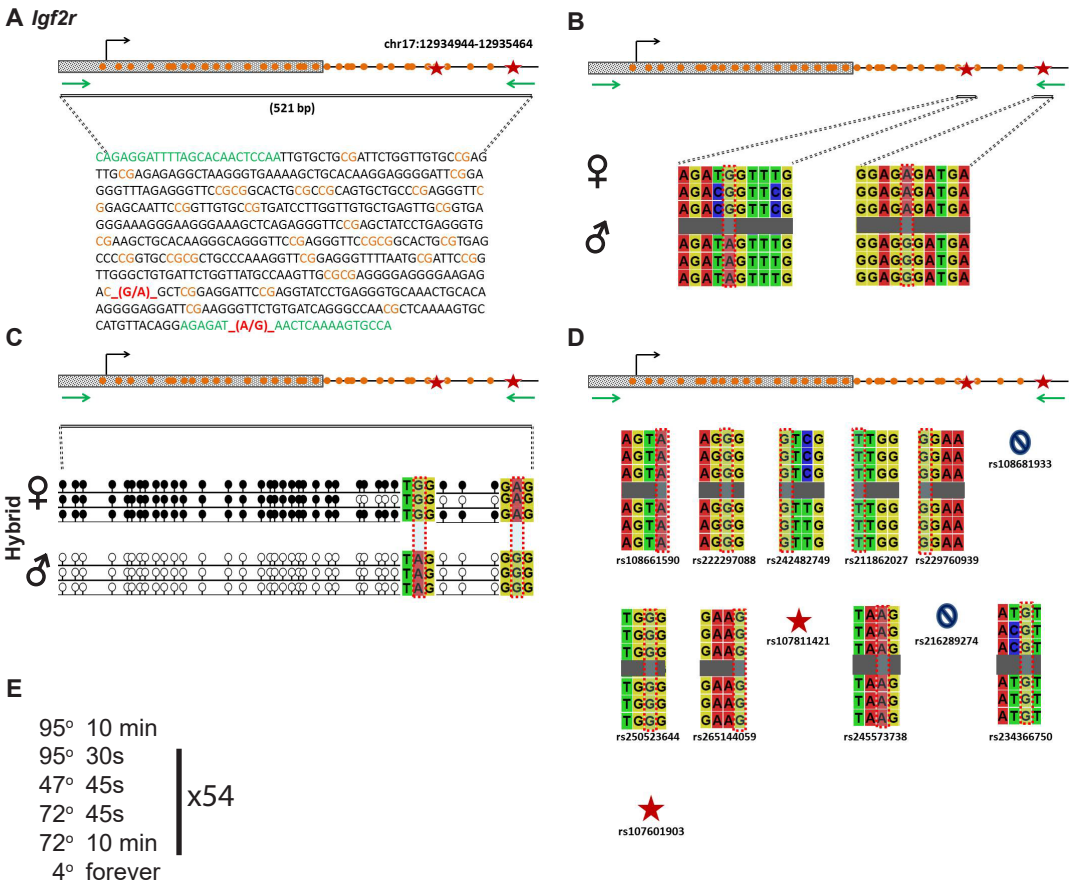


Figure 4: SNP verification within *Igf2r* ICR

A. Schematic of *Igf2r* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; Orange indicatees CpG dinucleotides; Red stars and bases indicate verified SNPs. B. Verified SNPs presented as sequences from B6 female and CAST male. G-to-A, and A-to-G SNPs are highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNPs highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red stars indicate validated SNP and blue crossed circle indicates C to T polymorphism that cannot be assayed in bisulfite analysis. E. Optimal PCR conditions for probed region.

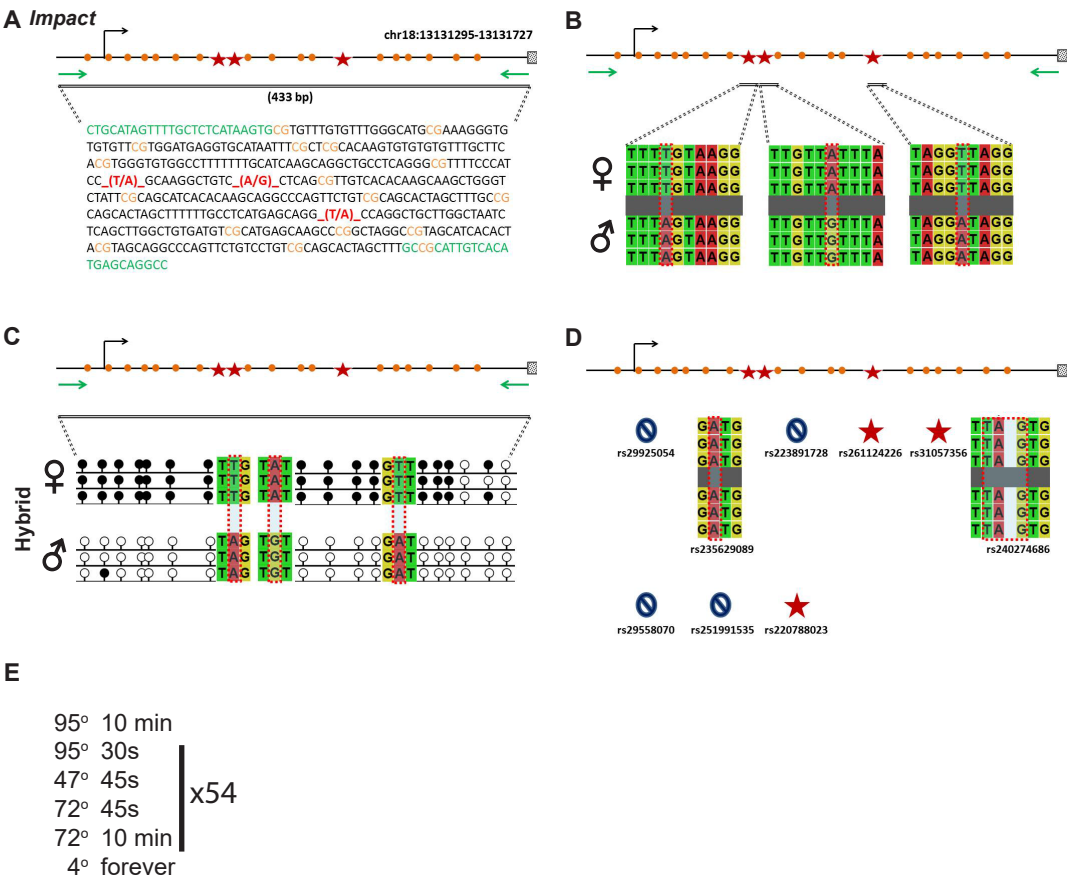


Figure 5: SNP verification within *Impact* ICR

A. Schematic of *Impact* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; Orange indicates CpG dinucleotides; Red stars and bases indicate verified SNP. B. Verified SNPs presented as sequences from B6 female and CAST male. T-to-A, A-to-G, and T-to-A SNPs are highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNPs highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red stars indicate validated SNP and blue crossed circle indicates C to T polymorphism that cannot be assayed in bisulfite analysis. E. Optimal PCR conditions for probed region.

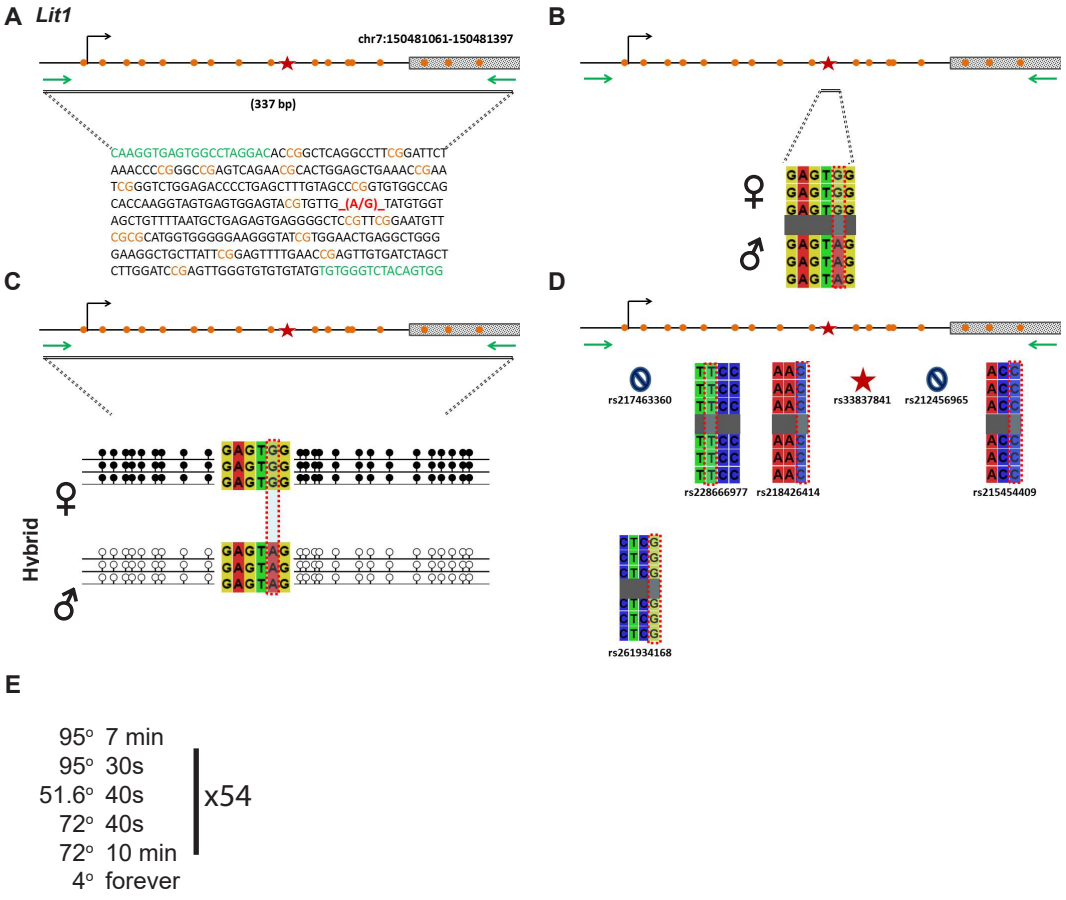


Figure 6: SNP verification within *Lit1* ICR

A. Schematic of *Lit1* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; Orange indicatees CpG dinucleotides; Red star and bases indicate verified SNP. B. Verified SNP presented as sequences from B6 female and CAST male. G-to-A SNP is highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue crossed circle indicates C to T polymorphism that cannot be assayed in bisulfite analysis. E. Optimal PCR conditions for probed region.

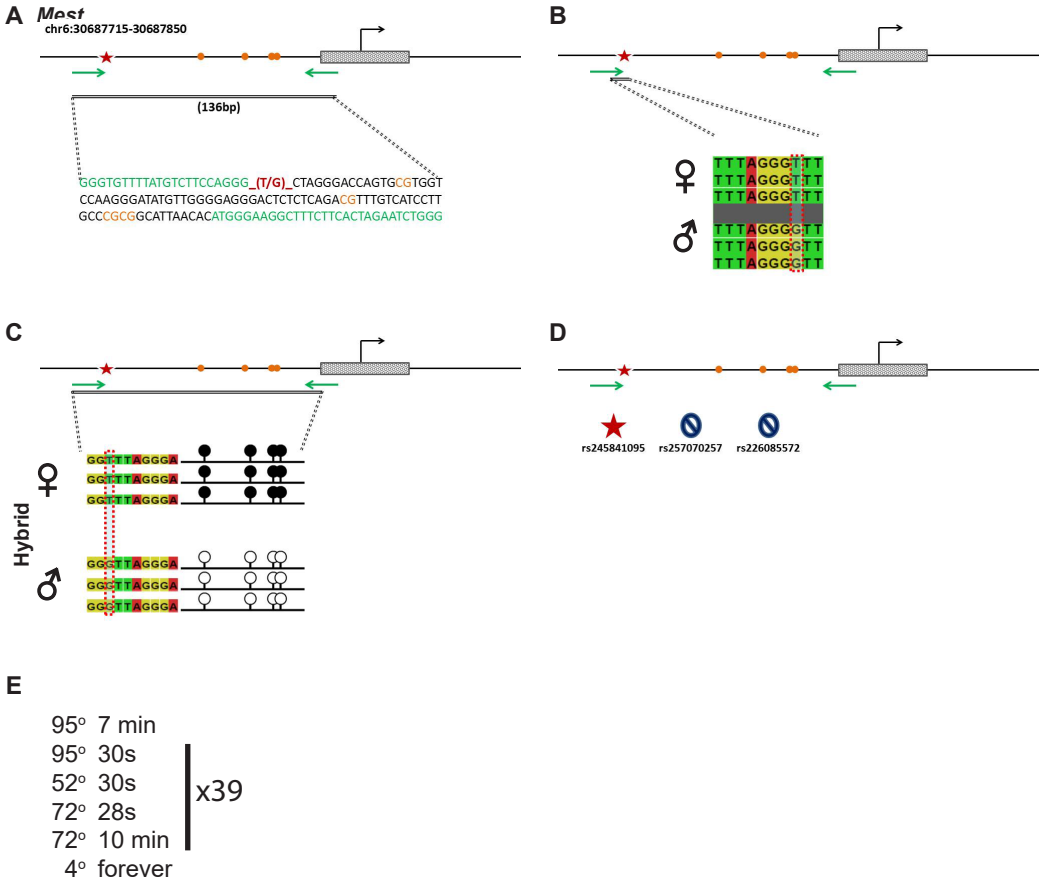


Figure 7: SNP verification within *Mest* ICR

A. Schematic of *Mest* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; Orange indicates CpG dinucleotides; Red star and bases indicate verified SNP. B. Verified SNP presented as sequences from B6 female and CAST male. T-to-G SNP is highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue crossed circle indicates C to T polymorphism that cannot be assayed in bisulfite analysis. E. Optimal PCR conditions for probed region.

