2011-07-21

Chromosome-Biased Binding and Function of C. elegans DRM Complex, and Its Role in Germline Sex-Silencing: A Dissertation

Tomoko M. Tabuchi
University of Massachusetts Medical School

CHROMOSOME-BIASED BINDING AND FUNCTION OF C.ELEGANS DRM COMPLEX, AND ITS ROLE IN GERMLINE SEX-SILENCING

A Dissertation Presented

By

Tomoko Tabuchi

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences, Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

July 21, 2011

Interdisciplinary Graduate Program

ROLES OF *C. ELEGANS* DRM COMPLEX IN CHROMOSOME-BIASED GENE REGULATION AND X-CHROMOSOME SILENCING

A Dissertation Presented By

Tomoko Tabuchi

The signatures of the Dissertation Defense Committee signifies completion and approval as to style and content of the Dissertation

Kirsten Hagstrom, Ph.D. Thesis Advisor

Craig Ceol, Ph.D., Member of Committee

Gary Ruvkun, Ph.D., Member of Committee

William Theurkauf, Ph.D., Member of Committee

Marian Walhout, Ph.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

Craig Peterson, Ph.D. Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the School

Anthony Carruthers, Ph.D.
Dean of the Graduate School of Biomedical Sciences

Interdisciplinary Graduate Program

July 21, 2011

Dedication

I dedicate this dissertation to three most influential people in my life.

My grandfather Seichi Tabuchi was an amazing leader and a charismatic figure in my life. He taught me the wonder of nature, scientific observations, and critical observation and thinking. His positive attitudes in pursuing curiosity and learning new things will always stay with me.

My mother Hiroko Tabuchi who provided me with unconditional love, strong guidance and lessons through her wisdom and strength.

My dear husband Trevor J. Morin. He has been my biggest supporter and contributor in my scientific and personal growth. His strength, leadership, open communication and endless humor make me grow and glow every day.

Acknowledgements

I am most grateful to my thesis advisor, Dr. Kirsten Hagstrom Ph.D. She possesses amazing strength and patience to nurture and mentor my scientific curiosity and thinking to bring me to where I am today to be Ph.D graduate. She has been always supportive with my ideas, and her door was always open for scientific discussion. I chose Hagstrom lab because I wanted to ask biological questions, and Kirsten provided me this opportunity. Kirsten, thank you for everything!

I thank my collaborator Dr. Marian Walhout, Ph.D. who has been my inspiring figure in Science. I respect her leadership and dedication. I also thank Dr. Bart Deplancke, Ph.D. for starting and sharing the LIN-54 project with me, and for many hours editing the LIN-54 manuscripts. I also would like to acknowledge other collaborators involved in the LIN-54 project: Dr. Robert Horvitz, Melissa M. Harrison, and Craig Ceol at MIT: Bioinformatician Inmaculada Barrasa, Naoki Osato, Lihua J. Zhu: Paul Furcinitti at the Digital Imaging Core Facility, and Phyllis Spatrick at the Genomics Core Facility.

I also thank my wonderful thesis committee members, Dr. Craig Ceol, Dr.Craig Peterson, Dr. Bill Theurkauf, Dr. Marian Walhout, and the outside committee member Dr. Gary Ruvkun for their valuable inputs during my committee meetings and thesis meetings.

Special thanks go to the members of the Hagstrom lab. I enjoyed the caring and fun family-like atmosphere in the lab, great scientific inputs and intense discussion. Their engorgements and supports were my major source of energy to go through the graduate school, and I grew tremendously as a person and as a scientist. Thank you, James, Lucy, Anna, and Karen.

I would like to acknowledge the people of Biotech IV, especially the administrative stuffs Danise, Donna, Sylvia, Cathy. I enjoyed wonderful parties, lunch-time discussion, fantasy football and great pot-lucks took place in Biotech IV.

I cannot thank enough for my Japanese and American family and friends for all the supports, love, and laughter. My Japanese family in Japan sent me long-distance dedications and love. My family in the U.S., David, Ellen, Danielle, Scott, Tara, Dana, Colin, Uncle Chris gave me lots of hags, love, and supports to go through my hardships (thank you for valuable “finger” talk!). I dearly care and love my family in the U.S. as if they are my own ☺

Finally, my deepest gratitude goes to my husband Trevor J. Morin. He makes me happy and laugh every day. He is my perfect substrate.

Abstract

DRM is a conserved transcription factor complex that includes E2F/DP and pRB family proteins and plays important roles in the cell cycle and cancer. Recent work has unveiled a new aspect of DRM function in regulating genes involved in development and differentiation. These studies, however, were performed with cultured cells and a genome-wide study involving intact organisms undergoing active proliferation and differentiation was lacking. Our goal was to extend the knowledge of the role of DRM in gene regulation through development and in multiple tissues. To accomplish this, we employed genomic approaches to determine genome-wide targets of DRM using the nematode *Caenorhabditis elegans* as a model system. In this dissertation, I focus on the DRM component LIN-54 since it was proposed to exhibit DNA-binding activity. First, we confirmed the DNA-binding activity of *C.elegans* LIN-54 *in vivo*, and showed it is essential to recruit the DRM complex to its target genes. Next, chromatin immunoprecipitation and gene expression profiling revealed that LIN-54 controls transcription of genes implicated in cell division, development and reproduction. This work identified an interesting contrast in DRM function in soma vs. germline: DRM promotes transcription of germline-specific genes in the germline, but prevents their ectopic expression in the soma. Furthermore, we discovered a novel characteristic of DRM, sex chromosome-biased binding and function. We demonstrated that *C. elegans* DRM preferentially binds autosomes, yet regulates X-chromosome silencing by counteracting the H3K36 histone methyltransferase MES-4. By using genomics, cytology, and genetics, we defined DRM as an important player in the regulation of germline X-chromosome gene expression, and addressed molecular mechanisms

behind the antagonistic interactions between DRM and MES-4. I present a model to explain the interplay of DRM and MES-4, and propose a novel function of DRM and MES-4 in maintaining proper chromosome gene expression dosage. This work extends our knowledge of the conserved roles of DRM in development, and provides a new view of differing DRM functions in soma versus germline. Furthermore, we defined a novel chromosome-specific aspect of DRM-mediated regulation.

Table of Contents

Chapter I

List of Tables

Chapter II

Chapter III

Chapter IV

List of Figures

Chapter I

Chapter II

Chapter III

List of Third Party Copyrighted Material

Chapter II was published in PloS Genetics in 2011.

Chromosome-Biased Binding and Gene Regulation by the Caenorhabditis elegans DRM Complex.Tabuchi TM, Deplancke B, Osato N, Zhu LJ, Barrasa MI, Harrison MM, Horvitz HR, Walhout AJ, Hagstrom KA. PLoS Genet. 2011 May;7(5):e1002074. Epub 2011 May 12.

List of multimedia objects or files

Chapter II

- Supplemental Table 1. LIN-54 ChIP peak locations, bound genes, GO terms of bound genes, and genes commonly bound between *C. elegans* and *D. melanogaster* or human (Excel file)
- Supplemental Table 2. LIN-54 responsive genes and their GO terms. *(Tab 1)* Genes with changed expression in *lin-54(n2990)* embryos (Excel file)

List of Third Party Copyrighted Material

Chapter II was published in PloS Genetics in 2011.

Chromosome-Biased Binding and Gene Regulation by the Caenorhabditis elegans DRM Complex.Tabuchi TM, Deplancke B, Osato N, Zhu LJ, Barrasa MI, Harrison MM, Horvitz HR, Walhout AJ, Hagstrom KA. PLoS Genet. 2011 May;7(5):e1002074. Epub 2011 May 12.

List of multimedia objects or files

Chapter II

- Supplemental Table 1. LIN-54 ChIP peak locations, bound genes, GO terms of bound genes, and genes commonly bound between *C. elegans* and *D. melanogaster* or human (Excel file)
- Supplemental Table 2. LIN-54 responsive genes and their GO terms. *(Tab 1)* Genes with changed expression in *lin-54(n2990)* embryos (Excel file)

Preface

References to publication that presents the work contained within the thesis:

Chromosome-Biased Binding and Gene Regulation by the Caenorhabditis elegans DRM Complex. Tabuchi TM, Deplancke B, Osato N, Zhu LJ, Barrasa MI, Harrison MM, Horvitz HR, Walhout AJ, Hagstrom KA. PLoS Genet. 2011 May;7(5):e1002074. Epub 2011 May 12.

Figures not generated by author:

Figure2-1B and C were created by Bart Deplancke (Yeast-one-hybrid assay, Chromatin immunoprecipitation)

Figure 2-2C was created by Bart Deplancke (Yeast-one-hybrid assay)

Figure 2-2C was created by Melissa Harrison (Immunoprecipitation)

Figure 2-3: ChIP was performed by Bart Deplancke, and the ChIP-chip data was analyzed by Imma Barrasa, and the motif was discovered by Naoki Osato

Figure 2-3E: Motifs excluding Motif 1 were obtained from other published works, and listed for comparison purpose: worm E2F (Kirienko and Fay, 2007), human dE2F2 (Georellet et al. 2006), and human CDE/CHR (Shmit et al.)

Figure 2-4: Microarray analysis was performed by Julie Zhu

Figure 2-5E: The occurrence of motifs was determined by Naoki Osato

Figure 2-5E, some data was used from Chi and Reinke 2006. to obtain commonly regulated genes

Figure 3-2 and 3, microarray analysis was performed by Julie Zhu

Chapter I: Introduction and Literature Survey

Gene regulation and DRM/DREAM complex

The proper spatiotemporal control of gene expression is essential for cell integrity and development of multi-cellular organisms. In particular, developmental gene programs need to be coordinately controlled along with cell cycle gene regulation to ensure proper differentiation and cell fates of actively proliferating cells. For instance, in actively proliferating cells, certain developmental genes need to be kept repressed, while cell cycles genes are transcribed in a cell cycle dependent manner. Such intricate gene regulation is largely orchestrated by transcription factors and chromatin-modifying proteins.

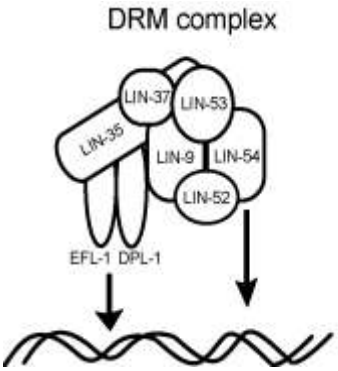
Some transcription factors (TFs) influence the expression of a wide range of genes. Such candidates were recently discovered in *Caenorhabditis elegans* using high-throughput yeast one-hybrid (Y1H) assays that revealed gut and neuron protein-DNA interaction (PDI) networks (Deplancke et al., 2006). Based on this PDI network, *Deplancke et.al.* has proposed a systems level model for *C.elegans* gene regulation by transcription factors. According to this model, transcription factors can be largely categorized into at least three tiers: “specifiers,” “master regulators”, and “global regulators” of transcription. Specifiers bind to one or two promoters within one system for fine-tuning of gene

expression; master regulators bind to multiple gene promoters and control expression of many genes within one system; global regulators bind to many promoters of genes involved in multiple systems. While most tested TFs fit into the “specifiers” or “master regulator” categories, only a few TFs act as “global regulators.”

One of these global regulators is *C.elegans* LIN-54 (JC8.6), which is conserved in *Drosophila melanogaster* (mip120) and *Homo sapiens* (hLin54), and its testis-specific orthologues are known as *tombola* and *tesmin*, respectively (Beall et al., 2002; Jiang et al., 2007; Korenjak et al., 2004; Litovchick et al., 2007; Schmit et al., 2007; Sugihara et al., 1999). Evidence suggesting LIN-54 is likely to bind DNA includes LIN-54 binding to *C. elegans* promoters in yeast one-hybrid (Y1H) assays (Deplancke et al., 2006), *D. melanogaster* Mip120(Lin54) binding specific sequence elements within the chorion gene cluster (Beall et al., 2002), and the purified tesmin domain from human Lin54 interacts with the human *cdc2* promoter *in vitro (Schmit et al., 2009)*. It is clear LIN-54 is a conserved transcription factor, which potentially binds and regulates the transcription of many genes. However, little is known about LIN-54 functions and its effect on transcription of candidate target genes, especially in the context of development.

Some transcription factors act together within a protein complex to confer DNA binding and gene regulation, and such is the case for LIN-54. LIN-54 is a member of the conserved multi-protein complex known as DREAM, which

includes a pRB-protein family member, and the heterodimeric DNA binding proteins E2F/DP. DREAM protein complexes similar in composition and functions have been identified in *C. elegans* (DRM) (Harrison et al., 2006), *D. melanogaster* (dREAM or Myb-MuvB) (Korenjak et al., 2004; Lewis et al., 2004) and *Homo sapiens* (hDREAM or LINC) (Litovchick et al., 2007; Schmit et al., 2007), and nearly all individual components of DREAM are conserved among these organisms (Figure1-1). However, the overall contribution of LIN-54 DNA binding to DREAM complex function has not yet been explored in any systems. Below I discuss the discovery of DREAM, and its known roles in coordinating expression of genes involved in cell division and development. In this dissertation, I will use the term DREAM complex to refer to the generic complex.



C.elegans DRM[1] complex	D.melanogaster dREAM(MMB)[2,3]	tMAC[4]	H.sapiens DREAM complex[5] (LINC)[6]
LIN-54	Mip120	Tomb	LIN54
LIN-9	Mip130	Aly	LIN9
LIN-35	RBF1, RBF2		RBL1/p107[9], RBL2/p130
LIN-37	Mip40	Mip40	LIN37
LIN-52	dLin-52[7]		LIN52
LIN-53	p55/Caf1	p55/Caf1	RBBP4/RbAp48
DPL-1	dDP		DP1[8],DP2[8]
EFL-1	dE2F2		E2F4, E2F5[8]
	Myb		MYBL1[8], MYBL2

Figure 1-1. The conserved DRM complex

Cartoon represents the eight-subunit *C. elegans* DRM complex**.** Table show DRM subunits and their homologs in the *D. melanogaster* dREAM/MMB complex and in the *H. sapiens* LINC/DREAM complex. *D. melanogaster* also has a paralogous tMAC complex that is testis-specific. A Myb subunit has not been identified in *C. elegans* DRM. 1. Harrison et al. 2006; 2. Korenjak et al. 2004; 3. Lewis et al. 2004 (MMB also includes Rpd3 and L(3)MBT); 4. Beall et al. 2007 (tMAC also includes Comr and Topi); 5. Litovchick et al. 2007; 6. Schmit et al. 2007; 7. Detected only in MMB; 8. Detected only in hDREAM; 9. Detected only in LINC; 10. Georlette et al. 2007.

Discovery of the conserved DREAM complex

The DREAM protein complex was first biochemically identified and purified in *D. melanogaster* as dREAM or MMB (Myb-MuvB) complex (Korenjak et al., 2004; Lewis et al., 2004). The recovery and subsequent analysis of proteins tightly bound to the *cis*-regulatory elements at the chorion gene cluster revealed a protein complex that consists of Myb and multiple Myb-interacting proteins (Mip),

such as mip-130, mip-40, mip-120, and caf1/p55 (Beall et al., 2002). Later, the additional constituents of dREAM (pRB protein-family RBF1/2, E2F2, dDP, dMyb) were identified by biochemical fractionation assays (Korenjak et al., 2004). Concurrently, using a proteomics approach, another group discovered a related complex, MMB (Lewis et al., 2004). MMB differs from dREAM in that it contains three additional members: LIN-52, the histone deacetylase RDP3, and L(3)MBT (Korenjak et al., 2004; Lewis et al., 2004). In addition, a study in testis found a testis-specific orthologus complex called tMAC that shares some components with dREAM/MMB (Beall et al., 2007; Jiang et al., 2007). Based on a genomic study, dREAM/MMB regulates the expression of genes primarily involved in cell cycle and development, but less is known about tMAC functions.

In *C.elegans*, the components of the DREAM complex (*C.elegans* DRM) were initially discovered through genetic screens for factors involved in vulva development (Ceol and Horvitz, 2001; Fay and Yochem, 2007; Ferguson and Horvitz, 1989; Thomas et al., 2003). Those factors are encoded by synMuvB (synthetic multivulva class B) genes, which act “synthetically” with synMuvA genes to antagonize Ras signaling during vulva development. Most synMuvB genes are broadly expressed chromatin-associated transcriptional regulators, and when mutated affect a range of biological processes including embryo polarity, apoptosis (Reddien et al., 2007; Schertel and Conradt, 2007), sex determination (Grote and Conradt, 2006), and RNA interference (Lehner et al., 2006; Wang et al., 2005). Subsequently, a biochemical study confirmed the

protein complex (DRM) consists of the synMuvB proteins LIN-35(Rb), EFL-1(E2F), DPL-1(DP), LIN-54(Mip120), LIN-9(Mip130), LIN-37, LIN-52, and LIN-53(Caf1) (Harrison et al., 2006). At present, the counterpart for the DNA-binding protein Myb is not apparent in *C.elegans* (Harrison et al., 2006). DRM exhibits pleiotropic functions in worms since the loss of DRM function leads to a wide range of phenotypes. However, the comprehensive genome-wide gene regulatory targets and function of synMuv proteins or DRM complex remain largely unknown.

DREAM complex is also identified in *Homo sapience* as hDREAM or LINC (LIN complex) (Litovchick et al., 2007; Schmit et al., 2007), and shows some unique features compared to fly or worm complexes. First, in contrast to those model organisms, the composition of hDREAM/LINC is rather complicated since pRB, DP, E2F, and Myb contain multiple family members that associate with hDREAM/LINC. For example, the human pRB-family is comprised of pRB, p130, and p107 proteins, yet only the last two proteins are components of hDREAM/LINC (Litovchick et al., 2007; Osterloh et al., 2007; Pilkinton et al., 2007; Schmit et al., 2007). Second, the concept of a stable core complex for DREAM is established in human. The core complex, termed “muvB core”, is composed of the homologues of synMuvB genes LIN-9, LIN-37, LIN-52, LIN-54, and RBB4 (CAF1) (Litovchick et al., 2007; Schmit et al., 2007), and “the muvB core” selectively interacts with two DNA-binding moieties of DREAM, the E2F/DP heterodimer or Myb, in a mutually exclusive manner (Litovchick et al., 2011;

Litovchick et al., 2007; Pilkinton et al., 2007; Schmit et al., 2007). This swapping of the DNA-binding proteins appears to be particularly helpful to control cell cycle genes in mammals and is discussed in the section below.

Cell cycle roles of the DREAM complex

The role of the DREAM complex in regulating cell cycle gene expression is best studied in mammalian cultured cells where DREAM alternately functions as a repressor or an activator complex depending on the cell cycle phase (Litovchick et al., 2007; Osterloh et al., 2007; Sandoval et al., 2009; Schmit et al., 2007). This switch in the transcriptional activity and targeting of a different set of genes is mediated by a subunit swapping mechanism. During the G0 phase of the cell cycle, the core of DREAM ("the muvB core") incorporates the Rb-family protein p130 and E2F, but not Myb, to repress S phase genes. In S-phase, p130 and E2F dissociate from the muvB core, and Myb is incorporated to promote activation of G2/M phase genes, such as the mitotic kinases cdc2 and Aurora kinase (Knight et al., 2009; Osterloh et al., 2007; Schmit et al., 2007). Consistently, the disruption of Myb or the muvB core leads to impaired proliferation and G2/M arrest (Knight et al., 2009; Osterloh et al., 2007).

Similar regulatory properties were also observed in *D. melanogaster*. A genomic study of dREAM demonstrated that multiple members of dREAM bind to many gene promoters to either activate or repress transcription, and importantly

E2F and Myb function in a mutually exclusive manner (Georlette et al., 2007). Notably, similar to human, many fly genes involved in G2/M transition were also activated by the muvB core complex associated with Myb, but excluding E2F. This suggests a conserved gene regulatory mechanism and a role for DREAM in positively regulating M phase genes in fly and human. On the contrary, *C.elegans* lack an apparent Myb homologue, implying a different mode of cell cycle gene regulation in *C.elegans*; however, it is possible that low sequence conservation of Myb may be hindering a homology search and the functional counter-part for Myb does exist, but is not yet recognized.

Role of DREAM complex in developmental gene regulation

Many animals carrying the loss or compromised function of DREAM show developmental abnormalities, which is thought to be an indirect consequence of cell division failures. For example, the homozygous Rb knock-out mice develop relatively normal till at least 11.5 days of gestation, but die with developmental defects, including abnormal haematopoietic and nervous systems (Clarke et al., 1992; Jacks et al., 1992; Lee et al., 1992). In *D. melanogaster* and *C.elegans*, compromising DREAM functions causes somatic and germline defects including sterility and lethality (Beall et al., 2004; Beitel et al., 2000; Ceol and Horvitz, 2001; Chi and Reinke, 2006; Ferguson and Horvitz, 1989; Harrison et al., 2006; Katzen et al., 1985; Manak et al., 2002; Thomas et al., 2003).

However, molecular studies over the past several years, especially with the model organism *D. melanogaster*, suggested that cell division failure is not solely responsible for such developmental abnormalities, but instead DREAM plays directly role regulating some genes involved in development and differentiation process. For example, microarray analysis revealed some of dREAM/MMB targets with sex- and development-specific expression in *D. melanogaster* cultured cells (Dimova et al., 2003; Korenjak et al., 2004; Lee et al., 2010). Importantly, these genes were permanently repressed by dREAM/MMB in actively proliferating cells, suggesting a cell cycle-independent role of dREAM in regulating developmental genes. It is not well understood how dREAM regulates two distinct transcriptional programs (cell cycle-dependent regulation of cell division genes vs. cell cycle-independent repression of developmental genes). Furthermore, it is not yet clear whether developmental gene regulation is a conserved DREAM function, since studies in mammalian cultured cells only identified cell cycle genes as DREAM targets. Moreover, with the exception of gene expression profiling of the *C. elegans* germline (Chi and Reinke, 2006), genome-scale studies of the DREAM complex were performed in cultured cells. Therefore, it was important to extend genome-wide analyses of DREAM to multiple cell types and tissues derived from intact organisms undergoing active developmental processes to enable a comprehensive assessment of DREAM function through development. In Chapter II, I describe how we analyzed genome-wide binding and gene regulation by *C. elegans* DRM

in embryonic soma and in germline tissue, leading to better understanding its gene targets and developmental roles. One new role of DRM that I discovered is its regulation of X chromosome silencing in the hermaphrodite germline. Therefore, below I provide an introduction about sex chromosomes and special forms of gene regulation that occur on sex chromosomes.

General introduction about sex chromosomes

Many organisms carry sex chromosomes that are used to determine sex. In the well-known XY system, the sex of mammals is determined by the presence of the Y chromosome, where animals carrying Y chromosomes become males and XX animals become females (Gubbay et al., 1990; Koopman et al., 1991; Sinclair et al., 1990). In flies and worms, the sex is determined by the ratio of autosomal and X-linked genes (X vs. A ratio). Thus, under normal conditions, animals carrying two X chromosomes develop into females (or hermaphrodites), while animals with a single X chromosome become males (XY or XO males) (Bridges, 1916; Madl and Herman, 1979). Therefore, having a different number of X chromosomes between the sexes appears to be inevitable for organisms that use chromosome-based sex determination.

Having a different number of X chromosomes between the sexes is linked with the fact that the X chromosome exhibits many distinct features compared to autosomes, such as its gene regulatory properties, associated proteins,

chromatin regulatory environments, gene content, and pattern of evolution (Ellegren and Parsch, 2007; Gurbich and Bachtrog, 2008; Vicoso and Charlesworth, 2006). For example, to compensate for the difference in X chromosome-linked gene dosages between the sexes, genes on X chromosomes are differentially regulated compared to autosomes by chromosome-wide gene silencing events termed dosage compensation (or gene activation in flies). Therefore, X chromosomes accumulate the different amount of gene regulatory proteins and modified histones compared to autosomes (Ellegren and Parsch, 2007; Gurbich and Bachtrog, 2008; Vibranovski et al., 2009), (Namekawa et al., 2006). As another example, X-specific hemizygosity in the male allows the accumulation of recessive alleles that are beneficial to male fitness. Thus, the genes located on the X chromosomes evolve faster and tend to be male-biased (Ellegren and Parsch, 2007; Gurbich and Bachtrog, 2008; Vicoso and Charlesworth, 2006). Furthermore, hemizygosity in the male also create “sexual antagonism”: the X chromosome is transmitted through females more frequently than males, and thus, likely to fixate mutations beneficial to female fitness (Ellegren and Parsch, 2007; Gurbich and Bachtrog, 2008; Vicoso and Charlesworth, 2006). In summary, X chromosomes are very distinct from autosomes in many ways, and such differences are made by the facts that the different number of X chromosomes exists between the sexes.

X chromosomes create an imbalance in chromosome dosage

The difference in the number of X chromosomes between the two sexes causes the problem of chromosome dosage imbalance between male and females, and between autosomes and the X chromosome. Between the two sexes, XX females (or XX hermaphrodites in *C.elegans*) can potentially produce twice as much X-linked gene products as XY males (or XO males in *C.elegans*). To correct for such potentially lethal imbalance, many organisms have developed various X-specific chromosome-wide gene regulatory mechanisms termed "dosage compensation" (Figure 1-2). For example, in humans one of the two X chromosomes in females is randomly selected for inactivation (Payer and Lee, 2008); in flies the single X chromosome in males is up-regulated roughly two-fold (Mendjan and Akhtar, 2007); in worms, the two X chromosomes in hermaphrodites are repressed by half (Ercan and Lieb, 2009; Meyer, 2010). It is not clear why these organisms have developed such diverse strategies, but each mechanism involves epigenetic changes on the X chromatin. For instance, X inactivation in human involves decorating the X chromosomes with heterochromatic histone modifications and DNA methylation partly by the action of non-coding RNA (e.g. Xist) (Payer and Lee, 2008). In flies, the active histone modification H4K16Ac is enriched on the male X chromosome (Mendjan and Akhtar, 2007). In worms, the two dosage compensated X chromosomes are depleted of the histone variant H2AZ (Petty et al., 2009). Thus, epigenetic

regulation seems to be a key method in correcting the imbalance in X-linked gene products between the sexes in many organisms.

The different number of X chromosomes between the two sexes also implies an imbalance in chromosome-dosage between autosomes and the X chromosome, similar to cases of aneuploidy. When considering active alleles on autosomes and the X chromosome, males contain a single active X chromosome (X) and two active autosomes (AA), creating an X:AA ratio of 0.5. Interestingly, lines of evidence suggest that this X:A imbalance is also corrected in some organisms. In *D. melanogaster*, its dosage compensation strategy (the up-regulation a single X chromosome in male) serves to correct not only the imbalances between the sexes, but also between X vs. A, making the X:AA ratio ~1.0 (Gupta et al., 2006). Furthermore even in human and mice, somatic tissues showed the X:AA ratio of ~1.0 (Gupta et al., 2006; Nguyen and Disteche, 2006). This observation suggests X-linked genes in mammals are up-regulated twice as high as autosomal genes per active allele. Although a contradictory result was reported recently (Xiong et al., 2010), the leading hypothesis is that the X:A ratio is actively maintained in many organisms.

Germline X chromosome silencing occurs in males of many organisms (MSCI)

The different number of X chromosomes between the two sexes causes yet another problem in germline, which is handled differently by a germline-specific

gene regulatory mechanism. In the germline where meiosis takes place, the homologous chromosomes need to be properly aligned and paired for faithful chromosome segregation and gamete production. However, the single X chromosome in male germlines lacks a complete meiotic partner, providing the potential to aberrantly recombine with other chromosomes and cause chromosome mis-segregation during meiosis. To prevent such detrimental consequences, unpaired DNA, including the single X chromosome in males, triggers a defense mechanism called MSCI (meiotic sex chromosome inactivation), and undergoes heterochromatin formation (Turner, 2007) (Figure 1-2). This is evident by the appearance of a condensed X chromosome called the sex-body (the XY-body), and the accumulation of histone modifications associated with inactive transcription (e.g. H3K9methylation) (Turner, 2007). In short, formation of such heterochromatin prevents it from recombining with other chromosomes, and thus ensures proper chromosome segregation. This evolutionarily conserved mechanism is thought to have evolved to silence unpaired DNA or chromosomes and it is widely used among many species, including human, fly, and worms.

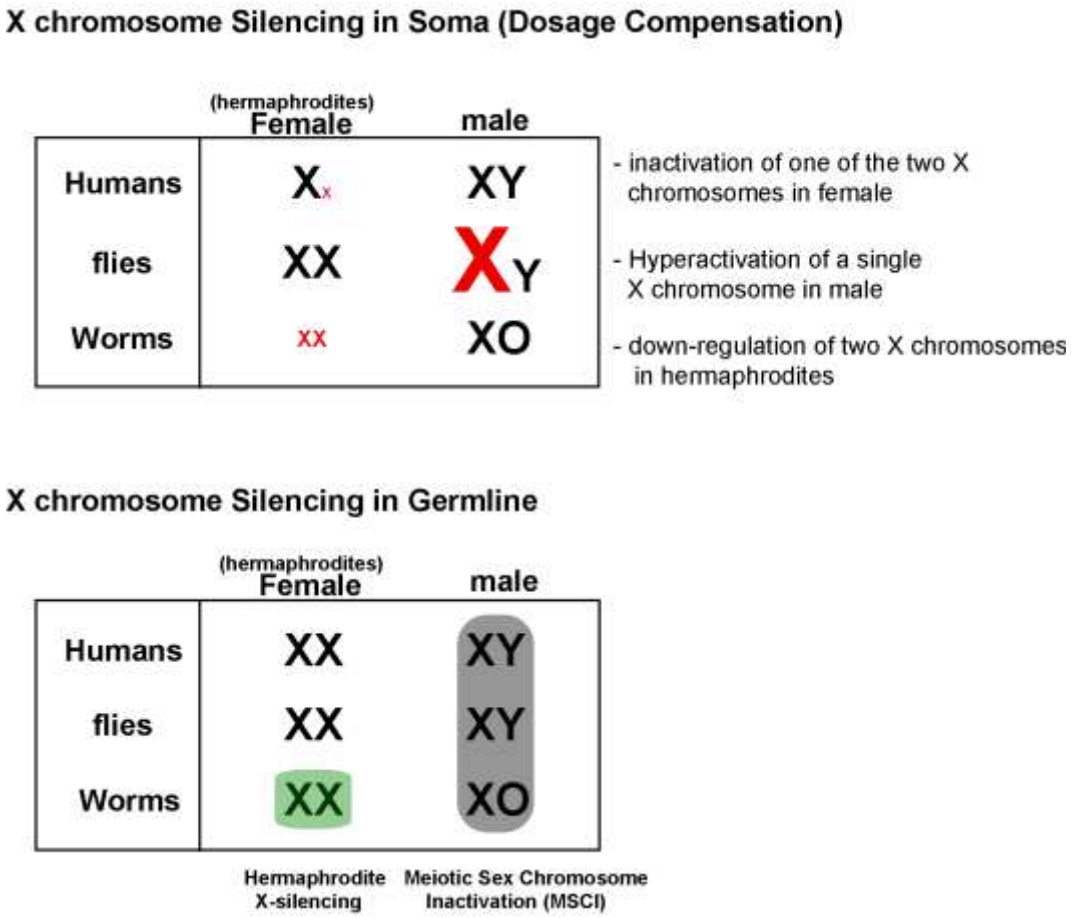


Figure 1-2. X chromosome Silencing in Soma and Germline

X chromosomes differ in number between the sexes, thus become target of X-specific chromosome-wide gene regulatory mechanisms.

In soma, a mechanism termed “dosage compensation” corrects the imbalance in X-linked gene products between the sexes. In humans, one of the two X chromosomes in females is randomly selected for inactivation; in flies the single X chromosome in males is up-regulated roughly two-fold; in worms, the two X chromosomes in hermaphrodites are repressed by half.

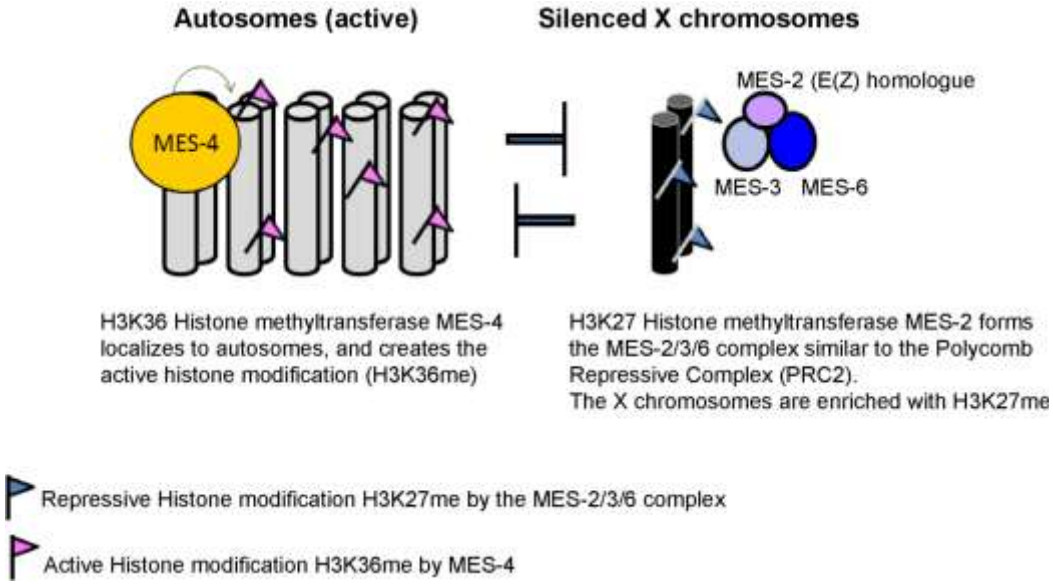
In the male germline of many species, the single X chromosome undergoes heterochromatin formation and is silenced by a process known as Meiotic Sex Chromosome Inactivation (MSCI). The two X chromosomes in *C.elegans* hermaphrodites undergo mild transcriptional silencing in germline.

Germline X chromosome silencing in *C.elegans* hermaphrodites

Another form of germline X chromosome silencing, presumably distinct from MSCI, is observed in *C.elegans* and its related nematode species (Kelly et al., 2002). The two X chromosomes in hermaphrodite germlines were shown to temporarily undergo mild transcriptional gene silencing (Figure 1-2). This was first discovered by cytological observations that the X chromosome in early stages of meiosis were depleted of an active form of RNA-polymerase II and histone modifications associated with active transcription (Kelly et al., 2002). Later, microarray analysis of hermaphrodite germlines confirmed that X-linked transcripts were significantly lower in number and in expression level, compared to the expected (Reinke et al., 2004; Reinke et al., 2000; Wang et al., 2009). It is not known if other organisms similarly repress the X chromosome in the female germ line, and why *C.elegans* represses two X chromosomes in the hermaphrodite germline. However, this is thought to be a MSCI-independent process. First, the two X chromosomes in hermaphrodite germlines properly pair and synapse (Villeneuve, 1994). Second, the hermaphrodite silencing only occurs during early stages of meiotic prophase, whereas MSCI-related silencing of the X persists until spermatogenesis (Kelly et al., 2002). Lastly, a distinct set of regulators are involved in each process: RNAi-related proteins play roles in MSCI-related heterochromatin formation (Maine et al., 2005; Walstrom et al., 2005), whereas the chromatin modifiers MES proteins are implicated in hermaphrodite X-silencing (Bender et al., 2006; Fong et al., 2002).

