Title

Sperm-Inherited histone marks shape offspring transcription and development in C. elegans

Permalink

https://escholarship.org/uc/item/3rc026rr

Author

Kaneshiro, Kiyomi R

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

SANTA CRUZ

**Sperm-inherited histone marks shape offspring transcription
and development in *C. elegans***

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MOLECULAR, CELLULAR AND
DEVELOPMENTAL BIOLOGY

by

Kiyomi Kaneshiro

September 2019

The Dissertation of Kiyomi Kaneshiro
is approved:

_____
Professor Susan Strome, Chair

_____
Professor John Tamkun

_____
Professor Needhi Bhalla

_____
Quentin Williams
Acting Vice Provost and Dean of Graduate Studies

Table of Contents

List of figures and tables

Chapter 1

Chapter 2

Abstract

Sperm-inherited histone marks shape offspring transcription

and development in *C. elegans*

Kiyomi Kaneshiro

Sperm and oocyte deliver all the information, both genetic and epigenetic, that will specify development of the next generation. Histone modifications are one form in which the gametes deliver epigenetic information to the embryo. Our understanding of how histone modifications inherited on parental genomes influences offspring transcription and development is poorly understood. I investigated a paradigm in which worms inherit the sperm genome lacking H3K27me3, a widely conserved mark of repression. I used genetic, genomic and microscopy approaches to investigate how distributions of histone marking are established in parental germlines, transmitted to offspring and influence offspring transcription and development. I demonstrate that altered inheritance of H3K27me3 results in misregulation of many genes both in the soma and in the germline of offspring. I show that there are developmental consequences to germline health and function as a result of this altered inheritance. My analysis reveals that the derepression of numerous genes in offspring is the result of increased transcription from sperm alleles specifically, which demonstrates that sperm-inherited histone marks can directly influence offspring transcription. I found that in the germlines of fathers lacking H3K27me3, H3K36me3, a mark associated with gene transcription, invades domains normally occupied by H3K27me3 in wild-type male germlines. These changes in H3K36me3 distribution in the germlines of

fathers correlate with changes in transcription that occur in offspring both in somatic cells and in the germline. This work demonstrates that gamete-inherited histone marks are one mechanism through which epigenetic information can be transmitted to the next generation to influence transcription and development in offspring. A deeper understanding of this process will inform our broader understanding of how changes to the parental epigenome may influence the development and health of future generations.

Chapter 1: Introduction

All multicellular organisms arise from an individual totipotent cell, the newly fertilized embryo. A fundamental question in biology is how that single cell can give rise to the diverse cell types that make up a complex, multicellular organism. This is especially perplexing given that each cell carries the same genetic material. If not genetically distinct, what differentiates one cell from another?

Unequal provisioning of maternal products plays an early and essential role in distinguishing one blastomere from another. However, a key point in lineage specification occurs when cellular products switch from being maternally provided to zygotically produced. As the zygotic genome is activated, lineage-specific transcriptional programs are launched and the chromatin, in which the genome is packaged, becomes epigenetically modified to reflect the distinct transcription patterns of different cell types. When passed to daughter cells, those chromatin states constrain transcription patterns in a lineage-specific manner. In this way, chromatin states can serve to transmit a memory of transcription patterns through cell division and, as this thesis will discuss, across generations as well.

I: Transcription shapes and is shaped by the chromatin landscape

Transcriptional patterns can both shape and be shaped by the chromatin landscape. Whether a given histone modification will encourage or discourage transcription depends on multiple factors. Some histone modifications, such as acetylation, can directly affect chromatin compaction by reducing the positive

charges on histone lysines and therefore the association between histone tails and the negatively charged DNA backbone (Castillo *et al.* 2017). For this reason, histone acetylation is most often associated with open chromatin states and transcriptional competency. Other modifications, such as methylation, primarily impact chromatin accessibility through the activity of effector proteins that 'read' the modification. Histone methylation can function to either increase or decrease chromatin compaction in a process that depends more on the activity of effector proteins rather than on a direct effect on histone association with DNA. In general, histone modifications are described by 1) which histone is modified 2) at what residue and 3) with what modification. This work will largely focus on the roles of two histone modifications, H3K36me3 and H3K27me3, and the enzymes that generate them in the model organism *C. elegans*.

H3K36me3 consists of a trimethyl group added to lysine 36 of histone H3 tails. This modification is present across metazoans and is a characteristic mark of active transcription. This is due to transcription-coupled deposition of H3K36me3 by SETD2 homologs. SETD2 associates with elongating RNA Polymerase II (PolII) and generates H3K36me3 in the body of genes, which in turn recruits histone deacetylases (HDACs) and promotes transcriptional fidelity by resetting chromatin in the wake of transcription by PolII (Huang and Zhu 2018). In addition to its role in transcriptional fidelity, this mark also functions in other aspects of chromatin biology, such as splicing and DNA repair (Huang and Zhu 2018). Although this mark is associated with active genes, it is not necessarily an 'activating' mark. However,

H3K36me3 has been shown to antagonize PRC2-mediated deposition of the repressive mark H3K27me3 both in vitro (Yuan *et al.* 2011; Schmitges *et al.* 2011) and in vivo (Gaydos *et al.* 2012; Kaneshiro *et al.* 2019). Consistent with this, H3K36me3 and H3K27me3 1) do not co-occur on histone tails (Voigt *et al.* 2012), 2) rarely occur within the same nucleosome (Voigt *et al.* 2012) and 3) form domains that are anti-correlated across the genome (Gaydos *et al.* 2012). This observation has led some to suggest that, in terms of transcriptional regulation, it may be more appropriate to think of H3K36me3 and the enzymes that generate it as anti-repressors rather than activators (Dorighi and Tamkun 2013; Klymenko and Muller 2004).

Like H3K36me3, H3K27me3 is a trimethyl modification on histone H3 tails at lysine 27 rather than lysine 36. H3K27me3 is a conserved mark of repression across metazoans, generated by the enzyme complex Polycomb Repressive Complex 2 (PRC2). Consistent with its repressive nature, this mark is found at repressed regions across the genome, both genic and intergenic regions (Pauler *et al.* 2009; Evans *et al.* 2016; Gaydos *et al.* 2012). Two key questions in the field are what recruits PRC2 to particular genomic locations and is H3K27me3 the cause or consequence of repression. First, I will discuss the role of H3K27me3 in the establishment and maintenance of gene repression. Then I will discuss H3K27me3 deposition as a consequence of gene repression.

Work in *Drosophila* identified sequence-specific elements that recruit PRC2 to specific loci in the genome (Simon *et al.* 1993). Targeting PRC2 via these Polycomb Response Elements (PREs) results in the deposition of H3K27me3 and the

stable repression of developmental genes in particular cell lineages throughout the life of the organism (Francis and Kingston 2001). This process is necessary for proper body plan patterning (Struhl 1981). As a result, defects in this process, lead to developmental abnormalities and malformations. Through experimental insertion and removal of PREs, researchers have shown that PREs recruit PRC2, which drives the deposition of H3K27me3 and repression of nearby reporter genes (Hansen *et al.* 2008). Upon removal of PREs, gene repression can persist through a few rounds of cell division and depends on the level of H3K27me3 that persists at the locus (Hansen *et al.* 2008; Laprell *et al.* 2017; Coleman and Struhl 2017). These findings indicate that PREs are sufficient to recruit PRC2 which establishes H3K27me3 and gene repression. However, once established, H3K27me3 can carry a memory of the repressed state through cell division independent of sequence-specific elements such as PREs. The discovery of PREs and their utility in probing the role of PRC2 in gene repression has been invaluable to the understanding of H3K27me3-mediated gene repression. However, sequence-specific elements, such as PREs, have not been found in other species. The question of how PRC2 is recruited to particular regions of the genome in species other than *Drosophila* remains an active area of research.

As discussed, PRC2 can be directly recruited to regions of the genome via sequence-specific elements such as PREs. However, such elements are not required to localize PRC2 to chromatin. In fact, PRC2 binds to DNA in a non-specific manner in which PRC2 primarily interacts with the DNA backbone (Wang *et al.* 2017; Poepsel *et al.* 2018). The current view is that PRC2-mediated deposition of H3K27me3 is

regulated at the level of where PRC2 binds the genome and also by interactions of PRC2 with local features that either inhibit or enhance its activity. Features of active genes such as the presence of nascent RNAs and active histone marks directly inhibit PRC2 enzymatic activity (Wang *et al.* 2017; Yuan *et al.* 2011; Schmitges *et al.* 2011). In contrast, features often associated with repressed chromatin such as densely packed nucleosomes and the presence of H3K27me3, stimulate its enzymatic activity (Yuan *et al.* 2012; Margueron *et al.* 2009; Poepsel *et al.*2018). In these manners, H3K27me3 deposition can occur as a consequence of transcriptional quiescence.

Evidence that the H3K27me3 mark itself is repressive comes from the finding that replacing lysine 27 on histone H3 tails with a non-methylatable amino acid causes the same loss of HOX gene repression observed in PRC2 mutants (Pengelly *et al.* 2013). Despite the many observations across diverse organisms that H3K27me3 is repressive, H3K27me3-enriched loci are not impervious to gene activation, as strong signaling cascades or tethering of strong activators can force transcription to occur at these locations (Hosogane *et al.* 2016). These events rewrite the local chromatin landscape, resulting in the removal of H3K27me3 and the accumulation of H3K36me3 (Hosogane *et al.* 2016). Thus, it is not surprising that in most circumstances, the current transcriptional states of cells are largely reflected in the distributions of these two marks. These observations suggest that once established, the distribution of H3K27me3 across the genome will maintain stable repressive states unless acted upon by strong drivers of transcription. In addition,

transcriptionally active domains will remain transcriptionally permissive unless factors directly recruit PRC2 to those regions.

II: Passage of chromatin landscapes through cell division

Can the chromatin landscape of a parent cell transmit a memory of transcriptional states to daughter cells? When a cell replicates its genome, marked histones are disassembled from the parental chromatin to allow for passage of the replication fork. These histones, in particular H3/H4 tetramers, are then distributed to each newly replicated daughter chromatid (Petryk *et al.* 2018; Yu *et al.* 2018). This results in newly replicated DNA that is packaged with a half dose of marked parental histones interspersed with newly synthesized histones. To accurately replicate the parental chromatin landscape requires that 1) marked parental histones are incorporated in a manner that preserves their genomic location and 2) newly synthesized histones become appropriately modified to restore full levels of marking.

Recent studies have shown that the genomic location of marked parental histones is retained in newly synthesized DNA regardless of whether these histones bear marks associated with active domains (H3K4me3 and H3K36me3) or repressed domains (H3K9me3 and H3K27me3) (Alabert *et al.* 2011; Reveron-Gomez *et al.* 2018; Escobar *et al.* (preprint)). The presence of marked parental histones is thought to be important for directing the modification of adjacent unmarked (newly synthesized) histones by histone modifying enzymes including PRC2, which generates H3K27me2/3, (Margueron *et al.* 2009; Poepsel *et al.* 2018) and NSD

6

homologs, which generate H3K36me2/3 (Vermeulen *et al.* 2010; Sankaran *et al.*

2016).

Recent studies have also shown that the level of histone marks in nascent

chromatin is restored by appropriately modifying newly synthesized histones,

although the rate at which the parental level of marking is restored depends on the

specific modification and its location (Alabert *et al.* 2011; Reveron-Gomez *et al.*

2018). In general, the active mark H3K4me3 is restored rapidly and prior to mitosis

(Reveron-Gomez *et al.* 2018). In contrast, the repressive mark H3K27me3 is among

the slowest to recover after genome duplication and is not fully restored until after

cell division (Alabert *et al.* 2011; Reveron-Gomez *et al.* 2018). These kinetics suggest

that the fidelity with which a parental chromatin landscape is restored during cell

division likely depends on the rate of proliferation and the activity of histone

modifying enzymes. These factors may in turn influence the stability-plasticity

balance of cell identity during cell division.

III: Passage of epigenetic information across generations

The preceding section focuses on transmission of transcriptional memory

from parent cells to daughter cells via chromatin states. Such a memory can preserve

the transcription pattern and therefore cellular identity when parent and daughter cells

are similar. The mechanism and utility of transmitting a memory of transcription

patterns across generations are less clear.

Sperm and oocytes are the conduits through which genetic and epigenetic material is passed from one generation to the next. There are three main forms in which epigenetic information has been shown to cross generational boundaries: 1) DNA methylation, 2) non-coding RNAs (ncRNAs), and 3) histone modifications. Gamete-inherited DNA methylation and ncRNAs have been demonstrated to influence offspring transcription with developmental consequences (Millership *et al.* 2019; Acharya *et al.* 2017). In contrast, similar evidence for gamete-inherited histone modifications influencing offspring transcription and development have remained elusive. As discussed earlier, the chromatin landscape can both shape and be shaped by transcription. This has made it difficult to interpret changes in histone marks as being the cause or the consequence of changes in transcription. Although some studies provide suggestive evidence that sperm-inherited histone marks may influence transcription in offspring (Teperek *et al.* 2016; Siklenka *et al.* 2015), the work discussed in this thesis demonstrates it (Kaneshiro *et al.* 2019). In brief, I have shown that sperm-inherited histone marks can directly influence offspring transcription with consequences to germline health and function.

IV: *C. elegans* as a model for studying histone-based epigenetic inheritance

The germline produces the specialized cells, sperm and oocyte, that carry epigenetic information from one generation to the next. *C. elegans* is a remarkably tractable and powerful system to study both germline development and epigenetic inheritance.

C. elegans is transparent at all life stages and has an invariant cell lineage (Sulston 1983). This allows the entire germ lineage to be tracked from the 1-cell embryo to the primordial germ cells (PGCs) of newly hatched larvae to the proliferating, gamete-producing germline of adults (Figure 1). In addition, these features enable for the direct visual assessment of fertility in live worms based on the presence or absence of embryos in the uterus.
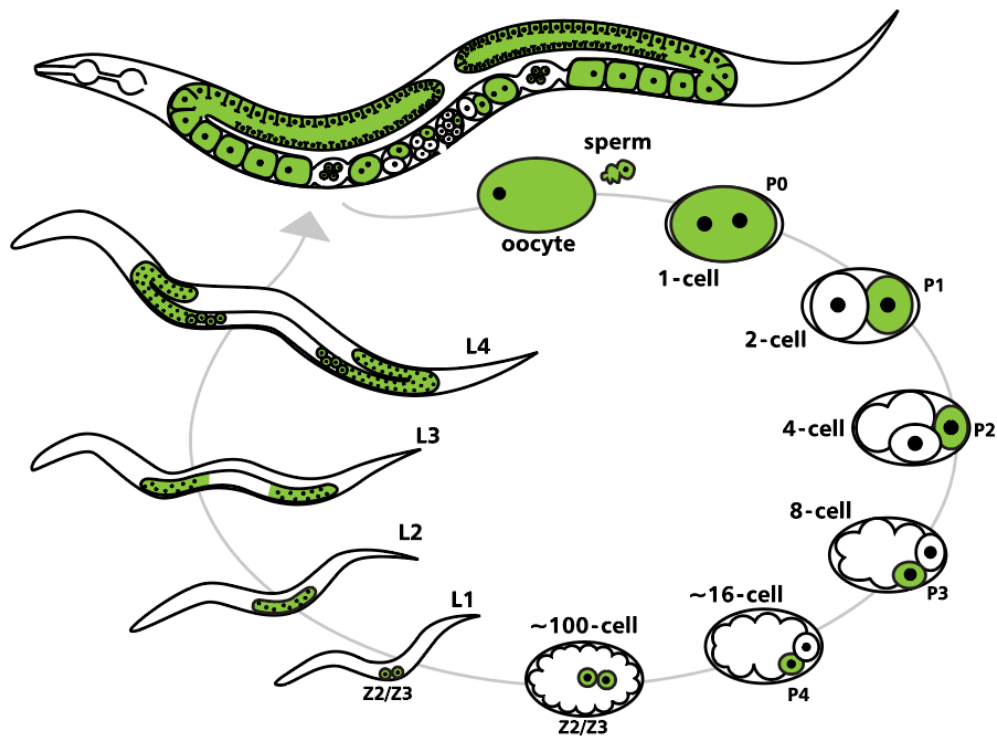
Figure 1. The *C. elegans* hermaphrodite life cycle
C. elegans hermaphrodites produce both sperm and oocytes, which unite to form the 1-cell zygote. Beginning at the 1st cell division, the germ lineage (in green) is asymmetrically segregated away from somatic blastomeres (in white). After 4 cell divisions, the germline blastomere P4 is born. Around the 100-cell stage, P4 divides to give rise to Z2/Z3, the primordial germ cells (PGCs) responsible for generating the entire germline.

Another powerful feature of this model system is the ability to propagate clonally. This enables the production of genetically identical worms, which provides the ideal backdrop to investigate consequences of altered epigenetic inheritance independent of genetic differences (Pembrey *et al.* 2014).

In addition to reproducing clonally, *C. elegans* can also reproduce via mating. Oocytes contribute many factors needed for embryonic development, making it difficult to resolve developmental consequences due to altered epigenetic inheritance versus due to altered maternal provisioning (Pembrey *et al.* 2014). For this reason, many studies in the field of epigenetic inheritance focus on the contributions of epigenetic information inherited through the paternal lineage via sperm (Pembrey *et al.* 2014). In mammals, including humans, sperm DNA is largely packaged with non-histone proteins called protamines (Bao and Bedford 2016). Although there is debate over the degree to which histones are retained in mammalian sperm (ranging from 1% to 10% histone retention), there is agreement that modified histones are retained at developmentally important loci (Yamaguchi *et al.* 2018; Yoshida *et al.* 2018; Jung *et al.* 2017). This observation has led many to hypothesize that marking of the sperm genome with histone modifications serves important roles in embryo development. However, the limited degree to which histones are present in mammalian sperm makes it a challenging system in which to investigate histone-based epigenetic inheritance. Unlike mammals, *C. elegans* sperm retain histones genome-wide (Tabuchi *et al.* 2018). Furthermore, the *C. elegans* sperm chromatin landscape has been thoroughly elucidated, including the histone modifications that are the focus of

this thesis (Tabuchi *et al.* 2018, Samson *et al.* 2014). These features make *C. elegans* a powerful model system to investigate how paternally inherited histone marking influences offspring transcription and development.

V: Model for the transmission of histone-based epigenetic memory across generations in *C. elegans*

The following section summarizes important findings that have contributed to a model that in *C. elegans* a histone-based memory of transcription patterns is transmitted from parent germ cells, sperm and oocyte, to the primordial germ cells of offspring, where the memory may promote appropriate transcription and prevent inappropriate transcription in the developing germline.

-H3K27me3 and H3K36me3 are inherited on sperm and oocyte genomes.

Sperm and oocyte chromosomes are delivered to the embryo bearing H3K27me3 and H3K36me3 (Gaydos *et al.* 2014; Kreher *et al.* 2018). These gamete-inherited marks are propagated to daughter chromatids during DNA replication, as judged by analyzing embryos that inherited marks on sperm chromosomes but that lacked the enzyme needed to generate new marks (Gaydos *et al.* 2014; Kreher *et al.* 2018). Each cell division reduced H3K27me3 levels by ~50%. This is consistent with the dilution of parental, in this case sperm-inherited, histones that occurs during DNA replication and suggests that, in early embryos, H3K27me3 is propagated through cell divisions where it can guide restoration of the H3K27me3 landscape after genome

duplication (Gaydos *et al.* 2014; Margueron *et al.* 2009; Poepsel *et al.* 2018). In contrast to H3K27me3, embryos that lacked the enzyme to generate new H3K36me3 showed a greater than 50% reduction in H3K36me3 with each cell division (Kreher *et al.* 2018). This suggests that in addition to replication-coupled dilution, sperm-inherited H3K36me3 is actively removed by histone turnover and/or the activity of demethylases.

-The inability to maintain H3K27me3 or H3K36me3 during embryogenesis results in sterility.

Maintenance of H3K27me3 and H3K36me3 during early embryogenesis requires the enzymes that generate them (Bender *et al.* 2004; Bender *et al.* 2006). H3K27me3 is generated by MES-2/3/6, which constitutes the worm version of PRC2, while H3K36me3 is generated by MES-4 (Korf *et al* 1998; Holdeman *et al.* 1998; Xu *et al.* 2001; Bender *et al.* 2004; Ketel *et al.* 2005; Bender *et al.* 2006). All of these MES (for Maternal Effect Sterile) proteins are provided maternally via the oocyte. If a mother fails to provide any one of these, the nascent germ cells in her offspring die and the offspring grow up to be sterile adults (Capowski *et al.* 1991; Paulsen *et al.* 1995).

-PRC2 and MES-4 propagate inherited states of H3K27me3 and H3K36me3 to the PGCs.

An elegant and helpful way to track the fate of sperm- vs oocyte-inherited chromosomes during cell divisions is to keep them in separate nuclei. I collaborated with researchers at the NIH to show that polo-like kinase 1 (PLK-1) is required for the merging of parental genomes at the first and subsequent mitoses (Rahman *et al.* 2015). As a result, *plk-1* mutant embryos keep sperm- and oocyte-inherited genomes in separate nuclei for many divisions. I took advantage of this genetic tool to demonstrate that sperm- and oocyte-inherited chromosomes maintain inherited states of H3K27me3 through cell divisions in the early embryo (Kaneshiro *et al.* 2019). To do this, I generated embryos that inherited H3K27me3 on oocyte-inherited but not sperm-inherited chromosomes. I refer to these embryos as *K27me3 M+P-* (M+ for maternal H3K27me3(+) and P- for paternal H3K27me3(-)). In *K27me3 M+P-* embryos, PRC2 maintains high levels of H3K27me3 on oocyte chromosomes without newly depositing H3K27me3 on sperm chromosomes (Figure 2) (Gaydos *et al.* 2014; Kaneshiro *et al.* 2019). MES-4 similarly maintains gamete-inherited states of H3K36me3 during embryogenesis (Kreher *et al.* 2018). Thus, these enzymes function to propagate inherited states of the marks they make to the PGCs.
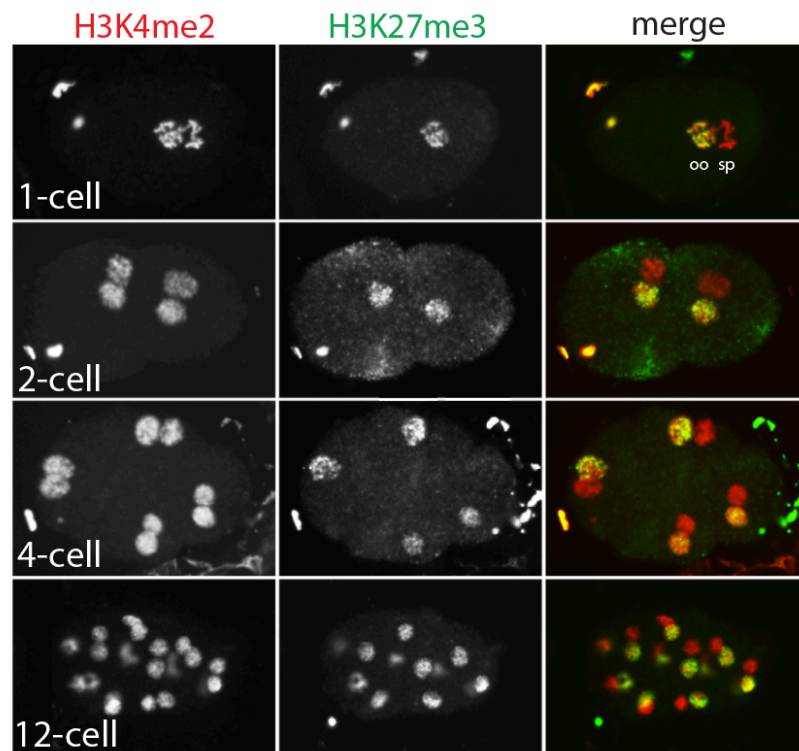
Figure 2. Gamete-inherited states of H3K27me3 are maintained during embryogenesis Immunofluorescence images of *K27me3 M+P-* embryos in a *plk-1* mutant background, which keeps sperm- and oocyte-inherited chromosomes in separate nuclei for many divisions. *K27me3 M+P-* embryos inherit H3K27me3 on oocyte chromosomes but not on sperm chromosomes. H3K4me2 (shown in red) stains both sperm-and oocyte-inherited chromosomes. H3K27me3 (shown in green) only stains oocyte-inherited, chromosomes. Top panel: oo and sp indicate the oocyte- and sperm-inherited pronucleus, respectively. Images of 1-cell and 4-cell embryos were previously published in Rahman *et al.* 2015.

The distributions of H3K27me3 and H3K36me3 in embryos reflect germline transcription patterns

In wild-type embryos, the distribution of H3K27me3 and H3K36me3 across the genome reflect transcription patterns of the parental germline (Gaydos *et al.* 2012; Rechtsteiner *et al.* 2010). For instance, H3K27me3 is enriched over regions of the genome that were repressed in the parental germline, while H3K36me3 is enriched

over genes that were expressed in the parental germline, even genes that were expressed *specifically* in the germline and are no longer being expressed in embryos.

Altogether these observations indicate that the chromatin landscape of sperm and oocyte genomes carry a memory of germline transcription patterns. This memory is epigenetically propagated to the PGCs by the activities of PRC2 and MES-4 in a process that is essential for germline survival and development.