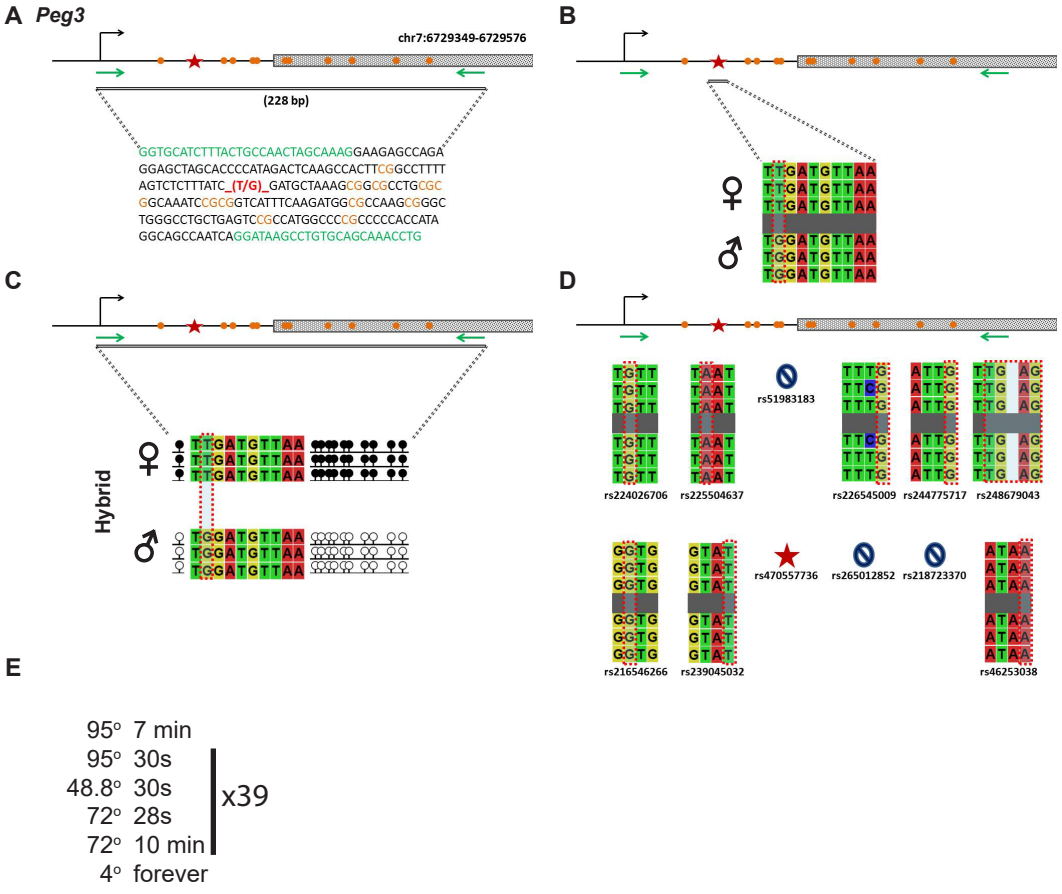


Figure 8: SNP verifcation within *Peg3* ICR

A. Schematic of *Peg3* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; Orange indicates CpG dinucleotides; Red star and bases indicate verified SNP. B. Verified SNP presented as sequences from B6 female and CAST male. T-to-G SNP is highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue crossed circle indicates C to T polymorphism that cannot be assayed in bisulfite analysis. E. Optimal PCR conditions for probed region.

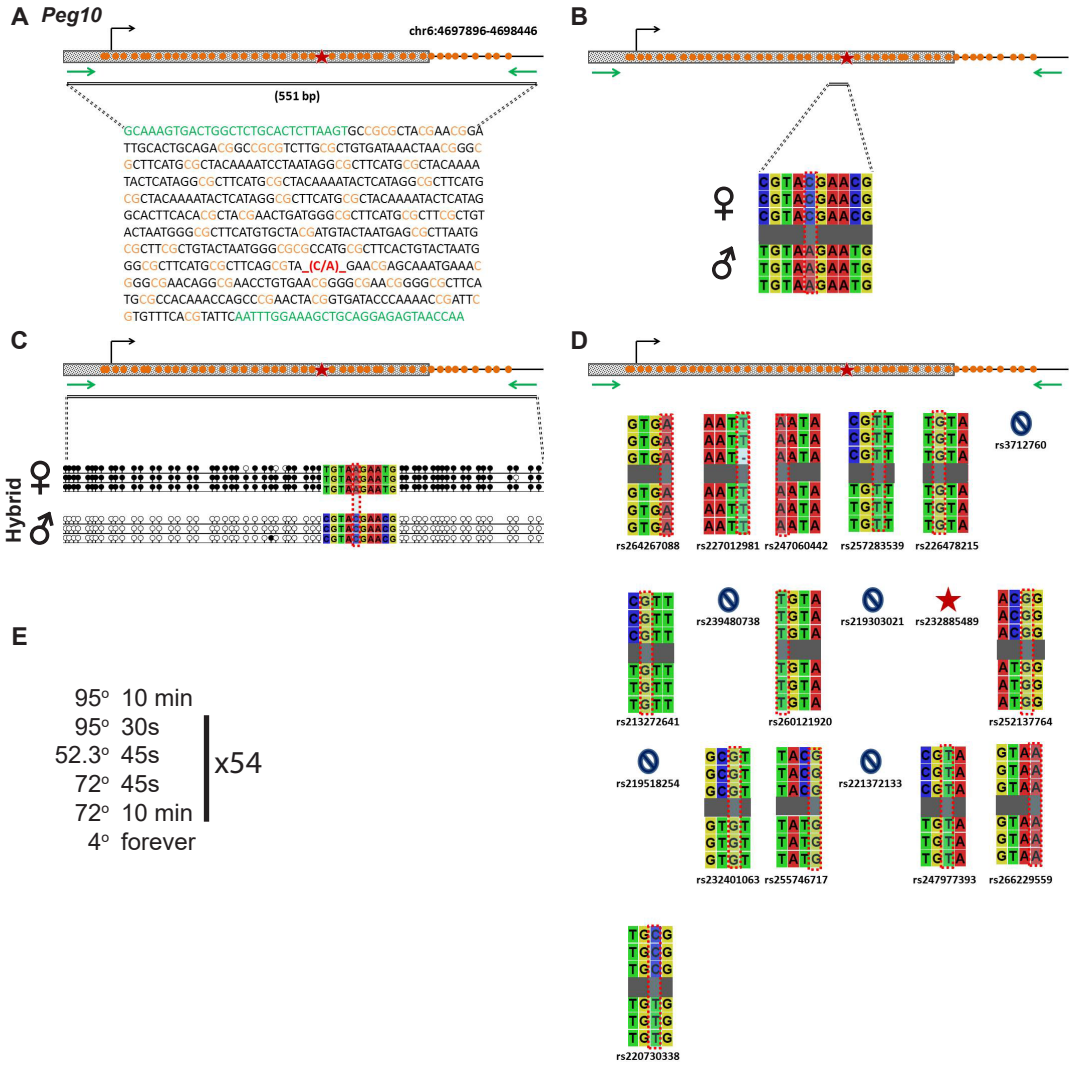


Figure 9: SNP verification within *Peg10* ICR

A. Schematic of *Peg10* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; Orange indicates CpG dinucleotides; Red star and bases indicate verified SNP. B. Verified SNP presented as sequences from B6 female and CAST male. C-to-A SNP is highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue crossed circle indicates C to T polymorphism that cannot be assayed in bisulfite analysis. E. Optimal PCR conditions for probed region.

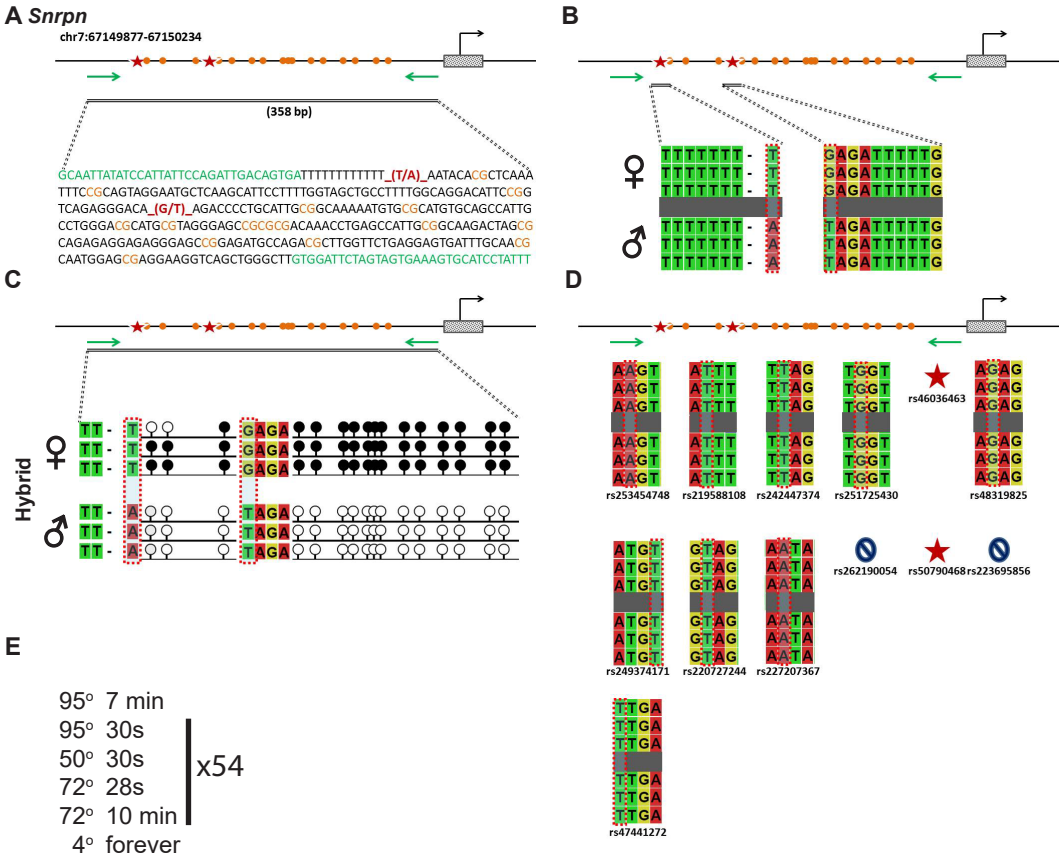


Figure 10: SNP verification within *Snrpn* ICR

A. Schematic of *Snrpn* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; Orange indicates CpG dinucleotides; Red star and bases indicate verified SNP. B. Verified SNPs presented as sequences from B6 female and CAST male. T-to-A, and G-to-T SNPs are highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP.Red star indicates validated SNP and blue crossed circle indicates C to T polymorphism that cannot be assayed in bisulfite analysis. E. Optimal PCR conditions for probed region.

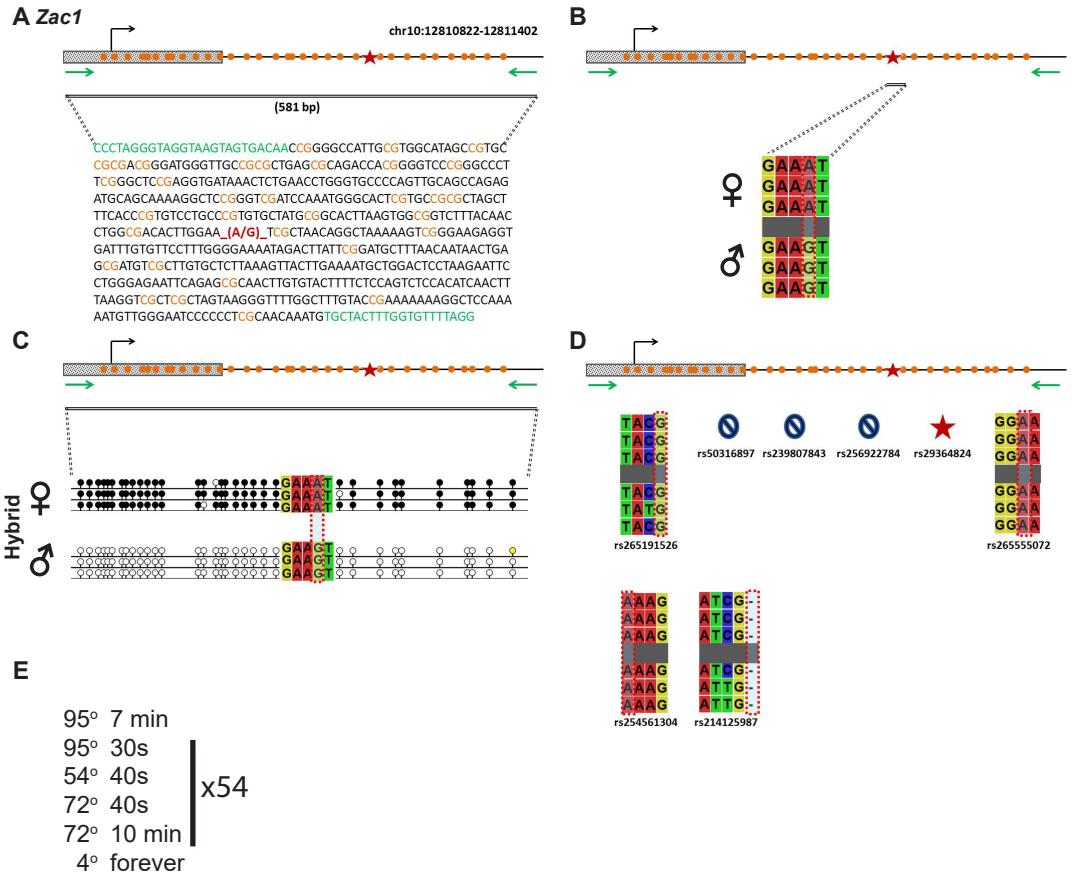


Figure 11: SNP verification within *Zac1* ICR

A. Schematic of *Zac1* imprinting control region. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; Orange indicates CpG dinucleotides; Red star and bases indicate verified SNP. B. Verified SNP presented as sequences from B6 female and CAST male. A-to-G SNP is highlighted by red dotted rectangle. C. Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. D. dbSNP reported SNPs that were invalidated within probed region highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue crossed circle indicates C to T polymorphism that cannot be assayed in bisulfite analysis. E. Optimal PCR conditions for probed region.

Concluding Remarks

This resource was designed to streamline molecular analyses at genomically imprinted loci. Imprinted genes have been demonstrated to be affected in many different disease states. For example, imprinted loci including *Zac1* and *Impact* have been demonstrated to be involved in neuronal development and have also been implicated in neurologically disorders[202,204–206]. Furthermore, imprinted genes disruptions have been associated with phenotypic alterations in mouse models. Upon deletion of *Dnmt3L,* the regulatory subunit of *de novo* DNA methyltransferases, there is loss of DNA methylation at a number of imprinted loci[143]. LSD2, a histone demethylase, results in a complete loss of DNA methylation at a subset of imprinted genes including *Zac1*, *Grb10* and *Mest*[40]. Loss of maternally contributed enzymes also results in DNA methylation defects in developing embryos. Both STELLA and ZFP57 act to protect imprinted loci from becoming demethylated during the epigenetic reprogramming that occurs at fertilization[153,167].

In addition, analyzing genomically imprinted loci can serve as a model of more general epigenetic perturbation. For example, DNA methylation patterns are significantly altered in cancers[90,92,242]. For example, in prostate cancer, the cyclin D2 promoter becomes hypermethylated which correlates with cancer progression[243]. As a result, the systematic analysis of multiple imprinted loci is increasingly important. The resource provided here will facilitate the analysis of some of the more commonly studied imprinted genes in any system.