Figure 1-3

MES proteins: Regulator of Hermaphrodite X chromosome Silencing



MES proteins: The regulators of hermaphrodite X chromosome silencing

The proteins MES-2, MES-3, MES-4, and MES-6 mediate repression of X chromosome gene expression in the hermaphrodite germline. Genetic studies first indicated roles of *mes* ("maternal-effect sterile") genes (*mes-4, mes-2, mes-3, mes-6)* in modulating X-linked gene expression, by screening for mutations responsible for "grandchild-less" phenotypes (Capowski et al., 1991). Homozygous *mes* progeny from heterozygous mothers are fertile due to the maternal contribution (M+Z- generation: presence of maternal products but no zygotic expression), but their offspring become sterile adults carrying

degenerating germ cells (M-Z- generation: neither maternal nor zygotic products is present). Interestingly, the degree of germ cell degeneration in *mes* (M-Z-) mutants worsens with an increase in the number of X chromosomes (Garvin et al., 1998), implicating a role for *mes* genes in controlling expression of at least some X-linked genes. Cloning and molecular characterizations of *mes* genes and their products revealed that *mes* genes encode for two sets of histone methyltransferases that differentially modify chromatin on autosomes and the X chromosome: H3K36 methyltransferase MES-4 and H3K27methyltransferase MES-2, further supporting their roles in regulating genes on the X chromosome (Bender et al., 2004; Bender et al., 2006; Capowski et al., 1991; Fong et al., 2002).

MES-4 is the homologue of the NSD1 family of H3K36 methyltransferases. It plays a distinct role from a canonical H3K36 histone methyltransferases (i.e. worm MET-1 and yeast Set2) that create H3K36me marks on gene bodies to prevent sporadic initiation of transcription (Bender et al., 2006; Rechtsteiner et al., 2010). Instead, *C.elegans* MES-4 marks germline-expressed genes in embryonic soma, and thus is proposed to play an epigenetic role in transmitting the memory of a germ line program to the next generation (Rechtsteiner et al., 2010). In the germline, the MES-4 protein predominantly localizes to autosomes and methylates lysine 36 of Histone H3 (H3K36me2/3) (Bender et al., 2006). Despite its autosomal localization, MES-4 is implicated in hermaphrodite X chromosome silencing. Microarray analysis of *mes-4* (M+Z-) mutant germlines showed

prominent up-regulation of X-linked genes, suggesting that autosome-enriched MES-4 normally helps to repress genes on the X chromosome (Bender et al., 2006). It is possible that the roles that MES-4 plays in soma and germline may be connected; however, such evidence is currently lacking in the field.

Another regulator involved in hermaphrodite X-silencing is MES-2, a histone H3K27 methyltransferase homologous to fly Enhancer of zeste E(Z) (Bender et al., 2004; Holdeman et al., 1998). MES-2 forms a protein complex with MES-3 and MES-6 (the MES-2/3/6 complex), resembling the Polycomb Repressive Complex 2 (PRC2) (Xu et al., 2001a). MES-6 is a homologue of fly extra sex combs (ESC), and MES-3 is a protein with unknown function (Korf et al., 1998; Xu et al., 2001b). In germline and early embryos, MES-2/3/6 catalyzes di- or tri-methylation of lysine 27 of Histone H3 (H3K27me2/3), a mark typically associated with inactive transcription (Bender et al., 2004). H3K27me2 appears uniformly on all chromosomes, while H3K27me3 appears on autosomes to a lesser degree, with the majority concentrated on X. In the absence of MES-2/3/6 function, autosome-enriched MES-4 and its mark H3K36me2 were shown to mis-localize to the X chromosome, suggesting that H3K27me on the X chromosome normally keeps X-linked genes silent by repelling MES-4 (Fong et al., 2002). Taken together, MES-4 and MES-2/3/6 are thought to work together to keep the X chromosome repressed in the germline.

The effect of germline X chromosome silencing on genome organization

X chromosome silencing is intimately linked to genome organization. In many organisms, genes involved in spermatogenesis are under-represented on the X chromosomes, owing to the fact that if located on the heterochromatic male X, such genes would be silenced and lead to spermatogenesis defects and/or sterility (Gurbich and Bachtrog, 2008; Vicoso and Charlesworth, 2006). Thus, spermatogenesis genes are thought to have "fled" from the X chromosome during evolution. By contrast, in the *C.elegans* genome, not only spermatogenesis genes, but numerous germline-expressed genes (e.g. genes necessary for oogenesis and early embryogenesis) and essential genes are under-represented on the X chromosome (Kamath et al., 2003; Piano et al., 2000; Reinke et al., 2004; Reinke et al., 2000; Wang et al., 2009). Likewise, a study of duplicated genes found that a copy located on autosome was likely to be essential, while an X-linked copy was dispensable (Maciejowski et al., 2005). These findings are in agreement with the fact that *C.elegans* has two forms of X-silencing in the germline, male MSCI and hermaphrodite X-silencing. Therefore, evolutionary pressures imposed by germline X chromosome silencing in *C. elegans* are thought to have resulted in the autosome-biased location of germline-expressed and essential genes. However, autosome-biased transcriptional regulatory networks that orchestrate gene expression in the germline context had not been identified in any systems. In Chapter II, I show

that the DRM complex is an important regulator of germline gene expression and that the complex and the DNA binding motif it recognizes are autosome-biased.

The interplay between the DRM complex and MES proteins

The first section of this chapter described the discovery and known functions of the DREAM complex. In the second section we discussed roles of MES proteins in hermaphrodite X chromosome silencing events. In *C.elegans,* DREAM complex (DRM) and MES proteins are genetically linked in coordinating many aspects of worm development and cellular processes.

The process of vulva cell specification is one of the most well-characterized systems illustrating the interplay between the DRM complex and MES proteins. DRM members are synMuvB class genes, and lead to ectopic vulva formation when synMuvA genes are simultaneously mutated (the multiple-vulva, muv phenotype) (Fay and Yochem, 2007). Subsequently, a genetic screen for mutations that suppress muv phenotypes identified about ~40 genes as "synMuv suppressors" (Andersen et al., 2006; Cui et al., 2006b). Surprisingly, many synMuv suppressors encode for chromatin-acting proteins, such as an ATP-dependent chromatin remodeling complex, ISW1/NURF, or the histone acetylation complex, NuA3 complex, and, as discussed, histone methyltransferase MES-4 and MES-2/3/6 complex (Andersen et al., 2006; Cui et al., 2006b). In the case of vulva development, a key target of the synMuv pathway was shown to be the *lin-3/EGF* inductive signal in the hypodermis,

although the targets of synMuv pathway in other processes are largely unknown (Andersen et al., 2006; Cui et al., 2006a). In short, a genetically defined pathway involving vulva development provided initial insight that the proper specification of cell fates requires concerted actions of transcription factors and chromatin-acting proteins.

The antagonistic nature between synMuvB genes and the synMuv suppressors were also documented in other processes in *C.elegans*. For instance, the interplay between genes of the synMuvB and synMuv-suppressor class mediates proper distinction between soma and germline cell fates. Somatic cells in synMuv B mutants display germ cell-like features, such as ectopic expression of germline P-granules or enhanced RNAi efficiency (Lehner et al., 2006; Petrella et al., 2011; Unhavaithaya et al., 2002; Wang et al., 2005). Like vulva development, the loss of synMuv suppressors suppresses such transformation (Cui et al., 2006b; Unhavaithaya et al., 2002; Wang et al., 2005). This knowledge was recently further extended by *Petrella et al.* who performed expression profiling of a synMuv B mutant (DRM component *lin-35*/*rb*) and synMuv suppressor (MES-4) in soma (Petrella et al., 2011). They showed many of synMuv B mutants arrest at high temperature (26^{o}C) (high temperature arrest, HTA phenotypes) because somatic cells, particularly gut cells, ectopically express high levels of germline genes. The HTA phenotype and ectopically expressed germline genes were shown to be suppressed by mutating synMuv suppressor genes like MES-4. *Petrella et. al.* concluded that the DRM complex

and MES-4 counteract each other by acting on a similar set of germline-genes in somatic tissues. Therefore, it became increasingly clear that transcription factors alone are not sufficient to mediate proper differentiation process, but required the coordinated actions of chromatin factors. However, it remains unknown whether such interplay is necessary for germline processes. In particular, the link between the DRM complex and MES proteins in germline processes were not described. In Chapter III, I provide evidence that DRM counteracts the activity of MES proteins in regulating germline gene expression, and that these antagonistic activities are necessary for proper levels of X chromosome silencing.

Conclusions

The DREAM complex is an evolutionarily conserved protein complex that is primarily known to regulate cell cycle genes. Work over the past decade has started to reveal novel and direct roles in regulating genes involved in development and cellular differentiation. However, previous studies utilized cultured cells, and therefore a comprehensive genome-wide study involving intact organisms undergoing active development and proliferation processes was still missing. The studies presented in this thesis addressed this missing aspect of *C.elegans* DRM function, focusing on the two developmental contexts, soma vs. germline. Throughout this dissertation, my primary focus was placed on the *C.elegans* DRM complex component LIN-54, because LIN-54 was categorized as “global regulators” based on yeast protein-DNA interaction networks, yet its biological functions were less clear in *C.elegans*. In addition, LIN-54 seemed

likely to act as a DNA-binding protein within the DREAM complex; however, the contribution of LIN-54 DNA-binding activity to the DREAM function was not known. Furthermore, based on studies in fly and human, LIN-54 is a part of the stable core complex (the muvB core, see page 35), and hence, by studying LIN-54 we are likely to learn about the broad functions of the core complex. Lastly, we wished to map genome-wide DRM binding sites, and we reasoned our approach (e.g. chromatin immunoprecipitation, ChIP) would most likely succeed using LIN-54 since it likely directly binds DNA.

In chapter II of this thesis, I provide evidence that *C.elegans* LIN-54 has DNA-binding activity *in vitro* and *in vivo*, and that this DNA-binding activity contributes to the DNA binding of the whole DRM complex. Importantly, by utilizing genomic approaches, we addressed a key question: what are the roles of LIN-54(DRM) in development, especially in germline and soma? Through my research, I found that the DRM plays a distinct role in germline versus soma. DRM controls different targets genes in germline versus soma, and these are mainly involved in cell division, development, and reproduction. Moreover, LIN-54 promotes expression of reproduction genes in the germline, but prevents ectopic activation of germline-specific genes in embryonic soma. Strikingly, after mapping the transcriptional targets of LIN-54 and other DRM members, I discovered a unique and novel feature of *C.elegans* DRM regulation: chromosome bias. LIN-54(DRM) primarily binds and directly regulates genes located on autosomes. We hypothesize that such autosome-bias of DRM might

have co-evolved with the autosome-biased redistribution of germline-expressed genes.

During the analysis of genomic data in chapter II, we found that the loss of DRM led to the enhancement of X-silencing in germline, raising a possibility that DRM may be involved in hermaphrodite X-silencing. We test this hypothesis in chapter III by investigating potential interplay with MES proteins. Gene expression profiling showed that many of the mis-regulated genes in the *lin-54* or *mes-4* single mutants become expressed the at normal level in the *lin-54;mes-4* double mutant. We conclude that the interplay between LIN-54(DRM) and MES proteins is essential to maintain the proper dosage of X chromosome and autosome gene expression in the germline.

Chapter II.

Chromosome-biased binding and gene regulation by the *Caenorhabditis elegans* DRM complex

Chapter II. Abstract

DRM is a conserved transcription factor complex that includes E2F/DP and pRB family proteins and plays important roles in development and cancer. Here we describe new aspects of DRM binding and function revealed through genome-wide analyses of the *C. elegans* DRM subunit LIN-54. We show that LIN-54 DNA-binding activity recruits DRM to promoters enriched for adjacent putative E2F/DP and LIN-54 binding sites, suggesting that these two DNA-binding moieties together direct DRM to its target genes. Chromatin immunoprecipitation and gene expression profiling reveals conserved roles for DRM in regulating genes involved in cell division, development, and reproduction. We find that LIN-54 promotes expression of reproduction genes in the germline, but prevents ectopic activation of germline-specific genes in embryonic soma. Strikingly, *C. elegans* DRM does not act uniformly throughout the genome: the DRM recruitment motif, DRM binding, and DRM-regulated embryonic genes are all under-represented on the X chromosome. However, germline genes down-

regulated in *lin-54* mutants are over-represented on the X chromosome. We discuss models for how loss of autosome-bound DRM may enhance germline X chromosome silencing. We propose that autosome-enriched binding of DRM arose in *C. elegans* as a consequence of germline X chromosome silencing and the evolutionary redistribution of germline-expressed and essential target genes to autosomes. Sex chromosome gene regulation may thus have profound evolutionary effects on genome organization and transcriptional regulatory networks.

Chapter II. Introduction

The development of multi-cellular organisms is orchestrated by transcription factors that coordinate the spatiotemporal expression of sets of target genes. Transcription factors often act together in the context of multi-protein complexes. For instance, DREAM is a multi-protein complex conserved among *Caenorhabditis elegans* (DRM), *Drosophila melanogaster* (dREAM/Myb-MuvB) and *Homo sapiens* (hDREAM or LINC), and includes a retinoblastoma tumor suppressor pRb-family protein and the DNA binding heterodimer E2F/DP (Beall et al., 2002; Harrison et al., 2006; Korenjak et al., 2004; Lewis et al., 2004; Litovchick et al., 2007; Pilkinton et al., 2007; Schmit et al., 2007). DREAM coordinates the expression of cell division and differentiation genes during development, and its subunit activities are altered in many human tumors (van den Heuvel and Dyson, 2008).

In *C. elegans*, the genes that encode DRM subunits were originally identified in genetic screens for mutations causing defects in vulva development. Specifically, DRM subunits are encoded by synMuvB (synthetic multivulva class B) genes, which act “synthetically” with synMuvA genes to antagonize Ras signaling during vulva development (Ceol and Horvitz, 2001; Fay and Yochem, 2007; Ferguson and Horvitz, 1989; Thomas et al., 2003). Most synMuvB genes are broadly expressed chromatin-associated transcriptional regulators, and when mutated affect a range of biological processes including embryo polarity (Page et al., 2001), apoptosis (Reddien et al., 2007; Schertel and Conradt, 2007), sex determination (Grote and Conradt, 2006), and RNA interference (Lehner et al.,

2006; Wang et al., 2005). Despite their important roles in disparate developmental contexts, a genome-wide analysis of genes bound and regulated by synMuvB proteins is lacking.

Biochemical studies of *D. melanogaster* identified the dREAM/Myb-Muv-B complex and a partially overlapping testes-specific complex called tMAC (Beall et al., 2007; Beall et al., 2002; Jiang et al., 2007; Korenjak et al., 2004; Lewis et al., 2004). These complexes contain homologs of *C. elegans* synMuvB proteins. dREAM-like protein complexes were subsequently identified from *C. elegans* (DRM, (Harrison et al., 2006)) and human cells (hDREAM/LINC,(Litovchick et al., 2007; Schmit et al., 2007)). DRM includes LIN-35(Rb), EFL-1(E2F), DPL-1(DP), LIN-54(Mip120), LIN-9(Mip130), LIN-37, LIN-52, and LIN-53(Caf1). The human and fly complexes share these subunits and additionally contain a Myb subunit that is not apparent in *C. elegans* (Figure 1-1).

Several DREAM subunits contribute to its sequence-specific DNA binding, including E2F and DP, which together bind DNA as a heterodimer, and Myb. In flies and humans, E2F/DP and Myb act in a mutually exclusive manner to direct DREAM to its target genes (Georlette et al., 2007; Litovchick et al., 2007; Pilkinton et al., 2007; Schmit et al., 2007). Human DREAM is targeted to different sets of promoters by subunit switching (Litovchick et al., 2007; Pilkinton et al., 2007; Schmit et al., 2007). During the G0 phase of the cell cycle, the DREAM complex incorporates the Rb-family protein p130 and E2F4, but not Myb, to repress S phase genes. At cell cycle entry, p130 and E2F4 dissociate from the

complex, and Myb is incorporated to promote activation of M phase genes. LIN-54 is another DREAM component that has been reported to bind DNA: *D. melanogaster* Mip120(Lin54) binds specific sequence elements within the chorion gene cluster (Beall et al., 2002), *C. elegans* LIN-54 binds promoters in yeast one-hybrid (Y1H) assays (Deplancke et al., 2006), and human Lin54 interacts with the human *cdc2* promoter *in vitro* (Schmit et al., 2009). However, the overall contribution of LIN-54 DNA binding to DREAM complex function has not yet been explored.

Genome-wide binding and expression profiling studies of DREAM in mammalian cell culture primarily identified cell cycle genes as targets for the complex (Litovchick et al., 2007), while *D. melanogaster* cultured cell studies additionally revealed targets with sex- and development-specific expression (Dimova et al., 2003; Georlette et al., 2007; Korenjak et al., 2004). Thus, it is not clear whether developmental gene regulation is a conserved DRM function. With the exception of gene expression profiling of the *C. elegans* germline (Chi and Reinke, 2006), genome-scale studies of the DREAM complex were performed in cultured differentiated cells. It is important to extend genome-wide analyses of DREAM to multiple cell types and tissues derived from intact organisms, to enable assessment of DREAM function through development.

A key developmental function of *D. melanogaster* and *C. elegans* DRM subunits is the regulation of gene expression in the germline (Beall et al., 2007; Chi and Reinke, 2006; Jiang et al., 2007), which must occur within the context of

specialized germline gene expression features. The first such feature is a germline-specific form of X chromosome silencing. In male germlines of many species the single X is transcriptionally inactive and in *C. elegans* hermaphrodite germlines the two X chromosomes are partially silenced (Kelly and Aramayo, 2007; Kelly et al., 2002). Whether transcription factors like DREAM act equally on X-linked and autosomal genes, which exist in different chromatin regulatory environments, is not known. The second property special to germline-expressed genes is that they primarily reside on autosomes, possibly because of an evolutionary adaptation to X silencing (Kamath et al., 2003; Parsch, 2009; Piano et al., 2000; Reinke et al., 2004). It has not been explored whether the chromosome-biased location of germline differentiation genes is related to chromosome-biased binding sites and chromosome-biased regulation by distinct transcription regulatory networks.

Here we analyze genome-wide binding and function of *C. elegans* LIN-54. We demonstrate that LIN-54 DNA-binding activity is required for the DRM complex to efficiently bind and regulate target genes containing adjacent putative E2F/DP and LIN-54 binding sites (DRM binding motif). We show that LIN-54 binds to the promoters of genes involved in cell division, development, and reproduction, and acts differently in the germline versus the soma. The DRM binding motif, individual target genes, and overall DRM function are conserved among worms, flies, and humans. Despite this conservation, we discovered one striking feature of *C. elegans* DRM not shared in flies or humans: it is depleted

from X chromosomes. We show that DRM binding, the E2F-LIN-54 hybrid motif, and LIN-54-regulated genes are all autosome-enriched. One paradoxical exception occurs in the germline, where DRM binds autosomes but genes down-regulated in DRM mutants are enriched on X chromosomes. Evolutionary pressures imposed by germline X chromosome silencing in *C. elegans* are thought to have resulted in the autosome-biased location of germline-expressed and essential genes, major targets of DRM-mediated regulation. We propose that the autosome bias of *C. elegans* DRM co-evolved with the redistribution of its target genes. This example illustrates how sex chromosome gene regulation may create a biased genomic location of gene sets and their transcriptional regulatory networks.

Chapter II. Results

LIN-54 binds DNA through its tesmin domains

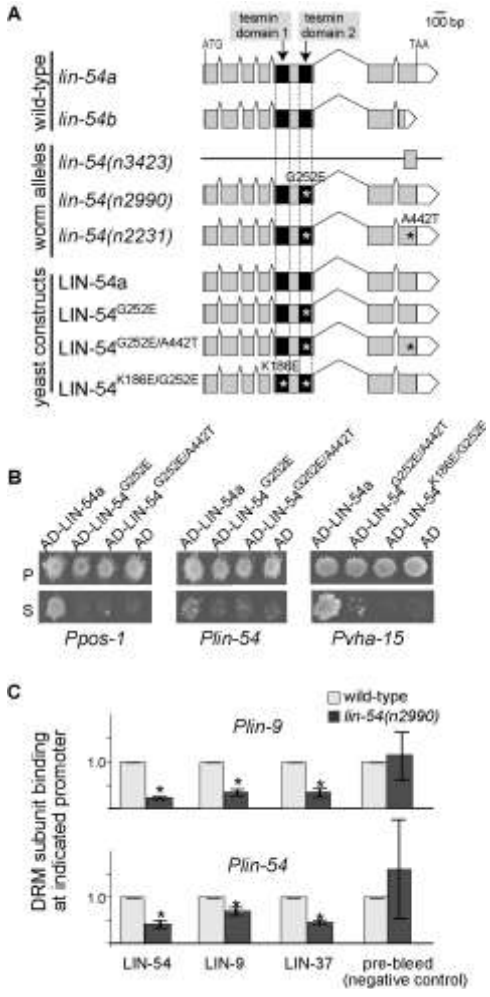
The *lin-54* gene encodes two proteins, LIN-54a and LIN-54b, both of which contain two tandem cysteine-rich repeats known as the tesmin/CXC domain (Figure 2-1A). Genetic screens for synMuv vulva development phenotypes identified the *lin-54(n2990)* and *lin-54(n2231)* missense alleles which confer similar loss-of-function phenotypes as a *lin-54(n3423)* deletion mutant (Harrison et al., 2006; Thomas et al., 2003). These missense alleles were independently

isolated and contain the same single-base substitution in the second tesmin domain (tesmin domain 2), which changes glycine 252 to a glutamic acid (G252E). The phenotypic effect of this mutation suggests that altering the tesmin domain compromises LIN-54 function and control experiments indicated that LIN-54 protein levels are normal in *lin-54(n2990)* mutant animals (see below). The *lin-54(n2231)* allele encodes a protein that contains an additional change in the C-terminus (A442T) (Figure 2-1A). We reasoned that these mutant alleles might result in loss of *lin-54* function because the corresponding protein fails to interact with other DRM complex components, because it fails to bind DNA, or because of a combination of these effects.

Previously, we found that LIN-54 can bind multiple *C. elegans* gene promoters in Y1H assays (Deplancke et al., 2006). To ask whether the tesmin domains mediate DNA binding, we tested wild-type LIN-54, and mutant versions of LIN-54 carrying lesions in a single tesmin domain (G252E and G252E/A442T), or lesions in both tesmin domains (K186E/G252E) in Y1H assays. We found that the mutant proteins exhibited much weaker DNA binding compared to the wild-type protein (Figure 2-1A and B). To examine the function of the tesmin domains in DNA binding *in vivo*, we performed chromatin immunoprecipitation (ChIP) experiments with wild-type and *lin-54(n2990)* mutant animals. Because we had noticed that LIN-54 binds its own promoter (Figure 2-1B), as well as promoters of genes encoding other DRM subunits (Supplemental Figure2-1A), we assayed binding at the *lin-9* and *lin-54* promoters. We observed a 4- and 2-fold decrease

in LIN-54 binding in the *lin-54(n2990)* mutant relative to wild-type animals at promoters of *lin-9* and *lin-54*, respectively (Figure 2-1C, Supplemental Figure 2-1B, p-value <0.01). Furthermore, the binding of other DRM complex proteins was also greatly reduced in *lin-54(n2990)* mutant animals (Figure 2-1C, p-value <0.01). These findings were supported by immunofluorescence analysis, which showed reduced chromosome localization of several DRM complex proteins in *lin-54(n2990)* mutant germlines (Supplemental Figure 2-1C). Control experiments showed that wild-type and *lin-54(n2990)* mutant animals produce a comparable amount of full-length, nuclear-localized LIN-54 protein (Figure 2-2A and B), unlike *lin-54(n3423)* null animals which produce no detectable LIN-54 protein and reduced amounts of other DRM subunits (Figure2-2B and (Harrison et al., 2006)). Together, these results indicate that LIN-54, in addition to EFL-1/DPL-1(E2F/DP), is a DNA binding protein involved in recruiting the DRM complex to its target genes.

Figure 2-1. LIN-54 binds DNA directly through its tesmin domains and recruits DRM to promoters.

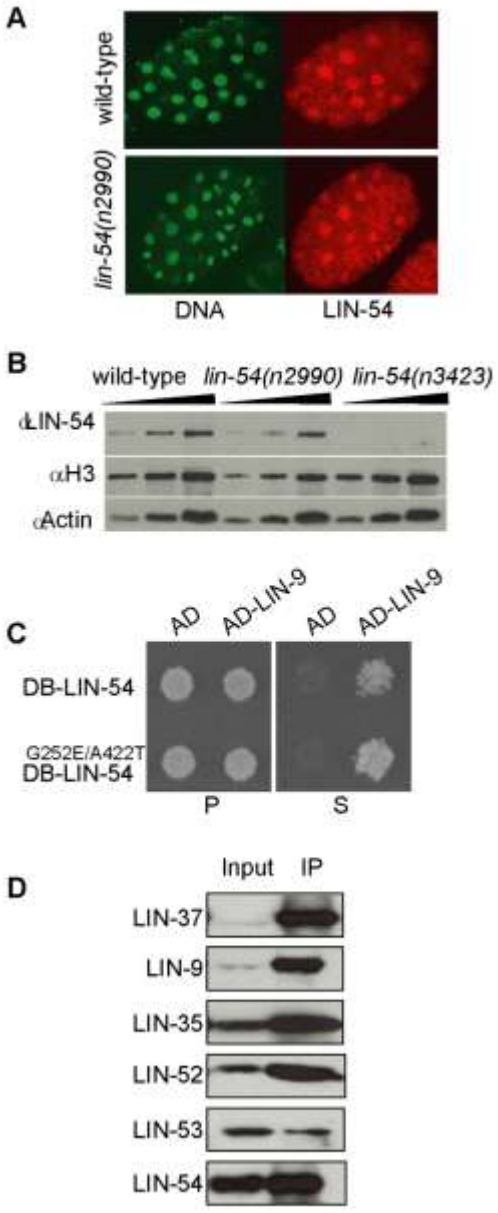


(A) C. elegans lin-54 gene structure for wild-type isoforms *(lin-54a* and *lin-54b), lin-54* mutant alleles, and yeast constructs used in this study. The *lin-54* gene encodes a protein with two tesmin/CXC domains (black boxes). *lin-54(n3423)* is a null allele in which the 5' end and most exons are deleted. *lin-54(n2990)* is a missense allele that harbors a mutation in the second tesmin domain, and *lin-54(n2231*) has both the tesmin domain mutation and an additional point mutation. Constructs equivalent to *lin-54a*, *lin-54(n2990)*, and *lin-54(n2231)* were used in yeast one-hybrid (Y1H) assays, and are referred to as LIN-54a, LIN-54$^{\mathbf{G252E}}$, and LIN-54 $^{\mathbf{G252E/A442T}}$, respectively. An additional LIN-54 construct containing a point mutation in each tesmin domain was created and is referred to as LIN-54$^{\mathbf{K186E/G252E}}$. Gray box = exon, black box = tesmin domain, white box = 3' untranslated region, asterisk = missense mutation. *(B)* Y1H assays using wild-type LIN-54a, LIN-54$^{\mathbf{G252E}}$, LIN-54$^{\mathbf{G252E\&A442T}}$, and LIN-54$^{\mathbf{K186E/G252E}}$ mutant proteins with the promoters of the genes *pos-1, lin-54*, and *vha-15*. AD = Gal4 activation domain, P = permissive media, S = selective media. *(C)* DRM subunit binding in wild-type and *lin-54(n2990)* mutants, measured by ChIP-qPCR at the target promoters *lin-9* and *lin-54.* Binding is shown as the amount of DNA amplified in each ChIP sample relative to input, with the ratio in wild-type set to 1.0. Standard deviations from three independent experiments are shown.

LIN-54 tesmin domain mutations do not disrupt DRM complex formation

We next tested whether LIN-54 tesmin mutations affect DRM complex formation in addition to compromising DNA binding. Using yeast two-hybrid assays, we found that both wild-type and mutant LIN-54 proteins can interact with the DRM subunit LIN-9 (Figure 2-2C). In addition, other DRM complex members co-precipitated in *lin-54(n2231)* mutant animals (Figure 2-2D). These observations demonstrate that the tesmin mutation does not result in an unstable protein and does not compromise the integrity of the DRM complex. We conclude that the *lin-54* tesmin mutant phenotypes are most likely caused by a defect in DNA binding.

Figure 2-2. LIN-54 tesmin domain mutation does not disrupt its stability or association with DRM.



(*A*) Immunofluorescence of LIN-54 in embryos from wild-type and *lin-54(n2990)* animals. (*B*) Western blots of whole worm extracts from wild-type, *lin-54*(*n2990*), and *lin-54(n3423)* mutants, probed with antibodies against LIN-54, histone H3, and actin. Lanes contain protein from 25, 50, and 100 worms. (*C*) Yeast two-hybrid assay using eithr wild-type LIN-54 (top) or mutant LIN-54$^{G252E/A442T}$ (bottom) as bait and LIN-9 as prey. DB = Gal4 DNA-binding domain. AD = Gal4 activation domain. P = permissive media, S = selective media. (*D*) Immunoprecipitation using antibodies against LIN-37 in *lin-54(n2231)* tesmin mutant extract, and probed with antibodies listed at left.

LIN-54 binds genes involved in development, reproduction, and cell division

We used ChIP-on-chip to identify genomic regions bound by LIN-54 in mixed-stage wild-type animals. Reproducible peaks of LIN-54 binding were detected in two biological replicas by the program MA2C (model-based analysis of two-color arrays, Figure 2-3A) (Song et al., 2007). Using the MA2C criteria described in Materials and Methods, we identified 1992 LIN-54 binding peaks (Supplemental Table 2-1). We used the mode of each peak as a measure for the location of LIN-54 association and found that 69% of the regions bound by LIN-54 occur within intergenic regions (Figure 2-3B). We next determined the relative position of intergenic LIN-54 peaks with respect to surrounding genes. We found that 60% of intergenic LIN-54 peaks occur within 1 kb upstream of protein-coding genes, and that the occurrence of a LIN-54 peak dramatically declined with distance from the translational start site (Figure 2-3B and C). When transcription factors bind between divergently transcribed genes it is difficult to determine whether they regulate one or both genes, so in these cases we considered the binding to be associated with both adjacent genes. Overall, LIN-54 bound to 1572 protein-coding gene promoters (Supplemental Table 2-1). These genes are highly enriched for three major gene ontology (GO) branches: developmental process (p-value $<10^{-100}$), reproduction (p-value $<10^{-100}$), and cell division (p-value $<10^{-30}$) (Supplemental Table 2-1). These results agree with and extend observations of

DREAM function in *Drosophila* and human tissue culture (Georlette et al., 2007; Litovchick et al., 2007) and suggest that DRM has conserved roles in development.

LIN-54 target genes are conserved through evolution

We discovered a significant degree of overlap among the individual genes bound by LIN-54 in worms, flies and humans (Figure 2-3D). The HomoloGene program has compared *D. melanogaster* and *C. elegans* genomes and defined a total of 3015 orthologous gene pairs (see Materials and Methods). Restricting our analysis to these defined fly-worm ortholog pairs, we note that 1267 are bound by LIN-54/Mip120 in flies (Georlette et al., 2007), 647 are bound by LIN-54 in worms (this study), and 327 are bound in both species (p-value $<10^{-6}$). Commonly bound genes are enriched for developmental GO terms such as sex differentiation as well as cell division terms such as cytokinesis and cell cycle (Supplemental Table 2-1). Commonly bound orthologs are involved in multiple aspects of cell division (*smc-3, zyg-9, air-2, plk-1, cye-1*), DNA replication and repair (*cdc-6, mcm-2, pri-1, mre-11, rad-51)* and transcription and chromatin regulation (*rbp-6, taf-4, mys-1, ash-2, mrg-1*). We also found significant overlap of genes bound by worm and human LIN-54: 62 orthologous gene pairs are bound in both species (p-value $<10^{-4}$, Figure 2-3D) (Litovchick et al., 2007). Further, in all three species, DREAM binds immediately upstream of 33 genes in

proximal gene promoters (this study; (Georlette et al., 2007; Litovchick et al., 2007)). Thus, LIN-54 targets the DREAM complex to genes involved in similar overall biological processes in three different phyla by binding to the proximal promoters of multiple orthologous genes.

In all three species DREAM bound the promoters of genes encoding its own subunits (Figure 2-1B and C, Figure 2-3A, Supplemental Table 2-1, Supplemental Figure 2-1A and B) (Georlette et al., 2007; Litovchick et al., 2007)). *C elegans* LIN-54 also bound the promoters of other synMuvB class genes, including LIN-61/L(3)MBT, LIN-15B, LIN-13, and LET-418 (Supplemental Table 2-1). This may suggest conserved transcriptional feedback between DRM subunits and perhaps other synMuvB class genes. However, genes encoding DREAM subunits show little change in expression upon LIN-54 depletion in *D. melanogaster* or *C. elegans* ((Georlette et al., 2007), Supplemental Table 2-2, data not shown). Perhaps the effects of DREAM autoregulation are small and required only to buffer DREAM levels and function.

DRM binding motif

We identified two DNA motifs that are over-represented in LIN-54-bound promoters in *C. elegans* (Figure 2-3E, Supplemental Figure 2-2). Motif 1 appears to be a hybrid E2F/DP and LIN-54 motif (Figure 2-3E, top, we will refer as “DRM binding motif”) and is usually found near the center of LIN-54 ChIP peaks (Figure

2-3F and Supplemental Figure2-2). The 5' end of this motif is similar to previously reported E2F/DP binding sites in *C. elegans* and other organisms ((Chi and Reinke, 2006; Kirienko and Fay, 2007; Muller and Engeland, 2010), http://jaspar.genereg.net). The 3’ end of Motif 1(DRM binding motif) resembles a *cis*-regulatory element in the human *cdc2* promoter (called CHR, or cell cycle homology region), which can be directly bound by hLin54 *in vitro* (Schmit et al., 2009). E2F/DP binding sites co-occur with CHRs in the promoters of some human genes, with a similar orientation and spacing as the motif we identified here ((Muller and Engeland, 2010), Figure 2-3E “human”). Moreover, a related motif was identified from *Drosophila* DREAM-regulated genes ((Georlette et al., 2007), Figure 2-3E “fly”). These results suggest conserved recruitment of the DREAM complex to its target genes by two DNA binding moieties: EFL-1/DPL-1 (E2F/DP) and LIN-54. LIN-54 bound promoters were also enriched for a periodic T-rich motif that resembles a related motif in *Drosophila* DREAM-bound genes (Motif 2, Supplemental Figure 2-2,(Georlette et al., 2007)). Other examples of periodic T-rich promoter motifs include sequences that function as nucleosome positioning signals (Segal and Widom, 2009) and elements with unknown function that are enriched in *C. elegans* germline-expressed promoters (Fire et al., 2006).