Chapter 2: Sperm-inherited H3K27me3 impacts offspring transcription and development in *C. elegans*

I: Introduction

There is growing awareness that development and health are influenced by epigenetic information passed from parents to offspring (Pembrey *et al.* 2014; Vagero *et al.* 2018). The gametes, sperm and oocyte, are the conduits through which epigenetic information is passed from one generation to the next. Investigating epigenetic inheritance via the maternal line is complicated by the many non-epigenetic factors the oocyte provides to the embryo (Pembrey *et al.* 2014). For this reason, the field of epigenetic inheritance largely focuses on transmission via sperm. DNA methylation and non-coding RNAs (ncRNAs) are epigenetic carriers that are transmitted by sperm to the embryo and that have been shown to influence gene expression in offspring (Adalsteinsson & Ferguson-Smith 2014; Gapp & Bohacek 2018). Modified histones can also be passed, but their role in intergenerational and transgenerational epigenetic inheritance is less clear. In mammals, the sperm genome

15

is primarily packaged with protamines and, while there is current debate about the degree to which the sperm genome retains histones, there is consensus that marked histones are found on developmentally important loci (Yamaguchi *et al.* 2018; Yoshida *et al.* 2018; Jung *et al.* 2017). Despite this observation and mounting suggestive evidence (Teperek *et al.* 2016; Siklenka *et al.* 2015; Klosin *et al.* 2017), establishing a causal role for sperm-inherited histone marks in regulating offspring transcription and development has been an ongoing challenge. We took advantage of features of *C. elegans* to tackle this issue.

We previously showed that *C. elegans* sperm retain nucleosomes and histone marking genome-wide (Tabuchi *et al.* 2018) and that *C. elegans* Polycomb Repressive Complex 2 (PRC2) maintains inherited states of H3K27me3 during embryogenesis (Gaydos *et al.* 2014*).* In wild-type embryos H3K27me3 is enriched over genes that were silent in the parental germline (Rechtsteiner *et al.* 2010; Gaydos *et al.* 2012). H2K27me3 marking inherited from hermaphrodite parent worms is essential for germline development in offspring, since hermaphrodite parents lacking PRC2 generate offspring in which the primordial germ cells die during early larval development (Capowski *et al.* 1991). We thus hypothesize that the transmission of chromatin states from parent germ cells via sperm and oocyte to the primordial germ cells in offspring protects germline-appropriate gene expression patterns in the developing and adult germline.

In this work we investigate how chromatin states inherited from parents are maintained in offspring and whether inherited states are important for offspring

transcription and development. We elucidate a mechanism through which an inherited H3K27me3(-) state is propagated from parent germ cells (sperm) to offspring germ cells. We show that inheriting a sperm genome lacking the repressive mark H3K27me3 results in derepression of many genes for somatic development, especially neuronal genes, in offspring germlines. This results in germ cells that in a sensitized genetic background lose their germ cell identity and adopt a neuronal fate. Taken together, these findings establish a cause-effect relationship between sperm-inherited histone marks and offspring transcription and development in *C. elegans*.

II: Results

Maintenance of an inherited H3K27me3(-) state requires MES-4.

When *C. elegans* embryos inherit some chromosomes with and some chromosomes without H3K27me3, PRC2 maintains inherited states by 1) restoring levels of H3K27me3 on H3K27me3(+) chromosomes after DNA replication and 2) failing to *de novo* methylate H3K27me3(-) chromosomes (Gaydos *et al.* 2014). The ability of PRC2 to restore levels of H3K27me3 after genome duplication is likely explained by the EED subunit of PRC2 (MES-6 in worms) binding to H3K27me3 and stimulating the methyltransferase activity of the EZH2 subunit (MES-2 in worms) (Margueron *et al.* 2009; Poepsel *et al.* 2018). How PRC2 is prevented from *de novo* methylating chromosomes inherited lacking H3K27me3 is less clear. One possibility is that chromosomes lacking H3K27me3 are unable to recruit and stimulate PRC2 activity. Another possibility is that these chromosomes bear an opposing mark that

antagonizes PRC2 activity. Likely candidates for antagonizing PRC2 activity are histone marks associated with gene expression, hereafter referred to as active marks, and their corresponding histone modifiers (Yuan *et al.* 2011; Schmitges *et al.* 2011).

To test if active marks prevent PRC2 from *de novo* methylating sperm-inherited H3K27me3(-) chromosomes during early embryogenesis, we monitored sperm chromosomes through rounds of cell division in embryos lacking a maternal load of MES-4 or SET-2, which generate the active marks H3K36me2/3 (Rechtsteiner *et al.* 2010; Bender *et al.* 2006) and H3K4me2/3 (Xiao *et al.* 2011), respectively. We generated embryos that inherited H3K27me3(+) oocyte chromosomes and H3K27me3(-) sperm chromosomes by mating wild-type females with PRC2 mutant males (Figure 1a). To increase the likelihood that sperm chromosomes completely lacked H3K27me3, we used PRC2 mutant males whose parents also lacked PRC2. We call the offspring of wild-type mothers and PRC2 mutant fathers *K27me3 M+P-* (M for maternal H3K27me3(+), P for paternal H3K27me3(-)). We assessed whether sperm-inherited chromosomes retained the H3K27me3(-) state or acquired H3K27me3, by tracking the status of H3K27me3 on sperm-inherited chromosomes during early embryogenesis. To facilitate analysis of sperm versus oocyte chromosomes, we performed this analysis in a temperature-sensitive *plk-1* mutant background that keeps sperm- and oocyte-inherited chromosomes in separate nuclei for many divisions (Rahman *et al.* 2015) (Figure 1b).
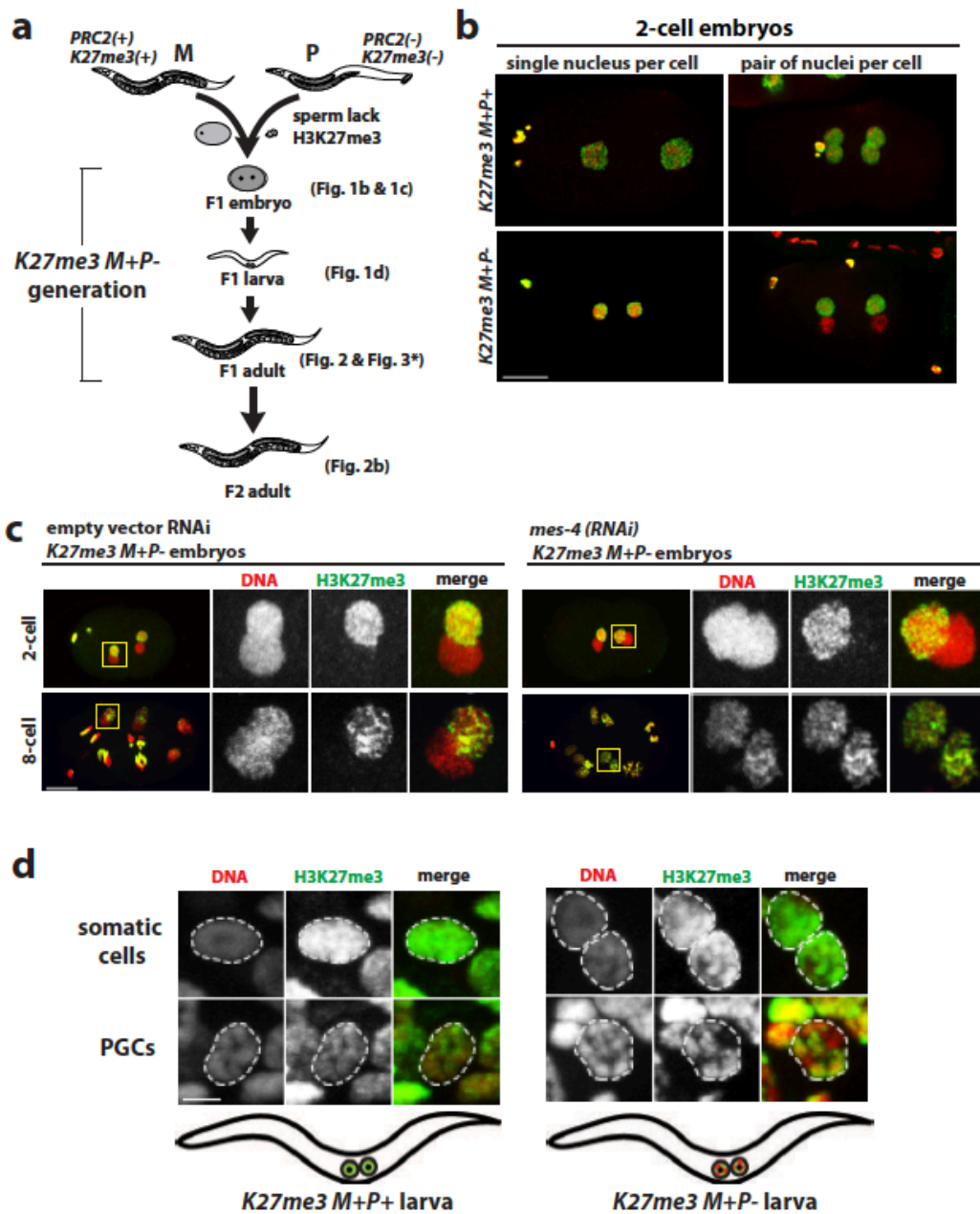
Figure 1. Maintenance of an inherited H3K27me3(-) chromatin state depends on antagonism of PRC2 by MES-4. a Diagram showing the cross between wild-type *PRC2(+) K27me3(+)* females and *mes-3/mes-3 PRC2(-) K27me3(-)* males that generates *K27me3 M+P-* F1 offspring. Maternal and paternal pronuclei are shown in the F1 1-cell embryo. The 2 primordial germ cells (PGCs) are shown in the newly hatched F1 larva. F1 hermaphrodites self-cross to produce F2 offspring. Figures that analyzed each generation/stage are indicated in parentheses. M=maternal; P=paternal; *= sensitized genetic background. Hybrid F1s were generated by mating Hawaii

mothers with Bristol fathers. **b** Images of wild-type 2-cell embryos (left panels) or *plk-1* 2-cell embryos (right panels) demonstrate that the H3K27me3(-) state of sperm chromosomes in *K27me3 M+P-* embryos (bottom panels) is easily monitored when sperm- and oocyte-inherited chromosomes are kept in separate nuclei in *plk-1* mutants. DAPI-stained DNA in red. H3K27me3 immunostaining in green. Scale bar represents 10μm. **c** 2-cell and 8-cell stage *K27me3 M+P-* embryos whose mothers were fed empty vector RNAi (control) or *mes-4(RNAi)* to knock down the maternal load of MES-4. DAPI-stained DNA in red. H3K27me3 immunostaining in green. Regions boxed in yellow in the left panels are shown at higher magnification in the right panels. Scale bar represents 10μm. **d** Somatic cells and PGCs in a *K27me3 M+P+* L1 (left) and a *K27me3 M+P-* L1 (right). L1 schematics show the observed staining pattern in PGCs for each genotype. DAPI-stained DNA in red. H3K27me3 immunostaining in green. Scale bar represents 2μm.

In all 2-cell *K27me3 M+P- plk-1* embryos, we observed that one nucleus in each cell contained all H3K27me3(+) chromosomes (oocyte-inherited) and the other nucleus contained H3K27me3(-) chromosomes (sperm-inherited). In control and SET-2-lacking embryos, this pattern persisted in all cells during the early divisions and in the P lineage (germ lineage) throughout embryogenesis (Figure 1c, 1d, Supplementary Figure 1). In contrast, in embryos depleted of MES-4, we observed gradual acquisition of H3K27me3 on the entire set of sperm chromosomes (Figure 1c, Supplementary Figure 1, Supplementary Table 1). At the 1-cell stage, all *K27me3 M+P-* embryos examined lacked detectable H3K27me3 on sperm chromosomes irrespective of maternal MES-4 (n=6 with maternal MES-4 and n=4 without maternal MES-4). In the presence of maternal MES-4, this state was maintained in all embryos up to the 4-cell stage (n=21) and in most embryos up to the 40-cell stage (n=15/19). In the absence of maternal MES-4, all 2-cell embryos examined acquired some H3K27me3 on sperm chromosomes (n=6). The levels of H3K27me3 on sperm and oocyte chromosomes became indistinguishable in embryos by the 4-cell stage (n=4)

and at later stages (n=15). This gradual acquisition of H3K27me3 matches the

gradual loss of H3K36me3 in the absence of maternal MES-4 (Kreher *et al.* 2018).

Thus, MES-4 inhibits PRC2 activity, presumably through MES-4-mediated

methylation of H3K36, to maintain the paternal H3K27me3(-) state. These findings

show that the H3K27me3(-) state is maintained by opposing PRC2 activity with

active marks and also suggest that PRC2 does not require H3K27me3 to seed its

activity, as observed in mouse embryonic stem cells (Hojfeldt *et al.* 2018).
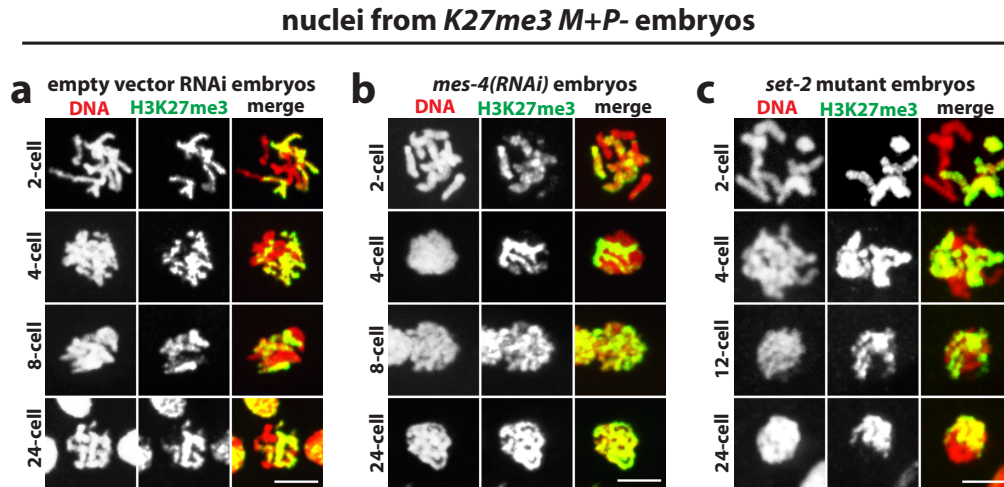
Supplementary Figure 1. Tracking the H3K27me3(-) state of sperm-inherited chromosomes in embryos depleted of the H3K36 methylator MES-4 or lacking the H3K4 methylator SET-2. Nuclei from *K27me3 M+P-* embryos from mothers treated with control RNAi (**a**), *mes-4* RNAi (**b**) or homozygous mutant for *set-2* (**c**). DAPI-stained DNA in red. H3K27me3 immunostaining in green. Scale bars represent 2μM.

Sperm chromatin states shape transcription in offspring.

Do gamete-inherited chromatin states influence gene expression in offspring?

This is a burning question in the field and one that remains elusive, in part because it

is difficult to determine cause versus consequence when comparing chromatin states

with transcriptional status. To establish a cause-effect relationship between inherited chromatin and offspring transcription requires that 1) genetically identical individuals display different transcriptional outcomes as a result of inheriting different chromatin states, 2) transcriptional changes must be demonstrated to occur in *cis*, thus eliminating alternative epigenetic carriers that function in *trans* (e.g. cytoplasmic ncRNAs), and 3) DNA methylation, which also functions in *cis,* must be eliminated as a potential mediator of the observed transcriptional changes. We met the third criterion by using the model system *C. elegans,* which lacks canonical DNA methylation. We met the remaining criteria by comparing transcription from sperm-inherited and oocyte-inherited alleles in genetically identical (hybrid *fem-2/+; mes-3/+*) worms (Figure 1a) that inherited the sperm genome with or without H3K27me3, which we call *K27me3 M+P+* and *K27me3 M+P-*.

We showed previously that *K27me3 M+P-* embryos have different chromatin states on their sperm- and oocyte-inherited genomes and that those states are maintained in all cells of early embryos (Gaydos *et al.* 2014). As embryos develop, the gamete-inherited chromatin states are rewritten in somatic cells but maintained through the five cell divisions that generate the two primordial germ cells (PGCs) present in newly hatched larvae (Gaydos *et al.* 2014) (Figure 1d). When newly hatched larvae begin to feed, the PGCs launch their transcriptional program and begin to proliferate. During larval development, H3K27me3 gradually begins to accumulate on sperm-inherited chromosomes (Gaydos *et al.* 2014). Absence of H3K27me3 on sperm chromosomes at the onset of germ cell transcription could result in the

activation of inappropriate genes. Once activated, a euchromatic state could be propagated at specific loci, sensitizing daughter germ cells to continued expression of inappropriate genes and potentially downstream target genes. We investigated whether global differences in inherited chromatin states in PGCs impact transcription in the developing germline. We used single-nucleotide polymorphisms (SNPs) between parental strains (from Bristol, England and from Hawaii) to compare transcription from sperm- and oocyte-inherited genomes. This allowed us to determine whether transcriptional changes occurred in *cis* (from sperm alleles specifically) or in *trans* (from both sperm and oocyte alleles).

Worms that inherited sperm chromosomes lacking H3K27me3 misexpressed genes in their germlines; 149 genes were upregulated and 116 genes were downregulated (FDR<0.1) in *K27me3 M+P-* compared to *K27me3 M+P+* worms (Figure 2a, Supplementary Figure 2, 3). Of the genes that were upregulated in *K27me3 M+P-* germlines, 96% are enriched for H3K27me3 in wild-type sperm (Tabuchi *et al* 2018) (Supplementary Figure 4). While *K27me3 M+P-* worms were typically fertile, they generated a significant proportion of sterile offspring (Figure 2b). These findings demonstrate that altering sperm-inherited chromatin states results in changes to germline transcription and function.

K27me3 M+P- vs M+P+ mRNA reads in germlines (a,c,d)

a read comparison of all coding genes (20,273)

149 genes sig up

116 genes sig down

b % sterile offspring from K27me3 M+P- vs M+P+ worms

p<0.001   p<0.001

M+P-   M+P+   M+P-   M+P+
n=5,142   n=4,228   n=4,143   n=5,550
replicate 1   replicate 2

c sperm vs oocyte read comparison of misregulated SNP-containing genes

29 genes

5 genes

26 genes

18 genes

d K27me3 M+P- vs M+P+ gamete-specific read comparison of misregulated genes (78)

sig up (34)   sp / oo $p = 1.5e^{-6}$

sig down (44)   sp / oo $p = 0.7$

Figure 2. Inheritance of sperm chromosomes lacking H3K27me3 alters germline transcription and function in offspring. a MA plot comparing germline transcripts (RNA-seq reads) from *K27me3 M+P-* hermaphrodites to genetically identical *K27me3 M+P+* hermaphrodites. Genes that are significantly upregulated or downregulated (p<0.1) are highlighted in red and blue, respectively. **b** Boxplots showing percent sterility of the F2-generation hermaphrodites from *K27me3 M+P-* and *K27me3 M+P+* F1 worms for each of 2 replicates. n=the number of F2 worms scored. Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend from the 2.5th to the 97.5th percentile; individual data points outside of these boundaries are shown as open circles. P-values were calculated using a Wilcoxon rank sum test. **c** Scatter plot comparing sperm- and oocyte-specific log2 fold changes of SNP-containing reads from significantly upregulated genes (in red) and downregulated genes (in blue). Genes upregulated in *cis* (>1.5 fold increase in reads from the sperm allele; <1.5 fold

24

increase from the oocyte allele) are plotted as dark red triangles. The numbers of misregulated genes above and below the diagonal line are indicated. **d** Boxplots showing sperm-specific (sp) and oocyte-specific (oo) $\log_2$ fold changes of significantly upregulated genes (34 genes sig up) and significantly downregulated genes (44 sig down). Downregulated genes (blue) are similarly downregulated from sperm and oocyte alleles, while upregulated genes (red) are significantly more upregulated from sperm alleles than oocyte alleles (i.e. pairwise comparisons of transcript reads from sperm versus oocyte alleles indicate that sperm allele reads are significantly more numerous than oocyte allele reads for upregulated genes but not for downregulated genes). Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend from the 2.5th to the 97.5th percentile; individual data points outside of these boundaries are shown as open circles. The waist indicates the 95% confidence interval for the medians. P values were generated using the paired student's t-test.
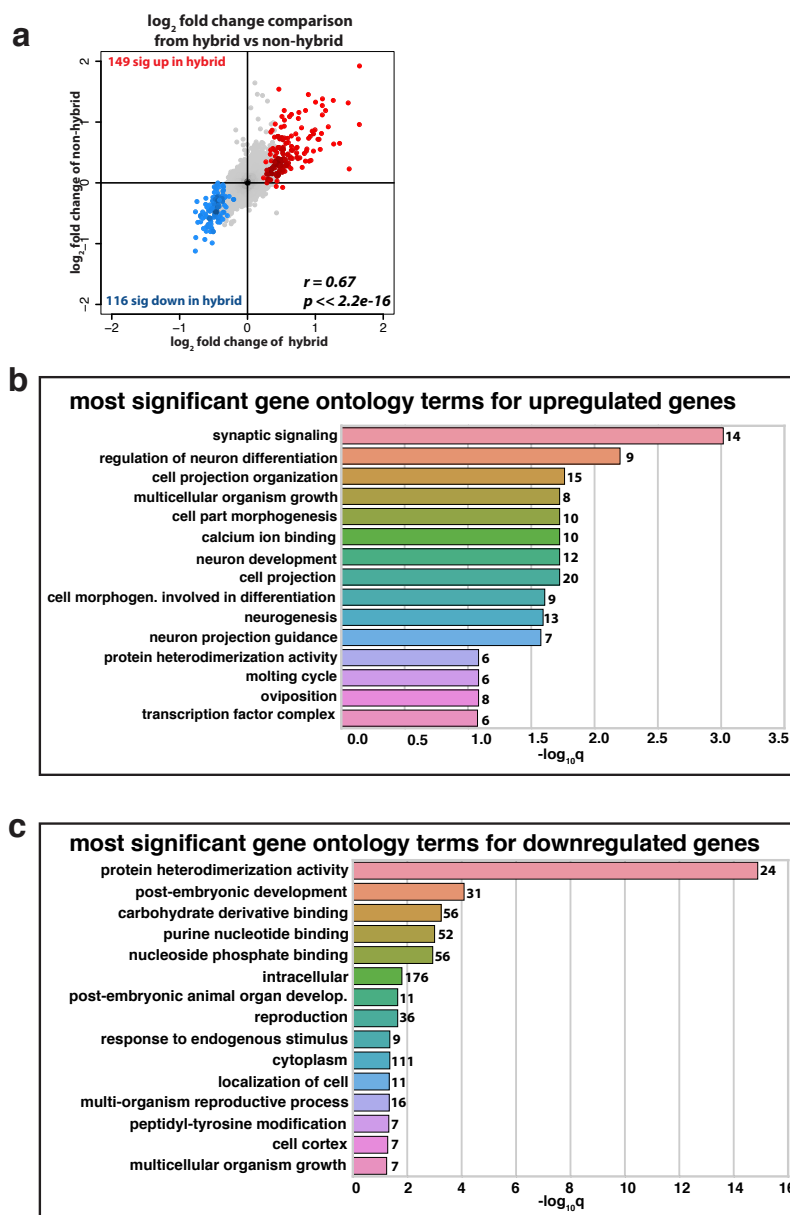
Supplementary Figure 2 Gene ontology analysis of significantly misregulated genes when sperm chromosomes are inherited lacking H3K27me3. **a** Log2 fold change comparison of *K27me3 M+P-* vs *K27me3 M+P+* germline transcripts from hybrid worms (mother and father are from different wild-type isolates) versus non-hybrid worms (both parents are from the Bristol wild-type isolate). **b, c** Most significant gene ontology terms of upregulated genes (**b**) and downregulated genes (**c**) from the germlines of non-hybrid *K27me3 M+P-* vs *K27me3 M+P+* worms. Numbers of genes in each category are listed to the right of the histogram bins.

Expression of genes assessed by reads overlapping SNPs between Bristol and Hawaii correlated well with expression assessed by using all reads mapping to a gene (Supplementary Figure 3a). Therefore, we could use SNPs to determine, for the 23% of misexpressed genes that contain a SNP, if transcription differences arose from the sperm allele, the oocyte allele, or both (Supplementary Data 1). Our analysis of SNP-containing transcript reads indicated that downregulated genes reflect decreased transcription from both the sperm and oocyte allele. However, upregulated genes reflect increased transcription primarily from the sperm allele (Figure 2c, d). Importantly, half of these genes were upregulated in *cis* with increased transcription from the sperm allele specifically (dark red triangles in Figure 2c). These findings demonstrate that differential marking on sperm- and oocyte-inherited genomes allows sperm- and oocyte-inherited alleles within the same cell to achieve different transcriptional outcomes in offspring. Given that *C. elegans* lacks canonical DNA methylation, this finding establishes a causal role for sperm-inherited H3K27me3 in regulating transcription in offspring. It also raises the possibility that environmentally induced changes to gamete marking could lead to transcriptional changes in offspring, with developmental consequences.