2.2 METHODS AND MATERIALS

Generation of LSD1 Oocyte-specific Mutant Females

The following mouse strains were used: *Kdm1a/Lsd1*$^{fl/fl}$ MGI: 3711205[184], *Gdf9*^{Cre} MGI: 3056522[186] , *Ddx4/Vasa*^{Cre} MGI: 3757577[185], and *Zp3*^{Cre} MGI: 2176187[187] animals. To generate *Lsd1* oocyte conditional knockout mice, *Kdm1a/Lsd1*$^{fl/fl}$ females were crossed to transgenic *Cre* males to generate *Kdm1a/Lsd1*$^{\lrcorner/+}$ F1 animals with *Cre*. F1 males were crossed to transgenic *Cre* females to generate *Kdm1a/Lsd1*$^{\lrcorner/+}$ F2 animals with *Cre* and F2 males were mated to *Kdm1a/Lsd1*$^{fl/fl}$ females to generate *Kdm1a/Lsd1*$^{\lrcorner/+}$ F2 animals with *Cre*. These maternally deleted *Kdm1a/Lsd1*$^{\lrcorner/\lrcorner}$ females were then mated to wild-type B6 or *M. castaneus* males to produce maternal effect progeny. Prior to the initial crosses, *Kdm1a/Lsd1*$^{fl/fl}$ females were mated to C57BL/6 mice for several generations so that the genetic background is mostly B6. For *Gdf9*^{Cre} and *Zp3*^{Cre}, *Kdm1a/Lsd1*$^{\lrcorner/+}$ F1 animals with *Cre* were mated directly with *Kdm1a/Lsd1*$^{fl/fl}$ females. All mouse work was performed under the approved guidelines of the Emory University, NIH and Salk Institute IACUC. Polymorphisms were identified through the UCSC Genome Browser and subsequently verified through sequencing of parental strains and hybrid progeny. (See Table 1)

Isolation of Pre-implantation Mouse Embryos

To establish phenotypes in LSD1 maternal mutants, timed matings were set up between control/mutant females and wild type males. Superovulation was found to have little effect on phenotypes, so we used superovulation to collect enough embryos for RNA-seq. For super-ovulation, PMSG was injected into sexually mature females on day 1. After 48 hours, females were injected with HcG and subsequently housed with stud males. Confirmation of natural or superovulatory matings on subsequent days was made via observation of a

Gene name	Use (Bisulfite or qPCR)	Verified SNP (rs#)	Primers
Zac1	Allele-Specific Bisulfite Analysis	A-B6/ G-Cast (rs29364824)	For- GGGTAGGTAAGTAGTGATAA
			Rev- CCTAAAACACCAAAATAACA
	qPCR	T-B6/C-Cast (rs33583472)	For- CATTTGTAGGCATGCCCGTC
			Rev- GTGGTAGCTGCATCTGGGGCTGGA
Impact	Allele-Specific Bisulfite Analysis	A-B6/G-Cast (rs31057356)	For- TTGTATAGTTTTGTTTTTATAAGTG
			Rev- AACCTACTCATATAACAATACAAC
	qPCR	A-B6/ G-Cast (rs31052361)	For- GAAGAAAACTGAAGAGGTTG
			Rev- GCATAGATGTTGTGGGTGGC
H19	Allele-Specific Bisulfite Analysis	G-B6/ A-Cast (verified from Bartolomei Lab)	For- ATTTATAAATGGTAATGTTGTGG
			Rev- CCTCATAAAACCCATAACTATAAAATC
	qPCR		For- CCACTACACTACCTGCCTCAGAATCTGC
			Rev- GGTGGGTACTGGGGCAGCATTG
Grb10	Allele-Specific Bisulfite Analysis	A-B6/ G-Cast (rs217648878)	ForMat- GAGAAAAAAGGTTTAGTTATTTTAGA ; ForPat- GAGAAAAAAGGTTTAGTTATTTTAGG
			Rev- TCACCTCCCAAAATCTACAATAATC
	qPCR		For- GCTTGATGATCCTGTGAGAC
			Rev- TGCTCCTGTACCAAAACTAT
Igf2r	Allele-Specific Bisulfite Analysis	G-B6/ A-Cast (rs107811421)	For- TAGAGGATTTTAGTATAATTTTAA
			Rev- TAACACTTTTAAATTCATCTCT

	qPCR		For- CTGGAGGTGATGAGTGTAGCTCTGGC
			Rev- GAGTGACGAGCCAACACAGACAGGTC
Mest	**Allele-Specific Bisulfite Analysis**	**G-B6/ T-Cast (rs245841095)**	ForMat- GGGTGTTTTATGTTTTTTAGGGT; ForPat- GGGTGTTTTATGTTTTTTAGGGG
			Rev- CCCAAATTCTAATAAAAAAAACCTTCCCAT
	qPCR		For- GCTGGGGAAGTAGCTCAGT
			Rev- TTTCTTCTTAGCAAGGGCCA
Snrpn	**Allele-Specific Bisulfite Analysis**	**A-B6/T-Cast (rs50790468)**	ForMat- GTAATTATATTTATTATTTTAGATTGATAGTGAT ; ForPat- GTAATTATATTTATTATTTTAGATTGATAGTGAG
			Rev- ATAAAATACACTTTCACTACTAAAATCC
	qPCR		For- TGCTCGTGTTGCTGCTACTG
			Rev- GCAGTAAGAGGGGTCAAAAGC
β-actin	**qPCR**		For- GTGACGAGGCCCAGAGCAAGAG
			Rev- CGTACATGGCTGGGGTGTTGAAGG

Table 1.

Allele-Specific Primers and Polymorphisms

Each primer and polymorphism used for allele-specific analysis bisulfite analysis and allele-specific expression analysis.

copulation plug. 1-2 cell embryos were flushed from ovarian tract at embryonic day 1.5, morulae on e2.5 and blastocysts on e3.5. Flushed embryos were categorized and imaged.

Immunofluorescence

Isolated mouse ovaries were fixed for 1 hr in 4% PFA on ice then washed with PBS multiple times over a 2hr period. Ovaries were allowed to sit in 30% sucrose solution at 4°C overnight and were then embedded in OCT compound. 10 micron cryosections were obtained for analysis. Immunostaining was performed using rabbit polyclonal anti-LSD1 (1:200, Abcam, ab17721) and Alexa fluor conjugated secondary antibodies.

Immunohistochemistry

Isolated mouse ovaries were fixed overnight in 4% PFA. Ovaries were then dehydrated in the following series of steps: 70% ethanol for 20 minutes 3 times; 85% ethanol for 45 minutes 2 times; 95% ethanol for 1 hr; 100% for 1 hr; xylenes overnight; xylenes:paraffin mix for 2 hrs twice under vacuum; paraffin for 4 hrs under vacuum; and paraffin under vacuum overnight. Oocytes were embedded in paraffin and 10 micron sections were taken for analysis. Immunostaining was performed using rabbit polyclonal anti-LSD1 (1:500, Abcam, ab17721). Oocytes were scored for presence of LSD1 signal with qualitative comparison to wild-type oocyte signal.

Bisulfite Analysis and Bisulfite-PCR optimization

DNA was isolated from sagittal sections of each perinatal pup. Mouse embryo DNA obtained from the DNA prep was quantified using a nanospectrometer. Using 400ng of DNA, bisulfite conversion was done according to Zymo EZ DNA Methylation Kit© protocol. Embryo DNA was amplified via PCR in a 15ul reaction, 3ul was saved for

subsequent TA cloning, and remaining reaction was run on a 1% agarose gel. Primer sets and polymorphisms used are listed in Table 1. A twelve-step optimization protocol was used determine optimal PCR conditions for each primer set according to conditions in Table 2. The TA cloning reaction was set up using 3ul of fresh PCR product, 1ul of salt solution, 1ul of sterile water and 1ul of TOPO vector and allowed to incubate for 5 minutes at room temperature. For each TA cloning reaction, 4ul of the reaction was added to one vial of One Shot© *E. coli* and incubated on ice for 5 minutes. The cells were then heat shocked for 30 seconds at 42°C and the reaction was put on ice for an additional 5 minutes. LB-Ampicillin plates were pre-warmed and plated with 80ul of Xgal as a β-galactosidase substrate. To each plate, 60ul of the TA cloning reaction was added and allowed to incubate overnight at 37°C. From each plate, 10-15 white colonies were picked and cultured in 3ml of LB Amp on a shaker overnight at 37°C. Cell cultures were miniprepped according to Qiagen QIAprep kit© protocol. A portion of each miniprep was digested with EcoR1 in a 15ul reaction. Each reaction was then run on a 1% agarose gel and examined for correct product. The BiQ Analyzer program was used in the analysis of bisulfite converted sequences[244].

Methylation-Sensitive Restriction Enzyme Digest for Gross Analysis of Global DNA methylation levels

Concentration of DNA from experimental animals was determined by nanospectrometer. 20μg of DNA was digested with either Msp (methylation insensitive restriction enzyme) to give view of complete DNA digestion, HpaII (methylation sensitive restriction enzyme to give relative methylation state, or no enzyme as a control. Digested DNA was ran on 0.8% agarose gel and images were obtained over a course of four hours. Gel images with analyzed using Kodak imager software. Each gel lane was defined with consistent "banded" regions.

1.5mM $MgCl_2$ 0% DMSO	1.5mM $MgCl_2$ 1.5% DMSO	1.5mM $MgCl_2$ 5% DMSO
2.5mM $MgCl_2$ 0% DMSO	2.5mM $MgCl_2$ 1.5% DMSO	2.5mM $MgCl_2$ 5% DMSO
3.5mM $MgCl_2$ 0% DMSO	3.5mM $MgCl_2$ 1.5% DMSO	3.5mM $MgCl_2$ 5% DMSO
4.5mM $MgCl_2$ 0% DMSO	4.5mM $MgCl_2$ 1.5% DMSO	4.5mM $MgCl_2$ 5% DMSO

Table 2: Twelve-Step PCR Optimization Protocol
Twelve different conditions with varying levels of DMSO and Magnesium Chloride used to determine optimal condition for maximum PCR primer efficiency.

Band intensity values were obtained for each lane. To get relative ratio of DNA methylation status, each no enzyme control band intensity was divided by HpaII band intensity. To get relative percent DNA methylation, each HpaII band intensity was divided by no enzyme control band intensity.

Quantitative Real-time PCR Analysis

RNA was isolated from sagittal sections of each perinatal pup using Trizol. SuperScript® III first-strand synthesis system was used to generate cDNA. Under the following cycling conditions: 95°C for 3min, 95°C for 15sec, 60°C for 30sec, 72°C for 30sec, 50 cycles.

Genome-wide Expression Analysis

RNA-seq on LSD1 mutant and wild type 2C embryos was performed as described[44]. Briefly embryos were lysed in Prelude Direct Lysis buffer (Nugen) and amplified cDNAs were prepared using the Ovation RNA-seq system V2 (Nugen). Paired end libraries were prepared according to the Tru-seq library construction protocol starting with Covaris fragmentation step. Libraries were then sequenced on an Illumina Hi-seq 2000. For expression analysis, only first mate pair was used and reads were trimmed from 3' end to 50bp. Trimmed reads were filtered to a minimum average base quality of 15. We combined the knownGene, ensGene and refGene annotations for mm10 (downloaded from UCSC Genome Browser) with the full RepeatMasker annotation, also from UCSC Genome Browser, to build a single gene annotation. Redundant transcripts were filtered out using the gffread utility packaged with Cufflinks. The annotation was modified to include a common gene_id value for same-name repeat elements resulting in a total of 875 repeats. Sequences for all annotated features were extracted from the mouse genome and RNA-Seq reads were aligned to them using BBMap with 95% identity, up to a single INDEL and up to 2000 equally best alignments per

read. Alignments to features were quantified by counting hits to gene loci and down-weighting read alignments that mapped to multiple gene loci by 1/N^2 where N is equal to the number of gene loci. Reads aligned to repeat elements were counted in the same way but instead of using gene loci the repeat name was used for binning hits. If a read mapped to multiple repeat features all of the same name then it was counted as 1 hit to that repeat name. Raw read counts were loaded into R and median normalized for differential expression analysis. For each pair-wise test genes with raw counts less than 10 in all conditions were not tested. Differential expression was performed by testing the null hypothesis that a gene's fold change between conditions is zero. The observed dispersion-mean relationship was fit using a method similar to that of DESeq and the predicted dispersions were used as the minimum dispersion per gene. Fold-changes between conditions were tested by monte-carlo simulation using the observed and estimated means and dispersions to generate random negative binomial distributed count values (the rnbinom method in R). In each monte-carlo iteration N samples were simulated per condition (N = to the number of samples in each condition), averaged into conditions and the log fold-change was calculated. The resulting simulated log fold-change distributions are normally distributed. Z-scores were computed for each observed log fold-change compared to its corresponding monte-carlo simulated fold-change distribution and translated to two-tail p-values against the normal probability distribution. Raw p-values were adjusted using the Benjamini & Hochberg correction. Genes and repeats were tested together and marked significant if the adjusted p-value was less than 0.05. RNA-seq data is deposited at GEO (GSE66547). Oocyte expression data used for comparisons was previously deposited at GEO (GSE33923.) For the heat map plot, each gene found to be significantly associated with PC1 is standardized relative to its mean and scaled relative to its variance (i.e. each gene

is scaled separately to improve visualization.) Each row is a gene which is standardized where mean is zero with standard deviation is 1. This makes each row relative to its mean. The scale is relative to its variance. Rows are sorted by a clustering of the genes. Color scale is relative to mean expression level per gene.

Characterization of Food-grinding Behavior in *Lsd1*^{Vasa} Adult Progeny

Each animal was placed in a mouse housing unit (32.8 x 18.6 x 13.6cm) with 5/8 inches of bedding. Over an 8 day period, the height of the bedding was measured in inches. In addition, 400 grams of standard mouse diet pellets was placed in each food hopper. The amount of food remaining in the food hopper was weighed each day over a 3 day period.

Marble Burying Assay

A clean, transparent plastic cage (32.8 x 18.6 x 13.6cm) was filled 4.5cm deep with corncob bedding material. 20 glass marbles (20mm diameter) were evenly spaced in a 5x4 grid on the surface of the bedding. During the testing phase, the mice were placed in the cage for 25 minutes and allowed to explore. At the end of the testing phase, mice were removed from the cage and the number of marbles that were buried 2/3 their height in the bedding were counted.

Open Field Test

Mice were placed in a clean, transparent plastic cage (55.5 x 32.5 x 19.5cm) with grid lines marked on bottom of cage. The center area was clearly marked for analysis (27 x 16.2cm). After 10 minutes, video was scored for number of times center area was crossed and amount of time spent in center area.