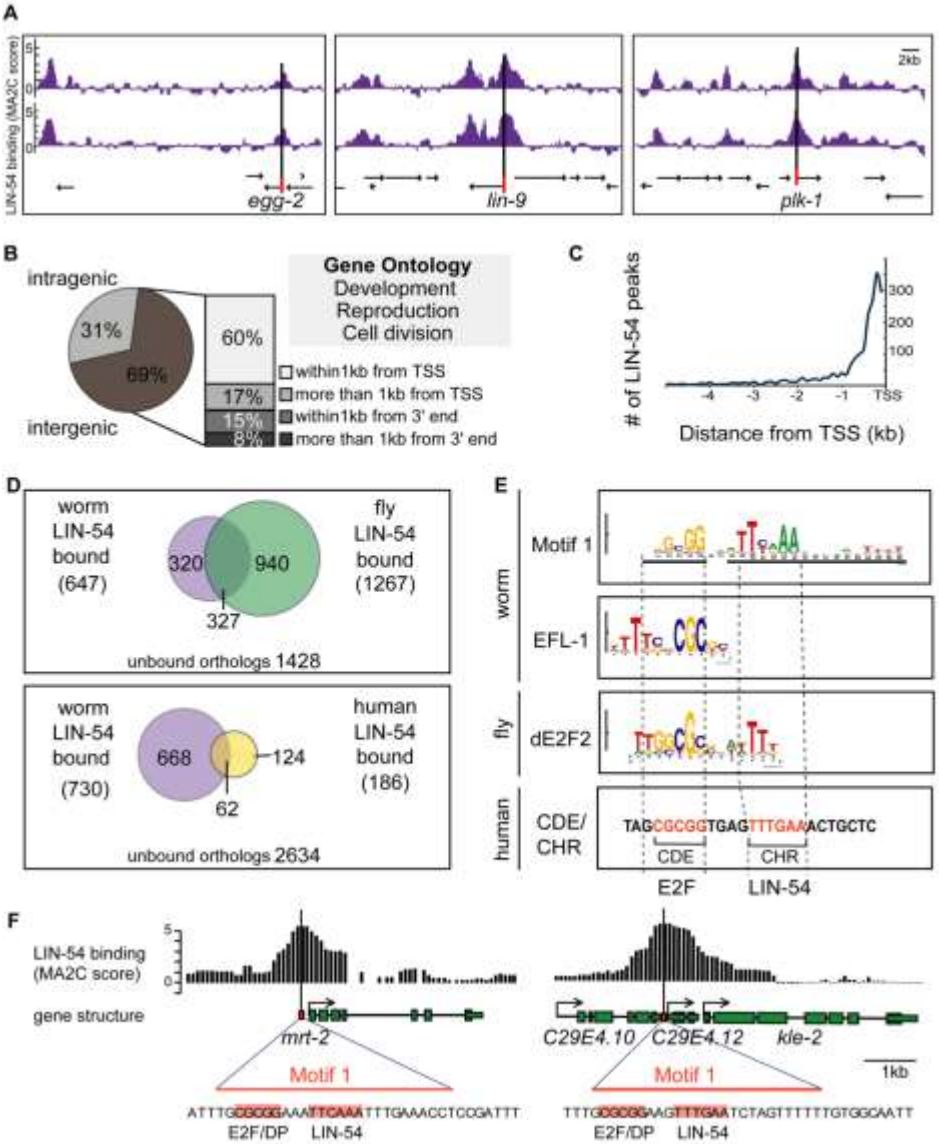


Figure 2-3. LIN-54 binding is enriched at promoters of genes involved in development, reproduction, and cell division that contain a putative E2F-LIN-54 binding motif.

Figure 2-3. LIN-54 binding is enriched at promoters of genes involved in development, reproduction, and cell division that contain a putative E2F-LIN-54 binding motif.

(A) Representative MA2C derived peaks from two biological replicates of LIN-54 ChIP-chip from mixed-stage worms. Arrows indicate genes and direction of transcription. *(B)* Relative locations of LIN-54 ChIP peaks. The distance between the mode of each LIN-54 ChIP peak and the translational start site (TSS) of neighboring genes was calculated, and the percentages of four classes of LIN-54 locations are indicated. Enriched gene ontology (GO) terms among genes with peaks within 1 kb of their TSS include development, reproduction, and cell cycle/cell division. *(C)* The numbers of intergenic LIN-54 peaks relative to their distance from the nearest TSS. *(D)* Conservation of orthologous LIN-54 binding targets between worms, flies, and humans. (*E)* An overrepresented motif in LIN-54-bound promoters (Motif 1, top). Aligned below are previously defined motifs: the *C. elegans* EFL-1 consensus (Kirienko and Fay, 2007), an extended *Drosophila* dE2F2 motif enriched among dE2F2, dLIN-9 and dLIN-54 co-regulated genes (Georlette et al., 2007) and the human CDE/CHR motif from the cdc2 promoter (Schmit et al., 2009). Dotted lines outline regions bound by human E2F4 and LIN-54 at cdc2 and their homologous motif sequences in other organisms. *(F)* Examples of LIN-54 binding (ChIP peaks shown by black bars representing MA2C score) and location of Motif 1 (orange square) at promoters of two genes (*mrt-2* and C29E4.12, arrows = TSS; green boxes = exons).

LIN-54 can activate or repress gene expression

Mutations in *lin-54* confer both germline and somatic abnormalities ((Harrison et al., 2006; Thomas et al., 2003), Supplemental Figure 2-3). To identify genes regulated by LIN-54 *in vivo*, we performed microarray expression profiling analysis of wild-type and *lin-54* mutant *C. elegans* embryos and of isolated germlines. We chose embryos because they consist primarily of somatic cells, at a developmental stage with both active cell divisions and dynamic developmental gene expression programs. Since *lin-54* null animals are sterile (Harrison et al., 2006), embryos were obtained from the *lin-54(n2990)* strain. *lin-54(n2990)* is a partial loss-of-function allele that causes the same spectrum of phenotypes as a null allele, albeit weaker, making it an appropriate strain in which to examine partial loss of *lin-54* function ((Harrison et al., 2006), Supplemental Figure. 2-3A). Germlines were dissected from *lin-54* null adults that lack detectable *lin-54* transcript and protein ((Harrison et al., 2006), Figure 2-2, and data not shown), exhibit reduced levels of other DRM complex proteins (Harrison et al., 2006), and exhibit reduced germline chromosome association of DRM complex proteins tested (Supplemental Figure 2-1C). We isolated the germline region from the tip until late pachytene stage of meiosis, because nuclei in this region are morphologically similar between wild-type and mutant (Supplemental Figure 2-

3B) and are undergoing X chromosome silencing (Kelly et al., 2002). While embryos contain a few primordial germ cells and dissected germlines contain some cells of the somatic gonad, the two samples predominantly represent somatic and germline tissue, respectively.

We identified 678 genes whose transcripts increased at least 1.5-fold in mutant embryos (Figure 2-4A, Supplemental Table 2-2). Of these, 119 (18%) were also bound by LIN-54 (Figure 2-4A). We note that ChIP was performed on mixed-stage animals to survey binding sites, while microarray was performed on a single stage, which may make it more difficult to identify all genes that are both bound and regulated. Nevertheless, this degree of overlap is similar to that observed in other ChIP and microarray studies (Georlette et al., 2007; Hu et al., 2007), and suggests that this gene set includes direct targets bound and regulated by LIN-54. GO analysis of up-regulated genes or of bound and up-regulated genes revealed over-represented terms related to development (Supplemental Table 2-2), terms that were also enriched among genes bound by LIN-54 (Supplemental Table 2-1). Fewer genes showed reduced expression in mutant embryos (299, Figure 2-4A). These genes showed no GO term overlap with LIN-54 bound genes, and only 2% (7/299) contained LIN-54 ChIP peaks at their promoters. This observation suggests that most of these genes are regulated indirectly. We conclude that LIN-54 predominantly functions as a transcriptional repressor in embryos (Figure 2-4B).

We noted that many up-regulated genes fell into discrete functional sub-categories related to development. Some of these gene sets might explain abnormalities of synMuvB mutant animals. For instance in *lin-54* mutant embryos, 18 up-regulated genes are involved in meiosis (GO term GO0001726) and overall, 11% of the up-regulated genes normally show germline-specific or enriched expression (Wang et al., 2009). Previously, mutations in synMuvB genes were shown to cause ectopic expression of certain germline P granule components in the soma, proposed to reflect soma to germline transformation (Unhavaithaya et al., 2002; Wang et al., 2005). Our genome-wide study strengthens this model by indicating that LIN-54 represses transcription of a variety of germline genes in embryo soma, including the P granule protein *glh-1*, the meiotic recombination protein *spo-11*, and the eggshell protein *cpg-2.* We also observed up-regulation of many RNA interference pathway genes in *lin-54* mutant embryos, including *ego-1, rde-4,* and *sago-2*. If these factors are normally limiting for a full RNAi response, their up-regulation might account for the enhanced RNAi phenotype that has been observed in synMuvB mutants (Lehner et al., 2006; Wang et al., 2005).

In the germline, 78 genes showed increased and 251 genes showed decreased expression in mutant relative to wild-type animals (Figure 2-4A, Supplemental Table 2-2). Both sets of genes exhibit overlap with LIN-54 ChIP peaks (18% and 12%, respectively) (Figure 2-4A). Further, both up-regulated and down-regulated germline genes are enriched for development GO terms, which

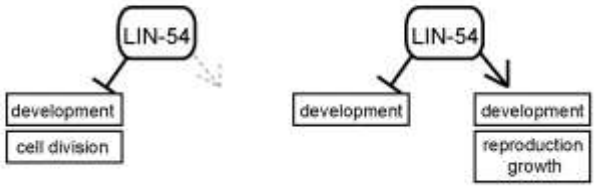
again overlaps with the terms found in the ChIP data (Figure 2-4B, Supplemental Table 2-2). These observations suggest that both up- and down-regulated germline genes could include targets directly regulated by LIN-54. While the development GO term is associated with both embryonic and germline LIN-54 target genes, reproduction and growth terms were only enriched in genes with decreased expression in the *lin-54* mutant germline. These reproduction genes that we presume are normally activated by LIN-54 include germline-produced transcripts required for meiosis, oogenesis and early embryogenesis, as observed previously for EFL-1/DPL-1 (Chi and Reinke, 2006). Thus in contrast to embryos, in the germline LIN-54 appears to both activate and repress gene expression, and activates a distinct set of reproduction and growth genes required for germline function.

Figure 2-4

A

	embryo		germline	
	up	down	up	down
# of LIN-54 responsive genes	678	299	78	251
% of genes bound	18 % (119/678)	2% (7/299)	18% (14/78)	12% (29/251)

B



LIN-54 can function as a transcriptional activator or repressor. *(A)* Microarray gene expression profiling analysis of *lin-54(n2990)* embryos and *lin-54(n3423)* germlines. Genes that change expression in *lin-54* mutant animals are grouped into four classes: “up in embryo”, “down in embryo”, “up in germline” and “down in germline”. Overlap with LIN-54 ChIP peaks is indicated. (*B*) Cartoon indicating the inferred regulation by wild-type LIN-54 in embryo (left) or germline (right) and the major Gene Ontology (GO) terms associated with each class of regulated genes. p-value <0.05 for all GO terms.

LIN-54 binding is under-represented on the X chromosome

We discovered a striking non-uniform distribution of LIN-54 binding across the *C. elegans* genome: X chromosomes had significantly fewer LIN-54 ChIP peaks than autosomes (p-value < 10^{-15}, Figure 2-5A). Each autosome had on average 369 LIN-54 ChIP peaks (23 peaks per Mb), whereas the X chromosome contained only 145 (8 peaks per Mb) (Figure 2-5B, Supplemental Table 2-3A). On average, 8% of autosomal gene promoters, but only 2% of X chromosome promoters, were bound by LIN-54 (Figure 2-5C, Supplemental Table 2-3A, p-value <10^{-41}). This analysis shows that LIN-54-bound promoters are significantly under-represented on the X chromosome, independent of chromosome size and gene density.

We also found that Motif 1 (DRM binding motif, Figure 2-3E), as well as the T-rich motif (Motif 2, Supplemental Figure 2-2A), were under-represented on X compared to autosome promoters (Figure 2-5D, Supplemental Figure 2-2B, Supplemental Table 2-3B, p-value < 10^{-13} for Motif 1). However, a published EFL-1 consensus site alone shows no bias against X chromosomes (Figure 2-5D, (Kirienko and Fay, 2007)). A uniform distribution was also observed for three additional transcription factors for which a consensus DNA binding motif has previously been determined (HLH-27, FLH-1, and NFI-1, Figure 2-5D) (Grove et al., 2009; Ow et al., 2008; Whittle et al., 2009). These results imply that the DRM

complex is recruited more frequently to autosomes than to the X chromosome through the combined DNA binding activities of LIN-54 and EFL-1.

lin-54 *mutants exhibit chromosome-biased gene expression changes*

We addressed whether the non-uniform binding of LIN-54 in the genome results in differential regulation of autosomal versus X-linked genes. LIN-54-responsive genes are distributed across all six *C. elegans* chromosomes (Supplemental Table 2-3), and we analyzed chromosome bias in two ways. First, to normalize for the variable number of genes on each chromosome, the percentage of LIN-54 responsive genes out of all genes per chromosome was calculated (Figure 2-5E). Second, to compare expected to observed distributions, we calculated the percent of all genes in the genome located on autosomes and compared that to the percent of LIN-54 responsive genes on autosomes (Figure 2-5F “all genes” versus “genes up in mut” or “genes down in mut”). Additionally, because the germline has an inherent autosomal bias in its expressed genes, we also calculated the percent of autosomal genes typically expressed in embryo or in germline as “expected” and compare that to the “observed” percent of LIN-54 responsive genes that reside on autosomes in each sample (Figure 5F “expressed genes” versus “genes up in mut” or “genes down in mut.”

Embryonic genes that were up-regulated in *lin-54* mutants are over-represented on autosomes (633/678, 93% observed versus 86% expected by chance, p-value $<10^{-8}$, Figure 2-5E and Figure 2-5F "embryo up"). This finding is consistent with the idea that LIN-54 is preferentially recruited to autosomes, and primarily acts as a repressor in the embryo. Embryonic genes down-regulated in *lin-54* mutants showed no significant chromosomal bias, consistent with our interpretation that these genes are mostly indirectly regulated (244/299, 82% versus 86% expected by chance, p-value = 0.03, Figure 2-5E and Figure 2-5F embryo down).

To our surprise, LIN-54 exhibited two different patterns of chromosome-biased gene regulation in the germline. Genes up-regulated in *lin-54* mutants were over-represented on autosomes, to a degree that is significantly different from all genes (77/78, 99% versus 86% expected by chance for all genes, p-value $<10^{-3}$, Figure 2-5E and Figure 2-5F), and comparable to the inherent bias of the germline (99% versus 93% expected by chance for germline-expressed genes, p-value = 0.06). This is consistent with the autosome-biased localization of LIN-54. LIN-54 is likely a direct repressor of at least some of these genes, since 18% overlap with LIN-54 ChIP peaks (Figure 2-5F). In striking contrast, germline genes that were down-regulated in *lin-54* mutants were located more frequently on the X chromosome than expected (64/251, 25% versus 14% expected by chance for all genes, p-value $<10^{-5}$, or versus 7% expected by

chance for all germline-expressed X-linked genes, p-value $<10^{-40}$, Figure 2-5E and Figure 2-5F, “germline down”).

It appears paradoxical that LIN-54 and its binding motif are preferentially located within autosomal gene promoters, yet in the absence of LIN-54 more genes on the X chromosome than on an average autosome decrease expression in the germline. One possibility is that LIN-54 affects these X-linked genes indirectly, which would predict less correlation between binding (ChIP peaks) and gene expression changes. Indeed, down-regulated X-linked genes overlap less frequently with LIN-54 ChIP peaks than down-regulated autosomal genes (6% versus 13% overlap, Figure 2-5E). Our interpretation of this observation is that LIN-54 is normally a direct activator of at least some autosomal genes that are down-regulated in the mutant, but that LIN-54 more indirectly regulates X-linked genes. Perhaps LIN-54 regulates an autosomal gene involved in X chromosome gene regulation, or prevents inappropriate spread of a repressor to the X chromosome (see Discussion).

Another apparent paradox is that LIN-54 loss leads to down-regulation of X-linked genes, when X chromosomes already undergo chromosome-wide silencing in the hermaphrodite germline. However, when we examined transcripts normally expressed in our wild-type germline samples using “present” calls from microarrays, we found that 15% of all X-linked genes are in fact expressed (376/2491 on array), consistent with published estimates from SAGE analysis (Materials and Methods, (Wang et al., 2009)). Of the 376 total germline-

expressed X-linked genes, 17% are down-regulated in the *lin-54* mutant (64/376) while only 4% of all germline-expressed autosomal genes are down-regulated (187/5097). The large percentage of total X-linked genes affected in the mutant may support models in which LIN-54 has chromosome-wide effects on X chromosome transcription (see Discussion). Thus on the X chromosome, the loss of LIN-54 function causes further silencing of X-linked genes.

Figure 2-5. LIN-54 shows autosome-enriched binding and chromosome-biased gene regulation.

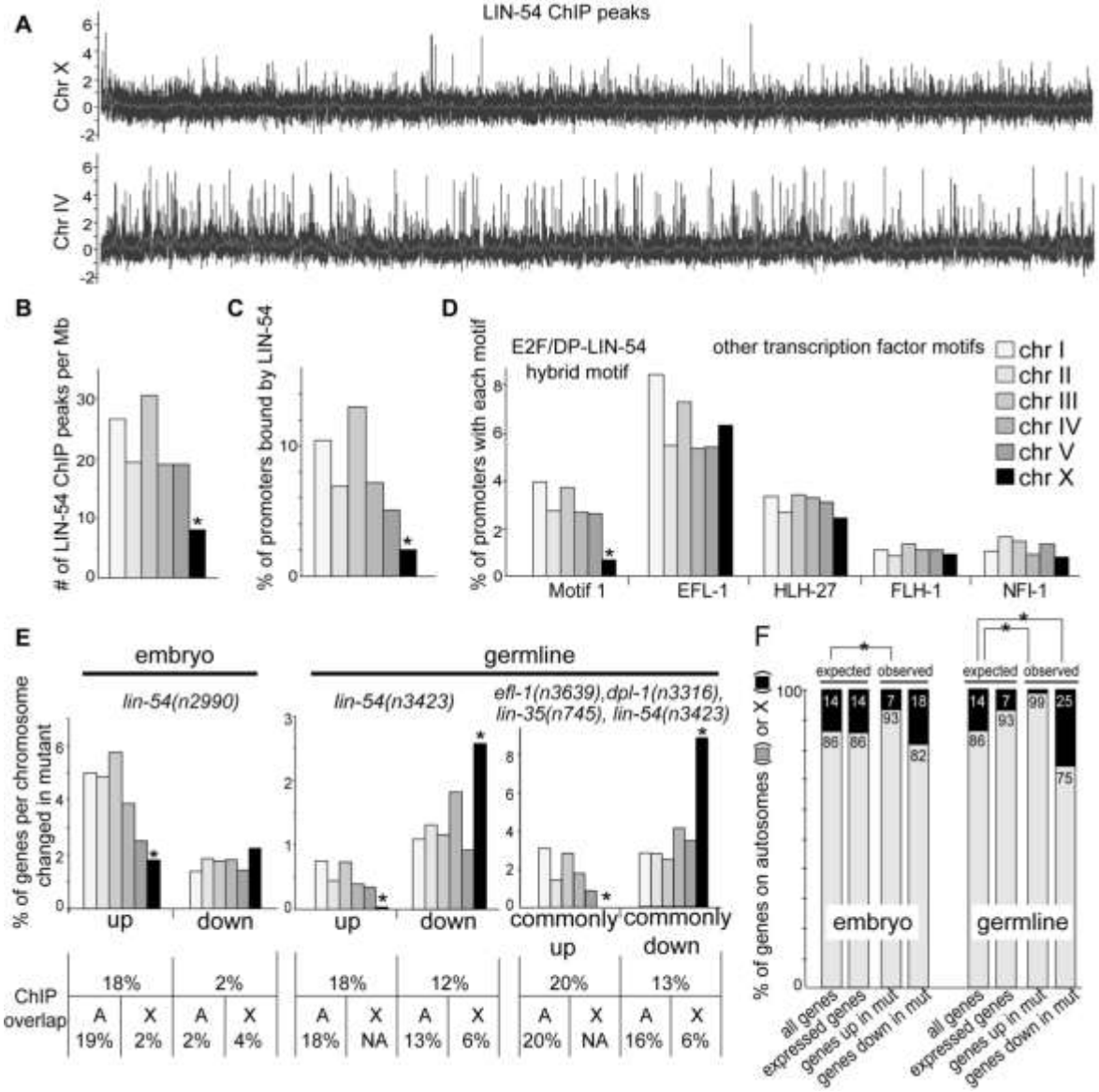


Figure 2-5. LIN-54 shows autosome-enriched binding and chromosome-biased gene regulation. *(A)* LIN-54 ChIP peaks along the entire X chromosome (top) and chromosome IV (bottom). *(B-C)* Number of LIN-54 ChIP peaks per mega base *(B)* and percentage of promoters bound by LIN-54 *(C)* on each *C. elegans* chromosome. LIN-54 ChIP peaks occur less frequently on the X chromosome, independent of chromosome size and gene density. *(D)* Occurrence of putative E2F/DP-LIN-54 binding Motif 1 and other transcription factor binding motifs in promoter regions (1kb upstream from translational start site) of autosomal genes and X-linked genes. Motif 1 is under-represented in X-linked promoters. *(E)* Chromosome distribution of genes up-regulated or down-regulated in *lin-54(n2990)* embryos (left), *lin-54(n3423)* germline (middle), or commonly co-regulated by cluster analysis of *lin-54(n3423), efl-1(n3639), dpl-1(n3316), and lin-35(n745)* germlines (right). Overlap with LIN-54 ChIP peaks for an average autosome or X chromosome is indicated below. Commonly up is group E, commonly down is group B from Supplemental Figure 2-4. Data for *efl-1, dpl-1,* and *lin-35* are from (Chi and Reinke, 2006). *(F)* The percentage of genes located on the five autosomes (gray) or the X chromosome (black). Expected values are presented both for all genes in the genome, and for all genes normally expressed (expressed genes) in embryo or germline, and compared to observed percentages of genes up-regulated (genes up in mut) or down-regulated (genes down in mut) in *lin-54* mutants. Asterisks indicate p-value $<10^{-3}$ by Fisher's Exact test or G-test.

The DRM complex preferentially localizes to germline autosomes

We wondered whether the chromosome-biased localization and function of LIN-54 are features shared by other members of the DRM complex. We first compared germline expression profiles of *lin-54(n3423)* with published germline expression profiles for *efl-1(n3639), dpl-1(n3316),* and *lin-35(n745)* mutant animals ((Chi and Reinke, 2006), Supplemental Figure 2-4). Genes commonly down-regulated in all four DRM mutants were more frequently located on X chromosomes than autosomes, consistent with observations in the *lin-54* mutant (Figure 2-5E, “commonly down” and Supplemental Figure 2-4, group B). Also consistent was the finding that commonly down-regulated X-linked genes overlapped less frequently with LIN-54 ChIP peaks than commonly down-regulated autosomal genes, again suggesting that more X-linked genes are regulated indirectly (6% versus 16% overlap, Figure 2-5E). Up-regulated genes common to all four mutants were more difficult to define. However, we did note that a commonly up-regulated group of genes primarily regulated in *lin-54(n3423)* (Figure 2-5E “commonly up”, Supplemental Figure 2-4 group E) and another cluster primarily up-regulated in *lin-35(n745)* (Supplemental Figure 2-4 group I) were each autosome-enriched, as observed for the *lin-54* mutant alone. These results show that similar patterns of chromosome-biased gene regulation are exhibited by multiple DRM subunits.

Next, we examined the chromosomal localization of DRM complex members in the germline by immunofluorescence. Figure 2-6 shows nuclei in the pachytene stage of meiotic prophase, when homologous chromosomes are paired and beginning to condense. LIN-54 (red) co-localized with DNA (green), with the exception of one prominent region (Figure 2-6A, arrowheads). We demonstrated that this region corresponds to the X chromosome in two different ways. First, LIN-54 colocalized with H4K12Ac (blue), a histone modification associated with actively transcribed regions, which is under-represented on the partially silenced X chromosome ((Kelly et al., 2002), Figure 2-6B). Second, LIN-54 did not co-localize with the H3K9me2-stained X chromosome in *him-8(e1489)* mutants (Figure 2-6C). In these mutants the X chromosomes do not pair during meiosis and therefore acquire this heterochromatic histone mark (Bean et al., 2004).

The DRM complex members LIN-9, LIN-35, LIN-37, LIN-52, and DPL-1 were also under-represented on the X chromosome in the germline (Figure 2-6D and E). Thus, most DRM complex members localize on autosomes. Only one DRM subunit was not autosome-enriched. The CAF1 homolog LIN-53, which participates in multiple complexes (Harrison et al., 2006), showed little localization to DNA during this stage of meiotic prophase (Figure 2-6E). It is interesting to note that despite the uniform genomic distribution of the EFL-1/DPL-1 motif, DPL-1 was enriched on the autosomes in the germline and co-localized with LIN-54 (Figure 2-6D). These results support the hypothesis derived

from our motif analysis that when EFL-1/DPL-1 and LIN-54 jointly bind Motif 1 (DRM binding motif), this complex disfavors the X chromosome. These results are also consistent with the finding that germline genes co-regulated by EFL-1/DPL-1 and LIN-54 share similar biases in chromosome location. We conclude that LIN-54 acts with other DRM complex members to govern chromosome-biased gene regulation in *C. elegans*.

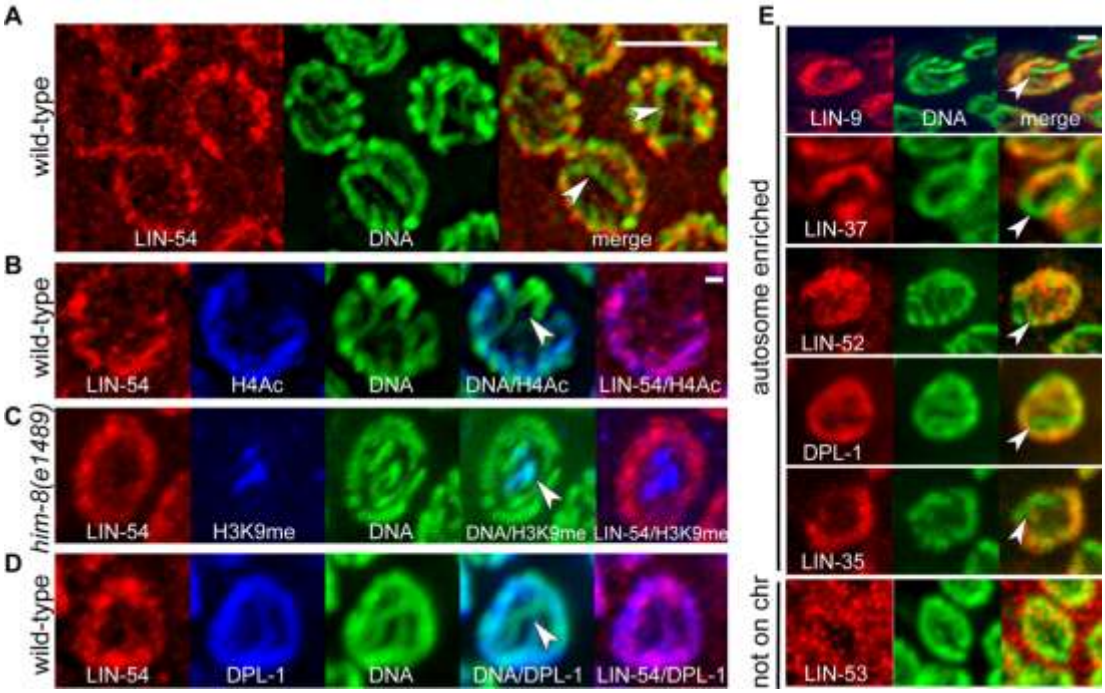


Figure 2-6. DRM complex members localize to germline autosomes. Shown are nuclei in the meiotic pachytene stage in the hermaphrodite germline. Arrowheads indicate a chromosome in the nucleus with less LIN-54 staining *(A-D)* or less staining of other DRM subunits *(E)*. *(A)* Immunofluorescence with anti-LIN-54 antibody (red) and DNA dye (green, merge in yellow). (*B*) Antibodies against a histone modification associated with active transcription (H4K12Ac, blue) show enrichment on autosomes, and co-localize with LIN-54 (red, DNA in green). (*C*) LIN-54 (red, DNA in green) staining in the *him-8(e1489)* mutant in which X chromosomes do not pair and acquire the histone modification H3K9me2 (blue). (*D*) Co-staining of LIN-54 (red) with DPL-1 (blue, DNA in green). Both are under-represented on the X chromosome (arrowhead). *(E)* Immunofluorescence of DRM complex subunits (red) on wild-type germline nuclei (DNA, green; merge yellow). Images in *A* and *B* represent deconvolved confocal stacks. Scale bar represents 5μm (A) or 1μm (B-E).

Chapter II. Discussion

Our genome-scale analyses of LIN-54 provide new insights into the binding and regulatory activities of the conserved transcription factor complex DRM. Our results in *C. elegans*, considered along with those available from *Drosophila* and human cells, highlight both conserved and non-conserved features of DRM. Conserved features include 1) DRM recruitment to promoters with a hybrid E2F/DP and LIN-54 binding motif (DRM binding motif), likely by the coordinated action of LIN-54 and E2F/DP, 2) its regulation of genes involved in cell cycle, development, and reproduction, and 3) its activity as both an activator or repressor. Through analysis of cells from a developing organism, we revealed conserved critical roles for DRM during animal development and showed that DRM activities vary in different tissues. Remarkably, we found that DRM binding and regulation are chromosome-biased in *C. elegans* but not *Drosophila* or humans, perhaps due to evolutionary pressures imposed by X chromosome silencing mechanisms.

Targeting the DREAM complex to promoters

Several members of the DREAM transcription factor complex have known or presumed DNA binding activity, but how they act in concert to direct promoter recognition was not well understood. Here we show that the DRM component

LIN-54 binds DNA directly, helps recruit DRM to promoters *in vivo*, and likely recognizes a hybrid E2F/DP and LIN-54 consensus motif (DRM binding motif). In *Drosophila* and humans, Myb is a DNA-binding component of the DREAM complex and it has been shown that Myb and E2F/DP function in a mutually exclusive manner (Georlette et al., 2007; Litovchick et al., 2007; Pilkinton et al., 2007; Schmit et al., 2007). We show that LIN-54 is another key DRM recruitment subunit and may function coordinately with E2F/DP: the E2F/DP and LIN-54 motifs co-occur in LIN-54 target genes and both components regulate a common set of genes. Our recognition that the *C. elegans* hybrid Motif 1, the CDE/CHR element of human cell cycle genes, and a motif identified in *Drosophila* DRM-bound genes are related elements suggests that coordinate binding by E2F/DP and LIN-54 is a conserved means of recruiting DRM to promoters (this study, (Georlette et al., 2007; Litovchick et al., 2007; Muller and Engeland, 2010)). It has been observed that the E2F binding motif is more widely distributed than E2F family protein binding in *vivo*, and E2F family members often rely on cooperating transcription factors bound to neighboring sites for specificity (Freedman et al., 2009). Simultaneous binding of adjacent sequence motifs by E2F/DP and LIN-54 might increase the affinity of DREAM for target sites and might provide increased selectivity for target gene recognition. Future studies will reveal if there is a Myb-like component in the *C. elegans* DRM complex, and whether other subunits contribute to DRM targeting to the genome.

Conservation of DREAM function

Genes bound and regulated by *C. elegans* LIN-54 predominantly function in development and differentiation, cell cycle and cell division, and in reproduction. Similar categories of regulated genes have been reported in genome-wide studies of *Drosophila* DREAM (Georlette et al., 2007). In human tissue culture cells, however, only cell cycle genes were enriched (Litovchick et al., 2007; Schmit et al., 2007). The similarities between *C. elegans* and *Drosophila* suggest broad conservation of DREAM function in both cell cycle and developmental gene regulation. Within the common GO term categories targeted by the DREAM complex, interesting functional subcategories were conserved. In all three organisms DREAM binds groups of genes involved in cell division processes such as sister chromatid cohesion, spindle assembly, and cytokinesis, as well as DNA replication and DNA repair. Both worm and fly DREAM bind and regulate genes involved in sex differentiation such as those required for genitalia formation, and genes required for germline functions including gametogenesis, fertilization, and meiosis. It seems likely that DREAM also regulates transcription of developmental and reproduction genes in mammalian systems, given known developmental roles of its individual subunits and the overall conservation of DREAM function. Perhaps developmental genes were not observed in mammalian studies because of the use of cultured cells derived from differentiated tissues. We find that similarities of DREAM function across species

lie not only at the overall level of biological processes: a remarkable degree of overlap exists among individual target genes. Further, the genes targeted by DREAM in all these organisms possess highly similar over-represented E2F/DP-LIN-54 motifs. Altogether, our results unveil an evolutionarily conserved mode of DNA binding that targets the DREAM complex to similar sets of functionally coherent target genes.

Different activities of DRM in the soma and germline

We demonstrate that DRM acts differently in the soma versus the germline. In embryos, LIN-54 appears to primarily repress genes (a majority of genes are up-regulated in the mutant, and up-regulated genes overlap with LIN-54 ChIP peaks and ChIP GO terms). In the germline, LIN-54 appears to primarily activate genes, yet may also serve as a repressor (a majority of genes are down-regulated in the mutant, and both up- and down-regulated genes overlap with ChIP peaks and ChIP GO terms). The target genes regulated in embryo versus germline are largely distinct, and fall into different enriched functional pathways (Figure 2-4, Supplemental Table 2-2). For example, in the germline LIN-54 promotes expression of genes required for germline functions like oogenesis, meiosis, and fertilization, as observed previously for EFL-1 and DPL-1 (Chi and Reinke, 2006). In the embryo, however, LIN-54 does just the opposite: it represses germline-specific genes to prevent their ectopic activation in the soma. Even patterns of

chromosome-biased gene regulation mediated by LIN-54 showed differences between soma and germline, as discussed below. Our results highlight how DRM may serve as either an activator or repressor. The mechanisms by which DRM may either activate or repress gene expression are at present not well understood, but may involve sub-complexes with different subunit composition or interactions with transcriptional co-factors such as chromatin modifiers. Importantly, our results provide the first genome-wide comparison of DRM function in two cell types isolated from whole animals, and indicate that DRM function differs depending on developmental context. Continued genome-wide analyses of DREAM binding and regulation in a variety of organisms, particularly using specific tissues isolated from animals, will further our understanding of how this key transcriptional complex functions during development and reproduction.

Why does C. elegans DRM avoid X chromosomes?

We discovered that *C. elegans* LIN-54 binding and gene regulation are autosome-enriched. This bias is likely a feature of the worm DRM complex as a whole, since the localization patterns of all but one DRM subunit are autosome-enriched, as are a class of germline genes co-regulated by multiple DRM subunits. Biased binding appears to be directed by a biased recruitment element, since the hybrid E2F/DP-LIN-54 recognition motif is also autosome-enriched in *C.*

elegans. However, when we examined the related hybrid motif in *Drosophila* (Figure 2-3E “fly”), and the published *Drosophila* and human DREAM ChIP profiles we found that they are evenly distributed between autosome and X chromosome promoters (data not shown, (Georlette et al., 2007; Litovchick et al., 2007)).

What evolutionary pressures might have driven the *C. elegans* DRM complex to disfavor the X chromosome? X chromosomes differ from autosomes in many aspects including histone variants and modifications, gene regulation, and rates of gene divergence and movement (Vicoso and Charlesworth, 2006). One possibility is that DRM targets are under-represented on the X chromosome because some aspect of this chromosomal environment is incompatible with DRM-mediated transcription regulation. A second possibility is that DRM localization and its differential regulation of autosomal and X-linked genes reflects some role in balancing autosome and X chromosome gene expression. Only a limited number of non-histone proteins have been shown to exhibit X chromosome- or autosome-biased localization, and these are involved in somatic dosage compensation or germline X chromosome silencing (Fong et al., 2002; Meyer, 2010; Takasaki et al., 2007). A third possibility is that the biased localization of DRM arose as a consequence of X chromosome silencing in the germline. The X chromosome is silenced in the germline by mechanisms that are distinct from somatic X chromosome silencing (Kelly and Aramayo, 2007). Germline-expressed genes and genes with essential functions are autosome

enriched, and thought to have “fled” the X chromosome to avoid being silenced (Kamath et al., 2003; Maciejowski et al., 2005; Piano et al., 2000; Reinke et al., 2004). One hypothesis is that the DNA-binding properties of the *C. elegans* DRM complex co-evolved with the redistribution of its germline-expressed and essential target genes across the genome, resulting in an autosomal bias. Silencing of the X chromosome has not been reported in *Drosophila* or mammalian female germlines, perhaps explaining why autosome bias is specific to *C. elegans* DRM. The regulation of sex chromosome gene expression, by processes that evolve rapidly and vary widely among organisms, may therefore have consequences on the genomic distribution of gene sets and, as shown here, their transcriptional regulatory networks.