K27me3 M+P- vs M+P+ mRNA reads in germlines

a read comparison of SNP-containing genes (2,606)

34 genes sig up

44 genes sig down

log$_2$ fold change based on all reads (Y-axis)

log$_2$ fold change based on SNP-containing reads

b sperm vs oocyte read comparison of SNP-containing genes (2,606)

34 genes sig up

44 genes sig down

log$_2$ fold change of reads from sperm-inherited genome

log$_2$ fold change of reads from oocyte-inherited genome
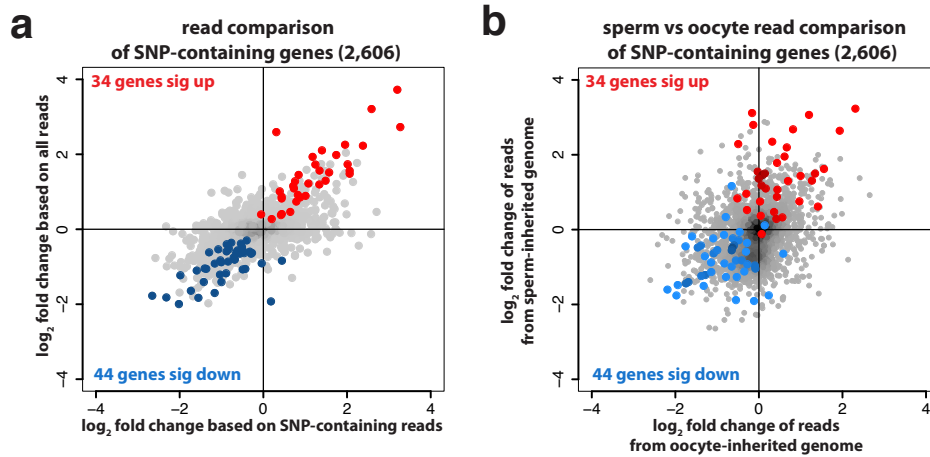
Supplementary Figure 3. Log2 fold change comparisons of SNP-containing genes. Log2 fold change comparison of *K27me3 M+P-* vs *K27me3 M+P+* germline transcripts for (**a**) SNP-containing reads (X-axis) vs all reads (Y-axis) and (**b**) reads that emanate from the oocyte-inherited (X-axis) vs sperm-inherited (Y-axis) genome. Genes that are significantly upregulated or downregulated (p<0.1) are highlighted in red and blue, respectively. Number of genes per category is shown in parentheses.

a log$_2$ z-scores of ChIP-seq reads from *C. elegans* sperm

76%   149 genes sig up   20%

0%   4%

H3K27me3

H3K36me3

b log$_2$ z-scores of ChIP-seq reads from *C. elegans* sperm

67%   15 genes sig up in *cis*   33%

0%   0%
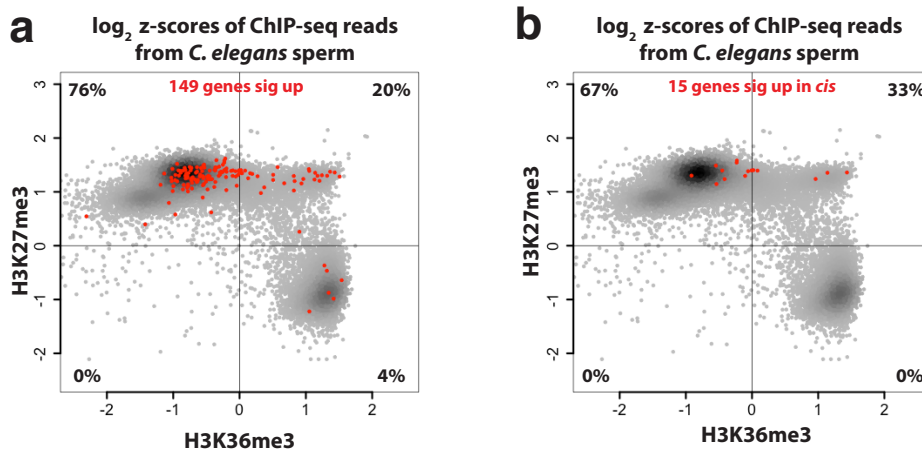
H3K27me3

H3K36me3

Supplementary Figure 4 Relative H3K27me3 versus H3K36me3 enrichment of upregulated genes in wildtype sperm. **a, b** Scatterplots of genes generated from published *C. elegans* sperm ChIP-seq experiments[10] highlighting genes that are significantly upregulated in *K27me3 M+P-* germlines. All genes that are significantly upregulated (FDR< 0.1) (**a**) and genes that are significantly upregulated in *cis* (dark

red triangles in Fig. 2c) (**b**) are highlighted in red. Percent of highlighted genes that fall within each quadrant is indicated. Axes are mean normalized gene-body ChIP-seq signals.

The genes upregulated from both sperm and oocyte alleles likely represent downstream targets of an initially misexpressed factor(s). The observation that these genes are more upregulated from sperm alleles than oocyte alleles led us to hypothesize that the presence of H3K27me3 on the oocyte genome acts as a barrier against transcriptional activation of germline-inappropriate genes. H3K27me3 on the oocyte-inherited genome would not be expected to protect it from factors that would shut down expression of germline-appropriate genes. This may explain the observation that downregulated genes were downregulated from both sperm and oocyte alleles. This led us to wonder if the global pattern of misexpression resulting from altered sperm chromatin was a concerted deviation from a germline program. To explore this, we performed gene ontology analysis on upregulated and downregulated genes. We found that downregulated genes are enriched for genes involved in reproduction and that upregulated genes are enriched for genes involved in somatic development, especially neuronal development (Supplementary Figure 2b, c). This suggests that *K27me3 M+P-* germlines are deviating from a germline program of gene expression toward a somatic program, perhaps favoring a neuronal program.

Sperm-inherited H3K27me3 protects germ cell fate in offspring.

To test whether germlines in *K27me3 M+P-* worms that inherited altered sperm chromatin are in fact transitioning toward a neuronal fate, we analyzed worms

carrying a pan-neuronal reporter gene, *unc-119::GFP*. We compared expression of

this reporter in genetically identical worms that inherited sperm chromatin with either

a wild-type pattern of H3K27me3, *K27me3 M+P+,* or lacking H3K27me3, *K27me3*

M+P-. As noted above, *K27me3 M+P-* and *K27me3 M+P+* worms are fertile*;* we

observed that fertile germlines do not detectably express *unc-119::GFP* (n>500) nor

do they significantly upregulate *unc-119* transcript. Performing this analysis in a

sensitized genetic background in which the maternal load of PRC2 was reduced by

half (*mes-3/+* heterozygous mother) rendered ~30% of *K27me3 M+P-* worms sterile

(n=267/894); <0.3% of *K27me3 M+P+* worms were rendered sterile (n=2/870). In

this background, we found that the *unc-119::GFP* neuronal reporter was expressed in

a significant percentage of sterile *K27me3 M+P-* adult germlines (Supplementary

Table 1). Notably, expression of the neuronal reporter increased with age, from 16%

in sterile day 1 adults to 31% in sterile day 2 adults. Expression of the neuronal

reporter was rarely detected (<0.1%) in germlines from late-stage larvae (L4)

(n=1/146) and never observed in the germlines from the rare sterile *K27me3 M+P+*

worms. These findings indicate that deviation from a germline program in *K27me3*

M+P- worms is progressive and strongly suggest that loss of H3K27me3 from the

sperm-inherited genome results in a failure to maintain germ cell identity rather than

a failure to establish it.

worms analyzed		GFP+
M+P- L4 (n=146)		**0.07%**
M+P- **sterile** (n=267/894)	**Day 1**	**16%**
	Day 2	**31%**
M+P+ **sterile** (n=2/870)	**Day 1**	**0%**
	Day 2	**0%**

Supplementary Table 1. Percentage of worms expressing the neuronal marker *unc-119::gfp* in their germlines. M+P- and M+P+ indicate *K27me3 M+P-* and *K27me3 M+P+* genotypes. Number of worms scored is shown in parentheses. Day 1 and Day 2 of the same genotype represent the same worms scored on day 1 and then again on day 2 of adulthood.
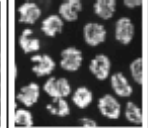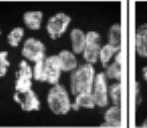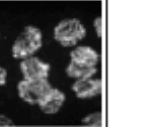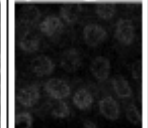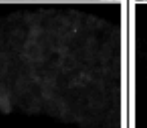
To test whether neuronal genes are derepressed from endogenous loci in *K27me3 M+P-* germlines, we stained UNC-119::GFP(+) *K27me3 M+P-* germlines for UNC-64, a neuronally expressed plasma membrane protein involved in synaptic vesicle fusion (Saifee *et al.* 1998). We detected this protein in all of the germlines analyzed (n=6) (Figure 3c). To test for loss of expression of genes involved in reproduction, we stained sterile *K27me3 M+P-* germlines for the germline proteins PGL-1 (a component of germ granules) (Kawasaki *et al.* 1998) and HTP-3 (a component of the synaptonemal complex) (MacQueen *et al.* 2005). We found that PGL-1 and HTP-3 were below detection (or nearly so) in 100% (n=27) and 28% (n=60) of sterile *K27me3 M+P-* germlines analyzed, respectively, regardless of whether or not they expressed the neuronal GFP reporter (Figure 3d, e). Although only 28% of sterile *K27me3 M+P-* germlines lacked detectable HTP-3 staining, the

31

remaining 72% of germlines analyzed displayed abnormal HTP-3 patterns (Figure

3d). Of note, germlines expressing the neuronal UNC-119::GFP reporter consistently

displayed more severe HTP-3 phenotypes than germlines that were UNC-119::GFP(-)

(Figure 3d). Consistent with *K27me3 M+P-* germlines adopting a neuronal identity,

we readily observed patches of highly concentrated GFP signal in UNC-119::GFP(+)

germlines, indicating that the normally syncytial germ cells had cellularized. In some

cases, we observed axo-dendritic structures emanating from these bright GFP(+) cells

(Figure 3b, c). Our findings demonstrate that absence of H3K27me3 from the sperm-

inherited genome results in the derepression of many somatic genes, especially

neuronal genes, predominantly from sperm-inherited alleles. In a sensitized

background, lack of H3K27me3 on the sperm-inherited genome results in germlines

that transition away from a germline program and toward a neuronal program.

a

DNA

UNC-119::GFP

DNA UNC-119::GFP

b

UNC-119::GFP

c

control germline   *M+P-* germline

UNC-119::GFP UNC-64 DNA

UNC-119::GFP

UNC-64

UNC-119::GFP UNC-64 DNA
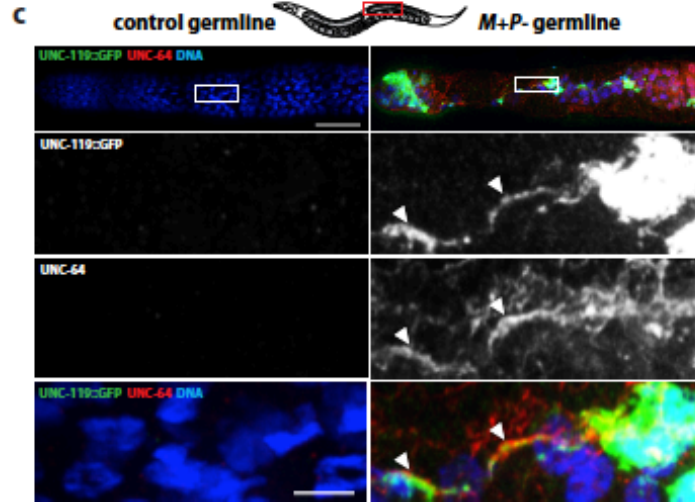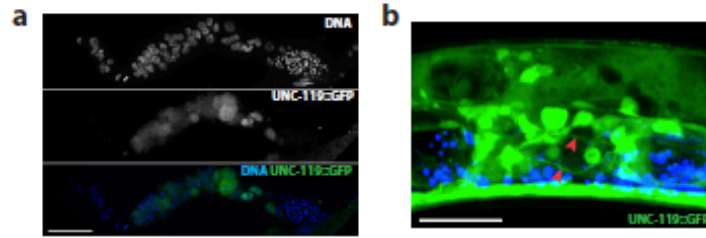
d percent of control and sterile *K27me3 M+P-* worms with different patterns of germline-specific protein HTP-3

	chromosomal	reduced chromosomal	not chromosomal	nearly below detection
control (n=31)	100%	-	-	-
M+P- GFP(-/+) (n=60)	-	22%	50%	28%
GFP(-) (n=31)	-	42%	35%	23%
GFP(+) (n=29)	-	-	66%	34%
pachytene germline nuclei	DNA / HTP-3			

e

100% control germlines (n=12)

DNA   PGL-1   DNA PGL-1

100% sterile M+P- germlines (n=27)
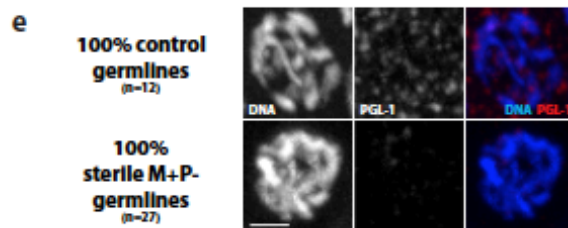
Figure 3. Inheritance of sperm chromosomes lacking H3K27me3 causes germ cells in offspring to transition toward a neuronal fate. a Immunofluorescence images of a sterile germline from a *K27me3 M+P-* worm in a sensitized genetic background (see text) and carrying an *unc-119::gfp* neuronal reporter transgene. DNA in blue. GFP in green. Scale bar represents 40μm. **b** Image of GFP in a live *K27me3 M+P-* worm carrying the *unc-119::gfp* reporter transgene. Arrowheads point to axo-dendritic processes extending from cells in the gonad with high levels of UNC-119::GFP. GFP in green. Blue is autofluorescent gut granules. Scale bar represents 10μm. **c** Immunofluorescence images of a control germline and a sterile germline from a *K27me3 M+P-* worm showing expression of the UNC-119::GFP reporter in green, immunostaining of the neuronally expressed protein UNC-64 in red, and DNA in blue. Worm schematic indicates the region of germline images. Arrowheads point to axo-dendritic processes co-stained for UNC-119::GFP and UNC-64 in the zoomed-in images. Scale bar in top panel represents 30μm. Scale bar in lower panel represents 5μm. **d, e** Summary of HTP-3 staining (**d**) and PGL-1 staining (**e**) observed in control germlines from fertile *K27me3 M+Z-; unc-119::gfp* worms and germlines from sterile *K27me3 M+P-* worms. For each staining category, representative immunofluorescence images of DNA and HTP-3 or PGL-1 are shown. Scale bar represents 10μm (d) and 2μm (e).

III: Discussion

Establishment and maintenance of chromatin states are critical during development and aging. During early development, inherited chromatin states must be rewritten to enable differentiation of different lineages. During aging, established chromatin states must be maintained to preserve cell-type-specific gene expression patterns and thus protect cell identity. Defects in these processes can lead to severe developmental defects and cancer. Whether chromatin states established in the parent and transmitted to offspring via sperm or oocyte influence the establishment and/or maintenance of cell fates in those offspring is a critical question in the field. Our findings demonstrate that in *C. elegans* sperm-inherited chromatin states influence germ cell identity in offspring. Because we observe reprogramming of germ cells toward neuronal cells in adult *K27me3 M+P-* germlines but only rarely in larval

34

K27me3 M+P- germlines, we conclude that offspring germ cells lose rather than fail to establish their germ cell identity. These findings reveal that in worms, sperm-inherited H3K27me3 is transmitted to offspring germ cells to prevent germline-inappropriate transcription and thus protect germ cell identity.

Altogether, our findings point to a model in which gamete-inherited patterns of H3K27me3 act as a barrier to inappropriate transcription; in germ cells, this barrier function is particularly important in suppressing transcription of factors that could drive germ cells toward alternative cell identities (Figure 4). In the germlines of worms that inherited the sperm genome lacking H3K27me3, we speculate that some genes become upregulated as a direct result of diminished H3K27me3 (Factor X in Figure 4), while other genes may be upregulated as a consequence of being the target of an upregulated gene. An alternative scenario is that Factor X genes are normally expressed in wild-type germ cells but prevented from activating germline-inappropriate target genes by H3K27me3 on those target genes. An example of the first scenario (Figure 4) is the *vab-7* gene, which encodes a neuronal transcription factor (Hobert 2013) *vab-7* is not expressed in wild-type germlines but becomes upregulated in *K27me3 M+P-* germlines. Examples of the alternative scenario are genes that encode neuronal specification factors (e.g. *die-1*, *nsy-7*, and *hif-1*) (Hobert 2013) that are highly expressed in both wild-type and *K27me3 M+P-* germlines. Regardless of the exact cascade of gene activation, our expression analysis of *K27me3 M+P-* germlines suggests a programmatic shift toward a neuronal program

and supports the hypothesis that an upstream factor(s) is altering gene expression in a concerted manner, especially from the H3K27me3-depleted sperm genome.
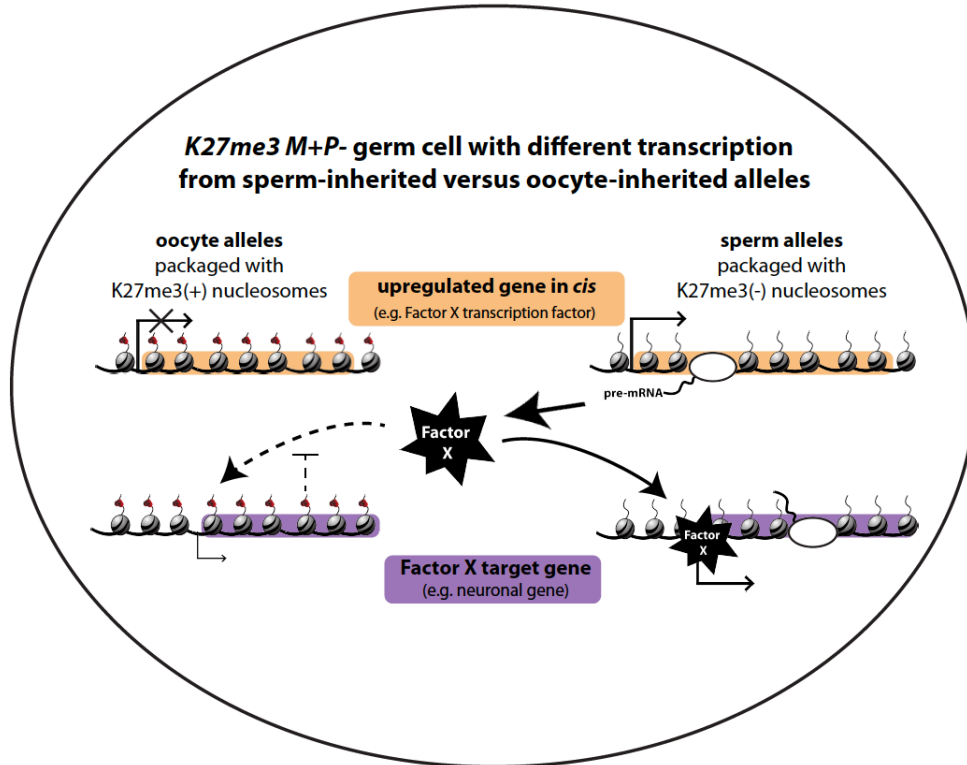
Figure 4. A model for transcriptional upregulation in germ cells that inherited sperm chromosomes lacking H3K27me3. The schematic is a germ cell nucleus from a *K27me3 M+P-* worm showing oocyte-inherited and sperm-inherited alleles for 2 representative genes: a transcription factor-encoding gene and a target gene of that transcription factor. Sperm-inherited alleles lack repressive H3K27me3 and as a result are prone to upregulate Transcription Factor X. If expressed, Transcription Factor X can then activate target genes strongly from H3K27me3(-) sperm alleles and weakly from H3K27me3(+) oocyte alleles. In this model, H3K27me3 would act as a barrier against 1) production of Factor X and 2) activation of Factor X target genes. In *K27me3 M+P+* worms, oocyte- and sperm-inherited alleles are similarly marked with H3K27me3.

Our findings establish a causal role for sperm-inherited histone marks influencing transcription in offspring and reveal their role in protecting cell identity. Together with observations that in mammals, sperm retain modified histones over

developmentally important genes (Yamaguchi *et al.* 2018; Yoshida *et al.* 2018; Jung *et al.* 2017), this work implicates sperm-inherited chromatin states as one mechanism through which environmentally induced changes to parental epigenomes can impact the development and health of future generations.

IV: Methods:

***C. elegans* growth conditions.** Worms were maintained on Nematode Growth Medium (NGM) agar plates spotted with *Escherichia coli* OP50 and maintained at 15˚C (*fem-2* and *plk-1* containing strains) or 20˚C (all other strains). To generate females, *fem-2* strains were raised at 24˚C until L4.

***C. elegans* strains.** All strains are in the N2 background except CB4856 and SS1292.

N2 wild-type isolate from Bristol, England

CB4856 wild-type isolate from Hawaii

SS1292 *fem-2(b245ts) III* - mutation CRISPR-engineered into CB4856

DH0245 *fem-2(b245ts) III*

SS1167 *mes-3(bn35) I/hT2*-GFP *(I;III); fem-2(b245ts) III*

SS1099 *mes-3(bn35) I/hT2*-GFP *(I;III); him-8(e1489) IV*

OCF65 *plk-1(or683) III; fog-1(e2121) unc-11(e47) I/hT2 [qIs48] (I:III);*
    itIs37[pAA64: pie-1p::mCHERRY::his-58 + unc-119(+)].

SS1424 *mes-3(bn35) I/hT2*-GFP*(I;III); edIs6[unc-119::GFP + rol-6(su1006)] IV*

SS1298 *set-2(bn129) fem-2(b245ts) III*

Fixing and staining embryos, larvae, and germlines. Specimens in the buffer (see below) used above were flattened between a slide and a coverslip by wicking away buffer. Slides were frozen in liquid nitrogen, coverslips were removed, then slides were fixed in methanol (10min) and acetone (10min) and allowed to air dry for up to 3 days. Slides were blocked (1.5% ovalbumin, 1.5% BSA in PBS) for 25min, then incubated with primary antibodies overnight at 4°C. Slides were washed 3X in PBST (0.1% tween in phosphate buffered saline pH7.2), blocked for 5min, incubated with secondary antibodies and DAPI at room temperature for 2hr, and washed 3X in PBST. Coverslips were mounted on slides with gelutol. Slides were allowed to dry for up to 3 days before being imaged.

Antibodies. Primary antibodies and their dilutions were as follows: 1:30,000 of 2.45mg/mL mouse anti-H3K27me3 (Kimura mAb 1E7 clone CMA323, Wako cat#309-95259), 1:3000 original stock solution of rabbit anti-H3K27me3 (C36B11 Cell Signaling MAb#9733 lot#C36B11), 1:4000 of 0.4mg/mL mouse anti-GFP (Roche cat#11 814 460 001 lot#14158300), 1:4000 rabbit anti-UNC-64 serum[22], 1:50,000 rabbit anti-PGL-1 serum[23], 1:500 guinea pig anti-HTP-3 serum[24]. Secondary antibodies conjugated to Alexa Fluor: 488 goat anti-mouse (Life cat#A11001), 488 goat anti-rabbit (Molecular Probes cat#A-11008), 594 anti-rabbit (Life cat#A11012), and 594 goat anti-guinea pig (Molecular Probes cat#A11076) were used at 1:300 with 0.05μg/ml 4',6-diamidino-2-phenylindole (DAPI).

Microscope set-up for immunocytochemistry. All images were acquired using the following: Yokogawa CSUX-1 spinning disk scanner, Nikon (Garden City, NY) TE2000-E inverted stand, Hamamatsu ImageEM X2 camera, solid state 405-nm, 488-nm and 561-nm laser lines, 435-485, 500-550 and 573-613 fluorescent filters, Nikon Plan Apo VC 60x/1.40 oil objective, Nikon Plan Apo 100x/1.40 oil objective, and Micro-Manager software.

Analysis of MES-4-depleted and SET-2 mutant embryos. *plk-1; fog-1 unc-11* females (Figure 1b, c) or *fem-2(ts)* or *set-2 fem-2(ts)* females (Supplementary Figure 1) were mated to *N2* (P+) or *mes-3 M-Z-; him-8* (P-) males (M- for no maternal load of MES-3, Z- for no zygotic synthesis of MES-3) on plates with *mes-4* RNAi, empty vector RNAi, or OP50 (Supplementary Figure 1c) and incubated overnight at 23˚C (Figure 1b, c) or 20˚C (Supplementary Figure 1). Embryos were dissected from mated females in egg buffer (25mM HEPES pH 7.4, 118mM NaCl, 48mM KCl, 2mM $CaCl_2$, 2mM $MgCl_2$) with 1mM levamisole on polylysine-coated slides. Embryos attached to slides were fixed and stained as described above. Stacks of optical sections were acquired on a Solamere spinning disk confocal system as described above. Stacks were collapsed into 2D images using Micro-Manager. Single-channel and merge images were generated in Photoshop.