CHAPTER 4:

LSD1 MATERNAL FUNCTION AND ITS EFFECT ON DEVELOPMENT

4.1 EPIGENETIC REPROGRAMMING AT FERTILIZATION AND THE ROLE OF LSD1 IN THIS PROCESS

In this dissertation, I have discussed how epigenetic information – information associated with DNA independent of sequence– can be inherited, and how its inheritance affects phenotypic outcomes from one generation to the next. At fertilization, maternal effect genes modulate this epigenetic information to create a competent chromatin environment to ensure that proper development occurs. Here, I described a maternal reprogramming function of the histone demethylase LSD1. Complete loss of the maternal pool of LSD1 (via GDF9-Cre) resulted in a two-cell arrest in resulting progeny. These arrested embryos are unable to undergo the maternal to zygotic transition. These results demonstrated that LSD1 is required for proper development. Partial loss of maternal LSD1 (via Vasa-Cre deletion) resulted in pre-implantation arrest embryos. We performed RNA-seq analysis on embryos that survived to the blastocyst stage and maintained expression on maternal genes. We further observed a small number of these embryos that survive to birth. Surviving animals exhibited higher levels of perinatal lethality which was associated with alterations in DNA methylation at imprinted loci. Animals that survive to adulthood exhibited abnormal behaviors including excessive digging and increased anxiety. These behaviors were reminiscent of psychiatric disorders including obsessive-compulsive disorder and autism.

Taken together, this work provides evidence that not only is LSD1 reprogramming essential for development, but that even slight disturbances in reprogramming result in far-reaching consequences in the life of an animal.

4.2 LSD1 IN THE OOCYTE

We have reported minimal oocyte defects based on our transcriptional analysis of wild type oocytes and *Lsd1* deficient oocytes: where only about 280 genes were upregulated in *Lsd1* mutant oocytes and about 190 were downregulated. This role of LSD1 in mouse oocytes has also been addressed in recent work from the Kelsey group and the Chen group. The Chen group reported defects in germinal vesicle breakdown and in meiotic progression[245]. They went on to further show that cell cycle regulators CDK1 and CDC25B were affected by loss of *Lsd1*, contributing to the meiotic phenotype. A transcriptional analysis of these oocytes reveals changes in transcription that are consistent with our data where they observe 367 genes upregulated in *Lsd1* deficient oocytes and 252 genes were downregulated. This demonstrates consistent changes resulting from *Lsd1* loss in mouse oocytes. It is possible that the effects we have observed could be due to compounding effects of slight oocyte defects and loss of LSD1 reprogramming function post-fertilization. However, there appears to be distinct differences between the role of LSD1 in oocytes and the role of LSD1 post-fertilization. In a report by Edith Heard's group, the developmental arrest in embryos derived from *Lsd1*-null oocytes was recapitulated by chemical inhibition of LSD1 post-fertilization *in vitro*[195]. This work, in addition to our work, demonstrates a clear role for LSD1 post-fertilization. It remains to be seen whether LSD1's function post-fertilization also influences its role during embryonic development.

4.3 LSD1 AND THE REGULATION OF DEVELOPMENTAL TIMING

Post-fertilization, LSD1 is not expressed until embryonic day 3.5, during the blastocyst stage when the first cell fates are being defined. Post-implantation, LSD1

expression is restricted to the epiblast, which specifies all three germ layers[246]. The maternal to zygotic transition begins in minor waves at the one-cell stage but continues until the blastocyst stage when all maternally derived transcripts have been cleared from the embryo. Based on the function of LSD1 post-fertilization, I propose that a reduction in this function at fertilization could alter the developmental timing of embryonic LSD1 expression in an auto-regulatory fashion, due to a less efficient reprograming from lower levels of LSD1 (Figure 1). In embryonic stem cells and *in vivo*, LSD1, in association with CoREST and HDAC1/2, regulates the timing of expression of key developmental regulators during early development. A few of these genes include Hox genes and genes involved in tissue specification[246]. In addition, LSD1 appears to regulate the stability of CoREST. Furthermore, LSD1 is involved in a regulatory feedback loop with miR-137 and TLX[50]. miR-137 represses LSD1 expression by binding to its mRNA while LSD1 in complex with TLX represses the expression miR-137[50]. Based on this data, I propose that LSD1 may regulate its own expression. In doing so, this creates a model where LSD1 may affect not only the developmental clock but can also have downstream effects on other targets of its binding partners (Figure 1).

Homozygous *Lsd1*-null mice die by embryonic day 6.5, indicating that the embryonic role of LSD1 occurs during the specification of the first cell fates[43,63]. LSD1 functions during cell fate transitions and is absolutely necessary for differentiation to occur. Since germ layer specification occurs during this window, we hypothesize that homozygous *Lsd1* null embryos die because of a differentiation failure. It would be intriguing to determine if partial loss of LSD1 function post-fertilization has an effect on the zygotic expression of LSD1. In reduced maternal LSD1 animals, I propose that developmental timing could be altered, thus also affecting differentiation of the germ layers. For example, the embryonic

lethality we observed in animals from reduced maternal LSD1 function could be explained by an inability to fully undergo germ layer differentiation, which might also explain the more subtle defects we observed. If LSD1 does in fact influence developmental delay at this stage, it can ultimately affect the timing or robustness of its own expression. Based on this, it is possible that defects in cell fate specification could arise from maternal LSD1 modulating the expression of embryonic LSD1.

4.4 IMPLICATIONS OF HYPOMORPHIC LSD1 FUNCTION

In this work, we examined the effects of reducing LSD1 function at fertilization, but did not address what the consequences would be of reducing LSD1 activity throughout development. In a 2011 paper, two point mutations present in LSD1 caused a significant reduction in LSD1 catalytic activity[247]. Cells expressing these LSD1 hypomorphs had significantly higher levels of H3K4me1/2 and LSD1 had a reduced ability to associate with both CoREST and HDAC1. When these mutations were studied *in vivo*, the consequences of reduced embryonic LSD1 function were elucidated. Animals expressing this mutated LSD1 die perinatally; none survived to adulthood[247]. This parallels what we observed with reduced maternal LSD1, where many of the animals that survived any early embryonic lethality eventually died perinatally (although we did see that a few animals survived to adulthood). In reduced embryonic LSD1, the animals that died perinatally exhibited severe heart defects, which we did not examine in our reduced maternal effect LSD1 animals. It would be interesting to compare the phenotypes observed in each of these reduced LSD1 function cases. Do the hypomorphic embryonic LSD1 animals exhibit any methylation defects? Alternatively, do our reduced maternal effect animals exhibit any of the same physiological phenotypes? Investigating these questions would give insights into how much of the

maternal effect phenotype affects the embryonic function of LSD1 itself, as discussed in section 4.2. The major difference between the phenotypes resulting from these studies could be explained based on the timing of LSD1 function. I propose that the major role for embryonic LSD1 is first in germ layer specification, and then in cellular differentiation. Thus, the phenotypes, including the embryonic lethality, observed in the hypomorphic embryonic LSD1 are due to a loss of this function. In addition, reduction of LSD1 maternally results in separate but similar phenotypes due to the function of LSD1 in epigenetic reprogramming at fertilization. I propose that this reprogramming is very sensitive to levels of LSD1 activity, which why we see a range of phenotypes in our reduced maternal LSD1 mutants that have varying levels of LSD1 knockdown. Animals with perinatal lethality presumably had enough LSD1 reprogramming at fertilization to survive to term, but misregulation at that stage propagated an altered epigenetic landscape that ultimately caused lethality. However, we cannot rule out the possibility that this altered epigenetic landscape could also affect the activity of embryonic LSD1, which I propose as an alternative mechanism for LSD1 reprogramming defects.

4.5 POTENTIAL HYPOMORPHIC EFFECTS OF LSD1 DEFECTS IN HUMAN PATIENTS

Recently, two papers were published discussing *de novo* mutations in LSD1 found in three human patients. These were the first mutations in *Lsd1* to be discovered in humans[248,249]. It is not surprising that human defects in *Lsd1* are rare, considering that mice homozygous for *Lsd1* die during embryogenesis, indicating that LSD1 is essential for proper development. Mice that are heterozygous for *Lsd1* exhibit no obvious phenotypes, suggesting that LSD1 function is not dosage dependent. Thus, these human mutations are unusual in that they

have resulted in viability.

Patients with *Lsd1* mutations exhibited severe developmental delay, changes in brain structure and cognitive function, and other minor abnormalities including and craniofacial abnormalities[248,249]. It must be noted that the minor phenotypes associated with the three patients did not affect all patients, indicating some pleiotropic effect of these *Lsd1* mutations. This is very similar to what we observed in our mice, including complete penetrance in behavioral changes, but variable alterations at the molecular level. Our work raises the possibility that the defects present in these patients could have been propagated from early embryonic development, even from defects in reprogramming at fertilization.

Each of the three mutations in *Lsd1* were all missense mutations resulting in amino acid substitutions at highly conserved residues[249]. Furthermore, all mutations were in the amine-oxidase domain of LSD1, which acts as the catalytic domain[249]. It is likely that the mutations were dominant negative in nature which affects the function of LSD1 that comes from the non-mutated allele. However, we do not know if heterozygous *Lsd1* humans are phenotypically normal like with heterozygous mice. A follow-up study by a group in Italy characterized the effects of these mutations on the function of LSD1[250]. LSD1 proteins containing each mutation were able to correctly fold and maintain stable interactions with CoREST though each had a reduced affinity for both HDAC1/2 and a histone H3 tail peptide. However, *in vivo,* it was observed that LSD1 stability was reduced by these mutations. Demethylase activity also differed between assays, where it was impaired *in vitro*, but still able to repress gene transcription *in vivo*, albeit at a lower level. This work suggests that in human patients, these mutations affect LSD1 in a hypomorphic fashion. In addition, it was observed that there were slight alterations in the ability of LSD1 to association with various transcription factors including Snai1, which is implicated to be involved in

differentiation of epithelial cells into mesenchymal cells. This reduced binding could explain the hypotonia (loss in muscle tone) present in each patient[248,249]. It would be intriguing to determine the effects of these mutations on the association of LSD1 with its other binding partners, especially those required at fertilization. In fact, many of LSD1's well characterized binding partners are expressed in the oocyte (Table 1). These mutations can likely affect not only LSD1 function post fertilization but could also exacerbate defects caused by a reduction in the ability of LSD1 to associate with its known binding partners. Furthermore, the reduction in affinity for LSD1's cognate substrate, the tail of histone H3, can explain the hypomorphic effects seen in the human patients. Timing of pre-implantation development is crucial in order for proper development to occur. If LSD1 enzyme kinetics are altered at this stage, this could lead to a propagation of defects post-implantation and well into development.

4.6 LSD1 AND DNA METHYLATION

The interaction between LSD1 and DNA methylation remains unclear. It has been reported that methylation at H3K4 blocks the acquisition of DNA methylation, as it prevents DNMT3L from docking on the histone tail to facilitate the methylation of DNA by DNMT3A or DNMT3B[251]. This work suggests that LSD1 is required to remove H3K4me1/2 methylation in order to silence chromatin. Studies in embryonic stem cells reveals that loss of LSD1 causes a global loss of DNA methylation correlating with a decrease in stability of DNMT1, the maintenance DNA methyltransferase[63]. The decreased stability of DNMT1 was reported to be due to a regulatory role of LSD1 in demethylating DNMT1, which is necessary for DNMT1 stabilization. Therefore, loss of LSD1 can have indirect effects on DNA methylation in addition to its direct effects on histone tails.

	wild-type oocyte expression	*Lsd1* null oocyte expression	relative fold change
CoREST (RCOR1)	601.64	563.78	-0.09
REST	3703.92	2055.38	-0.85
Supervillin (SVIL)	111.73	268.79	1.26
Androgen Receptor	(-)	(-)	(-)
Snai1	5.85	6.55	0.14
CtBP2	1073.54	635.89	-0.75
HDAC1	19.98	67.31	1.70
HDAC2	1446.97	2725.02	0.91
NuRD (snurf)	7.30	27.06	1.76
Blimp1	(-)	(-)	(-)
Braf35	(-)	(-)	(-)
PRKCB	74.40	4.01	-3.91
ASXL1	2561.26	1653.29	-0.63
TLX/NR2E1	6.576766667	200.4472333	4.194478293
Kdm1a	921.26	6.37	-6.97

Table 1: Expression of Known LSD1 binding partners in Mouse oocytes
We analyzed genome-wide expression data from both wild-type oocytes and *Lsd1* null oocytes for known LSD1 interactors. FPKM values are given for each gene with relative fold change in expression reported. (-) indicates gene not expressed.

Conversely, although LSD1 has been reported to remove H3K9 methylation when in complex with the androgen receptor or with Supervillin in a neuronal context, it can also perform an activating role[49,54]. In fact, the Heard group reported that loss of maternal LSD1 led to global increases in all species of both H3K4 methylation but also H3K9 methylation[195]. In oocytes lacking *Lsd1*, the Kelsey group reported both large scale increases and decreases in gene expression[252]. Furthermore, changes in DNA methylation correlated with changes in gene expression suggesting a primary role for LSD1 in regulating the oocyte transcriptome with alterations in DNA methylation being secondary effects.

Conversely, work with LSD2, a homolog of LSD1, has demonstrated that its loss results in complete loss of DNA methylation at a subset of imprinted genes and modest losses of methylation at most imprints[40,252]. These global decreases in DNA methylation upon loss of LSD2 highlight a more direct link between LSD2's function and DNA methylation, especially when compared with that of LSD1. Furthermore, in the progeny derived from *Lsd2* null females, the subset of imprinted loci that had a complete loss of DNA methylation were the same loci observed to have a complete loss in the oocytes. These data suggest a distinction between the role of LSD2, which appears to be more directly linked with DNA methylation acquisition and maintenance, and LSD1, which appears to function more as a transcriptional regulator and indirectly affect DNA methylation. I propose that the changes in DNA methylation at imprinted loci in the progeny of females with partial loss of LSD1 function is more a consequence of the epigenetic landscape being altered post-fertilization than from direct LSD1 action at these loci.

4.7 DNA METHYLATION AT IMPRINTED LOCI: DEFECTS AND DISORDERS

We report that partial loss of LSD reprogramming function at fertilization results in long range phenotypes, including an altered DNA methylation landscape at a subset of imprinted genes in animals that exhibited perinatal lethality. We cannot fully determine if the cause of lethality is due to alterations at these imprinted genes or due to more global molecular defects.

Altering a single imprinted locus often leads to defects on just growth and rarely lead to complete lethality– often, it takes alterations in several imprinted loci before embryonic lethality is observed. This has been demonstrated by loss of *Dnmt3b*, *Dnmt1*, *Trim28*, *Stella*, *Zfp57*, and *Lsd2*[124,153,167,168,253,39]. We observed both DNA methylation gains and losses at imprinted loci. The potential compensation from opposing effects on multiple imprinted loci could potentially explain why a few of our animals survive in part to birth and to adulthood.