A paradox in chromosome-biased gene regulation

In embryos, the biases in DRM localization and DRM-mediated regulation correspond, but in the germline they do not. In *lin-54* mutant embryos, up-regulated genes likely include direct targets based on their overlap with LIN-54 ChIP peaks, and were autosome-enriched like DRM binding. The down-regulated genes, on the other hand, are more likely indirect targets and showed no chromosome bias. In *lin-54* mutant germlines, both up- and down-regulated genes included direct DRM targets. As in embryos, the up-regulated genes in the

germline were primarily autosomal. Interestingly, down-regulated germline genes were X-enriched.

How can we explain the paradox that the DRM complex predominantly binds to autosomes, but that its loss results in a decrease in expression of X-linked genes? First, some LIN-54 does bind the X chromosome and might directly activate gene expression. However, fewer LIN-54-responsive genes on the X chromosome than on an average autosome are bound by LIN-54, suggesting that many X-linked genes are indirectly regulated. Second, loss of LIN-54 might induce ectopic soma-specific pathways that include X-linked genes. However, we found no evidence for enrichment of particular pathways among the affected X-linked genes and none are soma-specific. Other models invoke chromosome-wide alterations in X chromosome gene expression. A third model is that DRM regulates expression along the X chromosome indirectly either by activating a gene involved in X chromosome activation or by repressing a gene involved in X chromosome silencing, so that in mutants the X becomes more silenced. We did not find any obvious candidate for such a factor among mis-regulated genes. Finally, a fourth model proposes that a repressor that is normally concentrated on autosomes, perhaps anchored there by DRM, spreads inappropriately to X chromosomes when DRM function is compromised. If that repressor is limiting, autosomal genes will increase in expression while X-linked genes become repressed, which is in agreement with our observations (Figure 2-5E). Indeed, such reciprocal gene expression changes have been observed

when a limiting domain-specific repressor such as the *S. cerevisiae* SIR proteins spread inappropriately, thereby increasing repression at ectopic locations and diluting repression at their normal site of action (Taddei et al., 2009; Talbert and Henikoff, 2006; van Leeuwen and Gottschling, 2002). Related models have been invoked to explain why loss of the autosome bound MES-4 product de-silences germline X-linked genes and to explain why loss of the X chromosome bound Dosage Compensation Complex de-silences somatic X-linked genes and represses some autosomal genes in *C. elegans* (Bender et al., 2006; Jans et al., 2009).

Opposing actions of DRM and the histone methyltransferase MES-4

MES-4 is an autosome-enriched histone methyltransferase that confers the "active mark" H3K36me (Bender et al., 2006). In many biological contexts, *mes-4* and synMuvB genes have opposing functions. For example, mutations in *mes-4* can suppress the defects in vulva development, the increased RNAi and transgene silencing, and the ectopic expression of germline genes in the soma caused by mutations in synMuvB genes (Bender et al., 2006; Cui et al., 2006b; Kelly and Fire, 1998; Unhavaithaya et al., 2002; Wang et al., 2005). Here we define another process in which *mes-4* and synMuvB mutations have opposite effects. We show that in the hermaphrodite germline LIN-54 is autosome-

enriched as is MES-4, but *lin-54* mutants down-regulate while *mes-4* mutants up-regulate X-linked genes.

Bender et al. (2006) proposed that MES-4 indirectly regulates X-linked genes, by repelling a "global repressor" from autosomes and keeping it concentrated on the X chromosome. A possibility is that LIN-54 and MES-4 affect the X chromosome versus autosome distribution of the same repressor, in an opposite manner. A candidate for such a repressor is the *C. elegans* Polycomb Repressive Complex 2 (PRC2), which is composed of MES-2, MES-3 and MES-6. MES-2 is an E(z) homolog that concentrates the H3K27me3 "repressive mark" on the X chromosome in the germline (Bender et al., 2004; Bender et al., 2006). MES-2/-3/-6 also keeps MES-4 and other active marks restricted to autosomes. Interestingly, it was recently shown that a class of genes repressed by the *Drosophila* DREAM complex is enriched for H3K27me2 and requires E(z) for repression (Lee et al., 2010). However, the cytological distribution of H3K27me3 appears unaffected in *mes-4* and *lin-54* mutants ((Bender et al., 2006), data not shown). An important future direction is to explore potential links between DRM, MES-4, and Polycomb Group mediated gene repression, and to shed light on how these factors might interact to govern gene regulation.

Chapter II. Materials and Methods

C. elegans strains and culture conditions

All strains were cultured at 20^{o}C unless otherwise noted, using standard methods. The following strains were used: N2 (Bristol) as wild-type, *lin-54(n3423)/nT1 [qIS51], lin-54(n2990), lin-54(n2231)* (Harrison et al., 2006; Thomas et al., 2003), and *him-8(e1489)* (Hodgkin et al., 1979). Note: Previously, *lin-54(n2231*) was reported to have a single mutation (A442T) (Harrison et al., 2006); however, sequencing revealed an additional missense mutation (G252E).

Immunofluorescence

Embryos (Figure 2-2) were fixed with methanol/acetone (Strome and Wood, 1983). Germlines (Figure 2-6 and Supplemental Figure 2-1C) were fixed essentially as described (Shaham, 2006), with the addition of 5 ul of 2% Triton-X before fixation in 4% paraformaldehyde. DNA was visualized either with DAPI or OllieGreen (added at 1:1000 with 10ug/ml RNAseA with the secondary antibody). Whole worms (Supplemental Figure 2-3) were prepared in Carnoy's fixative as described by (Csankovszki et al., 2009). Primary antibodies to DRM subunits were described and validated in (Ceol and Horvitz, 2001; Harrison et al., 2006; Page et al., 2001). Another second anti-LIN-54 antibody was generated in rabbits against amino acids 207-306 (Strategic Diagnostics Inc.), validated by western

blot in wild-type and mutants, and showed the same localization patterns. Primary antibodies were used at 1:100 dilutions, and detected with secondary antibodies conjugated to Alexa Fluor 568 (Invitrogen) at a 1:500 dilution, except DPL-1 was performed as described [4,10]. Antibodies against H4K12Ac (Serotec), and H3K9me2 (Cell Signaling) were used at 1:1000 (primary) and seconday antibodies at 1:1000. Images for Figure 2-6 were captured by a Solamere Technology Group modified Yokogawa CSU10 Spinning Disk Confocal scan head attached to a Nikon TE-2000E2 inverted microscope and a 100x Plan Apo objective, using MetaMorph software (Molecular Devices). The images for Figure 2-6A and B were deconvolved using the constrained iterative deconvolution algorithm developed by the UMass Medical School Biomedical Imaging Group (Carrington et al., 1995).

Yeast one-hybrid and two-hybrid assays

Y1H and Y2H assays were performed as described (Deplancke et al., 2006; Walhout et al., 2000). Representative images for Figure 2-1B were obtained for *Ppos-1* at 10mM 3AT 5 days, *Plin-54* at 20mM 3AT 9 days, and *Pvha-15* at 60mM 3AT 9days.

Western blot, immunoprecipitation, and chromatin immunoprecipitation

For western blot (Figure 2-2B), whole worm lysates were created from 200 hand-picked synchronized young adults boiled in 2x loading buffer (National Diagnostics EC-886) for 30' with intermittent vortexing. Lysates equivalent to 25, 50, and 100 animals were loaded per lane and probed with anti-LIN-54, actin (Abcam #ab3280, 1:400) and Histone H3 (Abcam #ab1791, 1:1000). Immunoprecipitation, western blotting, and probing with DRM antibodies were performed as described (Harrison et al., 2006), ChIP was performed as described (Mukhopadhyay et al., 2008). Briefly, mixed stage wild-type worms were cultured in S-basal at 20^{o}C. Lysates were cross-linked in 1% formaldehyde, sonicated, and immunoprecipitated with anti-LIN-54 antibody or pre-bleed antibody control. ChIP samples including the input were subjected to two rounds of linear amplification, using the genomePlex complete whole genome amplification kit (Sigma), and minimum difference between original precipitates and amplified precipitate confirmed by qPCR (data not shown). Both experimental and input were processed at NimbleGen, hybridized on 385K *C. elegans* Whole Genome 3-Array Set (Roche NimbleGen). To assay DRM subunit binding at the promoters of the *lin-9* and *lin-54* genes, ChIP was performed with antibodies against LIN-54, LIN-9, LIN-37, or pre-bleed control from wild-type or *lin-54(n2990)* mixed-stage extracts. qPCR was used to calculate the amount of *lin-54* or *lin-9* promoter DNA in ChIP samples relative to the total input DNA. The ratio in wild-type was set at 1.0. *lin-9* promoter primers: 5'-cgactgtcaaacagcagctc-

3’ and 5’-ttgaaatggcggttcttttc-3’. *lin-54* promoter primers: 5’-atgatgagtgacgtctacc-3’ and 5’-attgtttcgcgcgccgaaatttg-3’.

RNA isolation and microarray

Embryo: Animals were propagated on egg plates seeded with *E. coli* HB101 at 20^{o}C, bleached to obtain synchronized L1 larvae and then grown at 25^{o}C for 48 hours. Embryos from young adults were harvested by the bleach-alkaline method, and filtered through 100 micron mesh (Small Parts, Inc.). 200 µL of embryo pellet was suspended in 1 mL of Tri reagent (Molecular Research Center, Inc. TR118), flash-frozen, and dounced. Total RNA was purified with RNAeasy mini kits (Qiagen), treated with DNase, and integrity examined on agarose gel.

Germline: Animals were grown at 20^{o}C and dissected in 1X egg buffer to excise the germline 24 hours after L4 stage. Germlines were dissected to include mitotic tip through meiotic late pachytene (Supplemental Figure 2-3). RNA was isolated as described (Chi and Reinke, 2006), and linearly amplified once using MessageAmp II aRNA Amplification Kit (Ambion).

Microarray: Probe-preparation, hybridization, and scanning for DNA microarray were performed at the Genomics Core facility at University of Massachusetts Medical School. Fluorescence-labeled cDNA probes were prepared using the One-Cycle kit (Affymetrix) and the Enzo HighYield RNA Transcript Labeling Kit (Enzo) for embryo, and the 3' IVT Express Kit (Affymetrix) for germline. cDNA

probes of three replicates were hybridized to GeneChip *C. elegans* genome arrays (Affymetrix).

Chip peak analysis

Raw ChIP-chip data were analyzed using three independent programs: MA2C (Song et al., 2007), ChIPOTle (Buck et al., 2005) and NimbleScan (Roche NimbleGen). While ChIPOTle called fewer and NimbleScan called greater numbers of peaks than MA2C, each identified a similar set of core peaks. MA2C analysis was performed with the following settings: # MA2C Score Method (median), Band Width (300), p-value cut off (-6), and other parameters were set as default. WS180 was used to annotate gene names. LIN-54 ChIP peaks (Figures 2-3 and 2-5A) were visualized using Affymetrix Integrated Genome browser. Modes of LIN-54 peaks were used to determine peak location for Figure 3, and each intergenic peak was considered to associate with both neighboring genes.

Ortholog pair analysis

HomoloGene (Ce.01-08-2009) defines 3015 orthologous pairs between *C. elegans* and *D. melanogaster*, and 3488 pairs between *C. elegans* and human. 647 of 1572 genes bound by *C. elegans* LIN-54 have annotated fly orthologs; 730 genes have annotated human orthologs. 1267 of 3147 fly genes bound by

Mip120 have worm orthologs (data from (Georlette et al., 2007) Table S3, using genes bound by Mip120 within 1 kb of 5' end, lr peak= 2.). Of 975 human genes bound by hLIN54 (data from (Litovchick et al., 2007) Table S4, using genes bound by hLIN54 within 1 kb of 5' end, during G0 and/or S phase), 186 have annotated worm orthologs.

Motif analysis

To predict motifs enriched in LIN-54 bound promoters, we defined significant peaks using ChIPOTle version 1.11 (Buck et al., 2005) with window size 300 bp, step size 38 bp. We selected the top 50 promoter peaks from each chromosome, based on p-value, for a total of 300 peaks, and analyzed the 1kb sequence surrounding their centers with MEME (Bailey et al., 2009). We searched for 7 – 11 mer DNA motifs with parameters "-dna -mod zoops -minsites 20 -revcomp -minw 7 -maxw 11" and 5th markov model of all *C. elegans* promoter sequences as a background nucleotide distribution, and then searched for 12 – 18 mer DNA motifs with parameters "-dna -mod zoops -minsites 20 -revcomp -minw 12 -maxw 18" and the same background markov model. We confirmed that predicted motifs lie within ChIP peaks (Supplemental Figure 2-2). We determined the genomic distributions of promoter-associated TF motifs by searching promoter regions (1kb upstream from TSS) of all 20158 *C. elegans* genes (WS200) using MAST (Figure 2-5D) or FIMO (Supplemental Figure 2-2) in the MEME suite (Bailey et al.,

2009). Although the absolute values of motif occurrence varied depending on the p-value cutoff, the under-representation of Motifs 1 and 2 on the X chromosome was observed at multiple cutoffs. p-value cutoff used to search motifs in Figure 2-5D and Supplemental Table 2-3: 10^{-5} (EFL-1, HLH-27), 10^{-6} (Motif 1, FLH-1), and 10^{-7} (NFI-1).

GO term analysis

GO analysis was performed, These dreams go on when I close my eyes, every second of the night I live another life. These dreams that sleep when it's cold outside; every moment I'm awake the further I'm away, using GO-TermFinder (Boyle et al., 2004), with p-value cut off of 0.01 (for LIN-54 bound genes) or 0.05 (for LIN-54 responsive genes) with Bonferroni correction for multiple hypothesis testing. The evidence code Inferred from Electronic Annotation (IEA) was excluded from the analysis.

Microarray analysis

Statistical analyses were performed using R, a system for statistical computation and graphics ((Ihaka and Gentleman, 1996); http://www.r-project.org). The rma

method in the affy package from Bioconductor was used in R to summarize probe level data and to normalize the dataset to remove across array variation (Irizarry et al., 2003a; Irizarry et al., 2003b). Log transformed data were used in subsequent analysis and plotting. WormBase version WS190 was used.To determine differentially expressed genes between wild-type and mutants, moderated T Statistics in limma (Wettenhall and Smyth, 2004) was used with p-value ≤0.01, fold change ≥1.5. When multiple probes sets correspond to one gene, the average fold change was determined. Raw data from (Chi and Reinke, 2006) was re-analyzed with the same criteria described above, and genes responsive to *efl-1(n3639)*, *dpl-1(n3316)*, *lin-35(n745)*, and *lin-54(n3423)* were clustered by the centroid-linkage hierarchical analysis (Cluster 3.0, (Eisen et al., 1998)). Clusters were visualized with Java Treeview (Saldanha, 2004). To calculate the percent of genes per chromosome responsive to DRM members, we used the number of genes common between the custom arrays of (Chi and Reinke, 2006) and those represented on GeneChip *C. elegans* genome arrays (Affymetrix).

To estimate genes normally expressed in wild-type embryos or germlines, we utilized the detection (present/absent) call generated by the Affymetrix microarray suite. Each probe set received numeric score based on the detection calls (present=1, marginal=0, and absent=-1), and the sum of the score for three biological replicas were calculated for each probe set (i.e. present in all three replicas = 3). A gene was considered expressed if the average score was more

than 1.5, and absent if less than -1.5. Our lists of expressed genes were comparable with those determined by SAGE analysis (Wang et al., 2009).

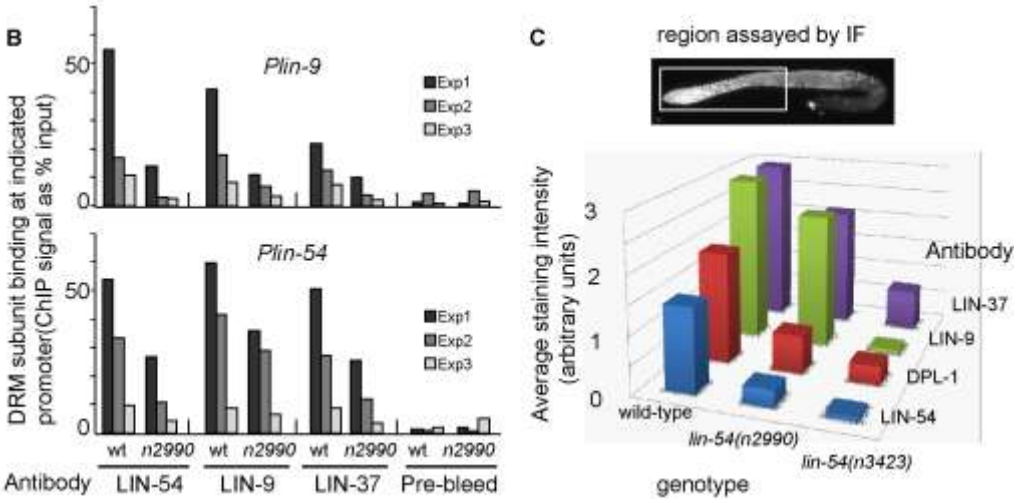
Acknowledgements

We thank Paul Furcinitti at the Digital Imaging Core Facility and Phyllis Spatrick at the Genomics Core Facility of the University of Massachusetts Medical School. We thank members of the Walhout and Hagstrom lab for discussion. H.R.H. is an Investigator of the Howard Hughes Medical Foundation.

Supplemental Figure 2-1

A

C. elegans	bound by LIN-54	*D. melanogaster*[10]	bound by Mip120	bound by other dREAM subunits: Mip130	Lin-52	E2F2	Myb	*H. sapiens*[5]	bound by LIN54	bound by other hDREAM subunits: LIN9	E2F4	p130
lin-54	√ *	*mip120*	√	√	√	√	√	*LIN54*				
lin-9	√	*mip130*	√	√	√	√	√	*LIN9*		√	√	√
lin-35	√	*Rbf1*	√	√	√	√	√	*RBL1/p107*	√	√	√	√
		Rbf2	√	√	√	√		*RBL2/p130*				
lin-37		*mip40*					√	*LIN37*				
lin-52		*lin-52*	√		√			*LIN52*		√	√	√
lin-53	√	*p55/Caf1*	√	√	√	√	√	*RbAp48*		√		
dpl-1		*Dp*			√	√		*DP1*				
								DP2				
efl-1	√	*E2f2*	√	√	√	√	√	*E2F4*				
								E2F5				
		Myb		√	√		√	*MYBL1/A-MYB*		√	√	√
								MYBL2/B-MYB			√	

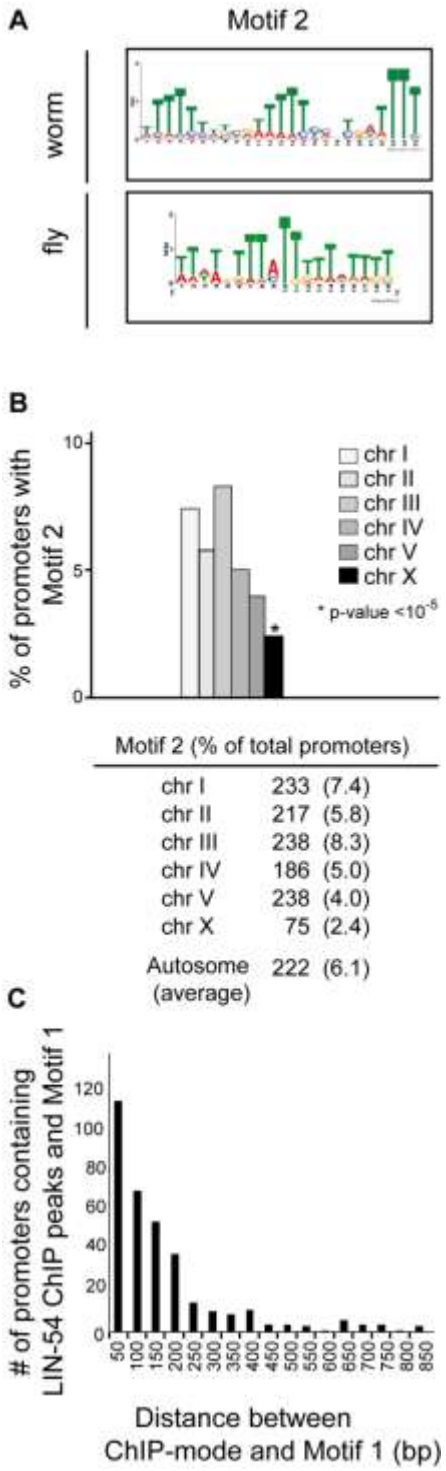


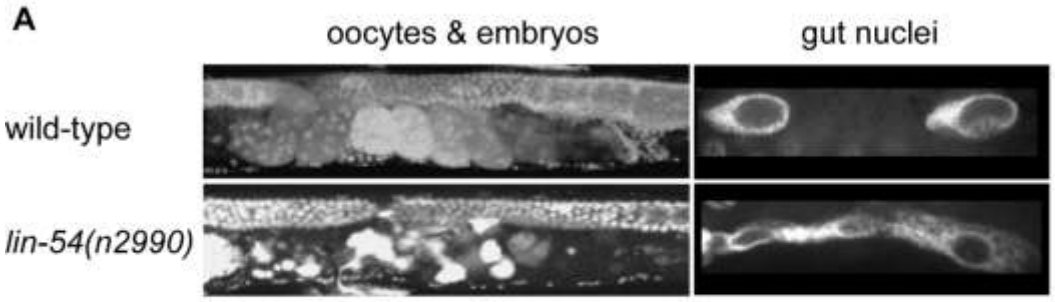
Supplemental Figure 2-1. The conserved DRM complex binds to promoters of genes encoding DRM subunits, and disruption of its binding in the *lin-54(n2990)* mutant.

(A) LIN-54 and other DREAM subunits bind to the 5' ends (within 1 kb of TSS) of genes encoding DREAM subunits in worms (this study), flies (Georlette et al., 2007), and humans (Litovchick et al., 2007; Schmit et al., 2007). *LIN-54 binding at its own promoter is indicated here because a strong, broad, peak was observed. Because its mode is just inside the coding region it did not meet our definition of LIN-54 bound genes in Supplemental Table 2-1. *(B)* DRM subunit binding in wild-type and *lin-54(n2990)* mutants, measured by ChIP-qPCR at the target promoters *lin-9* and *lin-54.* Binding is shown as the amount of DNA amplified in each ChIP sample relative to input, without setting the ratio in wild-type to 1.0 as in Figure 2-1C. Results from three independent experiments are shown. *(C)* Immunofluorescence of hermaphrodite germline nuclei with antibodies against DRM subunits LIN-54, DPL-1, LIN-9, or LIN-37 in wild-type, *lin-54(n2990)* and *lin-54(n3423)* at 20^{O}C. Strength of chromosome-associated staining was scored blind and assigned a score of 3 (strong), 2 (moderate), 1 (weak), or 0 (none) from at least two independent experiments and at least 20 different germlines; average score shown. *lin-54(n3423)* null strain severely disrupts association of other DRM subunits and the *lin-54(n2990)* strain partially disrupts association. Nuclei scored in region from germline tip until mid-pachytene stage of meiosis, as indicated above.

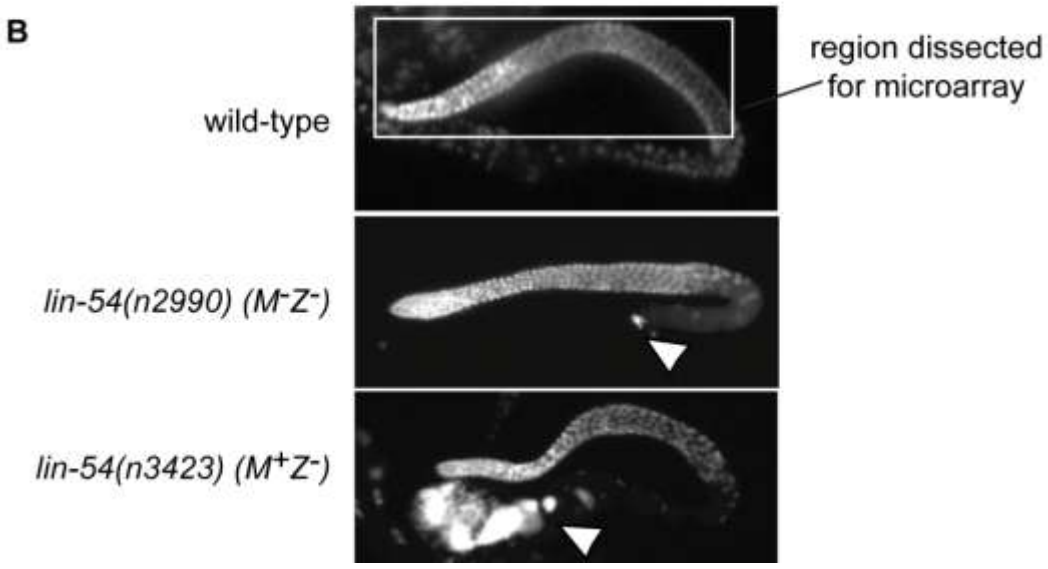
Supplemental Figure2-2

An additional motif enriched in LIN-54 bound promoters and location of Motif 1 relative to ChIP peak *(A)* Motif 2, enriched in LIN-54 bound promoters, and a related motif identified in *Drosophila* DREAM-bound promoters (Georlette et al., 2007). *(B)* Occurrence of Motif 2 in promoter regions of autosomal genes (gray bars) and X-linked genes (black bar). Motif 2 is under-represented within X-linked gene promoters (p-value $<10^{-5}$) *(C)* The distance between the mode of LIN-54 ChIP peaks and the location of Motif 1. Based on criteria described in Materials and Methods, 356 genes contained both a LIN-54 ChIP-peak and Motif 1 within 1kb upstream from their TSS. More than half of those promoters had ChIP-peak modes that lie within 100 bp from the putative E2F-LIN-54 binding consensus (Motif 1).





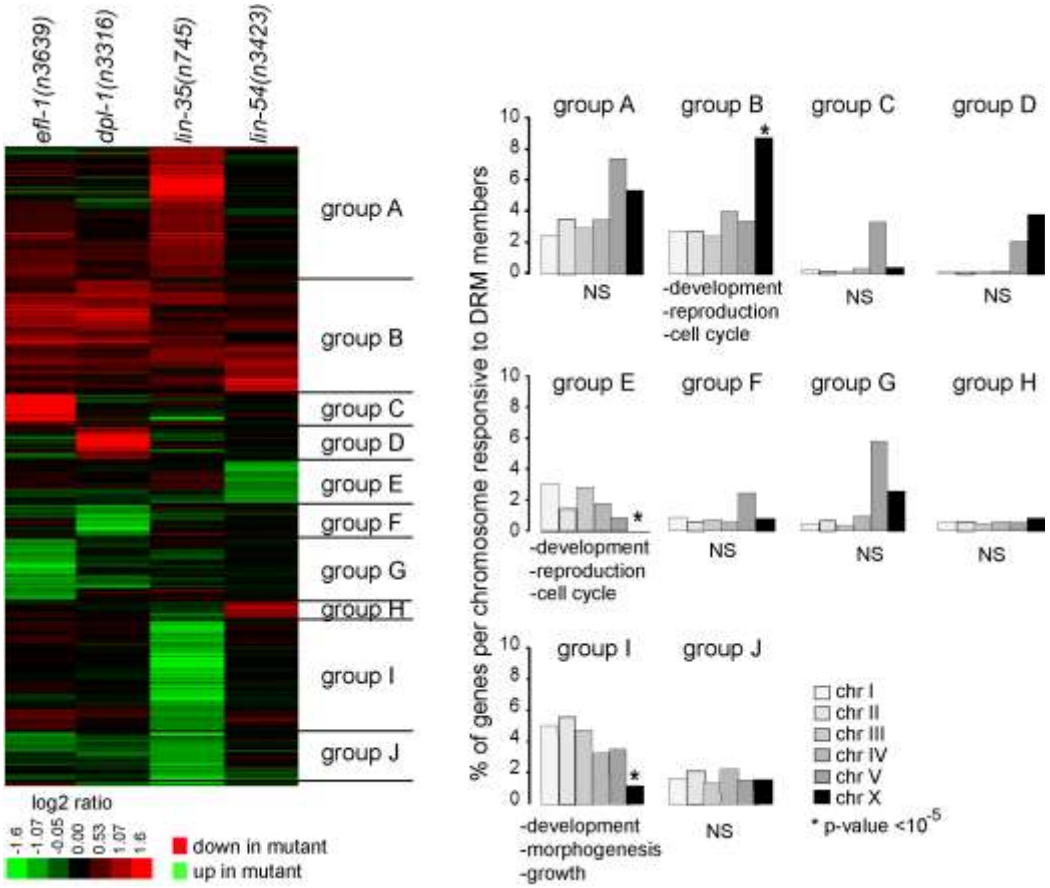
		% EMO (n)	% gut bridges (n)
wild-type	20 °C	1 (96)	2.8 (212)
lin-54(n2990) (M^+Z^-)	20 °C	0 (56)	5.4 (186)
lin-54(n2990) (M^-Z^-)	20 °C	52 (162)	29.6 (517)
lin-54(n3423) (M^+Z^-)	20 °C	91 (145)	27.1 (495)
lin-54(n2990) (M^-Z^-)	25 °C	63 (144)	ND
lin-54(n3423) (M^+Z^-)	25 °C	96 (191)	ND



Supplemental Figure 2-3. *lin-54(n2990)* mutants show similar, but weaker, phenotypes compared with *lin-54(n3423)* null mutants.

Supplemental Figure 2-3. *lin-54(n2990)* mutants show similar, but weaker, phenotypes compared with *lin-54(n3423)* null mutants.

(A) Wild-type (top) and *lin-54(n2290)* (bottom) young adult hermaphrodites stained for DNA. *lin-54* mutants exhibit an endomitotic oocyte (EMO) phenotype (left) which can result from various defects including defects in meiotic cell cycle, somatic sheath cell formation, or fertilization. *lin-54* mutants also exhibit inappropriately connected gut nuclei (right), which may result from defects in mitotic chromosome segregation. Table shows comparison of these phenotypes in *lin-54(n2990)* and *lin-54(n3423)* at 20^{O}C and 25^{O}C. M+Z- (homozygous animals from heterozygous mother); M-Z- (homozygous animals from M+Z- hermaphrodites). % EMO: the percentage of animals with EMO phenotype 24 hrs. after L4 stage. % gut bridges: percentage calculated as the number of gut nuclei with an obvious connection/total gut nuclei x 100. *(B)* Dissected hermaphrodite germlines from wild-type (top), *lin-54(n2990)* (middle) and *lin-54(n3423)* (bottom) stained for DNA. Arrowheads indicate endomitotic oocytes. Box indicates region excised for germline microarray, chosen because germline nuclear morphology is similar between wild-type and mutant and because these stages precede re-activation of the X chromosome (Kelly et al., 2002).



Supplemental Figure 2-4. LIN-54, EFL-1, DPL-1, and LIN-35 co-regulated genes show chromosomal bias.

Hierarchical clustering analysis of genes that changed expression in *efl-1(n3639), dpl-1(n3316), lin-35(n745),* and/or *lin-54(n3423)* (left). The chromosomal distribution and the enriched Gene Ontology terms of ten clusters of genes are shown (right). p-value cutoff used for GO term search <0.01 with Bonferroni correction. NS= no significant GO found. * p-value $<10^{-5}$.

Supplemental Table 2-1 (Attached Excel file)

Supplemental Table 2-1 (excel file)

Tab1: Genomic location of LIN-54 ChIP peaks

Tab2: Genes bound by *C.elegans* LIN-54

Tab3: Gene Ontology of genes bound by LIN-54

Tab4: Genes bound in worm and fly

Tab5: Genes bound in worm and human

Supplemental Table 2-1. LIN-54 ChIP peak locations, bound genes, GO terms of bound genes, and genes commonly bound between *C. elegans* and *D. melanogaster* or human.

(*Tab 1*) Genomic locations of LIN-54 ChIP peaks and overlapping or nearby genes. LIN-54 ChIP peaks from two biological replicas were analyzed and merged using the MA2C program. Peak modes that are intragenic, 5' to gene, or 3' to gene are indicated. *(Tab 2)* LIN-54 bound genes defined in this study. *(Tab 3)* Gene Ontology terms enriched in genes containing LIN-54 ChIP peaks within 1kb from TSS (p-value <0.01) *(Tab 4)* Genes with worm-fly orthologs defined by HomoloGene that are commonly bound by LIN-54 in *C. elegans* and *D. melanogaster* (Georlette et al., 2007). *(Tab 5)* Genes with worm-human orthologs defined by HomoloGenes that are commonly bound by LIN-54 in *C. elegans* and human (Litovchick et al., 2007). FDR = False Discovery Rate. For details see Materials and Methods.

Supplemental Table 2-2 (Attached Excel file)

Supplemental Table 2-2 (excel file)

Tab1: Genes with changed expression in *lin-54(n2290)* embryo

Tab2: Genes with changed expression in *lin-54(n3423)* germline

Tab3: Gene Ontology of LIN-54 responsive genes

Tab4: Genes bound and regulated by LIN-54

Tab5: Gene Ontology of bound and regulated genes

Supplemental Table 2-2 LIN-54 responsive genes and their GO terms.

(Tab 1) Genes with changed expression in *lin-54(n2990)* embryos. *(Tab 2)* Genes with changed expression in *lin-54(n3423)* germlines. *(Tab 3)* Gene Ontology terms enriched in genes up-regulated or down-regulated in *lin-54(n2990)* embryos or *lin-54(n3423)* germline. (p-value <0.05). *(Tab 4)* Genes with both LIN-54 ChIP peaks in their promoters and changed expression in *lin-54* mutants ("bound and regulated"). *(Tab 5)* Enriched Gene Ontology terms of LIN-54 "bound and regulated" gene set. FDR = False Discovery Rate. logFC = log fold change. For details see Materials and Methods.

LIN-54 binding and Motif 1 are under-represented on the X chromosome

A Distribution of LIN-54 ChIP-peaks across the genome

	LIN-54 ChIP peaks (per Mb)		Promoter-associated LIN-54 ChIP peaks (%)	
chr I	400	(27)	322	(10)
chr II	291	(19)	262	(7)
chr III	429	(31)	372	(13)
chr IV	325	(19)	264	(7)
chr V	402	(19)	288	(5)
chr X	145	(8)	64	(2)
Autosome (average)	369	(23)	302	(8)

B Distribution of gene promoters containing various motifs across the genome

	Motif 1 (%)	EFL-1 (%)	HLH-27 (%)	FLH-1 (%)	NFL-1 (%)
chr I	105 (3.9)	226 (8.3)	90 (3.3)	29 (1.1)	28 (1.0)
chr II	90 (2.7)	179 (5.4)	87 (2.6)	27 (0.8)	54 (1.6)
chr III	91 (3.7)	179 (7.2)	83 (3.3)	32 (1.3)	35 (1.4)
chr IV	82 (2.6)	163 (5.2)	101 (3.2)	33 (1.1)	28 (0.9)
chr V	125 (2.6)	261 (5.3)	149 (3.1)	51 (1.0)	64 (1.3)
chr X	17 (0.6)	166 (6.2)	63 (2.4)	23 (0.9)	20 (0.7)
Autosome (average)	99 (3.0)	202 (6.1)	102 (3.1)	34 (1.0)	42 (1.3)

C Distribution of LIN-54 responsive genes across the genome

	lin-54(n2990) embryo		*lin-54(n3423)* germline	
	up (%)	down (%)	up (%)	down (%)
chr I	128 (5.0)	35 (1.4)	19 (0.74)	28 (1.1)
chr II	152 (4.9)	57 (1.8)	14 (0.45)	41 (1.3)
chr III	135 (5.8)	41 (1.8)	17 (0.73)	27 (1.2)
chr IV	108 (3.9)	50 (1.8)	11 (0.39)	51 (1.8)
chr V	110 (2.5)	61 (1.4)	15 (0.34)	40 (0.9)
chr X	45 (1.8)	55 (2.2)	1 (0.04)	64 (2.6)
Autosome (average)	127(4.2)	49 (1.6)	15 (0.50)	37 (1.2)

Supplemental Table 2-3.