Analysis of *K27me3 M+P-* and *M+P+* L1 larvae. *fem-2(ts)* females were mated to *N2* (P+) or *mes-3 M-Z-; him-8* (P-) males for ~48 hours. Gravid females were transferred to 100uL S Basal (5.85g NaCl, 1g K_2HPO_4, 6g KH_2PO_4, 1ml cholesterol (5mg/mL in ethanol), H_2O to 1 liter) on gelatin chrom alum-coated slides and incubated overnight in a humid chamber at 20˚C. Starved *K27me3 M+P-* and *K27me3 M+P+* L1s were transferred to polylysine-coated slides. L1s attached to slides were fixed and stained as described above. Stacks of optical sections were acquired on a Solamere spinning disk confocal system as described above. Stacks were collapsed into 2D images using Micro-Manager. Single-channel and merge images were generated in Photoshop.

Analysis of sensitized *K27me3 M+P-* and *M+P+* germlines. *mes-3/hT2*-GFP*; fem-2(ts)* females were mated to either *mes-3/hT2-GFP; unc-119* (P+) or *mes-3 M+Z-; unc-119* (P-) males to generate *K27me3 M+P+* and *K27me3 M+P-* worms, respectively. 146 L4 *K27me3 M+P-* larvae were cut at the pharynx to extrude their anterior germline for assessment of UNC-119::GFP expression in germ cells. Sterile adults on day 1 of adulthood (n=267) were moved to new plates for live analysis to determine whether germ cells expressed UNC-119::GFP. Sterile day 1 adults that lacked germline expression of UNC-119::GFP were allowed to age at 20˚C for a repeat analysis of UNC-119::GFP expression in germ cells on day 2 of adulthood. L4s, day 1 sterile adults, and day 2 sterile adults were analyzed for germ cell expression of UNC-119::GFP on a Leica MZ16F microscope equipped with a GFP

filter. Germlines were imaged for GFP expression both by live imaging of whole worms and by fixing dissected germlines in egg buffer with 1mM levamisole and 0.5% tween on polylysine-coated slides and immunostaining for GFP, UNC-64, PGL-1, and/or HTP-3. Stacks of optical sections were acquired on a Solamere spinning disk confocal system controlled by Micro-Manager software as described above. Stacks were collapsed into 2D images using Micro-Manager and stitched together into composite images using Photoshop. Single-channel and merge images were generated in Photoshop. In some cases (Figure 3c), germlines were straightened using ImageJ.

Sterility analysis of *K27me3 M+P-* and *M+P+* offspring. *fem-2(ts)* females were mated to *mes-3/hT2-GFP; fem-2* (P+) or *mes-3 M+Z-; fem-2* (P-) males at 15˚C. *mes-3/+; fem-2 K27me3 M+P+* and *M+P-* L4s (n=30 replicate 1; n=40 replicate 2) were cloned to individual plates and allowed to lay eggs for ~24hr, then transferred to fresh plates each day until no more embryos were laid. Embryos from *K27me3 M+P+* and *M+P-* hermaphrodites were incubated at 15˚C until day 1 adulthood, then hermaphrodites were scored as sterile (no embryos in the uterus) or fertile (n=9,370 replicate 1; n=9,693 replicate 2).

RNA-seq of *K27me3 M+P-* and *M+P+* hybrid and non-hybrid germlines.
Feminized *fem-2* (CB4856 (hybrid) or N2 (non-hybrid) background) hermaphrodites were mated to either *mes-3/hT2-GFP; him-8* (P+) or *mes-3 M-Z-; him-8* males (P-) to

generate genetically identical *mes-3*/+; fem-2/+; *him-8*/+ *K27me3 M+P+* and *M+P-*
offspring, respectively. In addition, feminized *fem-2* (CB4856 background) were
mated to wild-type N2 males to generate "wild-type" hybrid offspring. *mes-3/+; fem-2/+; him-8/+* (*K27me3 M+P+ and M+P-*) and *fem-2/+* ("wild-type" hybrid) L4
offspring were picked and allowed to age overnight to generate *K27me3 M+P+,*
K27me3 M+P-, and "wild-type" hybrid day 1 adults. ~50 germlines were dissected
from day 1 adults in egg buffer with 1mM levamisole, 0.5% tween and transferred
into 300uL of ice-cold TRIzol reagent (Life Technologies), then stored at -80˚C.
Samples were subjected to 3 freeze-thaw cycles. 1uL linear polyacrylamide was
added to each sample, vortexed, then transferred to Phase Lock Gel-Heavy 2mL tubes
to extract the aqueous phase. 0.75 volume isopropanol was added to each sample and
incubated overnight at -20˚C. RNA was isolated via ethanol precipitation and
rehydrated in 14.5uL RNase-free H2O. 2.5uL were used to assess RNA quality
(Agilent RNA Nano Bioanalyzer Chip) and concentration (Qubit Quant-IT RNA
Assay Kit). The remaining 12uL were depleted of rRNA (NEBNext rRNA Depletion
Kit) and used for library preparation (NEBNext Ultra RNA Library Prep Kit).
Libraries were sequenced on an Illumina HiSeq4000 to acquire paired-end 100bp
reads. Four biological replicates each were generated for *K27me3 M+P-, K27me3*
M+P+, and "wild-type" hybrid samples.

Processing and analyzing RNA-seq data. For differential expression analysis, raw
sequences were mapped to transcriptome version WS220 using TopHat2[25]. Only

reads with one unique mapping were allowed. Otherwise default options were used. Reads mapping to ribosomal RNAs were removed from further analysis. HTSeq[26] was used to obtain read counts per transcript. DESeq2[27] was used to determine differentially expressed genes from HTSeq counts. A Benjamini–Hochberg multiple hypothesis corrected P-value cutoff of 0.1 was used as a significance cutoff to identify significantly differentially expressed genes between *K27me3 M+P-* and *M+P+* germlines.

To identify RNA-sequencing (RNA-seq) reads coming from the Hawaiian or Bristol parent in the hybrid offspring germlines, the annotated Single Nucleotide Polymorphisms (SNPs) of the Hawaiian strain CB4856 genome compared to the Bristol strain N2 WS220 genome were downloaded from WormBase. 126,615 unique SNPs were downloaded, of which 25,980 SNPs mapped to an exon in a total of 9,146 coding genes, according to WormBase. Henceforth, we refer to these SNPs as exonic SNPs.

RNA-seq reads from the Bristol strain, Hawaiian strain, and hybrids were mapped to the Bristol (N2) genome version WS220 using TopHat2, once allowing 1 mismatch in a read and a second time not allowing a mismatch. Reads overlapping exonic SNP positions were identified, and HTSeq was used to count these SNP-containing reads for each transcript. Counts per transcript were normalized by the number of total transcript reads, including non-SNP-overlapping reads, for each of the libraries

sequenced. An average count per transcript was obtained by averaging the replicates

for each of the Bristol, Hawaiian, and hybrid *K27me3 M+P-* and *M+P+* replicate

experiments. A further quality control step was imposed. For the Bristol and

Hawaiian RNA-seq data, we calculated a distance score between the mismatch 0 and

mismatch 1 read counts for each transcript. For each transcript i in the Hawaiian

RNA-seq data , a score hz_i was calculated: $hz_i = (hmm1_i - hmm0_i)/(hmm1_i + hmm0_i)$

where $hmm1_i$ and $hmm0_i$ denote the normalized read counts for transcript i from

mapping with a mismatch 1 and with a mismatch 0 to the Bristol genome,

respectively. In the same way, we defined bz_i, $bmm1_i$, and $bmm0_i$ for transcripts from

mapping our Bristol RNA-seq data. For Hawaiian transcript reads from exonic SNPs,

we expect hz_i to be close to 1; all reads should have mapped when allowing 1

mismatch, and none should have mapped when allowing no mismatches. For Bristol

transcript reads from exonic SNPs, we expect bz_i to be close to 0, as the number of

reads from both mappings should have been very similar. We removed transcripts

that did not follow these requirements. If a transcript i had a $hz_i < 0.75$ for the

Hawaiian RNA-seq data, or an absolute value $|bz_i| > 0.25$ for the Bristol RNA-seq

data, that transcript was removed from further analysis. These requirements removed

1,080 transcripts with exonic SNPs. For the remaining 8066 transcripts we obtained

from the hybrid *K27me3 M+P-* and *M+P+* RNA-seq data, we determined the

transcript reads from the sperm allele (Bristol) as reads that mapped when allowing 0

mismatches. We determined the transcript reads from the oocyte allele (Hawaii) as

reads that mapped when allowing 1 mismatch but that did not map when allowing 0

mismatches. We thus calculated transcript reads from the oocyte allele by subtracting

reads that mapped when allowing 0 mismatches (sperm allele reads) from reads that

mapped when allowing 1 mismatch (reads from both gamete alleles).

Many genes with exonic SNPs were not expressed in any of our germline RNA-seq

data. We filtered out these non-expressed transcripts from further analysis by

requiring the normalized read counts from the sperm allele or oocyte allele in at least

one condition, *K27me3 M+P-* or *M+P+*, to be at least 0.5. Because our goal was to

use SNP-containing transcript reads as a proxy for total transcript reads, we filtered

out transcripts that had a greater than 2.5-fold difference between the $\log_2$ of SNP-

containing reads versus the $\log_2$ of total reads for a given transcript in hybrid

germlines (Hawaii mother and Bristol father). This requirement removed 190 genes

and left 2,606 transcripts for further analysis. To avoid dividing by zero when

calculating log fold changes and to reduce the effect of noise on low read counts, we

added a pseudo-count of 0.5 to all remaining transcripts from oocyte alleles and

sperm alleles in both conditions.

GO analysis of significantly misregulated genes. Significantly upregulated genes

(361 genes) and downregulated genes (342 genes) in the "wild-type" hybrid vs

K27me3 M+P- comparison (FDR<0.05) were used to perform GO analysis

(Supplementary Figure 2b, c). GO analysis was performed using the tissue

enrichment tool provided by WormBase (version WS266).

Statistics. Sample sizes and statistical tests used are described in the figure legends. Paired student's t-tests and Wilcoxon rank sum tests were performed using R. Four biological replicates of sequencing data were used. No replicates were omitted. Differential expression analysis was performed using DESeq2, which uses negative binomial generalized linear models. Adjusted P-values were calculated using the Benjamini-Hochberg method performed by DESeq2.

Data availability. Raw sequence files and processed data sets generated in this study are available at the Gene Expression Omnibus (GEO) repository, accession GSE123415. Processed data sets available at GEO are as follows: RNA-seq read counts for *K27me3 M+P-, K27me3 M+P+* and "wild-type" hybrid containing 1) all read counts mapping to a transcript, 2) read counts overlapping SNPs mapped to the Bristol genome allowing 1 mismatch (Hawaii & Bristol reads), 3) read counts overlapping SNPs mapped to the Bristol genome allowing 0 mismatches (Bristol-specific reads), 4) differential expression analysis comparing transcript reads from *K27me3 M+P-* vs *M+P+*, and 5) differential expression analysis comparing transcript reads from *K27me3 M+P-* vs "wild-type" hybrid. Supplementary Data 1 consists of sperm-specific and oocyte-specific reads and the $\log_2$ fold changes between *K27me3 M+P-* and *M+P+* for significantly misregulated SNP-containing genes.

Chapter 3: Extended analysis of worms that inherit the sperm-inherited genome lacking H3K27me3

I: Introduction

In Chapter 1, I presented our lab's findings, based on ChIP-seq (Chromatin Immunoprecipitation followed by Sequencing) in early embryos, that the distributions of H3K36me3 and H3K27me3 across the genome of early embryos reflect germline transcription patterns, including genes that are specifically expressed and repressed in the germline and not in early embryos (Gaydos *et al.* 2012; Rechtsteiner *et al.* 2010). These findings led us to a model in which transcription patterns in the parental germline establish the distributions of these marks; we posited that the transcription-coupled activity of MET-1 deposits H3K36me3 over transcribed genes and PRC2 generates H3K27me3 at repressed regions across the genome. Embryos inherit these marks on the sperm and oocyte genomes, after which MES-4 maintains H3K36me3 and PRC2 maintains H3K27me3 in a transcription-independent manner during the cell divisions that generate the primordial germ cells (PGCs). We hypothesize that a chromatin landscape congruent with germline transcription patterns encourages appropriate transcription patterns and discourages inappropriate transcription patterns in the nascent and adult germline.

In Chapter 2, I presented evidence that the absence of H3K27me3 from the sperm-inherited genome results in the upregulation of many genes in the germline of their *K27me3 M+P-* offspring. Using single nucleotide polymorphisms (SNPs) between sperm- and oocyte-inherited genomes, I demonstrated that this upregulation

47

emanates primarily from sperm alleles. These findings support the model that gamete-inherited patterns of H3K27me3 and H3K36me3 can directly shape offspring transcription. This experimental system in which worms inherit a sperm genome lacking H3K27me3 is a powerful paradigm to test other aspects of our model.

Until recently, the ability to survey histone marks in isolated tissues was limited. This is because technologies such as ChIP-seq require large amounts of starting material. This made it unfeasible to profile the chromatin landscape of hand-dissected tissues, such as the germline. Cut and Run followed by sequencing (CRun-seq) is a new strategy for profiling chromatin associated marks and factors that, unlike ChIP-seq, requires very little starting material. This enabled us, for the first time, to profile histone marks in hand-dissected *C. elegans* germlines.

I took advantage of this new technology to probe the *K27me3 M+P*-paradigm. My aim was to determine the H3K27me3 and H3K36me3 landscape in wild-type male germlines and to assess how the H3K36me3 distribution shifts in fathers that lack PRC2 and thus H3K27me3. I used this analysis to investigate whether changes in paternal chromatin and gene expression shape histone marking and/or gene expression in somatic cells and the germline of offspring (Figure 1).
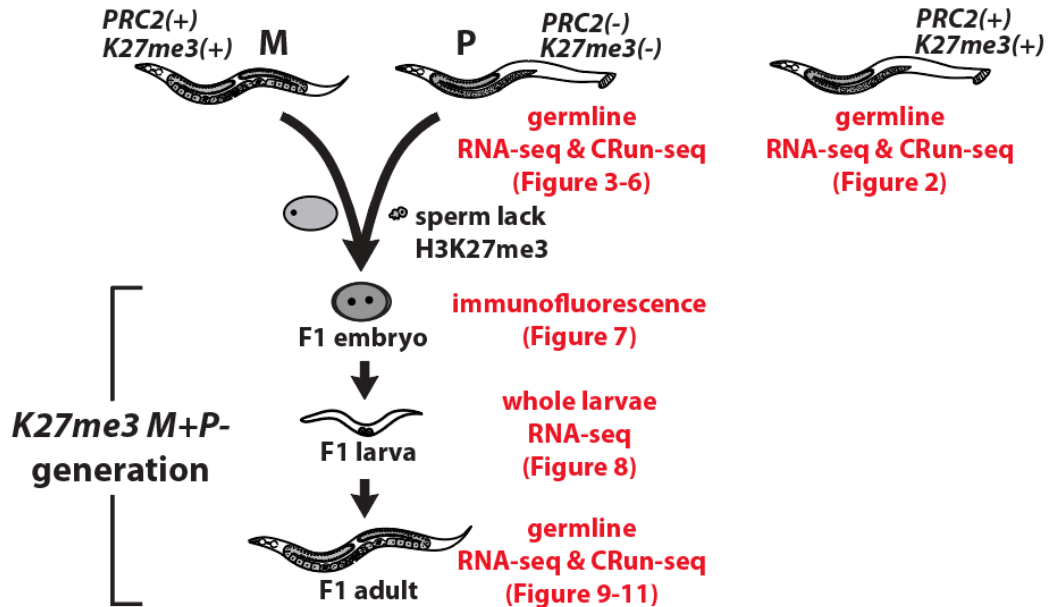
Figure 1. Diagram showing the cross between wild-type (WT) *PRC2(+) K27me3(+)* females and *mes-3/mes-3 PRC2(-) K27me3(-)* males that generates *K27me3 M+P-* F1 offspring. Maternal and paternal pronuclei are shown in the F1 1-cell embryo. The 2 primordial germ cells (PGCs) are shown in the newly hatched F1 larva. Figures that analyzed each generation/stage are indicated in parentheses. M=maternal; P=paternal. Hybrid F1s were generated by mating Hawaii mothers and Bristol fathers.

II: Results

H3K27me3 and H3K36me3 patterns in wild-type and *PRC2(-)* male germlines

To analyze the distributions of H3K36me3 and H3K27me3 in WT males and to investigate how the distribution of H3K36me3 changes in *PRC2(-)* males, Thea Egelhofer and I performed CRun-seq of *him-8* male germlines, which serve as wild-type (WT) controls, and *PRC2(-)* male germlines. In the following sections, I discuss our findings based on single replicates of each mark and compare our CRun-seq data to previously published RNA sequencing data from WT and *PRC2(-)* male germlines

(Tabuchi *et al.* 2018). We will perform additional replicates and re-analyze the data before publication.

As an overview, I found that in WT male germlines, H3K36me3 and H3K27me3 reflect transcriptional patterns on autosomes but not on the X chromosome. In *PRC2(-)* male germlines, H3K36me3 spreads to all regions of the genome, becoming preferentially enriched on genes that are occupied by H3K27me3 in WT. In most cases, the changes in H3K36me3 levels correlate with changes in transcription. However, the global increase in H3K36me3 across the genome and the finding that most genes are downregulated in *PRC2(-)* male germlines suggest that H3K36me3 is additionally generated via a transcription-independent mechanism. This is especially illustrated in the global acquisition of H3K36me3 across the X chromosome, despite the finding that X-linked genes are not upregulated and as a class are globally downregulated in *PRC2(-)* male germlines. These findings need to be validated by additional CRun-seq replicates, as this global increase in sequence reads from H3K36me3 CRun-seq could be due to alternative factors such as a global increase in chromatin accessibility that may accompany a lack of H3K27me3.

- H3K27me3 and H3K36me3 patterns reflect transcriptional patterns in male germlines.

Our model predicts that, in WT male germlines, transcriptional patterns are reflected in the distributions of H3K36me3 and H3K27me3. More specifically our model predicts that transcript levels should correlate with H3K36me3 and anti-

correlate with H3K27me3. To assess whether this is the case, I binned coding genes into quartiles based on transcript levels in WT male germlines. For each quartile, I determined the levels of H3K36me3 and H3K27me3 across gene bodies. The results of this analysis indicate that transcriptional patterns across the autosomes are largely reflected in the distributions of HK36me3 and H3K27me3 (Figure 2A,B). For instance, genes with the lowest transcript levels (lowest quartile of coding genes) have the lowest levels of H3K36me3 (red box in Figure 2A) and the highest levels of H327me3 (red box in Figure 2B). The reverse is true for genes that fall into the highest quartile of transcript levels (green boxes in Figure 2A,B). In contrast to autosomal genes, X-linked genes, regardless of transcript level, have low levels of H3K36me3 (Figure 2C) and high levels of H3K27me3 (Figure 2D). Therefore, in wild-type male germlines, transcript levels correlate with H3K36me3 and anti-correlate with H3K27me3 across the autosomes, while X-linked genes have high levels of H3K27me3 and low levels of H3K36me3 regardless of their transcriptional status.

autosomal genes (17,452 genes)

A transcript vs H3K36me3 levels in WT male germlines

B transcript vs H3K27me3 levels in WT male germlines
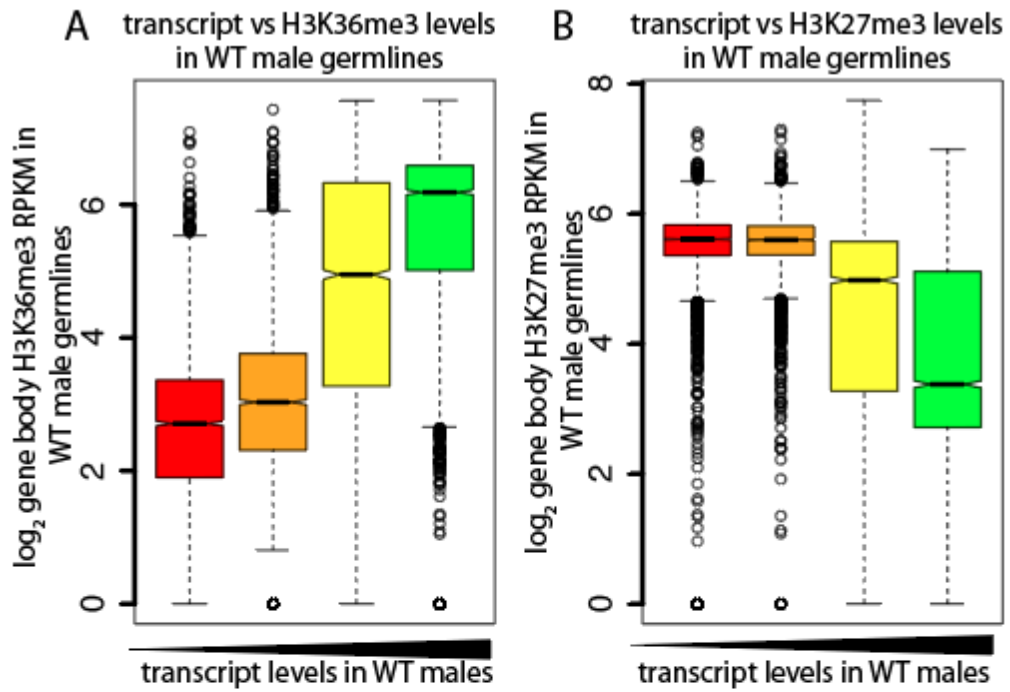
X-linked genes (2,808 genes)

C transcript vs H3K36me3 levels in WT male germlines

D transcript vs H3K27me3 levels in WT male germlines
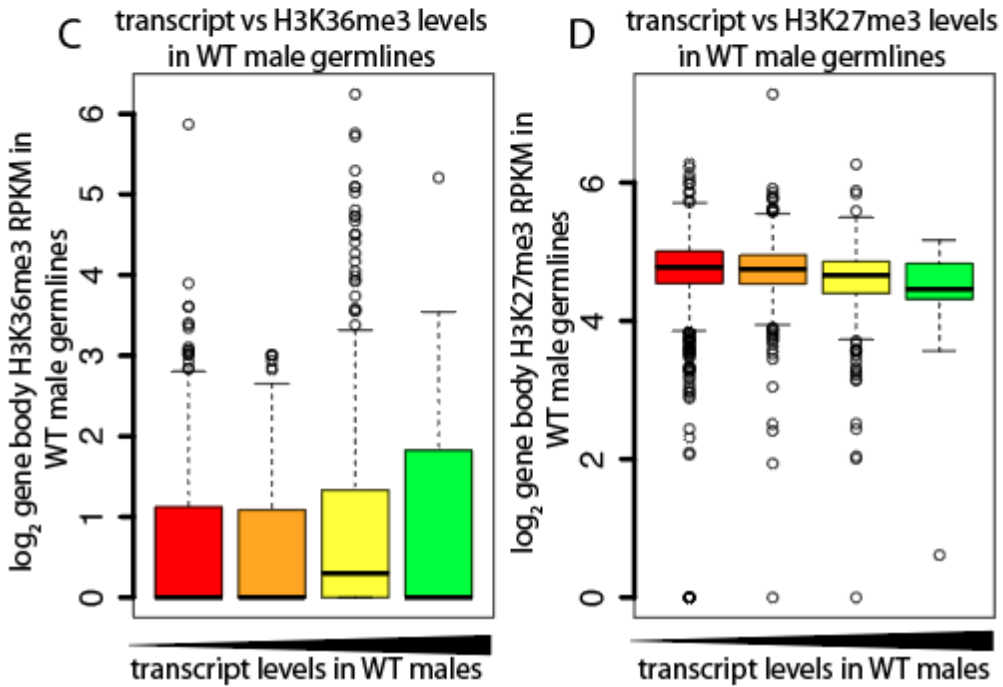
Figure 2. In WT male germlines, the distributions of H3K36me3 and H3K27me3 reflect transcription patterns on autosomes but not the X chromosome. Boxplots of log2(RPKM) across gene bodies for H3K27me3 and H3K36me3 from WT male germlines for autosomal (**A,B**) and X-linked (**C,D**) genes. Boxes represent genes that fall into increasing quartiles of transcript levels in WT male germlines. Red = lowest quartile (4,366 autosomal & 1,413 X-linked genes). Orange = 2nd lowest quartile (3,806 autosomal & 925 X-linked genes). Yellow = 2nd highest quartile (4,487 autosomal & 429 X-linked genes). Green = highest quartile (4,793 autosomal & 19 X-linked genes). The triangle below each plot indicates the direction of increasing transcript levels. Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend to the most extreme data point that is no more than 1.5 times the interquartile range from the box; individual data points outside of these boundaries are shown as open circles. The waist indicates the 95% confidence interval for the median.

- H3K36me3 is globally increased across the genome of PRC2(-) male germlines.