Mutations at imprinted loci cause imprinting disorders such as Prader-Willi Syndrome (PWS) and Angelman Syndrome (AS). PWS is characterized by intellectual disability, hypothalamic dysfunction, and obesity[254,255]. AS is characterized by developmental delay and poor motor skills[254,255]. In these cases, whole imprinting clusters are disrupted but development still occurs. The phenotypes associated these disorders, and with many others, strongly resemble the phenotypes seen in human patients with *Lsd1* mutations. Analysis of genomic imprinting in human *Lsd1* patients could provide insights into what defects may be present. Human *Lsd1* patients could have aberrant DNA methylation, which could lead to a wide range of phenotypes, including the observed intellectual disability and developmental delays. Our data highlight the pleiotropic effects of DNA methylation and reveal how the level of reprogramming at fertilization can influence the molecular landscape throughout development.

4.8 DNA METHYLATION, IMPRINTED GENES AND NEUROLOGICAL FUNCTION

In addition being required for embryonic growth and development, imprinted genes have a significant role in neural development and function. Many imprinted loci, including *Grb10,* have a different expression pattern in the brain than in other tissues[194]. *Grb10* is paternally expressed and maternally silenced in every tissue except in a subset of neurons. Paternal deletion of *Grb10* results in altered behaviors in mice including aggressive behaviors and altered social behaviors while *Peg3* has been implicated in sex-specific behaviors[256]. In our reduced *Lsd1* animals, we observed alterations in DNA methylation at *Grb10* in perinatal animals which could contribute to the changes in behavior seen in the adult animals. Further work addressing the imprinting status of *Grb10* and other loci will help determine this.

Although we mostly observed alterations at a few imprinted loci in our perinatal lethal animals, we also observed alterations in DNA methylation at a non-imprinted locus, *Oct4*. Furthermore, a cursory look at global DNA methylation in these animals with methylation-sensitive restriction enzymes reveals a slight change in global methylation. These data suggest that the DNA methylation landscape is altered. This finding has not been probed in the adult survivors but based on their behavioral abnormalities, I speculate that there may be alterations in these animals as well that contribute to the neuronal phenotype we observe.

4.9 IMPLICATIONS OF ALTERED LSD1 FUNCTION AND DNA METHYLATION IN NEUROLOGICAL DISORDERS

We reported changes in the expression pattern of many epigenetic modifiers including

Stella, *Dnmt1* and *Tet1* in the embryos that arrested at the two-cell stage due to complete loss of maternal *Lsd1*. In *Lsd1* partial loss of function animals, the epigenetic landscape has clearly been altered, as indicated by analysis of DNA methylation at imprinted genes. We speculate that the alterations in epigenetic modifiers are propagated throughout development and affect neuronal functioning in adults, causing them to exhibit obsessive-compulsive like behavior. This propagation is likely mediated by DNA methylation defects and may affect embryonic function of LSD1. Although we have not interrogated epigenetic modifiers that were altered at the two-cell stage in our hypomorphic *Lsd1* animals, we speculate that their transcription is also affected.

Many epigenetic modifying enzymes have roles in neuronal development. For example, Tet1 is required for the regulation of genes associated with normal neuronal activity[20]. Mutations in *Dnmt3b*, lead to Immunodeficiency, Centromere instability and Facial anomalies (ICF) syndrome[257,258], which presents similar phenotypes to autism spectrum disorders and the human *Lsd1* patients. Another factor, MECP2, a protein that binds to methylated cytosines, is required for normal neuronal development. Loss of MECP2 results in Rett Syndrome which causes autism-like behaviors[25,171,173,174]. This implicates a major role for DNA methylation in proper neuronal function.

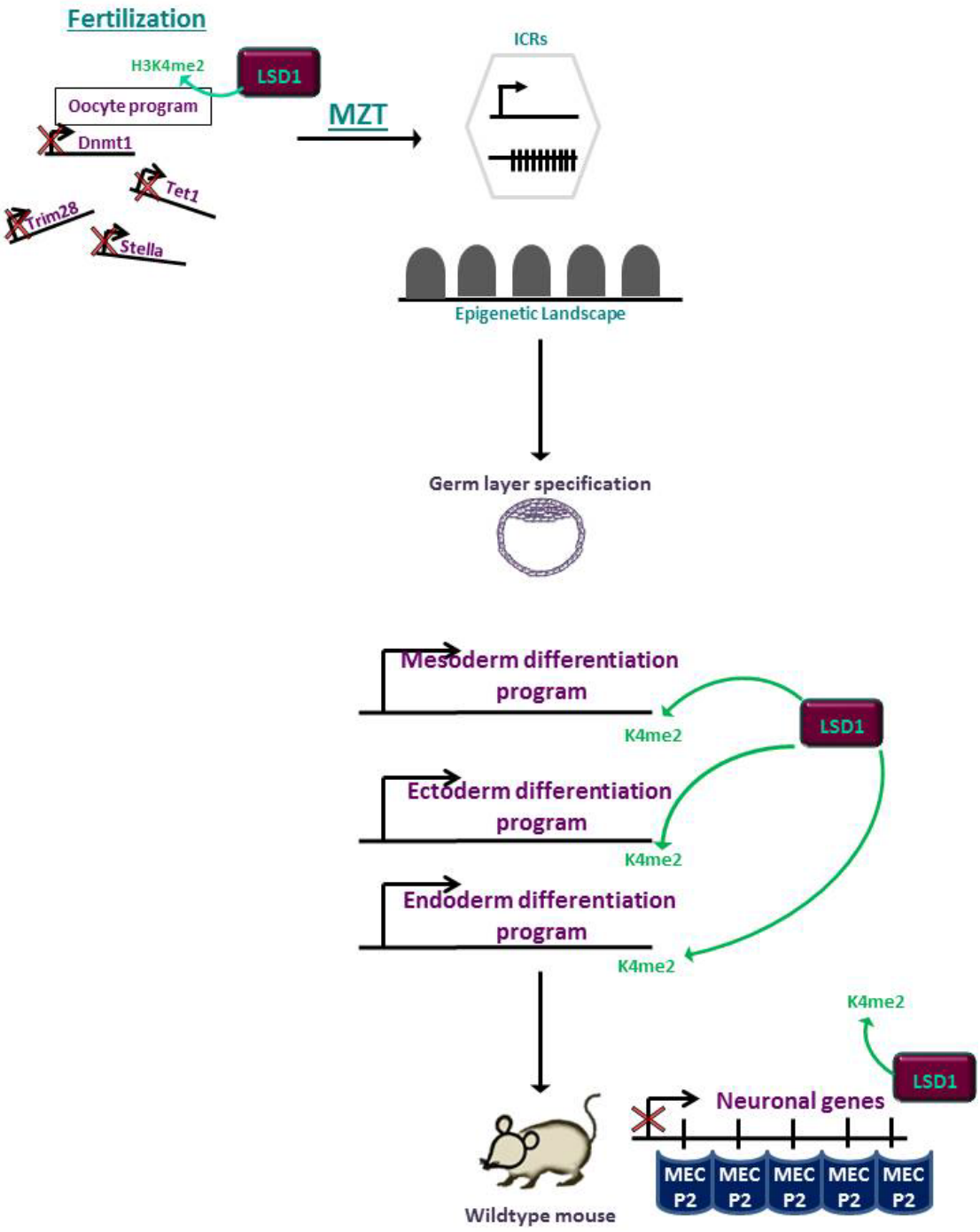
LSD1 is required for neural stem cell (NSC) proliferation and neurite morphogenesis[53]. Knockdown of the neuronal specific isoform of LSD1, nLSD1, *in vitro*, results in reduced neuronal complexity, while overexpression of LSD1 increases neuronal branching and complexity[53] . LSD1 has also been characterized as a member of a complex that contributes to X-linked intellectual disability[57]. Furthermore, in mouse models of epilepsy, LSD1 regulates neuronal excitability[259]. If LSD1-dependent epigenetic reprogramming affects LSD1 embryonic function, I propose that this could lead to defects in neurogenesis,

intellectual disability, and even psychiatric disorders. Furthermore, alterations in DNA methylation landscape in the adult survivors could explain the behavioral abnormalities observed.

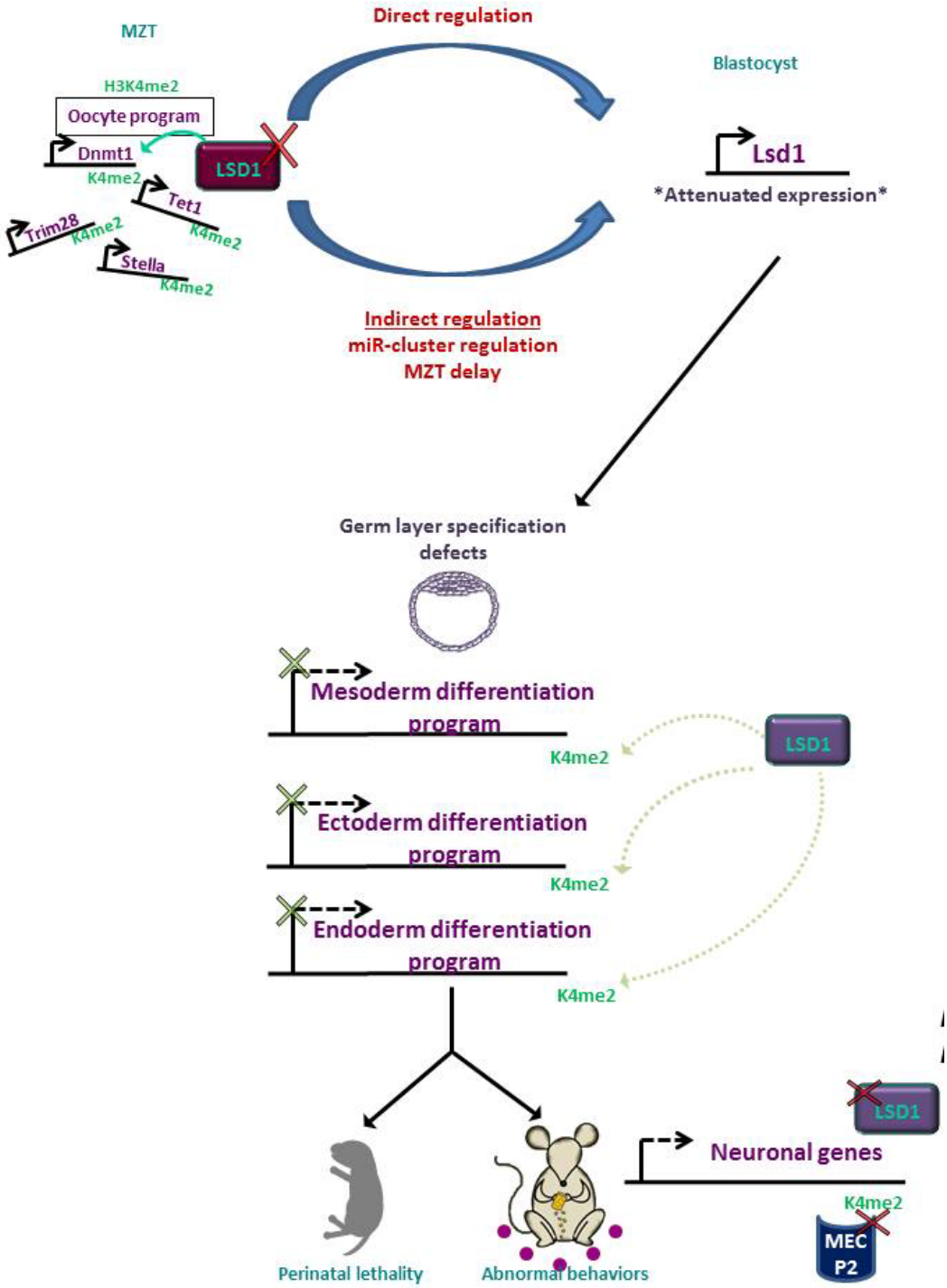
4.10 MODEL FOR RESULTING PHENOTYPES IN REDUCED MATERNAL LSD1 ANIMALS

LSD1 is absolutely required at fertilization in order to ensure that the maternal to zygotic transition occurs. It also appears to be involved in organizing the epigenetic landscape to help prime development. Perturbations to LSD1 reprogramming can even have consequences much later in an animal's life. This may be a result of inefficient reprogramming at fertilization, which could influence the timing or level of embryonic LSD1 expression and therefore propagate molecular defects throughout multiple cell fate decisions. Alternatively, alterations in LSD1 reprogramming at fertilization could affect the epigenetic landscape, whether directly through affecting histone modifications at imprinted loci or indirectly through regulating the expression of other epigenetic modifiers. Initial alteration in epigenetic landscape establishment is then propagated throughout development, perhaps by changes in DNA methylation. In both cases, changes in the epigenetic landscape lead to changes in neuronal function, potentially by affecting the activity of MECP2 or TET1, and ultimately resulting in the behaviors that we observed in the surviving adult animals (Model Figure 1). This work emphasizes the importance of proper reprogramming at fertilization in order to ensure proper development.