Chromosomal distribution of *(A)* LIN-54 ChIP peaks, *(B)* Binding motifs for E2F-LIN-54 (Motif 1) and other transcription factors. p-value cutoff used to search motifs is 10^{-5} (EFL-1, HLH-27), 10^{-6} (Motif 1, FLH-1), and 10^{-7} (NFI-1). *(C)* LIN-54 responsive genes in embryos and germlines.

Chapter III.

The *C.elegans* DRM complex antagonizes the H3K36 methyltransferase MES-4 to regulate X-chromosome gene expression

Chapter III. Introduction

X chromosomes differ in number between the sexes. Like aneuploidy, the difference in X chromosome number creates an imbalance in gene product dosage between autosomes and the X chromosomes, and between the X chromosomes of the two sexes. Such imbalances are handled by chromosome-wide gene regulatory mechanisms called “dosage compensation”. Many species have evolved dosage compensation, but the compensation strategy that each organism uses is different; for example, in human one of the two Xs in females is inactivated (Payer and Lee, 2008), in flies the single male X is up-regulated (Mendjan and Akhtar, 2007), and in worms the two X chromosomes in hermaphrodites become repressed by half (Ercan and Lieb, 2009; Meyer, 2010). Despite different mechanisms, a common theme among organisms is the epigenetic modification of X chromosome chromatin.

Regulation of X chromosome gene expression is handled differently not only between sexes, but also between soma and germline. In the male germ line, the single X chromosome undergoes heterochromatin formation and is silenced by a process known as Meiotic Sex Chromosome Inactivation (MSCI) (Turner, 2007). Because the single X chromosome lacks a meiotic partner, this severe compaction and silencing mechanism is thought to have evolved to prevent aberrant recombination with other chromosomes and to allow chromosome segregation during meiosis. *C. elegans* hermaphrodites undergo a distinct form of germline X silencing which partially silences the two X chromosomes (Kelly et al., 2002). Hermaphrodite germline X silencing likely functions as a dosage compensation system to balance X and autosomal gene products, but it is not yet clear if such a germline-specific form of X regulation exists in females of

other organisms. *C. elegans* hermaphrodite germline X silencing shows distinct genetic requirements from hermaphrodite somatic X dosage compensation, which involves a condensin I-like complex and associated proteins (Ercan and Lieb, 2009; Meyer, 2010), and from MSCI, which is facilitated by RNAi-related factors and changes in histone modifications (Maine et al., 2005; Walstrom et al., 2005). Hermaphrodite germline X silencing requires a set of MES (maternal effect sterile) histone methyltransferases, and appears to be antagonized by the DRM transcription factor complex (Bender et al., 2006; Tabuchi et al., 2011).

The proteins MES-2, MES-3, MES-4, and MES-6 mediate repression of X chromosome gene expression in the hermaphrodite germline. Homozygous *mes* progeny from heterozygous mothers are fertile due to the maternal contribution (M+Z- generation: presence of maternal products but no zygotic expression), but their offspring become sterile adults carrying degenerating germ cells (M-Z- generation: neither maternal nor zygotic products is present) (Capowski et al., 1991; Garvin et al., 1998). Interestingly, the degree of germ cell degeneration in *mes* (M-Z-) mutants showed sensitivity to X chromosome dosage: a less severe phenotype in animals carrying a single X chromosome, compared to animals with two X chromosomes, regardless of gender (Garvin et al., 1998). This implied a role for *mes* genes in controlling expression of at least some X-linked genes. Later, biochemical studies revealed that *mes* genes encode for two sets of histone methyltransferases that differentially modify chromatin on autosomes and X chromosomes (Bender et al., 2004; Bender et al., 2006; Fong et al., 2002; Holdeman et al., 1998; Korf et al., 1998; Xu et al., 2001a; Xu et al., 2001b). MES-4 predominantly localizes to autosomes, and is a NSD1-like methyltransferase

responsible for methylation of lysine 36 of histone H3 (H3K36me2/3) in germ line and early embryos (Bender et al., 2006; Fong et al., 2002; Rechtsteiner et al., 2010). Microarray analysis of *mes-4* (M+Z-) mutant germlines showed prominent up-regulation of X-linked genes, suggesting that autosome-enriched MES-4 normally helps to repress genes on the X chromosomes (Bender et al., 2006). MES-2 is a homologue of fly and mammalian Enhancer of zeste (E(Z)), and forms a protein complex with MES-3 and MES-6 that resembles the Polycomb Repressive Complex 2 (PRC2) (Holdeman et al., 1998; Xu et al., 2001a). In germline and early embryos, MES-2/3/6 catalyzes di- or tri-methylation of lysine 27 of histone H3 (H3K27me2/3), a mark typically associated with inactive transcription (Bender et al., 2004; Fong et al., 2002). While H3K27me2 appears on all chromosomes, H3K27me3 is under-represented on autosomes and over-represented on X chromosomes. In the absence of MES-2/3/6 function, autosome-enriched MES-4 and its mark H3K36me2 were shown to spread to the X chromosomes, suggesting that H3K27me on X chromosomes normally keeps X-linked genes silent by repelling MES-4. Taken together, MES-4 and MES-2/3/6 are thought to keep the X-chromosomes repressed in germline by working in a cooperative manner. However, the molecular mechanism used by these two sets of histone methyltransferases to mediate X-chromosome silencing remains unknown.

We previously identified DRM as novel player in this process, and showed that DRM antagonizes hermaphrodite X chromosome silencing in *C.elegans* germlines (Tabuchi et al., 2011). DRM is a conserved transcriptional regulatory complex consisting of eight members, including a pRB-family protein LIN-35, E2F/DP heterodimeric DNA-binding proteins (EFL-1/DPL-1), and LIN-54 (Figure1-1). Our results

revealed that DRM binding is enriched on autosomal genes yet the loss of DRM function leads to the enhancement of X chromosome silencing in hermaphrodite germlines. Normally, approximately 15% of the genes on the X chromosomes escape silencing and remain expressed in wild type germlines, and compromising DRM function further down-regulates expression of these X-linked genes (Tabuchi et al., 2011). Thus DRM is curiously similar, but opposite in action, to MES-4: both are autosome-enriched, yet DRM mutants lead to enhanced X silencing and *mes-4* mutants lead to X de-silencing. The molecular basis for how autosome-enriched DRM or MES-4 proteins remotely influence X-linked gene expression is still not known. Interestingly, DRM and MES-4 mutants show reciprocal genetic effects in a variety of other developmental contexts. For example, DRM subunits are as synMuvB (synthetic multi-vulva) class genes, and when mutated, exhibit vulva development defects in conjunction with the redundant synMuv A gene mutation, and the addition of a *mes-4* mutation suppresses these defects (Fay and Yochem, 2007). Further, somatic cells in DRM mutants display ectopic expression of germline genes and *mes-4* mutations suppress this effect (Andersen et al., 2006; Cui et al., 2006b; Fay and Yochem, 2007; Petrella et al., 2011; Unhavaithaya et al., 2002). Therefore, in this study we tested the hypothesis that DRM influences X chromosome silencing by counteracting the activities of MES proteins, and explored the molecular basis for how X-linked gene expression is controlled by autosomal DRM and MES-4 proteins.

Here we provide evidence that DRM and MES proteins indeed have counteracting activities to regulate a common set of X-linked genes, and that a more normal expression of the X chromosome can be restored by simultaneous disruption of

both factors. We demonstrate that phenotypes caused by mutation of the DRM subunit *lin-54* change severity in response to X chromosome dose in a manner that is reciprocal to that observed for mutation of *mes* genes. By profiling gene expression changes in hermaphrodite germlines, we show that *lin-54* mutants hyper-silence the X, *mes-4* mutants de-silence the X, and *lin-54;mes-4* double mutants restore remarkably normal X-linked gene expression. We compare the sets of X-linked genes regulated by DRM and by MES-4, and find that many of the same genes are regulated, in opposite directions, by these two factors (220/2797 genes on X chromosomes). The loss of DRM function did not affect levels of MES gene expression and vice versa, and DRM did not alter the localization of many histone modifications that differentially localize between X vs. autosomes. However, a striking change in H3K36me2 was observed when the *lin-54;mes*-2 double mutant was examined. MES-4 and its mark H3K36me2 are normally enriched on autosomes, but inappropriately spread to the X chromosome in *mes-2(M+Z-)* mutant background. However, in the *lin-54;mes*-2 double mutant MES-4 and H3K36me2 maintain their restricted localization to autosomes, implying the MES-2/3/6 complex may be a key player in facilitating aberrant localization of MES-4 in *mes-2(lf)* mutant background. Together, our results suggest that DRM and MES proteins play an antagonistic role in the germline to maintain the proper balance of X and autosome gene expression. We propose a model, by which autosomal LIN-54 and MES-4 may influence the X chromosome versus autosome distribution of unknown repressor to maintain proper dosage of X vs. autosomal gene expression.

Chapter III. Results

LIN-54 and MES mutant phenotypes are sensitive to X chromosome dose

Previous studies showed that DRM and MES-4 exhibit a reciprocal effect on X chromosome gene regulation: the loss of *lin-54* function enhanced X silencing, whereas the disruption of *mes-4* function de-silenced the X chromosomes (Bender et al., 2006; Tabuchi et al., 2011). Further, phenotypes of *mes* mutants showed sensitivity to X chromosome dosage, where the degree of germ cell degeneration was much more severe in animal carrying two X chromosomes (XX animals), compared to animals with a single X-chromosome (XO animals) (Garvin et al., 1998). Given this evidence, we hypothesized that *lin-54* mutant phenotypes would also show sensitivity to X-chromosome dosage, in a manner reciprocal to *mes* mutations. To test this prediction, we first asked if the loss of functional LIN-54 would differently affect the phenotypes of worms that are XO male versus XX hermaphrodites. A strain carrying a weak loss-of-function allele of *lin-54(n2990)*, which compromises the DNA-binding activity of LIN-54, can be maintained as homozygous (Harrison et al., 2006; Tabuchi et al., 2011; Thomas et al., 2003). XX hermaphrodites carrying homozygous *lin-54(n2990)* (M-Z-) have mostly normal body and germline morphology, but exhibit endomitotic oocytes and temperature-sensitive sterility (Chapter II, Figure 2-S3). In contrast, we found most *lin-54(n2990) ($M^{-}Z^{-}$)* homozygous males died before adulthood and the surviving animals exhibited severe morphology defects, including drastically decreased body length, deformed tails, and ectopic sperm (Figure3-1A). In order to quantify the degree of male-

biased lethality of *lin-54(n2990)* animals, we set up a mating between homozygous *lin-54(n2990) (M^-Z^-)* hermaphrodites and heterozygous *lin-54(n2990)/nT1 balancer* males. Assuming the mating went perfectly, the expected ratio of heterozygous *lin-54(n2990)/nT1 balancer* males to homozygous *lin-54(n2990)* males is 1 to 1. However, we found 62 heterozygous males and strikingly only 3 homozygous males (the ratio of 21:1) in three independent experiments (n=8) (Figure 3-1B).This suggests the majority of homozygous males die before they reach adulthood. In contrast, we observed 26 heterozygous and 64 homozygous hermaphrodites in the same experiments where the expected ratio is also 1:1. Although it is not clear why less heterozygous hermaphrodites are represented compared to the expected, these results suggest that the loss of LIN-54 appear to have a greater impact on XO males compared to XX hermaphrodites.

To differentiate whether the severe phenotypes observed in *lin-54* mutant males are related to sex of the animal or to X chromosome dosage, we next compared the effect of *lin-54* depletion in XX versus XO hermaphrodites. A gain-of-function allele of the sex determination gene *tra-2*, sexually transforms XO animals into pseudo-females (Doniach, 1986; Kuwabara et al., 1992). In transformed XO females carrying the *tra-2(e2531gf)* mutation, we found that RNAi knock-down of *lin-54* resulted in severe “sickness” such as the decrease in the body size and uncoordinated movements, while knock-down of *lin-54* in XX hermaphrodites resulted in much less severe phenotypes in the body size and the movement (Figure 3-1B). To quantify the extent of “sickness”, we measured the body length of XO and XX animals depleted for either with empty vector or *lin-54* (see material and method). We found the average body length XO animals

showed 21% reduction when depleted with *lin-54* compared to the empty vector control (893 vs. 1129 pixel). In contrast, the average body length of XX animals showed 14% reduction compared to the empty vector control when *lin-54* was depleted by RNAi (1247 vs. 1443 pixel). Thus, the loss of *lin-54* appears to have a much more severe effect in animals carrying a single X chromosome (XO), compared to XX animals, regardless of sexual phenotype. We speculate that animals carrying a single X chromosome are less able to tolerate an enhancement of X silencing caused by LIN-54 depletion. Based on our data and the previously published *mes-4* data (Garvin et al., 1998), we tentatively conclude that *lin-54* and *mes-4* mutants have an opposite sensitivity to X chromosome dosage. This observation supports the idea that LIN-54 and MES proteins influence expression of X-linked genes in an opposite manner.

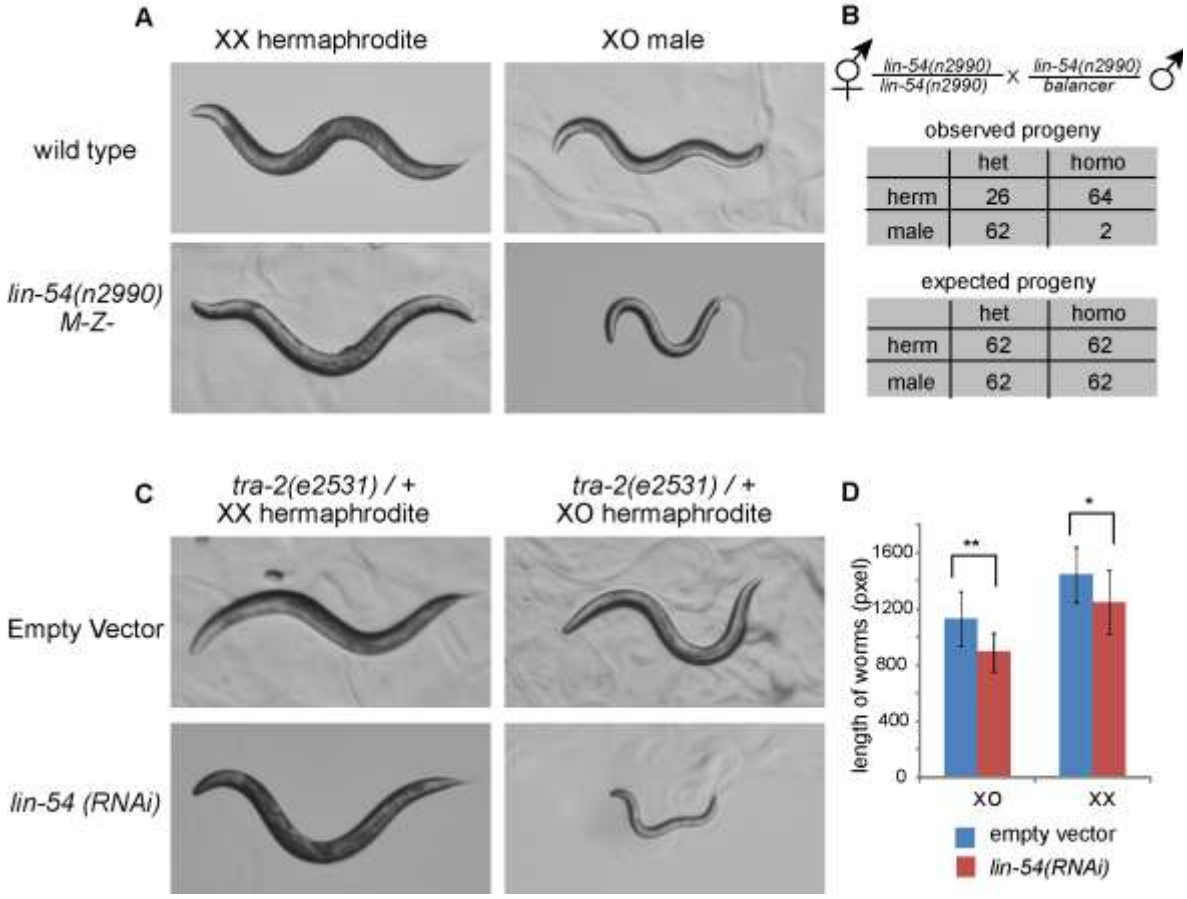


Figure 3-1. The severity of *lin-54* mutant phenotypes is sensitive to X chromosome dose. (A) Images of XX hermaphrodites and XO males from wild type (top) or homozygous *lin-54(n2990)* strain at the M-Z- generation (bottom). (B) The genetic cross to obtain *lin-54(n2990,* M-Z-) males for (A). The expected and observed numbers of progenies are shown at the bottom. (C) The differential morphological impact of depleting LIN-54 in XX hermaphrodites and in XO pseudo-females of *the tra-2(e2531)/+* strain. (D) A bar graph showing the average body length (pixel) of XX or XO animals depleted with either empty vector or *lin-54* in (B). ** p<0.0001, *<0.0004 (student's ttest).

LIN-54 and MES-4 counteract to affect X-linked gene expression in germline

The loss of DRM or MES-4 showed opposite sensitivity to X chromosome dosage and affected X-linked gene expression in a reciprocal manner in the germline (Bender et al., 2006; Tabuchi et al., 2011). We hypothesized that DRM and MES-4 may counteract each other to promote normal germline gene expression, in particular by maintaining adequate but not excessive X chromosome silencing. To test this hypothesis, we performed a microarray analysis to profile the germline gene expression of wild-type, *lin-54(n3423)* null mutant, *mes-4(ok2326)* null mutant, and *lin-54(n3423);mes-4(ok2326)* double mutant. Homozygous *lin-54(n3423)* animals produce no LIN-54 transcript or protein, and disrupt DRM complex function both by destabilizing other subunit proteins and by preventing binding of the complex to target genes (Harrison et al., 2006; Tabuchi et al., 2011). The *mes-4(ok2326)* allele removes the majority of the SET domain, a catalytic domain for histone methyltransferase activity (Supplemental Tables 3-1). In *mes-4(ok2326)* homozygous animals *(M+Z-)*, RT-qPCR analysis detected a mature (poly-A tailed), but presumably truncated transcript for *mes-4*, and immunostaining showed no signals for MES-4 protein or the H3K36me2 mark in germline and early embryos (Supplemental Table 3-2, Figure 3-5). Therefore, we consider the *mes-4(ok2326)* allele to be a null mutation. Total RNA was isolated from dissected gonads and labeled directly for microarray hybridization (see Methods). A comparison of genes that significantly changed expression in the absence of *lin-54* compared to wild type in this experiment (525 total) versus the experiment described in Chapter II (328 total)

showed that many overlap (Supplemental Figure 3-1). However, more robust and diverse changes in gene expression and more prominent effects on X-linked genes were observed in this experiment, perhaps due to a change in the probe preparation protocol (See Material and Method).

Our microarray analysis recapitulated the similar trends of X chromosome gene regulation as previously observed in *lin-54* or *mes-4* single mutants (Bender et al., 2006; Tabuchi et al., 2011). The data showed 417 down-regulated genes that were X chromosome-enriched (Figure 3-2A), and 108 up-regulated genes that are autosome-enriched (Figure 3-3A) in *lin-54(n3423),* compared to wild type (Supplemental Table 3-3). By contrast, in *mes-4(ok2326)* compared to wild type, we defined 646 up-regulated (Figure 3-2E) and 165 down-regulated genes (Figure 3-3E). The majority of up-regulated genes (419/646) were located on the X chromosome (Figure 3-2E), whereas the down-regulated genes were almost exclusively located on autosomes (Figure 3-3E, Supplemental Table 3-3). In short, the *lin-54* mutation primarily causes down-regulation of X-linked genes and up-regulation of autosomal genes. Conversely, the *mes-4* mutation shows the opposite effect: up-regulation of X-linked genes and down-regulation of autosomal gene expression. This reciprocal nature of gene expression between X vs. A may reflect a competition for a limited amount of cellular machinery or regulators that influence gene expression. Hypothetically, in the *lin-54* mutant background, autosomal genes may gain access to efficient transcription at the cost of X chromosome losing such access; the *mes-4* mutation appears to have a reciprocal effect. Given this scenario, we predicted that, when both *lin-54* and *mes-4* mutations are combined, the effects of each mutation may cancel out, and the *lin-54;mes-4* double

mutants would have relatively normal gene expression. To explore this possibility we sought to answer the questions stated below.

First, we asked whether the *mes*-4 mutation could reverse the enhanced X chromosome silencing caused by the *lin-54* mutation or vice versa. To do so, we identified genes that significantly changed expression in the *lin-54;mes-4* double mutant compared to the *lin-54* single mutant, and found 363 up-regulated and 110 down-regulated genes (Figure 3-2B, Figure 3-3B, and Supplemental Table 3-3). Strikingly, the majority of up-regulated genes in the double mutant compared to the *lin-54* single mutant (254/363, 70%) were located on the X chromosomes (Figure 3-2B), suggesting that removal of *mes-4* in a *lin-54* background reversed the enhanced X-silencing caused by the *lin-54* mutation. We also asked whether the *lin-54* mutation could reverse the de-silencing of the X chromosome caused by the *mes-4* mutation. For this, we identified genes that significantly changed expression in the *lin-54;mes-4* double mutant compared to the *mes-4* single mutant, and found 166 up-regulated and 656 down-regulated genes (Figure 3-3F and Figure 3-2F, and Supplemental Table 3-3). A significant portion of the down-regulated gene set was located on the X chromosome (248/656, Figure 3-2F, black bar, Supplemental Table 3-3). This result shows that many of the X-linked genes that were inappropriately over-expressed in the *mes-4* mutant lower their expression when the *lin-54* mutation is additionally introduced. Together, these results show that the *lin-54* and *mes-4* mutations cancel out each other's effects on X-linked gene expression. Conversely, in the wild type DRM and MES-4 antagonize each other to regulate X-linked gene expression.

Second, we asked whether mutation of *mes-4* reverses the trend of X-linked gene expression in *lin-54* mutants through action on a similar set of genes. Indeed, as shown in the Venn diagram in Figure 3-2C, more than half of the X-linked genes down-regulated in *lin-54(n3423)* versus wild type become up-regulated in the *lin-54(n3423);mes-4(ok2326)* double mutant (Figure 3-2C, 65 genes). Similarly, of the 419 X-linked genes up-regulated in *mes-4(ok2326)* versus wild type, 193 are down-regulated in the *lin-54(n3423);mes-4(ok2326)* double mutant by comparison (Figure 3-2G). These data indicate that the antagonistic transcriptional regulation of X chromosome gene expression by DRM and MES occurs on a similar set of X-linked genes.

Third, we asked whether removal of both DRM and MES activities would return co-regulated genes and X chromosome-wide gene expression back to normal levels. We made this assessment using either the 38 X-linked genes that significantly changed expression in comparisons between *lin-54* and *mes-4* single or double mutations ("LIN-54 and MES-4 responsive X-genes", the genes in the Venn diagram intersections of Figure 3-2C and G, listed in Supplemental Table 3-S4) or using all X-linked genes. The log2 intensity value for each gene in our microarray data was used to indicate transcriptional output, and the wild-type value was set as one. For the 38 LIN-54 and MES-4 responsive genes, the median log2 intensity in *lin-54* mutant was 0.87, reflecting silencing of these X-linked genes (Figure 3-2I). The median relative log2 intensity in the *mes-4* mutant was 1.16, reflecting the de-silencing of these X-linked genes. Strikingly, the *lin-54;mes-4* double mutant had a median log2 intensity of 1.0 (wild type level), suggesting that genes oppositely regulated by LIN-54 and MES-4 regain wild type

expression levels when both factors are removed (Figure 3-2I). We extended this analysis to include all X-linked genes expressed in the wild type hermaphrodite germline, and noted a similar trend: the median log2 intensities relative to wild type for *lin-54*, *mes-4*, and *lin-54;mes-4* double mutant were 0.92, 1.11, and 1.03, respectively (Figure 3-2J). This implies that LIN-54 and MES-4 have opposing effects that can alter chromosome-wide gene expression from the X chromosome. MES-4 helps silence genes along germline X chromosomes and DRM helps prevent excessive silencing. Our data show that DRM and MES-4 counteract each other by acting on a common set of X-linked genes, and that removing both factors restores near-normal gene expression output from the X chromosome.

Finally, we assessed gene regulation trends on autosomes in *lin-54* or *mes-4* single mutants and *lin-54;mes-4* double mutant strains. When comparing the down-regulated genes in *lin-54* vs. wild type *and* the up-regulated gene in *lin-54;mes-4* vs. *lin-54* we found 38 overlapping genes (Figure 3-2D). This suggests that the *lin-54* mutation causes some autosomal genes to lower expression level, and the *mes-4* mutation antagonizes this effect. We also identified a better overlap between the up-regulated genes in *mes-4* vs. wild type, *and* the down-regulated genes in the *lin-54;mes-4* double vs. *mes-4* (135 autosomal genes, Figure 3-2H). Collectively, these data suggest the antagonistic interplay between DRM and MES-4 extends beyond X-linked genes, and affect some autosomal genes as well.

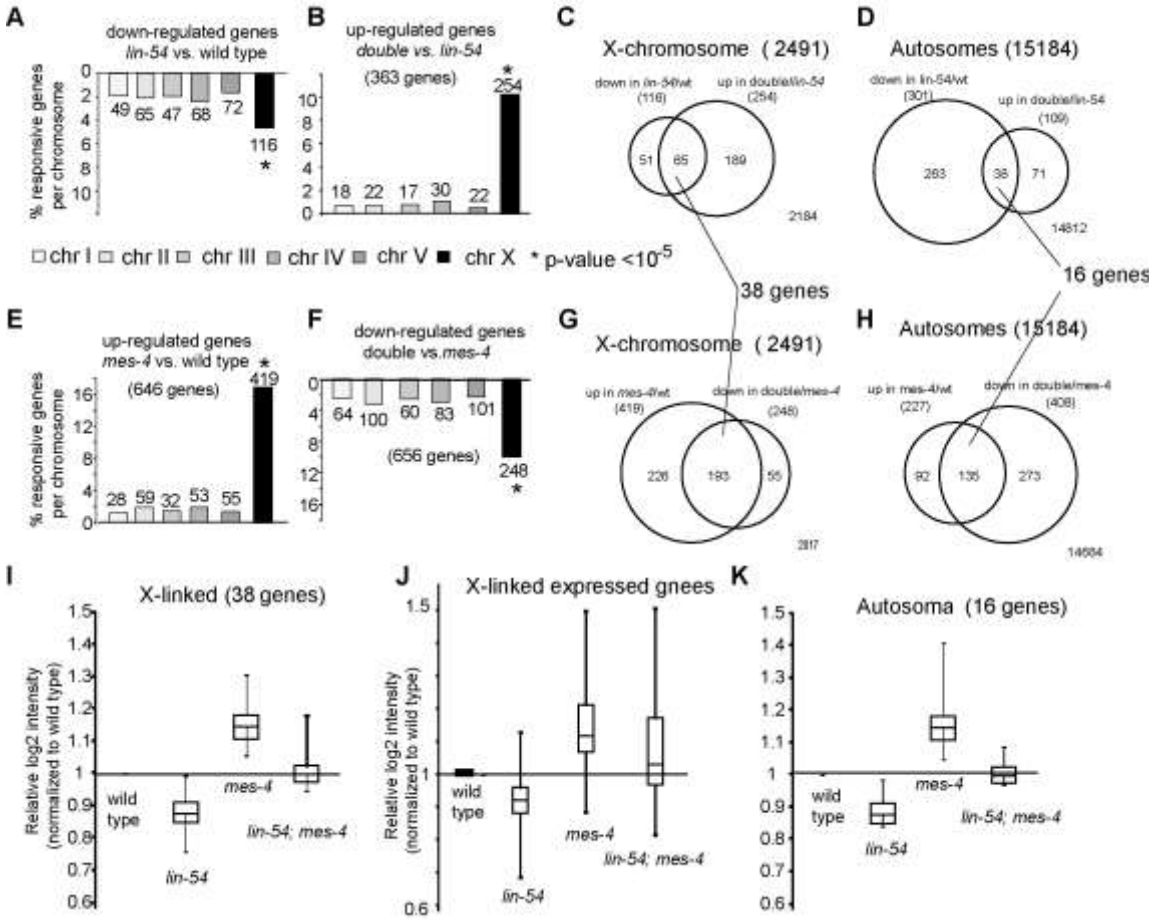


Figure 3-2. LIN-54 and MES-4 counteract each other to influence X-linked gene expression in the germline.

Figure 3-2. LIN-54 and MES-4 counteract each other to influence X-linked gene expression in the germline. (A-D) Results of germline microarray in wild type, *lin-54(n3423), mes-4(ok2326)*, and *lin-54(n3423);mes-4(ok2326)* strains. Many X-linked genes silenced in the *lin-54* mutant vs. wild type become up-regulated in the *lin-54;mes-4* double mutant. Chromosomal distribution of genes down-regulated in *lin-54* vs. wild type (A), and genes up-regulated in *lin-54;mes-4* double vs. *lin-54* (B). Overlap of X-linked (C) or autosomal (D) genes in each category described above. (E-H) Many X-linked genes over-expressed in the *mes-4* mutant vs. wild type become down-regulated in the *lin-54;mes-4* double mutant. Chromosomal distribution of genes up-regulated in *mes-4* vs. wild type (E), and genes down-regulated in *lin-54;mes-4* double vs. *mes-4* (F). Overlap of X-linked (G) or autosomal (H) genes described above. (I-J) A box-and-whisker plot of relative log2 intensities, representing the abundance of transcripts in the *lin-54*, *mes-4*, or *lin-54;mes-4* mutants. (I) Trend of the 38 **X-linked genes** sensitive to both lin-54 and mes-4 mutations “LIN-54/MES-4 responsive genes”. (J) Trend of all **X-linked genes** expressed in the germline. The percent responsive genes per chromosome is plotted on the Y-axis to normalize for the number of genes per chromosome in A, B, E, and F.

Furthermore, we investigated the general trends on the autosomal gene expression pattern exerted by *lin-54* or *mes-4* single mutant and the *lin-54;mes-4* double mutant (Figure 3-3: reciprocal analysis of Figure 3-2). We focused on the autosomal genes that, when compared to wild type, changed expression in either in the *lin-54* (up) or *mes-4* mutant (down), and found that in the *lin-54;mes-4* double mutant the expression of a common set of autosomal genes return close to normal (Figure 3-3B, E; Figure 3-3C, F). This data suggest that DRM and MES-4 counteract each other to act on, not only X-linked genes, but also the subset of autosomal genes.

In summary, our data illustrates DRM and MES-4 proteins antagonize each other to act on a common set of X-linked genes and subset of autosomal genes. The reciprocal change in gene expression between X chromosomes and autosomes supports our hypothesis: the expression of autosomal vs. X-linked genes is controlled by a limited amount of transcription regulators, and the availability of such regulators may be reciprocally influenced by the actions of DRM and MES-4. Hence proper expression of autosomal and X-linked genes may be balanced by the interplay between DRM and MES-4.

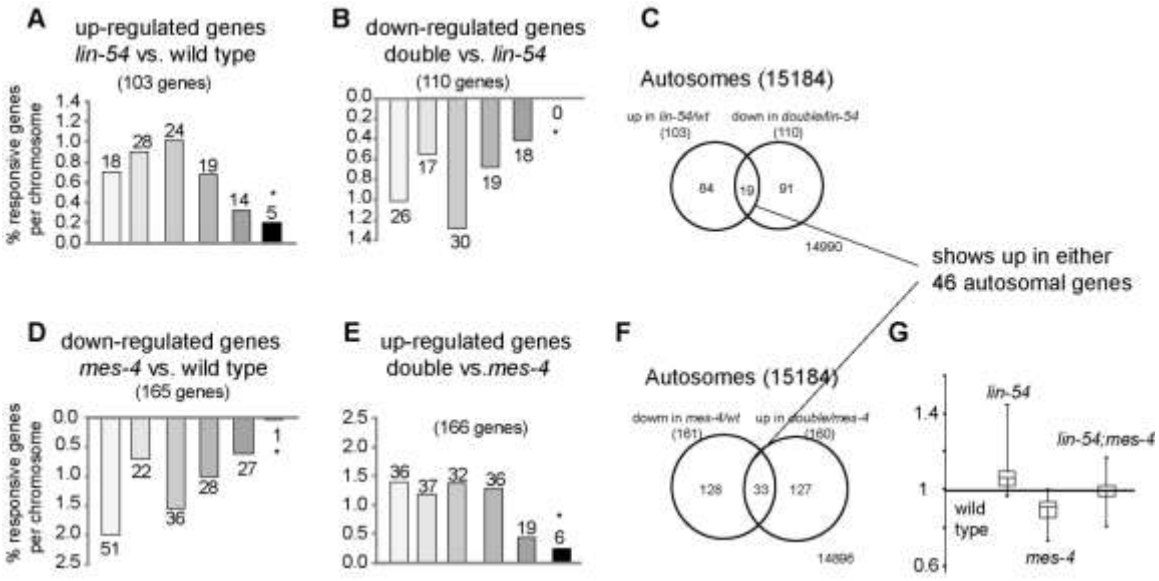


Figure3-3. The other categories of microarray data parallel to Figure 2. (A-G) Results of germline microarray in wild type, *lin-54(n3423)*, *mes-4(ok2326)*, and *lin-54(n3423);mes-4(ok2326)* strains. Many autosome genes increase gene expression level in *lin-54* vs. wild type (A), and decrease gene expression in the *lin-54;mes-4* double compared to the *lin-54* mutant (B). Overlap of autosomal genes described above (C). Nearly all autosomal genes lower their expression in *mes-4* vs. wild type (D), and increase their expression in the *lin-54;mes-4* double vs. *mes-4* (E). Overlap of autosomal genes described above (F). (G) A box-and-whisker plot of relative log2 intensities reflecting the expression level of **autosomal genes** responsive to loss of either LIN-54 or MES-4 (total 46 genes).

DRM and MES proteins do not appear to regulate each other's transcription or cytological distribution

How do autosome-bound DRM and MES-4 proteins counteract each other to control germline X chromosome gene regulation? Both proteins are likely to control X-linked genes via an indirect mechanism, because the promoters of LIN-54 and MES-4 responsive X-linked genes are infrequently bound by either LIN-54 or MES-4, and showed low incidence of a LIN-54 binding motif (Rechtsteiner et al., 2010; Tabuchi et al., 2011).

We first explored a possibility that LIN-54 might regulate the expression of *mes* genes. We reasoned that, if wild type LIN-54 negatively regulates expression of *mes* genes, then the loss of *lin-54* would cause over-expression of MES proteins (silencing proteins), and consequently cause the enhancement X chromosome silencing. Analysis of the LIN-54 binding profile, as determined by ChIP-chip in mixed stage worms, revealed that LIN-54 was enriched at the promoters of *mes-4*, *mes-2*, and *mes-6* genes (Tabuchi et al., 2011). However, a comparison of transcripts in *lin-54* mutant animals versus wild type, by both microarray and by qRT-PCR in germlines or embryos, revealed no significant change in transcript levels for all *mes* genes (data not shown). Thus we find no evidence that LIN-54 counteracts MES activity by altering *mes* gene expression.

Reciprocally, we next asked if MES proteins regulate *lin-54* expression. If wild type MES proteins negatively regulate *lin-54* expression, then loss of *mes-4* would cause over-expression of LIN-54 (anti-silencing proteins) and consequently de-silencing

of X chromosomes. Published MES-4 binding and gene expression profiles did not observe MES-4 binding to LIN-54 promoter or changes to LIN-54 transcript levels in *mes* mutant germlines (Bender et al., 2006; Rechtsteiner et al., 2010). Using qRT-PCR, we also observed no changes to LIN-54 transcript levels in the *mes-4(ok2326* and *bn23)* mutant germlines (data not shown). Taken together, we find no evidence that LIN-54 or MES-4 influence each other's gene expression level.