Another aspect of our model is that H3K36me3 and H3K27me3 antagonistically pattern the genome. In Chapter 2, I presented evidence that MES-4, likely through the mark it generates (H3K36me3), antagonizes PRC2 activity from depositing H3K27me3 onto sperm-inherited chromosomes lacking H3K27me3. This observation indicates that, at least at the level observable by immunofluorescence, the entire lengths of H3K27me3(-) chromosomes are recalcitrant to PRC2 activity and suggests that either 1) in *PRC2(-)* male germlines, H3K36me3 spreads to all regions of the genome or 2) in *K27me3 M+P-* embryos, MES-4 rapidly generates antagonistic H3K36me3 across H3K27me3(-) chromosomes. To distinguish between these possibilities, I compared levels of H3K36me3 detected by CRun-seq in *PRC2(-)* vs WT male germlines. In Figure 2, I analyzed genes divided into quartiles based on transcript levels in WT male germlines. In Figure 3, I assessed each quartile for either changes in H3K36me3 or changes in transcript levels that occur in *PRC2(-)* male

germlines compared to WT. This analysis indicates that in *PRC2(-)* male germlines

both transcript levels and H3K36me3 levels preferentially increase on genes that, in

WT males, have the lowest levels of transcription and therefore the highest levels of

H3K27me3 (red boxes in Figure 3B,D). On autosomes, this anti-correlation between

expression level in WT males and increases in H3K36me3 and transcription in

PRC2(-) males extends in a dose-dependent manner across all quartiles of WT

transcript level (Figure 3A,B). In contrast, X-linked genes had increased levels of

H3K36me3 across all quartiles of WT transcript level (Figure 3C). Despite this global

increase in H3K36me3, the vast majority of X-linked genes have reduced transcript

levels in *PRC2(-)* male germlines (Figure 3D). These findings suggest that the

autosomal genes, but not the X-linked genes, that gain the most H3K36me3 will also

have the greatest increase in transcript level. To directly compare changes in

transcript level with changes in H3K36me3, I binned all coding genes based on the

degree to which transcript levels change in *PRC2(-)* germlines and assessed

autosomal genes within each bin for changes in gene body H3K36me3 levels. This

analysis revealed a U-shaped relationship between changes in transcript level and

changes in H3K36me3 (Figure 3E). I observe the expected dose-dependent

relationship between increased transcription and increased gene body H3K36me3

levels (right half of boxes in Figure 3E). However, there is also a slight but still

unexpected dose-dependent relationship between *decreased* transcription and

increased gene body H3K36me3 levels (left half of boxes in Figure 3E). This

indicates that there is a general correlation between the degree to which a gene is

misregulated, whether up or down, and the degree to which H3K36me3 accumulates over the gene body. To investigate whether changes in H3K36me3 predict changes in transcription, I performed the reverse analysis. I binned all coding genes this time based on the degree to which H3K36me3 levels change in *PRC2(-)* germlines compared to WT and assessed autosomal genes within each bin for changes in transcript levels. In this case, changes in H3K36me3 across all bins predicted changes in transcription (Figure 3F). Together these findings suggest that changes in H3K36me3 levels over autosomal gene bodies are highly correlated with changes in transcription. However, this correlation does not hold for X-linked genes or for autosomal genes that become downregulated.

autosomal genes (3,183 genes)

A transcript levels in WT males vs changes in H3K36me3

B transcript levels in WT males vs changes in transcripts

X-linked genes (159 genes)

C transcript levels in WT males vs changes in H3K36me3

D transcript levels in WT males vs changes in transcripts

autosomal genes (3,183 genes)

E changes in transcript levels vs changes in H3K36me3
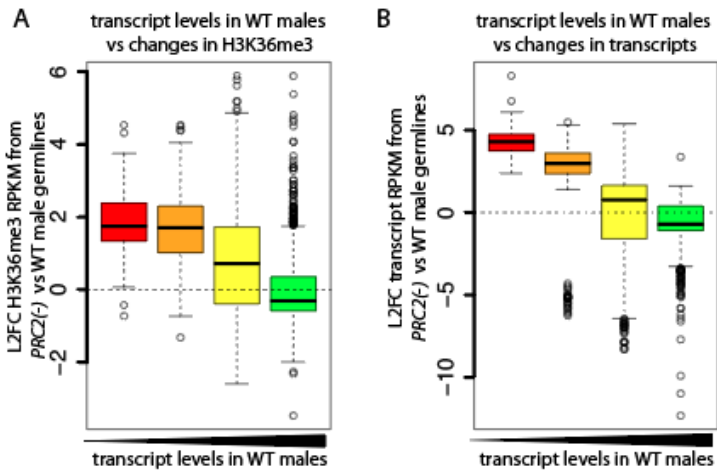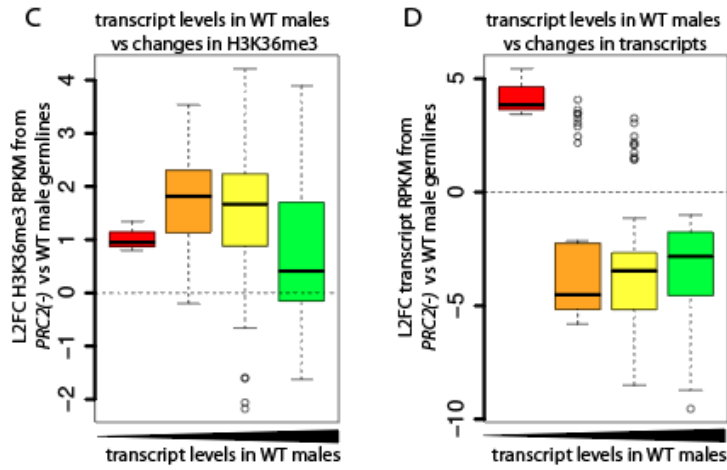
F changes in H3K36me3 vs changes in transcripts
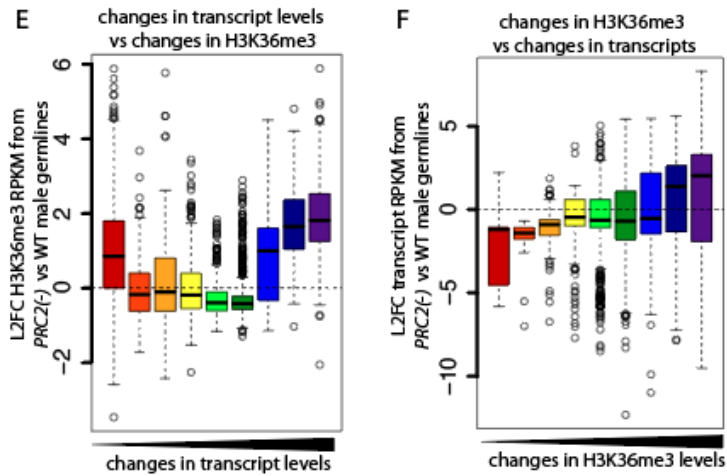
Figure 3. H3K36me3 is globally increased across the genome of *PRC2(-)* male germlines. Boxplots of $\log_2$ fold change (L2FC) in gene body H3K36me3 RPKM (**A,C,E**) or L2FC transcript RPKM (**B,D,F**) in *PRC2(-)* male germlines vs WT male germlines. Only genes with a p-value <0.1 based on differential analysis (Deseq2) of transcript reads from *PRC2(-)* vs WT germlines were analyzed. **A-D** Boxes represent genes that fall into increasing quartiles of transcript level in WT male germlines. Red = lowest quartile (90 autosomal & 3 X-linked genes). Orange = 2nd lowest quartile (255 autosomal & 40 X-linked genes). Yellow = 2nd highest quartile (997 autosomal & 105 X-linked genes). Green = highest quartile (1,841 autosomal & 11 X-linked genes). **E,F** L2FC H3K36me3 RPKM (**E**) or L2FC transcript RPKM (**F**). Boxes show progressive increases in L2FC gene body H3K36me3 RPKM (**F**) or transcript RPKM (**E**) in *PRC2(-)* vs WT male germlines and range from negative L2FC (less than -2 in red) to positive L2FC (greater than 2 in purple). Whether the plots display autosomal or X-linked genes is indicated, with the number of genes in each class indicated in parentheses. The triangle below each plot indicates the direction of increasing levels (**A-D**) and L2FC (**E,F**). Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend to the most extreme data point that is no more than 1.5 times the interquartile range from the box; individual data points outside of these boundaries are shown as open circles.

The striking difference between changes in autosomal versus X-linked genes in *PRC2(-)* male germlines led me to analyze whole classes of misregulated autosomal versus X-linked genes (Figure 4). To analyze a large set of genes within each class, I analyzed autosomal and X-linked genes that display a low-stringency change in transcript level in *PRC2(-)* male germlines compared to WT (p-value < 0.1). First, I compared changes in transcript levels with changes in H3K36me3 levels for each gene (Figure 4A). Strikingly, the vast majority of X-linked genes gain H3K36me3 and yet are downregulated (purple in Figure 4A-D). In contrast, autosomal genes split between gain and loss in H3K36me3 and transcriptional levels (Figure 4A). On average, transcriptional changes across autosomal genes falls just below zero (Figure 4C) while H3K36me3 changes average to zero (Figure 4D). Although X-linked genes gain H3K36me3, the absolute levels of H3K36me3 on X-

linked genes are still considerably lower than the absolute levels on autosomal genes

(dark gold vs dark purple in Figure 4B).

A L2FC H3K36me3 vs L2FC transcript from *PRC2(-)* vs WT male germlines

694 genes
1 genes
780 genes
21 genes
895 genes
12 genes
812 genes
120 genes

L2FC transcript RPKM

L2FC gene body H3K36me3 RPKM

B H3K36me3 levels in WT & *PRC2(-)* male germlines

log$_2$ gene body H3K36me3 RPKM in WT male germlines

WT auto
PRC2(-) auto
WT X chr
PRC2(-) X chr

PRC2(-) vs WT male germlines

C changes in transcript levels

autosomal genes
X-linked genes

L2FC transcript RPKM

D changes in H3K36me3 levels

autosomal genes
X-linked genes

L2FC gene body H3K36me3 RPKM

E L2FC transcript reads vs significance

auto (1,373)
X-linked (75)
auto (1,047)
X-linked (17)

-log$_{10}$ p-value

L2FC transcript reads

F changes in H3K36me3 levels on autosomal & X-linked genes

L2FC gene body H3K36me3 RPKM

(1,373)
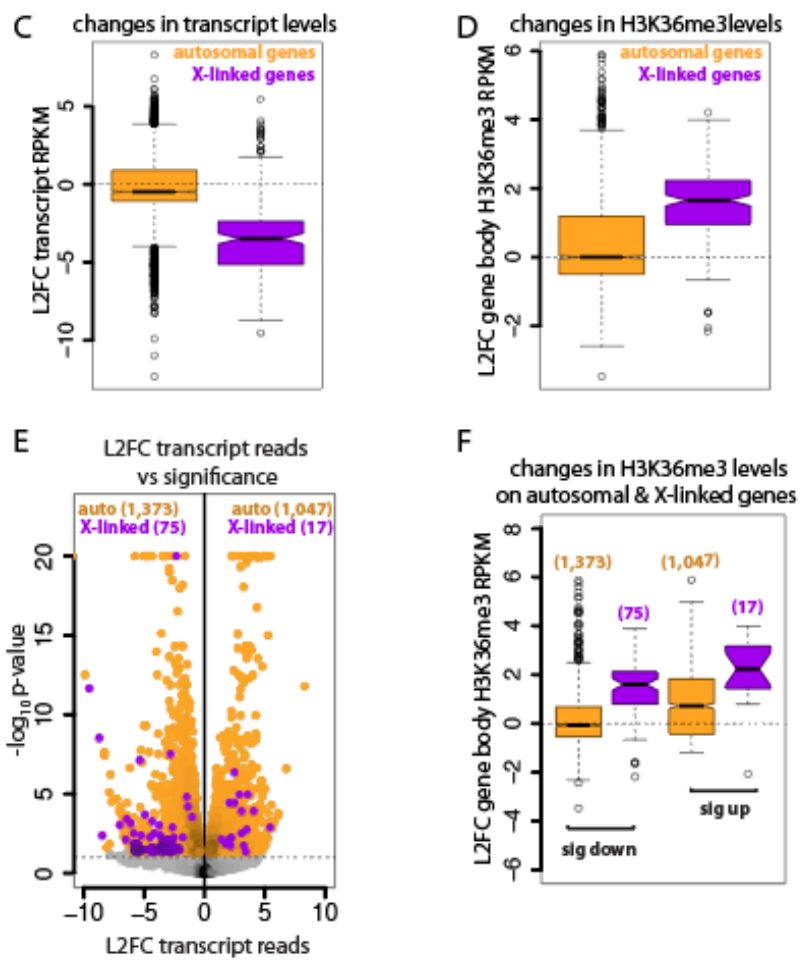(75)
(1,047)
(17)

sig down
sig up

59

Figure 4. Autosomal and X-linked genes differentially accumulate H3K36me3 in *PRC2(-)* male germlines. **A-D** Scatterplot and boxplots analyzing autosomal and X-linked genes with a p-value <0.1 based on differential analysis (Deseq2) of transcript reads from *PRC2(-)* vs WT germlines. **A** Scatterplot of log2 fold change (L2FC) of gene body H3K36me3 RPKM vs L2FC transcript RPKM in *PRC2(-)* vs WT male germlines. **B** Boxplots of $\log_2$ H3K36me3 RPKM in WT and *PRC2(-)* male germlines. **C,D** Boxplots of autosomal and X-linked genes showing L2FC transcript RPKM (**C**) or L2FC gene body H3K36me3 RPKM (**D**). **E,F** Volcano plot and boxplots analyzing significantly misregulated genes (p-value <0.05). **E** Volcano plot comparing L2FC transcript reads vs $-\log_{10}$ p-value from *PRC2(-)* vs WT male germlines. The number of upregulated or downregulated autosomal and X-linked genes are indicated in the upper right and left, respectively. **F** Boxplots of the L2FC of gene body H3K36me3 RPKM for significantly upregulated and downregulated autosomal and X-linked genes. The number of genes in each category is indicated in parentheses. Autosomal and X-linked are in orange and purple, respectively. In panel **B**, WT and *PRC2(-)* boxes for autosomal and X-linked genes are shown in light gold and purple or dark gold and purple, respectively. Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend to the most extreme data point that is no more than 1.5 times the interquartile range from the box; individual data points outside of these boundaries are shown as open circles. The waist indicates the 95% confidence interval for the median.

These broad trends in autosomal and X-linked genes also applies to a more stringent set of misregulated genes (p-value< 0.05) (Figure 4E,F). Significantly upregulated genes, whether autosomal or X-linked, gain H3K36me3 (Figure 4F). This is likely mediated through transcription-coupled deposition of H3K36me3 by MET-1. In contrast, significantly downregulated autosomal genes do not accumulate H3K36me3, while downregulated X-linked genes dramatically acquire H3K36me3 (Figure 4F). These findings indicate that, for X-linked genes, accumulation of H3K36me3 is decoupled from transcription and suggest that H3K36me3 generated on the X chromosome is mediated by MES-4, which is the only worm histone methyl transferase (HMT) known to generate H3K36me3 independently of transcription (Bender *et al.* 2006; Furuhashi *et al*. 2010; Kreher *et al.* 2018).

Global analysis of changes that occur across autosomal versus X-linked genes

indicates that X-linked genes are globally downregulated and yet gain H3K36me3.

Autosomal genes, on the other hand, do not uniformly gain nor lose H3K36me3 or

uniformly become up- or downregulated. This led me to delve more deeply into

changes that occur across autosomal genes. My aim was to develop a cohesive model

that would unite the seemingly disparate changes that occur on autosomal genes in

PRC2(-) male germlines.

To achieve this, I binned genes according to changes in transcript levels that

occur in *PRC2(-)* male germlines compared to WT (genes with a p-value < 0.1 were

binned based on log2 fold change in transcript RPKM). Then, for each bin, I assessed

WT levels of H3K27me3 (Figure 5A), transcripts (Figure 5B) and H3K36me3

(Figure 5C). For each bin, I also assessed H3K36me3 levels in *PRC2(-)* germlines

(Figure 5D). By comparing H3K36me3 levels in WT to *PRC2(-)* males, I was able to

assess what happens to the absolute levels of H3K36me3 when H3K27me3 is absent.

In brief, I found that some genes gain H3K36me3 independent of transcription (red

boxes in Figure 5). These genes have high levels of H3K27me3 and low to moderate

levels of H3K36me3 in WT germlines. The most highly transcribed genes in WT

males, have the lowest levels of H3K27me3 and the highest levels of H3K36me3 and

thus do not gain additional H3K36me3 when H3K27me3 is absent (orange, yellow

and green in Figure 5). Instead, these genes represent a smaller fraction of total

H3K36me3 CRun-seq and RNA-seq reads in *PRC2(-)* male germlines. Thus, this

class appears to have slightly reduced transcript and H3K36me3 levels. The set of genes that display the greatest gain in transcript levels have the highest levels of H3K27me3 and lowest levels of H3K36me3 in WT germlines (blue and purple in Figure 5). These genes gain the most H3K36me3 upon loss of H3K27me3. This may be the result of both transcription-dependent and transcription-independent mechanisms.
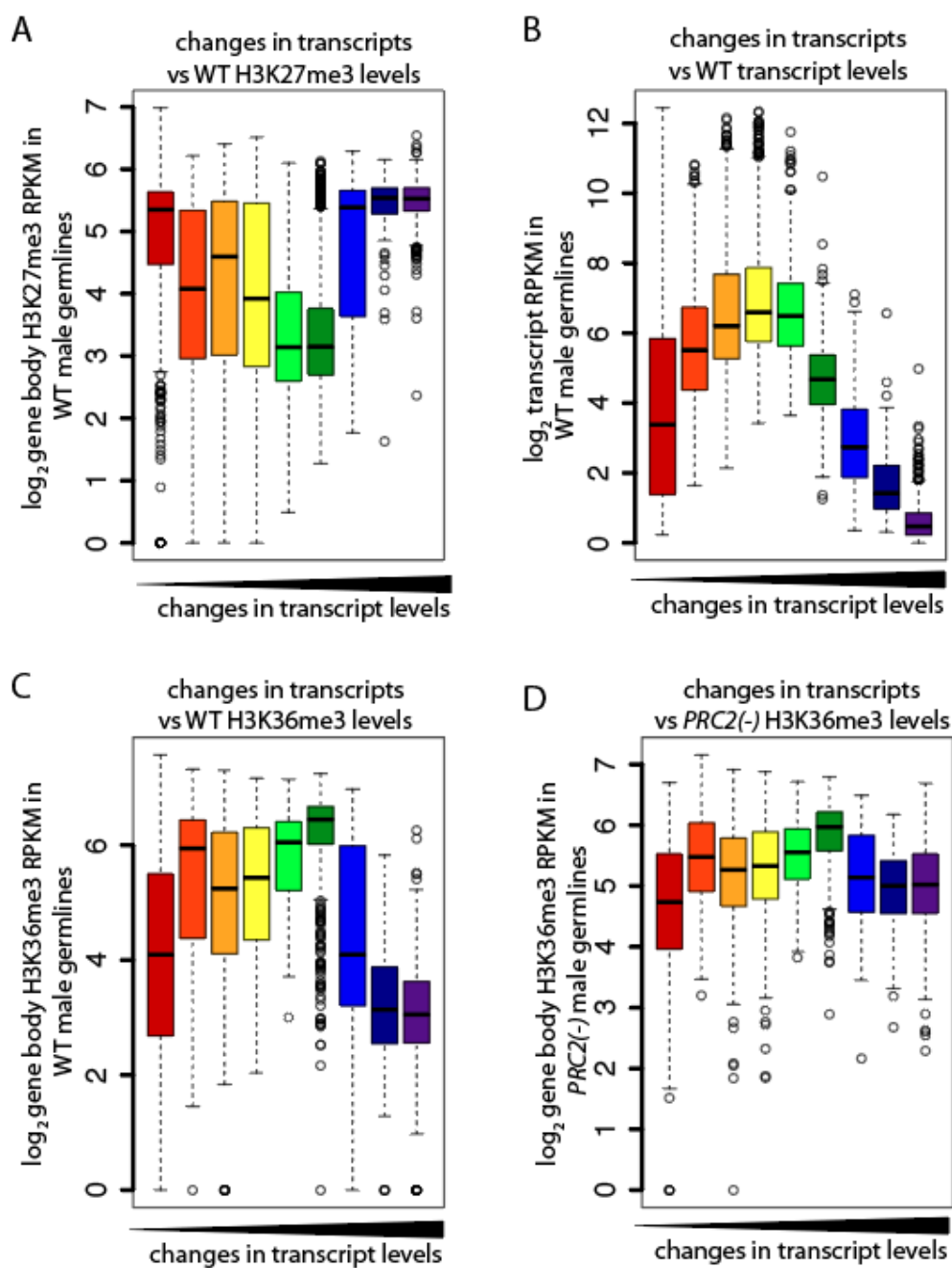
Figure 5. H3K36me3 is globally increased across the genome of *PRC2(-)* male germlines. Boxplots of log2 gene body H3K27me3 RPKM (**A**), H3K36me3 RPKM (**C,D**) or transcript RPKM (**B**) from WT (**A-C**) or *PRC2(-)* (**D**) male germlines. Boxes show progressive increase in log$_2$ fold change (L2FC) transcript RPKM from *PRC2(-)* vs WT male germlines and range from negative L2FC (less than -2 in red) to

63

positive L2FC (greater than 2 in purple). Only genes with a p-value <0.1 based on differential analysis (Deseq2) of transcript reads from *PRC2(-)* vs WT germline were analyzed. The triangle below plot indicates the direction of increasing L2FC. Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend to the most extreme data point that is no more than 1.5 times the interquartile range from the box; individual data points outside of these boundaries are shown as open circles.

Altogether these findings, if replicated, indicate that in the absence of H3K27me3, H3K36me3 spreads into genes that in WT male germlines are transcriptionally repressed and enriched for H3K27me3. This gain in H3K36me3 levels may or may not be associated with an increase in transcription. However, of the genes that gain H3K36me3, those associated with an increase in transcription have a greater increase in H3K36me3 levels, likely through the added deposition of transcription-coupled H3K36me3. Altogether, the changes in H3K36me3 levels that occur in *PRC2(-)* male germlines result in relatively high levels of H3K36me3 across coding genes (Figure 5D).

The finding that in the absence of H3K27me3, H3K36me3 can spread into H3K27me3-enriched genes independent of transcription suggests that this may not be limited to gene bodies. Indeed, visual analysis of normalized reads in the genome browser indicates that increased H3K36me3 is not restricted to coding regions of the genome (Figure 6). The entire length of the X chromosome appears to acquire H3K36me3 in *PRC2(-)* male germlines (Figure 6A). On the autosomes, H3K36me3 appears to accumulate across entire domains typically enriched for H3K27me3 in WT germlines (Figure 6B). If additional replicates and deeper analysis of H3K36me3 over intergenic regions of the genome validate these findings, it suggests that *PRC2(-)*

fathers contribute sperm chromosomes coated with high levels of H3K36me3. If

maintained by MES-4, these high levels of inherited H3K36me3 may serve to

antagonize PRC2 activity, which would allow the H3K27me3(-) state of sperm

chromosomes to be propagated to the PGCs in *K27me3 M+P-* embryos.
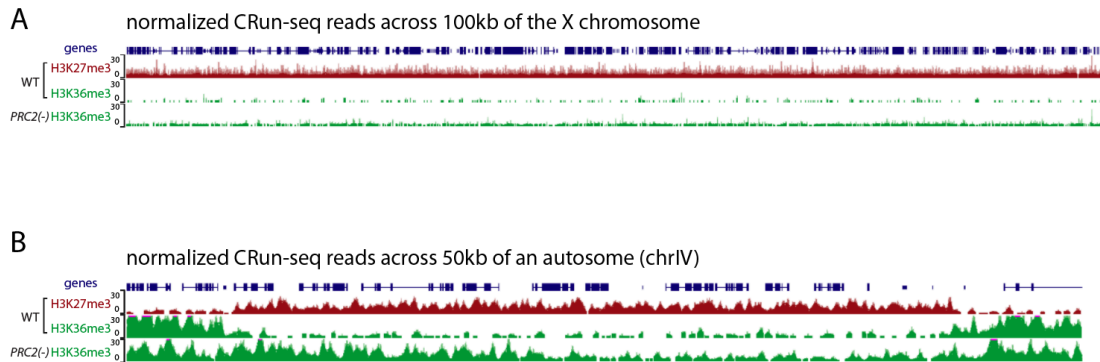
Figure 6. Genome browser view of H3K27me3 and H3K36me3 in WT and *PRC2(-)* male germlines. Genome browser view of a 100kb region of the X chromosome (**A**) and 50kb region of chr IV (**B**) showing normalized CRun-seq reads for H3K27me3 (red) and H3K36me3 (top green) from WT male germlines and H3K36me3 (bottom green track) from *PRC2(-) male* germlines. The Y-axis indicates normalized read counts. Genes are shown in dark blue.