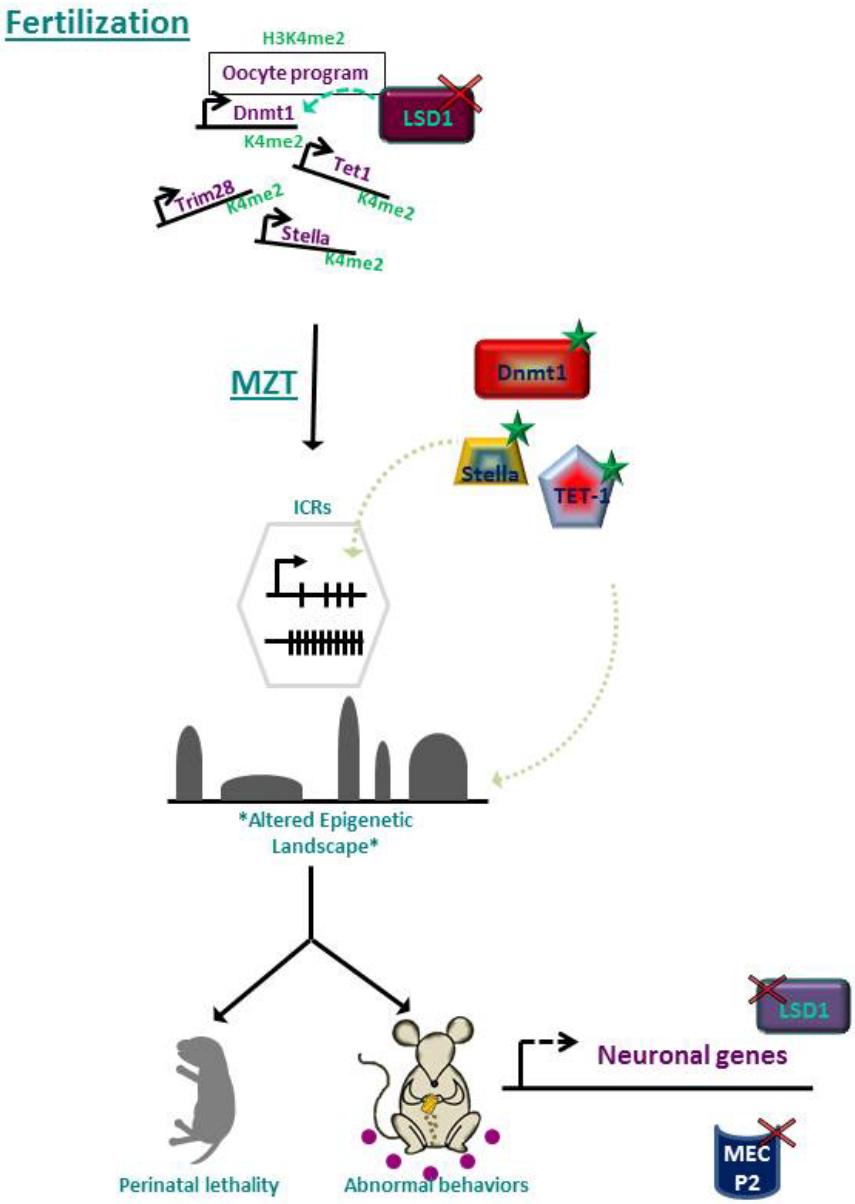
A. Wildtype


Fertilization
H3K4me2
LSD1
Oocyte program
Dnmt1
Tet1
Trim28
Stella
MZT
ICRs
Epigenetic Landscape
Germ layer specification
Mesoderm differentiation program
K4me2
LSD1
Ectoderm differentiation program
K4me2
Endoderm differentiation program
K4me2
K4me2
LSD1
Neuronal genes
MEC P2
MEC P2
MEC P2
MEC P2
MEC P2
Wildtype mouse

B. Maternal mutant model 1



C. Maternal mutant model 2



Model Figure 1:
Proposed model of LSD1 modulation of epigenetic landscape throughout development

A. In normal development, LSD1 functions to silence the maternal transcriptional program in order to silence epigenetic regulators such as Stella, TET1, DNMT1 and TRIM28, and to allow the embryonic program to become fully activated. Imprinted loci are maintained correctly through proper DNA methylation at imprinting control regions (ICRs) and the epigenetic landscape is regulated to ensure proper gene expression globally. During germ layer specification, embryonic LSD1 is expressed at normal levels and aids in the differentiation process through regulating H3K4me2 at germ layer specific genes. During neuronal development, LSD1 acts to remove H3K4me2 from neuronal genes in order to allow DNA acquisition at these loci. This allows MECP2 to bind to the DNA methylation at these loci, thus regulating neuronal development and ensuring proper neuronal functioning in adult animals. B. In model 1 of LSD1 modulation of development, reduction in LSD1 levels at fertilization affects the expression level of embryonic *Lsd1*. Maternal LSD1 can affect embryonic *Lsd1* either through directly regulating the locus, through regulating a microRNA cluster that regulates LSD1 mRNA expression, or through delaying the maternal to zygotic transition and thus development. This results in reduced embryonic LSD1 activity which affects germ layer specification. This can result in either perinatal lethality or altered behaviors in adult animals. C. In model 2 of LSD1 modulation of development, reduced maternal LSD1 results in a maintenance of expression of maternal epigenetic regulators. This maintenance of expression leads to these regulators have a gain of function at multiple loci including imprinting control regions (ICRs) and the epigenetic landscape in general. During neuronal development in both the indirect and direct models, reduced LSD1 results in a maintenance of H3K4me2 at neuronal genes which blocks the acquisition of DNA methylation at these loci. This prevents MECP2 action, this resulting in altered neuronal development and and neuronal functioning. Abnormal behaviors in adult animals would result.

REFERENCES

1. Bianconi, E. *et al.* An estimation of the number of cells in the human body. *Ann. Hum. Biol.* **40,** 463–471 (2013).
2. Waddington, C. H. Canalization of development and the inheritance of acquired characters. *Nature* (1942).
3. Berger, S. L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. An operational definition of epigenetics. *Genes Dev.* **23,** 781–783 (2009).
4. Zhou, V. W., Goren, A. & Bernstein, B. E. Charting histone modifications and the functional organization of mammalian genomes. *Nat. Publ. Gr.* **12,** 7–18 (2010).
5. Tessarz, P. & Kouzarides, T. Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* **15,** 703–708 (2014).
6. Cutter, A. R. & Hayes, J. J. A brief review of nucleosome structure. *FEBS Letters* **589,** 2914–2922 (2015).
7. Zink, L. M. & Hake, S. B. Histone variants: Nuclear function and disease. *Current Opinion in Genetics and Development* **37,** 82–89 (2016).
8. Maze, I., Noh, K.-M., Soshnev, A. A. & Allis, C. D. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nat. Rev. Genet.* **15,** 259–71 (2014).
9. Karch, K. R., DeNizio, J. E., Black, B. E. & Garcia, B. A. Identification and interrogation of combinatorial histone modifications. *Frontiers in Genetics* **4,** (2013).
10. Rothbart, S. B. & Strahl, B. D. Interpreting the language of histone and DNA modifications. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1839,** 627–643 (2014).
11. Dehennaut, V., Leprince, D. & Lefebvre, T. O-GlcNAcylation, an epigenetic mark. Focus on the histone code, TET family proteins, and polycomb group proteins. *Frontiers in Endocrinology* **5,** (2014).
12. Berndsen, C. E. & Denu, J. M. Catalysis and substrate selection by histone/protein lysine

acetyltransferases. *Current Opinion in Structural Biology* **18,** 682–689 (2008).

13. Du, J., Johnson, L. M., Jacobsen, S. E. & Patel, D. J. DNA methylation pathways and their crosstalk with histone methylation. *Nat. Rev. Mol. Cell Biol.* **16,** 519–532 (2015).
14. Mozzetta, C., Boyarchuk, E., Pontis, J. & Ait-Si-Ali, S. Sound of silence: the properties and functions of repressive Lys methyltransferases. *Nat. Rev. Mol. Cell Biol.* **16,** 499–513 (2015).
15. Whyte, W. A. *et al.* Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature* **482,** 221–5 (2012).
16. Benayoun, B. A. *et al.* H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* **158,** 673–688 (2014).
17. Ng, H. H., Robert, F., Young, R. A. & Struhl, K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11,** 709–719 (2003).
18. Mito, Y., Henikoff, J. G. & Henikoff, S. Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* **37,** 1090–1097 (2005).
19. Wu, H. & Zhang, Y. Mechanisms and functions of Tet protein- mediated 5-methylcytosine oxidation. *Genes Dev.* **25,** 2436–2452 (2011).
20. Rudenko, A. *et al.* Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. *Neuron* **79,** 1109–1122 (2013).
21. Dawlaty, M. M. *et al.* Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell* **24,** 310–323 (2013).
22. Qiao, Y., Yang, X. & Jing, N. Epigenetic regulation of early neural fate commitment. *Cellular and Molecular Life Sciences* **73,** 1399–1411 (2016).
23. Raquel Montalbán-Loro, Ana Domingo-Muelas, Alexandra Bizy, S. R. F. & Raquel. Epigenetic regulation of stemness maintenance in the neurogenic niches. *World Journal of Stem Cells* 700–710 (2015). doi:10.4252/wjsc.v7.i4.700
24. Fuks, F. *et al.* The methyl-CpG-binding protein MeCP2 links DNA methylation to histone

methylation. *J. Biol. Chem.* **278,** 4035–4040 (2003).

25. Song, C. *et al.* DNA methylation reader MECP2: cell type- and differentiation stage-specific protein distribution. *Epigenetics Chromatin* **7,** 17 (2014).

26. Herz, H. M., Garruss, A. & Shilatifard, A. SET for life: Biochemical activities and biological functions of SET domain-containing proteins. *Trends in Biochemical Sciences* **38,** 621–639 (2013).

27. Yeates, T. O. Structures of SET domain proteins: Protein lysine methyltransferases make their mark. *Cell* **111,** 5–7 (2002).

28. Andreu-Vieyra, C. V. *et al.* MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. *PLoS Biol.* **8,** 53–54 (2010).

29. Takeuchi, T., Watanabe, Y., Takano-Shimizu, T. & Kondo, S. Roles of jumonji and jumonji family genes in chromatin regulation and development. *Developmental Dynamics* **235,** 2449–2459 (2006).

30. Kidder, B. L., Hu, G. & Zhao, K. KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation. *Genome Biol.* **15,** R32 (2014).

31. Tachibana, M., Nozaki, M., Takeda, N. & Shinkai, Y. Functional dynamics of H3K9 methylation during meiotic prophase progression. **26,** 3346–3359 (2007).

32. Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H. & Shinkai, Y. G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J.* **27,** 2681–2690 (2008).

33. Hathaway, N. A. *et al.* Dynamics and memory of heterochromatin in living cells. *Cell* **149,** 1447–1460 (2012).

34. Ooi, S. K. T. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. **448,** (2007).

35. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119,** 941–953 (2004).

36. Karytinos, A. *et al.* A novel mammalian flavin-dependent histone demethylase. *J. Biol. Chem.* **284,** 17775–17782 (2009).

37. Dong, C., Zhang, H., Xu, C., Arrowsmith, C. H. & Min, J. Structure and function of dioxygenases in histone demethylation and DNA/RNA demethylation. *IUCrJ* **1,** 540–549 (2014).

38. Zhang, Q. *et al.* Structure-function analysis reveals a novel mechanism for regulation of histone demethylase LSD2/AOF1/KDM1b. *Cell Res.* **2,** 1–17 (2012).

39. Fang, R. *et al.* Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating intragenic H3K4me2 Methylation. *Mol. Cell* **39,** 222–233 (2010).

40. Ciccone, D. N. *et al.* KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* **461,** 415–418 (2009).

41. You, A., Tong, J. K., Grozinger, C. M. & Schreiber, S. L. CoREST is an integral component of the CoREST- human histone deacetylase complex. *Proc Natl Acad Sci U S A* **98,** 1454–1458 (2001).

42. Shi, Y. J. *et al.* Regulation of LSD1 histone demethylase activity by its associated factors. *Mol. Cell* **19,** 857–864 (2005).

43. Wang, J. *et al.* Opposing LSD1 complexes function in developmental gene activation and repression programmes. *Nature* **446,** 882–887 (2007).

44. Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487,** 57–63 (2012).

45. Macfarlan, T. S. *et al.* Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev* **25,** 594–607 (2011).

46. Rudolph, T. *et al.* Heterochromatin formation in Drosophila is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. *Mol Cell* **26,** 103–115 (2007).

47. Schneider, R. *et al.* LSD1 demethylates repressive histone marks to promote androgen-

receptor-dependent transcription. **437,** 25–28 (2005).

48. Wissmann, M. *et al.* Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. **9,** (2007).

49. Metzger, E. *et al.* LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437,** 436–439 (2005).

50. Sun, G. *et al.* miR-137 forms a regulatory loop with nuclear receptor TLX and LSD1 in neural stem cells. *Nat. Commun.* **2,** 529 (2011).

51. Grimaldi, P. *et al.* The faah gene is the first direct target of estrogen in the testis: Role of histone demethylase LSD1. *Cell. Mol. Life Sci.* **69,** 4177–4190 (2012).

52. Lim, S. *et al.* Lysine-specific demethylase 1 (LSD1) is highly expressed in ER-negative breast cancers and a biomarker predicting aggressive biology. *Carcinogenesis* **31,** 512–520 (2010).

53. Zibetti, C. *et al.* Alternative splicing of the histone demethylase LSD1/KDM1 contributes to the modulation of neurite morphogenesis in the mammalian nervous system. *J. Neurosci.* **30,** 2521–32 (2010).

54. Laurent, B. *et al.* A Specific LSD1/KDM1A Isoform Regulates Neuronal Differentiation through H3K9 Demethylation. *Mol. Cell* **57,** 957–970 (2015).

55. Wang, J. *et al.* LSD1n is an H4K20 demethylase regulating memory formation via transcriptional elongation control. *Nat. Neurosci.* **18,** 1256–64 (2015).

56. Shi, Y., Sawada, J., Sui, G. & Affar, E. B. Coordinated histone modifications mediated by a CtBP co-repressor complex. 735–738 (2003).

57. Hakimi, M. A., Dong, Y., Lane, W. S., Speicher, D. W. & Shiekhattar, R. A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. *J. Biol. Chem.* **278,** 7234–7239 (2003).

58. Humphrey, G. W. *et al.* Stable Histone Deacetylase Complexes Distinguished by the Presence of SANT Domain Proteins CoREST/kiaa0071 and Mta-L1. *J. Biol. Chem.* **276,** 6817–6824 (2001).

59. Hakimi, M.-A. *et al.* A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7420–7425 (2002).

60. Su, S. *et al.* Involvement of Histone Demethylase LSD1 in Blimp-1-Mediated Gene Repression during Plasma Cell Differentiation Involvement of Histone Demethylase LSD1 in Blimp-1-Mediated Gene Repression during Plasma Cell Differentiation □ . (2009). doi:10.1128/MCB.01158-08

61. Forneris, F., Binda, C., Battaglioli, E. & Mattevi, A. LSD1 : oxidative chemistry for multifaceted functions in chromatin regulation. 181–189 (2008). doi:10.1016/j.tibs.2008.01.003

62. Huang, J. *et al.* p53 is regulated by the lysine demethylase LSD1. *Nature* **449,** 105–108 (2007).

63. Wang, J. *et al.* The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat Genet* **41,** 125–129 (2009).

64. Althoff, K. *et al.* MiR-137 functions as a tumor suppressor in neuroblastoma by downregulating KDM1A. *Int. J. Cancer* **133,** 1064–1073 (2013).

65. Balaguer, F. *et al.* Epigenetic silencing of miR-137 is an early event in colorectal carcinogenesis. *Cancer Res.* **70,** 6609–6618 (2010).

66. Nam, H. J. *et al.* Phosphorylation of LSD1 by PKCalpha Is Crucial for Circadian Rhythmicity and Phase Resetting. *Mol. Cell* **53,** 791–805 (2014).

67. Nair, V. D. *et al.* Involvement of Histone Demethylase LSD1 in Short-Time-Scale Gene Expression Changes during Cell Cycle Progression in Embryonic Stem Cells. *Mol. Cell. Biol.* **32,** 4861–4876 (2012).