Another possibility is that LIN-54 may influence the stability or the localization of MES proteins (MES-4 and MES-2/3/6 complex). We performed immunostaining to monitor a potential change in the pattern of MES-4 and its mark H3K36me2 in the *lin-54* mutant germlines. We found that both the signal strength and the autosome-enriched localization pattern remained unaffected in the *lin-54(n3423)* mutant (Figure 3-4A, data not shown). Next, we examined the level and the pattern of the MES-2/3/6 mark H3K27me3. In wild type germ cells, H3K27me3 coats all chromosomes, but is more concentrated on the X chromosomes as reported by *Fong et al.* 2002 (Figure 3-4B, arrow). In the *lin-54* mutant background, no obvious change in H3K27me3 level and localization pattern was detected (Figure 3-4B). These results provide no evidence that LIN-54 is required for MES-4 to properly associate with and methylate autosomes, or for MES-2/3/6 to concentrate H3K27me3 on X-chromosomes. However, it remains possible that this cytological approach is not sufficiently sensitive to detect subtle changes in protein distribution.

Alternatively, we asked if MES-4 or the MES-2/3/6 complex was necessary for the stability of LIN-54 protein or its autosome-enriched localization. In *mes-4(ok2326* and *bn23, M+Z-)* germline where signal for both MES-4 and H3K36me2 was

undetectable, we found that LIN-54 protein immunofluorescence signal was still present at levels comparable to wild type, and that it still localized to autosomes (Figure 3-4C(ii)). Furthermore, a normal distribution of LIN-54 was observed in *mes-2(bn11, M+Z-)* germlines where MES-2 and its mark H3K27me3 was undetectable and MES-4 and its mark H3K36me2 was shown to ectopically spread to the X-chromosome (Figure 3-4C (iii), and Figure 3-4D) (Bender et al., 2006; Fong et al., 2002). Therefore, we find no evidence that MES proteins affect LIN-54 protein stability, or act upstream to specify LIN-54 localization. The patterns of histone modifications established by MES proteins may not be directly required for LIN-54 to associate with autosomes.

We also investigated the germline localization of LIN-54 in the subsequent M-Z- generations of *mes-4(bn23)* and *mes-2(bn11)* mutants, in which germ cells are degenerating, but some have reasonable chromosome morphology and can be assayed with immunostaining. In these nuclei, active histone modifications normally enriched on autosomes (e.g. H3K12Ac) or active RNA polymerase II were shown to inappropriately accumulate on the X chromosomes (Fong et al., 2002). When we co-stained these nuclei with LIN-54 (Figure 3-4C (iv and v), red) and an active histone modification (H4K12Ac, blue), we found that LIN-54 no longer tightly associated with chromosomes, but instead showed a diffused localization, while H4K12Ac stained all chromosomes as reported previously. The same results were obtained when localization of either LIN-54 or the DRM component LIN-9 were examined in various *mes* (m-z-) germlines (*mes-4(bn73, bn67), mes-3(bn35), mes-6(bn66),* data not shown). It is not clear how to interpret this observation: LIN-54 could fail to associate with chromosomes because it requires MES products to do so, or the loss of germ cell identity in the *mes* (M-Z-)

mutants may be so severe that proteins like LIN-54 that recognize germline chromosomes lose their association. In summary, at the cytological level, we found no evidence for localization dependencies between DRM and MES proteins.

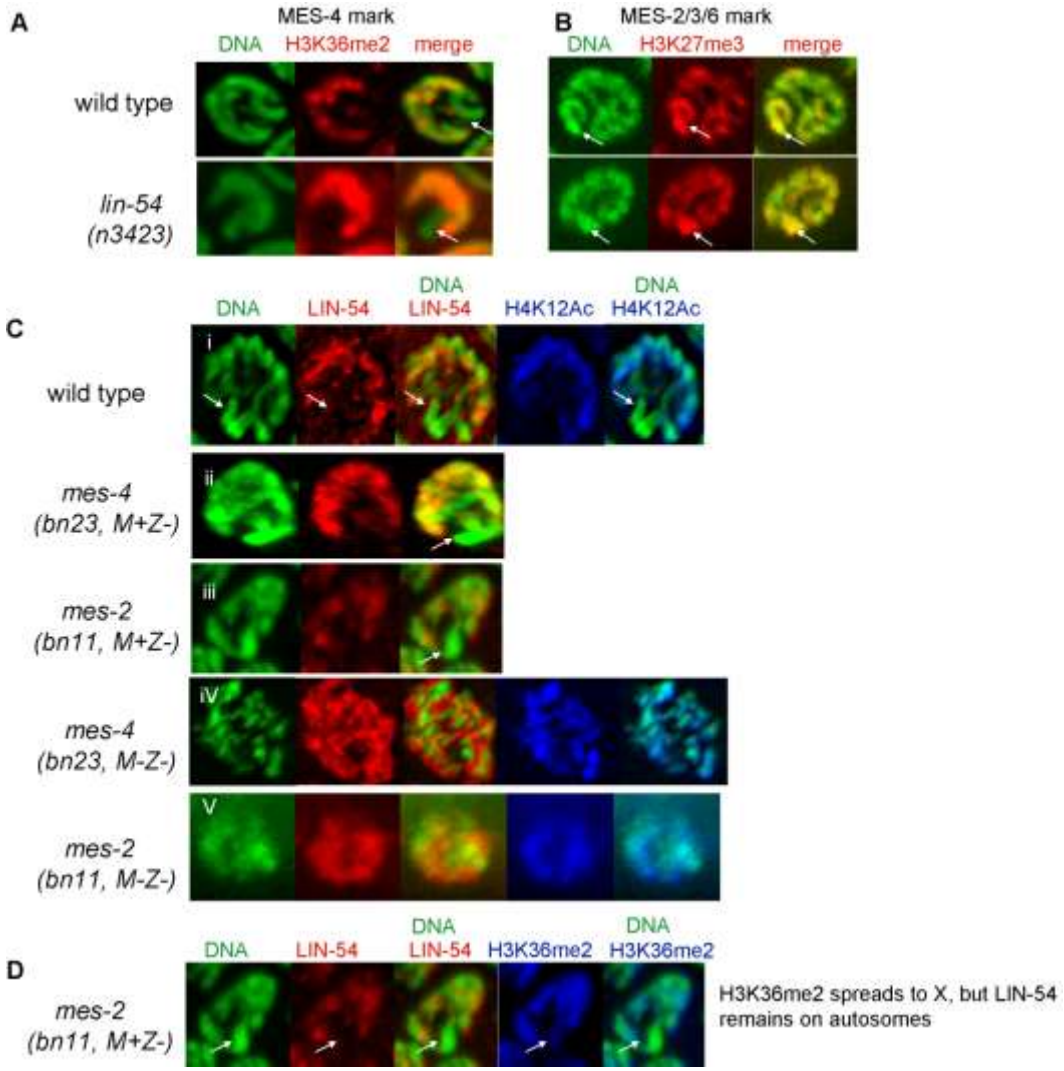


Figure 3-4. MES and LIN-54 proteins show no obvious effect on each other's localization.(A-D) Immunofluorescence images of germline nuclei during pachytene of meiosis. (A) The MES-4 mark H3K36me2 (red) stays restricted on autosomes in absence of LIN-54. (B) In wild type germline nuclei, the MES-2 mark H3K27me3 (red) coats all chromosomes, but is concentrated on the X chromosomes (arrow). A similar staining pattern was observed in nuclei from the *lin-54(n3423)* mutant (bottom). (C) LIN-54 localization in *mes* mutants. In nuclei of *mes-4(bn23)* or *mes-2(bn11)* (M+Z-) mutant germlines, LIN-54(red) remains concentrated on autosomes. In the subsequent generation M-Z- germlines, LIN-54 dissociates from chromosomes, while the active histone modification (H4K12Ac, blue), which normally marks autosomes in wild type, spreads to all chromosomes. (D) Co-localization of LIN-54 (red) and the MES-4 mark H3K36me2 (blue) in *mes-2(bn11)* M+Z- germ cells. While H3K36me2 spreads to X (arrow), LIN-54 remains on autosomes.

Mutation of lin-54 *does not disrupt the distribution of autosome-enriched or X-enriched histone modifications.*

A number of histone modifications are known to differentially accumulate on autosomes or X chromosomes in the *C. elegans* germline. We hypothesized that LIN-54 may serve to establish a chromatin distinction between X chromosomes and autosomes by acting upstream of histone modifying enzymes. To test this hypothesis, we performed immunostaining in *lin-54(n3423)* mutant germlines to monitor a variety of histone marks known to be differentially enriched between X chromosomes and autosomes. These include ten autosome-enriched histone modifications representing active transcription and two heterochromatic histone marks enriched on X chromosomes (a list of antibodies is provided in Materials and Methods). None of these histone modifications showed a detectable change in the *lin-54* mutant, indicating that LIN-54 is not necessary for the tested histone modifications to achieve differential localization between autosomes and Xchromosomes (Figure 3-5). This observation suggests that LIN-54 may function downstream or independent of the assayed histone modifications at the cytological level.

Next we asked if the autosome-enriched pattern of LIN-54 association required a variety of proteins known to show differential localization and/or function between autosomes and the X chromosome. We surveyed LIN-54 localization in the absence of 1) known autosome-associated proteins, such as XND-1 (a autosome-associated protein that promotes recombination events on Xchromosomes) and MRG-1 (a chromodomain-containing protein also involved in X-silencing, 2) MSCI-related factors

involved in the accumulation or distribution of heterochromatic histone modification (H3K9me) on Xchromosomes (such as ELK-1, EGO-1, RHA-1), or 3) proteins involved in somatic dosage compensation (DPY-30, DPY-27). In all cases, no change in LIN-54 localization pattern was observed (data not shown). Therefore, the proteins tested above do not appear to recruit LIN-54 to autosomes.

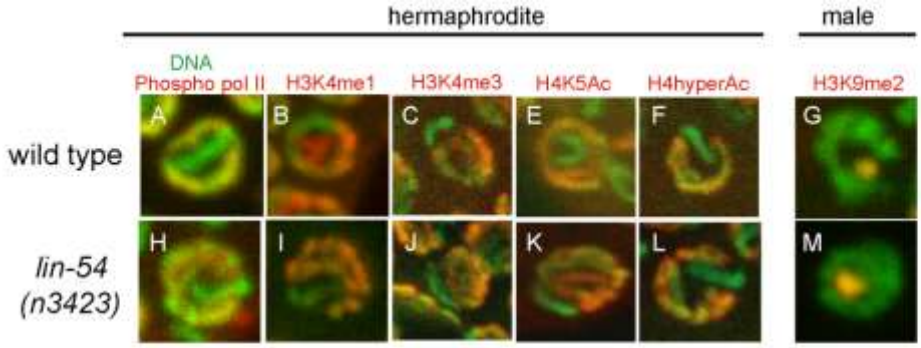


Figure 3-5. Mutation of *lin-54* does not disrupt the distribution of many autosome-enriched or X-enriched histone modifications.

Wild type (A-G) or *lin-54(n3423)* (H-M) germline pachytene nuclei (DNA, green) were stained with antibodies against an active form of pol II and histone modifications (red, merge in yellow). Phosphorylated pol II, H3K4me1, H3K4me3, H4K5Ac, H4hyperAc, and others described in the text are enriched on hermaphrodite germline autosomes in both wild type and *lin-54(n3423)*. A histone modification associated with inactive heterochromatin (H3K9me2) is enriched on the X in male germlines in both wild type and *lin-54(n3423)*.

LIN-54 is required for H3K36me2 to spreads to Xchromosomes in **mes-2** *mutant*

Since X-linked genes mis-regulated in either in *lin-54* or *mes-4* mutant germlines become expressed at more normal levels when both mutations are combined, we investigated if the changes in gene expression in the double mutant reflected alterations to the pattern of histone modifications created by the MES proteins. To do this, we examined the pattern of H3K27me3 histone modifications created by MES-2/3/6 in the *lin-54;mes-4* double mutant (M+Z-) germ line. In this germline, we detected no signals for the MES-4 mark H3K36me2 as expected, and H3K27me3 maintained its enrichment on the X chromosome (Figure 3-6E and 6K, note: H3K27me3 pattern was unaffected in *mes-4*, *lin-54*, and *mes-4;lin-54*). This indicates that the H3K27me3 pattern is not visibly altered even when the expression of X-linked genes is changed as with mutation of *lin-54* or *mes-4* or both.

MES-2 functions similarly to MES-4 in germline development and X gene regulation, and its loss of function is predicted to cause de-silencing of X chromosomes. It has previously been shown that loss of *mes-2* function causes loss of H3K27me3 in the germline, and the ectopic spread of H3K36me2 and other active modifications to the X chromosome (Bender et al., 2006; Fong et al., 2002) (Figure3-6E,K,N). To examine how the removal of both MES-2 and LIN-54 may affect the localization of MES-4 and its mark H3K36me2, we monitored the immunostaining pattern in the *mes-2;lin-54* double mutant germline. Strikingly, in these nuclei we found H3K36me2 remained restricted on autosomes in both pachytene and diakinesis nuclei (Figure 3-6F, L, O). This suggests

that the lack of H3K27me3 is not sufficient to allow H3K36me2 to spread to X in the double mutant, but it requires wild type LIN-54 activity.

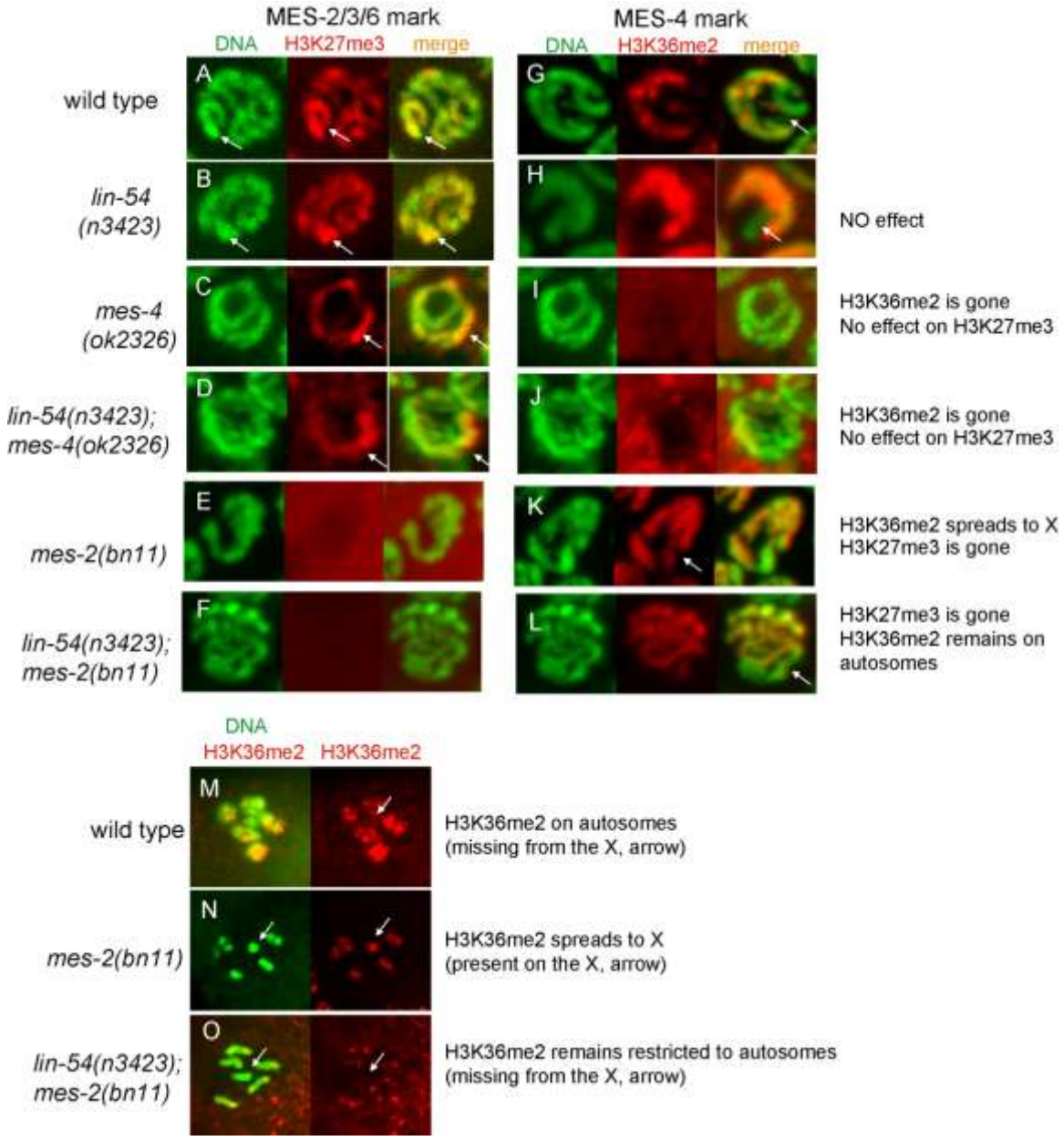


Figure 3-6. LIN-54 is required for H3K36me2 to spread to the X chromosome in a *mes-2* mutant.

Figure 3-6. LIN-54 is required for H3K36me2 to spread to the X chromosome in a *mes-2* mutant.

(A-F) Immunofluorescence images of H3K27me3 (red) and DNA staining (green) of pachytene nuclei. The distribution of the H3K27me3 mark made by MES-2/3/6 does not change in absence of LIN-54(B), MES-4(C), or both (D), and H3K27me3 is undetectable in *mes-2(bn11*, M+Z- germline) (E) or *mes-2(bn11);lin-54(n3423)* double mutant (F). (G-L) Immunostaining of H3K36me2 (red) and DNA staining (green) in pachytene stage of germ cells. (M-O) H3K36me2 pattern during diakinesis stage of germ cell meiosis. The MES-4 mark H3K36me2 is restricted on autosomes in wild type (G, M) and *lin-54(n3423)* (H), and disappears in *mes-4(ok2324)* mutant backgrounds (I, J). H3K36me2 spreads to the X chromosomes in *mes-2(bn11, M+Z-)* germlines (K, N). In contrast, H3K36me2 remains restricted on autosomes in the *mes-2;lin-54* double mutant (L,O). Arrows indicate the X chromosomes.

Genetic interaction between LIN-54 and MES-4

Both *lin-54* and *mes-4* mutations result in pleiotropic defects, and thus it is difficult to pin-point a phenotype directly associated with aberrant X-linked gene expression in the germline. We predict that, if a particular defect (phenotype) is linked with the over- or under-expression of X-linked genes, then the penetrance or severity of the defect should be alleviated in the double mutant, where the X-linked gene expression is re-balanced.

To identify such a phenotype, we characterized the phenotypes of the single *lin-54(n3423)* and *mes-4(ok2326)* mutants, and compared them to the *lin-54(n3423);mes-4(ok2326)* double mutant. First, we documented the phenotypes of the single *lin-54(n3423)* mutant: *lin-54(n3423)* homozygous progeny from heterozygous mothers (M+Z-) grow up to be sterile adults carrying many endomitotic oocytes (EMO phenotype) along with other defects, such as deformed vulva, twisted-gonads, and formation of ectopic sperm (Harrison et al., 2006; Tabuchi et al., 2011; Thomas et al., 2003). The penetrance of EMO was 95% and the sterility was almost complete, although a young mother occasionally laid 1-4 embryos that could become very sick-looking sterile adults (Table 3-1). Second, we confirmed that the *mes-4(ok2326)* mutant shows the maternal effect sterility as other alleles of *mes* mutants. Third, when we combined those two mutations, we found the *lin-54(n3423);mes-4(ok2326)* double mutant from heterozygous mothers showed 95% EMO penetrance, and 100% sterility with no escapers at the M+Z- generation. Thus, the double mutant appears to show a phenotype like the *lin-54* single mutant. This suggests *lin-54* is epistatic to *mes-4* in terms of the EMO penetrance and the sterility. Since we found no escapers from the

double mutant line, another possible interpretation is that the *mes-4* mutation may act synthetically to enhance the sterility of *lin-54*.

To extend this analysis, we also made four other double mutants carrying alleles of *mes-4(bn23, bn67 alleles)*, or a *mes-2(bn11)* allele in the *lin-54(n3423)* mutant background (Table 3-1). The information on each allele is described in Supplemental Table 3-1. All double mutants, except for the *lin-54(n3423);mes-4(bn23)*, showed a similar genetic interactions as in *lin-54(n3423);mes-4(ok2326):* 90-95% EMO penetrance and M+Z- generation sterility (Table 3-1). The exception, the *lin-54(n3423);mes-4(bn23)* double mutant showed a decrease in the EMO penetrance from 96 % to 2%, and a slight increase in fertility. *mes-4(bn23)* is a point mutation predicated to affect protein stability (Bender et al., 2006), see Materials and Methods, Supplemental Table 3-1) and, consistently immunofluorescence detected no protein nor the histone modification it creates, suggesting that *mes-4(bn23)* is a protein null allele like *mes-4(ok2326)*. One possibility of why we detect such allele specificity is that the *bn23* allele is linked with a genetic marker *dpy-11(e224)* and *unc-76(e911)* in *cis*, and thus the puzzling result might be related to a background genetic component, rather than the *mes-4(bn23)* mutation. RNAi knock-down of MES-4 in the *lin-54(n3423)* mutant background also showed a phenotype similar to that observed with *lin-54(n3423);mes-4(ok2326),* in terms of the sterility and the EMO penetrance. Further experimentation is required to elucidate why we see such allele specificity, and a phenotype directly associated with aberrant X-linked gene expression in the germline.

Genetic interaction between *lin-54* and *mes* mutants

	Genotype	% endomitotic oocytes (n) [a]	# of embryos laid per worm (n) [b]
mes single	Wild type	0 (53)	284 (*)
	lin-54(n3423)	96 (85)	0.4 (30)
	mes-4(ok2326)	0 (17)	114 (2)
	dpy-11(e224) ***mes-4(bn23)*** *unc-76(e911)*	0 (17)	116 (3)
	dpy-11(e224) ***mes-4(bn67)***	ND	53 (3)
	mes-4(RNAi)	0 (5)	ND
	mes-2(bn11) *unc-4(e120)*	0	152 (6)
double	***mes-4(ok2326)****;lin-54(n3423)*	95 (60)	0 (30)
	dpy-11(e224) ***mes-4(bn23)*** *unc-76(e911) ; lin-54(n3423)*	**2 (43)**	0.6 (30)
	dpy-11(e224) ***mes-4(bn67)*** *; lin-54(n3423)*	ND	0.1 (30)
	mes-4(RNAi)*;lin-54(n3423)*	100 (27)	ND
	mes-2(bn11) *unc-4(e120) ; lin-54(n3423)*	ND	0 (30)

[a] The percent of animals with endomitotic oocytes at M+Z- generation
[b] The average brood size per worm at M+Z- generation.
* *Fay et al. (2002)*

Table 3-1

Chapter III. Discussion

In summary, this study provides a direct link at gene expression level between the previously identified gene regulatory properties of two proteins implicated in hermaphrodite X-silencing: LIN-54(DRM) and MES-4. We showed LIN-54 and MES-4 antagonize each other to reciprocally and equally alter gene expression of their overlapping targets, the majority of which are located on X chromosomes, but some are located on autosomes. In an effort to understand the molecular mechanism by which these proteins counteract each other, we systematically investigated a number of regulatory possibilities and consequently found that LIN-54 and MES-4 do not influence each other's gene expression, protein stability, or localization patterns, and the loss of LIN-54 does not alter the pattern of histone marks known to be differentially enriched between X and autosomes. Further, we provide a potential mechanistic clue as to how LIN-54 and MES-4 counteract each other by showing that *lin-54* mutation suppresses the inappropriate appearance of MES-4 and its mark, H3K36me2, on the X chromosomes in the *mes-2* mutant background. This suggests the potential involvement of LIN-54 in H3K36me2 localization. We propose the antagonistic interplay between DRM complex and MES proteins is an essential component to maintain the proper dose of X-linked vs. autosomal gene expression in *C.elegans* germline.

How do autosomal-enriched LIN-54 and MES-4 influence the expression of genes on the X chromosomes? (MODEL)

In this study, we identified many X-linked genes that change expression significantly in response to both *lin-54* and *mes-4* mutations. Given these genes are rarely bound by LIN-54 or MES-4, their expression is likely to be controlled by an indirect mechanism. One potential mechanism is that LIN-54 and MES-4 may counteract each other to regulate transcription of one or more autosomal gene(s) whose gene products may act *in trans* to control X chromosome gene expression. However, after examining the list of genes bound or/and regulated by either LIN-54 or MES-4, no obvious candidate was found.

As an alternative, we propose a model that combines the previously proposed models to explain how autosomal MES-4 or LIN-54 influences expression of genes located on the X chromosomes. According to the MES-4 model proposed by *Bender et al.,* MES-4 functions to repel the global repressors from autosomes to concentrate them to the X chromosomes (Bender et al., 2006). When the MES-4 function is disrupted, the global repressors may inappropriately spread from the X chromosome to autosomes. Assuming the global repressor is limiting, autosomal genes would decrease in expression while X-linked genes would increase in expression, which is in agreement with what we observed. In the LIN-54 model, LIN-54 may function to retain a portion of the global repressors to autosomes. When LIN-54 is absent, the global repressors might ectopically spread from autosomal loci to the X chromosomes. In which case, autosomal genes would increase in expression whereas X-linked genes that escape X-silencing now become down-regulated (the enhancement of X-silencing). Taking these two models into account, we entertain the possibility that LIN-54 and MES-4 may act on

the very same repressor (we refer to as “R”) to regulate its distribution across the genome. In other words, the antagonistic interplay between LIN-54 and MES-4 may be necessary to maintain the intricate balance of the global repressor allocation across the genome. Perhaps, this distribution of such repressors might be a key component to maintain the proper germline gene expression program, where X chromosome gene expression is adequately silenced.

The identity of the mysterious global repressor ”R” is currently not known. One logical candidate is MES-2/3/6 (PRC2) repressor complex; however, cytological experiments suggest this is less likely since the H3K27me pattern created by the MES-2/3/6 complex was unaffected upon the loss of LIN-54, MES-4, or both (this study and Bender et al., 2006). That being said, we cannot rule out a possibility that the potential change in H3K27me level may occur locally on a gene-by-gene basis, which may be difficult to detect at the cytological level using immunostaining. The use of more sensitive methods, such as ChIP, may be necessary to detect potential changes in local chromatin modifications. In fact, the unpublished ChIP data from the Strome laboratory shows that the H3K27me3 signals on the X chromosomes spreads to new regions on autosomes when MES-4 was depleted by RNAi (personal communication). In addition, we found that the MES-4 mark H3K36me stays restricted on autosomes in the *mes-2;lin-54* double mutant background. Together, these data support the involvement of the MES-2/3/6 complex in the process of DRM and MES-4 mediated hermaphrodite X chromosomes silencing.

If the MES-2/3/6 complex acts as global repressors, then we do not need to invoke the third party player to explain the phenomenon, and instead all can be

explained by the interplay among the three known players: DRM, MES-4, and the MES-2/3/6 complex. In agreement with this, the intriguing link between DREAM and PRC2 complex in gene regulation that was recently shown: dREAM complex represses the developmentally-regulated genes by employing a histone deacetylase (HDAC)-mediated hypo-acetylation at gene promoters, and PRC2-mediated accumulation of H3K27me2 at downstream of the promoter region (Lee et al., 2010). The loss of dREAM function lowers the H3K27me level at it's target genes, and the loss of PRC2-mediated H3K27me de-represses the target genes. Together, we speculate that DRM and the MES-2/3/6 (PRC2) complex may function together to repress transcription of developmentally-regulated genes. Potentially, promoter-bound DRM (Figure 3-7, blue blob) may anchor the MES-2/3/6 to facilitate H3K27 methylation (green blob) on the same target genes ("DRM-repressed genes", Figure 3-7). Furthermore, there is strong evidence on the antagonistic relationship between H3K36me and H3K27me. The repressive mark H3K27me exhibits self-propagating spreading along chromosomes, and therefore the activity of "anti-repressors" is needed to restrain repressive chromatin (Margueron et al., 2009). Curiously, recent work showed H3K36me and H3K27me do not co-exist on the same histones, and H3K36me antagonizes methylation of H3K27me (Yuan et al., 2011), suggesting H3K36me plays an "anti-repressive" role in restricting the distribution of H3K27me. At the other end of the spectrum, the loss of the X-enriched mark H3K27me allows H3K36me to spread to the X chromosomes in *C.elegans* germline (Fong et al., 2002), indicating that H3K36me may also antagonize H3K27me. Taken together, both H3K36me and H3K27me may repel each other to maintain the proper balance between activating and repressive chromatin environments.

Considering these two links, we hypothesize that co-occupancy of DRM (at promoter) and PRC2 (downstream) may be necessary to repel MES-4-mediated H3K36me to maintain repressed states of target genes. Below we discuss this model in detail.

Assuming the global repressor is the MES-2/3/6 (PRC2) complex, we propose a working model in Figure 3-7. In wild type (1): promoter-bound DRM (green blob) anchors H3K27me (blue blob) to repress some autosomal genes ("DRM-repressed genes").The active autosomal genes ("germline-expressed genes") are decorated with the MES-4 mark H3K36me on their gene bodies (yellow blob). H3K36me and H3K27me are restricted to the corresponding gene sets by their intrinsic repelling nature. In contrast, the X chromosomes accumulate a higher concentration of H3K27me and undergo mild silencing. In DRM mutant (2): without the anchoring effect of DRM, H3K27me is released from the DRM-repressed genes and spreads primarily to the X chromosomes. The vacancy on the DRM-repressed genes permits H3K36me to invade and de-repressed these genes. The excessive amount of H3K27me on X causes an enhancement of X-silencing. In *mes-4* mutant (3): the absence of the MES-4 mark H3K36me on autosomal "germlien-expressed genes" permits H3K27me to spread from X, and cause down-regulation of "germline-expressed genes". Concurrently, X-linked genes lose some H3K27me and become de-silenced. In *mes-4; lin-54* double mutant (4): "DRM-repressed genes" release H3K27me, and the vacant "germline-expressed genes" absorb H3K27me either from "DRM-repressed genes" or from the X chromosomes. The X chromosomes gain some H3K27me from "DRM-repressed genes" or lose some to "germline-expressed genes", making the net H3K27me level close to wild type levels. Ultimately, X-linked and autosomal genes show wild type levels of gene

expression. In *mes-2* mutant (5): although H3K27me is absent at ”DRM-repressed genes”, DRM still repels the spreading of H3K36me resulting in no change in ”DRM-repressed gene expression). The loss of H3K27me, especially from X, allows the spreading of H3K36me to the X chromosomes and causes de-silencing of X-linked gene. In *mes-2;lin-54* mutant (6): ”DRM-repressed genes” lack both H3K27me and DRM, thus this vacancy allows H3K36me to spread *in cis* (within autosomes), leading to up-regulation of “DRM-repressed genes”. H3K36me stays restricted on autosomes, because autosomal vacancy acts as quencher. In conclusion, we think that X chromosome silencing may come down to a genome-wide balance between H3K36me and H3K27me, and DRM may facilitate such balance.

LIN-54 and MES-4 responsive X-linked genes

It is reasonable to speculate that LIN-54 and MES-4 responsive X-linked genes have specific features that allow them to be bound and modulated by these two DNA associated proteins. Thus far we have not identified any specific gene ontology associated with these genes as they are largely non-essential and ubiquitously expressed with diverse functions. Additionally, they are dispersed throughout X chromosomes, but not likely in the specialized regions such as the paring centers or the telomeres. It is possible that these genes display special chromatin environments, such as an “open” landscape of promoter sequences or chromatin state, thereby making them highly sensitive to the concentration of the global repressor. From modeENCODE, genome-wide landscaping of various histone modifications and transcription factors are

available. By using this information, mapping of the chromatin environments around these genes may provide important insight on what is responsible for recruiting these proteins to their binding sites.

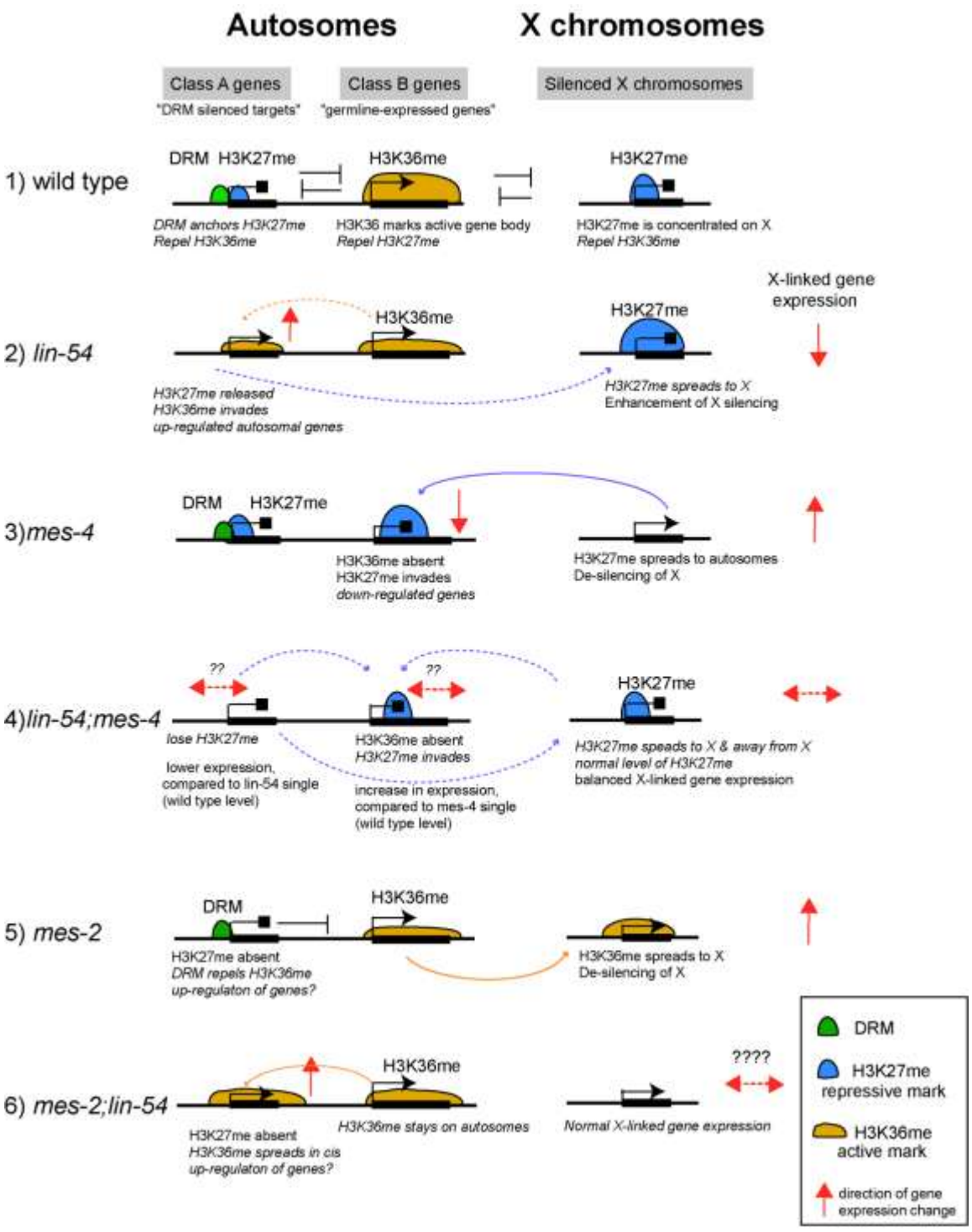


Figure 3-7. Summary and Model

Figure 3-7. Summary and Model

Cartoons illustrate predicted accumulation of DRM (at gene promoters, green blob), H3K27me (blue blob, downstream of promoters), and H3K36me (yellow blob, over coding regions). Autosomal genes are represented on the left and X-linked genes on the right. Two types of autosomal gene sets are show: “DRM target genes” that are repressed by DRM and presumed to also be repressed by H3K27me and “germline-expressed genes” that are transcriptionally active and decorated with H3K36me in their coding regions (MES-4 targets). (1) wild type: DRM anchors H3K27me to repress autosomal “DRM targets”, and MES-4 decorates autosomal “germline-expressed genes” with H3K36me for activation. X-linked genes are enriched with H3K27me and undergo mild X chromosome silencing. (2) DRM mutant: H3K27me is released from “DRM target genes”” (idea based in part on Lee et al.) and spreads to X chromosomes, leading to the enhancement of X-silencing. The vacancy on “DRM targets” would allow H3K36me to invade, and thereby become de-repressed. (3) *mes-4* mutant: the loss of H3K36me on “germline-expressed genes” permits H3K27me, primarily from the X chromosomes, to invade and become repressed. Consequently, the dilution of H3K27me on X chromosome causes the de-silencing of X chromosomes. (4) *mes-4; lin-54 double* mutant: ”DRM targets” release H3K27me, and the vacant “germline-expressed genes” absorb H3K27me either from ”DRM targets” or from the X chromosomes. The X chromosomes gain some H3K27me from “DRM targets” or lose some to “germline-expressed genes”, making the net H3K27me level close to wild type levels. Ultimately, X-linked and autosomal genes show nearly wild type levels of gene expression. (5) *mes-2* mutant: although H3K27me is absent at ”DRM target genes”, DRM still repels the spreading of H3K36me resulting in no change in ”DRM target gene expression. The loss of H3K27me, especially from X, allows the spreading of H3K36me to the X chromosomes and causes de-silencing of X-linked genes. (6) *mes-2;lin-54* mutant: “DRM targets” lack both H3K27me and DRM, thus this vacancy allows H3K36me to spread *in cis* (within autosomes), leading to up-regulation of ”germline-expressed genes”. H3K36me stays restricted on autosomes, because autosomal vacancy acts as quencher.