H3K27me3 accumulates on the sperm-inherited X chromosome in *K27me3 M+P-*

embryos

As discussed in Chapter 2, when sperm chromosomes are inherited lacking

H3K27me3, the H3K27me3(-) state of sperm-inherited chromosomes is maintained

through the embryonic cell divisions that generate the PGCs. I found that this

maintenance depends on the activity of MES-4; a maternal load of MES-4 prevents

PRC2 in embryos from methylating H3K27me3(-) sperm chromosomes. However, I

found an exception to this rule. Half the embryos that inherited H3K27me3(-) sperm

chromosomes acquired H3K27me3 on a single chromosome in the sperm-inherited

set of chromosomes (Figure 7A,B). As these embryos result from mating, half the embryos inherited an X chromosome from the sperm and developed into XX hermaphrodites. The other half only inherited autosomes from the sperm and developed into XO males. I interpret these findings to indicate that in XX embryos, the sperm-inherited X chromosome acquires H3K27me3. This sperm-inherited chromosome begins acquiring detectable H3K27me3 beginning at the 2-cell stage, and H3K27me3 staining increases in intensity with each cell division. By the 12-cell stage, the level of H3K27me3 on the sperm-inherited X chromosome is comparable in intensity to the level of H3K27me3 on oocyte-inherited chromosomes (Figure 7B). This observation is unexpected given the finding that, in *PRC2(-)* male germlines, H3K36me3 accumulates on the sperm-inherited X chromosome. Our model would predict that the presence of H3K36me3 on the X, if delivered to the embryo, would prevent PRC2 from *de novo* methylating the X chromosome. At least two possibilities may explain these seemingly contradictory observations. One possibility is that H3K36me3 is removed from the X chromosome prior to packaging into mature sperm. Another possibility is that the levels of H3K36me3 present on the sperm-inherited X chromosome are not sufficient to antagonize PRC2 activity.
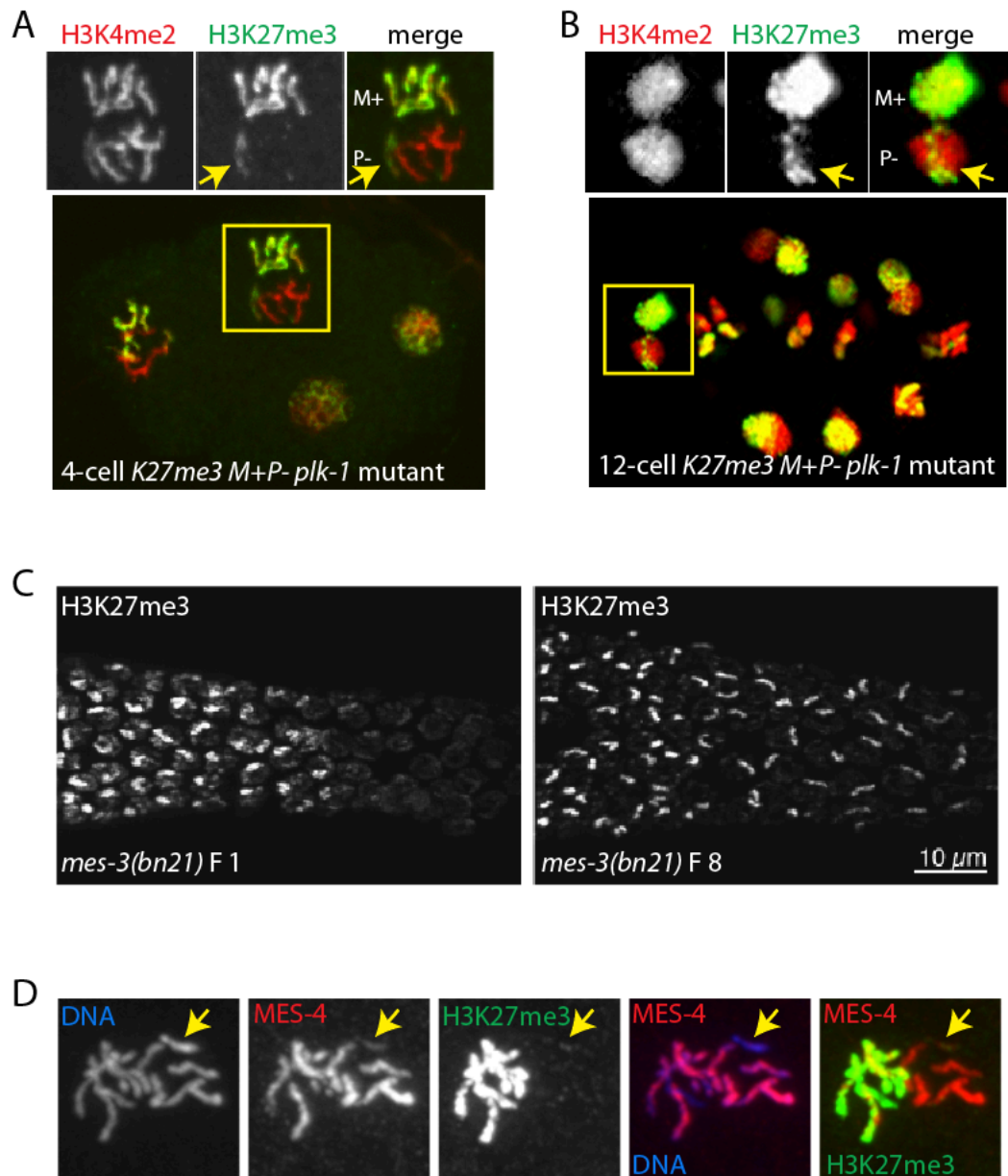
Figure 7. H3K27me3 accumulates on the X chromosomes in various genetic backgrounds. **A,B** Immunofluorescence images of a *K27me3 M+P-* 4-cell (**A**) and 12-cell (**B**) embryo in a *plk-1* mutant, which keeps sperm- and oocyte-inherited chromosomes in separate nuclei. The yellow boxes are the maternal (M+)/paternal (P-) pair of nuclei shown at higher magnification above. H3K27me3 immunostaining in green, and H3K4me2 in red. **C** Immunofluorescence images of a PRC2 hypomorpic mutant *(mes-3(bn21))* in which the X chromosomes maintain H3K27me3 while the autosomes progressively lose H3K27me3 across generations (images from Gaydos dissertation 2014). **D** immunofluorescence images of a 1-cell *K27me3 M+P-* embryo.

H3K27me3 immunostaining in green, MES-4 in red and DNA in blue. Yellow arrows point to the sperm-inherited X chromosome.

We assume that histone marks in the male germline accurately predict histone marks in mature sperm. However, there are several reasons why this may not be the case. Within a male germline there are many cells at various stages of germ cell development. At the distal-most end of the germline are mitotically cycling germline stem cells. As germ cells progress proximally, they also progress meiotically. As germ cells differentiate into mature sperm, their genome condenses which, in many species, involves repackaging of the genome with non-histone proteins called protamines. Although, *C. elegans* sperm retain histones genome-wide, the degree to which histones are replaced during spermatogenesis is not known. CRun-seq analysis of male germlines was performed on germ cells distal to this stage. Therefore, changes in marking that result from histone replacement would not be captured in my CRun-seq analysis of male germlines. For this reason, it is possible that the increased level of H3K36me3 on the X chromosome observed in *PRC2(-)* germlines is not retained in mature sperm.

If the X chromosome retains high levels of H3K36me3 in mature sperm of *PRC2(-)* males, there are several possibilities that may explain why, unlike autosomes, H3K36me3 on the X chromosome is not sufficient to antagonize PRC2 activity in *K27me3 M+P-* embryos. One possibility is that PRC2 is actively recruited to the X chromosome. As discussed in Chapter 1, some active marks antagonize PRC2 activity. However, experimentally introducing a strong PRC2 recruitment

element is sufficient to repress reporter gene expression and reset the locus to a repressed chromatin state (Hansen *et al.* 2008; Laprell *et al.* 2017; Coleman *et al.* 2017). This suggests that if the X chromosome has the intrinsic ability to recruit PRC2, then PRC2 may be able to reset the X chromosome to a repressed chromatin state regardless of its transcriptional history of expression or enrichment of active marks. In fact, the oocyte- but not the sperm-inherited X chromosome has a transcriptional history of expression. We think that it is for this reason that PRC2 is required for male fertility if they inherit their X chromosome from the oocyte but not if they inherit their X chromosome from the sperm (Gaydos *et al.* 2014). This suggests that PRC2 is required to reset the oocyte X chromosome to a repressed state. If the X chromosome is inherited without a transcriptional history of expression, PRC2 is not necessary. The possibility that the X chromosome intrinsically recruits PRC2 is further supported by the observation that in *mes-3* mutants in which PRC2 activity is compromised but not absent, H3K27me3 persists and may even accumulate to higher levels on the X chromosomes while autosomes become depleted of H3K27me3 (Figure 7C) (Delaney *et al.* 2019).

Another, not mutually exclusive, possibility is that MES-4 may not be sufficiently recruited to the X chromosome in *K27me3 M+P-* embryos. In fact, immunofluorescence images of 1-cell *K27me3 M+P-* embryos, indicates that, in contrast to sperm-inherited autosomes, very little MES-4 is detected on the sperm-inherited X chromosome (Figure 7D). MES-4 has been shown to be recruited to H3K36me3(+) but not H3K36me3(-) chromosomes in embryos. It could be that the

lower absolute levels of H3K36me3 on the sperm-inherited X chromosome in

K27me3 M+P- embryos is not sufficient to recruit MES-4 (Figure 4B).

Immunostaining *K27me3 M+P-* embryos at the 1-cell stage could shed light on this.

For example, if differences in H3K36me3 levels on autosomes versus the X

chromosome match differences in MES-4, this could suggest that lower absolute

levels may explain reduced MES-4 recruitment.

Increased H3K36me3 in germ cells of fathers predicts upregulation in the soma
of offspring

In Chapter 2, I discussed changes in germline gene expression and

development that result from inheriting the sperm genome lacking H3K27me3. I

found that *K27me3 M+P-* germlines were prone to upregulate genes primarily from

sperm alleles. That finding indicates that, relative to oocyte alleles, the absence of

H3K27me3 from sperm alleles during germline development made them particularly

susceptible to upregulation in the germline. This same logic can be extended to other

cell types, which also launch their transcriptional programs from sperm- and oocyte-

inherited genomes that differ in H3K27me3. To investigate whether somatic cells

may be similarly impacted by inheriting a sperm genome lacking H3K27me3, I

performed RNA-seq on L1 larvae that inherited the sperm genome with (*K27me3*

M+P+) or without (*K27me3 M+P-*) H3K27me3. *C. elegans* larvae hatch with

roughly 550 somatic cells and only 2 germ cells (Sulston & Horvitz 1977). Thus,

profiling RNA from whole L1 larvae is a proxy for assessing RNA content in somatic cells.

I found that inheriting the sperm genome lacking H3K27me3 resulted in some gene misregulation in the somatic cells of *K27me3 M+P-* L1s (Figure 8A). Most misregulated genes were autosomal and biased towards upregulation (Figure 8A & 8B). In contrast, X-linked genes showed no preference for upregulation vs downregulation (Figure 8A,B). To investigate whether this trend holds for a larger set of autosomal and X-linked genes, I analyzed a less-stringent set of misregulated genes (p-value < 0.4). I found that this larger set of genes displays a similar trend for autosomal genes, in which the average fold change is greater than 0. However, X-linked genes appeared to display a slight bias toward downregulation (Figure 8C). As discussed earlier, immunofluorescence images of *K27me3 M+P-* embryos indicate that sperm-inherited autosomes retain their H3K27me3(-) state, while the sperm-inherited X chromosome acquires a H3K27me3(+) state (Figure 7A,B). The comparatively less H3K27me3 on autosomes in *K27me3 M+P-* cells relative to *K27me3 M+P+* cells during early embryogenesis is consistent with a general upregulation of autosomal genes. The similar levels of H3K27me3 on the X chromosomes in *K27me3 M+P-* and *K27me3 M+P+* is also consistent with the lack of X-linked gene misregulation.

K27me3 M+P- vs K27me3 M+P+ L1 larvae

A L2FC transcript reads vs significance

sig auto (45)
sig X-linked (5)

sig auto (72)
sig X-linked (5)

−log₁₀ p-value

L2FC transcript reads

B change in transcripts for significantly misreg genes

L2FC transcript RPKM

sig auto (117)

sig X-linked (10)

C change in transcripts for low-stringency genes

L2FC transcript RPKM

auto (458)

X-linked (57)

D gamete-specific change in transcripts for signficantly misreg genes

gamete-specific L2FC in SNP-containing transcript reads

sig down (20)
oo   sp

oo   sp
sig up (6)

E gamete-specific change in transcripts for low-stringency genes

gamete-specific L2FC in SNP-containing transcript reads

down (81)
oo   sp

oo   sp
up (33)

F changes in H3K36me3 vs changes in transcripts

L2FC transcript RPKM

changes in H3K36me3 levels in *PRC2(-)* fathers

G changes in transcripts vs changes in transcripts

L2FC transcript RPKM

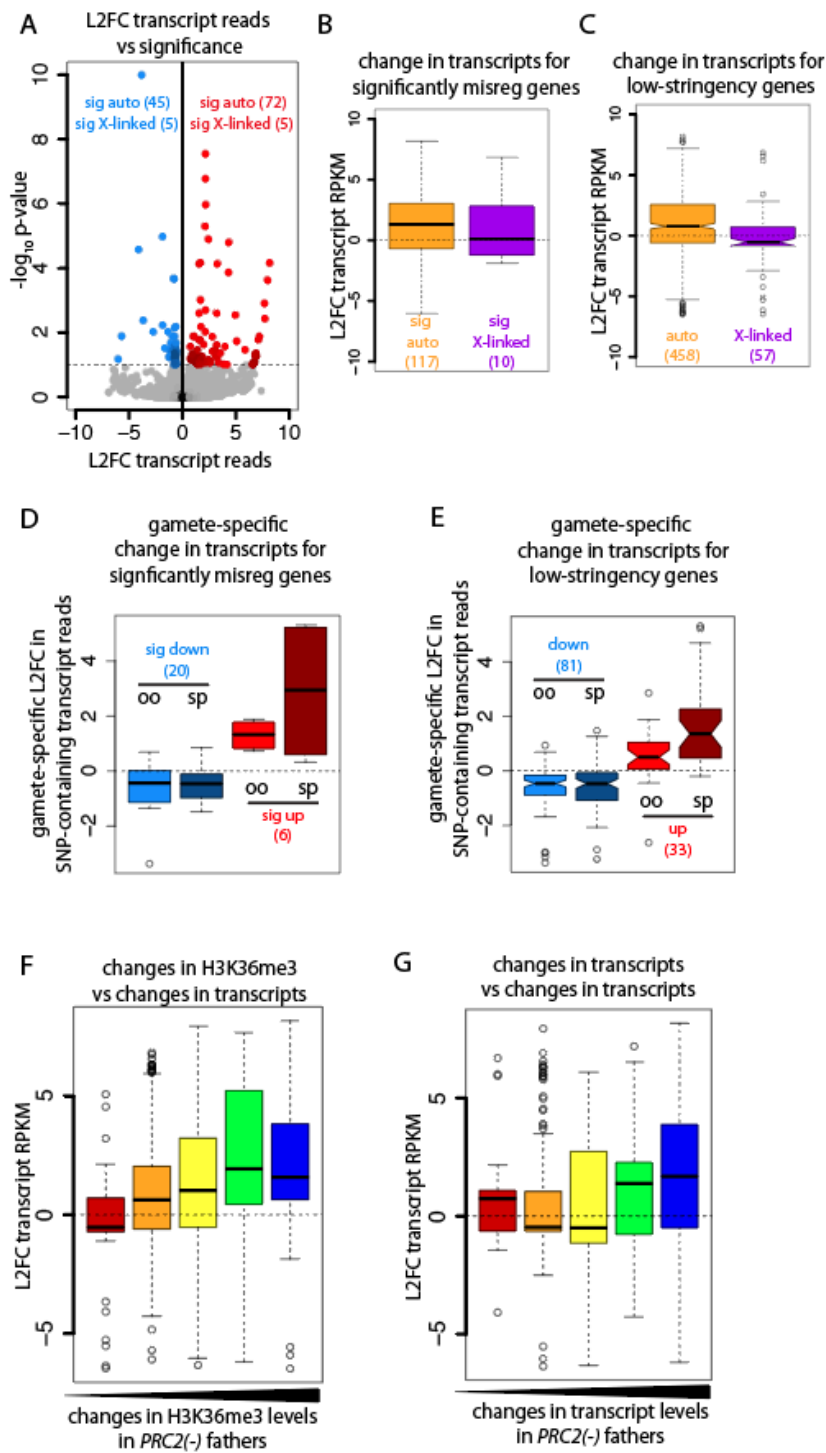changes in transcript levels in *PRC2(-)* fathers

72

Figure 8. Increases in H3K36me3 in the germline of *PRC2(-)* fathers correlates with increased transcript levels in the soma of offspring. **A** Volcano plot comparing $\log_2$ fold change (L2FC) in transcript reads vs $-\log_{10}$ p-value from *K27me3 M+P-* vs *K27me3 M+P+* L1 larvae. The number of upregulated or downregulated autosomal and X-linked genes are indicated in the upper right and left, respectively. Horizontal dashed line corresponds to a p-value of 0.1. **B-G** Boxplots of L2FC in transcript RPKM (**B,C,F,G**) or L2FC SNP-containing reads specific to oocyte or sperm alleles (**D,E**) from *K27me3 M+P- vs M+P+ L1s.* **B,C** Gold and purple boxes represent autosomal and X-linked genes, respectively. The number of upregulated or downregulated autosomal and X-linked genes are indicated below each box. Genes that are significantly misregulated between *K27me3 M+P-* and *M+P+* L1s (p-value <0.1) are shown in panel **A**; genes with a p-value < 0.4 are shown in panels **C,E,F,G**. Significantly up- and down-regulated genes (**A**) and boxes (**D**) are shown in red and blue, respectively. **D,E** The numbers of upregulated and downregulated SNP-containing genes analyzed in each plot are shown in parentheses. Sperm (sp) and oocyte (oo) comparisons are indicated. **F,G** Boxes show progressive increases in L2FC gene body H3K36me3 RPKM (**F**) and L2FC in transcript RPKM (**G**) from *PRC2(-)* vs WT male germlines. Red through blue boxes contain L2FC increments of 0.5 and range from L2FC -1 (red) to 2 (green). The last box contains genes with a L2FC greater than 2 (blue). The triangle below plot indicates the direction of increasing L2FC. Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend to the most extreme data point that is no more than 1.5 times the interquartile range from the box; individual data points outside of these boundaries are shown as open circles. The waist indicates the 95% confidence interval for the median.

In Chapter 2, I used single nucleotide polymorphisms (SNPs) between sperm and oocyte genomes to determine from which parental allele misexpression emanates in *K27me3 M+P-* germlines. I applied the same strategy to assess the transcriptional changes that occur in *K27me3 M+P-* L1 larvae. Similar to what I found in *K27me3 M+P-* germlines, I found that upregulated genes are upregulated to a greater degree from sperm alleles than from oocyte alleles (Figure 8D dark and light red boxes), while downregulated genes are equally downregulated from both sperm and oocyte alleles (Figure 8D blue boxes). I found that these relationships in misexpression from

sperm vs oocyte alleles persist in a larger more relaxed set of misregulated genes (p-value < 0.4) (Figure 8E).

Next, I determined whether there was a relationship between changes that occur in the germline of *PRC2(-)* fathers and changes that occur in the soma of their offspring. I focused my attention on autosomal genes since the sperm-inherited H3K27me3(-) state of the X chromosome is rapidly lost in early embryos (Figure 7A,B). To investigate this, I binned all coding genes based on the degree to which H3K36me3 levels change in the germlines of *PRC2(-)* fathers compared to WT. Then I assessed autosomal genes within each bin for changes in transcript levels in offspring. This analysis revealed a positive correlation between increases in H3K36me3 in the germlines of *PRC2(-)* fathers and increases in offspring somatic transcript levels (Figure 8F). The greater the gain in H3K36me3 in the germline of fathers, the greater the increase in transcript level in the soma of offspring. Next, I investigated whether there was also a relationship between changes in transcript levels in the germlines of fathers and changes in transcript levels in the soma of offspring. This time I binned all coding genes based on the degree to which transcript levels change in the germlines of *PRC2(-)* fathers compared to WT. Then I assessed autosomal genes within each bin for changes in transcript levels in offspring. This analysis revealed that transcript level changes in the germlines of fathers is not a robust predictor of transcript level changes in offspring (Figure 8G).

H3K27me3 and H3K36me3 patterns in *K27me3 M+P- vs M+P+* hermaphrodite germlines

In Chapter 2, I analyzed *K27me3 M+P-* adult hermaphrodites for consequences to germline health and function that result from inheriting the sperm genome lacking H3K27me3. In brief, I found that *K27me3 M+P-* worms, although fertile, misexpress genes in their germlines. Upregulated genes were primarily the result of increased transcription from sperm alleles, while downregulated genes were the result of reduced transcription from both sperm and oocyte alleles. I also presented several lines of evidence that this altered inheritance results in compromised germline health and function. For instance, while *K27me3 M+P-* worms are typically fertile, they give rise to a significant proportion of sterile offspring. Additionally, if *K27me3 M+P-* worms are born to *PRC2/+* heterozygous mothers, the reduced maternal load of PRC2 renders them sterile. These sterile *K27me3 M+P-* worms generate germ cells. However, the germ cells fail to express key germline markers, such as PGL-1 and HTP-3, and instead express neuronal markers. In this chapter, I again investigate the germlines of *K27me3 M+P-* hermaphrodite worms. This time, I investigate *K27me3 M+P-* germlines for changes in H3K27me3 and H3K36me3 states both at the global and allele-specific level. I then compare chromatin and transcriptional changes that occur in offspring germlines with those that occurred in the germlines of their *PRC2(-)* fathers. My aim is to investigate if the transcriptional and chromatin states in the germlines of fathers influence the chromatin states and thus transcriptional states in offspring.

***- K27me3 M+P- hermaphrodites upregulate autosomal genes and downregulate X-
linked genes in their germlines***

To investigate how the distributions of these two marks are altered in worms
that inherited the sperm genome lacking H3K27me3, Thea Egelhofer and I performed
CRun-seq analysis of *K27me3 M+P- and K27me3 M+P+* hermaphrodite germlines.
First, I analyzed *K27me3 M+P-* germlines for global trends in misexpression across
autosomal and X-linked genes. To do this, I generated a large set of misregulated
genes with a low-stringency cutoff (p-value < 0.4). Using a larger, low-stringency set
allowed me to analyze autosomal and X-linked misregulation at a broader level than a
smaller, higher-stringency set would. I assessed autosomal and X-linked genes in this
low-stringency set for changes in transcript, H3K27me3 and H3K36me3 levels. I
found that autosomal genes tend to gain H3K36me3 and lose H3K27me3 across their
gene bodies (gold dots in Figure 9A). Consistent with this, autosomal genes are, on
average, upregulated in the germlines of *K27me3 M+P-* hermaphrodites (gold box
Figure 9B). This upregulation is associated with an average gain in H3K36me3 levels
and reduction in H3K27me3 levels across the gene bodies of autosomal genes (gold
boxes in Figure 9C,D). In contrast, X-linked genes tend to lose H3K36me3 and gain
H3K27me3 (purple dots in Figure 9A). Consistent with this, X-linked genes are
globally downregulated (purple box in 9B). Illustrating this point is the observation
that, even with this relaxed set of misregulated genes, only a single X-linked gene
qualifies as upregulated while the remaining 184 qualify as downregulated. Assessing

these X-linked genes for changes in chromatin, I found that, consistent with their

downregulation, misregulated X-linked genes are associated with a reduction in

H3K36me3 levels and gain in H3K27me3 (purple boxes in Figure 9C,D).

misregulated autosomal and X-linked genes from
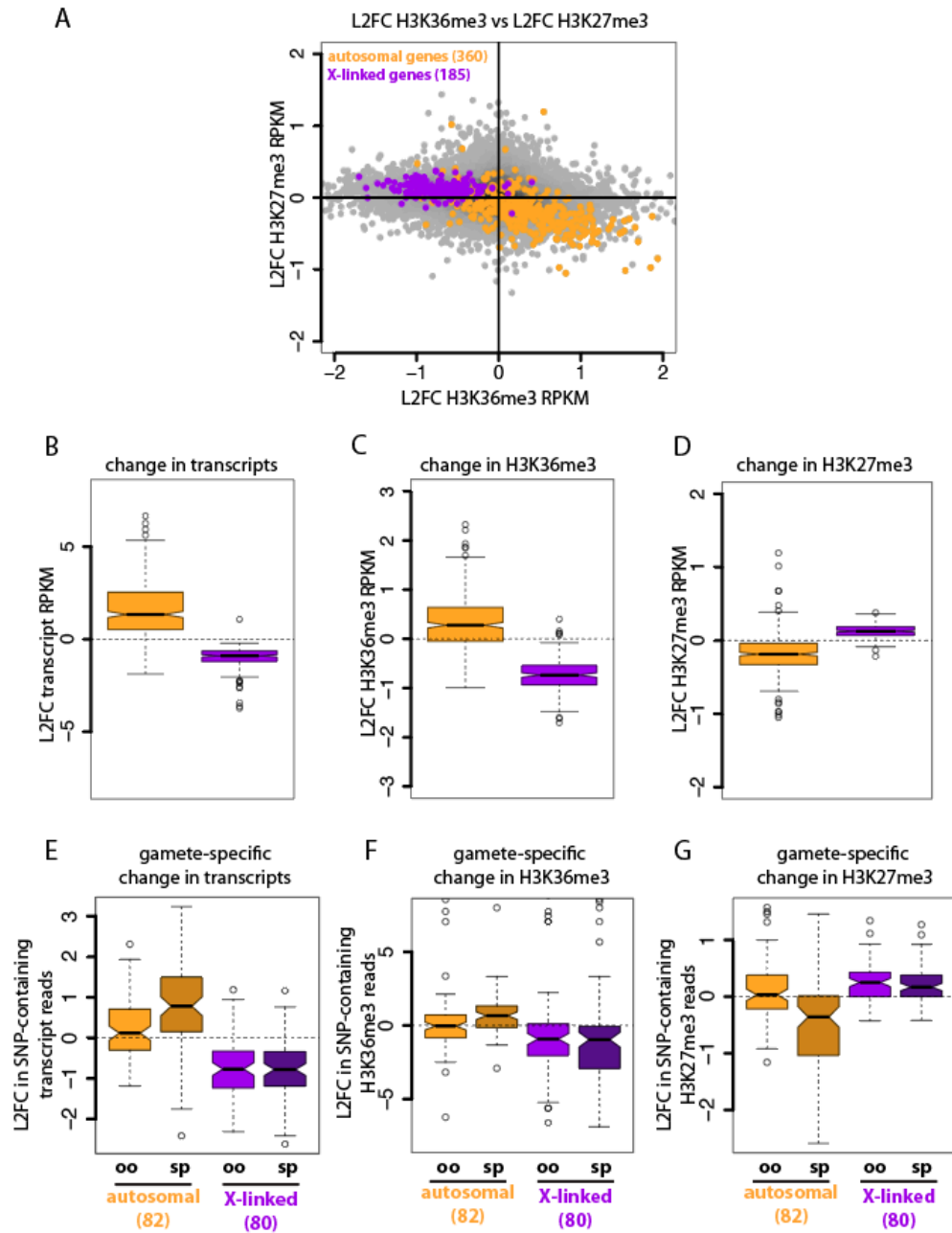***K27me3 M+P-* vs *K27me3 M+P+* germlines**

Figure 9. Sperm but not oocyte alleles have increased transcription, increased H3K36me3 and reduced H3K27me3 on autosomal genes in *K27me3 M+P-* germlines. **A-G** Scatterplot and boxplots analyzing autosomal and X-linked genes with a p-value <0.4 based on differential analysis (Deseq2) of transcript reads from *K27me3 M+P-* vs *K27me3 M+P+* germlines. **A** Scatterplot of log2 fold change (L2FC) of gene body H3K36me3 RPKM vs L2FC of gene body H3K27me3 RPKM from *K27me3 M+P-* vs *K27me3 M+P+* germlines. **B-D** Boxplots of L2FC transcript RPKM (**B**), L2FC gene body H3K36me3 RPKM (**C**), and L2FC gene body H3K27me3 RPKM (**D**). **E-G** Boxplots of L2FC SNP-containing transcript (**E**), H3K36me3 (**F**), and H3K27me3 (**G**) reads specific to oocyte or sperm. Autosomal and X-linked genes are shown in orange and purple, respectively. In panels **E-G**, oocyte and sperm allele-specific boxes for autosomal and X-linked genes are shown in light gold and purple or dark gold and purple, respectively. The number of autosomal and X-linked genes analyzed in each plot is indicated in parentheses. Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend to the most extreme data point that is no more than 1.5 times the interquartile range from the box; individual data points outside of these boundaries are shown as open circles or beyond the boundaries of the plot (**F**). The waist indicates the 95% confidence interval for the median.