68. Lin, Y. *et al.* The SNAG domain of Snail1 functions as a molecular hook for recruiting lysine-specific demethylase 1. *EMBO J.* **29,** 1803–1816 (2010).

69. Wu, Y. *et al.* The Deubiquitinase USP28 Stabilizes LSD1 and Confers Stem-Cell-like Traits to Breast Cancer Cells. *Cell Rep.* **5,** 224–236 (2013).

70. Bird, A. DNA methylation patterns and epigenetic memory. *Genes and Development* **16,** 6–21

(2002).

71. Ng, R. K. & Gurdon, J. B. Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. *Nat Cell Biol* **10,** 102–109 (2008).

72. Migicovsky, Z. & Kovalchuk, I. Epigenetic memory in mammals. *Frontiers in Genetics* **2,** (2011).

73. Aikawa, S., Kobayashi, M. J., Satake, A., Shimizu, K. K. & Kudoh, H. Robust control of the seasonal expression of the Arabidopsis FLC gene in a fluctuating environment. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 11632–7 (2010).

74. Berry, S. & Dean, C. Environmental perception and epigenetic memory: mechanistic insight through FLC. *Plant J.* **83,** 133–48 (2015).

75. Michaels, S. D. & Amasino, R. M. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* **11,** 949–56 (1999).

76. Ramsden, S. C., Clayton-smith, J., Birch, R. & Buiting, K. Practice guidelines for the molecular analysis of Prader-Willi and Angelman syndromes. (2010).

77. Sheldon, C. C. *et al.* The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. *Plant Cell* **11,** 445–458 (1999).

78. Crevill??n, P. & Dean, C. Regulation of the floral repressor gene FLC: The complexity of transcription in a chromatin context. *Current Opinion in Plant Biology* **14,** 38–44 (2011).

79. Yang, H., Howard, M. & Dean, C. Antagonistic roles for H3K36me3 and H3K27me3 in the cold-induced epigenetic switch at Arabidopsis FLC. *Curr. Biol.* **24,** 1793–1797 (2014).

80. Bastow, R. *et al.* Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* **427,** 164–167 (2004).

81. De Lucia, F., Crevillen, P., Jones, A. M. E., Greb, T. & Dean, C. A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 16831–16836 (2008).

82. Buzas, D. M., Robertson, M., Finnegan, E. J. & Helliwell, C. A. Transcription-dependence of

histone H3 lysine 27 trimethylation at the Arabidopsis polycomb target gene FLC. *Plant J.* **65,** 872–881 (2011).

83. Berry, S., Hartley, M., Olsson, T. S. G., Dean, C. & Howard, M. Local chromatin environment of a Polycomb target gene instructs its own epigenetic inheritance. *Elife* **4,** e07205 (2015).

84. Katz, D. J., Edwards, T. M., Reinke, V. & Kelly, W. G. A C. elegans LSD1 demethylase contributes to germline immortality by reprogramming epigenetic memory. *Cell* **137,** 308–320 (2009).

85. Di Stefano, L., Ji, J. Y., Moon, N. S., Herr, A. & Dyson, N. Mutation of Drosophila Lsd1 disrupts H3-K4 methylation, resulting in tissue-specific defects during development. *Curr Biol* **17,** 808–812 (2007).

86. Szabad, J., Reuter, G. & Schroder, M. B. The effects of two mutations connected with chromatin functions on female germ-line cells of Drosophila. *Mol Gen Genet* **211,** 56–62 (1988).

87. Wang, Z., Oron, E., Nelson, B., Razis, S. & Ivanova, N. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell* **10,** 440–454 (2012).

88. Wang, Z. X. *et al.* Oct4 and Sox2 directly regulate expression of another pluripotency transcription factor, Zfp206, in embryonic stem cells. *J. Biol. Chem.* **282,** 12822–12830 (2007).

89. Rizzino, A. Concise review: The Sox2-Oct4 connection: Critical players in a much larger interdependent network integrated at multiple levels. *Stem Cells* **31,** 1033–1039 (2013).

90. S??awek, S. *et al.* Pluripotency transcription factors in lung cancer???a review. *Tumor Biology* **37,** 4241–4249 (2016).

91. Müller, M. *et al.* The role of pluripotency factors to drive stemness in gastrointestinal cancer. *Stem Cell Research* **16,** 349–357 (2016).

92. Cairns, B. R. Emerging roles for chromatin remodeling in cancer biology. *Trends Cell Biol.* **11,**

S15–S21 (2001).

93. Wu, Y. & Zhou, B. P. Epigenetic regulation of LSD1 during mammary carcinogenesis. *Mol. Cell. Oncol.* **1,** e963426 (2014).

94. Abdel-Wahab, O. & Levine, R. L. Mutations in epigenetic modifiers in the pathogenesis and therapy of acute myeloid leukemia. *Blood* **121,** 3563–3572 (2013).

95. Wada, T., Koyama, D., Kikuchi, J., Honda, H. & Furukawa, Y. Overexpression of the shortest isoform of histone demethylase LSD1 primes hematopoietic stem cells for malignant transformation. *Blood* **125,** 3731–46 (2015).

96. Li, Y. *et al.* Dynamic interaction between TAL1 oncoprotein and LSD1 regulates TAL1 function in hematopoiesis and leukemogenesis. *Oncogene* 5007–5018 (2012). doi:10.1038/onc.2012.8

97. Andricovich, J., Kai, Y. & Tzatsos, A. Lysine-specific histone demethylases in normal and malignant hematopoiesis. *Exp. Hematol.* (2016). doi:10.1016/j.exphem.2016.05.006

98. Weismann, A., Parker, W. N. & Rönnfeldt, H. *The germ-plasm; a theory of heredity*. (Scribner's, 1893).

99. Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. **447,** 425–432 (2007).

100. Campbell, K. H., McWhir, J., Ritchie, W. A. & Wilmut, I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* **380,** 64–66 (1996).

101. Gurdon, J. B., Elsdale, T. R. & Fischberg, M. Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. *Nature* **182,** 64–65 (1958).

102. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126,** 663–676 (2006).

103. Hochedlinger, K. & Jaenisch, R. Nuclear transplantation: lessons from frogs and mice. *Curr Opin Cell Biol* **14,** 741–748 (2002).

104. Gurdon, J. B. The developmental capacity of nuclei taken from intestinal epithelium cells of

feeding tadpoles. *J Embryol Exp Morphol* **10,** 622–640 (1962).

105. Hochedlinger, K. & Jaenisch, R. Nuclear reprogramming and pluripotency. *Nature* **441,** 1061–1067 (2006).

106. Ng, R. K. & Gurdon, J. B. Epigenetic memory of active gene transcription is inherited through somatic cell nuclear transfer. *Proc Natl Acad Sci U S A* **102,** 1957–1962 (2005).

107. Byrne, J. A., Simonsson, S. & Gurdon, J. B. From intestine to muscle: nuclear reprogramming through defective cloned embryos. *Proc Natl Acad Sci U S A* **99,** 6059–6063 (2002).

108. Ruthenburg, A. J., Allis, C. D. & Wysocka, J. Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark. *Molecular Cell* **25,** 15–30 (2007).

109. Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128,** 707–719 (2007).

110. Ahmad, K. & Henikoff, S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9,** 1191–1200 (2002).

111. Muramoto, T., Muller, I., Thomas, G., Melvin, A. & Chubb, J. R. Methylation of H3K4 Is required for inheritance of active transcriptional states. *Curr Biol* **20,** 397–406

112. Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. & Cavalli, G. Genome regulation by polycomb and trithorax proteins. *Cell* **128,** 735–745 (2007).

113. Byrd, K. N. & Shearn, A. ASH1, a Drosophila trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *Proc Natl Acad Sci U S A* **100,** 11535–11540 (2003).

114. Cavalli, G. & Paro, R. Epigenetic inheritance of active chromatin after removal of the main transactivator. *Science (80-.).* **286,** 955–958 (1999).

115. Cavalli, G. & Paro, R. The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. *Cell* **93,** 505–518 (1998).

116. Brykczynska, U. *et al.* Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat Struct Mol Biol* **17,** 679–687

117. Hammoud, S. S. *et al.* Distinctive chromatin in human sperm packages genes for embryo development. *Nature* (2009).

118. Wu, S. *et al.* Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. (2011). doi:10.1101/gr.113167.110

119. Vastenhouw, N. L. *et al.* Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* **464,** 922–926 (2010).

120. Smallwood, S. A. *et al.* Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat Genet* **43,** 811–814 (2011).

121. Lindeman, L. C. *et al.* Prepatterning of developmental gene expression by modified histones before zygotic genome activation. *Dev Cell* **21,** 993–1004 (2011).

122. Yoder, J. A., Soman, N. S., Verdine, G. L. & Bestor, T. H. DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *J. Mol. Biol.* **270,** 385–395 (1997).

123. Kim, K.-H. & Lee, K.-A. Maternal effect genes: Findings and effects on mouse embryo development. *Clin. Exp. Reprod. Med.* **41,** 47–61 (2014).

124. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99,** 247–257 (1999).

125. Kaneda, M. *et al.* Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429,** 900–903 (2004).

126. Torres-Padilla, M. E., Bannister, A. J., Hurd, P. J., Kouzarides, T. & Zernicka-Goetz, M. Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int. J. Dev. Biol.* **50,** 455–461 (2006).

127. Erhardt, S. *et al.* Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. **2,** 4235–4248 (2003).

128. Bultman, S. J. *et al.* Maternal BRG1 regulates zygotic genome activation in the mouse. *Genes Dev* **20,** 1744–1754 (2006).

129. Wan, L.-B. *et al.* Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. *Development* **135,** 2729–38 (2008).

130. Kaneda, M., Tang, F., O'Carroll, D., Lao, K. & Surani, M. A. Essential role for Argonaute2 protein in mouse oogenesis. *Epigenetics Chromatin* **2,** 9 (2009).

131. Ramos, S. B. V. Characterization of DeltaN-Zfp36l2 mutant associated with arrest of early embryonic development and female infertility. *J. Biol. Chem.* **287,** 13116–27 (2012).

132. Tsukamoto, S., Kuma, A. & Mizushima, N. The role of autophagy during the oocyte-to-embryo transition. *Autophagy* **4,** 1076–1078 (2008).

133. Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts. *Development* **136,** 3033–3042 (2009).

134. Li, L., Lu, X. & Dean, J. The maternal to zygotic transition in mammals. *Molecular Aspects of Medicine* **34,** 919–938 (2013).

135. Adenot, P. G., Mercier, Y., Renard, J. P. & Thompson, E. M. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* **124,** 4615–4625 (1997).

136. Li, L., Zheng, P. & Dean, J. Maternal control of early mouse development. *Development* **137,** 859–870 (2010).

137. Flach, G., Johnson, M. H., Braude, P. R., Taylor, R. a & Bolton, V. N. The transition from maternal to embryonic control in the 2-cell mouse embryo. *EMBO J.* **1,** 681–686 (1982).

138. Howlett, S. K. & Bolton, V. N. Sequence and regulation of morphological and molecular events during the first cell cycle of mouse embryogenesis. *J. Embryol. Exp. Morphol.* **87,** 175–206 (1985).

139. Ooi, S. K. T., O'Donnell, A. H. & Bestor, T. H. Mammalian cytosine methylation at a glance. *J. Cell Sci.* **122,** 2787–2791 (2009).

140. Bartolomei, M. S. Genomic imprinting: Employing and avoiding epigenetic processes. *Genes and Development* **23,** 2124–2133 (2009).

141. Smith, Z. D. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484,** 339–344 (2012).

142. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69,** 915–926 (1992).

143. Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B. & Bestor, T. H. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294,** 2536–2539 (2001).

144. Farthing, C. R. *et al.* Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet* **4,** e1000116 (2008).

145. Imamura, M. *et al.* Transcriptional repression and DNA hypermethylation of a small set of ES cell marker genes in male germline stem cells. *BMC Dev Biol* **6,** 34 (2006).

146. Oswald, J. *et al.* Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* **10,** 475–478 (2000).

147. Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome. *Nature* **403,** 501–502 (2000).

148. Gu, T. P. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477,** 606–610 (2011).

149. Huangfu, D. *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotechnol* **26,** 795–797 (2008).

150. Lane, N. *et al.* Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35,** 88–93 (2003).

151. Howell, C. Y. *et al.* Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* **104,** 829–838 (2001).

152. Nakamura, T. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486,** 415–419 (2012).

153. Nakamura, T. *et al.* PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* **9,** 64–71 (2007).

154. Li, X. *et al.* A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev Cell* **15,** 547–557 (2008).

155. Quenneville, S. *et al.* Article In Embryonic Stem Cells , ZFP57 / KAP1 Recognize a Methylated Hexanucleotide to Affect Chromatin and DNA Methylation of Imprinting Control Regions. *Mol. Cell* **44,** 361–372 (2011).

156. S~Kageyama *et al.* Analysis of transcription factor expression during oogenesis and preimplantation development in mice. *Zygote %L Kageyama:2007fv* **15,** 117–128 (2007).

157. Carrozza, M. J. *et al.* Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123,** 581–592 (2005).

158. Bender, L. B. *et al.* MES-4: an autosome-associated histone methyltransferase that participates in silencing the X chromosomes in the C. elegans germ line. *Development* **133,** 3907–3917 (2006).

159. Andersen, E. C. & Horvitz, H. R. Two C. elegans histone methyltransferases repress lin-3 EGF transcription to inhibit vulval development. *Development* **134,** 2991–2999 (2007).

160. Furuhashi, H. *et al.* Trans-generational epigenetic regulation of C. elegans primordial germ cells. *Epigenetics Chromatin* **3,** 15 (2010).

161. Rechtsteiner, A. *et al.* The Histone H3K36 Methyltransferase MES-4 acts epigenetically to transmit the memory of germline gene expression to progeny. *PLoS Genet.* **6,** (2010).

162. Fong, Y., Bender, L., Wang, W. & Strome, S. Regulation of the different chromatin states of autosomes and X chromosomes in the germ line of C. elegans. *Science (80-.).* **296,** 2235–2238 (2002).

163. Richards, E. J. Inherited epigenetic variation--revisiting soft inheritance. *Nat Rev Genet* **7,** 395–401 (2006).

164. Roemer, I., Reik, W., Dean, W. & Klose, J. Epigenetic inheritance in the mouse. *Curr Biol* **7,**

277–280 (1997).

165. Anway, M. D., Cupp, A. S., Uzumcu, M. & Skinner, M. K. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science (80-.).* **308,** 1466–1469 (2005).

166. Morgan, H. D., Sutherland, H. G., Martin, D. I. & Whitelaw, E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* **23,** 314–318 (1999).

167. Li, X. *et al.* A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and Paternal Imprints. *Dev. Cell* **15,** 547–557 (2008).

168. Messerschmidt, D. M. *et al.* Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science (80-.).* **335,** 1499–1502 (2012).