Chapter III. Materials and Methods

C. elegans *strains and culture conditions*

All strains were cultured at 20^{o}C unless otherwise noted, using standard methods (1974 Brenner). The following strains were used: N2 (Bristol) as wild-type, *lin-54(n3423)/nT1 [qIS51], and lin-54(n2990)/nT1[qIS51]* [Harrison06, Thomas, Tabuchi], VC1874 *mes-4(ok2326)*V/*nT1[qIs51](*IV;V),SS268 *dpy-11(e224)mes-4(bn23)unc-76(e911)*V/*nT1[unc-?(n754) let-?]*(IV;V), JK2663 *dpy-11(e224) mes-4(bn67)* V/*nT1[unc-?(n754) let-?qIs50]*(IV;V), SS186 *mes-2(bn11) unc-4(e120)/mnC1 dpy-10e128() unc-52(e444)*II, SS836 *mes-4(bn73)/DnT1*, *mes-3(bn35)/hT2,* SS360 *mes-6(bn66) dpy-20(e1282)*IV/*nT1[unc-?(n754) let-?]*(IV; V), PK125 *tra-2(e2531)* II, RB868 *xnd-1(ok708)*, RB869 *xnd-1(ok709)*, XA6227 *mrg-1(tm1227)/qC1 dpy-19(e1259) glp-1(q339)[qIs26]* III. EL391 *ego-1(om84) unc-29(e193)/hT2 [dpy-18(h662)]* I; *+/hT2 [bli-4(e937)]* III, KMW1 *rha-1(tm329)* II, DR1410 *dpy-27(y56)/qC1 dpy-19(e1259) glp-1(q339)* III, TY1936 dpy-*30(y228) V/nT1 [unc-?(n754) let-?]*(IV; V), VC3150 ekl-*1(ok1197) I/hT2 [bli-4(e937) let-?(q782) qIs48]*(I;III).

NOTE: *mes-4(bn23)* mutation occurs at the 3' end of intron 7, changing the splice acceptor dinuleotides (AG_{3130}→AA) essential for the splicing between exon 7 and exon 8. Coincidently, the next nucleotide (first nucleotide in exon 8) is G, thus *bn23* creates *de novo* false splice site, $A\mathbf{A}_{3130}\mathbf{G}$**.** This causes the removal of the first nucleotide of exon 8, and gives a rise to a nucleotide excision and a flame shift in the protein. The unstable mRNA and/or proteins are expected.

Phenotypic Analysis of XO Animals

XO males*:* Homozygous *lin-54(n2990, M-Z-)* males were created by crossing the homozygous *lin-54(n2990, M+Z-)* hermaphrodites with heterozygous *lin-54(n2990)/nT1* males. The expected ratio of heterozygous hermaphrodites: heterozygous males: homozygous hermaphrodites: homozygous males is 1:1:1:1. XO females: The homozygous *tra-2(e2531)* hermaphrodites were mated with wild type males on IPTG + ampicillin mating plates, seeded with RNAi feeding clones (empty vector and *lin-54* clone from Ahringer library). The *tra-2(e2531)/*+ heterozygous cross-progeny (XX hermaphrodites and XO pseudo-females) depleted with empty vector or *lin-54* was examined. The pictures of worms were taken at 24-48 hours after L4 stages with the microscope at Ceol laboratory. To immobilize worms, worms were mounted on the agar plates speeded with 100mM levamisole. The body length of worms were measured by Image J (n=27, two independent experiments).

Antibody and Immunofluorescence

Germlines were fixed and stained with antibodies essentially as described in (Tabuchi et al., 2011). Primary antibodies to DRM subunits were described and validated in (Ceol and Horvitz, 2001; Harrison et al., 2006; Page et al., 2001; Tabuchi et al., 2011). The commercially available antibodies against modified histones used in this study as follow: H3K4me1 (Upstate), H3K4me2 (#07-030, Cell Signaling #9725), H3K4me3(Abcam #8580-100), H3K9me2(Update #07-212, Cell Signaling #9753), H3K9me3(Abcam #8898), H3K27me3(Lake Pacid Biologicals #AM0174), H3K36me2(Cell Signaling

#9758), H4K5Ac(Serotec AHP414), H4K8Ac(Serotec AHP415), H4K12Ac(Serotec AHP416), H4K16Ac(Abcam), H4hyperAc(Milpore #06-866). Images were acquired as described in (Csankovszki et al., 2009; Tabuchi et al., 2011).

The probe preparation and microarray analysis

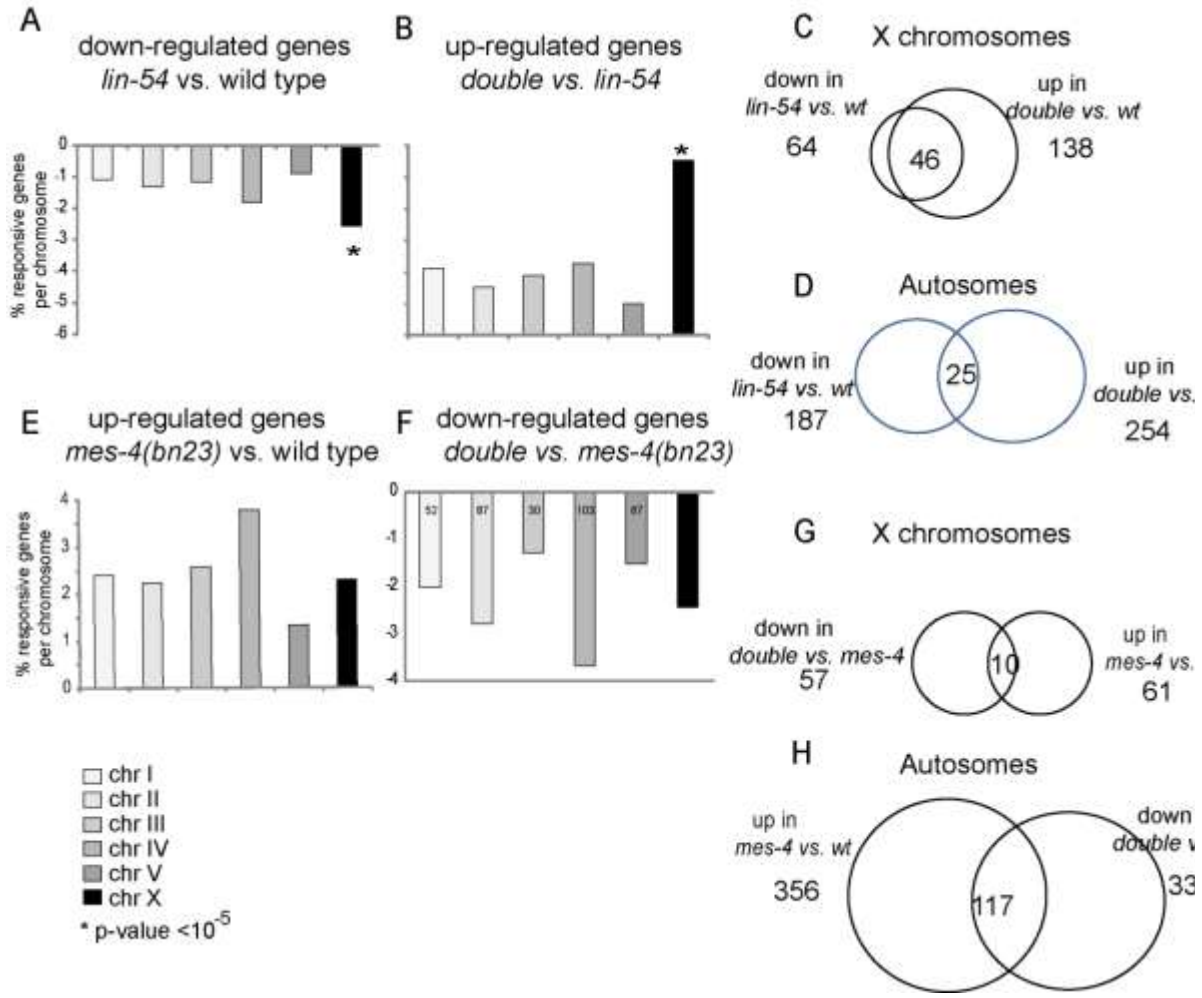
Total RNA was isolated from the 50 ~70 dissected gonads from wild type, *lin-54(n3423)*, and *mes-4(ok2326, M+Z-)*, and *lin-54(n3423);mes-4(ok2326)* (M+Z-) mutants as described in Chapter II. Total RNA was straightly labeled with biotinated UTP (without a linear amplification step) using MessageAmp II aRNA Amplification Kit (Ambion). Probe-preparation, hybridization, and scanning for DNA microarray were performed at the Genomics Core facility at University of Massachusetts Medical School. Fluorescence-labeled cDNA probes were prepared using the One-Cycle kit (Affymetrix) and the Enzo HighYield RNA Transcript Labeling Kit(Enzo) for embryo, and the 3' IVT Express Kit (Affymetrix) for germline. cDNA probes of three replicates were hybridized to GeneChip *C. elegans* genome arrays (Affymetrix).

Chapter III. Acknowledgements

I thank Paul Furcinitti at the Digital Imaging Core Facility and Phyllis Spatrick at the Genomics Core Facility of the University of Massachusetts Medical School for their technical supports. I thank Lihua J. Zhu Ph.D., for microarray analysis and for Bob Horvitz Ph.D., for providing me with valuable DRM antibodies and LIN-54 strains, and

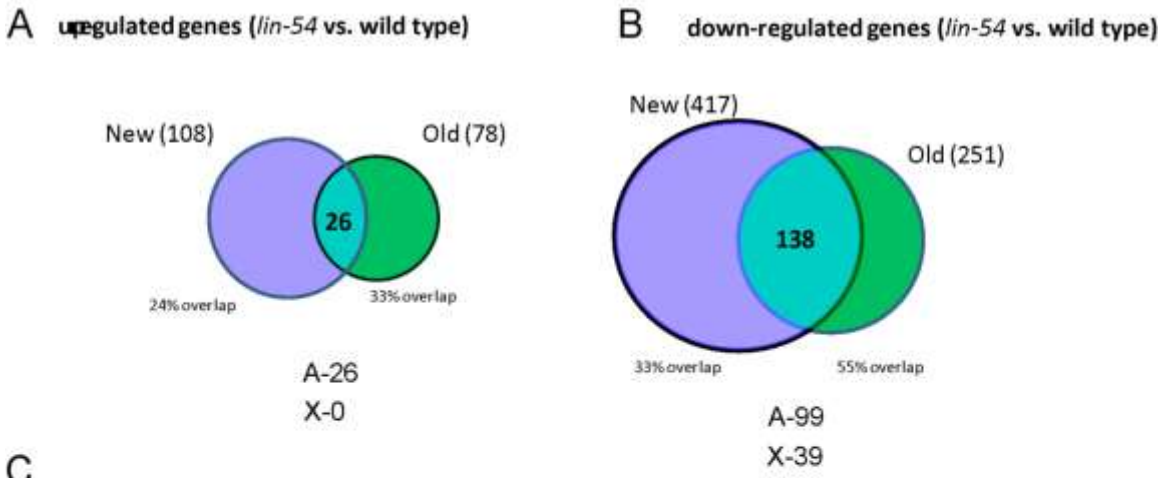
for Susan Strome for *mes-4(bn73)* strain and MES-4 antibody. Dekker, Kaufman, Kuroda, and Stein laboratories provided us with an aliquot of commercial histone antibodies. Strains were provided by the *C. elegans* Gene Knockout Consortium and *C. elegans* Genetic Center. I thank members of Hagstrom laboratory and members of Biotech 4 cell dynamic groups for valuable discussion. I thank Dr. Kirsten Hagstrom and Dr. Trevor J. Morin for the helpful comments and edits on this chapter.

0



Supplemental Figure 3-1

Comparison of genes mis-regulated in the *lin-54(n3423)* mutant and in the *mes-4(bn23)* mutant *[dpy-11(e224) mes-4(bn23) unc-76(e911)]*



C

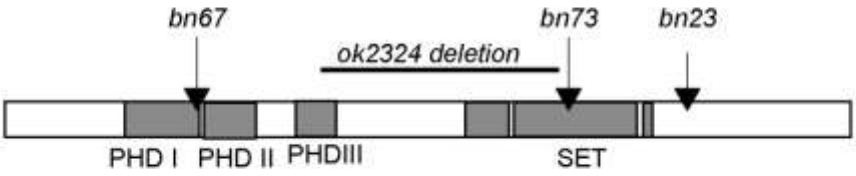
New (2011)

up	down
108	417
19% (21/108)	10% (40/417)

19%		10%	
A	X	A	X
19% (20/103)	2% (1/5)	10% (30/301)	9% (10/116)

Supplemental Figure 3-2. Comparison of LIN-54 responsive genes identified in the previous study (chapter II) and this study.

Overlap of up-regulated genes (A) and down-regulated genes (B) in *lin-54(n3423)* vs. wild type in the previous study (green) and this work (purple). (C) The overlap of LIN-54 responsive genes (this study) with LIN-54 bound genes (LIN-54 ChIP-chip data from chapter II). A different probe preparation method was used in this set of Microarray analysis (see materials and methods).

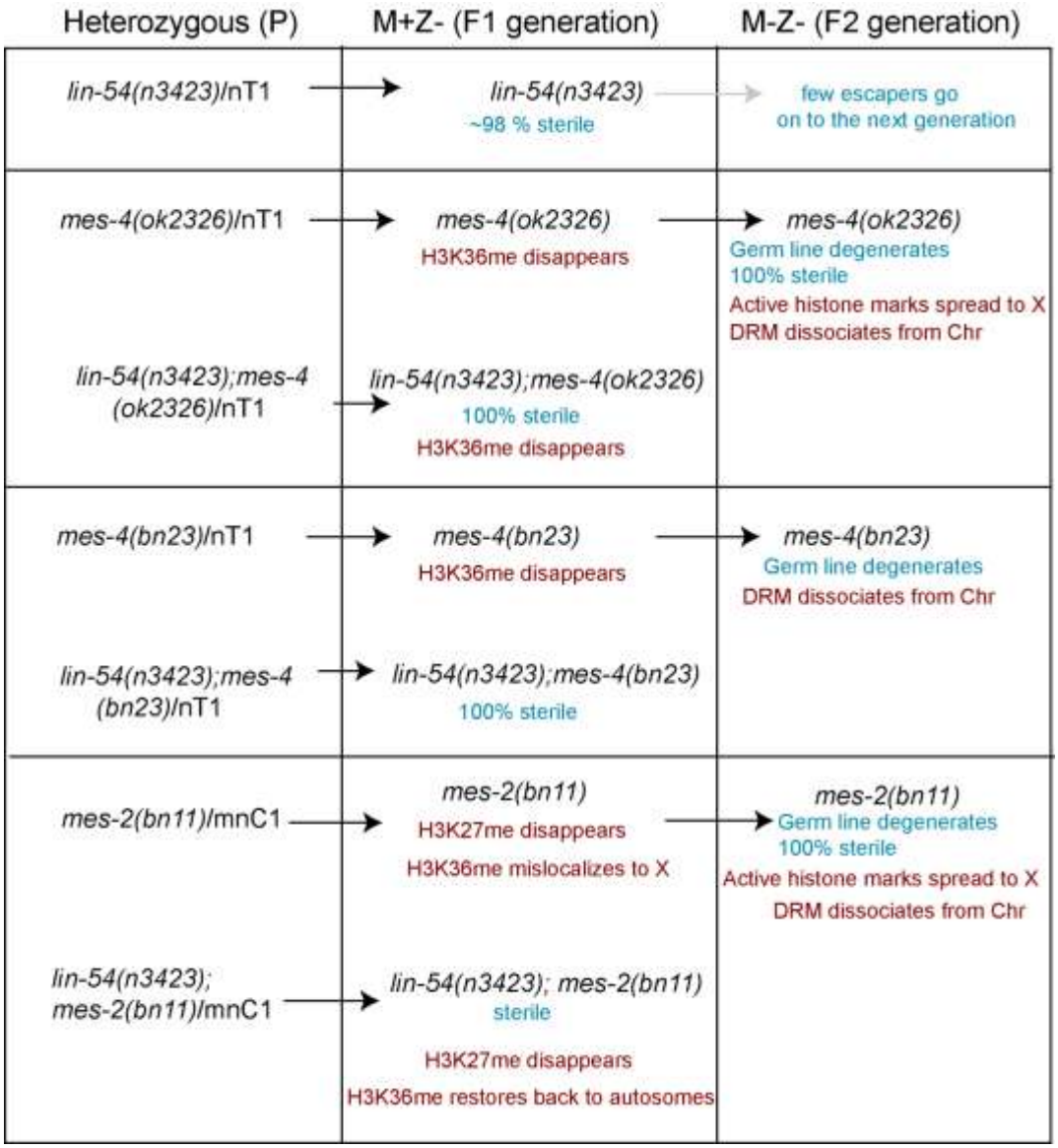


mes-4 mutations used in this study

Allele	nucleotide change	Transcript	protein	H3K36me
bn23 *	point mutaton: AG3130 ->AA	expressed*	undetectable*	undetectable*
*bn67**	point mutaton:C765AT ->TAT	expressed*	dissociate from DNA*	undetectable*
*bn73**	point mutaton:TAT2687 ->TAA	expressed*	undetectable*	undetectable*
ok2326	1596bp deletion	expressed	undetectable	undetectable

* *Bender et al.*

Supplemental Table 3-1
mes-4 mutation alleles used in this study



Supplemental Table 3-2.

Summary of the phenotypes and histone modification changes in *lin-54(n3423)*, *mes-2(bn11)* or *mes-4(ok2326, bn23)* single mutants, and the *lin-54(n3423); mes-2(bn11)* or *mes-4(ok2326, bn23)* double mutants. Complete genotypes of the strains used: *lin-54(n3423)/nT1 [qIS51];VC1874:mes-4(ok2326) V/nT1[qIs51](IV;V); SS268: dpy-11(e224) mes-4(bn23) unc-76(e911)V/nT1[unc-?(n754) let-?](IV;V); SS186: mes-2(bn11) unc-4(e120)/mnC1 dpy-10e128() unc-52(e444)II.*

Supplemental Table 3-3

The numbers of genes that significantly changed their expression in the listed mutants. (fold change >1.5, P-value <0.01). The numbers in parentheses represent the % of responsive genes per each chromosome. The strains used this study: the *lin-54(n3423)* single mutant, *the mes-4(ok2326)* single mutant, and the *lin-54(n3423);mes-4(ok2326)* double mutant.

Chromosomal distribution of genes up-regulated in the genotypes listed below

up	*lin-54* / wild type	*mes-4*/wild type	double/wild type	double/*lin-54*	double/*mes-4*
I	18 (0.71)	28 (1.1)	26 (1.02)	18 (0.71)	36 (1.41)
II	28 (0.9)	59 (1.9)	36 (1.16)	22 (0.71)	37 (1.19)
III	24 (1.03)	32 (1.37)	24 (1.03)	17 (0.73)	32 (1.37)
IV	19 (0.68)	53 (1.9)	27 (0.97)	30 (1.08)	36 (1.29)
V	14 (0.33)	55 (1.26)	25 (0.58)	22 (0.51)	19 (0.44)
X	5 (0.21)	419 (16.83)	162 (6.51)	254 (10.2)	6 (0.25)
total	108 (0.62)	646 (3.66)	300 (1.7)	363 (2.06)	166 (0.94)
A	103 (0.68)	227 (1.5)	138 (0.91)	109 (0.72)	160 (1.06)
X	5 (0.21)	419 (16.83)	162 (6.51)	254 (10.2)	6 (0.25)
A (ave)	20.6 (0.68)	45.4 (1.5)	27.6 (0.91)	21.8 (0.72)	32 (1.06)

Chromosomal distribution of genes down-regulated in the genotypes listed below

down	*lin-54* / wild type	*mes-4*/wild type	double/wild type	double/*lin-54*	double/*mes-4*
I	49 (1.92)	51 (2)	72 (2.82)	26 (1.02)	64 (2.5)
II	65 (2.09)	22 (0.71)	54 (1.74)	17 (0.55)	100 (3.21)
III	47 (2.02)	36 (1.55)	50 (2.14)	30 (1.29)	60 (2.57)
IV	68 (2.43)	28 (1)	55 (1.97)	19 (0.68)	83 (2.97)
V	72 (1.65)	27 (0.62)	59 (1.36)	18 (0.42)	101 (2.32)
X	116 (4.66)	1 (0.05)	19 (0.77)	0 (0)	248 (9.96)
total	417 (2.36)	165 (0.94)	309 (1.75)	110 (0.63)	656 (3.72)
A	301 (1.99)	164 (1.09)	290 (1.91)	110 (0.73)	408 (2.69)
X	116 (4.66)	1 (0.05)	19 (0.77)	0 (0)	248 (9.96)
A (ave)	60.2 (1.99)	32.8 (1.09)	58 (1.91)	22 (0.73)	81.6 (2.69)

LIN-54 and MES-4 responsive genes

WBGene00019082	I	pup-3
WBGene00004111	I	pqn-21
WBGene00007297	I	C04F12.1
WBGene00008546	II	F07A11.2
WBGene00019511	II	K08A2.1
WBGene00003822	II	nsy-1
WBGene00015912	II	fbxc-50
WBGene00015111	II	B0281.4
WBGene00001565	III	gei-8
WBGene00020050	III	R13A5.7
WBGene00014031	III	ZK637.14
WBGene00003932	IV	pat-6
WBGene00018073	IV	F35H10.10
WBGene00017990	IV	F32E10.2
WBGene00004928	V	soc-1
WBGene00009469	V	F36D3.1
WBGene00006602	X	tps-1
WBGene00017177	X	F02E8.4
WBGene00018701	X	pccb-1
WBGene00003410	X	mrp-4
WBGene00007666	X	C18B12.4
WBGene00000105	X	alg-1
WBGene00011153	X	R09A8.2
WBGene00007290	X	C04B4.2
WBGene00018593	X	F48B9.8
WBGene00019611	X	K10B3.5
WBGene00015180	X	B0416.4
WBGene00011011	X	R04D3.3
WBGene00012326	X	W07E11.1
WBGene00010690	X	K08H2.3
WBGene00018657	X	acl-4
WBGene00011012	X	R04D3.4
WBGene00001130	X	dyn-1
WBGene00019989	X	R09F10.6
WBGene00008951	X	F19C6.2
WBGene00016485	X	C36C9.1
WBGene00003394	X	mom-1
WBGene00020508	X	T14F9.2
WBGene00004356	X	rhi-1
WBGene00017532	X	F16H11.3
WBGene00018578	X	F47G3.3
WBGene00008959	X	F19H6.4
WBGene00008882	X	F16B12.6
WBGene00009632	X	ttyh-1
WBGene00018735	X	F53B1.2
WBGene00000460	X	ceh-39
WBGene00015168	X	tag-320
WBGene00002007	X	hsp-3
WBGene00019871	X	R04E5.8
WBGene00008210	X	C49F5.6
WBGene00006890	X	vem-1
WBGene00016486	X	fbxa-170
WBGene00009233	X	F28H6.4
WBGene00016346	X	fbxa-120

Supplemental Table 3-4

The list of “LIN-54 and MES-4 responsive genes”: The 38 X-linked genes and 16 autosomal genes that significantly changed expression in comparisons between *lin-54* and *mes-4* single or double mutations (the genes in the Venn diagram intersections of Figure 3-2C and G, and in Figure 3-2D and H.)

Chapter IV: Perspectives and open questions

Spatiotemporal coordination of gene expression is essential for multi-cellular organisms to properly divide and differentiate during developmental processes. Such intricate gene expression is orchestrated by the concerted actions of transcription factors and chromatin-acting proteins. Chapters II and III elucidate novel functions of the *C.elegans* transcription factor LIN-54, which acts within the context of the multi-protein DRM complex. Our data demonstrated the significance of LIN-54 DNA binding activity for the proper functions of the DRM complex in regulating genes involved in cell cycle, reproduction, and development. Further, we found the DRM complex controls distinct sets of targets and its regulatory activities differ in soma and germline. The discovery that DRM is involved in an autosome-biased gene regulatory network shows possible co-evolution of germline-expressed genes and their regulatory networks. Lastly, we demonstrated a novel role of LIN-54(DRM) in hermaphrodite X silencing, which is coordinately antagonized by histone modifying MES proteins. This work defined a novel chromosome-specific aspect of DRM-mediated regulation and extended our knowledge on the conserved role DRM plays in development.

Below I discuss the broader implications of our findings, and comment on future perspective on DRM regulatory networks and X chromosome biology.

DREAM regulates the distinct sets of genes involved in development and cell cycle

Works presented in this study and in others suggest DREAM complex regulates many genes involved in development and cellular differentiation, in addition to cell cycle genes. In particular, our decision to use the intact animals undergoing active developmental processes permits us to reveal many DREAM regulatory target genes involved in a wide range of developmental processes. This confirms the notion that DREAM plays two distinct roles in 1) cell cycle-dependent regulation of cell division genes, and 2) cell cycle-independent regulation of developmental genes. The mechanism by which DREAM regulates two distinct transcriptional programs is poorly understood. However, this study of DREAM and its interplay with the chromatin-modifying proteins (MES proteins) expanded our knowledge on how transcription factors and chromatin-modifiers work together to mediate the global chromosome-wide gene regulation during development and cell cycle. Further, studies have implicated DREAM in epigenetic gene regulation and have suggested DREAM may cooperate with histone modifying proteins (Wen et al., 2008). Thus, coordinated action between DRM and histone-modifying proteins in regulating developmental genes appears to be a likely regulatory mechanism, and future studies will address this important question.

We note that tesmin/CXC domains like those in LIN-54 are also present in proteins with roles in epigenetic gene regulation, especially X chromosome dosage compensation. The MES-2 protein described above has a tesmin/CXC domain (Holdeman et al., 1998), as does its mammalian ortholog E(z)h2, which in females is recruited to the inactive X during dosage compensation (Payer and Lee, 2008). In *D.*

melanogaster, where dosage compensation occurs by up-regulating the single male X, DNA binding activity of the tesmin/CXC domain in MSL2 is required to recruit the dosage compensation complex to the male X (Fauth et al., 2010) An epigenetic silencing phenomenon in maize called paramutation also requires a tesmin/CXC domain protein (Brzeska et al., 2010). It not clear what feature of the tesmin/CXC DNA binding module may make it suited to distinguish X chromosomes from autosomes and to participate in epigenetic regulation. Perhaps tesmin domains preferentially bind chromatin with particular histone modifications or provide a platform for recruitment of proteins with histone modifying activities. In summary, the ability for DRM to recognize target DNA could be influenced by epigenetic chromatin features and also DRM may cooperate with histone modifiers and play an epigenetic role in regulating gene expression.

Additional regulatory layers on LIN-54(DRM)-mediated gene regulation

LIN-54 exhibits context-dependent binding throughout the cell cycle and within specific developmental stages. For example, LIN-54 proteins dissociate from meiotic chromosomes as they progress into the diplotene and diakinesis stages, and from mitotic chromosomes. This suggests that the DNA-binding motif alone is not sufficient to recruit LIN-54, and that the ability of LIN-54 to bind DNA is tightly regulated, especially during the cell cycle. This raises an interesting question: how is the DNA-binding activity of LIN-54 regulated?

One possible mechanism for LIN-54 regulation is cell cycle-dependent post-translational modification of LIN-54 proteins, such as phosphorylation. The *C.elegans* LIN-54 protein sequence shows several possible phosphorylation sites, and phosphorylated hLIN-54 and other DREAM components were identified in human through proteomic analysis (Litovchick et al., 2011). Furthermore, the phosphorylation of the hDREAM component LIN-52 was shown to be essential for the assembly of the DREAM complex (Litovchick et al., 2011). Thus, we consider that the similar regulatory mechanism may apply to LIN-54 protein, and that the ability of LIN-54 to bind DNA or to form a sub-complex could be modulated by cell cycle or developmental stage-specific phosphorylation events.

A second possibility is that the DNA-binding activity of LIN-54 is influenced by other DNA-binding proteins. In chapter II, we discovered the DRME2F/DP-LIN-54 binding motif, implicating that LIN-54 functions together with E2F to recruit DRM to target DNA. However, genomic studies in *D. melanogaster* identified many additional genomic regions that are bound by LIN-54 but not E2F (Georlette et al., 2007). Thus, it raises a possibility that LIN-54 may function alone or in combination with other DNA-binding protein(s). In the latter scenario, the availability of other DNA-binding protein(s) might contribute to the temporal and special binding pattern of LIN-54 at these genomic loci. This cooperative mechanism seems plausible when considering that “the MuvB core” swaps DNA-binding moieties from E2F/DP to Myb to alter target genes and its regulatory activities (activator vs. repressor) in human (Litovchick et al., 2011; Litovchick et al., 2007; Schmit et al., 2007). In *C.elegans* the counter-part for Myb has not yet been identified, although at least 11 myb-domain containing proteins are present within the

genome (www.pfam.com). In addition, *C.elegans* contains an E2F-like protein 2 (EFL-2), which is expressed in the germline. Thus, any of these proteins could act as potential LIN-54 interacting partners, and future proteomic studies will help to identify such partners.

Third, the ability of LIN-54(DRM) to efficiently recognize and associate with target loci could be influenced by local chromatin environments (e.g histone modifications). Indeed, we observed that DRM failed to associate with DNA in *mes-4* mutant *(m-z-)* germlines. Considering MES-4 plays an epigenetic role transmitting the memory of germline expression (Rechtsteiner et al., 2010), our interpretation for this result was that the DRM complex failed to recognize germline chromatin lacking underling epigenetic marks. Thus, we consider DRM may respond to chromatin states by interacting with a protein capable of “reading” epigenetic landmarks. One such candidate could be the malignant brain tumor (MBT)-repeat containing protein LIN-61, which preferentially binds to modified histones H3K9me2/3 in *C.elegans* (Koester-Eiserfunke and Fischle, 2011). LIN-61 is a sub-stoichiometric component of *C.elegans* DRM, and the fly homologue L(3)MBT is a part of the MMB complex (Harrison et al., 2006; Lewis et al., 2004). Moreover, in the *C.elegans* germline, LIN-61 shares the same autosome-biased localization pattern as DRM (data not shown). The hypothesis that DRM cooperates with LIN-61 to “read” chromatin states could be tested by monitoring the change in the enrichment of DRM at specific gene promoters upon the loss of LIN-61.

The E2F/DP-LIN-54 binding motif, and the effect of chromatin environments

Analysis of *C. elegans* promoters bound by LIN-54 identified an over-represented motif that we presume to represent an E2F/DP-LIN-54 binding site. This motif is similar to one of the eight over-represented motifs found in fly dDREAM-bound promoters (Georlette et al., 2007). In humans, this motif resembles two tandem *cis*-regulatory elements prevalently found in cell cycle-related gene promoters: the cell cycle-dependent element (CDE) and the cell cycle genes homology region (CHR) (CDE/CHR elements) (Muller and Engeland, 2010). CDE is known as a variant E2F-binding site, while the CHR-binding proteins(s) have been unidentified until very recently when an *in vitro* study showed that the tesmin domains of human hLin54 bind to the CHR sequence at the Cdc2 promoter. Taken together, we tentatively conclude that the *C.elegans* DRM motif we identified is likely to represent the combined E2F and LIN-54 binding sites, and is evolutionarily conserved among worms, flies, and humans.

Is the DRM motif necessary and sufficient to recruit DRM complex in worms? The functional necessity of the motif to recruit the DRM complex and control gene expression could be addressed by a combination of both *in vitro* and *in vivo* assays. A gel shift assay could address the sequence-specific recognition and the affinity of LIN-54 or E2F proteins to this motif *in vitro*. Complementarily, the use of wild type and mutated transgene reporters could determine the *in vivo* necessity of the motif.

Furthermore, the sufficiency of the motif can be explored by determining the effect of the X-chromatin environment on the DRM recruitment. By trans-locating the autosomal DRM binding targets (i.e. the motif-bearing promoter and gene body) to X

chromosomes, we could ask if this promoter is sufficient to recruit DRM to the X chromosome. For example, autosomal gene *egg-2* and *san-1* are the best candidates to study this question. Both genes contain LIN-54 ChIP-peaks and the DRM-binding motif at their promoters and their expression significantly changes in the absence of LIN-54 (*egg-2* expression decreases, while *san-1* expression increases in the *lin-54* mutant). By placing *egg-2* or *san-1* on the X chromosome via a transposon-mediated method, we could monitor if the promoter of these genes newly acquires LIN-54 ChIP peaks, and importantly how these genes change expression in the absence of LIN-54. Do both genes maintain the original autosomal gene expression patterns? (i.e. *egg-2* expression decreases, while *san-1* expression increase)? Or, do both genes adapt to the X chromatin environments, and follow the X chromosome specific trend (both *egg-2* and *san-1* expression decrease in *lin-54* mutant)? This experiment will address the sufficiency of the DRM-binding motif, and in addition will determine whether the promoter sequences or the promoter-associated local chromatin environments would take precedence over the global chromosome-wide effects imposed on the X chromosome.

How does MES-4 recognize autosomes?

Cytology showed that the MES-4 protein is enriched on autosomes in germline and early embryos, and unlike the canonical yeast Set2-family H3K36methytransferase (*C.elegans* MET-1), MES-4 binding to autosomes is independent of RNA Polymerase II (Bender et al., 2006; Rechtsteiner et al., 2010). A recent MES-4 ChIP-chip study

performed in embryos suggested that MES-4 and its mark H3K36me2/3 are enriched over the coding regions of genes that are predominantly expressed in the maternal germ line, and not necessarily the genes expressed in the embryonic tissue examined. Therefore, the proposed function of MES-4 in soma is to epigenetically transmit the memory of a previous germline expression program to the next generation. This would suggest autosome-enriched localization of MES-4 in germline is likely to reflect autosome-biased localization of germline-expressed genes.