- Autosomal but not X-linked genes are misregulated in cis in K27me3 M+P-

germlines

To assess from which parental genome changes in transcription and marking

emanate in *K27me3 M+P- germlines*, I performed SNP analysis on the relaxed set of

misregulated autosomal and X-linked genes. This entailed comparing SNP-containing

reads that emanate from sperm and oocyte alleles in *K27me3 M+P-* relative to

K27me3 M+P+ germlines. The results of this analysis indicate that upregulated

autosomal genes have increased transcription specifically from sperm alleles in

K27me3 M+P- germlines (dark gold box in Figure 9E). Sperm alleles specifically

have increased levels of H3K36me3 and reduced levels of H3K27me3 (dark gold

boxes in Figure 9F,G). In contrast, oocyte alleles, on average, do not have increased

transcription of autosomal genes (light gold box in Figure 9E) nor do they differ in

H3K36me3 or H3K27me3 levels relative to *K27me3 M+P+* oocyte alleles (light gold boxes in Figure 9F,G). These findings demonstrate that differences in the H3K27me3 state of sperm-inherited alleles resulted in upregulation of autosomal genes in *cis.* In contrast, X-linked genes did not display this *cis* misregulation. Instead, X-linked genes had equivalently reduced transcription from sperm and oocyte alleles. This reduction in transcript levels was associated with equivalently reduced H3K36me3 levels and equivalently increased H3K27me3 levels from sperm and oocyte alleles. These findings are consistent with the observation that the sperm-inherited X chromosome acquires H3K27me3 during early embryogenesis. Based on this observation, the PGCs inherit X-linked genes that, regardless of their gamete source, are H3K27me3(+).

- ***Significantly misregulated genes retain differential marking on sperm vs oocyte alleles in K27me3 M+P- germlines***

In Chapter 2, I analyzed significantly misregulated genes in *K27me3 M+P-* germlines. In this chapter, I aimed to extend that analysis to encompass changes in H3K27me3 and H3K36me3 as well. I reanalyzed the RNA-seq data presented in Chapter 2 with an updated release of DESeq2. This reanalysis called a few more genes significantly misregulated relative to the analysis presented in Chapter 2. Nonetheless, as previously reported, most significantly misregulated genes are upregulated (Figure 10A,B). This upregulation is primarily due from increased transcription from sperm alleles specifically (dark red box in Figure 10E). Similar to

what was observed for X-linked genes, significantly downregulated genes, which are primarily X-linked, show reduced transcription from both sperm and oocyte alleles (blue boxes in Figure 10E). Next, I assessed significantly misregulated genes for changes in H3K27me3 and H3K36me3. I found that significantly upregulated genes have increased H3K36me3 and decreased H3K27me3 across their gene bodies (red boxes in Figure 10C,D). SNP analysis indicates that these differences in H3K36me3 and H3K27me3 are due to changes in these marks on the sperm genome specifically (red boxes in Figure 10F,G). Downregulated genes have reduced H3K36me3 and increased H3K27me3 across their gene bodies (blue boxes in Figure 10F,G). Unlike upregulated genes, the changes in these marks on downregulated genes is due to equivalent changes on sperm and oocyte alleles. These findings indicate that inheriting the sperm genome lacking H3K27me3 results in upregulation of genes in *cis* and downregulation of genes in *trans*.

Figure 10. Significantly misregulated genes have increased transcription primarily from sperm alleles and increased H3K36me3 and reduced H3K27me3 specifically from sperm alleles in *K27me3 M+P-* germlines. **A-G** Volcano and boxplots analyzing significantly (p-value<0.1) up- and down-regulated genes from *K27me3 M+P-* vs *K27me3 M+P+* germlines. **A** Volcano plot comparing $\log_2$ fold change (L2FC) in transcript reads vs $-\log_{10}$ p-value. The number of up- or down-regulated genes are indicated in the upper right and left, respectively. Horizontal dashed line corresponds to a p-value of 0.1. **B-D** Boxplots of L2FC transcript RPKM (**B**), L2FC gene body H3K36me3 RPKM (**C**), and L2FC gene body H3K27me3 RPKM (**D**). **E-G** Boxplots of L2FC SNP-containing transcript (**E**), H3K36me3 (**F**), and H3K27me3 (**G**) reads

81

specific to oocyte or sperm alleles. Significantly up- and down-regulated genes are shown in red and blue, respectively. In panels **E-G**, oocyte and sperm allele-specific boxes for up- and down-regulated genes are shown in light red and blue or dark red and blue, respectively. The number of up- and down-regulated genes analyzed in each plot is indicated in parentheses. Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend to the most extreme data point that is no more than 1.5 times the interquartile range from the box; individual data points outside of these boundaries are shown as open circles. The waist indicates the 95% confidence interval for the median.

- Transcriptional and H3K36me3 changes in the germlines of PRC2(-) fathers

correlates with changes in the germlines of their offspring

Next, I assessed whether changes in transcript or H3K36me3 levels in the germlines of *PRC2(-)* fathers correlate with changes in transcription in the germlines of offspring. To investigate this, I binned all coding genes based on the degree to which transcript levels change in the germlines of *PRC2(-)* fathers compared to WT and assessed each bin for changes in offspring transcription (Figure 11A). I found that changes in transcript levels in the germlines of fathers positively correlates with changes in transcript levels in the germlines of offspring (Figure 11A). I performed a similar analysis to determine if changes in H3K36me3 levels in the germlines of *PRC2(-)* fathers similarly predict changes in offspring transcription (Figure 11B). The results of this analysis indicate that changes in H3K36me3 levels in the germlines of fathers shows a similar albeit weaker correlation with changes in transcript levels in the germline of offspring (Figure 11B). These findings indicate that both the transcriptional and H3K36me3 changes that occur in the germlines of *PRC2(-)* fathers predict changes in transcription in the germlines of offspring.
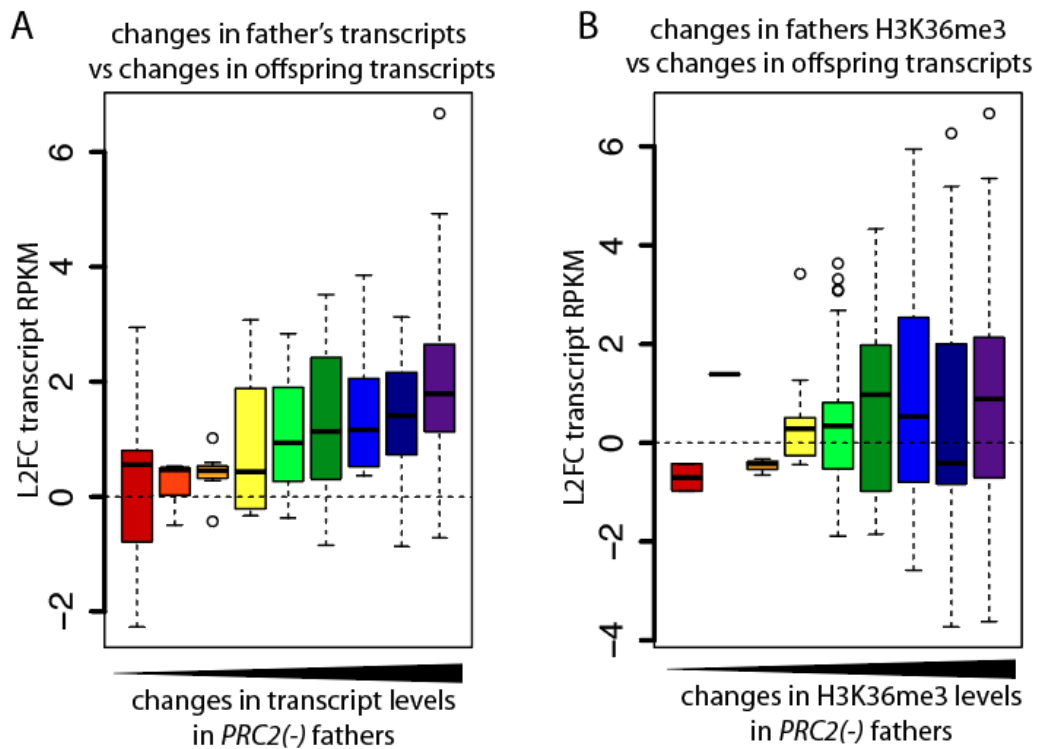
Figure 11. Transcriptional changes that occur in offspring germlines are predicted by transcriptional and H3K36me3 changes that occur in the germlines of fathers. **A,B** Boxplots of log2 fold change (L2FC) in transcript RPKM from *K27me3 M+P-* vs *K27me3 M+P+* germlines. Boxes show progressive increase in L2FC transcript RPKM (**A**) and L2FC gene body H3K36me3 RPKM (**B**) from *PRC2(-)* vs WT male germlines and range from negative L2FC (less than -2 in red) to positive L2FC (greater than 2 in purple). Triangle below plots indicates the direction of increasing L2FC. Only genes with a p-value <0.4 based on differential analysis (Deseq2) of transcript reads from *K27me3 M+P-* vs *K27me3 M+P+* (**B**) and additionally a p-value <0.1 based on differential analysis (Deseq2) of transcript reads from *PRC2(-)* vs WT male germlines (**A**) were analyzed. The triangle below plot indicates the direction of increasing L2FC. Each box extends from the 25th to the 75th percentile, with the median indicated by the horizontal line; whiskers extend to the most extreme data point that is no more than 1.5 times the interquartile range from the box; individual data points outside of these boundaries are shown as open circles.

III: Discussion

The distributions of H3K27me3 and H3K36me3 in the germlines of male *C.*

elegans largely reflect the transcriptional patterns of autosomal but not X-linked

genes. A fundamental question is whether these distributions are the cause or consequence of transcription? The work presented in my thesis provides evidence for both. In the absence of H3K27me3, many regions of the genome gain H3K36me3 and many genes become misregulated. Does the absence of repressive H3K27me3 or the gain in 'active' H3K36me3 *cause* transcriptional changes? In terms of H3K27me3 absence, *PRC2(-)* males lack H3K27me3 on all genes in their germlines and yet most genes do not become misregulated. This suggests that the absence of H3K27me3 is not sufficient to activate genes. However, the absence of H3K27me3 does render a subset of genes susceptible to upregulation. This upregulation is associated with a gain in H3K36me3. However, many genes that gain H3K36me3 show no changes in transcription and, in many cases, have reduced transcription. This suggests that the acquisition of H3K36me3 in the germlines of *PRC2(-)* males is also not sufficient to activate genes. If the absence of H3K27me3 and/or the gain in H3K36me3 is not sufficient to activate a gene, then what predisposes some genes to misregulation when H3K27me3 is absent?

I hypothesize that germline-expressed transcription factors have both germline-appropriate and germline-inappropriate targets. In wild-type (WT) germlines, germline-inappropriate targets are enriched for H3K27me3 and therefore resistant to activation. In the absence of H3K27me3, either from all alleles in the case of *PRC2(-)* male germlines or from sperm alleles specifically in *K27me3 M+P-* germlines, these germline-inappropriate targets can be activated. If some of these target genes are transcription factors themselves, then they can activate their own

targets on H3K27me3(-) and possibly H3K27me3(+) alleles. This may explain genes that become upregulated from both H3K27me3(+) oocyte and H3K27me3(-) sperm alleles in the soma and germline of *K27me3 M+P-* worms. Driving transcription factors in *K27me3 M+P-* germlines could reveal a differential susceptibility of sperm and oocyte alleles to transcriptional activation. If such a difference in susceptibility is found, it would provide support for this hypothesis.

The finding that H3K36me3 accumulates in regions of the genome that, in WT males, is highly enriched for H3K27me3 suggests that in WT males, H3K27me3 antagonizes H3K36me3 deposition by H3K36 HMTs. I provide evidence that in the germlines of *PRC2(-)* males, H3K36me3, in many cases, is deposited in a transcription-independent manner. This observation, if validated, implicates the H3K36 HMT MES-4, the only worm HMT identified to date that generates transcription-independent H3K36me3. MES-4 is highly expressed in the germline. However, its localization is not uniform. In wild-type hermaphrodite germlines, MES-4 protein is detected in the mitotically dividing cells of the distal germline and in the proximal cells undergoing oogenesis. This leaves the bulk of pachytene germ cells depleted of MES-4 activity. If these general distributions are observed in male germlines, this would suggest that the transcription-independent deposition of H3K36me3 in *PRC2(-)* male germlines primarily occurs in the distal most germ cells. Alternatively, the localization of MES-4 in *PRC2(-)* male germlines could be altered. Immunostaining of MES-4 in WT and *PRC2(-)* male germlines could distinguish between these possibilities.

How do changes in transcription or H3K36me3 levels in the germlines of
fathers influence transcription in the germline or soma of offspring? I already
discussed the hypothesis that the absence of H3K27me3 from sperm-inherited alleles
makes them susceptible to activation by factors already present in the germline or
somatic cells of offspring. In support of this, there is little correlation between the
genes that become misexpressed in the germlines of fathers and the genes that
become misexpressed in the soma of offspring. One exception to this is the
observation that the most highly upregulated genes in the germline of fathers are
prone to upregulation in somatic cells of offspring. This could be due to the higher
levels of H3K36me3 at these loci or the presence of other marks associated with
transcription that either directly predispose genes to transcriptional activation or
antagonize PRC2 activity during embryogenesis and thereby perpetuate the
H3K27me3(-) state of sperm-inherited alleles. Consistent with this, genes that gain
the most H3K36me3 in fathers tend to have increased transcription in the soma of
offspring. Similar to the most highly upregulated genes in the germlines of fathers,
these genes may be more likely to perpetuate the H3K27me3(-) state during
embryogenesis via H3K36me3 antagonism of PRC2. If these genes maintain the
H3K27me3(-) state later into embryogenesis, then it stands to reason that these genes
would be more susceptible to activation by factors in the somatic cells of offspring. In
summary, these findings indicate that misexpression of genes in the soma of *K27me3*
M+P- offspring is best predicted by H3K36me3 changes and not transcriptional
changes in the germlines of *PRC2(-)* fathers. Interestingly, the reverse is true for

misexpression of genes in the germlines of *K27me3 M+P-* offspring. These differences may be due to the high likelihood of shared transcription factors in the germlines of *K27me3 M+P-* offspring and the germlines of their *PRC2(-)* fathers. Shared factors would result in the activation, and possibly repression, of similar sets of genes. The added effect of increased transcription generating higher levels of H3K36me3 in the father's germline could enhance the susceptibility of these genes to activation in offspring, thus amplifying the positive correlation between increased transcription in the germlines of fathers with increased transcription in the germlines of offspring.

The somatic cells of offspring and the germlines of their fathers are likely to have a very different repertoire of transcription factors. For this reason, I hypothesize that the genes that become activated in the soma of *K27me3 M+P-* larvae are first and foremost determined by the factors present in the soma and secondly by the degree to which they retain the H3K27me3(-) state of sperm alleles during embryogenesis. The former is unlikely to be influenced by the changes that occur in the germlines of fathers. Therefore, the latter is more likely to link changes in the germlines of fathers with changes in the soma of offspring. The observation that H3K36me3 antagonizes PRC2 supports the likely correlation between genes with the highest H3K36me3 in the germlines of fathers and those that retain the H3K27me3(-) state later into embryogenesis. For this reason, misexpression of genes in the soma of *K27me3 M+P-* offspring is better predicted by changes in H3K36me3 levels rather than changes in transcript levels in the germlines of *PRC2(-)* fathers. I analyzed the

transcript levels of known and putative transcription factor-encoding genes in male germlines, hermaphrodite germlines and L1s. I found that there are many shared factors between male and hermaphrodite germlines and very few shared factors between germlines and L1s. These findings support the hypothesis that the genes that become misregulated in *K27me3 M+P-* worms is largely dictated by the transcription factors present in the tissues being profiled.

The work in this thesis presents evidence that histone marks on sperm- and likely oocyte-inherited genomes can directly influence the transcriptional patterns in the cells of offspring. I envision that the presence of H3K27me3 over a given allele provide a threshold that a transcription factor must overcome in order to activate its target gene. In the case of the germline, H3K36me3 and H3K27me3 pattern the autosomal genome in a manner congruent with germline gene expression patterns. These patterns are delivered to the embryo on sperm and egg genomes. Through their antagonistic activities, MES-4 and PRC2 perpetuate these patterns to the PGCs, where H3K27me3 levels set appropriate thresholds for gene activation in the nascent germline. For example, germline-expressed genes lack H3K27me3(-) because MES-4 maintains gamete-inherited H3K36me3 at these loci. On the other hand, germline-repressed genes are enriched for H3K27me3 because PRC2 perpetuates the inherited H3K27me3(+) state at these loci. In contrast, the X chromosome needs to be reset to a repressed state by the activity of PRC2 during embryogenesis. PRC2 may also continue to antagonize transcription of X-linked genes in the germline. It may be for this reason that, although some X-linked genes are highly expressed in WT male

germlines, they still retain high levels of H3K27me3 and low levels of H3K36me3. This hypothesis requires that X-linked genes required for germline function must have strong transcription factors capable of activating their target genes despite H3K27me3 antagonism.

Many aspects of the above model have yet to be tested. Identifying germline-expressed transcription factors and defining their target genes would be extremely useful in designing experiments to test this model. A first step could include an in-depth motif analysis of misregulated genes in the germlines and soma of *K27me3 M+P-* worms and in the germlines of their *PRC2(-)* fathers. As mentioned earlier, inducing expression in *K27me3 M+P-* germlines of a transcription factor with well-defined target genes could determine whether the differences in H3K27me3 states of sperm and oocyte alleles translate into a differential ability of those alleles to be activated. Refining this model through continued investigation will help us understand how histone marks inherited across generations and cell divisions can influence transcription. Differences in parental germline transcription or differences in the activity of chromatin modifiers can alter the distributions of histone marks on sperm and oocyte genomes. These events can occur in response to parental environment. For instance, there are several FDA-approved drugs that target chromatin modifiers, including PRC2, that are currently used in clinical trials for the treatment of cancer (Kong *et al.* 2014; Raynal *et al.* 2017). These drugs influence histone marks in somatic cells and therefore have the potential to modify histone marks in germ cells as well. This is a clear example of how histone marks in germ

cells could be affected by environmental influences. There may be more subtle ways in which the environment can alter germ cell histone marks. If histone marks inherited across generations and cell divisions indeed set thresholds for gene activation, then subtle changes in histone marking on sperm or oocyte genomes could alter transcription patterns in offspring. Depending on the time in development and the tissue affected, even slight alterations in gene activation thresholds could have important impacts on development.

In this thesis, I presented evidence that supports a model in which histone marks inherited on gamete genomes shape transcription in offspring by setting appropriate thresholds for gene activation. Inheriting the appropriate distributions of these marks is particularly important in the germline, as alterations in sperm marking compromises germline health and function. Refining this model through continued investigation will deepen our understanding of how epigenetic information transmitted across generations shapes development and how alterations to the epigenome can influence the development and health of future generations.

IV: Methods:

C. elegans **growth conditions.** Worms were maintained on Nematode Growth Medium (NGM) agar plates spotted with *Escherichia coli* OP50 and maintained at 15˚C (*fem-2, mes-3(bn21ts)* and *plk-1* containing strains) or 20˚C (all other strains). To generate females, *fem-2* strains were raised at 24˚C until L4.

C. elegans **strains.** All strains are in the N2 background except CB4856 and SS1292.

N2 wild-type isolate from Bristol, England

CB4856 wild-type isolate from Hawaii

SS1292 *fem-2(b245ts) III* - mutation CRISPR-engineered into CB4856

DH0245 *fem-2(b245ts) III*

SS1099 *mes-3(bn35) I/hT2-GFP (I;III); him-8(e1489) IV*

SS1199 *mes-3(bn21ts) I/hT2-GFP (I;III)*

OCF65 *plk-1(or683) III; fog-1(e2121) unc-11(e47) I/hT2 [qIs48] (I:III);*
     itIs37[pAA64: pie-1p::mCHERRY::his-58 + unc-119(+)].

CB1489 *him-8(e1489) IV*

Fixing and staining embryos, larvae, and germlines. Specimens in the buffer (see below) used above were flattened between a slide and a coverslip by wicking away buffer. Slides were frozen in liquid nitrogen, coverslips were removed, then slides were fixed in methanol (10min) and acetone (10min) and allowed to air dry for up to 3 days. Slides were blocked (1.5% ovalbumin, 1.5% BSA in PBS) for 25min, then incubated with primary antibodies overnight at 4˚C. Slides were washed 3X in PBST (0.1% tween in phosphate buffered saline pH7.2), blocked for 5min, incubated with secondary antibodies and DAPI at room temperature for 2hr, and washed 3X in PBST. Coverslips were mounted on slides with gelutol. Slides were allowed to dry for up to 3 days before being imaged.

Antibodies. Primary antibodies and their dilutions were as follows: 1:30,000 of

2.45mg/mL mouse anti-H3K27me3 (Kimura mAb 1E7 clone CMA323, Wako

cat#309-95259), 1:20,000 of stock concentration rabbit anti-H3K4me2 (Millipore 07-

030 lot# DAM1570816), 1:1000 rabbit anti-MES-4 purified serum (IU243 bleed 2

eluate 1). Secondary antibodies conjugated to Alexa Fluor: 488 goat anti-mouse (Life

cat#A11001), 488 goat anti-rabbit (Molecular Probes cat#A-11008) and 594 goat

anti-rabbit (Life cat#A11012) were used at 1:300 with 0.05μg/ml 4',6-diamidino-2-

phenylindole (DAPI).

Microscope set-up for immunocytochemistry. All images generated in this thesis

were acquired using the following: Yokogawa CSUX-1 spinning disk scanner, Nikon

(Garden City, NY) TE2000-E inverted stand, Hamamatsu ImageEM X2 camera, solid

state 405-nm, 488-nm and 561-nm laser lines, 435-485, 500-550 and 573-613

fluorescent filters, Nikon Plan Apo VC 60x/1.40 oil objective, Nikon Plan Apo

100x/1.40 oil objective, and Micro-Manager software.

Collection of *K27me3 M+P-* and *M+P+* L1 larvae. *fem-2(ts)* females were mated to

mes-/hT2-GFP; him-8 (P+) or *mes-3 M-Z-; him-8* (P-) males overnight at 20˚C.

Gravid females were transferred to fresh plates seeded with OP50 and allowed to lay

eggs overnight at 20˚C. The next day, late-stage embryos along with some bacteria

were transferred to 50uL drops of S Basal (5.85g NaCl, 1g K_2HPO_4, 6g KH_2PO_4, 1ml

cholesterol (5mg/mL in ethanol), H_2O to 1 liter) on gelatin chrom alum-coated slides.