169. Guo, J. U. *et al.* Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nat Neurosci* **14,** 1345–1351 (2011).

170. Laurent, B. *et al.* A Specific LSD1/KDM1A Isoform Regulates Neuronal Differentiation through H3K9 Demethylation. *Mol Cell* **57,** 957–970 (2015).

171. Amir, R. E. *et al.* Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23,** 185–188 (1999).

172. Ghosh, R. P., Horowitz-Scherer, R. A., Nikitina, T., Shlyakhtenko, L. S. & Woodcock, C. L. MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites. *Mol. Cell. Biol.* **30,** 4656–70 (2010).

173. McGraw, C. M., Samaco, R. C. & Zoghbi, H. Y. Adult neural function requires MeCP2. *Science* **333,** 186 (2011).

174. Gabel, H. W. *et al.* Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature* **522,** 89–93 (2015).

175. Aoki, F., Worrad, D. M. & Schultz, R. M. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev Biol* **181,** 296–307 (1997).

176. Arico, J. K., Katz, D. J., van der Vlag, J. & Kelly, W. G. Epigenetic patterns maintained in early Caenorhabditis elegans embryos can be established by gene activity in the parental germ

cells. *PLoS Genet* **7,** e1001391 (2011).

177. Burton, A. & Torres-Padilla, M. E. Epigenetic reprogramming and development: a unique heterochromatin organization in the preimplantation mouse embryo. *Br. Funct Genomics* **9,** 444–454 (2010).

178. Wu, S. F., Zhang, H. & Cairns, B. R. Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. *Genome Res* **21,** 578–589 (2011).

179. Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts. *Development* **136,** 3033–3042 (2009).

180. Hamatani, T., Carter, M. G., Sharov, A. A. & Ko, M. S. Dynamics of global gene expression changes during mouse preimplantation development. *Dev Cell* **6,** 117–131 (2004).

181. Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500,** 593–7 (2013).

182. Bartolomei, M. S. & Tilghman, S. M. Genomic imprinting in mammals. *Annu Rev Genet* **31,** 493–525 (1997).

183. Foster, C. T. *et al.* Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Mol Cell Biol* **30,** 4851–4863 (2010).

184. Wang, J. *et al.* Opposing LSD1 complexes function in developmental gene activation and repression programmes. **446,** 1–6 (2007).

185. Gallardo, T., Shirley, L., John, G. B. & Castrillon, D. H. Generation of a germ cell-specific mouse transgenic Cre line, Vasa-Cre. *Genesis* **45,** 413–417 (2007).

186. Lan, Z. J., Xu, X. & Cooney, A. J. Differential oocyte-specific expression of Cre recombinase activity in GDF-9-iCre, Zp3cre, and Msx2Cre transgenic mice. *Biol Reprod* **71,** 1469–1474 (2004).

187. de Vries, W. N. *et al.* Expression of Cre recombinase in mouse oocytes: a means to study maternal effect genes. *Genesis* **26,** 110–112 (2000).

188. Turgeon, B. & Meloche, S. Interpreting neonatal lethal phenotypes in mouse mutants:

insights into gene function and human diseases. *Physiol Rev* **89,** 1–26 (2009).

189. Shmelkov, S. V *et al.* Slitrk5 deficiency impairs corticostriatal circuitry and leads to obsessive-compulsive-like behaviors in mice. *Nat Med* **16,** 598–602, 1p following 602 (2010).

190. Gene, A. M. E. *et al.* Article Maintains Both Maternal and Paternal Imprints. *Dev. Cell* **15,** 547–557 (2008).

191. Yamaguchi, S., Shen, L., Liu, Y., Sendler, D. & Zhang, Y. Role of Tet1 in erasure of genomic imprinting. *Nature* **504,** 460–464 (2013).

192. Pradhan, S., Bacolla, A., Wells, R. D. & Roberts, R. J. Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *J Biol Chem* **274,** 33002–33010 (1999).

193. Vertino, P. M., Yen, R. W., Gao, J. & Baylin, S. B. De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5-)-methyltransferase. *Mol Cell Biol* **16,** 4555–4565 (1996).

194. Garfield, A. S. *et al.* Distinct physiological and behavioural functions for parental alleles of imprinted Grb10. *Nature* **469,** 534–538 (2011).

195. Ancelin, K. *et al.* Maternal LSD1/KDM1A is an essential regulator of chromatin and transcription landscapes during zygotic genome activation. *Elife* **5,** 1–24 (2016).

196. Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (80-.).* **317,** 1760–1764 (2007).

197. Quenneville, S. *et al.* In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell* **44,** 361–372 (2011).

198. Voon, H. P. *et al.* ATRX Plays a Key Role in Maintaining Silencing at Interstitial Heterochromatic Loci and Imprinted Genes. *Cell Rep* **11,** 405–418 (2015).

199. Cedar, H. & Bergman, Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* **10,** 295–304 (2009).

200. Sanli, I. & Feil, R. Chromatin mechanisms in the developmental control of imprinted gene expression. *Int J Biochem Cell Biol* **67,** 139–147 (2015).

201. Whitelaw, N. C. *et al.* Reduced levels of two modifiers of epigenetic gene silencing, Dnmt3a and Trim28, cause increased phenotypic noise. *Genome Biol* **11,** R111 (2010).

202. Chung, S. H. *et al.* Zac1 plays a key role in the development of specific neuronal subsets in the mouse cerebellum. *Neural Dev* **6,** 25 (2011).

203. Kosaki, K. *et al.* Human homolog of the mouse imprinted gene Impact resides at the pericentric region of chromosome 18 within the critical region for bipolar affective disorder. *Mol Psychiatry* **6,** 87–91 (2001).

204. Pereira, C. M. *et al.* IMPACT, a protein preferentially expressed in the mouse brain, binds GCN1 and inhibits GCN2 activation. *J Biol Chem* **280,** 28316–28323 (2005).

205. Roffe, M., Hajj, G. N., Azevedo, H. F., Alves, V. S. & Castilho, B. A. IMPACT is a developmentally regulated protein in neurons that opposes the eukaryotic initiation factor 2alpha kinase GCN2 in the modulation of neurite outgrowth. *J Biol Chem* **288,** 10860–10869 (2013).

206. Schmidt-Edelkraut, U., Daniel, G., Hoffmann, A. & Spengler, D. Zac1 regulates cell cycle arrest in neuronal progenitors via Tcf4. *Mol Cell Biol* **34,** 1020–1030 (2014).

207. Varrault, A. *et al.* Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth. *Dev Cell* **11,** 711–722 (2006).

208. Kelsey, G. & Bartolomei, M. S. Imprinted genes... and the number is? *PLoS Genet.* **8,** (2012).

209. DeVeale, B., van der Kooy, D. & Babak, T. Critical Evaluation of Imprinted Gene Expression by RNA–Seq: A New Perspective. *PLoS Genet* **8,** e1002600 (2012).

210. Gregg, C. *et al.* High resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science* **329,** 643–8 (2010).

211. Bartolomei, M. S. & Ferguson-Smith, A. C. Mammalian genomic imprinting. *Cold Spring Harb. Perspect. Biol.* **3,** 1–17 (2011).

212. McGrath, J. & Solter, D. Nuclear transplantation in mouse embryos. *J. Exp. Zool.* **228,** 355–362 (1983).

213. McGrath, J. & Solter, D. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37,** 179–183 (1984).

214. Surani, M. A. H. & Barton, S. C. Development of gynogenetic eggs in the mouse: Implications for parthenogenetic embryos. *Science (80-.).* **222,** 1034–1036 (1983).

215. Surani, M. A., Barton, S. C. & Norris, M. L. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nat. Educ.* **308,** 548–550 (1984).

216. Bartolomei, M. S., Zemel, S. & Tilghman, S. M. Parental imprinting of the mouse H19 gene. *Nature* **351,** 153–155 (1991).

217. Ferguson-Smith, A. C., Cattanach, B. M., Barton, S. C., Beechey, C. V & Surani, M. A. Embryological and molecular investigations of parental imprinting on mouse chromosome 7. *Nature* **351,** 667–670 (1991).

218. Barlow, D. P., Stöger, R., Herrmann, B. G., Saito, K. & Schweifer, N. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature* **349,** 84–87 (1991).

219. DeChiara, T. M., Robertson, E. J. & Efstratiadis, A. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64,** 849–859 (1991).

220. Begemann, M. *et al.* Genetic and epigenetic findings in Silver-Russell syndrome. *Pediatric Endocrinology Reviews* **8,** 86–93 (2010).

221. Weksberg, R., Shuman, C. & Beckwith, J. B. Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* **18,** 8–14 (2010).

222. Wan, L. Ben & Bartolomei, M. S. Chapter 7 Regulation of Imprinting in Clusters: Noncoding RNAs Versus Insulators. *Advances in Genetics* **61,** 207–223 (2008).

223. Ferguson-Smith, A. C. Genomic imprinting: the emergence of an epigenetic paradigm. *Nat. Rev. Genet.* **12,** 565–575 (2011).

224. Barlow, D. P. Genomic Imprinting : A Mammalian Epigenetic Discovery Model. (2011). doi:10.1146/annurev-genet-110410-132459

225. John, R. M. & Lefebvre, L. Developmental regulation of somatic imprints. *Differentiation* **81,** 270–280 (2011).

226. Guttman, M. & Rinn, J. L. Modular regulatory principles of large non-coding RNAs. *Nature* **482,** 339–46 (2012).

227. Reik, W. & Dean, W. DNA methylation and mammalian epigenetics. *Electrophoresis* **22,** 2838–2843 (2001).

228. Reik, W. & Walter, J. Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.* **2,** 21–32 (2001).

229. Reik, W., Santos, F. & Dean, W. Mammalian epigenomics : reprogramming the genome for development and therapy. **59,** 21–32 (2003).

230. Canovas, S. & Ross, P. J. Epigenetics in preimplantation mammalian development. *Theriogenology* (2016). doi:10.1016/j.theriogenology.2016.04.020

231. Hajkova, P. *et al.* Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* **117,** 15–23 (2002).

232. Hajkova, P., B, P. T. R. S. & Hajkova, P. Epigenetic reprogramming in the germline : towards the ground state of the epigenome Epigenetic reprogramming in the germline : towards the ground state of the epigenome. 2266–2273 (2011). doi:10.1098/rstb.2011.0042

233. Hata, K., Okano, M., Lei, H. & Li, E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129,** 1983–1993 (2002).

234. Hayatsu, H., Shiraishi, M. & Negishi, K. Bisulfite modification for analysis of DNA methylation. *Current Protocols in Nucleic Acid Chemistry* (2008). doi:10.1002/0471142700.nc0610s33

235. Laird, P. W. Principles and challenges of genome-wide DNA methylation analysis. *Nat. Rev.*

Genet. **11,** 191–203 (2010).

236. Beck, J. A. *et al.* Genealogies of mouse inbred strains. *Nat. Genet.* **24,** 23–25 (2000).

237. Frazer, K. A. *et al.* A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* **448,** 1050–1053 (2007).

238. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29,** 308–11 (2001).

239. Smigielski, E. M., Sirotkin, K., Ward, M. & Sherry, S. T. dbSNP: a database of single nucleotide polymorphisms. *Nucleic Acids Res.* **28,** 352–355 (2000).

240. Nekrutenko, A. & Taylor, J. Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. *Nat Rev Genet* **13,** 667–672 (2012).

241. Mitchell, A. A., Zwick, M. E., Chakravarti, A. & Cutler, D. J. Discrepancies in dbSNP confirmation rates and allele frequency distributions from varying genotyping error rates and patterns. *Bioinformatics* **20,** 1022–1032 (2004).

242. Egger, G., Liang, G., Aparicio, A. & Jones, P. a. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429,** 457–463 (2004).

243. Hsu, A. *et al.* Promoter de-methylation of cyclin D2 by sulforaphane in prostate cancer cells. *Clin. Epigenetics* **3,** 3 (2011).

244. Bock, C. *et al.* BiQ Analyzer: Visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* **21,** 4067–4068 (2005).

245. Kim, J. *et al.* LSD1 is essential for oocyte meiotic progression by regulating CDC25B expression in mice. *Nat. Commun.* **6,** 10116 (2015).

246. Foster, C. T. *et al.* Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Mol. Cell. Biol.* **30,** 4851–63 (2010).

247. Nicholson, T. B. *et al.* Defective heart development in hypomorphic LSD1 mice. 1–15 (2011). doi:10.1038/cr.2011.194

248. Tunovic, S., Barkovich, J., Sherr, E. H. & Slavotinek, A. M. De novo ANKRD11 and

KDM1A gene mutations in a male with features of KBG syndrome and Kabuki syndrome. *Am. J. Med. Genet. Part A* **164,** 1744–1749 (2014).

249. Chong, J. *et al.* Gene discovery for Mendelian conditions via social networking: de novo variants in KDM1A cause developmental delay and distinctive facial features. *bioRxiv* 028241 (2015). doi:10.1101/028241

250. Pilotto, S. *et al.* LSD1/KDM1A mutations associated to a newly described form of intellectual disability impair demethylase activity and binding to transcription factors. *Hum. Mol. Genet.* (2016). doi:10.1093/hmg/ddw120

251. Ooi, S. K. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448,** 714–717 (2007).

252. Stewart, K. R. *et al.* Dynamic changes in histone modifications precede de novo DNA methylation in oocytes. *Genes Dev.* **29,** 2449–2462 (2015).

253. Howell, C. Y. *et al.* Genomic Imprinting Disrupted by a Maternal Effect Mutation in the Dnmt1 Gene. **104,** 829–838 (2001).

254. Horsthemke, B. & Wagstaff, J. Mechanisms of imprinting of the Prader-Willi/Angelman region. *American Journal of Medical Genetics, Part A* **146,** 2041–2052 (2008).

255. Buiting, K. Prader-Willi syndrome and Angelman syndrome. *American Journal of Medical Genetics, Part C: Seminars in Medical Genetics* **154,** 365–376 (2010).

256. Kaneko-ishino, T., Kohda, T. & Ishino, F. The Regulation and Biological Significance of Genomic Imprinting in Mammals. **133,** 699–711 (2003).

257. Jin, B. *et al.* DNA methyltransferase 3B (DNMT3B) mutations in ICF syndrome lead to altered epigenetic modifications and aberrant expression of genes regulating development, neurogenesis and immune function. *Hum. Mol. Genet.* **17,** 690–709 (2008).

258. Xie, Z. H. *et al.* Mutations in DNA methyltransferase DNMT3B in ICF syndrome affect its regulation by DNMT3L. *Hum. Mol. Genet.* **15,** 1375–1385 (2006).

259. Rusconi, F. *et al.* LSD1 Neurospecific Alternative Splicing Controls Neuronal Excitability in

Mouse Models of Epilepsy. *Cereb. Cortex* **25,** 2729–2740 (2015).