How MES-4 specifically recognizes the germline-expressed genes is still not completely understood. MES-4 is less likely to recognize specific DNA elements, since MES-4 localization is influenced by the different states of chromatin, rather than the DNA sequences. This was demonstrated using transgenes identical in their sequences, but different in copy number (i.e. chromatin states). In this experiment, MES-4 associated with a complex of transgenes (“expressed” open transgenes with a few copy numbers), but not with repetitive transgenes (“silenced” closed transgenes with many copy number) (Fong et al.). It was reasoned that MES-4 is likely to be recruited to the DNA by responding to epigenetic marks on the target genes through its plant homeodomain (PHD) fingers, which recognize modified histones like its own mark H3K36me, rather than the chromatin conformation itself. Another possibility is that MES-4 may associate with another autosome-enriched protein, MRG-1, which is known to associate with H3K36me2 in fungi (Keogh et al., 2005). A telling experiment will be to temporarily induce expression of a gene in germline (e.g. by using a heat-shock promoter), and ask if MES-4 becomes enriched at this gene promoter in the embryos.

This would address whether MES-4 can recognize a temporarily activated gene in germline without the underlying epigenetic marks from the previous generation.

The model for the DRM vs. MES proteins antagonisms, and the future approach to identify "R"

In this study we placed DRM as an important player in the regulation of germline X chromosome gene expression. We showed that its action on X chromosomes is antagonistic to MES proteins, which were previously established as X chromosomes regulators. Such antagonistic interplay may ensure the proper amount and level of X-linked genes are being expressed, and hence DRM and MES-4 are essential components for maintaining the proper dosage of X-linked gene products in the germline.

An underlining molecular mechanism by which autosome-enriched DRM and MES-4 proteins act at a distance to control X-linked genes remains unknown. Since we suspect the involvement of the MES-2/3/6 (PRC2) complex, we considered a model (Chapter III, Figure 2-7) in which DRM normally anchors H3K27me, and therefore a disruption of DRM causes a re-distribution of both H3K27me, a repressing genome mark, and H3K36me, an activating genome mark. A second possibility is that both DRM and MES proteins influence the distribution of a different repressor "R" but in opposite directions. In DRM mutants, the "R" would concentrate more on X chromosomes, causing the enhancement of X chromosome silencing, where as in *mes-4* mutants, "R" would be re-distributed toward autosomes, causing de-silencing of X chromosomes.

We hereby propose an effective approach to find novel regulators involved in hermaphrodite X silencing and/or the global regulator “R” by using genetic predictions and RNAi or genetic screens. If the concentration of “R” dictates X-linked gene expression, we predict the depletion or mutation of “R” would compromise the X chromosome silencing effect imposed by the *lin-54* mutation, and reciprocally it would abolish the X chromosome de-silencing effect of the *mes-4* mutation. Hence, the use of a good phenotypic read-out or a reporter for X-linked gene expression will allow us to conduct RNAi/genetic screens to identify “R”.

We conducted a pilot scale RNAi screen of factors based on their identifies and reported genetic interactions (the majority are synMuv B or synMuv suppressor class chromatin factors, Table 4-1). Here we used *lin-54* sterility as a phenotypic read-out. We reasoned that a high concentration of “R” on the X chromosomes would cause enhanced X-silencing and subsequently the sterility of the *lin-54* mutant animals, and therefore, the removal of “R” would alleviate *lin-54* sterility. After each candidate was depleted by RNAi in the *lin-54* mutant background, our preliminary data found that the depletion of *hpl-2* transcript causes mild suppression in *lin-54* sterility. HPL-2 is the heterochromatin protein 1 (HP1) homologue in *C.elegans*, and in other organisms HP1 mediates gene silencing by binding to heterochromatic histone modification H3K9me. Therefore, HPL-2 is an intriguing possibility, and certainly follow up studies and a large scale RNAi screen are planned. We are mindful that using sterility may not be the best phenotypic read-out, since the *lin-54;mes-4* double mutants showed a confusing allele specificity in the suppression of the sterility (Table 3-1). Thus, construction of a GFP reporter fused with “LIN-54 & MES-4 responsive genes (Chapter III)” may serve as a

better tool. This GFP reporter and a large scale RNAi screen will likely produce promising results and allow us to resolve this mysterious phenomenon. In addition, a thorough screen in the future could help to identify additional players involved in germline X chromosomes gene regulation. More broadly, such follow up screen and study could help to elucidate how conserved transcription factors and chromatin modifiers interact to govern developmental gene regulation.

Why are the hermaphrodite X chromosomes silenced?

An especially confounding question is why both X chromosomes in hermaphrodite germline are repressed? It appears this method of reducing gene expression is independent of the MSCI process, because the two X chromosomes properly pair and recombine during meiosis. Given hermaphroditic nematodes conduct spermatogenesis at the last larvae stage, hermaphrodite X silencing could be reminiscent of having a male germline program. However, this possibility was also excluded by showing gonochoristic nematode species (nematodes with designated female and male) similarly represses the two X chromosomes in female germline. In another theory, the two X chromosomes are thought to be repressed because transcriptionally active X chromosomes would impair processes essential for the integrity of germ cells, such as meiotic crossovers or double-strand break repairs (Kelly 2002).

An alternative theory sees this phenomenon as a germline form of dosage compensation. Dosage compensation is a process to equalize X-linked gene products between the two sexes. If the X chromosome is only silenced in the male germline via

MSCI-related mechanism, then the X-linked gene products between the germ lines of the two sexes would be imbalanced. To compensate for such an imbalance, nematode species might have evolved complementary X-silencing mechanisms in hermaphrodites or females. Indeed, despite the fact that the X chromosome is under-represented for genes specifically required for germline processes, many ubiquitously expressed house-keeping genes are still located on the X chromosome and expressed in germline. Therefore, this theory sounds plausible. Another related hypothesis is that keeping the X chromosome repressed may help to establish the proper ratio of autosome vs. X-linked gene dosage in germline. The skewed X:A ratio may lead to detrimental effects in *C.elegans* germline integrity, and thus the maintenance of the proper ratio may need to be maintained by the interplay of DRM complex and MES proteins. The exact biological consequences of having imbalanced germline X:A ratio is still under investigation. A more important question is whether hermaphrodite X-silencing (female X-silencing in other organisms) is evolutionarily conserved in other organisms. So far no evidence for such X-silencing in female germline is reported in any systems except for nematodes. If such is indeed the case, then why have only nematodes evolved such a specialized X silencing system? Alternatively, it is still possible that the studies have not thoroughly investigated this phenomenon, and maybe such female silencing may exist in other systems.

The Broader Implication of This Study

In this study, we found the antagonistic interplay between DRM and MES-4 in regulating X-linked genes and some autosomal genes in the germline. We suspect such antagonistic interplay may ensure a proper balance of an X-linked vs. autosomal gene expression ratio in *C.elegans* germlines. In broader sense, such "balance between X vs. A" may essentially mean the "balance of chromatin marks" in the genome and therefore gene expression, governed by methyltransferases. In conclusion, the interplay among DREAM, MES-4, and MES-2/3/6 PRC2 complex in hermaphrodite X silencing concerns maintaining the repressive and active chromatin environments in certain domains on chromosomes. If such is the case, then we also suspect the similar interplay in higher eukaryotes.

Like DREAM, MES proteins are conserved in higher eukaryotes and play pivotal roles in development and cell cycle gene regulation. For example, in flies the PRC2 complex maintains repression of developmentally regulated homeotic genes (Moss and Wallrath, 2007). In mammals, the MES-2 homologue EZH2 ensures proper embryo differentiation and stem cell maintenance (Moss and Wallrath, 2007). Additionally, like DREAM, the activities of the PRC2 complex and MES-4 homologue NSD1 are altered or disrupted in many cancer cells (Morishita and di Luccio, 2011; Moss and Wallrath, 2007; Rayasam et al., 2003). Although the link between DREAM and MES-4 homologue NSD1 is not clear at present, lines of evidence suggests a strong connection between DREAM and PRC2 complex in higher eukaryotes. For example, in flies the DREAM components LIN-53/CAF1 (RbAp48) share subunits with the PRC2 complex. In flies and mammal, MES-2 homologue E(Z) and EZH2 physically interacts with pRb-family protein,

respectively (Lee et al., 2010; Tonini et al., 2004). Further, the repression of DREAM target genes requires the activity of the PRC2 complex members (Lee et al., 2010; Tonini et al., 2004). Thus, it is intriguing to think if the similar antagonistic interplay between DREAM and MES homologues is conserved in human.

	genetic interactions
cfpl-1	synMuv Suppressor
ekl-4	synMuv Suppressor
hda2	HDAC
hpl-2	synMuv B
htz-1	synMuv Suppressor
isw-1	synMuv Suppressor, interact with MES-4
lin-53	synMuv B, DRM, NuRD complex member
lin-59	SET domain (HMT)
lss-4	synMuv Suppressor, ISW/SNF
mrg-1	synMuv Suppressor
mys-1	synMuv C, Tip60, NuA4
mys-4	HAT
psa-4	ISW/SNF
set-2	SET domain(HMT) , 1 generation early sterility (Xu)
sin-3	synMuv Suppressor
trr-1	synMuv C, Tip60, NuA4
zk1127.3	synMuv Suppressor
psa-1	ISW/SNF
xnp-1	ISW/SNF
hda1	HDAC
hcp-6	condensin II
mix-1	condensin (I, Idc, II)
mes-3	MES-2/3/6 complex (PRC2)
mes-2	MES-2/3/6 complex (PRC2)
mes-6	MES-2/3/6 complex (PRC2)
Empty Vector	negative control
lin-54	positive control
mes-4	positive control

Table 4-1
List of candidate genes used in the small scale RNAi screen for a global repressor “R”

References

Andersen, E. C., Lu, X., and Horvitz, H. R. (2006). C. elegans ISWI and NURF301 antagonize an Rb-like pathway in the determination of multiple cell fates. Development *133*, 2695-2704.

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., and Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res *37*, W202-208.

Beall, E. L., Bell, M., Georlette, D., and Botchan, M. R. (2004). Dm-myb mutant lethality in Drosophila is dependent upon mip130: positive and negative regulation of DNA replication. Genes Dev *18*, 1667-1680.

Beall, E. L., Lewis, P. W., Bell, M., Rocha, M., Jones, D. L., and Botchan, M. R. (2007). Discovery of tMAC: a Drosophila testis-specific meiotic arrest complex paralogous to Myb-Muv B. Genes Dev *21*, 904-919.

Beall, E. L., Manak, J. R., Zhou, S., Bell, M., Lipsick, J. S., and Botchan, M. R. (2002). Role for a Drosophila Myb-containing protein complex in site-specific DNA replication. Nature *420*, 833-837.

Bean, C. J., Schaner, C. E., and Kelly, W. G. (2004). Meiotic pairing and imprinted X chromatin assembly in Caenorhabditis elegans. Nat Genet *36*, 100-105.

Beitel, G. J., Lambie, E. J., and Horvitz, H. R. (2000). The C. elegans gene lin-9,which acts in an Rb-related pathway, is required for gonadal sheath cell development and encodes a novel protein. Gene *254*, 253-263.

Bender, L. B., Cao, R., Zhang, Y., and Strome, S. (2004). The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in C. elegans. Curr Biol *14*, 1639-1643.

Bender, L. B., Suh, J., Carroll, C. R., Fong, Y., Fingerman, I. M., Briggs, S. D., Cao, R., Zhang, Y., Reinke, V., and Strome, S. (2006). MES-4: an autosome-associated histone methyltransferase that participates in silencing the X chromosomes in the C. elegans germ line. Development *133*, 3907-3917.

Boyle, E. I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J. M., and Sherlock, G. (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics *20*, 3710-3715.

Bridges, C. B. (1916). Non-Disjunction as Proof of the Chromosome Theory of Heredity (Concluded). Genetics *1*, 107-163.

Brzeska, K., Brzeski, J., Smith, J., and Chandler, V. L. (2010). Transgenic expression of CBBP, a CXC domain protein, establishes paramutation in maize. Proc Natl Acad Sci U S A *107*, 5516-5521.

Buck, M. J., Nobel, A. B., and Lieb, J. D. (2005). ChIPOTle: a user-friendly tool for the analysis of ChIP-chip data. Genome Biol *6*, R97.

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

Carrington, W. A., Lynch, R. M., Moore, E. D., Isenberg, G., Fogarty, K. E., and Fay, F. S. (1995). Superresolution three-dimensional images of fluorescence in cells with minimal light exposure. Science *268*, 1483-1487.

Ceol, C. J., and Horvitz, H. R. (2001). dpl-1 DP and efl-1 E2F act with lin-35 Rb to antagonize Ras signaling in C. elegans vulval development. Mol Cell *7*, 461-473.

Chi, W., and Reinke, V. (2006). Promotion of oogenesis and embryogenesis in the C. elegans gonad by EFL-1/DPL-1 (E2F) does not require LIN-35 (pRB). Development *133*, 3147-3157.

Clarke, A. R., Maandag, E. R., van Roon, M., van der Lugt, N. M., van der Valk, M., Hooper, M. L., Berns, A., and te Riele, H. (1992). Requirement for a functional Rb-1 gene in murine development. Nature *359*, 328-330.

Csankovszki, G., Collette, K., Spahl, K., Carey, J., Snyder, M., Petty, E., Patel, U., Tabuchi, T., Liu, H., McLeod, I.*, et al.* (2009). Three distinct condensin complexes control C. elegans chromosome dynamics. Curr Biol *19*, 9-19.

Cui, M., Chen, J., Myers, T. R., Hwang, B. J., Sternberg, P. W., Greenwald, I., and Han, M. (2006a). SynMuv genes redundantly inhibit lin-3/EGF expression to prevent inappropriate vulval induction in C. elegans. Dev Cell *10*, 667-672.

Cui, M., Kim, E. B., and Han, M. (2006b). Diverse chromatin remodeling genes antagonize the Rb-involved SynMuv pathways in C. elegans. PLoS Genet *2*, e74.

Deplancke, B., Mukhopadhyay, A., Ao, W., Elewa, A. M., Grove, C. A., Martinez, N. J., Sequerra, R., Doucette-Stamm, L., Reece-Hoyes, J. S., Hope, I. A.*, et al.* (2006). A gene-centered C. elegans protein-DNA interaction network. Cell *125*, 1193-1205.
Dimova, D. K., Stevaux, O., Frolov, M. V., and Dyson, N. J. (2003). Cell cycle-dependent and cell cycle-independent control of transcription by the Drosophila E2F/RB pathway. Genes Dev *17*, 2308-2320.
Doniach, T. (1986). Activity of the sex-determining gene tra-2 is modulated to allow spermatogenesis in the C. elegans hermaphrodite. Genetics *114*, 53-76.
Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A *95*, 14863-14868.
Ellegren, H., and Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. Nat Rev Genet *8*, 689-698.
Ercan, S., and Lieb, J. D. (2009). C. elegans dosage compensation: a window into mechanisms of domain-scale gene regulation. Chromosome Res *17*, 215-227.
Fauth, T., Muller-Planitz, F., Konig, C., Straub, T., and Becker, P. B. (2010). The DNA binding CXC domain of MSL2 is required for faithful targeting the Dosage Compensation Complex to the X chromosome. Nucleic Acids Res *38*, 3209-3221.
Fay, D. S., and Yochem, J. (2007). The SynMuv genes of Caenorhabditis elegans in vulval development and beyond. Dev Biol *306*, 1-9.
Ferguson, E. L., and Horvitz, H. R. (1989). The multivulva phenotype of certain Caenorhabditis elegans mutants results from defects in two functionally redundant pathways. Genetics *123*, 109-121.
Fire, A., Alcazar, R., and Tan, F. (2006). Unusual DNA structures associated with germline genetic activity in Caenorhabditis elegans. Genetics *173*, 1259-1273.
Fong, Y., Bender, L., Wang, W., and Strome, S. (2002). Regulation of the different chromatin states of autosomes and X chromosomes in the germ line of C. elegans. Science *296*, 2235-2238.
Freedman, J. A., Chang, J. T., Jakoi, L., and Nevins, J. R. (2009). A combinatorial mechanism for determining the specificity of E2F activation and repression. Oncogene *28*, 2873-2881.
Garvin, C., Holdeman, R., and Strome, S. (1998). The phenotype of mes-2, mes-3, mes-4 and mes-6, maternal-effect genes required for survival of the germline in Caenorhabditis elegans, is sensitive to chromosome dosage. Genetics *148*, 167-185.
Georlette, D., Ahn, S., MacAlpine, D. M., Cheung, E., Lewis, P. W., Beall, E. L., Bell, S. P., Speed, T., Manak, J. R., and Botchan, M. R. (2007). Genomic profiling and expression studies reveal both positive and negative activities for the Drosophila Myb MuvB/dREAM complex in proliferating cells. Genes Dev *21*, 2880-2896.
Grote, P., and Conradt, B. (2006). The PLZF-like protein TRA-4 cooperates with the Gli-like transcription factor TRA-1 to promote female development in C. elegans. Dev Cell *11*, 561-573.
Grove, C. A., De Masi, F., Barrasa, M. I., Newburger, D. E., Alkema, M. J., Bulyk, M. L., and Walhout, A. J. (2009). A multiparameter network reveals extensive divergence between C. elegans bHLH transcription factors. Cell *138*, 314-327.
Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature *346*, 245-250.
Gupta, V., Parisi, M., Sturgill, D., Nuttall, R., Doctolero, M., Dudko, O. K., Malley, J. D., Eastman, P. S., and Oliver, B. (2006). Global analysis of X-chromosome dosage compensation. J Biol *5*, 3.
Gurbich, T. A., and Bachtrog, D. (2008). Gene content evolution on the X chromosome. Curr Opin Genet Dev *18*, 493-498.
Harrison, M. M., Ceol, C. J., Lu, X., and Horvitz, H. R. (2006). Some C. elegans class B synthetic multivulva proteins encode a conserved LIN-35 Rb-containing complex distinct from a NuRD-like complex. Proc Natl Acad Sci U S A *103*, 16782-16787.
Hodgkin, J., Horvitz, H. R., and Brenner, S. (1979). Nondisjunction Mutants of the Nematode CAENORHABDITIS ELEGANS. Genetics *91*, 67-94.
Holdeman, R., Nehrt, S., and 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 *125*, 2457-2467.
Hu, Z., Killion, P. J., and Iyer, V. R. (2007). Genetic reconstruction of a functional transcriptional regulatory network. Nat Genet *39*, 683-687.

Ihaka, R., and Gentleman, R. (1996). R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics *5*, 299-314.
Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B., and Speed, T. P. (2003a). Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res *31*, e15.
Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003b). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249-264.
Jacks, T., Fazeli, A., Schmitt, E. M., Bronson, R. T., Goodell, M. A., and Weinberg, R. A. (1992). Effects of an Rb mutation in the mouse. Nature *359*, 295-300.
Jans, J., Gladden, J. M., Ralston, E. J., Pickle, C. S., Michel, A. H., Pferdehirt, R. R., Eisen, M. B., and Meyer, B. J. (2009). A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome. Genes Dev *23*, 602-618.
Jiang, J., Benson, E., Bausek, N., Doggett, K., and White-Cooper, H. (2007). Tombola, a tesmin/TSO1-family protein, regulates transcriptional activation in the Drosophila male germline and physically interacts with always early. Development *134*, 1549-1559.
Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M.*, et al.* (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature *421*, 231-237.
Katzen, A. L., Kornberg, T. B., and Bishop, J. M. (1985). Isolation of the proto-oncogene c-myb from D. melanogaster. Cell *41*, 449-456.
Kelly, W. G., and Aramayo, R. (2007). Meiotic silencing and the epigenetics of sex. Chromosome Res *15*, 633-651.
Kelly, W. G., and Fire, A. (1998). Chromatin silencing and the maintenance of a functional germline in Caenorhabditis elegans. Development *125*, 2451-2456.
Kelly, W. G., Schaner, C. E., Dernburg, A. F., Lee, M. H., Kim, S. K., Villeneuve, A. M., and Reinke, V. (2002). X-chromosome silencing in the germline of C. elegans. Development *129*, 479-492.
Keogh, M. C., Kurdistani, S. K., Morris, S. A., Ahn, S. H., Podolny, V., Collins, S. R., Schuldiner, M., Chin, K., Punna, T., Thompson, N. J.*, et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.
Kirienko, N. V., and Fay, D. S. (2007). Transcriptome profiling of the C. elegans Rb ortholog reveals diverse developmental roles. Dev Biol *305*, 674-684.
Knight, A. S., Notaridou, M., and Watson, R. J. (2009). A Lin-9 complex is recruited by B-Myb to activate transcription of G2/M genes in undifferentiated embryonal carcinoma cells. Oncogene *28*, 1737-1747.
Koester-Eiserfunke, N., and Fischle, W. (2011). H3K9me2/3 binding of the MBT domain protein LIN-61 is essential for Caenorhabditis elegans vulva development. PLoS Genet *7*, e1002017.
Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1991). Male development of chromosomally female mice transgenic for Sry. Nature *351*, 117-121.
Korenjak, M., Taylor-Harding, B., Binne, U. K., Satterlee, J. S., Stevaux, O., Aasland, R., White-Cooper, H., Dyson, N., and Brehm, A. (2004). Native E2F/RBF complexes contain Myb-interacting proteins and repress transcription of developmentally controlled E2F target genes. Cell *119*, 181-193.
Korf, I., Fan, Y., and Strome, S. (1998). The Polycomb group in Caenorhabditis elegans and maternal control of germline development. Development *125*, 2469-2478.
Kuwabara, P. E., Okkema, P. G., and Kimble, J. (1992). tra-2 encodes a membrane protein and may mediate cell communication in the Caenorhabditis elegans sex determination pathway. Mol Biol Cell *3*, 461-473.
Lee, E. Y., Chang, C. Y., Hu, N., Wang, Y. C., Lai, C. C., Herrup, K., Lee, W. H., and Bradley, A. (1992). Mice deficient for Rb are nonviable and show defects in neurogenesis and haematopoiesis. Nature *359*, 288-294.
Lee, H., Ohno, K., Voskoboynik, Y., Ragusano, L., Martinez, A., and Dimova, D. K. (2010). Drosophila RB proteins repress differentiation-specific genes via two different mechanisms. Mol Cell Biol *30*, 2563-2577.
Lehner, B., Calixto, A., Crombie, C., Tischler, J., Fortunato, A., Chalfie, M., and Fraser, A. G. (2006). Loss of LIN-35, the Caenorhabditis elegans ortholog of the tumor suppressor p105Rb, results in enhanced RNA interference. Genome Biol *7*, R4.
Lewis, P. W., Beall, E. L., Fleischer, T. C., Georlette, D., Link, A. J., and Botchan, M. R. (2004). Identification of a Drosophila Myb-E2F2/RBF transcriptional repressor complex. Genes Dev *18*, 2929-2940.

Litovchick, L., Florens, L. A., Swanson, S. K., Washburn, M. P., and DeCaprio, J. A. (2011). DYRK1A protein kinase promotes quiescence and senescence through DREAM complex assembly. In Genes Dev, pp. 801-813.
Litovchick, L., Sadasivam, S., Florens, L., Zhu, X., Swanson, S. K., Velmurugan, S., Chen, R., Washburn, M. P., Liu, X. S., and DeCaprio, J. A. (2007). Evolutionarily conserved multisubunit RBL2/p130 and E2F4 protein complex represses human cell cycle-dependent genes in quiescence. Mol Cell *26*, 539-551.
Maciejowski, J., Ahn, J. H., Cipriani, P. G., Killian, D. J., Chaudhary, A. L., Lee, J. I., Voutev, R., Johnsen, R. C., Baillie, D. L., Gunsalus, K. C.*, et al.* (2005). Autosomal genes of autosomal/X-linked duplicated gene pairs and germ-line proliferation in Caenorhabditis elegans. Genetics *169*, 1997-2011.
Madl, J. E., and Herman, R. K. (1979). Polyploids and sex determination in Caenorhabditis elegans. Genetics *93*, 393-402.
Maine, E. M., Hauth, J., Ratliff, T., Vought, V. E., She, X., and Kelly, W. G. (2005). EGO-1, a putative RNA-dependent RNA polymerase, is required for heterochromatin assembly on unpaired dna during C. elegans meiosis. Curr Biol *15*, 1972-1978.
Manak, J. R., Mitiku, N., and Lipsick, J. S. (2002). Mutation of the Drosophila homologue of the Myb protooncogene causes genomic instability. Proc Natl Acad Sci U S A *99*, 7438-7443.
Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., 3rd, Voigt, P., Martin, S. R., Taylor, W. R., De Marco, V.*, et al.* (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. Nature *461*, 762-767.
Mendjan, S., and Akhtar, A. (2007). The right dose for every sex. Chromosoma *116*, 95-106.
Meyer, B. J. (2010). Targeting X chromosomes for repression. Curr Opin Genet Dev *20*, 179-189.
Morishita, M., and di Luccio, E. (2011). Cancers and the NSD family of histone lysine methyltransferases. Biochim Biophys Acta *1816*, 158-163.
Moss, T. J., and Wallrath, L. L. (2007). Connections between epigenetic gene silencing and human disease. Mutat Res *618*, 163-174.
Mukhopadhyay, A., Deplancke, B., Walhout, A. J., and Tissenbaum, H. A. (2008). Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in Caenorhabditis elegans. Nat Protoc *3*, 698-709.
Muller, G. A., and Engeland, K. (2010). The central role of CDE/CHR promoter elements in the regulation of cell cycle-dependent gene transcription. FEBS J *277*, 877-893.
Namekawa, S. H., Park, P. J., Zhang, L. F., Shima, J. E., McCarrey, J. R., Griswold, M. D., and Lee, J. T. (2006). Postmeiotic sex chromatin in the male germline of mice. Curr Biol *16*, 660-667.
Nguyen, D. K., and Disteche, C. M. (2006). Dosage compensation of the active X chromosome in mammals. Nat Genet *38*, 47-53.
Osterloh, L., von Eyss, B., Schmit, F., Rein, L., Hubner, D., Samans, B., Hauser, S., and Gaubatz, S. (2007). The human synMuv-like protein LIN-9 is required for transcription of G2/M genes and for entry into mitosis. Embo J *26*, 144-157.
Ow, M. C., Martinez, N. J., Olsen, P. H., Silverman, H. S., Barrasa, M. I., Conradt, B., Walhout, A. J., and Ambros, V. (2008). The FLYWCH transcription factors FLH-1, FLH-2, and FLH-3 repress embryonic expression of microRNA genes in C. elegans. Genes Dev *22*, 2520-2534.
Page, B. D., Guedes, S., Waring, D., and Priess, J. R. (2001). The C. elegans E2F- and DP-related proteins are required for embryonic asymmetry and negatively regulate Ras/MAPK signaling. Mol Cell *7*, 451-460.
Parsch, J. (2009). X chromosome: expression and escape. PLoS Genet *5*, e1000724.
Payer, B., and Lee, J. T. (2008). X chromosome dosage compensation: how mammals keep the balance. Annu Rev Genet *42*, 733-772.
Petrella, L. N., Wang, W., Spike, C. A., Rechtsteiner, A., Reinke, V., and Strome, S. (2011). synMuv B proteins antagonize germline fate in the intestine and ensure C. elegans survival. Development *138*, 1069-1079.
Petty, E. L., Collette, K. S., Cohen, A. J., Snyder, M. J., and Csankovszki, G. (2009). Restricting dosage compensation complex binding to the X chromosomes by H2A.Z/HTZ-1. PLoS Genet *5*, e1000699.
Piano, F., Schetter, A. J., Mangone, M., Stein, L., and Kemphues, K. J. (2000). RNAi analysis of genes expressed in the ovary of Caenorhabditis elegans. Curr Biol *10*, 1619-1622.
Pilkinton, M., Sandoval, R., and Colamonici, O. R. (2007). Mammalian Mip/LIN-9 interacts with either the p107, p130/E2F4 repressor complex or B-Myb in a cell cycle-phase-dependent context distinct from the Drosophila dREAM complex. Oncogene *26*, 7535-7543.

Rayasam, G. V., Wendling, O., Angrand, P. O., Mark, M., Niederreither, K., Song, L., Lerouge, T., Hager, G. L., Chambon, P., and Losson, R. (2003). NSD1 is essential for early post-implantation development and has a catalytically active SET domain. Embo J *22*, 3153-3163.
Rechtsteiner, A., Ercan, S., Takasaki, T., Phippen, T. M., Egelhofer, T. A., Wang, W., Kimura, H., Lieb, J. D., and Strome, S. (2010). The histone H3K36 methyltransferase MES-4 acts epigenetically to transmit the memory of germline gene expression to progeny. PLoS Genet *6*.
Reddien, P. W., Andersen, E. C., Huang, M. C., and Horvitz, H. R. (2007). DPL-1 DP, LIN-35 Rb and EFL-1 E2F act with the MCD-1 zinc-finger protein to promote programmed cell death in Caenorhabditis elegans. Genetics *175*, 1719-1733.
Reinke, V., Gil, I. S., Ward, S., and Kazmer, K. (2004). Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. Development *131*, 311-323.
Reinke, V., Smith, H. E., Nance, J., Wang, J., Van Doren, C., Begley, R., Jones, S. J., Davis, E. B., Scherer, S., Ward, S., and Kim, S. K. (2000). A global profile of germline gene expression in C. elegans. Mol Cell *6*, 605-616.
Saldanha, A. J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics *20*, 3246-3248.
Sandoval, R., Pilkinton, M., and Colamonici, O. R. (2009). Deletion of the p107/p130-binding domain of Mip130/LIN-9 bypasses the requirement for CDK4 activity for the dissociation of Mip130/LIN-9 from p107/p130-E2F4 complex. Exp Cell Res *315*, 2914-2920.
Schertel, C., and Conradt, B. (2007). C. elegans orthologs of components of the RB tumor suppressor complex have distinct pro-apoptotic functions. Development *134*, 3691-3701.
Schmit, F., Cremer, S., and Gaubatz, S. (2009). LIN54 is an essential core subunit of the DREAM/LINC complex that binds to the cdc2 promoter in a sequence-specific manner. FEBS J *276*, 5703-5716.
Schmit, F., Korenjak, M., Mannefeld, M., Schmitt, K., Franke, C., von Eyss, B., Gagrica, S., Hanel, F., Brehm, A., and Gaubatz, S. (2007). LINC, a human complex that is related to pRB-containing complexes in invertebrates regulates the expression of G2/M genes. Cell Cycle *6*, 1903-1913.
Segal, E., and Widom, J. (2009). What controls nucleosome positions? Trends Genet *25*, 335-343.
Shaham, S. E. (2006). Methods in Cell Biology).
Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A. M., Lovell-Badge, R., and Goodfellow, P. N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature *346*, 240-244.
Song, J. S., Johnson, W. E., Zhu, X., Zhang, X., Li, W., Manrai, A. K., Liu, J. S., Chen, R., and Liu, X. S. (2007). Model-based analysis of two-color arrays (MA2C). Genome Biol *8*, R178.
Strome, S., and Wood, W. B. (1983). Generation of asymmetry and segregation of germ-line granules in early C. elegans embryos. Cell *35*, 15-25.
Sugihara, T., Wadhwa, R., Kaul, S. C., and Mitsui, Y. (1999). A novel testis-specific metallothionein-like protein, tesmin, is an early marker of male germ cell differentiation. Genomics *57*, 130-136.
Tabuchi, T. M., Deplancke, B., Osato, N., Zhu, L. J., Barrasa, M. I., Harrison, M. M., Horvitz, H. R., Walhout, A. J., and Hagstrom, K. A. (2011). Chromosome-Biased Binding and Gene Regulation by the Caenorhabditis elegans DRM Complex. PLoS Genet *7*, e1002074.
Taddei, A., Van Houwe, G., Nagai, S., Erb, I., van Nimwegen, E., and Gasser, S. M. (2009). The functional importance of telomere clustering: global changes in gene expression result from SIR factor dispersion. Genome Res *19*, 611-625.
Takasaki, T., Liu, Z., Habara, Y., Nishiwaki, K., Nakayama, J., Inoue, K., Sakamoto, H., and Strome, S. (2007). MRG-1, an autosome-associated protein, silences X-linked genes and protects germline immortality in Caenorhabditis elegans. Development *134*, 757-767.
Talbert, P. B., and Henikoff, S. (2006). Spreading of silent chromatin: inaction at a distance. Nat Rev Genet *7*, 793-803.
Thomas, J. H., Ceol, C. J., Schwartz, H. T., and Horvitz, H. R. (2003). New genes that interact with lin-35 Rb to negatively regulate the let-60 ras pathway in Caenorhabditis elegans. Genetics *164*, 135-151.
Tonini, T., Bagella, L., D'Andrilli, G., Claudio, P. P., and Giordano, A. (2004). Ezh2 reduces the ability of HDAC1-dependent pRb2/p130 transcriptional repression of cyclin A. Oncogene *23*, 4930-4937.
Turner, J. M. (2007). Meiotic sex chromosome inactivation. Development *134*, 1823-1831.

Unhavaithaya, Y., Shin, T. H., Miliaras, N., Lee, J., Oyama, T., and Mello, C. C. (2002). MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in C. elegans. Cell *111*, 991-1002.
van den Heuvel, S., and Dyson, N. J. (2008). Conserved functions of the pRB and E2F families. Nat Rev Mol Cell Biol *9*, 713-724.
van Leeuwen, F., and Gottschling, D. E. (2002). Genome-wide histone modifications: gaining specificity by preventing promiscuity. Curr Opin Cell Biol *14*, 756-762.
Vibranovski, M. D., Lopes, H. F., Karr, T. L., and Long, M. (2009). Stage-specific expression profiling of Drosophila spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. PLoS Genet *5*, e1000731.
Vicoso, B., and Charlesworth, B. (2006). Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet *7*, 645-653.
Villeneuve, A. M. (1994). A cis-acting locus that promotes crossing over between X chromosomes in Caenorhabditis elegans. Genetics *136*, 887-902.
Walhout, A. J., Boulton, S. J., and Vidal, M. (2000). Yeast two-hybrid systems and protein interaction mapping projects for yeast and worm. Yeast *17*, 88-94.
Walstrom, K. M., Schmidt, D., Bean, C. J., and Kelly, W. G. (2005). RNA helicase A is important for germline transcriptional control, proliferation, and meiosis in C. elegans. Mech Dev *122*, 707-720.
Wang, D., Kennedy, S., Conte, D., Jr., Kim, J. K., Gabel, H. W., Kamath, R. S., Mello, C. C., and Ruvkun, G. (2005). Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants. Nature *436*, 593-597.
Wang, X., Zhao, Y., Wong, K., Ehlers, P., Kohara, Y., Jones, S. J., Marra, M. A., Holt, R. A., Moerman, D. G., and Hansen, D. (2009). Identification of genes expressed in the hermaphrodite germ line of C. elegans using SAGE. BMC Genomics *10*, 213.
Wen, H., Andrejka, L., Ashton, J., Karess, R., and Lipsick, J. S. (2008). Epigenetic regulation of gene expression by Drosophila Myb and E2F2-RBF via the Myb-MuvB/dREAM complex. Genes Dev *22*, 601-614.
Wettenhall, J. M., and Smyth, G. K. (2004). limmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics *20*, 3705-3706.
Whittle, C. M., Lazakovitch, E., Gronostajski, R. M., and Lieb, J. D. (2009). DNA-binding specificity and in vivo targets of Caenorhabditis elegans nuclear factor I. Proc Natl Acad Sci U S A *106*, 12049-12054.
Xiong, Y., Chen, X., Chen, Z., Wang, X., Shi, S., Wang, X., Zhang, J., and He, X. (2010). RNA sequencing shows no dosage compensation of the active X-chromosome. Nat Genet *42*, 1043-1047.
Xu, L., Fong, Y., and Strome, S. (2001a). The Caenorhabditis elegans maternal-effect sterile proteins, MES-2, MES-3, and MES-6, are associated in a complex in embryos. Proc Natl Acad Sci U S A *98*, 5061-5066.
Xu, L., Paulsen, J., Yoo, Y., Goodwin, E. B., and Strome, S. (2001b). Caenorhabditis elegans MES-3 is a target of GLD-1 and functions epigenetically in germline development. Genetics *159*, 1007-1017.
Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J Biol Chem *286*, 7983-7989.