Drops were visualized to ensure that only embryos were transferred. After 2 hours at room temperature *K27me3 M+P-* and *K27me3 M+P+* (no balancer) L1s were transferred into 300uL of ice-cold TRIzol reagent (Life Technologies). This process of transferring embryos to S basal, incubating 2 hours and collecting L1s was repeated throughout the day. Samples in TRIzol were then stored at -80˚C. Samples were subjected to 3 freeze-thaw cycles. 1uL linear polyacrylamide was added to each sample, vortexed, then transferred to Phase Lock Gel-Heavy 2mL tubes to extract the aqueous phase. 0.75 volume isopropanol was added to each sample and incubated overnight at -20˚C. RNA was isolated via ethanol precipitation and rehydrated in 14.5uL RNase-free H2O. 2.5uL were used to assess RNA quality (Agilent RNA Nano Bioanalyzer Chip) and concentration (Qubit Quant-IT RNA Assay Kit). The remaining 12uL were polyA selected and used for library preparation (NEBNext Ultra RNA Library Prep Kit). Libraries were sequenced at UC Santa Cruz using the Illumina NextSeq platform (150 bp paired end reads). Three biological replicates each were generated for *K27me3 M+P-* and *K27me3 M+P+* hybrid L1 samples.

RNA-seq of *him-8 and mes-3(bn35) M+Z-* male germlines (Tabuchi *et al.* 2018).
Germlines were collected from CB1489 *him-8(e1489)* males and SS818 *mes-3(bn35)* homozygous M+ Z − males from heterozygous parents (M for maternal product and Z for zygotic product) germlines were dissected from day 1 adults as 20–100 distal gonad arms were cut at the late-pachytene gonad bend. Total RNA was extracted in TRIzol reagent (Invitrogen), ribosomal RNA was depleted using an NEBNext rRNA Depletion

kit (E6310), and libraries were prepared using an NEBNext Ultra RNA Library Prep Kit for Illumina (E7530). Libraries were sequenced at the Vincent J. Coates Genomics Sequencing Laboratory at University of California, Berkeley (50 bp single-end read sequencing). For each genotype, 2–4 biological replicates were obtained.

CUT&RUN on germlines

CUT&RUN was performed on gonads from young adult hermaphrodites or males following a protocol adapted from Skene *et al.* (2018). Gonads were dissected by picking 10-15 worms from a blank plate into egg buffer containing 0.5% Tween 20 and 1 mM Levamisole. Each worm was cut in half just below the pharynx using 30½ gauge needles. The extruded gonads were cut at the bend and the distal half was transferred through a drop of 100 ul GWB (20 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 0.05% BSA, protease inhibitor tablet), then transferred to a 1.5 ml epi tube on ice with 1 mL GWB. This process was repeated until each epi tube contained enough germlines for each sample: 16 for H3K27me3 and 50 for H3K36me3. 10 ul activated ConA beads (Bangs Laboratory cat. no. BP531) were added to each tube of germlines. ConA beads were activated by washing twice with 1.5 mL binding buffer (20 mM Hepes-KOH pH 7.9, 10 mM KCl, 1 mM CaCl2, 1 mM MnCl2), then resuspended in 10 ul binding buffer. For the remaining protocol, the beads/gonads were washed by doing a quick spin in a micro centrifuge, placed on magnet stand for 30 seconds, and supernatant was removed with a pipette. The activated beads and gonads were incubated 5 min at RT with gentle rotation then

washed twice with 1 mL GWB. Primary antibodies were added to 50 ul antibody

buffer (GWB with 2 mM EDTA and 0.1% digitonin [5% digitonin stock should be

made fresh, and can be kept for one week at 4°C]) at the following dilutions:

H3K27me3 (Cell Signaling Technology cat. no. 9733) at 1 ug, H3K36me3 (Wako

cat. no. 300-95289) at 0.5 ug. 50 ul of these diluted antibodies were added to each

sample and incubated at 4°C overnight with gentle rotation. Samples were washed

twice for 5 minutes with gentle rotation with 1 mL GWB-D (GWB with 0.1%

digitonin). Secondary Rabbit Anti-Mouse IgG antibody (Abcam cat. no. ab7070) was

added to 50 ul GWB-D at a final concentration of 1:100 to the H3K36me3 samples

and incubated at 4°C for 1 hour with gentle rotation. Samples were washed twice for

5 minutes with gentle rotation with 1 mL GWB-D and the beads were resuspended in

50 ul GWB-D. 2.5 ul of 1:10 dilution of pA-MNase (from Henikoff Lab) was added

to each sample. Tubes were placed on a rotator for 1 hour at room temperature, then

washed 3 times for 10 minutes with gentle rotation with 1 mL GWB-D. 100 ul GWB-

D was added to each tube, then incubated in cold metal block for 5 minutes. Targeted

digestion by pA-MNase was initiated by adding 2 ul 100 mM CaCl2 to each tube and

incubated in the cold block for 30 minutes. The digestion reaction was stopped by the

addition of 100 ul G2XSTOP buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA,

1% Triton-X, 5 ug RNAse A, 50 ug/mL glycogen, and 2 pg/mL spike-in DNA from

Henikoff Lab). The samples were mixed, incubated at 37°C for 30 minutes, spun

down for 5 minutes at 4°C at full speed, and placed on a magnet stand. The

supernatant was transferred to a new 1.5 mL tube containing 2 ul 10% SDS & 2.5 ul

Proteinase K (20 mg/ml) and incubated 10 minutes at 70°C. Each sample was phenol

chloroform extracted, ethanol precipitated overnight at -80°C, and the pellet was

resuspended in 13 ul water. 2 ul of each sample was run on High Sensitivity

TapeStation to check for cutting and the remaining 10 ul were used to make libraries

with the Ovation Ultralow System V2 kit from Tecan genomics (cat.no 0344NB-32).

Libraries were sequenced at UC Santa Cruz using the Illumina NextSeq platform (75

bp paired end).

Processing and analyzing RNA-seq & CRun-seq data. For differential expression

analysis, raw sequences were mapped to transcriptome version WS220 using

TopHat2 *(*Kim *et al.* 2013). Only reads with one unique mapping were allowed.

Otherwise default options were used. For RNA-depleted samples, reads mapping to

ribosomal RNAs were removed from further analysis. HTSeq (Anders *et al.* 2015)

was used to obtain read counts per transcript. DESeq2 (Love *et al.* 2014) was used to

determine differentially expressed genes from HTSeq counts. A Benjamini–Hochberg

multiple hypothesis corrected P-value cutoff of 0.1 was used as a significance cutoff

to identify significantly differentially expressed genes between *K27me3 M+P-* and

M+P+ germlines and L1s while a P-value cutoff of 0.05 was used for male

germlines.

To identify RNA-seq and CRun-seq reads coming from the Hawaiian or Bristol

parent in the hybrid offspring germlines, the annotated Single Nucleotide

Polymorphisms (SNPs) of the Hawaiian strain CB4856 genome compared to the Bristol strain N2 WS220 genome were downloaded from WormBase. 126,615 unique SNPs were downloaded, of which 25,980 SNPs mapped to an exon in a total of 9,146 coding genes, according to WormBase. Henceforth, we refer to these SNPs as exonic SNPs.

RNA-seq reads from the Bristol strain, Hawaiian strain, and hybrids were mapped to the Bristol (N2) genome version WS220 using TopHat2, once allowing 1 mismatch in a read and a second time not allowing a mismatch. Reads overlapping exonic SNP positions were identified, and HTSeq was used to count these SNP-containing reads for each transcript. Counts per transcript were normalized by the number of total transcript reads, including non-SNP-overlapping reads, for each of the libraries sequenced. An average count per transcript was obtained by averaging the replicates for each of the Bristol, Hawaiian, and hybrid *K27me3 M+P-* and *M+P+* replicate experiments. A further quality control step was imposed. For the Bristol and Hawaiian RNA-seq data, we calculated a distance score between the mismatch 0 and mismatch 1 read counts for each transcript. For each transcript i in the Hawaiian RNA-seq data , a score hz_i was calculated: $hz_i = (hmm1_i - hmm0_i)/(hmm1_i + hmm0_i)$ where $hmm1_i$ and $hmm0_i$ denote the normalized read counts for transcript i from mapping with a mismatch 1 and with a mismatch 0 to the Bristol genome, respectively. In the same way, we defined bz_i, $bmm1_i$, and $bmm0_i$ for transcripts from mapping our Bristol RNA-seq data. For Hawaiian transcript reads from exonic SNPs,

we expect hz_i to be close to 1; all reads should have mapped when allowing 1

mismatch, and none should have mapped when allowing no mismatches. For Bristol

transcript reads from exonic SNPs, we expect bz_i to be close to 0, as the number of

reads from both mappings should have been very similar. We removed transcripts

that did not follow these requirements. If a transcript i had a $hz_i < 0.75$ for the

Hawaiian RNA-seq data, or an absolute value $|bz_i| > 0.25$ for the Bristol RNA-seq

data, that transcript was removed from further analysis. These requirements removed

1,080 transcripts with exonic SNPs. For the remaining 8066 transcripts we obtained

from the hybrid *K27me3 M+P-* and *M+P+* RNA-seq data, we determined the

transcript reads from the sperm allele (Bristol) as reads that mapped when allowing 0

mismatches. We determined the transcript reads from the oocyte allele (Hawaii) as

reads that mapped when allowing 1 mismatch but that did not map when allowing 0

mismatches. We thus calculated transcript reads from the oocyte allele by subtracting

reads that mapped when allowing 0 mismatches (sperm allele reads) from reads that

mapped when allowing 1 mismatch (reads from both gamete alleles).

Many genes with exonic SNPs were not expressed in any of our germline RNA-seq

data. We filtered out these non-expressed transcripts from further analysis by

requiring the normalized read counts from the sperm allele or oocyte allele in at least

one condition, *K27me3 M+P-* or *M+P+*, to be at least 0.5. Because our goal was to

use SNP-containing transcript reads as a proxy for total transcript reads, we filtered

out transcripts that had a greater than 2.5-fold difference between the $\log_2$ of SNP-

containing reads versus the $\log_2$ of total reads for a given transcript in hybrid germlines (Hawaii mother and Bristol father). This requirement removed 190 genes and left 2,606 transcripts for further analysis. To avoid dividing by zero when calculating log fold changes and to reduce the effect of noise on low read counts, we added a pseudo-count of 0.5 to all remaining transcripts from oocyte alleles and sperm alleles in both conditions.

References

Acharya, S., Hartmann, M. & Erhardt, S. (2017). Chromatin-associated noncoding RNAs in development and inheritance. *Wiley Interdisciplinary Reviews: RNA*, *8*(6).

Adalsteinsson, B. T., & Ferguson-Smith, A. C. (2014). Epigenetic control of the genome—lessons from genomic imprinting. *Genes*, *5*(3), 635–655.

Alabert, C., Barth, T. K., Reverón-Gómez, N., Sidoli, S., Schmidt, A., Jensen, O.N. *et al.* (2015). Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes & Development*, *29*(6), 585–590.

Anders, S., Pyl, P. T. & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)*, *31*(2), 166–169.

Bao, J. & Bedford, M. T. (2016). Epigenetic regulation of the histone-to-protamine transition during spermiogenesis. *Reproduction*, *151*(5), R55–R70.

Bender, L.B., Cao, R., Zhang, Y. & Strome, S. (2004). The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in *C. elegans*. *Current Biology*, *93*(1), 43–46.

Bender, L. B., Suh, J., Carroll, C., Fong, Y., Fingerman, I. *et al.* (2006). MES-4: an autosome-associated histone methyltransferase that participates in silencing the X chromosomes in the *C. elegans* germ line. *Development*, *133*(19), 3907–3917.

Capowski, E. E., Martin, P., Garvin, C. & Strome, S. (1991). Identification of grandchildless loci whose products are required for normal germ-line development in the nematode *Caenorhabditis elegans. Genetics*, *129*(4), 1061–1072.

Castillo, J., López-Rodas, G. & Franco, L. (2017). Histone post-translational modifications and nucleosome organisation in transcriptional regulation: some open questions. *Adv Exp Med Biol - Protein Reviews,* 2:65–92

Coleman, R. T. & Struhl, G. (2017). Causal role for inheritance of H3K27me3 in maintaining the off state of a *Drosophila* HOX gene. *Science*, *356*(6333).

Delaney, K., Strobino, M., Wenda, J. M., Pankowski, A. & Steiner, F. A. (2019) H3.3K27M-induced chromatin changes drive ectopic replication through misregulation of the JNK pathway *in C. elegans*. *Nature Communications, 10*(1), 2529.

Dorighi, K. M. & Tamkun, J. W. (2013). The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in *Drosophila. Development (Cambridge, England)*, *140*(20), 4182–4192.

Escobar, A. T. M., Oksuz, O., Descostes, N., Bonasio, R. & Reinberg, D. (2018). Active and repressed chromatin domains exhibit distinct nucleosome segregation during DNA replication. bioRxiv, doi: http://dx.doi.org/10.1101/418707

Evans, K. J., Huang, N., Stempor, P., Chesney, M. A., Down, T. A. & Ahringer, J. (2016). Stable *Caenorhabditis elegans* chromatin domains separate broadly expressed and developmentally regulated genes. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(45), E7020–E7029.

Francis, N. J. & Kingston, R. E. (2001). Mechanisms of transcriptional memory. *Nature Reviews Molecular Cell Biology*, *2*(6), 409–421.

Furuhashi, H., Takasaki, T., Rechtsteiner, A., Li, T., Kimura, H. *et al.* (2010). Transgenerational epigenetic regulation of *C. elegans* primordial germ cells. *Epigenetics & Chromatin*, *3*(1), 15.

Gapp, K. & Bohacek, J. (2018). Epigenetic germline inheritance in mammals: looking to the past to understand the future. *Genes, Brain and Behavior*, *17*(3), e12407.

Gaydos, L. J., Rechtsteiner, A., Egelhofer, T. A., Carroll, C. R. & Strome, S. (2012). Antagonism between MES-4 and Polycomb Repressive Complex 2 promotes appropriate gene expression in *C. elegans* germ cells. *Cell Reports*, *2*(5), 1169–1177.

Gaydos, L. J., Wang, W. & Strome, S. (2014). H3K27me and PRC2 transmit a memory of repression across generations and during development. *Science*, *345*(6203), 1515–1518.

Hansen, K. H., Bracken, A. P., Pasini, D., Dietrich, N., Gehani, S. S., *et al.* (2008). A model for transmission of the H3K27me3 epigenetic mark. *Nature Cell Biology*, *10*(11), 1291–1300.

Hobert, O. (2013). Neurogenesis in the nematode *Caenorhabditis elegans*: *Patterning and Cell Type Specification in the Developing CNS and PNS. Comprehensive Developmental Neuroscience, Volume 1* (pp. 609–626).

Højfeldt, J. W., Laugesen, A., Willumsen, B. M., Damhofer, H., Hedehus, L. *et al.* (2018). Accurate H3K27 methylation can be established de novo by SUZ12-directed PRC2. *Nature Structural & Molecular Biology*, *25*(3), 225–232.

Holdeman, R., Nehrt, S. & Strome, S. (1998). MES-2, a maternal protein essential for viability of the germline in Caenorhabditis elegans, is homologous to a *Drosophila* Polycomb group protein. *Development (Cambridge, England)*, *125*(13), 2457–2467.

Hosogane, M., Funayama, R., Shirota, M. & Nakayama, K. (2016). Lack of transcription triggers H3K27me3 accumulation in the gene body. *Cell Reports*, *16*(3), 696–706.

Huang, C. & Zhu, B. (2018). Roles of H3K36-specific histone methyltransferases in transcription: antagonizing silencing and safeguarding transcription fidelity. *Biophysics Reports*, *4*(4), 170–177.

Jung, Y. H., Sauria, M. E. G., Lyu, X., Cheema, M. S., Ausio, J. *et al.* (2017). Chromatin states in mouse sperm correlate with embryonic and adult Regulatory landscapes. *CellReports*, *18*(6), 1366–1382.

Kaneshiro, K. R., Rechtsteiner, A. & Strome, S. (2019). Sperm-inherited H3K27me3 impacts offspring transcription and development in *C. elegans. Nature Communications*, *10*(1), 1271.

Kawasaki, I., Shim, Y. H., Kirchner, J., Kaminker, J., Wood, W. B. & Strome, S. (1998). PGL-1, a Predicted RNA-Binding Component of Germ Granules, Is Essential for Fertility in *C. elegans. Cell*, *94*(5), 635–645.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, *14*(4), R36.

Klosin, A., Casas, E., Hidalgo-Carcedo, C., Vavouri, T. & Lehner, B. (2017). Transgenerational transmission of environmental information in *C. elegans. Science*, *356*(6335), 320–323.

Klymenko, T. & Muller, J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Reports*, *5*(4), 373–377.

Kong, X., Chen, L., Jiao, L., Jiang, X., Lian, F. *et al.* (2014). Astemizole arrests the proliferation of cancer cells by disrupting the EZH2-EED interaction of polycomb repressive complex 2. *Journal of Medicinal Chemistry*, *57*(22), 9512–9521.

Korf, I., Fan, Y. & Strome, S. (1998). The Polycomb group in *Caenorhabditis elegans* and maternal control of germline development. *Development (Cambridge, England)*, *125*(13), 2469–2478.

Kreher, J., Takasaki, T., Cockrum, C., Sidoli, S., Garcia, B. A. *et al.* (2018). Distinct roles of two histone methyltransferases in transmitting H3K36me3-based epigenetic memory across generations in *Caenorhabditis elegans. Genetics*. *210*(3), 969-982.

Laprell, F., Finkl, K. & Müller, J. (2017). Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. *Science (New York, N.Y.)*, *356*(6333), 85–88.

Love, M. I., Huber, W. & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550.

MacQueen, A. J., Phillips, C. M., Bhalla, N., Weiser, P., Villeneuve, A. M. & Dernburg, A. F. (2005). Chromosome sites play dual roles to establish homologous synapsis during meiosis in *C. elegans. Cell*, *123*(6), 1037–1050.

Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J. *et al.* (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature*, *461*(7265), 762–767.

Millership, S. J., Van de Pette, M. & Withers, D. J. (2019). Genomic imprinting and its effects on postnatal growth and adult metabolism. *Cellular and Molecular Life Sciences*, 1–13.

Pauler, F. M., Sloane, M. A., Huang, R., Regha, K., Koerner, M. V, *et al.* (2009). H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies

a histone banding pattern on a mouse autosomal chromosome. *Genome Research*, *19*(2), 221–233.

Paulsen, J. E., Capowski, E. E. & Strome, S. (1995). Phenotypic and molecular analysis of *mes-3*, a maternal-effect gene required for proliferation and viability of the germ line in *C. elegans. Genetics*, *141*(4), 1383–1398.

Pembrey, M., Saffery, R. & Bygren, L. O. (2014). Human transgenerational responses to early-life experience: potential impact on development, health and biomedical research. *Journal of Medical Genetics*, *51*, 563–572.

Pengelly, A. R., Copur, Ö., Jäckle, H., Herzig, A. & Müller, J. (2013). A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science (New York, N.Y.)*, *339*(6120), 698–699.

Petryk, N., Dalby, M., Wenger, A., Stromme, C. B., Strandsby, *et al.* (2018). MCM2 promotes symmetric inheritance of modified histones during DNA replication. *Science (New York, N.Y.)*, eaau0294.

Poepsel, S., Kasinath, V. & Nogales, E. (2018). Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes. *Nature Structural & Molecular Biology*, *25*, 154-162.

Rahman, M. M., Munzig, M., Kaneshiro, K., Lee, B., Strome, S. *et al.* (2015). *Caenorhabditis elegans* polo-like kinase PLK-1 is required for merging parental genomes into a single nucleus. *Molecular Biology of the Cell*, *26*(25).

Raynal, N. J.-M., Da Costa, E. M., Lee, J. T., Gharibyan, V., Ahmed, S. *et al.* (2017). Repositioning FDA-approved drugs in combination with epigenetic drugs to reprogram colon cancer epigenome. *Molecular Cancer Therapeutics*, *16*(2), 397–407.

Rechtsteiner, A., Ercan, S., Takasaki, T., Phippen, T. M., Egelhofer, T. A. *et al.* (2010). The histone H3K36 methyltransferase MES-4 acts epigenetically to transmit the memory of germline gene expression to progeny. *PLoS Genetics*, *6*(9).

Reverón-Gómez, N., González-Aguilera, C., Stewart-Morgan, K. R., Petryk, N., Flury, V. *et al.* (2018). Accurate recycling of parental histones reproduces the histone modification landscape during DNA replication. *Molecular Cell*, *72*, 1-11.

Saifee, O., Wei, L. & Nonet, M. L. (1998). The *Caenorhabditis elegans unc-64* locus encodes a syntaxin that interacts genetically with synaptobrevin. *Molecular Biology of the Cell*, *9*(6), 1235–1252.

Samson, M., Jow, M. M., Wong, C. C. L., Fitzpatrick, C., Aslanian, A. *et al.* (2014). The specification and global reprogramming of histone epigenetic marks during gamete formation and early embryo development in *C. elegans*. *PLoS Genetics*, *10*(10), 17–21.

Sankaran, S. M., Wilkinson, A. W., Elias, J. E. & Gozani, O. (2016). A PWWP domain of histone-lysine N-methyltransferase NSD2 binds to dimethylated lys-36 of histone H3 and regulates NSD2 function at chromatin. *Journal of Biological Chemistry*, *291*(16), 8465–8474.

Schmitges, F. W., Prusty, A. B., Faty, M., Stützer, A., Lingaraju, G. M. *et al.* (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. *Molecular Cell*, *42*(3), 330–341.

Siklenka, K., Erkek, S., Godmann, M., Lambrot, R., McGraw, S. *et al.* (2015). Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. *Science, 350*(6261), aab2006.

Simon, J., Chiang, A., Bender, W., Shimell, M. J. & O'Connor, M. (1993). Elements of the *Drosophila* bithorax complex that mediate repression by Polycomb group products. *Developmental Biology*, *158*(1), 131–144.

Skene, P. J., Henikoff, J. G. & Henikoff, S. (2018). Targeted in *situ* genome-wide profiling with high efficiency for low cell numbers. *Nature Protocols*, *13*(5), 1006–1019.

Struhl, G. (1981). A gene product required for correct initiation of segmental determination in *Drosophila. Nature*, *293*(5827), 36–41.

Sulston, J. E. & Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans. Developmental Biology*, *56*(1), 110–156.

Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans. Developmental Biology*. *100*(1), 64-119.

Tabuchi, T. M., Rechtsteiner, A., Jeffers, T. E., Egelhofer, T. A., Murphy, C. T. & Strome, S. (2018). *Caenorhabditis elegans* sperm carry a histone-based epigenetic memory of both spermatogenesis and oogenesis. *Nature Communications*, *9*(1), 4310.

Teperek, M., Simeone, A., Gaggioli, V., Miyamoto, K., Allen, G. E. *et al.* (2016). Sperm is epigenetically programmed to regulate gene transcription in embryos. *Genome Research*, *26*(8), 1034–1046.

Vågerö, D., Pinger, P. R., Aronsson, V. & van den Berg, G. J. (2018). Paternal grandfather's access to food predicts all-cause and cancer mortality in grandsons. *Nature Communications*, *9*(1), 5124.

Vermeulen, M., Eberl, H. C., Matarese, F., Marks, H., Denissov, S. *et al.* (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell*, *142*(6), 967–980.

Voigt, P., LeRoy, G., Drury, W. J., Zee, B. M., Son, J. *et al.* (2012). Asymmetrically modified nucleosomes. *Cell*, *151*(1), 181–193.

Wang, X., Paucek, R. D., Gooding, A. R., Brown, Z. Z., Ge, E. J. *et al.* (2017). Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. *Nature Structural and Molecular Biology*, *24*(12), 1028–1038.

Xiao, Y., Bedet, C., Robert, V. J. P., Simonet, T., Dunkelbarger, S. *et al.* (2011). *Caenorhabditis elegans* chromatin-associated proteins SET-2 and ASH-2 are differentially required for histone H3 Lys 4 methylation in embryos and adult germ cells. *Proceedings of the National Academy of Sciences*, *108*(20), 8305–8310.

Xu, L., Fong, Y. & Strome, S. (2001). The *Caenorhabditis elegans* maternal-effect sterile proteins, MES-2, MES-3, and MES-6, are associated in a complex in embryos. *Proceedings of the National Academy of Sciences*, *98*(9), 5061–5066.

Yamaguchi, K., Hada, M., Fukuda, Y., Inoue, E., Makino, Y. *et al.* (2018). Re-evaluating the localization of sperm-retained histones revealed the modification-dependent accumulation in specific genome regions. *Cell Reports*, *23*(13), 3920–3932.

Yoshida, K., Muratani, M., Araki, H., Miura, F., Suzuki, T. *et al.* (2018). Mapping of histone-binding sites in histone replacement-completed spermatozoa. *Nature Communications*, *9*(1), 3885.

Yu, C., Gan, H., Serra-Cardona, A., Zhang, L., Gan, S. *et al.* (2018). A mechanism for preventing asymmetric histone segregation onto replicating DNA strands. *Science*, *8849*(August), eaat8849.

Yuan, W., Wu, T., Fu, H., Dai, C., Wu, H. *et al.* (2012). Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. *Science*, *337*(6097), 971–975.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S. & Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *Journal of Biological Chemistry*, *286*(10), 7983–7989.